Why We Should be Excited about the Cross-Cohort Collaboration?

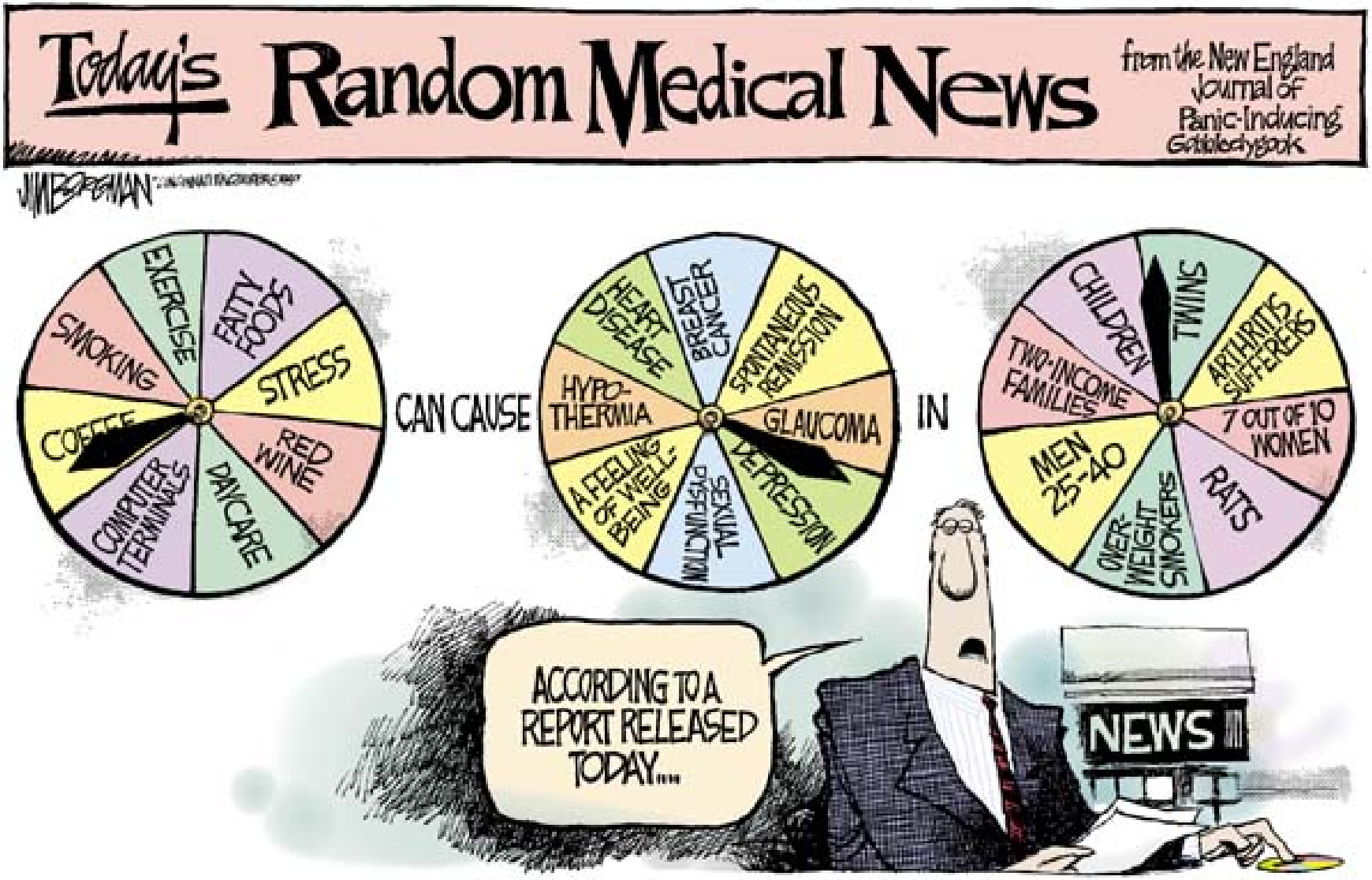
George Howard
for the REGARDS Investigators

School of Public Health
University of Alabama at Birmingham

What is to be Gained in the Cross-Cohort Collaboration

- Together, we can build a remarkable data resource
- But why?
- Statistical power to finally address opportunities to study:
  - Rare exposures
  - Rare outcomes
  - Challenging study designs and hypotheses
- Broadens opportunity for assessing environmental exposures (geographic heterogeneity)
- Avoid making mistakes (specifically – type 1 errors)

We have a (potentially well-earned) problem of reporting spurious results



Replication as Standard Approach in Genetics

- The challenge of spurious findings arising from multiple testing in genetics has lead to the standard approach of “learning” and “confirmation” cohorts
- The cross-cohort collaboration offers the opportunity for similar approach for more general epidemiological investigations

Approaches

- "Genetic" approach: easy confirmation of observations in other cohorts
- "Pooling" approach: use of split sample replication
  - Pool available data
  - Randomly split into learning and confirmation cohorts
  - Perform "aggressive" analysis in learning cohort (embracing data mining techniques)
  - But avoid spurious findings by confirmation in remaining patients

Examples of Opportunities

- Inclusion of broad spectrum of predictors in risk functions
- Assessing which component of exposure contributes to risk (diet exposures, medication responses, etc)
- Environmental exposures available at the county level

Planning

- Challenge will be deciding what items are in the “common data”
- Only when we get into analysis will we discover what we should be sharing
- Can we prospectively plan an iterative approach with process to update variables at regular interval?

NIH/NINDS U01 NS41588