

# A user-friendly Shiny web application for choosing pool sizes when testing pooled specimens

Christopher R. Bilder<sup>1</sup>, Brianna D. Hitt<sup>1</sup>, Joshua M. Tebbs<sup>2</sup>, Christopher S. McMahan<sup>3</sup>  
<sup>1</sup>University of Nebraska-Lincoln, <sup>2</sup>University of South Carolina, <sup>3</sup>Clemson University

## Introduction

### Abstract

**Background:** High volume screening of clinical specimens for infectious diseases is often made possible by a process known as pooled testing. This algorithmic process involves testing portions of specimens from separate individuals together as one unit (or “pool”) to detect infection. Follow-up retesting is performed on members of positive-testing pools to decode the positive members from the negative ones. An important decision needed prior to implementation of pooling is what pool size to use. Choosing too large of a pool size can lead to a large number of retests, perhaps even resulting in a total number of tests larger than what would occur by individually testing specimens. Choosing too small of a pool size can also lead to an overall larger number of tests than necessary.

**Methods:** To help laboratories choose a pool size, we developed a Shiny web application that leverages the power of the R statistical software package. This application uses analytical derivations and computer simulation methods to emulate the pooling process. A user-friendly web interface is provided so that users do not need experience with R to perform calculations.

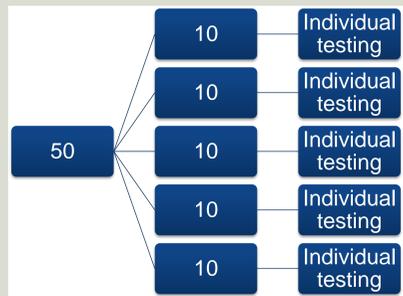
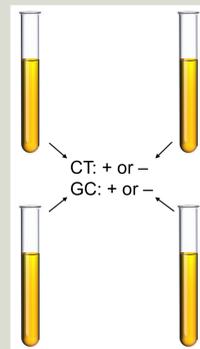
**Results and conclusions:** The application calculates the expected number of tests and the expected accuracy for commonly used pooling algorithms. This application can also determine the “optimal” pool size(s) based on minimizing the expected number of tests. Access to the application is available through our [www.chrisbilder.com/shiny](http://www.chrisbilder.com/shiny) website.

**Corresponding author:** Christopher R. Bilder, [chris@chrisbilder.com](mailto:chris@chrisbilder.com), [www.chrisbilder.com](http://www.chrisbilder.com)

This research was supported by Grant R01 AI121351 from the National Institutes of Health.

## What is pooled testing?

- Also known as *group testing* and *specimen pooling*
- Example #1: Chlamydia (CT) and gonorrhea (GC) testing with the Aptima Combo 2 Assay at the University of Iowa’s State Hygienic Laboratory (SHL)
  - An amalgamation of specimens from 4 individuals is a pool
  - If a pool tests negatively for both diseases, all individuals within it are declared disease free
  - If a pool tests positively for at least one disease:
    - Need to determine who is positive and who is negative for which diseases
    - SHL retests all pool members individually with the same assay; thus, a **2-stage hierarchical testing algorithm** overall
  - Estimated savings during a recent 5-year evaluation period was approximately \$3 million
- Example #2: HIV testing in San Francisco with a single-disease assay and a **3-stage hierarchical testing algorithm** (Sherlock et al. 2007)
  - Stage 1 pools are formed from 50 separate individuals
  - If a pool tests positively, 5 non-overlapping subpools are formed from its members, where each subpool has a size of 10
  - If a subpool tests positively, each member is tested individually
  - When a pool/subpool tests negatively, all of its members are declared disease free and their testing ends



## Expected number of tests

### Input

Enter information about assay and overall disease prevalence

**Pooled Testing**

Hierarchical testing

Calculate the operating characteristics for a given configuration

**Specifications**

How many diseases for the assay?  1  2

What is the sensitivity of the assay?

How many stages for the pooling algorithm?  2  3

What is the overall disease prevalence?

What is the specificity of the assay?

What is the initial pool size?  Note: The minimum size allowed is 3.

Operating Characteristics | Algorithm Diagram | Results

## Output

Select “Calculate” to produce the screen below

**Operating Characteristics** | Algorithm Diagram | Results

**6.70**  
Expected number of tests

**0.45**  
Expected number of tests per individual

The configuration has an initial pool size of 15 followed by individual testing.

The expected number of tests for the configuration is 6.70, and the expected number of tests per individual is 0.45.

For the *overall* implementation of the algorithm, the sensitivity is 0.9025, the specificity is 0.9819, the positive predictive value is 0.6063, and the negative predictive value is 0.9969.

The inputs used for these calculations were

- Diseases: 1
- Overall disease probability: 0.03
- Sensitivity: 0.95
- Specificity: 0.95
- Stages: 2
- Initial pool size: 15

## Optimal pool sizes

### Input/Output

Determine the pool sizes that minimize the expected number of tests (optimal testing configuration)

**Pooled Testing**

Hierarchical testing

Find the optimal testing configuration

**Specifications**

How many diseases for the assay?  1  2

What is the sensitivity of the assay?

How many stages for the pooling algorithm?  2  3

What is the overall disease prevalence?

What is the specificity of the assay?

Please specify a range of initial pool sizes.

OTC | Algorithm Diagram | Similar configurations | Results

**3.35**  
Expected number of tests

**0.28**  
Expected number of tests per individual

The optimal testing configuration (OTC) has an initial pool size of 12 and subsequent pool sizes of 4, 4, and 4.

OTC | Algorithm Diagram | Similar configurations | Results

Initial pool size	Second stage pool sizes	Expected number of tests	Expected number of tests per individual	Overall sensitivity	Overall specificity	Overall positive predictive value	Overall negative predictive value
12	4,4,4	3.3488	0.2791	0.8574	0.9955	0.855	0.9956
16	4,4,4,4	4.4756	0.2797	0.8574	0.9953	0.8499	0.9956
13	5,4,4	3.6451	0.2804	0.8574	0.995	0.8414	0.9956