# Clinstix

## Scott Vrieze

### March 17, 2016

# 1 Dawes et al

## 1.1 Clinical versus actuarial prediction

1. We *collect* data and *interpret* data. Dawes et al only care about *interpreting* data after it has been collected.

2. Data, however collected, can be combined in two ways, *clinically* or *mechanically* (aka, actuarially, statistically, etc.).

   (a) *Clinical prediction* involves combining data "in one's head" to arrive at a prediction or decision.

   (b) *Mechanical prediction* uses a formal algorithm to combine data, preferably based on known relationships between those data.

   (c) Given the same input data, mechanical prediction will arrive at the same outcome every time. If one takes the output of a mechanical algorithm and "tweaks" it based on their clinical judgment, that chain of interpretive events becomes a clinical prediction, by definition. There is no "hybrid".

## 1.2 The "Goldberg Rule"

1. Differential diagnosis of "psychoses" versus "neuroses".

2. Mechanical rule (add 3 scales and subract 2) versus 29 judges on the same 861 MMPIs, including so-called experts. Seven settings.

   (a) Goldberg Rule beat the judges, including the best judge.

   (b) Even after extensive practice the Rule beat the judges.

   (c) Even after the judges had access to the Rule result, the Rule beat the judges.

3. Goldberg then created models of the judges' predictions. On new data, the model of the judges beat the judges themselves.

## 1.3 Hodgkin's

193 individuals died of Hodgkin's disease. Pathologist ratings of biopsy along multiple dimensions, as well as global rating of severity. Actuarial formula based on these ratings and survival.

1. Pathologists' ratings showed no association with survival.

2. Actuarial formula based on pathologists' ratings did show associations.

## 1.4 Mechanical predicton advantages seem general, not specific

- "There is no controversy in social sience that shows such a large body of qualitatively diverse studies coming out so uniformly... as this one"

- The result is general, it holds across domains, it holds when the clinician is given an informational edge such as interview impressions.

- What about rare countervailing events? The broken leg problem? When clinicians are allowed to countervail the algorithmic output they perform worse.

- What about utilities? They can be incorporated into the algorithm, of course, but research (afaik) on a clinical advantage based on utility has not addressed this. I'd say the burden is on the individuals making the claim that utilities show a clinical advantage.

## 1.5 Why do actuarial predictions win out?

- Reliability

- Algorithms faithfully weight variables according to observed relationships, take into account regression to the mean, and can incorporate base rates.

- Difficult to learn from experience as clinician. Often lack outcome data, recall bias of correct versus incorrect predictions.

- Despite the claim of "clinical experience", their clinical experience is often severely limited by number, limited to individuals with psychopathology, sheer memory constraints, and memory bias, compared to a coded database of predictors and outcomes in a large sample.

- Judges are more confident in their conclusions than warranted.

- Humans are bad calculators. Well, most of them.

## 1.6 Why no adoption?

- Lack of familiarity with evidence?

- Dehumanizing?

- But my patient is an individual.

- But I know something others don't - I'm a real expert. Just look, I have a PhD and people are forced to listen to me and take me seriously. (Sound familiar, students?)

# 2 Grove et al

## 2.1 Mechanical prediction outperforms clinical prediction

1. 63 (47%) studies favor mechanical prediction.

2. 65 studies show similar performance

3. 8 (6%) studies favor clinical prediction ($\alpha = .05$?)

4. Including interviews increases superiority of mechanical prediction ($ES = .2$ with interview versus .07 without). prediction.

5. Training and experience have no measurable impact on superiority of mechanical prediction.

6. Amount of data available to clinician had no effect.

7. Only consistent feature in the 8 studies where clinical outperformed was that clinician's received more data than the algorithm. (Perhaps the algorithm was missing critical data and a much better algorithm could be constructed?)