Thick descriptions in psychopathology: Kraepelin meets Geertz

Hanna van Loo and Jan-Willem Romeijn

University of Groningen

Outline

1. Big data in psychiatry
2. The theory-ladenness of data
3. The thinness of research data versus the thickness of clinical data
4. Ways to enrich research data & limitations
5. Conclusions
6. Q&A

1. Big data in psychiatry

Pubmed search
"Data science" & "Psychiatry"
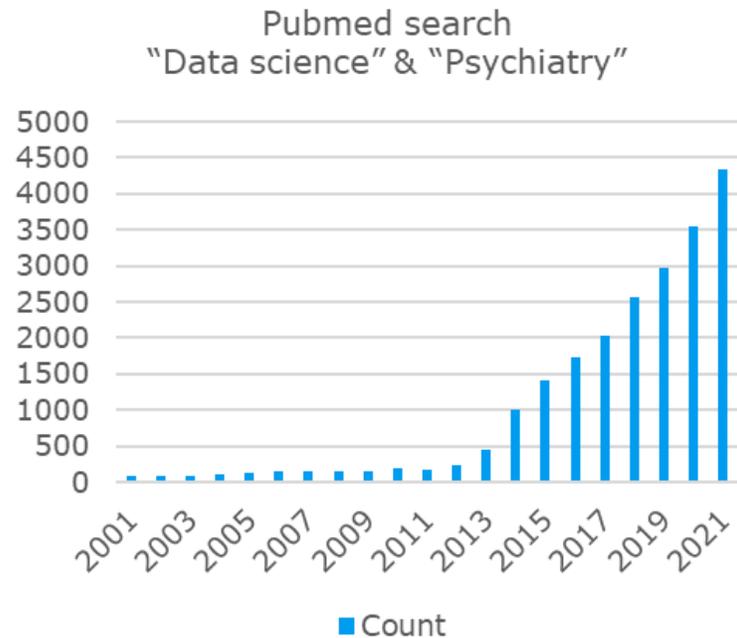
A rapid uptake of new data science methods in psychiatry.

- In response to the availability of large data sets.

- As a means to achieve personalized care.

"Gold rush in the data mine"

Popular media herald a new data science era.

- Theory-free prediction methods.

- Pattern recognition that far exceeds human cognitive capacities.

The data mine holds untold riches. But can data science live up to the promise?

Examples

Online predictions of treatment response for clinical use.

Examples

Data-driven diagnosis using automated clustering techniques.
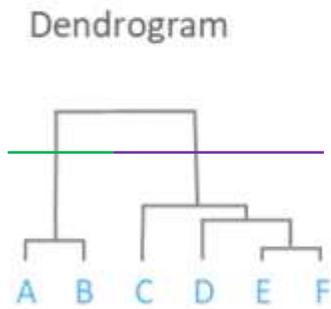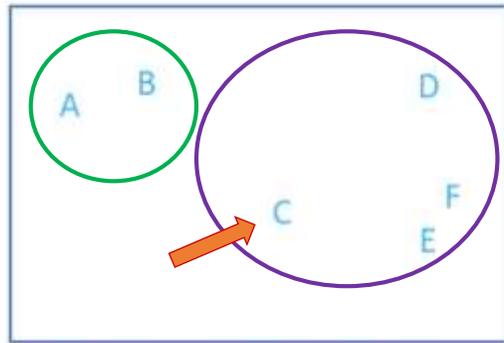
Examples

Early-warning systems using mobile phone data.

Emil Kraepelin

Data-driven psychiatry seems like the end point of a development towards empirical science.

- Observation and experimentation rather than reflection and speculation as an epistemic basis.

- Classification and prediction within patient groups instead of tracing individual narratives.

2. The theory-ladenness of data

There are well-known problems with the data-driven conception of psychiatric science: machine learning makes implicit choices.
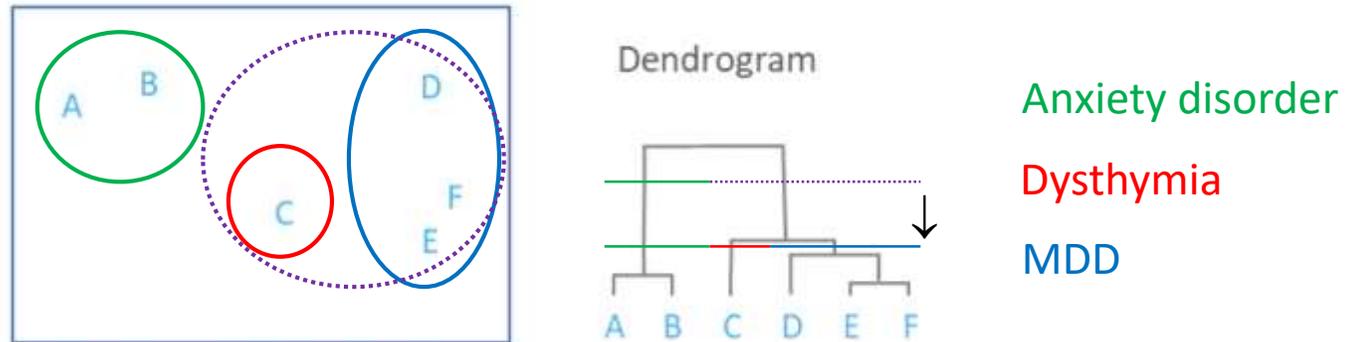
Pilot study

Beijers et al (2020) show that the parameter settings of the algorithms determine the clustering results.

"...1200 combinations of [settings]... no robust patterns of clustering"

Implicit theory

Clustering is steered by presuppositions that are implicit in the data construction, models and inference, and interpretations.
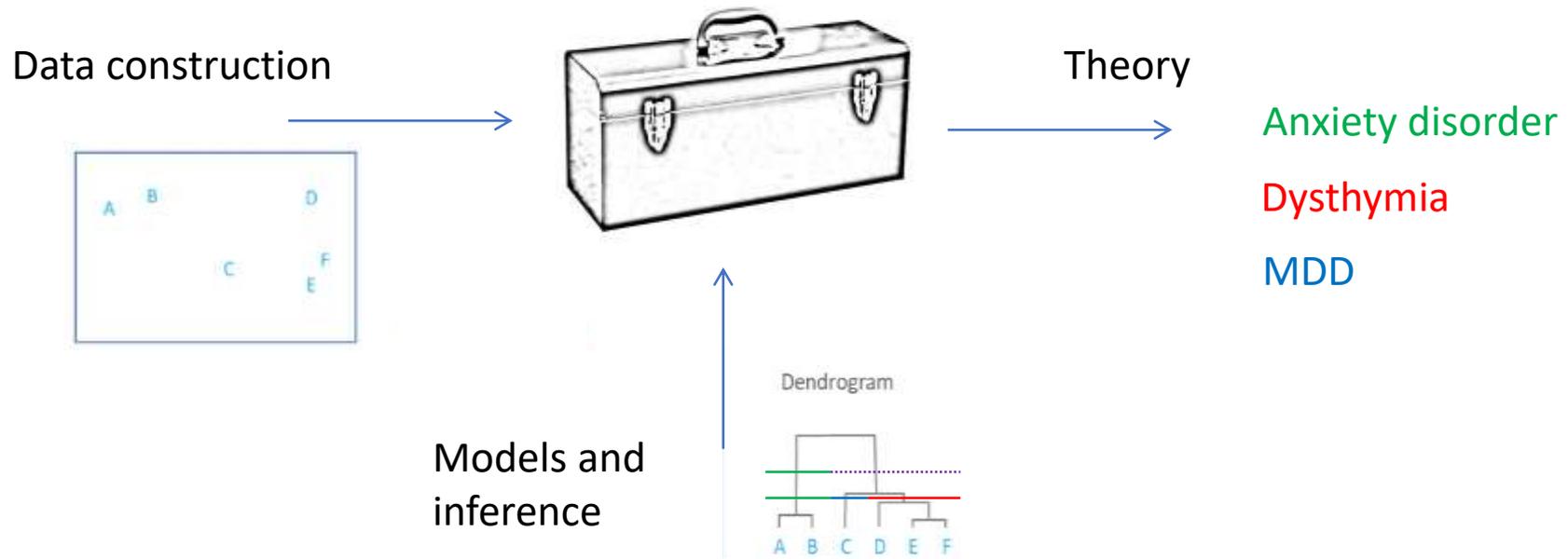
Too much but never enough

The data and methods are in a sense too rich but... they are also not rich enough.

- The data are not free from theory and the models and interpreted findings also hold presuppositions. This is an "evergreen" insight from the philosophy of science.

- Another "evergreen" insight is that we only find things where we care to look. The data are selected and curated, and the models and findings are similarly constrained.

Clifford Geertz
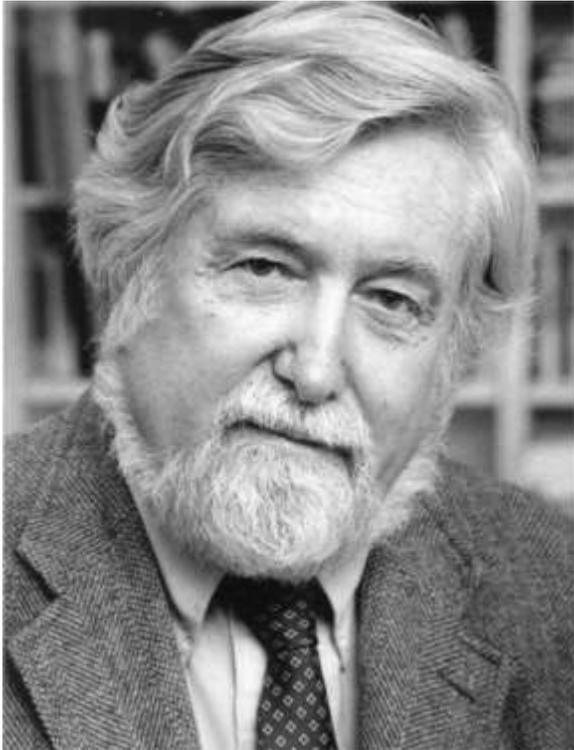
A truly data-driven human science accommodates and remedies this paucity of the empirical basis.

- Observation needs to be extensive and sensitive to the cultural context, rather than general and concise.

- Theoretical developments need to take the form of detailed narrative rather than generality and abstraction.

Thick descriptions

In data-driven psychiatry we need the figures of Kraepelin and Geertz to meet.

- The basis of psychiatric science in observation, classification and prediction needs to be retained.

- But in all these aspects of research we need to work with thick rather than concise and reduced descriptions.

3. The thinness of research data versus the thickness of clinical data

What goes missing if we compare clinical data with research data?

Research data

- Some of the richest datasets today:
  - Lifelines
  - Netherlands Study of Depression and Anxiety (NESDA)

Lifelines: N~165,000, covid barometer

Depression — Percentage of participants who report feeling continually low over the last 14 days.

1 March
Depression : 4.47%
18 - 30 year : 10.08%

Lifelines.nl
Coronabarometer.nl
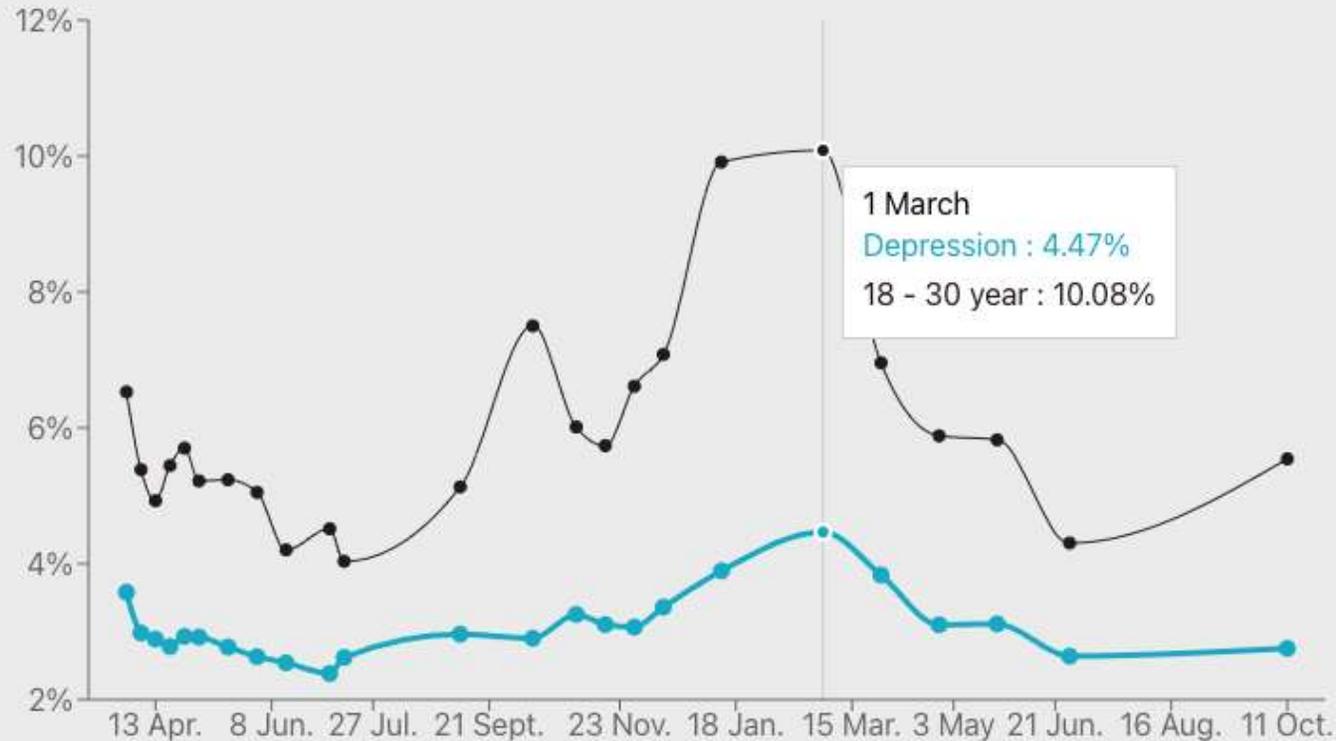
NESDA: N~3,000, 7 waves (2004–2022)

- Demographics
- Lifestyle and health behavior
- Psychology and personality
- Psychopathology
- Stressful life events / childhood trauma
- Somatic health / medication use
- Disability
- Biomaterials
- MRI
- Actigraphy

Topic	Instrument	Code Datadoc	Method	T0 B	FU1 W2	FU2 W3	FU4 W4	FU6 W5	FU9 W6	Sibs	Fu13 W7
Demographics, work and social environment											
Age, gender, ethnicity, education	Standard questions	N1_x100	Int	X	-	-	-	-	-	X	-
Partner, household status, marriage, sexual preference	Standard questions	N1_x101	Int	X	-	X	X	X	X	X	X
Work productivity	TIC-P work	N1_x102	Int	X	X	X	X	X	X	X	X
Partner characteristics	Standard questions	N1_x103	Int	X	-	X	X	X	X	X	X
Living situation, moving house	Standard questions	N1_x104	Int	X	-	X	X	X	X	X	X
Parents and family	Standard questions	N1_x105	Int	X	-	X	X	X	X	X	X
Socioeconomic status	Income	N1_x106	Int	X	-	X	X	X	X	X	X
Religion	Standard questions	N1_x108	SR	X	-	-	-	-	-	-	-
Work content/environment	Karasek questionnaire	N1_x109	Int	X	-	-	-	-	-	-	X
Family tree	Stamboom	N1_x110	Int	X	-	-	-	-	-	-	-
Children and household	Standard questions	N1_x111	Int	X	-	-	-	-	-	-	-
Differences in upbringing	Siblings	N1_x115	SR	-	-	-	-	-	X	X	-
Relationship with mother	Parental Bonding Inventory (PBI)	N1_x116	SR	-	-	-	-	-	X	X	-
Relationship with father	Parental Bonding Inventory (PBI)	N1_x117	SR	-	-	-	-	-	X	X	-
Interpersonal relationships	IIP	N1_x118	SR						X	X	X
Social support	Close Person Inventory (CPI)	N1_x239	SR	X	-	X	-	-	-	-	-
Close relationships	Experiences in Close Rel. (ESR)	N1_x119	SR						X	X	-
Relationships	Dyadic Adjustment Scale	N1_x120	SR						X	X	-
Sibling inventory	Standard questions	N1_x121	Int	-	-	-	-	-	X	-	-
SES during childhood	Lifebrain questions		SR	-	-	-	-	-	-	-	X
Inventory mobile phone use	Standard questions		Int	-	-	-	-	-	X	X	-
Kids	MARIO questions about kids		Int	-	-	-	-	-	-	-	X

Big and rich data!

- Very costly to collect all these data
- Challenge to retain participants and prevent selective drop-outs
- However, data are still relatively thin
- What are aspects that are missing in research data?

Missing data I: Thick descriptions

"Two boys fairly swiftly contract the eyelids of their right eyes. In the first boy this is only an involuntary twitch; but the other is winking conspiratorially to an accomplice.

At the lowest or the thinnest level of description the two contractions of the eyelids may be exactly alike. …

Yet there remains the immense .. difference between a twitch and a wink."

Gilbert Ryle, "The Thinking of Thoughts, What is 'le Penseur' doing?" 1968

Missing data I: Thick descriptions

- Life events are highly predictive of onset of psychiatric disorders, and are often included in research data

- However, the *meaning* of life events is highly individual

- Examples:
  - Death of a parent
  - Impact of covid lockdowns

- Clarifying the meaning of events for the individual is part of clinical interviews but is not part of research questionnaires

Missing data II: "5,000,000 things"

- Millions of behavioral variables may be relevant signs or symptoms
- These may be observed in clinical practice
- Impossible to capture them in questionnaires
- Simply too many
- Or a patient is not aware, observant is needed (e.g., relative / therapist, e.g. countertransference)
- Examples:
  - Patient strokes coat of therapist
  - Partner of patient wears a red dress with mushroom-print

Missing data III: Causal relations

In clinical practice, one not only records symptoms but also tries to identify potential risk or protective factors

- Predisposing factors (eg genetic predisposition, childhood emotional neglect)
- Precipitating factors (eg conflict at work)
- Perpetuating factors (eg loneliness)
- Protective factors (eg strong social network)

Example of a case formulation

She is a 33-year old married woman, HR manager, mother of 3 children, who was referred for an emergency visit at the psychiatric ward by her GP because of mood problems and loss of functioning that started 4 weeks ago after giving birth. Her chief complaints are low mood, loss of interest, insomnia, appetite loss, fatigue, feelings of guilt, suicidal ideation, nightmares, feeling anxious. The symptoms increased in severity over a couple of weeks. Currently, she is not able to take care of herself or her children, and is dependent on her husband and family for food intake and selfcare.

The symptoms started after a complicated delivery with an emergency caesarian section, which led to substantial blood loss and anemia.

She has no psychiatric history except a few days of elevated mood after giving birth to her second child.

She has a family history of possible borderline personality disorder (mother) and bipolar disorder (father from father).

Her parents divorced when she was 11 years old, and she grew up in an emotionally unstable family, and may have suffered emotional neglect and sexual abuse.

Current problems may be maintained by anemia; protective factors are her husband and his family who take care of her and the children.

Example case formulation

She is a 33-year old married woman, HR manager, mother of 3 children, who was referred for an emergency visit at the psychiatric ward by her GP because of mood problems and loss of functioning that started 4 weeks ago after giving birth. Her chief complaints are low mood, loss of interest, insomnia, appetite loss, fatigue, feelings of guilt, suicidal ideation, nightmares, feeling anxious. The symptoms increased in severity over a couple of weeks. Currently, she is not able to take care of herself or her children, and is dependent on her husband and family for food intake and selfcare.

Presenting problem, nature and severity

The symptoms started after a complicated delivery with an emergency caesarian section, which led to substantial blood loss and anemia.

Precipitating factor

She has no psychiatric history except a few days of elevated mood after giving birth to her second child.

She has a family history of possible borderline personality disorder (mother) and bipolar disorder (father from father).

Predisposing factors

Her parents divorced when she was 11 years old, and she grew up in an emotionally unstable family, and may have suffered emotional neglect and sexual abuse.

Current problems may be maintained by anemia; protective factors are her husband and his family who take care of her and the children.

Perpetuating and protective factors

Missing data III: Causal relations

How?

- By assessing **temporality / order** of events (conflict at work -> sleep problems -> mood problems -> major depression)

- By assessing the **severity** of the event - is it **proportional** to the onset of the disorder?

- One formulates a tentative hypothesis about the causality in collaboration with the patient, to target treatment (e.g. resolve work conflict)

Another example: case conceptualization in CBT
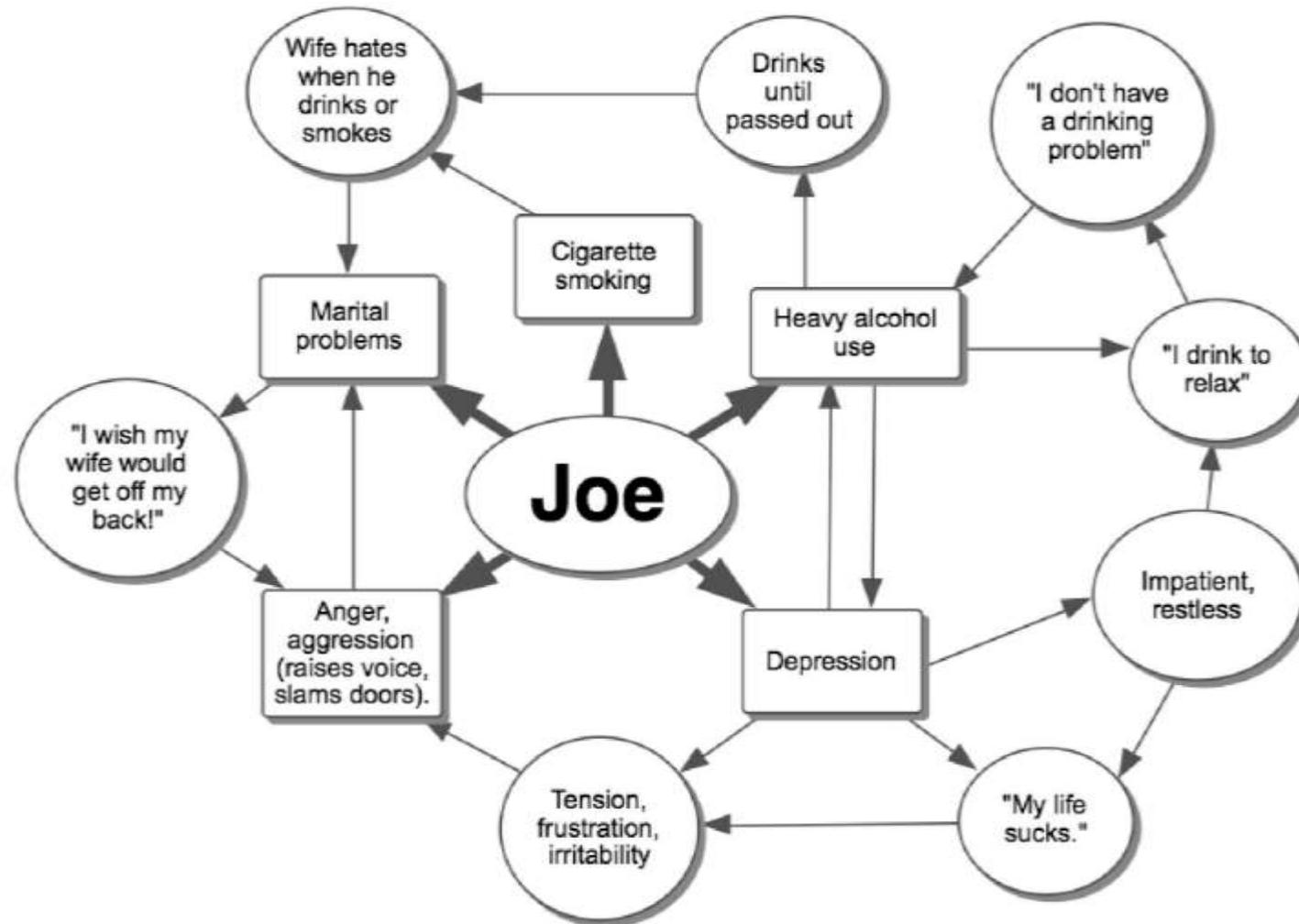
Missing data IV: Data may change due to increased self-understanding / reflection

- In a clinical interview, you don't take all answers for granted

- Example: Patient with depression denies impact of relationship problems on depressive symptoms.
  - Research shows that relationship problems often are associated with depression
  - By asking questions, patient starts to reconsider her own beliefs about impact of R on D.

- Or patient denies difficulties in the first place: "My parents are heroes"

SIGMUND **PETER DE WIT**

Papa en mama vinden me heel speciaal en zeggen dat ik superveel talent heb!

Nee hoor.

Je bent een doodgewoon meisje, niks bijzonders.

Later is ze me dankbaar.

4. Ways to enrich research data

1. Open questions
   - Meaning: "What did the covid lockdown mean to you?"
   - Can be analyzed with text mining: incl. type of words, number of words, intonation, speed

2. Video observations
   - E.g. behavior of parent during immunization child

3. Analyze behavior through smartphone applications (e.g. social media activity)

4. Establish temporal proximity and severity of life events
   - Causality

5. Use other informants (e.g. partner, relative)

Limitations in enriching research data

- Therapeutic relationship cannot be replaced
- Open questions expect certain degree of insight/self-awareness and ability to express this in written language
- Participants have limited time and energy
- Privacy issues (e.g. to use data from social media apps)
- Costly

5. Conclusions

- Data science methods hold promises for psychiatry

- However, these methods depend on the quality of the data

- The data and methods are in a sense too rich

- We argue that the data are also not rich enough: crucial information that can be obtained in a clinical interview is missing in research data

- This may limit the promises of these methods for psychiatry

Thank you for your attention

For questions and remarks:

j.w.romeijn@rug.nl

h.van.loo@umcg.nl