Group testing with incomplete block designs and imperfect tests

Pranta Das
University of Nebraska–Lincoln
Department of Statistics

Research is supported by NIH grant R01 AI121351

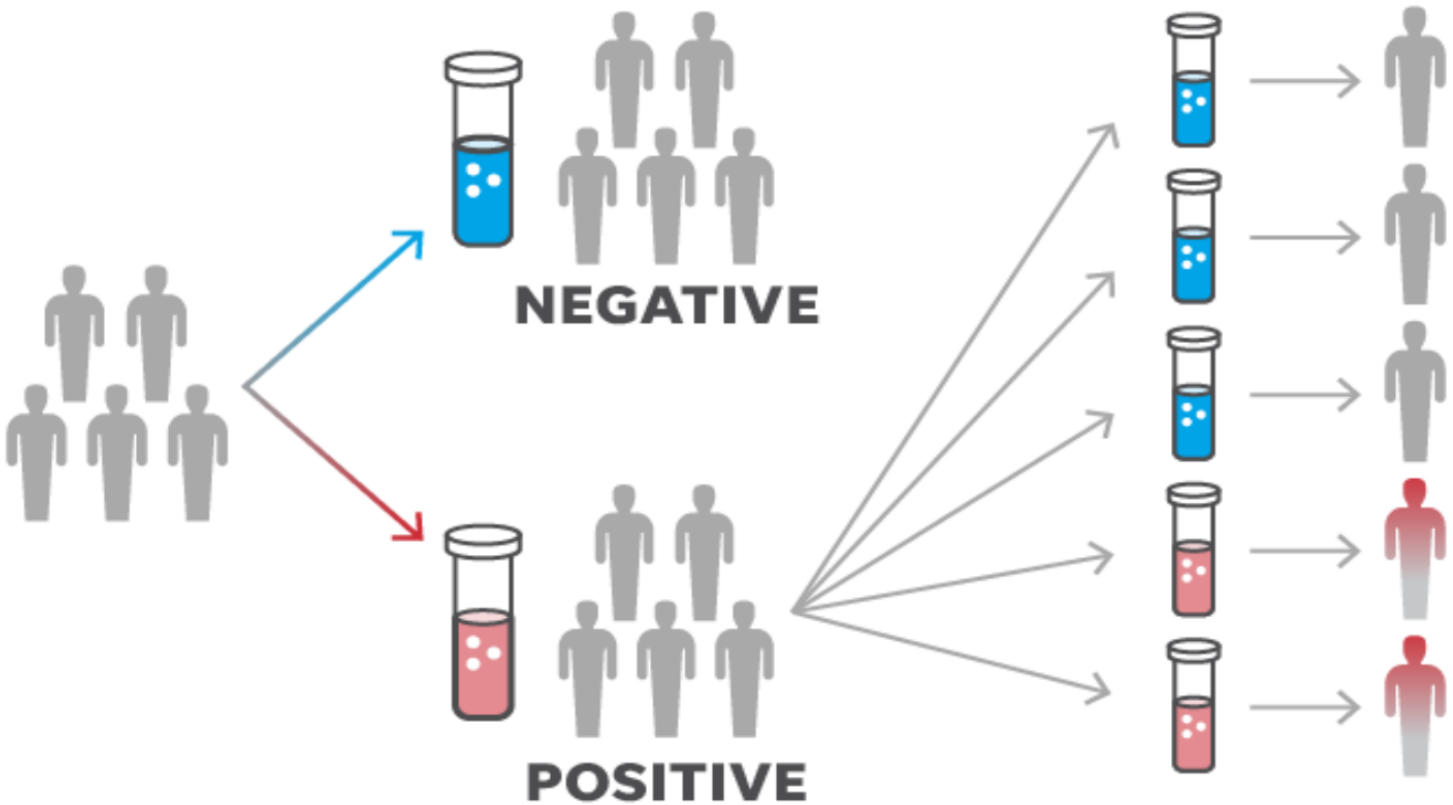
Joint work with
Christopher R. Bilder at University of Nebraska-Lincoln,
Joshua M. Tebbs at University of South Carolina, and
Christopher S. McMahan at Clemson University

June 10, 2024

pdas8@huskers.unl.edu

What is Group testing?

- Also known as "pooled testing"
- Screening for infectious diseases



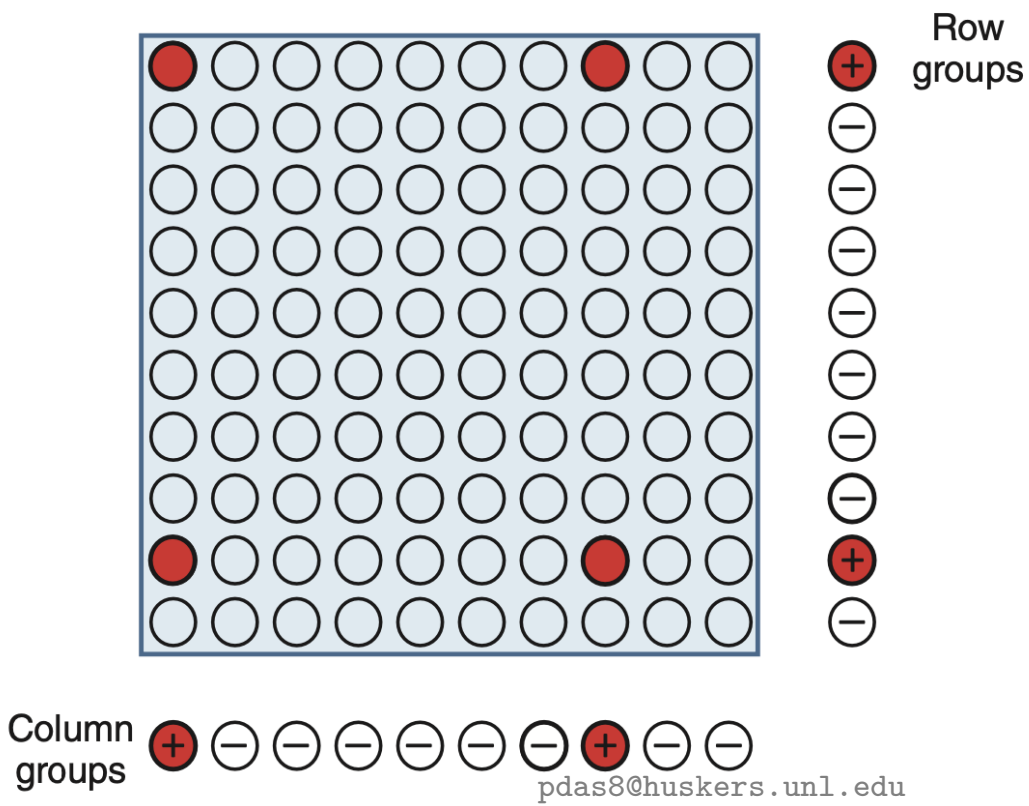
SOURCE USA TODAY research
Karl Gelles/USA TODAY

Importance

- Decrease number of tests
  - Increase in testing capacity
  - Works when disease prevalence is low
- A recent instance: COVID-19 pandemic
  - News media outlets: *The New York Times*, *The Washington Post*, ABC News, PBS NewsHour, USA Today, ...
  - Over 1 GB of papers published during the first 2 years of the pandemic!
    - Abdalhamid et al. (*American Journal of Clinical Pathology*, 2020)
    - Hogan et al. (*Journal of the American Medical Association*, 2020)
    - Lohse et al. (*Lancet Infectious Diseases*, 2020)
- Widely used elsewhere: Blood donation screening, computer network assessments, ...

Algorithms

- Hierarchical algorithm
  - Dorfman testing (two stages)
  - Three stages and higher are possible
- Non-hierarchical algorithm
  - Array testing



Algorithms

- Non-hierarchical algorithm (continued)
  - Incomplete block designs (IBDs)

	Treatments					
Blocks	1	2	3	4	5	6
1	1	0	1	0	1	0
2	0	1	0	1	1	0
3	0	1	1	0	0	1
4	1	0	0	1	0	1

  - Group testing setting

	Specimens					
Groups	1	2	3	4	5	6
1	1	0	1	0	1	0
2	0	1	0	1	1	0
3	0	1	1	0	0	1
4	1	0	0	1	0	1

Purpose

- Previous work using IBDs
  - Sudbury (*Biometrics*, 2010) and Eskridge et al. (*Biotechnology Progress*, 2018)
  - Did not incorporate **false positive** or **false negative**
- Our purpose
  - Examine IBDs when false positive or false negative possible
  - Compare with Dorfman testing and array testing
  - Use the expected number of tests as the performance measure

M matrix

- Matrix for specimen arrangements

M:

	Specimens					
Groups	1	2	3	4	5	6
1	1	0	1	0	1	0
2	0	1	0	1	1	0
3	0	1	1	0	0	1
4	1	0	0	1	0	1

- Referred as "**M** without immediate ambiguity"

M matrix

- With immediate ambiguity

M:

	Specimens						
Groups	1	2	3	4	5	6	7
1	1	0	1	0	1	0	1
2	0	1	1	0	0	1	1
3	0	1	0	1	1	0	0
4	1	0	0	1	0	1	0

  - Specimens 3 and 7 in same groups

Comparisons

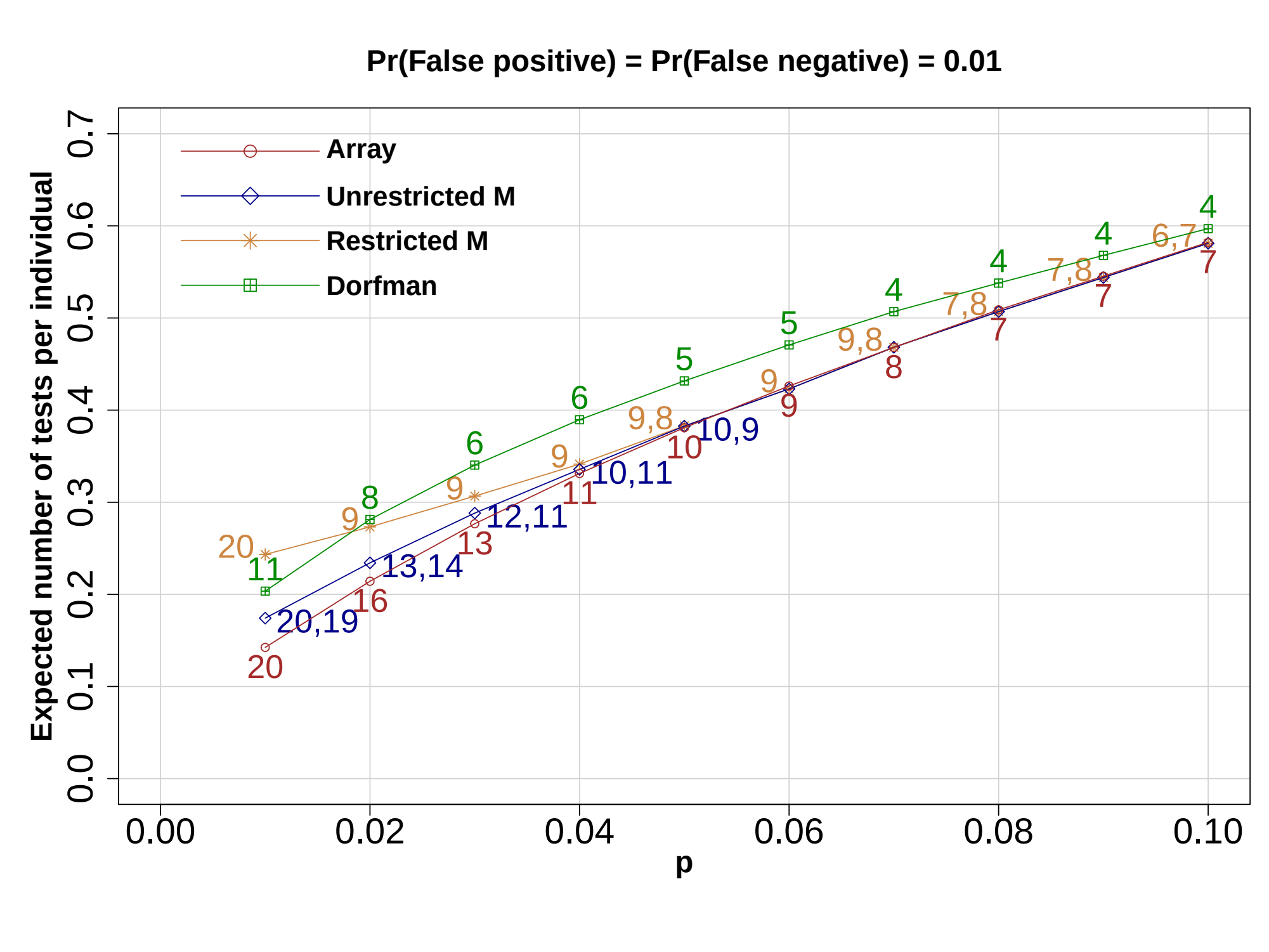
- Cases of **M** and other algorithms
  - IBD using restricted **M** (without ambiguity)
  - IBD using unrestricted **M** (with or without ambiguity)
  - Dorfman testing
  - Array testing
- Expected number of tests per individual
  - Let T be the number of tests for an initial set of I individuals
    - For Dorfman, stage 1 group size
    - For array, number of individuals in an array
    - For IBD, number of individuals represented by **M**
  - $\mathrm{E}(T)/I$ is the expected number of tests per individual
    - $\mathrm{E}(T)/I = 1$ when testing each individual separately
    - Want $\mathrm{E}(T)/I < 1$ for group testing
    - Example: $\mathrm{E}(T)/I = 0.25$ ...
  - Closed form expressions exist for Dorfman and array testing (Kim et al., *Biometrics*, 2007)

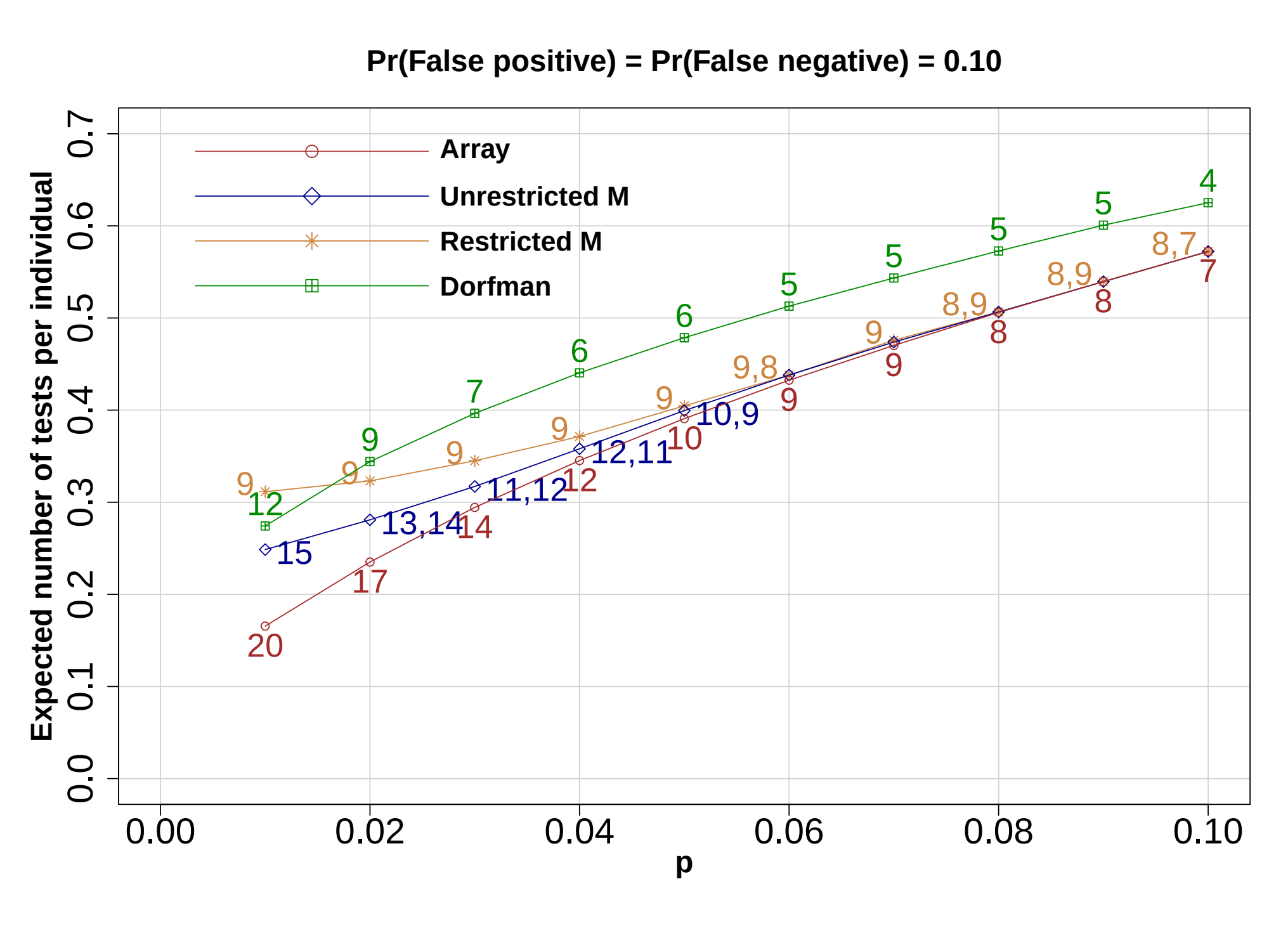
Comparisons

- A closed form expression for E(T)/I does not exist currently for IBDs
  - Simulate the testing process for an **M**
  - Use Monte Carlo simulation
  - Repeat this process 10,000 times and observe number of tests each time
  - Estimate E(T)/I
- **Best algorithm** has the smallest E(T)/I

Steps

1. Find optimal testing configuration (OTC) for each algorithm at a prevalence (p)
   - $p = 0.01, 0.02, ..., 0.10$
   - $Pr(\text{False positive}) = Pr(\text{False negative}) = 0.01, 0.10$
   - IBD
     - **Number of specimens:** 5 - 60
     - **Number of groups:** 4 - 10
     - **Number of groups a specimen can be in:** 2 - 4
   - Dorfman testing
     - **Group size:** 2 - 100
   - Array testing:
     - **Group size:** 2 - 20
2. Compare OTCs by expected number of tests per individual


Pr(False positive) = Pr(False negative) = 0.01
Array
Unrestricted M
Restricted M
Dorfman
Expected number of tests per individual
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.00
0.02
0.04
0.06
0.08
0.10
p
20
16
13
11
9
8
6
5
4
20,19
13,14
12,11
10,11
10,9
9,8
7,8
6,7
9
8
7


Pr(False positive) = Pr(False negative) = 0.10
Array
Unrestricted M
Restricted M
Dorfman
Expected number of tests per individual
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.00
0.02
0.04
0.06
0.08
0.10
p
20
17
14
12
9
8
7
15
13,14
11,12
12,11
10,9
9,8
8,9
8,7
12
9
9
9
9
9
9
9
9
7
6
6
5
5
5
5
4

Group testing with incomplete block designs and imperfect tests

Pranta Das
University of Nebraska–Lincoln
Department of Statistics

Research is supported by NIH grant R01 AI121351

Joint work with
Christopher R. Bilder at University of Nebraska-Lincoln,
Joshua M. Tebbs at University of South Carolina, and
Christopher S. McMahan at Clemson University

June 10, 2024

pdas8@huskers.unl.edu