Stability analysis of clustering of Norris' visual analogue scale: applying the consensus clustering approach

Guan, Z.; Chen, X.G.; Hay, J.; Gerven, J. van; Burggraaf, J.; Kam, M. de

Quality Improvement Study

Medicine®

OPEN

Stability analysis of clustering of Norris' visual analogue scale

Applying the consensus clustering approach

Zheng Guan, MSc[a,b], X. Gregory Chen, PhD[b], Justin Hay, PhD[a], Joop van Gerven, PhD[a,b], Jacobus Burggraaf, PhD[a,b], Marieke de Kam, MSc[a,*]

Abstract

Visual analogue scales are widely used to measure subjective responses. Norris' 16 visual analogue scales (N_VAS) measure subjective feelings of alertness and mood. Up to now, different scientists have clustered items of N_VAS into different ways and Bond and Lader's way has been the most frequently used in clinical research. However, there are concerns about the stability of this clustering over different subject samples and different drug classes. The aim of this study was to test whether Bond and Lader's clustering was stable in terms of subject samples and drug effects. Alternative clustering of N_VAS was tested.

Data from studies with 3 types of drugs: cannabinoid receptor agonist (delta-9-tetrahydrocannabinol [THC]), muscarinic antagonist (scopolamine), and benzodiazepines (midazolam and lorazepam), collected between 2005 and 2012, were used for this analysis. Exploratory factor analysis (EFA) was used to test the clustering algorithm of Bond and Lader. Consensus clustering was performed to test the stability of clustering results over samples and over different drug types. Stability analysis was performed using a three-cluster assumption, and then on other alternative assumptions.

Heat maps of the consensus matrix (CM) and density plots showed instability of the three-cluster hypothesis and suggested instability over the 3 drug classes. Two- and four-cluster hypothesis were also tested. Heat maps of the CM and density plots suggested that the two-cluster assumption was superior.

In summary, the two-cluster assumption leads to a provably stable outcome over samples and the 3 drug types based on the data used.

Abbreviations: BEN = benzodiazepine, CHDR = Centre of Human Drug Research, CM = consensus matrix, CNS = central nervous system, EFA = exploratory factor analysis, N_VAS = Norris' visual analogue scales, SCO = scopolamine, THC = delta-9-tetrahydrocannabinol.

Keywords: cluster analysis, consensus clustering, Norris' visual analogue scale, stability analysis

1. Introduction

Central nervous system (CNS) drugs suffer from a low clinical development success rate and they rank third to last according to Kola and Landis.[1] Other than pharmacokinetic issues and lack of predictivity of preclinical tests,[2] there are also difficulties in measuring the efficacy of CNS drugs. In clinical trials important outcome measures are the subjective feelings of the subject rather than, for example, objective biomarkers. Therefore, the correct interpretation and analysis of subjective rating scores are important to understanding drug effects.

Norris' visual analogue scales (N_VASs) are commonly used for the rating subjective feelings in clinical research. The scales were originally designed in the 1920s, and discovered to be sensitive to detect sedative drug effects in normal subjects.[3–7] A set of 16 scales, which were intuitively grouped into 4 classes (mental sedation, physical sedation, tranquilization, and other), were used to measure drug effects by Malpas and Norris.[7,8] Since then, VAS has been widely used in clinical studies in the fields of pain and psychology and improvement of methodology has been kept developing over time.[9–15] However, the large number of scales and dependence among the scores have caused difficulty in processing and interpretation of the scores. In order to assess the different drug effects, many recursive modelling techniques are available, in which the scores of 16 scales are placed as the response variables. These models can be set up and estimated either separately for each scale or for all scales simultaneously. In the former case, it is hard to interpret results with connections from 16 separated models, due to the dependence of 16 scales scores. In the latter case, it raises a

Editor: Elena M. Varoni.

The authors have no conflicts of interest to disclose.

Supplemental Digital Content is available for this article.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

[a] Centre for Human Drug Research, [b] Leiden University Medical Center, The Netherlands.

[*] Correspondence: Marieke de Kam, Zernikedreef 8, 2333 CL, Leiden, The Netherlands (e-mail: Mdekam@chdr.nl).

How to cite this article: Guan Z, Chen XG, Hay J, van Gerven J, Burggraaf J, de Kam M. Stability analysis of clustering of Norris' visual analogue scale: Applying the consensus clustering approach. Medicine 2021;100:17(e25363).

Received: 14 January 2020 / Received in final form: 25 January 2021 / Accepted: 11 March 2021

http://dx.doi.org/10.1097/MD.0000000000025363

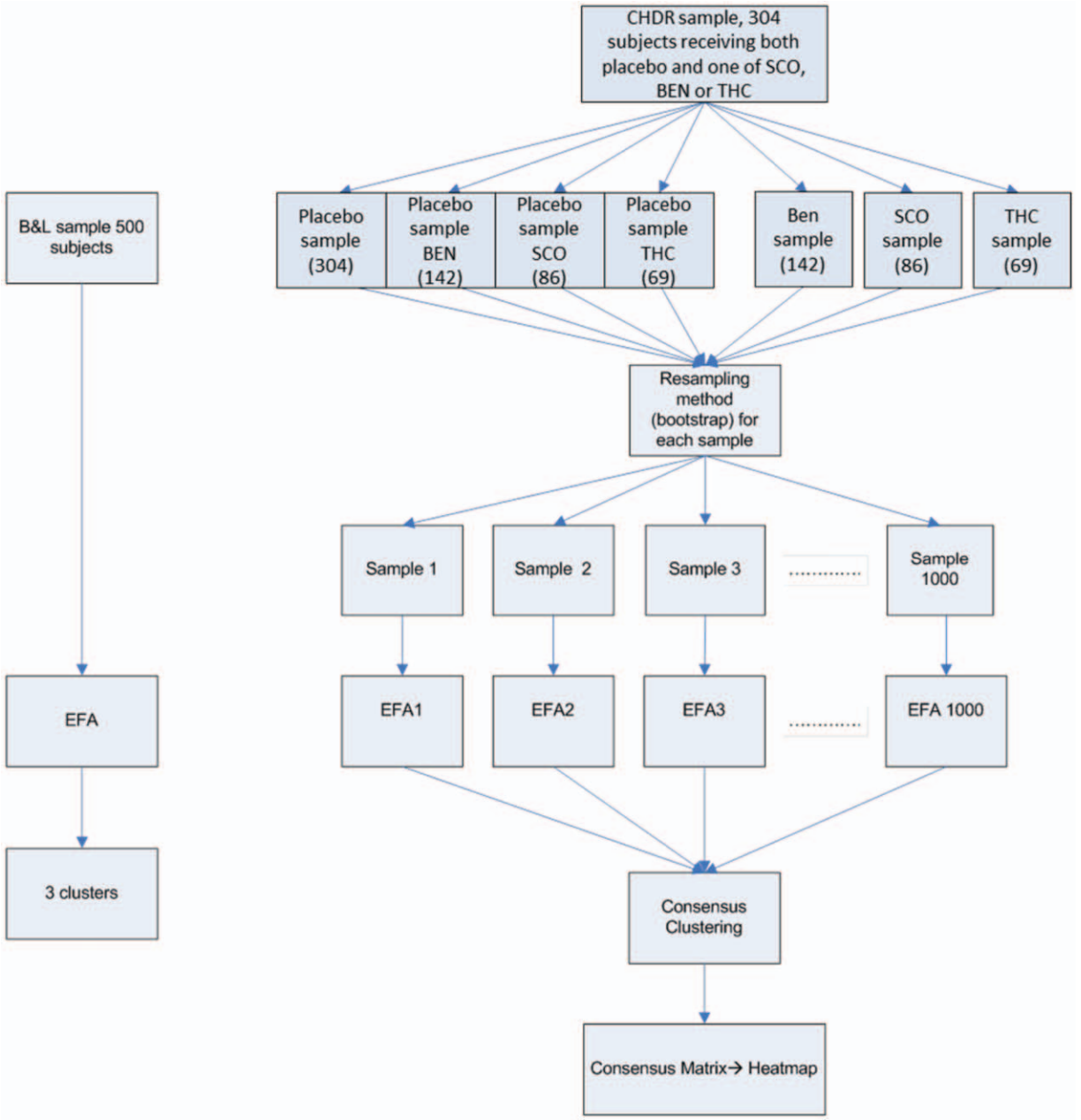

Figure 1. Research scheme: left is Bond and Lader's approach; right is the scheme of current study. BEN = benzodiazepine; SCO = scopolamine, THC = delta-9-tetrahydrocannabinol.

great computational challenge to build a linear (mixed) model with dependent multivariate response variables.

For the sake of simplicity and efficiency, reducing the dimensions of the scale scores is essential.

Typically, all 16 N_VASs are assigned to a small number of clusters that are mutually exclusive, so that scales in the same group are similar to each other yet distinct from ones in other groups. Then a representative score is calculated for each cluster, usually by taking the average of the scales in a cluster. Instead of theoretically clustering these 16 scales by their literal meaning, it is preferable to take a data-driven approach for objectiveness. Bond and Lader developed N_VAS into its present shape and clusters.[16] As shown on left side of Figure 1, Bond and Lader collected data from 500 subjects from technical colleges, universities, and hospitals who were not taking medication. One measurement on each of the 16 scales, was taken from each subject. Then that dataset was directly analyzed using factor analysis. The outcome was a clustering in 3 groups, as in Table 1.[16] This result is widely used by clinical researchers, yet the stability of this clustering should be validated over samples and for assessing the effects of different drugs.

After Norris, Herbert applied the N_VASs measuring subject feelings before and after sleep and also clustered the 16 items into 2 using principle-component analysis, but the results were limited in terms of small subject number and only 1 medication type was investigated.[17]

Table 1

Bond and Lader's 3 clusters of Norris' scales.

Categories	Original VAS scale	Item	Indicator in heat map
Alertness	Alert – Drowsy	VASBL01	1
	Strong – Feeble	VASBL03	3
	Muzzy – Clear-headed	VASBL04	4
	Well-coordinated – Clumsy	VASBL05	5
	Lethargic – Energetic	VASBL06	6
	Mentally slow – Quick-witted	VASBL09	9
	Attentive – Dreamy	VASBL11	11
	Incompetent – Proficient	VASBL12	12
	Interested – Bored	VASBL15	15
Mood	Contented – Discontented	VASBL07	7
	Troubled – Tranquil	VASBL08	8
	Happy – Sad	VASBL13	13
	Antagonistic – Amicable	VASBL14	14
	Withdrawn – Gregarious	VASBL16	16
Calm	Calm – Excited	VASBL02	2
	Tense – Relaxed	VASBL10	10

At the Centre of Human Drug Research (CHDR), many studies have been conducted with CNS drugs, using the 16 N_VAS scales to evaluate drug effects.[18–24] These data create the opportunity to test the robustness of the Bond and Lader clustering over samples and different drug classes using factor analysis approach. Exploratory factor analysis was the main factor analysis used in this study. Factor analysis covers a range of multivariate methods used to explain how underlying factors influence a set of observed variables. Exploratory factor analysis (EFA) is used to identify these underlying factors and has become an integral statistical method in the social, health, biological, and physical sciences.[25] It is also applied in clinical research such as in pulmonary disease, antidepressant drug study, and lab examination data analysis.[26–28]

2. Methods

2.1. Norris VAS

A computer-based method was used for collection of the N_VAS data. Sixteen lines, corresponding to 16 scales, were presented on a computer screen to subjects, and the lines consist of a line segments with opposite terms on both the extremities of the line. The scoring system involved 16 scales with continuous scaling from 0 to 100, and in each scale, 50 was presumed to be the score of no effect or feeling normal. Dutch language versions of the scales were used. Subjects were required to put a mark on the line that best represented their subjective state according to the tested condition, and the length between the left head of the line and the marked point is defined as the score of the scale. All participants were trained in-house before the studies.

2.2. Data collection

Data collected in this research groups the observations of studies performed at the CHDR between 2005 and 2012. In total, data of 304 healthy male subjects were included. All subjects received placebo and 297 out of 304 received a second treatment in a crossover manner: 142 benzodiazepines (midazolam, alprazolam, and lorazepam), 86 muscarinic receptor antagonist (scopolamine), and 69 a cannabinoid receptor agonist (delta-9-tetrahydrocannabinol [THC]), N_VAS was conducted once or twice before and usually 8 to 9 times after drug administration. Although all 304 subjects had repeated measures of the 16 VAS scales, the average over the time for each individual subject was used as a single observation.

These 3 classes of drugs were selected due to their different mechanisms of action and their common usage in pharmacodynamic challenge tests. Details of the clinical trial designs have been published previously.[21,22,29–33] All studies were performed at the Centre for Human Drug Research in Leiden, the Netherlands, and approved by the local ethics committee of Leiden University Medical Center (Leiden).[21,22,29–33] The subjects consented in writing to the study after full explanation of what was involved. All trials were double-blind, crossover, or partial crossover.

2.3. Analysis in R

The statistical software R version 3.0.2 (R Foundation for Statistical Computing, Vienna, Austria, 2010) was used for input file preparation, processing, and visualizing of the analysis.

2.4. Research scheme

The datasets used were: the total placebo set, 3 subsets of the placebo data based on the other drug given, the total set with drug treated data, the subset with scopolamine data, the subset with benzodiazepine data, and the subset with THC data. As shown in Figure 1 for all the mentioned datasets and subsets, bootstrapping was used to generate 1000 samples. By using a resampling approach (bootstrap), it was possible to generate multiple sub-datasets from the original dataset. In each sub-dataset, EFA was applied, at first with the fixed assumption of 3 clusters. The outcomes of the 1000 factor analyses were used to generate a consensus cluster that was visualized by heat maps. Different cluster numbers were further tested for the drug treated samples.

2.5. Exploratory factor analysis

EFA was used as the main step in the clustering algorithm. It is an analytical paradigm for modelling the relationships among sets of observed variable (i.e., manifest variables) and latent factors. The exploratory branch of this approach is designed for the situation that the prior knowledge of these relationships is unknown.[34] The latent factors are usually assumed to be mutually independent. EFA can be also seen as a dimension-reduction technique. As the number of latent factors is usually presumed to be much smaller than the number of observed variables, the leading latent factors (sorted by corresponding eigenvalues) are considered to retain major information of the data. The number of latent factors can be decided by performing a likelihood ratio test, or, in case of Bond and Lader, by setting a threshold on the corresponding eigenvalues. The reason for choosing an EFA was that this study was aligned with Bond and Lader's previous work[16] and as such the validation of the stability would not be impacted by a different clustering approach.

The performance of EFA was based on a publication of Hardle and Simar.[34] For each of the original 16 scales used, the factor loading reflects the magnitude of its relation with a latent factor. The 16 scales were assigned to the factor with the highest loading, and the factors become the clusters. The items in the same cluster are mainly influenced by the same factor, and the difference factors are uncorrelated under the EFA paradigm.

2.6. Consensus cluster

Consensus clustering was performed to test the stability of the clustering result over the 1000 bootstrapped samples and over different drug types.[35] It is a model-independent resampling-based methodology, so that it is not influenced by different approaches of clustering. The core of consensus clustering is a method to represent the consensus across multiple runs of a cluster algorithm. The methodology used was based on those of Monti and colleagues.[35]

The typical result of consensus clustering method is a consensus matrix (CM) and heat map visualization. The CM is an $N \times N$ matrix, with the original items in a pre-specified order or ordered automatically by the outcome of the clustering. Each element in the CM represents a ratio between the count of times that the 2 items were clustered together and the total number of times that 2 items are both in the resampled dataset. Every element of CM ranges between 0 and 1. A heat map provides visualization. In this study, the color "blue" represents 1 and "white" represents 0. If the sequence of the original items is sorted based on the clustering result, it is expected that a perfect clustering result is translated into some blue blocks in the diagonal of the heat map and purely white in other positions. The further the departure from this image, the poorer the quality of the clustering result. Additionally, each block corresponds to a different cluster, and the number of blocks is highly related to the true number of cluster in the data.

2.7. Number of clusters to be tested in CM

As mentioned, CM requests a prefixed number of clusters. All possible assumptions on the number of clusters for 16 N_VASs are 1 to 16. The assumption that all 16 N_VASs are from 1 single

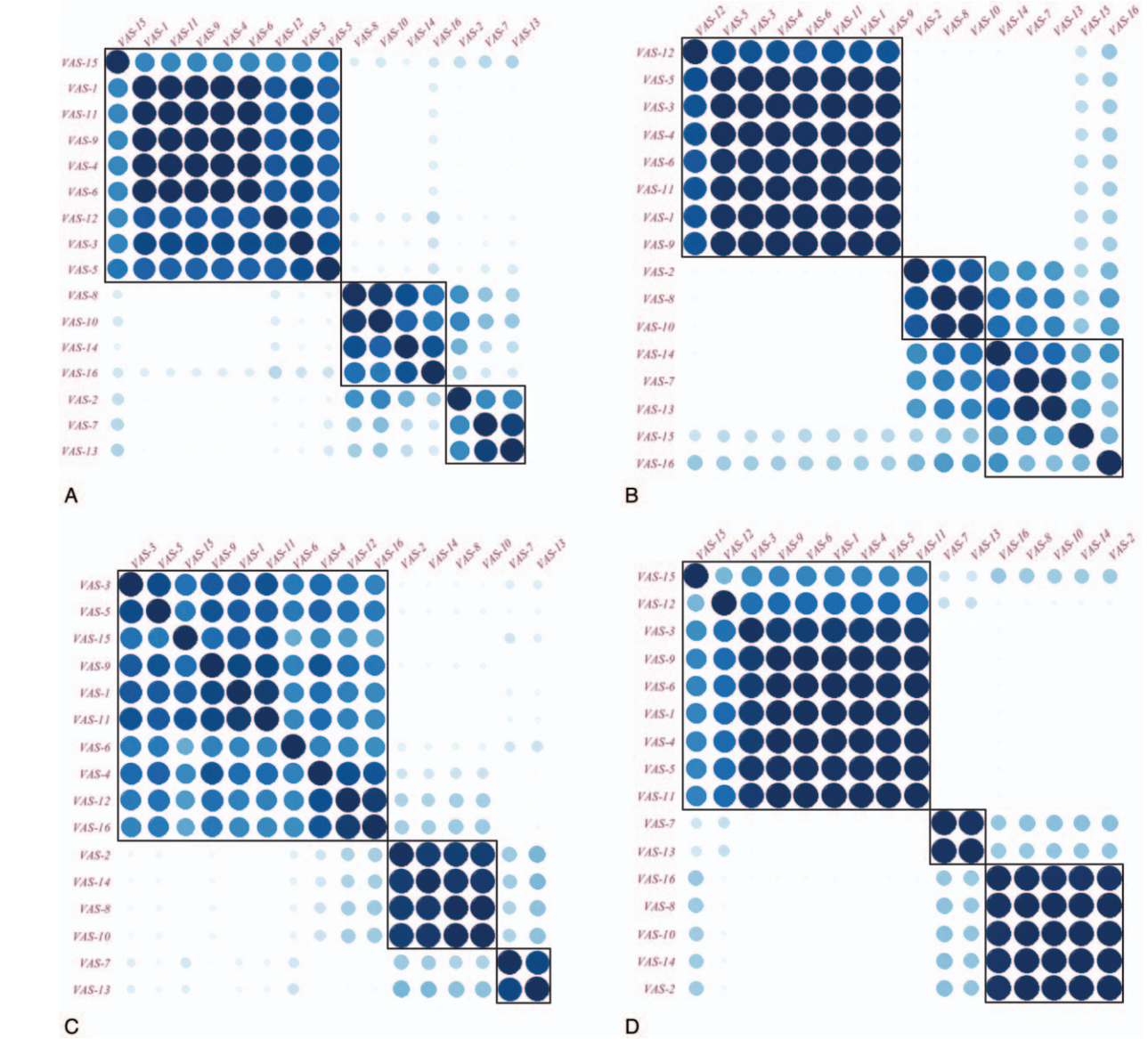

Figure 2. Bootstrap consensus matrix plots of all placebo data by prefixing the number of factors at 3. The sequence of N_VAS items is sorted by internal cluster analysis based on the consensus matrix. The blue color represents correlation = 1, while the white color represent = 0. Panels A, B, C, and D represent all data, THC, SCO, and BEN separately. BEN = benzodiazepine, N_VAS = Norris' visual analogue scales, SCO = scopolamine, THC = delta-9-tetrahydrocannabinol.

cluster was considered trivial and therefore not tested. On the other hand, more than 4 clusters were considered not worthwhile as a dimension-reduction method. Therefore, this study limited the analysis to 2, 3, and 4 clusters.

3. Results

3.1. Descriptive distribution

In total, mean observations of 304 subjects treated with placebo and 297 subjects treated with THC, or scopolamine or benzodiazepine were used. The total study design is depicted in Figure 1. The distribution of each item was plotted (Supplemental Digital Content [Appendix Figs. 1A and 1B], http://links.lww.com/MD2/A90).

3.2. Three-way clustering of scales in data of subjects treated with placebo

Data from the placebo group was first tested using the three-cluster assumption. As shown in Figure 2, even though the cluster number was fixed as 3, the block with items 8, 10, 14, and 16 and block with items 2, 7, and 13 could not be clearly identified. Also, if comparing the clustering result of all placebo data and result of 3 sub-datasets after categorizing by the co-treated drug in the same study, it shows that not only 3 clear sharp, blue blocks could

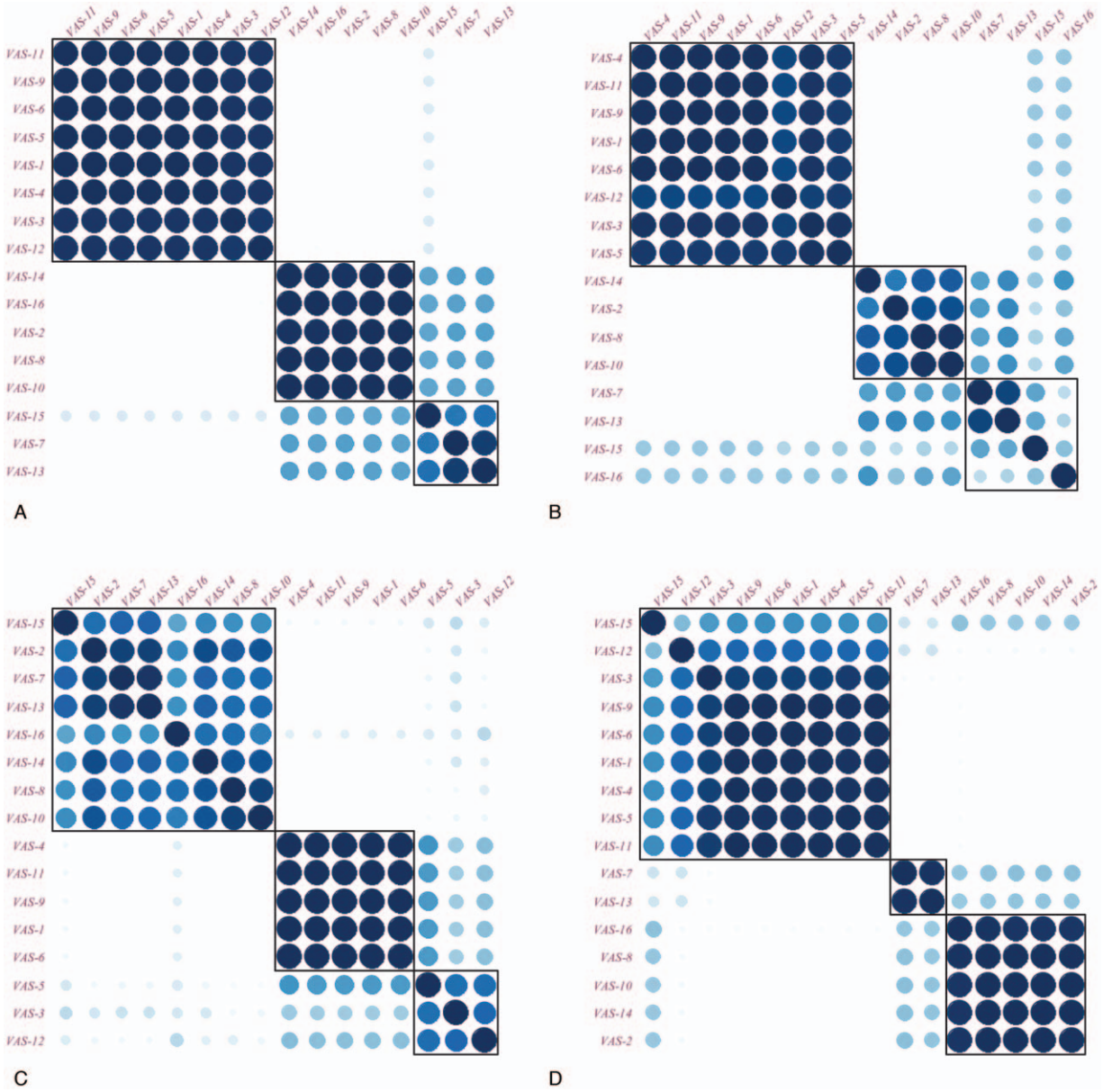

Figure 3. Bootstrap consensus matrix plots by prefixing the number of factors at 3. The sequence of N_VAS items is sorted by internal cluster analysis based on the consensus matrix. The blue color stands for correlation 1, while the white color stands for 0. Panels A, B, C, and D represent all data, THC, SCO, and BEN separately. BEN = benzodiazepine, SCO = scopolamine, THC = delta-9-tetrahydrocannabinol.

not be identified, but also that quite a few items were very unstable among different datasets, such as items 2, 7, 8, 10, 14, 15, and 16.

3.3. Three-way clustering of scales in data of subjects treated with medication

The stability analysis of the Bond and three-clusters was then carried out in data treated with medication. The analysis results are presented as a heat maps (Fig. 3). The heat maps present themselves with no clear sharp 3 blue blocks in the diagonal line. The clustering differs between each of the drugs. In other words, these data give no support for Bond and Lader's three-way clustering or any other hypothetical three-way clustering.

3.4. Two-way clustering of scales in data of subjects treated with medication

Two blue blocks in the diagonal line are adequately clear across the 4 sub-graphs in Figure 4, especially for the matrix based on merged data, which suggests that the two-cluster assumption is relatively well supported. The 4 cases generally agree on two-clusters: items 2, 7, 8, 10, 13, 14, 15, 16

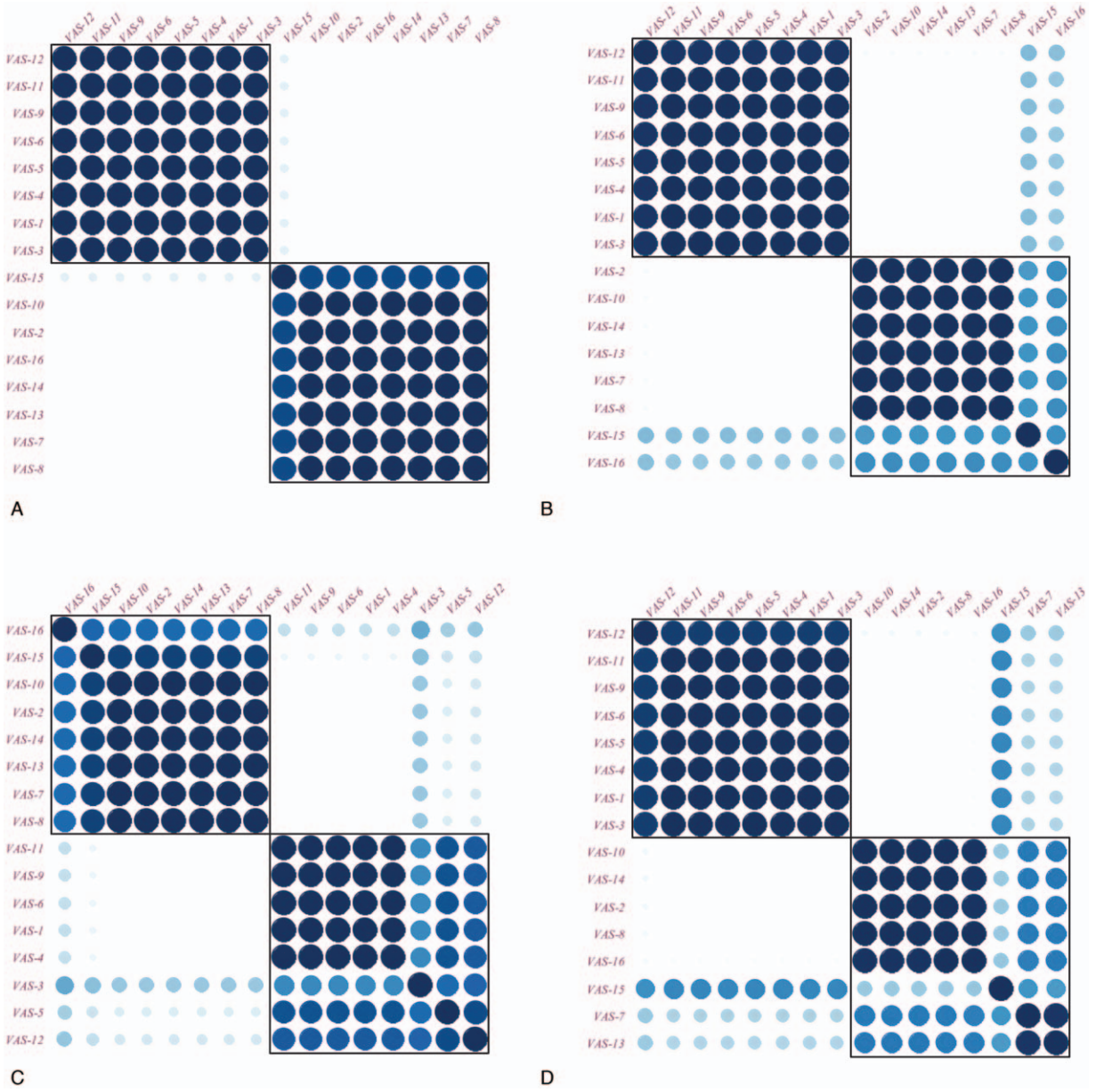

Figure 4. Bootstrap consensus matrix plots by prefixing the number of factors at 2. The sequence of N_VAS items is sorted by internal cluster analysis based on the consensus matrix. The blue color stands for correlation 1, while the white color stands for 0. Panels A, B, C, and D represent all data, THC, SCO, and BEN separately. BEN = benzodiazepine, N_VAS = Norris' visual analogue scales, SCO = scopolamine, THC = delta-9-tetrahydrocannabinol.

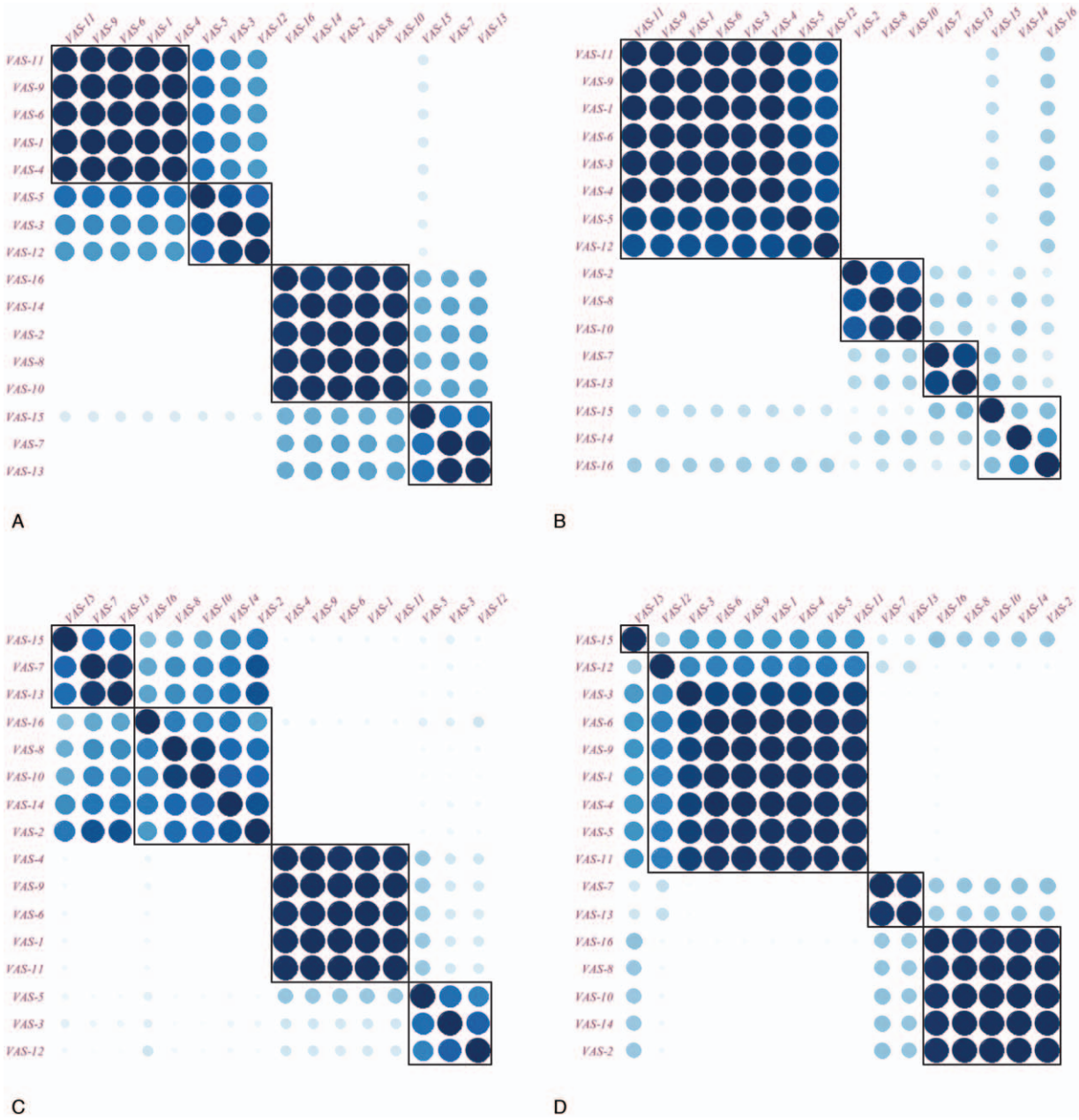

Figure 5. Bootstrap consensus matrix plots by prefixing the number of factors at 4. The sequence of N_VAS items is sorted by internal cluster analysis based on the consensus matrix. The blue color stands for correlation 1, while the white color stands for 0. Panels A, B, C, and D represent all data, THC, SCO, and BEN separately. BEN = benzodiazepine, N_VAS = Norris' visual analogue scales, SCO = scopolamine, THC = delta-9-tetrahydrocannabinol.

and items 1, 3, 4, 5, 6, 9, 11, 12. However, there is still some mild instability, mainly caused by 2 highly unstable items 15 items [Interested – Bored] and 16 [Withdrawn – Gregarious]. Moreover, item 3 [Strong – Feeble] shows mild instability for the scopolamine treatment.

3.5. Four-way clustering of scales in data of subjects treated with medication

Figure 5 shows that a four-cluster assumption is not supported in any of the cohorts or their combination. The clustering differs significantly between each drug treatment as well.

4. Discussion

The aim of this analysis was to test whether Bond and Lader's clustering was stable over samples and over drug effects.[3] If not, alternative clustering of N_VAS was to be tested.

The three-cluster presented by Bond and Lader[3] was not stable in general, neither over samples nor over drug effects. This is consistent for data of subjects treated with placebo and study drugs. The heat maps of the placebo data were different if categorized by co-treated drugs in the same trial. This suggests considerable inter-trial variability which may lead to instability based on Bond and Lader's three-way clustering.[3]

Table 2

Newly identified clusters of Norris' scales.

New categories	Original VAS scale	Item	Indicator in heat map
Sedation and impairment	Alert – Drowsy	VASBL01	1
	Strong – Feeble	VASBL03	3
	Muzzy – Clear-headed	VASBL04	4
	Well-coordinated – Clumsy	VASBL05	5
	Lethargic – Energetic	VASBL06	6
	Mentally slow – Quick-witted	VASBL09	9
	Attentive – Dreamy	VASBL11	11
	Incompetent – Proficient	VASBL12	12
Mood	Calm – Excited	VASBL02	2
	Contented – Discontented	VASBL07	7
	Troubled – Tranquil	VASBL08	8
	Tense – Relaxed	VASBL10	10
	Happy – Sad	VASBL13	13
	Antagonistic – Amicable	VASBL14	14
	Interested – Bored	VASBL15	15
	Withdrawn – Gregarious	VASBL16	16

As the three-cluster approach lacked supporting evidence, data reduction could be improved by other ways of clustering. To test this hypothesis, 2 and 4 clusters were then applied. The two-cluster assumption was proved to be sufficiently stable in general, compared to the three- and four-cluster assumption. However, the two-clusters resulted in minor instability across the drug effects. More specifically, the two-cluster result was the same for all combined data, but items 15 and 16 in THC, item 3, 5, and 12 in the scopolamine treatment, and items 7, 13, and 15 following benzodiazepine administration showed minor instability. Item 15 was found to be unstable in most of the clusters. The reason for the poor performance could be the difficulty of interpreting the question (interested – Bored), and hence a further instruction from the researcher administering the questionnaire could be beneficial.

As the two-cluster approach was most stable, the most appropriate description of the clusters was sought. One cluster was named sedation and impairment, which included items 1, 3, 4, 5, 6, 9, 11, and 12. These items are the first 2 types of Norris original design (see Table 2). In our study, mental and physical sedation could not be identified, nor intellectual, and bodily impairment. As a result, the cluster containing these items was called sedation and impairment in general. The remaining cluster includes items 2, 7, 8, 10, 13, 14, 15, and 16, which are collectively best described as mood. Current clustering results also agreed with Herbert's formal cluster[17] using principle-component analysis except item 15 (Interested – Bored), which is one of the most instable items among difference drugs.

One of the basic assumptions is that N_VAS scores, which basically are of an ordinal categorical nature, are turned into continuous variables.[7,8] Byrne presented an elegant summary on the topic of treating categorical data as continuous.[36] In short, the majority of the literature supports the opinion that ignoring the categorical nature of the data is negligible as the number of categories and the proximity of the data to a normal distribution increase. However, departure from normality, especially skewness, can cause decline of the reliability of the Chi-square-based test or estimation. Hence, the assumption of continuousness may not cause serious trouble in the exploratory factor analysis, which is the basis of our cluster algorithm, but the internal statistical tests should only be used with caution. Besides, even though 3 types of medications were tested, it is worth mentioning that this clustering was only limited to drugs with sedative properties. To evaluate drugs with stimulant properties, the conclusion from this study may not be applicable. For a specific new drug, the statistical approach proposed in this work can be applied for this drug, when the data is available, to ensure the most proper clustering number can be selected for it. Another limitation of this study was that the current approach was only aimed to cluster the 16 scales of N_VAS in a population manner but was unable to quantify variations between subjects. Statistical approach focusing on estimate inter-subject variability, such as non-linear mix effect model,[37] etc might be further explored on top of the clustering methods in future. In this retrospective study, only data of male subjects were available. Considering multiple reports on sex/gender impact on central nervous system,[38–40] further data should be collected for female subjects in future.

The analysis in this manuscript is mainly based on consensus clustering method of Monti et al,[35] with necessary adjustments for application in the drug development research. It is a model-independent resampling-based methodology of class discovery and cluster validation, which provides both numerical measurements and visualized tools for assessing the quality of the clustering results with respect to sensitiveness-to-noise and over-fitting. From the diverse number of resampling schemes, Monti et al proposed the appropriateness of subset sampling on features (comparable to the scales in our case) in gene expression studies due to the high dimensionality of the data.[35] However, this is not suitable for pharmaceutical/clinical studies, and the reason is two-fold. Compared to the normal number of features in gene studies (i.e., over 10,000), we only have 16 variables, which is undesirable for any resampling scheme with an adequate amount of runs. On the other hand, a clinical study is usually able to include much more samples than gene studies. Therefore, bootstrapping on samples was chosen for our study due to the parsimony of information, with the underlying assumption of independence between scales clustering and samples. Besides, bootstrapping produces perturbed data sets with the same size as the original, which can be one of the influential factors for determining the number of clusters.

It is inherent that 2 factor clustering is more stable than 3 or 4 and only a single factor analysis may already help to test the stability on the original dataset. Yet, the current study provided a statistical proof that is based on clinical data of multiple types of drugs. We applied the consensus cluster method to present direct and visible advantages of two-cluster way, which suggested that Bond and Lader's cluster method could be improved.

5. Conclusion

In summary, based on the consensus clustering method with EFA and bootstrapping, it can be concluded that a 2-cluster approach, best described as sedation/impairment and mood, proved to be sufficiently stable for Norris' VAS scale.

Acknowledgment

The authors wish to thank Prof Dr Theo Stijnen (Department of Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands) for his valuable input and contributions.

Author contributions

Conceptualization: Zheng Guan, X. Gregory Chen, Marieke de Kam.
Data curation: Zheng Guan, Justin Hay.
Formal analysis: X. Gregory Chen, Marieke de Kam.
Investigation: Joop van Gerven, Jacobus Burggraaf.
Methodology: Zheng Guan, X. Gregory Chen, Marieke de Kam.
Project administration: Zheng Guan.
Resources: Justin Hay, Joop van Gerven, Jacobus Burggraaf, Marieke de Kam.
Supervision: Justin Hay, Jacobus Burggraaf, Marieke de Kam.
Visualization: Zheng Guan, X. Gregory Chen.
Writing – original draft: Zheng Guan, X. Gregory Chen.
Writing – review & editing: Zheng Guan, Justin Hay, Joop van Gerven, Jacobus Burggraaf, Marieke de Kam.

References

[1] Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov 2004;3:711–5.
[2] Talevi A, Bellera CL, Di Ianni M, et al. CNS drug development – lost in translation? Mini Rev Med Chem 2012;12:959–70.
[3] Bond AJ, Lader MH. Residual effects of hypnotics. Psychopharmacologia 1972;25:117–32.
[4] Jarvis MJ, Lader MH. The effects of nitrous oxide on the auditory evoked response in a reaction time task. Psychopharmacologia 1971;20:201–12.
[5] Mathew H, Proudfoot A, Atiken R, et al. Nitrazepam: a safe hypnotic. Br Med J 1969;3:23–5.
[6] Montagu JD. Effects of quinalbarbitone (secobarbital) and nitrazepam on the EEG in man: quantitative investigations. Eur J Pharmacol 1971;14:238–49.
[7] Norris H. The action of sedatives on brain stem oculomotor systems in man. Neuropharmacology 1971;10:181–91.
[8] Malpas JF. The use of chlorine for water disinfection in industry. Chem Ind 1971;4:111–5.
[9] Breivik EK, Bjørnsson GA, Skovlund E. A comparison of pain rating scales by sampling from clinical data. Clin J Pain 2000;16:22–8.
[10] Hawker GA, Mian S, Kendzerska T, et al. Measures of adult pain: Visual Analog Scale for pain (VAS pain), Numeric Rating Scale for pain (NRS pain), Mcgill Pain Questionnaire (MPQ), Short-Form Mcgill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). Arthritis Care Res 2011;63: S240–52.
[11] Heller GZ, Manuguerr M, Chow R. How to analyze the visual analogue scale: myths, truths and clinical relevance. Scand J Pain 2016;13:67–75.
[12] Sung Y, Wu J. The visual analogue scale for rating, ranking and paired-comparison (VAS-RRP): a new technique for psychological measurement. Behav Res Methods 2018;50:1694–715.
[13] Brown A. Item response models for forced-choice questionnaires: a common framework. Psychometrika 2014;81:1–26.
[14] Sajatovic M, Chen P, Young RC. Rating Scales in Bipolar Disorder, Clinical Trial Design Challenges in Mood Disorders. 1st ed2015; Academic Press,
[15] Dauphin AP, Guillemin F, Virion JM, et al. Bias and precision in visual analogue scales: a randomized controlled trial. Am J Epidemiol 1999;150:1117–27.
[16] Bond A, Lader M. The use of analogue scales in rating subjective feelings. Br J Med Psychol 1974;47:211–8.
[17] Herbert M, Johns MW, Doré C. Factor analysis of analogue scales measuring sugjective feelings before and after sleep. Br J Med Psychol 1976;49:373–9.
[18] van der Post JP, de Visser SJ, Schoemaker RC, et al. Pharmacokinetic/pharmacodynamic assessment of tolerance to central nervous system effects of a 3 mg sustained release tablet of rilmenidine in hypertensive patients. J Psychopharmacol 2004;18:221–7.
[19] Liem-Moolenaar M, Gray FA, de Visser SJ, et al. Psychomotor and cognitive effects of a single oral dose of talnetant (SB223412) in healthy volunteers compared with placebo or haloperidol. J Psychopharmacol 2010;24:73–82.
[20] Liem-Moolenaar M, Zoethout RW, de Boer P, et al. The effects of the glycine reuptake inhibitor R213129 on the central nervous system and on scopolamine-induced impairments in psychomotor and cognitive function in healthy subjects. J Psychopharmacol 2010;24:1671–9.
[21] Liem-Moolenaar M, Zoethout RW, de Boer P, et al. The effects of a glycine reuptake inhibitor R231857 on the central nervous system and on scopolamine-induced impairments in cognitive and psychomotor function in healthy subjects. J Psychopharmacol 2010;24:1681–7.
[22] Liem-Moolenaar M, te Beek ET, de Kam ML, et al. Central nervous system effects of haloperidol on THC in healthy male volunteers. J Psychopharmacol 2010;24:1697–708.
[23] Kleinloog D, Roozen F, De Winter W, et al. Profiling the subjective effects of delta9-tetrahydrocannabinol using visual analogue scales. Int J Methods Psychiatr Res 2014;23:245–56.
[24] Muehlan C, Zuiker R, Peeters P, et al. Pharmacokinetics and pharmacodynamics of the dual orexin receptor antagonist daridorexant in Japanese and caucasian subjects. J Clin Psychopharmacol 2020; 40:157–66.
[25] Cudeck R. Cudeck R, MacCallum RC. Factor analysis in the year 2004: still spry at 100. Factor Analysis at 100: Historical Developments and Future Directions Mahwah, NJ: Lawrence Erlbaum; 2007;1–7.
[26] Paoletti M, Camiciottoli G, Meoni E, et al. Explorative data analysis techniques and unsupervised clustering methods to support clinical assessment of Chronic Obstructive Pulmonary Disease (COPD) phenotypes. J Biomed Inform 2009;42:1013–21.
[27] Feighner JP, Sverdlov L, Nicolau G. Cluster analysis of clinical data to identify subtypes within a study population following treatment with a new pentapeptide antidepressant. Int J Neuropsychopharmacol 2000;3: 237–42.
[28] Hirano S, Sun X, Tsumoto S. Comparison of clustering methods for clinical databases. Inf Sci 2004;159:155–65.
[29] Liem-Moolenaar M, Rad M, Zamuner S, et al. Central nervous system effects of the interaction between risperidone (single dose) and the 5-HT6 antagonist SB742457 (repeated doses) in healthy men. Br J Clin Pharmacol 2011;71:907–16.
[30] Klumpers LE, Fridberg M, de Kam ML, et al. Peripheral selectivity of the novel cannabinoid receptor antagonist TM38837 in healthy subjects. Br J Clin Pharmacol 2013;76:846–57.
[31] Liem-Moolenaar M, de Boer P, Timmers M, et al. Pharmacokinetic-pharmacodynamic relationships of central nervous system effects of scopolamine in healthy subjects. Br J Clin Pharmacol 2011;71:886–98.
[32] Klumpers LE, Roy C, Ferron G, et al. Surinabant, a selective CB(1) antagonist, inhibits THC-induced central nervous system and heart rate effects in humans. Br J Clin Pharmacol 2013;76:65–77.
[33] Baas JM, Mol N, Kenemans JL, et al. Validating a human model for anxiety using startle potentiated by cue and context: the effects of alprazolam, pregabalin, and diphenhydramine. Psychopharmacology (Berl) 2009;205:73–84.
[34] Hardle W, Simar L. Applied Multivariate Statistical Analysis. 2nd edBerlin Heidelberg: Springer-Verlag; 2007.
[35] Monti S, Tamayo P, Mesirov J, et al. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. Mach Learn 2013;52:91–118.
[36] Byrne BM. Structural Equation Modeling with EQS: Basic Concepts, Applications and Programming. 2006;Lawrence Erlbaum Associates,
[37] Davidian M, Giltinan DM. Nonlinear models for repeated measurement data: an overview and update. J Agric Biol Environ Stat 2003;8:387–419.
[38] de Vries GJ, Forger NG. Sex differences in the brain: a whole body perspective. Biol Sex Differ 2015;6:15.
[39] Brooks C, Clayton J. Sex/gender influences on the nervous system: basic steps toward clinical progress. J Neurosci Res 2017;95:14–6.
[40] Zagni E, Simoni L, Colombo D. Sex and gender differences in central nervous system-related disorders. Neurosci J 2016;2016:2827090.