



OPEN

Phenotyping chronic tinnitus patients using self-report questionnaire data: cluster analysis and visual comparison

Uli Niemann¹✉, Petra Brueggemann², Benjamin Boecking², Wilhelm Mebus², Matthias Rose³, Myra Spiliopoulou¹ & Birgit Mazurek²

Chronic tinnitus is a complex, multi-factorial symptom that requires careful assessment and management. Evidence-based therapeutic approaches involve audiological and psychological treatment components. However, not everyone benefits from treatment. The identification and characterisation of patient subgroups (or “phenotypes”) may provide clinically relevant information. Due to the large number of assessment tools, data-driven methods appear to be promising. The acceptance of these empirical results can be further strengthened by a comprehensive visualisation. In this study, we used cluster analysis to identify distinct tinnitus phenotypes based on self-report questionnaire data and implemented a visualisation tool to explore phenotype idiosyncrasies. 1228 patients with chronic tinnitus from the Charité Tinnitus Center in Berlin were included. At baseline, each participant completed 14 questionnaires measuring tinnitus distress, loudness, frequency and location, depressivity, perceived stress, quality of life, physical and mental health, pain perception, somatic symptom expression and coping attitudes. Four distinct patient phenotypes emerged from clustering: *avoidant group* (56.8%), *psychosomatic group* (14.1%), *somatic group* (15.2%), and *distress group* (13.9%). Radial bar- and line charts allowed for visual inspection and juxtaposition of major phenotype characteristics. The phenotypes differed in terms of clinical information including psychological symptoms, quality of life, coping attitudes, stress, tinnitus-related distress and pain, as well as socio-demographics. Our findings suggest that identifiable patient subgroups and their visualisation may allow for stratified treatment strategies and research designs.

Tinnitus, the perception of a phantom sound in absence of an external sound source, is a complex multi-factorially caused and maintained phenomenon. It is estimated to affect 10% and 15% of the adult population¹. The associated annual economic burden amounts to US\$19.4 billion in the United States², and €6.8 billion in the Netherlands alone³. Clinical assessment of tinnitus is challenging due to its various heterogeneities. Tinnitus patients can differ with respect to perception of tinnitus (laterality, pitch, sound characteristics, frequency, permanence, chronicity), risk factors (including hearing loss, temporomandibular joint disorder, aging), comorbidities (including hyperacusis, depression, sleep disorders), perceived distress, and treatment responses⁴. These differences make the identification of a suitable treatment difficult. Currently, there is no consensus on or gold standard for a therapy form that is effective for every patient. Sound therapy (masking) alone is not sufficient to significantly improve tinnitus loudness and severity⁵. Informational counselling (minimal contact education) was found to be effective for subgroups of patients^{6,7}. While some studies affirm the efficacy of cognitive behavioural therapy (CBT) in reducing tinnitus impairment and distress⁸, others conclude that CBT is not superior to other treatments or no treatment in improving subjective tinnitus loudness and quality of life⁹. Similarly, tinnitus retraining (TRT) was found to be helpful in reducing tinnitus impairment and quality of life in some studies^{10,11}, but not others¹². Due to the heterogeneous nature of the tinnitus symptom as well as the unclear evidence-base

¹Faculty of Computer Science, Otto von Guericke University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany. ²Tinnitus Center, Charité Universitaetsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany. ³Division of Psychosomatic Medicine, Medical Department, Charité Universitaetsmedizin Berlin, Hindenburgdamm 30, 12200 Berlin, Germany. ✉email: uli.niemann@ovgu.de

as to its treatment and management, the identification of patient subgroups is vital to stratify individual pathophysiology and treatment pathways^{13–15}.

Since clinically relevant subgroups have not been established yet, clustering emerges as a promising approach to identify distinct tinnitus *phenotypes* in a data-driven, hypothesis-free way. Clustering is the process of grouping subjects into multiple groups or clusters. The goal is that subjects within a group are similar to another and dissimilar to subjects in others groups. Previous studies found subgroups of tinnitus patients with cluster analysis based on a small number of audiometric features¹³, a combination of features extracted from self-reports, audiometry and psychoacoustics¹⁴, a subset of socio-demographics, tinnitus characteristics, self-reports and audiological measurements¹⁶ or neuroimaging data and socio-demographics¹⁷. Although each of these studies provided insights in tinnitus subgroup patterns we believe that to increase acceptance amongst medical practitioners, clustering results need to be presented with intuitive visualisations that show individual subgroup patterns and enable the visual juxtaposition of multiple subgroups with respect to multi-variate data. With this in mind, Schlee et al. proposed a compact radar chart visualisation that allows to juxtapose the degree of health burden between either individuals or subgroups based on multi-variate data¹⁸. While their visualisation could be applied to any disease domain, Schlee et al. demonstrated its efficacy showing subgroup differences with respect to measurements of tinnitus distress and associated comorbidities. However, Schlee et al. did not aim to visualise clustering results, but restricted themselves to pre-defined cohorts such as female vs male patients or patients with low vs high tinnitus frequency.

Hence, the goal of this study is to combine clustering for tinnitus phenotyping with visual cluster representation and juxtaposition. Firstly, we performed cluster analysis and identified distinct phenotypes in patients with chronic tinnitus based on 64 (sub-)scales and sociodemographics extracted from self-report questionnaires. Secondly, we created an intuitive radial chart visualisation to display multi-variate characteristics of a single cluster, and to facilitate comparison of multiple clusters. Prospectively, the results of this study could be used for stratified research designs and treatment approaches.

Materials and methods

Patients. Analyses were based on data from N = 1228 patients with chronic subjective tinnitus who had been treated at the Tinnitus Center of Charité Universitätsmedizin Berlin, Germany, between January 2011 and October 2015. All patients had been suffering from tinnitus for 3 months or longer and were 18 years of age or older. Exclusion criteria comprised the presence of acute psychotic illnesses or addictions, deafness and insufficient knowledge of the German language. Multimodal psychosomatic assessments were carried out by ENT, internal, psychosomatic and physical therapy specialists. Treatment comprised an intensive, multimodal 7-day program that included informational counselling, detailed ear-nose-throat (ENT) as well as medical and psychological diagnostics, cognitive-behaviour therapy interventions, auditory training, relaxation exercises, and physiotherapy. Patients who presented with objective tinnitus were excluded from the present therapy and treated medically, as applicable. Ethical approval was granted by Charité Universitätsmedizin Berlin ethics committee (reference number EA1/115/15). All relevant guidelines and regulations were followed. All patients gave informed written consent for data collection. Prior to the analyses, all data had been anonymised.

Features. All patients completed a routine questionnaire assessment battery. These questionnaires were selected to obtain a comprehensive tinnitus assessment, including tinnitus-related distress and the psychosomatic background of tinnitus with anxiety, depression, general quality of life and experienced physical impairments. For clustering, a total of 64 features from 14 questionnaires was used, comprising 49 compound scores and 15 single-item measurements:

- (a) The *Anamnestic Comparative Self-Assessment (ACSA)*¹⁹ is a visual analogue scale to measure the current quality of life (feature *ACSA_qualityoflife*).
- (b) The *General Depression Scale (Allgemeine Depressionsskala; ADSL)*^{20,21} comprises 20 items for self-assessment of depressive symptoms including insecurity, exhaustion, hopelessness, self-devaluation, dejection, loneliness, sadness, lack of drive, perceived rejection by others, crying, enjoyment, withdrawal, fear, happiness, lack of reactivity, sleep disorders, appetite disorders, concentration problems and pessimism. Each item has 4 response options on an ordinal scale: 0 = “rarely or none of the time (less than 1 day)”, 1 = “some or a little of the time (1–2 days)”, 2 = “occasionally or a moderate amount of time (3–4 days)”, and 3 = “most or all of the time (5–7 days)”. We used the total sum score *ADSL_depression*.
- (c) The *Berlin Complaint Inventory (Berliner Beschwerdeinventar; BI)*²² contains 57 items from the areas general well-being, autonomic nervous system, pain and emotionality. Two exemplary items are “I am bothered by feelings of weakness” and “I am bothered by nausea”. Respondents can answer on a 5-level ordinal scale with 0 = “not at all”, 1 = “hardly”, 2 = “somewhat”, 3 = “considerably” or 4 = “severely”. We used the 4 subscales exhaustion (*BI_fatigue*), abdominal symptoms (*BI_abdominalsymptoms*), limb pain (*BI_limbpain*), heart symptoms (*BI_heartsymptoms*) as well as the overall complaints sum score (*BI_overallcomplaints*).
- (d) The *Berlin Mood Questionnaire (Berliner Stimmungsfragebogen; BSF)*²³ constructs a multi-dimensional mood model from 30 items such as “I feel anxious” and “I feel belligerent” with 5 response options: 0 = “not at all”, 1 = “somewhat”, 2 = “rather”, 3 = “mainly” and 4 = “very much”. We used the following 6 subscales: fatigue (*BSF_fatigue*), apathy (*BSF_apathy*), anxious depression (*BSF_anxdepression*), anger (*BSF_anger*), positive mindset (*BSF_mindset*) and elevated mood (*BSF_elevatedmood*).
- (e) The *ICD-10 Symptom Rating (ISR)*²⁴ measures the severity of different mental disorders and comprises 36 items, including 29 single items on a 5-level ordinal scale: 0 = “does not apply”, 1 = “a little”, 2 = “quite a bit”,

- 3 = “to a great extent”, and 4 = “extremely”. We used the 6 subscales depressive syndrome (ISR_depression), anxiety syndrome (ISR_anxiety), obsessive-compulsive syndrome (ISR_compulsivesyn), somatoform syndrome (ISR_somatosyn), eating disorder syndrome (ISR_eatingdisorder), additional items score (ISR_additionalitems) and the total psychiatric syndrome score (ISR_totalpsychiatric-syn).
- (f) The (*short form*) *Patient Health Questionnaire (PHQK)*²⁵ measures the extent of depressive symptoms in the last 2 weeks and symptoms of anxiety in the last 4 weeks. The depressivity scale (PHQK_depression) is calculated as sum of 9 ordinal items (0 = “not at all”, 1 = “at some days”, 2 = “more than half of the days”, 3 = “almost every day”). The binary panic syndrome score (PHQK_panicyn) is equal to “1” if all of the associated 5 “yes”/“no”-items are answered with “yes”.
 - (g) The *Perceived Stress Questionnaire (PSQ)*²⁶ contains 20 items measuring subjective levels of stress within the last 4 weeks on a scale from 0 = “hardly ever” to 2 = “rarely”, 3 = “frequently” and 4 = “mostly”. We used the subscales demand (PSQ_demand), tension (PSQ_tension), joy (PSQ_joy), and worries (PSQ_worries) as well as the total perceived stress score (PSQ_stress).
 - (h) The *Pain Perception Scale (SES)*²⁷ quantifies affective pain (SES_affectivepain) and sensoric pain (SES_sensoricpain) components as sum scores aggregated from a total of 24 items where the response options range from 1 = “does not apply at all” to 4 = “strongly applies”.
 - (i) *Visual Analogue Scales Pain (SSKAL)* consists of the 3 scales pain impairment (SSKAL_painimpairment), pain frequency (SSKAL_painfrequency) and pain intensity (SSKAL_painintensity).
 - (j) The *Short Form-8 Health Survey (SF8)*²⁸ assesses 8 aspects of health-related quality of life. Each of the scores is composed of 8 items, each with a different set of 5 to 6 response options. We used the subscales bodily health (SF8_bodilyhealth), overall health (SF8_overallhealth), mental health (SF8_mentalhealth), physical functioning (SF8_physicalfunct), role emotional (SF8_roleemotional), role physical (SF8_rolephysical), social functioning (SF8_socialfunct), vitality (SF8_vitality), as well as the mental component summary score (SF8_mentalcomp) and the physical component summary score (SF8_physicalcomp).
 - (k) The *Self-Efficacy-Optimism-Pessimism Scale questionnaire (Selbstwirksamkeits-Optimismus-Pessimismus Skala; SWOP)*²⁹ comprises 9 items on a scale from 1 = “not true” to 4 = “exactly true”. We used the scores on self-efficacy (SWOP_selfefficacy), optimism (SWOP_optimism) and pessimism (SWOP_pessimism).
 - (l) *Visual analogue scales (TINSKAL)* measuring tinnitus loudness (TINSKAL_loudness), frequency (TINSKAL_frequency) and distress (TINSKAL_distress) within a range between 0 and 10, respectively.
 - (m) From the *Tinnitus Localization and Quality Questionnaire (TLQ)*³⁰, we extracted binary features indicating the location where the tinnitus is perceived as loudest (TINSKAL_01_leftear, TINSKAL_01_rightear, TINSKAL_01_bothears, TINSKAL_01_entirehead) and binary features on the sound that describe the tinnitus best (TINSKAL_02_whistling, TINSKAL_02_hissing, TINSKAL_02_ringing, TINSKAL_02_rustling).
 - (n) The German version of the *Tinnitus Questionnaire (TQ)*³¹ is an instrument to assess tinnitus-related distress and tinnitus severity. The questionnaire comprises 52 items with 3 levels each (0 = “true”, 1 = “partially true”, 2 = “not true”). We used the sum scores on auditory perceptual difficulties (TQ_auditoryperceptdiff), cognitive distress (TQ_cognitivedistress), emotional distress (TQ_emodistress), intrusiveness (TQ_intrusiveness), psychological distress (TQ_psychodistress), sleep disturbances (TQ_sleepdisturbances), somatic complaints (TQ_somaticcomplaints), as well as the total tinnitus distress score (TQ_distress).

Data from 2875 (70.1%) patients who did not complete all questionnaires were excluded due to missing values. Excluded patients were slightly, but significantly older than those included in the final sample ($\mu_{\text{excluded}} = 51.73$, $\sigma_{\text{excluded}} = 13.63$; $\mu_{\text{included}} = 50.00$, $\sigma_{\text{included}} = 11.91$; $t(2630.8) = -4.07$, $p < 0.01$). Since all features of the SF8 and some features of the BSE, SWOP and PSQ have higher scores with better quality of life, features with a positive wording were reversed (new value = maximum feature value – old value) so that the interpretation (higher scores represent higher burden) remained consistent. Hereafter, feature names with a *-suffix denote reversed features. Due to widely differing value ranges, each feature was standardised prior to cluster analysis via z-score normalisation to have a mean of 0 and standard deviation of 1.

Identification of tinnitus phenotypes using clustering. To identify a distinct set of tinnitus phenotypes, the clustering algorithm X-means was employed³². X-means is a parameter-free extension of the popular K-means clustering algorithm that incorporates the Bayesian Information Criterion³³ (BIC) to automatically find an appropriate number of clusters K by finding a good trade-off between high goodness of fit and a low number of clusters. Let \mathcal{D} be the dataset with d dimensions and D a subset of \mathcal{D} , i.e., $D \subseteq \mathcal{D}$. A K-means clustering on D yields the set of clusters $\mathcal{C} = \{C_1, \dots, C_k, \dots, C_K\}$, where c_k is the centroid of cluster k , r_k is the number of points in D assigned to c_k and $p = (d + 1) \cdot K$. The BIC of a cluster C_k using Schwarz criterion is calculated as $\text{BIC}(C_k) = \hat{l}_k(\mathcal{D}) - \frac{p_k}{2} \cdot \log |\mathcal{D}|$, where $\hat{l}_k(\mathcal{D})$ is the log-likelihood of \mathcal{D} according to C_k . The point probabilities are computed as $\hat{P}(x_i) = \frac{r_{(i)}}{|\mathcal{D}|} \cdot \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp\left(-\frac{1}{2\hat{\sigma}^2} \|x_i - c_{(i)}\|^2\right)$, where the maximum likelihood estimate for the variance (under the identical spherical Gaussian assumption) is $\hat{\sigma}^2 = \frac{1}{|\mathcal{D}| - K} \sum_{i=1}^{|\mathcal{D}|} (x_i - \mu_{(i)})^2$. The log-likelihood of \mathcal{D} according to \mathcal{C} is $l(\mathcal{D}) = \log \prod_{i=1}^{|\mathcal{D}|} P(x_i) = \sum_{i=1}^{|\mathcal{D}|} \left(\log \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} - \frac{1}{2\hat{\sigma}^2} \|x_i - c_{(i)}\|^2 + \log \frac{r_{(i)}}{|\mathcal{D}|} \right)$.

The X-means algorithm consists of 4 steps: (1) First, an initial K-means is run with $K = K_{\text{lower}}$. (2) Then each centroid is bisected into two children which are placed in opposite directions along a randomly chosen vector. (3) A “local” 2-means clustering is run for each pair of children and a BIC score is assigned to this new partitioning along a randomly chosen vector. (4) If the BIC score increases with bisection of a centroid, the respective

child centroids are kept, otherwise, the parent centroid is kept. Steps (2)–(4) are repeated until neither centroid's BIC score can be improved by bisection. We used the R implementation of the X-means algorithm of Ishioka and set K_{lower} to 2³⁴. Numbering of clusters as cluster 1, cluster 2, etc. was done arbitrarily.

Cluster visualisation. Visualisation of clusters in high dimensionality is challenging. The popular scatterplot matrices (SPLOMs) and their extensions intuitively represent the relation between pairs of features as matrix where each non-diagonal element is a two-dimensional scatterplot^{35,36}. However, the number of scatterplots grows quadratically with increasing dimensionality which leads to scalability problems such as overplotting. Hence, advanced visualisation techniques have been proposed as a remedy, e.g. density contours, hexagon binning, coloring, transparency, layers showing aggregated geometric characteristics (minimal spanning trees, alpha shape, convex hull), animation, or combinations of multiple techniques such as splatterplots³⁷. Still, SPLOMs and other traditional visualization techniques such as parallel coordinate charts³⁸ are more suited for rather low-dimensional data with only a handful of features.

In case the original data cannot be adequately displayed on low-dimensional projections, dimensionality reduction (DR) is often applied as preprocessing step in advance of visualisation. DR algorithms transform a high-dimensional feature space onto a low-dimensional (often 2D) projection. Ideally, the projection preserves important structures of the original data, such as clusters, outliers, correlations and other important structures. Principal component analysis (PCA) is a frequently used DR algorithm that generates a new coordinate system with orthogonal dimensions³⁹. The new dimensions (principal components, PC) are linear combinations of the original dimensions and are sorted according to variance. Each PC carries a loading that characterises how much variability of the data is explained. PCA is primarily suited for normally distributed data. However, PCA has problems with outliers and is incapable of capturing non-linear relationships. Another popular DR algorithm is multi-dimensional scaling (MDS) which puts emphasis on preserving distance. Points that are close in high-dimensional space should also be close in low-dimensional space. For complex (arbitrarily-shaped) structures, large distance is meaningless because of the curse of dimensionality, thus results may be unsatisfactory. *t*-stochastic neighbourhood embedding (*t*-SNE) is a non-linear dimensionality reduction technique that visualises a matrix of pairwise-similarities⁴⁰. The similarities are calculated in a way to both preserve global structures (clusters at different scales) and local structures (distances and neighbours). *t*-SNE does not allow for interpretation of the original dimensions. Further, the technique does not support to add a new observation to the existing projection without recalculation.

DR techniques cannot be applied here because even if the clustering structure is preserved in the data projection, the semantics of the original features will be lost. Discussions with domain experts led to the following requirements for a cluster visualisation:

- preservation of original features,
- intuitive cluster representation for multi-variate data with dozens of features,
- compact, at a glance comparison of multiple clusters,
- chart design that allows to juxtapose a cluster with the overall patient mean.

We therefore introduce a new radial bar chart visualisation as graphical representation of a single cluster, enriched with dedicated elements that satisfy the aforementioned requirements. In particular, the height of a bar depicts the average value of a feature over the patients assigned to that cluster. The radial spatial layout distributes the bars around a circle where each bar starts at the black 0 line which represents the feature average over all patients included in this study. Due to the scaling (z-score normalisation) of the features, bars inclined to the outside represent feature averages above the overall patient mean and bars inclined to the inside represent feature averages below the overall patient mean. This interpretation is visually supported by colour-coded bars using a sequential gradient from dark blue (low burden) to yellow (mean burden) to bright red (high burden). Feature names are shown on top of each bar. All values are depicted in terms of standard deviation away from the mean. For example, a value of -1 indicates that the cluster average is 1 standard deviation smaller than the overall patient average. Intra-cluster standard deviation are represented as grey error lines facing the coloured inner circle. To facilitate quick feature localisation, features were grouped into categories which are displayed in the inner circle, alongside the cluster name and the number of patients assigned to that cluster. These categories were (in clockwise order): tinnitus characteristics, physical quality of life, experiences of pain, somatic expressions, affective symptoms, tinnitus-related distress, internal resources, perceived stress, and mental quality of life.

To provide a graphical overview of all clusters at the same time, we designed a radar chart variant where feature averages are represented as points instead of bars which allows to show multiple clusters. Within each feature category, the points of a cluster are connected by line segments. Points and line segments are coloured by cluster.

Interactive components for cluster inspection. To provide a graphical overview, an interactive demo of the cluster solutions and the visualisations is available under <https://unmnn.de/phs/app/>. Radar charts were augmented with interactive components: by hovering over a bar or a feature label, additional cluster summaries and compact feature descriptions are shown as tooltips. Clicking on a feature invokes an additional chart which shows the (normalised) distribution of the selected feature stratified by cluster, and if selected, also after treatment. Continuous features are shown using semi-transparent boxplots placed on violin plot⁴¹ layers whereas for nominal features, category proportions alongside their 95% confidence intervals are displayed as points and error lines, respectively.

Statistical methods. Kruskal–Wallis test was used for statistical comparison of inter-phenotype differences for continuous features (like age), and Pearson’s chi-squared test was used for categorical features (like gender). Differences were considered significant if $p < 0.05$. Correction for multiple comparison was not conducted due to the exploratory nature of the study. Confidence intervals for the means were estimated using nonparametric basic bootstrap sampling⁴² with 2000 samples, respectively.

Results

According to X -means, four clusters (referred to as phenotypes hereafter) represent an optimal solution for the given dataset. The radial barcharts in Fig. 1 visualise phenotype–individual averages for all features. Graphical summaries of phenotype value distributions for all features on their original scales are provided in Supplementary Fig. S1. The radar plot in Fig. 2 shows average scores for each variable and for all clusters. While the four phenotypes are clearly distinguishable with respect to the psychosomatic and somatic variables, the line segments for most features of the group “tinnitus characteristics” are close to the overall patient average. Phenotype 1 (PT 1) represents the largest subgroup with 697 out of 1,228 patients (56.8%). This patient subgroup is characterised by ostensibly below-average symptom expression across tinnitus-related and broader psychosomatic symptom indices, including affective symptoms, perceived stress, tinnitus-related distress and somatic symptoms, as well as (above-average) quality of life and internal resources (Fig. 1). Due to their help-seeking behaviour, presentation in clinic and wish to participate in multimodal treatment, it can be assumed that this group of patients do experience psychological distress, however aim to present themselves as healthily as possible. We therefore label this phenotype “avoidant group”. Patients in this subgroup feature proportionately high levels of education, employment and low levels of leave of absence and psychotherapeutic treatment (Table 1). PT 2 comprised 173 patients (14.1%) who reported the highest emotional and somatic burden among all PTs (Fig. 1b). More specifically, PT 2 represents a patient subgroup with high psychosomatic-comorbidity and is thus labelled “psychosomatic group”. This patient subgroup shows high tinnitus burden alongside clinically relevant impairment across all affective indices including depression, anxiety, and perceived stress. These affective symptoms appear to align with somatoform expressions of distress including somatic symptoms. Patients of these subgroup report severely reduced quality of life and reduced coping opportunities with more pessimism, less experienced self-efficacy and optimism. Patients in this subgroup feature a high proportion of patients who live alone, are unemployed or show an overall lower educational status. Patients in this cluster further appear to consult more doctors, take more leave of absence and use more psychotherapy. PT 2 patients reported the tinnitus sound to be audible in the entire head to a greater portion than the other groups. PT 3 contained 187 associated patients (15.2%) characterised by above-average scores of features measuring somatic complaints and near-average scores for affective symptoms (Fig. 1c). Since pain scores of SF8_bodilyhealth* and SSKAL_painfrequency were similarly large as PT2, this patient subset was labelled “somatic group”. PT 3 represented the oldest subgroup, with the largest proportion of female patients and largest reported time period since tinnitus onset. Unlike PT3, PT 4 ($n = 171$; 13.9%) exhibited above-average scores for affective scores, components of quality of life and perceived stress (Fig. 1d), e.g. mental component summary score (SF8_mentalcomp*; 0.85) and anxious depression score (BSF_anxdepression; 0.79). Hence, we label PT 4 as “distress group”. PT 4 comprises the youngest of the 4 subgroups, with the largest fraction of male patients (Table 1).

Discussion

In this study, we combined data-driven clustering with a novel visualisation to identify and display distinct phenotypes in a large sample of patients with chronic tinnitus. Patient data were extracted from self-report questionnaires prior to starting a multimodal treatment program. Our analysis suggests four phenotypes of patients with chronic tinnitus.

PT 1 (*avoidant group*) represents a large proportion of patients. Apart from the tinnitus symptom, patients in this subgroup reported few other affective or psychosomatic symptoms, and the tinnitus is used as an index-representation of experienced distress. Due to these patients’ focused presentation (“*everything is okay were it not for the tinnitus*”), clinicians can easily be led to believe that potential other contributors to individual distress must not require assessment. However, clinical experience strongly suggests that a thorough assessment of broader psychosocial stressors is warranted in so far as it is feasible in clinical practice environments. The psychosocial resourcefulness of this subgroup enables patients to seek help quickly and in a solution-focused manner. Good tinnitus-specific counselling and individualised (online) therapy modules featuring audiological, psychological or relaxation procedures would possibly represent an adequate treatment strategy for this patient subgroup.

PT 2 (*psychosomatic group*) represents 15% of patients who showed high tinnitus burden alongside clinically relevant impairment across all affective indices including depression, anxiety, and perceived stress. These affective symptoms appear to align with somatoform expressions of distress including physical complaints and somatic symptoms. Patients of these subgroup report severely reduced quality of life and reduced coping opportunities with more pessimism, less experienced self-efficacy and optimism. There is the frequently asked question as to whether increased tinnitus-related distress contributes to increases in depression or vice versa. In this group, we consider depressive or anxious symptoms to be a crucial underlying factor for general symptom burden and treatment must begin with a focus on improving mood and relieving depression. Here, tinnitus-related distress needs to be seen within a broader context of medical and psychological influencing factors that require idiosyncratic conceptualisation. According to the socio-demographic variables, this patient subgroup features a higher proportion of women, and more patients who live alone, are unemployed or show an overall lower educational status. Patients in this cluster further appear to consult more doctors, take more leaves of absence and use more psychotherapy.

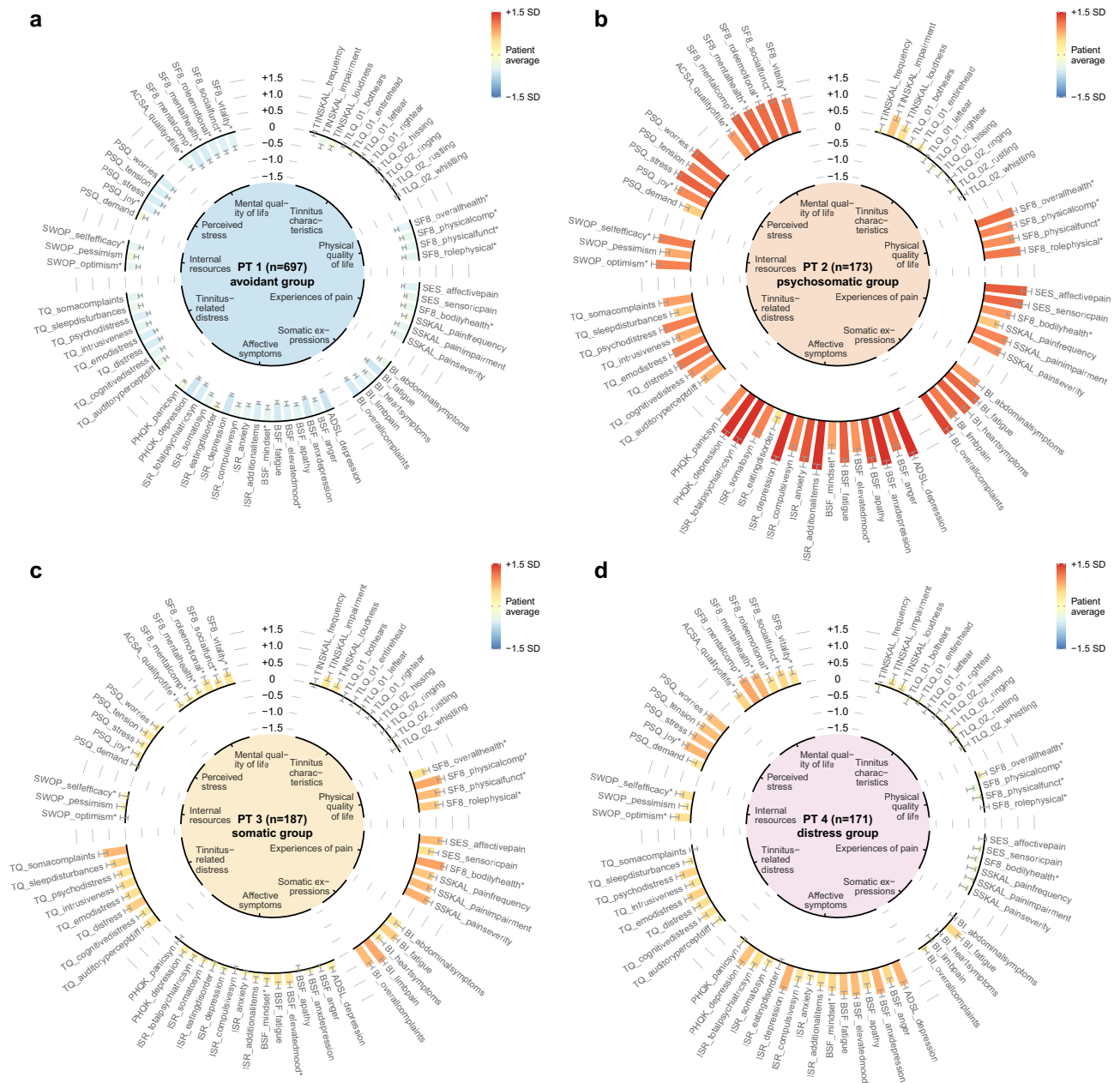


Figure 1. Radial barcharts visualizing the 4 phenotypes. **(a)** Phenotype 1 (PT1) characterises the patient subgroup with lowest health burden among all phenotypes. **(b)** PT2 represents the most suffering subgroup, with all of the psychosomatic and somatic measurement averages exceeding the population mean + 0.5 standard deviations (SD). PT3 **(c)** exhibits above population average scores for somatic indicators whereas PT4 **(d)** is characterised by increased distress scores, including subjective stress and perceived quality of life. Bars are arranged in a circular layout. The height of a bar shows a feature’s z-score normalised within-cluster average, and the grey line centred at the top of the bar illustrates the 95% confidence interval. The colour of a bar represents the difference of the within-cluster average from the overall patient average (PA), from - 1.5 SD below PA (dark blue) to PA (yellow) and + 1.5 SD above PA (bright red). Features were grouped into 9 categories defined by tinnitus experts. The categories are shown within the inner circle. See subsection Features for a description of each questionnaire and the extracted features.

PT 3 (*somatic group*) appears to represent a patient subgroup that is characterised by somatopsychic symptom expressions, i.e., physical symptoms that may reflect distress and/or underlying medical conditions. To adequately address the needs of this patient subgroup, multimodal interventions might include a proportion of body-oriented procedures such as relaxation exercises or physiotherapy whose effect, however, should be interpreted with regard to both direct and indirect psychological effects (e.g. through increased senses of well-being or others’ care).

Characteristic	Total ● N=1,228	PT 1 ● n=697	PT 2 ● n=173	PT 3 ● n=187	PT 4 ● n=171	p-value	Mean ± 95%-CI
Age (years)	50.0 [49.3,50.7]	50.3 [49.4,51.2]	49.1 [47.5,50.6]	52.2 [50.7,53.8]	47.3 [45.6,49.1]	<0.01*	
Male gender (%)	49.6 [46.8,52.4]	53.9 [50.5,57.7]	43.4 [36.4,50.9]	33.7 [26.7,40.6]	55.6 [48.0,63.2]	<0.01*	
German nationality (%)	95.2 [94.1,96.4]	96.7 [95.4,98.1]	91.3 [87.3,96.0]	93.6 [90.4,97.3]	94.7 [91.8,98.2]	0.02*	
In partnership (%)	72.5 [70.0,74.9]	74.5 [71.3,77.6]	65.3 [58.4,72.3]	74.3 [67.9,80.7]	69.6 [62.6,76.6]	0.08	
Abitur (%)	47.7 [45.0,50.5]	53.8 [50.1,57.5]	35.3 [27.7,42.2]	37.4 [30.5,44.4]	46.8 [39.8,54.4]	<0.01*	
Employed (%)	73.5 [71.1,75.9]	77.5 [74.3,80.6]	63.0 [55.5,69.9]	66.8 [60.4,73.8]	74.9 [68.4,81.3]	<0.01*	
Sick leave last year [months]	1.7 [1.6,1.9]	1.2 [1.1,1.4]	3.3 [2.8,3.8]	2.1 [1.7,2.4]	1.8 [1.5,2.2]	<0.01*	
Duration of psychotherapeutic treatment [months]	3.8 [3.5,4.0]	2.5 [2.2,2.8]	6.5 [5.8,7.2]	5.1 [4.4,5.8]	4.5 [3.7,5.2]	<0.01*	
Time since tinnitus onset [years]	2.7 [2.6,2.8]	2.6 [2.4,2.7]	2.8 [2.5,3.1]	3.1 [2.8,3.3]	2.8 [2.5,3.1]	0.01*	

Table 1. Inter-group comparison of socio-demographics. Summaries of socio-demographic features are given as means [95% confidence interval] for all patients and for each of the 4 phenotypes. An asterisk indicates whether $p < 0.05$.

Patients in PT 4 (*distress group*) reported above-average perceived stress, accompanied with physical exhaustion and anxious-depressive mood. This group includes rather younger, employed patients (more men), who indicated chronic distress, potentially being susceptible to burnout syndrome with subjective reduced mental capacity (“hamster wheel”), which is used as a description of the life situation even without tinnitus stress. In this subgroup, tinnitus might represent chronic stress associated with psychological vulnerabilities, environmental/work stressors and dysfunctional coping strategies. Multimodal therapy should initially focus on stress-regulation techniques, including relaxation or individually tailored behavioural modification approaches. Similar to the highly psychosomatically burdened PT 2, patients in PT 4 could also benefit from longer psychotherapeutic or multimodal treatment procedures (inpatient or rehabilitative).

The overview and juxtaposition of all clusters shows that some questionnaires and characteristics contribute a lot to differences among patient phenotypes. In particular, patient phenotypes differ substantially with respect to their coping attitudes, their stress and their perception of quality of life, as well as their tinnitus distress. Some of the questionnaire items separate well among some of the phenotypes, see e.g. the items on perceived pain and complaints. In contrast, patients do not seem to differ in their perception of tinnitus. These contributions of the questionnaires to the phenotypes indicate that phenotyping may be achievable also with less questionnaires, especially because some of the questionnaires are overlapping.

Previous studies also employed clustering algorithms to identify tinnitus subtypes^{13,14,16,17}. It is difficult to compare our findings to theirs because of the different set of available measurements: whereas the strength of our study was a large pool of self-report questionnaire data, Tyler et al. used both self-report data and audiometrics¹⁴, Schecklmann et al. used self-report data and cardiac imaging features¹⁷, and Langguth et al. used audiometric data¹³. Nevertheless, PT 2 (psychosomatic suffering group) appears to match the “constant distressing tinnitus” subgroup reported by Tyler et al.¹⁴, as average scores on features measuring tinnitus-related health burden were distinctly greater than in the other subgroups. Of course, the selection of meaningful features is pivotal for the efficacy of any cluster analysis. Schlee et al. argued against the usage of single-item features like visual analogue scale measurements because of their higher susceptibility to random measurement errors, lower test-retest reliability and higher vulnerability to unknown biases¹⁸. Due to the exploratory nature of our study, we decided to include both single-item measurements and compound scores. We assigned 11 out of 15 single-item measurements into the category “Tinnitus characteristics”. Figure 2 (top-right) shows their low discriminative power, as all phenotypes’ means are close to the population average, with the exception of TINSKAL_impairment and TINSKAL_loudness. Future research might focus on identifying a subset of key questionnaires alongside a simple computational tool that will enable clinicians to match individual patients with one or more of the here identified phenotypes.

Closest to our radial bar chart visualisation is the radar chart proposed by Schlee et al.¹⁸. Their solution indeed facilitates the comparison of two subgroups by comparing the areas of their associated polygons. However, there

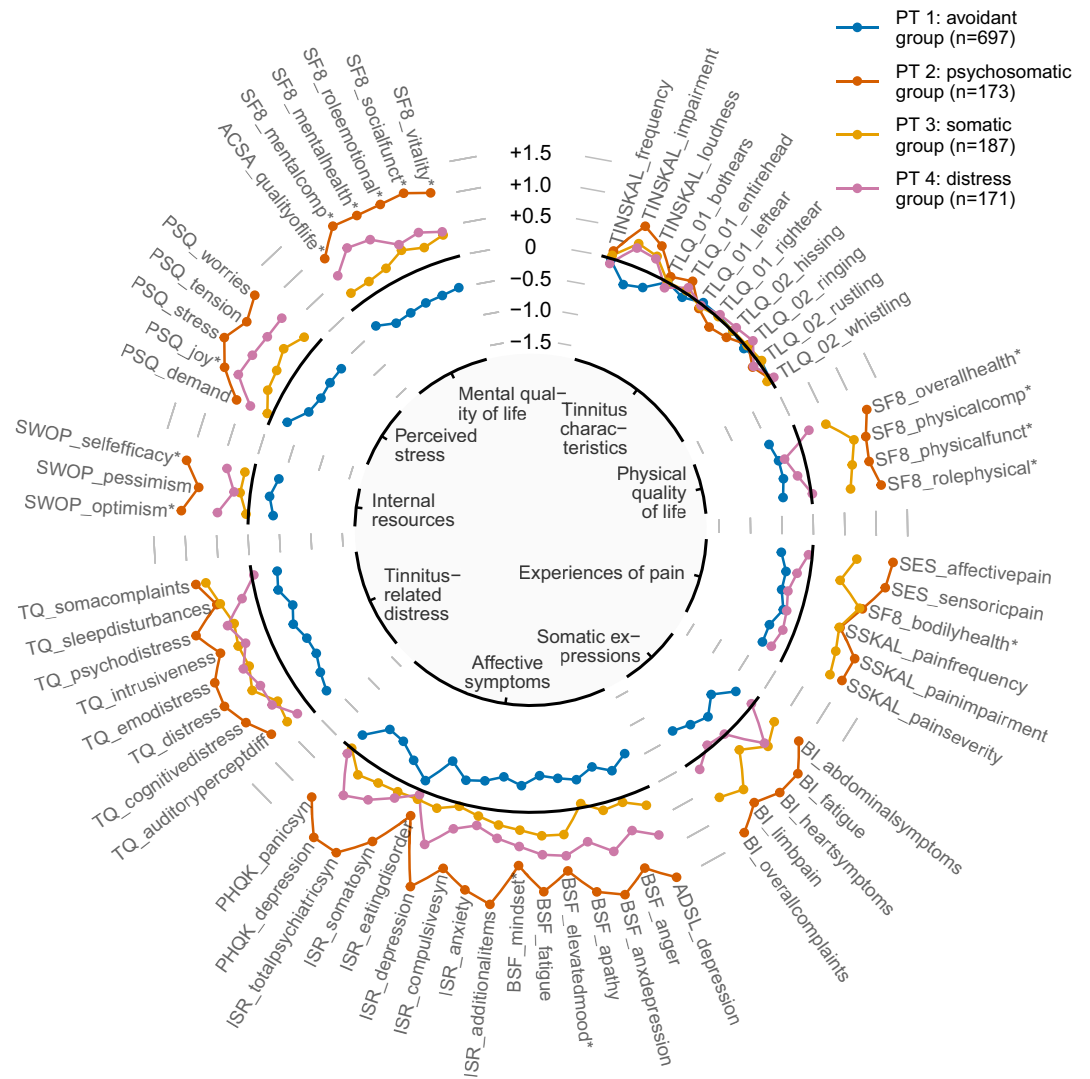


Figure 2. Radial line chart juxtaposing the 4 phenotypes. In the chart, a point shows a feature’s (z-score normalised) within-phenotype average. In each feature category (labels in inner circle), points are connected with line segments. Points and lines are coloured by cluster.

is still the potential problem of overplotting when one wishes to compare more than 2 subgroups as we do. Hence, we did not opt to fill up the areas spanned by the connected points with colour, to avoid a polygonal that fully overlays another one. Further, since the main criterion for comparison is the polygons’ shapes, inferences highly depend on the ordering of features around the plot which can be misleading. Schlee et al.¹⁸ tackled this problem by computing an ordering that yields areas that achieve maximum mean surface difference between subgroups and minimum surface variance within subgroups. This approach is feasible for up to a moderate (≈ 20) number of features. In our study with 64 features, we decided to arrange features based on semantic categories, e.g. quality of life. This allows to detect and track features easier. Further, our visualisation is not specific to tinnitus but could be used to present a compact visual summary of characteristics of any condition or index symptoms subgroups. Whether the visualisations will be adopted by clinicians for finding suitable tinnitus management strategies needs to be tested. Preliminarily, clinicians suggested that graphical summaries of possible patient subtypes may alleviate allocation of modular treatment strategies to specific combinations of symptom presentations.

A potential limitation of our analysis is the exclusion of patients who did not fill all questionnaires during admission. There are several reasons why a patient may not have filled all questionnaires, including unfamiliarity with the technical devices (the questionnaires must be filled electronically), loss of motivation due to the relatively large number of questionnaires and collision with baseline examinations in the lab. The exclusion of these patients may have led to a selection bias. Nonetheless, our analysis over all 15 questionnaires allowed us to acquire insights to the contribution of these questionnaires to phenotyping, so that eventually a reduction of questionnaires might become possible. Further, our analysis is a static snapshot of phenotypes at baseline. Hence, it might be sensitive to possible changes in tinnitus perception and associated health burden over time. It is possible that a patient will transition from one phenotype to another in later stages of her life or depending on her

tinnitus management. Thus, a next step would be to study the effects of treatment to these phenotypes and find whether some patient phenotypes benefit more than others. Another limitation comes from the heuristic choice of the number of phenotypes. Many works use the number of clusters as input parameter. Since this number is not known, we used the non-parametric X-Means clustering algorithm.

Data availability

The datasets for this article are not publicly available because no consent of the patients to publish their data was obtained. Notwithstanding, interested researchers can contact the directorate of the Tinnitus Center Charité Universitaetsmedizin Berlin with data access requests addressed at the senior author [birgit.mazurek@charite.de].

Received: 25 May 2020; Accepted: 16 September 2020

Published online: 02 October 2020

References

1. Baguley, D., McFerran, D. & Hall, D. Tinnitus. *Lancet* **382**, 1600–1607 (2013).
2. Bhatt, J. M., Bhattacharyya, N. & Lin, H. W. Relationships between tinnitus and the prevalence of anxiety and depression. *Laryngoscope* **127**, 466–469. <https://doi.org/10.1002/lary.26107> (2017).
3. Maes, I. H., Cima, R. F., Vlaeyen, J. W., Anteunis, L. J. & Joore, M. A. Tinnitus: A cost study. *Ear Hear.* **34**, 508–514. <https://doi.org/10.1097/aud.0b013e31827d113a> (2013).
4. Cederroth, C. R. *et al.* Towards an understanding of tinnitus heterogeneity. *Front. Aging Neurosci.* **11**, 53 (2019).
5. Hobson, J., Chisholm, E. & El Refaie, A. Sound therapy (masking) in the management of tinnitus in adults. *Cochrane Database Syst. Rev.* <https://doi.org/10.1002/14651858.CD006371.pub3> (2012).
6. Kröner-Herwig, B., Frenzel, A., Fritsche, G., Schilkowsky, G. & Esser, G. The management of chronic tinnitus: Comparison of an outpatient cognitive-behavioral group training to minimal-contact interventions. *J. Psychosom. Res.* **54**, 381–389. [https://doi.org/10.1016/S0022-3999\(02\)00400-2](https://doi.org/10.1016/S0022-3999(02)00400-2) (2003).
7. Henry, J. & Wilson, P. The psychological management of tinnitus: Comparison of a combined cognitive educational program, education alone and a waiting-list control. *Int. Tinnitus J.* **2**, 9–20 (1996).
8. Hesser, H., Weise, C., Westin, V. Z. & Andersson, G. A systematic review and meta-analysis of randomized controlled trials of cognitive-behavioral therapy for tinnitus distress. *Clin. Psychol. Rev.* **31**, 545–553. <https://doi.org/10.1016/j.cpr.2010.12.006> (2011).
9. Martinez-Devesa, P., Waddell, A., Perera, R. & Theodoulou, M. Cognitive behavioural therapy for tinnitus. *Cochrane Database Syst. Rev.* <https://doi.org/10.1002/14651858.CD005233.pub3> (2007).
10. Grewal, R., Spielmann, P., Jones, S. & Hussain, S. Clinical efficacy of tinnitus retraining therapy and cognitive behavioural therapy in the treatment of subjective tinnitus: A systematic review. *J. Laryngol. Otol.* **128**, 1028–1033. <https://doi.org/10.1017/S0022215114002849> (2014).
11. Jastreboff, P. J. *et al.* Tinnitus retraining therapy (TRT) as a method for treatment of tinnitus and hyperacusis patients. *J. Am. Acad. Audiol.* **11**, 162–177 (2000).
12. Phillips, J. S. & McFerran, D. Tinnitus retraining therapy (TRT) for tinnitus. *Cochrane Database Syst. Rev.* <https://doi.org/10.1002/14651858.CD007330.pub2> (2010).
13. Langguth, B. *et al.* Different patterns of hearing loss among tinnitus patients: A latent class analysis of a large sample. *Front. Neurol.* **8**, 46. <https://doi.org/10.3389/fneur.2017.00046> (2017).
14. Tyler, R. *et al.* Identifying tinnitus subgroups with cluster analysis. *Am. J. Audiol.* **17**, 176–184 (2008).
15. Landgrebe, M. *et al.* The Tinnitus Research Initiative (TRI) database: A new approach for delineation of tinnitus subtypes and generation of predictors for treatment outcome. *BMC Med. Inform. Decis. Mak.* **10**, 42. <https://doi.org/10.1186/1472-6947-10-42> (2010).
16. van den Berge, M. J. C. *et al.* Cluster analysis to identify possible subgroups in tinnitus patients. *Front. Neurol.* **8**, 115. <https://doi.org/10.3389/fneur.2017.00115> (2017).
17. Schecklmann, M. *et al.* Cluster analysis for identifying sub-types of tinnitus: A positron emission tomography and voxel-based morphometry study. *Brain Res.* **1485**, 3–9. <https://doi.org/10.1016/j.brainres.2012.05.013> (2012).
18. Schlee, W. *et al.* Visualization of global disease burden for the optimization of patient management and treatment. *Front. Med.* **4**, 86. <https://doi.org/10.3389/fmed.2017.00086> (2017).
19. Bernheim, J. L. & Buysse, M. The anamnestic comparative self-assessment for measuring the subjective quality of life of cancer patients. *J. Psychosoc. Oncol.* **1**, 25–38. https://doi.org/10.1300/J077v01n04_03 (1993).
20. Radloff, L. S. The CES-D scale: A self-report depression scale for research in the general population. *Appl. Psychol. Meas.* **1**, 385–401. <https://doi.org/10.1177/014662167700100306> (1977).
21. Hautzinger, M. & Bailer, M. ADS-Allgemeine Depressionsskala. In *Diagnostische Verfahren in der Psychotherapie* (eds Brähler, E. *et al.*) 25–27 (Hogrefe, Göttingen, 2002).
22. Hörhold, M. *et al.* Testing a screening strategy for identifying psychosomatic patients in gynecologic practice. *Psychother. Psychosom. Med. Psychol.* **47**, 156–162 (1997).
23. Hörhold, M., Klapp, B. & Schimmack, U. Testungen der Invarianz und der Hierarchie eines mehrdimensionalen Stimmungsmodells auf der Basis von Zweipunkterhebungen an Patienten- und Studentenstichproben. *Z. Med. Psychol.* **2**, 27–35 (1993).
24. Tritt, K. *et al.* Entwicklung des Fragebogens ICD-10-Symptom-Rating (ISR). *Z. Psychosom. Med. Psychother.* **54**, 409–418. <https://doi.org/10.13109/zptm.2008.54.4.409> (2008).
25. Spitzer, R. L. *et al.* Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. *JAMA* **282**, 1737–1744. <https://doi.org/10.1001/jama.282.18.1737> (1999).
26. Fliege, H. *et al.* The perceived stress questionnaire (PSQ) reconsidered: Validation and reference values from different clinical and healthy adult samples. *Psychosom. Med.* **67**, 78–88. <https://doi.org/10.1097/01.psy.0000151491.80178.78> (2005).
27. Geissner, E. The pain perception scale—A differentiated and change-sensitive scale for assessing chronic and acute pain. *Die Rehabil.* **34**, 35–43 (1995).
28. Bullinger, M. & Morfeld, M. Der SF-36 Health Survey. In *Gesundheitsökonomische Evaluationen* (eds Schöffski, O. & von der Schulenburg, J.-M.G.) 387–402 (Springer, Berlin, 2008). https://doi.org/10.1007/978-3-540-49559-8_15.
29. Scholler, G., Fliege, H. & Klapp, B. F. Fragebogen zu Selbstwirksamkeit, Optimismus und Pessimismus. *Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID)* **49**, 275–283. <https://doi.org/10.23668/psycharchives.337> (1999).
30. Goebel, G. & Hiller, W. Psychische Beschwerden bei chronischem Tinnitus: Erprobung und Evaluation des Tinnitus-Fragebogens (TF). *Verhaltenstherapie* **2**, 13–22. <https://doi.org/10.1159/000258202> (1992).
31. Goebel, G. & Hiller, W. *Tinnitus-Fragebogen:(TF); ein Instrument zur Erfassung von Belastung und Schweregrad bei Tinnitus; Handanweisung* (Hogrefe, Verlag für Psychologie, Göttingen, 1998).
32. Pelleg, D. *et al.* X-means: Extending k-means with efficient estimation of the number of clusters. *ICML* **1**, 727–734 (2000).
33. Schwarz, G. *et al.* Estimating the dimension of a model. *Ann. Stat.* **6**, 461–464. <https://doi.org/10.1214/aos/1176344136> (1978).

34. Ishioka, T. An expansion of x-means for automatically determining the optimal number of clusters. *Proc. Int. Conf. Comput. Intell.* **2**, 91–95 (2005).
35. Im, J.-F., McGuffin, M. J. & Leung, R. GPLOM: The generalized plot matrix for visualizing multidimensional multivariate data. *IEEE Trans. Vis. Comput. Graph.* **19**, 2606–2614. <https://doi.org/10.1109/TVCG.2013.160> (2013).
36. Klemm, P. *et al.* 3D regression heat map analysis of population study data. *IEEE Trans. Vis. Comput. Graph.* **22**, 81–90. <https://doi.org/10.1109/TVCG.2015.2468291> (2015).
37. Mayorga, A. & Gleicher, M. Splatterplots: Overcoming overdraw in scatter plots. *IEEE Trans. Vis. Comput. Graph.* **19**, 1526–1538. <https://doi.org/10.1109/TVCG.2013.65> (2013).
38. Hartigan, J. A. Printer graphics for clustering. *J. Stat. Comput. Simul.* **4**, 187–213. <https://doi.org/10.1080/00949657508810123> (1975).
39. Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **24**, 417. <https://doi.org/10.1037/h0071325> (1933).
40. Maaten, Lvd & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
41. Hintze, J. L. & Nelson, R. D. Violin plots: A box plot-density trace synergism. *Am. Stat.* **52**, 181–184 (1998).
42. Davison, A. C. & Hinkley, D. V. *Bootstrap Methods and Their Application* Vol. 1 (Cambridge University Press, Cambridge, 1997).

Acknowledgements

We thank Bernhard Preim for his insightful suggestions on an early manuscript draft.

Author contributions

U.N. designed and performed the data analysis and wrote the original draft. P.B., B.B., W.M., and B.M. augmented the medical and psychological interpretation and discussion of the phenotypes. B.B., P.B., M.S., and B.M. supervised the data analysis. B.B., P.B., M.S., and B.M. reviewed and edited the manuscript. P.B., B.B., M.R., and B.M. curated the datasets. M.S., M.R. and B.M. led the project.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-73402-8>.

Correspondence and requests for materials should be addressed to U.N.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020