Three distinct patterns of mental response after a sport accident

Manuscript

Katharina Hüfner, MD1,✉, Hanna Salvotti, MD1, and Piotr Tymoszuk, PhD2

2023-02-24

1 Department of Psychiatry, Psychotherapy, Psychosomatics and Medical Psychology, University Hospital for Psychiatry II, Medical University of Innsbruck, Innsbruck, Austria  
2 Data Analysis as a Service Tirol, Innsbruck, Austria

✉ Correspondence: [Katharina Hüfner, MD <Katharina.Huefner@i-med.ac.att>](mailto:Katharina Hüfner, MD <Katharina.Huefner@i-med.ac.att>)

# Abstract

**Background:** Long-term psychological consequences of sport accidents and sport accident victims at risk of mental deterioration are poorly characterized. We sought to identify patterns of mental response following leisure time sport accidents.

**Methods:** A cross-sectional survey study was conducted with 307 individuals treated at the Innsbruck University Hospital (Austria) following a non-professional sport accident. Measures of anxiety (GAD-7), depression, somatic symptoms and panic (PHQ), resilient comping (RS-13), sense of coherence (SOC-9L), quality of life (EUROHIS), post-traumatic growth (PTGI) and post-traumatic syndrome (PTSD) were recorded along with demographic, socioeconomic, medical history, accident- and injury-related data. Patterns of mental readouts were defined by semi-supervised clustering and factors affecting the cluster assignment were identified by statistical testing and conditional random forest machine learning.

**Results:** Participants were assigned to three mental clusters characterized by a neutral mental response, post-traumatic growth and symptoms of post-traumatic syndrome disorder, respectively. Overall differences in demographic, socioeconomic, injury- and accident-related variables between the clusters were low-to-moderate. The post-traumatic growth and post-traumatic syndrome clusters were characterized by relatively frequent hospitalization, surgery, somatic illness and female gender. The most predictive factors for the post-traumatic syndrome cluster were more cautious behavior and flashbacks during sport activity, low income, mental disorders, somatic consequences of the accident and subjective need for psychological support.

**Conclusion:** A subset of non-professional sport accident victims develops signs of mental health problems including symptoms of post-traumatic syndrome disorder. Patients with low socioeconomic status, somatic and mental conditions, and hospitalization represent a vulnerable group which may benefit most from a psychological intervention.

# Keywords

sport accident, mental health, post-traumatic syndrome disorder, post-traumatic growth

# Introduction

# Methods

## Study population

*Hanna and Katharina: could you please provide a description?*

## Procedures

The psychometric battery (**Supplementary Table S2**) consisted of German versions of assessment tools for anxiety (GAD-7: 7-item general anxiety disorder scale) (1), depression (PHQ: patient health questionnaire, PHQ-9) (2,3), panic (PHQ-panic module) (2,3), persistent somatic symptoms (PHQ-15) (4), resilient coping (RS13: 13-item resilience scale) (5), loss of sense of coherence (SOC-9L: Leipzig 9-item sense of coherence questionnaire) (6), loss of quality of life (EUROHIS-QOL 8: 8-item EUROHIS project quality of life scale) (7), post-traumatic growth (PTGI: post-traumatic growth inventory) (8) and post-traumatic syndrome disorder (PCL-5 DSM-5: PTSD checklist for DSM-5) (9). The tools displayed good-to-excellent consistency as measured by McDonald’s > 0.8 (10) (**Supplementary Table S3**).

Symptoms of anxiety were defined as GAD-7 10, symptoms of depression were defined as PHQ-9 11 (11), significant somatization/persistent somatic symptoms were defined as PHQ-15 11 (4). Resilient coping classes were defined as follows: low: RS-13 0 - 65, moderate: 66 - 72, high: 73 (5).

Separate scores were computed for each domain of the PTGI tool (I: relations, II: possibilities, III: personal strength, IV: spiritual strength, V: life appreciation) with each item scores as 0: none, 1: very little, 2: little, 3: moderate, 4: great, 5: extremely great. In addition, the total PTGI score was calculated as the sum of all items (8).

Separate scores were calculated for domains B, C, D and E of the PCL-5 DSM-5 tool along with the total score being the sum of all items. Participants positive for the B domain or C domain were identified by at least one item scored with ‘moderate’, participants positive for the D or E domain were identified by at least two item scored with ‘moderate’. Significant PTSD symptoms where assumed in participants screened positive for at least one of the PCL-5 DSM-5 domains (9).

Traumatic events prior to the sport accident were assessed with the DIA-X tool (12). Flashbacks during sport activity, were surveyed as a single yes/no item. Flashbacks frequency was assessed in the following categories: none, more than one per month and more than one per year. Smoking was surveyed as a single yes/no question. Substance abuse was investigated with the CAGE tool (13). Data on hospital treatment, hospitalization and surgery were extracted from electronic patient’s records. Injury severity was assessed with the abbreviated injury scale (AIS) (14).

Additional information on study variables and their stratification schemes are presented in **Supplementary Methods** and **Supplementary Table S1**.

The study participants were assigned to the training and test subsets with the 2:1 size ratio with a random procedure aiming at minimal differences in accident year, age, gender, somatic and mental illness rates, frequencies of prior traumatic events (12) and distribution of injury severity grades (14) (**Figure 1**). The training subset size (n = 204) was sufficient for clustering analysis as investigated by comparing Hopkins metrics of clustering tendencies (15) in random subsamples of the entire dataset of varying size (**Supplementary Methods**).

## Analysis endpoints

The primary analysis endpoint was identification of patterns of mental response by clustering of leisure time sport accident victims in respect to measures of anxiety, depression, panic, somatization, quality of life, post-traumatic growth and post-traumatic symptom disorder by clustering analysis. The secondary analysis endpoints were demographic, socioeconomic and clinical characteristic of the mental response clusters and development of a classification procedure to discern individuals at risk of deteriorated mental health following a sport accident.

## Statistical analysis

Statistical analysis was done with R version 4.2.0. Numeric variables are displayed as medians with interquartile ranges. Categorical variables are presented as percentages and counts within the complete observation set. The majority of numeric variables were non-normally distributed (Shapiro-Wilk test). Hence, differences in numeric variables between analysis groups were analyzed by Mann-Whitney test with r effect size statistic or Kruskal-Wallis test with effect size statistic (16). Differences in distribution of categorical variables were assessed by test with Cramer V effect size statistic (16).

The training subset was clustered in respect to normalized median-centered numeric psychometric scores by the partition around medoids algorithm with the cosine distance measure between observations (17,18). The clustering procedure was characterized a good explanatory performance (explained clustering variance: ratio of between-cluster sum of squares to total sum of squares) and the superior stability in 10-fold cross-validation (19) as compared with hierarchical and KMEANS algorithms. The cluster number was chosen based on the bend of the curve of within-cluster sum of squares and the peak of mean silhouette statistic (15). Three mental clusters were defined: the neutral, PTG (post-traumatic growth) and PTSD (post-traumatic syndrome disorder) cluster. Assignment of the training subset observations to the mental clusters was accomplished with an inverse distance weighted 7-nearest neighbor classifier. The quality of cluster predictions was assessed by comparing explained clustering variances in the training and test subset, visual inspection of UMAP (uniform manifold approximation and projection) layouts (20,21), of distance heat maps and heat maps of normalized clustering feature levels.

Univariable classifiers of the mental cluster assignment for demographic, socioeconomic, clinical and accident-related parameters were trained in the training subset of the study cohort by the one-rule algorithm (22,23). The multi-parameter classifier was developed in the training subset with the conditional random forest algorithm (1000 random trees, mtry = 5) (24–26). The cluster assignment was subsequently predicted for the test subset observations and accuracy and statistic for the predictions were computed (27). Conditional importance for explanatory variables in the random forest classifiers were expressed as accuracy loss (28).

# Results

# Characteristic of the study cohort

A total of 4559 patients aged 16 years treated at the emergency, surgery, orthopedics and trauma departments at the University Clinic of Innsbruck (Austria) following a non-professional sport accident between 2018 and 2020 were invited to participant in the study survey. Out of them, 370 completed the survey and 307 individuals with the complete psychometric parameter set were analyzed here amounting to an 83% overall response rate (**Figure 1**, **Supplementary Tables S1** - **S3**).

As compared with the study participants, the individuals excluded from the analysis had lower employment rates, household income and less chronic somatic conditions. The excluded collective tended towards more biking and less ski/snowboard accidents and less professional rescue need. Their injuries were on average less severe as measured by AIS (14) with more often upper limb injuries, and lower hospitalization and surgery rates. The excluded individuals tended suffered less often from somatic health consequences of the accident and tended towards lower ratings of post-traumatic syndrome disorder but also diminished post-traumatic growth and less sense of coherence. These differences were significant or near-significant with weak effects sizes (V 0.22 or r 0.15). This analysis suffered also from high data missingness in the excluded collective (**Supplementary Table S4**).

In the analyzed cohort, the accidents happened between 2016-01-05 and 2020-12-31 and were managed at the study site between 2018-01-01 and 2020-12-31. The median time between the ward management and survey completion was 1343 days (interquartile range: 804 - 1441). The participants were predominantly middle-aged (median 51 years) and 45% of them were females. The vast majority had secondary or tertiary education grade and two-third were professionally active. Sport and trauma-risk professions constituted only a minor fraction of the cohort (< 8%). Over 40% of participants had high annual household incomes of 45000 Euro. Less than 10% of participants were smokers or alcohol/drug consumers. Frequencies of somatic and mental conditions were 15% and 5.2%, respectively. The rates of smoking (Austria: 20.6%) and somatic (38.3%) illness were lower than in the general population (29,30). Four of ten participants were directly affected by or witnessed a traumatic event prior to the accident as gauged by the DIA-X questionnaire (12).

Nearly 40% of participants had sport accidents in the past. Concerning the recent accident, the almost two-third of them were injured during skiing or snowboarding followed by biking and mountain sport. One-third were alone during the accident and, in the vast majority of case, were the sole culprit and victim. Roughly the half of participants could rescue themselves, professional rescue team was involved in 29% of cases. In 35% of participants the injury grade was moderate (AIS 2) and in 28% severe to life threatening (AIS 3). Upper and lower limb injuries were the most frequent followed by head and face (**Supplementary Figure S1A**). Hospitalization and surgery rates were 26% and 14%, respectively. Psychological or psychiatric support was applied to 9.1% individuals. Among those who had not received it, 7.5% declared the need for psychological support concerning the accident. Despite somatic health consequences of the accident were reported by 37% and unwilling flashbacks during sport in 40%, the majority of participants returned to the accident sport type ( 85%). Yet, two-third described their behavior during sport activity as more cautious (**Table 2**).

Clinically relevant signs of anxiety (2.3%), depression (5.5%) and somatization (4.9%) were rare in the study cohort. The median loss of quality of life was scored with 1.6 EUROHIS QOL points, which was substantially more than in a generalized Western European population (7). The participants were most satisfied with their overall quality of life and housing (**Table 3**, **Supplementary Figure S1B**). Roughly two-third of participants were assigned to the high resilient coping class (5). The scores of overall post-traumatic growth (median: 32) and its specific domains were lower in the study cohort as compared with the figures originally reported for the PTGI tool (8) (**Table 3**, **Supplementary Figure S1C**). In 19% of participants symptoms of post-traumatic syndrome disorder defined as positivity for at least one PCL-5 DSM-5 tool domain (9) could be observed. Among them, domain B symptoms (11%) were the most frequent and domain D symptoms were the rarest (5.2%) (**Table 3** and **Supplementary Figure S1D**).

For sake of semi-supervised clustering and classification procedure development, the study cohort was split into the training (n = 204) and test subset (n = 103). There were no significant differences between the subsets in demographic, socioeconomic, clinical and accident-related parameters. The training subset tended towards more frequent face injuries and low resilient coping class individuals. Yet the effect size of those differences was weak (V < 0.14) (**Supplementary Table S5**).

# Three clusters of mental response in sport accident victims

To comprehend types of mental response to the sport accident, we subjected the study participants to semi-supervised clustering in respect to measures of anxiety, depression, panic, persistent somatic symptoms, resilience, sense of coherence, quality of life, post-traumatic growth and post-traumatic syndrome disorder. Among several clustering algorithms compared in the training subset, partition around medoids with cosine distance between observations (17,**Schuber2019?**) demonstrated good explanatory power and superior reproducibility in cross-validation (19) (**Supplementary Figure S2A**). Three clusters of participants named neutral, PTG (post-traumatic growth) and PTSD (post-traumatic syndrome disorder) cluster after the patterns of their characteristic mental features were identified in the training subset. Subsequently, the mental cluster assignment could be robustly validated in the test subset as evident from comparable fractions of explained clustering variance in both study cohort subsets (training: 0.54, test: 0.54) and comparable cluster sizes (**Supplementary Figures S1B** - **S4**).

Roughly one-third study participants assigned to the neutral clusters were characterized by low scores of anxiety, depression, panic and somatic symptoms along with high rating of coherence, resilient coping and quality of life. Scores of post-traumatic growth and post-traumatic syndrome were the lowest in the neutral cluster. The PTG cluster demonstrated similarly low score of major mental health problems as neutral cluster individuals. By contrast, the overall measures of post-traumatic growth and its domains were the highest in the PTG cluster. In addition, ratings of post-traumatic syndrome disorder symptoms were slightly elevated in the PTG as compared with the neutral cluster. The remaining participants subsumed in the PTSD cluster displayed the most intense signs of anxiety, depression, panic and somatic symptoms as well as poor sense of coherence, resilient coping and low quality of life as compared with the neutral and PTG cluster. PTSD cluster was hallmarked by the highest scores of post-traumatic syndrome disorder symptoms in every PCL-5 DSM-5 domain. Post-traumatic growth was also present in the PTSD cluster as evident from elevated PTGI scores (**Figure 2**, **Supplementary Figure S4**, **Supplementary Table S6**).

Along this line, clinically relevant symptoms of anxiety, depression and somatization were present virtually only in the PTSD clusters of the training and test subsets of the study cohort. Furthermore, frequencies of low and moderate resilient coping classes were by far the highest in the PTSD cluster. Finally, approximately 40% of PTSD cluster individuals were screened positive for at least one post-traumatic syndrome disorder symptom which was more than in the PTG (training: 17%, test: 12%) or neutral cluster (training: 3.2%, test: 5.6%). An analogical trend was observed for the remaining post-traumatic syndrome disorder domains (**Figure 3**).

Collectively, the PTG cluster may represent individuals with salutary reaction to their sport accident, whereas the PTSD cluster individuals may pose a group of mental health deterioration.

# Demographic, socioeconomic and clinical background of the mental clusters

Next, we sought to explore differences in demography, socioeconomic status, clinical background, accident details and consequences between the mental clusters. Among 50 investigated variables, only 6 parameters were found significant or near-significant (p < 0.1) in both the training and test subsets of the study cohort (**Supplementary Table S6**).

A substantial enrichment of females, low-to-middle income and non-tertiary education participants could be observed in the PTSD cluster. The PTG cluster had the highest share of middle-aged individuals. Frequency of somatic conditions were the highest in the PTG and PTSD clusters. Mental illness participants were in turn represented virtually only in the PTSD cluster. Although there were no differences in frequency of previous sport accidents between the clusters, PTG cluster individuals were found to be affected by traumatic events in the past at a higher percentage than neutral cluster individuals. These differences were, however, non or borderline significant and of weak effect size (**Figure 4**, **Supplementary Figure S5**). There were no consistent differences in accident sport type, rescue mode, injury severity or injured body parts between the mental clusters. Yet, participants assigned to the PTG and PTSD clusters were tended towards higher hospitalization and surgery rates (**Figure 5** and **Supplementary Figure S6**).

PTSD cluster participants reported somatic health consequences of the accident and flashbacks significantly more often than neutral cluster individuals. Additionally, PTG and PTSD cluster individuals tended to describe their behavior during sport activity as more careful as compared with the neutral cluster. Roughly one-sixth PTSD cluster participants declared need for psychological or psychiatric support as compared with none in the neutral clusters of the training or test subset. Only a small fraction the study participants received psychological support following the accident. Although such individuals tended to be enriched in the PTG cluster, their frequency (up to 16%) was low. Hence it is unlikely, that psychological support affected the mental cluster assignment (**Figure 6**).

In summary, differences in non-mental features between the clusters were weak. Yet, the PTSD cluster could be characterized by elevated fractions of female, low-to-middle income class individuals, increased rates of somatic and mental conditions, hospitalization and surgery. Furthermore, such individuals were disproportionately often affected by somatic health aftermath and flashbacks.

# Prediction of the mental cluster assignment by demographic, socioeconomic and accident-related factors

Finally, we intended to establish a mental cluster classification procedure based solely on demographic, socioeconomic, clinical and accident-related factors. As investigated by a simple one rule classification algorithm (22,23), predictive value of single explanatory parameters was low with cluster assignment accuracy 0.46 and 0.18. In such univariable analysis, flashback presence and frequency, somatic health consequences of the accident, age class, annual income and traumatic events in the past were identified as the most predictive factors (**Supplementary Figure S7**).

To account for cumulative effects of multiple weak predictors of the mental cluster assignment, we employed multi-parameter conditional random forest modeling (24–26). The conditional random forest classifier employing the full set of demographic, socioeconomic, clinical and accident-related explanatory factors could correctly assign 79% training subset observations to their mental clusters ( = 0.69). However, the predictive performance in the test subset of the study cohort was only fair with correct mental cluster predictions in 52% of cases ( = 0.28). The predictive performance was the best for PTG cluster individuals of whom >80% and > 62% were classified correctly in the training and test subset, respectively (**Figure 7A**). The most important explanatory variables for the mental cluster prediction by conditional random forests were presence and frequency of flashbacks, cautious behavior during sport activity after the accident, annual income, sex and somatic health consequences of the accident.

# Discussion

In individuals surveyed on average 3.5 years after the non-professional sport accident, some of its consequences like somatic health problems, unwilling flashbacks as well as phenomena of post-traumatic growth and symptoms of post-traumatic syndrome disorder were still tangible. By semi-supervised clustering we could discern three distinct patterns of mental health in this collective with each present in approximately one-third of participants: (1) a neutral mental response reaction, (2) salutatory reaction to the accident characterized by post-traumatic growth and (3) mental health deterioration hallmarked by symptoms of anxiety, depression, somatization, resilience loss and signs of post-traumatic syndrome disorder. Notably, separation of the PTG and PTSD clusters was in part blurry. Symptoms of post-traumatic syndrome were observed also in PTG cluster participants albeit less frequently than in the PTSD cluster. Conversely, mental health problems in PTSD cluster participants were paralleled by some signs of post-traumatic growth. In a detailed comparison of 50 factors between the clusters, no distinct demographic, socioeconomic, clinical or accident-related background of the mental clusters could be identified. We found a moderate enrichment of females, low-to-moderate income participants, non-academic graduates, chronic somatic and psychiatric conditions as well as individuals requiring hospitalization and surgery in the PTG and, in particular, in the PTSD cluster. Injury severity and localization or professional rescue rates indicative of potentially more severe accident were comparable between the clusters. Similarly, the cluster assignment was unlikely affected by psychological support after the accident. In turn, the PTSD and, to a lesser extent, the PTG cluster were characterized by somatic health consequences of the accident, flashbacks, more caution during sport activity and subjective need for psychological support. Finally, the lacking non-mental explanatory variables strongly associated with the cluster assignment likely resulted in sub-optimal performance of machine learning classifiers employed for prediction of the cluster assignment and identification of accident victims at risk of mental health problems.

*Hanna and Katharina: your turn now:)*

Our study bears limitations. First, although the key phenomena phenomena in the newly identified mental clusters, post-traumatic growth and post-traumatic syndrome disorder, are clearly associated with trauma (8,9), we could not exclude the mental cluster assignment to be affected by other events than the sport accident. To account for that, follow-up studies including a control group are needed. Second, our variable set misses potentially important explanatory factors for the mental cluster classification such as length of hospital stay, rehabilitation need or ability to work. Finally, effects of the accident, injury severity, hospitalization or surgery on mental health after an accident may have been obscured in the cross-sectional cohort and needs to be investigated in a more defined collective e.g. of high risk sport accident victims or hospitalized individuals.

# Conclusion

*Your part again please*

# Acknowledgments

# Author’s contribution

# Conflict of interest

# Data and code availability

An R data (RData) file with anonymized patient data will be made available upon request to the corresponding author. The study analysis pipeline is available at <https://github.com/PiotrTymoszuk/mental_accident>.

# Tables

Table 1: Demographic and socioeconomic characteristic of the study cohort. Numeric variables are presented as medians with interquartile ranges (IQR). Categorical variables are presented as percentages and counts within the complete observation set.

| **Variable** | **Statistic** |
| --- | --- |
| Participants, n | 307 |
| Ward – survey time, days | 1300 [IQR: 800 - 1400], range: 390 - 1600 |
| Age, years | 51 [IQR: 33 - 60], range: 18 - 82 |
| Age classa | young: 20% (n = 61) middle: 66% (n = 202) elderly: 14% (n = 44) |
| Sex | female: 45% (n = 137) male: 55% (n = 170) |
| Residence in the Alps | 73% (n = 225) |
| Education | primary/apprenticeship: 16% (n = 49) secondary: 38% (n = 115) tertiary: 45% (n = 136) |
| Employment | employed: 68% (n = 210) unemployed: 3.6% (n = 11) student: 10% (n = 32) retired: 18% (n = 54) |
| Sport profession | 5.2% (n = 16) |
| Trauma-risk profession | 7.2% (n = 22) |
| Income/year | no income: 21% (n = 63) < 30000 EUR: 18% (n = 56) 30000 - 45000 EUR: 19% (n = 59) ≥ 45000 EUR: 42% (n = 129) |
| Smoking | 7.8% (n = 24) |
| Substance abuse (CAGE ≥2)b | 9.4% (n = 29) |
| Somatic illness | 15% (n = 47) |
| Mental illness | 5.2% (n = 16) |
| Prior traumatic event/DIA-X | 40% (n = 124) |
| ayoung: 16 - 30, middle: 31 - 65, elderly: ≥ 66 years | |
| bCAGE: Cut/Annoyed/Guilty/Eye substance abuse scale | |

Table 2: Characteristic of the sport accident, injury, psychological management and accident consequences. Numeric variables are presented as medians with interquartile ranges (IQR). Categorical variables are presented as percentages and counts within the complete observation set.

| **Variable** | **Statistic** |
| --- | --- |
| Prior sport accidents | 38% (n = 118) n = 307 |
| Sport typea | ski/snowboard: 64% (n = 197) sledding: 3.9% (n = 12) mountain: 14% (n = 42) biking: 16% (n = 48) other: 2.6% (n = 8) n = 307 |
| Alone during the accident | 32% (n = 97) n = 307 |
| Accident culprit | self: 77% (n = 237) non-self: 23% (n = 70) n = 307 |
| Injured persons | only self: 64% (n = 195) self and partner: 3.6% (n = 11) 3+ persons: 1.3% (n = 4) no information: 32% (n = 97) n = 307 |
| Rescue | self: 50% (n = 155) partner/third party: 21% (n = 63) rescue team: 29% (n = 89) n = 307 |
| Injury severity, AIS | 1: 37% (n = 108) 2: 35% (n = 103) 3+: 28% (n = 83) n = 294 |
| Hospitalizedb | 26% (n = 80) n = 307 |
| Surgery | 14% (n = 43) n = 307 |
| Psychological support | 9.1% (n = 28) n = 307 |
| Psychological support need | 7.5% (n = 23) n = 307 |
| Somatic accident aftermath | 37% (n = 115) n = 307 |
| Returned to same sport | 85% (n = 262) n = 307 |
| Caution post accident | no change: 35% (n = 106) more cautious: 65% (n = 199) less cautious: 0.65% (n = 2) n = 307 |
| Flashback frequency | none: 60% (n = 185) > 1/month: 18% (n = 54) > 1/year: 22% (n = 68) n = 307 |
| aski/snowdoard: alpine skiing, snowboarding and cross-country skiing sledding: sledding or bobsled mountain: hiking, climbing and skitouring biking: mountainbike, tour and road cycling | |
| bAIS: abbreviated injury scale | |

Table 3: Mental health characteristic of the study participants at survey completion. Numeric variables are presented as medians with interquartile ranges (IQR). Categorical variables are presented as percentages and counts within the complete observation set.

| **Variablea** | **Statistic** |
| --- | --- |
| Participants, n | 307 |
| GAD-7 score | 1 [IQR: 0 - 3], range: 0 - 15 |
| Anxiety symptoms (GAD-7 ≥10) | 2.3% (n = 7) |
| PHQ-9 score | 2 [IQR: 1 - 5], range: 0 - 16 |
| Depression symptoms (PHQ-9 ≥11) | 5.5% (n = 17) |
| PHQ-15 score | 2 [IQR: 1 - 4], range: 0 - 23 |
| Somatization symptoms (PHQ-15 ≥11) | 4.9% (n = 15) |
| EUROHIS-QOL 8 score | 1.6 [IQR: 1.4 - 2], range: 1 - 4 |
| SOC-9L score | 19 [IQR: 16 - 25], range: 10 - 49 |
| RS13 score | 78 [IQR: 70 - 85], range: 15 - 91 |
| RS13 coping class | low: 18% (n = 56) moderate: 14% (n = 42) high: 68% (n = 209) |
| PTGI score | 32 [IQR: 16 - 48], range: 0 - 100 |
| PCL-5 DSM-5 score | 3 [IQR: 1 - 7], range: 0 - 44 |
| PTSD+ (at least one domain) | 19% (n = 58) |
| aGAD-7: 7-item general anxiety disorder scale; PHQ: patient health questionnaire; EUROHIS-QOL 8: 8-item EUROHIS project quality of life scale; SOC-9L: Leipzig 9-item sense of coherence questionnaire; RS13: 13-item resilience scale; PCL-5 DSM-5: PTSD checklist for DSM-5; PTGI: post-traumatic growth inventory; PTSD: post-traumatic syndrome disorder | |

# Figures

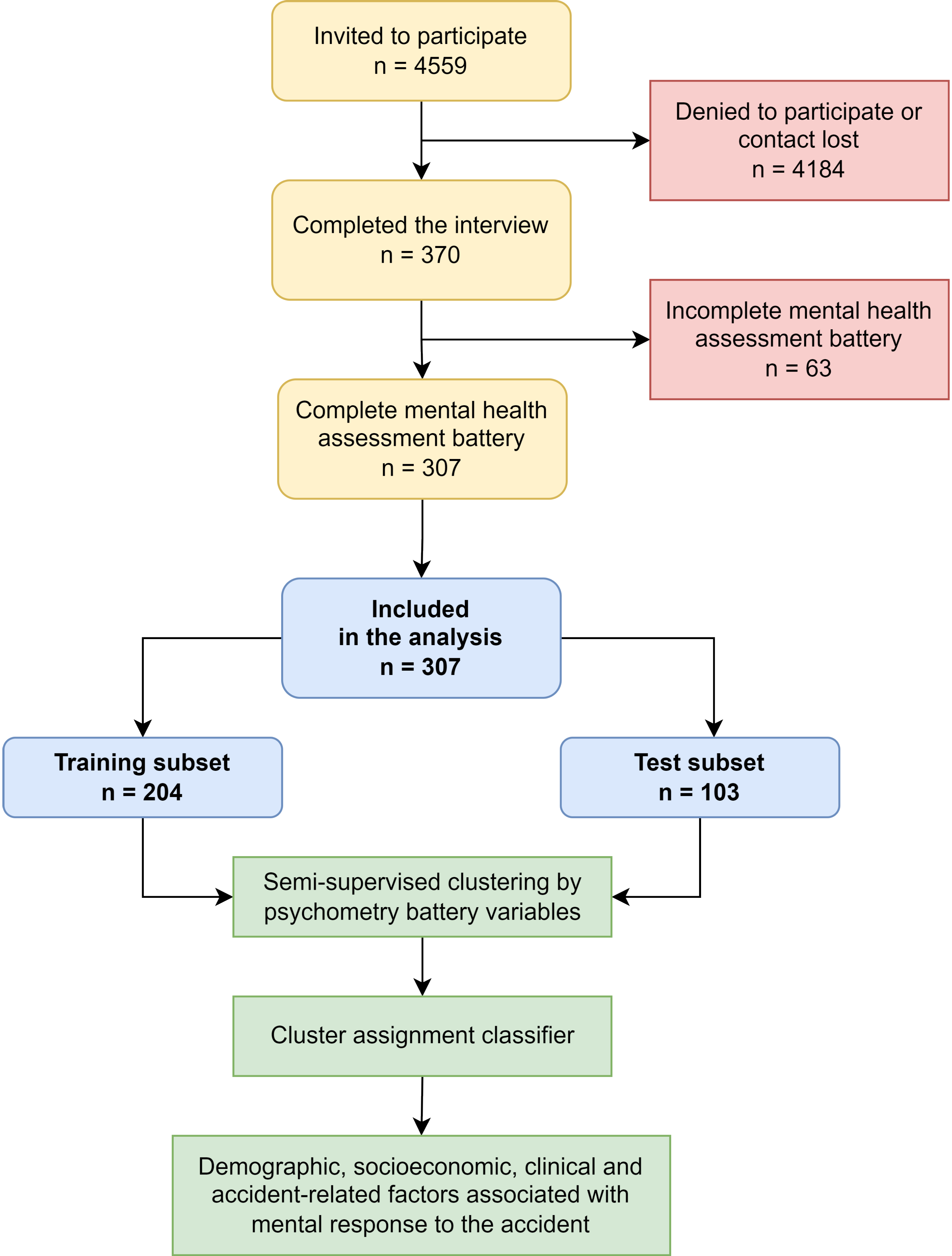


Figure 1: Flow diagram of the analysis inclusion process and the analysis strategy.

**Figure 1. Flow diagram of the analysis inclusion process and the analysis strategy.**

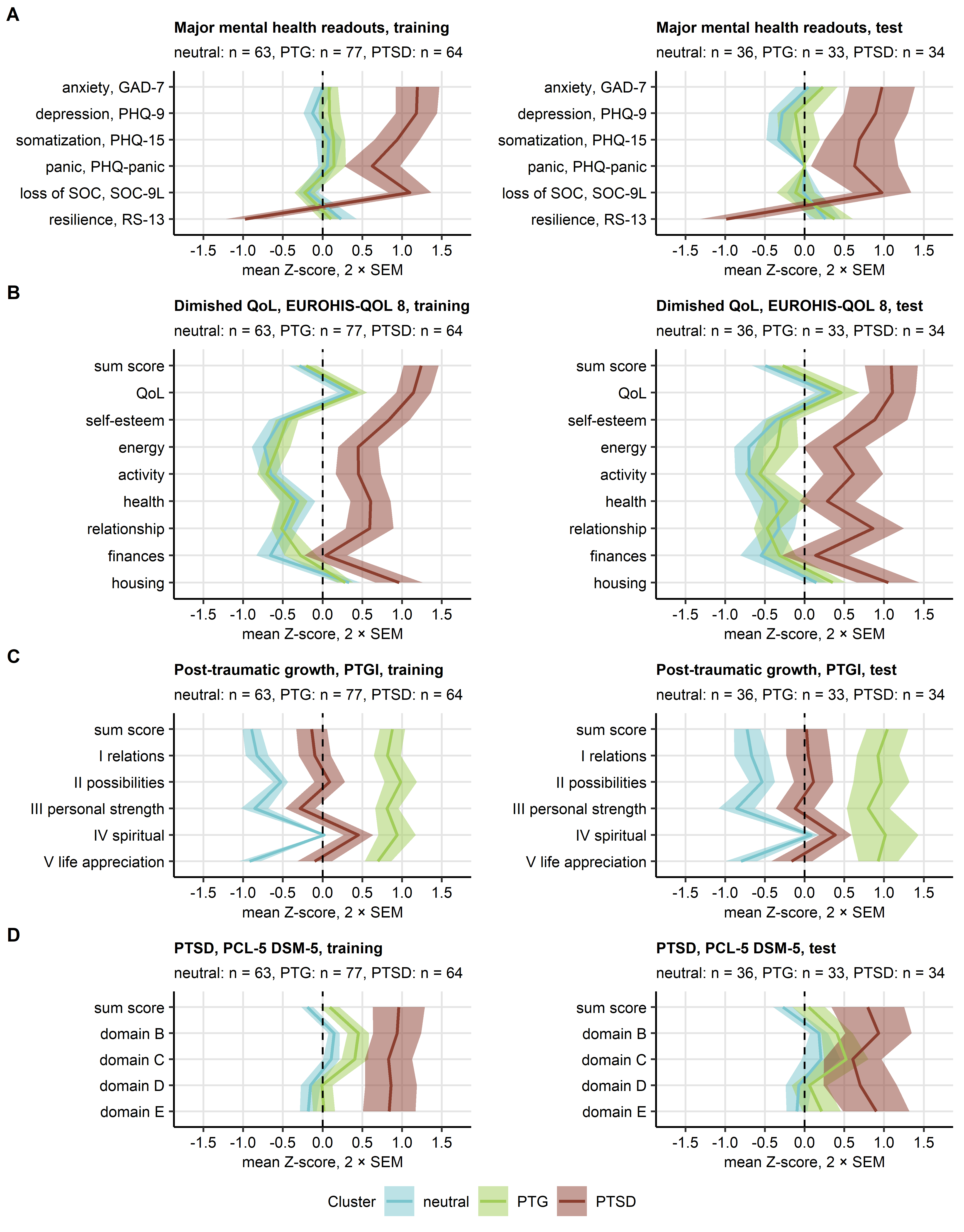


Figure 2: Scores of psychometry readouts in the mental clusters.

**Figure 2. Scores of psychometry readouts in the mental clusters.**

*The mental clusters were defined in respect to psychometric scoring in the training subset of the study cohort by PAM (partition around medoids) with cosine distance between the observations. Assignment of the test subset observations to the mental clusters was done with the inverse distance weighted 7-nearest neighbors classifier. Three mental clusters were identified: neutral, PTG (post-traumatic growth) and PTSD (post-traumatic syndrome disorder), All clustering parameters differed significantly between the clusters as assessed by Kruskal-Wallis test with effect size statistic. Mean normalized scores of major mental health readouts (A), diminished quality of life (B), post-traumatic growth (C) and post-traumatic symptom disorder (D) in the mental clusters of the training and test subsets of the study cohort are depicted as solid lines. Tinted regions represent 2 SEM (standard error of the mean). Numbers of observations in the clusters are displayed in the plot captions. GAD-7: 7-item general anxiety disorder scale; PHQ: patient health questionnaire; EUROHIS-QOL 8: 8-item EUROHIS project quality of life scale; SOC-9L: Leipzig 9-item sense of coherence questionnaire; RS13: 13-item resilience scale; PCL-5 DSM-5: PTSD checklist for DSM-5; PTGI: post-traumatic growth inventory; PTSD: post-traumatic syndrome disorder.*

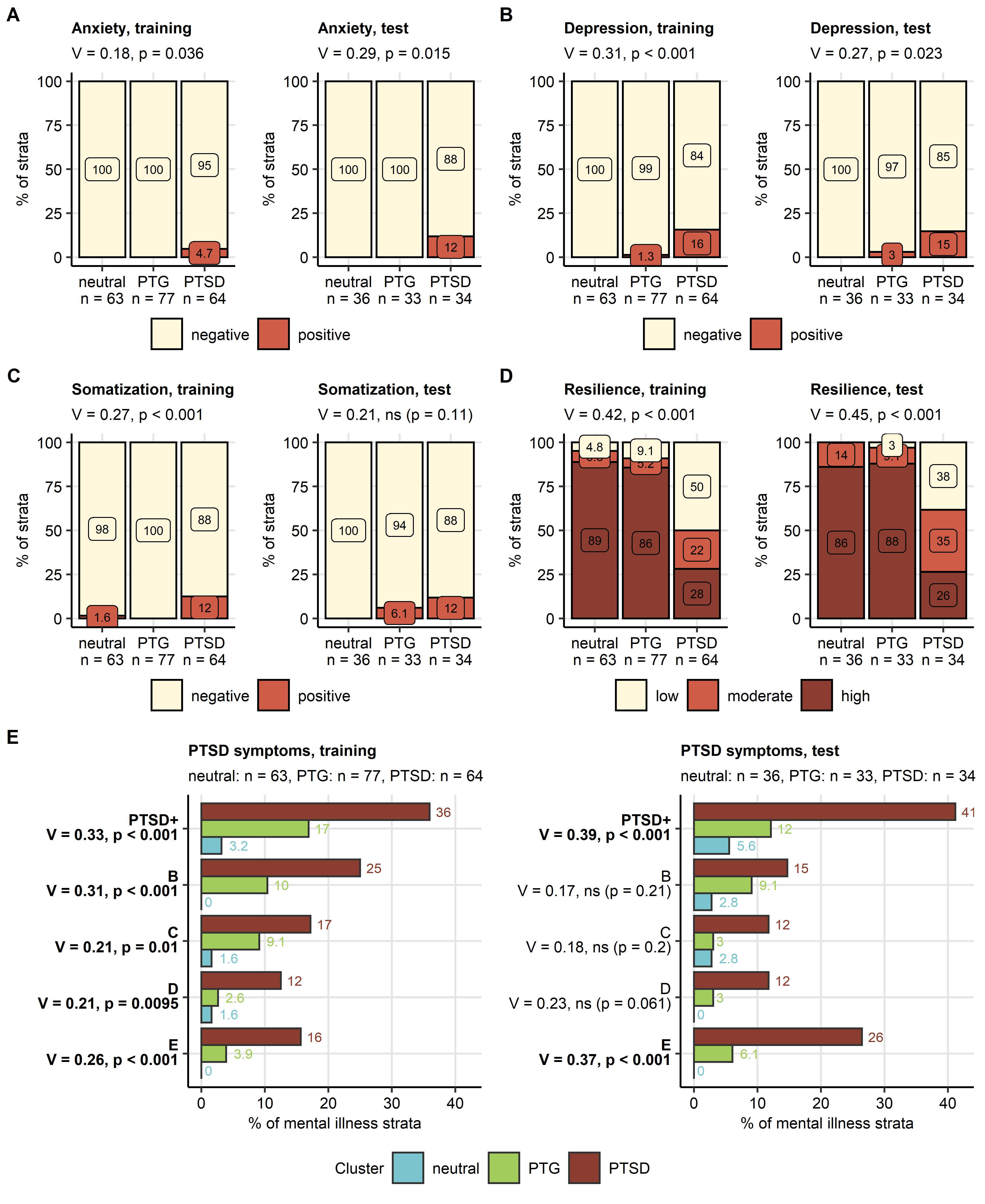


Figure 3: Signs of mental health problems in the mental clusters.

**Figure 3. Signs of mental health problems in the mental clusters.**

*Frequencies of symptoms of anxiety (A, GAD-7 ≥10), depression (B, PHQ-9 ≥ 11) and somatization (C, PHQ-15 ≥ 11), distribution of resilient coping classes (D, RS13) and symptoms of post-traumatic syndrome disorder (E, PTSD, PCL-5 DSM-5, positivity for PTSD domains and frequency of participants positive for at least one PTSD domain [PTSD+]) in the mental clusters. Statistical significance was determined by test with Cramer V effect size statistic. Fractions of symptom-positive and -negative participants in the mental clusters in the training and test subset of the study cohort are presented in stack and bar plots. Effect sizes and p-values are shown in the plot captions or Y axes. Numbers of observations in the clusters are shown in the X axes or plot captions. Significant effects in (E) are highlighted in bold. GAD-7: 7-item general anxiety disorder scale; PHQ: patient health questionnaire; RS13: 13-item resilience scale; PCL-5 DSM-5: PTSD checklist for DSM-5; PTSD: post-traumatic syndrome disorder.*

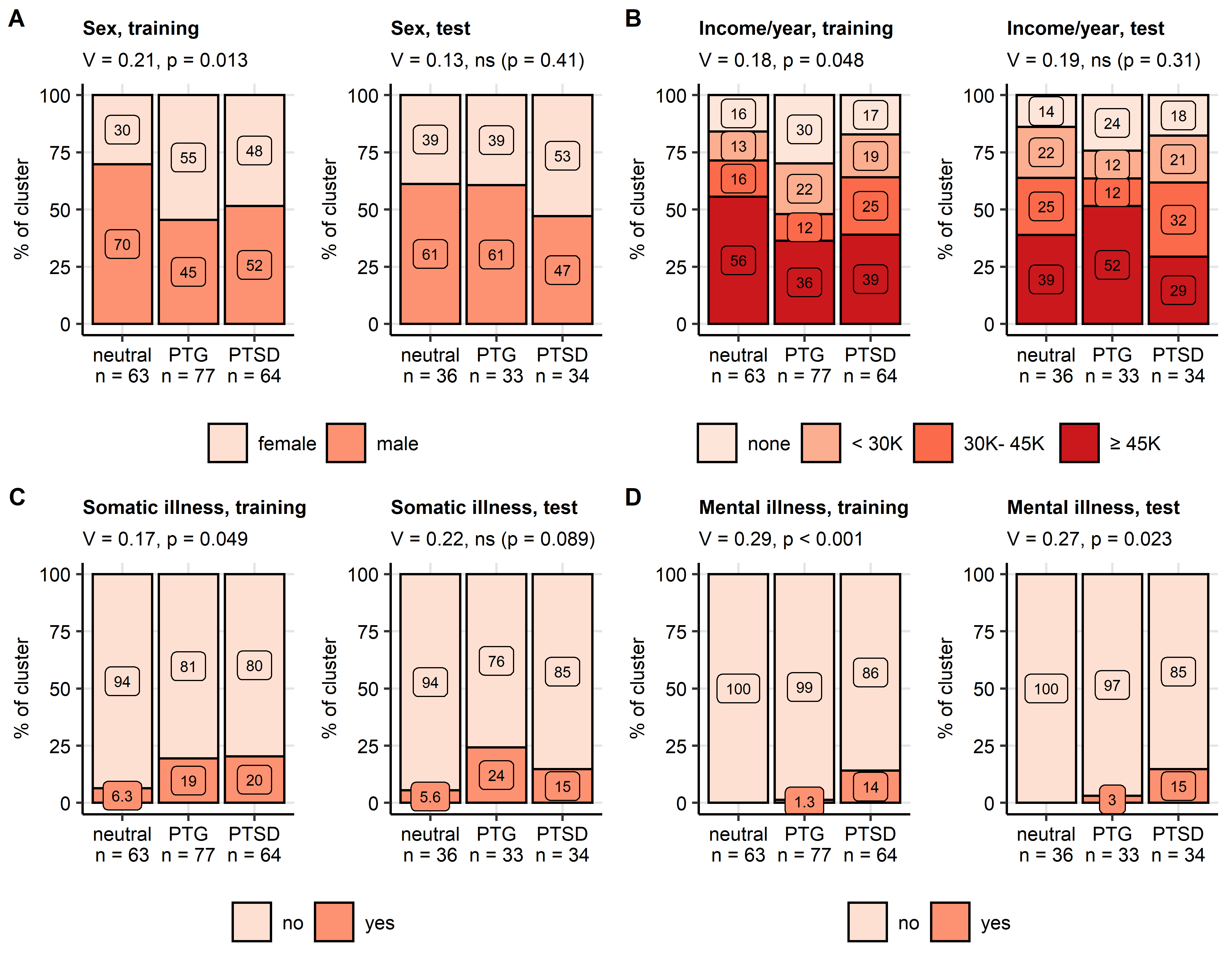


Figure 4: Demographic and clinical factors differing between the mental clusters of accident victims.

**Figure 4. Demographic and clinical factors differing between the mental clusters of accident victims.**

*Distribution of gender (A), annual household income (B), and frequencies of self-reported somatic (C) and mental illness (D) in the mental clusters. Statistical significance was determined by test with Cramer V effect size statistic. Percentages in the mental clusters in the training and test subset of the study cohort are presented in stack plots. Effect sizes and p-values are displayed in the plot captions. Numbers of observations in the clusters are presented in the X axes.*

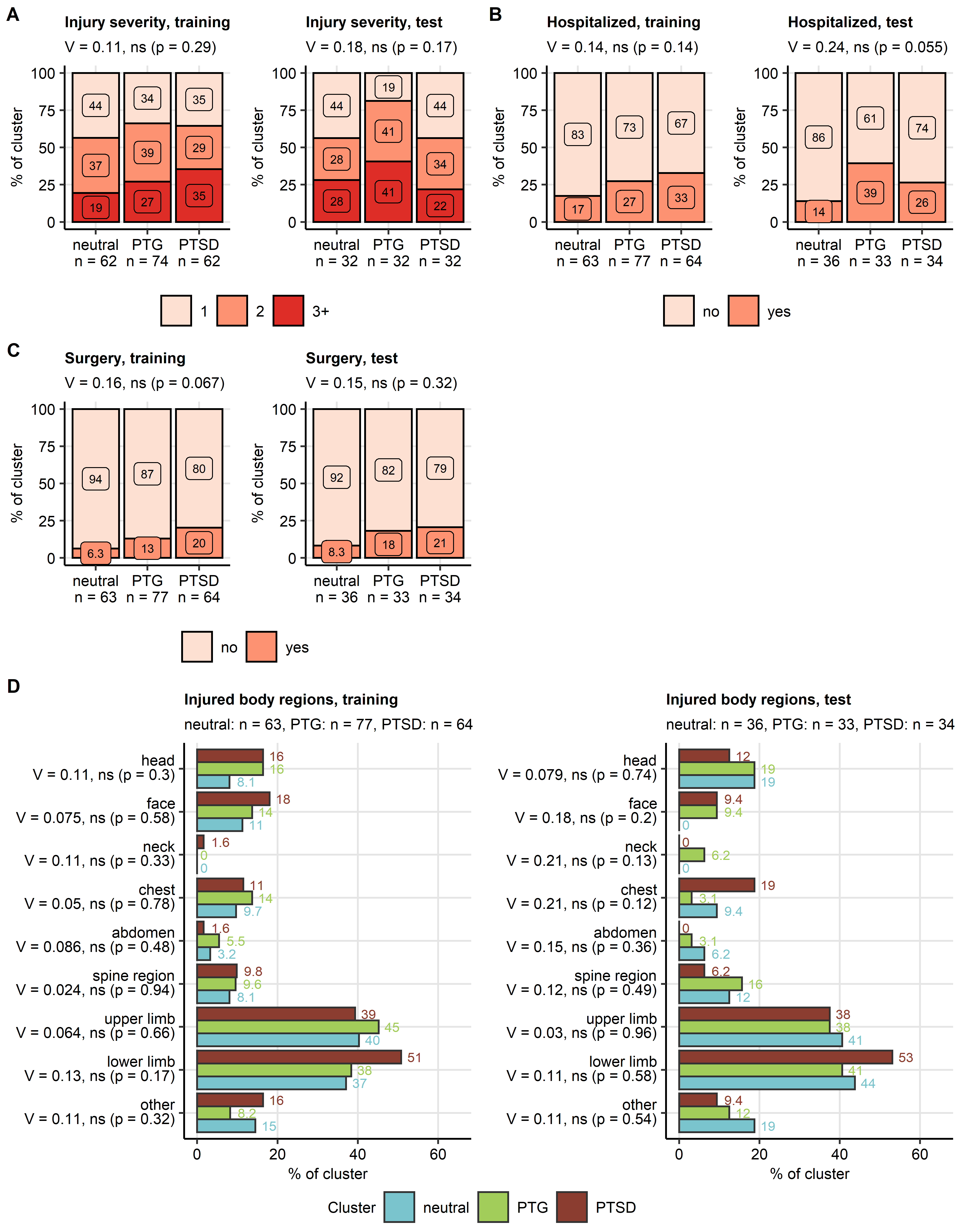


Figure 5: Injury severity and location in the mental clusters.

**Figure 5. Injury severity and location in the mental clusters.**

*Distribution of injury severity grades (A, AIS: abbreviate injury scale), hospitalization (B) and surgery rates (C), and injuries of the body parts (D) in the mental clusters. Statistical significance was determined by test with Cramer V effect size statistic. Percentages in the mental clusters in the training and test subset of the study cohort are presented in stack and bar plots. Effect sizes and p-values are displayed in the plot captions or Y axes. Numbers of observations in the clusters are presented in the X axes or plot captions.*

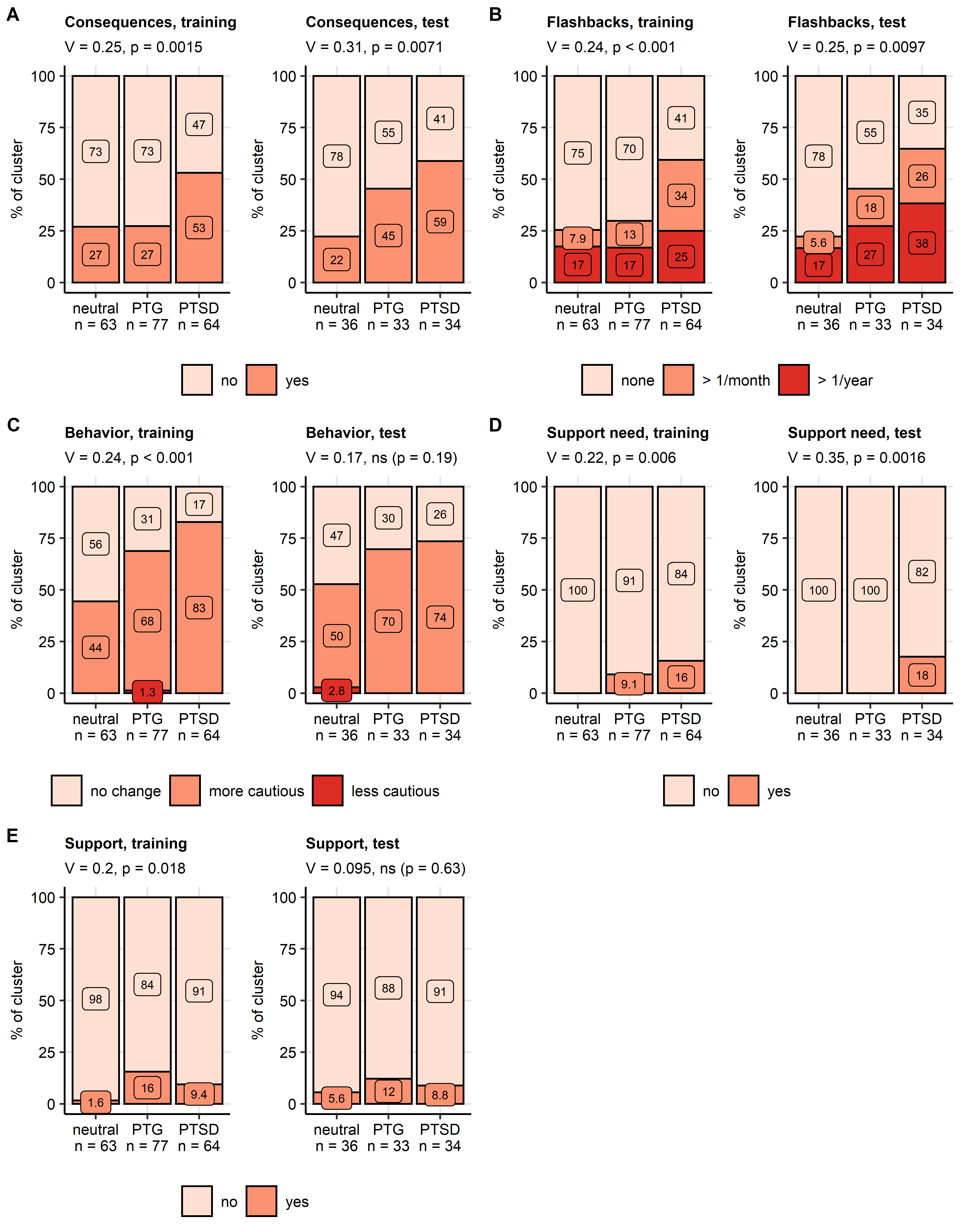


Figure 6: Consequences of the accident in the mental clusters.

**Figure 6. Consequences of the accident in the mental clusters.**

*Frequencies of self-reported somatic health consequences of the accident (A), flashbacks during sport (B), more or less cautious behavior during sport activity (C), self-reported psychological support need (D) and of a received psychological support (E) in the mental clusters. Statistical significance was determined by test with Cramer V effect size statistic. Percentages in the mental clusters in the training and test subset of the study cohort are presented in stack plots. Effect sizes and p-values are displayed in the plot captions. Numbers of observations in the clusters are presented in the X axes.*

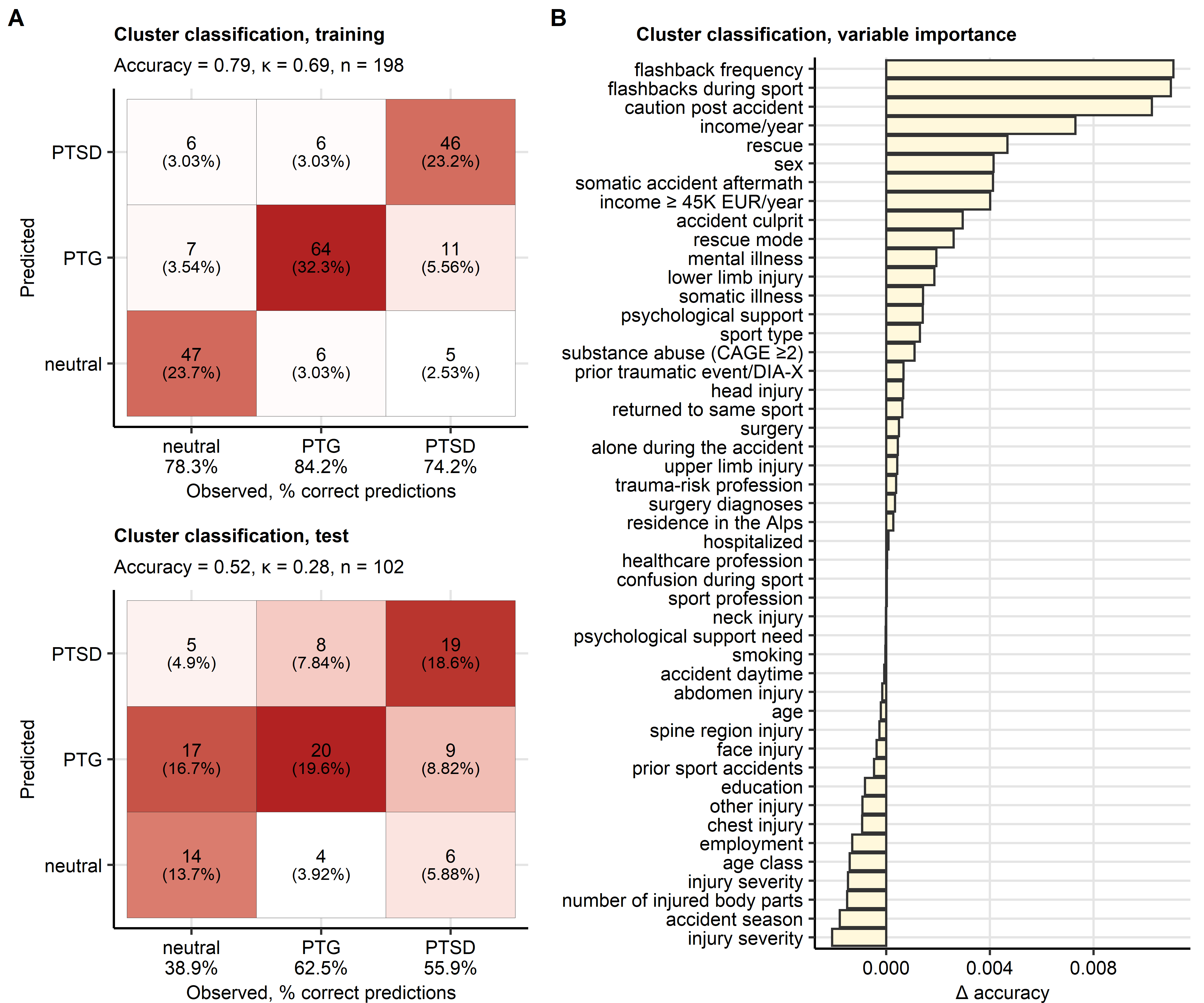


Figure 7: Assignment of accident victims to the mental clusters based on demographic, socioeconomic, clinical and accident-related factors.

**Figure 7. Assignment of accident victims to the mental clusters based on demographic, socioeconomic, clinical and accident-related factors.**

*A conditional random forest classifier for the mental cluster assignment based on demographic, socioeconomic, clinical and accident-related factors was trained in the training subset of the study cohort.*

*(A) Prediction accuracy of the classifier in the training and test subset of the study cohort presented in a heat map plot of the confusion matrices. Color codes for number of cases. The overall prediction accuracy, and observation numbers are displayed in the plot caption. Prediction accuracy for particular clusters is shown in the X axis.*

*(B) Conditional importance of the explanatory variables for prediction of the mental cluster assignment expressed as accuracy loss ( accuracy).*

# References

1. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Archives of Internal Medicine* (2006) 166:1092–1097. doi: [10.1001/ARCHINTE.166.10.1092](https://doi.org/10.1001/ARCHINTE.166.10.1092)

2. Löwe B, Spitzer RL, Zipfel S, Herzog W. Auflage Manual 17.07. (2002).

3. Gräfe K, Zipfel S, Herzog W, Löwe B. Screening psychischer störungen mit dem "Gesundheitsfragebogen für Patienten (PHQ-D)". Ergebnisse der Deutschen validierungsstudie. *Diagnostica* (2004) 50:171–181. doi: [10.1026/0012-1924.50.4.171](https://doi.org/10.1026/0012-1924.50.4.171)

4. Kroenke K, Spitzer RL, Williams JBW. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosomatic medicine* (2002) 64:258–266. doi: [10.1097/00006842-200203000-00008](https://doi.org/10.1097/00006842-200203000-00008)

5. Leppert K, Koch B, Brähler E, Und BS-KD, 2008 U. Die Resilienzskala (RS)–Überprüfung der Langform RS-25 und einer Kurzform RS-13. *Klinische Diagnostik und Evaluation* (2008) 1:226–243. <https://www.academia.edu/download/44388154/A_406.pdf>

6. Schumacher J, Wilz G, Gunzelmann T, Brähler E. Die sense of coherence scale von antonovsky: Teststatische überprüfung in einer repräsentativen bevölkerungsstichprobe und konstruktion einer kurzskala. *PPmP Psychotherapie Psychosomatik Medizinische Psychologie* (2000) 50:472–482. doi: [10.1055/s-2000-9207](https://doi.org/10.1055/s-2000-9207)

7. Schmidt S, Mühlan H, Power M. The EUROHIS-QOL 8-item index: psychometric results of a cross-cultural field study. *European Journal of Public Health* (2006) 16:420–428. doi: [10.1093/EURPUB/CKI155](https://doi.org/10.1093/EURPUB/CKI155)

8. Tedeschi RG, Calhoun LG. The Posttraumatic Growth Inventory: measuring the positive legacy of trauma. *Journal of traumatic stress* (1996) 9:455–471. doi: [10.1007/BF02103658](https://doi.org/10.1007/BF02103658)

9. Bovin MJ, Marx BP, Weathers FW, Gallagher MW, Rodriguez P, Schnurr PP, Keane TM. Psychometric properties of the PTSD Checklist for Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition (PCL-5) in veterans. *Psychological assessment* (2016) 28:1379–1391. doi: [10.1037/PAS0000254](https://doi.org/10.1037/PAS0000254)

10. McDonald RP. *Test theory: A unified treatment*. 1st Editio. New Yor: Psychology Press (1999). doi: [10.4324/9781410601087](https://doi.org/10.4324/9781410601087)

11. Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): A meta-analysis. *CMAJ* (2012) 184:E191. doi: [10.1503/CMAJ.110829/-/DC1](https://doi.org/10.1503/CMAJ.110829/-/DC1)

12. Maercker A, Bromberger F. Checklisten und Fragebogen zur Erfassung traumatischer Ereignisse in deutscher Sprache. *Trierer Psychologische Berichte* (2005) 32:

13. O’Brien CP. The CAGE Questionnaire for Detection of Alcoholism. *JAMA* (2008) 300:2054–2056. doi: [10.1001/JAMA.2008.570](https://doi.org/10.1001/JAMA.2008.570)

14. Gennarelli TA, Wodzin E. AIS 2005: A contemporary injury scale. *Injury* (2006) 37:1083–1091. doi: [10.1016/j.injury.2006.07.009](https://doi.org/10.1016/j.injury.2006.07.009)

15. Kassambara A, Mundt F. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. (2020) <https://cran.r-project.org/web/packages/factoextra/index.html>

16. Kassambara A. rstatix: Pipe-Friendly Framework for Basic Statistical Tests. (2021) <https://cran.r-project.org/package=rstatix>

17. Drost H-G. Philentropy: Information Theory and Distance Quantification with R. *Journal of Open Source Software* (2018) 3:765. doi: [10.21105/joss.00765](https://doi.org/10.21105/joss.00765)

18. Schubert E, Rousseeuw PJ. Faster k-Medoids Clustering: Improving the PAM, CLARA, and CLARANS Algorithms. *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*. Springer (2019). p. 171–187 doi: [10.1007/978-3-030-32047-8\_16](https://doi.org/10.1007/978-3-030-32047-8_16)

19. Lange T, Roth V, Braun ML, Buhmann JM. Stability-based validation of clustering solutions. *Neural Computation* (2004) 16:1299–1323. doi: [10.1162/089976604773717621](https://doi.org/10.1162/089976604773717621)

20. McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. (2018) <https://arxiv.org/abs/1802.03426v3 http://arxiv.org/abs/1802.03426>

21. Konopka T. umap: Uniform Manifold Approximation and Projection. (2022) <https://cran.r-project.org/web/packages/umap/index.html>

22. Jouanne-Diedrich H von. OneR: One Rule Machine Learning Classification Algorithm with Enhancements. (2017) <https://cran.r-project.org/web/packages/OneR/index.html>

23. Holte RC. Very Simple Classification Rules Perform Well on Most Commonly Used Datasets. *Machine Learning* (1993) 11:63–90. doi: [10.1023/A:1022631118932/METRICS](https://doi.org/10.1023/A:1022631118932/METRICS)

24. Hothorn T, Hornik K, Zeileis A. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* (2006) 15:651–674. doi: [10.1198/106186006X133933](https://doi.org/10.1198/106186006X133933)

25. Strobl C, Boulesteix AL, Zeileis A, Hothorn T. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics* (2007) 8:1–21. doi: [10.1186/1471-2105-8-25/FIGURES/11](https://doi.org/10.1186/1471-2105-8-25/FIGURES/11)

26. Hothorn T, Hornik K, Strobl C, Zeileis A. party: A Laboratory for Recursive Partytioning. (2022) <https://cran.r-project.org/web/packages/party/index.html>

27. Kuhn M. Building predictive models in R using the caret package. *Journal of Statistical Software* (2008) 28:1–26. doi: [10.18637/jss.v028.i05](https://doi.org/10.18637/jss.v028.i05)

28. Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A. Conditional variable importance for random forests. *BMC Bioinformatics* (2008) 9:1–11. doi: [10.1186/1471-2105-9-307/FIGURES/4](https://doi.org/10.1186/1471-2105-9-307/FIGURES/4)

29. Statistik Austria. Rauchen - STATISTIK AUSTRIA - Die Informationsmanager. (2022) <https://www.statistik.at/statistiken/bevoelkerung-und-soziales/gesundheit/gesundheitsverhalten/rauchen> [Accessed February 24, 2023]

30. Statistik Austria. Gesundheitszustand selbstberichtet - STATISTIK AUSTRIA - Die Informationsmanager. (2022) <https://www.statistik.at/statistiken/bevoelkerung-und-soziales/gesundheit/gesundheitszustand/gesundheitszustand-selbstberichtet> [Accessed February 24, 2023]