

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Faculty Publications, Department of Statistics

Statistics, Department of

3-2019

Informative Group Testing for Multiplex Assays

Christopher R. Bilder

Joshua M. Tebbs

Christopher S. McMahan

Follow this and additional works at: <https://digitalcommons.unl.edu/statisticsfacpub>



Part of the [Other Statistics and Probability Commons](#)

This Article is brought to you for free and open access by the Statistics, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, Department of Statistics by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.



Published in final edited form as:

Biometrics. 2019 March ; 75(1): 278–288. doi:10.1111/biom.12988.

Informative Group Testing for Multiplex Assays

Christopher R. Bilder^{1,*}, Joshua M. Tebbs², Christopher S. McMahan³

¹Department of Statistics, University of Nebraska-Lincoln, Lincoln, Nebraska 68583, U.S.A.

²Department of Statistics, University of South Carolina, Columbia, South Carolina 29208, U.S.A.

³Department of Mathematical Sciences, Clemson University, Clemson, South Carolina 29634, U.S.A.

Summary:

Infectious disease testing frequently takes advantage of two tools—group testing and multiplex assays—to make testing timely and cost effective. Until the work of Tebbs et al. (2013) and Hou et al. (2017), there was no research available to understand how best to apply these tools simultaneously. This recent work focused on applications where each individual is considered to be identical in terms of the probability of disease. However, risk-factor information, such as past behavior and presence of symptoms, is very often available on each individual to allow one to estimate individual-specific probabilities. The purpose of our paper is to propose the first group testing algorithms for multiplex assays that take advantage of individual risk-factor information as expressed by these probabilities. We show that our methods significantly reduce the number of tests required while preserving accuracy. Throughout this paper, we focus on applying our methods with the Aptima Combo 2 Assay that is used worldwide for chlamydia and gonorrhea screening.

Keywords

Case identification; Correlated binary data; Latent response; Pooled testing; Sensitivity; Specificity

1. Introduction

“Make do with less” is a constant mantra in today’s society. Governments and businesses all want to cut budgets without sacrificing services. Health care in particular is not immune to this philosophy despite its importance to personal well-being. Fortunately, infectious disease testing is one area of health care that is successfully making do with less. This is because laboratories are increasingly using group testing (also known as pooled testing) to provide reliable diagnostic testing in a timely manner and at lower costs. Group testing works by amalgamating specimens from individuals (e.g., blood or urine) into pools and performing

* chris@chrisbilder.com.

Supplementary materials

The referenced Web Appendices are included with this paper at the *Biometrics* website on Wiley Online Library. Our R programs are available from this website as well.

tests on these pools. When compared to testing each specimen individually, a substantial reduction in the number of tests occurs when a disease is rare, because multiple persons can be classified as disease free simultaneously when their group results in a negative outcome. Widespread health care applications of group testing include screening blood donations for infectious diseases (e.g., American Red Cross, 2018; Saá et al., 2018) and testing for chlamydia and gonorrhea as part of national sexually transmitted disease assessment programs (e.g., Lewis et al., 2012). Outside of health care, group testing is widely used in areas including infectious disease testing in animals (Nebraska Veterinary Diagnostic Center, 2018), plant disease assessment (Gildow et al., 2008), and drug discovery (Kainkaryam and Woolf, 2009).

Laboratories also make do with less by using multiplex assays. These assays test for multiple diseases in a single application rather than one-disease, one-test assays. Examples include the Procleix Ultrio Assay for HIV, hepatitis B, and hepatitis C and the Aptima Combo 2 Assay for chlamydia and gonorrhea. Multiplex assays used with group testing can be quite cost effective. For example, the State Hygienic Laboratory (SHL) at the University of Iowa uses group testing and the Aptima Combo 2 Assay with swab specimens. With thousands of individuals tested per year and each test costing about \$37, the SHL has estimated savings to be approximately 3 million dollars during a recent 5-year evaluation period.

The testing process used by the SHL is relatively similar to other laboratories. At the SHL, individual specimens are randomly assigned to groups of size four, and these groups are formed via a Tecan DTS robotic platform. The Aptima Combo 2 Assay is applied to each group to detect the ribosomal ribonucleic acid from the *Chlamydia trachomatis* and *Neisseria gonorrhoeae* bacteria. These bacteria are what lead to the chlamydia and gonorrhea diseases, respectively. If a group tests negatively for both diseases, all group members are declared negative. If a group tests positively for at least one disease, all group members are retested individually using the same multiplex assay. This same assay is used (rather than switching to single-disease assays) because it is much simpler for lab technicians to use one assay throughout the testing process.

Given the pervasiveness of group testing and multiplex assays in laboratories now, it is perhaps surprising that statistical research has not kept up in this area. Tebbs et al. (2013) were first to examine the use of group testing with multiplex assays. They focused on the two-stage algorithm used by the SHL, which represented a direct extension of the seminal Dorfman (1943) group-testing work for single-disease assays. Motivated by research that has shown other algorithms to often be more efficient (i.e., lower number of tests) than Dorfman's algorithm for single-disease assays, Hou et al. (2017) generalized Tebbs et al. (2013) to allow for more than two stages in a hierarchical manner (any positively testing group is divided into multiple subgroups to be tested in the next stage).

A restrictive assumption made by both Tebbs et al. (2013) and Hou et al. (2017) was that each individual tested had an equal probability of being positive for a particular disease. However, due to personal behavior or clinical observations, it is natural to think of some individuals as being at higher risk for disease than others. The purpose of our paper here is

to develop the first *informative group testing* algorithms to exploit this type of information with multiplex assays and to show these algorithms can lead to a substantial improvement in testing efficiency. Informative group testing was introduced by Bilder et al. (2010) but has only been examined in the context of single-disease assays (e.g., Black et al., 2015, Liu et al., 2017). The multiplex assay setting is much more complicated because disease statuses are likely to be correlated for each individual. Along with most specimens not being tested individually and assays being imperfect, this leads to unobservable, correlated binary random variables representing individual disease statuses as the underlying stochastic framework for this challenging problem.

We develop the first informative group testing algorithms for multiplex assays as follows. Section 2 develops a new algorithm that can be applied in three or more stages. We derive the expected number of tests and accuracy measures for the algorithm. These operating characteristics are used to determine how best to implement it for a specific application. Section 3 also develops a new informative group testing algorithm but for only two stages. We draw connections between this algorithm and the algorithm from Section 2, which will make clear why this order of development was chosen. Section 4 examines these algorithms through a Monte Carlo simulation study. We provide insight regarding under what conditions these new algorithms perform best, while also comparing to the *non-informative group testing* work of Tebbs et al. (2013) and Hou et al. (2017). Section 5 applies our new algorithms in the context of the Aptima Combo 2 Assay and its use at laboratories in Idaho, Iowa, and Oregon. We show that our new algorithms can substantially reduce the number of tests needed. Finally, Section 6 summarizes our work and offers ideas for extensions.

2. Hierarchical testing in three or more stages

Because group testing is used to reduce the time and cost of testing, choosing the exact algorithm to implement is an extremely important decision for a laboratory. Factors that affect time and cost include: 1) selection of an initial (stage 1) group size, 2) number of subgroups and their sizes at each stage, 3) number of stages, and 4) choosing what individuals are in which subgroups. These factors can lead to an extremely large number of possible testing configurations, which makes developing a general notation to be somewhat cumbersome. To address these issues, we introduce a *group membership matrix* \mathbf{M} as a simple tool to uniquely define a hierarchical group testing algorithm for an initial group of I individuals. The rows of this matrix correspond to stages $s = 1, \dots, S$, and the columns correspond to individuals $i = 1, \dots, I$. Cell values represent the group number of individual i at stage s . For example, consider a three-stage algorithm with a group size of $I = 30$ in stage 1, three groups of size 10 in stage 2, and individual testing in stage 3. This algorithm is described by Sherlock et al. (2007) for HIV testing in the Seattle area. The group membership matrix is

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 & 21 & 22 & 23 & 24 & 25 & 26 & 27 & 28 & 29 & 30 \end{bmatrix}.$$

Although this algorithm is easily described in words, others are not, especially when group sizes are unequal or when individuals are not tested in later stages. We provide examples of these more complicated testing configurations in Web Appendix A. It is also important to note that \mathbf{M} represents the totality of testing that would be performed if each group tests positively until individual testing. Therefore, actual implementation may result in fewer tests. For example, in the previous algorithm, testing would end at stage 1 if the initial group tests negatively for all diseases.

With so many testing configurations possible, it is desirable to implement one which is optimal in some manner. We choose to minimize the expected number of tests per individual and define the resulting configuration with its corresponding \mathbf{M} as the *optimal testing configuration* (OTC). Alternative optimal criteria have been suggested for single-disease group testing, and we provide a brief discussion of them in Web Appendix B.

2.1 Expected number of tests

Consider an initial group of I individuals. We represent the number of individuals within group j at stage s by I_{sj} ($I_{11} \equiv I$) and the number of possible groups at stage s by c_s . If group j at stage s tests positively for at least one disease (or a pathogen that can lead to a disease), it is divided into m_{sj} groups. Thus, in the context of the three-stage algorithm just described, $I_{11} = 30$, $c_1 = 1$, and $m_{11} = 3$ for stage 1. Continuing for stage 2, $I_{2j} = 10$, $c_2 = 3$, and $m_{2j} = 10$ for $j = 1, 2, 3$. Lastly, for stage 3, $I_{3j} = 1$, $c_3 = 30$, and $m_{3j} = 0$ for $j = 1, \dots, 30$. This precise representation of group membership is needed to provide a general expression for the expected number of tests.

Define T as a random variable denoting the total number of tests needed to determine which individuals are positive and negative in an initial group of size I . Disease statuses are regarded as being independent across individuals, but we allow for dependence among disease statuses for each individual. All individuals are hierarchically tested in successively smaller groups in up to S stages until they are tested alone or their group tests negatively for each disease. At each stage, an individual is a member of only one group. Define G_{sjk} as a binary random variable denoting the test result for the k th disease ($k = 1, \dots, K$) in group j at stage s , where $G_{sjk} = 1$ (0) denotes a positive (negative) test result. For example, the group at stage 1 may result in $(G_{111}, G_{112}) = (0, 1)$, meaning that the initial group tested negatively for the first disease and tested positively for the second disease. Following current group testing practice with multiplex assays, we assume that if a group tests positively for at least one disease, its members will be retested for all diseases again in the next stage.

To account for successive stage dependence, we define $G_{sjk}^{(t)}$ as the *ancestor group* result for G_{sjk} at stage $t \leq s$. These ancestor groups represent the groups that need to be tested at earlier stages in order for group j at stage s to be tested. Thus, individuals tested to produce the G_{sjk} response are a subset of the individuals that produce the $G_{sjk}^{(t)}$ response. This notation also means that $G_{sjk}^{(s)} \equiv G_{sjk}$, where the ancestor form is used at times for notational convenience. For example, $G_{32k} \equiv G_{32k}^{(3)}$ (group 2 in stage 3) for the previously given three-stage algorithm has ancestor pool results of $G_{32k}^{(2)} \equiv G_{21k}$ and $G_{32k}^{(1)} \equiv G_{21k}^{(1)} \equiv G_{11k}$. With this

definition of ancestor groups, the expected number of tests using group membership matrix \mathbf{M} can be expressed as

$$E(T | \mathbf{M}) = 1 + \sum_{s=1}^{S-1} \sum_{j=1}^{c_s} m_{sj} P(G_{sj+}^{(1)} > 0, \dots, G_{sj+}^{(s)} > 0), \quad (1)$$

where $G_{sj+}^{(t)} = G_{sj1}^{(t)} + \dots + G_{sjK}^{(t)}$ for $t = 1, \dots, s$. The joint probability in Equation (1) represents the successive testing that occurs up to stage s . For example, in the previously given three-stage algorithm, the second individual is tested at stage 3 with probability $P(G_{21+}^{(1)} > 0, G_{21+}^{(2)} > 0) = P(G_{11+} > 0, G_{21+} > 0)$. The leading 1 in Equation (1) is included because the group at stage 1 is always tested. This group tests positively for at least one disease with probability $P(G_{s1+}^{(1)} > 0) = P(G_{11+} > 0)$, and a positive outcome leads to m_{11} new tests at stage 2. Subsequent stages have m_{sj} new tests that occur with probability $P(G_{sj+}^{(1)} > 0, \dots, G_{sj+}^{(s)} > 0)$.

While Equation (1) is relatively simple in form, calculating the joint probability within it is challenging for a number of reasons. First, we need to account for potential testing error because assays are rarely 100% accurate. Thus, while we observe G_{sjk} , we would really like to observe \tilde{G}_{sjk} , say, denoting the true positive (1) and negative (0) group status. Second, these disease statuses for each group at each stage are correlated. For example, \tilde{G}_{111} and \tilde{G}_{112} are correlated random variables giving the initial group's statuses for two diseases. Third, due to the testing hierarchy, each response at stage s is dependent on its ancestor group response at stage $s-1$. Lastly and most importantly, each individual has potentially a different probability of being positive for the diseases. By defining the true status for individual i and disease k as $\tilde{Y}_{ik} = 1(0)$ for a positive (negative) status, we can denote these *joint probabilities of disease* for individual i as $P(\tilde{Y}_{i1} = \tilde{y}_1, \dots, \tilde{Y}_{iK} = \tilde{y}_K) = p_{i\tilde{y}_1 \dots \tilde{y}_K}$, where $\tilde{y}_1, \dots, \tilde{y}_K \in \{0, 1\}$. These individual-specific probabilities will be used for informative group testing in Section 2.2 to take advantage of the risk-factor information available.

For the above reasons, the derivation of $P(G_{sj+}^{(1)} > 0, \dots, G_{sj+}^{(s)} > 0)$ is quite long and requires special attention to understand its nuances. We provide this derivation in Web Appendix C.1. In summary, $P(G_{sj+}^{(1)} > 0, \dots, G_{sj+}^{(s)} > 0)$ can be shown to be equal to

$$\begin{aligned} & \sum_{\mathbf{g}_{sj}^{(1)}} \dots \sum_{\mathbf{g}_{sj}^{(s)}} \left[\prod_{s'=1}^s \left\{ 1 - \prod_{k=1}^K \left(1 - S_{e:sjk}^{(s')} \right)^{\tilde{g}_{sjk}^{(s')}} \left(S_{p:sjk}^{(s')} \right)^{1 - \tilde{g}_{sjk}^{(s')}} \right\} \right] P(\tilde{\mathbf{G}}_{sj}^{(1)} = \tilde{\mathbf{g}}_{sj}^{(1)}) \\ & \times \prod_{s'=1}^{s-1} P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = \tilde{\mathbf{g}}_{sj}^{(s'+1)} | \tilde{\mathbf{G}}_{sj}^{(s')} = \tilde{\mathbf{g}}_{sj}^{(s')}), \end{aligned} \quad (2)$$

where $\tilde{\mathbf{G}}_{sj}^{(t)} = (\tilde{G}_{sj1}^{(t)}, \dots, \tilde{G}_{sjK}^{(t)})$ is a vector of ancestor true positive/negative statuses for group j at stage t and $\tilde{\mathbf{g}}_{sj}^{(t)}$ is a vector of corresponding 0 and 1 potential realizations for each

$\tilde{G}_{sjk}^{(t)}, k = 1, \dots, K$. Expressions for $P(\tilde{G}_{sj}^{(1)} = \tilde{g}_{sj}^{(1)})$ and $P(\tilde{G}_{sj}^{(s'+1)} = \tilde{g}_{sj}^{(s'+1)} \mid \tilde{G}_{sj}^{(s')} = \tilde{g}_{sj}^{(s')})$ are given in Web Appendix C.1. It is important to note that these probability expressions are functions of $p_{i\tilde{y}_1 \dots \tilde{y}_K}$. Values for the sensitivity $S_{e:sjk}^{(s')} = P(G_{sjk}^{(s')} = 1 \mid \tilde{G}_{sjk}^{(s')} = 1)$ and the specificity $S_{p:sjk}^{(s')} = P(G_{sjk}^{(s')} = 0 \mid \tilde{G}_{sjk}^{(s')} = 0)$ are obtained from large validation trials performed by the assay manufacturer. Properly calibrated assays will typically have the same accuracies for each group at each stage, especially for nucleic acid amplification tests. However, our general notation allows for potential differences if needed, such as when dilution is a concern or when different assays are used within the testing algorithm.

We provide an R function in Web Appendix D to calculate $E(T|\mathbf{M})$. Being able to calculate $E(T|\mathbf{M})$ in closed form allows us to address optimal ways to apply group testing with multiplex assays, which we describe next.

2.2 Optimal testing configuration

The OTC is the most efficient group testing algorithm possible. It is given by a group membership matrix that minimizes $E(T|\mathbf{M})/I$ over all possible \mathbf{M} for a specified number of stages S and for a range of values for I to consider. Equipment constraints, assay detection thresholds, and time limitations all play important roles for laboratories when deciding on maximum values for I and S to include in a search for the OTC. Since the year 2000, we have seen at most 90 individuals pooled (Sherlock et al., 2007) and at most 4 stages used (Quinn et al., 2000) for human infectious disease testing. Most frequently, I is less than 50 and S is 2 or 3. For these reasons, one can most often enumerate over all possible \mathbf{M} and calculate $E(T|\mathbf{M})/I$ for each to find the OTC. Combinatorial optimization algorithms can be used instead in those rare cases when laboratories are without these types of constraints.

To search for the OTC, we adopt similar conventions as used by Hwang (1975), McMahan et al. (2012a), Black et al. (2012), and Black et al. (2015) for single-disease assays. In summary, they proposed to order individuals by their probabilities of disease (say, $p_{i\tilde{y}_1}$) and subsequently assign individuals in sequence (smallest to largest probabilities of being positive) to groups of the same or smaller size. For example, with $I = 7$, a testing configuration could have the individuals with the four smallest probabilities in the first group, the individuals with the next two smallest probabilities in the second group, and the individual with the largest probability tested alone. One can show mathematically that the OTC will always occur with this type of group assignment in special cases (no testing error, Hwang, 1975; halving of groups, Black et al., 2012). These conventions are also very intuitive toward the ultimate goal of minimizing the number of tests. For example, consider the most extreme situation when individuals either have a probability of 0 or 1 for being positive and assume there are no testing errors. The most efficient way to form groups is to assign all individuals with a probability of 0 to one group and all individuals with a probability of 1 to groups of size one. As values for these probabilities deviate from 0 and 1, one would naturally want to test those individuals with small (large) probabilities of disease in large (small) groups.

For multiplex assays, we now order individuals by their probabilities of being truly positive for at least one disease; e.g., when $K = 2$, form an ordering based on $1 - p_{\bar{0}0}$ values from smallest to largest. Individuals based on this ordering are assigned to the corresponding columns of \mathbf{M} . For subsequent stages, these ordered individuals are sequentially assigned again to groups of equal or smaller size when searching for the OTC. It is important to note that $E(T|\mathbf{M})$ is a function of all joint probabilities of disease for each individual i , e.g., $p_{\bar{0}0}$, $p_{\bar{0}1}$, $p_{\bar{1}0}$, and $p_{\bar{1}1}$ when $K = 2$. Thus, these probabilities play an important role in determining which configuration is best.

Of course, these probabilities of disease will not be known in actual application. Instead, they can be estimated by an appropriate regression model, with covariates incorporating the available risk-factor information. When prior testing data are available, a model can be trained upon these data and then applied to estimate probabilities for future individuals. We use this approach in Section 5 to obtain the OTC in our Aptima Combo 2 Assay application.

2.3 Accuracy

While $E(T|\mathbf{M})$ plays a critical role in determining which group testing algorithm to use, it is also important to characterize an algorithm's accuracy before recommendations are made for implementation. Of course, an algorithm with a small $E(T|\mathbf{M})$ may not be desirable for use without a sufficient level of accuracy. Two measures that describe accuracy are pooling sensitivity and pooling specificity. For individual i and disease k , the pooling sensitivity $PS_{e,ik}$ is the probability that the individual is diagnosed as positive through group testing given this individual is truly positive. Pooling specificity $PS_{p,ik}$ is defined analogously for a negative diagnosis given a truly negative individual.

Unlike in Tebbs et al. (2013) and Hou et al. (2017), these accuracy measures will likely be different for each individual due to the available risk-factor information. Therefore, in addition to characterizing an algorithm's accuracy before implementation, these measures could allow one to isolate incorrect diagnoses after implementation. For this purpose, we define the pooling positive predictive value $PPPV_{ik}$ as the probability that individual i is truly positive for disease k given a positive diagnosis. The pooling negative predictive value $PNPV_{ik}$ is defined analogously for a truly negative individual given a negative diagnosis. To illustrate the importance of predictive values, suppose individual i is diagnosed as positive for disease k and has $PPPV_{ik} = 0.25$. Arguably, a positive response is somewhat incongruous with this predictive value, and thus one may want to perform a confirmatory test. Although we do not develop a formal framework for this type of testing, it is an important potential advantage of our approach.

To express these accuracy measures, define $\tilde{\mathbf{G}}_{sj\bar{k}}$ as a vector of true positive/negative (1/0) statuses for group j at stage s , where \tilde{G}_{sjk} is omitted. Define $\tilde{\mathbf{g}}_{sj\bar{k}}$ in a similar manner for the potential realizations. Our derivation in Web Appendix C.2 shows the pooling sensitivity to be

$$\begin{aligned}
PS_{e:ik} &= P(G_{Lj+}^{(1)} > 0, \dots, G_{Lj+}^{(L-1)} > 0, G_{Ljk} = 1 \mid \tilde{Y}_{ik} = 1) \\
&= \sum_{\tilde{\mathbf{g}}_{Ljk}^{(1)}} \dots \sum_{\tilde{\mathbf{g}}_{Ljk}^{(L-1)}} S_{e:Ljk}^{(L)} \left[\prod_{s'=1}^{L-1} \left\{ 1 - \left(1 - S_{e:Ljk}^{(s')} \right. \right. \right. \\
&\quad \left. \left. \left. \prod_{q=1, q \neq k}^K \left(1 - S_{e:Ljq}^{(s')} \right)^{\tilde{\mathbf{g}}_{Ljq}^{(s')}} \left(S_{p:Ljq}^{(s')} \right)^{1 - \tilde{\mathbf{g}}_{Ljq}^{(s')}} \right\} \right] \\
&\quad \times P(\tilde{\mathbf{G}}_{Ljk}^{(1)} = \tilde{\mathbf{g}}_{Ljk}^{(1)} \mid \tilde{Y}_{ik} = 1) \prod_{s'=1}^{L-2} P(\tilde{\mathbf{G}}_{Ljk}^{(s'+1)} = \tilde{\mathbf{g}}_{Ljk}^{(s'+1)} \mid \tilde{\mathbf{G}}_{Ljk}^{(s')} = \tilde{\mathbf{g}}_{Ljk}^{(s')}, \tilde{Y}_{ik} = 1),
\end{aligned} \tag{3}$$

where L ($L \leq S$) denotes the stage at which individual i tests positively in the j th group that has a size of 1. The conditional probabilities in Equation (3), which are also provided in Web Appendix C.2, depend on the joint probabilities of disease for individual i , making $PS_{e:ik}$ likely unequal for each individual.

To find the pooling specificity, define Y_{ik} as a binary response denoting the positive (1) or negative (0) diagnosis for individual i with respect to disease k . We can simplify the process of finding the pooling specificity by writing it in terms of the pooling sensitivity:

$$PS_{p:ik} = P(Y_{ik} = 0 \mid \tilde{Y}_{ik} = 0) = 1 - \frac{P(Y_{ik} = 1) - PS_{e:ik}P(\tilde{Y}_{ik} = 1)}{1 - P(\tilde{Y}_{ik} = 1)},$$

where the derivation is given in Web Appendix C.2. The derivation for $P(Y_{ik} = 1)$ is very similar to the derivation of Equation (2), so we provide this in Web Appendix C.2 as well. The marginal probability $P(\tilde{Y}_{ik} = 1)$ is simply obtained by summing over the corresponding joint probabilities for $p_{i\tilde{y}_1 \dots \tilde{y}_K}$.

Predictive values are found through a standard application of Bayes' rule. The pooling positive predictive value for individual i and disease k is

$$PPPV_{ik} = P(\tilde{Y}_{ik} = 1 \mid Y_{ik} = 1) = \frac{P(\tilde{Y}_{ik} = 1)PS_{e:ik}}{P(\tilde{Y}_{ik} = 1)PS_{e:ik} + P(\tilde{Y}_{ik} = 0)(1 - PS_{p:ik})}.$$

The pooling negative predictive value is

$$PNPV_{ik} = P(\tilde{Y}_{ik} = 0 \mid Y_{ik} = 0) = \frac{P(\tilde{Y}_{ik} = 0)PS_{p:ik}}{P(\tilde{Y}_{ik} = 0)PS_{p:ik} + P(\tilde{Y}_{ik} = 1)(1 - PS_{e:ik})}.$$

3. Two-stage testing

Section 2 proposes an informative group testing algorithm for $S \geq 3$ hierarchical stages. By taking advantage of the hierarchy, one can incorporate a two-stage algorithm ($S = 2$) into this same framework with minor modifications. Specifically, two-stage testing is essentially three-stage testing with the initial group of the three stages not tested. Thus, for two-stage testing, a set of I individuals are immediately split into separate groups of size I_{1j} , $j = 1, \dots$,

c_1 , without presuming that at least one positive individual exists. This allows the I individuals to be optimally placed into groups by taking into account their joint probabilities of disease. Web Appendix A provides an example of a group membership matrix. Web Appendix F.2 provides an example of how to apply a group membership matrix for testing.

The expected number of tests for two-stage testing can be expressed as

$$E(T | \mathbf{M}) = c_1 + \sum_{j=1}^{c_1} m_{1j} P(G_{1j+} > 0).$$

Accuracy measures for two-stage testing follow from Section 2. The process to find the OTC also follows from Section 2.

4. Algorithm performance

We use Monte Carlo simulation to evaluate our proposed algorithms by calculating their expected number of tests and accuracy measures over a number of situations. Our focus is on $S = 3$ and $K = 2$ due to algorithm similarities for different S and K . For these investigations, suppose $\mathbf{P}_i = (P_{i00}, P_{i01}, P_{i10}, P_{i11}) \stackrel{iid}{\sim}$ Dirichlet($\boldsymbol{\alpha}$) represents probabilities of disease for individual i . Defining \mathbf{P}_i as a vector of random variables, rather than using the constant $p_{i\tilde{y}_1\tilde{y}_2}$ values as in previous sections, allows us to examine different combinations of probabilities of disease in a controlled manner and to emulate variation that would occur in practice from one individual to the next. For our simulations, we set $E(P_{i1+})$ and $E(P_{i+1})$, the expected values for the marginal probability of being positive for diseases 1 and 2, respectively, to particular fixed values. Similarly, we control the variability of these marginal probabilities with $Var(P_{i1+})$ and $Var(P_{i+1})$.

In our first set of investigations, we simulate \mathbf{P}_i with $E(P_{i1+}) = E(P_{i+1}) = 0.05$ to match low disease prevalence situations where group testing is most often used. There are an infinite number of $\boldsymbol{\alpha}$ values with these same expectations, so we begin with $\boldsymbol{\alpha} = (18.25, 0.75, 0.75, 0.25)$ to emulate variability levels among marginal probabilities that we have seen in practice. Note that large values for the first element of $\boldsymbol{\alpha}$ are needed because they lead to large values of p_{i00} (probability of being disease free) relative to the other joint probabilities.

Our investigations also include two other cases with $\mathbf{P}_i \stackrel{iid}{\sim}$ Dirichlet($4\boldsymbol{\alpha}$) and

$\mathbf{P}_i \stackrel{iid}{\sim}$ Dirichlet($\boldsymbol{\alpha}/4$) that result in $E(P_{i1+}) = E(P_{i+1}) = 0.05$. The former case reduces $Var(P_{i1+})$ and $Var(P_{i+1})$, while the latter case increases $Var(P_{i1+})$ and $Var(P_{i+1})$ relative to using $\mathbf{P}_i \stackrel{iid}{\sim}$ Dirichlet($\boldsymbol{\alpha}$). We simply refer to these three informative group testing cases as “low variability” ($4\boldsymbol{\alpha}$), “medium variability” ($\boldsymbol{\alpha}$), and “high variability” ($\boldsymbol{\alpha}/4$); summaries of these distributions are given in Web Appendix E. For each case, I random realizations of \mathbf{P}_i are obtained, and the OTC is found for these realizations using the methods described in Section 2.2. We repeat this process of simulating I random realizations of \mathbf{P}_i and finding the OTC for a total of 500 times. We summarize each set of simulations by averaging over the expected number of tests per individual. The assay sensitivity and specificity are set to 0.95 and 0.99, respectively, for both diseases and for all groups tested in each stage.

Our investigations also include comparisons to the non-informative group testing work of Hou et al. (2017). We refer to this case as “homogeneous” because all individuals would be assumed to have the same disease probabilities, so that $\text{Var}(P_{i1+}) = \text{Var}(P_{i+1}) = 0$. Therefore, we set $E(\mathbf{P}_j)$ with $\mathbf{a} = (18.25, 0.75, 0.75, 0.25)$ to be the realization of \mathbf{P}_j for each individual. The dependence between disease statuses for this case is given by an odds ratio, $OR = 8.11$. By defining a pseudo odds ratio for the three other cases as $\overline{OR} = E(P_{i00})E(P_{i11})/[E(P_{i01})E(P_{i10})]$, we see that $\overline{OR} = 8.11$ for each as well, so our comparisons are with similar amounts of dependence. With respect to finding the OTC for the algorithm of Hou et al. (2017), we make one small adjustment to it by allowing for unequal group sizes in stage 2. This adjustment enables comparisons to be made over a larger set of configurations.

Figure 1 displays the averaged $E(I|\mathbf{M})/I$ values at $I = 5, \dots, 20$ when \mathbf{M} is chosen to correspond to the OTC at that particular I . This figure demonstrates that the reduction in the number of tests for informative group testing can be very significant (up to 18% for high variability) when compared to non-informative group testing (homogeneous case). This reduction is further magnified when considering that thousands or even millions of individuals are tested in high-volume clinical specimen settings when group testing is necessary due to time and cost constraints. Figure 1 also shows that as the variability among P_{i1+} and P_{i+1} increases, informative group testing becomes more efficient. This happens because larger observed p_{i1+} and p_{i+1} may occur and informative group testing algorithms can exploit this better through using different group sizes in stage 2. Figure 2 illustrates these differently sized groups by plotting their average sizes in a stacked bar chart. The plot shows that more groups and larger variability in their sizes occur as the variability increases for P_{i1+} and P_{i+1} .

The clear advantage of using our proposals extends to other situations as well, including for $E(P_{i1+}) = E(P_{i+1}) = 0.01$ with $\overline{OR} = 43.67$ and for $E(P_{i1+}) = E(P_{i+1}) = 0.10$ with $\overline{OR} = 3.67$ that are shown in Web Appendix E. This web appendix also provides results where we allow \overline{OR} to vary among informative group testing cases, but keep $E(P_{i1+}) = E(P_{i+1})$ and $\text{Var}(P_{i1+}) = \text{Var}(P_{i+1})$ constant. These investigations show that as the strength of positive dependence ($\overline{OR} > 1$) increases, the average expected number of tests per individual decreases. This occurs due to positive individuals being more likely to have more than one disease as the dependence increases.

Figure 3 summarizes accuracy for $\mathbf{a} = (18.25, 0.75, 0.75, 0.25)$ through the use of box plots at each I . In total, there are $500I$ different accuracy measures summarized in each box plot, and these measures are only for disease $k = 1$ ($k = 2$ had similar results). Depending on which measure is summarized, there can be a small or large amount of variability, although each is reasonable in the context of what is calculated. For example, the pooling positive predictive values range from 0 to 1. Low values are due to low observed values for p_{i1+} that can occur when $E(P_{i1+})$ is low as well. Similarly, most pooling negative predictive values are close to 1. We also include in each plot a line connected across different I to represent the accuracy obtained by the homogeneous case, where again we allow for unequal group sizes in stage 2. Because the accuracy measures depend on stage 2 group sizes, this leads to

different values for some individuals, so we simply displayed the mean value. Overall, the accuracy of informative group testing and non-informative group testing (homogeneous case) is quite comparable when just focusing on the displayed means and medians. Additional plots in Web Appendix E show similar findings for $E(P_{I1+}) = E(P_{I+1}) = 0.01$ and for $E(P_{I1+}) = E(P_{I+1}) = 0.10$.

5. Aptima Combo 2 Assay Application

High-volume testing of clinical specimens for chlamydia and gonorrhea occurs at laboratories in each state of the US, where many laboratories have already transitioned to group testing with multiplex assays. A very frequently used multiplex assay in this setting is the Aptima Combo 2 Assay. To determine how to apply this assay, one would ideally implement different versions of a group testing algorithm (e.g., informative or non-informative, different I and S) simultaneously upon the same specimens and, for future use, choose the algorithm that performs best. Unfortunately, this is not practical due to the high cost of assays, the limited amount of material available for each specimen, and the exceedingly large overall implementation time. Therefore, we instead use statewide retrospective data for the years 2010 and 2011 from Idaho and Oregon and for the years 2013 and 2014 from Iowa as a basis to emulate how testing would be performed.

Details regarding the data content and emulation process are available in Web Appendix F. In summary, we use the former year as training data and the latter year as test data for each state. Separate training/test data sets are formed within states by gender due to how manufacturers test their assays and how laboratories often perform their testing. For each individual, the data consist of final chlamydia and gonorrhea diagnoses, age, personal behavior information (e.g., risk history, patient reported symptoms), and clinical observations (e.g., urethritis, cervicitis). Multinomial regression models are estimated to each training data to obtain probability estimates of chlamydia and/or gonorrhea (i.e., estimates of p_{D00} , p_{D01} , p_{D10} , and p_{D11}). Following the recommendations of Black et al. (2015) for informative group testing with single-disease assays, we use a “non-adaptive” process with these probability estimates to obtain one overall estimated OTC from each training data. To summarize, this process obtains mean estimates of p_{D00} , p_{D01} , p_{D10} , and p_{D11} over $i = 1, \dots, I$ sets and uses these estimates to find the most efficient testing configuration. Non-informative algorithms are implemented in the same manner to obtain estimated OTCs, except only the chlamydia and gonorrhea diagnoses are used to obtain the probability estimates.

These estimated OTCs are subsequently applied to I successive members of the test data to determine positive/negative outcomes for both diseases via group testing. To account for assay imperfection, we incorporate the assay’s sensitivity and specificity (manufacturer provided) into the analysis. We do this by treating the final diagnoses given in the test data as the “true” statuses, and then simulate the group and individual responses that could occur while implementing group testing. This process is necessary because true disease statuses are not observable. Each application of a group testing algorithm is simulated 500 times to account for assay imperfection. Our reported results are averaged over these 500 applications.

Table 1 displays the results. To help understand the table's contents, examine the entries for Idaho females tested over three stages. Informative group testing places the 4,168 individuals successively into non-overlapping initial groups of size $I = 8$ with an estimated OTC given by the group membership matrix of

$$\mathbf{M} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 2 & 2 & 2 & 3 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & \text{NA} \end{bmatrix}$$

as shown in Web Table 4. This \mathbf{M} was also provided as the first example in Web Appendix A. Table 1 shows that 1,927.1 tests on average were needed for informative group testing, while 2,029.9 tests on average were needed for non-informative group testing. Thus, the reduction in the number of tests is 5.1% when using informative rather than non-informative. Other state and gender combinations can result in a much greater reduction. For example, testing males in Idaho over two stages results in a 14.7% reduction in tests on average. In summary, the mean number of tests is reduced in all but one case when comparing informative to non-informative group testing, and this reduction is much greater for males than for females.

Performance differences between males and females in Table 1 are explained by examining the distributions of the joint probabilities of disease. We provide summaries of these distributions in Web Appendix F.1. Overall, the variability of these probabilities for males is much larger than for females, leading to a better algorithm performance for males. This variability result matches what was shown in Section 4 relative to $\text{Var}(P_{i1+})$ and $\text{Var}(P_{i+1})$. Also, it is interesting to note that the reduction in the mean number of tests is generally less for Iowa than for Idaho and Oregon. This occurs because there is less information available on individuals in Iowa than there is for those in Idaho and Oregon (see Web Tables 1 and 2).

Web Appendix F.4 provides the observed accuracy for the informative and non-informative algorithms. Overall, the numerical values are very similar among the algorithms, indicating that accuracy is not reduced by using informative group testing.

6. Discussion

Our research shows that substantial reductions in the number of tests can occur when using informative group testing with multiplex assays. These reductions are directly related to the joint probabilities of disease: as the variability in the marginal probabilities increases across individuals, the overall number of tests decreases. Also, these reductions occur without loss of accuracy in comparison to non-informative algorithms. To make our proposals readily available, we provide R functions to find $E(T|\mathbf{M})$, accuracy measures, and the OTC in Web Appendix D and at our research website www.chrisbilder.com/grouptesting. These functions can also be useful to decide if a potential reduction in tests from three or more stages is enough, relative to practicality of implementation, to forgo using only two stages.

The use of informative group testing with multiplex assays leads to a number of intriguing future research problems. First, McMahan et al. (2012b) proposed informative group testing algorithms for single-disease assays applied to specimens configured in a two-dimensional

array structure. We expect a similar development as with our hierarchical approach could be realized for multiplex assays applied in arrays. Second, there are multiple modeling methods that can be used to estimate the joint probabilities of disease. While we applied a multinomial regression model in Section 5, more flexible types of models, like generalized additive or random forests, could be used instead. One would expect better estimates of the disease probabilities; however, better estimates may not lead to significantly fewer tests. The reason is because informative group testing algorithms rely on the ordering of individuals, which may be the same regardless of the modeling method. Further research should explore the potential benefits from different types of models. Finally, multiplex assays can be discriminatory or non-discriminatory. Our paper examines discriminatory tests that produce positive/negative results for each disease. Non-discriminatory, such as the Cobas TaqScreen MPX Test, provide one positive/negative outcome overall regardless of the number of diseases. Future research should examine how best to use group testing with this type of assay.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Lucy DesJardin, Jeff Benfer, and Kristopher Eveland at the SHL for their consultation and desire to improve infectious disease testing through group testing. The authors also thank Cardea Services and the state public health laboratories in Idaho and Oregon for providing access to their testing data. This research was supported by Grant R01 AI121351 from the National Institutes of Health.

References

- American Red Cross (2018). Blood testing <https://www.redcrossblood.org/biomedical-services/blood-diagnostic-testing/blood-testing.html>. Retrieved September 20, 2018.
- Bilder C, Tebbs J, and Chen P (2010). Informative retesting. *Journal of the American Statistical Association* 105, 942–955. [PubMed: 21113353]
- Black M, Bilder C, and Tebbs J (2012). Group testing in heterogeneous populations by using halving algorithms. *Journal of the Royal Statistical Society: Series C* 61, 277–290.
- Black M, Bilder C, and Tebbs J (2015). Optimal retesting configurations for hierarchical group testing. *Journal of the Royal Statistical Society: Series C* 64, 693–710.
- Dorfman R (1943). The detection of defective members of large populations. *Annals of Mathematical Statistics* 14, 436–440.
- Gildow F, Shah D, Sackett W, Butzler T, Nault B, and Fleischer S (2008). Transmission efficiency of cucumber mosaic virus by aphids associated with virus epidemics in snap bean. *Phytopathology* 98, 1233–1241. [PubMed: 18943413]
- Hou P, Tebbs J, Bilder C, and McMahan C (2017). Hierarchical group testing for multiple infections. *Biometrics* 73, 656–665. [PubMed: 27657666]
- Hwang F (1975). A generalized binomial group testing problem. *Journal of the American Statistical Association* 70, 923–926.
- Kainkaryam R and Woolf P (2009). Pooling in high-throughput drug screening. *Current Opinion in Drug Discovery & Development* 12, 339. [PubMed: 19396735]
- Lewis J, Lockary V, and Kobic S (2012). Cost savings and increased efficiency using a stratified specimen pooling strategy for *Chlamydia trachomatis* and *Neisseria gonorrhoeae*. *Sexually Transmitted Diseases* 39, 46–48. [PubMed: 22183846]

- Liu T, Hogan J, Daniels M, Coetzer M, Xu Y, Bove G, DeLong A, Ledingham L, Orido M, Diero L, and Kantor R (2017). Improved HIV-1 viral load monitoring capacity using pooled testing with marker-assisted deconvolution. *Journal of Acquired Immune Deficiency Syndromes* 75, 580–587. [PubMed: 28489730]
- McMahan C, Tebbs J, and Bilder C (2012a). Informative Dorfman screening. *Biometrics* 68, 287–296. [PubMed: 21762119]
- McMahan C, Tebbs J, and Bilder C (2012b). Two-dimensional informative array testing. *Biometrics* 68, 793–804. [PubMed: 22212007]
- Nebraska Veterinary Diagnostic Center (2018). General policies and fee schedule. <http://vbms.unl.edu/VDC/Information/VDCFeeSchedule.pdf>. Retrieved September 20, 2018.
- Quinn T, Brookmeyer R, Kline R, Shepherd M, Paranjape R, Mehendale S, Gadkari D, and Bollinger R (2000). Feasibility of pooling sera for HIV-1 viral RNA to diagnose acute primary HIV-1 infection and estimate HIV incidence. *AIDS* 14, 2751–2757. [PubMed: 11125894]
- Saá P, Proctor M, Foster G, Kryzstof D, Winton C, Linnen J, Gao K, Brodsky J, Limberger R, Dodd R, and Stramer S (2018). Investigational testing for Zika virus among US blood donors. *New England Journal of Medicine* 378, 1778–1788.
- Sherlock M, Zetola N, and Klausner J (2007). Routine detection of acute HIV infection through RNA pooling: Survey of current practice in the United States. *Sexually Transmitted Diseases* 34, 314–316. [PubMed: 17483725]
- Tebbs J, McMahan C, and Bilder C (2013). Two-stage hierarchical group testing for multiple infections with application to the Infertility Prevention Project. *Biometrics* 69, 1064–1073. [PubMed: 24117173]

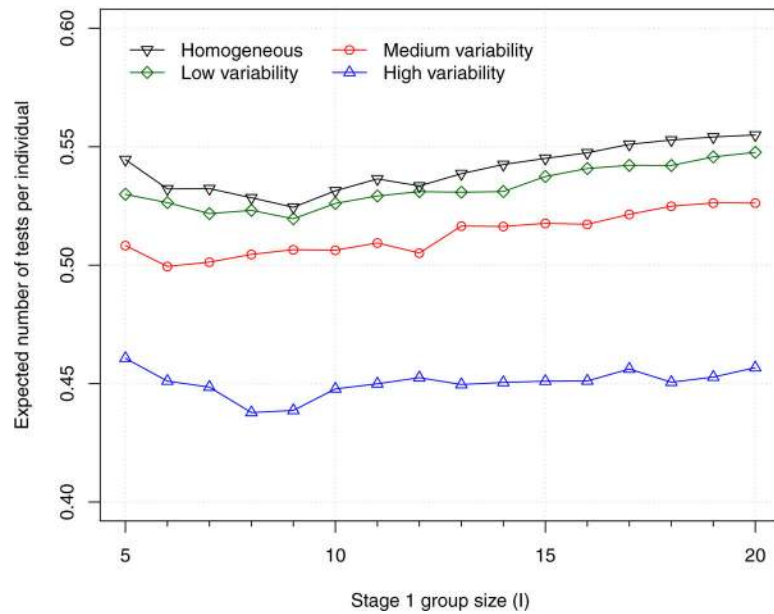
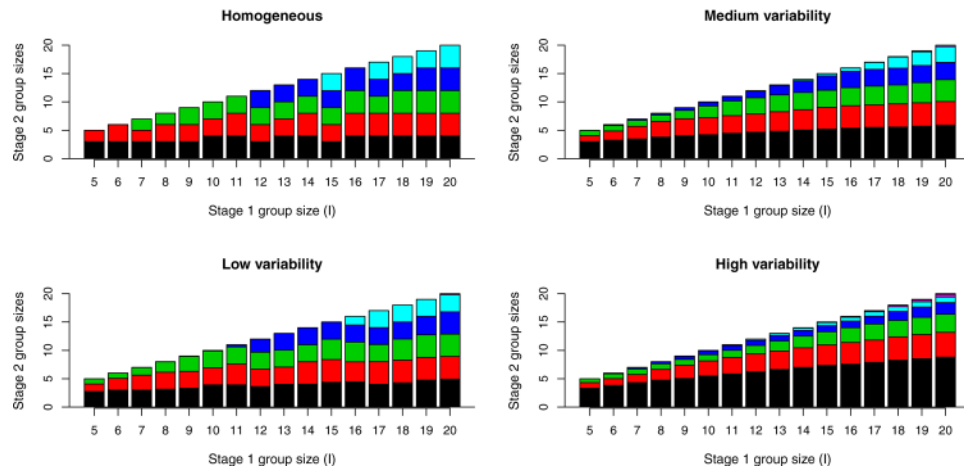


Figure 1.

Averaged expected number of tests per individual across 500 separate simulations for each I , where $\mathbf{P}_i \stackrel{iid}{\sim} \text{Dirichlet}(\cdot)$, $E(P_{i1+}) = E(P_{i+1}) = 0.05$, and $S = 3$. Results for informative group testing are given by low (Dirichlet($4\mathbf{a}$)), medium (Dirichlet(\mathbf{a})), and high (Dirichlet($\mathbf{a}/4$)) variability cases for \mathbf{P}_i with $\mathbf{a} = (18.25, 0.75, 0.75, 0.25)$. Results for non-informative group testing are given by the homogeneous case, where \mathbf{P}_i is set to $E(\mathbf{P}_i)$ for each individual and the Dirichlet parameter vector is the same as used for medium variability.

**Figure 2.**

Stacked bar chart of average stage 2 group sizes across 500 separate simulations for each I , where $\mathbf{P}_i \stackrel{iid}{\sim} \text{Dirichlet}(\cdot)$, $E(P_{i1+}) = E(P_{i+1}) = 0.05$, and $S = 3$. From the bottom to the top of each bar, the average group size is given for the first group containing the largest $p_{\lambda 0}$ values to the last group containing the smallest $p_{\lambda 0}$ values. Please see Figure 1 for a description of the homogeneous (non-informative group testing) and variability cases.

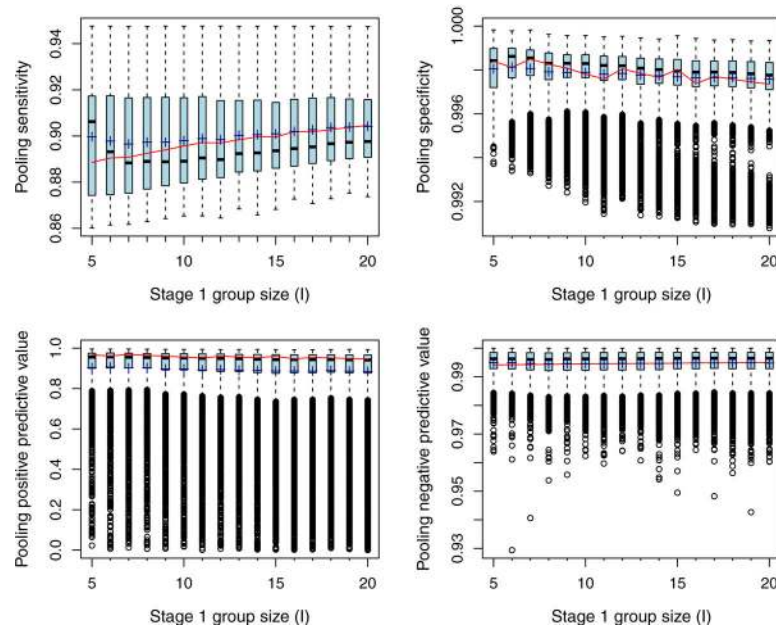


Figure 3.

Box plots of accuracy measures at each I for $P_i \stackrel{iid}{\sim} \text{Dirichlet}(\alpha)$, $E(P_{i|+}) = E(P_{i+1}) = 0.05$, $S = 3$, and the medium variability case. Note that the y-axis scale is different for each accuracy measure. The plus symbol represents the mean value for an accuracy measure in the simulations. The location of the mean value for the homogeneous case is depicted by a line across the different I .

Table 1

Mean number of tests and standard deviation (SD) for the number of tests when applying the group testing algorithms to data from 2011 for Idaho and Oregon and to data from 2014 for Iowa.

State	Gender	Number of individuals	Stages	Mean (SD) number of tests		
				Non-informative	Informative	Reduction
Idaho	Female	4168	2	2211.0 (23.6)	2105.9 (24.9)	4.8%
			3	2029.9 (27.4)	1927.1 (25.4)	5.1%
	Male	2545	2	2014.7 (12.4)	1717.8 (12.5)	14.7%
			3	2103.9 (26.8)	1831.7 (24.8)	12.9%
Iowa	Female	4351	2	2460.7 (22.3)	2459.1 (22.2)	0.1%
			3	2305.2 (29.0)	2350.0 (27.6)	-1.9%
	Male	4358	2	3419.9 (15.2)	3201.6 (16.4)	6.4%
			3	3588.3 (26.6)	3214.6 (18.3)	10.4%
Oregon	Female	8381	2	4408.5 (30.5)	4250.2 (32.4)	3.6%
			3	4000.5 (37.9)	3948.8 (37.6)	1.3%
	Male	6865	2	5478.6 (19.9)	4936.4 (19.2)	9.9%
			3	5574.8 (40.6)	5059.9 (39.0)	9.2%

Web-based Supplementary Materials for “Informative Group Testing for Multiplex Assays”

Christopher R. Bilder

Department of Statistics, University of Nebraska-Lincoln

Joshua M. Tebbs

Department of Statistics, University of South Carolina

Christopher S. McMahan

Department of Mathematical Sciences, Clemson University

September 24, 2018

Web Appendices

A	Group membership matrix examples	3
B	Optimality criteria used to find the OTC	4
C	Derivations for Section 2	5
C.1	Expected number of tests	5
C.2	Accuracy	11
C.3	Discussion	19
D	R functions	20
E	Algorithm performance	21
F	Aptima Combo 2 Assay application	30
F.1	Estimated joint probabilities of disease	30
F.2	Optimal testing configurations	33
F.3	Assay imperfection	34
F.4	Observed accuracy	37

A Group membership matrix examples

A group membership matrix used in Section 5 of the paper is

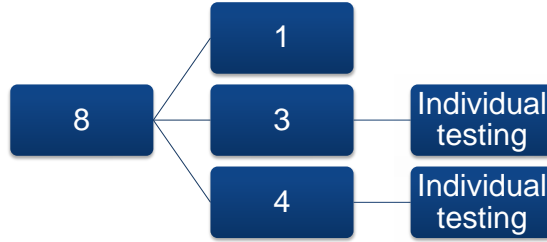
$$\mathbf{M} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 2 & 2 & 2 & 3 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & \text{NA} \end{bmatrix}. \quad (1)$$

This matrix represents a three-stage algorithm applied to an initial group of size $I = 8$ individuals. If the initial group tests positively, $m_{11} = 3$ new groups are formed for stage 2. These stage 2 groups have sizes of $I_{21} = 4$, $I_{22} = 3$, and $I_{23} = 1$. If group 1 in stage 2 tests positively, $m_{21} = 4$ individual tests are performed in stage 3 ($I_{31} = I_{32} = I_{33} = I_{34} = 1$). If group 2 in stage 2 tests positively, $m_{22} = 3$ individual tests are performed in stage 3 ($I_{35} = I_{36} = I_{37} = 1$). Finally, because group 3 in stage 2 contains only one individual, no more tests are performed after it is tested once ($m_{23} = 0$). An “NA” is included in the row 3 and column 8 cell for this individual to represent that no more testing is possible. Web Figure 1 provides a pictorial representation of the group testing algorithm.

Another group membership matrix used in Section 5 of the paper is

$$\mathbf{M} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 3 & 3 & 3 & 3 & 4 & 4 & 4 & 4 & 5 & 5 & 5 & 6 & 6 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 & 21 & 22 & 23 & \text{NA} & \text{NA} & \text{NA} & \text{NA} & \text{NA} & \text{NA} \end{bmatrix}.$$

This matrix represents a two-stage algorithm (see Section 3 of the paper) applied to $I = 29$ individuals. There are 12 initial groups. The last six groups have groups of size 1 ($I_{17} = \dots = I_{1,12} = 1$), so no further testing is performed on those individuals after the initial tests ($m_{17} = \dots = m_{1,12} = 0$). For the first six groups, group sizes are $I_{11} = 5$, $I_{12} = 4$, $I_{13} = 4$, $I_{14} = 4$, $I_{15} = 3$, and $I_{16} = 3$. If one of these groups tests positively, individual testing is performed on its members in stage 2. For example, the first group in stage 1 has $m_{11} = 5$ that leads to groups of size $I_{21} = I_{22} = I_{23} = I_{24} = I_{25} = 1$ in stage 2.



Web Figure 1: Diagram of the group testing algorithm given by Equation (1). Group sizes are provided within nodes.

B Optimality criteria used to find the OTC

We minimize $E(T|\mathbf{M})/I$ to find the OTC in our paper. This criteria is by far the most commonly used one in research. In our own experiences with laboratories, we have found this to be the case as well because expected costs can often be expressed as being proportional to the expected number of tests. In a recent article, [Malinovsky et al. \(2016\)](#) proposed minimizing $E(T|\mathbf{M})/E(C|\mathbf{M})$ instead, where C is the number of correct classifications for a group of size I . This alternative criteria allows one to simultaneously take into account the number of tests and the accuracy of the algorithm when choosing an OTC. [Malinovsky et al. \(2016\)](#) also considered other functions that included accuracy, but concluded this particular criteria was best.

Rejoinders to [Malinovsky et al. \(2016\)](#) were given by [Hudgens \(2016\)](#) and by us in [McMahan et al. \(2016\)](#). These rejoinders were in agreement that accuracy needs to be part of the decision process, but not necessarily in the manner proposed by [Malinovsky et al. \(2016\)](#). In particular, accuracy measures need to be examined for any testing configuration no matter what optimality criteria is chosen. This will help the laboratory determine if they are comfortable with the accuracy to be expected. If necessary, other testing configurations, even those from other group testing algorithms, can be chosen with a more desirable accuracy. Also, [McMahan et al. \(2016\)](#) showed by a few examples that using either $E(T|\mathbf{M})/I$ or $E(T|\mathbf{M})/E(C|\mathbf{M})$ will likely lead to the same OTC (or a very similar OTC) when using Dorfman’s algorithm for single-disease assays. Subsequent research by [Hou et al. \(2017\)](#) and [Hitt et al. \(2018\)](#) reached the same conclusions for other commonly used group testing algorithms.

Taking into account the discussion above and additional results given by [Hitt et al. \(2018\)](#), we chose to focus only on minimizing $E(T|\mathbf{M})/I$ to find the OTC for our current research.

C Derivations for Section 2

C.1 Expected number of tests

Section 2 of the paper gives the expected number of tests using group membership matrix \mathbf{M} as

$$E(T|\mathbf{M}) = 1 + \sum_{s=1}^{S-1} \sum_{j=1}^{c_s} m_{sj} P(G_{sj+}^{(1)} > 0, \dots, G_{sj+}^{(s)} > 0),$$

where $P(G_{sj+}^{(1)} > 0, \dots, G_{sj+}^{(s)} > 0)$ is

$$\begin{aligned} & \sum_{\tilde{\mathbf{g}}_{sj}^{(1)}} \dots \sum_{\tilde{\mathbf{g}}_{sj}^{(s)}} \left[\prod_{s'=1}^s \left\{ 1 - \prod_{k=1}^K (1 - S_{e:sjk}^{(s')})^{\tilde{g}_{sjk}^{(s')}} (S_{p:sjk}^{(s')})^{1-\tilde{g}_{sjk}^{(s')}} \right\} \right] P(\tilde{\mathbf{G}}_{sj}^{(1)} = \tilde{\mathbf{g}}_{sj}^{(1)}) \times \\ & \prod_{s'=1}^{s-1} P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = \tilde{\mathbf{g}}_{sj}^{(s'+1)} | \tilde{\mathbf{G}}_{sj}^{(s')} = \tilde{\mathbf{g}}_{sj}^{(s')}). \end{aligned} \quad (2)$$

The purpose of this subsection is to show how the expression for Equation (2) is found.

We can re-write $P(G_{sj+}^{(1)} > 0, \dots, G_{sj+}^{(s)} > 0)$ by using the theorem of total probability to incorporate each true status:

$$\begin{aligned} & P(G_{sj+}^{(1)} > 0, \dots, G_{sj+}^{(s)} > 0) \\ &= \sum_{\tilde{\mathbf{g}}_{sj}^{(1)}} \dots \sum_{\tilde{\mathbf{g}}_{sj}^{(s)}} P(G_{sj+}^{(1)} > 0, \dots, G_{sj+}^{(s)} > 0, \tilde{\mathbf{G}}_{sj}^{(1)} = \tilde{\mathbf{g}}_{sj}^{(1)}, \dots, \tilde{\mathbf{G}}_{sj}^{(s)} = \tilde{\mathbf{g}}_{sj}^{(s)}) \\ &= \sum_{\tilde{\mathbf{g}}_{sj}^{(1)}} \dots \sum_{\tilde{\mathbf{g}}_{sj}^{(s)}} P(G_{sj+}^{(1)} > 0, \dots, G_{sj+}^{(s)} > 0 | \tilde{\mathbf{G}}_{sj}^{(1)} = \tilde{\mathbf{g}}_{sj}^{(1)}, \dots, \tilde{\mathbf{G}}_{sj}^{(s)} = \tilde{\mathbf{g}}_{sj}^{(s)}) P(\tilde{\mathbf{G}}_{sj}^{(1)} = \tilde{\mathbf{g}}_{sj}^{(1)}, \dots, \tilde{\mathbf{G}}_{sj}^{(s)} = \tilde{\mathbf{g}}_{sj}^{(s)}), \end{aligned} \quad (3)$$

The summations over $\tilde{\mathbf{g}}_{sj}$ represent summing over all 2^K possible 0-1 combinations for the corresponding group statuses.

The large conditional probability in Equation (3) can be rewritten as the product of assay sensitivities and specificities:

$$\begin{aligned} & P(G_{sj+}^{(1)} > 0, \dots, G_{sj+}^{(s)} > 0 | \tilde{\mathbf{G}}_{sj}^{(1)} = \tilde{\mathbf{g}}_{sj}^{(1)}, \dots, \tilde{\mathbf{G}}_{sj}^{(s)} = \tilde{\mathbf{g}}_{sj}^{(s)}) \\ &= \prod_{s'=1}^s \left[1 - P(G_{sj+}^{(s')} = 0 | \tilde{\mathbf{G}}_{sj}^{(s')} = \tilde{\mathbf{g}}_{sj}^{(s')}) \right] \\ &= \prod_{s'=1}^s \left[1 - \prod_{k=1}^K (1 - S_{e:sjk}^{(s')})^{\tilde{g}_{sjk}^{(s')}} (S_{p:sjk}^{(s')})^{1-\tilde{g}_{sjk}^{(s')}} \right], \end{aligned}$$

where we make the assumption of conditional independence among the group responses once the true statuses are known. This assumption is used throughout group testing research (e.g., [Kim et al., 2007](#)), and empirical evidence of its validity was given by [Litvak et al. \(1994\)](#). In related research, [McMahan et al. \(2017\)](#) recently showed that small violations of this assumption will have little practical effect when group testing observations are used for parameter estimation.

Because each stage result is dependent only on the previous stage, we can rewrite the joint probability in Equation (3) as a product of one-stage dependent probabilities:

$$P(\tilde{\mathbf{G}}_{sj}^{(1)} = \tilde{\mathbf{g}}_{sj}^{(1)}, \dots, \tilde{\mathbf{G}}_{sj}^{(s)} = \tilde{\mathbf{g}}_{sj}^{(s)}) = P(\tilde{\mathbf{G}}_{11} = \tilde{\mathbf{g}}_{11}) \prod_{s'=1}^{s-1} P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = \tilde{\mathbf{g}}_{sj}^{(s'+1)} | \tilde{\mathbf{G}}_{sj}^{(s')} = \tilde{\mathbf{g}}_{sj}^{(s')}), \quad (4)$$

where we use the convention throughout our paper that an empty product has a value of 1. Because there are K diseases each with two possible outcomes, there are 2^{2K} conditional probabilities over each of the $s-1$

stages in Equation (4). For this reason, we focus on examining the $K = 2$ case, which is the same as for the Aptima Combo 2 Assay, for the remainder of this subsection. We discuss $K \geq 3$ in Section C.3.

By using the joint probabilities of disease, $P(\tilde{Y}_{i1} = \tilde{y}_1, \tilde{Y}_{i2} = \tilde{y}_2) = p_{i\tilde{y}_1\tilde{y}_2}$, we obtain $P(\tilde{\mathbf{G}}_{11} = \tilde{\mathbf{g}}_{11})$ to be equal to be the following:

- $\tilde{\mathbf{G}}_{11} = (0, 0)$

$$P(\tilde{\mathbf{G}}_{11} = (0, 0)) = \prod_{i=1}^I p_{i00}$$

- $\tilde{\mathbf{G}}_{11} = (0, 1)$

$$\begin{aligned} P(\tilde{\mathbf{G}}_{11} = (0, 1)) &= P(\tilde{G}_{111} = 0) - P(\tilde{\mathbf{G}}_{11} = (0, 0)) \\ &= \left(\prod_{i=1}^I p_{i0+} \right) - \left(\prod_{i=1}^I p_{i00} \right) \end{aligned}$$

- $\tilde{\mathbf{G}}_{11} = (1, 0)$

$$\begin{aligned} P(\tilde{\mathbf{G}}_{11} = (1, 0)) &= P(\tilde{G}_{112} = 0) - P(\tilde{\mathbf{G}}_{11} = (0, 0)) \\ &= \left(\prod_{i=1}^I p_{i+0} \right) - \left(\prod_{i=1}^I p_{i00} \right) \end{aligned}$$

- $\tilde{\mathbf{G}}_{11} = (1, 1)$

$$\begin{aligned} P(\tilde{\mathbf{G}}_{11} = (1, 1)) &= P(\tilde{G}_{111} = 1) - P(\tilde{\mathbf{G}}_{11} = (1, 0)) \\ &= 1 - P(\tilde{G}_{111} = 0) - P(\tilde{G}_{112} = 0) - P(\tilde{\mathbf{G}}_{11} = (0, 0)) \\ &= 1 - \left(\prod_{i=1}^I p_{i0+} \right) - \left(\prod_{i=1}^I p_{i+0} \right) + \left(\prod_{i=1}^I p_{i00} \right) \end{aligned}$$

where we use “+” in a subscript to denote summing over a particular index (e.g., $p_{i0+} = p_{i00} + p_{i01}$).

For the conditional probability in Equation (4), there are 16 different values that need to be calculated for each of the $s - 1$ stages that the product is taken over. Many of these quantities are simple to determine. For example, $P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = \tilde{\mathbf{g}}_{sj}^{(s'+1)} | \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0)) = 0$ for any stage $s' + 1$ status that contains at least one positive—once all true negatives occur for a group, subsequent groups cannot contain any positives. Furthermore, this leads to $P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0) | \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0)) = 1$. Other probabilities are more difficult to determine. Define B_{sj} as the set of all individuals that belong to group j at stage s ; i.e., those individuals who contribute to the group response G_{sjk} . Similarly, define ancestor group membership as $B_{sj}^{(t)}$ for those individuals who contribute to the group response $G_{sjk}^{(t)}$, and define $\bar{B}_{sj}^{(t)}$ as those individuals who contribute to the group response $G_{sjk}^{(t-1)}$ but not to $G_{sjk}^{(t)}$. Using this additional notation, we can derive expressions for the remaining conditional probabilities. Below are all 16 conditional probabilities:

- $\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0)$

Because $\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0) \Rightarrow \tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0)$,

$$P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0) | \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0)) = 1.$$

- $\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 1), \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0)$

A positive status for disease 2 cannot occur at stage $s' + 1$ if the disease 2 status is negative at stage s' . Therefore,

$$P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 1) | \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0)) = 0.$$

- $\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0)$

A positive status for disease 1 cannot occur at stage $s' + 1$ if the disease 1 status is negative at stage s' . Therefore,

$$P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0) | \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0)) = 0.$$

- $\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 1), \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0)$

A positive status for diseases 1 and 2 cannot occur at stage $s' + 1$ if the statuses are negative at stage s' . Therefore,

$$P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 1) | \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0)) = 0.$$

- $\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 1)$

$$\begin{aligned} P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0) | \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 1)) &= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 1))}{P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 1))} \\ &= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))}{P(\tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\ &= \frac{P(\tilde{G}_{sj2}^{(s'+1)} = 0, \tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))}{P(\tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\ &= \frac{\left(\prod_{i \in B_{sj}^{(s'+1)}} p_{i00} \right) \left(\prod_{i \in \bar{B}_{sj}^{(s'+1)}} p_{i0+} \right) - \left(\prod_{i \in B_{sj}^{(s')}} p_{i00} \right)}{\left(\prod_{i \in B_{sj}^{(s')}} p_{i0+} \right) - \left(\prod_{i \in B_{sj}^{(s')}} p_{i00} \right)}, \end{aligned}$$

The key for finding this and many of the other conditional probability expressions is to re-express probabilities like $P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 1))$ in terms of marginal probabilities and probabilities for all negative statuses. Thus, $P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 1)) = P(\tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0)) = (\prod_{i \in B_{sj}^{(s')}} p_{i0+}) - (\prod_{i \in B_{sj}^{(s')}} p_{i00})$.

The denominator for the above fraction represents the probability of having a negative true status for the first disease while having a positive true status for the second. While this probability is not guaranteed to be non-zero, the occurrence of a null probability is unrealistic. If the probability was somehow equal to 0, it would call into question the reasoning for why a multiplex assay would be used in this scenario.

- $\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 1), \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 1)$

$$\begin{aligned}
& P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 1) | \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 1)) \\
&= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 1), \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 1))}{P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 1))} \\
&= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 1), \tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 1), \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))}{P(\tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\
&= \frac{P(\tilde{G}_{sj1}^{(s'+1)} = 0, \tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{G}_{sj1}^{(s')} = 0) - 0}{P(\tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\
&= \frac{P(\tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{G}_{sj2}^{(s'+1)} = 0, \tilde{G}_{sj1}^{(s')} = 0)}{P(\tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\
&= \frac{\left(\prod_{i \in B_{sj}^{(s')}} p_{i0+} \right) - \left(\prod_{i \in B_{sj}^{(s'+1)}} p_{i00} \right) \left(\prod_{i \in \bar{B}_{sj}^{(s'+1)}} p_{i0+} \right)}{\left(\prod_{i \in B_{sj}^{(s')}} p_{i0+} \right) - \left(\prod_{i \in B_{sj}^{(s')}} p_{i00} \right)}
\end{aligned}$$

- $\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 1)$

A positive status for disease 1 cannot occur at stage $s' + 1$ if the disease 1 status is negative at stage s' . Therefore,

$$P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0) | \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 1)) = 0.$$

- $\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 1), \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 1)$

A positive status for disease 1 cannot occur at stage $s' + 1$ if the disease 1 status is negative at stage s' . Therefore,

$$P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 1) | \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 1)) = 0.$$

- $\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0)$

$$\begin{aligned}
P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0) | \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0)) &= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0))}{P(\tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0))} \\
&= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{G}_{sj2}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))}{P(\tilde{G}_{sj2}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\
&= \frac{P(\tilde{G}_{sj1}^{(s'+1)} = 0, \tilde{G}_{sj2}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))}{P(\tilde{G}_{sj2}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\
&= \frac{\left(\prod_{i \in B_{sj}^{(s'+1)}} p_{i00} \right) \left(\prod_{i \in \bar{B}_{sj}^{(s'+1)}} p_{i+0} \right) - \left(\prod_{i \in B_{sj}^{(s')}} p_{i00} \right)}{\left(\prod_{i \in B_{sj}^{(s')}} p_{i+0} \right) - \left(\prod_{i \in B_{sj}^{(s')}} p_{i00} \right)}
\end{aligned}$$

- $\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 1), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0)$

A positive status for disease 2 cannot occur at stage $s' + 1$ if the disease 2 status is negative at stage s' . Therefore,

$$P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 1) | \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0)) = 0.$$

- $\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0)$

$$\begin{aligned} P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0) | \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0)) &= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0))}{P(\tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0))} \\ &= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0), \tilde{G}_{sj2}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))}{P(\tilde{G}_{sj2}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\ &= \frac{P(\tilde{G}_{sj2}^{(s'+1)} = 0, \tilde{G}_{sj2}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{G}_{sj2}^{(s')} = 0) - 0}{P(\tilde{G}_{sj2}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\ &= \frac{P(\tilde{G}_{sj2}^{(s')} = 0) - P(\tilde{G}_{sj1}^{(s'+1)} = 0, \tilde{G}_{sj2}^{(s')} = 0)}{P(\tilde{G}_{sj2}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\ &= \frac{\left(\prod_{i \in B_{sj}^{(s')}} p_{i+0} \right) - \left(\prod_{i \in B_{sj}^{(s'+1)}} p_{i00} \right) \left(\prod_{i \in \bar{B}_{sj}^{(s'+1)}} p_{i+0} \right)}{\left(\prod_{i \in B_{sj}^{(s')}} p_{i+0} \right) - \left(\prod_{i \in B_{sj}^{(s')}} p_{i00} \right)} \end{aligned}$$

- $\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 1), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0)$

A positive status for disease 2 cannot occur at stage $s' + 1$ if the disease 2 status is negative at stage s' . Therefore,

$$P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 1) | \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0)) = 0.$$

- $\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 1)$

$$\begin{aligned} P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0) | \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 1)) &= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 1))}{P(\tilde{\mathbf{G}}_{sj}^{(s')} = (1, 1))} \\ &= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{G}_{sj1}^{(s')} = 1) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0))}{P(\tilde{G}_{sj1}^{(s')} = 1) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0))} \\ &= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0)) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{G}_{sj2}^{(s')} = 0) + P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))}{1 - P(\tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{G}_{sj2}^{(s')} = 0) + P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\ &= \frac{\left(\prod_{i \in B_{sj}^{(s'+1)}} p_{i00} \right) - \left(\prod_{i \in B_{sj}^{(s'+1)}} p_{i00} \right) \left(\prod_{i \in \bar{B}_{sj}^{(s'+1)}} p_{i0+} \right) - \left(\prod_{i \in B_{sj}^{(s'+1)}} p_{i00} \right) \left(\prod_{i \in \bar{B}_{sj}^{(s'+1)}} p_{i+0} \right) + \left(\prod_{i \in B_{sj}^{(s')}} p_{i00} \right)}{1 - \left(\prod_{i \in B_{sj}^{(s')}} p_{i0+} \right) - \left(\prod_{i \in B_{sj}^{(s')}} p_{i+0} \right) + \left(\prod_{i \in B_{sj}^{(s')}} p_{i00} \right)} \end{aligned}$$

- $\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 1), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 1)$

$$\begin{aligned}
& P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 1) | \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 1)) \\
&= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 1), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 1))}{P(\tilde{\mathbf{G}}_{sj}^{(s')} = (1, 1))} \\
&= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 1), \tilde{G}_{sj1}^{(s')} = 1) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 1), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0))}{P(\tilde{G}_{sj1}^{(s')} = 1) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0))} \\
&= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 1)) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 1), \tilde{G}_{sj1}^{(s')} = 0) - 0}{1 - P(\tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{G}_{sj2}^{(s')} = 0) + P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\
&= \frac{P(\tilde{G}_{sj1}^{(s'+1)} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0)) - P(\tilde{G}_{sj1}^{(s'+1)} = 0, \tilde{G}_{sj1}^{(s')} = 0) + P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{G}_{sj1}^{(s')} = 0)}{1 - P(\tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{G}_{sj2}^{(s')} = 0) + P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\
&= \frac{\left(\prod_{i \in B_{sj}^{(s'+1)}} p_{i0+} \right) - \left(\prod_{i \in B_{sj}^{(s'+1)}} p_{i00} \right) - \left(\prod_{i \in B_{sj}^{(s')}} p_{i0+} \right) + \left(\prod_{i \in B_{sj}^{(s'+1)}} p_{i00} \right) \left(\prod_{i \in \bar{B}_{sj}^{(s'+1)}} p_{i0+} \right)}{1 - \left(\prod_{i \in B_{sj}^{(s')}} p_{i0+} \right) - \left(\prod_{i \in B_{sj}^{(s')}} p_{i+0} \right) + \left(\prod_{i \in B_{sj}^{(s')}} p_{i00} \right)}
\end{aligned}$$

- $\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 1)$

$$\begin{aligned}
& P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0) | \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 1)) \\
&= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 1))}{P(\tilde{\mathbf{G}}_{sj}^{(s')} = (1, 1))} \\
&= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0), \tilde{G}_{sj2}^{(s')} = 1) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0), \tilde{\mathbf{G}}_{sj}^{(s')} = (0, 1))}{P(\tilde{G}_{sj1}^{(s')} = 1) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0))} \\
&= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0)) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0), \tilde{G}_{sj2}^{(s')} = 0) - 0}{1 - P(\tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{G}_{sj2}^{(s')} = 0) + P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\
&= \frac{P(\tilde{G}_{sj2}^{(s'+1)} = 0) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0)) - P(\tilde{G}_{sj2}^{(s'+1)} = 0, \tilde{G}_{sj2}^{(s')} = 0) + P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0), \tilde{G}_{sj2}^{(s')} = 0)}{1 - P(\tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{G}_{sj2}^{(s')} = 0) + P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\
&= \frac{\left(\prod_{i \in B_{sj}^{(s'+1)}} p_{i+0} \right) - \left(\prod_{i \in B_{sj}^{(s'+1)}} p_{i00} \right) - \left(\prod_{i \in B_{sj}^{(s')}} p_{i+0} \right) + \left(\prod_{i \in B_{sj}^{(s'+1)}} p_{i00} \right) \left(\prod_{i \in \bar{B}_{sj}^{(s'+1)}} p_{i+0} \right)}{1 - \left(\prod_{i \in B_{sj}^{(s')}} p_{i0+} \right) - \left(\prod_{i \in B_{sj}^{(s')}} p_{i+0} \right) + \left(\prod_{i \in B_{sj}^{(s')}} p_{i00} \right)}
\end{aligned}$$

- $\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 1), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 1)$

$$\begin{aligned}
& P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 1) | \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 1)) \\
&= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 1), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 1))}{P(\tilde{\mathbf{G}}_{sj}^{(s')} = (1, 1))} \\
&= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 1), \tilde{G}_{sj1}^{(s')} = 1) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 1), \tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0))}{P(\tilde{G}_{sj1}^{(s')} = 1) - P(\tilde{\mathbf{G}}_{sj}^{(s')} = (1, 0))} \\
&= \frac{P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 1)) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 1), \tilde{G}_{sj1}^{(s')} = 0) - 0}{1 - P(\tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{G}_{sj2}^{(s')} = 0) + P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\
&= \frac{P(\tilde{G}_{sj1}^{(s'+1)} = 1) - P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (1, 0)) - 0}{1 - P(\tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{G}_{sj2}^{(s')} = 0) + P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\
&= \frac{1 - P(\tilde{G}_{sj1}^{(s'+1)} = 0) - P(\tilde{G}_{sj2}^{(s'+1)} = 0) + P(\tilde{\mathbf{G}}_{sj}^{(s'+1)} = (0, 0))}{1 - P(\tilde{G}_{sj1}^{(s')} = 0) - P(\tilde{G}_{sj2}^{(s')} = 0) + P(\tilde{\mathbf{G}}_{sj}^{(s')} = (0, 0))} \\
&= \frac{1 - \left(\prod_{i \in B_{sj}^{(s'+1)}} p_{i0+} \right) - \left(\prod_{i \in B_{sj}^{(s'+1)}} p_{i+0} \right) + \left(\prod_{i \in B_{sj}^{(s'+1)}} p_{i00} \right)}{1 - \left(\prod_{i \in B_{sj}^{(s')}} p_{i0+} \right) - \left(\prod_{i \in B_{sj}^{(s')}} p_{i+0} \right) + \left(\prod_{i \in B_{sj}^{(s')}} p_{i00} \right)}
\end{aligned}$$

C.2 Accuracy

The pooling sensitivity for individual i with respect to disease k is

$$\begin{aligned}
PS_{e:ik} &= P(G_{Lj+}^{(1)} > 0, \dots, G_{Lj+}^{(L-1)} > 0, G_{Ljk} = 1 | \tilde{Y}_{ik} = 1) \\
&= \sum_{\tilde{\mathbf{g}}_{Lj\bar{k}}^{(1)}} \dots \sum_{\tilde{\mathbf{g}}_{Lj\bar{k}}^{(L-1)}} P(G_{Lj+}^{(1)} > 0, \dots, G_{Lj+}^{(L-1)} > 0, G_{Ljk} = 1, \tilde{\mathbf{G}}_{Lj\bar{k}}^{(1)} = \tilde{\mathbf{g}}_{Lj\bar{k}}^{(1)}, \dots, \tilde{\mathbf{G}}_{Lj\bar{k}}^{(L-1)} = \tilde{\mathbf{g}}_{Lj\bar{k}}^{(L-1)} | \tilde{Y}_{ik} = 1) \\
&= \sum_{\tilde{\mathbf{g}}_{Lj\bar{k}}^{(1)}} \dots \sum_{\tilde{\mathbf{g}}_{Lj\bar{k}}^{(L-1)}} P(G_{Lj+}^{(1)} > 0, \dots, G_{Lj+}^{(L-1)} > 0, G_{Ljk} = 1 | \tilde{\mathbf{G}}_{Lj\bar{k}}^{(1)} = \tilde{\mathbf{g}}_{Lj\bar{k}}^{(1)}, \dots, \tilde{\mathbf{G}}_{Lj\bar{k}}^{(L-1)} = \tilde{\mathbf{g}}_{Lj\bar{k}}^{(L-1)}, \tilde{Y}_{ik} = 1) \times \\
&\quad P(\tilde{\mathbf{G}}_{Lj\bar{k}}^{(1)} = \tilde{\mathbf{g}}_{Lj\bar{k}}^{(1)}, \dots, \tilde{\mathbf{G}}_{Lj\bar{k}}^{(L-1)} = \tilde{\mathbf{g}}_{Lj\bar{k}}^{(L-1)} | \tilde{Y}_{ik} = 1) \\
&= \sum_{\tilde{\mathbf{g}}_{Lj\bar{k}}^{(1)}} \dots \sum_{\tilde{\mathbf{g}}_{Lj\bar{k}}^{(L-1)}} P(G_{Ljk} = 1 | \tilde{Y}_{ik} = 1) \left\{ \prod_{s'=1}^{L-1} P(G_{Lj+}^{(s')} > 0 | \tilde{\mathbf{G}}_{Lj\bar{k}}^{(s')} = \tilde{\mathbf{g}}_{Lj\bar{k}}^{(s')}, \tilde{Y}_{ik} = 1) \right\} \times \\
&\quad P(\tilde{\mathbf{G}}_{Lj\bar{k}}^{(1)} = \tilde{\mathbf{g}}_{Lj\bar{k}}^{(1)} | \tilde{Y}_{ik} = 1) \prod_{s'=1}^{L-2} P(\tilde{\mathbf{G}}_{Lj\bar{k}}^{(s'+1)} = \tilde{\mathbf{g}}_{Lj\bar{k}}^{(s'+1)} | \tilde{\mathbf{G}}_{Lj\bar{k}}^{(s')} = \tilde{\mathbf{g}}_{Lj\bar{k}}^{(s')}, \tilde{Y}_{ik} = 1) \\
&= \sum_{\tilde{\mathbf{g}}_{Lj\bar{k}}^{(1)}} \dots \sum_{\tilde{\mathbf{g}}_{Lj\bar{k}}^{(L-1)}} S_{e:Ljk}^{(L)} \left[\prod_{s'=1}^{L-1} \left\{ 1 - P(G_{Lj+}^{(s')} = 0 | \tilde{\mathbf{G}}_{Lj\bar{k}}^{(s')} = \tilde{\mathbf{g}}_{Lj\bar{k}}^{(s')}, \tilde{Y}_{ik} = 1) \right\} \right] \times \\
&\quad P(\tilde{\mathbf{G}}_{Lj\bar{k}}^{(1)} = \tilde{\mathbf{g}}_{Lj\bar{k}}^{(1)} | \tilde{Y}_{ik} = 1) \prod_{s'=1}^{L-2} P(\tilde{\mathbf{G}}_{Lj\bar{k}}^{(s'+1)} = \tilde{\mathbf{g}}_{Lj\bar{k}}^{(s'+1)} | \tilde{\mathbf{G}}_{Lj\bar{k}}^{(s')} = \tilde{\mathbf{g}}_{Lj\bar{k}}^{(s')}, \tilde{Y}_{ik} = 1)
\end{aligned}$$

$$= \sum_{\tilde{\mathbf{g}}_{Lj\bar{k}}^{(1)}} \cdots \sum_{\tilde{\mathbf{g}}_{Lj\bar{k}}^{(L-1)}} S_{e:Ljk}^{(L)} \left[\prod_{s'=1}^{L-1} \left\{ 1 - (1 - S_{e:Ljk}^{(s')}) \prod_{q=1, q \neq k}^K (1 - S_{e:Ljq}^{(s')})^{\tilde{g}_{Ljq}^{(s')}} (S_{p:Ljq}^{(s')})^{1 - \tilde{g}_{Ljq}^{(s')}} \right\} \right] \times$$

$$P(\tilde{\mathbf{G}}_{Lj\bar{k}}^{(1)} = \tilde{\mathbf{g}}_{Lj\bar{k}}^{(1)} | \tilde{Y}_{ik} = 1) \prod_{s'=1}^{L-2} P(\tilde{\mathbf{G}}_{Lj}^{(s'+1)} = \tilde{\mathbf{g}}_{Lj}^{(s'+1)} | \tilde{\mathbf{G}}_{Lj}^{(s')} = \tilde{\mathbf{g}}_{Lj\bar{k}}^{(s')}, \tilde{Y}_{ik} = 1).$$

Focusing on $K = 2$ diseases again, probabilities for $P(\tilde{G}_{Lj\bar{k}}^{(1)} = \tilde{g}_{Lj\bar{k}}^{(1)} | \tilde{Y}_{ik} = 1)$ are

- $\tilde{G}_{Lj\bar{k}}^{(1)} = 0$

For $k = 1, \bar{k} = 2$:

$$P(\tilde{G}_{Lj2}^{(1)} = 0 | \tilde{Y}_{i1} = 1) = \frac{p_{i10} \prod_{b=1, b \neq i}^I p_{b+0}}{p_{i1+}}$$

For $k = 2, \bar{k} = 1$:

$$P(\tilde{G}_{Lj1}^{(1)} = 0 | \tilde{Y}_{i2} = 1) = \frac{p_{i01} \prod_{b=1, b \neq i}^I p_{b0+}}{p_{i+1}}$$

- $\tilde{G}_{Lj\bar{k}}^{(1)} = 1$

For $k = 1, \bar{k} = 2$:

$$P(\tilde{G}_{Lj2}^{(1)} = 1 | \tilde{Y}_{i1} = 1) = \frac{p_{i1+} - p_{i10} \prod_{b=1, b \neq i}^I p_{b+0}}{p_{i1+}}$$

For $k = 2, \bar{k} = 1$:

$$P(\tilde{G}_{Lj1}^{(1)} = 1 | \tilde{Y}_{i2} = 1) = \frac{p_{i+1} - p_{i01} \prod_{b=1, b \neq i}^I p_{b0+}}{p_{i+1}}$$

Also for $K = 2$ diseases, probabilities for $P(\tilde{G}_{Lj\bar{k}}^{(s'+1)} = \tilde{g}_{Lj\bar{k}}^{(s'+1)} | \tilde{G}_{Lj\bar{k}}^{(s')} = \tilde{g}_{Lj\bar{k}}^{(s')}, \tilde{Y}_{ik} = 1)$ are

- $\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 0, \tilde{G}_{Lj\bar{k}}^{(s')} = 0$

Because $\tilde{G}_{Lj}^{(s')} = 0 \Rightarrow \tilde{G}_{Lj}^{(s'+1)} = 0$,

$$P(\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 0 | \tilde{G}_{Lj\bar{k}}^{(s')} = 0, \tilde{Y}_{ik} = 1) = 1.$$

- $\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 0, \tilde{G}_{Lj\bar{k}}^{(s')} = 1$

$$P(\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 0 | \tilde{G}_{Lj\bar{k}}^{(s')} = 1, \tilde{Y}_{ik} = 1) = \frac{P(\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 0, \tilde{G}_{Lj\bar{k}}^{(s')} = 1, \tilde{Y}_{ik} = 1)}{P(\tilde{G}_{Lj\bar{k}}^{(s')} = 1, \tilde{Y}_{ik} = 1)}$$

$$= \frac{P(\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 0, \tilde{Y}_{ik} = 1) - P(\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 0, \tilde{G}_{Lj\bar{k}}^{(s')} = 0, \tilde{Y}_{ik} = 1)}{P(\tilde{Y}_{ik} = 1) - P(\tilde{G}_{Lj\bar{k}}^{(s')} = 0, \tilde{Y}_{ik} = 1)}$$

$$= \frac{P(\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 0, \tilde{Y}_{ik} = 1) - P(\tilde{G}_{Lj\bar{k}}^{(s')} = 0, \tilde{Y}_{ik} = 1)}{P(\tilde{Y}_{ik} = 1) - P(\tilde{G}_{Lj\bar{k}}^{(s')} = 0, \tilde{Y}_{ik} = 1)}$$

where for $k = 1, \bar{k} = 2$,

$$P(\tilde{G}_{Lj2}^{(s'+1)} = 0 | \tilde{G}_{Lj2}^{(s')} = 1, \tilde{Y}_{i1} = 1) = \frac{p_{i10} \left(\prod_{b \in B_{Lj}^{(s'+1)}, b \neq i} p_{b+0} \right) - p_{i10} \left(\prod_{b \in B_{Lj}^{(s')}, b \neq i} p_{b+0} \right)}{p_{i1+} - p_{i10} \left(\prod_{b \in B_{Lj}^{(s')}, b \neq i} p_{b+0} \right)}$$

where for $k = 2, \bar{k} = 1$,

$$P(\tilde{G}_{Lj1}^{(s'+1)} = 0 | \tilde{G}_{Lj1}^{(s')} = 1, \tilde{Y}_{i2} = 1) = \frac{p_{i01} \left(\prod_{b \in B_{Lj}^{(s'+1)}, b \neq i} p_{b0+} \right) - p_{i01} \left(\prod_{b \in B_{Lj}^{(s')}, b \neq i} p_{b0+} \right)}{p_{i+1} - p_{i01} \left(\prod_{b \in B_{Lj}^{(s')}, b \neq i} p_{b0+} \right)}$$

- $\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 1, \tilde{G}_{Lj\bar{k}}^{(s')} = 0$

A positive status for disease \bar{k} cannot occur at stage $s' + 1$ if the disease \bar{k} status is negative at stage s' . Therefore,

$$P(\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 1 | \tilde{G}_{Lj\bar{k}}^{(s')} = 0, \tilde{Y}_{ik} = 1) = 0.$$

- $\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 1, \tilde{G}_{Lj\bar{k}}^{(s')} = 1$

$$\begin{aligned} & P(\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 1 | \tilde{G}_{Lj\bar{k}}^{(s')} = 1, \tilde{Y}_{ik} = 1) \\ &= \frac{P(\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 1, \tilde{G}_{Lj\bar{k}}^{(s')} = 1, \tilde{Y}_{ik} = 1)}{P(\tilde{G}_{Lj\bar{k}}^{(s')} = 1, \tilde{Y}_{ik} = 1)} \\ &= \frac{P(\tilde{G}_{Lj\bar{k}}^{(s')} = 1, \tilde{Y}_{ik} = 1) - P(\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 0, \tilde{G}_{Lj\bar{k}}^{(s')} = 1, \tilde{Y}_{ik} = 1)}{P(\tilde{Y}_{ik} = 1) - P(\tilde{G}_{Lj\bar{k}}^{(s')} = 0, \tilde{Y}_{ik} = 1)} \\ &= \frac{P(\tilde{Y}_{ik} = 1) - P(\tilde{G}_{Lj\bar{k}}^{(s')} = 0, \tilde{Y}_{ik} = 1) - P(\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 0, \tilde{Y}_{ik} = 1) + P(\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 0, \tilde{G}_{Lj\bar{k}}^{(s')} = 0, \tilde{Y}_{ik} = 1)}{P(\tilde{Y}_{ik} = 1) - P(\tilde{G}_{Lj\bar{k}}^{(s')} = 0, \tilde{Y}_{ik} = 1)} \\ &= \frac{P(\tilde{Y}_{ik} = 1) - P(\tilde{G}_{Lj\bar{k}}^{(s')} = 0, \tilde{Y}_{ik} = 1) - P(\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 0, \tilde{Y}_{ik} = 1) + P(\tilde{G}_{Lj\bar{k}}^{(s')} = 0, \tilde{Y}_{ik} = 1)}{P(\tilde{Y}_{ik} = 1) - P(\tilde{G}_{Lj\bar{k}}^{(s')} = 0, \tilde{Y}_{ik} = 1)} \\ &= \frac{P(\tilde{Y}_{ik} = 1) - P(\tilde{G}_{Lj\bar{k}}^{(s'+1)} = 0, \tilde{Y}_{ik} = 1)}{P(\tilde{Y}_{ik} = 1) - P(\tilde{G}_{Lj\bar{k}}^{(s')} = 0, \tilde{Y}_{ik} = 1)} \end{aligned}$$

where for $k = 1, \bar{k} = 2$,

$$P(\tilde{G}_{Lj2}^{(s'+1)} = 1 | \tilde{G}_{Lj2}^{(s')} = 1, \tilde{Y}_{i1} = 1) = \frac{p_{i1+} - p_{i10} \prod_{b \in B_{Lj}^{(s'+1)}, b \neq i} p_{b+0}}{p_{i1+} - p_{i10} \prod_{b \in B_{Lj}^{(s')}, b \neq i} p_{b+0}}$$

where for $k = 2, \bar{k} = 1$,

$$P(\tilde{G}_{Lj1}^{(s'+1)} = 1 | \tilde{G}_{Lj1}^{(s')} = 1, \tilde{Y}_{i2} = 1) = \frac{p_{i+1} - p_{i01} \prod_{b \in B_{Lj}^{(s'+1)}, b \neq i} p_{b0+}}{p_{i+1} - p_{i01} \prod_{b \in B_{Lj}^{(s')}, b \neq i} p_{b0+}}$$

We discuss $K \geq 3$ in Section C.3.

To find the pooling specificity, we first write it in terms of the pooling sensitivity, the probability of a true positive individual, and the probability of a positive individual diagnosis from group testing:

$$\begin{aligned} PS_{p:ik} &= P(Y_{ik} = 0 | \tilde{Y}_{ik} = 0) \\ &= 1 - P(Y_{ik} = 1 | \tilde{Y}_{ik} = 0) \\ &= 1 - \frac{P(Y_{ik} = 1, \tilde{Y}_{ik} = 0)}{P(\tilde{Y}_{ik} = 0)} \\ &= 1 - \frac{P(Y_{ik} = 1) - P(Y_{ik} = 1, \tilde{Y}_{ik} = 1)}{1 - P(\tilde{Y}_{ik} = 1)} \\ &= 1 - \frac{P(Y_{ik} = 1) - P(Y_{ik} = 1 | \tilde{Y}_{ik} = 1)P(\tilde{Y}_{ik} = 1)}{1 - P(\tilde{Y}_{ik} = 1)} \\ &= 1 - \frac{P(Y_{ik} = 1) - PS_{e:ik}P(\tilde{Y}_{ik} = 1)}{1 - P(\tilde{Y}_{ik} = 1)} \end{aligned}$$

The probability of a true positive individual is found by taking into account which disease k is of interest. For example, when $k = 1$ and $K = 2$, we have $P(\tilde{Y}_{ik} = 1) = p_{i1+}$. To find $P(Y_{ik} = 1)$, we proceed in a similar manner as when $P(G_{sj+}^{(1)} > 0, \dots, G_{sj+}^{(s)} > 0)$ was found earlier:

$$\begin{aligned} P(Y_{ik} = 1) &= P(G_{Lj+}^{(1)} > 0, \dots, G_{Lj+}^{(L-1)} > 0, G_{Ljk} = 1) \\ &= \sum_{\tilde{\mathbf{g}}_{Lj}^{(1)}} \dots \sum_{\tilde{\mathbf{g}}_{Lj}^{(L-1)}} \sum_{\tilde{g}_{Ljk}} P(G_{Lj+}^{(1)} > 0, \dots, G_{Lj+}^{(L-1)} > 0, G_{Ljk} = 1, \tilde{\mathbf{G}}_{Lj}^{(1)} = \tilde{\mathbf{g}}_{Lj}^{(1)}, \dots, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = \tilde{\mathbf{g}}_{Lj}^{(L-1)}, \tilde{G}_{Ljk} = \tilde{g}_{Ljk}) \\ &= \sum_{\tilde{\mathbf{g}}_{Lj}^{(1)}} \dots \sum_{\tilde{\mathbf{g}}_{Lj}^{(L-1)}} \sum_{\tilde{g}_{Ljk}} P(G_{Lj+}^{(1)} > 0, \dots, G_{Lj+}^{(L-1)} > 0, G_{Ljk} = 1 | \tilde{\mathbf{G}}_{Lj}^{(1)} = \tilde{\mathbf{g}}_{Lj}^{(1)}, \dots, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = \tilde{\mathbf{g}}_{Lj}^{(L-1)}, \tilde{G}_{Ljk} = \tilde{g}_{Ljk}) \times \\ &\quad P(\tilde{\mathbf{G}}_{Lj}^{(1)} = \tilde{\mathbf{g}}_{Lj}^{(1)}, \dots, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = \tilde{\mathbf{g}}_{Lj}^{(L-1)}, \tilde{G}_{Ljk} = \tilde{g}_{Ljk}) \\ &= \sum_{\tilde{\mathbf{g}}_{Lj}^{(1)}} \dots \sum_{\tilde{\mathbf{g}}_{Lj}^{(L-1)}} \sum_{\tilde{g}_{Ljk}} P(G_{Ljk} = 1 | \tilde{G}_{Ljk} = \tilde{g}_{Ljk}) \prod_{s'=1}^{L-1} P(G_{Lj+}^{(s')} > 0 | \tilde{\mathbf{G}}_{Lj}^{(s')} = \tilde{\mathbf{g}}_{Lj}^{(s')}) \times \\ &\quad P(\tilde{\mathbf{G}}_{Lj}^{(1)} = \tilde{\mathbf{g}}_{Lj}^{(1)}) P(\tilde{G}_{Ljk} = \tilde{g}_{Ljk} | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = \tilde{\mathbf{g}}_{Lj}^{(L-1)}) \prod_{s'=1}^{L-2} P(\tilde{\mathbf{G}}_{Lj}^{(s'+1)} = \tilde{\mathbf{g}}_{Lj}^{(s'+1)} | \tilde{\mathbf{G}}_{Lj}^{(s')} = \tilde{\mathbf{g}}_{Lj}^{(s')}) \\ &= \sum_{\tilde{\mathbf{g}}_{Lj}^{(1)}} \dots \sum_{\tilde{\mathbf{g}}_{Lj}^{(L-1)}} \sum_{\tilde{g}_{Ljk}} S_{e:Ljk}^{\tilde{g}_{Ljk}} (1 - S_{p:Ljk})^{1-\tilde{g}_{Ljk}} \prod_{s'=1}^{L-1} \left[1 - P(G_{Lj+}^{(s')} = 0 | \tilde{\mathbf{G}}_{Lj}^{(s')} = \tilde{\mathbf{g}}_{Lj}^{(s')}) \right] \times \\ &\quad P(\tilde{\mathbf{G}}_{Lj}^{(1)} = \tilde{\mathbf{g}}_{Lj}^{(1)}) P(\tilde{G}_{Ljk} = \tilde{g}_{Ljk} | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = \tilde{\mathbf{g}}_{Lj}^{(L-1)}) \prod_{s'=1}^{L-2} P(\tilde{\mathbf{G}}_{Lj}^{(s'+1)} = \tilde{\mathbf{g}}_{Lj}^{(s'+1)} | \tilde{\mathbf{G}}_{Lj}^{(s')} = \tilde{\mathbf{g}}_{Lj}^{(s')}) \\ &= \sum_{\tilde{\mathbf{g}}_{Lj}^{(1)}} \dots \sum_{\tilde{\mathbf{g}}_{Lj}^{(L-1)}} \sum_{\tilde{g}_{Ljk}} S_{e:Ljk}^{\tilde{g}_{Ljk}} (1 - S_{p:Ljk})^{1-\tilde{g}_{Ljk}} \prod_{s'=1}^{L-1} \left[1 - \prod_{k=1}^K (1 - S_{e:Ljk}^{(s')})^{\tilde{g}_{Ljk}^{(s')}} (S_{p:Ljk}^{(s')})^{1-\tilde{g}_{Ljk}^{(s')}} \right] \times \\ &\quad P(\tilde{\mathbf{G}}_{Lj}^{(1)} = \tilde{\mathbf{g}}_{Lj}^{(1)}) P(\tilde{G}_{Ljk} = \tilde{g}_{Ljk} | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = \tilde{\mathbf{g}}_{Lj}^{(L-1)}) \prod_{s'=1}^{L-2} P(\tilde{\mathbf{G}}_{Lj}^{(s'+1)} = \tilde{\mathbf{g}}_{Lj}^{(s'+1)} | \tilde{\mathbf{G}}_{Lj}^{(s')} = \tilde{\mathbf{g}}_{Lj}^{(s')}) \end{aligned}$$

Note that $P(\tilde{\mathbf{G}}_{Lj}^{(1)} = \tilde{\mathbf{g}}_{Lj}^{(1)})$ is equivalently expressed as $P(\tilde{\mathbf{G}}_{11} = \tilde{\mathbf{g}}_{11})$ because there is only one group at stage 1. Also, note that we use $S_{e:Ljk} \equiv S_{e:Ljk}^{(L)}$ and $S_{p:Ljk} \equiv S_{p:Ljk}^{(L)}$ in the above equation for notational convenience.

We focus on $K = 2$ diseases for the remainder of this subsection and discuss $K \geq 3$ in Section C.3. Probabilities for $P(\tilde{\mathbf{G}}_{Lj}^{(s'+1)} = \tilde{\mathbf{g}}_{Lj}^{(s'+1)} | \tilde{\mathbf{G}}_{Lj}^{(s')} = \tilde{\mathbf{g}}_{Lj}^{(s')})$ and $P(\tilde{\mathbf{G}}_{Lj}^{(1)} = \tilde{\mathbf{g}}_{Lj}^{(1)})$ are already given in Section C.1 when finding $E(T|\mathbf{M})$. To find $P(\tilde{G}_{Ljk} = \tilde{g}_{Ljk} | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = \tilde{\mathbf{g}}_{Lj}^{(L-1)})$, we examine the following cases:

- $\tilde{G}_{Ljk} = 0, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0)$

Because $\tilde{G}_{Ljk}^{(L-1)} = 0 \Rightarrow \tilde{G}_{Ljk} = 0$,

$$P(\tilde{G}_{Ljk} = 0 | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0)) = 1.$$

- $\tilde{G}_{Ljk} = 1, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0)$

A positive status cannot occur at stage L if the statuses are negative at stage $L - 1$. Therefore,

$$P(\tilde{G}_{Ljk} = 1 | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0)) = 0.$$

- $\tilde{G}_{Ljk} = 0, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 1)$

For $k = 1$: Because $\tilde{G}_{Lj1}^{(L-1)} = 0 \Rightarrow \tilde{G}_{Lj1} = 0$,

$$P(\tilde{G}_{Ljk} = 0 | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 1)) = 1.$$

For $k = 2$:

$$\begin{aligned} P(\tilde{G}_{Lj2} = 0 | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 1)) &= \frac{P(\tilde{G}_{Lj2} = 0, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 1))}{P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 1))} \\ &= \frac{P(\tilde{G}_{Lj2} = 0, \tilde{G}_{Lj1}^{(L-1)} = 0) - P(\tilde{G}_{Lj2} = 0, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))}{P(\tilde{G}_{Lj1}^{(L-1)} = 0) - P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))} \\ &= \frac{p_{i00} \left(\prod_{b \in B_{Lj}^{(L-1)} b \neq i} p_{b0+} \right) - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)}{\left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b0+} \right) - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)} \end{aligned}$$

- $\tilde{G}_{Ljk} = 1, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 1)$

For $k = 1$: A positive status for disease 1 cannot occur at stage L if the disease 1 status is negative at stage $L - 1$. Therefore,

$$P(\tilde{G}_{Lj1} = 1 | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 1)) = 0.$$

For $k = 2$:

$$\begin{aligned}
P(\tilde{G}_{Lj2} = 1 | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 1)) &= \frac{P(\tilde{G}_{Lj2} = 1, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 1))}{P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 1))} \\
&= \frac{P(\tilde{G}_{Lj2} = 1, \tilde{G}_{Lj1}^{(L-1)} = 0) - P(\tilde{G}_{Lj2} = 1, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))}{P(\tilde{G}_{Lj1}^{(L-1)} = 0) - P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))} \\
&= \frac{P(\tilde{G}_{Lj1}^{(L-1)} = 0) - P(\tilde{G}_{Lj2} = 0, \tilde{G}_{Lj1}^{(L-1)} = 0) - 0}{\left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b0+} \right) - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)} \\
&= \frac{\left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b0+} \right) - p_{i00} \left(\prod_{b \in B_{Lj}^{(L-1)} b \neq i} p_{b0+} \right)}{\left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b0+} \right) - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)}
\end{aligned}$$

- $\tilde{G}_{Ljk} = 0, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0)$

For $k = 1$:

$$\begin{aligned}
P(\tilde{G}_{Lj1} = 0 | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0)) &= \frac{P(\tilde{G}_{Lj1} = 0, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0))}{P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0))} \\
&= \frac{P(\tilde{G}_{Lj1} = 0, \tilde{G}_{Lj2}^{(L-1)} = 0) - P(\tilde{G}_{Lj1} = 0, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))}{P(\tilde{G}_{Lj2}^{(L-1)} = 0) - P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))} \\
&= \frac{p_{i00} \left(\prod_{b \in B_{Lj}^{(L-1)} b \neq i} p_{b+0} \right) - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)}{\left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b+0} \right) - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)}
\end{aligned}$$

For $k = 2$: Because $\tilde{G}_{Lj2}^{(L-1)} = 0 \Rightarrow \tilde{G}_{Lj2} = 0$,

$$P(\tilde{G}_{Lj2} = 0 | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0)) = 1.$$

- $\tilde{G}_{Ljk} = 1, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0)$

For $k = 1$:

$$\begin{aligned}
P(\tilde{G}_{Lj1} = 1 | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0)) &= \frac{P(\tilde{G}_{Lj1} = 1, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0))}{P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0))} \\
&= \frac{P(\tilde{G}_{Lj1} = 1, \tilde{G}_{Lj2}^{(L-1)} = 0) - P(\tilde{G}_{Lj1} = 1, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))}{P(\tilde{G}_{Lj2}^{(L-1)} = 0) - P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))} \\
&= \frac{P(\tilde{G}_{Lj2}^{(L-1)} = 0) - P(\tilde{G}_{Lj1} = 0, \tilde{G}_{Lj2}^{(L-1)} = 0) - 0}{\prod_{b \in B_{Lj}^{(L-1)}} p_{b+0} - \prod_{b \in B_{Lj}^{(L-1)}} p_{b00}} \\
&= \frac{\left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b+0} \right) - p_{i00} \left(\prod_{b \in B_{Lj}^{(L-1)} b \neq i} p_{b+0} \right)}{\left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b+0} \right) - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)}
\end{aligned}$$

For $k = 2$: A positive status for disease 2 cannot occur at stage L if the disease 2 status is negative at stage $L - 1$. Therefore,

$$P(\tilde{G}_{Lj2} = 1 | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0)) = 0.$$

- $\tilde{G}_{Ljk} = 0, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 1)$

For $k = 1$:

$$\begin{aligned}
P(\tilde{G}_{Lj1} = 0 | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 1)) &= \frac{P(\tilde{G}_{Lj1} = 0, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 1))}{P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 1))} \\
&= \frac{P(\tilde{G}_{Lj1} = 0, \tilde{G}_{Lj1}^{(L-1)} = 1) - P(\tilde{G}_{Lj1} = 0, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0))}{P(\tilde{G}_{Lj1}^{(L-1)} = 1) - P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0))} \\
&= \frac{P(\tilde{G}_{Lj1} = 0) - P(\tilde{G}_{Lj1} = 0, \tilde{G}_{Lj1}^{(L-1)} = 0) - P(\tilde{G}_{Lj1} = 0, \tilde{G}_{Lj2}^{(L-1)} = 0) + P(\tilde{G}_{Lj1} = 0, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))}{1 - P(\tilde{G}_{Lj1}^{(L-1)} = 0) - P(\tilde{G}_{Lj2}^{(L-1)} = 0) + P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))} \\
&= \frac{p_{i0+} - P(\tilde{G}_{Lj1}^{(L-1)} = 0) - p_{i00} \left(\prod_{b \in B_{Lj}^{(L-1)} b \neq i} p_{b+0} \right) + P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))}{1 - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b0+} \right) - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b+0} \right) + \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)} \\
&= \frac{p_{i0+} - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b0+} \right) - p_{i00} \left(\prod_{b \in B_{Lj}^{(L-1)} b \neq i} p_{b+0} \right) + \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)}{1 - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b0+} \right) - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b+0} \right) + \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)}
\end{aligned}$$

For $k = 2$:

$$\begin{aligned}
& P(\tilde{G}_{Lj2} = 0 | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 1)) \\
&= \frac{P(\tilde{G}_{Lj2} = 0, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 1))}{P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 1))} \\
&= \frac{P(\tilde{G}_{Lj2} = 0, \tilde{G}_{Lj1}^{(L-1)} = 1) - P(\tilde{G}_{Lj2} = 0, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0))}{P(\tilde{G}_{Lj1}^{(L-1)} = 1) - P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0))} \\
&= \frac{P(\tilde{G}_{Lj2} = 0) - P(\tilde{G}_{Lj2} = 0, \tilde{G}_{Lj1}^{(L-1)} = 0) - P(\tilde{G}_{Lj2} = 0, \tilde{G}_{Lj2}^{(L-1)} = 0) + P(\tilde{G}_{Lj2} = 0, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))}{1 - P(\tilde{G}_{Lj1}^{(L-1)} = 0) - P(\tilde{G}_{Lj2}^{(L-1)} = 0) + P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))} \\
&= \frac{p_{i+0} - p_{i00} \left(\prod_{b \in B_{Lj}^{(L-1)} b \neq i} p_{b0+} \right) - P(\tilde{G}_{Lj2}^{(L-1)} = 0) + P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))}{1 - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b0+} \right) - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b+0} \right) + \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)} \\
&= \frac{p_{i+0} - p_{i00} \left(\prod_{b \in B_{Lj}^{(L-1)} b \neq i} p_{b0+} \right) - \left(\prod_{b \in B_{Lj}^{(L-1)} b \neq i} p_{b+0} \right) + \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)}{1 - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b0+} \right) - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b+0} \right) + \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)}
\end{aligned}$$

- $\tilde{G}_{Ljk} = 1, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 1)$

For $k = 1$:

$$\begin{aligned}
& P(\tilde{G}_{Lj1} = 1 | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 1)) \\
&= \frac{P(\tilde{G}_{Lj1} = 1, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 1))}{P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 1))} \\
&= \frac{P(\tilde{G}_{Lj1} = 1, \tilde{G}_{Lj1}^{(L-1)} = 1) - P(\tilde{G}_{Lj1} = 1, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0))}{P(\tilde{G}_{Lj1}^{(L-1)} = 1) - P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0))} \\
&= \frac{P(\tilde{G}_{Lj1} = 1) - P(\tilde{G}_{Lj1} = 1, \tilde{G}_{Lj1}^{(L-1)} = 0) - P(\tilde{G}_{Lj1} = 1, \tilde{G}_{Lj2}^{(L-1)} = 0) + P(\tilde{G}_{Lj1} = 1, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))}{1 - P(\tilde{G}_{Lj1}^{(L-1)} = 0) - P(\tilde{G}_{Lj2}^{(L-1)} = 0) + P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))} \\
&= \frac{1 - P(\tilde{G}_{Lj1} = 0) - 0 - P(\tilde{G}_{Lj2}^{(L-1)} = 0) + P(\tilde{G}_{Lj1} = 0, \tilde{G}_{Lj2}^{(L-1)} = 0) + 0}{1 - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b0+} \right) - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b+0} \right) + \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)} \\
&= \frac{1 - p_{i0+} - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b+0} \right) + p_{i00} \left(\prod_{b \in B_{Lj}^{(L-1)} b \neq i} p_{b+0} \right)}{1 - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b0+} \right) - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b+0} \right) + \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)}
\end{aligned}$$

For $k = 2$:

$$\begin{aligned}
& P(\tilde{G}_{Lj2} = 1 | \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 1)) \\
&= \frac{P(\tilde{G}_{Lj2} = 1, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 1))}{P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 1))} \\
&= \frac{P(\tilde{G}_{Lj2} = 1, \tilde{G}_{Lj1}^{(L-1)} = 1) - P(\tilde{G}_{Lj2} = 1, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0))}{P(\tilde{G}_{Lj1}^{(L-1)} = 1) - P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (1, 0))} \\
&= \frac{P(\tilde{G}_{Lj2} = 1) - P(\tilde{G}_{Lj2} = 1, \tilde{G}_{Lj1}^{(L-1)} = 0) - P(\tilde{G}_{Lj2} = 1, \tilde{G}_{Lj2}^{(L-1)} = 0) + P(\tilde{G}_{Lj2} = 1, \tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))}{1 - P(\tilde{G}_{Lj1}^{(L-1)} = 0) - P(\tilde{G}_{Lj2}^{(L-1)} = 0) + P(\tilde{\mathbf{G}}_{Lj}^{(L-1)} = (0, 0))} \\
&= \frac{1 - P(\tilde{G}_{Lj2} = 0) - P(\tilde{G}_{Lj1}^{(L-1)} = 0) + P(\tilde{G}_{Lj2} = 0, \tilde{G}_{Lj1}^{(L-1)} = 0) - 0 + 0}{1 - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b0+} \right) - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b+0} \right) + \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)} \\
&= \frac{1 - p_{i+0} - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b0+} \right) + p_{i00} \left(\prod_{b \in B_{Lj}^{(L-1)} b \neq i} p_{b0+} \right)}{1 - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b0+} \right) - \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b+0} \right) + \left(\prod_{b \in B_{Lj}^{(L-1)}} p_{b00} \right)}
\end{aligned}$$

C.3 Discussion

Parts of our derivations focused on the $K = 2$ disease setting. Direct extensions for $K \geq 3$ are definitely possible for those cases. However, the number of terms needed for probabilities grows quickly, as was shown for Equation (4) with 2^{2K} separate conditional probabilities needed for each stage. Also, we are aware only of multiplex assays up to $K = 3$ being used in current infectious disease applications. Still, due to time and cost efficiencies, we expect in the future that multiplex assays will continue to be developed to test for more diseases. For example, the American Red Cross currently screens blood donations for seven separate diseases, but only three of these diseases are screened via one mutiplex assay ([American Red Cross, 2018](#)). Overall, infectious disease testing could greatly benefit by increasing the number of diseases per assay and by the development of the corresponding statistical methodology to show how to best take advantage of these assays.

D R functions

Programs that can be used to reproduce the research in this paper are available with our supplementary materials and on our research website at www.chrisbilder.com/grouptesting. In summary, we provide programs with R functions that

- Calculate $E(T)$ and accuracy measures
- Create Figures 1, 2, and 3 of the paper
- Emulate two-stage and three-stage informative group testing

E Algorithm performance

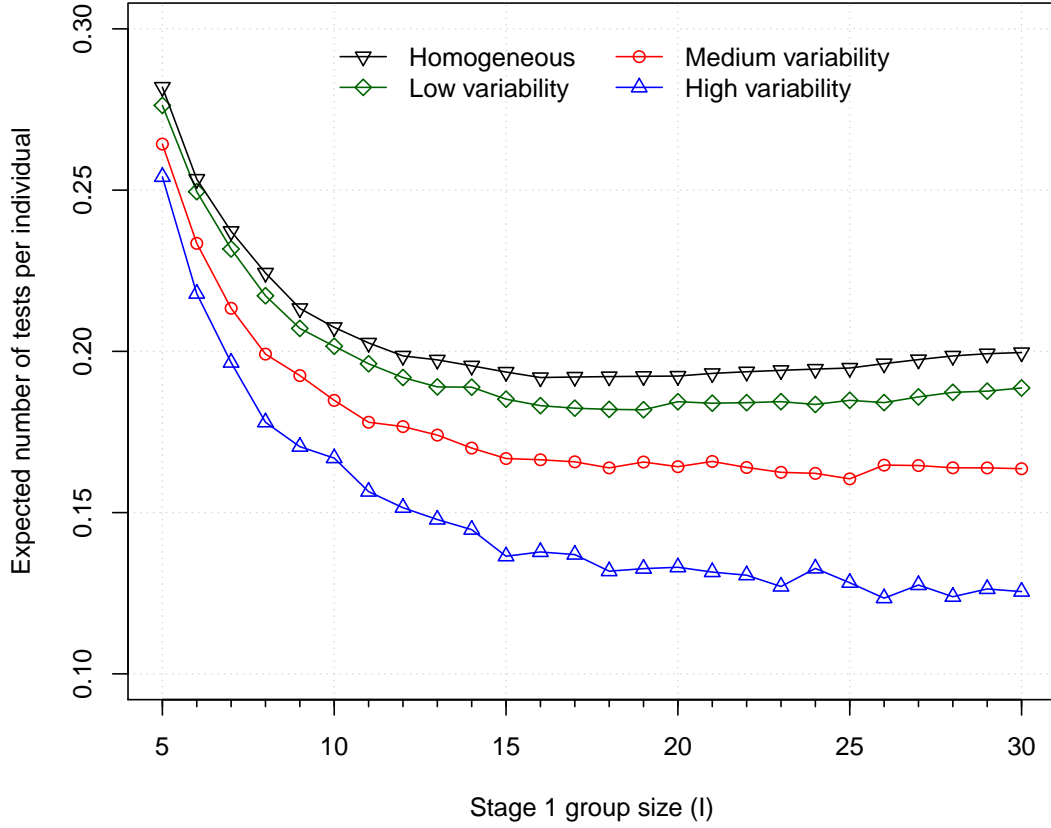
Below is additional information regarding the Dirichlet distributions discussed in Section 4 of the paper. After the bullets, we discuss Web Figures 2 - 7 of the supplementary materials.

- Figures 1 and 2
 - $E(P_{i1+}) = E(P_{i+1}) = 0.05$
 - $E(P_{i00}) = 0.9125$, $E(P_{i01}) = E(P_{i10}) = 0.0375$, and $E(P_{i11}) = 0.0125$
 - $\widetilde{OR} = 8.11$
 - $\alpha = (18.25, 0.75, 0.75, 0.25)$ for medium variability; 4α for low variability and $\alpha/4$ for high variability
 - $Var(P_{i1+}) = Var(P_{i+1})$: 0.0005864 for low variability, 0.002262 for medium variability, and 0.007917 for high variability
- Web Figures 2 and 3
 - $E(P_{i1+}) = E(P_{i+1}) = 0.01$
 - $E(P_{i00}) = 0.9825$, $E(P_{i01}) = E(P_{i10}) = 0.0075$, and $E(P_{i11}) = 0.0025$
 - $\widetilde{OR} = 43.67$
 - $\alpha = (19.65, 0.15, 0.15, 0.05)$ for medium variability; 4α for low variability and $\alpha/4$ for high variability
 - $Var(P_{i1+}) = Var(P_{i+1})$: 0.0001222 for low variability, 0.0004714 for medium variability, and 0.001650 for high variability
- Web Figures 4 and 5
 - $E(P_{i1+}) = E(P_{i+1}) = 0.10$
 - $E(P_{i00}) = 0.8250$, $E(P_{i01}) = E(P_{i10}) = 0.0750$, and $E(P_{i11}) = 0.0250$
 - $\widetilde{OR} = 3.67$
 - $\alpha = (16.50, 1.50, 1.50, 0.50)$ for medium variability; 4α for low variability and $\alpha/4$ for high variability
 - $Var(P_{i1+}) = Var(P_{i+1})$: 0.001111 for low variability, 0.004286 for medium variability, and 0.01500 for high variability
- Web Figures 6 and 7
 - $E(P_{i1+}) = E(P_{i+1}) = 0.05$
 - \widetilde{OR} , α , and $E(\mathbf{P}_i)$:

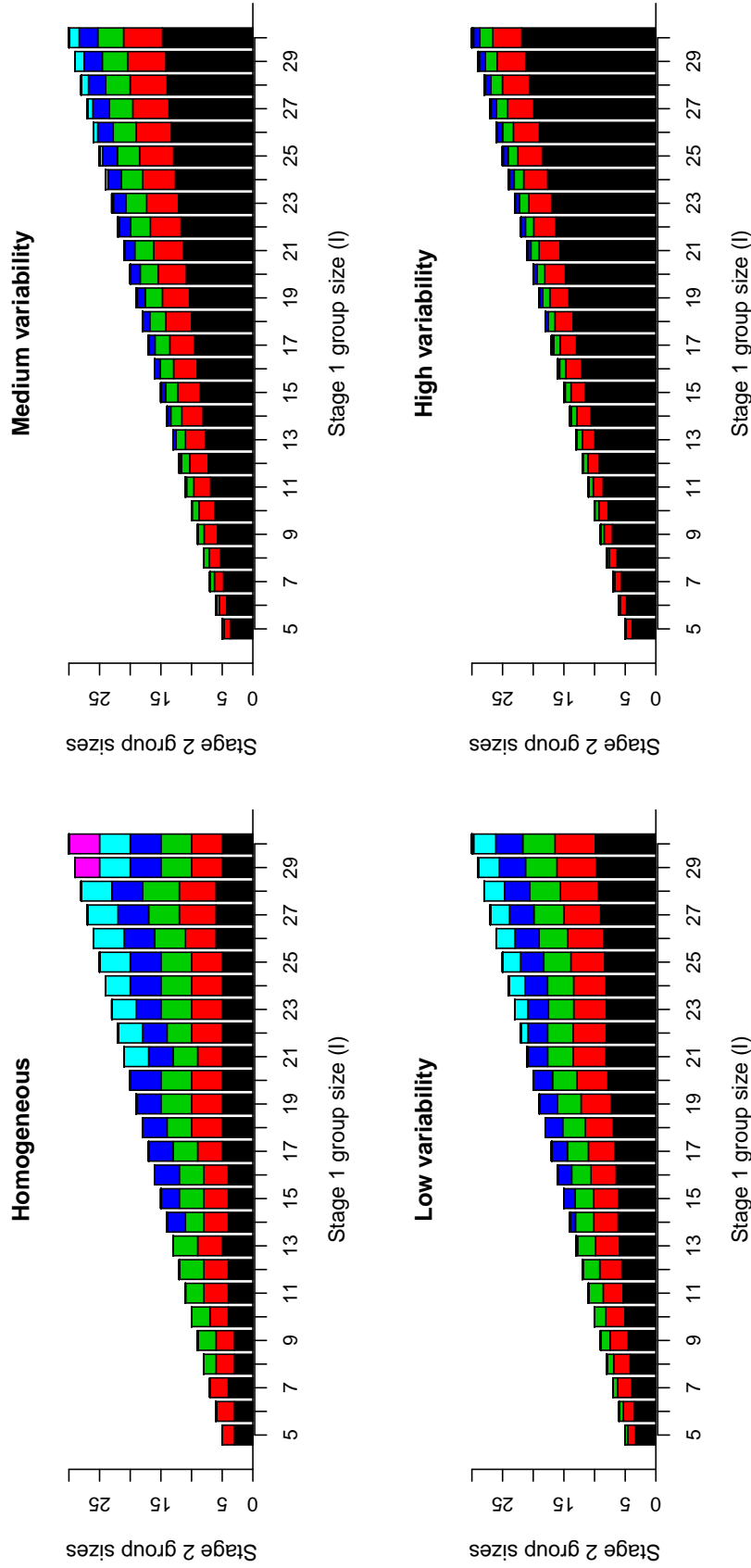
\widetilde{OR}	α	$E(\mathbf{P}_i)$
225	(18.75, 0.25, 0.25, 0.75)	(0.9375, 0.0125, 0.0125, 0.0375)
37.0	(18.5, 0.5, 0.5, 0.5)	(0.9250, 0.0250, 0.0250, 0.0250)
8.11	(18.25, 0.75, 0.75, 0.25)	(0.9125, 0.0375, 0.0375, 0.0125)
1.00	(18.05, 0.95, 0.95, 0.05)	(0.9025, 0.0475, 0.0475, 0.0025)
0.18	(18.01, 0.99, 0.99, 0.01)	(0.9005, 0.0495, 0.0495, 0.0005)
0.018	(18.001, 0.999, 0.999, 0.001)	(0.90005, 0.04995, 0.04995, 0.00005)

Web Figures 2 - 5 illustrate the same behaviors as shown in Figures 1 and 2 of the paper. Specifically, as the variability in the marginal probabilities increases: 1) the average expected number of tests per individual decreases and 2) the variability in the average group sizes increases. The average number of groups change as well, but not in a consistent manner across the three marginal expected probability cases.

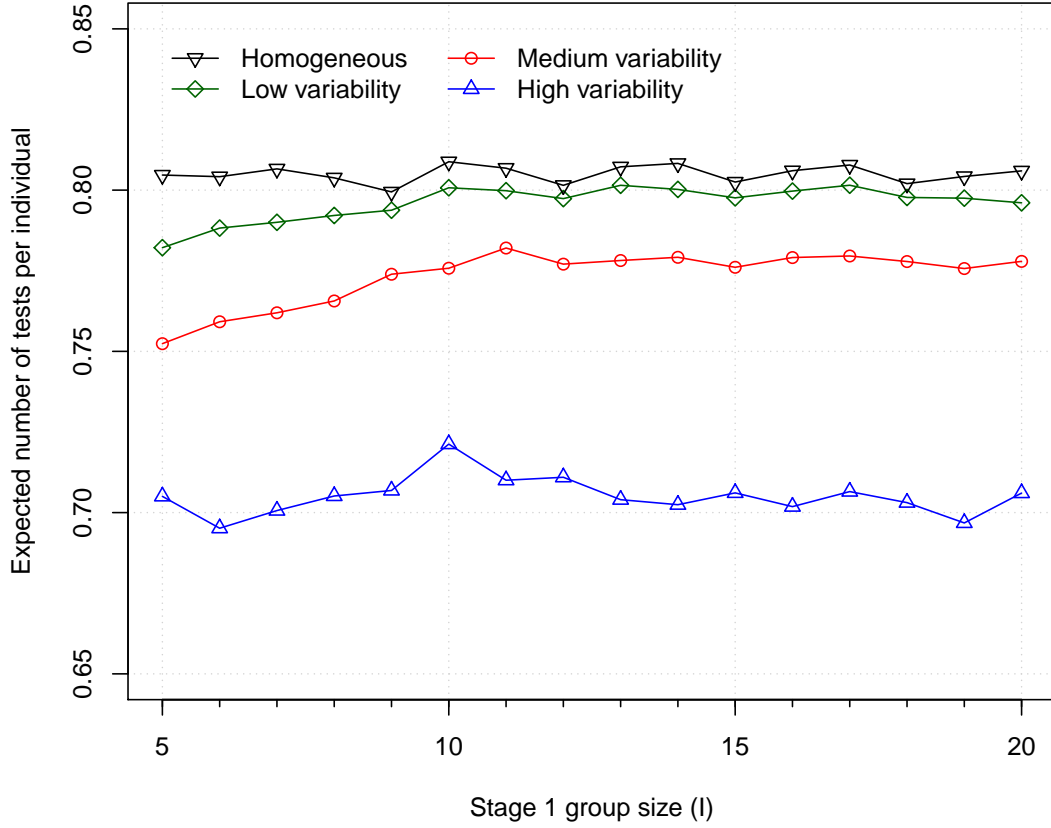
Web Figures 6 and 7 examine the average expected number of tests per individual and the average group sizes, respectively, for $E(P_{i1+}) = E(P_{i+1}) = 0.05$ and $Var(P_{i1+}) = Var(P_{i+1}) = 0.002262$. The value of α



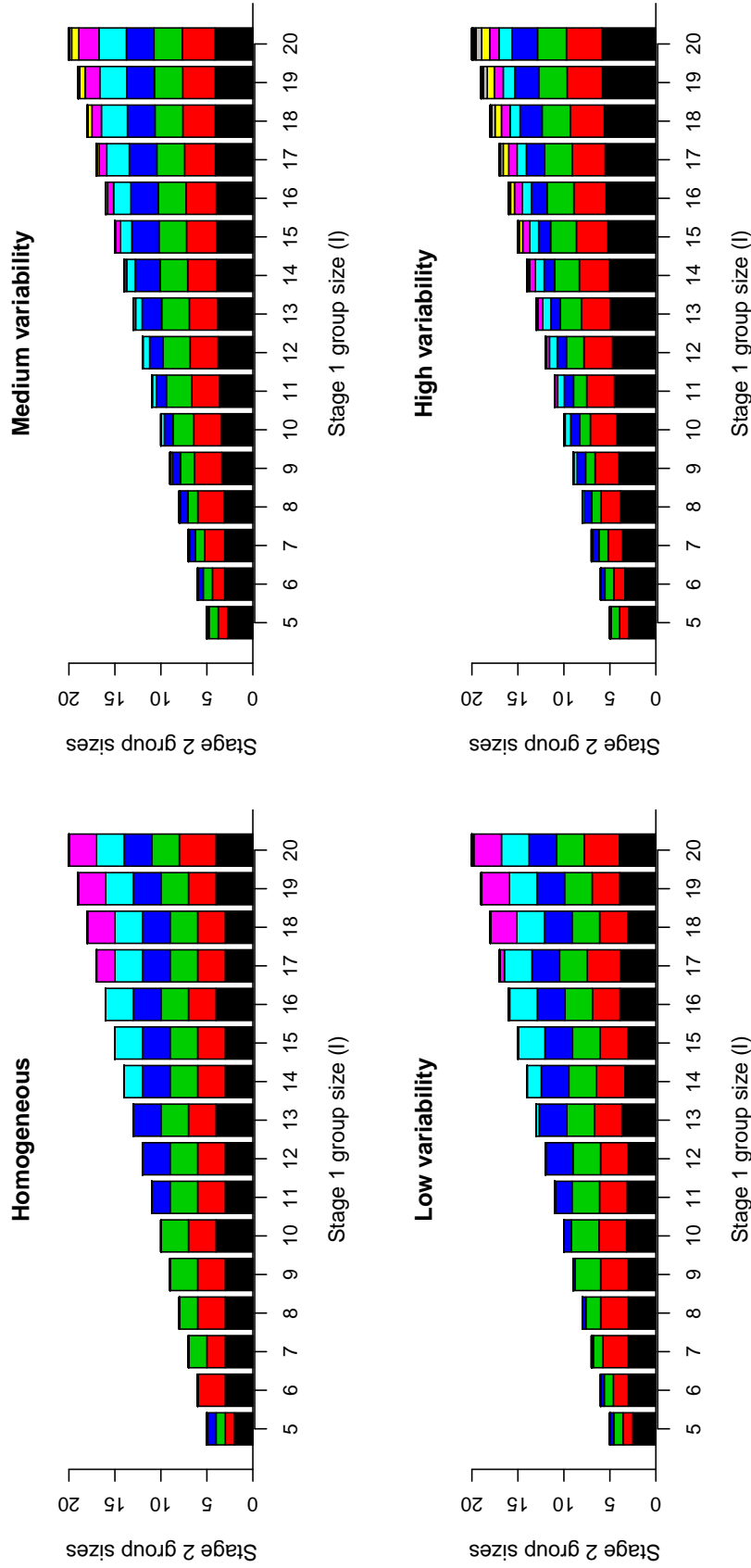
Web Figure 2: Averaged expected number of tests per individual across 500 separate simulations for each I , where $\mathbf{P}_i \stackrel{iid}{\sim} \text{Dirichlet}(\cdot)$, $E(P_{i1+}) = E(P_{i+1}) = 0.01$, and $S = 3$. Results for informative group testing are given by low ($\text{Dirichlet}(4\alpha)$), medium ($\text{Dirichlet}(\alpha)$), and high ($\text{Dirichlet}(\alpha/4)$) variability cases for \mathbf{P}_i with $\alpha = (19.65, 0.15, 0.15, 0.05)$. Results for non-informative group testing are given by the homogeneous case, where \mathbf{P}_i is set to $E(\mathbf{P}_i)$ for each individual and the Dirichlet parameter vector is the same as used for medium variability.



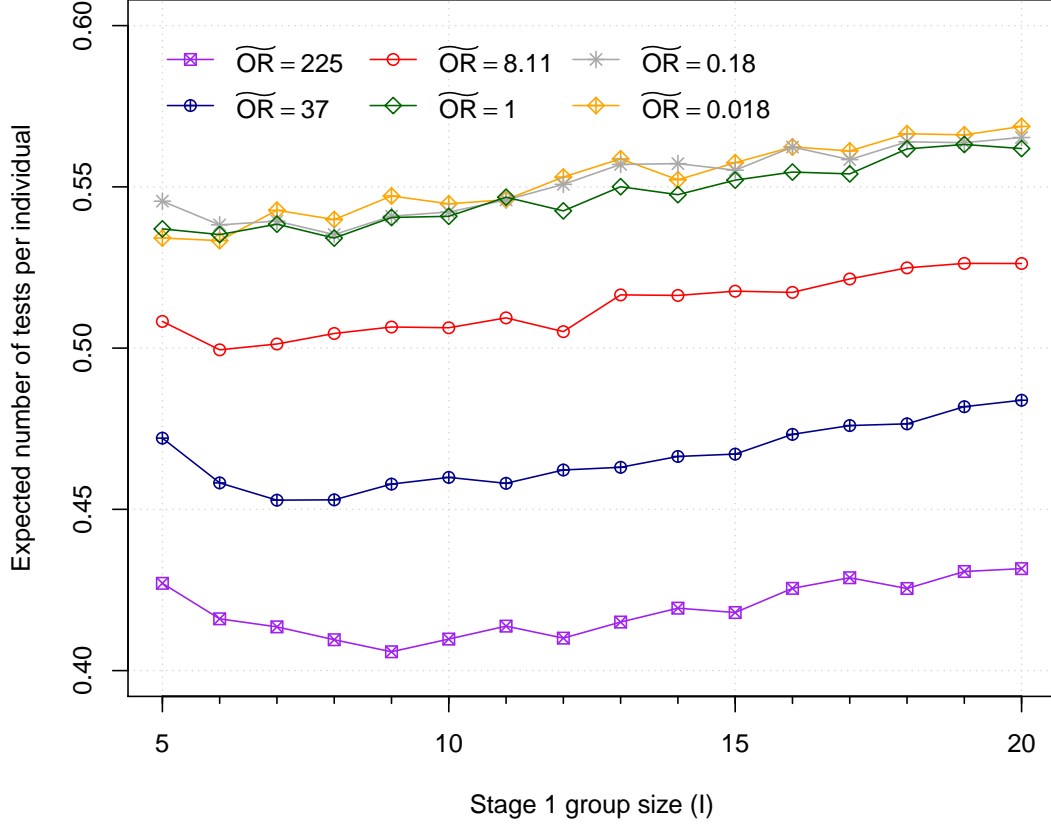
Web Figure 3: Stacked bar chart of average stage 2 group sizes across 500 separate simulations for each l , where $\mathbf{P}_i \stackrel{iid}{\sim} \text{Dirichlet}(\cdot)$, $E(P_{i+1}) = E(P_{i+1}) = 0.01$, and $S = 3$. From the bottom to the top of each bar, the average group size is given for the first group containing the largest p_{i00} values to the last group containing the smallest p_{i00} values. Please see Web Figure 2 for a description of the homogeneous (non-informative group testing) and variability cases.



Web Figure 4: Averaged expected number of tests per individual across 500 separate simulations for each l , where $\mathbf{P}_i \stackrel{iid}{\sim} \text{Dirichlet}(\cdot)$, $E(P_{i1+}) = E(P_{i+1}) = 0.10$, and $S = 3$. Results for informative group testing are given by low ($\text{Dirichlet}(4\alpha)$), medium ($\text{Dirichlet}(\alpha)$), and high ($\text{Dirichlet}(\alpha/4)$) variability cases for \mathbf{P}_i with $\alpha = (16.50, 1.50, 1.50, 0.50)$. Results for non-informative group testing are given by the homogeneous case, where \mathbf{P}_i is set to $E(\mathbf{P}_i)$ for each individual and the Dirichlet parameter vector is the same as used for medium variability.



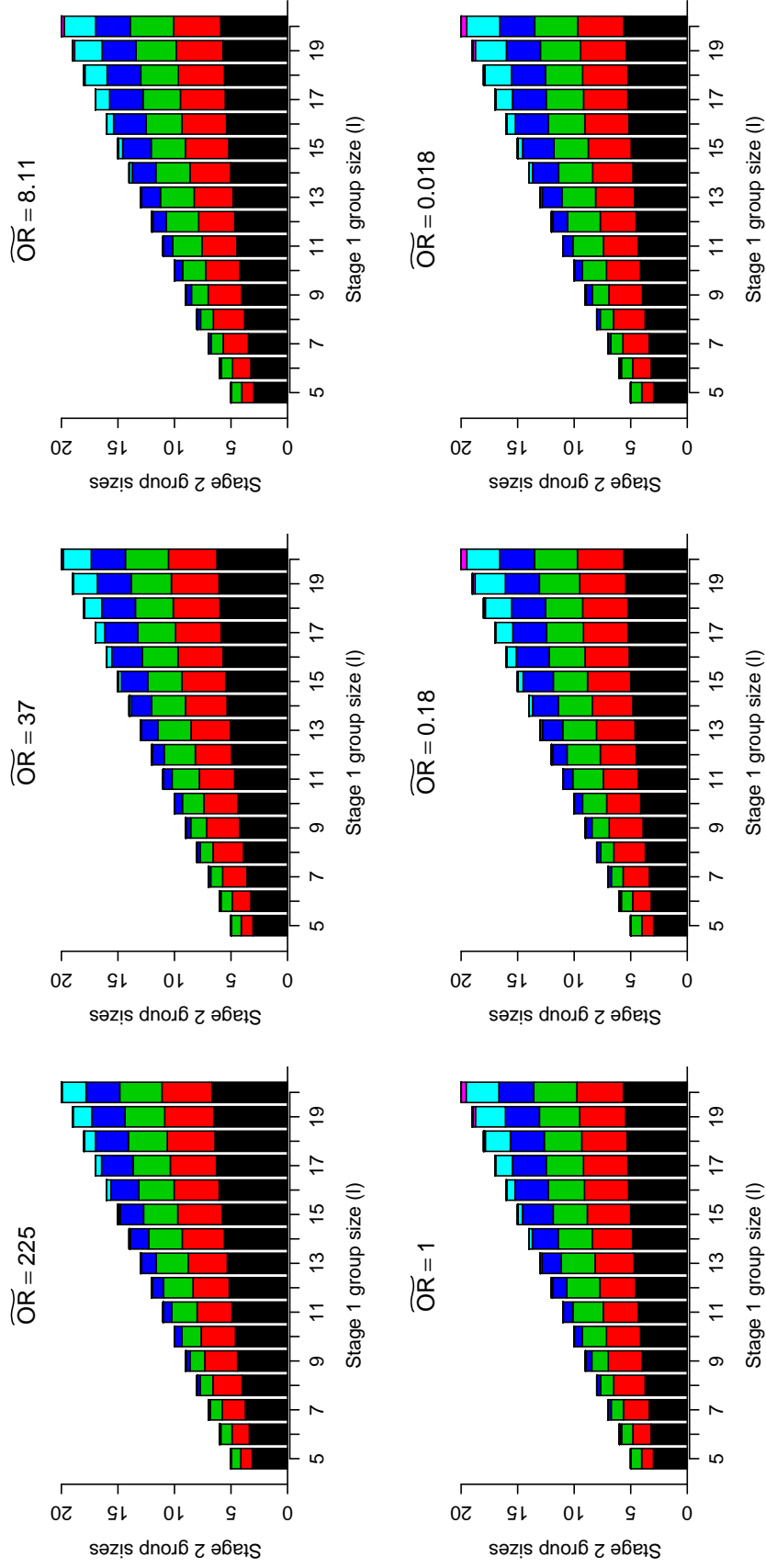
Web Figure 5: Stacked bar chart of average stage 2 group sizes across 500 separate simulations for each l , where $\mathbf{P}_i \stackrel{iid}{\sim} \text{Dirichlet}(\cdot)$, $E(P_{i+1}) = E(P_{i+1}) = 0.10$, and $S = 3$. From the bottom to the top of each bar, the average group size is given for the first group containing the largest p_{100} values to the last group containing the smallest p_{100} values. Please see Web Figure 4 for a description of the homogeneous (non-informative group testing) and variability cases.



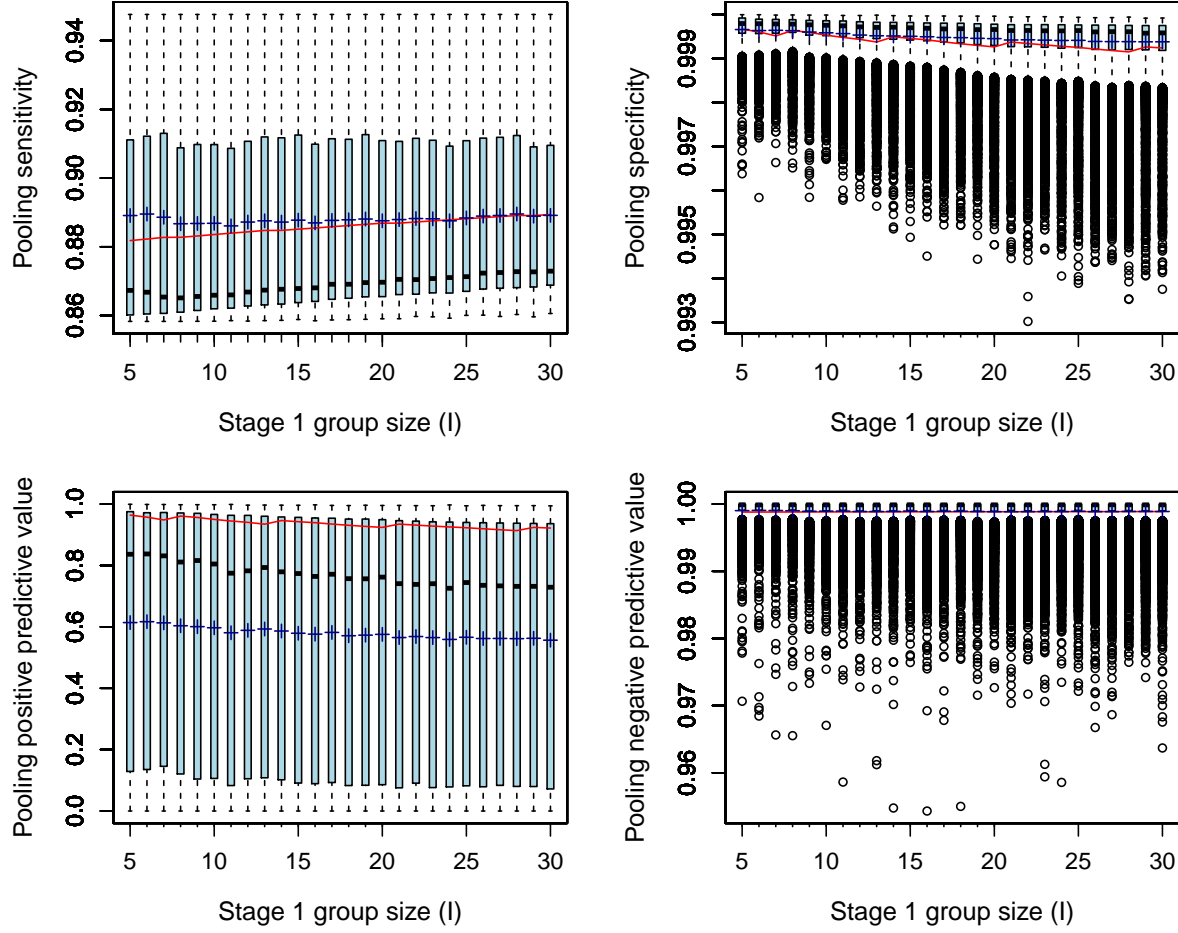
Web Figure 6: Averaged expected number of tests per individual across 500 separate simulations for each I , where $\mathbf{P}_i \stackrel{iid}{\sim} \text{Dirichlet}(\boldsymbol{\alpha})$, $E(P_{i1+}) = E(P_{i+1}) = 0.05$, $S = 3$, and \widetilde{OR} changing. The $\boldsymbol{\alpha}$ values are chosen so that $\text{Var}(P_{i1+}) = \text{Var}(P_{i+1}) = 0.002262$.

is chosen to allow \widetilde{OR} to vary among cases. Web Figure 6 shows that as the strength of positive dependence ($\widetilde{OR} > 1$) increases, the average expected number of tests per individual decreases. Web Figure 7 shows that only small differences in the average group sizes occur as the dependence decreases, where there is a propensity for slightly more groups of smaller sizes. For example, the average group size for the first stage 2 group with $\widetilde{OR} = 225$ is 6.7 at $I = 20$. For $\widetilde{OR} = 0.018$, this average group size decreases to 5.6. Similarly, the average group size for the seventh stage 2 group with $\widetilde{OR} = 225$ is 0.002 and with $\widetilde{OR} = 0.018$ is 0.014.

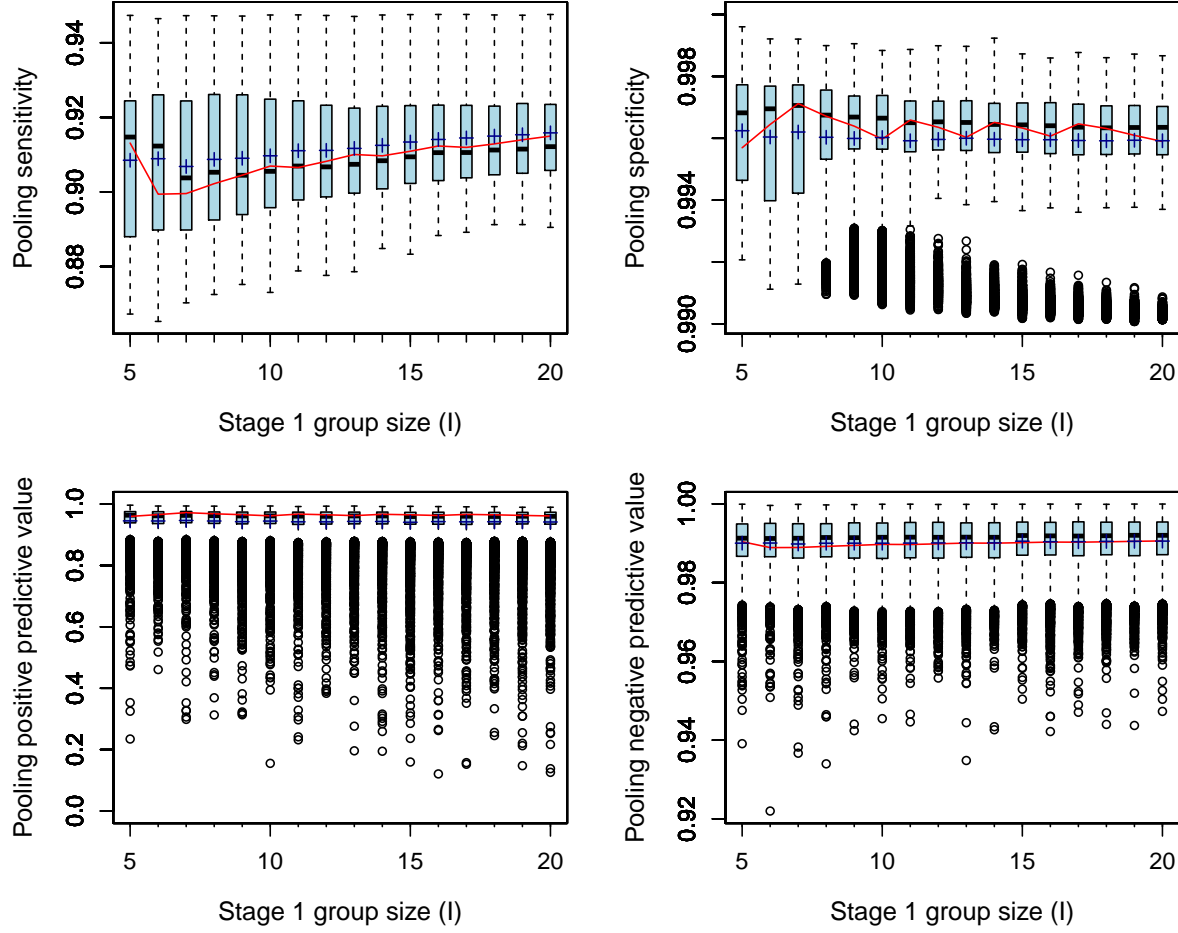
Web Figures 8 and 9 summarize the accuracy measures for the medium variability cases with $E(P_{i1+}) = E(P_{i+1}) = 0.01$ and $E(P_{i1+}) = E(P_{i+1}) = 0.10$, respectively. Like for Figure 3 in the paper, these plots are for disease $k = 1$ only. Overall, the means and medians for the pooling sensitivity and pooling specificity coincide with the means for non-informative group testing (homogeneous case). The separation of outliers from the lower whisker endpoint in some pooling specificity box plots is due to individuals not being tested in all three stages. This occurs when expected marginal probabilities are larger, which leads to some individuals being tested alone in stage 2. The plots also demonstrate the variability among the accuracy measures in comparison to what occurs for non-informative group testing. This variation is expected because each individual has potentially different probabilities of disease. The variability amount is largest for the pooling positive predictive value, where variability increases as the marginal probability of disease decreases. Very low PPP_{i1} values occur because very low p_{i1+} values occur (p_{i1+} is in the numerator of PPP_{i1}). Taking the average or even median of very low PPP_{i1} values can lead to a distorted measure of centrality.



Web Figure 7: Stacked bar chart of average stage 2 group sizes across 500 separate simulations for each I , where $\mathbf{P}_i \stackrel{iid}{\sim} \text{Dirichlet}(\alpha)$, $E(P_{i+}) = E(P_{i+1}) = 0.05$, $S = 3$, and \overline{OR} changing. The α values are chosen so that $\text{Var}(P_{i+}) = \text{Var}(P_{i+1}) = 0.002262$.



Web Figure 8: Box plots of the accuracy measures at each I for $\mathbf{P}_i \stackrel{iid}{\sim} \text{Dirichlet}(\boldsymbol{\alpha})$, $E(P_{i1+}) = E(P_{i+1}) = 0.01$, $S = 3$, and the medium variability case. Note that the y-axis scale is different for each accuracy measure. The plus symbol represents the mean value for an accuracy measure in the simulations. The location of the mean value for non-informative group testing is depicted by a line across the different I .



Web Figure 9: Box plots of the accuracy measures at each I for $\mathbf{P}_i \stackrel{iid}{\sim} \text{Dirichlet}(\boldsymbol{\alpha})$, $E(P_{i1+}) = E(P_{i+1}) = 0.10$, $S = 3$, and the medium variability case. Note that the y-axis scale is different for each accuracy measure. The plus symbol represents the mean value for an accuracy measure in the simulations. The location of the mean value for non-informative group testing is depicted by a line across the different I .

F Aptima Combo 2 Assay application

The Aptima Combo 2 Assay allows for a number of different types of specimens—urine, vaginal swab, rectal swab, etc.—that can be collected on individuals. In our analysis, we focus only on urine specimens for brevity reasons and because it clearly shows when there are performance differences (informative algorithms work better with larger variability in probabilities of disease).

Below is a summary of the steps followed to implement the informative group testing algorithms:

1. A multinomial regression model is fit to the training data.
2. Estimated chlamydia and/or gonorrhea disease probabilities are found using the model for every observation in the training data
3. These estimated joint probabilities of disease are treated as the true values in $E(T|\mathbf{M})$ to obtain a “best guess” for a training data OTC.
4. The training data model is used to estimate the probability of having at least one disease for individuals in the test data.
5. Monte Carlo simulation is used to simulate 500 separate data sets based on the test data so that assay testing error can be accounted for.
6. Group testing is applied to each of the 500 separate data sets using the OTC found from the training data.
7. The number of tests and accuracy measures are averaged over the 500 separate data sets to summarize the results.

The application of the non-informative group testing algorithms follow the same steps as given in their corresponding papers. These steps are essentially the same as above with the exception of using the observed disease prevalences from the training data rather than a multinomial regression model to estimate individual joint probabilities of disease.

The remaining text in this section provides more details regarding these steps. Section F.1 describes the multinomial regression models used with the training and test data. Special attention is given to the variability among the probabilities of having no diseases, because these probabilities play a critical role in understanding the performance of the informative group testing algorithms. Section F.2 shows how an OTC is estimated from the training data. Section F.3 explains how Monte Carlo simulation is used to account for the inherent testing error in the assay. Finally, Section F.4 summarizes the observed accuracy of the group testing algorithms.

F.1 Estimated joint probabilities of disease

Define \tilde{y}_1 as the positive (1) or negative (0) true status for chlamydia, and define \tilde{y}_2 in a similar manner for gonorrhea. The combination of $(\tilde{y}_1, \tilde{y}_2)$ is treated as a four-category response for a multinomial regression model (e.g., see Chapter 3 of Bilder and Loughin, 2014). We estimate these models separately for each state and gender combination in the training data and assume the observed disease outcomes to be the true disease statuses (we discuss reasons for this assumption in Section F.3). Due to differences in the information collected by each state and differences among genders, covariates may differ for each model. For this reason, we only provide detailed information regarding the covariates used for Idaho males and Iowa females. These two cases represent the largest and smallest reductions in the number of tests, respectively, when the informative algorithms are compared to the non-informative algorithms. Web Tables 1 and 2 provide the covariates and likelihood ratio test (LRT) p-values corresponding to their regression parameters. These LRTs are performed assuming that all other covariates given in the table are included in the model. All covariates are categorical except for age, which is included as a numerical value.

Once the models are obtained, we estimate the joint probabilities $\hat{p}_{\tilde{y}_1 \tilde{y}_2}$ for each individual in the test data (we omit the subscript i on $\hat{p}_{\tilde{y}_1 \tilde{y}_2}$ in this section for convenience). Summaries of these estimated probabilities are given in Web Table 3. This table includes maximum likelihood estimates for the parameters of a Dirichlet

Web Table 1: Covariates and LRT p-values for male urine specimens in the Idaho 2010 data. Note that the LRT for Age does not include Age² in the model.

Covariate	P-value
Age - Numerical age of individual	0.7544
Age ² - Squared value of age	0.2018
Likely chlamydia exposure	<0.0001
Likely gonorrhea exposure	<0.0001
Likely other sexually transmitted disease (STD) exposure	0.9334
Rescreening for chlamydia	0.5623
Routine visit	0.3585
Rescreening for gonorrhea	0.9450
Multiple partners during the last 60 days	0.0342
New partner during the last 60 days	0.9201
Symptomatic partner during the last 90 days	0.0081
Concurrent partner during last 12 months	0.0886
Positive chlamydia result in the last 12 months	0.9910
Condom use during last encounter	0.7775
Sexual relations with males only, females only, or both	0.0032
Presumptive chlamydia treatment given	<0.0001
Urethritis clinical observation	0.0849
Epididymitis clinical observation	0.4006

Web Table 2: Covariates and LRT p-values for female urine specimens in the Iowa 2013 data.

Covariate	P-value
Age - Numerical age of individual	0.0001
Visit reason (family planning, STD screen, prenatal, rescreen, intrauterine device insertion)	0.5460
New partner during the last 90 days	0.7565
Multiple partners during the last 90 days	0.6296
Likely contact with a sexually transmitted disease	<0.0001
Symptoms (yes or no as reported by patient)	0.0413
Cervical friability clinical observation	0.5097
Cervicitis clinical observation	0.4582
Pelvic inflammatory disease clinical observation	0.8204
Urethritis clinical observation	0.7953

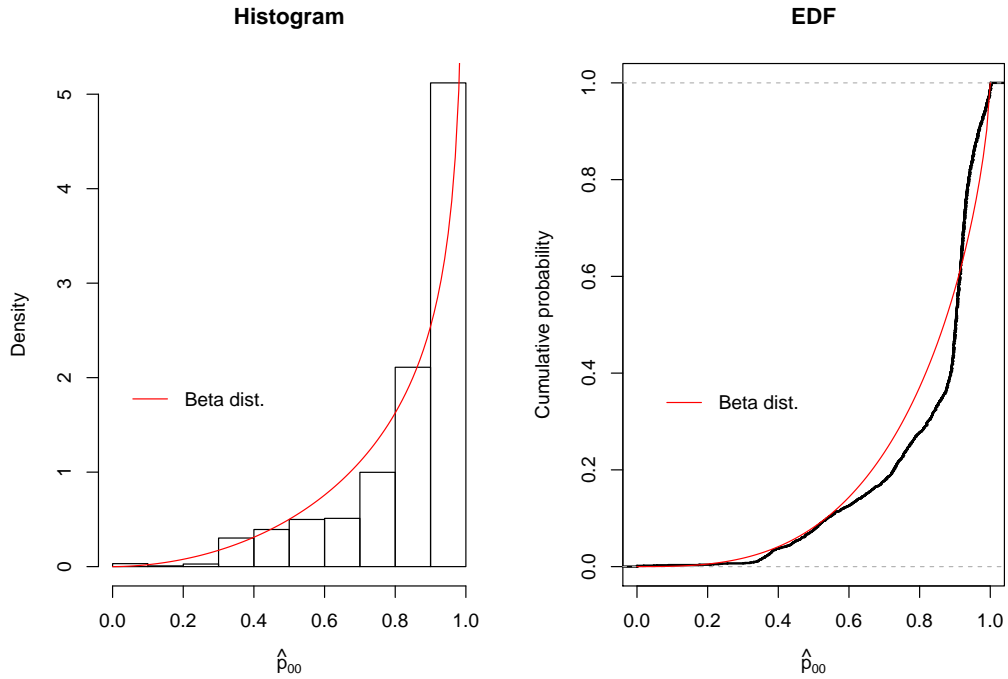
distribution as an overall summary measure. Also, the table includes estimates of the marginal probabilities averaged over all individuals and the variance of these marginal probabilities. Finally, the table includes an estimate of the psuedo odds ratio defined in Section 4 of the paper. This estimate is obtained by using $\widehat{OR} = \tilde{p}_{00}\tilde{p}_{11} / (\tilde{p}_{01}\tilde{p}_{10})$, where $\tilde{p}_{\tilde{y}_1\tilde{y}_2}$ is the mean $\hat{p}_{\tilde{y}_1\tilde{y}_2}$ value over all individuals in the corresponding test data.

Web Table 3 shows that males have much larger variability among marginal probabilities than females. This variability coincides with the group testing algorithm performance—larger (smaller) variability corresponds to lower (higher) number of tests. Also, this result is similar to what was found in Section 4 of the paper. To further understand this variability, Web Figures 10 and 11 show histograms and empirical distribution functions (EDFs) of \hat{p}_{00} for Idaho males and Iowa females. These types of plots are extremely important because the informative algorithms order individuals by $1 - \hat{p}_{00}$ (estimated probability of having at least one disease). The plots again show that the Idaho males have much more variability than the Iowa females, which leads to the performance differences for the informative and non-informative algorithms.

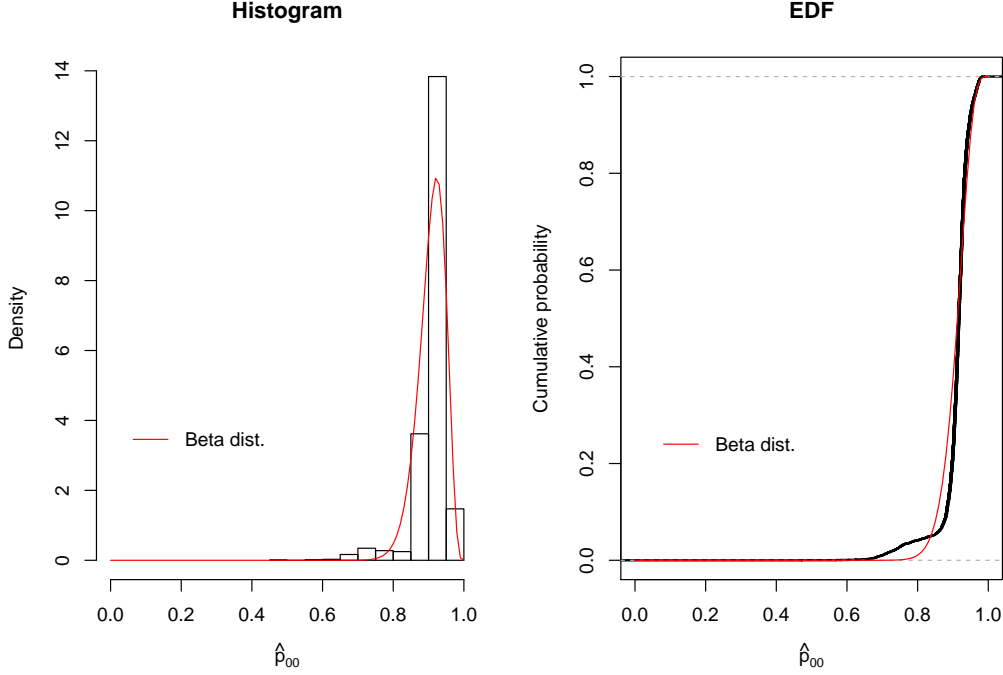
Web Table 3: Summaries of $\hat{p}_{\tilde{y}_1\tilde{y}_2}$ for the Idaho and Oregon 2011 data and the Iowa 2014 data.

State	Gender	Dirichlet MLEs				Mean		Variance		\widehat{OR}
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$	\hat{p}_{1+}	\hat{p}_{+1}	\hat{p}_{1+}	\hat{p}_{+1}	
Idaho	Female	27.48	0.21	2.00	0.20	0.0709	0.0042	0.0026	0.0009	3.9533
	Male	6.05	0.19	1.10	0.15	0.1638	0.0275	0.0110	0.0041	2.7010
Iowa	Female	68.38	0.41	6.26	0.29	0.0879	0.0019	0.0069	0.0001	8.6840
	Male	10.99	0.18	2.04	0.31	0.1670	0.0152	0.0233	0.0012	3.6557
Oregon	Female	33.70	0.16	2.72	0.13	0.0772	0.0045	0.0038	0.0002	8.7945
	Male	5.44	0.19	1.10	0.16	0.1801	0.0376	0.0188	0.0039	2.7358

Web Figure 10: Histogram and EDF of \hat{p}_{00} for male urine specimens in the Idaho 2011 data. A beta distribution fit to these estimated joint probabilities is also included on the plots.



Web Figure 11: Histogram and EDF of \hat{p}_{00} for female urine specimens in the Iowa 2014 data. A beta distribution fit to these estimated joint probabilities is also included on the plots.



F.2 Optimal testing configurations

Based on the recommendation of [Black et al. \(2015\)](#) for single-disease assays, we apply a “non-adaptive” process to choose one testing configuration that can be implemented on all test data individuals for a particular state and gender combination. This process works by first obtaining the estimated probability of having a disease $\hat{p}_{\tilde{y}_1 \tilde{y}_2}$ for individuals in the training data through a multinomial regression model as described in Section F.1. These individuals are arranged into initial groups of size I by the order of the date that they were actually tested within the laboratory. We denote the total number of these initial groups by H . We will discuss shortly how to account for situations when the total number of individuals is not evenly divisible by I . Individuals in these H initial groups are subsequently ordered within a group by their estimated probability of testing positive for at least one disease ($1 - \hat{p}_{00}$).

Each of these ordered individuals has four probabilities associated with them: $\hat{p}_{(i)h00}$, $\hat{p}_{(i)h10}$, $\hat{p}_{(i)h01}$, and $\hat{p}_{(i)h11}$, where we use the subscript “ $(i)h$ ” now to denote the i th ordered individual within the h th group for $h = 1, \dots, H$. These probabilities are averaged over the H initial groups to form $\tilde{p}_{(i)00}$, $\tilde{p}_{(i)10}$, $\tilde{p}_{(i)01}$, and $\tilde{p}_{(i)11}$ for $i = 1, \dots, I$. Next, we find an estimate of $E(T|\mathbf{M})$ by substituting these average values for the true probabilities in the expected value. Additional tests are added to $E(T|\mathbf{M})$ to account for individual testing of those individuals not in a group of size I (the not evenly divisible situation described in the previous paragraph). An OTC is found then as described in Section 2.2 of the paper using complete enumeration over possible \mathbf{M} . We limit the OTC to $I \leq 20$ for three-stage group testing because larger group sizes are not used in practice. We also limit the OTC to $I \leq 30$ for two-stage group testing for practicality considerations, while continuing to limit any group to be of size 20 individuals or less ($I_{1j} \leq 20$).

Web Tables 4 and 5 provide the group membership matrices obtained from the non-adaptive process. These corresponding testing configurations are used with the test data. Note that only two- and three-stage applications are considered for logistical reasons at the laboratories. To help understand the contents in these tables, examine the entries for Idaho females tested over two stages (a three-stage example is given in Section 5). Informative group testing places the 4,168 individuals successively into non-overlapping sets of size $I = 30$ individuals. Groups of size $I_{11} = 7$, $I_{12} = 6$, $I_{13} = 5$, $I_{14} = 5$, $I_{15} = 4$, and $I_{16} = 3$ are formed

upon individuals in each set for stage 1 testing. These individuals are ordered by their estimated probability of being positive for at least one disease. Thus, the individuals with the smallest probabilities are in the first group, and the individuals with the largest probabilities are in the last group. For any group that tests positively, individual testing is performed upon its group members in stage 2. Non-informative group testing proceeds in a similar manner, but always using groups of size 5 for stage 1 testing.

In summary, the group membership matrices are almost always the same across the states for non-informative group testing within each number of stages and gender combination. The informative group testing algorithms often have different group membership matrices across these combinations. The informative group testing algorithms also often do not have some individuals tested in the last possible stage due to individual testing occurring in the previous stage.

F.3 Assay imperfection

The chlamydia and gonorrhea responses in the retrospective data sets are the diagnostic outcomes from the Aptima Combo 2 Assay and its application by laboratories. Information on individual follow-up to confirm a positive/negative response is not available to a laboratory, even if this follow-up occurred. This unavailability is for a variety of reasons, including: antibiotics prescribed prior to diagnosis could skew the confirmation, the asymptomatic nature of the diseases, an individual failing to seek further medical attention, and a negative testing individual who is assumed to be truly negative.

The statistical evaluation of any new testing algorithm can be challenging due to these true disease statuses being unknown. The ideal process would be to apply each testing algorithm on the same specimens and include confirmatory testing. Unfortunately, actual evaluation in this format is generally not realistic for a number of reasons, where we discuss two here. First, implementing multiple testing algorithms is cost prohibitive. For example, the SHL estimates the cost per test to be \$37 and thousands of tests are performed at the laboratory each year. Second, the volume per specimen is insufficient to allow for multiple testing algorithms along with confirmatory testing. For example, at the SHL, individual specimens are approximately 2.9mL in volume and 700 μ L from each specimen is used per group.

Due to these challenges, we apply our new testing algorithms using a process similar to what was implemented by other papers on informative group testing (e.g., Bilder et al., 2010, McMahan et al., 2012; and Black et al., 2015). For each state and gender combination, we obtain testing responses as follows:

1. For an initial group of size I as found in Section F.2, determine the positive/negative *true* statuses of the group by using the actual observed chlamydia ($k = 1$) and gonorrhea ($k = 2$) responses in the test data. For example, if at least one individual is positive for a disease in the group, the group response is positive for that disease.
2. Obtain chlamydia and gonorrhea testing responses for the group as follows:
 - (a) For a true positive disease status, simulate the group outcome by generating a binary response from a Bernoulli($S_{e:11k}$) distribution.
 - (b) For a true negative disease status, simulate the group outcome by generating a binary response from a Bernoulli($1 - S_{p:11k}$) distribution.

Aggregate responses for chlamydia and gonorrhea together to form $\mathbf{g}_{11} = (g_{111}, g_{112})$.

3. If there is at least one positive response in \mathbf{g}_{11} , proceed to stage 2 of the emulated group testing algorithm in a similar manner as in 2., but now for each subgroup. Otherwise, the test performed in 2. is the only test performed and each individual within the group is declared to be negative for chlamydia and gonorrhea.
4. Stage 3 and higher stage testing continues as determined by the testing configuration whenever a subgroup tests positive.

Web Table 4: Group membership matrices for informative group testing. An “NA” cell entry denotes that no further testing is performed for a particular individual due to a group of size 1 (individual testing) occurring for this individual in the previous stage.

[illegible]

Web Table 5: Group membership matrices for non-informative group testing.

State	Stages	Gender	M
Idaho	2	Female	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}$
		Male	$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}$
	3	Female	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{bmatrix}$
		Male	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 & 4 & 4 & 4 & 5 & 5 & 5 & 6 & 6 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 \end{bmatrix}$
Iowa	2	Female	$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix}$
		Male	$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}$
	3	Female	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{bmatrix}$
		Male	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 & 4 & 4 & 4 & 5 & 5 & 5 & 6 & 6 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 \end{bmatrix}$
Oregon	2	Female	$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix}$
		Male	$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}$
	3	Female	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{bmatrix}$
		Male	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 & 4 & 4 & 4 & 5 & 5 & 5 & 6 & 6 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 \end{bmatrix}$

Web Table 6: Sensitivity and specificity levels for the Aptima Combo 2 Assay (available from pages 19 and 22 at <https://www.fda.gov/downloads/BiologicsBloodVaccines/SafetyAvailability/TissueSafety/ucm100275.pdf>). The levels used in the paper for the Aptima Combo 2 Assay application are the same for each stage ($s = 1, \dots, S$) and group ($j = 1, \dots, c_s$).

Gender	Chlamydia		Gonorrhea	
	$S_{e:sj1}$	$S_{p:sj1}$	$S_{e:sj2}$	$S_{p:sj2}$
Female	0.947	0.989	0.913	0.993
Male	0.979	0.985	0.985	0.996

Individuals are declared to be positive for a disease only when an individual test outcome is positive. Individuals are declared to be negative otherwise. Steps 1. - 4. are repeated across I successive members of the test data. When the total number of individuals is not evenly divisible by I , individual testing is performed on the remaining individuals.

As indicated above, our application assumes that the observed disease outcomes in the test data are the true disease statuses. This was done to obtain an assessment of accuracy for the group testing algorithms. To obtain estimates of the joint probabilities $\hat{p}_{\tilde{y}_1\tilde{y}_2}$ for the test data, we then also need to assume the observed disease outcomes in the training data are the true disease statuses, as first mentioned in Section F.1. This allows us to estimate a regression model with the training data that can be used directly with the test data for this type of application.

The values used for $S_{e:sjk}$ and $S_{p:sjk}$ are provided in Web Table 6. We assume these values are the same for every stage ($s = 1, \dots, S$) and every group ($j = 1, \dots, c_s$), which is reasonable for properly calibrated assays.

F.4 Observed accuracy

Web Table 7 provides the observed accuracy for the informative and non-informative group testing algorithms. The columns are abbreviated with PSe for pooling sensitivity, PSp for pooling specificity, PPPV for pooling positive predictive value, and PNPV for pooling negative predictive value. The accuracy measures are calculated for each of the 500 group testing applications to a data set. We report in the table the mean over these applications.

Web Table 7: Mean accuracy values of the group testing algorithms for the Idaho and Oregon 2011 data and the Iowa 2014 data.

State	Gender	Stages	Method	Chlamydia				Gonorrhea			
				PSe	PSp	PPPV	PNPV	PSe	PSp	PPPV	PNPV
Idaho	Female	2	Informative	0.9009	0.9972	0.9658	0.9914	0.8753	0.9979	0.5563	0.9996
			Non-informative	0.8975	0.9969	0.9622	0.9912	0.8865	0.9977	0.5354	0.9997
		3	Informative	0.8640	0.9982	0.9766	0.9883	0.8680	0.9984	0.6281	0.9996
			Non-informative	0.8522	0.9984	0.9784	0.9873	0.8508	0.9985	0.6369	0.9996
	Male	2	Informative	0.9693	0.9947	0.9764	0.9930	0.9817	0.9981	0.8291	0.9998
			Non-informative	0.9590	0.9949	0.9774	0.9907	0.9793	0.9981	0.8332	0.9998
		3	Informative	0.9523	0.9948	0.9766	0.9892	0.9793	0.9981	0.8322	0.9998
			Non-informative	0.9422	0.9950	0.9773	0.9870	0.9813	0.9982	0.8412	0.9998
	Female	2	Informative	0.9014	0.9971	0.9675	0.9907	0.8820	0.9978	0.8184	0.9987
			Non-informative	0.9003	0.9971	0.9678	0.9906	0.8816	0.9978	0.8196	0.9987
Iowa	Female	3	Informative	0.8678	0.9978	0.9745	0.9876	0.8618	0.9982	0.8412	0.9985
			Non-informative	0.8602	0.9981	0.9768	0.9869	0.8634	0.9984	0.8606	0.9985
	Male	2	Informative	0.9663	0.9941	0.9697	0.9934	0.9768	0.9943	0.8532	0.9992
			Non-informative	0.9615	0.9948	0.9733	0.9925	0.9727	0.9935	0.8344	0.9991
	Female	3	Informative	0.9564	0.9954	0.9762	0.9915	0.9772	0.9984	0.9548	0.9992
			Non-informative	0.9508	0.9948	0.9727	0.9904	0.9779	0.9982	0.9490	0.9993
		2	Informative	0.8988	0.9975	0.9682	0.9915	0.8790	0.9980	0.6451	0.9995
			Non-informative	0.8981	0.9976	0.9692	0.9914	0.8788	0.9981	0.6533	0.9995
	Male	3	Informative	0.8643	0.9981	0.9751	0.9886	0.8533	0.9984	0.6837	0.9994
			Non-informative	0.8541	0.9985	0.9794	0.9878	0.8515	0.9986	0.7165	0.9994
Oregon	Male	2	Informative	0.9676	0.9940	0.9740	0.9925	0.9830	0.9980	0.9222	0.9996
			Non-informative	0.9600	0.9948	0.9772	0.9907	0.9773	0.9982	0.9290	0.9995
	Female	3	Informative	0.9531	0.9941	0.9742	0.9891	0.9814	0.9981	0.9257	0.9996
			Non-informative	0.9454	0.9949	0.9775	0.9874	0.9763	0.9982	0.9294	0.9994
		2	Informative								
			Non-informative								

References

- American Red Cross (2018). Blood testing. <https://www.redcrossblood.org/biomedical-services/blood-diagnostic-testing/blood-testing.html>. Retrieved September 20, 2018.
- Bilder, C. and Loughin, T. (2014). *Analysis of Categorical Data with R*. CRC Press.
- Bilder, C., Tebbs, J., and Chen, P. (2010). Informative retesting. *Journal of the American Statistical Association*, 105:942–955.
- Black, M., Bilder, C., and Tebbs, J. (2015). Optimal retesting configurations for hierarchical group testing. *Journal of the Royal Statistical Society: Series C*, 64:693–710.
- Hitt, B., Bilder, C., Tebbs, J., and McMahan, C. (2018). The optimal group size controversy for group testing: Much ado about nothing?! Technical report. Department of Statistics, University of Nebraska-Lincoln, <http://www.chrisbilder.com/grouptesting>. Retrieved September 20, 2018.
- Hou, P., Tebbs, J., Bilder, C., and McMahan, C. (2017). Hierarchical group testing for multiple infections. *Biometrics*, 73:656–665.
- Hudgens, M. (2016). Rejoinder to “A note on the evaluation of group testing algorithms in the presence of misclassification”. *Biometrics*, 72:304.
- Kim, H., Hudgens, M., Dreyfuss, J., Westreich, D., and Pilcher, C. (2007). Comparison of group testing algorithms for case identification in the presence of test error. *Biometrics*, 63:1152–1163.
- Litvak, E., Tu, X., and Pagano, M. (1994). Screening for the presence of a disease by pooling sera samples. *Journal of the American Statistical Association*, 89:424–434.
- Malinovsky, Y., Albert, P., and Roy, A. (2016). Reader reaction: A note on the evaluation of group testing algorithms in the presence of misclassification. *Biometrics*, 72:299–302.
- McMahan, C., Tebbs, J., and Bilder, C. (2012). Informative Dorfman screening. *Biometrics*, 68:287–296.
- McMahan, C., Tebbs, J., and Bilder, C. (2016). Rejoinder to “A note on the evaluation of group testing algorithms in the presence of misclassification”. *Biometrics*, 72:303–304.
- McMahan, C., Tebbs, J., Hanson, T., and Bilder, C. (2017). Bayesian regression for group testing data. *Biometrics*, 73:1443–1452.