university of groningen

Psychiatric disease classification: a reference class problem

Institute for Psychiatric and Behavioral Genetics
Virginia Commonwealth University
28 August 2017

Jan-Willem Romeijn
Faculty of Philosophy
University of Groningen

Disease classification

Classification schemes for mental illness serve a large variety of goals.

- Medical doctors use classification schemes to design and apply treatments.
- Researchers employ them to design studies and carry them out.
- Patients and their families and friends fall back on classifications for explanation and understanding.

How can we best serve these goals? When is a classification scheme “good”?

This talk

I argue that we can gain some understanding of what makes classification good by framing it in a particular way.

› The problem of finding good classifications is a version of the so-called reference class problem.

› This insight points us to a particular research area in statistics, namely model selection.

› Similarly it leads us to consider interventions in terms of so-called causal networks.

A better understanding of what validates a classification might improve research efforts and clinical practice.

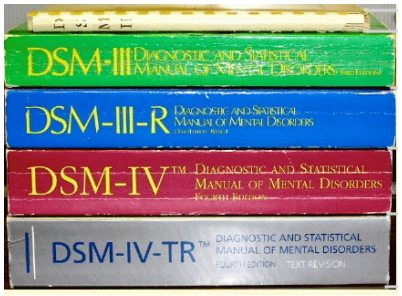
Chances

At the heart of my discussion of psychiatric disease classification are insights into the nature of chance.

The topic touches on a venerable philosophical discussion concerning chances for the single-case.

Disorders as conventions

Much of this paper is set against the background of work with Hanna van Loo.



We proposed a form of *conventionalism* about mental disorders. The disorders are real but at the same time they have a perspectival character.

Freedom to speculate

We thereby steer a course between two opposing views in the so-called science wars: realism and constructivism.

Conventionalism allows for speculation about classification while retaining a form of realism. Asking when schemes are good is the natural next step.

Plan of talk

① Reference class problem

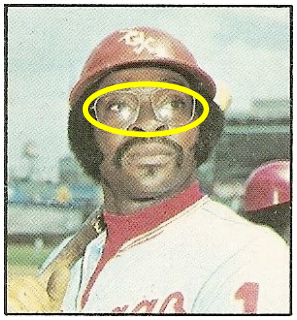
② Model selection

③ Causal modelling

④ Conclusions

① Reference class problem

What is the chance that Dick Allen lives up to the age of 80? Well… he will or he will not.



We can list numerous characteristics of Dick Allen and determine the statistics in the groups defined by them. But what are the *salient* characteristics?

Chances in psychiatry

Statistics can only be determined relative to a group.

	Country	N	DSM	12 mth any dx (in%)	1 dx	2 dx	3 dx	> 3 dx	% pts > 1dx
Bijl 1998	NL	7076	III-R	23,3	15,3	4,4	1,9	1,9	35
Jacobi 2004	BRD	4181	IV	31,1	18,8	6,3	2,8	3,2	40
Kessler 2005	USA	9282	IV	26,2	14,4	5,8	6,0		45

But what group is the salient one for chance ascriptions to a particular individual?

Psychiatric reference classes

My claim is that often a good classification scheme is one that generates the right reference classes.

- We want to select characteristics of individuals, i.e., identify specific groups, that allow us to reliably determine chances pertaining to those individuals.
- Such characteristics arguably provide us with a natural disease classification, and thereby with an understanding of the nature of psychiatric diseases.
- On this basis we may hope to perform interventions that give robust chances of success.

② Model selection

When viewed in this way, the problem of psychiatric disease classification becomes within the reach of statistics.

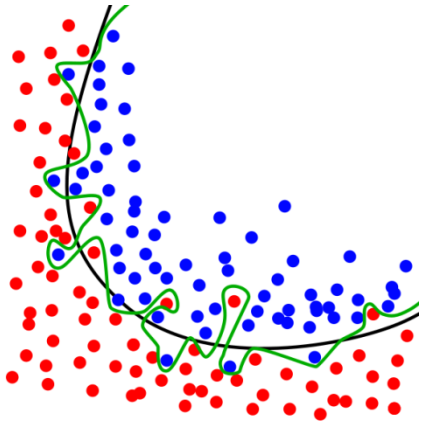
› The search for salient characteristics comes down to the choice for a distinct set of statistical variables, and hence the determination of a statistical model.

› In statistical model selection, the choice of variables is regulated by expected predictive performance. The choice of a model is therefore by-and-large data-driven.

The link with model selection offers a particular grip on disease classification.

Overfitting

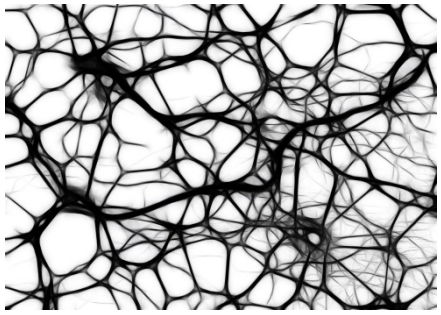
An important property of model selection tools is that they systematically guard against so-called overfitting.



Increasing the number of characteristics may improve the fit to data, but it will make predictions less reliable.

Relations to anti-reductionism

In this empiricist view it is perfectly possible that the classification employs characteristics from a several different levels of description.



This offers an alternative to classifications that are based on an assumed metaphysics, e.g., strictly neuro-scientific.

③ Causal modeling

Perhaps the most important application of classification is in designing treatment programs and allocating individuals to them.

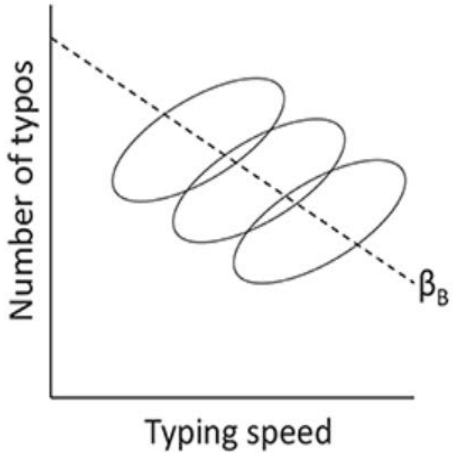
› We want to define mental disorders in such a way that they facilitate maximally effective clinical interventions.

› In psychiatry, as in the medical sciences generally, interventions are mostly stochastic, i.e., they merely raise the chance of some desired outcome.

How can we tailor the classification of disorders to the goal of facilitating interventions with good stochastic properties?

Simpson's paradox

We can look for correlation patterns in the population that correspond to controllable mechanisms in individuals.

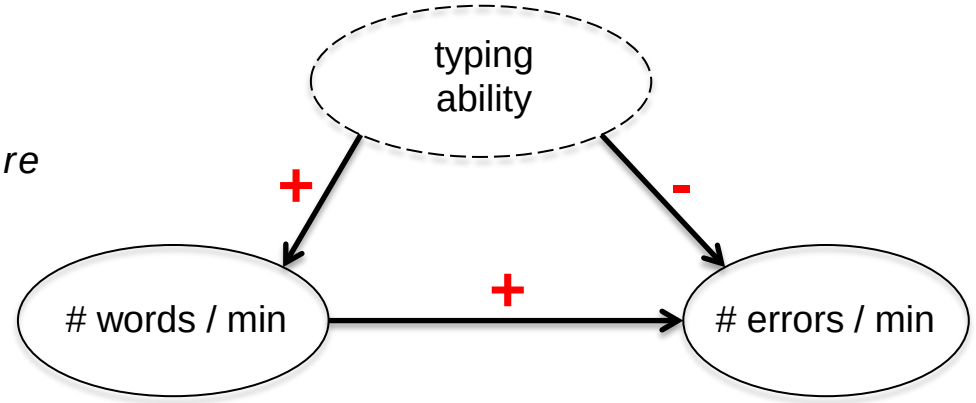


Simpson's paradox makes very vivid that correlation patterns do not always indicate such mechanisms.

Confounders

The mismatch between correlation and causation is due to the existence of hidden factors, so-called confounders.

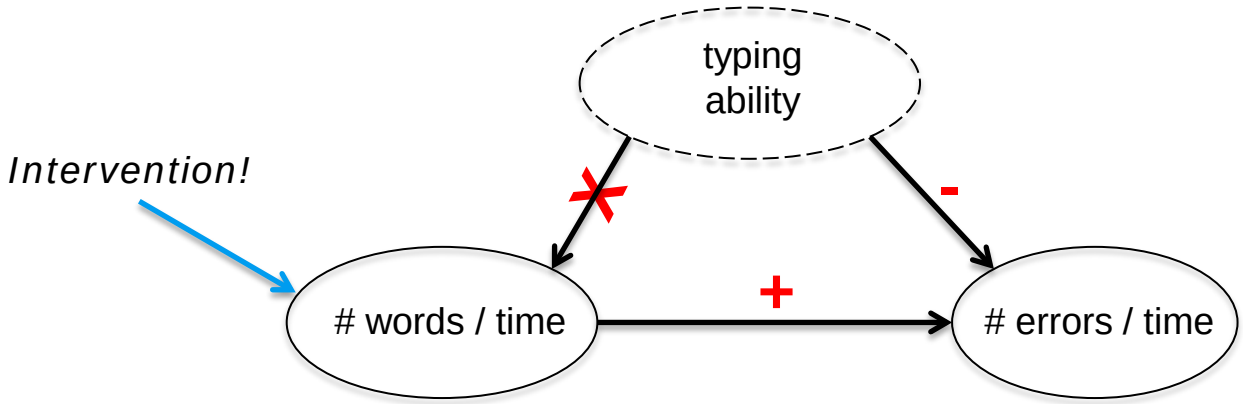
Causal structure



Bayesian causal networks offer a highly insightful account of confounding.

Interventions

We can systematically search for these confounders by means of interventions.

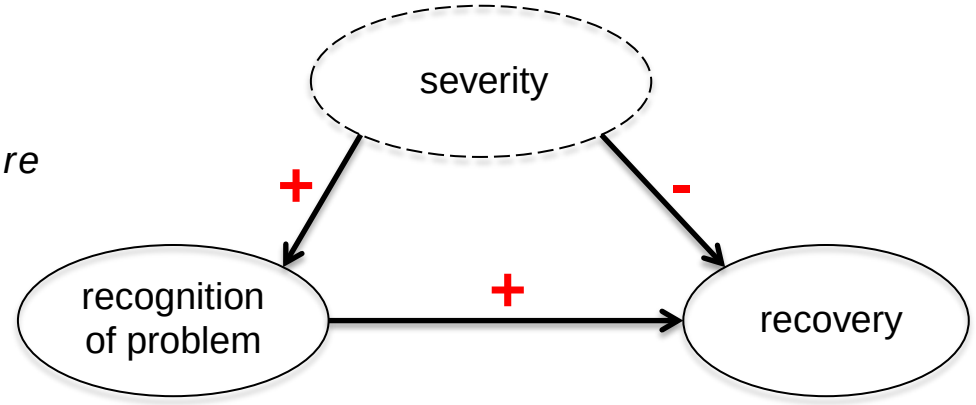


Intervention-free causal inference is not hopeless: we can prepare observational data such that it simulates an experiment.

Causally relevant classifications

Searching for confounders helps to determine the classifications that support better predictions and interventions.

Causal structure



This quality criterion for classification again seems applicable across a variety of validators and levels of description.

④ Conclusions

I hope that the above insights can be of use in the hunt for improvements in disease classification.

- A *conventionalist* view clears the way for trying out revisions to disease classification.
- Viewing disease classification as a reference class problem invites an *empiricist*, *pragmatic* and *pluralist* view on the task of finding good classifications.
- It directs us to the use of *model selection* and *causal modeling*, and suggests that we search for *robust chance* ascriptions.

Some topics for discussion

It is hardly believable that statistical tools alone will deliver answers to the fundamental questions.

- For the purpose of long-term improvements it may be beneficial to adhere to a metaphysics, or a specific disciplinary matrix, when designing a classification.
- Admittedly classification schemes serve many different goals. I have focused on prediction and manipulation as key objectives but this presents a substantive choice.
- The statistical methods and tools that I advertized are too generic and abstract. They need to be tailored to the case at hand.

Thanks for your attention

Some references to papers with Hanna:

- "Measuring and defining: the double role of the DSM-criteria for psychiatric disorders", with H. van Loo, *Psychological Medicine*, 2017.
- "Psychiatric comorbidity does not only depend on diagnostic thresholds: an illustration with major depressive disorder and generalized anxiety disorder", with H. van Loo, P. de Jonge, K.S. Kendler, and R.S. Schoevers, *Depression and Anxiety*, DOI 10.1002/da.22453, 2015.
- "Comorbidity: fact or artefact?", with H. van Loo, *Theoretical Medicine and Bioethics* 36(1), pp. 41-60, 2015.
- "Psychiatric comorbidity and causal disease models", with H.M. van Loo, P. de Jonge, R.A. Schoevers, *Preventive Medicine*, 57(6), pp. 748-752, 2013.
- "Data-driven subtypes of major depressive disorder: a systematic review", with H.M. van Loo, P. de Jonge, R.C. Kessler, and R.A. Schoevers, BMC medicine 10: 156, 2012.