Aus der Klinik und Poliklinik für Psychiatrie und Psychotherapie der

Ludwig-Maximilians-Universität zu München

Dissertation

zum Erwerb des Doctor of Philosophy (Ph.D.) an

der

Medizinischen Fakultät der

Ludwig-Maximilians-Universität zu München

Investigation of Childhood Trauma as a transdiagnostic risk factor using Multimodal

Machine Learning

vorgelegt von:

Dr. med. David Popovic

aus:

München, Deutschland

Jahr:

2023

Mit Genehmigung der Medizinischen Fakultät der

Ludwig-Maximilians-Universität zu München

First supervisor: *Prof. Dr. med. Nikolaos Koutsouleris*

Second supervisor*:* *Prof. Dr. med. Andrea Schmitt*

Third supervisor: *Prof. Dr. Alon Chen*

Dean: **Prof. Dr. med. Thomas Gudermann**

Datum der Verteidigung:

24.01.2023

Affidavit

| LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN | Promotionsbüro Medizinische Fakultät | MMRS | |

Affidavit

Popovic, David

_____

Surname, first name

Nussbaumstrasse 7

_____

Street

80336 Munich, Germany

_____

Zip code, town, country

I hereby declare that the submitted thesis entitled:

'Investigation of Childhood Trauma as a transdiagnostic risk factor using Multimodal Machine Learning'

is my own work. I have only used the sources indicated and have not made unauthorized use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.

I further declare that the submitted thesis or parts thereof have not been presented as part of an examination degree to any other university.

Munich, 10.02.2023   David Popovic
_____   _____
Place, date   Signature doctoral candidate

Confirmation of congruency

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN		Promotionsbüro Medizinische Fakultät	MMRS	

Confirmation of congruency between printed and electronic version of the doctoral thesis

Popovic, David
_____
Surname, first name

Nussbaumstrasse 7
_____
Street

80336 Munich, Germany
_____
Zip code, town, country

I hereby declare that the submitted thesis entitled:

'Investigation of Childhood Trauma as a transdiagnostic risk factor using Multimodal Machine Learning'

is congruent with the printed version both in content and format.

Munich, 10.02.2023   David Popovic
_____   _____
Place, date   Signature doctoral candidate

Table of content

List of abbreviations

AUC   Area under the curve

BAC   Balanced Accuracy

CHR   Clinical high-risk

CT   Childhood Trauma

CV   Cross-Validation

DALY   Disability-adjusted life-year

GDP   Gross domestic product

GMV   Grey matter volume

HC   Healthy control individuals

HPA axis   Hypothalamic-pituitary-adrenal axis

LV   Latent variable

ML   Machine Learning

MR(I)   Magnetic resonance (imaging)

PAT   Patients

PCA   Principal component analysis

PLS   Partial Least Squares

PRONIA   Personalised Prognostic Tools for Early Psychosis Management

ROD   Recent onset of depression

ROI   Region of Interest

ROP   Recent onset of psychosis

SPLS   Sparse Partial Least Squares

List of tables

List of publications

Paper I

Popovic, D., Ruef, A., Dwyer, D. B., Antonucci, L. A., Eder, J., Sanfelici, R., Kambeitz-Ilankovic, L., Oeztuerk, O. F., Dong, M. S., Paul, R., Paolini, M., Hedderich, D., Haidl, T., Kambeitz, J., Ruhrmann, S., Chisholm, K., Schultze-Lutter, F., Falkai, P., Pergola, G., … PRONIA Consortium. (2020). Traces of trauma: A multivariate pattern analysis of childhood trauma, brain structure, and clinical phenotypes. Biological Psychiatry, 88(11), 829–842. https://doi.org/10.1016/j.biopsych.2020.05.020

Paper II

Haidl, T. K., Hedderich, D. M., Rosen, M., Kaiser, N., Seves, M., Lichtenstein, T., Penzel, N., Wenzel, J., Kambeitz-Ilankovic, L., Ruef, A., **Popovic, D.**, Schultze-Lutter, F., Chisholm, K., Upthegrove, R., Salokangas, R. K. R., Pantelis, C., Meisenzahl, E., Wood, S. J., Brambilla, P., … Koutsouleris, N. (2021). The non-specific nature of mental health and structural brain outcomes following childhood trauma. Psychological Medicine, 1–10. https://doi.org/10.1017/S0033291721002439

1. Your contribution to the publications

1.1 Contribution to Paper I

Popovic, D., Ruef, A., Dwyer, D. B., Antonucci, L. A., Eder, J., Sanfelici, R., Kambeitz-Ilankovic, L., Oeztuerk, O. F., Dong, M. S., Paul, R., Paolini, M., Hedderich, D., Haidl, T., Kambeitz, J., Ruhrmann, S., Chisholm, K., Schultze-Lutter, F., Falkai, P., Pergola, G., … PRONIA Consortium. (2020). Traces of trauma: A multivariate pattern analysis of childhood trauma, brain structure, and clinical phenotypes. Biological Psychiatry, 88(11), 829–842. https://doi.org/10.1016/j.biopsych.2020.05.020

The analysis strategy for Paper I was conceived by Prof. Dr. Nikolaos Koutsouleris and myself. Specifically, it was the inspiration of Prof. Koutsouleris to implement the Sparse Partial Least Squares (SPLS) algorithm by Monteiro et al. into a MATLAB-based toolbox and then to use this toolbox as the analysis tool for this study. Therefore, the SPLS Toolbox was incepted, developed, and upgraded during the various analyses for Paper I. As the analyses for Paper I progressed, so did the toolbox. Therefore, the analysis for Paper I can be regarded both as a research article on the neuroanatomical correlates of childhood trauma and as a method paper for the SPLS Toolbox. Hence, I was involved in all aspects of Paper I, ranging from conception of the scientific idea, creating the analysis tool with which to conduct the analysis, visualization of the results, interpreting the results, and the writing of the manuscript. While the conception of the study idea was equally undertaken by Prof. Koutsouleris and me, the other parts that I mentioned were my main responsibility and driven by me.

The section 'Acknowledgments and Disclosures' of Paper I gives an even more detailed overview of the individual contributions (DP, David Popovic; NK, Nikolaos Koutsouleris; the other initials reflect the author line):

DP and NK had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. DP, NK, LK-I, SR, JK, PF, RU, EM, SJW, PB, SB, and CP were involved in concept and design. DP, NK, LK-I, SR, AR, DBD, RS, MSD, JE, MP, KC, JK, TH, FS-L, GB, AB, RU, CP, SJW, PB, and SB were involved in acquisition, analysis, or interpretation of data. DP, AR, DBD, LAA, and NK were involved in drafting of the manuscript. DP, NK, LK-I, SR, AR, DBD, LAA, RS, OFO, RP, MP, KC, JK, TH, FS-L, PF, RU, GP, AB, RKRS, CP, EM, SJW, PB, and SB were involved in critical revision of the manuscript for important intellectual content. DP,

AR, and NK were involved in statistical analysis. DP, NK, LK-I, SR, RKRS, CP, PB, SB, and SJW were involved in obtaining funding. NK, AR, MP, KC, TH, DH, RU, EM, AB, PB, and SB were involved in administrative, technical, or material support. NK, SR, FS-L, PF, SJW, PB, AB, RL, RU, SB, UD, and CP were involved in supervision. PRONIA consortium members listed here performed the screening, recruitment, rating, examination, and follow-up of the study participants and were involved in implementing the examination protocols of the study, setting up its information technological infrastructure, and organizing the flow and quality control of the data analyzed in this article between the local study sites and the central study database.

1.2   Contribution to Paper II

Haidl, T. K., Hedderich, D. M., Rosen, M., Kaiser, N., Seves, M., Lichtenstein, T., Penzel, N., Wenzel, J., Kambeitz-Ilankovic, L., Ruef, A., **Popovic, D.**, Schultze-Lutter, F., Chisholm, K., Upthegrove, R., Salokangas, R. K. R., Pantelis, C., Meisenzahl, E., Wood, S. J., Brambilla, P., … Koutsouleris, N. (2021). The non-specific nature of mental health and structural brain outcomes following childhood trauma. Psychological Medicine, 1–10. https://doi.org/10.1017/S0033291721002439

My contribution to Paper II included clinical discussion and conceptualization, methodological supervision, and manuscript writing. First, I was involved in the clinical discussion about this analysis. Specifically, the study was conceptualized to investigate the role of childhood trauma as a transdiagnostic risk factor, increasing the likelihood of various affective, psychotic, and other disorders through subtle, distributed, overarching brain changes. This clinical discussion influenced the methodological approach we took for this study. Hence, I participated in creating the different prediction sets (HC vs. PAT, HC vs. ROD, HC vs. CHR, HC vs. ROP) to investigate the predictive utility of childhood trauma signatures in either delineating transdiagnostic or diagnosis-specific disease patterns. Furthermore, I was involved in the critical revision of the manuscript, regarding machine learning strategies and interpretation of the findings, specifically concerning the transdiagnostic nature of childhood trauma and the impact of emotional trauma on clinical phenotypes.

2.   Introductory Summary

2.1   Childhood Trauma

2.1.1   Conceptualization & Operationalization of Childhood Trauma

The terms 'childhood trauma', 'childhood maltreatment', and 'adverse childhood experiences' are often used synonymously (1,2). In line with Papers I and II, the term childhood trauma (CT) will be used throughout this thesis, encompassing the nomenclature of childhood maltreatment and adverse childhood experiences. The most serious acts of CT, which are directly committed against a child such as physical, emotional, and sexual abuse or neglect, occur most frequently within the nuclear family (1). In fact, in about 80% of abuse or neglect cases, the perpetrator is the child's parent or family member (3,4). Furthermore, there are peer- or community-based forms of CT, such as bullying and discrimination (1). Beyond this, CT also refers to events that are not directly committed against children but to incidents that can still severely affect a child's mental and physical health. Among these, community deprivation and violence, disasters, witnessing intimate partner violence between caregivers, separation from caregivers, and refugee or migration experience are the most common ones (3). In psychiatric research, many instruments evaluate CT in an operationalized manner (5). Among self-reports, the Childhood Trauma Questionnaire (6) is one of the most frequently used and investigated (7). It is also the measure used in Papers I and II (8,9). The Childhood Trauma Questionnaire consists of 28 items and assesses CT along the domains of physical abuse and neglect, emotional abuse, and neglect, as well as sexual abuse. However, there are many other self-reports with varying types of CT assessment, such as the Early Trauma Inventory Self-Report (10,11), the Child Sexual Assaults Scale (12,13), and the Traumatic Life Events Questionnaire (14). Although self-reports are more common, they may suffer from validity issues (15). Hence, observer ratings might be a viable alternative. A very commonly used observer rating is the Early Trauma Inventory, which examines physical, emotional, and sexual abuse, as well as general trauma (16).

2.1.2   The individual and socioeconomic impact of childhood trauma

Decades of research have established CT as a risk factor for physical and mental health (3,17). CT increases the risk of a large array of disorders, such as respiratory diseases, diabetes, obesity, ischemic heart disease, gastrointestinal diseases, chronic pain, headaches, hypertension, stroke, and cancer, thereby heavily afflicting public health. Regarding lifestyle patterns, CT leads to an increased risk of tobacco, alcohol, and illicit

drug abuse, risky sexual behavior, decreased quality of life, behavioral problems, and victimization through violence (18). Therefore, CT is a predictor of psychosis (5,19), affective disorders (20,21), substance abuse (22–25) and borderline personality disorder (26–28) as well as post-traumatic stress disorder (29,30), generalized anxiety disorder (31), obsessive-compulsory disorder (32) and phobias (31).

CT not only disrupts the path of life of the affected individual but is also found very frequently in our societies. In North America, the pooled prevalence for at least one form or instance of CT was calculated at 58.4%, and for two or more it was 35.0% (33). Furthermore, 37.7% of all individuals in 28 European countries reported at least one form or instance of CT, while 15.2% reported two or more (34). The burden of CT can be assessed using the disability-adjusted life-year measure (DALY) (35). DALY is the loss of the equivalent of one year of full health due to mortality, disability, or other disease-related reasons. As Papers I and II investigated the PRONIA cohort (36), which recruited individuals in Germany, Switzerland, Italy, Finland, and the UK, Table 1 (adapted from 34) highlights the individual and socioeconomic impact of CT in the these countries.

Table 1: Annual CT-attributable DALY and costs in PRONIA recruiting countries

Country	CT-attributable DALY (thousands)	CT-attributable costs (US$ billion)	Equivalent % of GDP
Germany	2,796.6	129.4	3.4
Switzerland	250.5	20.5	2.9
Italy	916.2	30.4	1.5
Finland	225.2	11.0	4.1
UK	1,858.7	78.6	2.8

2.1.3   Neurobiological effects of childhood trauma

Considering these profound individual and societal consequences, great efforts have been made to better understand CT and its impact on affected individuals, specifically on a neurobiological level. The neurobiology of CT is related to stress mechanisms (37). Through the cortical pathway, the signal from a perceived stressor first passes the thalamus before entering higher-order cortical structures for further processing. However, there is also an additional subcortical pathway, in which the stimulus directly conveys information to the amygdala and other subcortical structures, eliciting a stress response without apparent awareness by activating the hypothalamic-pituitary-adrenal axis (HPA) (38–40). This cascade is vital to survive critical situations; however, in prolonged critical situations, such as CT, these mechanisms can have a detrimental effect (41). An allostatic overload of the HPA axis, specifically during sensitive periods of development such as

childhood, can lead to long-term maladaptive stress responses through epigenetic alterations and neural adaptations in stress-responsive brain regions (42) and ultimately contribute to the development of mental disorders (37).

For years, neuroimaging studies have employed a region-of-interest (ROI) approach, while whole brain analyses have only recently emerged (37). ROI analyses focused on the hippocampus, which is known to be vulnerable to the effects of stress (43), and the amygdala, which plays a major role in emotion processing, regulation and awareness (44–46). Therefore, the most well-established CT-related functional brain alterations have been found in the amygdala, the hippocampus/parahippocampal gyrus, and associated structures such as the insula and the dorsolateral prefrontal cortex (47–49). Furthermore, CT has been associated with reduced fractional anisotropy in the fornix and in the anterior and posterior thalamic radiations (50,51). The fornix is the main output tract of the hippocampus, while anterior and posterior thalamic radiations link the frontal and occipital lobes to the thalamus (51,52). Grey matter volume (GMV) studies have most consistently established a link between CT and reduced hippocampal GMV (53–55). Beyond that, however, the findings have been heterogeneous. The relationship between CT and amygdala GMV remains elusive; some studies reported decreased GMV and others reported unchanged GMV (56,57). Moreover, of the seven available meta-analyses, only two integrated data from whole brain analyses (54,58), while the others relied on ROI studies (53,56,57,59,60). These whole brain meta-analyses reported varying degrees of reduced GMV in prefronto-temporo-limbic regions, including the hippocampus, the dorsolateral prefrontal cortex, the postcentral gyrus, the superior temporal gyrus, the amygdala, the insula, and the parahippocampal gyrus (54,58).

2.1.4   Rationale for the investigation of childhood trauma by machine learning

Taking into account the current body of evidence on CT and GMV, there are several issues to be addressed (61). First, future analyses must adequately consider the most important moderators of CT. Age is one of the strongest factors that influence GMV throughout life (62). Furthermore, psychiatric disorders are influenced by age, such as the age of onset (100), the duration of the untreated illness (63,64) and the accelerated ageing of the brain in disease progression (65). GMV also shows distinct sexual dimorphisms (66–69), while most mental disorders display sexual differences with respect to the severity, symptomatology, and outcome of the disease (70). There is also some evidence that stress response pathways may be sensitive to sexual differences (47,71,72). However,

the vast majority of studies to date have not examined the effects of age or sex in the context of CT and neuroimaging (53,54,56–60,73). Although CT is a well-established transdiagnostic risk factor, most CT studies have restricted their analyses to on one or two diagnoses, separately investigating populations of major depressive disorder, post-traumatic stress disorder, or schizophrenia (5,19,37,74). Thus, a continued struggle to disentangle the effects of disease from CT-related effects is evident in most studies, which is further exacerbated by more chronically ill and often medicated study cohorts (57,75). Furthermore, most neuroimaging studies still investigate CT using voxel-wise mass-univariate analysis, which assumes highly localized functional specialization and statistical independence of voxels (76). This approach does not reflect the state-of-the-art understanding of neuroanatomical signatures encoded along distributed, interrelated and interdependent clusters of voxels, cortical regions, and brain systems (61,77,78). Considering the clinical complexity of CT and the heterogeneity of neurobiological evidence, this appears to be specifically the case for CT, whose neuroanatomical representations are most likely subtle, distributed, and multifaceted (79).

Therefore, in Papers I and II we investigated the relationship between structural brain data and CT 1) using the large multi-center PRONIA cohort, comprising young, minimally medicated individuals, 2) following a comprehensive, transdiagnostic approach, including important cofactors such as age, sex and diagnoses (80–83), and 3) employing advanced machine learning (ML), which are better suited to capture the complexity of CT and potentially associated structural brain surrogates (84).

2.2   Introduction to Machine Learning

2.2.1   General Principles

In contrast to other medical specialties, psychiatry faces the problem that the diagnostic catalog and psychiatric thinking itself are based on normative definitions, some of which were defined several decades ago and have changed only marginally since then (85). This is the starting point for modern ML methods, which can integrate different qualities of information and high-dimensional data sets to reveal new clinical and neurobiological insights and break up old structures. In ML, a cohort is called a sample, an individual is called a case, variables are called features, and the target that the algorithm is supposed to predict is called the label (86). ML algorithms can be separated into supervised techniques where the cases are labeled (e.g., diagnostic groups), unsupervised techniques where the goal is to divide an unlabeled sample into groups of related cases, and semi-

supervised techniques, containing labeled and unlabeled cases (86,87). Semi-supervised learning requires advanced knowledge of ML and is beyond the scope of this thesis (88). The goal of ML studies is to train an algorithm so that the resulting model is as good as possible at performing a certain task. Assuming that a binary classification is performed, where the positive label (1) reflects a disease case and the negative label (0) an unaffected healthy case, sensitivity defines the proportion of affected cases with a true positive test result with reference to all affected cases. Specificity defines the proportion of nonaffected cases with a true negative test result in reference to all nonaffected cases. However, the most important ML metric is balanced accuracy (BAC) (86,89). BAC is the mean of sensitivity and specificity and reflects the accuracy in terms of true positive and negative cases according to the sample size of the positive and negative groups. BAC is a robust measure, as it accurately reflects the performance of an algorithm in balanced and unbalanced label distributions (86).

The optimization of a learning algorithm usually takes place within k-fold cross-validation (CV) structures (90). These structures serve two purposes: 1) to maximize the number of cases the algorithm learns from, and 2) to assess the generalizability of the resulting models by applying them to previously unseen cases. Training of the algorithm, i.e., optimization of its hyperparameters, is conducted on the inner folds of the CV structure, the CV1 level. Testing of the generalizability of the model takes place in the outer folds of the CV structure, the CV2 level (see Paper I Supplementary Methods – Machine Learning Framework). In k-fold CV, the sample is divided into k subsets of individuals called folds on the CV1 and CV2 level. An entire CV2 fold is left out, while the algorithm learns on the CV1 training folds. The resulting model is then tested on the left-out CV2 fold. This process is repeated k times until all CV2 folds have been held out at least once for testing. This results in more stable estimates of generalizability, since training groups are more variable and more individuals are used in the CV2 test folds (86,90,91). The most common CV structures are (repeated) nested CV, leave-one-site-out CV, and out-of-sample CV. In nested CV, the cases are assigned to the folds based on a certain stratification, e.g., diagnosis, sex, or sites, so that each fold is representative of the entire sample, thus avoiding batch effects (92). Nested CV ensures that all cases are drawn into the folds without replacement, so that there is no double assignment of cases in the CV2 folds. This procedure can be further augmented when, after all iterations have finished, the cases are re-allocated to the folds via a new distribution, and then the training process is repeated. This approach is called repeated nested CV and is among the

most robust CV setups (93). If the folds are created based on site distribution, then this procedure is called leave-one-site-out CV, as in each iteration, an entire site is left out for testing, while the algorithm is trained on the remaining sites (94,95). In out-of-sample CV, also called external validation, the final model is applied to a new sample from a different cohort, which assesses generalizability even better (96).

2.2.2   Supervised Machine Learning

In supervised learning, classification models predict a discrete outcome (e.g., healthy group vs. patient group) (97), while regression models predict a continuous outcome (e.g., disease severity) (98). It is beyond the scope of this thesis to cover all supervised learning algorithms, however, L2-regularized logistic regression is a simple and frequently used classification method (99–101) and the main statistical tool of Paper II (9,102). In logistic regression, the labels are categorical or binary. This is accomplished by introducing the step function, fitting a smoothed sigmoid curve to the data, which does not incrementally in- or decrease depending on the features but outputs the value 1 above the decision boundary and 0 below the decision boundary (103). Since the step function does not differentiate between more extreme observations, it is robust against outliers. Logistic regression predicts labels using a linear combination of regression parameters and features. Therefore, logistic regression is an 'open book' algorithm with a high degree of transparency and interpretability.

Multicollinearity among features is a challenge in logistic regression. When faced with highly correlated features, optimization of the regression parameters becomes unstable since multiple linear combinations of features can predict the label. Thus, the algorithm is no longer able to converge on a unique linear combination of weighted features. The resulting model suffers from two main issues: 1) the regression parameters of the collinear features tend to be very high in absolute value and 2) these regression parameters tend to show a very high variance (103,104). While issue 1) leads to overfitting, that is, fitting the model too tightly to the dataset, thereby sacrificing generalizability, issue 2) leads to unstable models. To combat this issue, the L2-regularization, or Tikhonov regularization, is employed in the regression model (104,105). When L2-regularization is introduced into the model, the optimization function is updated by adding a shrinkage penalty. This penalty ensures that features that are not relevant to the prediction will become very low, while relevant features will retain moderate to high values, so that their predictive power is indicated. Hence, L2-regularization 1) increases generalizability, and 2) ensures stable

models, even when faced with multicollinearity (104,106). In summary, L2-regularized logistic regression produces robust results in an easy-to-understand and computationally efficient manner, even when faced with high-dimensional and multicollinear features, which is frequent in neuroscience (107,108).

2.2.3   Unsupervised Machine Learning

Unsupervised learning techniques are used to analyze and group unlabeled data sets. These algorithms discover hidden patterns or data groupings without or with minimal human intervention (86). This ability makes them very suitable for exploratory data analysis, segmentation of (sub)populations or patient cohorts, as well as image recognition (109). Among the most used unsupervised learning techniques is clustering. Clustering algorithms group unlabeled data based on their similarities or differences (110). Another important aspect of unsupervised learning is dimensionality reduction, which is used to gain a better understanding of the most relevant or principal features or components of high-dimensional and complex datasets by decomposing them into lower-dimensional components, while preserving the integrity of the dataset (111). A prominent example of such a dimensionality reduction technique is principal component analysis (PCA). PCA reduces redundancies and compresses data sets through feature extraction, thus leading to a set of 'principal components', which maximize the explained variance (111). Linked to PCA is an approach called singular value decomposition, which is the starting point of an unsupervised learning algorithm called Partial Least Squares (PLS)(112). PLS and its sparse version are at the center of my Ph.D. and feature prominently in Paper I.

2.3   Sparse Partial Least Squares Toolbox

2.3.1   Background: Sparse Partial Least Squares

PLS was developed in 1982 by Herman Wold (1908 – 1992) primarily for econometric purposes (113), while his son Svante Wold (1941 – 2022) adapted it for processes in the chemical industry (114). Since then, PLS has been used to investigate neuroanatomical correlates of observable clinical phenotypes (115). The first major application of PLS was in PET data and face recognition tasks (116). More recently, it has been used to assess the relationship between inflammation and treatment resistance in major depressive disorder (117) and clinical-anatomical dimensions of schizophrenia (118).

In general, PLS takes two different data matrices from a cohort (e.g., clinical and neuroimaging data) and then tries to extract as many multivariate associative effects as possible from these two matrices. These associative effects are called latent variables (LV); and each LV contains a pair of weight vectors, which place weights on every feature in the respective matrix, so that covariance between the two matrices is maximized. Hence, PLS examines which features in the matrices are covarying with each other, as well as the direction and the strength of that covariation. After computing such a weight vector pair, the effect explained by these vectors is removed from the data matrices via matrix deflation, before PLS tries to find the next associative effect (119). This allows the detection of multiple layers of associative effects within a dataset, possibly unmasking otherwise hidden associations (119). In a second step, the weight vectors of each LV can be projected back onto their respective data matrix, resulting in two new vectors, called latent scores. These latent scores represent the loading of each individual onto the respective LV. Correlating the latent scores with each other allows for the computation of their correlation coefficient, as well as R2, reflecting the explained covariance of this LV (115,119). Those latent scores can also be used for post hoc analyses (119).

One challenge of PLS applications can arise in high-dimensional datasets. In such situations, the interpretability of the resulting PLS models can be difficult, as PLS outputs pairs of non-sparse weight vectors (119). Considering that, for example, a vectorized 3 mm resliced T1-weighted MR image can have anywhere between 40,000 and 60,000 features, the weight vector can be challenging to interpret. To overcome this issue and to remove noisy features, a sparse variant of PLS, called sparse PLS (SPLS), has emerged (119,120). SPLS selects a subset of only the most key features to be used in the model. The number of features that are included in the model, i.e., the sparsity of the model, is controlled by a pair of hyperparameters (one for each weight vector) (119). Thus, SPLS combines the power of PLS to deconstruct complex multimodal data sets while providing a built-in feature selection that increases clarity and interpretability. The details of the algorithm can be found in the main text of Paper I (Methods and Materials - SPLS Analysis & Assessment of Generalizability and Significance) and its Supplement (Supplementary Methods - Sparse partial least squares algorithm & Machine learning framework).

Regarding the main ML programming languages (MATLAB, R, Python), there is no (free or proprietary) toolbox available that performs SPLS analysis. Therefore, a substantial part of my Ph.D. project was dedicated to creating an open source, user-friendly, and state-of-the-art SPLS toolbox in MATLAB. From its inception in October 2017 and its

first deployment in March 2020 to its current version, the SPLS toolbox has undergone many changes. The latest version is available free of charge at:
https://github.molgen.mpg.de/DavidPopovic/SPLS_Toolbox_2022/tree/main/SPLS_Toolbox_Dev_2022.

2.3.2 Software architecture

The SPLS Toolbox is coded in MATLAB 2020b (The MathWorks, Inc., Natick, MA, USA) and consists of the master module running the main analysis pipeline (dp_spls_standalone) as well as the three slave modules feeding into the master module: hyperopt (hyperparameter optimization), permutation (permutation testing of the LV) and bootstrapping (bootstrapping the features of the LV). All modules are compiled using the built-in compiling application of MATLAB to run on the Sun Grid Engine batch queueing system (Oracle Corporation, Santa Clara, California, USA). The four modules are deployed using adaptive Bash scripts (Bourne-again shell, General Public License, https://www.gnu.org/software/bash/). The main module dp_spls_standalone executes overarching computation steps, deploys the slave modules, and integrates their results. The slave modules are computed as job arrays, so that their high-computation tasks are divided into smaller parallel jobs, usually between 40 and 60. The SPLS Toolbox consists of 85 functions, of which 44 were coded by me, 33 functions were taken from MATLAB built-in functions or third-party sources, and 8 were adapted from the NeuroMiner Software of Nikolaos Koutsouleris (http://proniapredictors.eu/neurominer/index.html).

2.3.3 Machine Learning Framework

The ML framework of the SPLS Toolbox includes fully cross-validated pre-processing, hyperparameter optimization, permutation testing, and bootstrap sampling.

Cross-Validation: The user can choose between (repeated) nested CV, random holdout CV, leave-one-site-out CV, and random split-half CV (86). In addition, the user can define the number of CV1 and CV2 folds and add permutation steps at both levels. Furthermore, the toolbox also allows for the extraction of a separate validation set. Thus, a percentage of individuals (e.g., 10%, 25%) does not enter the CV sequence and is kept only for application of the final model. The stratification of cases into CV folds can be specified based on diagnoses, sites, sex, or any other criterion.

Pre-processing: The pre-processing of the features, including scaling and covariate correction, is carried out separately in the CV1 training and test folds, thus avoiding data

leakage (121). For scaling, mean centering or rescaling between 0 and 1 can be chosen. The user can choose whether to remove covariate effects from none, one, or both data matrices. The user can also define whether to correct across all individuals in the sample or to correct based on the covariate effect in a certain subpopulation. If the latter is chosen, the covariate effects are calculated in a chosen subpopulation and then removed from the entire sample, including the subpopulation. For example, the effect of ageing on MR images can be calculated only within the healthy control population, and, afterwards, this 'healthy' ageing effect is removed from the entire sample. This avoids deleting effects that might be collinear with age, but which shall be retained after age correction. The correction is based on univariate partial correlation analysis (122,123).

Hyperparameter Optimization: This module performs a grid search of all possible hyperparameter combinations. The lower and upper limits of the grid can be adapted to search for particularly dense or sparse solutions; however, it is recommended to use the entire grid by default. The number of data points within this grid can be defined by the user; most commonly, 20 or 40 data points are chosen for each hyperparameter, leading to 400 or 1600 hyperparameter combinations, respectively.

Permutation Testing: The performance of the optimal model on the CV2 level is compared against a certain number of models, retrained with the optimal hyperparameter combinations on distorted, permuted datasets. The number of permutations can be chosen by the user; the default setting is 5000. In the next step, the P value of the final model can be calculated in two ways: 1) counting, 2) area under the curve (AUC). The counting method counts how often the models trained on the permuted CV1 datasets performed better than the optimal model, divided by the number of permutations. The AUC method creates a null distribution of the performances of the models trained on the permuted CV1 datasets and then computes the AUC for the optimal model.

LV Selection: The P-values of all CV2 models are adjusted for multiple tests. The options for this are as follows: Bonferroni, Sidak, Holm-Bonferroni, Benjamini-Hochberg, Benjamini-Yekutieli, Storey, Fisher (124). The model with the lowest adjusted P value is chosen as the winning model and becomes the LV of this iteration. If the adjusted P-value of this model is below the significance threshold, then, according to the omnibus hypothesis, the LV is significant (119,125).

Bootstrapping: Bootstrap samples are drawn from the pooled CV1 folds and used to retrain the model a certain number of times using the optimal hyperparameter combination (115). Bootstrap ratios and confidence intervals are then calculated for each

feature weight, giving the user two different measures of feature weight stability. These can be used to prune the model post hoc. The default setting is 500 bootstrap samples; however, the user is free to choose any number of bootstrap samples.

Deflation/Iteration: If the LV is significant, its effect is removed from the matrices by projection deflation. The user can set the limit of how many nonsignificant LV are allowed. The default setting is one, meaning that as soon as one nonsignificant LV is detected, the SPLS analysis ends. If an LV is not significant, it represents a random effect. Continuing the analysis would mean deflating the matrices by this random effect and therefore arbitrarily changing them and compromising all the following LV. However, another perspective is that these non-significant LV are not 'random' effects, but noise within the data. Deflating the data matrices from these nonsignificant LV would be like removing noise from the dataset. Therefore, this kind of noise reduction could enable the detection of further significant LV. Thus, the stop criterion can be chosen individually.

2.3.4   Model Visualization

The final model with all its significant LV can be visualized in a visualization module. The module is a noncompiled MATLAB function that contains 190 other functions. Among these, 56 functions were taken from the SPM12 toolbox (https://www.fil.ion.ucl.ac.uk/spm/software/spm12/), 18 functions were coded by me and 13 stemmed from the NeuroMiner Software. The rest of the functions came from third-party sources. Phenotypic vectors or other types of vector data can be automatically exported as bar plots or sunburst charts. The coloring is adaptive and can be changed according to the MATLAB color maps or external color palettes. 3D brain images are reconstructed from brain vectors and written as NIFTI files (126). These files can then be imported into MRI visualization tools such as the Connectome Workbench (https://humanconnectome.org/software/connectome-workbench). The module can also overlay brain vectors onto atlases, calculate the percentage of positive and negative weights assigned to each individual brain region or brain network, and visualize them as spider plots. Currently, the Brainnetome (127) and Diedrichsen atlas (128) are used for parcellation of GMV regions in the cerebrum and cerebellum, respectively. The Yeo (129) and Buckner atlas (130) are used for parcellation of functional brain networks in the cerebrum and cerebellum, respectively.

3. Paper I: Traces of Trauma: A Multivariate Pattern Analysis of Childhood Trauma, Brain Structure, and Clinical Phenotypes

4. Paper II: The non-specific nature of mental health and structural brain outcomes following childhood trauma

References

1. Pearce J, Murray C, Larkin W. Childhood adversity and trauma: experiences of professionals trained to routinely enquire about childhood adversity. Heliyon [Internet]. 2019 Jul;5(7):e01900. Available from: http://dx.doi.org/10.1016/j.heliyon.2019.e01900

2. Sherfinski HT, Condit PE, Williams Al-Kharusy SS, Moreno MA. Adverse childhood experiences: Perceptions, practices, and possibilities. WMJ [Internet]. 2021 Oct;120(3):209–17. Available from: https://www.ncbi.nlm.nih.gov/pubmed/34710303

3. Hughes K, Bellis MA, Hardcastle KA, Sethi D, Butchart A, Mikton C, et al. The effect of multiple adverse childhood experiences on health: a systematic review and meta-analysis. Lancet Public Health [Internet]. 2017 Aug;2(8):e356–66. Available from: http://dx.doi.org/10.1016/S2468-2667(17)30118-4

4. Gilbert R, Widom CS, Browne K, Fergusson D, Webb E, Janson S. Burden and consequences of child maltreatment in high-income countries. Lancet [Internet]. 2009 Jan;373(9657):68–81. Available from: http://dx.doi.org/10.1016/s0140-6736(08)61706-7

5. Popovic D, Schmitt A, Kaurani L, Senner F, Papiol S, Malchow B, et al. Childhood Trauma in Schizophrenia: Current Findings and Research Perspectives. Front Neurosci [Internet]. 2019 Mar 21;13:274. Available from: http://dx.doi.org/10.3389/fnins.2019.00274

6. Bernstein DP, Ahluvalia T, Pogge D, Handelsman L. Validity of the Childhood Trauma Questionnaire in an adolescent psychiatric population. J Am Acad Child Adolesc Psychiatry [Internet]. 1997;36(3):340–8. Available from: https://www.ncbi.nlm.nih.gov/pubmed/9055514

7. Viola TW, Salum GA, Kluwe-Schiavon B, Sanvicente-Vieira B, Levandowski ML, Grassi-Oliveira R. The influence of geographical and economic factors in estimates of childhood abuse and neglect using the Childhood Trauma Questionnaire: A worldwide meta-regression analysis. Child Abuse Negl [Internet]. 2016;51:1–11. Available from: https://www.ncbi.nlm.nih.gov/pubmed/26704298

8. Popovic D, Ruef A, Dwyer DB, Antonucci LA, Eder J, Sanfelici R, et al. Traces of trauma: A multivariate pattern analysis of childhood trauma, brain structure, and clinical phenotypes. Biol Psychiatry [Internet]. 2020 Dec 1;88(11):829–42. Available from: http://dx.doi.org/10.1016/j.biopsych.2020.05.020

9. Haidl TK, Hedderich DM, Rosen M, Kaiser N, Seves M, Lichtenstein T, et al. The non-specific nature of mental health and structural brain outcomes following childhood trauma. Psychol Med [Internet]. 2021 Jul 6;1–10. Available from: http://dx.doi.org/10.1017/S0033291721002439

10. Bremner JD, Bolus R, Mayer EA. Psychometric properties of the Early Trauma Inventory-Self Report. J Nerv Ment Dis [Internet]. 2007;195(3):211–8. Available from: https://www.ncbi.nlm.nih.gov/pubmed/17468680

11. Plaza A, Torres A, Martin-Santos R, Gelabert E, Imaz ML, Navarro P, et al. Validation and test-retest reliability of Early Trauma Inventory in Spanish postpartum women. J Nerv Ment Dis [Internet]. 2011;199(4):280–5. Available from: https://www.ncbi.nlm.nih.gov/pubmed/21451355

12. Koss MP, Gidycz CA, Wisniewski N. The scope of rape: incidence and prevalence of sexual aggression and victimization in a national sample of higher education students. J Consult Clin Psychol [Internet]. 1987;55(2):162–70. Available from: https://www.ncbi.nlm.nih.gov/pubmed/3494755

13. Yampolsky L, Lev-Wiesel R, Ben-Zion IZ. Child sexual abuse: is it a risk factor for pregnancy? J Adv Nurs [Internet]. 2010;66(9):2025–37. Available from: https://www.ncbi.nlm.nih.gov/pubmed/20636469

14. Kubany ES, Haynes SN, Leisen MB, Owens JA, Kaplan AS, Watson SB, et al. Development and preliminary validation of a brief broad-spectrum measure of trauma exposure: the Traumatic Life Events Questionnaire. Psychol Assess [Internet]. 2000;12(2):210–24. Available from: https://www.ncbi.nlm.nih.gov/pubmed/10887767

15. Dang J, King KM, Inzlicht M. Why Are Self-Report and Behavioral Measures Weakly Correlated? Trends Cogn Sci [Internet]. 2020 Apr;24(4):267–9. Available from: http://dx.doi.org/10.1016/j.tics.2020.01.007

16. Bremner JD, Vermetten E, Mazure CM. Development and preliminary psychometric properties of an instrument for the measurement of childhood trauma: the Early Trauma Inventory. Depress Anxiety [Internet]. 2000;12(1):1–12. Available from: https://www.ncbi.nlm.nih.gov/pubmed/10999240

17. McKay MT, Cannon M, Chambers D, Conroy RM, Coughlan H, Dodd P, et al. Childhood trauma and adult mental disorder: A systematic review and meta-analysis of longitudinal cohort studies. Acta Psychiatr Scand [Internet]. 2021 Mar;143(3):189–205. Available from: http://dx.doi.org/10.1111/acps.13268

18. Petruccelli K, Davis J, Berman T. Adverse childhood experiences and associated health outcomes: A systematic review and meta-analysis. Child Abuse Negl [Internet]. 2019 Nov;97:104127. Available from: http://dx.doi.org/10.1016/j.chiabu.2019.104127

19. Stanton KJ, Denietolis B, Goodwin BJ, Dvir Y. Childhood Trauma and Psychosis: An Updated Review. Child Adolesc Psychiatr Clin N Am [Internet]. 2020 Jan;29(1):115–29. Available from: http://dx.doi.org/10.1016/j.chc.2019.08.004

20. Jaworska-Andryszewska P, Rybakowski JK. Childhood trauma in mood disorders: Neurobiological mechanisms and implications for treatment. Pharmacol Rep [Internet]. 2019 Feb;71(1):112–20. Available from: http://dx.doi.org/10.1016/j.pharep.2018.10.004

21. Peyrot WJ, Van der Auwera S, Milaneschi Y, Dolan CV, Madden PAF, Sullivan PF, et al. Does Childhood Trauma Moderate Polygenic Risk for Depression? A Meta-analysis of 5765 Subjects From the Psychiatric Genomics Consortium. Biol Psychiatry [Internet]. 2018 Jul 15;84(2):138–47. Available from: http://dx.doi.org/10.1016/j.biopsych.2017.09.009

22. Zhang S, Lin X, Liu J, Pan Y, Zeng X, Chen F, et al. Prevalence of childhood trauma measured by the short form of the Childhood Trauma Questionnaire in people with substance use disorder: A meta-analysis. Psychiatry Res [Internet]. 2020 Dec;294:113524. Available from: http://dx.doi.org/10.1016/j.psychres.2020.113524

23. Grummitt L, Kelly E, Barrett E, Keyes K, Newton N. Targets for intervention to prevent substance use in young people exposed to childhood adversity: A systematic review. PLoS One [Internet]. 2021 Jun 7;16(6):e0252815. Available from: http://dx.doi.org/10.1371/journal.pone.0252815

24. Leza L, Siria S, López-Goñi JJ, Fernández-Montalvo J. Adverse childhood experiences (ACEs) and substance use disorder (SUD): A scoping review. Drug Alcohol Depend [Internet]. 2021 Apr 1;221(108563):108563. Available from: http://dx.doi.org/10.1016/j.drugalcdep.2021.108563

25. Edalati H, Krank MD. Childhood maltreatment and development of substance use disorders: A review and a model of cognitive pathways. Trauma Violence Abuse [Internet]. 2016 Dec;17(5):454–67. Available from: http://dx.doi.org/10.1177/1524838015584370

26. Bozzatello P, Rocca P, Baldassarri L, Bosia M, Bellino S. The Role of Trauma in Early Onset Borderline Personality Disorder: A Biopsychosocial Perspective. Front Psychiatry [Internet]. 2021 Sep 23;12:721361. Available from: http://dx.doi.org/10.3389/fpsyt.2021.721361

27. Porter C, Palmier-Claus J, Branitsky A, Mansell W, Warwick H, Varese F. Childhood adversity and borderline personality disorder: a meta-analysis. Acta Psychiatr Scand [Internet]. 2020 Jan;141(1):6–20. Available from: http://dx.doi.org/10.1111/acps.13118

28. Kleindienst N, Vonderlin R, Bohus M, Lis S. Childhood adversity and borderline personality disorder. Analyses complementing the meta-analysis by Porter et al. (2020). Acta Psychiatr Scand [Internet]. Wiley; 2021 Feb;143(2):183–4. Available from: http://dx.doi.org/10.1111/acps.13256

29. Yehuda R, Hoge CW, McFarlane AC, Vermetten E, Lanius RA, Nievergelt CM, et al. Post-traumatic stress disorder. Nat Rev Dis Primers [Internet]. 2015 Oct 8;1:15057. Available from: http://dx.doi.org/10.1038/nrdp.2015.57

30. Daniels JK, Lamke J-P, Gaebler M, Walter H, Scheel M. White matter integrity and its relationship to PTSD and childhood trauma--a systematic review and meta-analysis. Depress Anxiety [Internet]. 2013 Mar;30(3):207–16. Available from: http://dx.doi.org/10.1002/da.22044

31. Juruena MF, Eror F, Cleare AJ, Young AH. The role of early life stress in HPA axis and anxiety. Adv Exp Med Biol [Internet]. 2020;1191:141–53. Available from: http://dx.doi.org/10.1007/978-981-32-9705-0_9

32. Destrée L, Brierley M-EE, Albertella L, Jobson L, Fontenelle LF. The effect of childhood trauma on the severity of obsessive-compulsive symptoms: A systematic review. J Psychiatr Res [Internet]. 2021 Oct;142:345–60. Available from: http://dx.doi.org/10.1016/j.jpsychires.2021.08.017

33. Bellis MA, Hughes K, Ford K, Ramos Rodriguez G, Sethi D, Passmore J. Life course health consequences and associated annual costs of adverse childhood experiences across Europe and North America: a systematic review and meta-analysis. Lancet Public Health [Internet]. 2019 Oct;4(10):e517–28. Available from: http://dx.doi.org/10.1016/S2468-2667(19)30145-8

34. Hughes K, Ford K, Bellis MA, Glendinning F, Harrison E, Passmore J. Health and financial costs of adverse childhood experiences in 28 European countries: a systematic review and meta-analysis. Lancet Public Health [Internet]. 2021 Nov;6(11):e848–57. Available from: http://dx.doi.org/10.1016/S2468-2667(21)00232-2

35. GBD 2019 Viewpoint Collaborators. Five insights from the Global Burden of Disease Study 2019. Lancet [Internet]. 2020 Oct 17;396(10258):1135–59. Available from: http://dx.doi.org/10.1016/S0140-6736(20)31404-5

36. Koutsouleris N, Kambeitz-Ilankovic L, Ruhrmann S, Rosen M, Ruef A, Dwyer DB, et al. Prediction Models of Functional Outcomes for Individuals in the Clinical High-Risk State for Psychosis or With Recent-Onset Depression: A Multimodal, Multisite Machine Learning Analysis. JAMA Psychiatry [Internet]. 2018 Nov 1;75(11):1156–72. Available from: http://dx.doi.org/10.1001/jamapsychiatry.2018.2165

37. Hakamata Y, Suzuki Y, Kobashikawa H, Hori H. Neurobiology of early life adversity: A systematic review of meta-analyses towards an integrative account of its neurobiological trajectories to mental disorders. Front Neuroendocrinol [Internet]. 2022 Apr;65:100994. Available from: http://dx.doi.org/10.1016/j.yfrne.2022.100994

38. LeDoux JE. Emotion circuits in the brain. Annu Rev Neurosci [Internet]. 2000;23(1):155–84. Available from: http://dx.doi.org/10.1146/annurev.neuro.23.1.155

39. Tamietto M, de Gelder B. Neural bases of the non-conscious perception of emotional signals. Nat Rev Neurosci [Internet]. 2010 Oct;11(10):697–709. Available from: http://dx.doi.org/10.1038/nrn2889

40. Gunnar MR, Frenn K, Wewerka SS, Van Ryzin MJ. Moderate versus severe early life stress: associations with stress reactivity and regulation in 10-12-year-old children. Psychoneuroendocrinology [Internet]. 2009 Jan;34(1):62–75. Available from: http://dx.doi.org/10.1016/j.psyneuen.2008.08.013

41. McEwen BS. Stress and the individual. Arch Intern Med [Internet]. 1993 Sep 27;153(18):2093. Available from: http://dx.doi.org/10.1001/archinte.1993.00410180039004

42. Gray JD, Kogan JF, Marrocco J, McEwen BS. Genomic and epigenomic mechanisms of glucocorticoids in the brain. Nat Rev Endocrinol [Internet]. 2017 Nov;13(11):661–73. Available from: http://dx.doi.org/10.1038/nrendo.2017.97

43. McEwen BS, Nasca C, Gray JD. Stress effects on neuronal structure: Hippocampus, amygdala, and prefrontal cortex. Neuropsychopharmacology [Internet]. 2016 Jan;41(1):3–23. Available from: http://dx.doi.org/10.1038/npp.2015.171

44. Giotakos O. Neurobiology of emotional trauma. Psychiatrike [Internet]. 2020 Apr;31(2):162–71. Available from: http://dx.doi.org/10.22365/jpsych.2020.312.162

45. Craig ADB. How do you feel--now? The anterior insula and human awareness. Nat Rev Neurosci [Internet]. 2009 Jan;10(1):59–70. Available from: http://dx.doi.org/10.1038/nrn2555

46. Etkin A, Egner T, Kalisch R. Emotional processing in anterior cingulate and medial prefrontal cortex. Trends Cogn Sci [Internet]. 2011 Feb;15(2):85–93. Available from: http://dx.doi.org/10.1016/j.tics.2010.11.004

47. Mothersill O, Donohoe G. Neural effects of social environmental stress – an activation likelihood estimation meta-analysis. Psychol Med [Internet]. 2016 Jul;46(10):2015–23. Available from: http://dx.doi.org/10.1017/s0033291716000477

48. Heany SJ, Groenewold NA, Uhlmann A, Dalvie S, Stein DJ, Brooks SJ. The neural correlates of Childhood Trauma Questionnaire scores in adults: A meta-analysis and review of functional magnetic resonance imaging studies. Dev Psychopathol [Internet]. 2018 Oct;30(4):1475–85. Available from: http://dx.doi.org/10.1017/s0954579417001717

49. Kraaijenvanger EJ, Pollok TM, Monninger M, Kaiser A, Brandeis D, Banaschewski T, et al. Impact of early life adversities on human brain functioning: A coordinate-based meta-analysis. Neurosci Biobehav Rev [Internet]. 2020 Jun;113:62–76. Available from: http://dx.doi.org/10.1016/j.neubiorev.2020.03.008

50. Benear SL, Ngo CT, Olson IR. Dissecting the fornix in basic memory processes and neuropsychiatric disease: A review. Brain Connect [Internet]. 2020 Sep;10(7):331–54. Available from: http://dx.doi.org/10.1089/brain.2020.0749

51. Lim L, Howells H, Radua J, Rubia K. Aberrant structural connectivity in childhood maltreatment: A meta-analysis. Neurosci Biobehav Rev [Internet]. 2020 Sep;116:406–14. Available from: http://dx.doi.org/10.1016/j.neubiorev.2020.07.004

52. van Velzen LS, Kelly S, Isaev D, Aleman A, Aftanas LI, Bauer J, et al. White matter disturbances in major depressive disorder: a coordinated analysis across 20 international cohorts in the ENIGMA MDD working group. Mol Psychiatry

[Internet]. 2020 Jul;25(7):1511–25. Available from: http://dx.doi.org/10.1038/s41380-019-0477-2

53. Ahmed-Leitao F, Spies G, van den Heuvel L, Seedat S. Hippocampal and amygdala volumes in adults with posttraumatic stress disorder secondary to childhood abuse or maltreatment: A systematic review. Psychiatry Res Neuroimaging [Internet]. 2016 Oct;256:33–43. Available from: http://dx.doi.org/10.1016/j.pscychresns.2016.09.008

54. Paquola C, Bennett MR, Lagopoulos J. Understanding heterogeneity in grey matter research of adults with childhood maltreatment-A meta-analysis and review. Neurosci Biobehav Rev [Internet]. 2016 Oct;69:299–312. Available from: http://dx.doi.org/10.1016/j.neubiorev.2016.08.011

55. Woon FL, Hedges DW. Hippocampal and amygdala volumes in children and adults with childhood maltreatment-related posttraumatic stress disorder: a meta-analysis. Hippocampus [Internet]. 2008;18(8):729–36. Available from: http://dx.doi.org/10.1002/hipo.20437

56. Logue MW, van Rooij SJH, Dennis EL, Davis SL, Hayes JP, Stevens JS, et al. Smaller hippocampal volume in posttraumatic stress disorder: A multisite ENIGMA-PGC study: Subcortical volumetry results from posttraumatic stress disorder consortia. Biol Psychiatry [Internet]. 2018 Feb;83(3):244–53. Available from: http://dx.doi.org/10.1016/j.biopsych.2017.09.006

57. Calem M, Bromis K, McGuire P, Morgan C, Kempton MJ. Meta-analysis of associations between childhood adversity and hippocampus and amygdala volume in non-clinical and general population samples. NeuroImage Clin [Internet]. 2017;14:471–9. Available from: http://dx.doi.org/10.1016/j.nicl.2017.02.016

58. Lim L, Radua J, Rubia K. Gray matter abnormalities in childhood maltreatment: a voxel-wise meta-analysis. Am J Psychiatry [Internet]. 2014 Aug;171(8):854–63. Available from: http://dx.doi.org/10.1176/appi.ajp.2014.13101427

59. Riem MME, Alink LRA, Out D, Van Ijzendoorn MH, Bakermans-Kranenburg MJ. Beating the brain about abuse: Empirical and meta-analytic studies of the association between maltreatment and hippocampal volume across childhood and adolescence. Dev Psychopathol [Internet]. 2015 May;27(2):507–20. Available from: http://dx.doi.org/10.1017/S0954579415000127

60. Woon FL, Sood S, Hedges DW. Hippocampal volume deficits associated with exposure to psychological trauma and posttraumatic stress disorder in adults: A meta-analysis. Progress in Neuro-Psychopharmacology and Biological Psychiatry [Internet]. 2010 Oct;34(7):1181–8. Available from: http://dx.doi.org/10.1016/j.pnpbp.2010.06.016

61. Woo C-W, Chang LJ, Lindquist MA, Wager TD. Building better biomarkers: brain models in translational neuroimaging. Nat Neurosci [Internet]. 2017;20(3):365. Available from: https://www.nature.com/articles/nn.4478.pdf?origin=ppub

62. Irimia A. Cross-sectional volumes and trajectories of the human brain, gray matter, white matter and cerebrospinal fluid in 9473 typically aging adults. Neuroinformatics [Internet]. 2021 Apr;19(2):347–66. Available from: http://dx.doi.org/10.1007/s12021-020-09480-w

63. Ghio L, Gotelli S, Marcenaro M, Amore M, Natta W. Duration of untreated illness and outcomes in unipolar depression: a systematic review and meta-analysis. J Affect Disord [Internet]. 2014 Jan;152–154:45–51. Available from: http://dx.doi.org/10.1016/j.jad.2013.10.002

64. Murru A, Carpiniello B. Duration of untreated illness as a key to early intervention in schizophrenia: A review. Neurosci Lett [Internet]. 2018 Mar 16;669:59–67. Available from: http://dx.doi.org/10.1016/j.neulet.2016.10.003

65. Franke K, Gaser C. Ten years of BrainAGE as a neuroimaging biomarker of brain aging: What insights have we gained? Front Neurol [Internet]. 2019 Aug 14;10:789. Available from: http://dx.doi.org/10.3389/fneur.2019.00789

66. Kaczkurkin AN, Raznahan A, Satterthwaite TD. Sex differences in the developing brain: insights from multimodal neuroimaging. Neuropsychopharmacology [Internet]. 2019 Jan;44(1):71–85. Available from: http://dx.doi.org/10.1038/s41386-018-0111-z

67. Sacher J, Neumann J, Okon-Singer H, Gotowiec S, Villringer A. Sexual dimorphism in the human brain: evidence from neuroimaging. Magn Reson Imaging [Internet]. 2013 Apr;31(3):366–75. Available from: http://dx.doi.org/10.1016/j.mri.2012.06.007

68. Ruigrok ANV, Salimi-Khorshidi G, Lai M-C, Baron-Cohen S, Lombardo MV, Tait RJ, et al. A meta-analysis of sex differences in human brain structure. Neurosci Biobehav Rev [Internet]. 2014 Feb;39:34–50. Available from: http://dx.doi.org/10.1016/j.neubiorev.2013.12.004

69. Kiesow H, Dunbar R, Kable JW, Kalenscher T, Vogeley K, Schilbach L, et al. 000 Social Brains: Sex Differentiation in Human Brain Anatomy. Science Advances. 2020;10.

70. Singh RS, Singh KK, Singh SM. Origin of sex-biased mental disorders: An evolutionary perspective. J Mol Evol [Internet]. 2021 Jun;89(4–5):195–213. Available from: http://dx.doi.org/10.1007/s00239-021-09999-9

71. Bunea IM, Szentágotai-Tătar A, Miu AC. Early-life adversity and cortisol response to social stress: a meta-analysis. Transl Psychiatry [Internet]. 2017 Dec;7(12). Available from: http://dx.doi.org/10.1038/s41398-017-0032-3

72. Kuhlman KR, Horn SR, Chiang JJ, Bower JE. Early life adversity exposure and circulating markers of inflammation in children and adolescents: A systematic review and meta-analysis. Brain Behav Immun [Internet]. 2020 May;86:30–42. Available from: http://dx.doi.org/10.1016/j.bbi.2019.04.028

73. Tiwari A, Gonzalez A. Biological alterations affecting risk of adult psychopathology following childhood trauma: A review of sex differences. Clin Psychol Rev [Internet]. 2018 Dec;66:69–79. Available from: http://dx.doi.org/10.1016/j.cpr.2018.01.006

74. Tozzi L, Garczarek L, Janowitz D, Stein DJ, Wittfeld K, Dobrowolny H, et al. Interactive impact of childhood maltreatment, depression, and age on cortical brain structure: mega-analytic findings from a large multi-site cohort. Psychol Med [Internet]. 2020 Apr;50(6):1020–31. Available from: http://dx.doi.org/10.1017/S003329171900093X

75. Bromis K, Calem M, Reinders AATS, Williams SCR, Kempton MJ. Meta-analysis of 89 structural MRI studies in posttraumatic stress disorder and comparison with major depressive disorder. Am J Psychiatry [Internet]. 2018 Oct 1;175(10):989–98. Available from: http://dx.doi.org/10.1176/appi.ajp.2018.17111199

76. Logothetis NK. What we can do and what we cannot do with fMRI. Nature [Internet]. 2008;453(7197):869–78. Available from: https://www.ncbi.nlm.nih.gov/pubmed/18548064

77. Kamitani Y, Tong F. Decoding the visual and subjective contents of the human brain. Nat Neurosci [Internet]. 2005;8(5):679–85. Available from: https://www.ncbi.nlm.nih.gov/pubmed/15852014

78. Kriegeskorte N, Cusack R, Bandettini P. How does an fMRI voxel sample the neuronal activity pattern: compact-kernel or complex spatiotemporal filter? Neuroimage [Internet]. 2010;49(3):1965–76. Available from: https://www.ncbi.nlm.nih.gov/pubmed/19800408

79. Davatzikos C. Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. Neuroimage [Internet]. 2004;23(1):17–20. Available from: https://www.ncbi.nlm.nih.gov/pubmed/15325347

80. Yahata N, Kasai K, Kawato M. Computational neuroscience approach to biomarkers and treatments for mental disorders. Psychiatry Clin Neurosci [Internet]. 2017;71(4):215–37. Available from: https://www.ncbi.nlm.nih.gov/pubmed/28032396

81. Freedman R, Lewis DA, Michels R, Pine DS, Schultz SK, Tamminga CA, et al. The initial field trials of DSM-5: new blooms and old thorns. Am J Psychiatry [Internet]. 2013;170(1):1–5. Available from: https://www.ncbi.nlm.nih.gov/pubmed/23288382

82. Khan A, McCormack HC, Bolger EA, McGreenery CE, Vitaliano G, Polcari A, et al. Childhood Maltreatment, Depression, and Suicidal Ideation: Critical Importance of Parental and Peer Emotional Abuse during Developmental Sensitive Periods in Males and Females. Front Psychiatry [Internet]. 2015;6:42. Available from: https://www.ncbi.nlm.nih.gov/pubmed/25870565

83. Whittle S, Simmons JG, Dennison M, Vijayakumar N, Schwartz O, Yap MB, et al. Positive parenting predicts the development of adolescent brain structure: a longitudinal study. Dev Cogn Neurosci [Internet]. 2014;8:7–17. Available from: https://www.ncbi.nlm.nih.gov/pubmed/24269113

84. Jollans L, Whelan R. Neuromarkers for Mental Disorders: Harnessing Population Neuroscience. Front Psychiatry [Internet]. 2018;9:242. Available from: https://www.ncbi.nlm.nih.gov/pubmed/29928237

85. Lilienfeld SO, Treadway MT. Clashing Diagnostic Approaches: DSM-ICD Versus RDoC. Annu Rev Clin Psychol [Internet]. 2016 Feb 3;12:435–63. Available from: http://dx.doi.org/10.1146/annurev-clinpsy-021815-093122

86. Dwyer DB, Falkai P, Koutsouleris N. Machine Learning Approaches for Clinical Psychology and Psychiatry. Annu Rev Clin Psychol [Internet]. 2018 May 7;14:91–118. Available from: http://dx.doi.org/10.1146/annurev-clinpsy-032816-045037

87. Zhou Z, Wu T-C, Wang B, Wang H, Tu XM, Feng C. Machine learning methods in psychiatry: a brief introduction. Gen Psychiatr [Internet]. 2020 Feb 3;33(1):e100171. Available from: http://dx.doi.org/10.1136/gpsych-2019-100171

88. Cheplygina V, de Bruijne M, Pluim JPW. Not-so-supervised: A survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. Med Image Anal [Internet]. 2019 May;54:280–96. Available from: http://dx.doi.org/10.1016/j.media.2019.03.009

89. Bzdok D, Meyer-Lindenberg A. Machine Learning for Precision Psychiatry: Opportunities and Challenges. Biol Psychiatry Cogn Neurosci Neuroimaging [Internet]. 2018 Mar;3(3):223–30. Available from: http://dx.doi.org/10.1016/j.bpsc.2017.11.007

90. Lever J, Krzywinski M, Altman N. Model selection and overfitting. Nat Methods [Internet]. 2016 Aug 30 [cited 2022 Jul 16];13(9):703–4. Available from: https://www.nature.com/articles/nmeth.3968

91. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. New York: Springer; 2009.

92. Soneson C, Gerster S, Delorenzi M. Batch effect confounding leads to strong bias in performance estimates obtained by cross-validation. PLoS One [Internet]. 2014 Jun 26;9(6):e100335. Available from: http://dx.doi.org/10.1371/journal.pone.0100335

93. Ruschhaupt M, Huber W, Poustka A, Mansmann U. A compendium to ensure computational reproducibility in high-dimensional classification tasks. Stat Appl

Genet Mol Biol [Internet]. 2004 Dec 19;3:Article37. Available from: http://dx.doi.org/10.2202/1544-6115.1078

94. Koutsouleris N, Dwyer DB, Degenhardt F, Maj C, Urquijo-Castro MF, Sanfelici R, et al. Multimodal machine learning workflows for prediction of psychosis in patients with clinical high-risk syndromes and recent-onset depression. JAMA Psychiatry [Internet]. 2020 Dec 2; Available from: http://dx.doi.org/10.1001/jamapsychiatry.2020.3604

95. Koutsouleris N, Kahn RS, Chekroud AM, Leucht S, Falkai P, Wobrock T, et al. Multisite prediction of 4-week and 52-week treatment outcomes in patients with first-episode psychosis: a machine learning approach. Lancet Psychiatry [Internet]. 2016 Oct;3(10):935–46. Available from: http://dx.doi.org/10.1016/S2215-0366(16)30171-7

96. Tsamardinos I, Greasidou E, Borboudakis G. Bootstrapping the out-of-sample predictions for efficient and accurate cross-validation. Mach Learn [Internet]. 2018 May 9;107(12):1895–922. Available from: http://dx.doi.org/10.1007/s10994-018-5714-4

97. Zarogianni E, Storkey AJ, Borgwardt S, Smieskova R, Studerus E, Riecher-Rössler A, et al. Individualized prediction of psychosis in subjects with an at-risk mental state. Schizophr Res [Internet]. 2019 Dec;214:18–23. Available from: http://dx.doi.org/10.1016/j.schres.2017.08.061

98. Portugal LCL, Schrouff J, Stiffler R, Bertocci M, Bebko G, Chase H, et al. Predicting anxiety from wholebrain activity patterns to emotional faces in young adults: a machine learning approach. Neuroimage Clin [Internet]. 2019 Apr 3;23:101813. Available from: http://dx.doi.org/10.1016/j.nicl.2019.101813

99. Megjhani M, Terilli K, Weiss M, Savarraj J, Chen LH, Alkhachroum A, et al. Dynamic Detection of Delayed Cerebral Ischemia: A Study in 3 Centers. Stroke [Internet]. 2021 Apr;52(4):1370–9. Available from: http://dx.doi.org/10.1161/STROKEAHA.120.032546

100. Topçuoğlu BD, Lesniak NA, Ruffin MT 4th, Wiens J, Schloss PD. A Framework for Effective Application of Machine Learning to Microbiome-Based Classification Problems. MBio [Internet]. 2020 Jun 9;11(3). Available from: http://dx.doi.org/10.1128/mBio.00434-20

101. Zhang X, Hu B, Ma X, Xu L. Resting-state whole-brain functional connectivity networks for MCI classification using L2-regularized logistic regression. IEEE Trans Nanobioscience [Internet]. 2015 Mar;14(2):237–47. Available from: http://dx.doi.org/10.1109/TNB.2015.2403274

102. Fan R-E, Chang K-W, Hsieh C-J, Wang X-R, Lin C-J. LIBLINEAR: A Library for Large Linear Classification. J Mach Learn Res [Internet]. 2008 Jun 1;9:1871–4. Available from: https://dl.acm.org/doi/10.5555/1390681.1442794

103. Lever J, Krzywinski M, Altman N. Logistic regression. Nat Methods [Internet]. 2016 Jul;13(7):541–2. Available from: http://dx.doi.org/10.1038/nmeth.3904

104. Lever J, Krzywinski M, Altman N. Regularization. Nat Methods [Internet]. 2016 Sep 29 [cited 2022 Jul 18];13(10):803–4. Available from: https://www.nature.com/articles/nmeth.4014

105. Arashi M, Roozbeh M, Hamzah NA, Gasparini M. Ridge regression and its applications in genetic studies. PLoS One [Internet]. 2021 Apr 8;16(4):e0245376. Available from: http://dx.doi.org/10.1371/journal.pone.0245376

106. Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc Series B Stat Methodol [Internet]. 2005 Apr;67(2):301–20. Available from: http://doi.wiley.com/10.1111/j.1467-9868.2005.00503.x

107.   Jollans L, Boyle R, Artiges E, Banaschewski T, Desrivières S, Grigis A, et al. Quantifying performance of machine learning methods for neuroimaging data. Neuroimage [Internet]. 2019 Oct 1;199:351–65. Available from: http://dx.doi.org/10.1016/j.neuroimage.2019.05.082

108.   Iacobucci D, Schneider MJ, Popovich DL, Bakamitsos GA. Mean centering, multicollinearity, and moderators in multiple regression: The reconciliation redux. Behav Res Methods [Internet]. 2017 Feb;49(1):403–4. Available from: http://dx.doi.org/10.3758/s13428-016-0827-9

109.   McCue C. Chapter 7 - Identification, Characterization, and Modeling. In: McCue C, editor. Data Mining and Predictive Analysis (Second Edition) [Internet]. Boston: Butterworth-Heinemann; 2015. p. 137–55. Available from: https://www.sciencedirect.com/science/article/pii/B9780128002292000079

110.   Marquand AF, Wolfers T, Mennes M, Buitelaar J, Beckmann CF. Beyond Lumping and Splitting: A Review of Computational Approaches for Stratifying Psychiatric Disorders. Biol Psychiatry Cogn Neurosci Neuroimaging [Internet]. 2016 Sep;1(5):433–47. Available from: http://dx.doi.org/10.1016/j.bpsc.2016.04.002

111.   Pang R, Lansdell BJ, Fairhall AL. Dimensionality reduction in neuroscience. Curr Biol [Internet]. 2016 Jul 25;26(14):R656-60. Available from: http://dx.doi.org/10.1016/j.cub.2016.05.029

112.   Liu C, Zhang X, Nguyen TT, Liu J, Wu T, Lee E, et al. Partial least squares regression and principal component analysis: similarity and differences between two popular variable reduction approaches. Gen Psychiatr [Internet]. 2022 Jan 27;35(1):e100662. Available from: http://dx.doi.org/10.1136/gpsych-2021-100662

113.   Wold HOA. Soft Modeling: The Basic Design and Some Extensions. In: Jöreskog KG, Wold HOA, editors. Systems Under Indirect Observations: Part II. Amsterdam; 1982. p. 1–54.

114.   Eriksson L, Antti H, Gottfries J, Holmes E, Johansson E, Lindgren F, et al. Using chemometrics for navigating in the large data sets of genomics, proteomics, and metabonomics (gpm). Anal Bioanal Chem [Internet]. 2004 Oct;380(3):419–29. Available from: http://dx.doi.org/10.1007/s00216-004-2783-y

115.   Krishnan A, Williams LJ, Mcintosh AR, Abdi H. Partial least squares (PLS)methods for neuroimaging: a tutorial and review. Neuroimage. 2011;56(2):455–75.

116.   McIntosh AR, Bookstein FL, Haxby JV, Grady CL. Spatial pattern analysis of functional brain images using partial least squares. Neuroimage [Internet]. 1996 Jun;3(3 Pt 1):143–57. Available from: http://dx.doi.org/10.1006/nimg.1996.0016

117.   Chamberlain SR, Cavanagh J, de Boer P, Mondelli V, Jones DNC, Drevets WC, et al. Treatment-resistant depression and peripheral C-reactive protein. Br J Psychiatry [Internet]. 2019 Jan;214(1):11–9. Available from: http://dx.doi.org/10.1192/bjp.2018.66

118.   Kirschner M, Shafiei G, Markello RD, Makowski C, Talpalaru A, Hodzic-Santor B, et al. Latent Clinical-Anatomical Dimensions of Schizophrenia. Schizophr Bull [Internet]. 2020 Dec 1;46(6):1426–38. Available from: http://dx.doi.org/10.1093/schbul/sbaa097

119.   Monteiro JM, Rao A, Shawe-Taylor J, Mourão-Miranda J. A multiple hold-out framework for Sparse Partial Least Squares. J Neurosci Methods [Internet]. 2016 Sep;271:182–94. Available from: http://dx.doi.org/10.1016/j.jneumeth.2016.06.011

120.   Liang W, Ma S, Zhang Q, Zhu T. Integrative sparse partial least squares. Stat Med [Internet]. 2021 Apr;40(9):2239–56. Available from: http://dx.doi.org/10.1002/sim.8900

121. Kaufman S, Rosset S, Perlich C. Leakage in data mining: formulation, detection, and avoidance. In: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining [Internet]. New York, NY, USA: Association for Computing Machinery; 2011 [cited 2022 Jul 18]. p. 556–63. (KDD '11). Available from: https://doi.org/10.1145/2020408.2020496

122. Koutsouleris N, Meisenzahl EM, Borgwardt S, Riecher-Rossler A, Frodl T, Kambeitz J, et al. Individualized differential diagnosis of schizophrenia and mood disorders using neuroanatomical biomarkers. Brain [Internet]. 2015;138(Pt 7):2059–73. Available from: https://www.ncbi.nlm.nih.gov/pubmed/25935725

123. Dukart J, Schroeter ML, Mueller K, Alzheimer's Disease Neuroimaging, Initiative. Age correction in dementia--matching to a healthy brain. PLoS One [Internet]. 2011;6(7):e22193. Available from: https://www.ncbi.nlm.nih.gov/pubmed/21829449

124. Goeman JJ, Solari A. Multiple hypothesis testing in genomics. Stat Med [Internet]. 2014 May 20;33(11):1946–78. Available from: http://dx.doi.org/10.1002/sim.6082

125. Nichols TE, Holmes AP. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum Brain Mapp [Internet]. 2002 Jan;15(1):1–25. Available from: http://dx.doi.org/10.1002/hbm.1058

126. Larobina M, Murino L. Medical image file formats. J Digit Imaging [Internet]. 2014 Apr;27(2):200–6. Available from: http://dx.doi.org/10.1007/s10278-013-9657-9

127. Fan L, Li H, Zhuo J, Zhang Y, Wang J, Chen L, et al. The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. Cereb Cortex [Internet]. 2016 Aug;26(8):3508–26. Available from: http://dx.doi.org/10.1093/cercor/bhw157

128. Diedrichsen J, Balsters JH, Flavell J, Cussans E, Ramnani N. A probabilistic MR atlas of the human cerebellum. Neuroimage [Internet]. 2009 Jul;47:S122. Available from: http://dx.doi.org/10.1016/s1053-8119(09)71166-8

129. Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol [Internet]. 2011 Sep;106(3):1125–65. Available from: http://dx.doi.org/10.1152/jn.00338.2011

130. Buckner RL, Krienen FM, Castellanos A, Diaz JC, Yeo BTT. The organization of the human cerebellum estimated by intrinsic functional connectivity. J Neurophysiol [Internet]. 2011 Nov;106(5):2322–45. Available from: http://dx.doi.org/10.1152/jn.00339.2011

Acknowledgements

First, I thank Prof. Dr. Nikolaos Koutsouleris for taking me under his wing five years ago and enabling me to have a career in science. I saw him give a lecture in 2016 and immediately knew that his approach to science and psychiatry, his expertise in data as the backbone of all scientific efforts, and his vision for the future of psychiatric care would guide and inspire me for a very long time. Beyond that, his care on a human level, whether with colleagues or patients, inspired me to become a clinician scientist. Furthermore, he helped me through tough times as well; he knew when to put on the brakes and when to challenge me. I have grown not only as a clinician or a scientist, but also as a human being, and for that I thank Nikos.

Additionally, I thank Prof. Dr. Kolja Schiltz for allowing me to conduct my studies and giving me scientific freedom over the past three years. Without this, neither Paper I nor this thesis would have been possible. Beyond that, he inspired me to follow my passion for forensic science, to hone my clinical skills in this field, and to transfer my methodological knowledge to the forensic field.

Moreover, Prof. Dr. Peter Falkai already believed in me while I was still a medical student and, to this day, has continuously supported me in all my scientific and clinical endeavors. He allowed me to intern at his clinic during medical school back in 2013 and gave me my first job as a medical doctor in February 2016. Afterwards, he supported me in applying for the joint residency/PhD-Program from the Else-Kröner-Fresenius Foundation and the International Max Planck Research School for Translational Psychiatry (IMPRS-TP), which has led to this thesis. He also believed in me by allowing me to switch to forensic psychiatry in 2019 and has continuously supported me throughout the past six years. Closely tied to this is, of course, Prof. Dr. Andrea Schmitt, who helped me navigate the challenges of being a young psychiatrist and a scientist from the first day I stepped into this hospital. I fondly remember our many talks, phone and zoom calls, as well as constant reflections and discussions about my progress and my career.

Altogether, I am deeply indebted to Prof. Falkai, Prof. Koutsouleris, Prof. Schiltz, and Prof. Schmitt for all the support, guidance, and inspiration, and I could not wish for any better supervisors and mentors.

Beyond that, I am grateful to the entire team at the Koutsouleris Laboratory, specifically to Anne Ruef, who supervised me during the first years of my thesis. She taught me to code, introduced me to PLS, and did not let me get away with sloppy analyses or unprecise statements. Dominic Dwyer and Riya Paul have provided constant guidance on

I dedicate this thesis to my late uncles Karl-Heinz and Radovan.

Archival Report

Traces of Trauma: A Multivariate Pattern Analysis of Childhood Trauma, Brain Structure, and Clinical Phenotypes

David Popovic, Anne Ruef, Dominic B. Dwyer, Linda A. Antonucci, Julia Eder, Rachele Sanfelici, Lana Kambeitz-Ilankovic, Omer Faruk Oztuerk, Mark S. Dong, Riya Paul, Marco Paolini, Dennis Hedderich, Theresa Haidl, Joseph Kambeitz, Stephan Ruhrmann, Katharine Chisholm, Frauke Schultze-Lutter, Peter Falkai, Giulio Pergola, Giuseppe Blasi, Alessandro Bertolino, Rebekka Lencer, Udo Dannlowski, Rachel Upthegrove, Raimo K.R. Salokangas, Christos Pantelis, Eva Meisenzahl, Stephen J. Wood, Paolo Brambilla, Stefan Borgwardt, Nikolaos Koutsouleris, and the PRONIA Consortium

ABSTRACT

BACKGROUND: Childhood trauma (CT) is a major yet elusive psychiatric risk factor, whose multidimensional conceptualization and heterogeneous effects on brain morphology might demand advanced mathematical modeling. Therefore, we present an unsupervised machine learning approach to characterize the clinical and neuroanatomical complexity of CT in a larger, transdiagnostic context.

METHODS: We used a multicenter European cohort of 1076 female and male individuals (discovery: $n = 649$; replication: $n = 427$) comprising young, minimally medicated patients with clinical high-risk states for psychosis; patients with recent-onset depression or psychosis; and healthy volunteers. We employed multivariate sparse partial least squares analysis to detect parsimonious associations between combinations of items from the Childhood Trauma Questionnaire and gray matter volume and tested their generalizability via nested cross-validation as well as via external validation. We investigated the associations of these CT signatures with state (functioning, depressivity, quality of life), trait (personality), and sociodemographic levels.

RESULTS: We discovered signatures of age-dependent sexual abuse and sex-dependent physical and sexual abuse, as well as emotional trauma, which projected onto gray matter volume patterns in prefronto-cerebellar, limbic, and sensory networks. These signatures were associated with predominantly impaired clinical state- and trait-level phenotypes, while pointing toward an interaction between sexual abuse, age, urbanicity, and education. We validated the clinical profiles for all three CT signatures in the replication sample.

CONCLUSIONS: Our results suggest distinct multilayered associations between partially age- and sex-dependent patterns of CT, distributed neuroanatomical networks, and clinical profiles. Hence, our study highlights how machine learning approaches can shape future, more fine-grained CT research.

Keywords: Childhood trauma, Machine learning, Morphometry, MRI, Sparse partial least squares, Transdiagnostic

https://doi.org/10.1016/j.biopsych.2020.05.020

Childhood trauma (CT) is defined as any act that results in harm, potential harm, or threat of harm to a child (1) and is generally operationalized along the dimensions of physical abuse or neglect, sexual abuse, and emotional abuse or neglect (2). CT acts as a transdiagnostic risk factor for a variety of psychiatric disorders (3–5), reduces an individual's quality of life (6), impairs levels of functioning (7), and is associated with dysfunctional personality development (8,9). Furthermore, neuroimaging studies have suggested associations between CT and gray matter volume (GMV), reporting alterations in the subcortical, temporal, and frontal regions (10–13). Yet, these findings have been highly heterogeneous, and so far, neither a

distinct correlate of CT (14–19) nor a link between CT-related brain changes and observable clinical phenotypes has been established (20,21).

A better neurobiological understanding of CT is important, as it could mitigate the long-term adverse effects of CT through early recognition and targeted multimodal intervention programs (22,23). Still, most studies investigating CT use voxelwise mass-univariate strategies, which assume highly localized functional specialization and statistical independence of voxels (24). This approach does not reflect the state-of-the-art understanding of neuroanatomical variation being encoded along distributed clusters of voxels, cortical regions, and brain

systems (25–27), potentially leading to subtle and distributed effects of CT on brain morphology (28). The diverse effects of CT might be better understood in a larger context by investigating the more generalized, transdiagnostic effects of CT, and CT's important interactions with age and sex (29–32). Therefore, advanced methods are needed to capture the complexity of CT and potentially associated structural brain surrogates (33).

We took an in-depth approach to better characterize the complex neuroanatomy of CT by investigating the relationship between structural brain data and CT in the multicenter, European PRONIA (Personalized Prognostic Tools for Early Psychosis Management) study cohort (https://www.pronia.eu/). Following a transdiagnostic, data-driven study design, we applied the multivariate sparse partial least squares (SPLS) algorithm to identify parsimonious and interpretable phenotype-brain signatures (34). Specifically, we used the strength of SPLS to model complex patterns of interactions between CT-related phenotypic features and brain voxels, possibly yielding new and distinct CT signatures. Finally, we wanted to examine the clinical and sociodemographic implications of these novel CT dimensions by performing correlation analyses between participants' loadings onto the CT signatures and measures of functioning, depressivity, quality of life, personality, and sociodemographic information. We expected to find transdiagnostic CT signatures linked to clinical and sociodemographic characteristics, providing further insights into the multidimensional fingerprints of CT.

METHODS AND MATERIALS

Study Participants

The PRONIA study cohort includes healthy control (HC) subjects, participants with recent-onset depression (ROD) or recent-onset psychosis (ROP), and patients with clinical high-risk states for psychosis (CHR). The cohort is divided into a discovery sample for model generation and a replication sample for model validation [Supplemental Methods and Koutsouleris *et al.* (35)]. Data from 649 participants from the discovery sample (264 HC subjects, 129 participants with ROD, 132 participants with ROP, 124 patients with CHR) (Table 1) and 427 individuals from the replication sample (135 HC subjects, 96 participants with ROD, 92 participants with ROP, 104 patients with CHR) (Table S6) were obtained for the analysis.

CT and Clinical and Sociodemographic Features Assessment

CT was measured using the Childhood Trauma Questionnaire (CTQ) (36,37). The CTQ is a 28-item self-report questionnaire that assesses 5 types of maltreatment—emotional, physical, and sexual abuse as well as emotional and physical neglect—and contains an additional denial measure. A 5-point Likert-type scale is used to record responses ranging from 1 ("Never True") to 5 ("Very Often True"). Internal consistency scores of the CTQ subscales range from 0.66 (physical neglect) to 0.94 (sexual abuse), while the test-retest coefficient over a 3.5-month period was calculated at 0.80 (36–38).

Functioning was evaluated using the Global Assessment of Functioning Symptom Scale (GAF:S) and Global Assessment of Functioning Disability/Impairment Scale (GAF:D/I) (39) and the Global Functioning: Social Scale (GF:S) and Global Functioning: Role Scale (GF:R) (40), while depressive symptoms were quantified using the Beck Depression Inventory (BDI) (41). The World Health Organization Quality of Life Short Version (WHOQOL-BREF) was applied to measure individual perception of quality of life (42). Personality domains were assessed using the Neuroticism-Extraversion-Openness Five-Factor Inventory (NEO-FFI), quantifying personality traits along 5 domains: openness, conscientiousness, extraversion, agreeableness, and neuroticism (43).

Sociodemographic features were assessed along the domains of the participant's ethnicity, urbanicity, religion, parental education background, family and relationship status, and education level and employment status.

Magnetic Resonance Imaging Data Acquisition and Preprocessing

T1-weighted structural magnetic resonance imaging data were acquired from the study participants (Supplemental Methods). All images were examined for artifacts, gross anatomical abnormalities, and signs of neurological disease by trained clinical neuroradiologists. Structural magnetic resonance imaging data were preprocessed using the CAT12 toolbox (version 1206; http://www.neuro.uni-jena.de/cat/), an extension of the SPM12 software (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/), and final GMVs were corrected for total intracranial volume.

SPLS Analysis

We used phenotypic and brain data as input for the SPLS algorithm. Our phenotypic dataset contained all 28 CTQ items, age and sex as well-established modulators of CT (31,32,44,45), and study group. The brain dataset consisted of vectorized whole-brain GMV (resliced to 3 mm) for all individuals. Given these two datasets, SPLS uses a singular value decomposition to compute a latent variable (LV) capturing a specific associative effect between phenotypic and brain data. For each dataset, the LV contains a vector with feature weights (values ranging from −1 to 1) measuring the covariance between the two datasets. Therefore, the LV consists of paired multivariate profiles measuring how the phenotypic features (phenotypic pattern) relate to the brain features (brain pattern) (Supplemental Methods). Another characteristic of SPLS is the enforcement of sparsity, whereby weights of zero are assigned to features that did not yield any relevant association. The process of weighting and selecting features according to their covariance is accomplished via l_1- and l_2-norm constraints, similar to elastic net regularization (46), and is controlled by a pair of hyperparameters. Additionally, every participant can be assigned a pair of latent phenotypic and brain scores. These latent phenotypic and brain scores indicate how strong a participant loads on the phenotypic and brain patterns of the LV, respectively, with greater latent score values reflecting higher individual loading and vice versa. We used these latent scores for post hoc correlation analyses to investigate clinical and sociodemographic aspects of the LV signatures (34).

Neuroanatomical Signatures of Childhood Adversity

Table 1. Clinical and Demographic Characteristics of the Discovery Sample

	All	HC Group	ROD Group	CHR Group	ROP Group	H/χ^2 (df)	p Value
Age, Years	28.39 (6.00)	28.50 (6.45)	29.09 (6.21)	27.02 (4.84)	28.73 (5.63)	8.98 (2)[a]	.011[b]
Sex, Female %	53%	62%	54%	48%	38%	7.41 (1)[c]	.024[b]
Education, Years	14.77 (3.25)	15.69 (3.17)	14.70 (3.16)	13.78 (3.03)	13.93 (3.15)	5.56 (2)[a]	.062
GAF:S	65.15 (21.12)	86.52 (6.51)	55.76 (12.48)	54.95 (11.00)	41.13 (13.22)	86.63 (2)[a]	1.55×10^{-19b}
GAF:D/I	65.57 (20.1)	85.16 (5.86)	56.36 (14.42)	55.93 (13.94)	44.44 (12.23)	59.82 (2)[a]	1.02×10^{-13b}
GF:S	7.15 (1.67)	8.51 (0.84)	6.47 (1.34)	6.51 (1.36)	5.68 (1.47)	28.11 (2)[a]	7.86×10^{-07b}
GF:R	6.97 (1.90)	8.56 (0.75)	6.23 (1.69)	6.18 (1.44)	5.24 (1.65)	29.66 (2)[a]	3.62×10^{-07b}
Right-Handed, %	91%	94%	90%	88%	90%	0.41 (1)[c]	.82
PANSS Total	55.97 (18.83)	NA	47.55 (10.91)	50.57 (13.23)	69.29 (21.92)	87.93 (2)[a]	8.07×10^{-20b}
PANSS Positive	11.92 (6.00)	NA	7.67 (1.24)	10.23 (2.96)	17.68 (6.50)	204.19 (2)[a]	4.59×10^{-45b}
PANSS Negative	13.77 (6.40)	NA	12.56 (4.98)	12.53 (5.88)	16.14 (7.37)	21.62 (2)[a]	2.02×10^{-05b}
PANSS General	30.25 (9.38)	NA	27.31 (6.73)	27.78 (6.90)	35.47 (11.23)	50.54 (2)[a]	1.06×10^{-11b}
BDI	15.78 (14.62)	3.73 (5.27)	26.23 (13.82)	25.49 (12.24)	21.05 (12.49)	11.05 (2)[a]	.004[b]
Study Center, n						149.87 (6)[c]	8.23×10^{-30b}
Munich	181	58	44	38	41		
Basel	84	37	15	17	15		
Cologne	131	59	24	20	28		
Birmingham	80	43	14	13	10		
Milan	37	13	6	7	11		
Turku	74	23	12	17	22		
Udine	62	31	14	12	5		
Total	649	264	129	124	132		

Values are mean (SD), except where noted. The p values are stated after false discovery rate correction for multiple testing.

BDI, Beck Depression Inventory; CHR, clinical high-risk states for psychosis; GAF:D/I, Global Assessment of Functioning Disability/Impairment Scale; GAF:S, Global Assessment of Functioning Social Scale; GF:R, Global Functioning: Role Scale; GF:S, Global Functioning: Social Scale; HC, healthy control; NA, not available; PANSS, Positive and Negative Syndrome Scale; ROD, recent-onset depression; ROP, recent-onset psychosis.

[a]Kruskal-Wallis test (H test).
[b]Significant value.
[c]χ^2 test.

Assessment of Generalizability and Significance

We implemented a nested cross-validation framework, which robustly prevents information leakage between participants used for training and validating the models (47,48) (see Figure S2). We used 10 inner folds for hyperparameter optimization of the l_1- and l_2-norm constraints and 10 outer folds to test the optimized model against a previously held-out dataset. Before entering the SPLS analysis, z-transformation models were generated in the training data and then applied to the test data within the nested cross-validation structure. Significance testing of each LV was done by comparing the performance of the optimized model against 5000 permutations of the dataset. If an LV proved significant, the respective covariance component was removed from the two datasets via projection deflation, and the next LV was computed on the deflated datasets. This process was repeated until an LV failed to reach significance, thus generating several layers of significant, associative effects. LVs are labeled according to the order of their computation (Supplemental Methods). The generalizability of the CT model was further validated by applying data from the replication sample onto the phenotypic and neuroanatomic patterns of all its LVs, thus generating latent phenotypic and brain scores in the replication sample. These latent scores were correlated to our predefined set of clinical and sociodemographic parameters. Univariate partial correlation analysis between the 7 study sites and the input

datasets was used within the nested cross-validation scheme to correct for site effects (49,50).

Univariate Analysis

Group-level sociodemographic and clinical differences were assessed using nonparametric tests (Kruskal-Wallis H test, Wilcoxon-Mann-Whitney test, Dunn's post hoc multiple comparison test, χ^2 test). Latent trauma and brain scores were correlated to clinical and sociodemographic features using Spearman's correlation coefficient (ρ). Analyses were false discovery rate–corrected for multiple testing at a significance threshold of $q < .05$ (51).

RESULTS

Group-Level Differences at Baseline

The clinical study populations (ROD, CHR, and ROP participants) revealed significant differences with respect to age, sex, GAF, GF, Positive and Negative Syndrome Scale, and BDI (Table 1, Tables S7 and S8). Furthermore, a significant difference for the recruitment of study groups across sites was found (Table 1, Table S9). The clinical study populations also displayed significant differences in antidepressant, antipsychotic, and sedative treatment (Tables S10 and S11). Moreover, the clinical study populations of the discovery and replication samples did not reveal any significant differences

with regard to CTQ total or subscale scores (Table 2, Table S12).

SPLS Results: Association Between Phenotypic and Brain Data

SPLS analysis of all 649 discovery sample subjects yielded 5 significant LVs (LV1–LV5), representing different layers of association between phenotypic and brain patterns (see Tables S13 and S14 for CTQ item list and atlas readouts and Figure S20 for visualization of phenotype-brain correlations).

LV1: Age (p value = 1.9 $\times$ 10^{-04}). Phenotypic pattern (Figure S6A): Age received the strongest positive weight, whereas further positive weights were assigned to male sex and ROP status and to the subscales of sexual abuse (5 items), physical abuse (4 items), emotional abuse (1 item), and physical neglect (1 item). Smaller negative weights were distributed to emotional abuse (1 item), denial (1 item), and female sex. Brain pattern (Figure S6B): GMV was widely negatively weighted across the frontal, temporal, parietal, and occipital regions as well as across the subcortical areas. Positive GMV weights were sparsely found in the thalamic region.

LV2: Sexual Abuse and Age (p = 1.9 $\times$ 10^{-04}). Phenotypic pattern (Figure 1A): Two questions from the sexual abuse subscale were positively weighted, while age was negatively weighted. Brain pattern (Figure 1B): GMV was

assigned negative weights bilaterally in the prefrontal cortex (PFC), particularly in the dorsolateral PFC (DLPFC) and medial PFC regions. Further negative weights were found bilaterally in the superior and middle temporal gyrus as well as unilaterally in the left angular gyrus. Positive weighting was detected bilaterally in the cerebellum, premotor cortex, cuneus, lingual gyrus, and the ganglia.

LV3: Sex (p = 1.9 $\times$ 10^{-04}). Phenotypic pattern (Figure S7A): The strongest positive and negative weights were detected for male and female sex, respectively. Moreover, positive weights were assigned to emotional abuse (1 item), physical abuse (2 items), sexual abuse (3 items), emotional neglect (1 item), and physical neglect (2 items), while smaller negative weights were distributed to age, sexual abuse (1 item), and denial (1 item). Brain pattern (Figure S7B): GMV was positively weighted in the occipital, parietal, and frontal areas, particularly in the precuneus region, and negatively bilaterally weighted in the prefrontal, hippocampal, and parietal areas.

LV4: Physical/Sexual Abuse and Sex (p = 1.2 $\times$ 10^{-03}). Phenotypic pattern (Figure 2A): Physical (3 items) and sexual abuse (4 items) received positive weights, while male and female sex were weighted inversely. Brain pattern (Figure 2B): The most profound effect was detected in bilateral positive weighting of GMV in the primary somatosensory cortex, basal ganglia, and cuneus as well as in unilaterally reduced GMV in the left fusiform gyrus and right DLPFC. GMV was also positively weighted bilaterally in the occipital gyrus, angular

Table 2. Group-Level Statistics for CTQ Differences Between the Discovery and Replication Samples

CTQ Scale	Sample	All	HC Group	ROD Group	CHR Group	ROP Group	H (df)	p
Total	Discovery	37.0 (12.1)	30.8 (5.8)	40.0 (14.6)	41.8 (13.1)	41.9 (12.5)	5.08 (2)	.55[a]
	Replication	38.3 (13.1)	31.0 (6.9)	40.6 (11.9)	42.6 (13.7)	41.8 (15.8)	1.20 (2)	.76[a]
	p	.50[b]	.91[b]	.59[b]	.84[b]	.61[b]		
Emotional Abuse	Discovery	8.4 (4.0)	6.5 (2.4)	9.2 (4.5)	10.2 (4.4)	9.8 (4.4)	5.20 (2)	.52[a]
	Replication	9.0 (4.5)	6.4 (2.0)	9.4 (4.1)	10.8 (4.9)	10.1 (5.2)	3.70 (2)	.50[a]
	p	.50[b]	.71[b]	.69[b]	.72[b]	.97[b]		
Physical Abuse	Discovery	6.0 (2.5)	5.4 (1.0)	6.5 (3.3)	6.5 (3.1)	6.5 (2.9)	1.33 (2)	.95[a]
	Replication	6.2 (2.6)	5.5 (1.5)	6.3 (2.4)	6.6 (3.0)	6.6 (3.3)	0.25 (2)	.98[a]
	p	.56[b]	.77[b]	.64[b]	.72[b]	.89[b]		
Sexual Abuse	Discovery	5.7 (2.4)	5.2 (0.9)	5.9 (2.8)	6.0 (2.8)	6.3 (3.1)	2.84 (2)	.50[a]
	Replication	5.8 (2.6)	5.1 (0.9)	5.9 (2.9)	6.1 (2.9)	6.3 (3.2)	2.39 (2)	.60[a]
	p	.95[b]	.71[b]	.76[b]	.92[b]	.87[b]		
Emotional Neglect	Discovery	10.0 (4.4)	7.9 (3.0)	11.3 (5.1)	11.8 (4.5)	11.4 (4.4)	1.73 (2)	.80[a]
	Replication	10.4 (4.6)	8.0 (3.2)	11.8 (4.8)	11.7 (4.4)	11.1 (5.0)	1.46 (2)	.72[a]
	p	.54[b]	.95[b]	.61[b]	.86[b]	.70[b]		
Physical Neglect	Discovery	6.8 (2.4)	5.8 (1.4)	7.1 (2.9)	7.3 (2.6)	7.8 (2.8)	9.70 (2)	.05[a]
	Replication	6.9 (2.5)	5.9 (1.6)	7.1 (2.3)	7.4 (2.6)	7.6 (3.1)	0.19 (2)	.99[a]
	p	.63[b]	.74[b]	.62[b]	.99[b]	.51[b]		
Denial	Discovery	0.6 (0.9)	0.7 (1.0)	0.4 (0.8)	0.4 (0.8)	0.5 (0.9)	1.22 (2)	.99[a]
	Replication	0.6 (0.9)	0.8 (1.1)	0.4 (0.8)	0.3 (0.8)	0.6 (0.9)	7.73 (2)	.15[a]
	p	.85[b]	.65[b]	.88[b]	.82[b]	.51[b]		

Values are mean (SD). The p values are stated after false discovery rate correction for multiple testing.

CHR, clinical high-risk states for psychosis; CTQ, Childhood Trauma Questionnaire; df, degrees of freedom; HC, healthy control; ROD, recent-onset depression; ROP, recent-onset psychosis.

[a]Kruskal-Wallis test (H test).

[b]Wilcoxon-Mann-Whitney test.

Neuroanatomical Signatures of Childhood Adversity
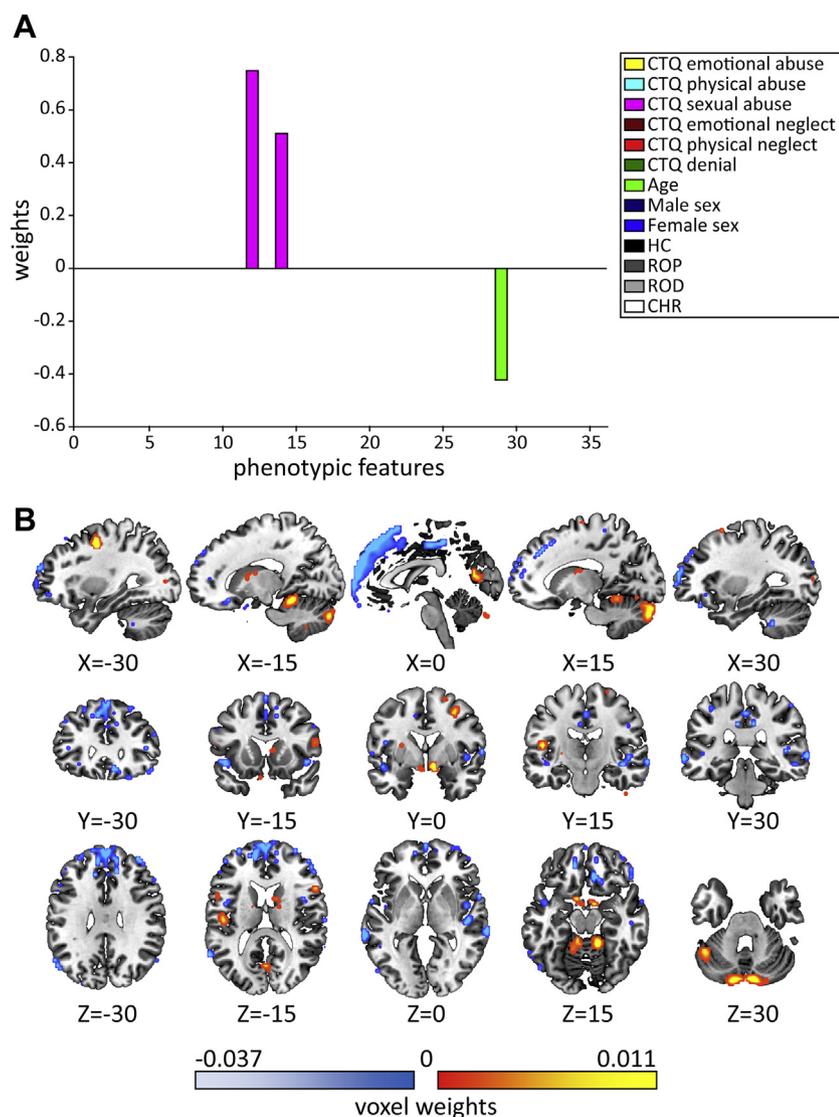
Figure 1. Age-dependent sexual abuse signature of latent variable 2 (LV2). **(A)** The barplot visualizes the direction and values of the weights included in the phenotypic pattern of LV2. Two questions from the Childhood Trauma Questionnaire (CTQ) sexual abuse subscale (CTQ21, CTQ24) received a positive weight, while age received a negative weight. **(B)** Depicted is the brain pattern of LV2, with positive weighting of voxels displayed in the red color scale and negative weighting displayed in the blue color scale. CHR, clinical high-risk states for psychosis; HC, healthy control; ROD, recent-onset depression; ROP, recent-onset psychosis.

and supramarginal gyrus, and thalamus. Smaller clusters of negative GMV weights were discovered bilaterally in the superior and middle temporal gyrus, cingulate gyrus, (para)hippocampus, precuneus, and right PFC.

LV5: Emotional Abuse/Neglect ($p = 1.9 \times 10^{-04}$). Phenotypic pattern (Figure 3A): Emotional abuse (3 items) and neglect (3 items) were weighted positively. Brain pattern (Figure 3B): The largest effects were found in bilateral positive GMV weights in in the cuneus and left primary somatosensory cortex as well as in bilateral negative weights in the cingulate. Smaller positive weights were found in the right occipital region and left DLPFC, whereas negative weighting was detected in the left insula, right caudate nucleus, left supramarginal gyrus, right hippocampus, and bilaterally in the fusiform gyrus.

In summary, LV1 and LV3 represented mostly patterns of age- and sex-related brain maturation processes,

respectively, whereas the other 3 LVs were more trauma specific, with LV2 reflecting an age-informed sexual abuse pattern, LV4 displaying a sex-dependent signature of physical and sexual abuse, and LV5 containing an emotional trauma pattern.

SPLS Results: Correlation Between Latent Scores and Clinical Domains

In the discovery sample, correlation analyses between clinical domains and latent scores yielded several significant results for all 3 CT-specific LVs (Tables 3 and 4) and for LV1 and LV3 as well (Tables S15 and S16).

LV2: Sexual Abuse and Age. Phenotypic scores: Negative correlations were observed for GF:S, GF:R, GAF:S, GAF:D/I, and WHOQOL-BREF, as well as for the NEO-FFI agreeableness and conscientiousness domains ($\rho = -.09$ to $-.17$, $p = .04$ to 1.3×10^{-05}). Positive correlations were detected for

Biological Psychiatry
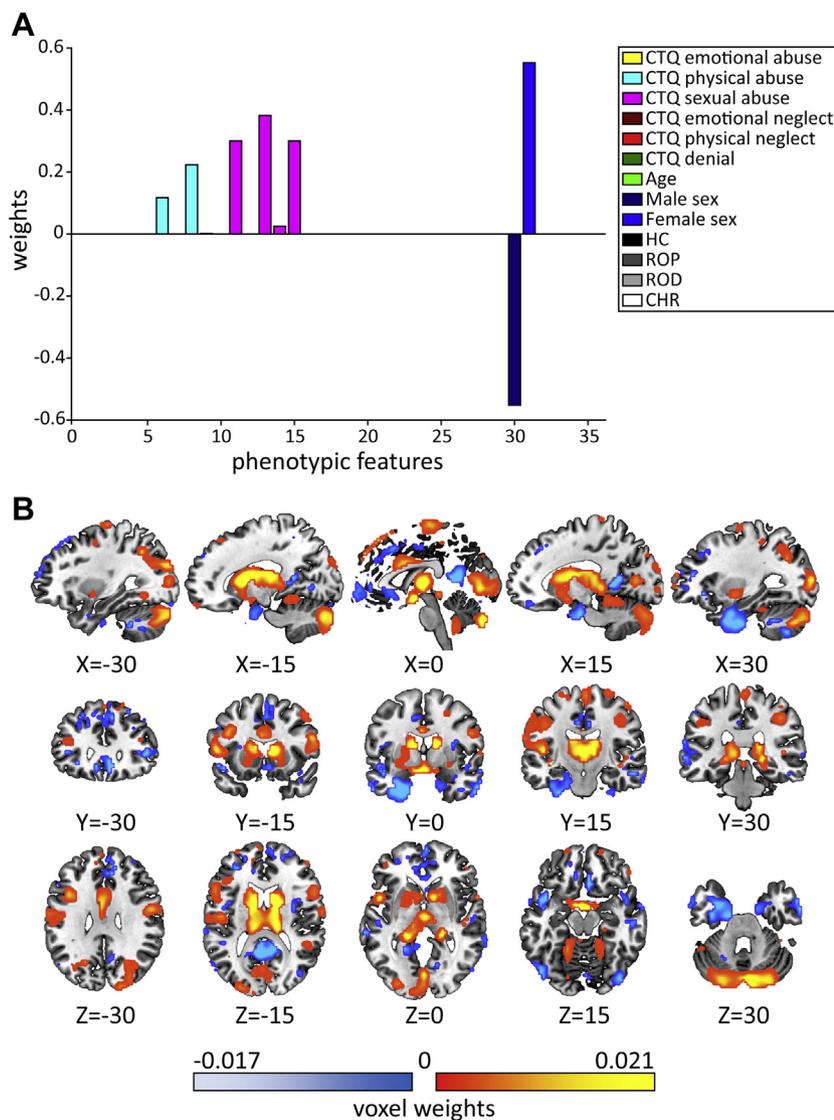
Figure 2. Sex-dependent sexual and physical abuse signature of latent variable 4 (LV4). **(A)** The barplot visualizes the direction and values of the weights included in the phenotypic pattern of LV4. Three questions from the Childhood Trauma Questionnaire (CTQ) physical abuse subscale (CTQ09, CTQ12, CTQ15) and 4 questions from the sexual abuse subscale (CTQ20, CTQ23, CTQ24, CTQ27) received positive weights. Male sex received a negative weight, and female sex received a positive weight. **(B)** Depicted is the brain pattern of LV4, with positive weighting of voxels displayed in the red color scale and negative weighting displayed in the blue color scale. CHR, clinical high-risk states for psychosis; HC, healthy control; ROD, recent-onset depression; ROP, recent-onset psychosis.

NEO-FFI neuroticism and BDI scores (ρ = .11 to .15, p = .01 to 3.3×10^{-04}). Brain scores: No significant associations were detected.

LV4: Sexual/Physical Abuse and Sex. Phenotypic scores: We detected negative correlations for most GAF, GF, and WHOQOL-BREF domains as well as for the NEO-FFI domains of extraversion and conscientiousness (ρ = −.08 to −.20, p = .04 to 4.5×10^{-07}). Positive associations were found for NEO-FFI neuroticism and BDI total scores (ρ = .18 to .21, p = 3.0 $\times 10^{-06}$ to 1.2×10^{-07}). Brain scores: Negative correlations were detected for GAF:S and WHOQOL-BREF (ρ = −.11 to −.17), p = .05 to 2.4×10^{-05}). We observed a positive association with the NEO-FFI neuroticism domain (ρ = .11, p = .05).

LV5: Emotional Abuse/Neglect. Phenotypic scores: Negative correlations were detected for all GAF, GF, and

WHOQOL-BREF domains as well as for the NEO-FFI extraversion, agreeableness, and conscientiousness domains (ρ = −.22 to −.47, p = .05 to 1.2×10^{-29}). Positive correlations were found for the BDI and NEO-FFI neuroticism domain levels (ρ = .44 to .48, p = 1.6×10^{-25} to 1.0×10^{-30}). Brain scores: Negative correlations were found for GAF, GF, and WHOQOL-BREF domains as well as for the NEO-FFI extraversion and conscientiousness domains (ρ = −.09 to −.18, p = .04 to 4.8×10^{-06}). Positive correlations were observed for the BDI and NEO-FFI neuroticism domain (ρ = .13 to .19, p = 2.1×10^{-03} to 3.7×10^{-06}).

External Clinical Validation of the SPLS Trauma Model

Fifty-nine of 84 (70%) significant clinical associations from the discovery sample were successfully validated in the replication sample, whereby 48 of 61 (79%) phenotype-level correlations and 11 of 23 (48%) brain-level correlations were replicated. Two phenotypic and 18 brain-level associations were

Neuroanatomical Signatures of Childhood Adversity
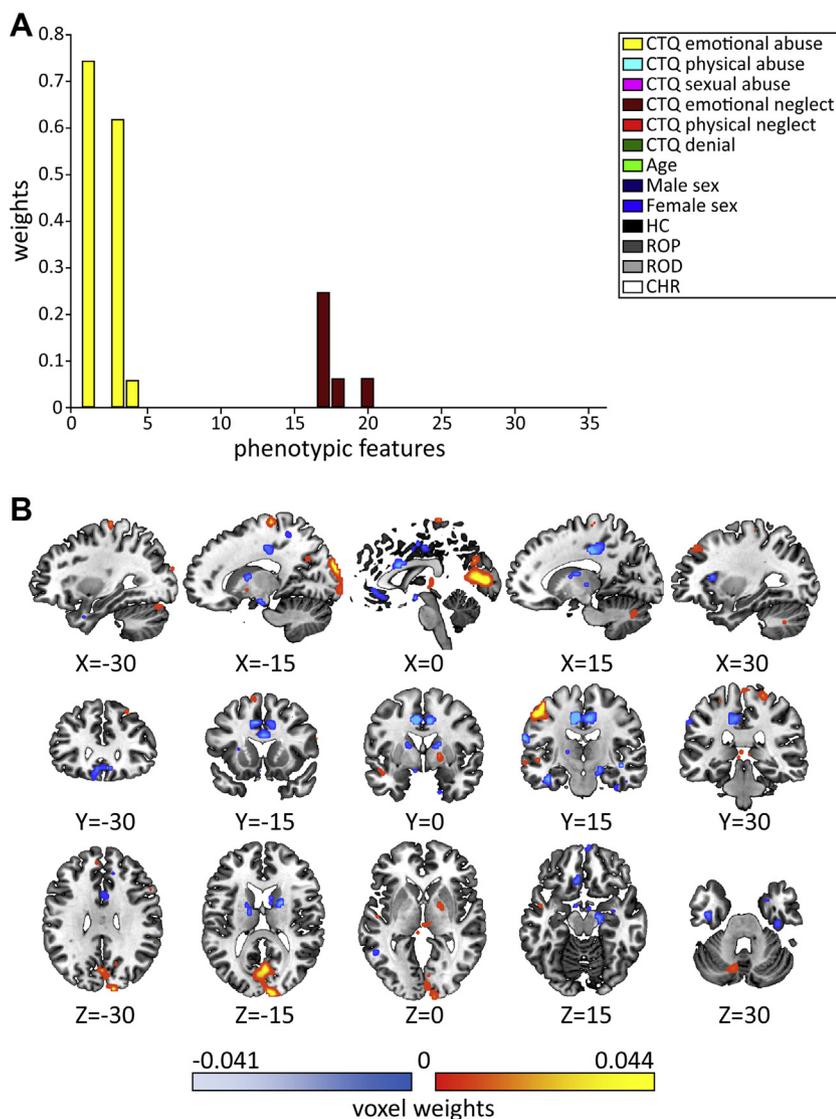
Figure 3. Emotional trauma signature of latent variable 5 (LV5). **(A)** The barplot visualizes the direction and values of the weights included in the phenotypic pattern of LV5. Three questions each from the Childhood Trauma Questionnaire (CTQ) subscales of emotional abuse (CTQ03, CTQ14, CTQ18) and emotional neglect (CTQ07, CTQ13, CTQ28) received positive weights. **(B)** Depicted is the brain pattern of LV5, with positive weighting of voxels displayed in the red color scale and negative weighting displayed in the blue color scale. CHR, clinical high-risk states for psychosis; HC, healthy control; ROD, recent-onset depression; ROP, recent-onset psychosis.

additionally detected, amounting to a total of 79 significant clinical associations (50 phenotypic, 29 brain-level) in the replication sample. Moreover, none of the significant correlations changed their orientation (Tables 3 and 4).

LV2: Sexual Abuse and Age. Phenotypic scores: A total of 12 of 18 (67%) associations were replicated. Additional significant associations were found for the GAF:S past month ($\rho = -.19$, $p = 1.0 \times 10^{-03}$) and NEO-FFI extraversion ($\rho = -.18$, $p = 1.1 \times 10^{-03}$) domains. Brain scores: Additional significant, positive associations were detected for 8 GAF and GF measures ($\rho = .13$ to $.20$, $p = .03$ to 2.4×10^{-04}).

LV4: Sexual/Physical Abuse and Sex. Phenotypic scores: A total of 13 of 20 (65%) associations were replicated, whereas additional correlations were not found. Brain scores: A total of 3 of 3 (100%) correlations were replicated, while further

correlations were found for the GAF and GF, NEO-FFI extraversion domain, and WHOQOL-BREF physical domain ($\rho = -.12$ to $-.19$, $p = .04$ to 7.5×10^{-03}) as well as for the BDI ($\rho = .18$, $p = 1.1 \times 10^{-03}$).

LV5: Emotional Abuse/Neglect. Phenotypic scores: A total of 23 of 23 (100%) associations were replicated, and no additional correlations were detected. Brain scores: A total of 8 of 20 (40%) associations were replicated, and 1 additional correlation was detected for the GAF:S lifetime domain ($\rho = -.15$, $p = .01$).

Sociodemographic Exploration of the SPLS Trauma Model

Correlation analyses between individual latent scores of LV2, LV4, and LV5 and sociodemographic features yielded several significant results (Tables S17–S24).

Table 3. Spearman's Correlation Analyses Between Latent Scores and Clinical Domains of Functioning in the Discovery and Replication Samples

Clinical Domain	Sample	LV2 Sexual Abuse and Age		LV4 Sexual/Physical Abuse and Sex		LV5 Emotional Abuse/Neglect	
		Phenotypic Score	Brain Score	Phenotypic Score	Brain Score	Phenotypic Score	Brain Score
GAF:S							
Lifetime	Discovery	$-.17$ $(3.7 \times 10^{-05})^a$	.01 (.99)	$-.15$ $(1.4 \times 10^{-04})^a$	$-.13$ $(.01)^a$	$-.24$ $(4.6 \times 10^{-03})^a$	$-.05$ (.32)
	Replication	$-.20$ $(2.4 \times 10^{-04})^a$	.07 (.52)	$-.17$ $(2.7 \times 10^{-03})^a$	$-.24$ $(1.4 \times 10^{-05})^a$	$-.29$ $(5.8 \times 10^{-08})^a$	$-.15$ $(.01)^a$
Past Year	Discovery	$-.13$ $(1.3 \times 10^{-03})^a$	.03 (.70)	$-.13$ $(1.5 \times 10^{-03})^a$	$-.09$ (.18)	$-.32$ $(1.5 \times 10^{-10})^a$	$-.09$ $(.03)^a$
	Replication	$-.17$ $(2.1 \times 10^{-03})^a$	.12 (.07)	$-.20$ $(2.6 \times 10^{-04})^a$	$-.13$ $(.03)^a$	$-.38$ $(2.0 \times 10^{-13})^a$	$-.05$ (.70)
Past Month	Discovery	$-.07$ (.15)	.10 (.33)	$-.09$ $(.03)^a$	$-.02$ (.73)	$-.36$ $(9.4 \times 10^{-15})^a$	$-.11$ $(.01)^a$
	Replication	$-.19$ $(1.0 \times 10^{-03})^a$	.15 $(.01)^a$	$-.19$ $(9.9 \times 10^{-04})^a$	$-.15$ $(.01)^a$	$-.38$ $(9.6 \times 10^{-14})^a$	$-.12$ $(.04)^a$
GAF:D/I							
Lifetime	Discovery	$-.17$ $(1.3 \times 10^{-05})^a$	.02 (.80)	$-.14$ $(4.2 \times 10^{-04})^a$	$-.10$ (.08)	$-.29$ $(3.4 \times 10^{-07})^a$	$-.18$ $(5.0 \times 10^{-06})^a$
	Replication	$-.19$ $(5.2 \times 10^{-04})^a$	.05 (.90)	$-.14$ (.02)	$-.17$ $(2.3 \times 10^{-03})^a$	$-.28$ $(1.4 \times 10^{-07})^a$	$-.16$ $(.01)^a$
Past Year	Discovery	$-.16$ $(7.5 \times 10^{-05})^a$	.04 (.64)	$-.14$ $(2.8 \times 10^{-04})^a$	$-.08$ (.30)	$-.35$ $(4.9 \times 10^{-13})^a$	$-.16$ $(5.4 \times 10^{-05})^a$
	Replication	$-.14$ $(.02)^a$	.13 $(.03)^a$	$-.14$ (.02)	$-.08$ (.32)	$-.36$ $(3.5 \times 10^{-12})^a$	$-.07$ (.44)
Past Month	Discovery	$-.09$ (.05)	.08 (.75)	$-.10$ $(.01)^a$	$-.05$ (.55)	$-.38$ $(1.9 \times 10^{-16})^a$	$-.15$ $(2.0 \times 10^{-03})^a$
	Replication	$-.10$ (.14)	.16 $(4.2 \times 10^{-03})^a$	$-.11$ (.11)	$-.09$ (.19)	$-.35$ $(2 \times 10^{-11})^a$	$-.13$ $(.03)^a$
GF:S							
Current	Discovery	$-.11$ $(.01)^a$	.10 (.30)	$-.12$ $(3.5 \times 10^{-03})^a$	.01 (.99)	$-.35$ $(8.4 \times 10^{-14})^a$	$-.12$ $(4.0 \times 10^{-03})^a$
	Replication	$-.10$ (.17)	.16 $(.01)^a$	$-.13$ $(.04)^a$	$-.10$ (.12)	$-.37$ $(8.5 \times 10^{-13})^a$	$-.10$ (.12)
Low Past Year	Discovery	$-.10$ $(.02)^a$	.07 (.52)	$-.12$ $(3.8 \times 10^{-03})^a$	.02 (.83)	$-.34$ $(2.5 \times 10^{-12})^a$	$-.11$ $(.01)^a$
	Replication	$-.08$ (.31)	.17 $(3.6 \times 10^{-03})^a$	$-.11$ (.09)	$-.06$ (.68)	$-.38$ $(1 \times 10^{-13})^a$	$-.07$ (.37)
High Past Year	Discovery	$-.15$ $(2.2 \times 10^{-04})^a$	.04 (.64)	$-.15$ $(1.4 \times 10^{-04})^a$	$-.04$ (.62)	$-.31$ $(4.1 \times 10^{-09})^a$	$-.09$ $(.04)^a$
	Replication	$-.10$ (.14)	.11 (.11)	$-.11$ (.09)	$-.14$ $(.02)^a$	$-.31$ $(3.2 \times 10^{-09})^a$	$-.09$ (.19)
High Lifetime	Discovery	$-.15$ $(1.6 \times 10^{-04})^a$	.06 (.55)	$-.14$ $(5.5 \times 10^{-04})^a$	$-.08$ (.43)	$-.30$ $(1 \times 10^{-07})^a$	$-.15$ $(3.4 \times 10^{-04})^a$
	Replication	$-.13$ $(.03)^a$	.02 (.76)	$-.09$ (.18)	$-.14$ $(.02)^a$	$-.22$ $(6 \times 10^{-05})^a$	$-.1$ (.16)
GF:R							
Current	Discovery	$-.09$ $(.04)^a$	.11 (.09)	$-.08$ (.05)	.01 (.99)	$-.38$ $(4.1 \times 10^{-17})^a$	$-.18$ $(4.8 \times 10^{-06})^a$
	Replication	$-.11$ (.08)	.19 $(8.0 \times 10^{-04})^a$	$-.11$ (.09)	$-.11$ (.08)	$-.30$ $(1.7 \times 10^{-08})^a$	$-.15$ $(.01)^a$
Low Past Year	Discovery	$-.07$ (.13)	.10 (.25)	$-.07$ (.08)	.02 (.75)	$-.37$ $(6.6 \times 10^{-16})^a$	$-.18$ $(7.6 \times 10^{-06})^a$
	Replication	$-.09$ (.22)	.20 $(2.4 \times 10^{-04})^a$	$-.10$ (.15)	$-.08$ (.28)	$-.32$ $(1.5 \times 10^{-09})^a$	$-.14$ $(.01)^a$
High Past Year	Discovery	$-.14$ $(7.3 \times 10^{-04})^a$	.08 (.50)	$-.09$ $(.02)^a$	$-.02$ (.79)	$-.30$ $(6.9 \times 10^{-08})^a$	$-.15$ $(2.5 \times 10^{-04})^a$
	Replication	$-.13$ $(.04)^a$	.15 $(.01)^a$	$-.09$ (.22)	$-.04$ (.50)	$-.25$ $(2.3 \times 10^{-06})^a$	$-.08$ (.30)
High Lifetime	Discovery	$-.13$ $(1.0 \times 10^{-03})^a$	.10 (.21)	$-.08$ $(.05)^a$	$-.05$ (.53)	$-.22$ $(.05)^a$	$-.14$ $(8.1 \times 10^{-04})^a$
	Replication	$-.19$ $(1.0 \times 10^{-03})^a$	.04 (.52)	$-.11$ (.12)	$-.12$ $(.05)^a$	$-.16$ $(.01)^a$	$-.12$ $(.04)^a$

Values are ρ (p). All p values were false discovery rate–corrected for multiple testing (family of tests with Table 4).

GAF:D/I, Global Assessment of Functioning Disability/Impairment Scale; GAF:S, Global Assessment of Functioning Social Scale; GF:R, Global Functioning: Role Scale; GF:S, Global Functioning: Social Scale; LV, latent variable.

[a]Significant value.

Table 4. Spearman's Correlation Analyses Between Latent Scores and Clinical Domains of Depressivity, Personality, and Quality of Life in the Discovery and Replication Samples

	Sample	LV2 — Sexual Abuse and Age		LV4 — Sexual/Physical Abuse and Sex		LV5 — Emotional Abuse/Neglect	
		Phenotypic Score	Brain Score	Phenotypic Score	Brain Score	Phenotypic Score	Brain Score
BDI							
Total score	Discovery	.11 (.01)[a]	−.08 (.84)	.18 (3.0×10^{-06})[a]	.09 (.25)	.48 (1.0×10^{-30})[a]	.19 (3.7×10^{-06})[a]
	Replication	.21 (1.2×10^{-04})[a]	−.08 (.32)	.3 (1.6×10^{-08})[a]	.18 (1.1×10^{-03})[a]	.48 (1.0×10^{-30})[a]	.14 (.02)[a]
NEO-FFI							
Neuroticism	Discovery	.15 (3.3×10^{-04})[a]	−.01 (.90)	.21 (1.2×10^{-07})[a]	.11 (.05)[a]	.44 (1.6×10^{-25})[a]	.13 (2.1×10^{-03})[a]
	Replication	.17 (2.2×10^{-03})[a]	.01 (.99)	.29 (7.2×10^{-08})[a]	.23 (1.9×10^{-05})[a]	.43 (1.5×10^{-25})[a]	.05 (.86)
Extraversion	Discovery	−.04 (.45)	.05 (.58)	−.08 (.05)[a]	.01 (.84)	−.3 (4.0×10^{-08})[a]	−.12 (.01)[a]
	Replication	−.18 (1.1×10^{-03})[a]	−.01 (.98)	−.21 (1.8×10^{-04})[a]	−.17 (3.9×10^{-03})[a]	−.33 (4.6×10^{-10})[a]	−.06 (.63)
Openness	Discovery	−.08 (.07)	−.02 (.81)	−.06 (.19)	−.04 (.61)	.02 (.50)	.06 (.27)
	Replication	.01 (.92)	−.02 (.69)	.01 (.98)	.01 (.88)	−.07 (.47)	.07 (.46)
Agreeableness	Discovery	−.16 (1.3×10^{-04})[a]	−.07 (.51)	−.07 (.11)	.06 (.50)	−.23 (.01)[a]	.02 (.50)
	Replication	−.11 (.11)	.02 (.73)	.02 (.84)	.01 (.99)	−.15 (.01)[a]	.01 (.99)
Conscientiousness	Discovery	−.17 (2.8×10^{-05})[a]	−.05 (.59)	−.10 (.01)[a]	.03 (.71)	−.33 (5.6×10^{-11})[a]	−.10 (.02)[a]
	Replication	−.30 (2.7×10^{-08})[a]	−.07 (.47)	−.2 (4.9×10^{-04})[a]	−.07 (.51)	−.32 (1.9×10^{-09})[a]	−.01 (.50)
WHOQOL-BREF							
Physical	Discovery	−.09 (.04)[a]	.03 (.68)	−.15 (1.2×10^{-04})[a]	−.07 (.54)	−.44 (1.7×10^{-25})[a]	−.12 (.01)[a]
	Replication	−.12 (.05)[a]	.10 (.18)	−.22 (8.4×10^{-05})[a]	−.15 (.01)[a]	−.45 (1.3×10^{-25})[a]	−.13 (.03)[a]
Psychosocial	Discovery	−.13 (2.0×10^{-03})[a]	.03 (.71)	−.20 (4.5×10^{-07})[a]	−.11 (.05)[a]	−.47 (1.2×10^{-29})[a]	−.12 (3.4×10^{-03})[a]
	Replication	−.21 (1.1×10^{-04})[a]	.05 (.80)	−.30 (2.9×10^{-08})[a]	−.19 (7.5×10^{-03})[a]	−.45 (1.3×10^{-29})[a]	−.11 (.09)
Social Relationships	Discovery	−.11 (.01)[a]	.07 (.52)	−.11 (.01)[a]	−.01 (.85)	−.41 (3.2×10^{-20})[a]	−.11 (.01)[a]
	Replication	−.09 (.18)	.07 (.41)	−.15 (.01)[a]	−.07 (.55)	−.41 (1.0×10^{-15})[a]	−.10 (.15)
Environment	Discovery	−.08 (.08)	.01 (.92)	−.05 (.54)	−.17 (2.4×10^{-05})[a]	−.45 (3.6×10^{-26})[a]	−.06 (.28)
	Replication	−.04 (.50)	.11 (.10)	−.06 (.68)	−.10 (.12)	−.36 (2.7×10^{-12})[a]	−.06 (.66)

Values are ρ (p). All p values were false discovery rate–corrected for multiple testing (family of tests with Table 3).

BDI, Beck Depression Inventory; LV, latent variable; NEO-FFI, Neuroticism-Extraversion-Openness Five-Factor Inventory; WHOQOL-BREF, World Health Organization Quality of Life Short Version.

[a]Significant value.

Discovery Sample. LV2: Sexual Abuse and Age. Positive associations were found between brain scores and population size at place of living (ρ = .28, p = .01), whereas negative correlations were detected between phenotypic scores and number of offspring, marital status, and years of education (ρ = −.29 to −.32, p = .01 to 1.78×10^{-03}).

LV4: Physical/Sexual Abuse and Sex. Phenotypic scores were negatively associated with years of education (ρ = −.29, p = .04).

LV5: Emotional Abuse/Neglect. Brain scores were negatively correlated with population at place of living (ρ = −.26, p = .04), while phenotypic scores were positively associated with lower education of the mother (ρ = .27, p = .03).

Replication Sample. No significant correlations were detected.

DISCUSSION

The goal of this study was a novel, comprehensive investigation of CT using a naturalistic and transdiagnostic machine learning approach. We performed SPLS analysis of CT-related phenotypic data and GMV in order to generate a transdiagnostic and multilayered CT model. We explored the clinical validity and sociodemographic ramifications of this CT model and confirmed the majority of our findings in a prospectively acquired replication sample.

We found 5 significant LVs, of which 3 (LV2, LV4, LV5) were more specifically linked to CT, while the other 2 LVs (LV1, LV3) represented predominantly age- and sex-related effects (Supplemental Results). Because all 3 CT-specific LVs did not contain any weighting for study group, they can be regarded as transdiagnostic signatures.

The highly parsimonious signature of LV2 links sexual abuse in younger individuals to GMV alterations along the prefronto-thalamo-cerebellar axis. Further GMV variation associated with CT involved the temporal and angular gyrus as well as the basal ganglia and the cuneus region. While the PFC has been among the most well-established GMV correlates of CT, the other brain regions in this signature have not yet been consistently associated with CT (20,52,53). Instead, the prefronto-thalamo-cerebellar axis has been implicated in various aspects of (social) cognition (54,55) and associative learning (56). Additionally, it has been proposed as a key system involved in psychiatric disorders, including affective (57,58) and nonaffective psychoses (59–61). Hence, the LV2 signature may point to disease-connected alterations in the prefronto-thalamo-cerebellar axis associated with sexual trauma experiences.

In LV4, a pattern of sexual and physical abuse was associated with a dense GMV signature involving the postcentral gyrus, hippocampus, and PFC (20) as well as the limbic brain regions associated with emotional learning and social cognitive processes (62,63). This signature was inversely expressed in male and female individuals. This supports previous studies, which reported contrary volumetric and connectivity changes in the PFC, hippocampus, amygdala, and anterior cingulate

cortex for male and female individuals after exposure to CT (44). Moreover, the LV4 trauma signature aligns with a recent study reporting an interaction between CT and sex on hippocampal volume, which could be predicted by neglect in male individuals and abuse in female individuals (45). This evidence emphasizes that the limbic system and key CT-associated regions are inversely affected by abuse in men and women and highlights the paramount need for further sex-specific CT research and sex-tailored therapeutic approaches in traumatized individuals.

The patterns observed in LV2 and LV4 further reflect previous findings concerning brain development, which showed differential developmental trajectories for female and male brains (64,65). The brain signature of LV2 comprises specifically the medial PFC, i.e., a cortical region that fully develops during adolescence (64), while the LV4 signature covers the temporal, prefrontal, and occipital lobes—regions in which sex has been shown to have a nonlinear relationship with age (65). Thus, sex exerts a modulating influence on cortical development from childhood to adulthood. The strong covariation of the age and sex effects on CT signatures might be explained in a developmental framework not only in which male and female individuals react differently to trauma, but also in which male and female individuals' brains may develop differently as a result of CT.

LV5 links emotional abuse and neglect to a brain pattern consisting of diverse GMV changes. First, emotional trauma is connected to brain regions responsible for sensory processing via the postcentral gyrus and occipital lobe (66,67). Second, associations with the DLPFC, insula, and cingulate gyrus relate emotional trauma to key brain systems subserving emotional processing (68–70), memory formation (71,72), and risk for psychiatric disorders (73–75). These findings support the hypothesis that trauma experience is connected to sensory and perceptive dysregulations, which could also be accessed therapeutically (76–78).

All 3 CT-specific signatures yielded significant correlations with clinical measures, which were largely validated in the replication sample. The phenotypic scores of the age-dependent sexual abuse signature (LV2) revealed strong connections to an impaired clinical phenotype in the discovery and replication samples. The brain scores appeared dissociated from that in both populations, yielding no significant associations in the discovery sample and positive associations with GAF and GF in the replication sample. One possible interpretation might be that the signature of LV2 had been influenced by unaccounted resilience dynamics, in which neurobiological adaptations compensate for the phenomenological trauma load, thus maintaining levels of functioning (79,80). Additional analyses revealed a positive correlation between LV2 brain scores and population size at the place of living as well as inverse associations between LV2 phenotypic scores and number of offspring, marital status, and years of education in the discovery sample. These findings suggest a possible connection between resilience-conferring brain adaptations and urbanicity as well as higher sexual trauma loadings and social (offspring, marriage) and educational status. Moreover, LV4 and LV5 revealed the most extensive significant associations with functioning, depressivity, personality domains,

Neuroanatomical Signatures of Childhood Adversity

and quality of life in the discovery and replication samples. Both trauma and brain scores of LV4 and LV5 were significantly correlated with lower levels of social and role functioning, more pronounced symptom severity, increased impairment, and higher levels of depressivity and reduced quality of life. Additionally, we found a strong connection between individual trauma loads and higher levels of neuroticism as well as between lower levels of extraversion, conscientiousness, agreeableness, and openness. Finally, phenotypic loading of LV4 was associated with lower educational status, whereas LV5 loading was connected to a less urban environment (phenotypic scores) and lower maternal educational status (brain scores). These findings confirm and extend the current body of literature on the negative clinical implications and complex sociodemographic constellations of CT. It has been well established that CT has a broad negative impact on mental health, including a higher vulnerability for mental disorders, the presence of maladaptive personality traits, and decreased psychosocial functioning and quality of life (21). Nonetheless, beyond these general associations, very few studies have investigated more domain-specific aspects of CT (81–83). Thus, our results provide more extensive evidence for a differential neurobiological, clinical, and sociodemographic imprint of CT. Moreover, the connection between the CT signatures and the presence of vulnerability-conferring personality domains provides novel neurobiological evidence for the long-standing and still controversially discussed hypothesis that adverse childhood experiences lead to the development of dysfunctional personality structures (9,84,85).

Because 70% of these clinical associations were successfully validated in the replication sample and 20 additional significant clinical correlations (18 on the brain level) emerged, the multilayered SPLS trauma model appears robustly generalizable at both the phenotypic and neuroanatomical levels. Furthermore, it emphasizes the validity and paramount clinical relevance of the multidimensional CT concept across a broad diagnostic spectrum in two large-scale international samples of young adults and adolescent individuals.

Potential limitations of the study need to be considered. Some of the brain variance might be attributed to psychopharmacological treatment. Yet, our transdiagnostic study design should provide a robust framework against such confounders. Moreover, some LV signatures were partly associated with magnetic resonance imaging data quality, albeit the impact was minimal. Additional SPLS analyses further supported the main results (Supplemental Results). Furthermore, the associative nature of our results should not lead to causal assumptions. Directed network analysis and supervised machine learning could help elucidate the inner workings of CT and assess their predictive value for psychiatric disorders.

To our knowledge, this is the first study that investigated CT in a transdiagnostic sample of young adults using a data-driven machine learning approach and a comprehensive, multidimensional framework for CT operationalization. Our novel approach confirms that CT is composed of distinct phenotypic-neuroanatomical dimensions that may have complex ramifications into clinically relevant phenotypes. We

found CT signatures of sexual, physical, and emotional trauma with distinct neuroanatomic correlates in the prefronto-thalamo-cerebellar, limbic, and sensory networks. Furthermore, sex-dependent combined sexual and physical abuse as well as emotional trauma appeared to be specifically predictive of relevant clinical state and trait phenotypes, whereas the age-dependent sexual abuse signature may have been further influenced by neurobiological resilience pathways and interacted with modulating factors such as urbanicity, education, and family status. As these results were largely validated in a large replication sample, our findings demonstrate that machine learning tools can generate new and generalizable insights into complex human phenomena such as CT and might help to develop superior treatments targeting CT and its psychiatric consequences at short- to long-term time scales.

ARTICLE INFORMATION

From the Department of Psychiatry and Psychotherapy (DP, AR, DBD, LAA, JE, RS, LK-I, OFO, MSD, RP, PF, NK) and Department of Radiology (MP), Ludwig Maximilian University of Munich; International Max Planck Research School for Translational Psychiatry (DP, OFO, NK), Max Planck Society; Max Planck Institute of Psychiatry (RP), Max Planck Schools; Department of Diagnostic and Interventional Neuroradiology (DH), Technical University of Munich, Munich; Max Planck School of Cognition (RS), Max Planck Schools, Leipzig; Department of Psychiatry and Psychotherapy (LK-I, TH, JK, SR), University of Cologne, Faculty of Medicine and University Hospital Cologne, Cologne; Department of Psychiatry and Psychotherapy (FS-L, EM), Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf; and Department of Psychiatry and Psychotherapy (RL, UD), University of Münster, Münster, Germany; Department of Education, Psychology and Communication (LAA); and Group of Psychiatric Neuroscience (GP, GB, AB), Department of Basic Medical Science, Neuroscience and Sense Organs, University of Bari Aldo Moro, Bari; and Department of Neurosciences and Mental Health (PB), Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, University of Milan, Milan, Italy; School of Psychology (KC, RU, SJW) and Institute for Mental Health (RU), University of Birmingham; and Department of Psychology (KC), School of Life and Health Sciences, Aston University, Birmingham, United Kingdom; Department of Psychiatry (RKRS), University of Turku, Turku, Finland; Melbourne Neuropsychiatry Centre (CP) and Centre for Youth Mental Health (SJW), University of Melbourne; Melbourne Health (CP); and Orygen (SJW), the National Centre of Excellence for Youth Mental Health, Melbourne, Victoria, Australia; and the Neuropsychiatry and Brain Imaging Group (SB), Department of Psychiatry, University of Basel, Basel, Switzerland.

PRONIA Consortium**: Department of Psychiatry and Psychotherapy, Ludwig-Maximilian-University, Munich, Bavaria, Germany:** Mark Sen Dong, M.Sc.; Anne Erkens; Eva Gussmann, M.Sc.; Shalaila Haas, Ph.D.; Alkomiet Hasan, M.D.; Claudius Hoff, M.D.; Ifrah Khanyaree, B.Sc.; Aylin Melo, M.Sc.; Susanna Muckenhuber-Sternbauer, M.D.; Janis Köhler, Ömer Faruk Öztürk, M.D.; Nora Penzel, M.Sc.; Adrian Rangnick, B.Sc.; Sebastian von Saldern, M.D.; Rachele Sanfelici, M.Sc.; Moritz Spangemacher; Ana Tupac, M.Sc.; Maria Fernanda Urquijo, M.Sc.; Johanna Weiske, M.Sc.; Julian Wenzel, M.Sc.; and Antonia Wosgien. **University of Cologne, North Rhineland–Westphalia, Germany:** Linda Betz, M.Sc.; Karsten Blume; Mauro Seves, M.Sc.; Nathalie Kaiser, M.Sc.; Thorsten Lichtenstein, M.D.; and Christiane Woopen, M.D. **Psychiatric University Hospital, University of Basel, Basel, Switzerland:** Christina Andreou, M.D., Ph.D.; Laura Egloff, Ph.D.; Fabienne Harrisberger, Ph.D.; Claudia Lenz, Ph.D.; Letizia Leanza, M.Sc.; Amatya Mackintosh, M.Sc.; Renata Smieskova, Ph.D.; Erich Studerus, Ph.D.; Anna Walter, M.D.; and Sonja Widmayer, M.Sc. **Institute for Mental Health, University of Birmingham, Birmingham, United Kingdom:** Chris Day, B.Sc.; Sian Lowri Griffiths, Ph.D.; Mariam Iqbal, B.Sc.; Mirabel Pelton, M.Sc.; Pavan Mallikarjun, M.B.B.S., D.P.M., M.R.C.Psych., Ph.D.; Alexandra Stainton, M.Sc.; and Ashleigh Lin, Ph.D. **Department of Psychiatry, University of Turku, Turku, Finland:** Alexander Denissoff, M.D.; Anu Ellilä, R.N.; Tiina From, M.Sc.; Markus Heinimaa, M.D., Ph.D.; Tuula Ilonen, Ph.D.; Päivi Jalo, R.N.; Heikki Laurikainen, M.D.; Maarit Lehtinen, R.N.; Antti Luutonen, B.A.; Akseli Mäkela, B.A.; Janina Paju, M.Sc.; Henri Pesonen, Ph.D.; Reetta-Liina Armio (Säilä), M.D.; Elina Sormunen, M.D.; Anna Toivonen, M.Sc.; and Otto Turtonen, M.D. **General Electric Global Research Inc, Munich, Germany:** Ana Beatriz Solana, Ph.D.; Manuela Abraham, M.B.A.; Nicolas Hehn, Ph.D.; and Timo Schirmer, Ph.D. **Workgroup of Paolo Brambilla, M.D., Ph.D.; University of Milan, Milan, Italy:** Department of Neuroscience and Mental Health, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, University of Milan, Milan, Italy: Carlo Altamura, M.D.; Marika Belleri, Psych.D.; Francesca Bottinelli, Psych.D.; Adele Ferro, Psych.D., Ph.D.; and Marta Re, Ph.D. Programma2000, Niguarda Hospital, Milan: Emiliano Monzani, M.D.; Mauro Percudani, M.D.; and Maurizio Sberna, M.D. San Paolo Hospital, Milan: Armando D'Agostino, M.D.; and Lorenzo Del Fabro, M.D. Villa San Benedetto Menni, Albese con Cassano: Giampaolo Perna, M.D.; Maria Nobile, M.D., Ph.D.; and Alessandra Alciati, M.D. **Workgroup of Paolo Brambilla, M.D., Ph.D.; University of Udine, Udine, Italy:** Department of Medical Area, University of Udine: Matteo Balestrieri, M.D.; Carolina Bonivento, Psych.D., Ph.D.; Giuseppe Cabras, Ph.D.; and Franco Fabbro, M.D., Ph.D. IRCCS

Scientific Institute "E. Medea," Polo FVG, Udine: Marco Garzitto, Psych.D., Ph.D.; and Sara Piccin, PsychD, Ph.D.

Address correspondence to Nikolaos Koutsouleris, M.D., Department of Psychiatry and Psychotherapy, Ludwig-Maximilian-University, Nussbaumstraße 7, D-80336 Munich, Germany; E-mail: nikolaos.koutsouleris@med.uni-muenchen.de.

REFERENCES

1. Arias I, Leeb RT, Melanson C, Paulozzi LJ, Simon TR (2008): Child Maltreatment Surveillance: Uniform Definitions for Public Health and Recommended Data Elements. Atlanta, GA: U.S. Department of Health and Human Services and Centers for Disease Control and Prevention.
2. Trottier K, MacDonald DE (2017): Update on psychological trauma, other severe adverse experiences and eating disorders: State of the research and future research directions. Curr Psychiatry Rep 19:45.
3. Walsh K, McLaughlin KA, Hamilton A, Keyes KM (2017): Trauma exposure, incident psychiatric disorders, and disorder transitions in a longitudinal population representative sample. J Psychiatr Res 92:212–218.
4. Isvoranu AM, van Borkulo CD, Boyette LL, Wigman JT, Vinkers CH, Borsboom D, et al. (2017): A network approach to psychosis: Pathways between childhood trauma and psychotic symptoms. Schizophr Bull 43:187–196.
5. Xie P, Wu K, Zheng Y, Guo Y, Yang Y, He J, et al. (2018): Prevalence of childhood trauma and correlations between childhood trauma, suicidal ideation, and social support in patients with depression, bipolar disorder, and schizophrenia in southern China. J Affect Disord 228:41–48.
6. Andrianarisoa M, Boyer L, Godin O, Brunel L, Bulzacka E, Aouizerate B, et al. (2017): Childhood trauma, depression and negative symptoms are independently associated with impaired quality of life in schizophrenia. Results from the national FACE-SZ cohort. Schizophr Res 185:173–181.
7. Kraan T, van Dam DS, Velthorst E, de Ruigh EL, Nieman DH, Durston S, et al. (2015): Childhood trauma and clinical outcome in patients at ultra-high risk of transition to psychosis. Schizophr Res 169:193–198.
8. Pos K, Boyette LL, Meijer CJ, Koeter M, Krabbendam L, de Haan L, et al. (2016): The effect of childhood trauma and Five-Factor Model personality traits on exposure to adult life events in patients with psychotic disorders. Cogn Neuropsychiatry 21:462–474.
9. Li X, Wang Z, Hou Y, Wang Y, Liu J, Wang C (2014): Effects of childhood trauma on personality in a sample of Chinese adolescents. Child Abuse Negl 38:788–796.
10. Ahmed-Leitao F, Spies G, van den Heuvel L, Seedat S (2016): Hippocampal and amygdala volumes in adults with posttraumatic stress disorder secondary to childhood abuse or maltreatment: A systematic review. Psychiatry Res Neuroimaging 256:33–43.
11. Carballedo A, Lisiecka D, Fagan A, Saleh K, Ferguson Y, Connolly G, et al. (2012): Early life adversity is associated with brain changes in subjects at family risk for depression. World J Biol Psychiatry 13:569–578.
12. Chaney A, Carballedo A, Amico F, Fagan A, Skokauskas N, Meaney J, et al. (2014): Effect of childhood maltreatment on brain structure in adult patients with major depressive disorder and healthy participants. J Psychiatry Neurosci 39:50–59.
13. Cancel A, Comte M, Truillet R, Boukezzi S, Rousseau PF, Zendjidjian XY, et al. (2015): Childhood neglect predicts disorganization in schizophrenia through grey matter decrease in dorsolateral prefrontal cortex. Acta Psychiatr Scand 132:244–256.
14. Andersen SL, Tomada A, Vincow ES, Valente E, Polcari A, Teicher MH (2008): Preliminary evidence for sensitive periods in the effect of childhood sexual abuse on regional brain development. J Neuropsychiatry Clin Neurosci 20:292–301.
15. Van Dam NT, Rando K, Potenza MN, Tuit K, Sinha R (2014): Childhood maltreatment, altered limbic neurobiology, and substance use relapse

Neuroanatomical Signatures of Childhood Adversity

severity via trauma-specific reductions in limbic gray matter volume. JAMA Psychiatry 71:917–925.

16. Baker LM, Williams LM, Korgaonkar MS, Cohen RA, Heaps JM, Paul RH (2013): Impact of early vs. late childhood early life stress on brain morphometrics. Brain Imaging Behav 7:196–203.

17. Aas M, Navari S, Gibbs A, Mondelli V, Fisher HL, Morgan C, et al. (2012): Is there a link between childhood trauma, cognition, and amygdala and hippocampus volume in first-episode psychosis? Schizophr Res 137:73–79.

18. Kuhn M, Scharfenort R, Schumann D, Schiele MA, Munsterkotter AL, Deckert J, et al. (2016): Mismatch or allostatic load? Timing of life adversity differentially shapes gray matter volume and anxious temperament. Soc Cogn Affect Neurosci 11:537–547.

19. Baldacara L, Zugman A, Araujo C, Cogo-Moreira H, Lacerda AL, Schoedl A, et al. (2014): Reduction of anterior cingulate in adults with urban violence-related PTSD. J Affect Disord 168:13–20.

20. Paquola C, Bennett MR, Lagopoulos J (2016): Understanding heterogeneity in grey matter research of adults with childhood maltreatment-A meta-analysis and review. Neurosci Biobehav Rev 69:299–312.

21. Herzog JI, Schmahl C (2018): Adverse childhood experiences and the consequences on neurobiological, psychosocial, and somatic conditions across the lifespan. Front Psychiatry 9:420.

22. Popovic D, Schmitt A, Kaurani L, Senner F, Papiol S, Malchow B, et al. (2019): Childhood trauma in schizophrenia: Current findings and research perspectives. Front Neurosci 13:274.

23. Oral R, Ramirez M, Coohey C, Nakada S, Walz A, Kuntz A, et al. (2016): Adverse childhood experiences and trauma informed care: The future of health care. Pediatr Res 79:227–233.

24. Logothetis NK (2008): What we can do and what we cannot do with fMRI. Nature 453:869–878.

25. Kamitani Y, Tong F (2005): Decoding the visual and subjective contents of the human brain. Nat Neurosci 8:679–685.

26. Kriegeskorte N, Cusack R, Bandettini P (2010): How does an fMRI voxel sample the neuronal activity pattern: Compact-kernel or complex spatiotemporal filter? Neuroimage 49:1965–1976.

27. Woo CW, Chang LJ, Lindquist MA, Wager TD (2017): Building better biomarkers: Brain models in translational neuroimaging. Nat Neurosci 20:365–377.

28. Davatzikos C (2004): Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. Neuroimage 23:17–20.

29. Yahata N, Kasai K, Kawato M (2017): Computational neuroscience approach to biomarkers and treatments for mental disorders. Psychiatry Clin Neurosci 71:215–237.

30. Freedman R, Lewis DA, Michels R, Pine DS, Schultz SK, Tamminga CA, et al. (2013): The initial field trials of DSM-5: New blooms and old thorns. Am J Psychiatry 170:1–5.

31. Khan A, McCormack HC, Bolger EA, McGreenery CE, Vitaliano G, Polcari A, et al. (2015): Childhood maltreatment, depression, and suicidal ideation: Critical importance of parental and peer emotional abuse during developmental sensitive periods in males and females. Front Psychiatry 6:42.

32. Whittle S, Simmons JG, Dennison M, Vijayakumar N, Schwartz O, Yap MB, et al. (2014): Positive parenting predicts the development of adolescent brain structure: A longitudinal study. Dev Cogn Neurosci 8:7–17.

33. Jollans L, Whelan R (2018): Neuromarkers for mental disorders: Harnessing population neuroscience. Front Psychiatry 9:242.

34. Monteiro JM, Rao A, Shawe-Taylor J, Mourao-Miranda J; Alzheimer's Disease Initiative (2016): A multiple hold-out framework for sparse partial least squares. J Neurosci Methods 271:182–194.

35. Koutsouleris N, Kambeitz-Ilankovic L, Ruhrmann S, Rosen M, Ruef A, Dwyer DB, et al. (2018): Prediction models of functional outcomes for individuals in the clinical high-risk state for psychosis or with recent-onset depression: A multimodal, multisite machine learning analysis. JAMA Psychiatry 75:1156–1172.

36. Bernstein DP, Ahluvalia T, Pogge D, Handelsman L (1997): Validity of the Childhood Trauma Questionnaire in an adolescent psychiatric population. J Am Acad Child Adolesc Psychiatry 36:340–348.

37. Fink LA, Bernstein D, Handelsman L, Foote J, Lovejoy M (1995): Initial reliability and validity of the childhood trauma interview: A new multidimensional measure of childhood interpersonal trauma. Am J Psychiatry 152:1329–1335.

38. Bernstein DP, Fink L, Handelsman L, Foote J, Lovejoy M, Wenzel K, et al. (1994): Initial reliability and validity of a new retrospective measure of child abuse and neglect. Am J Psychiatry 151:1132–1136.

39. American Psychiatric Association (2000): Diagnostic and Statistical Manual of Mental Disorders, 4th ed, text rev Washington, DC: American Psychiatric Association.

40. Cornblatt BA, Auther AM, Niendam T, Smith CW, Zinberg J, Bearden CE, et al. (2007): Preliminary findings for two new measures of social and role functioning in the prodromal phase of schizophrenia. Schizophr Bull 33:688–702.

41. Beck AT, Steer RA (1984): Internal consistencies of the original and revised Beck Depression Inventory. J Clin Psychol 40:1365–1367.

42. Skevington SM, Lotfy M, O'Connell KA, Group W (2004): The World Health Organization's WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 13:299–310.

43. Costa PT, McCrae RR (1992): Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) Professional Manual. Odessa, FL: Psychological Assessment Resources.

44. Helpman L, Zhu X, Suarez-Jimenez B, Lazarov A, Monk C, Neria Y (2017): Sex Differences in trauma-related psychopathology: A critical review of neuroimaging literature (2014-2017). Curr Psychiatry Rep 19:104.

45. Teicher MH, Anderson CM, Ohashi K, Khan A, McGreenery CE, Bolger EA, et al. (2018): Differential effects of childhood neglect and abuse during sensitive exposure periods on male and female hippocampus. Neuroimage 169:443–452.

46. Zou H, Hastie T (2005): Regularization and variable selection via the elastic net. J R Stat Soc Series B Stat Methodol 67:301–320.

47. Ruschhaupt M, Huber W, Poustka A, Mansmann U (2004): A compendium to ensure computational reproducibility in high-dimensional classification tasks. Stat Appl Genet Mol Biol 3:Article37.

48. Dwyer DB, Falkai P, Koutsouleris N (2018): Machine learning approaches for clinical psychology and psychiatry. Annu Rev Clin Psychol 14:91–118.

49. Koutsouleris N, Meisenzahl EM, Borgwardt S, Riecher-Rossler A, Frodl T, Kambeitz J, et al. (2015): Individualized differential diagnosis of schizophrenia and mood disorders using neuroanatomical biomarkers. Brain 138:2059–2073.

50. Dukart J, Schroeter ML, Mueller K; Alzheimer's Disease Neuroimaging Initiative (2011): Age correction in dementia–Matching to a healthy brain. PLoS One. 6:e22193.

51. Benjamini Y, Hochberg Y (1995): Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 57:289–300.

52. Lu S, Xu R, Cao J, Yin Y, Gao W, Wang D, et al. (2019): The left dorsolateral prefrontal cortex volume is reduced in adults reporting childhood trauma independent of depression diagnosis. J Psychiatr Res 112:12–17.

53. Heyn SA, Keding TJ, Ross MC, Cisler JM, Mumford JA, Herringa RJ (2019): Abnormal prefrontal development in pediatric posttraumatic stress disorder: A longitudinal structural and functional magnetic resonance imaging study. Biol Psychiatry Cogn Neurosci Neuroimaging 4:171–179.

54. Diamond A (2000): Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. Child Dev 71:44–56.

55. Van Overwalle F, Marien P (2016): Functional connectivity between the cerebrum and cerebellum in social cognition: A multi-study analysis. Neuroimage 124:248–255.

56. Taylor JA, Ivry RB (2014): Cerebellar and prefrontal cortex contributions to adaptation, strategies, and reinforcement learning. Prog Brain Res 210:217–253.

57. Samara Z, Evers EAT, Peeters F, Uylings HBM, Rajkowska G, Ramaekers JG, et al. (2018): Orbital and medial prefrontal cortex functional connectivity of major depression vulnerability and disease. Biol Psychiatry Cogn Neurosci Neuroimaging 3:348–357.

58. Bersani FS, Minichino A, Bernabei L, Spagnoli F, Corrado A, Vergnani L, et al. (2017): Prefronto-cerebellar tDCS enhances neurocognition in euthymic bipolar patients. Findings from a placebo-controlled neuropsychological and psychophysiological investigation. J Affect Disord 209:262–269.

59. Andreasen NC, Paradiso S, O'Leary DS (1998): "Cognitive dysmetria" as an integrative theory of schizophrenia: A dysfunction in cortical-subcortical-cerebellar circuitry? Schizophr Bull 24:203–218.

60. Lungu O, Barakat M, Laventure S, Debas K, Proulx S, Luck D, et al. (2013): The incidence and nature of cerebellar findings in schizophrenia: A quantitative review of fMRI literature. Schizophr Bull 39:797–806.

61. Andreasen NC, Nopoulos P, Magnotta V, Pierson R, Ziebell S, Ho BC (2011): Progressive brain change in schizophrenia: A prospective longitudinal study of first-episode schizophrenia. Biol Psychiatry 70:672–679.

62. Rolls ET (2015): Limbic systems for emotion and for memory, but no single limbic system. Cortex 62:119–157.

63. Catani M, Dell'acqua F, Thiebaut de Schotten M (2013): A revised limbic system model for memory, emotion and behaviour. Neurosci Biobehav Rev 37:1724–1737.

64. Gogtay N, Thompson PM (2010): Mapping gray matter development: Implications for typical development and vulnerability to psychopathology. Brain Cogn 72:6–15.

65. Gennatas ED, Avants BB, Wolf DH, Satterthwaite TD, Ruparel K, Ciric R, et al. (2017): Age-related effects and sex differences in gray matter density, volume, mass, and cortical thickness from childhood to young adulthood. J Neurosci 37:5065–5073.

66. Nauhaus I, Nielsen KJ (2014): Building maps from maps in primary visual cortex. Curr Opin Neurobiol 24:1–6.

67. Brecht M (2017): The body model theory of somatosensory cortex. Neuron 94:985–992.

68. Dixon ML, Thiruchselvam R, Todd R, Christoff K (2017): Emotion and the prefrontal cortex: An integrative review. Psychol Bull 143:1033–1081.

69. Gasquoine PG (2014): Contributions of the insula to cognition and emotion. Neuropsychol Rev 24:77–87.

70. Vogt BA (2014): Submodalities of emotion in the context of cingulate subregions. Cortex 59:197–202.

71. Brunoni AR, Vanderhasselt MA (2014): Working memory improvement with non-invasive brain stimulation of the dorsolateral prefrontal cortex: A systematic review and meta-analysis. Brain Cogn 86:1–9.

72. Leech R, Sharp DJ (2014): The role of the posterior cingulate cortex in cognition and disease. Brain 137:12–32.

73. Zhou Y, Fan L, Qiu C, Jiang T (2015): Prefrontal cortex and the dysconnectivity hypothesis of schizophrenia. Neurosci Bull 31:207–219.

74. Namkung H, Kim SH, Sawa A (2017): The insula: An underestimated brain area in clinical neuroscience, psychiatry, and neurology. Trends Neurosci 40:200–207.

75. Downar J, Blumberger DM, Daskalakis ZJ (2016): The neural crossroads of psychiatric illness: An emerging target for brain stimulation. Trends Cogn Sci 20:107–120.

76. Price CJ, Hooven C (2018): Interoceptive awareness skills for emotion regulation: Theory and approach of mindful awareness in body-oriented therapy (MABT). Front Psychol 9:798.

77. Clancy KJ, Albizu A, Schmidt NB, Li W (2020): Intrinsic sensory disinhibition contributes to intrusive re-experiencing in combat veterans. Sci Rep 10:936.

78. Iyadurai L, Visser RM, Lau-Zhu A, Porcheret K, Horsch A, Holmes EA, et al. (2019): Intrusive memories of trauma: A target for research bridging cognitive science and its clinical application. Clin Psychol Rev 69:67–82.

79. Teicher MH, Samson JA, Anderson CM, Ohashi K (2016): The effects of childhood maltreatment on brain structure, function and connectivity. Nat Rev Neurosci 17:652–666.

80. Gupta A, Love A, Kilpatrick LA, Labus JS, Bhatt R, Chang L, et al. (2017): Morphological brain measures of cortico-limbic inhibition related to resilience. J Neurosci Res 95:1760–1775.

81. Norman RE, Byambaa M, De R, Butchart A, Scott J, Vos T (2012): The long-term health consequences of child physical abuse, emotional abuse, and neglect: A systematic review and meta-analysis. PLoS Med 9:e1001349.

82. Upthegrove R, Chard C, Jones L, Gordon-Smith K, Forty L, Jones I, et al. (2015): Adverse childhood events and psychosis in bipolar affective disorder. Br J Psychiatry 206:191–197.

83. Thompson AD, Nelson B, Yuen HP, Lin A, Amminger GP, McGorry PD, et al. (2014): Sexual trauma increases the risk of developing psychosis in an ultra high-risk "prodromal" population. Schizophr Bull 40: 697–706.

84. de Carvalho HW, Pereira R, Frozi J, Bisol LW, Ottoni GL, Lara DR (2015): Childhood trauma is associated with maladaptive personality traits. Child Abuse Negl 44:18–25.

85. Baryshnikov I, Joffe G, Koivisto M, Melartin T, Aaltonen K, Suominen K, et al. (2017): Relationships between self-reported childhood traumatic experiences, attachment style, neuroticism and features of borderline personality disorders in patients with mood disorders. J Affect Disord 210:82–89.

Traces of Trauma – A Multivariate Pattern Analysis of Childhood Trauma, Brain Structure and Clinical Phenotypes

Supplementary Information

Supplementary Methods

PRONIA study design

The entire PRONIA cohort consists of a discovery sample for model generation and a replication sample for model validation. The 1076 study participants (discovery sample, n=649, replication sample, n=427) analyzed in the present study were recruited following the standardized recruitment and ascertainment protocol (Figure S1, Table S1) of the PRONIA study (Personalized Prognostic Tools for Early Psychosis Management, https://www.pronia.eu/). The observational study protocol involved follow-up examinations every three months after the index ascertainment and was implemented by the ten PRONIA sites. As described in the main text, only data from the baseline observation was used for our analysis. Upon study enrolment, the participants were pseudonymized twice, locally at each site and centrally at the level of the PRONIA portal. The PRONIA portal consists of a multi-user database hosting the clinical and neurocognitive information, and defaced MR images obtained from the study participants. The data are organized into digital questionnaires, visits, and cases. The portal provides the case managers with a controlled web-based interface to enter and upload the different data into the respective questionnaires. Furthermore, the PRONIA consortium has implemented a PRONIA@home mobile device interface that allows the study participants to securely log into the portal and fill out the self-rating questionnaires of given visit. Upon completion of the data entry across all questionnaires of a given visit, the data are checked by an automatic quality control procedure which executes approximately 1600 data integrity and dependency rules. These rules include 1) basic checking of missing data and data ranges, 2) checking of dependency within one questionnaire, 3) dependencies between two questionnaires within one visit, and 4) dependencies between two consecutive visits (such as consistency of dates). Detected errors are fed back to the respective case

managers allowing for a manual correction of the respective issues. This process is re-iterated until the quality of the clinical questionnaires in the given visit is sufficient for the entire visit to be locked.

The clinical data analyzed in the present study consist of the quality-checked and locked information of the study participants recruited until the 1st of May 2016, who had 1) received a structural MRI scan at baseline (Table S2), and 2) could be assessed using the global functioning scales (GAF, GF) at least on one occasion between the month-3 and month-12 follow-up visits (Figure S1).

A comprehensive battery of ascertainment tools was used within a longitudinal observational study design to generate a multi-modal phenotypic profile of each study participant (Figure S1). The clinical part of the battery compiled questionnaires that capture sociodemographic, somatic, environmental, diagnostic, psychopathological, functional and quality-of-life related variables in the PRONIA study population. This battery was complemented by multi-domain neurocognitive and neuroimaging examinations as well as blood sampling for later genetic characterization, which were carried out at the baseline and 9-month follow-up timepoints (see initial PRONIA publication by Koutsouleris et al. (1) for further information on the PRONIA study).

Detailed sample determination

From the PRONIA discovery sample, data from 264 healthy controls (HC), 124 patients with clinical high-risk states for psychosis (CHR), 132 patients with recent-onset of psychosis (ROP) and 129 patients with recent-onset of depression (ROD), recruited at seven sites in five countries (Munich, Basel, Cologne, Birmingham, Turku, Udine and Milan), were obtained for this study (Table 1 in Main Text). From the PRONIA replication sample, data from 135 healthy controls (HC), 104 patients with clinical high-risk states for psychosis (CHR), 92 patients with recent-onset of psychosis (ROP) and 96 patients with recent-onset of depression (ROD), recruited at ten sites in five countries (discovery sample sites and additionally in Muenster, Duesseldorf and Bari), were obtained for this study (Table S6). All adult participants provided their written informed consent prior to study inclusion. Minor participants (defined at all sites as those younger than 18 years) provided written informed assent and their

guardians, written informed consent. The study was registered at the German Clinical Trials Register (DRKS00005042) and approved by the local research ethics committees in each location. General inclusion criteria were age between 15 and 40 years, sufficient language skills for participation as well as capacity to provide informed consent/assent. General exclusion criteria were an IQ below 70, current or past head trauma with loss of consciousness (> 5 minutes), current or past known neurological or somatic disorders potentially affecting the structure or functioning of the brain, current or past alcohol dependence, or polysubstance dependence within the past six months, and any medical indication against MRI. The CHR state was defined by either 1) cognitive disturbances (COGDIS) criteria assessed using the Schizophrenia Proneness Instrument (SPI-A, (2)) and/or 2) ultra-high-risk (UHR) criteria for psychosis based on the Structured Interview for Prodromal Syndromes (SIPS, (3)). CHR exclusion criteria were 1) antipsychotic medication for > 30 days (cumulative number of days) at or above minimum dosage of the "1st episode psychosis" range of DGPPN S3 ("Deutsche Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde e. V.", German Association for Psychiatry, Psychotherapy and Psychosomatics) guidelines (4) and 2) any intake of antipsychotic medication within the past 3 months before clinical baseline assessments at or above minimum dosage of the "1st episode psychosis" range of DGPPN S3 guidelines (4). ROP participants had to meet the following criteria: 1) DSM-IV-TR affective or non-affective psychotic episode (lifetime), 2) criteria for DSM-IV-TR (Diagnostic and Statistical Manual of Mental Disorders, Text Revision) affective or non-affective psychotic episode fulfilled within past 3 months and 3) onset of psychosis within past 24 months. ROP exclusion criterion was antipsychotic medication longer than 90 days (cumulative number of days) with a daily dose rate at or above minimum dosage in the "1st episode psychosis" range of the DGPPN S3 guideline (4). ROD patients were identified by 1) DSM-IV-TR major depressive episode (lifetime), 2) major depressive disorder criteria fulfilled within past three months and 3) duration of first depressive episode no longer than 24 months. Specific ROD exclusion criteria were: 1) more than 1 major depressive episode, 2) antipsychotic medication for > 30 days (cumulative number of days) at or above minimum dosage of the "1st episode psychosis" range of the DGPPN S3 guidelines and 3) any intake of antipsychotic medication within the past 3 months before psychopathological

baseline assessments at or above minimum dosage of the "1st episode psychosis" range of the DGPPN

S3 guidelines (4).

MRI harmonization and data acquisition

When setting up the PRONIA study, we decided to generate an MRI database that would represent the

MR scanner sequence heterogeneity encountered in clinical real-world. The aim of this strategy was

to strengthen the generalizability and clinical applicability of the predictive models developed by our

machine learning analyses. Thus, we agreed on a minimal harmonization protocol that required the

PRONIA sites to only 1) acquire isotropic or nearly isotropic voxel sizes of preferably 1 mm resolution,

2) set the Field Of View (FOV) parameters accordingly to guarantee the full 3D coverage of the brain

including all parts of the cerebellum, and 3) define the relaxation time (TR) and echo time (TE) as well

as other imaging parameters in a way that would maximize the contrast between cortical ribbon and

the white matter and enhance the signal-to-noise ratio in the images. Table S2 lists the parameters

defining the structural MR sequences used to examine in the PRONIA discovery sample participants.

MRI preprocessing pipeline

The manual of the CAT12 toolbox (http://www.neuro.uni-jena.de/cat12/CAT12-Manual.pdf) details

the processing steps applied to the structural images. These steps consist of:

1) A 1st denoising step based on Spatially Adaptive Non-Local Means (SANLM) filtering (5).

2) An Adaptive Maximum A Posteriori (AMAP) segmentation technique, which models local
    variations of intensity distributions as slowly varying spatial functions and thus achieves a
    homogeneous segmentation across cortical and subcortical structures (6).

3) A 2nd denoising step using Markov Random Field approach which incorporates spatial prior
    information of adjacent voxels into the segmentation estimation generated by AMAP (6).

4) A Local Adaptive Segmentation (LAS) step, which adjusts the images for white matter (WM)
    inhomogeneities and varying gray matter (GM) intensities caused by differing iron content in e.g.
    cortical and subcortical structures. The LAS step is carried out before the final AMAP
    segmentation.

5) A Partial Volume Segmentation algorithm that is capable of modeling tissues with intensities between GM and WM, as well as GM and cerebrospinal fluid (CSF) and is applied to the AMAP-generated tissue segments.

6) A high-dimensional DARTEL registration of the image to an MNI-template generated from the MRI data of 555 healthy controls in the IXI database (http://www.braindevelopment.org)

MRI data quality assurance

To inspect homogeneity of the acquired MRI scans and assure a high standard of MRI data quality, we employed the homogeneity check option of the CAT12 toolbox. As part of the preprocessing, CAT12 writes out several individual quality measures for each MRI scan: NCR (Noise Contrast Ratio), ICR (Inhomogeneity Contrast Ratio) and RES (RMS resolution). The resulting ratings are then combined into the weighted average image quality rating (IQR). These quality ratings are scaled from 0.5 to 10.5, where 0.5 is a "perfect/excellent" score and 10.5 is deemed "unacceptable/failed". Values around 1 and 2 describe "(very) good" image quality and values around 5 and higher indicate problematic images (7). The data quality features were entered into the CAT12 "check homogeneity" module along with modulated (m) normalized (w) GM segments (p1). We then computed the Mahalanobis distance between the mean correlation and weighted overall image quality. While mean correlation measures the homogeneity of all selected MRI data used for statistical analysis and is therefore a measure of image quality after pre-processing, the weighted overall image quality combines measurements of noise and spatial resolution of the images before pre-processing. Hence, calculating the Mahalanobis distance between these two measurements quantifies image quality both before and after pre-processing. Following this approach, we only included cases with an overall image quality rating (IQR) of "good" to "very good" (Figure S4). This led to the exclusion of four cases, which deviated from the rest of the sample (n=649) by more than two standard deviations. This protocol closely follows the general recommendation as given in the official CAT12-Manual (http://www.neuro.uni-jena.de/cat12/CAT12-Manual.pdf).

As clinical phenotypes and possibly childhood trauma might impact MRI data quality, we extracted the

main CAT12 data quality measure (IQR) for all 649 cases and investigated it with both univariate and

multivariate analysis (see Additional Analysis: Investigation of MRI image quality in the main study

sample).

Sparse partial least squares algorithm

The Sparse Partial Least Squares (SPLS) algorithm used in this analysis follows the original publication

of Monteiro et al. (8). Like Partial Least Squares (PLS), SPLS requires two data matrices X and Y as inputs.

In our study, X contains neuroimaging information (structural MRI data), while Y contains trauma

information and related features (trauma questionnaire items, age, sex, diagnoses). n is the number of

samples; p is the number of voxels and q is the number of trauma features. PLS provides insights into the

brain's mechanisms by finding relationships between different measures (i.e. views) from the same

participants, i.e. between neuroimaging and trauma data, in a clinical population. PLS identifies a

projection or latent space containing the relevant information in both views by finding pairs of weight

vectors (generally called u and v) which maximize the covariance between the projections of the two

views (9):

$$1) \quad maximize_{\|u\|_2=\|v\|_2=1} \, Cov(Xu, Yv) = \, maximize_{\|u\|_2=\|v\|_2=1} \, u^T X^T Y v$$

The weight vector pair is also called a latent variable (LV) as it explains one specific associative effect

between the two different views. More specifically, the weight vectors place weights on each feature in

the trauma and the neuroimaging dataset, thus visualizing which features are associated with each other

as well as the direction and the strength of this multivariate association. Hence, by studying this latent

space, one can learn about the underlying relationship between trauma information and brain measures

(8).

In contrast to regular PLS, SPLS enforces sparsity on the weight vectors u and v through hyperparameters

c_u and c_v. c_u and c_v are the regularization hyperparameters that control the l_1-norm constraints of u

and v, respectively. The l_1-norm constraints impose sparsity, which means that the lower the values of

c_u and c_v are, the higher the sparsity in the respective view is (10). This leads to the following optimization problem:

2) $maximize_{u,v}\ u^T X^T Y v\ subject\ to\ \|u\|_2^2 \le 1, \|v\|_2^2 \le 1, \|u\|_1 \le c_u, \|v\|_1 \le c_v$

Yet, this type of constraint can only select up to n features if $p > n$. Furthermore, it will remove features which might be relevant for the model but are correlated with other features which are already included. Zou and Hastie addressed this issue by adding the l_2-norm constraints (11). For both l_1-norm and l_2-norm constraints to be active, the values of the hyperparameters must be between 1 and the square root of the number of features in the respective matrices. Therefore, the hyperparameter space is updated:

3) $1 \le c_u \le \sqrt{p}, 1 \le c_v \le \sqrt{q}$

Using the hyperparameter space of equation 3) and solving the optimization problem of equation 2) according to Witten et al. (10, 12) leads to the following SPLS algorithm steps as described in Monteiro et al. (8):

1. Let $C \leftarrow X^T Y$
2. Initialize v to have $\|v\|_2 = 1$
3. Repeat until convergence:
    a) Update u:
        i. $u \leftarrow Cv$
        ii. $u \leftarrow \frac{S(u, \Delta_u)}{\|S(u, \Delta_u)\|_2}$, where $\Delta_u = 0$ if this results in $\|u\|_1 \le c_u$, otherwise Δ_u is set to be a positive constant such that $\|u\|_1 = c_u$
    b) Update v:
        i. $v \leftarrow C^T u$
        ii. $v \leftarrow \frac{S(v, \Delta_v)}{\|S(v, \Delta_v)\|_2}$, where $\Delta_v = 0$ if this results in $\|v\|_1 \le c_v$, otherwise Δ_v is set to be a positive constant such that $\|v\|_1 = c_v$
4. If convergence is not reached after the iteration limit (default: 1000), return non-sparse weight vectors u and v

After a weight vector pair (h) is found by SPLS, its effect needs to be removed from the data, to look for the next possible weight vector pair ($h + 1$). This process is called matrix deflation. In this setup, projection deflation is used as it has been shown to outperform the classic Hoteling's deflation, which is

also used in Principal Component Analysis (8, 13, 14). For matrices X and Y, the deflation process from iteration h to iteration $h + 1$ is therefore computed as follows:

$$X_{h+1} \leftarrow X_h - (X_h u_h)u_h^T$$

$$Y_{h+1} \leftarrow Y_h - (Y_h v_h)v_h^T$$

The algorithm then uses the deflated matrices and looks for the next associative effect, i.e. the next LV. This way, SPLS iteratively provides latent variables consisting of sparse weight vector pairs (u, v), uncovering several layers of associative effects within the dataset.

The second step of the SPLS algorithm involves the creation of latent scores. For every LV, weight vectors u and v are projected onto the matrixes X and Y, thus generating latent scores ε and ω.

$$\varepsilon_h = X u_h$$

$$\omega_h = Y v_h$$

These latent scores are finite numerical values, which represent the loading of each individual on these weight vectors, i.e. how high a subject scored on specific trauma questions or how high their gray matter probability in certain voxels is. Therefore, every individual can be represented within each LV space with its latent trauma and brain scores. These specific scores can then be used for post-hoc analyses to investigate the meaning and relevance of these individual loadings.

Machine learning framework

The models were generated and tested in a nested cross-validation framework with 10 outer (X2, Y2) and 10 inner folds (X1, Y1) (Figure S2). Individuals were randomly assigned to the fold structure, while a stratification according to diagnoses was maintained. Therefore, all inner and outer loops contained an equal distribution of all 4 study groups (HC, ROD, CHR, ROP) in order to avoid training on diagnosis-related effects or indirectly on site-related effects (as diagnoses were unevenly distributed across sites, see Main Text Table 1). Within the inner folds, a 40x40 point grid search of both hyperparameters was

conducted covering the entire hyperparameter space, in which both l_1- and l_2-norm constraints are fulfilled: $1 \leq c_u \leq \sqrt{p}$, $1 \leq c_v \leq \sqrt{q}$ (with p features in matrix X and q features in matrix Y). Lower c_u and c_v values lead to a sparser solution, whereas higher c_u and c_v values amount to a denser solution. At the upper limit, the maximum values of hyperparameters are: $c_u = \sqrt{p}$, $c_v = \sqrt{q}$. A SPLS analysis with c_u and c_v reaching these maximum values is equal to a regular PLS analysis, where every feature receives a weight and no feature is removed, i.e. no zero weights are given. Hence, our hyperparameter grid search includes the computation of one regular non-sparse PLS model (with c_u and c_v at the maximum limits) and an array of sparse PLS versions as lower c_u and c_v values are tested. In order to illustrate that the densest hyperparameter combination in our grid search resembles a regular non-sparse PLS, Supplementary Figure S5 shows an exemplary pair of brain and trauma vectors that was trained on our dataset using the maximum values of c_u and c_v. Therefore, in this framework, the non-sparse regular PLS solution competes against the sparse PLS solution in the hyperparameter optimization process. The weight vector pairs were generated using the training folds in the inner loop ($X1_{train}$, $Y1_{train}$):

$$(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{spls}(\boldsymbol{X1_{train}}, \boldsymbol{Y1_{train}}, \boldsymbol{c_u}, \boldsymbol{c_v})$$

The model fit of the weight vector pair was then assessed by projecting them onto the testing folds ($\boldsymbol{X1_{test}}$, $\boldsymbol{Y1_{test}}$) in the inner loop and computing Spearman's correlation coefficient between the projections of the weight vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ onto their respective data matrices $X1_{test}$ and $Y1_{test}$:

$$\boldsymbol{\rho} = |\boldsymbol{Corr}(\boldsymbol{X1_{test}u}, \boldsymbol{Y1_{test}v})|$$

This approach delivers a simple and transparent measure of how well the weight vectors align the matrices to each other, i.e. how well they can maximize the covariance. The median correlation coefficient was computed for each hyperparameter combination in the inner loop. Afterwards, the best hyperparameter combinations ($\boldsymbol{c_{u-top}}, \boldsymbol{c_{v-top}}$) with the highest median correlation coefficients

(ρ_{top}) were retrained on the entirety of all 10 folds of the inner loop to increase the sample size for training once more:

$$(u_{top}, v_{top}) = spls(X2_{train}, Y2_{train}, c_{u-top}, c_{v-top})$$

The generalizability of the weight vector pairs ($\mathbf{u_{top}}, \mathbf{v_{top}}$) was tested by assessing the fit of their projections onto the previously held-out fold in the outer loop and thus computing the corresponding correlation coefficients (ρ_{max}).

$$\rho_{max} = |Corr(X2_{test}u_{opt}, Y2_{test}v_{opt})|$$

Significance testing of this weight vector pair was achieved by permutation testing against $\boldsymbol{B}$ permutations. Within the fold structure of the outer loop, $\boldsymbol{B}$ permutated datasets were created by randomly reshuffling the order of participants in one matrix ($\boldsymbol{Yb2}$) thus destroying relationship between the two matrices. The final model with the optimized hyperparameters (c_{u-opt}, c_{v-opt}) was then retrained and tested in each of the $\boldsymbol{B}$ permuted datasets, thus generating weight vectors u_b, v_b:

$$(u_b, v_b) = spls(X2_{train}, Yb2_{train}, c_{u-opt}, c_{v-opt})$$

$$\rho_b = |Corr(X2_{test}u_b, Y2_{test}v_b)|$$

Significance testing of the LV was done by assessing how often the model based on the permuted dataset performed better or equal to the model trained on the original dataset:

$$p = \frac{1 + \sum_{b=1}^{B} \mathbf{1}_{\rho_b \geq \rho_{max}}}{B + 1}$$

As our framework consisted of 10 outer folds, this approach led to 10 different models (i.e. 10 weight vector pairs u and v) for each latent variable iteration. Of these 10 different models, we selected the one model with the best performance as measured by means of permutation testing, i.e. the model that exhibited the lowest P value. If this optimal model passed significance testing against the FDR-corrected *P* value for multiple testing (10 models = 10 tests), the latent variable was deemed significant and the next latent variable was computed. This concept is known as the omnibus hypothesis, which was also applied in the original method paper of the SPLS algorithm (8). The SPLS algorithm is an

iterative process, in which based on hyperparameters c_u and c_v, the weight vectors u and v are computed in dependence of each other (Methods – Sparse Partial Least Squares Algorithm). First u and v are initialized as non-sparse weight vectors based on regular singular value decomposition. Then an iterative process is set in motion, where first an enforcement of sparsity is attempted on weight vector u in dependence of weight vector v. Then sparsity is enforced on v, based on the previously computed weight vector u. This iterative process is repeated, where u and v are sequentially updated based on each other's previous modification until convergence between the vectors is reached. Hence, every hyperparameter setup c_u and c_v leads to a unique process of finding converging weight vectors u and v that were generated in a dialectic manner. Thus, the multivariate information is contained in this highly specific combination of weight vectors u and v, with both vectors containing mathematical information of the other. This, in turn, makes weight vectors u and v from different models, such as in our 10x10 fold nested cross-validation, not suitable for usual merging techniques (weighted mean/mean/median merging or majority voting) as every vector u is dependent on the corresponding vector v. Therefore, we used the omnibus hypothesis to determine our final LV model out of the 10 computed within the NCV structure of each LV iteration. Using our 10x10-fold outer and inner cross-validation loops can lead to high variance in the results. After training on the inner loops and then testing on the outer loops, 10 models with 10 P values are obtained. A criterion is then needed to determine whether any statistically significant effects were indeed found. For this, we used the omnibus hypothesis, where a statistical test is performed j-times to test a null-hypothesis H_j. Following the omnibus approach, the combined hypothesis H_R over all tests j is: "All the hypothesis H_j are true". This hypothesis will be rejected if any of the H_j hypothesis is rejected (15). In our specific case, the omnibus hypothesis states that if any of the 10 p-values (obtained in the 10 outer folds) is statistically significant (corrected for multiple testing j-times), then the omnibus hypothesis will be rejected and the detected effect will be deemed significant. Therefore, the omnibus hypothesis will be rejected if any of the 10 splits generates a P value below .05 (adjusted for multiple testing). Of all significant splits, the model with the lowest P value will be determined as the final LV model (8). The computation end as soon as none of the 10 splits of the LV iteration did not pass the test for significance, which renders

the entire LV not significant. Since deflating the data matrices of non-significant effects would be not justified, the analysis pipeline stops after the first non-significant LV was detected.

External clinical validation of the SPLS trauma model

In order to assess the clinical real-world validity and generalizability of our trauma model (LV1-LV5), we extracted 427 additional individuals from the PRONIA replication sample for external validation (Table S6). These 427 individuals were recruited at 10 different sites (PRONIA discovery sites and additionally in three new sites: Duesseldorf and Muenster, Germany and Bari, Italy) during the second (replication) phase of the PRONIA study. We acquired structural MRI data as well as CTQ, age, sex and diagnostic information of these 427 individuals and applied our childhood trauma model onto this external validation sample. For each latent variable (LV1-LV5), describing a particular layer of trauma-brain association, we projected the corresponding weight vectors u (brain vector) and v (trauma vector) onto every individual's brain and trauma information, thus generating their individual loading onto these patterns, i.e. their latent brain (epsilon) and trauma scores (omega). We investigated the impact of these latent brain and trauma scores on individual clinical phenotypes by correlating them to the same set of clinical variables, we already used for the discovery sample: GAF, GF, BDI, NEO-FFI and WHOQOL-BREF. Hence, the aim was the assessment of the real-world clinical validity of our multi-layered trauma model in a temporally and partially geographically separated sample.

Visualization and atlas mapping of neuroanatomic weight vectors

The neuroanatomic weight vectors of the LV were visualized as 3D MRI images using the SPM12 software (Wellcome Department of Cognitive Neurology, London, UK; http://www.fil.ion.ucl.ac.uk/spm/software/spm12/) and the Connectome Workbench v1.3.2 (https://humanconnectome.org/software/connectome-workbench). Readouts of specific atlas regions (Table S14) were attained using the Brainnetome atlas (http://atlas.brainnetome.org/index.html, (16)) and the cerebellar atlas by Diedrichsen (http://www.diedrichsenlab.org/imaging/propatlas.htm, (17)).

Supplementary Results

Main Analysis - SPLS results: correlation between latent scores and clinical domains

Phenotypic and brain scores of LV1 and LV3 were part of the same correlation analysis as the other three CT-specific LV (LV2, LV4, LV5). We report the results of the correlation analyses for both the discovery and the replication sample of the more age- and sex-dependent signatures of LV1 and LV3 below (Table S15 and Table S16).

<u>Discovery Sample:</u>

LV1 (age). Phenotypic scores: Positive, significant associations were detected for measures of symptom severity (GAF:S Past Month), disability (GAF:D/I Past Month), social functioning (GF:S Current, Low Past Year) and role functioning (GF:R Current, Low Past Year) (ρ-range: 0.11-0.16, *P*-range: <10^{-3}-.04). Further positive, significant associations were detected for the WHOQOL-BREF physical, social relationships and environment domains (ρ-range: 0.09-0.15, *P*-range: <10^{-3}-.05). **Brain scores:** No significant associations were found.

LV3 (sex). Phenotypic scores: Significant, negative correlations were detected for all domains of GAF:S, GAF:D/I and GF:S, whereas further negative correlations were also found for several dimensions of GF:R, NEO-FFI (extraversion, openness, conscientiousness, agreeableness) and quality of life (WHOQOL-BREF: physical, psychosocial, social relationships, environment)(ρ: -0.11-(-0.29, *P*: <10^{-3}-.04). Significant, positive associations were observed for BDI and NEO-FFI neuroticism levels (ρ: 0.18-0.20, *P*: <10^{-3}). **Brain scores:** We detected significant, negative associations for GAF:D/I Lifetime and two NEO-FFI domains (agreeableness, conscientiousness) (ρ: -0.13-0.19, *P*: <10^{-3}-.04).

<u>Replication Sample:</u>

LV1 (age). Phenotypic scores: Negative, significant correlations were observed for GF:S Low Past Year and WHOQOL-BREF environment (ρ-range: -0.14-(-0.16), *P*-range: .03-.01). Positive, significant correlations were detected for NEO-FFI conscientiousness (ρ = 0.15, *P*=.01). **Brain scores:** Significant,

negative correlations were detected for several GAF:S, GAF:D/I, GF:R and GF:S domains as well as WHOQOL-BREF environment (ρ: -0.13-(-0.21), *P*: <10^{-3}-.04).

LV3 (sex). Phenotypic scores: Significant, negative correlations were detected for all domains of GAF:S, GAF:D/I and GF:S, whereas further negative correlations were also found for several dimensions of GF:R, NEO-FFI (extraversion, conscientiousness, agreeableness) and quality of life (WHOQOL-BREF: psychosocial, social relationships)(ρ: -0.11-(-0.33, *P*: <10^{-3}-.04). A significant, positive association was observed for BDI (ρ=0.17, P=.01). **Brain scores:** We detected one significant, negative association for NEO-FFI conscientiousness (ρ=-0.13, *P*=.05).

Main Analysis – Discussion of LV1 and LV3

For a thorough overview of the SPLS analysis results, all significant LV, which were not discussed in detail in the main text, are now mentioned with regards to their trauma and brain patterns (Table S15 and Table S16).

LV1 contained a phenotypic pattern, which was heavily dominated by age (Figure S6A). The associated brain pattern consisted of negatively weighted grey matter volume (GMV) across various cortical and subcortical brain areas (Figure S6B). These findings are in line with the normal development in young adults, mainly stemming from a microglia-induced pruning processes (18). Furthermore, the latent trauma and brain scores of LV1 did not yield any significant correlations with our predefined clinical domains (Table S15), thus further indicating that this LV reflects a physiological aging and pruning process.

LV3 yielded a phenotypic pattern (Figure S7A), which contained weights for age and several CTQ questions. Nonetheless, it was most strongly informed by sex, thus mainly separating male from female participants. The corresponding brain pattern is spread widely across the precuneus, the parietal and temporal lobe, the basal ganglia, the insula, the hippocampus, the thalamus and the frontal gyrus (Figure S7B). Hence. This signature predominantly reflects the well-established sexual GMV dimorphism, including temporal and parietal areas as well as the hippocampus, the planum

temporale and the insula (19, 20). The latent trauma and brain scores of LV3 yielded significant correlations with lower levels of functioning and maladaptive personality traits (Table S15). This association could be heavily informed by certain characteristics of our study sample. The ROP patients in our study generally have much lower levels of functioning and are predominantly male, while the HC population shows higher levels of functioning and is predominantly female. This might, at least in part, explain the association between the mainly sex-informed signature and the functioning measures (Table 1 in Main Text and Supplementary Table S6 and Table S7). Nonetheless, further studies should follow up these sex-specific effects and explore their association with underlying pathologies and diagnostic categories.

Additional Analysis: random split-half SPLS analysis

In order to assess other methods for ensuring generalizability and validity of generated machine learning models, we conducted a random split-half experiment, where a random selection of one half of the main study population was used for training in a nested cross-validation framework and the other half was used for testing the models. Following this approach, we detected two significant latent variables (LV).

LV1 (*P* value = 9x10^{-4}). Phenotypic pattern (Figure S8A): Age received a strong positive weight, far outweighing all other features. Further weights were assigned to the subscales of emotional (1 item) and physical neglect (1) as well as denial (3). **Brain pattern** (Figure S8B): GMV was widely negatively weighted across the superior and middle frontal and temporal gyrus, the inferior parietal lobe as well as in cingulate and orbital regions.

LV2 (*P* value = 3x10^{-4}). Phenotypic pattern (Figure S9A): Except for items CTQ08 (emotional abuse) and CTQ28 (emotional neglect), the entire feature space was represented in this phenotypic pattern. Items from emotional, physical and sexual abuse received consistent positive weighting, while emotional neglect yielded several negative weights. Items from the physical neglect subscale received both positive and negative weights and the denial items were all positively weighted. Further positive weights were assigned to age, male sex and ROP and ROD status, whereas female sex, HC and CHR

status were negatively weighted. **Brain pattern** (Figure S9B): GMV was widely negatively weighted bilaterally in superior, middle and orbitofrontal brain regions as well as in the superior, middle and inferior temporal gyri. Furthermore, strong negative weights were assigned to the cingulate and insular cortex as well as the fusiform gyrus. Positive GMV weights were detected in the right anterior superior temporal sulcus and bilaterally in the DLPFC.

Hence, the random split-half approach yielded a similar age-dependent signature as described in our main analysis (LV1) and further generated one highly dense, trauma-specific signature. In this signature, a pattern of abuse (emotional, physical, sexual) is associated with widespread negative GMV weights in pre-/orbitofrontal, temporal and subcortical brain regions as well as positive GMV weights mainly located in the anterior superior temporal sulcus. For the neglect items, this brain signature is either inversed (emotional neglect) or not precisely attributable due to heterogeneous weighting (physical neglect). Moreover, this brain pattern is expressed in older, male individuals from the ROP and ROD categories, i.e. the participants that suffer from a full-blown psychiatric disorder. For younger, female individuals from the HC or CHR category, this brain pattern is inversed, leading to the assumption that the two main dimensions of the CTQ, abuse and neglect, exert a differential effect on an individual's brain along the dichotomies of young vs. old, male vs. female and HC/CHR vs. ROD/ROP. LV2 of the RSH approach shows evenly distributed brain and phenotypic overlaps with all LV of the main SPLS model (Table S25 for mean squared error comparisons). These findings suggest that the RSH model can be interpreted as a condensed version of the main SPLS model. Due to the reduced sample size for the training sample (n=325), it can be speculated that the training input was not sufficient for the algorithm to further specify and separate the different trauma-brain layers in a more fine-grained way. In summary, the RSH approach indeed yielded intriguing results with regards to a differential trauma-brain connection of abuse and neglect CTQ dimensions, which should be followed up with future analyses.

Additional Analysis: Leave-one-site-out cross-validation (LOSOCV) analysis

To assess further methods for validating our model's generalizability to new datasets and populations, we conducted a leave-one-site-out cross-validation (LOSOCV) analysis. Age, sex and diagnoses were still included in the analysis pipeline, while site correction was now omitted due to the LOSOCV approach. Following this strategy, we detected five significant latent variables.

LV1 (P = 9x10^{-4}). Phenotypic pattern (Figure S10A): Positive weights were assigned most strongly to age as well as to items from the CTQ subscales of emotional abuse (1 item), physical abuse (1), sexual abuse (5) and inversely to male and female sex. **Brain pattern** (Figure S10B): GMV was globally and bilaterally negatively weighted across mostly frontal, temporal and parietal regions. Positive GMV weights were highly sparse and only found in the thalamus region.

LV2 (P = 3x10^{-3}). Phenotypic pattern (Figure S11A): All five questions from the CTQ sexual abuse subscale received positive weights, whereas age was negatively and sex inversely weighted. **Brain pattern** (Figure S11B): Positive GMV weights were found extensively bilaterally in the pre- and postcentral gyrus, the superior and middle frontal gyrus as well as in the inferior parietal lobule and the subcortical areas (basal ganglia, thalamus). Negative GMV weights were detected predominantly in the left hippocampus and parahippocampus, the left fusiform gyrus and bilaterally in the superior and middle frontal gyrus as well as the middle temporal gyrus.

LV3 (P = 9x10^{-4}). Phenotypic pattern (Figure S12A): Questions from the sexual abuse (2 items) received a positive weight, while sex was strongly and inversely weighted. **Brain pattern** (Figure S12B): We detected positive GMV weights mostly bilaterally in the precuneus as well as the superior and middle frontal gyrus, the inferior parietal lobule and further in the cingulate and subcortical areas (mostly basal ganglia). Negative GMV weights were found sparsely in bilateral reductions in the hippocampus and the DLPFC.

LV4 (P = 9x10^{-4}). Phenotypic pattern (Figure S13A): Questions from the emotional abuse subscale (2 items) were negatively weighted and questions from the sexual abuse (4) were mostly positively

weighted. Age received a strong positive and sex a strong inverse weighting. **Brain pattern** (Figure S13B): Positive GMV weights were mainly localized bilaterally in temporal regions and orbitofrontal regions as well as precuneus and the basal ganglia. Negative GMV weights were detected sparsely in bilateral reductions in the DLPFC as well as in the precentral gyrus and the thalamus.

LV5 (*P* = 9x10⁻⁴). Phenotypic pattern (Figure S14A): Questions from the emotional abuse subscale (2 items) were negatively weighted and one question from the sexual abuse (1) was positively weighted. Age received a positive and sex a strong inverse weighting. **Brain pattern** (Figure S14B): Positive GMV weights were mainly localized bilaterally in the inferior parietal lobule, the superior temporal sulcus and the cerebellum as well as the orbitofrontal and precuneus regions. Negative GMV weights were detected sparsely in the postcentral gyrus, the superior frontal gyrus and the thalamus bilaterally.

Hence, the LOSOCV approach yielded a highly similar age-dependent signature as described in our main analysis (Table S26 for mean squared error comparisons). Furthermore, it yielded two signatures (LV3, LV5) that display high phenotypic similarities with the sex-dependent LV3 signatures of our main model. On a neuroanatomic level, these two signatures were most similar to LV1 (age) and LV3 (sex) of our main model. It should be noted, that all LOSOCV LV showed similarity to LV1 (age) of the main model, since this pattern covers most of the brain and therefore always leads to a certain kind of overlap. Thus, it can be stated that apart from the global LV1 main model signatures, LV3 and LV5 of the LOSOCV model most closely resemble the sex-dependent LV3 signature of the main model, both phenotypically and neuroanatomically. Taking a similar approach, the LOSOCV pattern of LV4 bore the highest phenotype- and brain-level similarities to LV1 (age) and LV4 (combined sexual/physical abuse & sex) of the main model. Hence, the LOSOCV approach yielded a multi-layered model, that closely resembled LV1-LV4 of the main model. While LV1 of the LOSOCV model corresponded tightly to LV1 of the main model, LV3 and LV5 showed high similarities to LV3 and LV4 contained similar phenotypic and neuroanatomic patterns as LV1 and LV4 of the main model. Solely the emotional trauma signature of LV5 from the main model could not be specifically found with the LOSOCV approach and featured only very sparsely across all LV. Furthermore, it does appear that in the LOSOCV approach the sex

signature, contained in a single LV in the main model, was split up into two sparser signatures (LV3, LV5). This might indeed be an effect of the severe imbalances of male and female individuals across sites, so that the algorithm was not able to uncover a more global sex signature, but rather detected it in subsequent steps. Yet, it should be further investigated whether an added value can be obtained from these sparser signatures as well.

In summary, the LOSOCV approach yielded results, which were quite comparable to our main study. The most apparent difference arose from the missing emotional trauma signature and the split-up sex signature of the LOSOCV approach. In summary, we believe that the LOSOCV approach mostly confirmed our main results and offers intriguing options for further analyses.

Additional Analysis: Investigation of MRI image quality in the main study sample

To investigate MRI image quality in in our main study sample, we conducted univariate and multivariate analysis using the image quality rating (IQR) as our main measure for overall image quality.

Group-level statistics

Considering IQR as a possible additional phenotypic input feature, we performed group-level statistical tests and did not find any significant differences across study groups (Table S3). Furthermore, the overall distribution of IQR scores revealed "good" to "very good" MRI data quality for the entire sample (Figure S4).

Linear regression analysis

Aiming to assess the specific association between the clinical characteristics of our study sample and overall MRI data quality, we undertook linear regression analysis. We used the entire phenotypic input feature dataset containing all CTQ items as well as age, sex and study group and predicted individual IQR scores. The overall linear model proved significant, while the R-squared value showed that only 5.6% of the IQR variance could be explained by the phenotypic input features (

Table **S4**, adjusted R-Squared: 0.0558, F-statistic vs. constant model: 2.16, *P* Value = 2.3×10^{-4}). Furthermore, we found that three CTQ items belonging to emotional abuse (CTQ08) and emotional

neglect (CTQ13, 19) (*P* Value range: 1.95×10^{-2}-4.09×10^{-2}) as well as age (*P*=1.55×10^{-4}) and sex (*P*=1.08×10^{-5}) were significantly associated with IQR in the linear model.

In a second step, we investigated the association between IQR scores and the five CT signatures as detected in our main analysis (LV1-LV5). We conducted linear regression analyses using IQR scores to predict latent phenotypic and brain scores independently for each subject (Table S5). We detected significant associations between IQR scores and the phenotypic scores of LV1-LV3 (*P* Value range: 3.28×10^{-2}-1.01×10^{-5}) as well as the brain scores of LV4-LV5 (*P* Value range: 2.20×10^{-2}-2.30×10^{-3}). Yet, we did not detect a significant association between IQR scores and both layers (neuroanatomic, phenotypic) of an LV signature. Thus, the quality measure was never connected to the entire signature, but rather to isolated domains of it. Aiming to quantify the impact of these IQR measures on the latent scores, we additionally computed R-squared values and found that the IQR scores explained between 0.2-3.6% of the latent score variance.

In summary, these results suggest a significant association between the phenotypic input feature dataset–especially for three emotional trauma CTQ items, age and sex–and MRI data quality (IQR). Furthermore, we found a significant association between the IQR scores and isolated parts of the SPLS CT signatures. Yet, this significance might be attributed to the large sample size (n=649), while the effect sizes appear to be minimal. The phenotypic input features explain only about 5% of the IQR variance, while the IQR scores explain 0.2-3.6% of the variance of the latent phenotypic and brain scores. Hence, it can be assumed that the impact of our phenotypic input features on MRI data quality and in turn the influence of MRI data quality on our SPLS results is very limited.

SPLS analysis: IQR as additional phenotypic input feature

We performed an SPLS analysis within the same nested cross-validation framework as the main analysis and added the IQR scores as an additional phenotypic input feature. Following this strategy, we detected nine significant latent variables.

LV1 (*P* = 9x10⁻⁴). Phenotypic pattern (Figure S15, LV1-A): Positive weights were assigned to sexual abuse (4 items) and age. **Brain pattern** (Figure S15, LV1-B): GMV was globally and bilaterally negatively weighted across mostly frontal, temporal, parietal and orbital regions. Positive GMV weights were only found in the thalamus region.

LV2 (*P* = 9x10⁻⁴). Phenotypic pattern (Figure S15, LV2-A): Positive weights were distributed to sexual abuse (4 items) and female sex. Negative weights were found for emotional abuse (1), age and male sex. **Brain pattern** (Figure S15, LV2-B): Positive GMV weights were found bilaterally in the inferior parietal, postcentral and occipital gyrus as well as the cerebellum. Further positive GMV weighting was detected in the precentral gyrus, the basal ganglia and the orbital gyrus. Negative GMV weights were detected mostly bilaterally in the DLPFC, left fusiform gyrus and the right cingulate gyrus.

LV3 (*P* = 9x10⁻⁴). Phenotypic pattern (Figure S16 LV3-A): Positive weights were assigned to sexual abuse (3 items), emotional (1) and physical neglect (1) as well as male sex and IQR. Negative weighting was found for female sex. **Brain pattern** (Figure S16 LV3-B): We detected positive GMV weights predominantly bilaterally in the superior and medial frontal gyrus, the precuneus, the cingulate gyrus, the cerebellum and the occipital cortex. Negative GMV weights were found bilaterally in the DLPFC, the hippocampus and the left thalamus.

LV4 (*P* = 9x10⁻⁴). Phenotypic pattern (Figure S16 LV4-A): Age and female sex received positive weights, whereas male sex was negatively weighted. **Brain pattern** (Figure S16 LV4-B): Positive GMV weights were sparsely localized in the DLPFC. Negative GMV weights were detected in the right caudate nucleus and the right thalamus.

LV5 (*P* = 3x10⁻³). Phenotypic pattern (Figure S17 LV5-A): Emotional (4 items), physical (5) and sexual abuse (2) as well as physical neglect (4), age, male sex, ROP and ROD status were positively weighted. Negative weights were assigned to sexual abuse (1), emotional neglect (1), female sex, HC status and IQR. **Brain pattern** (Figure S17 LV5-B): Positive GMV weights were mainly localized bilaterally in the DLPFC, the basal ganglia and the thalamus. Negative GMV weights were detected widely in the inferior

parietal lobule, the orbital gyrus, the medial frontal gyrus, the cerebellum and the superior and middle temporal gyrus.

LV6 (P = 2.4x10^{-2}). Phenotypic pattern (Figure S17 LV6-A): Age and male sex were positively weighted, while sexual abuse (1 item) and female sex received a negative weighting. **Brain pattern** (Figure S17 LV6-B): Positive GMV weights were densely distributed across the inferior parietal lobule, the fusiform gyrus, the cingulate gyrus, the superior frontal and temporal gyrus as well as the cerebellum, precuneus and the orbital gyrus. Negative GMV weights were detected mostly in the anterior temporal sulcus, the DLPFC, the cerebellum and further superior frontal and temporal regions as well as orbital, cingulate and fusiform gyrus.

LV7 (P = 2x10^{-3}). Phenotypic pattern (Figure S18 LV7-A): Positive weights were found for emotional (2 items) and sexual abuse (1) as well as age. **Brain pattern** (Figure S18 LV7-B): Positive GMV weights were sparsely localized in the left cuneus and the left precentral gyrus. Negative GMV weights were detected similarly sparsely in the orbital gyrus and the cingulate gyrus.

LV8 (P = 9x10^{-4}). Phenotypic pattern (Figure S18 LV8-A): Positive weights were found for emotional (5 items), physical (3) and sexual abuse (3), emotional (5) and physical neglect (5) as well as male sex, ROP, CHR status and IQR. Negative weights were assigned to physical (2) and sexual abuse (2), denial (3), age, female sex, HC and ROD status. **Brain pattern** (Figure S18 LV8-B): Dense, positive GMV weights were localized bilaterally in the DLPFC, the precuneus, the anterior temporal sulcus, the fusiform, orbital and superior frontal gyrus as well as the cerebellum. Negative GMV weights were detected in the precuneus, cuneus and lingual gyrus, the cerebellum, the DLPFC, the polar cortex and further frontal and parietal regions.

LV9 (P = 2x10^{-2}). Phenotypic pattern (Figure S19A): Positive weights were distributed to emotional neglect (4 items), age and female sex, while negative weights were found for emotional (2) and sexual abuse (1) as well as male sex and IQR. **Brain pattern** (Figure S19B): Positive GMV weights were located in the right DLPFC. Negative GMV weights were detected in the right thalamus.

LV1 and LV4 of the IQR-addition SPLS model closely resembled the age-dependent signature of LV1 from the main analysis, both on a brain and on a phenotypic level (Table S27 for mean squared error comparisons). Furthermore, LV2 contained a sexual abuse signature, dependent on age and sex, which mostly resembled the sex-dependent physical and sexual abuse signature of LV4 (phenotype- and brain-level) as well as the age-dependent sexual abuse signature of LV2 of the main model (phenotype-level). LV3 yielded a sex-dependent signature, which showed a high neuroanatomic and phenotypic similarity to the strongly sex-informed physical and sexual abuse signature of LV3 of the main model. LV5 contained a dense age-, sex-, study-group and IQR-informed signature of predominantly emotional, sexual and physical abuse. It bore a close resemblance to LV3 of the main model. Moreover, LV6 featured age and sex almost exclusively and thus was highly similar to the aging signature of LV1 and the heavily sex-informed signature of LV3 of the main model. LV7 represented a highly sparse signature of emotional trauma and age, which, due to the strong age influence, was most similar to the aging pattern of LV1 of the main model. LV8 contained a highly dense pattern, which included weights for all phenotypic input features, but was most strongly dominated by emotional trauma. Due to its very dense nature, both on a neuroanatomic and phenotypic level, it yielded high similarities to LV1, LV3, LV4 and LV5 of the main model, while it was particularly closely associated to the emotional trauma signature of LV5 of the main model (phenotype-level). LV9 was dominated by IQR weighting, but further contained weighting for emotional trauma, age and sex. It mostly resembled LV1, LV3 and LV4 of the main model, presumably due to the age and sex weighting of these LV.

Conclusively, all five signatures from the main model were detected in the IQR-addition SPLS model. The aging signature of LV1 and the sex signature of LV3 of the main model were contained in five new LV (LV1, LV3, LV4, LV5 and LV6). Furthermore, the age-dependent sexual abuse signature of LV2 was found partly in the new LV2. The sex-dependent physical and sexual abuse signature of LV4 was mostly detected in the new LV2 and LV8 signatures, whereas the emotional trauma signature of LV5 was partly replicated in the new LV7 and LV8 signatures. Specifically, LV8 represented a new finding as it contained a very global CT signature, which featured the entire phenotypic input feature dataset. This

signature linked predominantly emotional trauma and, to a lesser degree, physical and sexual trauma to a widely distributed GMV pattern across cortical and subcortical areas. It contained a negative weighting of age as well as an inverse weighting of male and female sex. Additionally, it was inversely expressed between the HC/ROD population and the ROP/CHR population. Hence, LV8 can be summarized as a global CT signature (predominantly emotional trauma), which was inversely expressed in the study population along the following lines: psychosis spectrum (CHR/ROP) vs. healthy/affective spectrum (HC/ROD), young vs. old and male vs. female. LV9 featured IQR as the most strongly weighted feature and contained further weightings for emotional and sexual trauma as well as age and sex. The phenotypic pattern corresponded to a highly sparse brain pattern, which featured mostly the right DLPFC and the right thalamus. These findings suggest that a very particular subset of phenotype-brain association, which was strongly confounded by IQR, was extracted as a single LV.

Further regarding the impact of data quality, a total of four LV were partly influenced by IQR scores (LV3, LV5, LV8, LV9). In LV3, IQR received only a minimal weighting, so that a relevant influence of data quality cannot be assumed. In LV5 (IQR ranked 10th out of 25 weighted features) and LV8 (21st out of 36) IQR was moderately weighted within highly dense phenotypic patterns that contained weights for 69% (25/36) and 100% (36/36) of the phenotypic input feature dataset.

Within the LV5 signature, IQR scores were weighted congruently to female sex and HC status and inversely to most of the CTQ features, as well as age, male sex and ROP/ROD status. Contrary to these findings, in LV8, IQR scores were weighted in the same direction as most CTQ items, male sex and CHR/ROP status, and inversely to female sex and HC/ROD status. In LV9, IQR weights were oriented congruently to emotional and sexual abuse and male sex. The weighting was inverse for emotional neglect, age and female sex. Thus, in our multivariate setup, the relationship between IQR and CT loading, sex and study group appeared highly complex, so that a clearly interpretable, linear assumption could not be made. This might be due to high-dimensional interactions and non-linear effects within the phenotypic input features and the MRI data. Yet, a stable relationship was found between age and IQR scores, which was consistent between LV5, LV8 and LV9. It suggested that

younger age was associated with higher ("worse") IQR scores and vice versa, thus reflecting the current body of literature (21, 22).

In summary, the addition of IQR as a phenotypic input feature yielded three main findings: 1) The results from our main study were largely replicated as all five LV were found distributed or condensed across the newly acquired LV signatures; 2) A dimensional impact of IQR on the dataset does exist, yet its contribution to the detected phenotypic and neuroanatomic patterns is very limited. Thus, the multivariate SPLS analysis supported the findings of our univariate analysis approach, revealing only minimal contributions of IQR scores to the variance of the CT patterns (see Linear regression analysis); 3) A clear-cut connection between IQR and CT, sex and study group could not be drawn from our SPLS analysis as the relationship appeared to be highly complex and possibly driven by non-linear interactions between phenotypic and neuroanatomic data. Contrary to that, a consistent, inverse relationship was found between age and IQR, thus reflecting current knowledge on MRI data quality and supporting the validity of our data quality assessment and assurance protocol.

Additional Analysis: Visualization of correlation between phenotypic and brain weighting

To increase the interpretability of the phenotypic and brain patterns, we have conducted a visualization experiment between weighted features in the phenotypic and the brain pattern of LV1 in the discovery sample. We computed the mean voxel values for the most strongly positively (lateral prefrontal thalamus, lPFtha) and negatively weighted voxel clusters (medial superior frontal gyrus, SFG, A9m) in LV1 and then correlated these voxel values with age, which was the most strongly (positively) weighted feature in LV1 (Table S14, Figure S20). We performed Spearman's correlation analysis and detected an inverse association between age and GMV in the chosen cluster in the left medial superior frontal gyrus ($\rho=-0.34$, $P=1.24 \times 10^{-18}$) and a positive association between age and the chosen cluster in the left lateral prefrontal thalamus ($\rho=0.19$, $P=1.04 \times 10^{-06}$).

Supplementary Tables

Table S1: Characteristics of the recruiting institutions in the PRONIA consortium. Shorter version without replication sample sites (Muenster, Duesseldorf, Bari) previously published in Koutsouleris et al. (1) and reprinted with permission.

PRONIA Site	Institution Name	Country	Type of Service	Catchment Population	Screening population / year
Munich	Department of Psychiatry and Psychotherapy, Ludwig-Maximilian-University Munich	DE	Academic outpatient services including specialized service for early recognition of psychosis; tertiary care academic hospital	1,200,000	700
Basel	Department of Psychiatry and Psychotherapy, University of Basel	CH	Academic inpatient and outpatient services including specialized service for early recognition and intervention of psychosis; tertiary care academic hospital	500,000	200
Milan Niguarda	Department of Pathophysiology and Transplantation, University of Milan. Four recruitment hospitals: Niguarda, Policlinico, San Paolo, Villa San Benedetto Menni in Albese con Cassano	IT	Psychiatric outpatient services including specialized services for early recognition of psychosis and persons at high risk; Academic hospital, providing psychiatric inpatient services, psychiatric outpatient services and local services	600,000	1,000
Cologne	Department of Psychiatry and Psychotherapy, University of Cologne	DE	Academic outpatient services including specialized service for early recognition of psychosis; tertiary care academic hospital	1,000,000	600
Birmingham	The University of Birmingham	UK	Academic specialized Early Intervention Service for Psychosis covering Birmingham and Solihull. Community and Inpatient	1,200,000	800
Turku	Department of Psychiatry, University of Turku	FI	Psychiatric outpatient and hospital services responsible for treatment of psychiatric patients in their catchment areas in the South-Western Finland	284,000	2,300
Udine	Department of Psychiatry, University of Udine	IT	Psychiatric outpatient services, academic hospital and local services. Tertiary care neuropsychiatric service	600,000	500
Muenster	Department of Mental Health, University of Muenster, Muenster, Germany	DE	Academic in- and outpatient services including service for early recognition of psychosis; tertiary care academic hospital	300,000	400
Duesseldorf	Department of Psychiatry and Psychotherapy, Heinrich-Heine University Duesseldorf	DE	Academic outpatient services including specialized service for early recognition of psychosis; tertiary care academic hospital	650,000	6,500
Bari	Department of Basic Medical Sciences, Neuroscience and Sense Organs University of Bari 'Aldo Moro', Bari	IT	Academic inpatient and outpatient services including specialized service for early recognition of psychosis.	2,123,324	452

Table S2: MR scanner systems and structural MRI sequence parameters used at the respective PRONIA sites. Shorter version without replication sample sites (Muenster, Duesseldorf, Bari) previously published in Koutsouleris et al. (1) and reprinted with permission.

PRONIA Site	Model	Field Strength	Coil Channels	Flip Angle	TR [ms]	TE [ms]	Voxel Size [mm]	FOV	Slice Number
Munich	Philips Ingenia	3T	32	8	9.5	5.5	0.97 x 0.97 x 1.0	250 x 250	190
Milan Niguarda	Philips Achieva Intera	1.5T	8	12	Shortest (8.1)	Shortest (3.7)	0.93 x 0.93 x 1.0	240 x 240	170
Basel	SIEMENS Verio	3T	12	8	2000	3.4	1.0 x 1.0 x 1.0	256 x 256	176
Cologne	Philips Achieva	3T	8	8	9.5	5.5	0.97 x 0.97 x 1.0	250 x 250	190
Birmingham	Philips Achieva	3T	32	8	8.4	3.8	1.0 x 1.0 x 1.0	288 x 288	175
Turku	Philips Ingenuity	3T	32	7	8.1	3.7	1.0 x 1.0 x 1.0	256 x 256	176
Udine	Philips Achieva	3T	8	12	Shortest (8.1)	Shortest (3.7)	0.93 x 0.93 x 1.0	240 x 240	170
Muenster	Siemens Magnetom PRISMA-FIT	3T	20	8	2130	2,28	1x1x1	256	192
Duesseldorf	Siemens Prisma	3T	32	8	2000	3.37	1.0x1.0x1.0	256x256	176
Bari	Philips Ingenia	3T	32	8	8.1	3.7	1.0 x 1.0 x 1.0	256 x 256	180

Table S3: Group-level statistics for IQR differences. Abbreviations: IQR, image quality rating; HC, healthy control; ROD, recent-onset of depression; CHR, clinical high-risk state; ROP, recent-onset of psychosis.; LB, Lower Boundary; Diff., Difference; UB, Upper Boundary.

Kruskal-Wallis test statistics

	All	HC	ROD	CHR	ROP	χ2	*P* Value
IQR, mean	1.93	1.93	1.93	1.92	1.95	1.00	.80
SD	0.15	0.16	0.14	0.12	0.18		

Dunn's test statistics for multiple comparisons

Group 1	Group 2	LB	Diff.	UB	*P* Value
ROD	CHR	-73.48	-11.44	50.60	.99
ROD	ROP	-82.98	-21.90	39.17	.92
ROD	HC	-69.18	-16.19	36.81	.96
CHR	ROP	-72.15	-10.46	51.23	.99
CHR	HC	-58.45	-4.74	48.96	.99
ROP	HC	-46.87	5.72	58.30	.99

Table S4: Linear regression model using the phenotypic input feature dataset to predict IQR (image quality rating). Number of observations: 649; Error degrees of freedom: 615; Root Mean Squared Error: 0.145; R-squared: 0.104; Adjusted R-Squared: 0.0558; F-statistic vs. constant model: 2.16; P Value = 2.3×10^{-4}. Abbreviations: Estimate, Coefficient estimates for each corresponding term in the model; SE, Standard error of the coefficients; TS, t-statistic for each coefficient. Significant *P* values are highlighted in bold.

	Estimate	SE	TS	*P* Value
(Intercept)	**2.12**	**4.32×10^{-2}**	**49.1**	**1.22×10^{-214}**
CTQ_03	-1.20×10^{-2}	7.67×10^{-3}	-1.57	1.17×10^{-1}
CTQ_08	**1.93×10^{-2}**	**8.90×10^{-3}**	**2.17**	**3.04×10^{-2}**
CTQ_14	1.19×10^{-3}	8.04×10^{-3}	1.49×10^{-1}	8.82×10^{-1}
CTQ_18	-5.50×10^{-3}	8.18×10^{-3}	-6.72×10^{-1}	5.02×10^{-1}
CTQ_25	-1.42×10^{-3}	7.80×10^{-3}	-1.82×10^{-1}	8.56×10^{-1}
CTQ_09	6.83×10^{-3}	1.46×10^{-2}	4.68×10^{-1}	6.40×10^{-1}
CTQ_11	-1.99×10^{-2}	1.15×10^{-2}	-1.73	8.46×10^{-2}
CTQ_12	1.93×10^{-2}	1.06×10^{-2}	1.82	6.93×10^{-2}
CTQ_15	1.53×10^{-2}	1.33×10^{-2}	1.15	2.50×10^{-1}
CTQ_17	-6.01×10^{-3}	1.96×10^{-2}	-3.07×10^{-1}	7.59×10^{-1}
CTQ_20	1.22×10^{-2}	1.87×10^{-2}	6.55×10^{-1}	5.13×10^{-1}
CTQ_21	-6.06×10^{-3}	1.97×10^{-2}	-3.08×10^{-1}	7.58×10^{-1}
CTQ_23	2.21×10^{-3}	1.66×10^{-2}	1.33×10^{-1}	8.94×10^{-1}
CTQ_24	-1.43×10^{-3}	1.97×10^{-2}	-7.27×10^{-2}	9.42×10^{-1}
CTQ_27	-2.60×10^{-2}	1.55×10^{-2}	-1.68	9.35×10^{-2}
CTQ_05	-8.24×10^{-3}	7.39×10^{-3}	-1.12	2.65×10^{-1}
CTQ_07	1.74×10^{-3}	9.41×10^{-3}	1.85×10^{-1}	8.53×10^{-1}
CTQ_13	**-2.00×10^{-2}**	**8.55×10^{-3}**	**-2.34**	**1.95×10^{-2}**
CTQ_19	**1.62×10^{-2}**	**7.91×10^{-3}**	**2.05**	**4.09×10^{-2}**
CTQ_28	-1.04×10^{-2}	9.00×10^{-3}	-1.15	2.49×10^{-1}
CTQ_01	-1.26×10^{-2}	7.73×10^{-3}	-1.62	1.05×10^{-1}
CTQ_02	8.51×10^{-3}	8.57×10^{-3}	9.93×10^{-1}	3.21×10^{-1}
CTQ_04	-2.34×10^{-3}	9.61×10^{-3}	-2.44×10^{-1}	8.08×10^{-1}
CTQ_06	-1.72×10^{-3}	1.46×10^{-2}	-1.17×10^{-1}	9.07×10^{-1}
CTQ_26	3.07×10^{-3}	7.27×10^{-3}	4.22×10^{-1}	6.73×10^{-1}
CTQ_10	1.47×10^{-2}	1.86×10^{-2}	7.92×10^{-1}	4.29×10^{-1}
CTQ_16	-7.79×10^{-3}	1.92×10^{-2}	-4.05×10^{-1}	6.85×10^{-1}
CTQ_22	-2.27×10^{-2}	1.87×10^{-2}	-1.21	2.26×10^{-1}
Age	**-3.80×10^{-3}**	**9.98×10^{-4}**	**-3.81**	**1.55×10^{-4}**
Male sex	0	0	NaN	NaN
Female sex	**-5.39×10^{-2}**	**1.21×10^{-2}**	**-4.44**	**1.08×10^{-5}**
HC	-1.59×10^{-2}	1.75×10^{-2}	-9.12×10^{-1}	3.62×10^{-1}
ROP	0	0	NaN	NaN
ROD	-2.18×10^{-2}	1.88×10^{-2}	-1.16	2.48×10^{-1}
CHR	-1.62×10^{-2}	1.88×10^{-2}	-8.59×10^{-1}	3.91×10^{-1}

Table S5: Linear regression using IQR to predict latent phenotypic and brain scores of the main study model. Depicted are the results of the linear regression analysis using the IQR measures of the participants to predict their phenotypic and brain scores of all five LV. Abbreviations: R^2, adjusted R-squared (percentage of explained variance). Significant *P* values are highlighted in bold font (after FDR-correction for multiple testing).

	Phenotypic score		Brain score	
	R^2	*P* Value	R^2	*P* Value
LV1	1.2%	**3.28×10^{-2}**	0.4%	8.04×10^{-1}
LV2	1.2%	**3.27×10^{-2}**	0.2%	5.21×10^{-1}
LV3	3.6%	**1.01×10^{-5}**	1.1%	5.00×10^{-2}
LV4	0.2%	9.9×10^{-1}	1.3%	**2.20×10^{-2}**
LV5	0.3%	5.00×10^{-1}	2.0%	**2.30×10^{-3}**

Table S6: Clinical and demographic characteristics of the replication sample. Abbreviations: HC, healthy control; ROD, recent-onset of depression; CHR, clinical high-risk state; ROP, recent-onset of psychosis; SD, standard deviation; NA, not available; GAF:S, Global Assessment of Functioning Social Scale; GAF:D/I, GAF Disability/Impairment Scale; GF:S, Global Functioning Social Scale; GF:R, GF Role Scale; PANSS, Positive and Negative Symptom Scale; BDI, Beck Depression Inventory. Significant *P* values are highlighted in bold font (after FDR-correction for multiple testing).

	All	HC	ROD	CHR	ROP	H/χ^2	*P* Value
Age, years, mean	24.82	24.74	25.35	23.35	26.03	10.30 [a]	.05
SD	5.80	5.04	6.19	5.41	6.44		
Sex, women/men, %	46	42	51	43	50	1.43 [b]	.77
Years of education, years, mean	14.65	15.74	13.90	13.49	15.09	1.02 [a]	.85
SD	6.24	3.12	2.95	2.37	12.23		
GAF:S Past Month, mean	60.33	87.75	51.07	48.51	42.58	23.06 [a]	**8.47x10-05**
SD	22.02	5.67	13.10	10.45	15.28		
GAF:D/I Past Month, mean	61.29	87.70	52.32	48.27	46.11	14.40 [a]	**.01**
SD	21.50	5.33	14.31	10.02	15.65		
GF:S Current, mean	6.76	8.53	6.31	5.88	5.56	13.90 [a]	**.01**
SD	1.78	0.77	1.29	1.50	1.57		
GF:R Current, mean	6.41	8.59	5.66	5.40	5.05	5.26 [a]	.62
SD	2.23	0.66	1.82	1.97	2.01		
Handedness, right-handed, %	87	89	88	87	84	0.42 [b]	.90
PANSS total, mean	58.34	NaN	51.22	58.32	66.18	31.94 [a]	**9.96x10-07**
SD	16.75	NaN	12.45	15.53	18.61		
PANSS positive, mean	12.85	NaN	8.33	12.27	18.47	140.75 [a]	**2.35x10-30**
SD	5.88	NaN	1.99	3.91	5.99		
PANSS negative, mean	14.74	NaN	13.48	15.15	15.65	3.87 [a]	.52
SD	6.75	NaN	5.73	7.16	7.10		
PANSS general, mean	30.79	NaN	29.34	30.92	32.23	4.28 [a]	.50
SD	8.74	NaN	7.87	8.41	9.72		
BDI total, mean	18.64	3.86	28.22	27.33	21.47	14.34 [a]	**.01**
SD	15.06	6.03	12.26	12.06	12.91		
Study center						410.65 [b]	**7.30x10-83**
Munich	107	5	25	39	38		
Basel	21	17	4	0	0		
Cologne	33	22	2	6	3		
Birmingham	25	6	5	6	8		
Milan	34	8	8	13	5		
Turku	50	28	3	9	10		
Udine	37	29	2	3	3		
Bari	32	11	5	10	6		
Duesseldorf	26	0	15	6	5		
Muenster	62	9	27	12	14		
Total	427	135	96	104	92		

[a] Kruskal-Wallis-Test (H test), [b] χ^2-test

Table S7: Group-level Dunn's tests for sociodemographic and clinical differences. Abbreviations ROD, recent-onset of depression; CHR, clinical high-risk state; ROP, recent-onset of psychosis; GAF:S, Global Assessment of Functioning Social Scale; GAF:D/I, GAF Disability/Impairment Scale; GF:S, Global Functioning Social Scale; GF:R, GF Role Scale; PANSS, Positive and Negative Symptom Scale; BDI, Beck Depression Inventory; LB, Lower Boundary; Diff., Difference; UB, Upper Boundary. Significant *P* values are highlighted in bold font (after FDR-correction for multiple testing).

Group 1	Group 2	LB	Diff.	UB	*P* Value
Age					
ROD	CHR	4.79	38.14	71.48	**.02**
ROD	ROP	-28.86	3.97	36.80	.99
CHR	ROP	-67.33	-34.17	-1.01	**.04**
Years of education					
ROD	CHR	-32.08	0.76	33.59	.99
ROD	ROP	-60.66	-27.64	5.38	.13
CHR	ROP	-60.96	-28.40	4.17	.11
GAF:S					
ROD	CHR	-28.58	4.60	37.78	.98
ROD	ROP	80.37	113.16	145.95	**<10⁻³**
CHR	ROP	75.44	108.56	141.68	**<10⁻³**
GAF:D/I					
ROD	CHR	-29.84	3.35	36.55	.99
ROD	ROP	61.05	93.85	126.65	**<10⁻³**
CHR	ROP	57.37	90.50	123.63	**<10⁻³**
GF:S					
ROD	CHR	32.80	64.92	97.05	**<10⁻³**
ROD	ROP	-24.84	7.47	39.78	.93
CHR	ROP	-89.45	-57.45	-25.46	**<10⁻³**
GF:R					
ROD	CHR	27.20	59.48	91.77	**<10⁻³**
ROD	ROP	-39.78	-7.31	25.16	.93
CHR	ROP	-98.95	-66.79	-34.64	**<10⁻³**
PANSS total					
ROD	CHR	-130.95	-98.10	-65.25	**<10⁻³**
ROD	ROP	-11.02	22.02	55.06	.30
CHR	ROP	87.54	120.12	152.70	**<10⁻³**
PANSS positive					
ROD	CHR	-140.19	-107.91	-75.63	**<10⁻³**
ROD	ROP	51.01	83.48	115.95	**<10⁻³**
CHR	ROP	159.31	191.39	223.47	**<10⁻³**
PANSS negative					
ROD	CHR	-91.41	-58.64	-25.88	**<10⁻³**
ROD	ROP	-41.21	-8.25	24.71	.91
CHR	ROP	17.82	50.39	82.96	**<10⁻³**
PANSS general					
ROD	CHR	-113.55	-80.72	-47.89	**<10⁻³**
ROD	ROP	-26.86	6.16	39.18	.96
CHR	ROP	54.32	86.88	119.44	**<10⁻³**
BDI					
ROD	CHR	-33.90	-0.88	32.14	.99
ROD	ROP	6.30	38.80	71.29	**.01**
CHR	ROP	6.84	39.67	72.50	**.01**

Table S8: Individual χ2-comparisons of differences in sex and handedness. Abbreviations: ROD, recent-onset of depression; CHR, clinical high-risk state; ROP, recent-onset of psychosis. Significant *P* values are highlighted in bold font (after FDR-correction for multiple testing).

Group 1	Group 2	χ²	*P* Value
Sex			
ROD	CHR	0.87	.35
ROD	ROP	7.05	**.02**

CHR	ROP	2.88	.13

Handedness			
ROD	CHR	0.19	.80
ROD	ROP	0.04	.85
CHR	ROP	0.38	.80

Table S9: Individual χ2-comparisons of recruitment per site. Significant *P* values are highlighted in bold font (after FDR-correction for multiple testing).

Group 1	Group 2	χ^2	*P* Value
Munich	Basel	35.51	**<10^{-3}**
Munich	Cologne	8.01	**<10^{-3}**
Munich	Birmingham	39.08	**<10^{-3}**
Munich	Milan	95.12	**<10^{-3}**
Munich	Turku	44.90	**<10^{-3}**
Munich	Udine	58.28	**<10^{-3}**
Basel	Cologne	10.27	**<10^{-3}**
Basel	Birmingham	0.10	.75
Basel	Milan	18.26	**<10^{-3}**
Basel	Turku	0.63	.43
Basel	Udine	3.32	.07
Cologne	Birmingham	12.33	**<10^{-3}**
Cologne	Milan	52.60	**<10^{-3}**
Cologne	Turku	15.85	**<10^{-3}**
Cologne	Udine	24.67	**<10^{-3}**
Birmingham	Milan	15.80	**<10^{-3}**
Birmingham	Turku	0.23	.63
Birmingham	Udine	2.28	.13
Milan	Turku	12.33	**<10^{-3}**
Milan	Udine	6.31	**.01**
Turku	Udine	1.06	.30

Table S10: Group-level Kruskal-Wallis statistics for psychopharmacological differences. Abbreviations: HC, healthy control; ROD, recent-onset of depression; CHR, clinical high-risk state; ROP, recent-onset of psychosis. Depicted is the count of members of the respective study population that received at least one substance from the psychopharmacological class. Significant *P* values are highlighted in bold font (after FDR-correction for multiple testing).

	HC	ROP	ROD	CHR	χ^2	*P* Value
Antidepressants	2	33	95	65	61.97	**<10^{-3}**
Antipsychotics	0	99	25	29	104.44	**<10^{-3}**
Sedatives	4	54	36	20	19.23	**<10^{-3}**
Anticonvulsants	0	5	7	3	1.53	.47

Table S11: Group-level Dunn's tests for psychopharmacological differences. Abbreviations: ROD, recent-onset of depression; CHR, clinical high-risk state; ROP, recent-onset of psychosis; LB, Lower Boundary; Diff., Difference; UB, Upper Boundary. Significant *P* values are highlighted in bold font (after FDR-correction for multiple testing).

Group 1	Group 2	LB	Diff.	UB	*P* Value
Antidepressants					
ROP	ROD	-122.13	-93.64	-65.15	**<10⁻³**
ROP	CHR	-81.56	-52.78	-24.00	**<10⁻³**
ROD	CHR	11.92	40.86	69.80	**<10⁻³**
Antipsychotics					
ROP	ROD	79.19	107.07	134.95	**<10⁻³**
ROP	CHR	71.19	99.36	127.52	**<10⁻³**
ROD	CHR	-36.04	-7.71	20.61	.89
Sedatives					
ROP	ROD	-0.71	25.03	50.77	.06
ROP	CHR	21.70	47.70	73.70	**<10⁻³**
ROD	CHR	-3.48	22.67	48.82	.11
Anticonvulsants					
ROP	ROD	-14.18	-3.15	7.87	.87
ROP	CHR	-8.50	2.63	13.77	.92
ROD	CHR	-5.41	5.79	16.99	.52

Table S12: Group-level Dunn's tests for CTQ differences. Abbreviations: ROD, recent-onset of depression; CHR, clinical high-risk state; ROP, recent-onset of psychosis; LB, Lower Boundary; Diff., Difference; UB, Upper Boundary. Significant *P* values are highlighted in bold font (after FDR-correction for multiple testing).

Group 1	Group 2	LB	Diff.	UB	*P* Value
CTQ emotional abuse					
ROD	CHR	-58.13	-24.73	8.67	.21
ROD	ROP	-61.67	-28.80	4.08	.11
CHR	ROP	-37.28	-4.07	29.14	.99
CTQ physical abuse					
ROD	CHR	-64.40	-31.17	2.06	.07
ROD	ROP	-53.04	-20.33	12.38	.36
CHR	ROP	-22.20	10.84	43.88	.82
CTQ sexual abuse					
ROD	CHR	-37.50	-8.45	20.59	.86
ROD	ROP	-42.28	-13.70	14.89	.58
CHR	ROP	-34.12	-5.24	23.64	.96
CTQ emotional neglect					
ROD	CHR	-31.52	-6.44	18.64	.90
ROD	ROP	-41.93	-17.24	7.45	.26
CHR	ROP	-35.75	-10.81	14.13	.66
CTQ physical neglect					
ROD	CHR	-51.69	-18.36	14.97	.47
ROD	ROP	-41.72	-8.90	23.91	.89
CHR	ROP	-23.69	9.46	42.61	.87
CTQ denial					
ROD	CHR	-51.71	-18.97	13.76	.42
ROD	ROP	-74.18	-41.96	-9.73	**.01**
CHR	ROP	-55.54	-22.99	9.57	.25
CTQ total					
ROD	CHR	-16.88	8.55	33.97	.81
ROD	ROP	-27.73	-2.70	22.32	.99
CHR	ROP	-36.53	-11.25	14.03	.64

Table S13: Detailed list of weighted phenotypic features from LV1-LV5.

LV	CT domain	CTQ item
LV1	emotional abuse	I thought that my parents wished I hadn't been born.
		I believe that I was emotionally abused.
	physical abuse	I got hit so hard by someone in my family that I had to see a doctor or go to the hospital.
		People in my family hit me so hard that it left me with bruises or marks.
		I was punished with a belt, a board, a cord, or some other hard object.
		I believe that I was physically abused.
	sexual abuse	Someone tried to touch me in a sexual way. Or tried to make me touch them.
		Someone threatened to hurt me or tell lies about me unless I did something sexual with them.
		Someone tried to make me do sexual things or watch sexual things.
		Someone molested me.
		I believe that I was sexually abused.
	physical neglect	My parents were too drunk or high to take care of the family.
	denial	I had the best family in the world.
	other	Age, Male sex, Female sex, ROP
LV2	sexual abuse	Someone threatened to hurt me or tell lies about me unless I did something sexual with them.
		Someone molested me.
	other	Age
LV3	emotional abuse	I believe that I was emotionally abused.
	physical abuse	I got hit so hard by someone in my family that I had to see a doctor or go to the hospital.
		I was punished with a belt, a board, a cord, or some other hard object.
	sexual abuse	Someone tried to touch me in a sexual way. Or tried to make me touch them.
		Someone threatened to hurt me or tell lies about me unless I did something sexual with them.
		Someone tried to make me do sexual things or watch sexual things.
		I believe that I was sexually abused.
		There was someone in my family who helped me feel that I was important or special. (inverted question)
	emotional neglect	
		My parents were too drunk or high to take care of the family.
		There was someone to take me to the doctor if I needed it. (inverted question)
	denial	I had the best family in the world.
	other	Age, Male sex, Female sex
LV4	physical abuse	I got hit so hard by someone in my family that I had to see a doctor or go to the hospital.
		I was punished with a belt, a board, a cord, or some other hard object.
		I believe that I was physically abused.
	sexual abuse	Someone tried to touch me in a sexual way. Or tried to make me touch them.
		Someone tried to make me do sexual things or watch sexual things.
		Someone molested me.
		I believe that I was sexually abused.
	other	Male sex, Female sex
LV5	emotional abuse	People in my family called me things like ''stupid'', ''lazy'' or ''ugly''.
		People in my family said hurtful or insulting things to me.
		I felt that someone in my family hated me.
	emotional neglect	I felt loved. (inverted question)
		People in my family looked out for each other. (inverted question)
		My family was a source of strength and support. (inverted question)

Table S14: Atlas readouts from LV1-LV5. Depicted are the brain patterns of LV1-LV5, read out by the Brainnetome atlas and the cerebellar atlas by Diedrichsen. For each LV, the brain pattern is specified by the percentage of total voxels in a certain region (%) in bilateral hemispheres (left/right) being either positively (+) or negatively (-) weighted. For better clarity, blank spaces are used to indicate that a region did not receive any weights in this instance.

Latent variables		LV1				LV2				LV3				LV4				LV5			
	Hemispheres	left		right		left		right		left		right		left		right		left		right	
	Weights	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-
Brainnetome atlas regions	Percentages	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)
Superior Frontal Gyrus	Total voxels																				
A8m, medial area 8	241		100		97	22		4		2	15	17	12		13		34	1	5		5
A8dl, dorsolateral area 8	232		99		100	4		2		32	3	14	2	2	11	9	2	2		3	
A9l, lateral area 9	224		100		100	21		24			43	5	16	11		8	8				
A6dl, dorsolateral area 6	184		75		87	3		8		90		96		1		1					
A6m, medial area 6	211		98		84	2				28	7	37	3		11	7	3		17		11
A9m,medial area 9	230		100		100	43		32		3	31	4	15	4	5	1	47	8	3		
A10m, medial area 10	266		100		100	46		23		7	37	4	15	9	1	1	20				
Middle Frontal Gyrus																					
A9/46d, dorsal area 9/46	301		100		94	8		3		13	8	19	4			9	8	9			
IFJ, inferior frontal junction	196		93		83	1	2	1		40	4	46	2	25	3	14	1	1			
A46, area 46	241		100		100	11		9		1	13	10	7	1	10		9				
A9/46v, ventral area 9/46	279		100		98	8		9		3	15		12	9	1	6		1			
A8vl, ventrolateral area 8	248		100		95	2		1		58		32	3		5	2	1	2			
A6vl, ventrolateral area 6	168		87		88	1		10		95		86				3					
A10l, lateral area10	220		100		100	2		7		20	5		12	4	1	4					
Inferior Frontal Gyrus																					
A44d, dorsal area 44	94		100		98		1	1			5	3	2	54		43		1		1	
IFS, inferior frontal sulcus	101		100		100		3		10		24		31	31			4	2			
A45c, caudal area 45	88		100		98	1		3	1		1	14	7	49		31	3	1			
A45r, rostral area 45	105		100		100	2		4			17		24	11	1	11	6				
A44op, opercular area 44	148		87		100	1		1		49	4	9	7	7	2	4	3		19		
A44v, ventral area 44	84		100		99	24	1	19		5	9	19	4	73		64					1
Orbital Gyrus																					
A14m, medial area 14	165		100		100	1			6		9		6		6		37				7
A12/47o, orbital area 12/47	137		100		100				12	11	8	13	38		2	1	1				
A11l, lateral area 11	308		100		100				2	3	18	6	20		3	2	1				
A11m, medial area 11	158		100		99				1		10		8		1	2	11				13
A13, area 13	236		99		100	2			3	32	2	5	11	6	6		11		16		
A12/47l, lateral area 12/47	153		98		100				6	22	10	28	22	2	1	4	1				
Precentral Gyrus																					
A4hf, area 4(head and face region)	180		71		89	1		3		15	13	13	27	25		42		8			2
A6cdl, caudal dorsolateral area 6	226		79		95	1	20			53		61		27		20					
A4ul, area 4(upper limb region)	166		91		86	1	1				13	2	33	1		25		10		10	
A4t, area 4(trunk region)	103		70		63			2		50		52		4		14		4		23	
A4tl, area 4(tongue and larynx region)	118		100		97				7	16	4	13	10	28	2	3	5				
A6cvl, caudal ventrolateral area 6	199		98		97	9				45		42		29	1	29					
Paracentral Lobule																					

A1/2/3ll, area1/2/3 (lower limb region)	106	85	77	2	2	3				2					12	13
A4ll, area 4, (lower limb region)	153	67	60	4	2	23	3	23	1	22		37	2	1	2	4
Superior Temporal Gyrus																
A38m, medial area 38	194	15	26			49		4			34		4			8
A41/42, area 41/42	123	67	91		1	63		34	2		7	1	1			
TE1.0 and TE1.2	218	79	100	3	9	12	35	15	3	26	8	6	1	8		
A22c, caudal area 22	123	91	92	1	9	32	2	59	2		16		4	1		
A38l, lateral area 38	165	90	94	6	1	38	1	6	1		26		17	2		
A22r, rostral area 22	141	93	96	1		30	1	32	5		6		2	3		
Middle Temporal Gyrus																
A21c, caudal area 21	164	100	100	14	4	22	9	1	10		1		13			
A21r, rostral area 21	252	81	99		2	66		38			19		35			
A37dl, dorsolateral area37	172	97	97	7	2	7	16	40	2		2		7		3	
aSTS, anterior superior temporal sulcus	286	100	100	8	8	20	13	26	14		10		11	1		1
Inferior Temporal Gyrus																
A20iv, intermediate ventral area 20	70	46	33		2			41							16	2
A37elv, extreme lateroventral area37	68	81	91			23	1	46			9		2	1		
A20r, rostral area 20	129	3	37		1	76		39			24		9			1
A20il, intermediate lateral area 20	121	91	88		1	78		19	1				8			7
A37vl, ventrolateral area 37	100	91	85	6	1		15	60			7		8			
A20cl, caudolateral of area 20	124	100	100	1		33	1	1	2	6			1			
A20cv, caudoventral of area 20	148	56	93			24					1		1			
Fusiform Gyrus																
A20rv, rostroventral area 20	297	33	33			9	1	12		8	52		5		8	11
A37mv, medioventral area37	235	97	94	2		15		40		9	5	5	2			
A37lv, lateroventral area37	262	86	81		1	37		61			9			1		
Parahippocampal Gyrus																
A35/36r, rostral area 35/36	46							14			100		2			
A35/36c, caudal area 35/36	48	53	51			4		16	10	11	7	2	2			2
TL, area TL (lateral PPHC, posterior parahippocampal gyrus)	46	80	75			12	8	10	3	14	4					
A28/34, area 28/34 (EC, entorhinal cortex)	45	2						13		22	77		54			
TI, area TI(temporal agranular insular cortex)	29										85		47			41
TH, area TH (medial PPHC)	40	100	100		2	13	3				31		12			
Posterior Superior Temporal Sulcus																
rpSTS, rostroposterior superior temporal sulcus	105	100	100		6	1	46		35		7	16	3			
cpSTS, caudoposterior superior temporal sulcus	91	99	100	1	1	2	30	1	4			27	3			
Superior Parietal Lobule																
A7r, rostral area 7	135	53	81			25		4				1				
A7c, caudal area 7	135	83	74			35		31		5		3			1	
A5l, lateral area 5	117	70	93				3	5	15	7		3				
A7pc, postcentral area 7	133	58	61			3		4		3		4	2			4
A7ip, intraparietal area 7(hIP3)	128	74	90		1	24			7	16		12				
Inferior Parietal Lobule																
A39c, caudal area 39(PGp)	296	83	78	1		9	10	35		9		1	6			
A39rd, rostrodorsal area 39(Hip3)	252	86	87			3	12	6	2	16	1	35				
A40rd, rostrodorsal area 40(PFt)	280	97	96	1		4	8		30	18	1	15			5	1

	N														
A40c, caudal area 40(PFm)	274	100	100	2		2	10	1	3		1	31		4	
A39rv, rostroventral area 39(PGa)	396	100	99	6	1	16	11	19	2	2	3	18	1		1
A40rv, rostroventral area 40(PFop)	324	85	86			63			1		1	1		2	1
Precuneus															
A7m, medial area 7(PEp)	140	94	100			21	1	19	2						
A5m, medial area 5(PEm)	176	81	75			4	4	3	2						9
dmPOS, dorsomedial parietooccipital sulcus(PEr)	267	73	70			39		71		1	9	1	11	3	3
A31, area 31 (Lc1)	228	95	100			94		89			16		21	3	
Postcentral Gyrus															
A1/2/3ulhf, area 1/2/3(upper limb, head, and face region)	251	94	86	3			37		33	51		54		28	
A1/2/3tonla, area 1/2/3(tongue and larynx region)	158	99	99	15		39		1	5	59		11			7
A2, area 2	212	98	88	3			9	5	8	22	2	24		2	
A1/2/3tru, area1/2/3(trunk region)	181	82	68					2	2	6		1		4	11
Insular Gyrus															
G, hypergranular insula	85	82	97			34	2		4	4					
vIa, ventral agranular insula	64	100	98	1		4	1	26			1				
dIa, dorsal agranular insula	71	87	100	3	4	73	4	43	3		7	1		13	
vId/vIg, ventral dysgranular and granular insula	78	93	76			12	5		6	28	1				
dIg, dorsal granular insula	84	100	100	1	6		7		4	5	5		2		
dId, dorsal dysgranular insula	96	100	100	1	1	29	9		8	4			1		
Cingulate Gyrus															
A23d, dorsal area 23	129	99	100			94		66			8		3		
A24rv, rostroventral area 24	61	98	100			7		8		74	2	27	27	17	15
A32p, pregenual area 32	122	100	100		3	1			15		5	7	9	2	8
A23v, ventral area 23	88	70	78			92		85			86		57		
A24cd, caudodorsal area 24	92	100	100	1	1		8		10	20	9	34	3	45	54
A23c, caudal area 23	147	100	100	5	14	1	6		8	1	16		13	51	41
A32sg, subgenual area 32	157	100	100	1	1	23	5		9		24		9	17	15
Medioventral Occipital Cortex															
cLinG, caudal lingual gyrus	156	32	26	3		62		62		21		2	1		7
rCunG, rostral cuneus gyrus	249	52	42	7	5	50		68		48		29		35	44
cCunG, caudal cuneus gyrus	166	28	10		2	8		42		59		23		1	46
rLinG, rostral lingual gyrus	243	64	85	4	2	49		41		15		9			
vmPOS,ventromedial parietooccipital sulcus	268	19	57	10		56		30		28	6			9	3
Lateral Occipital Cortex															
mOccG, middle occipital gyrus	211	85	49	8	6	31	2	34		47		32	1		6
V5/MT+, area V5/MT+	209	91	73		3	17	4	34		1	7				4
OPC, occipital polar cortex	221	61	2			4		19		22		5			28
iOccG, inferior occipital gyrus	259	81	53	1	2	14	1	65		1	5		16		2
msOccG, medial superior occipital gyrus	161	34	7			2		13		34		29			13
lsOccG, lateral superior occipital gyrus	182	43	41			1	3	5		24		42		1	
Amygdala															
mAmyg, medial amygdala	50	7	12						12		7	3	7		3
lAmyg, lateral amygdala	31		8	8					35		50	3	8		
Hippocampus															
rHipp, rostral hippocampus	149	3	1	1			34		69	2	39	2	15		10
cHipp, caudal hippocampus	167	13	19				33	2	40	20		16			2

Basal Ganglia / Thalamus / Cerebellar regions

Region	Total voxels	G1 +	G1 +c	G1 −	G1 −c	G2 +	G2 +c	G2 −	G2 −c	G3 +	G3 +c	G3 −	G3 −c	G4 +	G4 +c	G4 −	G4 −c	G5 +	G5 +c	G5 −	G5 −c
Basal Ganglia																					
vCa, ventral caudate	103	82		83	1					19	6	15	5	62		65					
GP, globus pallidus	60	64		75		2				84		78		56		31				11	
NAC, nucleus accumbens	112	85		87	7	2	1			53	1	27	3	32		15	6				
vmPu, ventromedial putamen	78	92		74						61		81		71		57				9	
dCa, dorsal caudate	128	81		72	5	14				10	1	7	4	100		84		1		11	
dlPu, dorsolateral putamen	132	91		93	6					46		49		34		9	1				
Thalamus																					
mPFtha, medial pre-frontal thalamus	46	100		100						21		7		92		74					
mPMtha, pre-motor thalamus	17		24	24								53				76					
Stha, sensory thalamus	24	33	20	28	25					33		16		80		53		13			
rTtha, rostral temporal thalamus	59	91		98						4		25		92		85				3	
PPtha, posterior parietal thalamus	55	45		75						2		2		91		80					
Otha, occipital thalamus	56	80		95						4				94		78		1		1	
cTtha, caudal temporal thalamus	44	70		78										96		95					
lPFtha, lateral pre-frontal thalamus	69	20	42	66						33		19		87		88		5		3	
Cerebellar atlas regions by Diedrichsen																					
Cerebellum																					
Cerebellum 1-4	263	57		65		15		13		39		47	1							1	
Cerebellum 5	281	78		80		10	1	8	1	36	3	42	4		2		4			1	
Cerebellum 6	318	82		87		13	2	13	1					4		1		1			
Cerebellum 7b	540	94		89								15		11		2					
Cerebellum 8a	400	70		74						36		51		4							
Cerebellum 8b	302									46		55									
Cerebellum 9	331	88		81								29									
Cerebellum 10	127																				
Crus 1	148	73		60		14		4		26		30		41	1	40		1		3	
Crus 2	186	84		83		16		12		3				45		49					1

		medial		medial		medial		medial		medial	
		+	−	+	−	+	−	+	−	+	−
	Total voxels	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)
Vermis Crus 1	254		75			20	2	16	3		
Vermis Crus 2	272		87	7		1	1	51			1
Vermis 6	281		57	2		10		5		1	
Vermis 7b	280		88			6		10	2		
Vermis 8a	474		86			9		2	1		
Vermis 8b	162		66			45		3			
Vermis 9	423		82			22		29			
Vermis 10	656		71					86			

Table S15: Spearman's correlation analyses between latent scores and clinical domains of functioning in the discovery and replication sample. Results are states as correlation coefficient ρ, followed by its *P* value in brackets: ρ (*P* value). Abbreviations: D, Discovery Sample; R, Replication Sample; GAF:S, Global Assessment of Functioning Social Scale; GAF:D/I, GAF Disability/Impairment Scale; GF:S, Global Functioning Social Scale; GF:R, GF Role Scale. Significant *P* values are highlighted in bold font (after FDR-correction for multiple testing).

		LV1		LV3	
		Age		Sex	
		Phenotypic score	Brain score	Phenotypic score	Brain score
GAF:S					
Lifetime	D	0.01 (.96)	0.01 (.99)	**-0.24 (<10⁻³)**	-0.06 (.54)
	R	0.04 (.6)	-0.1 (.2)	**-0.2 (<10⁻³)**	0.01 (.99)
Past Year	D	0.07 (.2)	0.02 (.75)	**-0.25 (<10⁻³)**	-0.06 (.55)
	R	-0.04 (.6)	**-0.14 (.03)**	**-0.16 (.01)**	0.05 (.54)
Past Month	D	**0.16 (<10⁻³)**	0.08 (.5)	**-0.19 (<10⁻³)**	-0.04 (.7)
	R	-0.05 (.57)	**-0.17 (<10⁻³)**	**-0.17 (<10⁻³)**	0.06 (.53)
GAF:D/I					
Lifetime	D	0.01 (.99)	0.01 (.95)	**-0.29 (<10⁻³)**	**-0.13 (.05)**
	R	0.03 (.77)	-0.08 (.48)	**-0.24 (<10⁻³)**	-0.01 (.93)
Past Year	D	0.04 (.77)	0.02 (.77)	**-0.29 (<10⁻³)**	-0.09 (.85)
	R	-0.03 (.73)	**-0.14 (.02)**	**-0.19 (<10⁻³)**	0.05 (.56)
Past Month	D	**0.15 (<10⁻³)**	0.06 (.54)	**-0.22 (<10⁻³)**	-0.07 (.5)
	R	-0.06 (.88)	**-0.18 (<10⁻³)**	**-0.14 (.03)**	0.06 (.5)
GF:S					
Current	D	**0.11 (.01)**	0.08 (.5)	**-0.26 (<10⁻³)**	-0.06 (.57)
	R	-0.11 (.13)	**-0.16 (.01)**	**-0.13 (.05)**	0.08 (.51)
Low Past Year	D	**0.12 (.01)**	0.03 (.68)	**-0.23 (<10⁻³)**	-0.07 (.5)
	R	**-0.14 (.03)**	**-0.17 (<10⁻³)**	**-0.13 (.04)**	0.1 (.21)
High Past Year	D	0.04 (.73)	0.03 (.69)	**-0.28 (<10⁻³)**	-0.06 (.54)
	R	-0.08 (.51)	**-0.14 (.03)**	**-0.14 (.03)**	0.06 (.93)
High Lifetime	D	0.01 (.84)	0.03 (.69)	**-0.25 (<10⁻³)**	-0.07 (.51)
	R	0.02 (.86)	-0.04 (.59)	**-0.14 (.03)**	-0.02 (.81)
GF:R					
Current	D	**0.14 (<10⁻³)**	0.09 (.42)	**-0.22 (<10⁻³)**	-0.04 (.72)
	R	-0.02 (.81)	**-0.19 (<10⁻³)**	-0.12 (.11)	0.07 (.69)
Low Past Year	D	**0.16 (<10⁻³)**	0.08 (.51)	**-0.21 (<10⁻³)**	-0.04 (.66)
	R	-0.06 (.5)	**-0.21 (<10⁻³)**	-0.12 (.09)	0.1 (.24)
High Past Year	D	0.04 (.5)	0.06 (.55)	**-0.27 (<10⁻³)**	-0.06 (.57)
	R	0 (.99)	**-0.17 (<10⁻³)**	**-0.18 (<10⁻³)**	0.08 (.57)
High Lifetime	D	0.01 (.91)	0.07 (.51)	**-0.25 (<10⁻³)**	-0.05 (.63)
	R	0.09 (.31)	-0.05 (.52)	**-0.23 (<10⁻³)**	-0.02 (.87)

Table S16: Spearman's correlation analyses between latent scores and clinical domains of depressivity, personality and quality of life in the discovery and replication sample. Results are states as correlation coefficient ρ, followed by its *P* value in brackets: ρ (*P* value). Abbreviations: D, Discovery Sample; R, Replication Sample; BDI, Beck Depression Inventory; NEO-FFI, Neuroticism-Extraversion-Openness (NEO) Five-Factor Inventory; WHOQOL-BREF, World Health Organization Quality of Life Short Version. Significant *P* values are highlighted in bold font (after FDR-correction for multiple testing).

		LV1		LV3	
		Age		Sex	
		Phenotypic score	Brain score	Phenotypic score	Brain score
BDI					
Total score	D	-0.09 (.06)	-0.09 (.64)	**0.18 (<10⁻³)**	0.02 (.98)
	R	0.03 (.77)	0.1 (.21)	**0.17 (.01)**	-0.04 (.63)
NEO-FFI					
Neuroticism	D	-0.01 (.89)	-0.01 (.88)	**0.2 (<10⁻³)**	0.06 (.56)
	R	-0.01 (.94)	0.04 (.65)	0.1 (.24)	-0.04 (.67)
Extraversion	D	0.07 (.24)	0.04 (.64)	**-0.12 (.04)**	-0.02 (.99)
	R	-0.01 (.91)	-0.03 (.71)	**-0.2 (<10⁻³)**	-0.04 (.69)
Openness	D	-0.07 (.18)	-0.02 (.79)	**-0.11 (.05)**	-0.06 (.55)
	R	-0.09 (.28)	0.03 (.71)	-0.04 (.65)	-0.07 (.73)
Agreeableness	D	-0.05 (.61)	-0.1 (.28)	**-0.29 (<10⁻³)**	**-0.19 (<10⁻³)**
	R	-0.03 (.7)	-0.05 (.55)	**-0.17 (.01)**	-0.04 (.64)
Conscientiousness	D	-0.05 (.59)	-0.04 (.63)	**-0.29 (<10⁻³)**	**-0.17 (<10⁻³)**
	R	**0.15 (.01)**	0.05 (.56)	**-0.33 (<10⁻³)**	**-0.13 (.05)**
WHOQOL-BREF					
Physical	D	**0.11 (.02)**	0.04 (.65)	**-0.17 (<10⁻³)**	-0.05 (.63)
	R	-0.08 (.44)	-0.11 (.12)	-0.1 (.25)	0.05 (.56)
Psychosocial	D	0.06 (.41)	0.04 (.63)	**-0.17 (<10⁻³)**	-0.04 (.67)
	R	0.01 (.93)	-0.07 (.68)	**-0.16 (.01)**	0.03 (.73)
Social Relationships	D	**0.09 (.05)**	0.07 (.51)	**-0.25 (<10⁻³)**	-0.06 (.53)
	R	-0.1 (.2)	-0.08 (.46)	**-0.17 (.01)**	-0.02 (.84)
Environment	D	**0.15 (<10⁻³)**	0.02 (.81)	**-0.17 (<10⁻³)**	-0.04 (.67)
	R	**-0.16 (.01)**	**-0.13 (.05)**	-0.09 (.35)	0.07 (.65)

Table S17: Spearman's correlation analyses between socioeconomic factors and latent scores in the discovery sample (Ethnicity, Urbanicity, Religion). Abbreviations: PL = Place of Living, PB = Place of Birth, Practice = Extent to which religion is actively practiced. All *P* values FDR-corrected for multiple testing (family of tests: Table S17, Table S18, Table S19, Table S20).

		Ethnicity				Urbanicity			Religion				
		Caucasian	Asian	Mixed	Other	Population PL	Density PL	Density PB	None	Christian	Muslim	Other	Practice
brain score LV 1	ρ	-0.04	0.04	0.01	0.02	**0.30**	-0.23	-0.26	0.15	-0.12	-0.09	-0.04	-0.02
	P	.83	.80	.98	.92	**.01**	.20	.05	.52	.55	.62	.82	.91
phenotypic score LV 1	ρ	-0.07	0.17	-0.11	0.06	0.04	0.06	0.03	-0.01	-0.02	-0.02	0.12	0.06
	P	.67	.50	.58	.72	.81	.74	.85	.97	.93	.88	.56	.73
brain score LV 2	ρ	-0.08	0.05	0.03	0.05	**0.28**	-0.19	-0.19	0.15	-0.14	-0.04	0.01	-0.01
	P	.66	.77	.86	.76	**.01**	.56	.56	.51	.52	.79	.95	.96
phenotypic score LV 2	ρ	-0.11	0.17	-0.10	0.11	0.10	-0.04	-0.07	0.00	-0.02	-0.01	0.09	-0.01
	P	.57	.50	.59	.57	.59	.80	.70	.99	.90	.96	.64	.97
brain score LV 3	ρ	0.01	0.01	-0.02	0.01	0.18	-0.17	-0.21	0.17	-0.14	-0.08	-0.07	-0.13
	P	.98	.98	.90	1.00	.50	.51	.49	.51	.53	.65	.71	.54
phenotypic score LV 3	ρ	-0.10	0.12	-0.04	0.08	0.09	-0.04	-0.09	0.06	-0.08	-0.01	0.07	-0.08
	P	.60	.56	.82	.66	.63	.82	.64	.71	.66	.97	.68	.67
brain score LV 4	ρ	-0.14	0.03	0.07	0.12	-0.20	0.08	-0.01	0.02	-0.07	0.10	0.15	-0.04
	P	.54	.87	.71	.56	.50	.68	.94	.92	.69	.60	.53	.84
phenotypic score LV 4	ρ	-0.20	0.17	-0.06	0.20	0.07	-0.02	-0.05	0.01	-0.04	0.09	0.06	0.01
	P	.50	.51	.73	.50	.70	.91	.76	.97	.84	.64	.74	.99
brain score LV 5	ρ	0.09	-0.03	-0.15	-0.01	**-0.26**	0.09	0.03	-0.07	0.03	0.07	0.09	-0.09
	P	.61	.87	.52	.97	**.04**	.63	.87	.70	.85	.70	.61	.61
phenotypic score LV 5	ρ	-0.04	0.03	0.01	0.04	-0.03	-0.06	-0.09	-0.02	0.03	0.08	-0.09	-0.13
	P	.80	.87	.98	.83	.88	.75	.61	.91	.87	.65	.61	.53

Table S18: Spearman's correlation analyses between socioeconomic factors and latent scores in the discovery sample (Parental Education Background). Abbreviations: Secondary School = Parent finished Secondary School; Primary School = Parent only finished Primary School; Highest Occupation = Highest Occupation every achieved; M = Mother, F = Father. All *P* values FDR-corrected for multiple testing (family of tests: Table S17, Table S18, Table S19, Table S20).

		Secondary School (M/F)		Primary School (M/F)		No graduation (M/F)		University degree (M/F)		Highest Occupation (M/F)	
brain score LV 1	ρ	-0.03	-0.01	0.05	-0.05	-0.07	0.07	-0.05	-0.04	-0.22	-0.10
	P	.86	.95	.78	.78	.69	.70	.78	.81	.24	.60
phenotypic score LV 1	ρ	0.03	-0.02	0.07	0.10	-0.22	-0.17	0.19	0.07	0.01	0.08
	P	.88	.92	.68	.60	.27	.50	.88	.71	.98	.66
brain score LV 2	ρ	0.01	-0.03	0.06	0.03	-0.09	-0.03	0.00	-0.07	-0.20	-0.11
	P	.95	.86	.73	.85	.61	.85	.99	.70	.45	.58
phenotypic score LV 2	ρ	-0.01	-0.06	0.12	0.10	-0.18	-0.10	0.12	-0.01	-0.07	-0.01
	P	.93	.72	.55	.60	.83	.60	.54	.97	.68	.96
brain score LV 3	ρ	-0.05	-0.04	0.12	0.06	-0.07	-0.04	-0.06	-0.09	-0.18	-0.13
	P	.79	.80	.56	.73	.68	.80	.73	.64	.50	.55
phenotypic score LV 3	ρ	-0.04	-0.07	0.15	0.11	-0.17	-0.10	0.04	-0.08	-0.05	-0.07
	P	.82	.68	.53	.58	.51	.59	.80	.67	.79	.70
brain score LV 4	ρ	0.10	-0.02	-0.08	-0.01	-0.10	0.05	0.12	0.01	-0.05	0.01
	P	.61	.92	.67	.98	.62	.77	.57	.95	.76	.96
phenotypic score LV 4	ρ	-0.05	-0.04	0.09	0.03	-0.05	0.05	0.07	-0.01	-0.15	-0.03
	P	.78	.82	.62	.85	.76	.79	.70	.94	.53	.87
brain score LV 5	ρ	0.05	-0.04	0.11	0.17	-0.20	-0.10	0.06	0.09	0.02	0.10
	P	.76	.81	.57	.50	.53	.59	.71	.62	.91	.60
phenotypic score LV 5	ρ	-0.07	-0.14	**0.27**	0.20	-0.20	0.01	-0.11	-0.04	-0.18	-0.12
	P	.70	.53	**.03**	.46	.53	.96	.57	.82	.83	.56

Table S19: Spearman's correlation analyses between socioeconomic factors and latent scores in the discovery sample (Family & Relationships). Abbreviations: Nr. Sibling = Number of Siblings, Nr. Children = Number of own children, Nr. People in household = Number of people living in the same household, Partnership>1yr = has the participant ever been in a partnership over one year. All *P* values FDR-corrected for multiple testing (family of tests: Table S17, Table S18, Table S19, Table S20).

		Family				Relationship status last 12 months			
		Nr. Siblings	Nr. Children	Nr. people in household	Partnership>1yr	Single	Married	Partnership	Separated
brain score LV 1	ρ	0.04	-0.05	0.14	0.03	-0.01	0.01	0.05	-0.08
	P	.81	.78	.52	.86	.96	.95	.78	.65
phenotypic score LV 1	ρ	0.04	**-0.33**	0.08	-0.15	0.21	**-0.33**	0.07	-0.14
	P	.80	**1.18x10⁻⁰³**	.67	.52	.33	**1.15x10⁻⁰³**	.67	.53
brain score LV 2	ρ	-0.06	-0.07	0.09	0.06	0.02	-0.03	0.05	-0.08
	P	.74	.68	.62	.74	.89	.86	.75	.64
phenotypic score LV 2	ρ	0.11	**-0.29**	0.12	-0.15	0.23	**-0.32**	0.03	-0.04
	P	.57	**.01**	.55	.52	.16	**1.78x10⁻⁰³**	.84	.79
brain score LV 3	ρ	0.12	-0.03	0.14	0.04	-0.07	-0.02	0.12	-0.04
	P	.56	.86	.54	.82	.71	.91	.56	.80
phenotypic score LV 3	ρ	0.10	-0.20	0.14	-0.16	0.21	-0.27	-0.01	-0.01
	P	.61	.85	.53	.51	.50	.05	.95	.94
brain score LV 4	ρ	-0.12	-0.13	-0.09	-0.06	0.00	-0.07	0.05	0.11
	P	.57	.56	.64	.73	.99	.69	.78	.59
phenotypic score LV 4	ρ	0.10	-0.28	0.04	-0.10	0.13	-0.25	0.03	0.13
	P	.62	.05	.83	.61	.55	.23	.85	.55
brain score LV 5	ρ	0.14	-0.09	-0.16	-0.01	0.07	-0.26	0.07	-0.03
	P	.53	.62	.50	.97	.67	.05	.67	.87
phenotypic score LV 5	ρ	0.15	0.03	0.01	-0.08	0.14	-0.16	-0.09	0.14
	P	.52	.88	.94	.67	.53	.50	.62	.52

Table S20: Spearman's correlation analyses between socioeconomic factors and latent scores in the discovery sample (Education, Work). Abbreviations: Secondary School = finished Secondary School; Primary School = only finished Primary School; Highest Occupation = Highest Occupation every achieved; All *P* values FDR-corrected for multiple testing (family of tests: Table S17, Table S18, Table S19, Table S20).

		Education					Work				
		Years of Education	Secondary School	Primary School	No graduation	University Degree	Highest Occupation	Full-time current	Part-time current	Full-time last 12 months	Part-time last 12 months
brain score LV 1	ρ	-0.11	-0.05	0.05	0.08	-0.01	-0.21	-0.04	0.04	-0.01	0.01
	P	.58	.78	.79	.65	.96	.32	.82	.82	.94	.94
phenotypic score LV 1	ρ	-0.23	-0.08	0.01	0.04	-0.08	-0.21	-0.05	0.05	-0.10	0.10
	P	.20	.65	.99	.83	.66	.41	.76	.76	.60	.60
brain score LV 2	ρ	-0.09	0.01	0.01	0.03	0.03	-0.13	-0.01	0.01	0.02	-0.02
	P	.61	.98	.99	.84	.87	.54	.96	.96	.93	.93
phenotypic score LV 2	ρ	**-0.30**	-0.19	0.12	-0.02	-0.14	-0.25	-0.08	0.08	-0.13	0.13
	P	**4.97x10⁻⁰³**	.58	.56	.92	.52	.05	.66	.66	.53	.53
brain score LV 3	ρ	-0.24	-0.10	0.06	0.00	-0.07	**-0.27**	0.11	-0.11	0.10	-0.10
	P	.19	.60	.72	.99	.68	**.04**	.59	.59	.61	.61
phenotypic score LV 3	ρ	**-0.33**	-0.20	0.14	-0.04	-0.15	**-0.29**	0.01	-0.01	-0.07	0.07
	P	**1.51x10⁻⁰³**	.91	.54	.81	.52	**.02**	.97	.97	.69	.69
brain score LV 4	ρ	-0.07	-0.06	-0.06	0.11	0.02	0.08	0.04	-0.04	0.07	-0.07
	P	.69	.74	.73	.59	.92	.67	.82	.82	.71	.71
phenotypic score LV 4	ρ	**-0.29**	-0.20	0.12	-0.04	-0.13	-0.19	-0.10	0.10	-0.11	0.11
	P	**.04**	.50	.56	.83	.56	.50	.60	.60	.60	.60
brain score LV 5	ρ	-0.05	-0.09	0.02	-0.04	0.01	0.04	0.16	-0.16	0.13	-0.13
	P	.77	.61	.91	.80	.95	.82	.51	.51	.53	.53
phenotypic score LV 5	ρ	-0.10	-0.15	0.13	-0.08	-0.05	-0.10	0.07	-0.07	0.01	0.01
	P	.60	.51	.54	.64	.75	.60	.69	.69	.99	.99

Table S21: Spearman's correlation analyses between socioeconomic factors and latent scores in the replication sample (Ethnicity, Urbanicity, Religion). Abbreviations: PL = Place of Living, PB = Place of Birth, Practice = Extent to which religion is actively practiced, NaN = Not available due to too few cases. All *P* values FDR-corrected for multiple testing. All *P* values FDR-corrected for multiple testing (family of tests: Table S21, Table S22, Table S23, Table S24).

		Ethnicity				Urbanicity			Religion				
		Caucasian	Asian	Mixed	Other	Population PL	Density PL	Density PB	None	Christian	Muslim	Other	Practice
brain score LV 1	ρ	0.11	NaN	NaN	-0.11	-0.23	0.28	0.15	-0.30	0.28	NaN	0.07	0.37
	P	.67	NaN	NaN	.67	.51	.58	.59	.38	.73	NaN	.78	.05
phenotypic score LV 1	ρ	-0.09	NaN	NaN	0.09	-0.14	0.22	0.02	0.01	-0.02	NaN	0.05	-0.03
	P	.73	NaN	NaN	.73	.61	.52	.96	.99	.94	NaN	.86	.93
brain score LV 2	ρ	-0.18	NaN	NaN	0.18	0.31	-0.36	-0.25	0.32	-0.27	NaN	-0.14	-0.34
	P	.56	NaN	NaN	.56	.65	.17	.51	.51	.50	NaN	.61	.25
phenotypic score LV 2	ρ	0.12	NaN	NaN	-0.12	0.05	-0.15	0.01	0.09	-0.09	NaN	0.01	0.00
	P	.66	NaN	NaN	.66	.83	.59	.97	.72	.71	NaN	.97	.99
brain score LV 3	ρ	-0.10	NaN	NaN	0.10	0.25	-0.32	-0.21	0.39	-0.34	NaN	-0.14	**-0.40**
	P	.68	NaN	NaN	.68	.50	.38	.52	.05	.24	NaN	.62	**.04**
phenotypic score LV 3	ρ	0.01	NaN	NaN	-0.01	0.01	-0.09	0.01	0.10	-0.16	NaN	0.17	-0.10
	P	.98	NaN	NaN	.98	.96	.73	.96	.70	.58	NaN	.56	.70
brain score LV 4	ρ	-0.15	NaN	NaN	0.15	0.05	-0.05	0.04	-0.15	0.18	NaN	-0.07	0.11
	P	.60	NaN	NaN	.60	.85	.85	.88	.59	.56	NaN	.76	.68
phenotypic score LV 4	ρ	0.16	NaN	NaN	-0.16	0.00	-0.09	0.05	0.13	-0.16	NaN	0.11	-0.02
	P	.58	NaN	NaN	.58	.99	.73	.84	.63	.58	NaN	.68	.92
brain score LV 5	ρ	0.01	NaN	NaN	-0.01	0.01	-0.08	0.02	-0.11	0.14	NaN	-0.11	0.09
	P	.98	NaN	NaN	.98	.97	.72	.94	.66	.58	NaN	.65	.70
phenotypic score LV 5	ρ	-0.05	NaN	NaN	0.05	0.00	-0.02	-0.06	0.09	-0.10	NaN	0.01	-0.11
	P	.81	NaN	NaN	.81	.99	.93	.80	.69	.68	NaN	.95	.65

Table S22: Spearman's correlation analyses between socioeconomic factors and latent scores in the replication sample (Parental Education Background). Abbreviations: Secondary School = Parent finished Secondary School; Primary School = Parent only finished Primary School; Highest Occupation = Highest Occupation every achieved; M = Mother, F = Father. All *P* values FDR-corrected for multiple testing (family of tests: Table S21, Table S22, Table S23, Table S24).

		Secondary School (M/F)		Primary School (M/F)		No graduation (M/F)		University degree (M/F)		Highest Occupation (M/F)	
brain score LV 1	ρ	0.22	-0.04	-0.25	-0.02	0.01	0.03	0.25	0.18	-0.22	0.07
	P	.52	.88	.50	.96	.98	.90	.50	.55	.52	.79
phenotypic score LV 1	ρ	0.28	-0.08	-0.07	0.05	-0.11	0.08	0.04	-0.03	-0.02	0.06
	P	.69	.76	.79	.83	.67	.76	.87	.90	.93	.81
brain score LV 2	ρ	-0.21	0.02	0.26	0.03	-0.12	0.02	-0.19	-0.16	0.20	-0.11
	P	.53	.95	.50	.90	.65	.92	.55	.58	.54	.68
phenotypic score LV 2	ρ	-0.25	0.14	0.18	-0.11	0.10	-0.07	-0.14	-0.07	0.03	-0.14
	P	.50	.61	.55	.67	.70	.76	.61	.77	.92	.62
brain score LV 3	ρ	-0.07	0.20	0.23	0.00	-0.15	-0.16	-0.11	-0.10	0.05	-0.22
	P	.76	.53	.51	.99	.60	.58	.68	.70	.82	.52
phenotypic score LV 3	ρ	-0.13	0.14	-0.01	-0.08	0.13	-0.02	-0.03	-0.03	-0.05	-0.13
	P	.62	.60	.96	.75	.62	.92	.90	.92	.83	.62
brain score LV 4	ρ	-0.27	-0.08	0.15	0.02	-0.06	0.00	-0.10	-0.04	-0.03	0.09
	P	.50	.75	.60	.95	.81	.99	.69	.86	.91	.73
phenotypic score LV 4	ρ	-0.28	0.00	0.26	-0.05	-0.04	0.17	-0.21	-0.22	0.17	-0.07
	P	.50	.99	.51	.83	.87	.57	.53	.52	.58	.77
brain score LV 5	ρ	0.13	0.31	-0.10	-0.20	-0.07	-0.18	0.19	0.19	-0.32	-0.17
	P	.60	.64	.66	.50	.76	.52	.51	.51	.54	.53
phenotypic score LV 5	ρ	-0.16	-0.19	0.15	0.12	-0.18	0.10	-0.17	-0.19	0.06	0.12
	P	.56	.51	.57	.62	.52	.68	.53	.51	.80	.62

Table S23: Spearman's correlation analyses between socioeconomic factors and latent scores in the replication sample (Family & Relationships). Abbreviations: Nr. Sibling = Number of Siblings, Nr. Children = Number of own children, Nr. People in household = Number of people living in the same household, Partnership>1yr = has the participant ever been in a partnership over one year. All *P* values FDR-corrected for multiple testing (family of tests: Table S21, Table S22, Table S23, Table S24).

		Family				Relationship status last 12 months			
		Nr. Siblings	Nr. Children	Nr. people in household	Partnership>1yr	Single	Married	Partnership	Separated
brain score LV 1	ρ	-0.03	0.30	0.13	-0.16	0.20	0.01	0.15	0.12
	P	.93	.41	.64	.58	.53	.99	.60	.65
phenotypic score LV 1	ρ	0.02	**0.40**	**0.39**	-0.35	0.33	0.10	0.20	-0.21
	P	.96	**.02**	**.03**	.10	.17	.70	.54	.52
brain score LV 2	ρ	0.03	-0.38	-0.07	0.12	-0.30	0.08	-0.05	-0.07
	P	.90	.09	.78	.66	.84	.76	.83	.78
phenotypic score LV 2	ρ	0.00	-0.36	-0.23	0.30	-0.21	-0.14	-0.14	0.08
	P	.99	.14	.52	.77	.53	.61	.61	.74
brain score LV 3	ρ	0.06	-0.25	-0.08	0.06	-0.17	0.07	-0.14	0.00
	P	.80	.50	.75	.80	.57	.77	.61	.99
phenotypic score LV 3	ρ	-0.12	-0.32	-0.17	0.11	-0.19	0.00	0.03	0.01
	P	.65	.43	.56	.67	.54	.99	.91	.97
brain score LV 4	ρ	0.02	-0.34	-0.08	0.14	-0.30	0.06	-0.08	0.05
	P	.93	.25	.74	.61	.84	.81	.74	.85
phenotypic score LV 4	ρ	0.08	-0.27	-0.10	0.19	-0.11	-0.16	0.17	-0.02
	P	.76	.50	.70	.55	.67	.58	.56	.95
brain score LV 5	ρ	-0.03	-0.07	-0.17	0.06	-0.03	0.01	-0.17	0.18
	P	.88	.75	.54	.80	.88	.98	.53	.52
phenotypic score LV 5	ρ	-0.06	0.02	0.13	-0.06	0.03	0.02	0.09	-0.13
	P	.78	.93	.61	.80	.90	.94	.69	.61

Table S24: Spearman's correlation analyses between socioeconomic factors and latent scores in the replication sample (Education, Work). Abbreviations: Secondary School = finished Secondary School; Primary School = only finished Primary School; Highest Occupation = Highest Occupation every achieved. All *P* values FDR-corrected for multiple testing (family of tests: Table S21, Table S22, Table S23, Table S24).

|		Education					Work				
		Years of Education	Secondary School	Primary School	No graduation	University Degree	Highest Occupation	Full-time current	Part-time current	Full-time last 12 months	Part-time last 12 months
brain score LV 1	ρ	0.14	-0.05	-0.08	0.07	-0.12	0.20	-0.20	-0.19	0.19	-0.20
	P	.61	.84	.75	.78	.65	.53	.54	.55	.55	.53
phenotypic score LV 1	ρ	0.09	0.06	-0.06	0.08	0.15	0.20	-0.09	0.09	-0.09	-0.02
	P	.72	.83	.82	.76	.59	.54	.72	.73	.73	.96
brain score LV 2	ρ	-0.08	-0.06	0.14	-0.06	0.07	-0.11	0.08	0.09	-0.09	0.19
	P	.74	.81	.61	.80	.77	.67	.74	.73	.73	.54
phenotypic score LV 2	ρ	-0.19	0.10	-0.07	0.04	0.04	-0.31	0.10	-0.10	0.10	0.02
	P	.55	.70	.79	.88	.88	.69	.70	.69	.69	.95
brain score LV 3	ρ	-0.14	-0.05	0.13	-0.13	0.14	-0.09	0.14	0.13	-0.13	0.24
	P	.61	.84	.64	.64	.62	.72	.60	.63	.63	.51
phenotypic score LV 3	ρ	-0.14	0.08	-0.15	0.17	-0.08	-0.25	0.12	0.04	-0.04	0.14
	P	.62	.75	.60	.57	.75	.50	.65	.88	.88	.61
brain score LV 4	ρ	0.04	0.09	0.05	0.00	-0.12	-0.06	0.04	0.05	-0.05	0.06
	P	.87	.72	.84	.99	.65	.81	.86	.83	.83	.81
phenotypic score LV 4	ρ	-0.10	0.15	-0.06	0.02	0.06	-0.28	0.08	-0.04	0.04	0.03
	P	.71	.60	.80	.92	.80	.50	.74	.86	.86	.89
brain score LV 5	ρ	-0.02	0.10	-0.02	-0.10	-0.08	0.01	0.08	-0.01	0.01	0.06
	P	.95	.66	.92	.66	.74	.95	.72	.98	.98	.80
phenotypic score LV 5	ρ	-0.17	0.14	-0.19	-0.03	0.14	-0.20	0.14	0.06	-0.06	-0.02
	P	.54	.58	.51	.88	.58	.50	.58	.79	.79	.92

Table S25: Mean squared error (MSE) overlap of main SPLS model and RSH SPLS model. Depicted is the MSE between the phenotypic and brain patterns of the main SPLS model and the RSH SPLS model. As LV1 of the main SPLS model contains a global brain signature, leading to high overlaps with any other brain signature, the two minimum MSE values for each column are highlighted in bold to include similarities with other LV signatures as well.

Main SPLS Model	RSH SPLS Model	
Phenotypic Pattern		
	LV1	LV2
LV1	**4.85×10^{-3}**	**3.04×10^{-2}**
LV2	3.55×10^{-1}	5.69×10^{-2}
LV3	1.35×10^{-1}	**4.31×10^{-2}**
LV4	**1.33×10^{-1}**	6.06×10^{-2}
LV5	1.67×10^{-1}	4.85×10^{-2}
Brain Pattern		
	LV1	LV2
LV1	**2.64×10^{-5}**	**2.93×10^{-5}**
LV2	1.99×10^{-4}	1.14×10^{-4}
LV3	**1.07×10^{-4}**	**7.96×10^{-5}**
LV4	1.37×10^{-4}	8.83×10^{-5}
LV5	2.95×10^{-4}	1.18×10^{-4}

Table S26: Mean squared error (MSE) overlap of main SPLS model and LOSOCV SPLS model. Depicted is the MSE between the phenotypic and brain patterns of the main SPLS model and the LOSOCV SPLS model. As LV1 of the main SPLS model contains a global brain signature, leading to high overlaps with any other brain signature, the two minimum MSE values for each column are highlighted in bold to include similarities with other LV signatures as well.

Main SPLS Model	LOSOCV SPLS Model				
Phenotypic Pattern					
	LV1	LV2	LV3	LV4	LV5
LV1	**1.98×10^{-3}**	1.43×10^{-1}	**9.52×10^{-2}**	**5.20×10^{-2}**	**8.58×10^{-2}**
LV2	2.23×10^{-1}	**5.88×10^{-2}**	2.49×10^{-1}	1.94×10^{-1}	2.47×10^{-1}
LV3	1.38×10^{-1}	1.52×10^{-1}	**2.77×10^{-2}**	2.13×10^{-1}	**2.33×10^{-2}**
LV4	1.42×10^{-1}	**7.72×10^{-2}**	3.46×10^{-1}	**6.91×10^{-2}**	2.71×10^{-1}
LV5	**1.24×10^{-1}**	1.43×10^{-1}	2.00×10^{-1}	1.43×10^{-1}	1.82×10^{-1}
Brain Pattern					
	LV1	LV2	LV3	LV4	LV5
LV1	**5.28×10^{-6}**	**5.47×10^{-5}**	**5.79×10^{-5}**	**4.61×10^{-5}**	**5.39×10^{-5}**
LV2	**5.32×10^{-5}**	9.34×10^{-5}	3.60×10^{-4}	3.84×10^{-4}	3.67×10^{-4}
LV3	6.37×10^{-5}	7.76×10^{-5}	**6.42×10^{-5}**	**1.07×10^{-4}**	**1.10×10^{-4}**
LV4	6.33×10^{-5}	**4.86×10^{-5}**	1.91×10^{-4}	1.66×10^{-4}	1.44×10^{-4}
LV5	8.32×10^{-5}	1.37×10^{-4}	4.30×10^{-4}	5.07×10^{-4}	4.23×10^{-4}

Table S27: Mean squared error (MSE) overlap of main SPLS model and SPLS model with IQR addition. Depicted is the MSE between the phenotypic and brain patterns of the main SPLS model and the SPLS model with added IQR scores. As LV1 of the main SPLS model contains a global brain signature, leading to high overlaps with any other brain signature, the two minimum MSE values for each column are highlighted in bold to include similarities with other LV signatures as well.

Main SPLS Model	SPLS Model with IQR addition								
Phenotypic Pattern									
	LV1	LV2	LV3	LV4	LV5	LV6	LV7	LV8	LV9
LV1	$\mathbf{2.91\times10^{-3}}$	1.46×10^{-1}	$\mathbf{7.86\times10^{-2}}$	$\mathbf{9.61\times10^{-3}}$	$\mathbf{3.44\times10^{-2}}$	$\mathbf{3.86\times10^{-2}}$	$\mathbf{1.77\times10^{-2}}$	6.85×10^{-2}	$\mathbf{3.68\times10^{-2}}$
LV2	4.36×10^{-1}	$\mathbf{6.40\times10^{-2}}$	1.19×10^{-1}	5.66×10^{-1}	9.87×10^{-2}	4.15×10^{-1}	5.29×10^{-1}	5.29×10^{-2}	2.02×10^{-1}
LV3	$\mathbf{1.51\times10^{-1}}$	1.71×10^{-1}	$\mathbf{4.04\times10^{-2}}$	1.89×10^{-1}	$\mathbf{2.82\times10^{-2}}$	$\mathbf{5.07\times10^{-2}}$	$\mathbf{1.33\times10^{-1}}$	$\mathbf{3.99\times10^{-2}}$	1.26×10^{-1}
LV4	1.75×10^{-1}	$\mathbf{6.03\times10^{-2}}$	2.42×10^{-1}	$\mathbf{1.69\times10^{-1}}$	9.57×10^{-2}	2.97×10^{-1}	1.66×10^{-1}	6.82×10^{-2}	$\mathbf{6.72\times10^{-2}}$
LV5	1.67×10^{-1}	1.33×10^{-1}	1.54×10^{-1}	2.22×10^{-1}	5.48×10^{-2}	2.26×10^{-1}	1.58×10^{-1}	$\mathbf{3.13\times10^{-2}}$	1.77×10^{-1}
Brain Pattern									
	LV1	LV2	LV3	LV4	LV5	LV6	LV7	LV8	LV9
LV1	$\mathbf{1.25\times10^{-5}}$	$\mathbf{6.08\times10^{-5}}$	$\mathbf{5.76\times10^{-5}}$	$\mathbf{4.88\times10^{-5}}$	$\mathbf{3.60\times10^{-5}}$	$\mathbf{5.43\times10^{-5}}$	$\mathbf{3.65\times10^{-5}}$	$\mathbf{4.71\times10^{-5}}$	$\mathbf{4.31\times10^{-5}}$
LV2	8.29×10^{-5}	1.73×10^{-4}	1.73×10^{-4}	8.90×10^{-4}	1.19×10^{-4}	7.64×10^{-5}	5.58×10^{-4}	6.84×10^{-5}	6.41×10^{-4}
LV3	$\mathbf{8.22\times10^{-5}}$	1.01×10^{-4}	$\mathbf{3.84\times10^{-5}}$	1.16×10^{-4}	7.12×10^{-5}	6.88×10^{-5}	8.93×10^{-5}	6.47×10^{-5}	$\mathbf{1.02\times10^{-4}}$
LV4	8.98×10^{-5}	$\mathbf{6.46\times10^{-5}}$	1.31×10^{-4}	2.13×10^{-4}	8.51×10^{-5}	9.15×10^{-5}	1.31×10^{-4}	$\mathbf{6.17\times10^{-5}}$	1.42×10^{-4}
LV5	1.42×10^{-4}	2.39×10^{-4}	2.24×10^{-4}	9.35×10^{-4}	1.30×10^{-4}	1.30×10^{-4}	7.01×10^{-4}	1.00×10^{-4}	7.68×10^{-4}

Supplementary Figures

Figure S1: Observational study design of PRONIA. Colored boxed indicate type of assessment conducted in each of the study groups: Healthy controls (HC, green), patients with recent-onset depression (ROD, yellow), persons with a clinical high-risk for psychosis (CHR, orange), patients with recent-onset psychosis (ROP, red). Previously published in Koutsouleris et al. (1) and reprinted with permission.
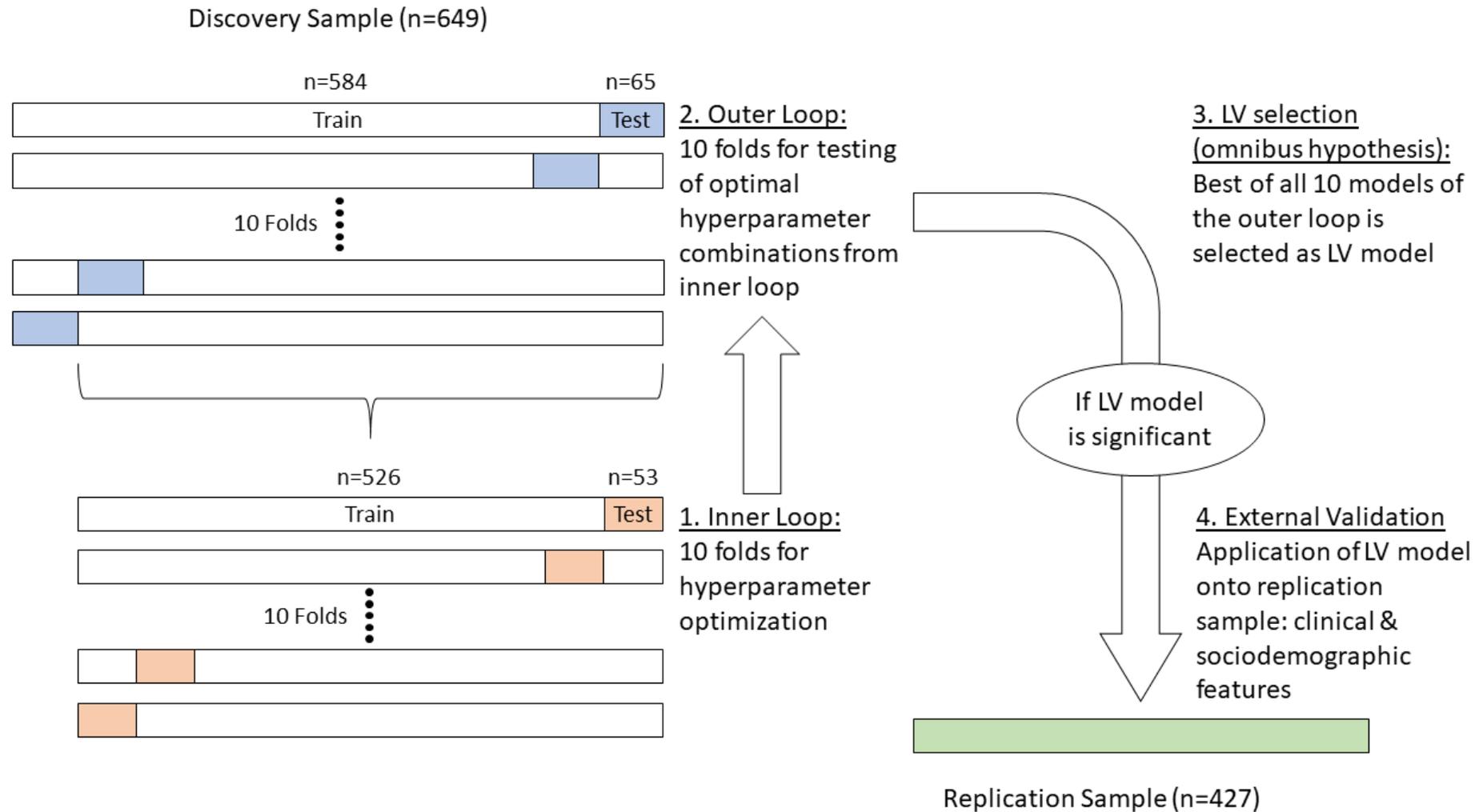
Figure S2: Nested cross-validation and external validation framework. Depicted is the nested cross-validation (NCV) framework with 10x10 folds on the outer and inner loop as well as the external validation of the final LV model. Hyperparameter optimization of cu and cv is performed in the inner loop, whereas testing of the optimized model is done on the outer loop. The best model of all 10 outer folds iterations is chosen as the LV model (in accordance with the omnibus hypothesis) and then applied on the replication sample for external validation.

| CTQ	03	08	14	18	25	09	11	12	15	17	20	21	23	24	27	05	07	13	19	28	01	02	04	06	26	10	16	22	age	male	female	HC	ROP	ROD	CHR
03	1.00	0.48	0.65	0.49	0.44	0.15	0.35	0.26	0.35	0.25	0.20	0.13	0.19	0.17	0.18	0.33	0.46	0.35	0.34	0.45	0.12	0.36	0.19	0.22	0.23	-0.17	-0.22	-0.21	-0.05	0.02	-0.02	-0.33	0.15	0.07	0.19
08	0.48	1.00	0.46	0.56	0.46	0.12	0.37	0.30	0.44	0.24	0.24	0.17	0.24	0.30	0.26	0.33	0.47	0.42	0.38	0.47	0.06	0.45	0.21	0.29	0.35	-0.11	-0.19	-0.16	-0.07	-0.05	0.05	-0.27	0.13	0.05	0.16
14	0.65	0.46	1.00	0.58	0.52	0.13	0.41	0.34	0.37	0.27	0.23	0.11	0.23	0.25	0.23	0.31	0.48	0.41	0.39	0.51	0.11	0.38	0.23	0.24	0.25	-0.23	-0.27	-0.27	-0.02	-0.06	0.06	-0.32	0.11	0.09	0.20
18	0.49	0.56	0.58	1.00	0.53	0.20	0.41	0.33	0.51	0.31	0.31	0.22	0.27	0.31	0.33	0.31	0.42	0.46	0.42	0.49	0.14	0.38	0.23	0.29	0.32	-0.17	-0.23	-0.24	-0.04	-0.04	0.04	-0.26	0.10	0.06	0.15
25	0.44	0.46	0.52	0.53	1.00	0.12	0.35	0.27	0.49	0.31	0.35	0.19	0.34	0.36	0.40	0.34	0.47	0.42	0.38	0.54	0.12	0.40	0.28	0.33	0.32	-0.14	-0.24	-0.25	0.09	0.00	0.00	-0.34	0.20	0.10	0.12
09	0.15	0.12	0.13	0.20	0.12	1.00	0.42	0.31	0.34	0.51	0.14	0.05	0.15	0.13	0.13	0.05	0.10	0.11	0.12	0.10	0.08	0.13	0.16	0.09	0.07	0.03	-0.05	-0.05	0.05	0.07	-0.07	-0.07	0.03	0.02	0.04
11	0.35	0.37	0.41	0.41	0.35	0.42	1.00	0.65	0.60	0.57	0.23	0.05	0.21	0.22	0.25	0.22	0.35	0.28	0.28	0.36	0.10	0.29	0.24	0.19	0.23	-0.08	-0.16	-0.18	0.06	0.00	0.00	-0.18	0.06	0.08	0.08
12	0.26	0.30	0.34	0.33	0.27	0.31	0.65	1.00	0.58	0.53	0.28	0.19	0.29	0.26	0.29	0.18	0.30	0.32	0.29	0.31	0.15	0.26	0.27	0.24	0.18	-0.08	-0.14	-0.13	0.09	-0.02	0.02	-0.20	0.09	0.07	0.08
15	0.35	0.44	0.37	0.51	0.49	0.34	0.60	0.58	1.00	0.53	0.45	0.32	0.42	0.44	0.48	0.21	0.35	0.33	0.29	0.36	0.17	0.32	0.29	0.43	0.23	-0.05	-0.14	-0.15	0.05	0.01	-0.01	-0.21	0.14	0.06	0.05
17	0.25	0.24	0.27	0.31	0.31	0.51	0.57	0.53	0.53	1.00	0.25	0.10	0.22	0.13	0.17	0.15	0.26	0.26	0.25	0.27	0.14	0.25	0.32	0.23	0.23	-0.02	-0.09	-0.07	0.03	0.02	-0.02	-0.17	0.04	0.09	0.08
20	0.20	0.24	0.23	0.31	0.35	0.14	0.23	0.28	0.45	0.25	1.00	0.61	0.72	0.75	0.75	0.18	0.20	0.16	0.17	0.23	0.14	0.24	0.24	0.30	0.19	-0.06	-0.10	-0.12	0.04	-0.10	0.10	-0.17	0.12	0.01	0.08
21	0.13	0.17	0.11	0.22	0.19	0.05	0.05	0.19	0.32	0.10	0.61	1.00	0.61	0.60	0.53	0.05	0.09	0.12	0.12	0.09	0.08	0.15	0.15	0.26	0.13	-0.05	-0.03	-0.05	0.02	-0.01	0.01	-0.12	0.13	0.01	0.00
23	0.19	0.24	0.23	0.27	0.34	0.15	0.21	0.29	0.42	0.22	0.72	0.61	1.00	0.73	0.67	0.12	0.21	0.15	0.15	0.20	0.11	0.25	0.15	0.25	0.18	-0.05	-0.05	-0.07	0.05	-0.03	0.03	-0.16	0.10	0.05	0.05
24	0.17	0.30	0.25	0.31	0.36	0.13	0.22	0.26	0.44	0.13	0.75	0.60	0.73	1.00	0.71	0.15	0.20	0.16	0.15	0.24	0.10	0.29	0.19	0.23	0.16	-0.05	-0.08	-0.08	0.04	-0.10	0.10	-0.18	0.09	0.08	0.06
27	0.18	0.26	0.23	0.33	0.40	0.13	0.25	0.29	0.48	0.17	0.75	0.53	0.67	0.71	1.00	0.17	0.24	0.21	0.20	0.26	0.17	0.28	0.22	0.34	0.22	-0.04	-0.08	-0.11	0.01	-0.04	0.04	-0.18	0.12	0.03	0.06
05	0.33	0.33	0.31	0.31	0.34	0.05	0.22	0.18	0.21	0.15	0.18	0.05	0.12	0.15	0.17	1.00	0.65	0.47	0.50	0.57	0.09	0.46	0.22	0.10	0.31	-0.17	-0.26	-0.29	0.00	0.06	-0.06	-0.33	0.17	0.06	0.17
07	0.46	0.47	0.48	0.42	0.47	0.10	0.35	0.30	0.35	0.26	0.20	0.09	0.21	0.20	0.24	0.65	1.00	0.53	0.57	0.65	0.14	0.62	0.23	0.25	0.35	-0.18	-0.29	-0.35	0.02	0.03	-0.03	-0.35	0.12	0.11	0.21
13	0.35	0.42	0.41	0.46	0.42	0.11	0.28	0.32	0.33	0.26	0.16	0.12	0.15	0.16	0.21	0.47	0.53	1.00	0.67	0.65	0.17	0.49	0.30	0.29	0.39	-0.22	-0.30	-0.38	0.00	0.01	-0.01	-0.30	0.16	0.08	0.14
19	0.34	0.38	0.39	0.42	0.38	0.12	0.28	0.29	0.29	0.25	0.17	0.12	0.15	0.15	0.20	0.50	0.57	0.67	1.00	0.65	0.16	0.47	0.30	0.23	0.35	-0.30	-0.38	-0.46	0.03	-0.02	0.02	-0.31	0.11	0.15	0.13
28	0.45	0.47	0.51	0.49	0.54	0.10	0.36	0.31	0.36	0.27	0.23	0.09	0.20	0.24	0.26	0.57	0.65	0.65	0.65	1.00	0.15	0.58	0.29	0.27	0.43	-0.28	-0.36	-0.42	0.06	0.01	-0.01	-0.33	0.09	0.16	0.16
01	0.12	0.06	0.11	0.14	0.12	0.08	0.10	0.15	0.17	0.14	0.14	0.08	0.11	0.10	0.17	0.09	0.14	0.17	0.16	0.15	1.00	0.18	0.17	0.28	0.16	-0.07	-0.13	-0.07	0.02	0.04	-0.04	-0.13	0.13	0.05	-0.02
02	0.36	0.45	0.38	0.38	0.40	0.13	0.29	0.26	0.32	0.25	0.24	0.15	0.25	0.29	0.28	0.46	0.62	0.49	0.47	0.58	0.18	1.00	0.22	0.26	0.41	-0.16	-0.21	-0.26	-0.02	0.05	-0.05	-0.30	0.18	0.05	0.15
04	0.19	0.21	0.23	0.23	0.28	0.16	0.24	0.27	0.29	0.32	0.24	0.15	0.15	0.19	0.22	0.22	0.23	0.30	0.30	0.29	0.17	0.22	1.00	0.30	0.17	-0.08	-0.13	-0.12	0.12	-0.03	0.03	-0.16	0.12	0.02	0.05
06	0.22	0.29	0.24	0.29	0.33	0.09	0.19	0.24	0.43	0.23	0.30	0.26	0.25	0.23	0.34	0.10	0.25	0.29	0.23	0.27	0.28	0.26	0.30	1.00	0.23	-0.05	-0.12	-0.11	0.03	0.02	-0.02	-0.17	0.15	0.03	0.02
26	0.23	0.35	0.25	0.32	0.32	0.07	0.23	0.18	0.23	0.23	0.19	0.13	0.18	0.16	0.22	0.31	0.35	0.39	0.35	0.43	0.16	0.41	0.17	0.23	1.00	-0.11	-0.12	-0.18	-0.03	0.03	-0.03	-0.23	0.12	0.07	0.09
10	-0.17	-0.11	-0.23	-0.17	-0.14	0.03	-0.08	-0.08	-0.05	-0.02	-0.06	-0.05	-0.05	-0.05	-0.04	-0.17	-0.18	-0.22	-0.30	-0.28	-0.07	-0.16	-0.08	-0.05	-0.11	1.00	0.36	0.41	-0.04	0.04	-0.04	0.12	-0.03	-0.05	-0.07
16	-0.22	-0.19	-0.27	-0.23	-0.24	-0.05	-0.16	-0.14	-0.14	-0.09	-0.10	-0.03	-0.05	-0.08	-0.08	-0.26	-0.29	-0.30	-0.38	-0.36	-0.13	-0.21	-0.13	-0.12	-0.12	0.36	1.00	0.59	0.00	-0.06	0.06	0.13	-0.05	-0.05	-0.06
22	-0.21	-0.16	-0.27	-0.24	-0.25	-0.05	-0.18	-0.13	-0.15	-0.07	-0.12	-0.05	-0.07	-0.08	-0.11	-0.29	-0.35	-0.38	-0.46	-0.42	-0.07	-0.26	-0.12	-0.11	-0.18	0.41	0.59	1.00	-0.06	-0.03	0.03	0.16	-0.03	-0.06	-0.11
age	-0.05	-0.07	-0.02	-0.04	0.09	0.05	0.06	0.09	0.05	0.03	0.04	0.02	0.05	0.04	0.01	0.00	0.02	0.00	0.03	0.06	0.02	-0.02	0.12	0.03	-0.03	-0.04	0.00	-0.06	1.00	0.03	-0.03	0.02	0.03	0.06	-0.11
male	0.02	-0.05	-0.06	-0.04	0.00	0.07	0.00	-0.02	0.01	0.02	-0.10	-0.01	-0.03	-0.10	-0.04	0.06	0.03	0.01	-0.02	0.01	0.04	0.05	-0.03	0.02	0.03	0.04	-0.06	-0.03	0.03	1.00	-1.00	-0.16	0.15	-0.01	0.05
female	-0.02	0.05	0.06	0.04	0.00	-0.07	0.00	0.02	-0.01	-0.02	0.10	0.01	0.03	0.10	0.04	-0.06	-0.03	-0.01	0.02	-0.01	-0.04	-0.05	0.03	-0.02	-0.03	-0.04	0.06	0.03	-0.03	-1.00	1.00	0.16	-0.15	0.01	-0.05
HC	-0.33	-0.27	-0.32	-0.26	-0.34	-0.07	-0.18	-0.20	-0.21	-0.17	-0.17	-0.12	-0.16	-0.18	-0.18	-0.33	-0.35	-0.30	-0.31	-0.33	-0.13	-0.30	-0.16	-0.17	-0.23	0.12	0.13	0.16	0.02	-0.16	0.16	1.00	-0.42	-0.41	-0.40
ROP	0.15	0.13	0.11	0.10	0.20	0.03	0.06	0.09	0.14	0.04	0.12	0.13	0.10	0.09	0.12	0.17	0.12	0.16	0.11	0.09	0.13	0.18	0.12	0.15	0.12	-0.03	-0.05	-0.03	0.03	0.15	-0.15	-0.42	1.00	-0.25	-0.25
ROD	0.07	0.05	0.09	0.06	0.10	0.02	0.08	0.07	0.06	0.09	0.01	0.01	0.05	0.08	0.03	0.06	0.11	0.08	0.15	0.16	0.05	0.05	0.02	0.03	0.07	-0.05	-0.05	-0.06	0.06	-0.01	0.01	-0.41	-0.25	1.00	-0.24
CHR	0.19	0.16	0.20	0.15	0.12	0.04	0.08	0.08	0.05	0.08	0.08	0.00	0.05	0.06	0.06	0.17	0.21	0.14	0.13	0.16	-0.02	0.15	0.05	0.02	0.09	-0.07	-0.06	-0.11	-0.11	0.05	-0.05	-0.40	-0.25	-0.24	1.00

Figure S3: Autocorrelation Spearman's coefficients of phenotypic input feature space. The CTQ items are sorted according to the subscales of emotional, physical, and sexual abuse as well as emotional and physical neglect and denial. After the CTQ items, the other phenotypic input features (age, male, female sex, HC, ROP, ROD, CHR) are stated. All phenotypic input features are stated in the same order as they were entered into the SPLS analysis. Warm colors indicate positive and cold colors indicate negative values.
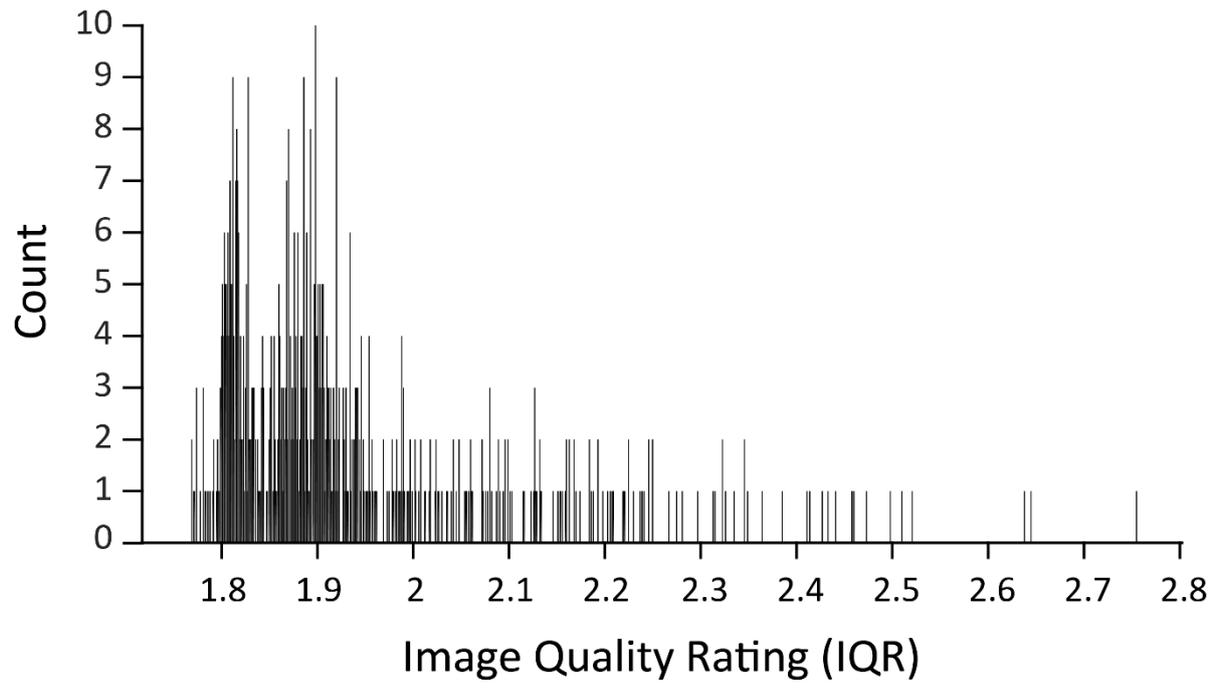
Figure S4: Histogram of IQR scores of the main study sample. Depicted are the IQR scores of the main study sample.

A)

B)
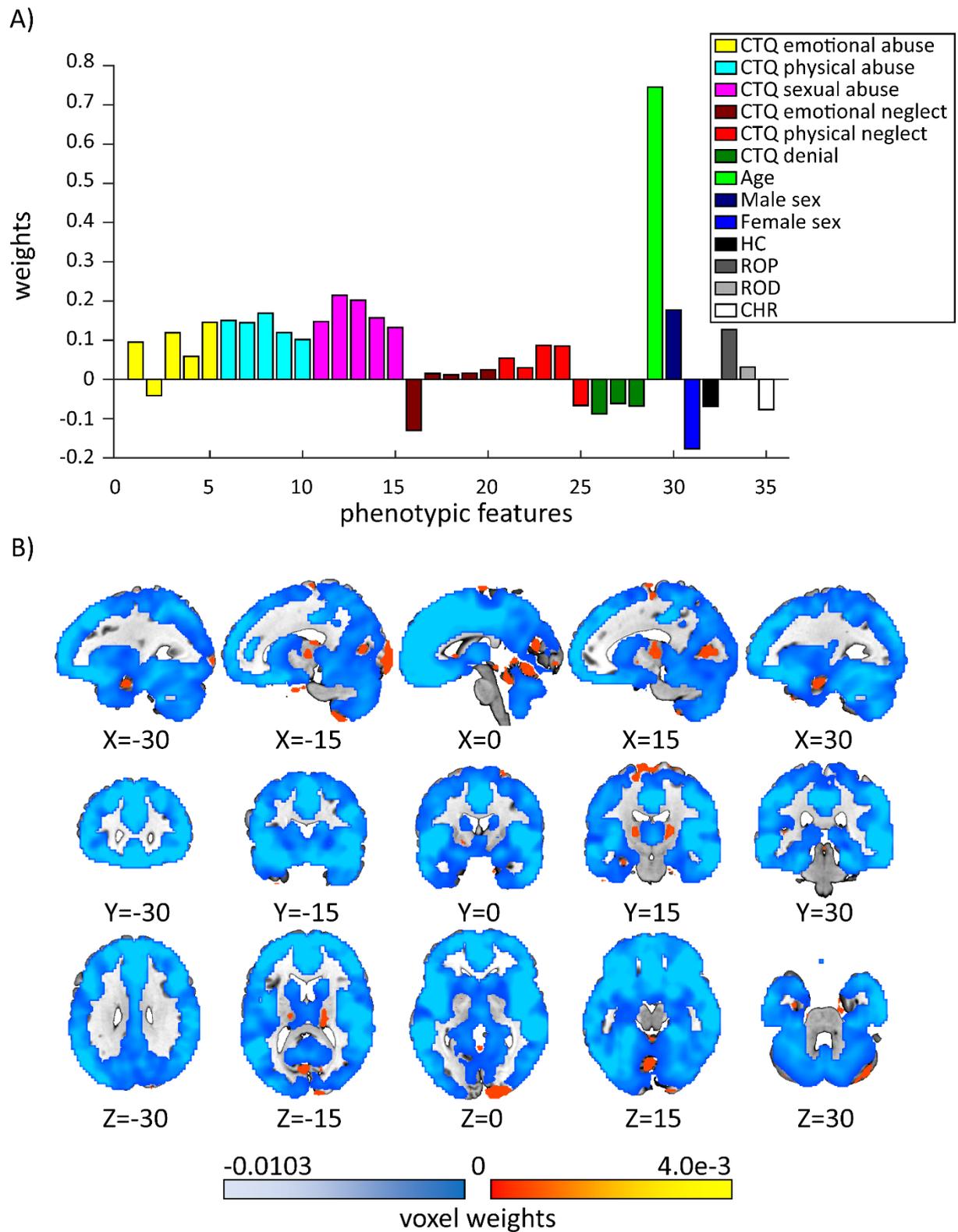
Figure S5: Exemplary illustration of non-sparse PLS solution with c_u and c_v at maximum values. This figure illustrates the weight vector solution u and v when c_u and c_v are tested at the maximum values of our predefined hyperparameter grid. It resembles a non-sparse, regular PLS analysis, which is therefore part of our SPLS hyperparameter grid search.

A)

B)
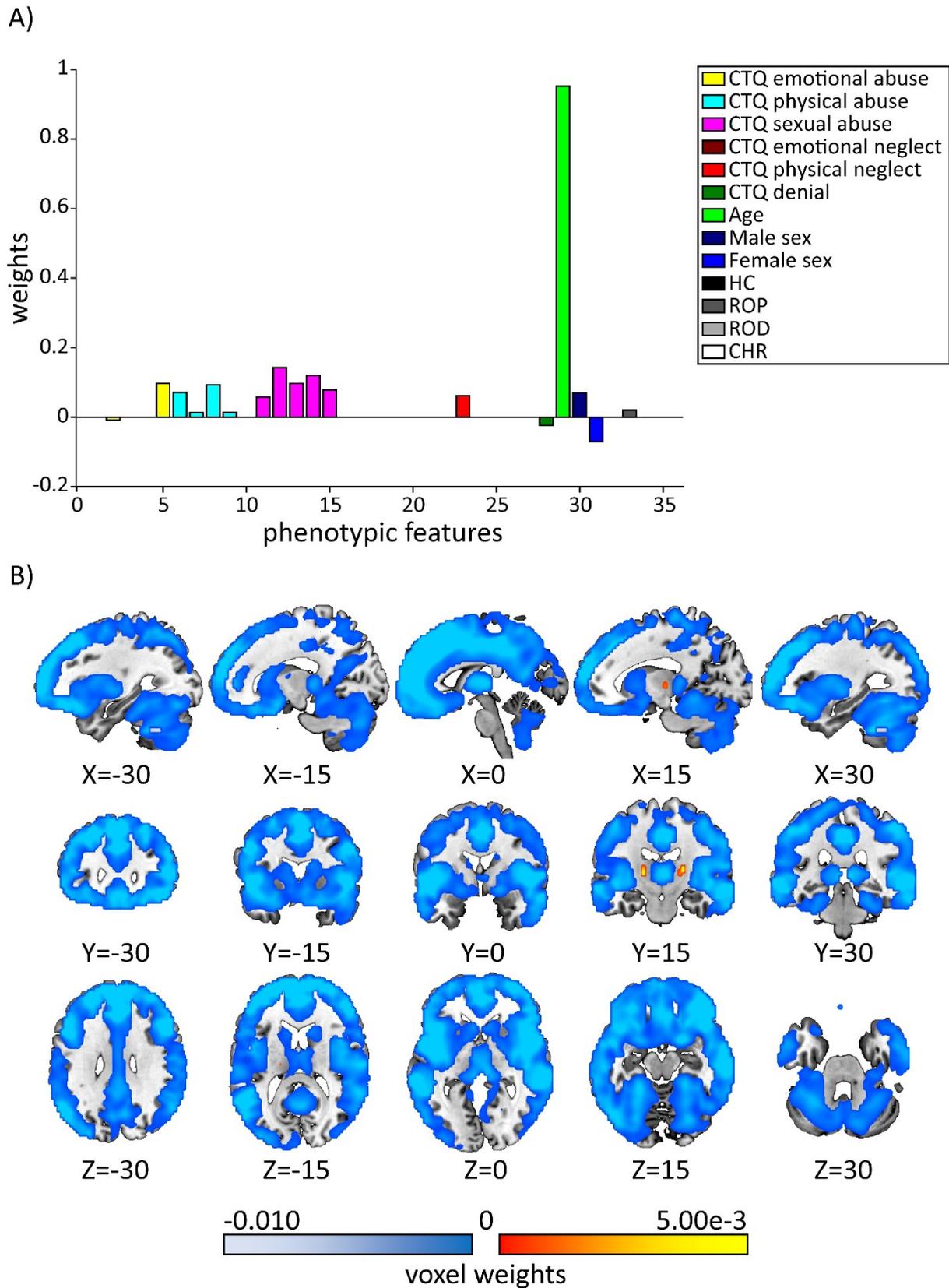
Figure S6: Predominantly age-informed signature of LV1. A) The barplot visualizes the direction and the values of the weights included in the phenotypic pattern of LV1. Positive weights were assigned to questions from the CTQ subscales of emotional abuse (CTQ25), physical abuse (CTQ09, CTQ11, CTQ12, CTQ15), sexual abuse (CTQ20, CTQ21, CTQ23, CTQ24, CTQ27) and physical neglect (CTQ04) as well as to age, male sex and ROP diagnosis. Negative weights were assigned to questions from the CTQ subscales of emotional abuse (CTQ08) and denial (CTQ22) as well as female sex. B) Depicted is the brain pattern of LV1, with positive weighting of voxels displayed in red and negative weighting displayed in blue color scale.

A)

B)
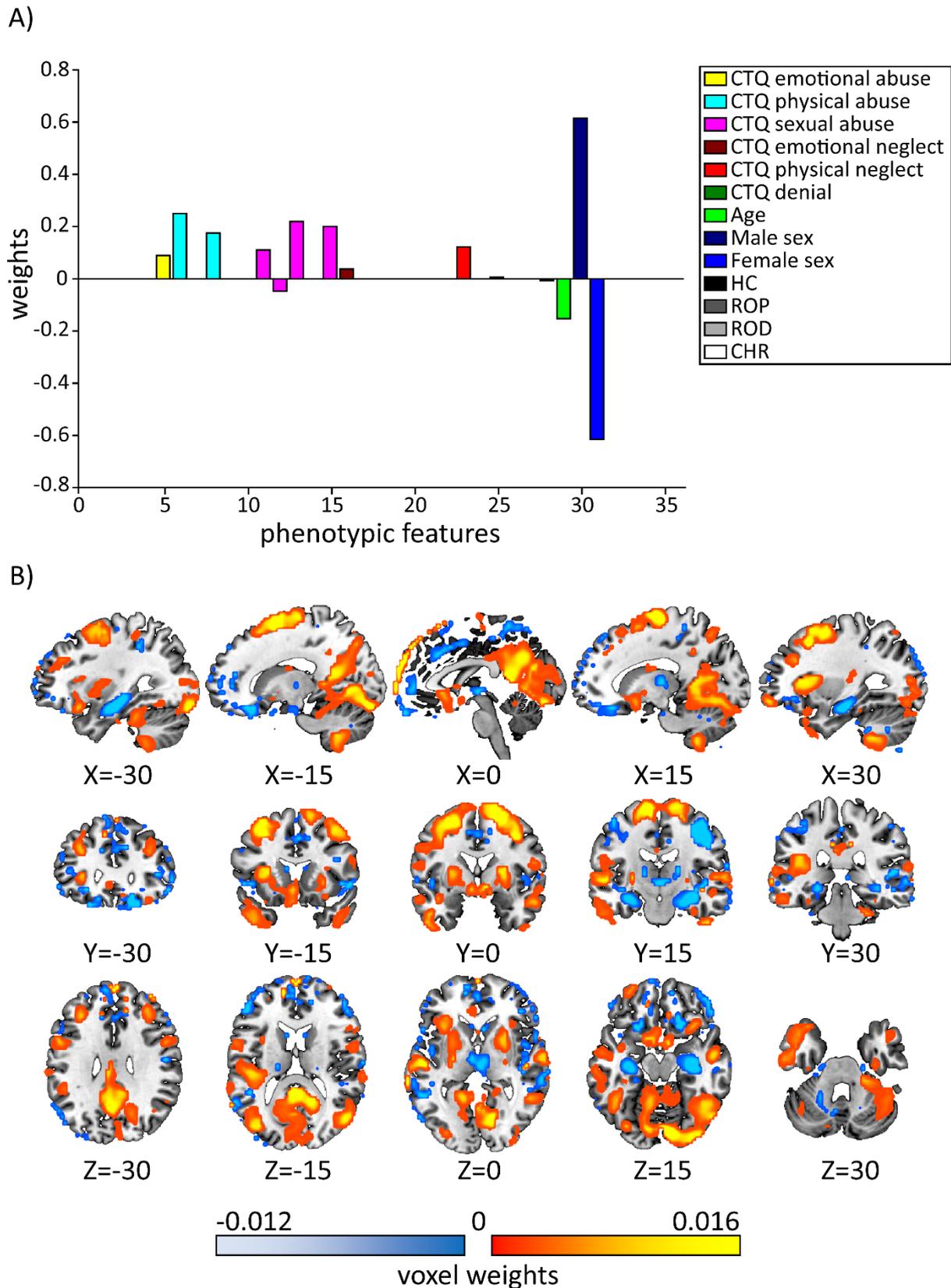
Figure S7: Predominantly sex-informed signature of LV3. A) The barplot visualizes the direction and the values of the weights included in the phenotypic pattern of LV3. One question from emotional abuse (CTQ25), two from physical abuse (CTQ09, CTQ12), four from sexual abuse (CTQ20, CTQ21, CTQ23, CTQ27), one from emotional neglect (CTQ05), two from physical neglect (CTQ04, CTQ26) and one from denial (CTQ22). Furthermore, age as well as female and male sex were included. B) Depicted is the brain pattern of LV3, with positive weighting of voxels displayed in red and negative weighting displayed in blue color scale.

A)

B)

X=-30   X=-15   X=0   X=15   X=30

Y=-30   Y=-15   Y=0   Y=15   Y=30

Z=-30   Z=-15   Z=0   Z=15   Z=30

-0.0475   0   2.1e-3

voxel weights
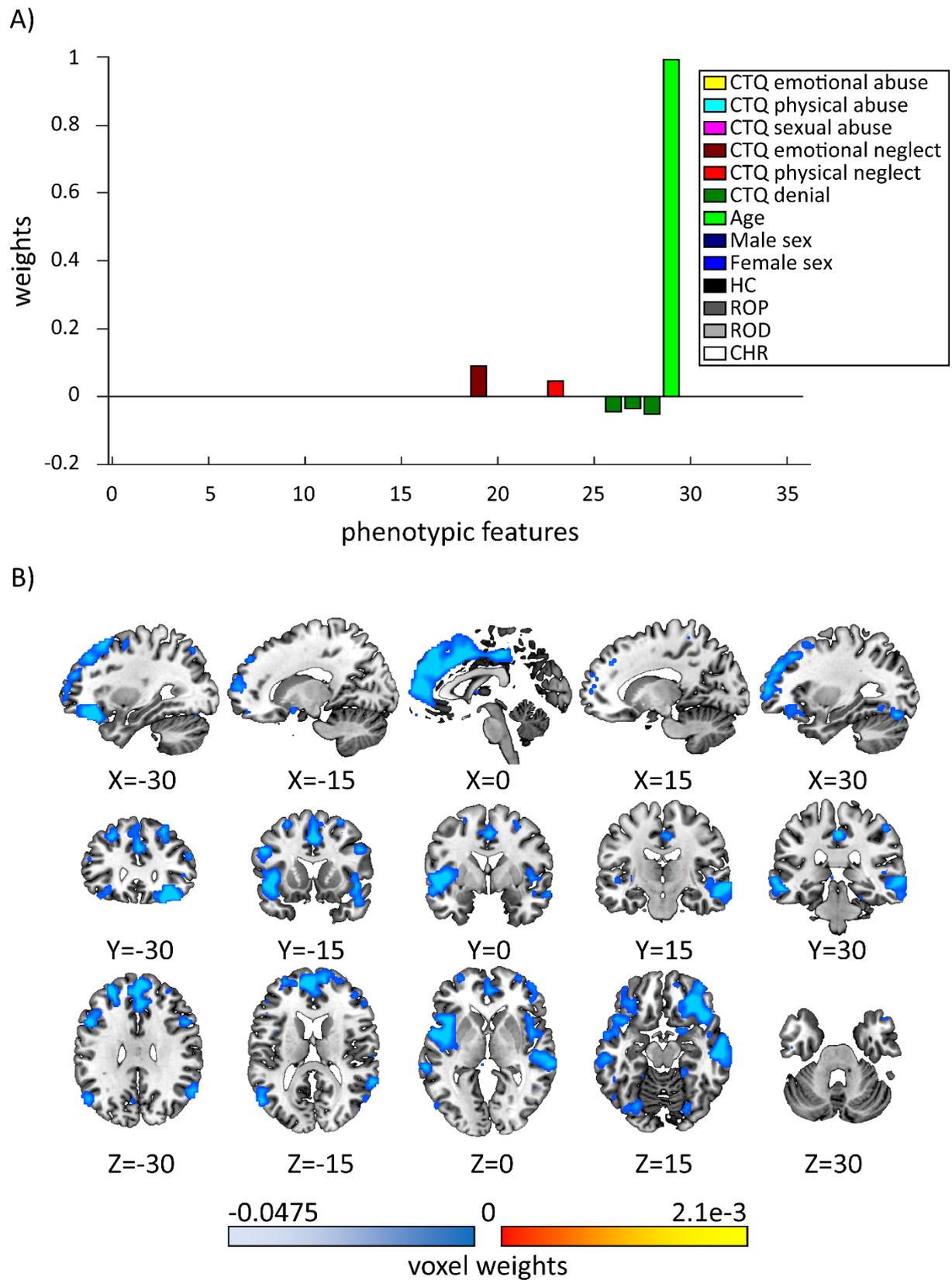
Figure S8: Random split-half analysis: Predominantly age-informed signature of LV1: A) The barplot visualizes the direction and the values of the weights included in the phenotypic pattern of LV1. Age received the strongest positive weighting, while further positive weights were assigned to questions from the CTQ subscales of emotional (CTQ19) and physical neglect (CTQ04) whereas negative weights were assigned to items from the denial subscale (CTQ10, 16, 22). B) Depicted is the brain pattern of LV1, with positive weighting of voxels displayed in red and negative weighting displayed in blue color scale.

A)

B)
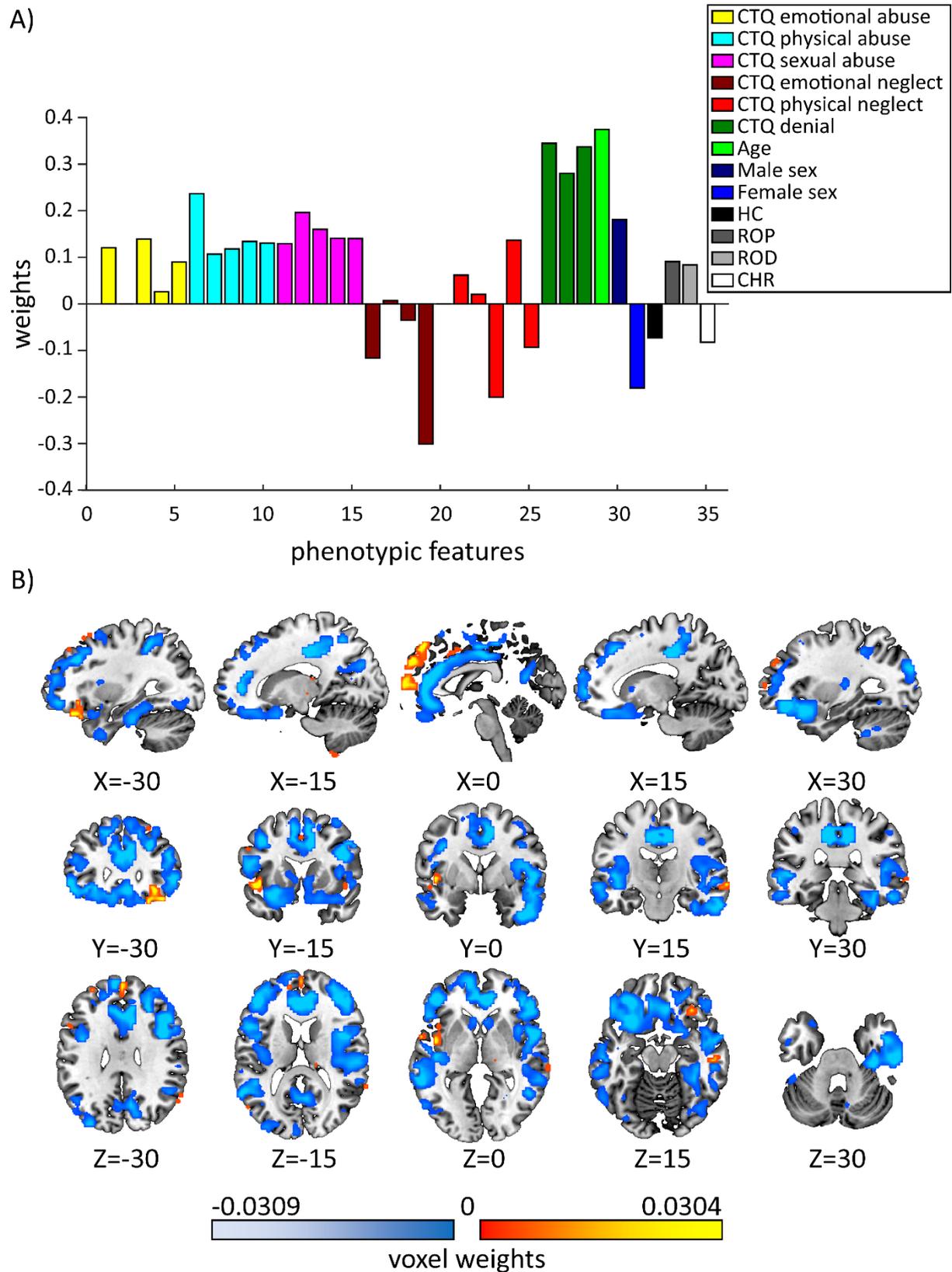
Figure S9: Random split-half analysis: Highly dense trauma signature of LV2: A) The barplot visualizes the direction and the values of the weights included in the phenotypic pattern of LV2. The dense phenotypic pattern featured weighting of most of the phenotypic input feature space except for one item each from the emotional abuse (CTQ08) and the emotional neglect subscale (CTQ28). B) Depicted is the brain pattern of LV2, with positive weighting of voxels displayed in red and negative weighting displayed in blue color scale.

A)

B)
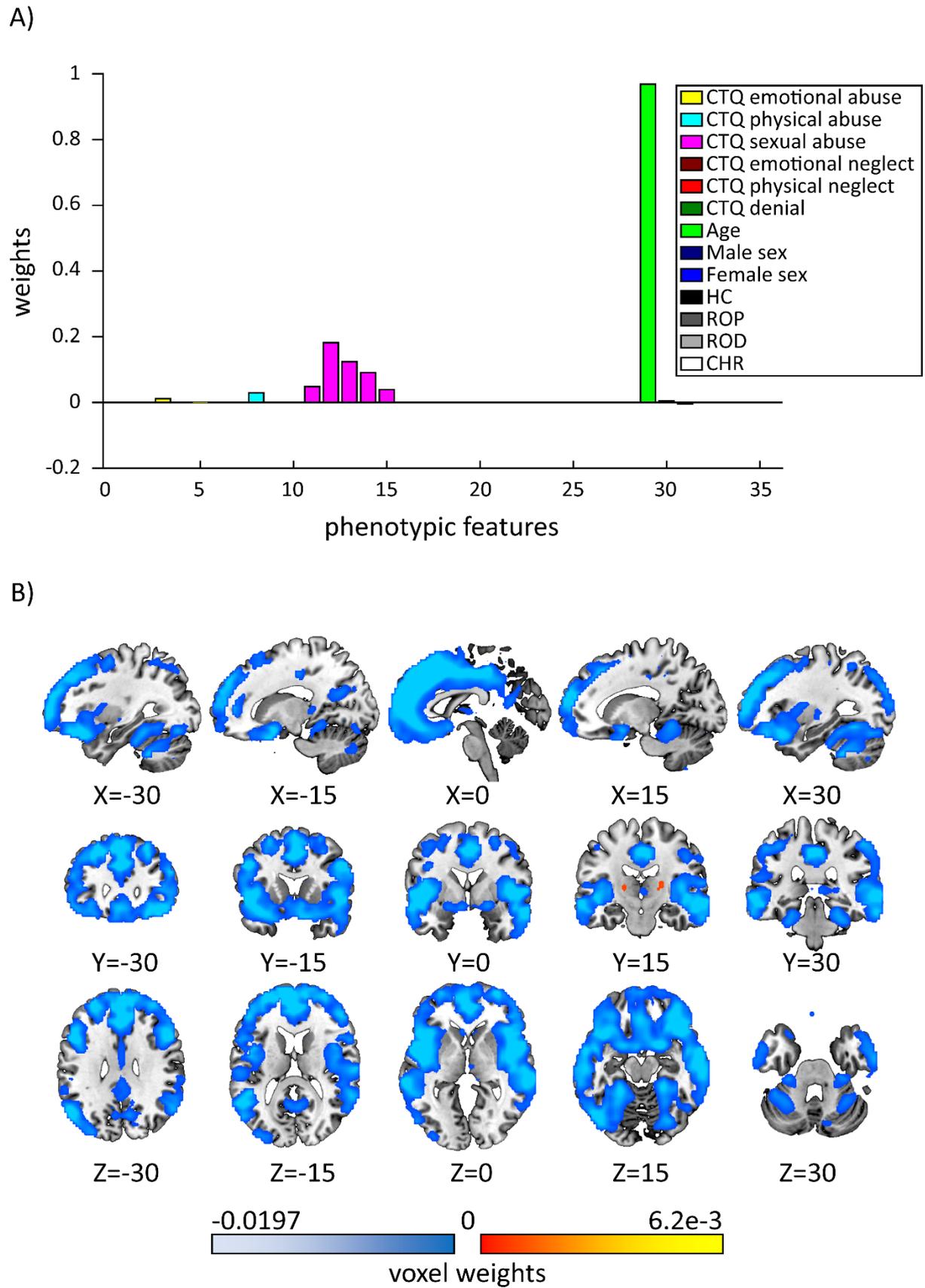
Figure S10: Predominantly age-informed LV1 of LOSOCV analysis. A) The barplot visualizes the direction and the values of the weights included in the phenotypic pattern of LV1. Positive weights were assigned most strongly to age as well as to questions from the CTQ subscales of emotional abuse (CTQ14), physical abuse (CTQ12), sexual abuse (CTQ20, CTQ21, CTQ23, CTQ24, CTQ27) and inversely to male and female sex. B) Depicted is the brain pattern of LV1, with positive weighting of voxels displayed in red and negative weighting displayed in blue color scale.
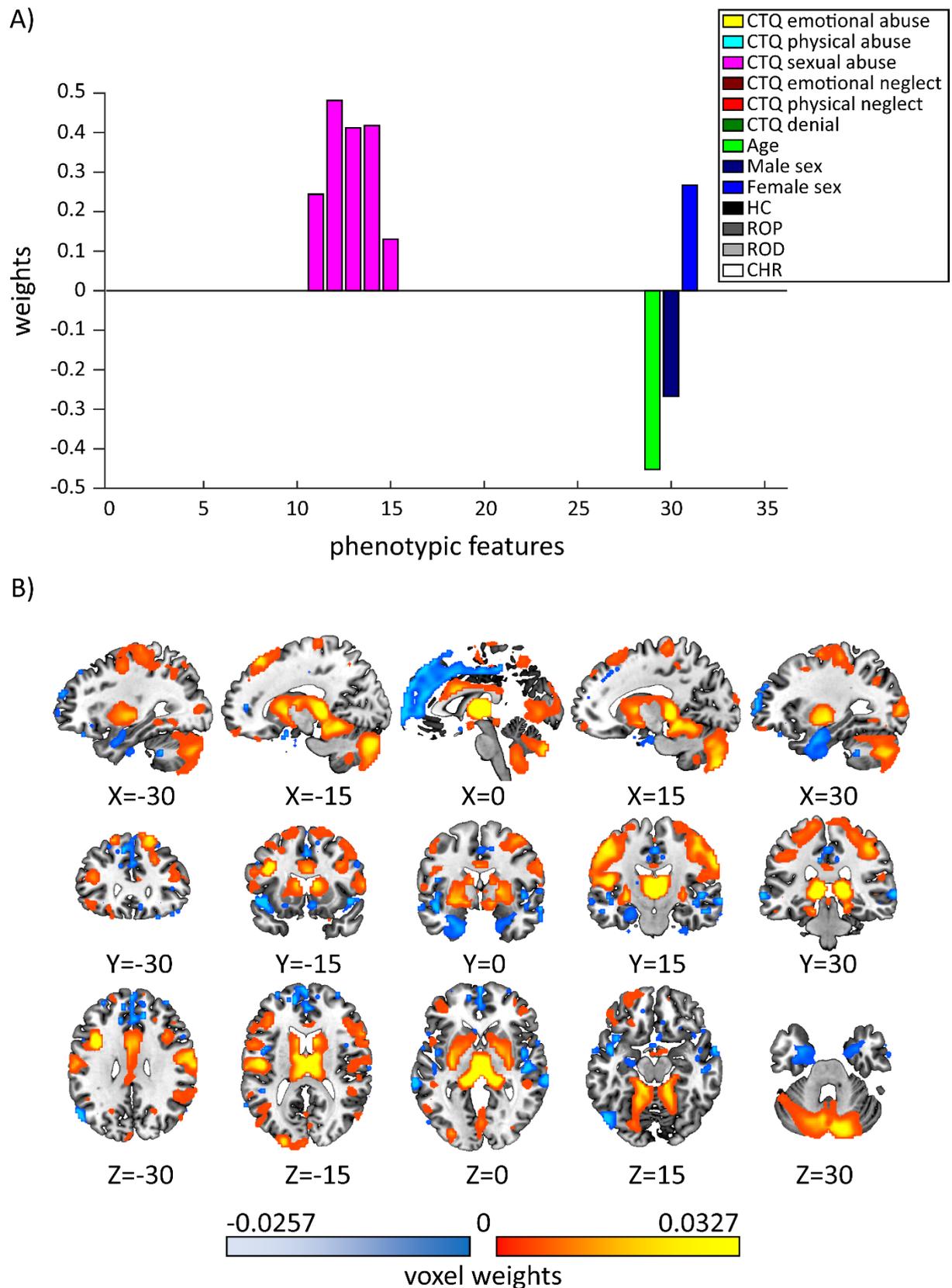
Figure S11: age- and sex-dependent sexual abuse signature of LV2 of the LOSOCV analysis. A) The barplot visualizes the direction and the values of the weights included in the phenotypic pattern of LV2. All five questions from the CTQ sexual abuse subscale (CTQ20, CTQ21, CTQ23, CTQ24, CTQ27) received positive weights, whereas age was negatively and sex inversely weighted. B) Depicted is the brain pattern of LV2, with positive weighting of voxels displayed in red and negative weighting displayed in blue color scale.

A)

B)
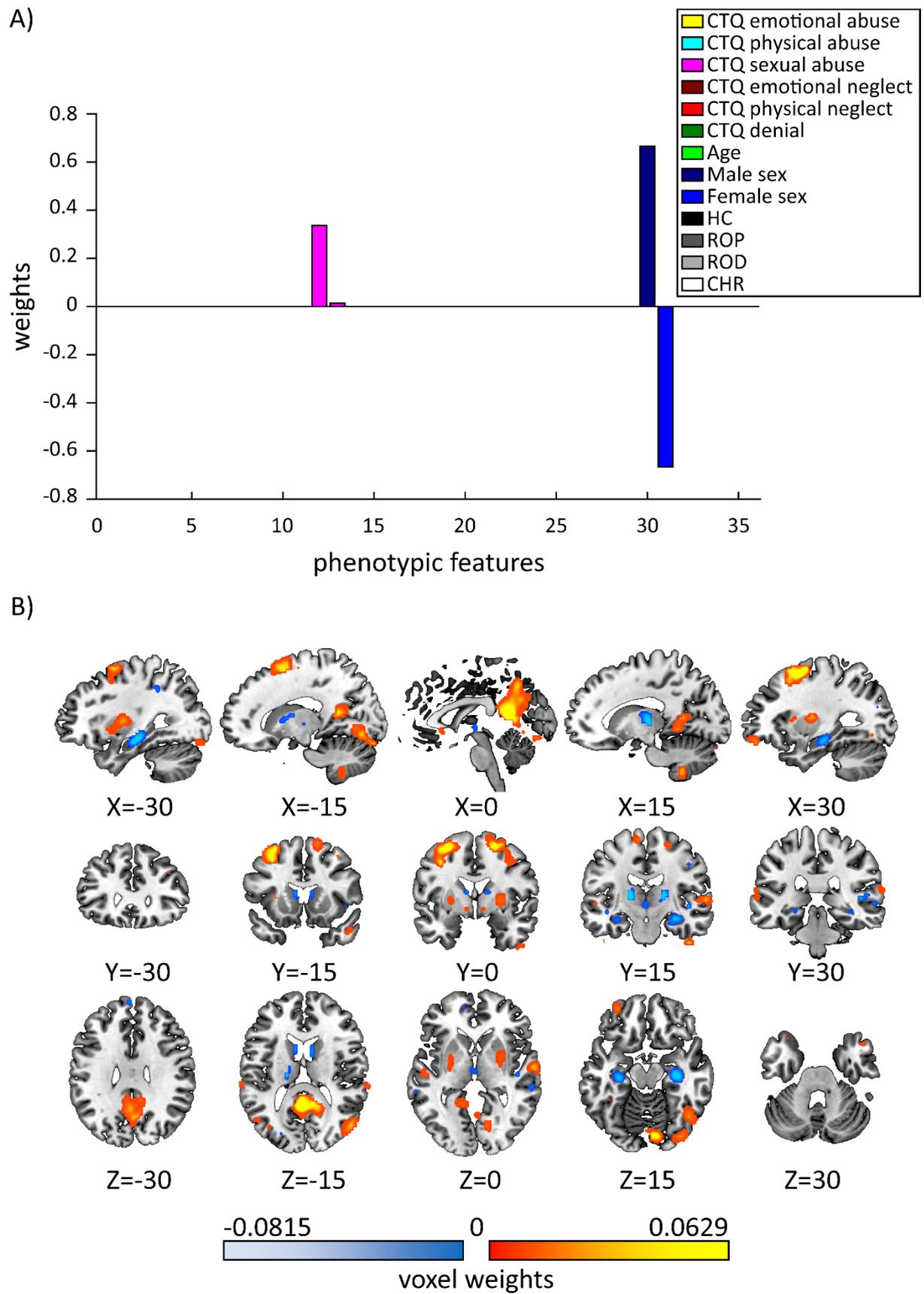
Figure S12: Predominantly sex-dependent signature of LV3 of LOSOCV analysis. A) The barplot visualizes the direction and the values of the weights included in the phenotypic pattern of LV3. Two questions from the sexual abuse (CTQ21, CTQ23) received a positive weight, while sex was strongly and inversely weighted. B) Depicted is the brain pattern of LV3, with positive weighting of voxels displayed in red and negative weighting displayed in blue color scale.

A)

B)

X=-30   X=-15   X=0   X=15   X=30

Y=-30   Y=-15   Y=0   Y=15   Y=30

Z=-30   Z=-15   Z=0   Z=15   Z=30

-0.0944   0   1.8e-3

voxel weights
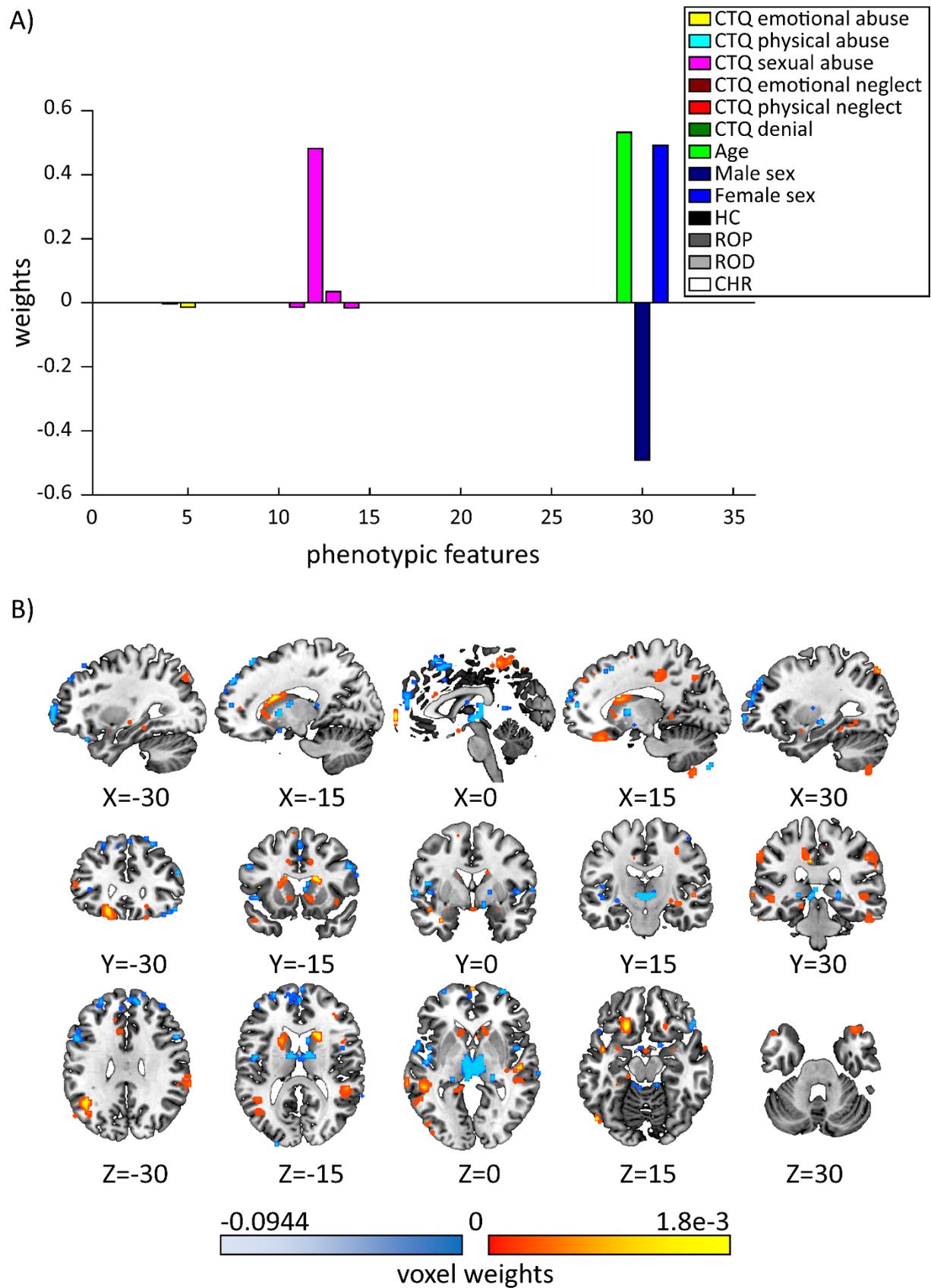
Figure S13: age- and sex-dependent sexual abuse signature of LV4 of the LOSOCV analysis. A) The barplot visualizes the direction and the values of the weights included in the phenotypic pattern of LV4. Two questions from the emotional abuse subscale (CTQ18, CTQ25) were negatively weighted and four questions from the sexual abuse (CTQ20, CTQ21, CTQ23, CTQ24) were mostly positively weighted. Age received a strong positive and sex a strong inverse weighting. B) Depicted is the brain pattern of LV4, with positive weighting of voxels displayed in red and negative weighting displayed in blue color scale

A)

B)
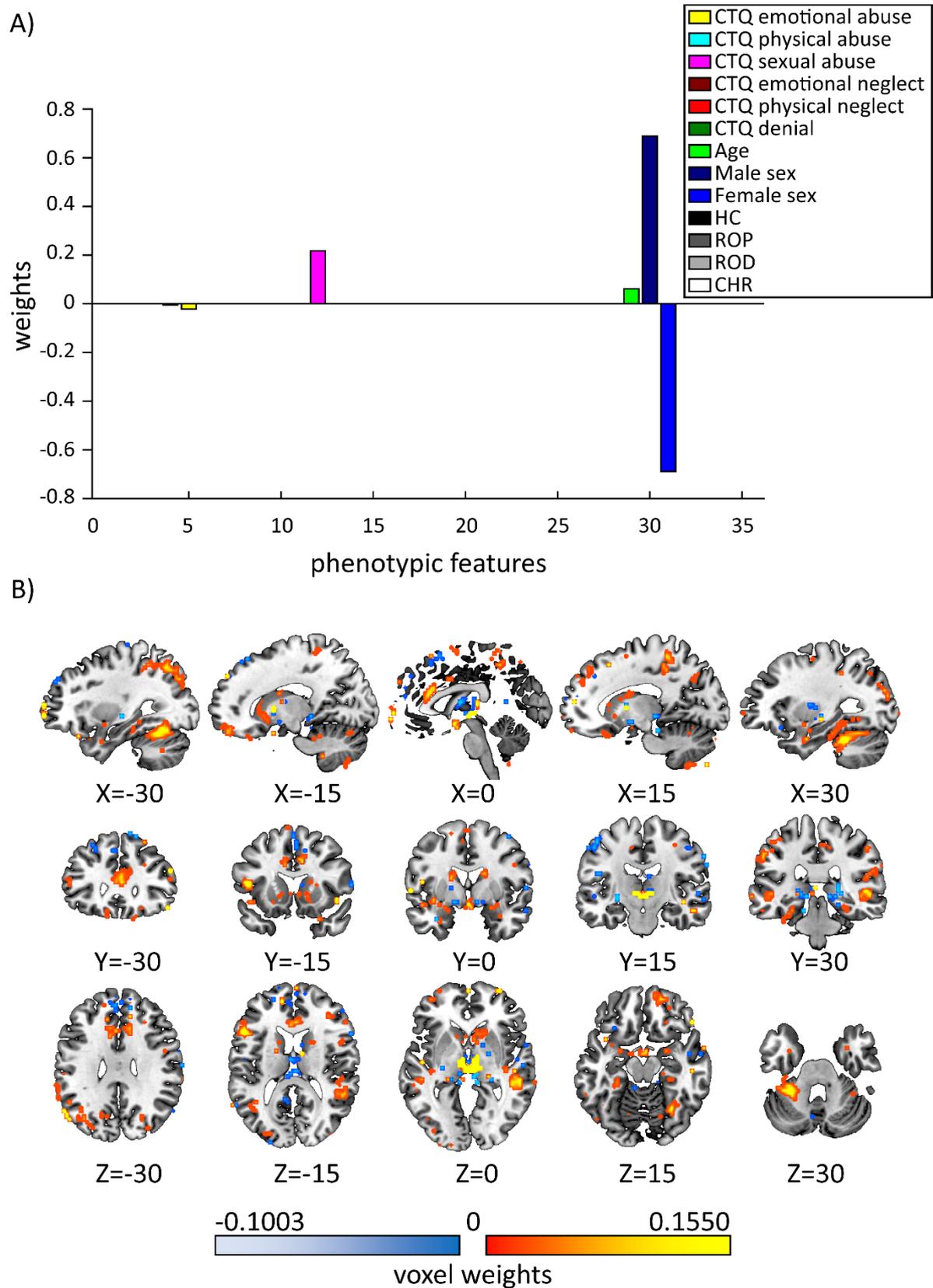
Figure S14: Predominantly sex-dependent signature of LV5 of the LOSOCV analysis. A) The barplot visualizes the direction and the values of the weights included in the phenotypic pattern of LV5. Two questions from the emotional abuse subscale (CTQ18, CTQ25) were negatively weighted and four questions from the sexual abuse (CTQ20, CTQ21, CTQ23, CTQ24) were mostly positively weighted. Age received a strong positive and sex a strong inverse weighting. B) Depicted is the brain pattern of LV5, with positive weighting of voxels displayed in red and negative weighting displayed in blue color scale.
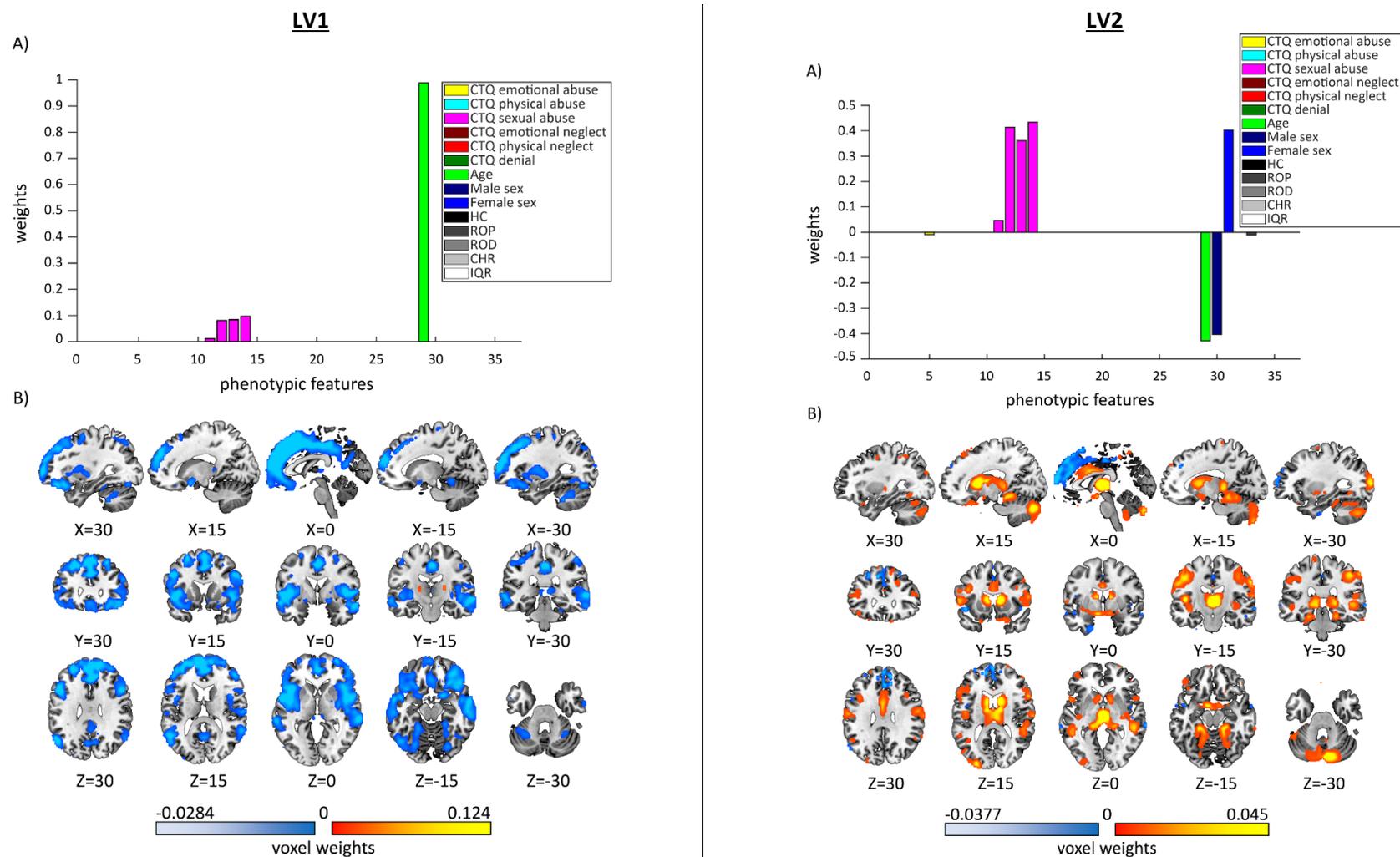
Figure S15: LV1 and LV2 signatures of SPLS analysis with IQR addition. LV1: A) Phenotypic pattern: Positive weights: sexual abuse (CTQ20, 21, 23, 24) and age. B) Brain pattern: positive voxel weights displayed in red and negative weights in blue color scale. LV2: A) Phenotypic pattern: Positive weights: sexual abuse (CTQ20, 21, 23, 24) and female sex. Negative weights: emotional abuse (CTQ25), age and male sex. B) Brain pattern: positive voxel weights displayed in red and negative weights in blue color scale.
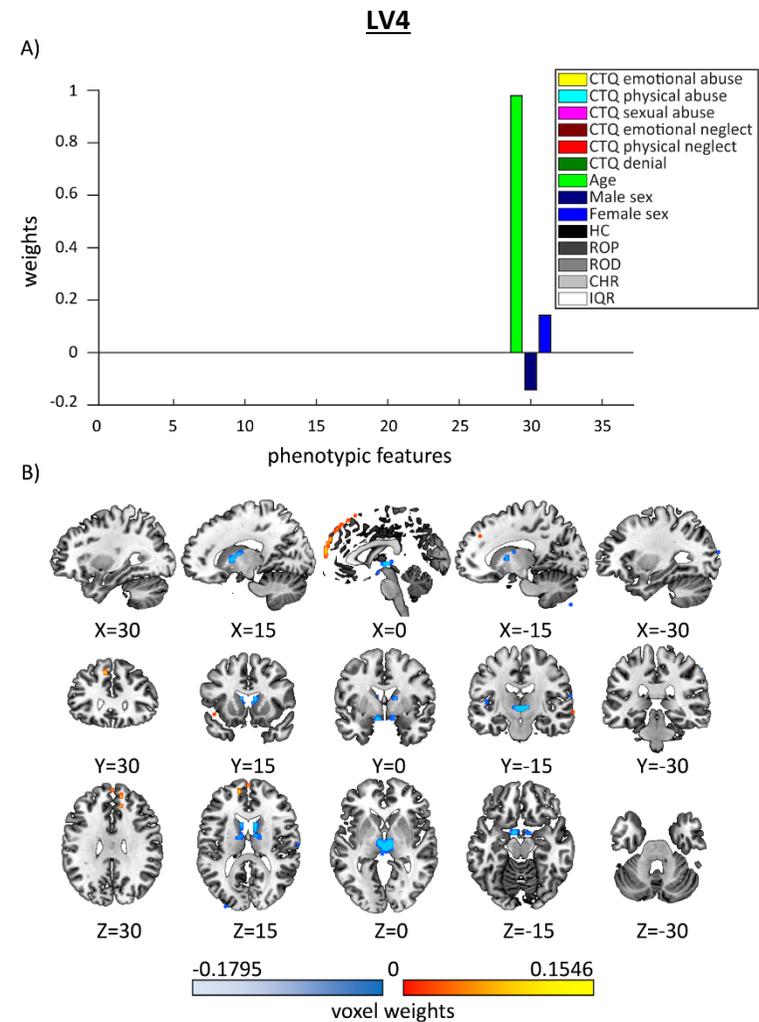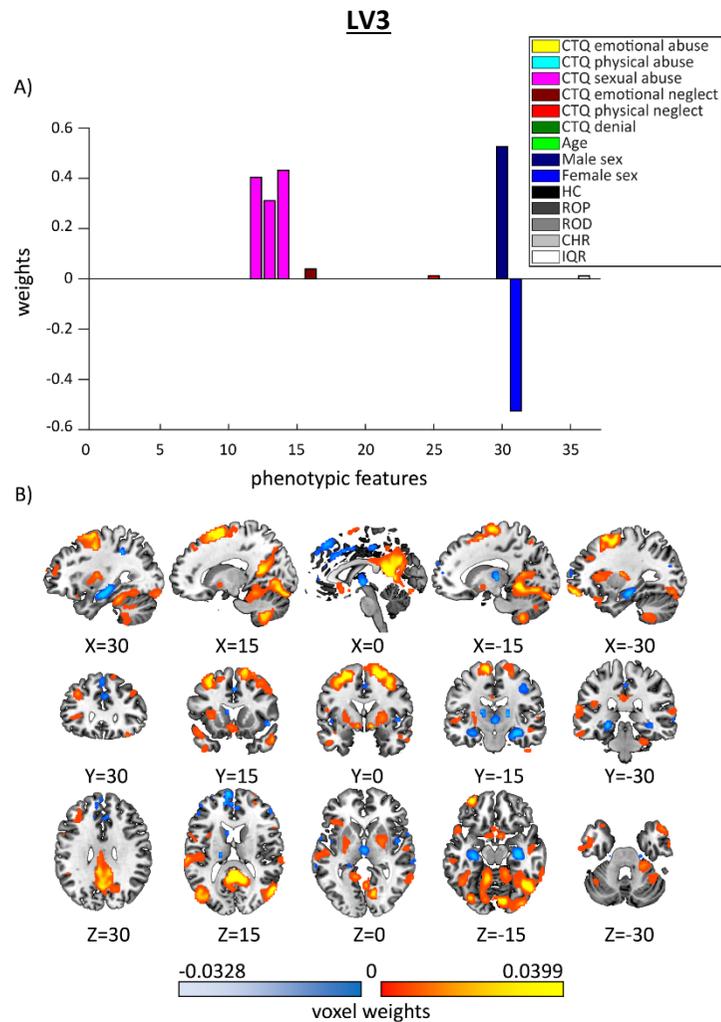
Figure S16: LV3 and LV4 signatures of SPLS analysis with IQR addition. LV3: A) Phenotypic pattern: Positive weights: sexual abuse (CTQ21, 23, 24), emotional (CTQ05) and physical neglect (CTQ26), male sex and IQR. Negative weights: female sex. B) Brain pattern: positive weighting of voxels in red and negative weighting in blue color scale. LV4: A) Phenotypic pattern: Positive weights: age and female sex; negative weight: male sex. B) Brain pattern: positive weighting of voxels in red and negative weighting in blue color scale.
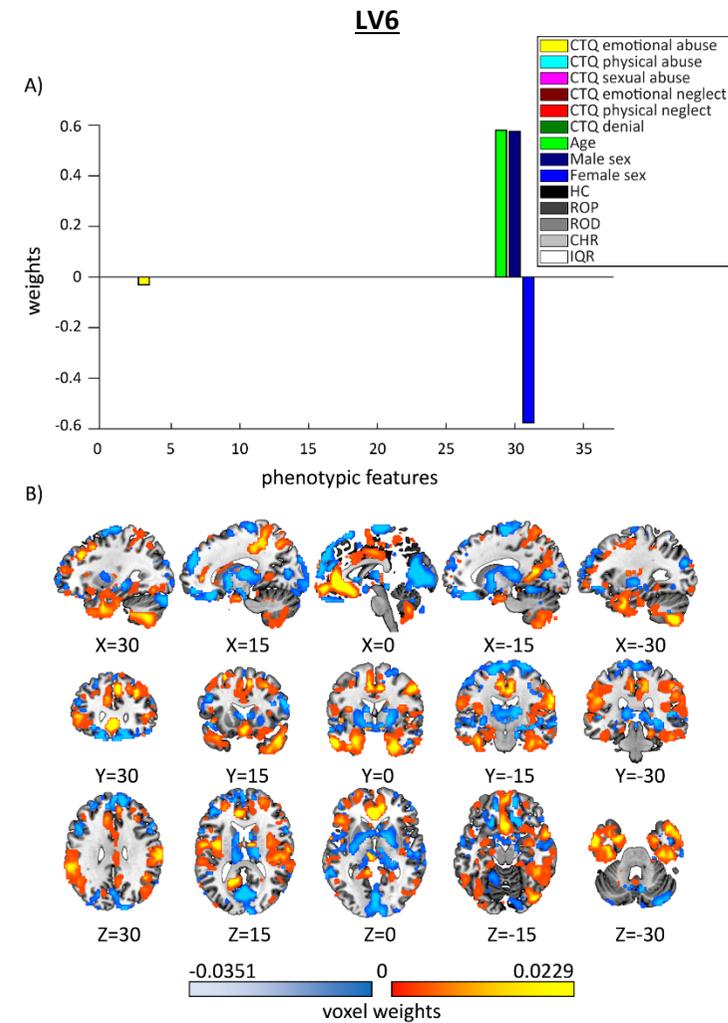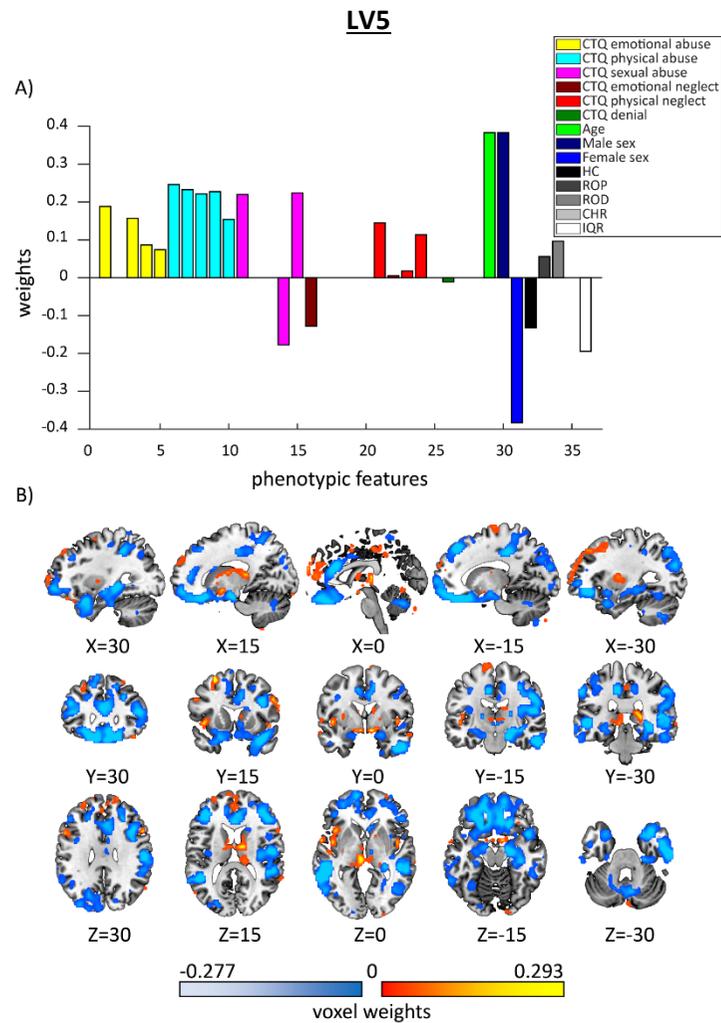
Figure S17: LV5 and LV6 signatures of SPLS analysis with IQR addition. LV5: A) Phenotypic pattern: Positive weights: emotional (CTQ03, 14, 18, 25), physical (CTQ09, 11, 12, 15, 17) and sexual abuse (CTQ20, 27), physical neglect (CTQ01, 02, 04, 06), age, male sex, ROP and ROD status. Negative weights: sexual abuse (CTQ24), emotional neglect (CTQ05), denial (CTQ10), female sex, HC status and IQR. B) Brain pattern: positive voxel weights displayed in red and negative weights in blue color scale. LV6: A) Phenotypic pattern: Positive weights: age and male sex. Negative weights: emotional abuse (CTQ14) and female sex. B) Brain pattern: positive weighting of voxels in red and negative weighting in blue color scale.

LV7

A)

B)

LV8

A)

B)
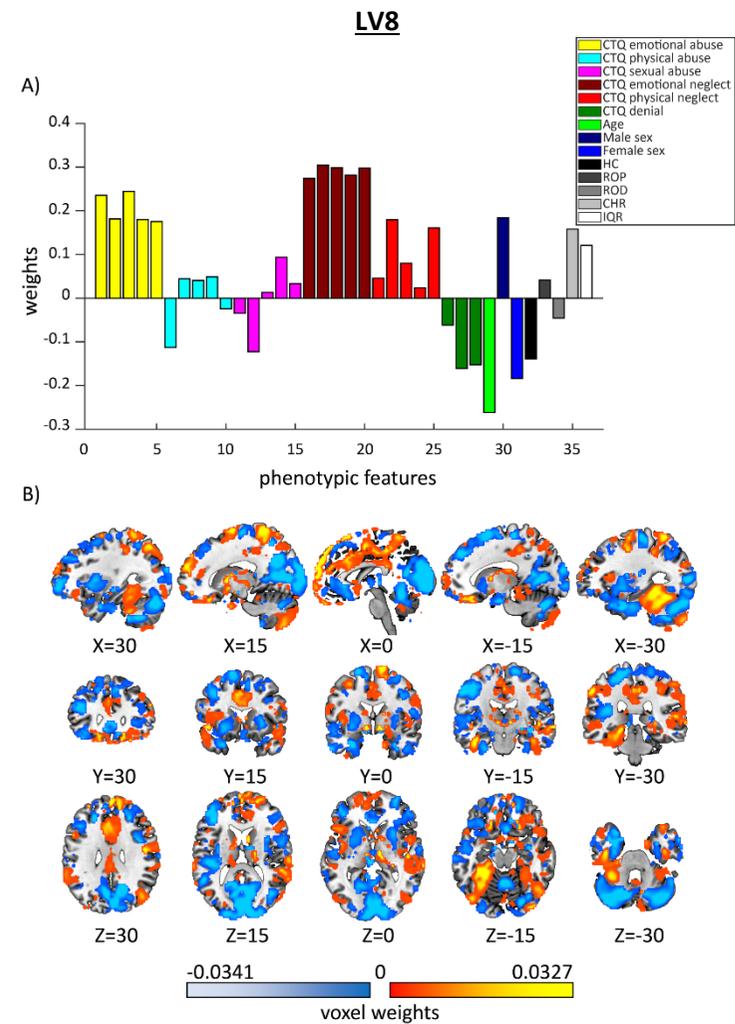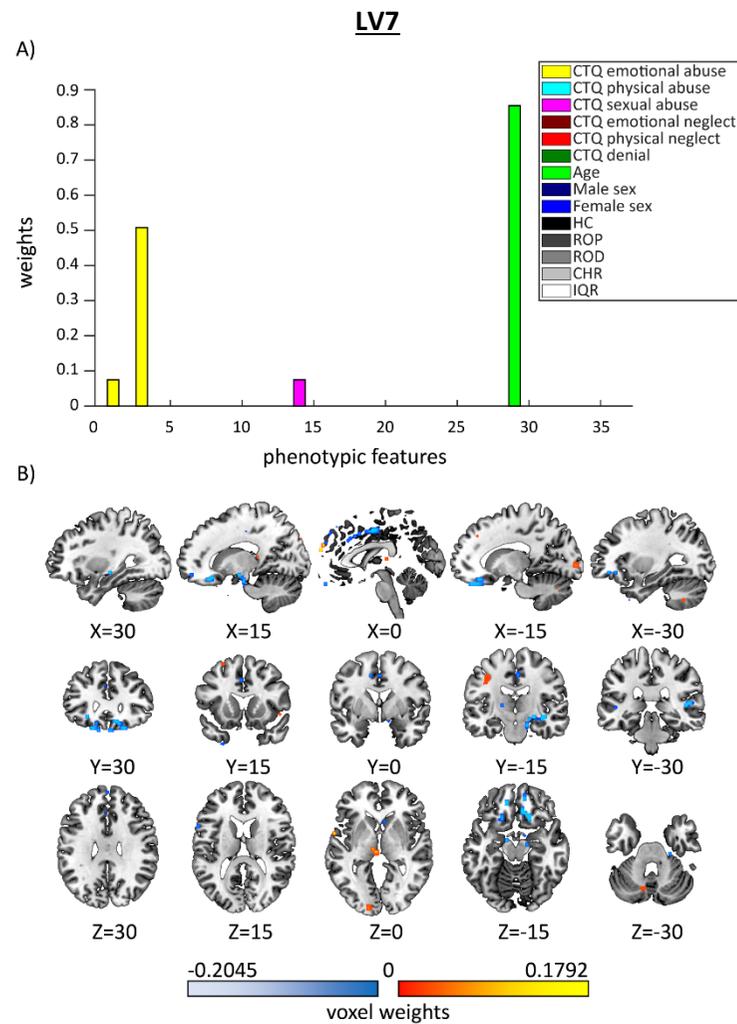
Figure S18: LV7 and LV8 signatures of SPLS analysis with IQR addition. <u>LV7</u>: A) Phenotypic pattern: Positive weights: emotional abuse (CTQ03, CTQ14), sexual abuse (CTQ24) and age. B) Brain pattern: positive voxel weights displayed in red and negative weights in blue color scale. <u>LV8</u>: A) Phenotypic pattern: Positive weights: emotional (CTQ03, 08, 14, 18, 25), physical (CTQ11, 12, 15) and sexual abuse (CTQ23, 24, 27), emotional (CTQ05, 07, 13, 19, 28) and physical neglect (CTQ01, 02, 04, 06, 26), male sex, ROP, CHR status and IQR. Negative weights: physical (CTQ09, 17) and sexual abuse (CTQ20, 21), denial (CTQ10, 16, 22), age, female sex, HC, and ROD status. B) Brain pattern: positive weighting of voxels in red and negative weighting in blue color scale.
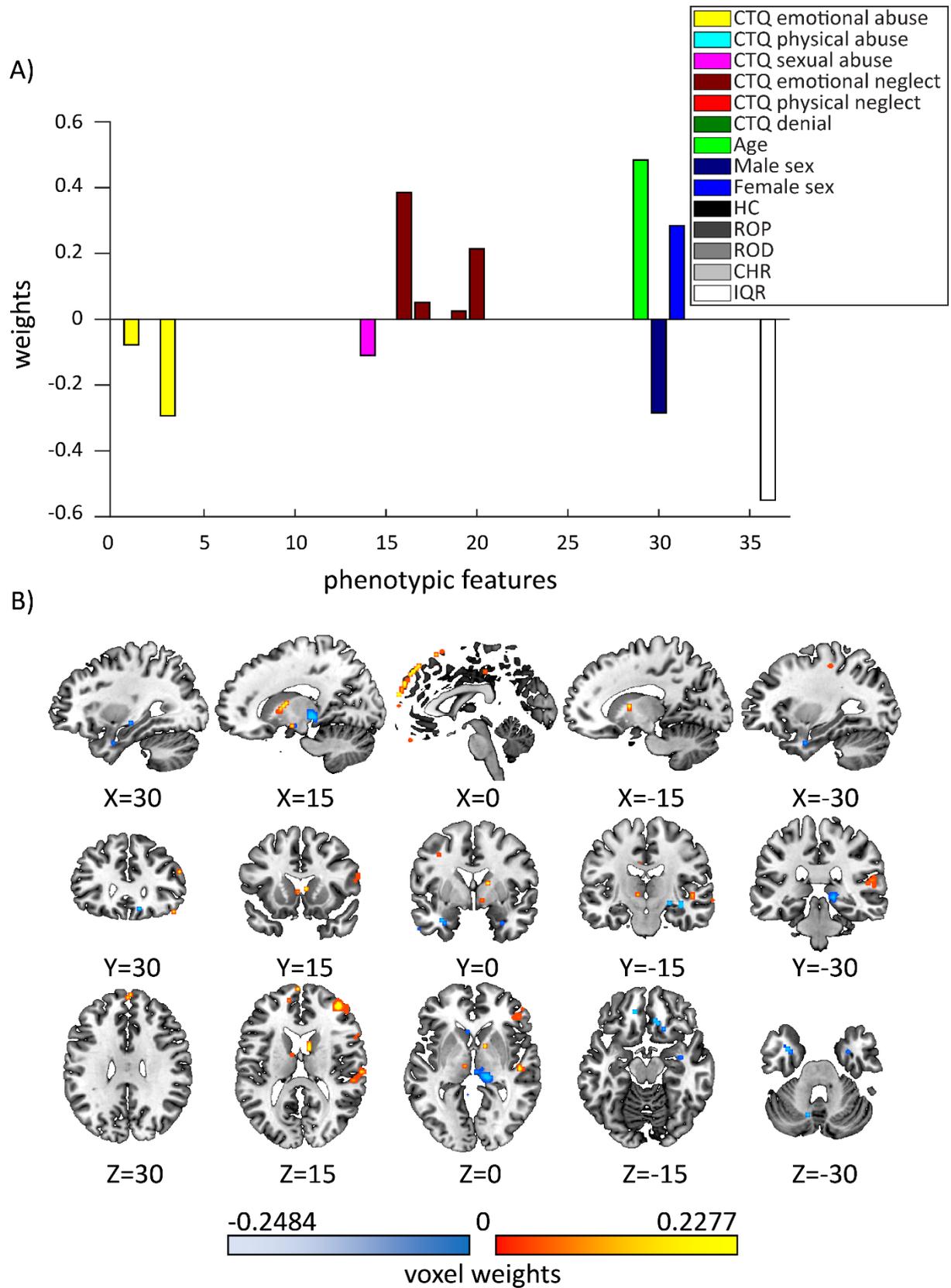
Figure S19: LV9 signature of SPLS analysis with IQR addition. A) Phenotypic pattern: Positive weights: emotional neglect (CTQ05, 07, 19, 28), age and female sex. Negative weights: emotional (CTQ03, 14) and sexual abuse (CTQ24), male sex and IQR. B) Brain pattern: positive voxel weights displayed in red and negative weights in blue color scale.

A)

B)
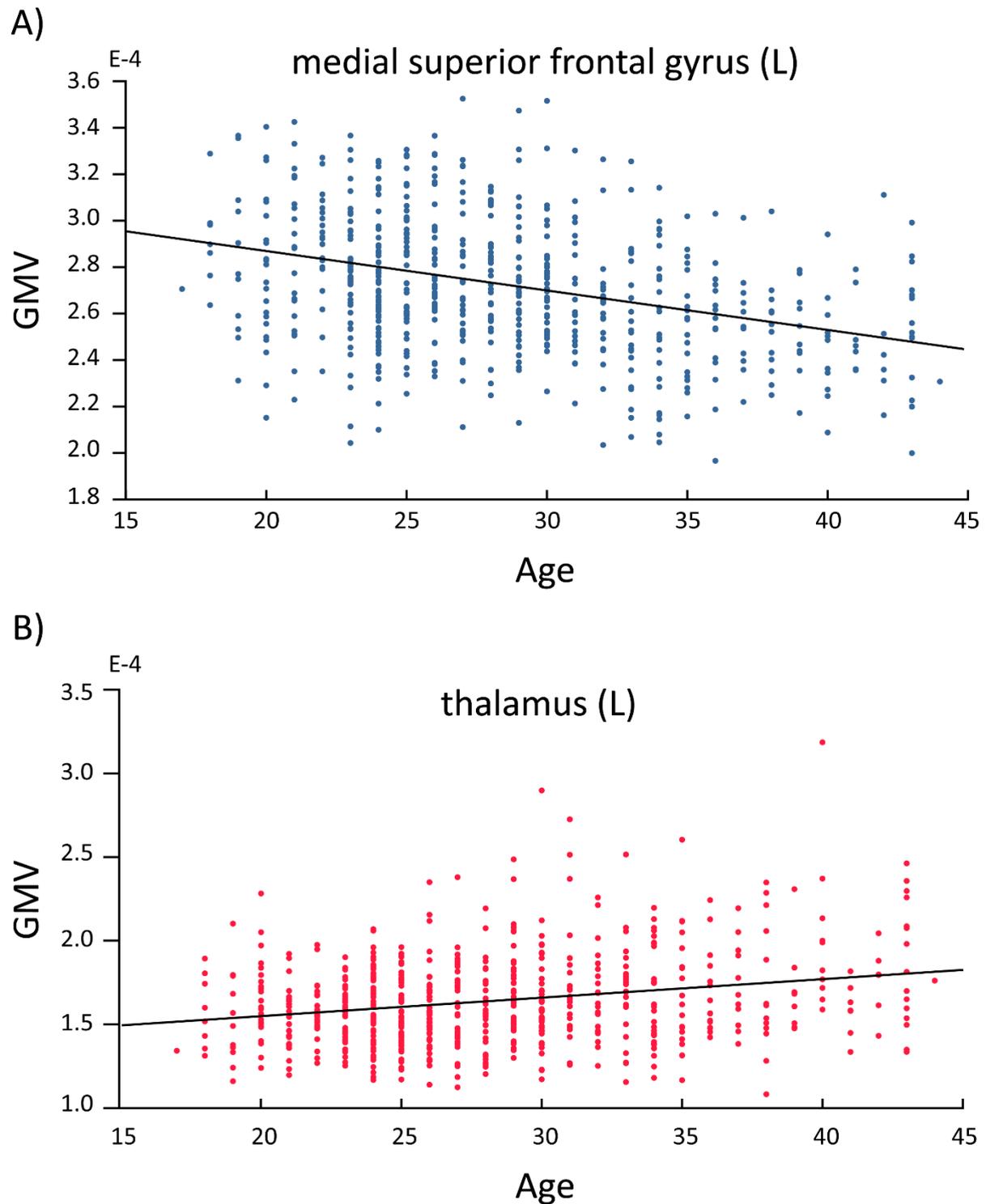
Figure S20: Exemplary illustration of phenotypic and brain correlations in LV1 of main SPLS model. A) Depicted is the correlation between age and mean grey matter volume (GMV) in the most strongly negatively weighted cluster in the left medial superior frontal gyrus (SFG). The left SFG was negatively weighted, whereas age was positively weighted in LV1. B) Depicted is the correlation between age and mean GMV in the most strongly positively weighted cluster in the left thalamus. Both age and the left lateral pre-frontal thalamus voxels were positively weighted in LV1.

Supplementary References

1. Koutsouleris N, Kambeitz-Ilankovic L, Ruhrmann S, Rosen M, Ruef A, Dwyer DB, et al. (2018): Prediction Models of Functional Outcomes for Individuals in the Clinical High-Risk State for Psychosis or With Recent-Onset Depression: A Multimodal, Multisite Machine Learning Analysis. *JAMA Psychiatry*. 75:1156-1172.

2. Schultze-Lutter F, Addington J, Ruhrmann S, Klosterkötter J (2007): Schizophrenia Proneness Instrument – Adult Version. Rome, Italy: Giovanni Fioriti Editore.

3. Miller TJ, McGlashan TH, Rosen JL, Cadenhead K, Cannon T, Ventura J, et al. (2003): Prodromal assessment with the structured interview for prodromal syndromes and the scale of prodromal symptoms: predictive validity, interrater reliability, and training to reliability. *Schizophr Bull*. 29:703-715.

4. Gaebel W, Hasan A, Falkai P (2019): *S3-Leitlinie Schizophrenie*. Berlin: Deutsche Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde e. V. (DGPPN).

5. Manjon JV, Tohka J, Garcia-Marti G, Carbonell-Caballero J, Lull JJ, Marti-Bonmati L, et al. (2008): Robust MRI brain tissue parameter estimation by multistage outlier rejection. *Magn Reson Med*. 59:866-873.

6. Rajapakse JC, Giedd JN, Rapoport JL (1997): Statistical approach to segmentation of single-channel cerebral MR images. *IEEE Trans Med Imaging*. 16:176-186.

7. Collins DL, Zijdenbos AP, Kollokian V, Sled JG, Kabani NJ, Holmes CJ, et al. (1998): Design and construction of a realistic digital brain phantom. *IEEE Transactions on Medical Imaging*. 17:463-468.

8. Monteiro JM, Rao A, Shawe-Taylor J, Mourao-Miranda J, Alzheimer's Disease I (2016): A multiple hold-out framework for Sparse Partial Least Squares. *J Neurosci Methods*. 271:182-194.

9. Wegelin JAJUoW, Tech. Rep (2000): A survey of Partial Least Squares (PLS) methods, with emphasis on the two-block case.

10. Witten DM, Tibshirani R, Hastie TJB (2009): A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. 10:515-534.

11. Zou H, Hastie T (2005): Regularization and variable selection via the elastic net. 67:301-320.

12. Witten DM, Tibshirani RJ (2009): Extensions of sparse canonical correlation analysis with applications to genomic data. *Stat Appl Genet Mol Biol*. 8:Article28.

13. Mackey L (2008): Deflation Methods for Sparse PCA. *Advances in Neural Information Processing Systems 21 - Proceedings of the 2008 Conference*. 21:1017-1024.

14. Monteiro JM, Rao A, Ashburner J, Shawe-Taylor J, Mourão-Miranda J (2016): Leveraging Clinical Data to Enhance Localization of Brain Atrophy. Cham: Springer International Publishing, pp 60-68.

15. Nichols TE, Holmes AP (2002): Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum Brain Mapp*. 15:1-25.

16. Fan L, Li H, Zhuo J, Zhang Y, Wang J, Chen L, et al. (2016): The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. *Cereb Cortex*. 26:3508-3526.

17. Diedrichsen J, Balsters JH, Flavell J, Cussans E, Ramnani N (2009): A probabilistic MR atlas of the human cerebellum. *Neuroimage*. 46:39-46.

18. Paolicelli RC, Bolasco G, Pagani F, Maggi L, Scianni M, Panzanelli P, et al. (2011): Synaptic pruning by microglia is necessary for normal brain development. *Science*. 333:1456-1458.

19. Ruigrok AN, Salimi-Khorshidi G, Lai MC, Baron-Cohen S, Lombardo MV, Tait RJ, et al. (2014): A meta-analysis of sex differences in human brain structure. *Neurosci Biobehav Rev*. 39:34-50.

20. Sacher J, Neumann J, Okon-Singer H, Gotowiec S, Villringer A (2013): Sexual dimorphism in the human brain: evidence from neuroimaging. *Magn Reson Imaging*. 31:366-375.

21. Satterthwaite TD, Wolf DH, Loughead J, Ruparel K, Elliott MA, Hakonarson H, et al. (2012): Impact of in-scanner head motion on multiple measures of functional connectivity: relevance for studies of neurodevelopment in youth. *Neuroimage*. 60:623-632.

22. Rosen AFG, Roalf DR, Ruparel K, Blake J, Seelaus K, Villa LP, et al. (2018): Quantitative assessment of structural image quality. *Neuroimage*. 169:407-418.

The non-specific nature of mental health and structural brain outcomes following childhood trauma

Author for correspondence:
Theresa K. Haidl,
E-mail: theresa.haidl@uk-koeln.de

Theresa K. Haidl[1,][*] [ORCID], Dennis M. Hedderich[2,3,][*], Marlene Rosen[1], Nathalie Kaiser[1], Mauro Seves[1], Thorsten Lichtenstein[2], Nora Penzel[1], Julian Wenzel[1], Lana Kambeitz-Ilankovic[1,4], Anne Ruef[4], David Popovic[4,5], Frauke Schultze-Lutter[6,7,8], Katharine Chisholm[9,10], Rachel Upthegrove[9,19], Raimo K. R. Salokangas[11], Christos Pantelis[12], Eva Meisenzahl[6], Stephen J. Wood[9,13,14], Paolo Brambilla[15,16], Stefan Borgwardt[17,18], Stephan Ruhrmann[1], Joseph Kambeitz[1,][†] and Nikolaos Koutsouleris[4,20,21,][†]

[1]Department of Psychiatry and Psychotherapy, Faculty of Medicine and University Hospital of Cologne, University of Cologne, Cologne, Germany; [2]Department of Diagnostic and Interventional Radiology, Faculty of Medicine and University Hospital of Cologne, University of Cologne, Cologne, Germany; [3]Department of Diagnostic and Interventional Neuroradiology, Klinikum rechts der Isar, School of Medicine Technical University of Munich, Munich, Germany; [4]Department of Psychiatry and Psychotherapy, Ludwig-Maximilians-University Munich, Munich, Germany; [5]International Max Planck Research School for Translational Psychiatry (IMPRS-TP), Munich, Germany; [6]Department of Psychiatry and Psychotherapy, Medical Faculty, Heinrich-Heine University, Düsseldorf, Germany; [7]Department of Psychology and Mental Health, Faculty of Psychology, Airlangga University, Surabaya, Indonesia; [8]University Hospital of Child and Adolescent Psychiatry and Psychotherapy, University of Bern, Bern, Switzerland; [9]Institute of Mental Health, University of Birmingham, Birmingham, UK; [10]Department of Psychology, Aston University, Birmingham, UK; [11]Department of Psychiatry, University of Turku, Turku, Finland; [12]Melbourne Neuropsychiatry Centre, University of Melbourne & Melbourne Health, Victoria, Australia; [13]Orygen, Melbourne, Australia; [14]Centre for Youth Mental Health, University of Melbourne, Melbourne, Australia; [15]Department of Neurosciences and Mental Health, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy; [16]Department of Pathophysiology and Transplantation, University of Milan, Milan, Italy; [17]Department of Psychiatry (Psychiatric University Hospital, UPK), University of Basel, Basel, Switzerland; [18]Department of Psychiatry and Psychotherapy, University of Lübeck, Lübeck, Germany; [19]Early Intervention Service, Birmingham Womens and Childrens NHS Foundation Trust, UK; [20]Institute of Psychiatry, Psychology and Neuroscience, London, UK and [21]Max-Planck Institute of Psychiatry Munich, Munich, Germany

Abstract

Background. Childhood trauma (CT) is associated with an increased risk of mental health disorders; however, it is unknown whether this represents a diagnosis-specific risk factor for specific psychopathology mediated by structural brain changes. Our aim was to explore whether (i) a predictive CT pattern for transdiagnostic psychopathology exists, and whether (ii) CT can differentiate between distinct diagnosis-dependent psychopathology. Furthermore, we aimed to identify the association between CT, psychopathology and brain structure.
Methods. We used multivariate pattern analysis in data from 643 participants of the Personalised Prognostic Tools for Early Psychosis Management study (PRONIA), including healthy controls (HC), recent onset psychosis (ROP), recent onset depression (ROD), and patients clinically at high-risk for psychosis (CHR). Participants completed structured interviews and self-report measures including the Childhood Trauma Questionnaire, SCID diagnostic interview, BDI-II, PANSS, Schizophrenia Proneness Instrument, Structured Interview for Prodromal Symptoms and structural MRI, analyzed by voxel-based morphometry.
Results. (i) Patients and HC could be distinguished by their CT pattern with a reasonable precision [balanced accuracy of 71.2% (sensitivity = 72.1%, specificity = 70.4%, $p \leq 0.001$]. (ii) Subdomains 'emotional neglect' and 'emotional abuse' were most predictive for CHR and ROP, while in ROD 'physical abuse' and 'sexual abuse' were most important. The CT pattern was significantly associated with the severity of depressive symptoms in ROD, ROP, and CHR, as well as with the PANSS total and negative domain scores in the CHR patients. No associations between group-separating CT patterns and brain structure were found.
Conclusions. These results indicate that CT poses a transdiagnostic risk factor for mental health disorders, possibly related to depressive symptoms. While differences in the quality of CT exposure exist, diagnostic differentiation was not possible suggesting a multi-factorial pathogenesis.

CAMBRIDGE
UNIVERSITY PRESS

CrossMark

Introduction

Childhood trauma (CT) is a frequent form of maltreatment comprising sexual, physical, and emotional dimensions. In Western countries, 30–40% of the adult population reported experiences with at least some form of maltreatment during childhood (Scher, Forde, McQuaid, & Stein, 2004). CT was revealed to influence the further course of life of the affected individuals, frequently leading to psychological symptoms and impairment in adulthood (Kessler et al., 2010; Scott, McLaughlin, Smith, & Ellis, 2012). It has been shown to be associated with an increased risk for psychiatric disorders such as major depression, anxiety disorders, addiction, post-traumatic stress disorder and psychosis, including patients at clinical high-risk for psychosis (CHR) (Kessler et al., 2010; Palmier-Claus, Berry, Bucci, Mansell, & Varese, 2016; Sahin et al., 2013; Scott et al., 2012; Varese et al., 2012). Even in the general population, CT seems to have long-standing effects on individuals' social perception (Salokangas, From, Luutonen, & Hietala, 2018). Due to its high prevalence and detrimental effects on both, mental health and associated socioeconomic costs (Fang, Brown, Florence, & Mercy, 2012), a better understanding of CT as a risk factor is essential. Furthermore, the fact that CT occurs during a period of important neurodevelopmental steps underlines the potential for prevention or better care for CT victims to contribute to lower lifetime burden of psychiatric disorders (Mikton & Butchart, 2009).

The sum of trauma exposure during childhood has been established as an important risk factor for mental health disorders. However, this has not been investigated in detail, although CTQ covers five different subcategories of different trauma exposure. These are in detail physical abuse (PA), physical neglect (PN), emotional abuse (EA), emotional neglect (EN), and sexual abuse (SA) (Bernstein & Fink, 1998). A promising approach to investigate the complex granularity of CT as a risk factor is multivariate pattern analysis (MVPA) which was previously shown to identify neuropsychiatric conditions based on, e.g. neuroimaging data (Kambeitz et al., 2015). The initial publication from the PRONIA study was on the prediction of functional and treatment outcomes based on clinical baseline data across multiple sites (Koutsouleris et al., 2018). Furthermore, two publications from the PRONIA consortium focused on different aspects of CT: Popovic et al. (2020) identified distinct volumetric brain patterns associated with single dimensions of CT (in particular physical and sexual abuse and emotional trauma) in a transdiagnostic approach. Salokangas et al. (2021) focused on CT in smaller patient groups and specifically investigated differences with respect to frontal lobe and hippocampal-amygdala complex volumes. In contrast, our study focuses on the potential ability of separating healthy controls (HC) and patient groups using machine learning techniques, and to identify potential clinical and volumetric brain correlates of CT in the entire cohort.

To answer these questions, the present study first investigated the discriminative value of CT for the individualized identification of transdiagnostic and diagnosis-specific psychiatric disorders using MVPA. In a second step, we examined whether the found CT patterns correlate with the measures of psychopathology and/or altered brain structure. The investigation was carried out in the PRONIA database ('Personalized Prognostic Tools for Early Psychosis Management'; www.pronia.eu), a large, multisite European cohort consisting of patients with recent onset depression (ROD), recent onset psychosis (ROP), CHR, and HC.

Aims of the study

We aimed to investigate whether (i) a predictive pattern of CT for transdiagnostic psychopathology exists, and whether (ii) CT can differentiate between distinct diagnosis-dependent psychopathology. Moreover, our aim was to identify associations between CT, psychopathology, and brain structure.

Methods

Participants

For the quality assurance of our proceedings, we followed the 'Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis' (TRIPOD) checklist for prediction model development and validation (Collins, Reitsma, Altman, & Moons, 2015).

All participants were recruited within the PRONIA project ('Personalized Prognostic Tools for Early Psychosis Management'). PRONIA is a multisite observational study funded by the European Union under the 7th Framework Programme (grant agreement n° 602152). Seven clinical centers in five European countries participated in the evaluation of patients with ROD, ROP, CHR, and HC. Within a longitudinal study design, a comprehensive battery of clinical assessment tools was used every 3 months over 18 months (see online Supplementary Fig. S1). Neuroimaging examinations were carried out at the baseline and the 9-month follow-up points. The entire study design has been previously described in detail by Koutsouleris et al. (2018).

All adult participants provided their written informed consent prior to study inclusion. Minors provided written informed assent and guardians written informed consent. The study was registered at the German Clinical Trials Register (DRKS00005042). The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. All procedures involving human subjects/patients were approved by the local research ethics committees.

Inclusion and exclusion criteria

The included persons were aged between 15 and 40 years and recruited into the study between 1 February 2014 and 1 May 2016. Patients with CHR were included by Cognitive Disturbances (COGDIS) criteria, assessed by the Schizophrenia Proneness Instrument (SPI-A) (Schultze-Lutter, Addington, Ruhrmann, & Klosterkötter, 2007), and/or UHR criteria (Phillips, Yung, & McGorry, 2000), assessed using a modified version of the Structured Interview for Prodromal Syndromes (SIPS) (McGlashan, Walsh, & Woods, 2010). For ROD, specific inclusion criteria were having a DSM-IV (American Psychiatric Association, 2000) Major Depressive Episode that was present within the past 3 months and did not last longer than 24 months. ROP fulfilled DSM-IV criteria for affective or non-affective psychosis within the last 24 months and not before. General inclusion and exclusion criteria have been described in detail in Koutsouleris et al. (2018) and were detailed as depicted in online Supplementary Table S1.

Procedure and instruments

The data used in this study were all acquired at baseline. As mentioned above, psychopathology of CHR patients was assessed

using SIPS and SPI-A. ROP and ROD were diagnosed by DSM-IV. Depressive syndrome severity was additionally measured using the Beck-Depression-Inventory II (BDI-II) (Hautzinger, Bailer, Worall, & Keller, 1995). Positive and negative symptoms were assessed by the Positive and Negative Syndrome Scale (PANSS) (Kay, Fiszbein, & Opler, 1987). For the assessment of CT, the Childhood Trauma Questionnaire (CTQ), developed by Bernstein and Fink (1998), was used. The CTQ is a self-assessment tool for the retrospective recording of mistreatment and neglect in childhood. It consists of 28 items, whereby three items (10, 16, 22) are used to determine denial and trivialization. It includes five subscales; emotional abuse (EA), physical abuse (PA), sexual abuse (SA), emotional neglect (EN), and physical neglect (PN). Rating was carried out on a five-point Likert scale (0 = never to 4 = very often). The convergent and discriminative validity has been reported as being good (Bernstein & Fink, 1998). In addition, the cumulative sum of the equivalent doses received until T0 was calculated for SSRIs (Hayasaka et al., 2015), chlorpromazine (Leucht, Samara, Heres, & Davis, 2016), olanzapine (Leucht et al., 2016), and benzodiazepines (diazepam) (Clinical Guidelines on Drug Misuse and Dependence Update 2017 Independent Expert Working Group, 2017).

MRI acquisition, preprocessing, and analysis

Participants underwent a comprehensive imaging protocol at seven sites respecting a minimal harmonization protocol including high-resolution 3D T1-weighted imaging. Detailed scanner and sequence specifications for all sites can be found in the online Supplementary Table S2. All images underwent quality control and were preprocessed using the CAT12 toolbox (version r1155; http://dbm.neuro.uni-jena.de/cat12/), an extension of SPM 12 as described previously (Koutsouleris et al., 2018). Images were smoothed with 10 mm before entering the subsequent analysis steps. The Quality Assurance framework of CAT12 was used to empirically check the quality of the GMV maps.

By computing the correlation of each image to all other 592 images, we found 11 (1.9%) images whose correlation exceeded two standard deviations from the sample mean. These images were inspected and nine were removed because of MRI artifacts. Thus, 583 persons could be included in the VBM analysis (109 CHR patients, 115 ROD patients, 110 ROP patients, and 249 HC). Notably, 98.47% of the images achieved a good overall weighted quality (B), and 83.0% of the data quality was rated with a B+ as provided by the internal quality assessment of CAT12 (Gaser & Dahnke, 2016). For analysis of brain structure and associations with CT, voxel-based morphometry (VBM) was employed. Preprocessed data entered a full-factorial general linear model design as implemented in SPM. Sex, site (coded as dummy regressors), and age were used as covariates of no interest to correct for potential confounds for VBM analyses. In order to investigate possible sex differences, male and female participants were also analyzed separately. Global proportional scaling for total intracranial volume was used to adjust for different global brain volume differences. Contrasts were defined for main-effects and interaction analyses to assess differences in mean and slope effects of associations between CTQ-based decision scores (DS) and local GM. Threshold-free cluster enhancement (TFCE) was used as implemented in the TFCE toolbox for SPM with 5000 permutations (Smith & Nichols, 2009). Significance threshold was set at $p < 0.05$, family-wise error corrected.

Machine learning strategies

To investigate discriminative patterns of CT experience in HC *v.* the combined three patient groups (PAT), we used an L2-regularized logistic regression (L2-LR) as provided by the LIBLINEAR library (Fan, Chang, Hsieh, Wang, & Lin, 2008), which offers methods for classifying individuals instead of describing statistical group differences.

We used our open-source machine learning toolkit NeuroMiner (https://github.com/neurominer-git/) to implement a fully automated machine learning pipeline. We trained different models to predict psychiatric disorders based on the single CTQ items;

1. PAT *v.* HC
2. HC *v.* CHR; HC *v.* ROD; HC *v.* ROP
3. ROD *v.* CHR; ROD *v.* ROP; CHR *v.* ROP

We followed the internal–external validation approach recommended for the assessment of model generalizability in multi-site studies (Steyerberg & Harrell, 2016) and validated our models using nested leave-one-site-out cross-validation (LOSOCV) (see in detail Supplementary Methods).

To compare the multivariate *v.* univariate methods, we repeated the HC *v.* PAT analysis after replacing the L2-LR (Fan et al., 2008) algorithm with a univariate logistic regression model (uLR) in NeuroMiner. Algorithm performance was measured using the balanced accuracy (BAC) of the out-of-training (OOT) group membership predictions and assessed for significance using 1000 random label permutations (Golland & Fischl, 2003). Predictive features for each L2-LR model were compared by their mean weights.

A further validation analysis assessed whether our model generalized across study groups. Therefore, we used LOSOCV to train and cross-validate three binary L2-LR-based diagnostic classifiers (HC *v.* CHR; HC *v.* ROD; HC *v.* ROP) using the identical algorithmic setup described above. Each trained classification ensemble was then applied to the CTQ data of the other two clinical study groups following an out-of-sample cross-validation (OOCV) approach. Class membership probabilities/DS of the patients in the held-back study groups were computed for the OOT predictions.

These main analyses were supplemented by an investigation of univariate associations between measures of *current* psychopathology and the OOT DS of clinical participants produced by the L2-LR algorithms, which were trained in the HC *v.* CHR, HC *v.* ROD, and HC *v.* ROP comparisons. For the ROD, ROP, and CHR groups, the correlations of the CTQ-based DS with the BDI-II, SIPS-P, SIPS-N, SIPS-D, SIPS-G, PANSS total, positive, negative, and general domain scores were examined, respectively. Furthermore, the relationship between the equivalent doses of the individual drug classes (neuroleptics, SSRIs, benzodiazepines) and the CTQ-based DS was calculated for each group. In order to exclude recall bias in older participants (with longer time spans between CT and study inclusion), we performed correlation analyses between CTQ-based DS and age at study inclusion as a control analysis.

Results

Study group characteristics

In total, 643 subjects (57.2% male, mean age 27.69 ± 5.99 years) were included in the analysis. These consisted of $n = 262$ (40.7%) HC, $n = 122$ (19.0%) CHR, $n = 130$ (20.2%) ROD, and

Table 1. Sociodemographic data and general psychopathology

	HC	PAT	U/χ^{2a}	p^a	CHR	ROD	ROP
n total (%)	262 (40.7%)	381 (59.3)	n.a.	n.a.	122 (19.0%)	130 (20.2%)	129 (20.1%)
Age (y), M (s.d.)	27.75 (6.41)	27.64 (5.68)	49 854	0.932	26.26 (4.9)	28.54 (6.14)	28.02 (5.68)
Sex (♀) F (%)	164 (62.6%)	204 (53.5)	16.23	<0.001	58 (47.5%)	70 (53.8%)	49 (38%)
Psychopathology [mean (s.d.)]							
BDI-II	3.76 (5.27)	24.12 (13.04)	6273.5	<0.001	24.89 (12.16)	26.71 (13.91)	20.79 (12.30)
CTQ	30.88 (6.4)	40.38 (12.64)	20 806.5	<0.001	41.28 (12.73)	39.33 (13.66)	40.57 (11.42)
PANSS total	n.a.	56.32 (18.90)	n.a.	n.a.	50.74 (13.11)	47.80 (11.33)	70.11 (21.70)
PANSS negative	n.a.	13.83 (6.39)	n.a.	n.a.	12.54 (5.83)	12.60 (5.00)	16.27 (7.38)
PANSS positive	n.a.	11.97 (6.05)	n.a.	n.a.	10.27 (2.95)	7.71 (1.39)	17.84 (6.52)
PANSS general	n.a.	30.49 (9.39)	n.a.	n.a.	27.83 (6.88)	27.48 (6.97)	35.99 (11.07)
Childhood trauma							
Emotional abuse	6.56 (2.42)	9.64 (4.37)	24 784.0	<0.001	10.16 (4.43)	9.2 (4.36)	9.62 (4.29)
Physical abuse	5.39 (1.0)	6.52 (3.08)	38 805.5	<0.001	6.56 (3.11)	6.45 (3.21)	6.56 (3.0)
Sexual abuse	5.2 (1.1)	6.04 (2.95)	39 828.5	<0.001	5.97 (2.77)	5.88 (2.84)	6.28 (3.22)
Emotional neglect	7.93 (3.14)	11.47 (4.58)	25 248.5	<0.001	11.78 (4.45)	11.25 (4.88)	11.4 (4.42)
Physical neglect	5.87 (1.51)	7.41 (2.73)	29 420.0	<0.001	7.35 (2.6)	7.08 (2.81)	7.79 (2.76)
Distribution across sites (total/%)							
Munich	59 (22.5)	125 (32.8)	n.a.	n.a.	40 (32.8)	44 (33.8)	41 (31.8)
Basel	44 (16.8)	51 (13.4)	n.a.	n.a.	18 (14.8)	17 (13.1)	16 (12.4)
Cologne	56 (21.4)	69 (18.1)	n.a.	n.a.	18 (14.8)	25 (19.2)	26 (20.2)
Birmingham	42 (16.0)	34 (8.9)	n.a.	n.a.	13 (10.7)	12 (9.2)	9 (7.0)
Turku	19 (7.3)	45 (11.8)	n.a.	n.a.	14 (11.5)	11 (8.5)	20 (15.5)
Udine	31 (11.8)	31 (8.1)	n.a.	n.a.	12 (9.8)	14 (10.8)	5 (3.9)
Milan	11 (4.2)	26 (6.8)	n.a.	n.a.	7 (5.7)	7 (5.4)	12 (9.3)

U, Mann–Whitney *U* test; χ^2, chi-squared test, M, mean; s.d., standard deviation; PAT, patients including ROP, ROD, and CHR; HC, healthy controls; CHR, clinical high-risk state; ROD, recent onset depression; ROP, recent onset psychosis; CTQ, Childhood Trauma Questionnaire, PANSS, Positive and Negative Syndrom Scale; BDI-II, Beck Depression Inventory II.
Statistical comparisons: sex with χ^2 statistics; age, BDI-II, and CTQ with Mann–Whitney *U* test.
Comparison between healthy controls and patients.
[a]Comparison only between PAT and HC.

n = 129 (20.1%) ROP. CTQ total scores and subdomain scores were significantly different between PAT and HC. No group differences were found between the PAT groups ROP, ROD, and CHR regarding the total CTQ score. Please see Table 1 and online Supplementary Table S3 for details. The online Supplementary Table S4 shows the mean values of the drug equivalent doses that have been taken cumulatively so far. As expected, the highest equivalent doses for antipsychotics were found in ROP patients (chlorpromazine = 8072.71 mg/d, olanzapine = 280.02 mg/d), followed by CHR patients (chlorpromazine = 1025.06 mg/d, olanzapine = 42.04 mg/d). Surprisingly, the highest equivalent doses for SSRIs were found in CHR patients (3864.95 mg/d), followed by ROD patients (2630.83). We did not find an association between patient age and DS, making an age-dependent recall bias unlikely to have influenced our results (see online Supplementary Table S5).

Childhood trauma profiles predict general psychopathology

The classifier distinguishing HC from PAT performed with a BAC of 71.2% (sensitivity: 72.1%, specificity: 70.4%). Leave-site-out

validation yielded good generalizability of the CTQ-based discriminative model (see Table 2). In order to deduct a CT profile predictive of general psychopathology, weights of CTQ single items from the MVPA were recorded and are depicted in Fig. 1. It must be emphasized that the resulting values do not allow to conclude on the direction of the prediction. The highest weights related to items within the subdomains EN and EA, namely: CTQ Item 5; 'There was someone in my family who helped me feel that I was important or special', CTQ Item 14; 'People in my family said hurtful or insulting things to me', and CTQ Item 13; 'People in my family looked out for each other'. The uLR analyses for the same classification (HC *v.* PAT) led to a BAC of 67.1% (sensitivity: 66.0%, specificity: 68.2%). For detailed results, please see Table 2.

Childhood trauma profiles for diagnosis-specific psychopathology

Classifying the three diagnostic groups within the PAT cohort, namely CHR, ROD, and ROP did not perform above chance level (CHR *v.* ROD: BAC = 46.1%, sensitivity = 35.8%, specificity

Table 2. Multivariate analyses

Classifier	TP	TN	FP	FN	Sens	Spec	BAC	PPV	NPV	PSI	AUC	*p*
Leave-site-out performance												
HC *v.* PAT (L2LR)	186	266	112	72	72.1	70.4	71.2	62.4	78.7	41.1	0.77	<0.001
HC *v.* PAT (GLM)	173	260	121	89	66	68.2	67.1	58.8	74.5	33.3	0.74	<0.001
ROD *v.* ROP	69	55	65	50	58.0	45.3	51.9	51.5	52.4	3.9	0.49	0.358
CHR *v.* ROP	51	62	58	69	42.5	51.7	47.1	46.8	47.3	−5.9	0.48	0.866
CHR *v.* ROD	43	67	52	77	35.8	56.3	46.1	45.3	46.5	−8.2	0.43	0.923
HC *v.* ROD	198	76	53	64	75.6	58.9	67.2	78.9	54.3	33.2	0.69	<0.001
HC *v.* CHR	189	84	33	72	72.4	71.8	72.1	85.1	53.8	39.0	0.72	<0.001
HC *v.* ROP	195	86	42	67	74.4	67.2	70.8	82.3	56.2	38.5	0.75	<0.001
Leave-group-out performance												
HC *v.* CHR OOCV ROP	249	51	78	13	95	39.5	67.3	76.1	79.7	55.8	0.79	<0.001
HC *v.* CHR OOCV ROD	249	38	92	13	95	29.2	62.1	73.0	74.5	47.5	0.74	<0.001
HC *v.* ROD OOCV ROP	249	51	78	13	95	39.5	67.3	76.1	79.7	55.8	0.76	<0.001
HC *v.* ROD OOCV CHR	249	48	74	13	95	39.3	67.2	77.1	78.7	55.8	0.77	<0.001
HC *v.* ROP OOCV CHR	248	49	73	14	94.7	40.2	67.4	77.3	77.8	55.0	0.79	<0.001
HC *v.* ROP OOCV ROD	248	45	85	14	94.7	34.6	64.6	74.5	76.3	50.7	0.71	<0.001

TP, true positive; TN, true negative; FP, false positive; FN, false negative; Sens, sensitivity; Spec, specificity; BAC, balanced accuracy; PPV, positive predictive value; NPV, negative predictive value; PSI, prognostic summary index; AUC, area-under-the-curve; HC, healthy controls; PAT, patients including ROP, ROD, and CHR; ROD, recent onset depression; ROP, recent onset psychosis; CHR, clinically high-risk; OOCV, out-of-sample cross-validation.
All analyses were single item based.

= 56.3%; CHR *v.* ROP: BAC = 47.1%, sensitivity = 42.5%, specificity = 51.7%; ROD *v.* ROP: BAC = 51.9 sensitivity = 58.0%, specificity = 45.3%). However, classifiers separating between HC and individual PAT groups performed well (HC *v.* ROD: BAC = 67.2%, sensitivity = 75.6%, specificity = 58.9%; HC *v.* CHR: BAC = 72.1%, sensitivity = 72.4%, specificity = 71.8%; HC *v.* ROP: BAC = 70.8%, sensitivity = 74.4%, specificity = 67.2%; please see Table 2).

Regarding the differentiation of HC *v.* CHR, highest weights belonged to items of the subdomains EA and EN (see Fig. 1); CTQ Item 14; 'People in my family said hurtful or insulting things to me', CTQ Item 13; 'People in my family looked out for each other', and CTQ Item 28; 'My family was a source of strength and support'.

Analyzing the profile of HC *v.* ROD revealed the highest weights in items of the subdomains PA, SA, and EA (see Fig. 1); CTQ Item 17; 'I got hit or beaten so badly that it was noticed by someone like a teacher, neighbor, or doctor', CTQ Item 24; 'Someone molested me', and CTQ Item 14; 'People in my family said hurtful or insulting things to me'.

Describing the profile which is distinguishing HC *v.* ROP, items of the subdomains EA, EN, and PN were most predictive (see Fig. 1); CTQ Item 25; 'I believe that I was emotionally abused', CTQ Item 13; 'People in my family looked out for each other', and CTQ Item 2; 'I knew that there was someone to take care of me and protect me'.

Correlation between childhood trauma and psychopathology

Across all groups, correlations between the CTQ-based DS and GAF symptoms ($r = 0.388$, $p \leq 0.01$) as well as disability and impairment ($r = 0.412$, $p \leq 0.01$) were moderate to strong. In the CHR group, there were no associations between the CTQ-based DS and any SIPS domain, but a weak correlation between the DS and the BDI total score was observed ($r = -0.175$, $p = 0.028$). Moreover, a weak correlation between the PANSS total ($r = -0.191$, $p = 0.038$) and the PANSS negative domain score ($r = -0.196$, $p = 0.033$) was seen in the CHR patients. Regarding the ROD group, a moderate association between the CTQ-based DS and the BDI total score was found ($r = -0.278$, $p = 0.001$). In the ROP group, there was no significant correlation between the PANSS scores and the CTQ-based DS but a moderate association between the BDI total score and the CTQ-based DS ($r = -0.246$, $p = 0.003$). For details, please see Table 3.

Correlation between childhood trauma and medication

Across all groups, weak negative correlations were found between the CTQ-based DS and all types of medication [chlorpromazine $r = -0.213$, $p \leq 0.001$, olanzapine $r = -0.213$, $p \leq 0.001$, SSRI $r = -0.193$, $p \leq 0.001$, benzodiazepine (diazepam) $r = -1.28$, $p = 0.001$]. Interestingly, however, no significant correlations were found in the individual groups, except for a weak positive correlation with benzodiazepine in HC individuals. For details, please see online Supplementary Table S6.

Correlation between childhood trauma and brain structure

Despite several methodological approaches and adjusted statistical thresholds, we did not find any associations between CTQ-based DS and brain morphology in our cohort. Additionally, there were no significant associations between DS and brain morphology when examining male and female participants separately, also suggesting no sex-specific brain alterations associated with CTQ-based DS.
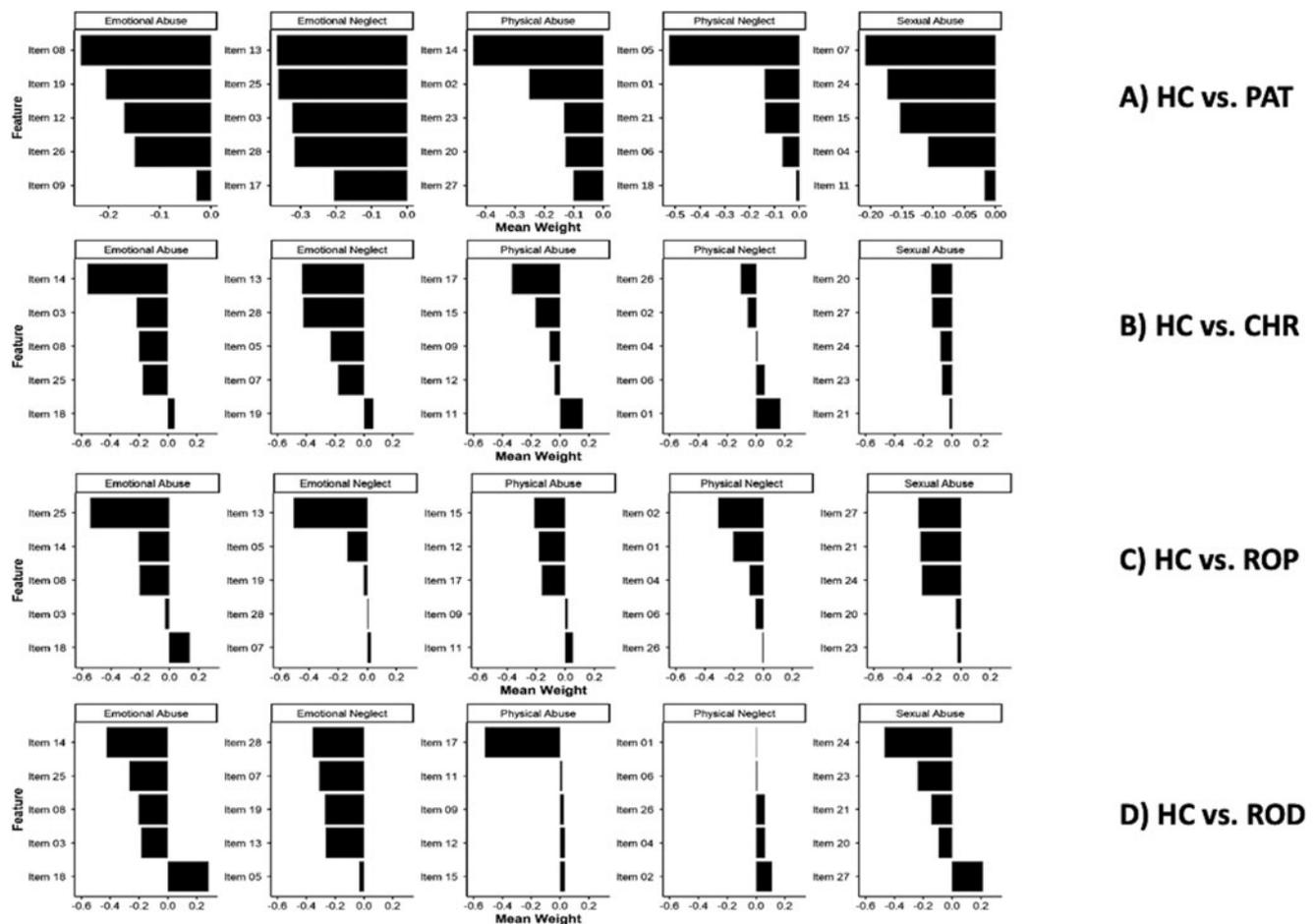
Fig. 1. Predictive pattern of the single CTQ items and related subdomains in the different diagnostic groups.

Discussion

We investigated CT and psychopathology in a large cohort of HC and patients with ROD, ROP, and CHR using MVPA. We found that CT significantly predicted transdiagnostic psychopathology using MVPA, while separation of diagnosis-specific psychopathology was not achieved. Qualitative analysis of CT patterns emphasized the importance of EN and EA for ROP and CHR identification while PA and SA yielded importance in ROD patients. The CTQ-based DS was significantly associated with the current severity of depressive symptoms in the ROD, ROP, and CHR group. Moreover, a correlation between the CTQ-based DS and the PANSS total and negative domain score was found in CHR patients. However, no further associations with psychopathology or structural brain alterations were found. Weak correlations between CTQ-based DS and medication were discovered across all groups, while no correlations were observed in the single groups, except for a weak positive correlation with benzodiazepine in HC individuals. The latter might reflect negative consequences of CT at a subthreshold level, resulting in higher tension and anxiety treated with benzodiazepine.

In order to investigate the association between CT and psychopathology, we tested whether PAT and HC could be separated based on CTQ information using a machine-learning model. We found that this distinction could be made with acceptable accuracy on the individual level and that the highest weights were assigned to domains pertaining to EA and EN. CT has

been associated with several specific psychiatric disorders such as psychosis (Varese et al., 2012), unipolar depression (Rubino, Nanni, Pozzi, & Siracusano, 2009), and bipolar disorder (Palmier-Claus et al., 2016) and has been posited as a general risk factor for their development. Recent reviews and meta-analyses have shown that each subdomain of the CTQ is by itself significantly associated with the occurrence of psychiatric illness (Lindert et al., 2014; Nelson, Klumparendt, Doebler, & Ehring, 2017; Varese et al., 2012). These results agree with our findings showing that CT is globally associated with early-stage psychiatric disease phenotypes but predictive of these illnesses from an individualized transdiagnostic perspective.

In order to test whether CTQ profiles also allow for diagnosis-specific prediction of early mental health disorders, we applied the same machine learning model to separate CHR, ROP, and ROD. In these analyses, we found that it was not possible to distinguish reliably between the three diagnostic groups based on trauma exposure patterns. This is in line with studies describing increased rates of CT in psychiatric patients, irrespective of the exact diagnosis (Kessler et al., 2010; Palmier-Claus et al., 2016; Sahin et al., 2013; Scott et al., 2012; Varese et al., 2012). However, other studies exist describing distinct forms of early adversity in specific patient groups. Particularly, Bruni et al. found escape from home, cannabis abuse, psychological abuse, physical abuse, and loneliness to be more frequent in patients with schizophrenic spectrum disorder than in patients with

Table 3. CTQ-class probabilities associations with psychopathology

	r_s	p
All groups		
GAF symptoms	0.388	<0.001[a]
GAF disability/impairment	0.412	<0.001[a]
CHR		
SIPS-P	−0.103	0.259
SIPS-N	−0.116	0.206
SIPS-D	−0.014	0.882
SIPS-G	0.026	0.777
BDI-II	−0.175	0.028[b]
PANSS total	−0.191	0.038[b]
PANSS positive	−0.127	0.168
PANSS negative	−0.196	0.033[b]
PANSS general	−0.126	0.174
ROD		
SIPS positive	0.018	0.839
SIPS negative	0.010	0.908
SIPS disorganizing	−0.124	0.166
SIPS general	−0.106	0.236
BDI-II	−0.278	0.001[a]
PANSS total	−0.017	0.854
PANSS positive	−0.120	0.182
PANSS negative	−0.025	0.782
PANSS general	−0.008	0.928
ROP		
SIPS positive	−0.042	0.638
SIPS negative	−0.230	0.10
SIPS disorganizing	−0.139	0.125
SIPS general	−0.050	0.581
BDI-II	−0.246	0.003[a]
PANSS-T	−0.107	0.117
PANSS-P	−0.097	0.141
PANSS-N	−0.115	0.100
PANSS-G	−0.083	0.177

CHR, clinical high-risk state; ROD, recent onset depression; ROP, recent onset psychosis; PANSS, Positive and Negative Syndrom Scale; BDI-II, Beck Depression Inventory II; SIPS, Structured Interview for Prodromal Symptoms; GAF, Global Assessment of Functioning; r_s, Spearman's correlation coefficient.
[a]Significant at the level of <0.01.
[b]Significant at the level of 0.05.

major depression or bipolar disorder (Bruni et al., 2018). Contrary to these results, our findings suggest that CT exposure is not associated with specific disorders but instead poses a rather general and transdiagnostic risk factor for early psychiatric disorders, which is also in line with an earlier study of our group (Popovic et al., 2020).

Regarding the individual CT patterns, we performed a qualitative comparison of the three CTQ questions which were assigned

the highest weights. We identified the subdomains EN and EA playing the most important role across all groups. On the single item level, especially items that reflect the family climate showed the highest predictive power. These results are in line with a recent structure equation model analysis of Salokangas et al. (2019), which indicated that subdomains EN and PA had the strongest association with depression and psychosis. Furthermore, in our analysis, EN and EA were most predictive in CHR patients, while PN was additionally predictive in psychosis. In contrast, an earlier work by Trauelsen et al. (2015) showed beside EA and EN, PA to be significantly associated with psychotic disorders. Other works revealed specific associations between SA and psychosis (Bentall, Wickham, Shevlin, & Varese, 2012) and hallucinations (Upthegrove et al., 2015). Interestingly, in ROD patients, besides EA, SA and PA were particularly predictive for a later depressive illness. In line with these observations, a meta-analysis of Lindert et al. (2014) pointed out that especially SA and PA are strongly associated with later depression and anxiety disorders. Although this meta-analysis identified SA and PA as the most important risk factors of depression and anxiety disorder, which are also common in CHR patients (Albert, Tomassi, Maina, & Tosato, 2018), we found EN and EA to play the most important role across all groups. One reason for this discrepancy might be the lower frequency of SA and PA compared to other CT domains in our sample that might have led to an underestimation of their role in our cohort. Thus, our results provide more comprehensive evidence for a differentiated neurobiological imprint of the CT in different psychiatric disorders, while at the same time highlighting emotional trauma as particularly relevant to a person's clinical phenotype.

Furthermore, we found evidence that the participants' CTQ-based DS was significantly associated with the current severity of depressive symptoms but not with psychotic symptoms (positive, negative, and general) in the ROD, ROP, and CHR groups. Moreover, pre-psychotic symptoms measured by the SIPS were not correlated with the DS in the CHR group but a weak relationship was detected between the CTQ-based DS and the PANSS total and negative domain scores. These results support the hypothesis that CT constitutes a dimension of vulnerability that is dependent on the current depressive state of the patients. This observation is in keeping with previous work of our group showing that an emotional trauma signature was significantly correlated with higher depression scores, lower levels of functioning, decreased quality of life, and maladaptive personality traits (Popovic et al., 2020). In the past, depressiveness has also been shown to be a mediating factor in the effect of CT on alcohol consumption (Salokangas, From, Luutonen, Salokangas, & Hietala, 2018) and suicidal thoughts (Salokangas et al., 2019).

No associations were found between the CTQ-based DS and brain structure. It can be assumed that the changes at the single item level of the CTQ are too subtle for individual prediction of disease. In a recent publication from our group, we performed a data-driven analysis of brain structure and phenotypic data including CT exposure and found three latent signatures specifically associated with CT. In this previous paper, and latent representations of brain–phenotype associations, SA was associated with aberrant volumes in the prefrontal cortex, the hippocampus, and occipital lobe. EA and EN were associated with volumetric alterations in the occipital lobe and postcentral regions associated with sensory processing. No associations between specific diagnostic groups and CT exposure were found, which is in line with the absence of diagnosis-specific associations between CT

and early mental health diseases, and in keeping with the current analysis (Popovic et al., 2020). In another previous mediation analysis of our group, PA was shown to be associated in particular with reduced volumes of the gray and white matter of the frontal lobe and amygdala-hippocampal complex in ROD and CHR patients (Salokangas et al., 2021). In addition, it was shown that the effect of PA on social anxiety in CHR patients was mediated by a reduced volume of gray matter in the frontal lobe. Since this was methodologically a mediation analysis and not a machine learning approach, these results should not be regarded as contradictory.

Limitations

Limitations of our study include the observational, retrospective, and cross-sectional character of the study. As with most CT assessments, the CTQ assesses trauma retrospectively, thus, running the risk of a 'recall bias' depending on the individual's current mental health situation, including the influence of depression severity (Colman et al., 2016). Another possible limitation is the non-assessment of factors such as the age at onset, the frequency, and the extent of the suffering associated with exposure to CT. It must be critically taken into account that despite diverse adverse experiences, many victims of CT show no or only minor long-term psychological impairment, suggesting that resilience factors appear to be important mediating variables as well (Lee, Yu, & Kim, 2020). Therefore, in the future, suitable methods and longitudinal population data utilizing methods such as structure equation models could be used to investigate the exact relationship between CT and functional or school outcome, against the background of the above-mentioned mediating variables.

Conclusions

In summary, our work has demonstrated that CT constitutes a discriminative transdiagnostic fingerprint of at-risk mental states and early-stage mental disorders. Focusing on the most predictive items of our analyses, we were able to show that a violence-free, supportive family environment as well as protection are important aspects for good mental health in later life. Our findings support the conclusions of a paper by Hudziak (2009) who called for a routine evaluation of CT history in persons presenting to mental health services in order to identify those who may need more intensive support and additional treatment. In line with that, Marshall, Shannon, Meenagh, Mc Corry, and Mulholland (2018) emphasized the importance of special preventive measures, such as therapeutic intervention aimed at sufferers of past abuse, neglect, and poor parenting to prevent 'trans-generational patterns' continuing with their own children. In the future, further analyses of the longitudinally administered PRONIA sample should investigate whether there are differences in the course of the diseases related to CT experiences. Furthermore, suitable methods, such as structural equation models, should be used to highlight the exact relationship between CT and mental illness against the background of mediating variables and resilience factors.

The PRONIA consortium members listed here performed the screening, recruitment, rating, examination, and follow-up of the study participants and were involved in implementing the examination protocols of the study, setting up its information technological infrastructure, and organizing the flow and quality control of the data analyzed in this article between the local study sites and the central study database.

Department of Psychiatry and Psychotherapy, Ludwig-Maximilian-University, Munich, Bavaria, Germany: Mark Sen Dong, M.Sc., Anne Erkens, Eva Gussmann, M.Sc., Shalaila Haas, Ph.D., Alkomiet Hasan, M.D., Claudius Hoff, M.D., Ifrah Khanyaree, B.Sc., Aylin Melo, M.Sc., Susanna Muckenhuber-Sternbauer, M.D., Janis Kohler, Ömer Faruk Öztürk, M.D., Adrian Rangnick, B.Sc., Sebastian von Saldern, M.D., Rachele Sanfelici, M.Sc., Moritz Spangemacher, Ana Tupac, M.Sc., Maria Fernanda Urquijo, M.Sc., Johanna Weiske, M.Sc., and Antonia Wosgien.

University of Cologne, North Rhineland–Westphalia, Germany: Karsten Blume, Tanja Pilgram, M.Sc., and Christiane Woopen, M.D.

Psychiatric University Hospital, University of Basel, Basel, Switzerland: Christina Andreou, M.D., Ph.D., Laura Egloff, Ph.D., Fabienne Harrisberger, Ph.D., Claudia Lenz, Ph.D., Letizia Leanza, M.Sc., Amatya Mackintosh, M.Sc., Renata Smieskova, Ph.D., Erich Studerus, Ph.D., Anna Walter, M.D., and Sonja Widmayer, M.Sc.

Institute for Mental Health, University of Birmingham, Birmingham, United Kingdom: Chris Day, B.Sc., Sian Lowri Griffiths, Ph.D., Mariam Iqbal, B.Sc., Mirabel Pelton, M.Sc., Pavan Mallikarjun, M.B.B.S., D.P.M., MRCPsych, Ph.D., Ashleigh Lin, Ph.D., Paris A Lalousis MSc, Alexandra Stainton PhD

Department of Psychiatry, University of Turku, Turku, Finland: Alexander Denissoff, M.D., Anu Ellila, RN, Tiina From, M.Sc., Markus Heinimaa, M.D., Ph.D., Tuula Ilonen, Ph.D., Paivi Jalo, RN, Heikki Laurikainen, M.D., Maarit Lehtinen, RN, Antti Luutonen, B.A., Akseli Makela, B.A., Janina Paju, M.Sc., Henri Pesonen, Ph.D., Reetta-Liina Armio (Saila), M.D., Elina Sormunen, M.D., Anna Toivonen, M.Sc., and Otto Turtonen, M.D.

General Electric Global Research Inc, Munich, Germany: Ana Beatriz Solana, Ph.D., Manuela Abraham, M.B.A., Nicolas Hehn, Ph.D., and Timo Schirmer, Ph.D.

Workgroup of Paolo Brambilla, M.D., Ph.D., University of Milan, Milan, Italy: Department of Neuroscience and Mental Health, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, University of Milan, Milan, Italy: Carlo Altamura, M.D., Marika Belleri, PsychD, Francesca Bottinelli, PsychD, Adele Ferro, PsychD, Ph.D., and Marta Re, Ph.D. Programma 2000, Niguarda Hospital, Milan: Emiliano Monzani, M.D., Mauro Percudani, M.D., and Maurizio Sberna, M.D. San Paolo Hospital, Milan: Armando D'Agostino, M.D., and Lorenzo Del Fabro, M.D. Villa San Benedetto Menni, Albese con Cassano: Giampaolo Perna, M.D., Maria Nobile M.D., Ph.D., and Alessandra Alciati, M.D.

Workgroup of Paolo Brambilla, M.D., Ph.D., University of Udine, Udine, Italy: Department of Medical Area, University of Udine: Matteo Balestrieri, M.D., Carolina Bonivento, PsychD, Ph.D., Giuseppe Cabras, Ph.D., and Franco Fabbro, M.D., Ph.D. IRCCS Scientific Institute 'E. Medea', Polo FVG, Udine: Marco Garzitto, PsychD, Ph.D. and Sara Piccin, PsychD, Ph.D. **Workgroup of Professor Alessandro Bertolino, University of Bari Aldo Moro, Italy:** Professor Giuseppe Blasi; Professor Linda A. Antonucci, Professor Giulio Pergola, Grazia Caforio, Ph.D., Leonardo Faio, Ph.D., Tiziana Quarto, Ph.D., Barbara Gelao, Ph.D., Raffaella Romano, Ph.D., Ileana Andriola, M.D., Andrea Falsetti, M.D., Marina Barone, M.D., Roberta Passatiore, M.Sc., Marina Sangiuliano, M.D.

Department of Psychiatry and Psychotherapy, Westfaelische Wilhelms-University Muenster, North Rhineland–Westphalia, Germany: Marian Surman, M.Sc., Olga Bienek, M.D., Georg Romer, M.D., Udo Dannlowski, M.D., Ph.D.

Department of Psychiatry and Psychotherapy of the University Düsseldorf, North Rhineland–Westphalia, Germany: Christian Schmidt-Kraepelin, M.D., Susanne Neufang, Ph.D., Alexandra Korda, Ph.D., Henrik Rohner, M.D.

References

Albert, U., Tomassi, S., Maina, G., & Tosato, S. (2018). Prevalence of non-psychotic disorders in ultra-high risk individuals and transition to psychosis: A systematic review. *Psychiatry Research*, 270, 1–12. doi: 10.1016/j.psychres.2018.09.028.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV* (4th ed., text rev.). Washington, DC.

Bentall, R. P., Wickham, S., Shevlin, M., & Varese, F. (2012). Do specific early-life adversities lead to specific symptoms of psychosis? A study from the 2007 the Adult Psychiatric Morbidity Survey. *Schizophrenia Bulletin*, 38(4), 734–740. doi: 10.1093/schbul/sbs049.

Bernstein, D., & Fink, L. (1998). *Childhood Trauma Questionnaire: A retrospective self-report*. San Antonio: The Psychological Corporation.

Bruni, A., Carbone, E. A., Pugliese, V., Aloi, M., Calabro, G., Cerminara, G., … De Fazio, P. (2018). Childhood adversities are different in schizophrenic spectrum disorders, bipolar disorder and major depressive disorder. *BMC Psychiatry*, 18(1), 391. doi: 10.1186/s12888-018-1972-8.

Clinical Guidelines on Drug Misuse and Dependence Update 2017 Independent Expert Working Group (2017). *Drug misuse and dependence: UK guidelines on clinical management*. London.

Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. (2015). Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD). *Annals of Internal Medicine*, 162(10), 735–736. doi: 10.7326/L15-5093-2.

Colman, I., Kingsbury, M., Garad, Y., Zeng, Y., Naicker, K., Patten, S., … Thompson, A. H. (2016). Consistency in adult reporting of adverse childhood experiences. *Psychological Medicine*, 46(3), 543–549. doi: 10.1017/S0033291715002032.

Fan, R., Chang, K., Hsieh, C., Wang, X., & Lin, C. (2008). LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research*, 9(9), 1871–1874. doi: 10.1145/1390681.1442794.

Fang, X., Brown, D. S., Florence, C. S., & Mercy, J. A. (2012). The economic burden of child maltreatment in the United States and implications for prevention. *Child Abuse & Neglect*, 36(2), 156–165. doi: doi:10.1016/j.chiabu.2011.10.006.

Gaser, C., & Dahnke, R. (2016). CAT – A computational anatomy toolbox for the analysis of structural MRI data. Retrieved from http://www.neuro.uni-jena.de/hbm2016/GaserHBM2016.pdf.

Golland, P., & Fischl, B. (2003). Permutation tests for classification: Towards statistical significance in image-based studies. *Information Processing in Medical Imaging*, 18, 330–341. doi: 10.1007/978-3-540-45087-0_28.

Hautzinger, M., Bailer, M., Worall, H., & Keller, F. (1995). *Beck-Depressions-Inventor* (BDI). *Testhandbuch. 2. Auflage*. Bern: Hans Huber.

Hayasaka, Y., Purgato, M., Magni, L. R., Ogawa, Y., Takeshima, N., Cipriani, A., … Furukawa, T. A. (2015). Dose equivalents of antidepressants: Evidence-based recommendations from randomized controlled trials. *Journal of Affective Disorders*, 180, 179–184. doi: 10.1016/j.jad.2015.03.021.

Hudziak, J. (2009). *Developmental Psychopathology and Wellness: Genetic and Environmental Influences* (1st ed., pp. 267–287). Washington, DC: American Psychiatric Publishing.

Kambeitz, J., Kambeitz-Ilankovic, L., Leucht, S., Wood, S., Davatzikos, C., Malchow, B., … Koutsouleris, N. (2015). Detecting neuroimaging biomarkers for schizophrenia: A meta-analysis of multivariate pattern recognition studies. *Neuropsychopharmacology*, 40(7), 1742–1751. doi: 10.1038/npp.2015.22.

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, 13(2), 261–276.

Kessler, R. C., McLaughlin, K. A., Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., … Williams, D. R. (2010). Childhood adversities and adult psychopathology in the WHO World Mental Health Surveys. *British Journal of Psychiatry*, 197(5), 378–385. doi: 10.1192/bjp.bp.110.080499.

Koutsouleris, N., Kambeitz-Ilankovic, L., Ruhrmann, S., Rosen, M., Ruef, A., Dwyer, D. B., … Consortium, P. (2018). Prediction models of functional outcomes for individuals in the clinical high-risk state for psychosis or with recent-onset depression: A multimodal, multisite machine learning analysis. *JAMA Psychiatry*, 75(11), 1156–1172. doi: 10.1001/jamapsychiatry.2018.2165.

Lee, D., Yu, E. S., & Kim, N. H. (2020). Resilience as a mediator in the relationship between posttraumatic stress and posttraumatic growth among adult accident or crime victims: The moderated mediating effect of childhood trauma. *European Journal of Psychotraumatology*, 11(1), 1704563. doi: 10.1080/20008198.2019.1704563.

Leucht, S., Samara, M., Heres, S., & Davis, J. M. (2016). Dose equivalents for antipsychotic drugs: The DDD method. *Schizophrenia Bulletin*, 42(Suppl 1), S90–S94. doi: 10.1093/schbul/sbv167.

Lindert, J., von Ehrenstein, O. S., Grashow, R., Gal, G., Braehler, E., & Weisskopf, M. G. (2014). Sexual and physical abuse in childhood is associated with depression and anxiety over the life course: Systematic review and meta-analysis. *International Journal of Public Health*, 59(2), 359–372. doi: 10.1007/s00038-013-0519-5.

Marshall, M., Shannon, C., Meenagh, C., Mc Corry, N., & Mulholland, C. (2018). The association between childhood trauma, parental bonding and depressive symptoms and interpersonal functioning in depression and bipolar disorder. *Irish Journal of Psychological Medicine*, 35(1), 23–32. doi: 10.1017/ipm.2016.43.

McGlashan, T., Walsh, B., & Woods, S. (2010). *The psychosis-risk syndrome: Handbook for diagnosis and follow-up*. Oxford: Oxford University Press.

Mikton, C., & Butchart, A. (2009). Child maltreatment prevention: A systematic review of reviews. *Bulletin of the World Health Organization*, 87(5), 353–361. doi: 10.2471/blt.08.057075.

Nelson, J., Klumparendt, A., Doebler, P., & Ehring, T. (2017). Childhood maltreatment and characteristics of adult depression: Meta-analysis. *British Journal of Psychiatry*, 210(2), 96–104. doi: 10.1192/bjp.bp.115.180752.

Palmier-Claus, J. E., Berry, K., Bucci, S., Mansell, W., & Varese, F. (2016). Relationship between childhood adversity and bipolar affective disorder: Systematic review and meta-analysis. *British Journal of Psychiatry*, 209(6), 454–459. doi: 10.1192/bjp.bp.115.179655.

Phillips, L. J., Yung, A. R., & McGorry, P. D. (2000). Identification of young people at risk of psychosis: Validation of personal assessment and crisis evaluation clinic intake criteria. *Australian and New Zealand Journal of Psychiatry*, 34(Suppl), S164–S169. doi: 10.1080/000486700239.

Popovic, D., Ruef, A., Dwyer, D., Antonucci, L., Eder, J., Sanfelici, R., … Koutsouleris, N. (2020). Traces of trauma – a multivariate pattern analysis of childhood trauma, brain structure and clinical phenotypes. *Biological Psychiatry*, 88(11), 829–842. doi: 10.1016/j.biopsych.2020.05.020

Rubino, I. A., Nanni, R. C., Pozzi, D. M., & Siracusano, A. (2009). Early adverse experiences in schizophrenia and unipolar depression. *Journal of Nervous and Mental Disease*, 197(1), 65–68. doi: 10.1097/NMD.0b013e3181925342.

Sahin, S., Yuksel, C., Guler, J., Karadayi, G., Akturan, E., Gode, E., … Ucok, A. (2013). The history of childhood trauma among individuals with ultra high risk for psychosis is as common as among patients with first-episode schizophrenia. *Early Intervention in Psychiatry*, 7(4), 414–420. doi: 10.1111/eip.12022.

Salokangas, R. K. R., From, T., Luutonen, S., & Hietala, J. (2018). Adverse childhood experiences leads to perceived negative attitude of others and the effect of adverse childhood experiences on depression in adulthood is mediated via negative attitude of others. *European Psychiatry*, 54, 27–34. doi: 10.1016/j.eurpsy.2018.06.011.

Salokangas, R. K. R., From, T., Luutonen, S., Salokangas, H. R. W., & Hietala, J. (2018). Effect of childhood adversities on alcohol problems is mainly mediated by depression. *The American Journal on Addictions*, 27, 391–399. 10.1111/ajad.12734..

Salokangas, R. K. R., Hietala, J., Armio, R. L., Laurikainen, H., From, T., Borgwardt, S., … Koutsouleris, N. (2021). Effect of childhood physical abuse on social anxiety is mediated via reduced frontal lobe and amygdala-hippocampus complex volume in adult clinical high-risk subjects. *Schizophrenia Research*, 227, 101–109. doi: 10.1016/j.schres.2020.05.041.

Salokangas, R. K. R., Patterson, P., Hietala, J., Heinimaa, M., From, T., Ilonen, T., … Ruhrmann, S. (2019). Childhood adversity predicts persistence of suicidal thoughts differently in females and males at clinical high-risk patients of psychosis. Results of the EPOS project. *Early Intervention Psychiatry*, 13 (4), 935–942. doi: 10.1111/eip.12714.

Scher, C. D., Forde, D. R., McQuaid, J. R., & Stein, M. B. (2004). Prevalence and demographic correlates of childhood maltreatment in an adult community sample. *Child Abuse & Neglect*, 28(2), 167–180. doi: 10.1016/j.chiabu.2003.09.012.

Schultze-Lutter, F., Addington, J., Ruhrmann, S., & Klosterkötter, J. (2007). *Schizophrenia Proneness Instrument, Adult Version (SPI-A)*. Rom: Giovanni Fioriti Editore.

Scott, K. M., McLaughlin, K. A., Smith, D. A., & Ellis, P. M. (2012). Childhood maltreatment and DSM-IV adult mental disorders: Comparison of prospective and retrospective findings. *British Journal of Psychiatry*, 200 (6), 469–475. doi: 10.1192/bjp.bp.111.103267.

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage*, 44(1), 83–98. doi: 10.1016/j.neuroimage.2008.03.061.

Steyerberg, E. W., & Harrell, F. E. Jr. (2016). Prediction models need appropriate internal, internal-external, and external validation. *Journal of Clinical Epidemiology*, 69, 245–247. 10.1016/j.jclinepi.2015.04.005.

Trauelsen, A. M., Bendall, S., Jansen, J. E., Nielsen, H. G., Pedersen, M. B., Trier, C. H., … Simonsen, E. (2015). Childhood adversity specificity and dose-response effect in non-affective first-episode psychosis. *Schizophrenia Research*, 165(1), 52–59. doi: 10.1016/j.schres.2015.03.014.

Upthegrove, R., Chard, C., Jones, L., Gordon-Smith, K., Forty, L., Jones, I., & Craddock, N. (2015). Adverse childhood events and psychosis in bipolar affective disorder. *British Journal of Psychiatry*, 206(3), 191–197. doi: 10.1192/bjp.bp.114.152611.

Varese, F., Smeets, F., Drukker, M., Lieverse, R., Lataster, T., Viechtbauer, W., … Bentall, R. P. (2012). Childhood adversities increase the risk of psychosis: A meta-analysis of patient-control, prospective- and cross-sectional cohort studies. *Schizophrenia Bulletin*, 38(4), 661–671. doi: 10.1093/schbul/sbs050.

Supplement Table 1: Study inclusion / exclusion criteria of the study.

Group Inclusion Criteria	Group Exclusion Criteria	General Inclusion / Exclusion / Drop-out
Clinical High-Risk Group (CHR)		**Inclusion Criteria:**
Psychosis-risk syndrome defined: **EITHER** by *Attenuated Positive Symptoms (APS)*, as measured by the SIPS (requires 1 of 5 attenuated psychotic symptoms: unusual thought content/ delusional ideas, suspicious-ness/persecutory ideas, grandiosity, perceptual abnormalities/hallucinations, and disorganized communication) with a moderate to severe, but not psychotic, severity (SIPS score 3-5) that (1) began with-in the past year or was rated one or more scale points higher compared to 12 month ago, **AND** (2) occurred at an average frequency of at least once per week for at least several minutes per event in the past month **OR:** by *Brief Intermittent Psychotic Symptoms (BLIPS)*, as measured by the SIPS (as defined by one of the symptoms listed above (1) reaching a psychotic level of intensity in each of the past 3 months for at least several minutes per day, **OR** (2) reaching a psychotic level of intensity in the past month, occurring at an average frequency of at least once per week for at least several minutes per event in the past month, or occurring at least for a cumulative period of more than one hour within the past month, **AND** (1+2) remitting spontaneously within one week (i.e. without antipsychotic medication) **OR:** by a *Genetic Risk and Functional Decline Psychosis-Risk Syndrome (GRFD)* defined by a current 30% or greater reduction in the functional disability score of the split version of the Global Assessment of Functioning Scale (GAF-F) compared with the highest lifetime level of functioning, **AND** (having a first-degree relative with a history of any psychotic disorder, **OR** having a DSM-IV-TR schizotypal personality disorder). **OR:** by a *Cognitive Disturbance Syndrome* (COGDIS) as measured by the SPI-A (requires at least 2 of 9 cognitive basic symptoms with at least weekly occurrence (score ≥3) during the last 3 months)	1. Any intake of antipsychotic medication for more than 30 cumulative days at or above the minimum dosage threshold defined by the DGPPN S3 Guidelines for the treatment of first-episode psychosis[1] 2. Any intake of antipsychotic drugs within the past 3 months before psychopathological baseline assessments at or above the minimum dosage threshold. 3. Occurrence of the CHR syndrome is better explained by other DSM-IV disorder	1. Age 15 to 40 years 2. Language skills sufficient for participation 3. Able to provide to consent / assent **Exclusion Criteria:** 1. IQ below 70 2. Hearing is not sufficient for neuro-cognitive testing 3. Current or past head trauma with loss of consciousness (> 5 min) 4. Current or past known neurological disorder of the brain 5. Current or past known somatic disorder potentially affecting the structure or functioning of the brain 6. Current or past alcohol dependence 7. Current poly-substance dependence or within the past six months (Note: any combination with E.6. led to exclusion) 8. Any contra-indication for MRI **Exclusion criteria for healthy controls:** 1. Any current or past DSM-IV axis disorder 2. A positive familial history (1st degree relatives) for affective or non-affective psychoses or major affective disorders; and 3. An intake of psychotropic medications or drugs more than 5 times/year and in the month before study inclusion. **Drop-out criteria:** 1. No follow-up examination after the 6-months follow-up examination (IV6) 2. Withdrawn consent / assent
Recent-Onset Depression (ROD)		
Recent-onset Depression as defined by DSM-IV-TR + ALL of the following criteria: 1. First life-time depressive episode, 2. Duration of current depressive episode no longer than 24 months, 3. Diagnostic criteria fulfilled within past three months	1. Occurrence of the major depressive episode is better explained by other DSM-IV disorder 2. See CHR exclusion criteria	
Recent-Onset Psychosis (ROP)		
Recent-onset Psychosis as defined by DSM-IV-TR (affective and non-affective) + ALL of the following criteria: 1. First life-time psychotic episode, 2. Duration of current psychotic episode no longer than 24 months, **3.** Diagnostic criteria fulfilled within past three months	1. Occurrence of the psychotic episode is better explained by other DSM-IV disorder 2. Antipsychotic medication for more than 90 days at or above the minimum dosage defined by the DGPPN S3 Guidelines for the treatment of first-episode psychosis[1]	

[1] Deutsche Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde. DGPPN S3 Treatment Guideline Schizophrenia / Psychotic Disorders. *AWMF* 2006.

Supplement Table 2: MR scanner systems and structural MRI sequence parameters used at the respective PRONIA sites

PRONIA Site	Model	Field strength [3T]	Coil channels	Flip angle [deg]	TR [ms]	TE [ms]	Voxel size [mm]	FOV	Slice number
Munich	Philips Ingenia	3	32	8	9.5	5.5	0.97x0.97x1.0	250 x 250	190
Milan Niguarda	Philips Achieva Intera	1.5	8	12	Shortest (8.1)	Shortest (3.7)	0.93x0.93x1.0	240 x 240	170
Basel	Siemens Verio	3	12	8	2000	3.4	1.0x1.0x1.0	256 x 256	176
Cologne	Philips Achieva	3	8	8	9.5	5.5	0.97x0.97x1.0	250 x 250	190
Birmingham	Philips Achieva	3	32	8	8.4	3.8	1.0x1.0x1.0	288 x 288	175
Turku	Philips Ingenuity	3	32	7	8.1	3.7	1.0x1.0x1.0	256 x 256	176
Udine	Philips Achieva	3	8	12	Shortest (8.1)	Shortest (3.7)	0.93x0.93x1.0	240 x 240	170

Supplement Table 3: Comparison of CTQ subscales across groups

CTQ Subscale			Average difference	Standard error	p
Emotional Abuse					
	HC	CHR	-3.60*	0.41	<.001
		ROD	-2.6*	0.4	<.001
		ROP	-3.06*	0.4	<.001
	CHR	HC	3.60*	0.41	<.001
		ROD	0.99	0.47	0.15
		ROP	0.53	0.47	0.67
	ROD	HC	2.6*	0.4	<.001
		CHR	-0.99	0.47	0.15
		ROP	-0.46	0.46	0.76
	ROP	HC	3.06*	0.4	<.001
		CHR	-0.53	0.47	0.67
		ROD	0.46	0.46	0.76
Physical Abuse					
	HC	CHR	-1.17*	0.27	<.001
		ROD	-1.05*	0.27	<.001
		ROP	-1.17*	0.27	<.001
	CHR	HC	1.17*	0.27	<.001
		ROD	0.11	0.31	0.98
		ROP	-0.004	0.31	1.00
	ROD	HC	1.05*	0.27	<.001
		CHR	-0.11	0.31	0.98
		ROP	-0.12	0.31	0.98
	ROP	HC	1.17*	0.27	<.001
		CHR	0.004	0.31	1.00
		ROD	0.12	0.31	0.98
Sexual Abuse					
	HC	CHR	-0.77*	0.26	0.02
		ROD	-0.69*	0.26	0.04
		ROP	-1.09*	0.26	<.001
	CHR	HC	0.77*	0.26	0.02
		ROD	0.08	0.3	0.99
		ROP	-0.32	0.3	0.72
	ROD	HC	0.69*	0.26	0.04
		CHR	-0.08	0.3	0.99
		ROP	-0.4	0.3	0.53
	ROP	HC	1.09*	0.26	<.001
		CHR	0.32	0.3	0.72
		ROD	0.4	0.3	0.53
Emotional Neglect					
	HC	CHR	-3.85*	0.45	<.001
		ROD	-3.31*	0.44	<.001
		ROP	-3.47*	0.44	<.001
	CHR	HC	3.85*	0.45	<.001
		ROD	0.54	0.52	0.73
		ROP	0.39	0.52	0.88
	ROD	HC	3.31*	0.44	<.001
		CHR	-0.54	0.52	0.73
		ROP	-0.15	0.51	0.99
	ROP	HC	3.47*	0.44	<.001
		CHR	-0.38	0.52	0.88
		ROD	0.15	0.51	0.99

Physical Neglect

	CHR	-1.48*	0.26	<.001
HC	ROD	-1.21*	0.25	<.001
	ROP	-1.92*	0.25	<.001
	HC	1.48*	0.26	<.001
CHR	ROD	0.27	0.29	0.79
	ROP	-0.43	0.29	0.45
	HC	1.21*	0.25	<.001
ROD	CHR	-0.27	0.29	0.79
	ROP	-0.71	0.29	0.07
	HC	1.92*	0.25	<.001
ROP	CHR	0.43	0.29	0.45
	ROD	0.71	0.29	0.07

Abbreviations: HC, Healthy Controlls; CHR, Clinical High-Risk state; ROD, Recent Onset Depression; ROP, Recent Onset Psychosis; CTQ, childhood trauma questionnaire
*Significant at the level of 0.05

Supplement Table 4: *Mean values of medication dose equivalents taken cumulatively over lifetime (mg/d)*

		CHLORPROMAZINE	OLANZAPINE	SSRI	BENZODIAZEPINE
ALL GROUPS	M	1919.98	68.16	1617.69	121.52
	N	623	623	623	623
	SD	8906.37	307.86	6417.95	651.69
HC	M	0	0	0	16.10
	N	250	250	250	250
	SD	0	0	0	232.24
ROD	M	504.47	19.05	2630.83	142.59
	N	129	129	129	129
	SD	2033.96	78.42	6682.12	556.91
CHR	M	1025.06	42.04	3864.95	123.49
	N	119	119	119	119
	SD	5463.24	235.20	8371.53	649.41
ROP	M	8072.71	280.02	1668.10	308.74
	N	125	125	125	125
	SD	17798.21	598.88	9090.62	1114.60

M, mean; SD, standard deviation; HC, Healthy Controlls; CHR, Clinical High-Risk state; ROD, Recent Onset Depression; ROP, Recent Onset Psychosis; Chlorpromazine; Chlorpromazine equivalent, Olanzapine; Olanzapine equivalent, SSRI; SSRI equivalent, Benzodiazepine; Benzodiazepine (Diazepam) equivalent

Supplement Table S5: *CTQ-class probabilities associations with age*

	r_s	p
All groups	-.018	.648
HC	.029	.640
CHR	-.157	0.83
ROD	-0.034	.703
ROP	-.002	.986

CHR, Clinical High-Risk state; ROD, Recent Onset Depression; ROP, Recent Onset Psychosis; r_s, Spearmans Correlation Coefficient

Supplement Table S6: *CTQ-class probabilities associations with medication dose equivalents taken cumulatively over lifetime*

	r_s	p
All groups		
Chlorpromazine	-.213	<0.001
Olanzapine	-.213	<0.001
SSRI	-.193	<0.001
Benzodiazepine	-.128	0.001
HC		
Chlorpromazine	n.a.	n.a.
Olanzapine	n.a.	n.a.
SSRI	n.a.	n.a.
Benzodiazepine	.141	0.013
CHR		
Chlorpromazine	-0.045	0.314
Olanzapine	-0.043	0.319
SSRI	0.015	0.437
Benzodiazepine	-0.038	0.341
ROD		
Chlorpromazine	0.058	0.255
Olanzapine	0.059	0.254
SSRI	0.047	0.298
Benzodiazepine	0.058	0.258
ROP		
Chlorpromazine	-0.002	.491
Olanzapine	-0.006	.472
SSRI	0.004	.481
Benzodiazepine	0.029	.374

CHR, Clinical High-Risk state; ROD, Recent Onset Depression; ROP, Recent Onset Psychosis; r_s, Spearmans Correlation Coefficient, Chlorpromazine; Chlorpromazine equivalent, Olanzapine; Olanzapine equivalent, SSRI; SSRI equivalent, Benzodiazepine; Benzodiazepine (Diazepam) equivalent

Supplement Method: Nested leave-site-out cross-validation

On the outer LOSOCV cycle (CV_2), the entire population was split into the seven sites. Each of these samples was iteratively held back as validation data, while the six remaining samples entered the inner CV loop. Hence, this outer CV loop provided a robust and unbiased estimate of the classification generalizability because all validation samples were strictly separated from the entire training process taking place at the inner loop (CV_1). A 10-fold CV with 10 repetitions was used at this inner loop, to generate classifier ensembles and the outer loop was repeated 5 times to further increase robustness of the generalizability assessments.

Specifically, in each of these training partitions, the CTQ-items were scaled feature-wise to a range of [0, 1]. Because of missing values (3.8% missing), we used a nearest neighbor-based imputation approach employing the Hamming distance[1] suitable for ordinal data. Then, the scaled and imputed data matrix was z-normalized to the training sample's means and standard deviations before it entered sequential backward elimination (SBE) algorithm that employed L_2-regularized logistic regression (L2-LR)[2] as provided by the LIBLINEAR library[2] in NeuroMiner. The SBE algorithm iteratively removed CTQ items from the item pool that decreased average model performance in the CV_1 training and CV_1 test data. An early stopping criterion at 50% of the variables remaining in the pool was introduced to avoid an overfitting of the algorithm. To further increase feature extraction stability, a probabilistic feature extraction step identified those CTQ items that were selected by at least 90% of the CV_1 models in the given CV_2 training partition. CTQ items not meeting this criterion were pruned from the feature pool and models were retrained with the remaining features using the entire CV_1 data partition.

To predict the group membership of unknown individuals in the CV_2 validation partitions, the scaling, imputation and z-normalization models developed in the training sample were first applied to these cases, followed by the computation of class membership probabilities by means of the trained L2-LR models. The class membership predictions produced by these CV_1 models for the unseen validation cases in each held-back site were bagged into an classification ensemble by means of averaging and majority voting[3]. Thus, an average CTQ-based class probability / decision score (DS) was calculated for each individual, predicting its out-of-training (OOT) group membership.

References

1. Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, Botstein D, Altman RB. Missing value estimation methods for DNA microarrays. *Bioinformatics* Jun 2001;17(6):520-525.
2. Fan RC, K; Hsieh, C; Wang, X; Lin, C; . LIBLINEAR: A Library for Large Linear Classification. 2008.

3.   Polikar R, Topalis A, Green D, Kounios J, Clark CM. Comparative multiresolution wavelet analysis of ERP spectral bands using an ensemble of classifiers approach for early diagnosis of Alzheimer's disease. *Comput Biol Med* Apr 2007;37(4):542-558.

a

b

c

d

Ψ   Ψ   Ψ   Ψ   Ψ   Ψ   Ψ

0   3   6   9   12   15   18
BL   T1   [months]   T2

Supplement Figure 2