TAKING DIAGNOSTIC REASONING TO VEGAS: APPLICATION AND EVALUATION OF PROBABILISTIC THINKING USING A NOVEL EDUCATIONAL GAME

Nathan Stehouwer, MD

Lukasz Weiner, MD

Collin Swafford, MD

Keith Ponitz, MD

Andrea Scioscia, MD

Michael Dell, MD

University Hospitals
Cleveland | Ohio

Rainbow Babies
& Children's Hospital

Case Western Reserve University EST. 1826

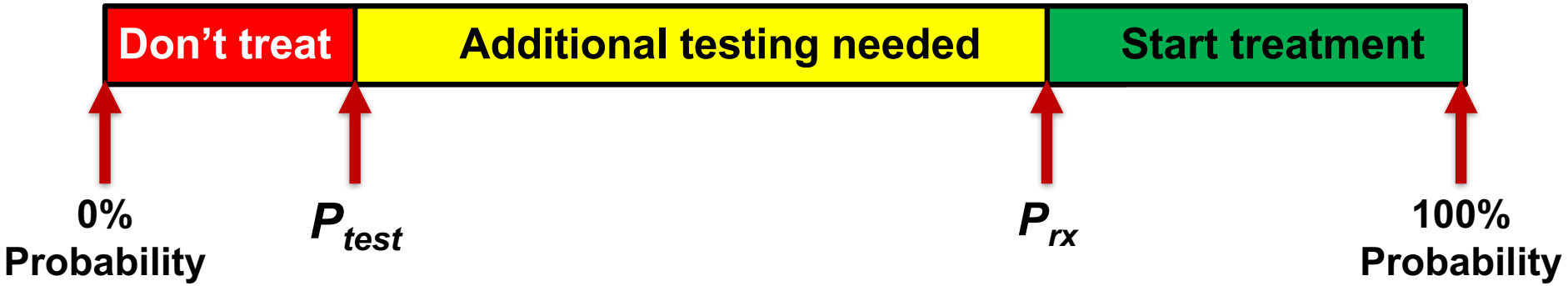
Goals of this workshop:

- List at least three educational benefits of teaching formats which encourage learners to consider probabilistic reasoning.
- Describe and apply a scoring system for probabilistic reasoning which rewards accurate appraisals of uncertainty.
- Implement a game of diagnostic reasoning, in which learners rate the probability of each item on their differential diagnosis, at didactic conferences at their home institutions.

Agenda

- Introduction to probabilistic reasoning
- Discuss Brier scoring and its application to differential diagnosis
- Participate in diagnostic reasoning game
- Debrief: Discuss benefits, challenges, and applications of this methodology

Decision-Making Model of Diagnostic Reasoning



How well do physicians think probabilistically?

- Unconscious probabilistic reasoning underlies “intuition” (Kreiter, 2017)
- Our usual preferred terms (probable, possible, unlikely, frequent) are unclear and may be interpreted differently by colleagues and by patients (Lurie, 1999)
- In one survey, **only 2 out of 300** clinicians surveyed reported thinking in terms of numerical probability (Grimes, 2005)

How well do physicians think

What is the probability (in percents) in your opinion, that the patient has:

- Acute coronary syndrome
- Dissecting aortic aneurysm
- Reflux esophagitis
- Biliary colic
- Anxiety disorder?
- Others

Source: Cahan, 2003

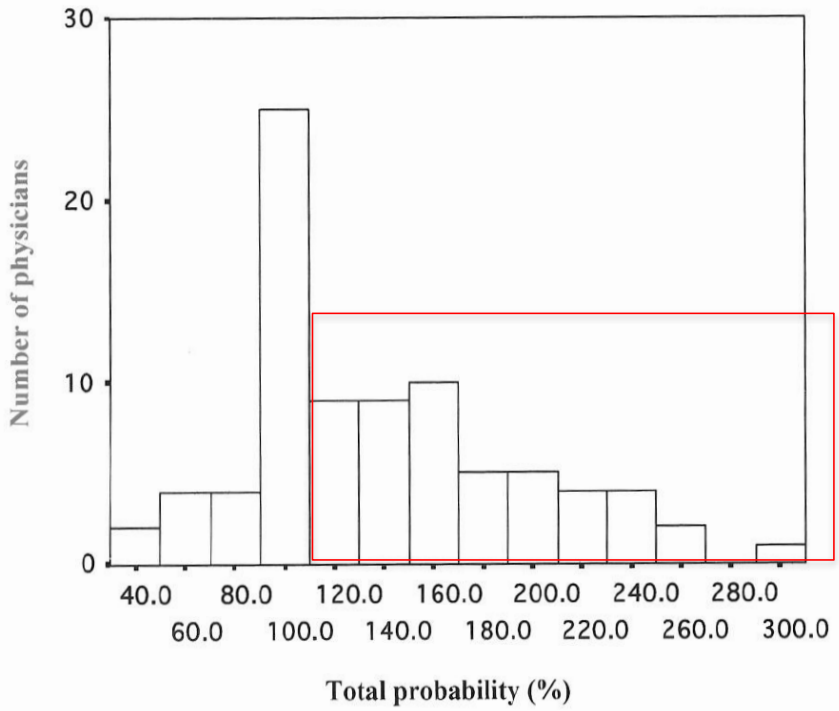


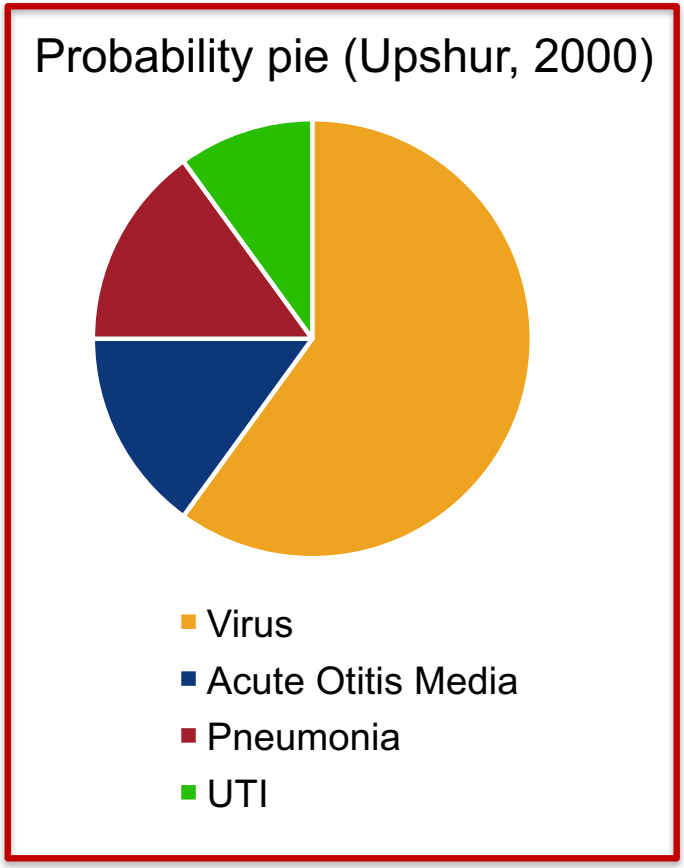
Figure 2. Frequency distribution of the total probabilities assigned by participants. The mean total probability was 136.7% (±53.9%). Sixty-five percent of participants had a total probability >100% (i.e. exhibited subadditivity).

Possible benefits of thinking probabilistically

- Appropriately determine when testing or treatment are indicated
- Improved interpretation of diagnostic test results (Bayesian inference)
  - Accurate pretest probability assessment largely negates the effect of false positive and negative results (Diamond, 1980)
- Improved risk/benefit assessments (Lurie, 1999)
- Make clinical uncertainty explicit and acceptable
  - Physicians may overstate diagnostic certainty, which appears to be negatively correlated with experience (Schoenherr, 2018)
- Encourage metacognition – corrective to System 1 biases

Can we teach probabilistic thinking?

- Decision tools
  - Students taught a prediction rule for estimating pretest probability of coronary artery disease (CAD) outperformed students given more traditional reading on CAD (Hickam, 1987)
  - Such tools limited by being:
    - Binary
    - Tied to prevalence in past population studied
    - Specific to one question at a time
- Bayesian inference supports sound clinical reasoning (Kreiter, 2017) but is challenging to objectively assess



FiveThirtyEight

how well you did against FiveThirtyEight and everybody else playing the game.

Your picks | Your results | Leaderboards

Week 20

2/2 forecasts completed

Sunday, Jan. 20 at 3:05 p.m.

L.A. Rams (Away) — Your forecast 81% — 64% FiveThirtyEight's forecast — **New Orleans** (Home)

You'll lose **81.2** points if the L.A. Rams win.
You'll gain **42.8** points if **New Orleans** wins.

Sunday, Jan. 20 at 6:40 p.m.

New England — Your forecast 55% — 61% FiveThirtyEight's forecast — **Kansas City**

You'll lose **10.4** points if New England wins.
You'll gain **9.6** points if **Kansas City** wins.

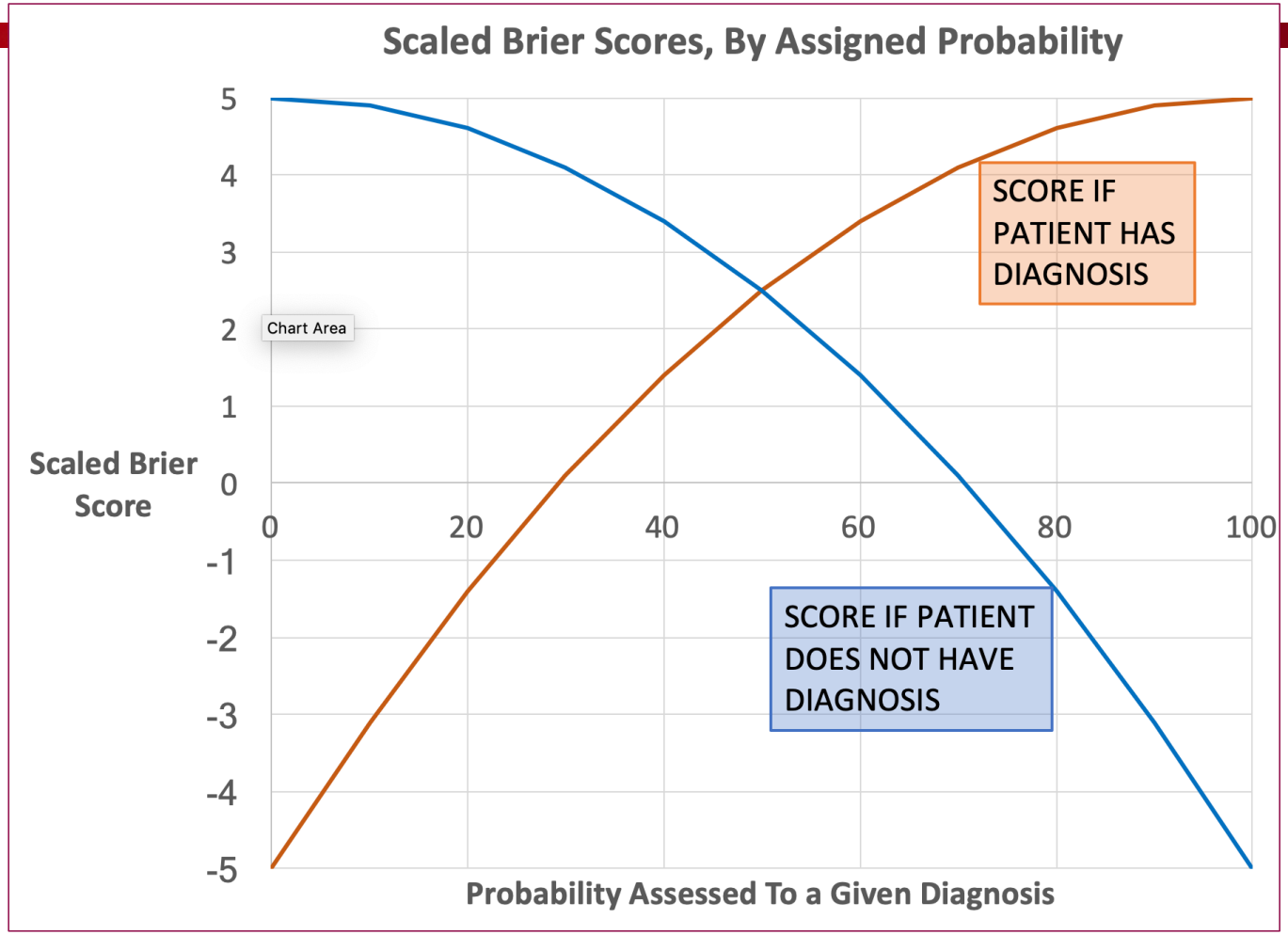
Source: FiveThirtyEight.com

Key Innovation – Application of Brier Scoring to Differential Diagnosis

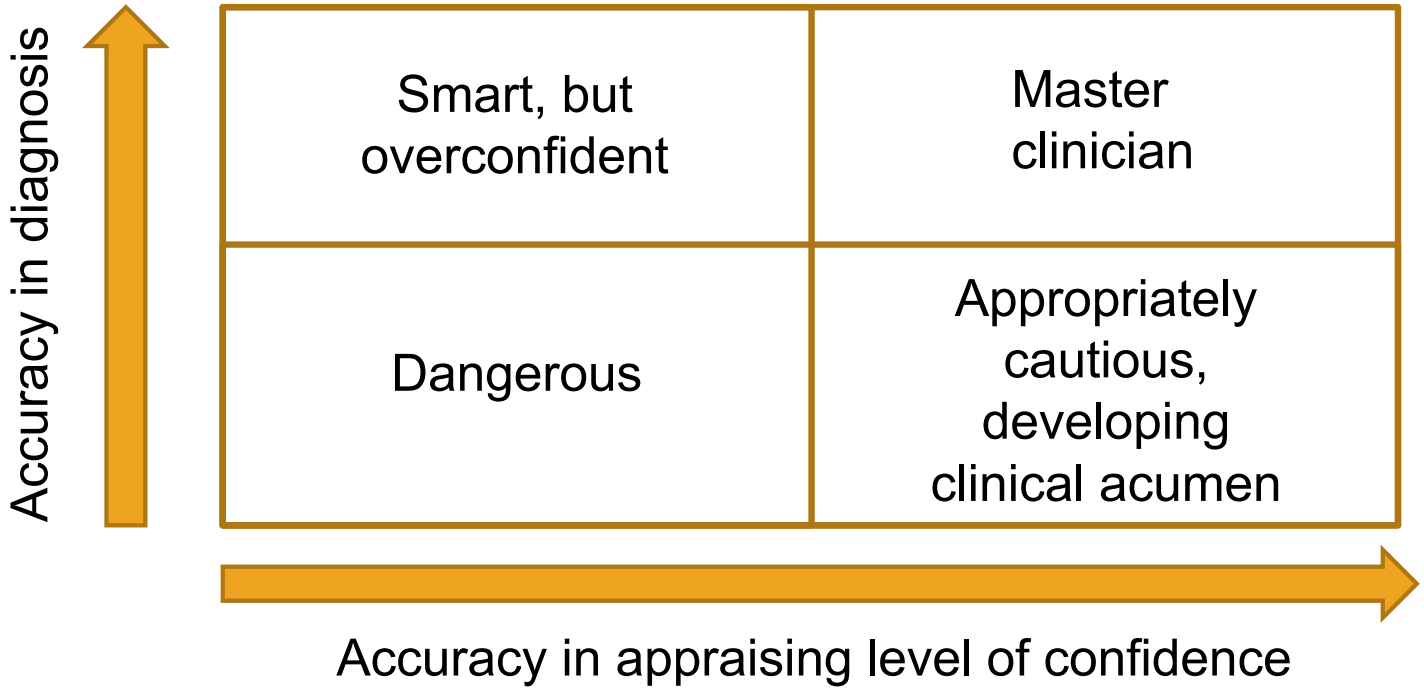
- Sum of the squares of the difference between predicted and actual outcome

$$BS = \frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{R}(f_{ti} - o_{ti})^2$$

- Score is affected by:
  - Uncertainty: Is the event easy to predict *a priori* (the child is very unlikely to have coronary artery disease) or difficult to predict (the child might or might not have pneumonia)
  - Reliability: Do I predict bronchiolitis at an overall rate similar to what is observed?
  - Resolution: How well are predictions adjusted to particular circumstances?


Scaled Brier Scores, By Assigned Probability
Scaled Brier Score
5
4
3
2
1
0
-1
-2
-3
-4
-5
0
20
40
60
80
100
Chart Area
SCORE IF PATIENT HAS DIAGNOSIS
SCORE IF PATIENT DOES NOT HAVE DIAGNOSIS
Probability Assessed To a Given Diagnosis

The score is a function of BOTH picking the right diagnosis, AND assessing appropriate level of confidence



Example: Card Drawing Predictions

Probability of…	Result	Player 1	Score	Player 2	Score	Player 3	Score	Player 4	Score	True probability	Score
Red card in 1 draw	YES	50%	2.5	50%	2.5	50%	2.5	50%	2.5	50%	2.5
One 2, 3, or 4 card in 1 draw	YES	20%	-1.4	23%	-0.9	20%	-1.4	22%	-1.1	23%	-0.9
At least 1 face card in 1 draw	NO	50%	2.5	23%	4.5	40%	3.4	45%	3.0	23%	4.5
2 Face Cards in 2 draws	NO	12%	4.9	5%	5.0	20%	4.6	8%	4.9	5%	5.0
2 Black Cards in 2 draws	NO	25%	4.4	25%	4.4	25%	4.4	25%	4.4	25%	4.4
One black, one red in 2 draws	YES	25%	-0.6	25%	-0.6	20%	-1.4	50%	2.5	51%	2.6
exactly one red card in 3 draws	YES	12%	-2.7	50%	2.5	60%	3.4	25%	-0.6	38%	1.2
At least one 5 in 3 draws	NO	20%	4.6	8%	4.9	10%	4.9	10%	4.9	22%	4.5
One pair in 5 card draw	YES	28%	-0.2	6%	-3.8	1%	-4.8	15%	-2.2	42%	1.6
Two pair in 5 card draw	NO	10%	4.9	0.5%	5.0	0.5%	5.0	5%	5.0	5%	5.0
No face cards or aces	NO	3%	5.0	75%	-0.6	1%	5.0	3%	5.0	15%	4.8
		Player 1 Total:	**23.8**	**Player 2 Total:**	**22.7**	**Player 3 Total:**	**25.6**	**Player 4 Total**	**28.2**	**Score**	**35.1**

On to the fun stuff!

- Get in teams of 5-6 people
- Designate one team leader who will enter answers
- Use computer (ideal) or phone and access the Google sheet which was sent out in advance
- You will be given a short description of an **actual case**
- You will have only about 2 minutes per case – please discuss, then enter estimated probabilities of each diagnosis in percentage.
- In these cases, **only one diagnosis is correct, so probabilities should add up to 100%**
- “Other” may well be the correct diagnosis

CASES

- Attendees will participate in a game of probabilistic diagnostic reasoning
- Cases will not be distributed in advance and are not included in this powerpoint

Benefits of this teaching tool?

Applications

Keys to to the probabilistic reasoning game conference

- Use actual case series – not invented cases
- Choose representative cases - do not select for unexpected twists (not morning report cases”)
- Need a high volume of cases for scoring to be valid – quick discussion on each, try to pare down case presentation to minimum necessary information
- Low/high probability events require a greater number of cases
- Diagnostic categories need to be broad enough that one can reasonably narrow down to 4-5 likely categories (i.e., “pneumonia” rather than “pneumococcal pneumonia.”)

Future questions

- Does this exercise improve assessment of confidence?
- What methods (for example, emphasizing epidemiology, or teaching “anchor and adjust” method) can improve assessment of pretest probability?
- Can ability to accurately describe probabilities translate into improved clinical decision making?

“Though some have argued that doctors are not good gamblers, by more clearly identifying the issues and sources of uncertainty and using probabilistic thinking to guide reasoned bets, physicians can become better consumers of data” (Jenssen et al, 2015)

References

- Cahan A, Gilon D, Manor O, Paltiel O. Probabilistic reasoning and clinical decision-making: do doctors overestimate diagnostic probabilities? Q J Med 2003; 96:763-769.
- Diamond GA, Forrester JS, Hirsch M, Staniloff HM, Vas R, Berman DS, Swan HJ. Application of conditional probability analysis to the clinical diagnosis of coronary artery disease. J Clin Invest 1980; 65(5):1210-21.
- "2018 NFL Predictions." FiveThirtyEight.com. Accessed January 18, 2019. https://projects.fivethirtyeight.com/2018-nfl-predictions/
- Grimes DA, Schulz KF. Refining clinical diagnosis with likelihood ratios: Lancet 2005; 365: 1500-1505.
- Hickam DH, Sox HC. Teaching Medical Students to Estimate Probability of Coronary Artery Disease. JGIM 1987; 2:73-77.
- Jenssen BP, Kenyon CC. Probability, Uncertainty, and Value in Inpatient Diagnosis: Connecting the Dots. Hospital Pediatrics 2015; 5(7): 403-405.
- Kreiter C. A Bayesian perspective on constructing a written assessment of probabilistic clinical reasoning in experienced clinicians. Journal of Evaluation in Clinical Practice 2017; 23:44-48.
- Lurie JD, Sox HD. Principles of Medical Decision Making. Spine 1999; 24(5): 493-498.
- Schoenherr JR, Waechter J, Millington SJ. Subjective awareness of ultrasound expertise development: individual experience as a determinant of overconfidence. Adv in Health Sci Educ 2018; 23: 749-765.
- Upshur REG. Two Techniques for Teaching the Estimation of Prior Probabilities. Teach Learn Med 2000; 12(3): 141-4