



**COMMENTARY**

**Open Access**

# Diagnostic heterogeneity in psychiatry: towards an empirical solution

Klaas J Wardenaar and Peter de Jonge\*

## Abstract

The launch of the 5th version of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) has sparked a debate about the current approach to psychiatric classification. The most basic and enduring problem of the DSM is that its classifications are heterogeneous clinical descriptions rather than valid diagnoses, which hampers scientific progress. Therefore, more homogeneous evidence-based diagnostic entities should be developed. To this end, data-driven techniques, such as latent class- and factor analyses, have already been widely applied. However, these techniques are insufficient to account for all relevant levels of heterogeneity, among real-life individuals. There is heterogeneity across persons (p:for example, subgroups), across symptoms (s:for example, symptom dimensions) and over time (t:for example, course-trajectories) and these cannot be regarded separately. Psychiatry should upgrade to techniques that can analyze multi-mode (p-by-s-by-t) data and can incorporate all of these levels at the same time to identify optimal homogeneous subgroups (for example, groups with similar profiles/connectivity of symptomatology and similar course). For these purposes, Multimode Principal Component Analysis and (Mixture)-Graphical Modeling may be promising techniques.

**Keywords:** DSM-5, Heterogeneity, Data-driven techniques, Cattell's cube

## Introduction

With the launch of the fifth version of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), the debate about current psychiatric diagnostics has come into the limelight again, focusing on specific alterations in the DSM-5, such as the deletion of pervasive developmental disorder not otherwise specified (PDD-NOS) and Asperger's Disorder [1,2] and the inclusion of mourning in major depressive disorder (MDD). However, more fundamental topics, such as the medicalization of normal behavior [3] and the categorical approach to continuous phenomena, are also debated [4]. Perhaps the most important criticism of the DSM-5 regards the poor validity of its classification. Several researchers have even stressed that the DSM-5 hampers research into the underlying mechanisms in the etiology of psychopathology and that the current state of affairs is one of scientific stagnation [5]. We argue that the development of more valid psychiatric classifications is important in

order to link mental states to specific causes in scientific research, and that this process should be evidence-based. Decreasing the amount of diagnostic heterogeneity is central in this process.

## The problem of diagnostic heterogeneity

Current psychopathological concepts are heterogeneous by default which restricts their usefulness for research [6,7]. In the past, evidence-based attempts to decrease heterogeneity have been made. For depression, for instance, subtypes have been identified with latent class analyses (LCA) [8,9], symptom-dimensions with factor analyses (FA) [10,11] and course-trajectory groups with mixture growth analyses (MGA) [12,13]. Unfortunately, these studies tackle only one aspect of heterogeneity at a time. LCA focuses on person (p)-level heterogeneity, but does not account for within-class symptom and course variations. FA tackles symptom (s)-level heterogeneity, but assumes stability across persons and time. MGA describes temporal (t) heterogeneity, but does not account for s-level heterogeneity. Not surprisingly, these approaches have led to artificial models with limited replicability [11].

\* Correspondence: peter.de.jonge@umcg.nl  
Interdisciplinary Center Psychopathology and Emotion Regulation (ICPE),  
University of Groningen, University Medical Center Groningen (UMCG),  
Hanzeplein 1, Groningen 9713 GZ, The Netherlands

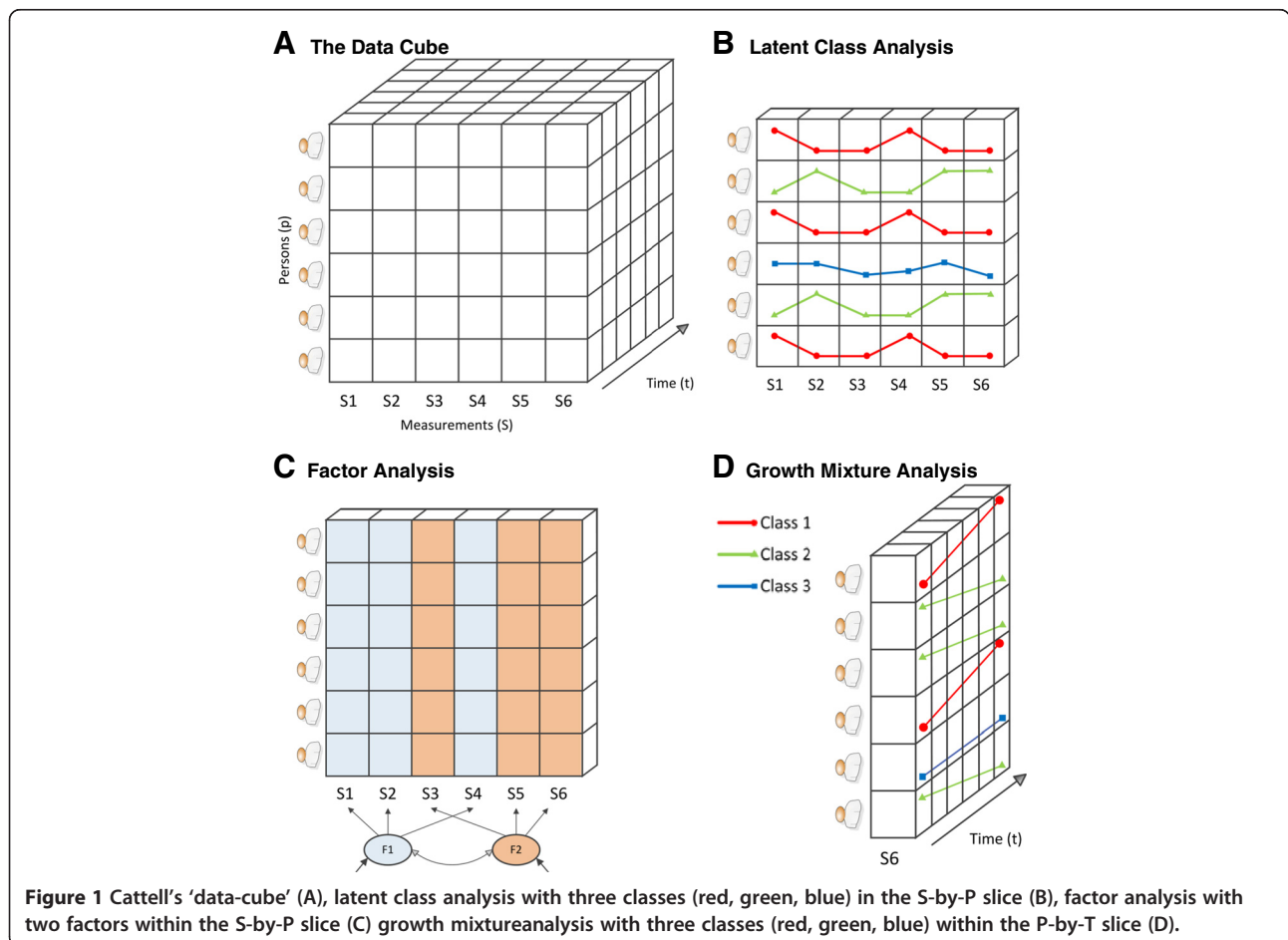
**The solution: simultaneous heterogeneity reduction**

If homogeneous diagnoses are what psychiatry aims for, a data-driven approach should be designed to minimize heterogeneity on each level simultaneously. To enable reduction of p-, s- and t-level heterogeneity, three-mode data are needed, visualized by Cattell's data cube [14] (Figure 1A). The cube consists of measured data (s-axis) for n individuals (p-axis) at k time-points (t-axis). For each combination of axes (slices), different statistical techniques apply. Cross-sectional studies of heterogeneity apply to the p-by-s slice: LCA divides the p-axis into classes (Figure 1B) and FA divides the s-axis into factors (Figure 1C). To model heterogeneity of the whole slice, model combinations (for example, factor mixture models) [15] can be used. Longitudinal studies of heterogeneity (for example, MGA) apply to the p-by-t slice, modeling classes-based temporal trajectories on one or more variables (Figure 1D). Although incomplete, this summary shows that none of these models incorporate all three sources of variation. If we look to other fields (for example, psychometrics, mathematics), we can see that statistical advances have reached the point where

'three-dimensional models' are a possibility. Here, we briefly discuss two candidates.

**The latent variable approach: three-mode principal component analysis (3MPCA)**

3MPCA [16] is an exploratory technique, designed to decompose the latent structure of three-dimensional data by identifying the number of components that make up each of the axes. Investigation of the interactions between the modes can yield insights into the latent structure of three-dimensional data as a whole. In anxiety patients, for instance, 3MPCA showed that patients could be divided into subgroups (p-component) with different clusters of symptoms (s-component) in different situations (t-component) [17]. Such an approach can be extended to a broader range of psychopathological phenomena. 3MPCA does have its limitations: it requires subjective judgments to enable model selection and can yield hard-to-interpret results. However, it is a fully developed technique that can be used to explore three-dimensional psychopathology data for more homogeneous diagnostic entities.



### The network approach: (mixture) graphical analysis

Traditional concepts of psychopathology (diagnoses, subtypes, dimensions) lean heavily on the assumption that corresponding latent constructs exist. Unfortunately, it is uncertain to what extent this is a realistic assumption [18]. Rather than assuming that different symptoms (energy loss, suicidal ideation) are caused by one underlying disease (for example, depression), one could instead look at how symptoms interact, amplify and sustain each other over time in a network of symptoms (nodes) and causal links (edges) [18,19], using graphical model methodology, developed in biostatistics [20]. Such patient-descriptions are highly personalized: they take homogeneity to the extreme, both at the s- and p-level. Within three-dimensional data the s-axis is completely subdivided down to its smallest components (for example, symptoms). On the p-axis, for each person, the repeatedly measured symptoms are incorporated in a personalized network model. On the p-level, such an approach could lead to an indefinite number of possible network configurations, leaving us without any common denominators. However, subgroups with common network characteristics can be identified by mixture/latent class analyses on network model parameters. Such an approach can yield subtypes that are not merely defined by common symptomatology, but particularly by their observed interconnectedness.

### Conclusions

The development of evidence-based diagnoses in psychiatry is bound to require the use of data-driven techniques. In order for the resulting diagnostic models to optimally reflect real-world variation among patients, multiple sources of heterogeneity should be simultaneously evaluated. Although complex, and dependent upon the data quality, such methods are a necessity when psychiatric diagnosis seeks an empirical basis.

### Competing interests

The authors declare that they have no competing interests.

### Acknowledgements

PdJ and KJW are supported by a VICI grant (no: 91812607) from the Netherlands Research Foundation (NWO-ZonMW).

Received: 13 June 2013 Accepted: 13 June 2013

Published: 12 September 2013

### References

1. Huerta M, Bishop SL, Duncan A, Hus V, Lord C: Application of DSM-5 criteria for autism spectrum disorder to three samples of children with DSM-IV diagnoses of pervasive developmental disorders. *Am J Psychiatry* 2012, **169**:1056–1064.
2. Ritvo ER, Ritvo RA: Commentary on the application of DSM-5 criteria for autism spectrum disorder. *Am J Psychiatry* 2013, **170**:444–445.
3. Frances A: *Saving Normal: An Insider's Revolt Against Out-of-Control Psychiatric Diagnosis, DSM-5, Big Pharma, and the Medicalization of Ordinary Life*. New York: William Morrow; 2013.

4. Wakefield JC: The DSM-5 debate over the bereavement exclusion: psychiatric diagnosis and the future of empirically supported treatment. *Clin Psychol Rev* 2013. doi:10.1016/j.cpr.2013.03.007.
5. Shorter E, Tyrer P: Separation of anxiety and depressive disorders: blind alley in psychopharmacology and classification of disease. *BMJ* 2003, **327**:158–160.
6. Widiger TA, Clark LA: Toward DSM-V and the classification of psychopathology. *Psychol Bull* 2000, **126**:946–963.
7. Widiger TA, Samuel DB: Diagnostic categories or dimensions? A question for the diagnostic and statistical manual of mental disorders—fifth edition. *J Abnorm Psychol* 2005, **114**:494–504.
8. Eaton WW, Dryman A, Sorenson A, McCutcheon A: DSM-III major depressive disorder in the community. A latent class analysis of data from the NIMH epidemiologic catchment area programme. *Br J Psychiatry* 1989, **155**:48–54.
9. Sullivan PF, Kessler RC, Kendler KS: Latent class analysis of lifetime depressive symptoms in the national comorbidity survey. *Am J Psychiatry* 1998, **155**:1398–1406.
10. van Loo HM, de Jonge P, Romeijn JW, Kessler RC, Schoevers RA: Data-driven subtypes of major depressive disorder: a systematic review. *BMC Med* 2012, **10**:156.
11. Shafer AB: Meta-analysis of the factor structures of four depression questionnaires: Beck, CES-D, Hamilton, and Zung. *J Clin Psychol* 2006, **62**:123–146.
12. Byers AL, Vittinghoff E, Lui LY, Hoang T, Blazer DG, Covinsky KE, Ensrud KE, Cauley JA, Hillier TA, Fredman L, Yaffe K: Twenty-year depressive trajectories among older women. *Arch Gen Psychiatry* 2012, **69**:1073–1079.
13. Rhebergen D, Lamers F, Spijker J, de Graaf R, Beekman AT, Penninx BW: Course trajectories of unipolar depressive disorders identified by latent class growth analysis. *Psychol Med* 2012, **42**:1383–1396.
14. Cattell RB: The data box: its ordering of total resources in terms of possible relational systems. In *Handbook of Multivariate Experimental Psychology*. Edited by Cattell RB. Chicago: Rand-McNally; 1966:67–128.
15. Lubke GH, Muthén B: Investigating population heterogeneity with factor mixture models. *Psychol Methods* 2005, **10**:21–39.
16. Kroonenberg PM, De Leeuw J: Principal component analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika* 1980, **45**:69–97.
17. Kiers HAL, van Mechelen I: Three-way component analysis: principles and illustrative application. *Psychol Methods* 2001, **6**:84–110.
18. Borsboom D, Cramer AO: Network analysis: an integrative approach to the structure of psychopathology. *Annu Rev Clin Psychol* 2013, **9**:91–121.
19. Cramer AO, Waldorp LJ, van der Maas HL, Borsboom D: Comorbidity: a network perspective. *Behav Brain Sci* 2010, **33**:137–150.
20. Abegaz F, Wit E: Sparse time series chain graphical models for reconstructing genetic networks. *Biostatistics* 2013, **14**:586–599.

doi:10.1186/1741-7015-11-201

Cite this article as: Wardenaar and de Jonge: Diagnostic heterogeneity in psychiatry: towards an empirical solution. *BMC Medicine* 2013 **11**:201.

Submit your next manuscript to BioMed Central  
and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

