OPERATING CHARACTERISTICS OF GROUP TESTING ALGORITHMS FOR CASE IDENTIFICATION IN THE PRESENCE OF TEST ERROR

by
Hae-Young Kim

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Public Health in the Department of Biostatistics, School of Public Health.

Chapel Hill
2007

Approved by:

Dr. Michael G. Hudgens, Advisor
Dr. Joseph G. Ibrahim, Co-Advisor
Dr. Vidyadhar G. Kulkarni, Reader
Dr. Chirstopher D. Pilcher, Reader
Dr. Paul W. Stewart, Reader

ABSTRACT
HAE-YOUNG KIM: OPERATING CHARACTERISTICS OF GROUP TESTING ALGORITHMS FOR CASE IDENTIFICATION IN THE PRESENCE OF TEST ERROR.
(Under the direction of Dr. Michael G. Hudgens.)

Pooling of specimens to increase efficiency of screening individuals for rare diseases has a long history, dating back to screening for syphilis in military inductees in the 1940s. Subsequently, specimen pooling has been applied to screening for many other infectious diseases and has also found broader application in entomology, screening for genetic mutations and many other areas. Currently the North Carolina Department of Public Health and investigators from the University of North Carolina at Chapel Hill have developed quick, cost effective methods to screen over 120,000 people annually for recent HIV infection using highly sensitive, automated specimen pooling algorithms as part of the Screening and Tracing Active Transmission (STAT) program. In this dissertation, we research group testing methodology to help optimize the pooling algorithm used in the STAT program and to assist in extending this innovative approach to other settings or detection of other infectious diseases where the overriding issues are identical but the specific conditions (e.g., disease prevalence) vary considerably.

This dissertation is comprised of three papers. In the first paper, we derive and compare the operating characteristics of hierarchical and two-dimensional array-based testing algorithms for case identification in the presence of testing error. The operating characteristics investigated include efficiency (i.e., expected number of tests per specimen) and error rates (i.e., sensitivity, specificity, positive and negative predictive values, per-family error rate, and per-comparison error rate). In the second paper, we extend two-dimensional array to three-dimensional array-based algorithms when there exist test errors. Efficiency and pooling measurement error rates of three-dimensional

array-based algorithms are compared with hierarchical and two-dimensional array-based algorithms in the presence of test error. In both the first and second papers, the methodology is illustrated by comparing different pooling algorithms for the detection of individuals recently infected with HIV in North Carolina and Malawi. In the third paper, the optimal configuration of a two-dimensional array-based pooling algorithm is considered. We derive $p*$, the highest prevalence, where pooling with this algorithm is better than individual testing. For the given prevalence less than $p*$, we determine the optimal algorithm configuration which minimizes the expected number of tests per specimen.

ACKNOWLEDGMENTS

I would like to gratefully and sincerely thank Dr. Michael Hudgens, my advisor, for his insightful guidance, advice, encouragement, and support at all stages of this dissertation. I have been amazingly fortunate to have an advisor who gave me the freedom to explore on my own, and at the same time the guidance to recover when my steps faltered. His patience and support helped me overcome many difficult situations and finish the dissertation. I owe many thanks to Dr. Joseph Ibrahim, my co-advisor, for kind guidance and enormously helpful discussions. I would also like to thank the other members of my committee, Dr. Vidyadhar Kulkarni, Dr. Chirstopher Pilcher, and Dr. Paul Stewart, for their willingness to serve in those roles. Their helpful comments have improved the quality of this work.

Many others have contributed to my professional growth. Among these are my colleagues in the Centers for Disease Control and GlaxoSmithKline. I have enjoyed the many rewarding opportunities of working with them along the way.

Finally, I would like to express my heart-felt gratitude to my family. None of this would have been possible without the love and patience of my family. Especially, I warmly appreciate the support and understanding of my parents throughout the dissertation process.

TABLE OF CONTENTS

LIST OF FIGURES ix

LIST OF TABLES xi

1. INTRODUCTION 1

2. BACKGROUND 4

2.1. Fundamentals 4

2.2. Motivating Example: Acute HIV Detection 4

2.3. Definitions and Operating Characteristics 6

2.4. Dorfman Algorithm 7

2.5. Hierarchical Algorithms 9

2.6. Square Array Algorithms 11

2.7. Multidimensional Array Algorithms 12

2.8. Error Rates 14

2.9. Optimal Pool Size 16

3. HIERARCHICAL AND SQUARE ARRAY ALGORITHMS 17

3.1. Introduction 17

3.2. Assumptions 18

3.3. Hierarchical Algorithm 19

3.3.1. Efficiency 20

3.3.2. Error Rates 21
3.4. Square Array Without Master Pool Testing 22
3.4.1. Efficiency 22
3.4.2. Error Rates 24
3.5. Square Array With Master Pool Testing 26
3.5.1. Efficiency 26
3.5.2. Error Rates 28
3.6. Assessing Variability 28
3.7. Application 30
3.8. Discussion 33
4. THREE-DIMENSIONAL ARRAY ALGORITHMS 40
4.1. Introduction 40
4.2. Preliminaries 41
4.2.1. Notation 41
4.2.2. Assumptions 41
4.3. Three-Dimensional Array Without Master Pool 41
4.3.1. Efficiency 42
4.3.2. Error Rates 45
4.4. Three-Dimensional Array With Master Pool 48
4.4.1. Efficiency 49
4.4.2. Error Rates 50
4.5. *A3PM*$_2$: An alternative to *A3PM* 51
4.6. Simulation Results 52

4.7. Application 53

4.8. Discussion 55

5. OPTIMAL CONFIGURATION OF A SQUARE ARRAY GROUP TESTING ALGORITHM 62

5.1. Introduction 62

5.2. Hierarchical Algorithms 63

5.2.1. Turner et al (1988) 63

5.2.2. Samuels (1978) 64

5.2.3. Wu and Zhao (1994) 64

5.3. Square Array-Based Algorithm 65

5.3.1. Lower and Upper bounds for q^* 65

5.3.2. Determining q^* 68

5.3.3. Determining n^* 73

5.4. Discussion 77

APPENDIX A. Per-Family Error Rates 84

APPENDIX B. Efficiency Variance 85

B.1. Dorfman algorithm (*D2*) 85

B.2. Three-stage hierarchical algorithm (*D3*) 85

B.3. Square array without master pool (*A2*) 86

B.4. Square array with master pool (*A2M*) 89

APPENDIX C. CDFs for Pooling Error Rates 91

C.1. Pooling Sensitivity and Specificity 91

C.2. Pooling *PPV* and *NPV* 92

REFERENCES 94

LIST OF FIGURES

3.1 (a) Expected number of tests per specimen, (b) pooling specificity, (c) pooling sensitivity, (d) pooling PPV, (e) pooling $PFER$, and (f) pooling $PFER_2$ for different algorithms assuming test sensitivity $S_e = 0.9$, test specificity $S_p = 0.99$, and prevalence $p = 0.0002$. The ▲ denotes the three stage hierarchical pooling algorithm employed in the NC STAT Program. Note pooling specificity for individual testing equals $S_p = 0.99$ and is not shown in panel (b). 36

3.2 (a) Expected number of tests per specimen, (b) pooling specificity, (c) pooling sensitivity, (d) pooling PPV, (e) pooling $PFER$, and (f) pooling $PFER_2$ for different algorithms assuming test sensitivity $S_e = 0.9$, test specificity $S_p = 0.99$, and prevalence $p = 0.045$. The ▲ denotes the three stage hierarchical pooling algorithm employed in Malawi. 37

3.3 (a) Expected number of tests per specimen, (b) pooling specificity, (c) pooling sensitivity, and (d) pooling PPV for optimally efficient configurations of $D3$ and $A2M$ assuming test sensitivity $S_e = 0.9$, test specificity $S_p = 0.99$, and a maximum allowable pool size of 100. 38

3.4 Contour plots of (a) expected number of tests per specimen, (b) pooling sensitivity, and (c) pooling PPV for $D3(16:4:1)$ assuming $p = 0.045$ as a function of test sensitivity (S_e) and test specificity (S_p). The • denotes the values of S_e and S_p assumed in Table 3.1. 39

4.1 Three dimensional planar algorithm ($A3P$) with $L = M = N = 3$. The black dots denote 9 ($= MN$) specimens in one of L planar slices. 59

4.2 Results of simulation study. (a) Expected number of tests per specimen, (b) pooling specificity, (c) pooling sensitivity for $A3P$ assuming $L = M = N = 3$, test sensitivity $S_e = 0.88$, test specificity $S_p = 0.92$, and prevalence $p \in \{0.0001, 0.001, 0.01, 0.1\}$. 60

4.3 (a) Expected number of tests per specimen, (b) optimal master pool size, (c) pooling sensitivity, (d) pooling specificity, (e) pooling PPV and (f) pooling NPV for optimally efficient configurations of $D3$, $A2M$, and $A3PM$ assuming test sensitivity $S_e = 0.9$, test specificity $S_p = 0.9$, and a maximum allowable pool size of 100. 61

5.1 $g(q, n)$ vs. n of $A2$ for $q = (1 - \sqrt{\frac{1}{2}})^{\frac{1}{4}}$, $q = 0.7502$, and $q = (\frac{1-\sqrt{\frac{1}{3}}}{2})^{\frac{1}{6}}$. . 78

5.2 $\Delta(q, n)$ vs. q of $A2$ for $n = 2, 3, 4, 5, 6$. 79

5.3 $\Delta(q, n)$ vs. q of $A2$ for $n = 5, 6, 7, 8, 9$. 80

LIST OF TABLES

3.1 Comparison of operating characteristics for individual testing and four potential pooling algorithms to be used in Malawi for detection of acute HIV. 35

4.1 Description of group testing algorithms. 56

4.2 Comparison of operating characteristics for the most efficient $D3$, $A2M$, $A3PM$, and $A3PM_2$ to be used in NC STAT for detection of acute HIV assuming pool size less than 100, test sensitivity $S_e = 0.9$, test specificity $S_p = 0.99$, and prevalence $p = 0.0002$. 57

4.3 Comparison of operating characteristics for the most efficient $D3$, $A2M$, $A3PM$, and $A3PM_2$ to be used in Malawi for detection of acute HIV assuming pool size less than 50, test sensitivity $S_e = 0.9$, test specificity $S_p = 0.99$, and prevalence $p = 0.045$. 58

5.1 The behavior of $f(q, n)$ and optimal pool size $n*$ for a given q 81

5.2 Optimal array size $n*$ of $A2$ for given prevalence p. $n*$ is bounded between $\lfloor \mu(q) \rfloor$ and $\lceil \rho(q) \rceil$ according to Theorems 2 and 3. 82

5.3 The optimal array size $n*$ of $A2$ for given prevalence p 83

CHAPTER 1

INTRODUCTION

Pooling of specimens to increase efficiency of screening individuals for rare diseases has a long history, dating back to screening for syphilis in military inductees in the 1940s (Dorfman, 1943). Subsequently, specimen pooling or *group testing* has been applied to screening for many other infectious diseases (Quinn et al., 2000; Pilcher et al., 2002; Pilcher et al., 2004; Mine et al., 2003; Kacena et al., 1998; Centers for Disease Control and Prevention, 2003) and has also found broader application in entomology (Venette et al., 2002), screening for genetic mutations (Gastwirth, 2000), the blood bank and pharmaceutical industries (Xie et al., 2001; Jones and Zhigljavsky, 2001), analytical chemistry (Woodbury et al., 1995), information theory (Wolf, 1985) and many other areas.

In the context of infectious diseases, group testing is typically used for (i) case identification, i.e., detecting all individuals having the disease of interest, and (ii) prevalence estimation, i.e., estimating the proportion of individuals in the population having a particular disease. This thesis is primarily motivated by examples of the former. For example, currently the North Carolina Department of Public Health and investigators from the University of North Carolina (UNC) at Chapel Hill have developed quick, cost effective methods to screen over 120,000 people annually for recent HIV infection using highly sensitive, automated specimen pooling algorithms as part of the Screening and

Tracing Active Transmission (STAT) program (Pilcher et al., 2002; Pilcher et al., 2005). Likewise, the Center for HIV/AIDS Vaccine Immunology plans to employ similar testing procedures as part of a global attempt to identify acute infections (NIAID Office of Communications and Public Liaison, 2005). A similar specimen pooling strategy have also been used to identify acute HIV in antibody negative males attending STD clinics in Malawi.

In these applications, the problem is how to detect very rare cases of HIV infection that elude detection by routine, standard antibody testing assays (Pilcher et al., 2005) because they are in the pre-antibody "acute" phase of infection. The PCR-based nucleic acid amplification assays that detect these persons are highly sensitive but (compared to antibody tests) are expensive, time consuming, and have inadequate specificity (Daar et al., 2001). In this case, group testing is used to enhance testing efficiency and accuracy of high throughput screening for rare cases of acute HIV.

Here we explore aspects of group testing methodology to help optimize the pooling algorithm used in the STAT program and to assist in extending this innovative approach to other settings (e.g. other US states, Africa) or detection of other infectious diseases (e.g., Hepatitis) where the overriding issues are identical but the specific conditions (e.g., disease prevalence) vary considerably.

The rest of this thesis is organized as follows. Chapter 2 comprises a review of a variety of topics related to group testing. These include acute HIV detection as a motivating example; operating characteristics; Dorfman, hierarchical, square array and multidimensional array algorithms; modeling of several pooling error rates; and optimal pool size determination. In Chapter 3, we derive the efficiencies and error rates of two dimensional square array pooling algorithms with and without master pool testing. Chapter 4 provides the efficiencies and error rates of three dimensional array pooling algorithms with and without master pool testing. We derive the optimal configuration

of a square-array algorithm in Chapter 5.

CHAPTER 2

BACKGROUND

2.1 Fundamentals

The pooling of individual serum samples as a cost saving method for diagnosis of infectious disease was first used for identifying individuals with syphilis. The basic idea behind group testing is that for rare diseases, efficiency can be gained by pooling specimens (e.g., urine, sera, or plasma) and testing these pools rather than individual specimens. Here, efficiency is defined in the sense of minimizing the expected number of tests required. A positive test suggests at least one of the specimens in the pool is positive, while a negative test suggests that all specimens in the pool are negative. If the group tests gives a positive result, individual sample is tested. If the group test has a negative result, no more tests are needed. Thus, time and cost can be saved with testing pooled samples rather than individual samples, especially when prevalence rate is very low.

2.2 Motivating Example: Acute HIV Detection

Group testing has been applied to screening for many other infectious diseases including HIV (Mine et al., 2003; Quinn et al., 2000; Pilcher et al., 2002; Pilcher et al., 2004;

Pilcher et al., 2005). The recent availability of highly sensitive nucleic acid amplification tests (NAATs) has created important opportunities for both surveillance and clinical testing of infectious diseases. The possible improvement that NAATs can provide for infectious disease diagnosis is exemplified by HIV. While health authorities have repeatedly emphasized that a goal of HIV testing and surveillance should be to increase detection of early infection, HIV testing has until very recently relied exclusively on antibody tests that miss all *acute* infections, i.e., infections occurring within approximately the last three months during which time individuals are antibody negative and believed to be highly contagious (Pilcher et al., 2004; Wawer et al., 2005). By detecting acute HIV infections and distinguishing them from older infections, NAATs allow for: early treatment of HIV, a strategy which might possibly ameliorate disease pathogenesis; interventions to prevent secondary transmission; better understanding of host-virus dynamics; and improved surveillance (Pilcher et al., 2005).

Unfortunately NAATs are usually considered insufficient for general use due to expense and unacceptably high false positive rates (Pilcher et al., 2002; Klausner, 2004). This lack of specificity results in extremely low positive predictive value in low prevalence settings such as acute HIV. However, NAATs in conjunction with specimen pooling can dramatically improve efficiency, specificity, and positive predictive value (Quinn et al., 2000; Pilcher et al., 2005). For instance, the NC STAT program has now successfully demonstrated that a simple group testing strategy renders NAAT screening much more accurate and cost efficient than testing of individual specimens for acute HIV (Pilcher et al., 2002; Pilcher et al., 2005). This specimen pooling strategy have also been used to identify acute HIV in antibody negative males attending STD clinics in Malawi (Pilcher et al., 2004). Pilcher et al. (2004) concluded that the specimen pooling algorithm has the advantage of being both highly cost efficient and highly specific when used in populations with low expected HIV prevalence.

2.3 Definitions and Operating Characteristics

In this section, we introduce several definitions and operating characteristics of testing algorithms for case identification in the presence of test error. First, in the group testing problem, *stage* refers to the number of sequential steps required by a particular algorithm to identify all positive specimens. In this thesis, we will consider only 2 and 3 stage algorithms.

Define *efficiency* of a particular pooling algorithm to be the expected number of tests per specimen required to identify all positive specimens. For a group testing algorithm $\mathcal{A}$, denote the efficiency by $E(\mathcal{A})$.

Define *pooling specificity* to be the probability an individual is categorized as negative by a particular pooling procedure given that individual is truly negative (Litvak et al. 1994; Johnson et al. 1991). Similarly, define the *pooling sensitivity* to be the probability an individual is categorized as positive by a particular pooling procedure given that individual is truly positive. Denote the pooling specificity and sensitivity for $\mathcal{A}$ by $S_p(\mathcal{A})$ and $S_e(\mathcal{A})$. For example, if $\mathcal{I}$ denotes individual testing, then $S_p(\mathcal{I}) = S_p$ and $S_e(\mathcal{I}) = S_e$.

Define the *pooling positive predictive value* of $\mathcal{A}$ to be the probability an individual is truly positive given they are categorized as positive by $\mathcal{A}$. Likewise, define the *pooling negative predictive value* of $\mathcal{A}$ to be the probability an individual is truly negative given they are categorized as negative by $\mathcal{A}$. Denote the pooling positive and negative predictive values of $\mathcal{A}$ by $PPV(\mathcal{A})$ and $NPV(\mathcal{A})$. The predictive values are simple functions of $S_e(\mathcal{A})$ and $S_p(\mathcal{A})$:

$$PPV(\mathcal{A}) = \frac{pS_e(\mathcal{A})}{q[1 - S_p(\mathcal{A})] + pS_e(\mathcal{A})},$$

and

$$NPV(\mathcal{A}) = \frac{qS_p(\mathcal{A})}{p[1 - S_e(\mathcal{A})] + qS_p(\mathcal{A})}.$$

Additionally, we consider other error rates that have been proposed in the multiple comparisons literature (Cook and Dunnett, 1998), but to our knowledge have not been considered in the context of group testing. These quantities provide alternative metrics for quantifying the degree of misclassification of negative individuals as positive and positive individuals as negative.

Define the *per-family error rate* ($PFER$) to be the expected number of false positive classifications and the *per-comparison error rate* ($PCER$) to be the expected number of false positive classifications divided by the total number of specimens (Hochberg and Tamhane, 1987). In Appendix A, we show that $PFER(\mathcal{A}) = nq\{1 - S_p(\mathcal{A})\}$ for any pooling algorithm $\mathcal{A}$, which immediately gives $PCER(\mathcal{A})$ since $PFER(\mathcal{A}) = nPCER(\mathcal{A})$ where n is the number of specimen to be tested.

Define the *type II per-family error rate* ($PFER_2$) to be the expected number of false negative classifications and the *type II per-comparison error rate* ($PCER_2$) to be the expected number of false negative classifications divided by the total number of specimens. In Appendix A, we also show that $PFER_2(\mathcal{A}) = np\{1 - S_e(\mathcal{A})\}$ for any pooling algorithm $\mathcal{A}$, which gives $PCER_2(\mathcal{A})$ since $PFER_2(\mathcal{A}) = nPCER_2(\mathcal{A})$.

2.4 Dorfman Algorithm

Dorfman (1943) studied the application of a group testing procedure for screening men called up for induction into the army for presence of syphilitic antigen. Case identification (or *classification*) was the original motivation behind group testing, as proposed by Dorfman (1943). Dorfman's algorithm entailed pooling together biological specimens from several individuals and testing these pools of specimens rather than

testing each individual specimen. If a pool tested negative, all specimens in that pool were declared negative. Otherwise, further testing on individual specimen from the pool were employed to identify all positive individuals.

He derived the expected number of tests which is equal to one master pool plus the number of individuals in the master pool which require individual testing. Using notation from the Section 2.3, it is easy to see that the expected number of tests per specimen is

$$E\left(\frac{T}{n}\right) = \frac{1}{n} + 1 - q^n, \tag{2.1}$$

where T is the number of tests and n is the pool size. Dorfman's original pooling algorithm required simply testing all individual specimens within positive master pool without considering classification errors. This pooling procedure is appealing in that, for rare diseases, fewer number of tests are required on average to identify all cases compared to individual testing. His method was applied to blood tests for syphilis and found optimal group sizes, which minimize the expected number of tests per specimen, for various fixed values of p and concluded that his method is more efficient when compared to individual testing, for small values of p ranging from .01 to .15. Note that optimal size $n = \frac{1}{\sqrt{p}}$ of Dorfman's procedure was obtained by Feller (1957), Wilks (1962), Samuels (1978) and Turner et al. (1988).

Dorfman's method has been modified and analyzed extensively. Sterrett (1957) modified Dorfman's original procedure by proposing that samples from a positive group be tested individually only until the first positive sample is identified. Any remaining samples at that time are tested as a group and this procedure continues until all samples from the group that was first classified positive are exhausted. He used numerical analysis to show that his method is more efficient than Dorfman's by 6%, if the prevalence rate is 0.01. Sobel and Groll (1959) suggested that once a positive group is obtained, it might be more efficient to test a subgroup formed with samples from the original group,

rather than perform individual testing as both Dorfman and Sterrett procedures advocate. If the subgroup tests positive, the remaining samples from the original positive group are treated as if they were never tested. This is statistically accurate so long as the population of items being tested is binomial. If the subgroup tests negative, it is completely classified, and a new subgroup is drawn from the remaining members of the original group. Sobel and Groll (1959)'s procedure is complicated since optimal group (subgroup) sizes have to be determined at each stage of testing. Also, in its worst case, it can require a very large number of tests to classify the entire population. Hwang (1972) moved away from the assumption that the number of defectives is binomial and developed a procedure based on an upper bound for the number of defectives or its probability distribution. Johnson et al. (1991) generalized Dorfman's pooling algorithm in the presence of test error and extended to hierarchical algorithm. Lancaster and Keller-McNulty (1998) and Venette et al. (2002) reviewed sampling method and analyzed generalizations of Dorfman's pooling algorithm.

2.5 Hierarchical Algorithms

A simple extension of Dorfman's original procedure entails repeatedly dividing positive pools into smaller, non-overlapping subpools until eventually all positive specimens are individually tested. We refer to this approach as a *hierarchical* group testing algorithm (Finucan, 1964; Johnson et al., 1991; Litvak et al., 1994). For example, the NC STAT program (Pilcher et al., 2002) employs a three stage hierarchical algorithm as follows. First, disjoint master pools of 90 specimens are tested. Second, positive master pools are divided into subpools of 10 specimens each and these subpools are tested. Third, specimens from positive subpools are individually tested. Similar hierarchical pooling algorithms have been used for the detection of recent HIV infections in Malawi (Pilcher

et al., 2004) and India (Quinn et al., 2000).

Johnson et al. (1991) considered hierarchical group testing algorithms in the presence of test errors. They derived the expected number of tests, pooling sensitivity and pooling specificity for a hierarchical algorithm. They also allow sensitivity and specificity to be dependent on the number of pool size of each stage.

Litvak et al. (1994) generalized some commonly used pooling procedures. They presented group testing methods in which positive groups are split into several subgroups of almost equal size. Any subgroups that test positive are further split into even smaller subgroups and this process continues until items are tested individually. They called this method T_2 and extended it as ${T_2}^+$. ${T_2}^+$ is same as T_2 except that each group that produces a negative outcome is retested at most $r-1$ times where r is the maximum number of times a group will be retested before being classified or split further into smaller groups. If all $r-1$ tests are negative, the group is classified as negative ($r = 2$ in Litvak et al. (1994)). Otherwise, it is classified as positive. Litvak et al. (1994) compared several such procedures with each other and with Dorfman's procedure for different values of test reliability. They concluded that T_2 and ${T_2}^+$ are better than Dorfman's procedure. They also concluded that if the purpose of screening is to determine the infectious status of individuals and estimate prevalence, then T_2 is more competitive in this screening situation. However, if the goal is to screen donated blood, ${T_2}^+$ should be chosen, since this is the only procedure that reduces the $FNPV$.

Amos et al. (2000) and Gastwirth (2000) also re-derived efficiency in the presence of test error for the hierarchical group testing algorithm. They discussed pooling methods for identifying rare mutations and provided formulas for the optimal pool size as a function of the mutation frequency and the specificity.

2.6 Square Array Algorithms

Array based specimen pooling is an alternative to hierarchical group testing which uses overlapping pools. While the corresponding pooling algorithms are more complex, array based group testing can be very efficient. This approach is frequently employed in genetics (Bruno et al., 1995; Amemiya et al., 1992; Barillot et al., 1991; Zwaal et al., 1993), but remains largely underutilized in the infectious disease setting. In its simplest form, n^2 specimens are placed on an $n \times n$ matrix. Pools of size n are made from all samples in the same row or in the same column. These $2n$ pools are then tested and, assuming no false negative tests, all positive specimens will lie at the intersection of a positive row pool and a positive column pool. Any ambiguities are resolved by testing all specimens at these intersections.

Phatarfod and Sudbury (1994) evaluated the use of a two-dimensional array scheme. Their idea was based on the observation that blood samples are often placed in a square tray, and one could exploit this arrangement by compositing across rows and columns. Tests were conducted on the row and column composite samples, and confirmatory tests were conducted on sample units at the intersection of positive row and column. They called this scheme as $SA1$ and derived the expected number of tests per specimen:

$$E(SA1) \quad = \tfrac{2}{n} + 1 - 2q^n + q^{2n-1}. \tag{2.2}$$

A second scheme, $SA2$ which was proposed by Phatarfod and Sudbury (1994), proceeds as for $SA1$ expect that if all rows test negative no further tests are done (n tests in all), and if exactly one row tests positive then each individual in that row is tested ($2n$ tests in all). While $SA2$ requires one more stage than $SA1$, Phatarfod and Sudbury (1994) showed $SA2$ requires fewer tests than $SA1$ and the Dorfman scheme for p ranging from .000001 to .05. Phatarfod and Sudbury (1994) also considered rectangular $m \times n$ array

schemes, e.g., the classical 8×12 array and concluded that these schemes also require fewer tests than the Dorfman scheme.

Woodbury et al. (1995) noted that testing row and column pools, without further testing, results in ambiguities in assigning true sample locations. In order to resolve these ambiguities, they proposed pooling samples along diagonals of the array, wherein the diagonals are wrapped around the array so that each diagonal pool contains the same number of samples as each row and column pool.

Unfortunately, there are limitations in directly applying the array based procedures proposed in the literature to the infectious disease setting. For example, several proposed array pooling algorithms do not guarantee that all cases will be unambiguously identified (Woodbury et al., 1995; Bruno et al., 1995). For instance, diagonal pooling as suggested by Woodbury et al. (1995) will not resolve all ambiguities if all specimens in the first row and first column are positive. Perhaps most importantly, with the exception of Section 5 of Phatarfod and Sudbury (1994), the array based group testing literature does not consider test error, i.e., false positive and false negative test results.

2.7 Multidimensional Array Algorithms

Berger et al. (2000) extended Phatarfod and Sudbury (1994) to higher dimensional arrays. They proposed two natural three-dimensional extensions of the basic matrix method to an $L \times M \times N$ cubic array, the three-dimensional planar method and the three-dimensional linear method.

For the three-dimensional planar method (called $3P$), in stage 1, each of the L planar slices from front to back, the M planar slices from top to bottom, and the N planar slices from left to right should be tested. In stage 2, each specimen located at the intersections of three planes that all tested positive in stage 1 is tested individually.

The efficiency of the $3P$ method is the sum of four terms, namely, the reciprocals of LM, LN, MN and the quantity $(1 - q^{LM} - q^{LN} - q^{MN} + q^{L(M+N-1)} + q^{M(L+N-1)} + q^{N(L+M-1)} - q^{LM+LN+MN-L-M-N+1})$. For example, if $L = M = N$, then the expected number of tests per specimen of algorithm $3P$ equals

$$\tfrac{3}{N^2} + 1 - 3q^{N^2} + 3q^{N(2N-1)} - q^{3N^2-3N+1}. \tag{2.3}$$

For the three-dimensional linear method (called $3L$), in stage 1, for each of the MN sites on the front face of the cube, test the group of size L on the line perpendicular to this face through this site. Repeat for the LN groups of size M lying on similar lines perpendicular to the top and the bottom faces and for the LM groups of size N located on lines perpendicular to the left and right faces. In stage 2, test individually each specimen located at the intersections of three lines that all tested positive in stage 1. They showed the extension of $3L$ to multiple dimensions, denoted by ML, yields the family of maximally efficient positive testing procedures and the most efficient multidimensional parallelepiped is a generalized cube. Berger et al. (2000) proved that the efficiency of ML for a d- dimensional cube whose side have length s is

$$ds^{-1} + p + q(1 - q^{s-1})^d. \tag{2.4}$$

The optimal procedures proposed by Berger et al. (2000) are not necessarily clinically feasible due to the requirement of extremely large numbers of specimens. For example, the parameters $s = 74$ and $d = 6$ that maximize efficiency when $p = .01$ require $s^d = 74^6 \approx 164$ billion items to be in the population to be screened.

Schuster (2002) indicated there exists a theoretical connection between the work of Berger et al. (2000) and design theory. Design theory is the study of combinatorial design, which is an arrangement of the elements of a finite set into subsets or arrays

which satisfy certain regularity conditions. For example, thirty-six officers are given, belonging to six regiments and holding six ranks. One wants to know that the officers can be paraded in a 6×6 array so that, in any line (row or column) of the array, each regiment and each rank occurs exactly once. Design theory deals with this kind of problem of arranging objects according to certain rules. Schuster (2002) showed how design theory can be employed to determine optimal array algorithms which require much smaller numbers of specimen than reported in Berger et al. (2000).

Roederer and Koup (2003) used various multidimensional array pooling algorithms to identify T cell immune responses to vaccines or pathogens and used a Monte Carlo simulation to optimize the construction of peptide pools that could identify responses to individual peptide using the fewest numbers of assays and patient material. They found that the number of assays required to deconvolute a pool increases by the logarithm of the number of peptides within the pool. However, the optimum configuration of pools changes dramatically according to the number of responses to individual peptides that are expected to be in the sample.

2.8 Error Rates

A majority of the group testing literature assumes testing is done without error, i.e., assumes 100% sensitivity and specificity. For instance, Dorfman (1943) does not consider classification errors. Phatarfod and Sudbury (1994) developed two dimensional array methods without test errors except briefly in Section 5.

Some work has been done on group testing when there exists test error. Johnson et al. (1991) considered test errors such as false negative and false positive rates to generalize hierarchical group testing algorithms and developed pooling sensitivity and specificity of hierarchical group testing algorithms. They also allowed sensitivity and

specificity to depend on pool size.

Litvak et al. (1994) focused on HIV antibody tests and considered sensitivity and specificity as inherent characteristics of the test kit that influence how accurately an individual sample can be classified. They assumed the underlying test kits all have the same sensitivity and the same specificity at each step of the testing and for all samples. High incidence of false positives is commonly reported in HIV antibody screening of blood samples. Litvak et al. (1994) calculated that when prevalence rate is 4 per 10,000 people and a test kit having a sensitivity and a specificity of 98% is used, the false positive predictive value of test can be as high as 0.98.

Wein and Zenios (1996) attempted to capture the dilution effects. A *dilution effects* is defined as the failure of the test to detect an infected sample when it is mixed with a large number of negative samples. They developed a generalized linear model which connects optical density or OD levels (the attribute that is observed and measured in an ELISA test) with the antibody concentration in the pooled serum and later used it in a dynamic programming algorithm to produce a testing strategy which minimizes the total expected cost. Total cost is the sum of testing cost and the cost of incorrect classifications. It should be noted that, unlike others, these authors allowed three possibilities at each stage of testing: declare the group negative; require further testing, and declare all members of the group positive.

Gupta and Malina (1999) modified Dorfman and Sterrett's group testing protocols in the presence of classification errors. They begin with a highly sensitive test to achieve virtually 100% sensitivity. Specimens are tested in groups and individually as prescribed by the modified Dorfman and Sterrett methods. They controlled the incidence of false positives by repeating tests of grouped and individual samples that are initially reactive. The number of retests is chosen carefully so as to bring the overall incidence of false positives below some desired level, usually set close to 0. They

also compared the modified Dorfman and the modified Sterrett procedures with their modified individual testing and showed that the modified Dorfman and the modified Sterrett are more efficient than the modified individual testing in the presence of test errors and the modified Sterrett is slightly efficient than the modified Dorfman, but harder to implement.

2.9 Optimal Pool Size

Feller (1957) and Wilks (1962) have presented text-book exercises which showed the optimal pool size, $n = \frac{1}{\sqrt{p}}$ of Dorfman algorithm. Finucan (1964) gave an approximation to the optimal hierarchical scheme when there exist no test errors. Samuels (1978) provided a simple method to obtain the optimal pool size for the Dorfman algorithm without test errors, while correctly indicating the fault of the assumption of unimodality used by Finucan and others. Turner et al. (1988) mentioned neither Wilks' method nor Feller's approximation always lead to the optimal pool size. They used a calculus based approach and Wilks' suggestion to get the optimal pool size. Wu and Zhao (1994) provided a concrete procedure to determine the precise optimal plan which minimizes the expected number of tests per specimen for hierarchical group testing problem when there exist test errors such as false negative and false positive.

CHAPTER 3

HIERARCHICAL AND SQUARE ARRAY ALGORITHMS

3.1 Introduction

The focus of this chapter is to research the utility of two-dimensional array-based group testing algorithms for case identification in the presence of test error.

Array-based specimen pooling is an alternative to hierarchical group testing which uses overlapping pools. In its simplest form, n^2 specimens are placed on an $n \times n$ matrix. Pools of size n are made from all samples in the same row or in the same column. These $2n$ pools are then tested and, assuming no false negative tests, all positive specimens will lie at the intersection of a positive row pool and a positive column pool. Any ambiguities are resolved by individually testing all specimens at these intersections. Phatarfod and Sudbury (1994) derived the expected number of tests for two-dimensional array (i.e., matrix) group testing procedures. This approach is also employed in genetics (Bruno et al. 1995; Amemiya et al. 1992; Barillot, Lacroix, and Cohen 1991) . However, with the exception of Section 5 of Phatarfod and Sudbury (1994), the array-based group testing literature does not consider test error.

In this chapter, we derive and compare the operating characteristics of hierarchical and square array based testing algorithms for case identification in the presence of testing error. We assume constant sensitivity and specificity for each test independent of pool size for this chapter. The operating characteristics investigated include efficiency (i.e., expected number of tests per specimen) and error rates (i.e., sensitivity, specificity, positive and negative predictive values, per-family error rate, and per-comparison error rate). The methodology is illustrated by comparing different pooling algorithms for the detection of individuals recently infected with HIV in North Carolina and Malawi.

3.2 Assumptions

In order to derive operating characteristics of the various pooling algorithms considered, we make the assumptions enumerated below. These assumptions are general enough to apply to both the hierarchical and square array algorithms.

Assumption 1 *All specimens are independent and identically distributed with probability p of being positive.*

We refer to p as the *prevalence* and let $q = 1 - p$.

Assumption 2 *Given a pool $\mathcal{P}$ containing at least one positive specimen is tested, the probability $\mathcal{P}$ tests positive equals S_e.*

We refer to S_e as the *test sensitivity*. Assumption 2 implies that the test sensitivity is independent of the number of specimens within a pool and the number of positive specimens therein. It follows that S_e is also the sensitivity for a test of an individual specimen, i.e., a pool of size 1. We do not consider here more elaborate models where test sensitivity depends on the number of specimens within a pool, i.e., dilution effects (Hwang, 1976; Wein and Zenios, 1996). The practical implication of Assumption 2 is

that the results in this paper are applicable only to settings where the largest pool size is not believed to suffer appreciable dilution effects. Determination of the maximum allowable pool size will be application specific. For example, Litvak et al. (1994) consider pool sizes up to 15 specimens when using an ELISA test to detect individuals with HIV antibodies. Another example is given by detection of acute HIV in antibody negative populations where NAATs are thought to be sufficiently sensitive such that pools of size 90 or 100 are often used in practice (Quinn et al., 2000; Pilcher et al., 2002; Pilcher et al., 2005).

Assumption 3 *Given a pool $\mathcal{P}$ containing no positive specimens is tested, the probability $\mathcal{P}$ tests positive equals $1 - S_p$.*

We refer to S_p as the *test specificity*. Assumption 3 implies test specificity is independent of pool size.

3.3 Hierarchical Algorithm

In this section, we present the efficiency and error rates of a hierarchical group testing algorithm with S stages under Assumption 1-3. These results first appeared in Johnson et al. (1991), but often go unrecognized in the literature and thus are restated here.

Consider a hierarchical algorithm where a master pool of size $n_1 = k_1 k_2 \cdots k_{S-1}$ is tested at first stage where $k_1 k_2 \cdots k_{S-1}$ are positive integers. If the master pool tests positive, then k_1 pools of size $n_2 = k_2 \cdots k_{S-1}$ are tested, and so forth. Denote this algorithm by $\mathcal{A} = DS(n_1 : n_2 : n_3 : \ldots : n_{S-1} : n_S)$ where $n_1 = k_1 k_2 \cdots k_{S-1}$, $n_2 = k_2 \cdots k_{S-1}$, $\cdots$, $n_{S-1} = k_{S-1}$, $n_S = 1$. Let X_{si} be random variables that equal 1 if the i^{th} pool in s^{th} stage **tests** positive, and 0 otherwise for $s = 1, \ldots, S$ and $i = 1, \ldots, n_1/n_s$.

3.3.1 Efficiency

Let the number of tests be $T = T_1 + T_2 + T_3 + \ldots + T_S$ where T_s is the number of tests at the s^{th} stage for $s = 1, 2, 3, \ldots, S$. Then the expected number of tests given n_1 specimens equals

$$E(T) \quad = 1 + k_1 E(X_{1i}) + k_1 k_2 E(X_{2i}) + \ldots + n_1 E(X_{S-1,i}).$$

From equation (6.18) of Johnson et al. (1991),

$$E(X_{si}) \quad = q^{n_1}(1 - S_p)^s + \textstyle\sum_{j=1}^{s-1}(q^{n_{j+1}} - q^{n_j}) S_e^j (1 - S_p)^{s-j} + (1 - q^{n_s}) S_e^s, \tag{3.1}$$

for $s = 1, 2, \ldots, S - 1$.

From equation (3.1), it follows that for a two-stage hierarchical algorithm, i.e., $S = 2$, the expected number of tests per specimen for $D2$ is

$$E(D2) \equiv E\left(\frac{T}{n_1}\right) = \frac{1}{n_1} + f(n_1), \tag{3.2}$$

where $f(n) \equiv (1 - S_p)q^n + S_e(1 - q^n)$, is the probability a pool of size n tests positive without any knowledge of the true status of the pool. For notational simplicity, we suppress the dependence of f on the parameters p, S_e, and S_p. For given values of p, S_e, and S_p, the optimally efficient two stage procedure is determined by the value of n that minimizes (3.2). Special cases of (3.2) were also derived by Litvak et al. (1994) and Gastwirth (2000). In particular, for $n_1 = 15$, (3.2) reduces to the first equation in Section 2.4 of Litvak et al. (1994). Similarly, if $S_e = 1$, (3.2) reduces to equation (2) of Gastwirth (2000).

For a three-stage hierarchical algorithm, i.e., $S = 3$, it follows from equation (3.1)

that the expected number of tests per specimen for $D3$ is

$$E(D3) \equiv E\left(\frac{T}{n_1}\right) = \frac{1}{n_1} + \frac{f(n_1)}{k_2} + {S_e}^2(1-q^{k_2}) + (1-S_p)f(n_1-k_2)q^{k_2}. \tag{3.3}$$

Finucan (1964) showed that, in the absence of test error, $n_2 = \sqrt{n_1}$ is approximately optimal with regards to minimizing the expected number of tests per specimen. Thus in the Application below (Section 3.7) we primarily consider configurations of $D3$ where $n_2 = \sqrt{n_1}$.

3.3.2 Error rates

From equations (6.15) and (6.16) of Johnson et al. (1991), the pooling sensitivity and specificity of DS are

$$S_e(DS) \;\; = S_e^S, \tag{3.4}$$

and

$$S_p(DS) \;\; = 1 - \textstyle\sum_{s=1}^{S}(q^{n_s-1} - q^{n_{s-1}-1})S_e^{s-1}(1-S_p)^{S-(s-1)}, \tag{3.5}$$

where $q^{n_0-1} \equiv 0$. It follows that the pooling sensitivity and specificity of the Dorfman procedure (i.e., $S = 2$) are

$$S_e(D2) \;\; = S_e^2, \tag{3.6}$$

and

$$S_p(D2) = 1 - (1-S_p)f(n_1-1). \tag{3.7}$$

For $n_1 = 15$, (3.6) and (3.7) reduce to equations (1) and (4) of Litvak et al. (1994). For $S = 3$, the pooling sensitivity and specificity are $S_e(D3) = S_e^3$, and $S_p(D3) = 1 - [(1-S_p)\{(1-S_p)f(n_1-k_2)q^{k_2-1} + S_e^2(1-q^{k_2-1})\}]$. Note because $S_e^3 \leq S_e^2 \leq S_e$, neither $D2$ nor $D3$ will ever improve over individual testing in terms of sensitivity.

3.4 Square Array without Master Pool Testing

In this section, we derive the operating characteristics of a two-stage square array testing algorithm in the presence of testing error, denoted by $A2(n:1)$. In particular, consider the $n \times n$ square array set-up of Phatarfod and Sudbury (1994) where n^2 specimen are placed on an $n \times n$ matrix. Pools are then made from all samples in the same row or in the same column. These $2n$ pools (n row pools and n column pools) are then tested and, assuming no test error, all positive specimens will lie at the intersection of a positive row pool and a positive column pool. Therefore all specimens at the intersection of a positive row and a positive column are subsequently tested. However, when there is testing error, one must allow for the possibility of positive row pools and no positive column pools (or vice-versa). Thus we assume that the $(i,j)^{th}$ sample is tested individually if either both the i^{th} row and the j^{th} column test positive, the i^{th} row tests positive but all columns test negative, or the j^{th} column tests positive but all rows test negative. Let R_i represent the test outcome of the i^{th} row, and C_j represent the test outcome of the j^{th} column. Similarly, denote the true values by R_i^T and C_j^T. Let X_{ij} denote the test outcome of individual (i,j) and Y_{ij} denote the true status of individual (i,j) such that $R_i^T = I[\sum_j Y_{ij} > 0]$ and $C_j^T = I[\sum_i Y_{ij} > 0]$. We will make the following additional assumption to derive efficiencies and error rates for array pooling algorithms:

Assumption 4 *Given the true status of i^{th} row and j^{th} column, the i^{th} row and j^{th} column tests are conditionally independent of each other.*

3.4.1 Efficiency

Let the number of tests be $T = T_1 + T_2$, where $T_1 = 2n$ corresponds to the row and column testing and $T_2 = \sum_{i,j} T_{2ij}$ corresponds to the possible subsequent individual

testing where

$$T_{2ij} = \begin{cases} 1 & \text{if } R_i = 1 \text{ and } C_j = 1 \\ 1 & \text{if } R_i = 1 \text{ and } \sum C_j = 0 \\ 1 & \text{if } \sum R_i = 0 \text{ and } C_j = 1 \\ 0 & \text{otherwise} \end{cases}. \tag{3.8}$$

To derive the expected number of tests, by symmetry we write

$$E(T_{2ij}) = \Pr[R_i = 1, C_j = 1] + 2\Pr[R_i = 1, \sum_j C_j = 0]. \tag{3.9}$$

The first component of the right side of (3.9) equals

$$\Pr[R_i = 1, C_j = 1] \quad = \textstyle\sum_{r,c=0}^{1} \Pr[R_i = 1, C_j = 1 | R_i^T = r, C_j^T = c] \Pr[R_i^T = r, C_j^T = c].$$

Under Assumptions 1 - 4, direct substitution and some algebra yields

$$\begin{aligned} \Pr[R_i = 1, C_j = 1] \quad &= (1 - S_p)^2 q^{2n-1} + 2S_e(1 - S_p)(1 - q^{n-1})q^n + S_e^2(1 - 2q^n + q^{2n-1}) \\ &= S_e f(n) + (1 - S_p - S_e)q^n f(n-1) \equiv g(n). \end{aligned}$$

The second component of the right side of (3.9) can be written as

$$\sum_{r=0}^{1} \sum_{c=0}^{n} \Pr[R_i = 1, \sum_j C_j = 0 | R_i^T = r, \sum_j C_j^T = c] \Pr[R_i^T = r, \sum_j C_j^T = c]. \tag{3.10}$$

For $c \in \{0, \ldots, n\}$, let

$$\beta_0(c) \equiv \Pr[R_i^T = 0, \sum_j C_j^T = c] = \binom{n}{c} (q^{n^2 - cn + c})(1 - q^{n-1})^c, \tag{3.11}$$

$$\beta_1(c) \equiv \Pr[R_i^T = 1, \sum_j C_j^T = c] = \binom{n}{c} q^{n^2 - cn} (1 - q^n)^c - \beta_0(c), \tag{3.12}$$

$$\gamma_0(c) \equiv \Pr[R_i = 1, \sum_j C_j = 0 | R_i^T = 0, \sum_j C_j^T = c] = (1 - S_p)(1 - S_e)^c S_p^{n-c},$$

and

$$\gamma_1(c) \equiv \Pr[R_i = 1, \sum_j C_j = 0 | R_i^T = 1, \sum_j C_j^T = c] = S_e S_p^{n-c}(1 - S_e)^c,$$

where we define $\gamma_1(0) \equiv 0$. Then it follows that

$$\Pr[R_i = 1, \sum_j C_j = 0] = \sum_{c=0}^{n} \{\gamma_0(c)\beta_0(c) + \gamma_1(c)\beta_1(c)\} \equiv h(n),$$

such that the expected number of tests per specimen for $A2$ is

$$E(A2) \equiv E\left(\frac{T}{n^2}\right) = \frac{2}{n} + g(n) + 2h(n). \tag{3.13}$$

Note if $S_e = S_p = 1$, then $h(n) = 0$, $g(n) = f(n) - q^n f(n-1)$ and $f(n) = 1 - q^n$ such that $E(A2) = 2/n + 1 - 2q^n + q^{2n-1}$, which equals equation (2) of Phatarfod and Sudbury (1994).

3.4.2 Error rates

For $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, n\}$,

$$\begin{aligned}
&1 - S_p(A2) = \Pr[X_{ij} = 1 | Y_{ij} = 0] \\
&= \Pr[R_i = 1, C_j = 1, X_{ij} = 1 | Y_{ij} = 0] + 2\Pr[R_i = 1, \textstyle\sum_k C_k = 0, X_{ij} = 1 | Y_{ij} = 0] \\
&= \Pr[X_{ij} = 1 | Y_{ij} = 0, R_i = 1, C_j = 1] \Pr[R_i = 1 | Y_{ij} = 0] \Pr[C_j = 1 | Y_{ij} = 0] \\
&+2\Pr[X_{ij} = 1 | Y_{ij} = 0, R_i = 1, \textstyle\sum_k C_k = 0] \Pr[R_i = 1, \textstyle\sum_k C_k = 0 | Y_{ij} = 0] \\
&= (1 - S_p) f(n-1)^2 + 2(1 - S_p) \Pr[R_i = 1, \textstyle\sum_k C_k = 0 | Y_{ij} = 0],
\end{aligned} \tag{3.14}$$

where the second equality is from the definition of algorithm $A2$ given by (3.8) and the third equality holds by Assumption 4. Note that $\Pr[R_i = 1 | Y_{ij} = 0] = \Pr[C_j = 1 | Y_{ij} =$

$0] = f(n-1)$. Next let

$$\begin{aligned} h(n|y) &\equiv \Pr[R_i = 1, \textstyle\sum_k C_k = 0 | Y_{ij} = 0] \\ &= \textstyle\sum_{r=0}^{1} \sum_{c=0}^{n} \left\{ \Pr[R_i = 1, \sum_k C_k = 0 | R_i^T = r, \sum_k C_k^T = c, Y_{ij} = 0] \right. \\ &\left. \times \Pr[R_i^T = r, \textstyle\sum_k C_k^T = c | Y_{ij} = 0] \right\}, \end{aligned} \tag{3.15}$$

$$\beta_0(c|y) \equiv \Pr[R_i^T = 0, \textstyle\sum_k C_k^T = c | Y_{ij} = 0] = \frac{\beta_0(c)}{q},$$

and

$$\begin{aligned} \beta_1(c|y) &\equiv \Pr[R_i^T = 1, \textstyle\sum_k C_k^T = c | Y_{ij} = 0] = \Pr[\textstyle\sum_k C_k^T = c | Y_{ij} = 0] - \beta_0(c|y) \\ &= \tbinom{n-1}{c}(1-q^n)^c q^{n^2-nc-1} + \tbinom{n-1}{c-1}(1-q^n)^{c-1} q^{n^2-nc}(1-q^{n-1}) - \beta_0(c|y), \end{aligned}$$

for $c \in \{0, ..., n\}$ where $\beta_1(0|y) \equiv 0$. Then

$$h(n|y) = \sum_{r=0}^{1} \sum_{c=0}^{n} \beta_r(c|y)\gamma_r(c), \tag{3.16}$$

since $\gamma_r(c) \equiv \Pr[R_i = 1, \sum_k C_k = 0 | R_i^T = r, \sum_k C_k^T = c, Y_{ij} = 0]$. Thus, (3.14) is equivalent to

$$S_p(A2) = 1 - \{(1-S_p)f(n-1)^2 + 2(1-S_p)h(n|y)\}. \tag{3.17}$$

We can derive pooling sensitivity using several applications of Assumptions 1 and 2:

$$
\begin{aligned}
S_e(A2) &= \Pr[X_{ij}=1|Y_{ij}=1] \\
&= \Pr[R_i=1, C_j=1, X_{ij}=1|Y_{ij}=1] + 2\Pr[R_i=1, \textstyle\sum_k C_k=0, X_{ij}=1|Y_{ij}=1] \\
&= S_e^3 + 2S_e^2 \Pr[C_j=0|Y_{ij}=1] \textstyle\prod_{k\neq j} \Pr[C_k=0] \\
&= S_e^3 + 2S_e^2(1-S_e)\Pr[C_k=0]^{n-1} \\
&= S_e^3 + 2S_e^2(1-S_e)\{1-f(n)\}^{n-1}.
\end{aligned}
\tag{3.18}
$$

3.5 Square Array with Master Pool Testing

We also consider a three-stage square array testing where we first test a master pool of size n^2. If the master pool tests negative, the procedure stops. Otherwise, the procedure continues as in $A2$. Denote this square array testing procedure with master pool testing by $A2M(n^2 : n : 1)$.

3.5.1 Efficiency

Let the number of tests be $T = T_0 + T_1 + T_2$ where $T_0 = 1$ corresponds to testing the master pool, T_1 corresponds to possible row and column testing and T_2 corresponds to possible individual testing. To compute the efficiency of $A2M$, let X_0 be a random variable that equals 1 if the master pool tests positive and 0 otherwise such that $T_1 = 2nX_0$ and $E(T_1) = 2nf(n^2)$. Next write $T_2 = \sum_{i,j} T_{2ij}$ where

$$
T_{2ij} = \begin{cases}
1 & \text{if } X_0 = 1, R_i = 1 \text{ and } C_j = 1 \\
1 & \text{if } X_0 = 1, R_i = 1 \text{ and } \sum C_j = 0 \\
1 & \text{if } X_0 = 1, \sum R_i = 0 \text{ and } C_j = 1 \\
0 & \text{otherwise}
\end{cases}
$$

such that the expected number of tests per specimen for $A2M$ is $E(A2M) \equiv \frac{1}{n^2} + \frac{2}{n} f(n^2) + E(T_{2ij})$, where

$$E(T_{2ij}) = \Pr[X_0 = 1, R_i = 1, C_j = 1] + 2\Pr[X_0 = 1, R_i = 1, \sum_j C_j = 0]. \tag{3.19}$$

It is straightforward to show the first part of the right side of (3.19) equals

$$\Pr[X_0 = 1, R_i = 1, C_j = 1] = (1 - S_p)^2 q^{n^2} (1 - S_p - S_e) + S_e g(n). \tag{3.20}$$

Likewise, the second part of the right side of (3.19), i.e., $\Pr[X_0 = 1, R_i = 1, \sum_j C_j = 0]$, can be written as

$$\begin{aligned} \sum_{r=0}^{1} \sum_{c=0}^{n} \Big\{ & \Pr[R_i = 1, \textstyle\sum_j C_j = 0 | X_0 = 1, R_i^T = r, \sum_j C_j^T = c] \\ & \times \Pr[X_0 = 1 | R_i^T = r, \textstyle\sum_j C_j^T = c] \Pr[R_i^T = r, \sum_j C_j^T = c] \Big\}, \end{aligned}$$

which implies

$$\begin{aligned} \Pr[X_0 = 1, R_i = 1, \textstyle\sum_j C_j = 0] &= (1 - S_p)\gamma_0(0)\beta_0(0) + S_e \textstyle\sum_{c=1}^{n} \{\gamma_0(c)\beta_0(c) + \gamma_1(c)\beta_1(c)\} \\ &= (1 - S_p - S_e)\gamma_0(0)\beta_0(0) + S_e h(n). \end{aligned}$$

Therefore the expected number of tests per specimen for $A2M$ is

$$E(A2M) = \frac{1}{n^2} + (1 - S_p - S_e) q^{n^2} \left\{ \frac{2}{n} + (1 - S_p)^2 + 2(1 - S_p) S_p^n \right\} + S_e E(A2).$$

3.5.2 Error rates

To derive the pooling specificity of $A2M$, write

$$\begin{aligned} 1 - S_p(A2M) &= \Pr[X_{ij} = 1 | Y_{ij} = 0] \\ &= (1 - S_p)\mu(n|y) + 2(1 - S_p)\nu(n|y) \end{aligned} \tag{3.21}$$

where

$$\begin{aligned} \mu(n|y) &= \Pr[X_0 = 1, R_i = 1, C_j = 1 | Y_{ij} = 0] \\ &= f(n^2 - 2n + 1)\{(1 - S_p)q^{n-1}\}^2 + S_e^2(1 - q^{n-1})\{(1 - S_p)q^{n-1} + f(n - 1)\}, \end{aligned}$$

and

$$\begin{aligned} \nu(n|y) &= \Pr[X_0 = 1, R_i = 1, \textstyle\sum_k C_k = 0 | Y_{ij} = 0] \\ &= (1 - S_p)\beta_0(0|y)\gamma_0(0) + S_e \textstyle\sum_{r=0}^{1} \sum_{c=1}^{n} \beta_r(c|y)\gamma_r(c), \end{aligned}$$

since $\gamma_r(c) = \Pr[R_i = 1, \sum_k C_k = 0 | R_i^T = r, \sum_k C_k^T = c, Y_{ij} = 0, X_0 = 1]$.

Applying Assumptions 1 and 2 several times, pooling sensitivity equals

$$\begin{aligned} S_e(A2M) &= \Pr[X_{ij} = 1, R_i = 1, C_j = 1, X_0 = 1 | Y_{ij} = 1] \\ &\quad + 2\Pr[X_{ij} = 1, R_i = 1, \textstyle\sum_k C_k = 0, X_0 = 1 | Y_{ij} = 1] \\ &= S_e^4 + 2S_e^3(1 - S_e)\{1 - f(n)\}^{n-1}. \end{aligned} \tag{3.22}$$

3.6 Assessing Variability

Thus far we have presented expressions for different group testing algorithm parameters such as the expected number of tests, pooling sensitivity and pooling specificity. Formulae for pooling predictive values (PPV and NPV) and the other error rates described in Section 2.3 follow immediately. However, for a particular set of specimens, the *observed* number of tests may differ from the expected number of tests due to sampling variability. While the error rates for a particular set of specimens are not directly

observable in the absence of a gold standard test, clearly these latent "observed" error rates may also differ from the corresponding parameter values due to sampling variability. Quantifying this uncertainty associated with the different operating characteristics can be helpful in comparing different pooling strategies. For example, if two group testing procedures have comparable expected number of tests, the one with smaller variance may be preferable operationally.

Explicit expressions for the variance associated with the number of tests per specimen for $D2$, $D3$, $A2$, and $A2M$ are derived in the Appendix B. In turn, large sample probability intervals (PIs) for the observed number of tests per specimen can be computed by appealing to the Central Limit Theorem. For instance, suppose N samples are to be tested using $D3(n : \sqrt{n} : 1)$ with N sufficiently larger than n. Then there is an approximate $(1 - \alpha)$ probability that the observed number of tests per specimen required to identify all N individuals as positive or negative will be in the PI

$$E(D3) \pm z_{1-\alpha/2}\sqrt{\frac{Var(D3)}{N/n}},$$

where $z_{1-\alpha/2}$ is the $1-\alpha/2$ quantile of the standard normal distribution, $E(D3)$ is given by equation (3.3) and $Var(D3)$ follows from equation (B.2) in Appendix B.2. Similar reasoning can be employed to obtain PIs for the observed number of tests using $D2$, $A2$, and $A2M$.

The cumulative distribution functions (CDFs) for the observed pooling sensitivity, specificity, and predictive values are derived in Appendix C. Using these CDFs, corresponding PIs can easily be determined. For instance, we show that for any pooling algorithm $\mathcal{A}$ applied to N specimens, the CDF of the observed sensitivity ($S_e^O(\mathcal{A})$)

equals

$$\Pr[S_e^O(\mathcal{A}) \leq s] = \sum_{y=0}^{N} \sum_{x=0}^{\lfloor sy \rfloor} \binom{y}{x} S_e(\mathcal{A})^x (1 - S_e(\mathcal{A}))^{y-x} \binom{N}{y} p^y (1-p)^{N-y},$$

where $\lfloor sy \rfloor$ denotes the largest integer less than or equal to sy. Using this expression, an approximate $(1-\alpha)$ PI for the observed sensitivity is given by (s_L, s_U) where s_L is the largest value such that $\Pr[S_e^O(\mathcal{A}) \leq s_L] \leq \alpha/2$ and s_U is the smallest value such that $\Pr[S_e^O(\mathcal{A}) \geq s_U] \leq \alpha/2$. PIs for pooling specificity and predictive values follow analogously using the corresponding CDFs given in Appendix C.

3.7 Application

Using the results derived above, in this section we explore the operating characteristics of individual testing, $D2$, $D3$, $A2$, and $A2M$ for identification of acute HIV using NAATs. For our first example, we consider a setting similar to the NC STAT program. First, we assume prevalence of acute HIV is $p = 0.0002$ (Pilcher et al., 2005) and NAAT has a 99% test specificity (Hecht et al., 2002) and 90% test sensitivity. Suppose further we are limited to a master pool size of 100 due to dilution effects (Quinn et al., 2000). Under these assumptions, Figure 3.1 depicts the efficiency, pooling sensitivity, specificity, PPV and $PFER$s of individual testing, $D2(n:1)$, $D3(n:\sqrt{n}:1)$ (here $k_2 = \sqrt{n}$), $A2M(n:\sqrt{n}:1)$, and $A2(\sqrt{n}:1)$ as a function of the number of specimens n. (Recall that $D2(n:1)$ denotes two-stage hierarchical testing with pools of size n in the first stage; $D3(n:\sqrt{n}:1)$ denotes three-stage hierarchical testing with pools of size n at the first stage and pools of size $\sqrt{n}$ at the second stage; $A2(\sqrt{n}:1)$ denotes two-stage $\sqrt{n} \times \sqrt{n}$ array testing; and $A2M(n:\sqrt{n}:1)$ denotes three-stage $\sqrt{n} \times \sqrt{n}$ array testing wherein the first stage entails testing a master pool of size n.) The panel (a) of Figure 3.1 indicates the two most efficient algorithms are $D3$ and $A2M$ when

$n = 100$. The expected number of tests per specimen for $D3(90:10:1)$, the algorithm currently employed by the NC STAT program, is 0.016. For $N = 8,505$ total specimens (Pilcher et al., 2002), the 95% PI for efficiency of $D3(90:10:1)$ is (0.010, 0.021), which contains the observed rate of 0.018 reported by Pilcher et al. (2002). Under these same conditions, the expected number of tests per specimen is also 0.016 (95% PI: 0.008 to 0.024) for $A2M(100:10:1)$. Panels (b) and (d) of Figure 3.1 indicate $D3$ and $A2M$ are also preferable with regards to pooling specificity and, especially, PPV. However these two algorithms are also the least sensitive as depicted in panel (c) of Figure 3.1. Panel (e) of Figure 3.1 indicates $D3$ and $A2M$ are preferable with regards to pooling $PFER$. We also see that other algorithms are better than individual testing with regards to the efficiency, pooling specificity, PPV and $PFER$. Overall, these results suggest by moving from $D3(90:10:1)$ to $A2M(100:10:1)$, could improve pooling specificity, sensitivity, PPV, $PCER$s, and $PFER$s without sacrificing efficiency. In particular, given the NC STAT program processes 120,000 specimens per year, this change in pooling algorithm would result in a decrease in $PFER_2$ from 6.5 to 5.1. In other words, on average 1-2 additional acute HIV cases would be detected each year, representing a 5-10% increase over the current detection rate (Pilcher et al., 2005).

As a second motivating example, we consider a setting similar to that described by Pilcher et al. (2004), who employed $D3(50:10:1)$ to identify acute HIV in Malawi. They found 4.5% of antibody negative males attending STD clinics to be NAAT positive. Assuming $S_p = 0.99$ and $S_e = 0.9$ as before and $p = 0.045$, the expected number of tests per specimen of $D3(50:10:1)$ is 0.40. For $N = 1,361$ total specimens (Pilcher et al., 2004), the 95% PI for efficiency is (0.32, 0.49). However as seen in panel (a) of Figure 3.2, there are more efficient algorithms in this setting. For example, using $A2M(100:10:1)$ results in 0.31 (95% PI: 0.23 to 0.38) tests per specimen on average while $D3(16:4:1)$ results in 0.32 (95% PI: 0.26 to 0.38) expected tests per specimen.

Based on the other panels of Figure 3.2, we see the choice between $A2M(100:10:1)$ and $D3(16:4:1)$ represents a trade-off in pooling sensitivity, specificity, PPV and $PFER$. Table 3.1 provides a closer look at the operating characteristics of individual testing, $D3(50:10:1)$ and three alternative algorithms. Arguably $D3(16:4:1)$ provides the best balance of efficiency and error rates while being less susceptible to dilution effects.

The results above suggest either $D3$ or $A2M$ are generally preferable for detection of acute HIV. Thus, we also consider how $D3$ and $A2M$ compare for prevalences ranging from 10^{-5} to 10^{-1}. As in the examples above, we assume $S_e = 0.9$, $S_p = 0.99$, and the maximum allowable pool size is 100 due to dilution effects. Under these assumptions, for each prevalence we found the values of $n \in \{4, 9, 16, 25, \ldots, 100\}$ that minimize $E\{D3(n:\sqrt{n}:1)\}$ and $E\{A2M(n:\sqrt{n}:1)\}$. The expected numbers of tests per specimen, pooling specificities, sensitivities, and PPVs at these optimal values of n are depicted in Figure 3.3. These results indicate $A2M$ is generally the preferable algorithm for prevalence less than 0.01.

The examples above assume that prevalence, test sensitivity, and test specificity are known exactly, which will rarely be the case in practice. To account for such uncertainty one can use a Bayesian analysis where priors are placed on p, S_e and S_p (e.g., see Dendukuri and Joseph (2001)). Alternatively one can employ a "sensitivity analysis" wherein the operating characteristics of the pooling algorithms are examined over a range of values for p, S_e and/or S_p. For example, suppose investigators in Malawi are interested in the effect of the assumed values of S_e and S_p on the efficiency, pooling sensitivity and pooling PPV of $D3(16:4:1)$. Then a graphical display such as Figure 3.4 can be used to show $E(D3)$, $S_e(D3)$ and $PPV(D3)$ over a bivariate range of values of S_e and S_p. Note Figure 3.4 includes the special case of no test error, i.e., $S_e = S_p = 1$. For this case, panel (a) demonstrates that $D3$ will actually be less efficient than in the

presence of test error. Panels (b) and (c) demonstrate the more general phenomenon that pooling sensitivity, specificity and predictive values equal 1 when $S_e = S_p = 1$.

3.8 Discussion

We derived several operating characteristics of hierarchical and square array-based testing algorithms for case identification in the presence of testing error. Using these results, we showed that the NC STAT program's currently implemented pooling algorithm $D3(90:10:1)$ is approximately optimal among the algorithms considered here with respect to the expected number of tests per specimen. However, moving to the array-based pooling algorithm $A2M(100:10:1)$ would generally improve pooling error rates, in particular the expected number of false negatives ($PFER_2$), without resulting in a decrease in efficiency. Our results also suggest moving from $D3(50:10:1)$ to $D3(16:4:1)$ in the Malawi setting would lead to substantial improvement in efficiency as well as pooling error rates.

There are several areas of potential future research related to this work. First, we assume constant sensitivity and specificity independent of pool size. This assumption is not necessarily realistic and is highly dependent on individual disease and assay characteristics. For HIV, the sensitivity of NAAT is likely inversely related to the number of specimens per pool; as such, the applications above should be interpreted with caution. Incorporating previously proposed methods of Hwang (1976) and Wein and Zenios (1996) that allow sensitivity and specificity to be a function of pool size could be considered for future research. Second, we have only considered two dimensional square array algorithms. Generalizations to $n \times m$ arrays with $m \neq n$ should be straightforward, however Berger et al. (2000) suggest $m = n$ will be most efficient in the absence of test error. Likewise, the proposed square array methods could be

extended to higher dimensional arrays in the presence of testing error. The efficiency of higher dimensional arrays should simplify to the results given in Berger et al. (2000) when $S_e = S_p = 1$.

In summary, our results confirm that group testing algorithms can increase efficiency and confer remarkable accuracy (predictive value) across a broad range of prevalence in the presence of testing error. This ability of group testing algorithms to enhance the accuracy of low prevalence disease screening is likely under-appreciated by clinical laboratories. Of course, in clinical laboratory practice, functional constraints may affect the pooling algorithm choice. For example, both the master pool size and the number of test stages can affect the turn-around time for results. For investigators attempting to evaluate the potential use of group testing algorithms for a particular application, results such as ours can help estimate the trade-offs to be expected in terms of efficiency, accuracy and turn-around time.

TABLE 3.1: Comparison of operating characteristics for individual testing and four potential pooling algorithms to be used in Malawi for detection of acute HIV.

$\mathcal{A}$	$E(\mathcal{A})$	$S_p(\mathcal{A})$	$PCER(\mathcal{A})$	$S_e(\mathcal{A})$	$PCER_2(\mathcal{A})$	$PPV(\mathcal{A})$	$NPV(\mathcal{A})$
IT*	1.00	0.9900	0.0096	0.9000	0.0045	0.8092	0.9953
$D3(50)^\sharp$	0.40	0.9972	0.0028	0.7290	0.0122	0.9247	0.9874
$D3(16)^\dagger$	0.32	0.9989	0.0010	0.7290	0.0122	0.9696	0.9874
$A2M(49)^\ddagger$	0.34	0.9995	0.0005	0.6810	0.0144	0.9836	0.9852
$A2M(100)^\S$	0.31	0.9991	0.0009	0.6596	0.0153	0.9721	0.9842

*IT: Individual Test
$^\sharp D3(50) : D3(50 : 10 : 1)$
$^\dagger D3(16) : D3(16 : 4 : 1)$
$^\ddagger A2M(49) : A2M(49 : 7 : 1)$
$^\S A2M(100) : A2M(100 : 10 : 1)$

FIGURE 3.1: (a) Expected number of tests per specimen, (b) pooling specificity, (c) pooling sensitivity, (d) pooling PPV, (e) pooling $PFER$, and (f) pooling $PFER_2$ for different algorithms assuming test sensitivity $S_e = 0.9$, test specificity $S_p = 0.99$, and prevalence $p = 0.0002$. The ▲ denotes the three stage hierarchical pooling algorithm employed in the NC STAT Program. Note pooling specificity for individual testing equals $S_p = 0.99$ and is not shown in panel (b).

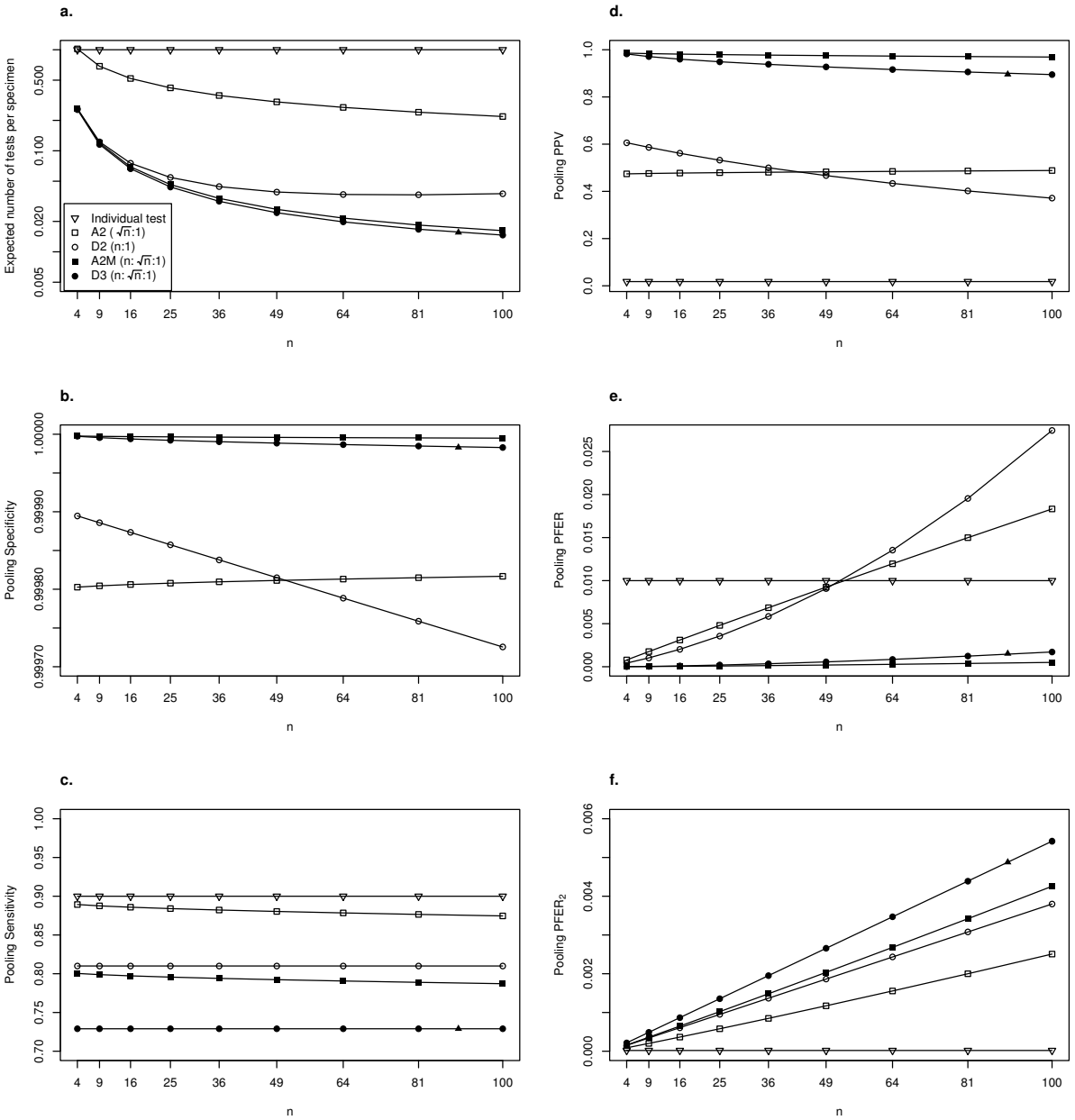


FIGURE 3.2: (a) Expected number of tests per specimen, (b) pooling specificity, (c) pooling sensitivity, (d) pooling PPV, (e) pooling $PFER$, and (f) pooling $PFER_2$ for different algorithms assuming test sensitivity $S_e = 0.9$, test specificity $S_p = 0.99$, and prevalence $p = 0.045$. The ▲ denotes the three stage hierarchical pooling algorithm employed in Malawi.

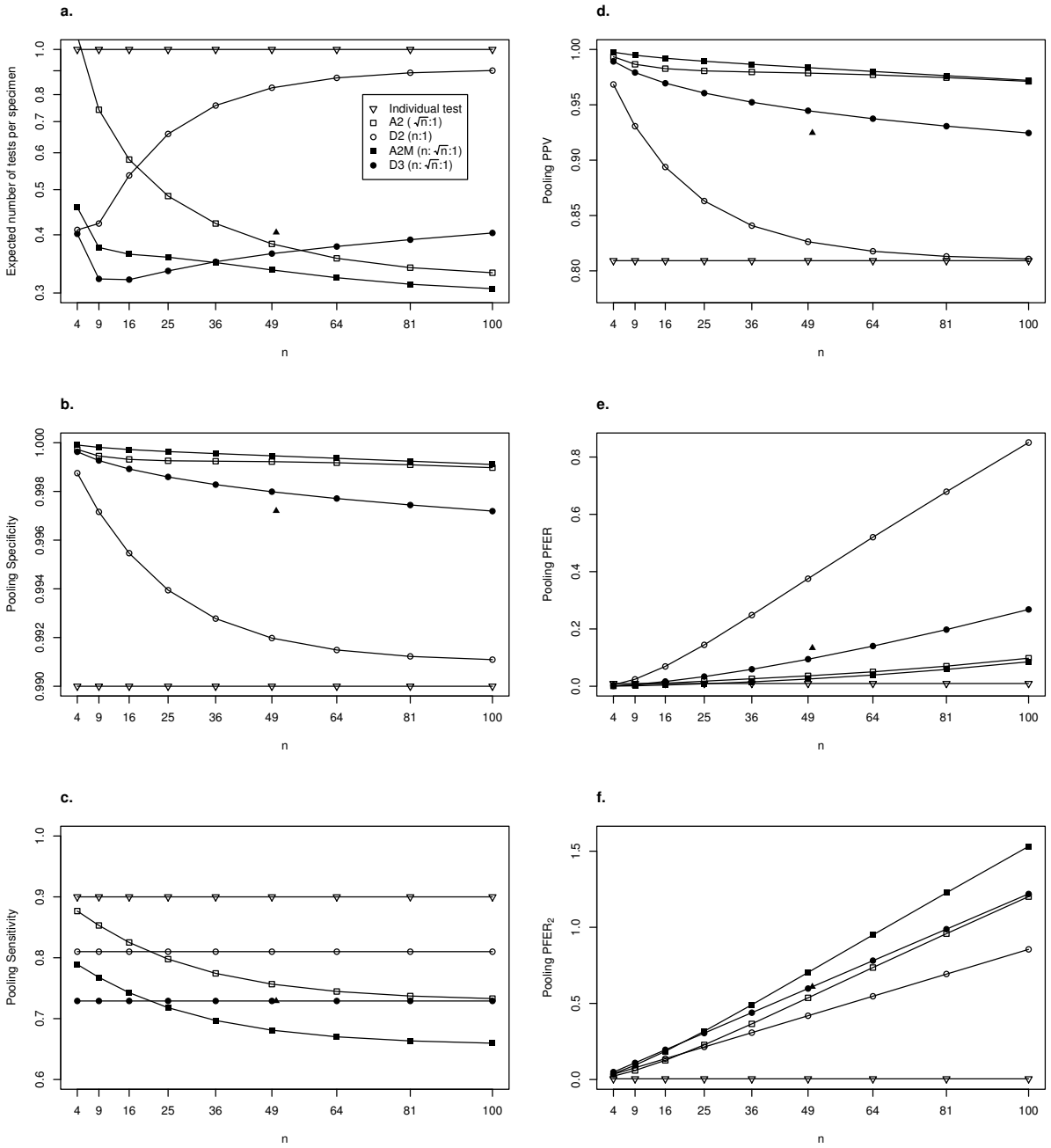


FIGURE 3.3: (a) Expected number of tests per specimen, (b) pooling specificity, (c) pooling sensitivity, and (d) pooling PPV for optimally efficient configurations of $D3$ and $A2M$ assuming test sensitivity $S_e = 0.9$, test specificity $S_p = 0.99$, and a maximum allowable pool size of 100.

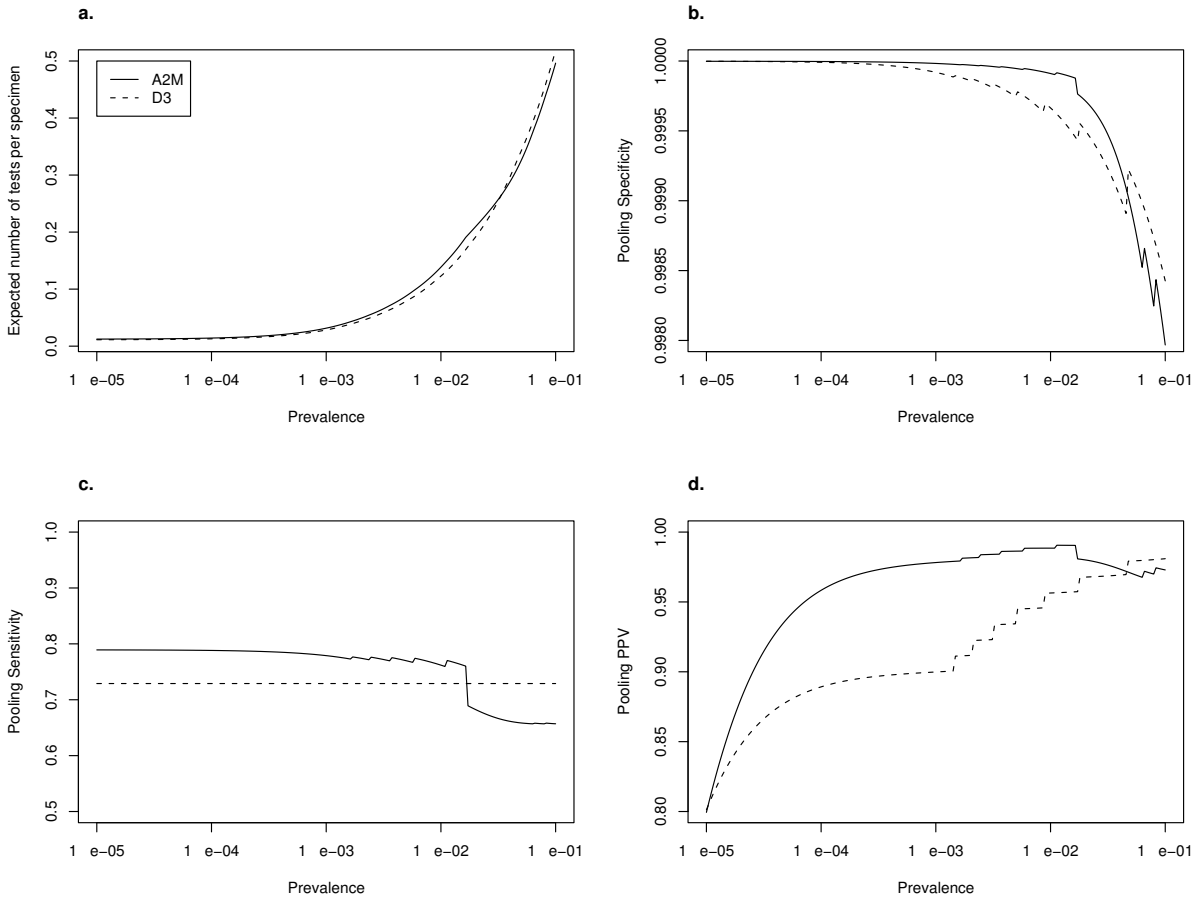
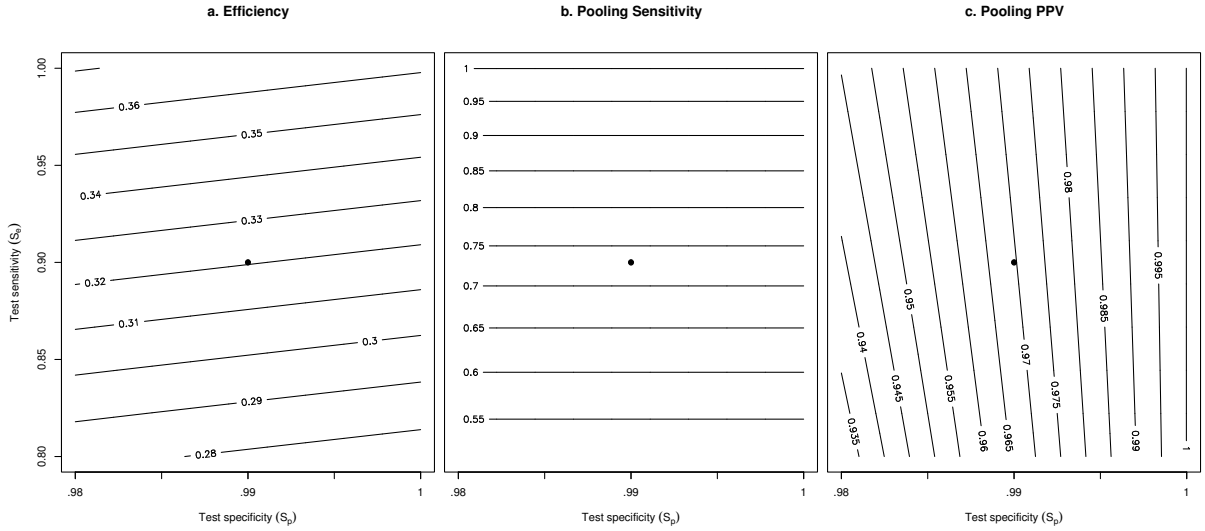


FIGURE 3.4: Contour plots of (a) expected number of tests per specimen, (b) pooling sensitivity, and (c) pooling PPV for $D3(16:4:1)$ assuming $p = 0.045$ as a function of test sensitivity (S_e) and test specificity (S_p). The $\bullet$ denotes the values of S_e and S_p assumed in Table 3.1.



CHAPTER 4

THREE-DIMENSIONAL ARRAY ALGORITHMS

4.1 Introduction

The focus of this chapter is to research aspects of three-dimensional array-based group testing algorithms for case identification in the presence of test error.

Array-based algorithms were proposed by several researchers. Phatarfod and Sudbury (1994) derived the expected number of tests for two-dimensional array (i.e., matrix) group testing procedures. Berger et al. (2000) extended this work to higher dimensional arrays assuming no test errors, i.e., no false negative or false positive tests.

In this chapter, we extend Berger et al. (2000)'s results to allow for imperfect testing. We derive efficiency and pooling measurement error rates such as specificity and sensitivity for three dimensional array-based pooling algorithms when there exist test errors. Algorithms with and without master pool testing are considered. Our results are compared with previously derived operating characteristics for hierarchical and two-dimensional array-based group testing algorithms in the presence of test errors.

4.2 Preliminaries

4.2.1 Notation

In addition to the notation introduced in Section 2.3 and Section 3, we need to define the following notation in order to derive the operating characteristics of three-dimensional array-based algorithms.

Let X_{i_1,i_2,i_3} denote the test outcome of individual (i_1, i_2, i_3) and Y_{i_1,i_2,i_3} denote the true status of individual (i_1, i_2, i_3) for $i_1 = 1, \ldots, L$, $i_2 = 1, \ldots, M$ and $i_3 = 1, \ldots, N$. For $i_1 = 1, \ldots, L$, let X_{i_1++} denote the test outcome for the pool of size MN corresponding to the ${i_1}^{th}$ planar slice from front to back. Define X_{+i_2+} for $i_2 = 1, \ldots, M$ and X_{++i_3} for $i_3 = 1, \ldots, N$ similarly. Denote the corresponding true values by Y_{i_1++}, Y_{+i_2+} and Y_{++i_3}.

4.2.2 Assumptions

In addition to Assumptions 1-3 in Section 3.2, we make the following additional assumption:

Assumption 5 *X_{i_1++}, X_{+i_2+} and X_{++i_3} are conditionally independent of each other given the true status of Y_{i_1++}, Y_{+i_2+} and Y_{++i_3}.*

4.3 Three-Dimensional Array Without Master Pool

In this section, we derive the efficiency and error rates of a three-dimensional $(L \times M \times N)$ array-based testing algorithm in the presence of test errors. This algorithm entails planar slices of a three-dimensional array and thus is denoted $A3P([L : M : N] : 1)$. For this method, in stage 1, each of the L planar slices from front to back, the M planar slices from top to bottom, and the N planar slices from left to right are tested.

In stage 2, each specimen is tested individually if at least two of the three planar slices containing this sample test positive. Figure 4.1 shows $A3P$ with $L = M = N = 3$. In this example, the total number of specimens is 27 and the black dots denote 9 ($= MN$) specimens in one of L planar slices.

4.3.1 Efficiency

Let the number of tests of $A3P$ be $T = T_1 + \sum_{i_1,i_2,i_3} T_{(i_1,i_2,i_3)}$ where $T_1 = L + M + N$ corresponds to pool testing (i.e., planar slices) and

$$T_{(i_1,i_2,i_3)} = \begin{cases} 1 & \text{if } X_{i_1++} = 1 \text{ and } X_{+i_2+} = 1 \text{ and } X_{++i_3} = 1 \\ 1 & \text{if } X_{i_1++} = 1 \text{ and } X_{+i_2+} = 1 \text{ and } \sum_{i_3=1}^{N} X_{++i_3} = 0 \\ 1 & \text{if } X_{i_1++} = 1 \text{ and } \sum_{i_2=1}^{M} X_{+i_2+} = 0 \text{ and } X_{++i_3} = 1 \\ 1 & \text{if } \sum_{i_1=1}^{L} X_{i_1++} = 0 \text{ and } X_{+i_2+} = 1 \text{ and } X_{++i_3} = 1 \\ 0 & \text{otherwise} \end{cases}$$

corresponds to the possible subsequent test of the $(i_1, i_2, i_3)^{th}$ sample. In other words, the $(i_1, i_2, i_3)^{th}$ sample is tested individually if at least two of the three planar slices containing this sample test positive.

The efficiency, or expected number of tests per specimen, of $A3P$ equals

$$E(A3P) = \frac{L + M + N}{LMN} + E(T_{(i_1,i_2,i_3)}),$$

where

$$E(T_{(i_1,i_2,i_3)}) \quad = g_{A3P}(L, M, N) + \sum_{j\in\{L,M,N\}} h_{A3P}(L, M, N; j), \tag{4.1}$$

$$g_{A3P}(L,M,N) \equiv \Pr[X_{i_1++}=1, X_{+i_2+}=1, X_{++i_3}=1],$$
$$h_{A3P}(L,M,N;N) \equiv \Pr[X_{i_1++}=1, X_{+i_2+}=1, \textstyle\sum_{i_3} X_{++i_3}=0],$$
$$h_{A3P}(L,M,N;M) \equiv \Pr[X_{i_1++}=1, \textstyle\sum_{i_2} X_{+i_2+}=0, X_{++i_3}=1],$$

and

$$h_{A3P}(L,M,N;L) \equiv \Pr[\textstyle\sum_{i_1} X_{i_1++}=0, X_{+i_2+}=1, X_{++i_3}=1].$$

Below we derive explicit forms for g_{A3P} and h_{A3P}. First let $q_1 = q^{LM} + q^{LN} + q^{MN}$, $q_2 = q^{L(M+N-1)} + q^{M(L+N-1)} + q^{N(L+M-1)}$, and $q_3 = q^{LM+MN+LN-(L+M+N)+1}$. Then

$$\begin{aligned} g_{A3P}(L,M,N) &= \sum_{l=0}^{1}\sum_{m=0}^{1}\sum_{n=0}^{1}\{\Pr[Y_{i_1++}=l, Y_{+i_2+}=m, Y_{++i_3}=n] \\ &\times \Pr[X_{i_1++}=1, X_{+i_2+}=1, X_{++i_3}=1 | Y_{i_1++}=l, Y_{+i_2+}=m, Y_{++i_3}=n]\} \\ &= (1-S_p)^3 q_3 + (1-S_p)^2 S_e (q_2 - 3q_3) + (1-S_p) S_e^2 (q_1 - 2q_2 + 3q_3) \\ &+ S_e^3 (1 - q_1 + q_2 - q_3). \end{aligned}$$

To derive $h_{A3P}(L,M,N;N)$, we use the following decomposition.

$$\begin{aligned} h_{A3P}(L,M,N;N) &= \sum_{c_1=0}^{1}\sum_{c_2=0}^{1}\sum_{r=0}^{N}\{\Pr[Y_{i_1++}=c_1, Y_{+i_2+}=c_2, \sum_{i_3} Y_{++i_3}=r] \\ &\times \Pr[X_{i_1++}=1, X_{+i_2+}=1, \sum_{i_3} X_{++i_3}=0 | Y_{i_1++}=c_1, Y_{+i_2+}=c_2, \sum_{i_3} Y_{++i_3}=r]\}. \end{aligned}$$

For $r \in 0, \cdots, N$, let

$$\begin{aligned} \beta_{00}(L,M,N;r) &\equiv \Pr[Y_{i_1++}=0, Y_{+i_2+}=0, \textstyle\sum_{i_3} Y_{++i_3}=r] \\ &= \binom{N}{r}(q^{LMN-r(LM-L-M+1)})(1-q^{LM-L-M+1})^r, \end{aligned}$$

$$\begin{aligned} \beta_{01}(L,M,N;r) &\equiv \Pr[Y_{i_1++}=0, Y_{+i_2+}=1, \textstyle\sum_{i_3} Y_{++i_3}=r] \\ &= \binom{N}{r}(q^{LMN-r(LM-M)})(1-q^{r(L-1)})^r, \end{aligned}$$

$$\begin{aligned}\beta_{10}(L,M,N;r) &\equiv \Pr[Y_{i_1++}=1, Y_{+i_2+}=0, \textstyle\sum_{i_3} Y_{++i_3}=r] \\ &= \tbinom{N}{r}(q^{LMN-r(LM-L)})(1-q^{r(M-1)})^r,\end{aligned}$$

$$\begin{aligned}\beta_{11}(L,M,N;r) &\equiv \Pr[Y_{i_1++}=1, Y_{+i_2+}=1, \textstyle\sum_{i_3} Y_{++i_3}=r] \\ &= \tbinom{N}{r}(q^{LMN-r(LM)})(1-q^{LM})^r \\ &\quad -\{\beta_{00}(L,M,N;r)+\beta_{01}(L,M,N;r)+\beta_{10}(L,M,N;r)\},\end{aligned}$$

$$\begin{aligned}\gamma_{00}(N;r) &\equiv \Pr[X_{i_1++}=1, X_{+i_2+}=1, \textstyle\sum_{i_3} X_{++i_3}=0 | Y_{i_1++}=0, Y_{+i_2+}=0, \sum_{i_3} Y_{++i_3}=r] \\ &= (1-S_p)^2(1-S_e)^r S_p^{N-r},\end{aligned}$$

$$\begin{aligned}\gamma_{01}(N;r) &\equiv \gamma_{10}(N;r) \\ &\equiv \Pr[X_{i_1++}=1, X_{+i_2+}=1, \textstyle\sum_{i_3} X_{++i_3}=0 | Y_{i_1++}=0, Y_{+i_2+}=1, \sum_{i_3} Y_{++i_3}=r] \\ &= (1-S_p)S_e(1-S_e)^r S_p^{N-r},\end{aligned}$$

and

$$\begin{aligned}\gamma_{11}(N;r) &\equiv \Pr[X_{i_1++}=1, X_{+i_2+}=1, \textstyle\sum_{i_3} X_{++i_3}=0 | Y_{i_1++}=1, Y_{+i_2+}=1, \sum_{i_3} Y_{++i_3}=r] \\ &= S_e^2(1-S_e)^r S_p^{N-r},\end{aligned}$$

where $\beta_{01}(L,M,N;0) \equiv 0$, $\beta_{10}(L,M,N;0) \equiv 0$, $\beta_{11}(L,M,N;0) \equiv 0$, $\gamma_{01}(N;0) \equiv 0$, $\gamma_{10}(N;0) \equiv 0$, and $\gamma_{11}(N;0) \equiv 0$. Then it follows that

$$h_{A3P}(L,M,N;N) = \textstyle\sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{r=0}^{N}\gamma_{ij}(N;r)\beta_{ij}(L,M,N;r). \tag{4.2}$$

Similarly one can show

$$h_{A3P}(L,M,N;L) = \textstyle\sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{r=0}^{L}\gamma_{ij}(L;r)\beta_{ij}(M,N,L;r),$$

and

$$h_{A3P}(L,M,N;M) = \textstyle\sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{r=0}^{M}\gamma_{ij}(M;r)\beta_{ij}(L,N,M;r).$$

Note if $S_e = S_p = 1$, then $\gamma_{ij}(N;r) = \gamma_{ij}(M;r) = \gamma_{ij}(L;r) = 0$ for $i,j \in \{(0,0),(0,1),(1,0),(1,1)\}$. Therefore,

$$E(A3P) \ = \tfrac{L+M+N}{LMN} + 1 - q_1 + q_2 - q_3,$$

which is equivalent to the result given in Section 2.2.1 of Berger et al. (2000). Note if $L = M = N$, then

$$E(A3P) = \frac{3}{N^2} + 1 - 3q^{N^2} + 3q^{N(2N-1)} - q^{3N^2-3N+1}.$$

which equals equation (2.3).

4.3.2 Error Rates

In this section, we derive the pooling specificity and pooling sensitivity of $A3P$. For $i_1 \in \{1,\ldots,L\}$, $i_2 \in \{1,\ldots,M\}$ and $i_3 \in \{1,\ldots,N\}$, the pooling false positive rate (i.e., 1-pooling specificity) is

$$\begin{aligned}
1 - S_p(A3P) &= \Pr[X_{i_1 i_2 i_3} = 1 | Y_{i_1 i_2 i_3} = 0] \qquad (4.3)\\
&= \Pr[X_{i_1++} = 1, X_{+i_2+} = 1, X_{++i_3} = 1, X_{i_1 i_2 i_3} = 1 | Y_{i_1 i_2 i_3} = 0]\\
&+ \Pr[X_{i_1++} = 1, X_{+i_2+} = 1, \sum_{j=1}^{N} X_{++j} = 0, X_{i_1 i_2 i_3} = 1 | Y_{i_1 i_2 i_3} = 0]\\
&+ \Pr[X_{i_1++} = 1, \sum_{j=1}^{M} X_{+j+} = 0, X_{++i_3} = 1, X_{i_1 i_2 i_3} = 1 | Y_{i_1 i_2 i_3} = 0]\\
&+ \Pr[\sum_{j=1}^{L} X_{j++} = 0, X_{+i_2+} = 1, X_{++i_3} = 1, X_{i_1 i_2 i_3} = 1 | Y_{i_1 i_2 i_3} = 0].
\end{aligned}$$

Therefore by Assumption 5,

$$1 - S_p(A3P) = (1 - S_p)\{g_{A3P|y}(L, M, N) + \textstyle\sum_{j\in\{L,M,N\}} h_{A3P|y}(L, M, N; j)\},$$

where

$$g_{A3P|y}(L, M, N) \equiv \Pr[X_{i_1++} = 1, X_{+i_2+} = 1, X_{++i_3} = 1|Y_{i_1i_2i_3} = 0],$$
$$h_{A3P|y}(L, M, N; N) \equiv \Pr[X_{i_1++} = 1, X_{+i_2+} = 1, \textstyle\sum_{j=1}^{N} X_{++j} = 0|Y_{i_1i_2i_3} = 0],$$
$$h_{A3P|y}(L, M, N; M) \equiv \Pr[X_{i_1++} = 1, \textstyle\sum_{j=1}^{M} X_{+j+} = 0, X_{++i_3} = 1|Y_{i_1i_2i_3} = 0],$$

and

$$h_{A3P|y}(L, M, N; L) \equiv \Pr[\textstyle\sum_{j=1}^{L} X_{j++} = 0, X_{+i_2+} = 1, X_{++i_3} = 1|Y_{i_1i_2i_3} = 0].$$

$g_{A3P|y}(L, M, N)$ can be expressed as

$$\begin{aligned} g_{A3P|y}(L, M, N) &= (1 - S_p)^3 q_{3b} + S_e(1 - S_p)^2(q_{2b} - 3q_{3b}) \\ &\quad + S_e^2(1 - S_p)(q_{1b} - 2q_{2b} + 3q_{3b}) + S_e^3(1 - q_{1b} + q_{2b} - q_{3b}), \end{aligned}$$

where $q_{1b} = q^{LM-1} + q^{LN-1} + q^{MN-1}$, $q_{2b} = q^{L(M+N-1)-1} + q^{M(L+N-1)-1} + q^{N(L+M-1)-1}$, and $q_{3b} = q^{LM+MN+LN-(L+M+N)}$.

To derive $h_{A3P|y}(L, M, N; N)$, we use the following decomposition.

$$\begin{aligned} h_{A3P|y}(L, M, N; N) = \sum_{c_1=0}^{1}\sum_{c_2=0}^{1}\sum_{r=0}^{N}\{&\Pr[Y_{i_1++} = c_1, Y_{+i_2+} = c_2, \sum_{j=1}^{N} Y_{++j} = r|Y_{i_1i_2i_3} = 0] \\ &\times \Pr[X_{i_1++} = 1, X_{+i_2+} = 1, \sum_{j=1}^{N} X_{++j} = 0|Y_{i_1++} = c_1, Y_{+i_2+} = c_2, \sum_{j=1}^{N} Y_{++j} = r, Y_{i_1i_2i_3} = 0]\}. \end{aligned}$$

Define

$$\beta_{00|y}(L,M,N;r) \equiv \Pr[Y_{i_1++}=0, Y_{+i_2+}=0, \sum_{j=1}^{N} Y_{++j}=r | Y_{i_1 i_2 i_3}=0] = \frac{\beta_{00}(L,M,N;r)}{q},$$

$$\beta_{01|y}(L,M,N;r) \equiv \Pr[Y_{i_1++}=0, Y_{+i_2+}=1, \sum_{j=1}^{N} Y_{++j}=r | Y_{i_1 i_2 i_3}=0] = \frac{\beta_{01}(L,M,N;r)}{q},$$

$$\beta_{10|y}(L,M,N;r) \equiv \Pr[Y_{i_1++}=1, Y_{+i_2+}=0, \sum_{j=1}^{N} Y_{++j}=r | Y_{i_1 i_2 i_3}=0] = \frac{\beta_{10}(L,M,N;r)}{q},$$

$$\begin{aligned}
&\beta_{11|y}(L,M,N;r) \equiv \Pr[Y_{i_1++}=1, Y_{+i_2+}=1, \sum_{j=1}^{N} Y_{++j}=r | Y_{i_1 i_2 i_3}=0] \\
&= \binom{N-1}{r}(q^{LMN-r(LM)-1})(1-q^{LM})^r \\
&+ \binom{N-1}{r-1}(q^{LMN-r(LM)})(1-q^{LM})^{r-1}(1-q^{LM-1}), \\
&- \{\beta_{00|y}(L,M,N;r) + \beta_{01|y}(L,M,N;r) + \beta_{10|y}(L,M,N;r)\}
\end{aligned}$$

where $\beta_{01|y}(L,M,N;0) \equiv 0$, $\beta_{10|y}(L,M,N;0) \equiv 0$, and $\beta_{11|y}(L,M,N;0) \equiv 0$. Then it follows that

$$h_{A3P|y}(L,M,N;N) = \sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{r=0}^{N} \gamma_{ij}(N;r)\beta_{ij|y}(L,M,N;r).$$

Similarly one can show

$$h_{A3P|y}(L,M,N;L) \;\; = \textstyle\sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{r=0}^{L} \gamma_{ij}(L;r)\beta_{ij|y}(M,N,L;r)$$

and

$$h_{A3P|y}(L,M,N;M) \;\; = \textstyle\sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{r=0}^{M} \gamma_{ij}(M;r)\beta_{ij|y}(L,N,M;r).$$

Pooling sensitivity can be derived as follows.

$$
\begin{aligned}
S_e(A3P) &= \Pr[X_{i_1i_2i_3}=1|Y_{i_1i_2i_3}=1] \\
&= \Pr[X_{i_1++}=1, X_{+i_2+}=1, X_{++i_3}=1, X_{i_1i_2i_3}=1|Y_{i_1i_2i_3}=1] \\
&+ \Pr[X_{i_1++}=1, X_{+i_2+}=1, \sum_{j=1}^{N} X_{++j}=0, X_{i_1i_2i_3}=1|Y_{i_1i_2i_3}=1] \\
&+ \Pr[X_{i_1++}=1, \sum_{j=1}^{M} X_{+j+}=0, X_{++i_3}=1, X_{i_1i_2i_3}=1|Y_{i_1i_2i_3}=1] \\
&+ \Pr[\sum_{j=1}^{L} X_{j++}=0, X_{+i_2+}=1, X_{++i_3}=1, X_{i_1i_2i_3}=1|Y_{i_1i_2i_3}=1] \\
&= S_e^4 + S_e^3\{\Pr[X_{++i_3}=0|Y_{i_1i_2i_3}=1] \prod_{k\neq i_3} \Pr[X_{++k}=0] \\
&+ \Pr[X_{+i_2+}=0|Y_{i_1i_2i_3}=1] \prod_{k\neq i_2} \Pr[X_{+k+}=0] \\
&+ \Pr[X_{i_1++}=0|Y_{i_1i_2i_3}=1] \prod_{k\neq i_1} \Pr[X_{k++}=0]\} \\
&= S_e^4 + S_e^3(1-S_e)\{\Pr[X_{++k}=0]^{N-1} + \Pr[X_{+k+}=0]^{M-1} + \Pr[X_{k++}=0]^{L-1}\} \\
&= S_e^4 + S_e^3(1-S_e)[\{1-f(LM)\}^{N-1} + \{1-f(LN)\}^{M-1} + \{1-f(MN)\}^{L-1}],
\end{aligned}
$$

where $f(n) \equiv (1-S_p)q^n + S_e(1-q^n)$.

4.4 Three-Dimensional Array With Master Pool

In this section we derive the efficiency and error rates of a three-dimensional array-based testing algorithms where we first test a master pool containing all LMN samples. If the master pool tests negative, the procedure stops. Otherwise, the procedure continues as in $A3P$. We denoted this algorithm by $A3PM(LMN : [L : M : N] : 1)$.

4.4.1 Efficiency

Let the number of tests be $T = T_0 + T_1 + T_2$ where $T_0 = 1$ corresponds to testing the master pool, T_1 corresponds to testing of planar slice pools, and T_2 corresponds to the possible subsequent individual testing. To compute the efficiency of $A3PM$, let X_0 be a random variable that equals 1 if master pool tests positive, and 0 otherwise such that $T_1 = (L + M + N)X_0$ and $E(T_1) = (L + M + N)f(LMN)$. Next write $T_2 = \sum_{i_1,i_2,i_3} T_{(i_1,i_2,i_3)}$ where

$$T_{(i_1,i_2,i_3)} = \begin{cases} 1 & \text{if } X_0 = 1, X_{i_1++} = 1 \text{ and } X_{+i_2+} = 1 \text{ and } X_{++i_3} = 1 \\ 1 & \text{if } X_0 = 1, X_{i_1++} = 1 \text{ and } X_{+i_2+} = 1 \text{ and } \sum_{i_3=1}^{N} X_{++i_3} = 0 \\ 1 & \text{if } X_0 = 1, X_{i_1++} = 1 \text{ and } \sum_{i_2=1}^{M} X_{+i_2+} = 0 \text{ and } X_{++i_3} = 1 \\ 1 & \text{if } X_0 = 1, \sum_{i_1=1}^{L} X_{i_1++} = 0 \text{ and } X_{+i_2+} = 1 \text{ and } X_{++i_3} = 1 \\ 0 & \text{otherwise} \end{cases}.$$

The expected number of tests per specimen for $A3PM$ is

$$E(A3PM) = \frac{1}{LMN} + \frac{L + M + N}{LMN} f(LMN) + E(T_{(i_1,i_2,i_3)}),$$

where

$$E(T_{(i_1,i_2,i_3)}) \quad = g_{A3PM}(L, M, N) + \sum_{j\in\{L,M,N\}} h_{A3PM}(L, M, N; j), \tag{4.4}$$

$$g_{A3PM}(L, M, N) \equiv \Pr[X_0 = 1, X_{i_1++} = 1, X_{+i_2+} = 1, X_{++i_3} = 1],$$
$$h_{A3PM}(L, M, N; N) \equiv \Pr[X_0 = 1, X_{i_1++} = 1, X_{+i_2+} = 1, \textstyle\sum_{i_3} X_{++i_3} = 0],$$
$$h_{A3PM}(L, M, N; M) \equiv \Pr[X_0 = 1, X_{i_1++} = 1, \textstyle\sum_{i_2} X_{+i_2+} = 0, X_{++i_3} = 1],$$

and

$$h_{A3PM}(L,M,N;L) \equiv \Pr[X_0 = 1, \textstyle\sum_{i_1} X_{i_1++} = 0, X_{+i_2+} = 1, X_{++i_3} = 1].$$

It is straightforward to show

$$g_{A3PM}(L,M,N) = (1-S_p)^3 q^{LMN}(1-S_p-S_e) + S_e g_{A3P}(L,M,N),$$
$$h_{A3PM}(L,M,N;N) = (1-S_p-S_e)\gamma_{00}(N;0)\beta_{00}(L,M,N;0) + S_e h_{A3P}(L,M,N;N),$$
$$h_{A3PM}(L,M,N;L) = (1-S_p-S_e)\gamma_{00}(L;0)\beta_{00}(M,N,L;0) + S_e h_{A3P}(L,M,N;L),$$

and

$$h_{A3PM}(L,M,N;M) = (1-S_p-S_e)\gamma_{00}(M;0)\beta_{00}(L,N,M;0) + S_e h_{A3P}(L,M,N;M).$$

4.4.2 Error Rates

In this section, we derive the pooling specificity and pooling sensitivity of $A3PM$.

To derive the pooling specificity of $A3PM$, let

$$\begin{aligned}
&1 - S_p(A3PM) = \Pr[X_{i_1i_2i_3} = 1 | Y_{i_1i_2i_3} = 0] \\
&= \Pr[X_0 = 1, X_{i_1++} = 1, X_{+i_2+} = 1, X_{++i_3} = 1, X_{i_1i_2i_3} = 1 | Y_{i_1i_2i_3} = 0] \\
&+ \Pr[X_0 = 1, X_{i_1++} = 1, X_{+i_2+} = 1, \textstyle\sum_{j=1}^{N} X_{++j} = 0, X_{i_1i_2i_3} = 1 | Y_{i_1i_2i_3} = 0] \\
&+ \Pr[X_0 = 1, X_{i_1++} = 1, \textstyle\sum_{j=1}^{M} X_{+j+} = 0, X_{++i_3} = 1, X_{i_1i_2i_3} = 1 | Y_{i_1i_2i_3} = 0] \\
&+ \Pr[X_0 = 1, \textstyle\sum_{j=1}^{L} X_{j++} = 0, X_{+i_2+} = 1, X_{++i_3} = 1, X_{i_1i_2i_3} = 1 | Y_{i_1i_2i_3} = 0].
\end{aligned} \tag{4.5}$$

Therefore by Assumption 5,

$$1 - S_p(A3PM) = (1-S_p)\{g_{A3PM|y}(L,M,N) + \textstyle\sum_{j\in\{L,M,N\}} h_{A3PM|y}(L,M,N;j)\},$$

where

$$g_{A3PM|y}(L,M,N) = \Pr[X_0 = 1, X_{i_1++} = 1, X_{+i_2+} = 1, X_{++i_3} = 1 | Y_{i_1 i_2 i_3} = 0],$$
$$h_{A3PM|y}(L,M,N;N) = \Pr[X_0 = 1, X_{i_1++} = 1, X_{+i_2+} = 1, \textstyle\sum_{j=1}^{N} X_{++j} = 0 | Y_{i_1 i_2 i_3} = 0],$$
$$h_{A3PM|y}(L,M,N;M) = \Pr[X_0 = 1, X_{i_1++} = 1, \textstyle\sum_{j=1}^{M} X_{+j+} = 0, X_{++i_3} = 1 | Y_{i_1 i_2 i_3} = 0],$$

and

$$h_{A3PM|y}(L,M,N;L) = \Pr[X_0 = 1, \textstyle\sum_{j=1}^{L} X_{j++} = 0, X_{+i_2+} = 1, X_{++i_3} = 1 | Y_{i_1 i_2 i_3} = 0].$$

One can show

$$g_{A3PM|y}(L,M,N) = (1-S_p)^3 q^{LMN-1}(1-S_p-S_e) + S_e g_{A3P|y}(L,M,N),$$
$$h_{A3PM|y}(L,M,N;N) = (1-S_p-S_e)\gamma_{00}(N;0)\beta_{00|y}(L,M,N;0) + S_e h_{A3P|y}(L,M,N;N),$$
$$h_{A3PM|y}(L,M,N;L) = (1-S_p-S_e)\gamma_{00}(L;0)\beta_{00|y}(M,N,L;0) + S_e h_{A3P|y}(L,M,N;L),$$

and

$$h_{A3PM|y}(L,M,N;M) = (1-S_p-S_e)\gamma_{00}(M;0)\beta_{00|y}(L,N,M;0) + S_e h_{A3P|y}(L,M,N;M).$$

Pooling sensitivity of $A3PM$ equals

$$S_e(A3PM) = \Pr[X_{i_1 i_2 i_3} = 1 | Y_{i_1 i_2 i_3} = 1] = \Pr[X_0 = 1, X_{i_1 i_2 i_3} = 1 | Y_{i_1 i_2 i_3} = 1] = S_e S_e(A3P).$$

4.5 $A3PM_2$: An alternative to $A3PM$

In this section, we consider an alternative algorithm to $A3PM$, which we denote as $A3PM_2$. The difference between $A3PM$ and $A3PM_2$ is the definition of the possible subsequent test of the $(i_1, i_2, i_3)^{th}$ sample.

For $A3PM_2$, $T_{(i_1,i_2,i_3)}$ is defined as

$$T_{(i_1,i_2,i_3)} = \begin{cases} 0 & \text{if } X_0 = 0 \text{ or } (\sum_{i_1} X_{i_1++} = 0 \text{ and } \sum_{i_2} X_{+i_2+} = 0 \text{ and } \sum_{i_3} X_{++i_3} = 0) \\ 1 & \text{otherwise} \end{cases} .$$

In other words, the $(i_1, i_2, i_3)^{th}$ sample is tested individually if the master pool tests positive and at least one of the planar slices containing that specimen tests positive. Intuitively, $A3PM_2$ would be expected to be less efficient and specific but more sensitive than $A3PM$, since individual specimens will be tested more often under $A3PM_2$ than $A3PM$. Derivations of efficiency and error rates of $A3PM_2$ are not shown in this paper, but the corresponding R programs are available from the authors.

4.6 Simulation Results

We conducted a simulation study to confirm the derived efficiencies and error rates of $A3P$ and $A3PM$. Number of simulation was 10,000,000. For each simulation, we assumed LMN samples to be tested using $A3P([L : M : N] : 1)$ (or $A3PM(LMN : [L : M : N] : 1)$). The $L \times M \times N$ array Y was simulated by generating independent Bernoulli random deviates with success probability p for each entry in the array. For $i_1 = 1, \cdots, L$, Y_{i_1++} was created from Y by taking the maximum of the MN elements of the corresponding planar slice. Test outcomes of each pool were then generated by letting,

$$X_{i_1++} = Y_{i_1++}\delta_{i_1} + (1 - Y_{i_1++})\gamma_{i_1},$$

where $\delta_{i_1} \sim Bernoulli(S_e)$ and $\gamma_{i_1} \sim Bernoulli(1-S_p)$. Y_{+i_2+}, Y_{++i_3}, X_{+i_2+}, and X_{++i_3} were created similarly. The observed efficiency and pooling error rates were computed from the simulated data.

Figure 4.2 summarizes the first set of simulation results of $A3P$ with $L = M =$

$N = 3$, $S_e = 0.88$ and $S_p = 0.92$. A second set of simulations of $A3P$ with $L = M = 3$, $N = 4$, $S_e = 0.90$ and $S_p = 0.90$ and a third set of simulations of $A3P$ with $L = 4$, $M = N = 5$, $S_e = 0.95$ and $S_p = 0.80$ yielded similar results (not shown). We also conducted three sets of simulation study of $A3PM$ with same values as $A3P$ and the derived results were close to the simulated results in all cases (results not shown). These results demonstrate excellent agreement between the derived operating characteristics and those observed from simulated data.

4.7 Application

Using the results derived above, in this section we compare the operating characteristics of three-dimensional array-based algorithms with two-dimensional array and hierarchical algorithms for identification of acute HIV using nucleic acid amplification tests (NAATs) in conjunction with specimen pooling. A brief description of the different algorithms considered is given in Table 4.1.

As a first motivating example, we consider a setting similar to the NC STAT program. We assume prevalence of acute HIV is $p = 0.0002$ (Pilcher et al., 2005) and NAAT has a 99% test specificity (Hecht et al., 2002) and 90% test sensitivity. We also assume the master pool size is less than or equal to 100 due to dilution effects (Quinn et al., 2000). For each algorithm in Table 4.1, the optimal configuration was selected that minimizes the expected number of tests per specimen. For instance, the optimal configurations of $A3PM$ and $A3PM_2$ were determined by computing the efficiency for all possible positive integers (L, M, N) such that $8 \leq L \times M \times N \leq 100$. For both $A3PM$ and $A3PM_2$ the most efficient configuration is $(L, M, N) = (4, 5, 5)$. Similarly, $D3(100 : 10 : 1)$ and $A2M(100 : 10 : 1)$ were determined to be the optimal configurations of $D3(N^2 : N : 1)$ and $A2M(N^2 : N : 1)$ for $2 \leq N \leq 100$. Table 4.2 shows

the operating characteristics of the optimal configurations of each algorithm as well as $D3(90:10:1)$, the algorithm employed by Pilcher et al. (2005). These results suggest moving from $D3(90:10:1)$ to $A3PM(100:[4,5,5]:1)$ or $A3PM_2(100:[4,5,5]:1)$ would improve efficiency, pooling specificity, sensitivity, PPV and NPV of the NC STAT HIV detection program.

For our second example, we consider a setting similar to that described by Pilcher et al. (2004), who employed $D3(50:10:1)$ to identify acute HIV in Malawi. They found 4.5% of antibody negative males attending STD clinics to be NAAT positive. Assuming $S_p = 0.99$ and $S_e = 0.9$ as before and $p = 0.045$, the expected number of tests per specimen for $D3(50:10:1)$ is 0.40. Table 4.3 shows the most efficient configuration for each algorithm in Table 4.1 given the master pool size can be no larger than 50. For example, using $A3PM(48:[4,4,3]:1)$ results in 0.33 tests per specimen on average while $D3(16:4:1)$ results in 0.32 expected tests per specimen. Based on the results in Table 4.3, $D3(16:4:1)$ appears to yield the best balance of efficiency and error rates while at the same time being least susceptible to dilution effects.

Figure 4.3 shows the operating characteristics of the optimal configurations of $D3$, $A2M$, and $A3PM$ for prevalences ranging from 10^{-5} to 10^{-1}. For this example, we assume $S_e = 0.9$, $S_p = 0.9$, and the maximum allowable pool size is 100. Under these assumptions, for each prevalence we found the values of n that minimize $E\{D3(n:\sqrt{n}:1)\}$, $E\{A2M(n:\sqrt{n}:1)\}$, and $E\{A3PM(n:[L,M,N]:1)\}$, where $n = LMN$ for $A3PM$. The expected numbers of tests per specimen, optimal master pool size, pooling sensitivities, specificities, PPVs, and NPVs at these optimal values of n are depicted in Figure 4.3. These results indicate $A3PM$ is generally the preferable algorithm with regards to efficiency and PPV for prevalence less than 0.001.

4.8 Discussion

We derived several operating characteristics of three-dimensional array-based testing algorithms for case identification in the presence of testing error. Using these results, we showed that moving the NC STAT algorithm, from $D3(90:10:1)$ to $A3PM(100:[4,5,5]:1)$ or $A3PM_2(100:[4,5,5]:1)$ would improve efficiency, pooling specificity, sensitivity, PPV and NPV. For the Malawi example, $D3(16:4:1)$ is most efficient and sensitive. This result shows that moving from $D3(50:10:1)$ to $D3(16:4:1)$ in the Malawi setting would improve efficiency and pooling error rates. It indicates that the choice of optimal algorithm will be context specific and that no single algorithm can be universally recommended.

There are several areas of potential future research related to this work. First, we have considered three-dimensional array-based algorithms. The proposed methods could be extended to higher dimensional arrays in the presence of test errors. Second, we derived the operating characteristics of a three-dimensional planar algorithm. However, Berger et al. (2000) also proposed a multi-dimensional linear method where pools are formed along lines instead of planar slices. The higher dimensional linear array-based algorithms could also be generalized to allow for test error.

TABLE 4.1: Description of group testing algorithms.

Algorithm	*Description*
DS	S stage hierarchical algorithm
$A2$	Two dimensional square array without master pool
$A2M$	Two dimensional square array with master pool
$A3P$	Three dimensional planar array without master pool
$A3PM$	Three dimensional planar array with master pool
$A3PM_2$	Three dimensional planar array with master pool: Definition of the possible subsequent test of the $(i_1, i_2, i_3)^{th}$ sample is different from the $A3PM$

TABLE 4.2: Comparison of operating characteristics for the most efficient $D3$, $A2M$, $A3PM$, and $A3PM_2$ to be used in NC STAT for detection of acute HIV assuming pool size less than 100, test sensitivity $S_e = 0.9$, test specificity $S_p = 0.99$, and prevalence $p = 0.0002$.

$\mathcal{A}$	$E(\mathcal{A})$	$S_p(\mathcal{A})$	$S_e(\mathcal{A})$	$PPV(\mathcal{A})$	$NPV(\mathcal{A})$
$D3(90:10:1)$*	0.016	0.999983	0.729	0.896	0.999946
$D3(100:10:1)$	0.015	0.999983	0.729	0.895	0.999946
$A2M(100:10:1)$	0.016	0.999995	0.787	0.969	0.999957
$A3P([4,5,5]:1)$	0.014	0.999995	0.864	0.972	0.999973
$A3PM(100:[4,5,5]:1)$	0.014	0.999998	0.778	0.988	0.999956
$A3PM_2(100:[4,5,5]:1)$	0.015	0.999996	0.797	0.975	0.999960

* $D3(90:10:1)$ is the algorithm used in Pilcher et al. (2002).

TABLE 4.3: Comparison of operating characteristics for the most efficient $D3$, $A2M$, $A3PM$, and $A3PM_2$ to be used in Malawi for detection of acute HIV assuming pool size less than 50, test sensitivity $S_e = 0.9$, test specificity $S_p = 0.99$, and prevalence $p = 0.045$.

$\mathcal{A}$	$E(\mathcal{A})$	$S_p(\mathcal{A})$	$S_e(\mathcal{A})$	$PPV(\mathcal{A})$	$NPV(\mathcal{A})$
$D3(50:10:1)$*	0.40	0.9972	0.729	0.925	0.987
$D3(16:4:1)$	0.32	0.9989	0.729	0.970	0.987
$A2M(49:7:1)$	0.34	0.9995	0.681	0.984	0.985
$A3P([4,4,3]:1)$	0.36	0.9990	0.710	0.970	0.987
$A3PM(48:[4,4,3]:1)$	0.33	0.9991	0.639	0.970	0.983
$A3PM_2(48:[4,4,3]:1)$	0.33	0.9991	0.640	0.970	0.983

* $D3(50:10:1)$ is the algorithm used in Pilcher et al. (2004).

FIGURE 4.1: Three dimensional planar algorithm ($A3P$) with $L = M = N = 3$. The black dots denote 9 ($= MN$) specimens in one of L planar slices.

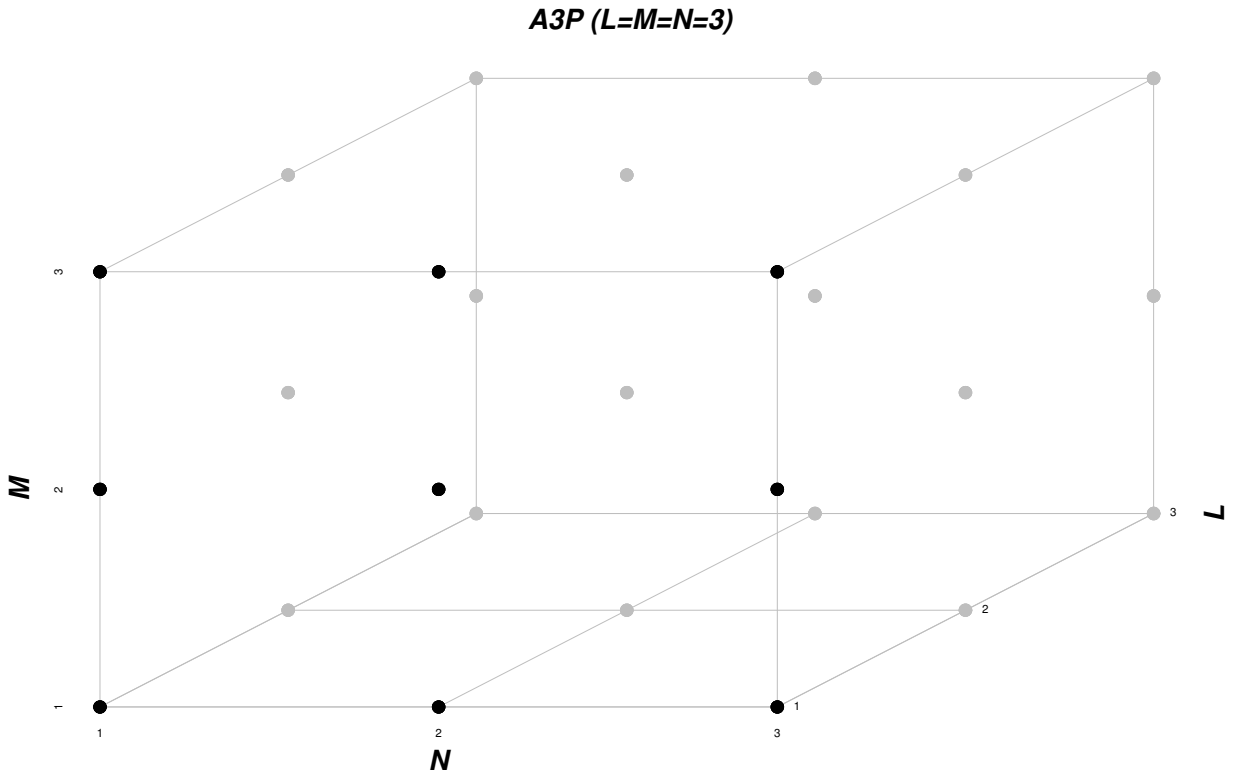


FIGURE 4.2: Results of simulation study. (a) Expected number of tests per specimen, (b) pooling specificity, (c) pooling sensitivity for $A3P$ assuming $L = M = N = 3$, test sensitivity $S_e = 0.88$, test specificity $S_p = 0.92$, and prevalence $p \in \{0.0001, 0.001, 0.01, 0.1\}$.

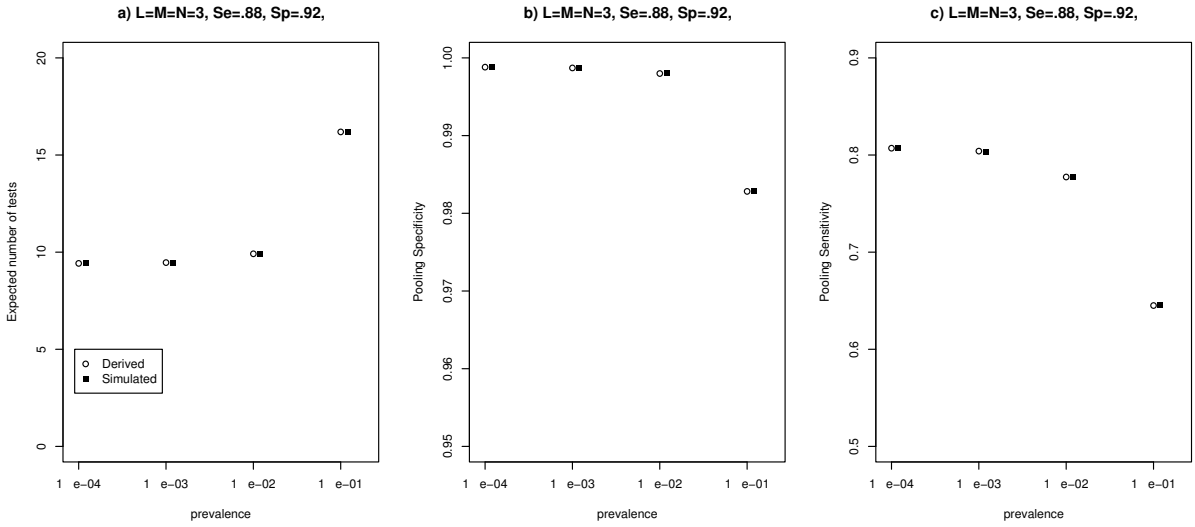
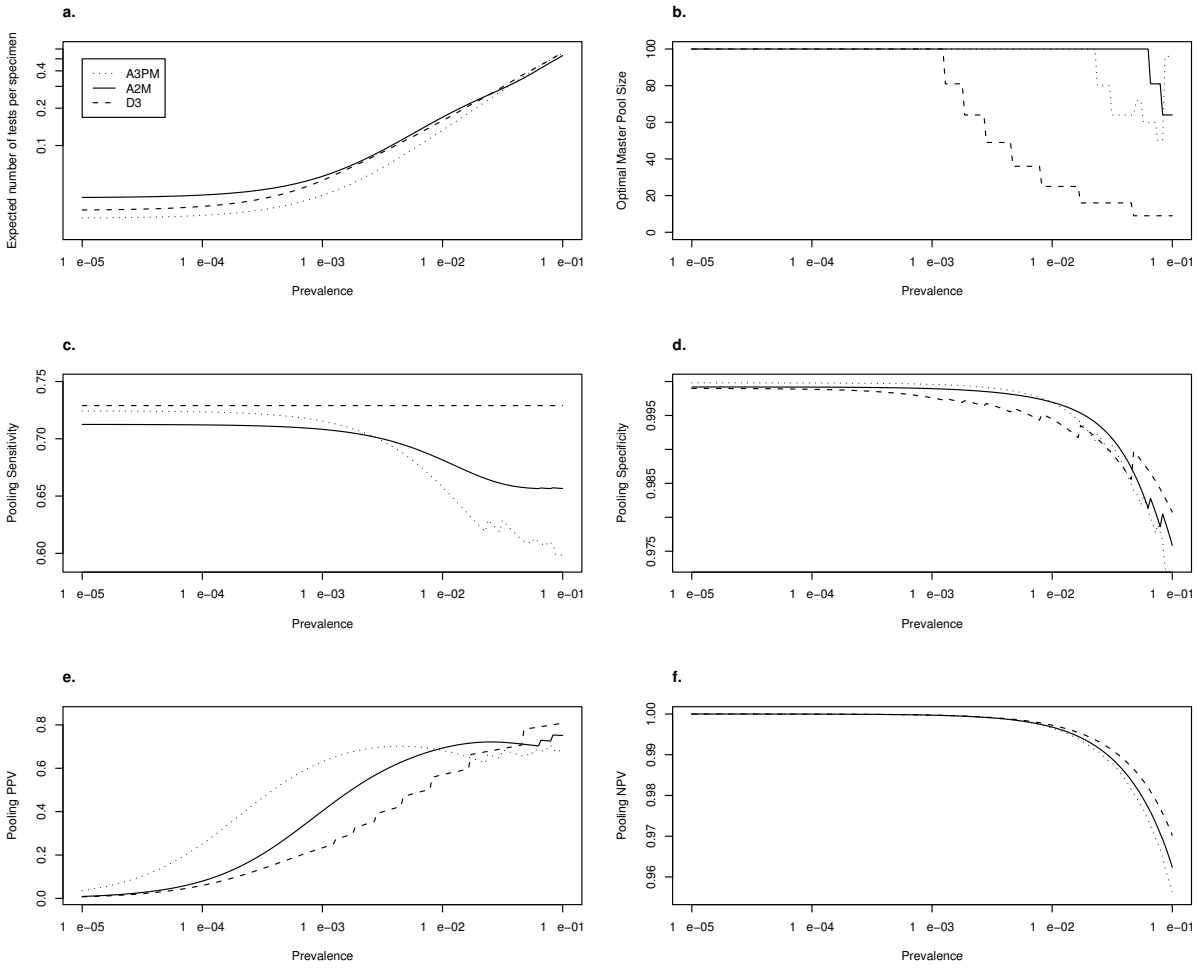


FIGURE 4.3: (a) Expected number of tests per specimen, (b) optimal master pool size, (c) pooling sensitivity, (d) pooling specificity, (e) pooling PPV and (f) pooling NPV for optimally efficient configurations of $D3$, $A2M$, and $A3PM$ assuming test sensitivity $S_e = 0.9$, test specificity $S_p = 0.9$, and a maximum allowable pool size of 100.



CHAPTER 5

OPTIMAL CONFIGURATION OF A SQUARE ARRAY GROUP TESTING ALGORITHM

5.1 Introduction

Pooling of blood samples as a cost saving method was first used for screening of syphilis in 1943 by Dorfman's two-stage hierarchical pooling algorithm (Dorfman, 1943). Subsequently, multistage hierarchical (Finucan 1964; Johnson et al. 1991; Litvak et al. 1994) and array-based (Phatarfod and Sudbury 1994; Berger et al. 2000) pooling algorithms were proposed to detect all individuals having the disease of interest. In order to minimize the expected number of tests per specimen, it is necessary to know the optimal configuration of these algorithms.

Several researchers have derived the optimal pool size of two-stage hierarchical pooling algorithms to minimize the expected number of tests per specimen (Feller, 1957; Wilks, 1962; Samuels, 1978; Turner et al., 1988). Finucan (1964) and Wu and Zhao (1994) derived the optimal number of stages and the optimal pool size at each stage for

multistage hierarchical pooling algorithms. Wu and Zhao (1994) also considered the presence of test errors when they derived the optimal pool size and number of stages for hierarchical pooling algorithm. To our knowledge, no analogous research has been done on the optimal configuration of array-based algorithms.

The outline of this chapter is as follows. In Section 5.2, we summarize Turner et al. (1988), Samuels (1978) and Wu and Zhao (1994)'s results on the optimal pool size of hierarchical algorithms. In Section 5.3, we derive the optimal pool size of a two-dimensional array-based algorithm for a given prevalence. We conclude with a short discussion in Section 5.4.

5.2 Hierarchical algorithms

5.2.1 Turner et al (1988)

Turner et al. (1988) proposed a calculus based approach to determine the optimal pool size of a two-stage hierarchical algorithm assuming no test error. They considered minimizing the expected number of tests per specimen of algorithm $D2$ without testing error. Mathematically, this problem can be stated as follows: For a given real number $p \in (0, 1)$, find the positive integer n that minimizes the function

$$f_{D2}(q, n) = \begin{cases} 1 & \text{if } n = 1 \\ \frac{1}{n} + 1 - q^n & \text{if } n > 1 \end{cases} \tag{5.1}$$

where $q = 1 - p$. For simplicity, Turner et al. (1988) considered the related problem of finding the positive integer n that minimizes $g_{D2}(q, n) = \frac{1}{n} - q^n$. Turner et al. (1988) proved three theorems concerning the function g_{D2}:

Theorem The equation $g_{D2}(q, n) = 0$ has no solution if $q < c$, one solution if $q = c$, and two solutions if $q > c$, where $c = \frac{1}{e}^{\frac{1}{e}}$.

Theorem The equation $g'_{D2}(q,n) = \frac{\partial g_{D2}(q,n)}{\partial n} = 0$ has no solution if $q < b$, one solution if $q = b$, and two solutions if $q > b$, where $b = e^{-4e^{-2}}$.

Theorem If $q < \frac{1}{3}^{\frac{1}{3}}$, there is no positive integer n for which $g_{D2}(q,n) < 0$.

Based on these three theorems, they made the following conclusions:

- Individual testing is more efficient than $D2$ if $0 < q < \frac{1}{3}^{\frac{1}{3}}$.
- $g_{D2}(q,3) = 0$ and $g_{D2}(q,n) > 0$ for any positive integer $n \neq 3$ if $q = \frac{1}{3}^{\frac{1}{3}}$. Thus, f is minimized for either $n = 3$ or $n = 1$ when $q = \frac{1}{3}^{\frac{1}{3}}$.
- g_{D2} has local minimum near $p^{-\frac{1}{2}}$ if $q > \frac{1}{3}^{\frac{1}{3}} \approx 0.693361$.

5.2.2 Samuels (1978)

Samuels (1978) also derived optimal pool size of $D2$. In contrast to Turner et al. (1988), Samuels' approach relied on studying the difference between the expected number of tests per specimen for two successive group sizes. He also showed that for all $p < 1 - \frac{1}{3}^{\frac{1}{3}}$, the optimal group size is either $1 + \lfloor p^{-1/2} \rfloor$ or $2 + \lfloor p^{-1/2} \rfloor$, where $\lfloor x \rfloor$ denotes the integer part of x, and, when the optimal group size is used, the expected number of tests per specimen is between $2p^{1/2} - p/2$ and $2p^{1/2} + 4p^{3/2}$. He also showed that for every value of p, $f_{D2}(q,3) < f_{D2}(q,2)$, i.e., pools of size 2 are never optimal.

5.2.3 Wu and Zhao (1994)

Wu and Zhao (1994) extended the results of Turner et al. (1988) and Samuels (1978) to multistage hierarchical algorithms with test error. Using an approach similar to Samuels, Wu and Zhao provided a simple procedure to determine the optimal number of stages, number of sub-pools per stage, and pool size for each sub-pool.

5.3 Square array-based algorithm

In this section, we derive the optimal pool size of a two-stage array-based algorithm. In particular, consider the $n \times n$ square array set-up of Phatarfod and Sudbury (1994), denoted by $A2$, where n^2 specimen are placed on an $n \times n$ matrix. Pools are then made from all samples in the same row or in the same column. These $2n$ pools (n row pools and n column pools) are then tested and, assuming no test error, all positive specimens will lie at the intersection of a positive row pool and a positive column pool. Therefore all specimens at the intersection of a positive row and a positive column are subsequently tested.

Denote the expected number of tests per specimen of algorithm $A2$ by

$$f(q,n) \equiv \begin{cases} 1 & \text{if } n = 1 \\ \frac{2}{n} + 1 - 2q^n + q^{2n-1} & \text{if } n > 1 \end{cases} \tag{5.2}$$

where $q = 1 - p$, and prevalence rate, $p \in (0,1)$. The problem is to find the positive integer n that minimizes $f(q,n)$ for given real number $q \in (0,1)$. Moreover, we would like to find $q*$ and $n*$ such that:

[1] If $0 < q < q*$, then $f(q,n) > 1$ for all integers $n > 1$, indicating individual testing is more efficient than $A2$.

[2] If $q > q*$, then $A2$ is more efficient than individual testing and $f(q,n)$ has a global minimum at the positive integer $n*$ for fixed q.

5.3.1 Lower and Upper bounds for $q*$

In this section, we derive lower and upper bounds of $q*$ using a calculus approach similar to Turner et al. (1988). These bounds are then used in Section 5.3.2 to determine $q*$.

The problem of interest is to find smallest $q \in (0,1)$, say $q*$, where $A2$ is more effi-

cient than individual testing, i.e., find minimum value of $q \in (0,1)$, such that $g(q,n) < 0$ for some integer $n \geq 2$, where

$$g(q,n) \equiv f(q,n) - 1 = \frac{2}{n} - 2q^n + q^{2n-1}.$$

Lemma 1 *For given integer $n \geq 3$, there exists a unique $q \in (0,1)$ such that $g(q,n) = 0$.*

Proof: The lemma follows by noting that $g(0,n) = \frac{2}{n} > 0$, $g(1,n) = \frac{2}{n} - 1 < 0$, and $\frac{\partial g(q,n)}{\partial q} = q^{n-1}\{(2n-1)q^{n-1} - 2n\} < 0$ for $n \geq 3$. ■

Lemma 2 *$n = 2$ is never optimal, i.e., $n* \neq 2$ for all $q \in (0,1)$.*

Proof: The lemma follows by noting that $g(1,2) = 0$ and $\frac{\partial g(q,2)}{\partial q} < 0$ for $q \in (0,1)$. In other words, $f(q,2) > 1$ for all $q \in (0,1)$, then a 2×2 square array is never more efficient than individual testing. ■

Lemma 3 *For a given integer $n \geq 3$, $q_L(n) = \{1 - \sqrt{1 - \frac{2}{n}}\}^{\frac{1}{n}}$ is a lower bound to the solution of $g(q,n) = 0$.*

Proof: We know $q^{2n-1} > q^{2n}$, therefore $g(q,n) > \frac{2}{n} - 2q^n + q^{2n}$. Next, solve $\frac{2}{n} - 2q^n + q^{2n} = 0$. By the quadratic formula, the positive solution is $q^n = \frac{-b-\sqrt{b^2-4ac}}{2a}$, where $a = 1, b = -2$, and $c = \frac{2}{n}$. Therefore, $q_L(n)^{2n} - 2q_L(n)^n + \frac{2}{n} = 0$, where $q_L(n) = \{1 - \sqrt{1 - \frac{2}{n}}\}^{\frac{1}{n}}$. ■

Lemma 4 *$\{q_L(n) : n = 3, 4, \cdots\}$ has a minimum value at $q_L(4)$.*

Proof: By direct evaluation, $q_L(3) > q_L(4) < q_L(5) < q_L(6)$, i.e., $q_L(4)$ is a local minimum. Below we show $q_L(4)$ is the global minimum by showing $\frac{\partial q_L(n)}{\partial n} > 0$ for $n \geq 6$. Let $l = \log q_L(n)$, $a(n) = \sqrt{1 - \frac{2}{n}}$ such that

$$\frac{\partial l}{\partial n} = -\frac{1}{n^2}\log(1 - a(n)) - \frac{1}{n^3}\frac{1}{(1 - a(n))a(n)}.$$

Then $\frac{\partial l}{\partial n} > 0$ if and only if

$$-n\log(1-a(n)) > \frac{1}{(1-a(n))a(n)}. \tag{5.3}$$

Below we show (5.3) holds for $n \geq 6$. First, we use the fact that

$$-\tfrac{3}{2}x > \log(1-x) \tag{5.4}$$

for $x > 0.5828$ (Wolfram, 1998). For $n \geq 6$, $a(n) \geq \sqrt{\frac{2}{3}} = 0.8165$. Therefore, by (5.4),

$$-n\log(1-a(n)) > \frac{3}{2}a(n)n.$$

Thus it is sufficient to show

$$\frac{3}{2}a(n)n > \frac{1}{(1-a(n))a(n)} \tag{5.5}$$

for $n \geq 6$. Noting that $1-a(n) = \frac{2/n}{1+\sqrt{1-2/n}}$, equation (5.5) is equivalent to

$$6 - \tfrac{12}{n} > 2(1+\sqrt{1-\tfrac{2}{n}}) \tag{5.6}$$

for $n \geq 6$, which holds since $6 - \frac{12}{n} \geq 4$ and $2(1+\sqrt{1-\frac{2}{n}}) < 4$ for $n \geq 6$. Thus $\frac{\partial l}{\partial n} > 0$ for $n \geq 6$. Therefore, $\frac{\partial q_L(n)}{\partial n} = q_L(n)\frac{\partial l}{\partial n} > 0$ for $n \geq 6$, since $q_L(n)$ is always greater than 0. Therefore, $\min\{q_L(n) : n = 3, 4, \cdots\} = q_L(4)$. ■

Lemma 5 *For a given n, $q_U(n) = \{\frac{1-\sqrt{1-\frac{4}{n}}}{2}\}^{\frac{1}{n}}$ is an upper bound to the solution of $g(q,n) = 0$.*

Proof: By Lemmas 3 and 4, a lower bound of $q*$ is $q_L(4) = \{1-\sqrt{1-\frac{1}{2}}\}^{\frac{1}{4}} \approx 0.7357$. Therefore, the range of an upper bound of $q*$ is $0.7357 < q < 1$. We know $q^{2n-1} < 2q^{2n}$

for $0.7357 < q < 1$, since $q^{2n-1}(1-2q) < 0$ for $0.5 < q < 1$. Therefore $g(q,n) < \frac{2}{n} - 2q^n + 2q^{2n}$ for $0.7357 < q < 1$. Next, solve $\frac{2}{n} - 2q^n + 2q^{2n} = 0$. By the quadratic formula, the positive solution is $q^n = \frac{-b-\sqrt{b^2-4ac}}{2a}$, where $a = 2, b = -2$, and $c = \frac{2}{n}$. Therefore, $2q_U(n)^{2n} - 2q_U(n)^n + \frac{2}{n} = 0$, where $q_U(n) = \{\frac{1-\sqrt{1-\frac{4}{n}}}{2}\}^{\frac{1}{n}}$. ■

Lemma 6 *$\{q_U(n) : n = 3, 4, \cdots\}$ has a minimum value at $q_U(6)$*

Proof: The proof parallels the proof of Lemma 4. ■

Let $q_L \equiv q_L(4) = \{1 - \sqrt{1 - \frac{1}{2}}\}^{\frac{1}{4}} \approx 0.7357$ and $q_U \equiv q_U(6) = \{\frac{1-\sqrt{1-\frac{4}{6}}}{2}\}^{\frac{1}{6}} \approx 0.7718$. From Lemma 3 - 6 it follows that $q_L < q* < q_U$. Figure 5.1 depicts $g(q,n)$ for $q = q_L$, $q = q_U$, and $q = q*$, with the value of $q*$ based on the results in the section below.

5.3.2 Determining $q*$

In this section, we determine $q*$ by studying the difference between the expected number of tests per specimen for two successive group size. First, we define

$$\Delta(q,n) = f(q,n+1) - f(q,n) = 2q^n(1-q) - q^{2n-1}(1-q^2) - \frac{2}{n(n+1)} \tag{5.7}$$

for $n \geq 2$. The purpose of $\Delta(q,n)$ is to compare the efficiencies (the expected numbers of tests per specimen) at $n+1$ and n for a given q. If we can determine n such that $\Delta(q,n-1) < 0$ and $\Delta(q,n) > 0$, then we know n is a local minimum of $f(q,n)$. Differentiating $\Delta(q,n)$ with respect to q

$$\begin{aligned} \frac{\partial\Delta(q,n)}{\partial q} &= 2nq^{n-1}(1-q) - 2q^n - (2n-1)q^{2n-2}(1-q^2) + 2q^{2n} \\ &= q^{n-1}\{2n - 2(n+1)q - (2n-1)q^{n-1} + (2n+1)q^{n+1}\}, \end{aligned} \tag{5.8}$$

and setting $\frac{\partial\Delta(q,n)}{\partial q} = 0$ yields

$$2n - 2(n+1)q - (2n-1)q^{n-1} + (2n+1)q^{n+1} = 0. \tag{5.9}$$

We use (5.9) to prove that $n = 3$ is never optimal.

Lemma 7 *$n = 3$ is never optimal, i.e., $n* \neq 3$ for all $q \in (0, 1)$.*

Proof: If $n = 3$, then (5.9) equals

$$6 - 8q - 5q^2 + 7q^4 = (q - 1)(7q^3 + 7q^2 + 2q - 6) = 0 \tag{5.10}$$

which has two real solutions, $q = 1$ and $q \approx 0.6409$ (by cubic formula). Equation (5.10) is positive for $q \in (0, 0.6409)$ and negative for $q \in (0.6409, 1)$. Therefore, $\Delta(q, 3)$ has a maximum at $q = 0.6409$ for $0 < q < 1$. Since $\Delta(0.6409, 3) < 0$, it follows that $f(q, 4) < f(q, 3)$ for all $q \in (0, 1)$. That is, 4×4 square arrays are always more efficient than 3×3 arrays. ■

Figures 5.2 and 5.3 show $\Delta(q, n)$ of $A2$ for $0 < q < 1$. That $n = 2$ and $n = 3$ are never optimal can be seen in Figure 5.2.

Lemma 8 *For fixed integer n, $\Delta(q, n)$ is unimodal, the roots $r_1(n)$ and $r_2(n)$ of $\Delta(q, n) = 0$ exist if $n \geq 4$ and these two roots satisfy $r_1(n) < q_{max,n} < r_2(n)$, where $q_{max,n}$ is the value of q that maximizes $\Delta(q, n)$ for given n.*

Proof: From equation (5.8) it follows that $\frac{\partial \Delta(q,n)}{\partial q} = 0$ has one solution at $q = 0$. Other solutions must satisfy

$$e(q) \equiv 2n - 2(n + 1)q - (2n - 1)q^{n-1} + (2n + 1)q^{n+1} = 0. \tag{5.11}$$

Now $e(q)$ has one root at $q = 1$, i.e., $e(1) = 0$. We also know $e(0) = 2n$ and

$$e'(q) = \frac{\partial e(q)}{\partial q} = -2(n + 1)q - (2n - 1)(n - 1)q^{n-2} + (2n + 1)(n + 1)q^n. \tag{5.12}$$

In particular, $e'(1) = 2(2n-1) > 0$. Thus $e(q)$ is positive at $q = 0$, equals zero at $q = 1$ and is increasing at $q = 1$, implying there exists at least one $q \in (0,1)$ such that $e(q) = 0$. Below, we show that there is exactly one such q. First, note that

$$e''(q) = \frac{\partial^2 e(q)}{\partial q^2} = -(2n-1)(n-1)(n-2)q^{n-3} + (2n+1)(n+1)nq^{n-1} > 0, \qquad (5.13)$$

if and only if

$$q > \sqrt{\frac{(2n-1)(n-1)(n-2)}{(2n+1)(n+1)n}} = c.$$

Thus $e(q)$ is strictly concave for $q < c$ and strictly convex for $q > c$. Since $e'(0) < 0$ and $e'(1) > 0$, this implies $e(q) = 0$ has only one solution for $q \in (0,1)$. Therefore, for fixed n, $\Delta(q,n)$ is unimodal. We also know $\Delta(0,n) < 0$ and $\Delta(1,n) < 0$. It can also be shown that $\Delta(q = \frac{n}{n+2}, n) > 0$ for $n \geq 5$, implying an existence of $q_{max,n}$. Therefore, $\Delta(q,n)$ has a maximum at $q_{max,n}$ and two roots of $\Delta(q,n) = 0$ exist for $0 < q < 1$. ■

Lemma 9 *$r_2(n)$ of A2 is increasing for $n \geq 4$.*

Proof: If $\Delta(q,n) = 0$, then $2q^n(1-q) = q^{2n-1}(1-q^2) + \frac{2}{n(n+1)}$ and

$$\Delta(q,n+1) = q^{2n}(1-q^2)(1-q) + \frac{2(q - \frac{n}{(n+2)})}{n(n+1)}. \qquad (5.14)$$

By using (5.8), we can show that $\Delta'(q = \frac{n}{n+2}, n) > 0$, implying $\frac{n}{n+2} < r_2(n)$ for $n \geq 4$. Therefore, $\Delta(q = r_2(n), n+1) > 0$, which implies $r_2(n)$ is increasing for $n \geq 4$. ■

Lemma 10 *$r_1(n)$ of A2 is increasing for $n \geq 5$.*

Proof: From Lemma 9, we know $\frac{n}{n+2} < r_2(n)$. First note

$$\Delta(q = \tfrac{n}{n+2}, n) = \tfrac{2}{n(n+1)}[(\tfrac{n}{n+2})^{n+1}2(n+1)\{1 - (\tfrac{n}{n+2})^n(\tfrac{n+1}{n})\} - 1]. \qquad (5.15)$$

It can be shown that $\Delta(q = \frac{n}{n+2}, n) > 0$ for $n \geq 5$, which implies $r_1(n) < \frac{n}{n+2}$. Therefore, $\Delta(q = r_1(n), n+1) < 0$ by (5.14), indicating $r_1(n)$ is increasing for $n \geq 5$. ■

In the lemma below, for a fixed q, we say $f(q,n)$ is a decreasing function of n if $\Delta(q,n) < 0$, i.e., $f(q,n) > f(q,n+1)$. Likewise, we say $f(q,n)$ is an increasing function of n if $\Delta(q,n) > 0$.

Lemma 11 *For a given q, let u be the smallest integer $n \geq 4$ for which $r_2(n) > q$, and v be the smallest integer $n \geq 4$ for which $r_1(n) > q$. (A) If $q > r_2(4)$, then $f(q,n)$ is decreasing for $4 \leq n \leq u$, increasing for $u \leq n \leq v$, and decreasing for $n \geq v$. (B) If $r_1(4) < q < r_2(4)$, then $f(q,n)$ is increasing for $4 \leq n \leq v$, and decreasing for $n \geq v$. (C) $r_1(5) < q < r_1(4)$, then $f(q,n)$ is decreasing for $n = 4$, increasing for $n = 5$ and decreasing for $n \geq 6$. (D) If $q < r_1(5)$, then $f(q,n)$ is decreasing for $n \geq 4$.*

Proof: Lemma 9 and Lemma 10 prove the following results. (A) If $r_2(4) < q$, then $r_2(4) < q < r_2(u) \leq r_1(v)$ or $r_2(4) < q < r_1(v) \leq r_2(u)$ by definition of u and v. Therefore, $\Delta(q,n) < 0$ for $4 \leq n \leq u$ and $\Delta(q,n) > 0$ for $u \leq n \leq v$ and $\Delta(q,n) < 0$ for $n \geq v$. (B) If $r_1(4) < q < r_2(4)$, then $r_1(4) < q < r_1(v) \leq r_2(4)$ or $r_1(4) < q < r_2(4) \leq r_1(v)$ by definition of u and v. Therefore, $\Delta(q,n) > 0$ for $4 \leq n \leq v$ and $\Delta(q,n) < 0$ for $n \geq v$. (C) If $r_1(5) < q < r_1(4)$, then $\Delta(q,n) < 0$ for $n = 4$, $\Delta(q,n) > 0$ for $n = 5$ and $\Delta(q,n) < 0$ for $n \geq 6$. (D) If $q < r_1(5)$, then $\Delta(q,n) < 0$ for $n \geq 4$. ■

Lemma 12 $r_2(4) < q_L < q* < q_U < r_2(5)$.

Proof: It can be shown that $\Delta(q = q_L, 4) < 0$ and $\Delta(q = q_U, 5) > 0$, implying $r_2(4) < q_L$ and $q_U < r_2(5)$. By Lemma 3 - Lemma 6, we know $q_L < q* < q_U$. Therefore, $r_2(4) < q_L < q* < q_U < r_2(5)$. ■

Lemma 13 *$n = 4$ is never optimal, i.e., $n* \neq 4$ for all $q \in (0, 1)$.*

Proof: By Lemma 12, $\Delta(q, 4) < 0$ for all $q \geq q*$, implying 5×5 square arrays are more efficient than 4×4. If $q < q*$, then individual testing is more efficient than $A2$ for any configuration. Therefore, 4×4 square arrays are never optimal. ■

Lemma 14 *Table 5.1 is valid.*

Proof: Table 5.1 follows immediately from Lemma 11 and Lemma 12. ■

The Contraction Mapping Theorem is a useful fixed point theorem. We apply this theorem to prove Theorem 1 below.

Contraction Mapping Theorem (Devaney 1992; Devaney 2003) : Let I be a closed real interval, i.e., I has one of the following forms: $[a, b]$, $[a, \infty)$, $(-\infty, b]$, or $(-\infty, \infty)$. A contraction mapping on I is a function $f : I \to I$ such that $|f'(x)| \leq K < 1$ for some contraction constant K. For any contraction mapping I, (i) f has a unique fixed point s in I; and (ii) for any $x_0 \in I$, the simple iteration $x_{n+1} = f(x_n)$ gives a sequence converging to s.

Lemma 15 *(Devaney 1992; Devaney 2003) . Let G be a contraction mapping on the interval [0,1] with contraction constant K. Then*

$$|q_c - q_\infty| \leq \frac{K^c}{1 - K}|q_1 - q_0|,$$

for any sequence $\{q_c\}$ where $q_c = G(q_{c-1})$ for $c \in \{1, 2, 3, \cdots\}$ and $q_0 \in [0, 1]$.

Lemma 15 is used in the proof of Theorem 1.

Theorem 1 $q* \in (0.7502000 \pm 10^{-7})$.

Proof: Consider $g(q,n)$ for a given n. Now $g(q,n)=0$ if and only if

$$q=\{\frac{2}{n}(2-q^{n-1})^{-1}\}^{1/n}\equiv G(q),$$

with derivative $G'(q)\equiv\frac{\partial G(q)}{\partial q}=2^{1/n}(n-1)q^{n-2}(n(2-q^{n-1}))^{-(n+1)/n}$. We have $0<G(q)<1$ for all $q\in[0,1]$, so G maps $[0,1]$ into itself. Since $\frac{\partial G'(q)}{\partial q}>0$ for $q\in(0,1)$, $|G'(q)|\leq 2^{1/n}(n-1)n^{-(n+1)/n}<1$ on $[0,1]$. Thus the function G is a contraction mapping on $[0,1]$ with contraction constant $K=2^{1/n}(n-1)n^{-(n+1)/n}\in[0,1]$. Therefore, we can use the Contraction Mapping Theorem for the function G. If $q_0=0$, then $q_1=G(q_0)=\frac{1}{n}^{1/n}$. By Lemma 15, $|q_c-q_\infty|\leq\frac{K^c}{1-K}(\frac{1}{n})^{1/n}$. Suppose now we require a bound on the number of iterations needed to determine q_∞ to six decimal places, i.e., we want $|q_c-q_\infty|\leq 10^{-7}=0.0000001$, i.e, $q_\infty\in(q_c\pm 0.0000001)$. By the above, it is enough to ensure that $\frac{K^c}{1-K}(\frac{1}{n})^{1/n}<10^{-7}$, that is

$$c>\frac{\log\{10^{-7}(\frac{1}{n})^{-1/n}(1-K)\}}{\log K}.$$

By Table 5.1 and Lemma 13, $n*=5$ when $q=q*$. Then $K\approx 0.6660$ and $c>41.57$, thus 42 iterations will be enough. For $q_0=0$, $q_{42}=0.7502000$. ■

5.3.3 Determining $n*$

In this section, we determine the array size $n*$ that minimizes the expected number of tests per specimen for a given $q>q*$. In other words, our goal is to find $n*$ such that

$$n*=\min_{n\in\{1,2,3,\cdots\}}f(q,n)\text{ for fixed }q\in(q*,1).\tag{5.16}$$

From Table 5.1, $f(q*,5)=1$. Since $f(q,n)$ is a decreasing function of q for fixed $n>1$, $f(q,5)<1$ for $q>q*$. Therefore, $f(q,n*)<1$ for $q>q*$. Also, by Table 5.1 we

know $f(q,n)$ decreases from $n=2$ to $n=n*$, then increases to $n=v$ and decreases thereafter. Since $f(q,n*)<1$ and $\lim_{n\to\infty} f(q,n)=1$, it follows that $n*$ from Table 5.1 is the solution of (5.16), i.e., $n*$ is the global minimum of $f(q,n)$ if we treat $f(q,n)$ as a function of the positive integers. Instead, if we think of $f(q,n)$ as a function of the positive real numbers and solve

$$f'(q,n)=\frac{\partial f(q,n)}{\partial n}=-\frac{2}{n^2}-2q^n\ln q+2q^{2n-1}\ln q=0, \tag{5.17}$$

then $n*$ and v will be close to the two positive solutions to this equation. In particular, if $\tilde{n}$ is the smallest positive real solution to (5.17), then $n*=\lfloor\tilde{n}\rfloor$ or $n*=\lceil\tilde{n}\rceil$, where $\lceil\tilde{n}\rceil$ denotes the smallest integer greater than or equal to $\tilde{n}$. However, there does not appear to be a closed form for $\tilde{n}$. Therefore, we use an approximation of $f'(q,n)$ to get the lower and upper bounds of $n*$.

Theorem 2 *For fixed $q\in[0.98,1)$, an upper bound for $n*$ is given by $\lceil\rho(q)\rceil$ where $\rho(q)$ is the smallest real root of the following quartic equation*

$$\begin{aligned}&\{(\tfrac{4}{q}-1)(1-q)^2\ln q\}n^4+\\&\{(2(1-q)+(1-q)^2-\tfrac{4}{q}(1-q)-\tfrac{2}{q}(1-q)^2)\ln q\}n^3+\\&\{(\tfrac{1}{q}-1)2\ln q\}n^2-2=0.\end{aligned} \tag{5.18}$$

For fixed $q\in(q*,0.98)$, an upper bound for $n*$ is given by $\lceil\rho(.98)\rceil$.

Proof: Recall Newton's generalized binomial theorem

$$(1+y)^k=\textstyle\sum_{c=0}^{\infty}\binom{k}{c}y^c=1+ky+\frac{k(k-1)}{2!}y^2+\cdots.$$

for $|y| < 1$. Let $y = -p$ and $k = n$, then

$$q^n = (1-p)^n \approx 1 - np + \tfrac{n(n-1)}{2}p^2.$$

Thus, $f'(q,n)$ can be approximated by

$$\begin{aligned} f'_{approx1}(q,n) &= -\tfrac{2}{n^2} - 2(\ln q)\{1 - np + \tfrac{n(n-1)}{2}p^2\} + 2(\ln q)\{1 - (2n-1)p + \tfrac{(2n-1)(2n-2)}{2}p^2\} \\ &= -\tfrac{2}{n^2} - 2(\ln q)\{(n-1)p - \tfrac{(n-1)(3n-2)}{2}p^2\} \end{aligned} \tag{5.19}$$

Let $\rho(q)$ be the smallest positive real solution to $f'_{approx1}(q,n) = 0$, which can be determined by applying the quartic formula to (5.18). Below, we show $f'(q,n) \geq f'_{approx1}(q,n)$ for given q, indicating $n* \leq \lceil \rho(q) \rceil$. Let $A(q) = B(q) - \frac{(n-1)(3n-2)}{2}$ for $n \geq 5$, where $B(q) = q^{n-1}(1 + q + q^2 + \cdots + q^{n-2})$. Then,

$$\begin{aligned} \tfrac{\partial\{qA(q)\}}{\partial q} &= A(q) + qA'(q) \\ &= q^{n-1}(n + (n+1)q + (n+2)q^2 + \cdots + (2n-2)q^{n-2}) - \tfrac{(n-1)(3n-2)}{2} \\ &= C(q) - \tfrac{(n-1)(3n-2)}{2} \end{aligned}$$

where $C(q) \leq \frac{(n-1)(3n-2)}{2}$. Therefore, $\frac{\partial\{qA(q)\}}{\partial q} \leq 0$ implying $qA(q)$ is a decreasing function of q. At $q = 1$, $qA(q) = -\frac{(n-1)(3n-4)}{2}$. Therefore,

$$-\frac{(n-1)(3n-4)}{2} \leq qA(q) = q\{B(q) - \frac{(n-1)(3n-2)}{2}\},$$

for $q \in (0,1)$. Thus

$$q\{B(q) - \frac{(n-1)(3n-2)}{2}\} + \frac{(n-1)(3n-4)}{2} \geq 0,$$

implying

$$qB(q) = \frac{q^n(1-q^{n-1})}{1-q} \geq q\frac{(n-1)(3n-2)}{2} - \frac{(n-1)(3n-4)}{2}.$$

Therefore,

$$q^n(1-q^{n-1}) = q^n - q^{2n-1} \geq (n-1)p - \frac{(n-1)(3n-2)}{2}p^2$$

implying $f'(q,n) \geq f'_{approx1}(q,n)$. We note that $f'_{approx1}(q,n) = 0$ does not have a real solution for $p \geq 0.03$. In particular, numerical evaluation of the discriminant indicates that all roots of the quartic equation (5.18) are complex for $p \geq 0.03$. Therefore, for $p \geq 0.03$, the upper bound of $n*$ is given by $\lceil \rho(.98) \rceil$. ■

Theorem 3 *For fixed $q \in (q*, 1)$, a lower bound for $n*$ is given by $\lfloor \mu(q) \rfloor$ where $\mu(q)$ is the smallest real root of the following cubic equation.*

$$\{2(1-q)\ln q\}n^3 + (2\ln q)n^2 + 2q\ln q - 2 = 0$$

Proof: Again we apply Newton's generalized binomial theorem, but now we use the approximation $q^n \approx 1 - np$, which yields

$$\begin{aligned} f'_{approx2}(q,n) &= -\tfrac{2}{n^2} - 2\ln q\{1-np\} + 2\ln q\{1-(2n-1)p\} \\ &= -\tfrac{2}{n^2} - 2\ln q\{(n-1)(1-q)\}. \end{aligned} \tag{5.20}$$

Let $\mu(q)$ be the smallest positive real solution to $f'_{approx2}(q,n) = 0$, which can be determined by the cubic formula. Below we show $f'(q,n) \leq f'_{approx2}(q,n)$ for given q, indicating $\lfloor \mu(q) \rfloor \leq n*$. We know

$$\frac{q^n(1-q^{n-1})}{1-q} = q^n(1+q+q^2+\cdots+q^{n-2}) \leq n-1.$$

Therefore,

$$q^n - q^{2n-1} \leq (n-1)(1-q),$$

and $f'(q,n) \leq f'_{approx2}(q,n)$, since $-2\ln q > 0$. ■

Table 5.2 shows $\mu(q)$, $n*$, and $\rho(q)$ for given values of p. Values for $n*$ in Table 5.2 were determined by a finite search over the integers $\lfloor \mu(q) \rfloor$ to $\lceil \rho(q) \rceil$. From this table, we conjecture without proof that $n* = \lfloor \rho(q) \rfloor$ for $p \leq 0.01$. The results in Table 5.2 agree with those in Table 1 of Phatarfod and Sudbury (1994), except $n* = 476$ (475 in Phatarfod and Sudbury) for $p = 0.0001$ and $n* = 751$ (750 in Phatarfod and Sudbury) for $p = 0.00005$. By Theorem 2, we know that $f'_{approx1}(q,n) = 0$ does not have a real solution for $p \geq 0.03$. Therefore, for $p \geq 0.03$, $n*$ can be determined by a finite search over the integers $\lfloor \mu(q) \rfloor$ to 18, where 18 is the value of $\lceil \rho(.98) \rceil$. Table 5.3 gives $n*$ for $p \in [0.03, p*)$.

5.4 Discussion

In this chapter, we studied the optimal configuration of $A2$. In particular we showed for prevalence greater than 0.2498, individual testing is more efficient than $A2$. For prevalence less than 0.2498, the optimal pool size for $A2$ can be determined by a simple finite search between the lower and upper bounds given in Theorems 2 and 3. We also showed that 2×2, 3×3, and 4×4 arrays are never optimal.

FIGURE 5.1: $g(q,n)$ vs. n of $A2$ for $q = (1-\sqrt{\frac{1}{2}})^{\frac{1}{4}}$, $q = 0.7502$, and $q = (\frac{1-\sqrt{\frac{1}{3}}}{2})^{\frac{1}{6}}$.

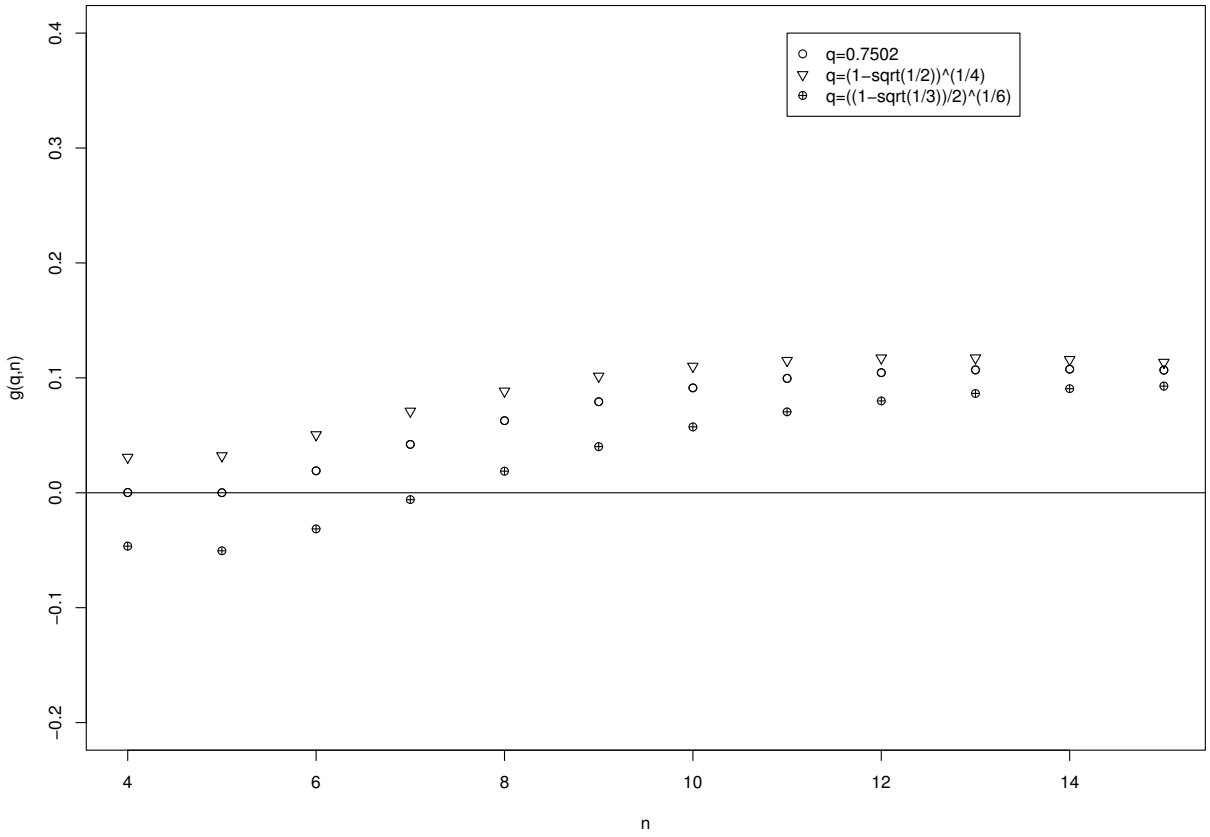


FIGURE 5.2: $\Delta(q, n)$ vs. q of $A2$ for $n = 2, 3, 4, 5, 6$.

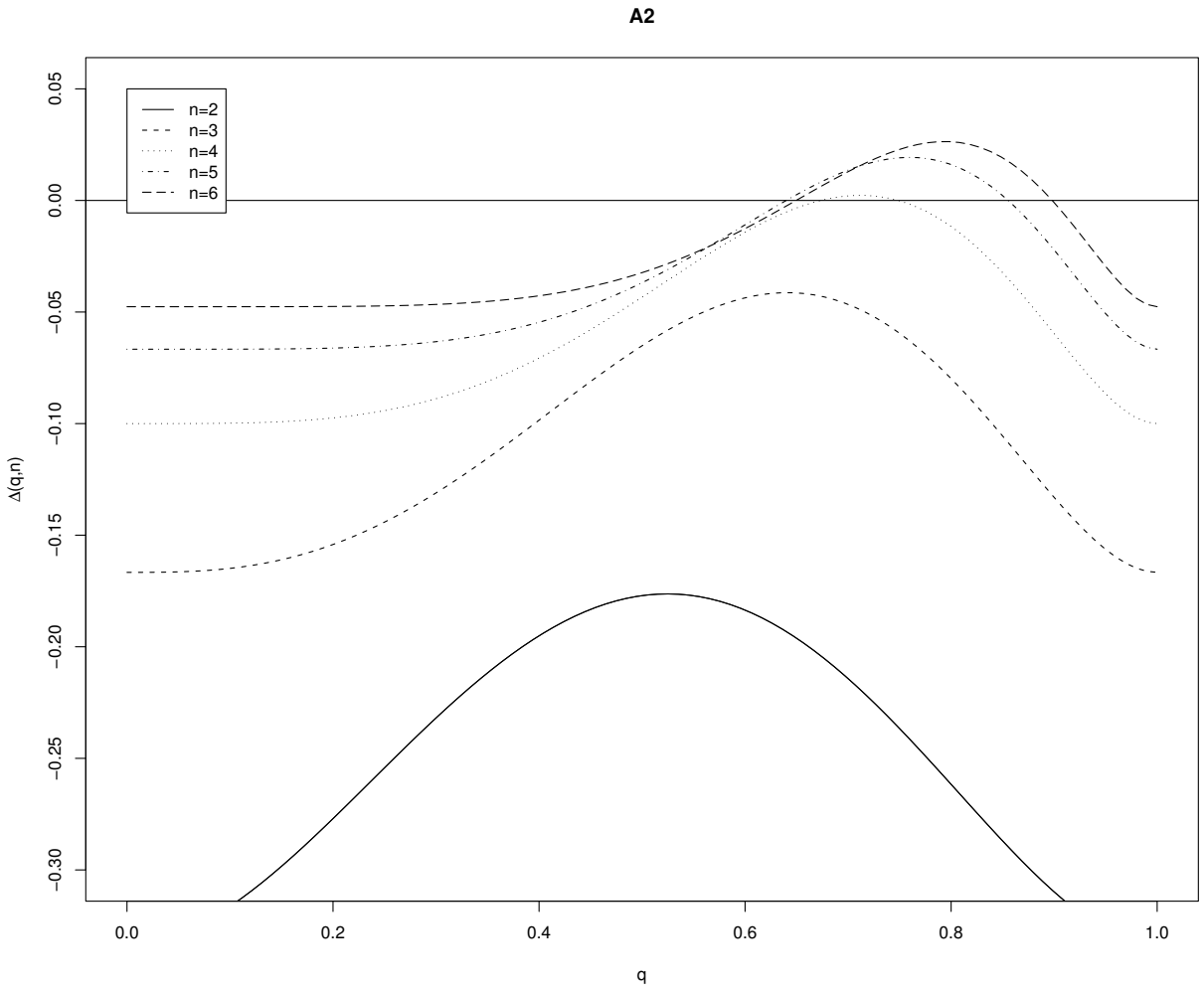


FIGURE 5.3: $\Delta(q, n)$ vs. q of $A2$ for $n = 5, 6, 7, 8, 9$.

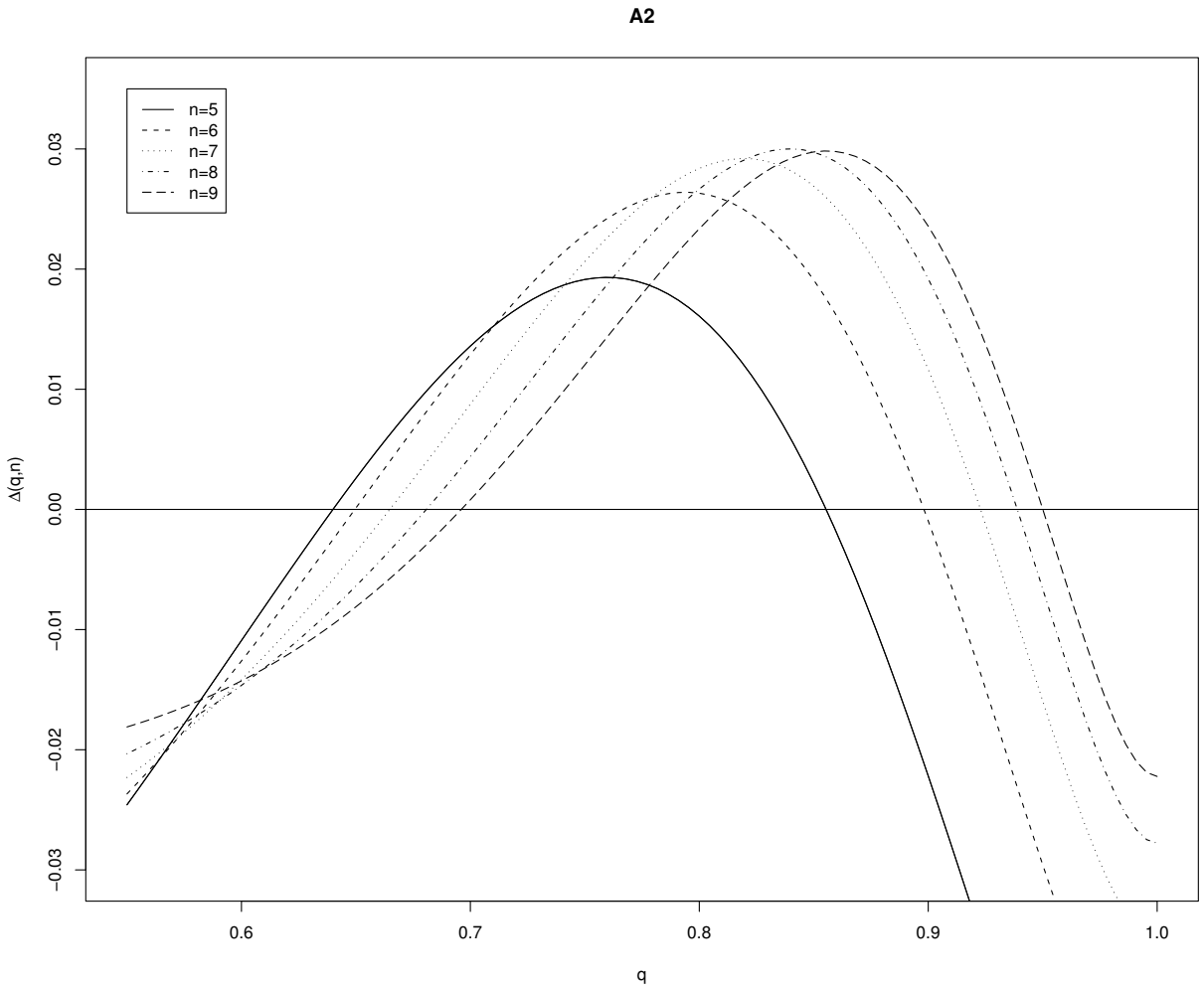


TABLE 5.1: The behavior of $f(q, n)$ and optimal pool size $n*$ for a given q

q	$f(q, n)$	$n*$
$q* < q$	$\uparrow$ to 2, then $\downarrow$ to $n*$, then $\uparrow$ to v, then $\downarrow$ to $f(q, n) = 1$	≥ 5
$r_2(4) < q < q*$	$\uparrow$ to 2, then $\downarrow$ to 5, then $\uparrow$ to v, then $\downarrow$ to $f(q, n) = 1$	1
$r_1(4) < q < r_2(4)$	$\uparrow$ to 2, then $\downarrow$ to 4, then $\uparrow$ to v, then $\downarrow$ to $f(q, n) = 1$	1
$r_1(5) < q < r_1(4)$	$\uparrow$ to 2, then $\downarrow$ to 5, then $\uparrow$ to v, then $\downarrow$ to $f(q, n) = 1$	1
$q < r_1(5)$	$\uparrow$ to 2, then $\downarrow$ to $f(q, n) = 1$	1

v is the smallest $n \geq 4$ for which $r_1(n) > q$.
$q* = 0.7502000$.

TABLE 5.2: Optimal array size $n*$ of $A2$ for given prevalence p. $n*$ is bounded between $\lfloor \mu(q) \rfloor$ and $\lceil \rho(q) \rceil$ according to Theorems 2 and 3.

p	$\mu(q)$	$n*$	$\rho(q)$
0.02	13.868	16	17.663
0.01	21.847	25	25.536
0.005	34.508	38	38.572
0.001	100.318	106	106.241
0.0005	159.061	166	166.239
0.0001	464.485	476	476.075
0.00005	737.134	751	751.508
0.00001	2154.76	2178	2178.75

TABLE 5.3: The optimal array size $n*$ of $A2$ for given prevalence p

p	$n*$
$> .2498$	1
$[0.1447, 0.2498]$	5
$[0.1019, 0.1446]$	6
$[0.0772, 0.1018]$	7
$[0.0611, 0.0771]$	8
$[0.0500, 0.0610]$	9
$[0.0419, 0.0499]$	10
$[0.0357, 0.0418]$	11
$[0.0309, 0.0356]$	12
$[0.0300, 0.0308]$	13
< 0.0300	$\lfloor \mu(q) \rfloor \leq n* \leq \lceil \rho(q) \rceil$

APPENDIX A

Per-Family Error Rates

In this section, we prove the following lemma concerning the per-family error rates ($PFER$ and $PREF_2$). This result holds for any pooling algorithm.

Lemma: For any pooling algorithm $\mathcal{A}$, $PFER(\mathcal{A}) = nq\{1-S_p(\mathcal{A})\}$ and $PFER_2(\mathcal{A}) = np\{1 - S_e(\mathcal{A})\}$.

Proof: Suppose the pooling algorithm $\mathcal{A}$ is applied to n specimens. For $i = 1, \ldots, n$, let $Y_i = 1$ if the i^{th} specimen is positive and 0 otherwise. Let $X_i = 1$ if the i^{th} specimen tests positive by the algorithm $\mathcal{A}$, and 0 otherwise. Then $W = \sum_i (1 - X_i)Y_i$ is the number of false negatives and $Z = \sum_i X_i(1 - Y_i)$ is the number of false positives such that

$$\begin{aligned} PFER(\mathcal{A}) &= E\{\textstyle\sum_i X_i(1 - Y_i)\} \\ &= nE\{X_i(1 - Y_i)\} \\ &= n \Pr[X_i(1 - Y_i) = 1] \\ &= n \Pr[X_i = 1, Y_i = 0] \\ &= n \Pr[X_i = 1 | Y_i = 0] \Pr[Y_i = 0] \\ &= nq\{1 - S_p(\mathcal{A})\}. \end{aligned}$$

Similarly, one can show $PFER_2(\mathcal{A}) = np\{1 - S_e(\mathcal{A})\}$. $\square$

APPENDIX B

Efficiency Variance

In this section, we derive $E(T^2)$ for $D2$, $D3$, $A2$ and $A2M$, where T is the number of tests required to classify all specimens as positive or negative. In turn, one can compute $Var(T) = E(T^2) - \{E(T)\}^2$ using $E(T)$ as given in Sections 3.3, 3.4, and 3.5 of the main paper.

B.1 Dorfman algorithm ($D2$)

Letting $T = 1 + T_2$ for $D2$, then $E(T^2)$ can be expressed as

$$\begin{aligned} E(T^2) &= 1 + 2E(T_2) + E({T_2}^2) \\ &= 1 + 2n_1 \Pr[X_{1i} = 1] + {n_1}^2 \Pr[X_{1i} = 1] \\ &= 1 + 2n_1 f(n_1) + {n_1}^2 f(n_1) \\ &= 1 + (2n_1 + {n_1}^2) f(n_1). \end{aligned} \tag{B.1}$$

B.2 Three-stage hierarchical algorithm ($D3$)

Letting $T = 1 + T_2 + T_3$ for $D3$, then $E(T^2)$ can be written as

$$\begin{aligned} E(T^2) &= 1 + 2E(T_2) + 2E(T_3) + 2E(T_2T_3) + E({T_2}^2) + E({T_3}^2) \\ &= 1 + 2k_1 f(n_1) + 2n_1 \Pr[X_{2i} = 1] + 2E(T_2T_3) + {k_1}^2 f(n_1) + E({T_3}^2). \end{aligned} \tag{B.2}$$

We may express $E(T_2T_3)$ and $E(T_3^2)$ as

$$E(T_2T_3) \quad = n_1k_1\{f(n_1-k_2)(1-S_p)q^{k_2}+S_e^2(1-q^{k_2})\}$$

and

$$E(T_3^2) \quad = k_2^2\{k_1\Pr[X_{2i}=1]+k_1(k_1-1)\Pr[X_{2i}=1]\Pr[X_{2j}=1]\},$$

where $\Pr[X_{2i}=1]=\{S_e^2(1-q^{k_2})+(1-S_p)f(n_1-k_2)q^{k_2}\}$.

B.3 Square array without master pool ($A2$)

Letting $T=2n+T_2$ for $A2$ where $T_2=\sum_{i,j}T_{2ij}$, then $E(T^2)$ can be written as

$$\begin{aligned}E(T^2) &= E(4n^2+4nT_2+{T_2}^2)\\ &= 4n^2+4n^3E(T_{2ij})\\ &+n^2E({T_{2ij}}^2)+[n^4-\{n^2+n^2(n-1)^2\}]E(T_{2ij}T_{2ij'})+\{n^2(n-1)^2\}E(T_{2ij}T_{2i'j'}).\end{aligned} \tag{B.3}$$

Note that $E(T_{2ij})=E({T_{2ij}}^2)=g(n)+2h(n)$. $E(T_{2ij}T_{2ij'})$ and $E(T_{2ij}T_{2i'j'})$ of equation (B.3) can be expressed as

$$\begin{aligned}E(T_{2ij}T_{2ij'}) &= \Pr[R_i=1,C_j=1,C_{j'}=1]+\Pr[R_i=1,\textstyle\sum_j C_j=0]\\ &+\Pr[\textstyle\sum_i R_i=0,C_j=1,C_{j'}=1]\end{aligned} \tag{B.4}$$

and

$$E(T_{2ij}T_{2i'j'}) \quad = \Pr[R_i=1,R_{i'}=1,C_j=1,C_{j'}=1]+2\Pr[\textstyle\sum_i R_i=0,C_j=1,C_{j'}=1]. \tag{B.5}$$

The first parts of the right sides of equations (B.4) and (B.5) equal

$$
\begin{aligned}
&\Pr[R_i = 1, C_j = 1, C_{j'} = 1] \\
&\quad = \sum_{c_1=0}^{1} \sum_{c_2=0}^{1} \sum_{r=0}^{1} \Pr[R_i = 1, C_j = 1, C_{j'} = 1 | R_i^T = r, C_j^T = c_1, C_{j'}^T = c_2] \\
&\qquad \times \Pr[R_i^T = r, C_j^T = c_1, C_{j'}^T = c_2] \\
&\qquad = (1 - S_p)^3 q^{3n-2} + (1 - S_p)^2 S_e q^{2n} (1 - q^{n-2}) + 2(1 - S_p)^2 S_e q^{2n-1} (1 - q^{n-1}) \\
&\qquad + 2(1 - S_p) S_e^2 q^n (1 - q^{n-1} - q^n + q^{2n-2}) + (1 - S_p) S_e^2 q^n (1 - q^{n-1})^2 \\
&\qquad + S_e^3 (1 - 3q^n + q^{2n} + 2q^{2n-1} - q^{3n-2})
\end{aligned}
$$

and

$$
\begin{aligned}
&\Pr[R_i = 1, R_{i'} = 1, C_j = 1, C_{j'} = 1] \\
&\quad = \sum_{c_1=0}^{1} \sum_{c_2=0}^{1} \sum_{r_1=0}^{1} \sum_{r_2=0}^{1} \Pr[R_i = 1, R_{i'} = 1, C_j = 1, C_{j'} = 1 | R_i^T = r_1, R_{i'}^T = r_2, C_j^T = c_1, C_{j'}^T = c_2] \\
&\qquad \times \Pr[R_i^T = r_1, R_{i'}^T = r_2, C_j^T = c_1, C_{j'}^T = c_2] \\
&\qquad = (1 - S_p)^4 q^{4n-4} + 4(1 - S_p)^3 S_e q^{3n-2} (1 - q^{n-2}) + 2(1 - S_p)^2 S_e^2 q^{2n} (1 - q^{n-2})^2 \\
&\qquad + 4(1 - S_p)^2 S_e^2 q^{2n-1} (1 - 2q^{n-1} + q^{2n-3}) \\
&\qquad + 4(1 - S_p) S_e^3 q^n (1 - q^n - 2q^{n-1} + 3q^{2n-2} - q^{3n-4}) \\
&\qquad + S_e^4 (1 - 4q^n + 4q^{2n-1} + 2q^{2n} - 4q^{3n-2} + q^{4n-4}).
\end{aligned}
$$

$\Pr[R_i = 1, \sum_j C_j = 0]$ of equation (B.4) equals $h(n)$ in Section 3.4.1. The last parts of the right sides of equations (B.4) and (B.5) equal

$$\begin{aligned}
&\Pr[\sum_i R_i = 0, C_j = 1, C_{j'} = 1] \\
&= \sum_{c_1=0}^{1}\sum_{c_2=0}^{1}\sum_{r=0}^{n} \Pr[\sum R_i = 0, C_j = 1, C_{j'} = 1 | \sum R_i^T = r, C_j^T = c_1, C_{j'}^T = c_2] \\
&\quad \times \Pr[\sum R_i^T = r, C_j^T = c_1, C_{j'}^T = c_2].
\end{aligned}$$

For $r \in 0, \cdots, n$, let

$$\beta_{00}(r) \equiv \Pr[\sum R_i^T = r, C_j^T = 0, C_{j'}^T = 0] = \binom{n}{r}(q^{n^2-r(n-2)})(1-q^{n-2})^r,$$

$$\beta_{01}(r) \equiv \Pr[\sum R_i^T = r, C_j^T = 0, C_{j'}^T = 1] = \binom{n}{r}(q^{n^2-r(n-1)})(1-q^r-q^{n-1}+q^{n-2+r})^r,$$

$$\beta_{11}(r) \equiv \Pr[\sum R_i^T = r, C_j^T = 1, C_{j'}^T = 1] = \binom{n}{r}(q^{n^2-rn})(1-q^n)^r - (\beta_{00}(r)+2\beta_{01}(r)),$$

$$\begin{aligned}
\gamma_{00}(r) &\equiv \Pr[\sum R_i = 0, C_j = 1, C_{j'} = 1 | \sum R_i^T = r, C_j^T = 0, C_{j'}^T = 0] \\
&= (1-S_p)^2(1-S_e)^r S_p^{n-r},
\end{aligned}$$

$$\begin{aligned}
\gamma_{01}(r) &\equiv \Pr[\sum R_i = 0, C_j = 1, C_{j'} = 1 | \sum R_i^T = r, C_j^T = 0, C_{j'}^T = 1] \\
&= (1-S_p)S_e(1-S_e)^r S_p^{n-r},
\end{aligned}$$

and

$$\gamma_{11}(r) \equiv \Pr[\sum R_i = 0, C_j = 1, C_{j'} = 1 | \sum R_i^T = r, C_j^T = 1, C_{j'}^T = 1] = S_e^2(1-S_e)^r S_p^{n-r},$$

where we define $\beta_{01}(0) \equiv 0$, $\beta_{11}(0) \equiv 0$, $\gamma_{01}(0) \equiv 0$ and $\gamma_{11}(0) \equiv 0$. Then it follows that

$$\Pr[\sum_i R_i = 0, C_j = 1, C_{j'} = 1] = \sum_{r=0}^{n}\{\gamma_{00}(r)\beta_{00}(r) + 2\gamma_{01}(r)\beta_{01}(r) + \gamma_{11}(r)\beta_{11}(r)\}.$$

B.4 Square array with master pool ($A2M$)

Letting $T = 1 + 2nX_0 + T_2$ for $A2M$ where $T_2 = \sum_{i,j} T_{2ij}$, then $E(T^2)$ can be written as

$$\begin{aligned} E(T^2) &= E(1 + 4n^2 {X_0}^2 + {T_2}^2 + 4nX_0 + 2T_2 + 4nX_0T_2) \\ &= 1 + 4n^2 f(n^2) \\ &+ n^2 E({T_{2ij}}^2) + [n^4 - \{n^2 + n^2(n-1)^2\}]E(T_{2ij}T_{2ij'}) + \{n^2(n-1)^2\}E(T_{2ij}T_{2i'j'}) \\ &+ 4nf(n^2) + 2n^2 E(T_{2ij}) + 4n^3 E(T_{2ij}). \end{aligned} \tag{B.6}$$

Note that $E(T_{2ij})$ and $E({T_{2ij}}^2)$ are same as equation (3.19) in Section 3.5.1. $E(T_{2ij}T_{2ij'})$ and $E(T_{2ij}T_{2i'j'})$ of equation (B.6) can be expressed as

$$\begin{aligned} E(T_{2ij}T_{2ij'}) &= \Pr[X_0 = 1, R_i = 1, C_j = 1, C_{j'} = 1] + \Pr[X_0 = 1, R_i = 1, \textstyle\sum_j C_j = 0] \\ &+ \Pr[X_0 = 1, \textstyle\sum_i R_i = 0, C_j = 1, C_{j'} = 1] \end{aligned} \tag{B.7}$$

and

$$\begin{aligned} E(T_{2ij}T_{2i'j'}) &= \Pr[X_0 = 1, R_i = 1, R_{i'} = 1, C_j = 1, C_{j'} = 1] \\ &+ 2\Pr[X_0 = 1, \textstyle\sum_i R_i = 0, C_j = 1, C_{j'} = 1]. \end{aligned} \tag{B.8}$$

The first parts of the right sides of equations (B.7) and (B.8) equal

$$\begin{aligned} &\Pr[X_0 = 1, R_i = 1, C_j = 1, C_{j'} = 1] \\ &\quad = (1 - S_p)^3 q^{n^2}(1 - S_p - S_e) + S_e \Pr[R_i = 1, C_j = 1, C_{j'} = 1] \end{aligned}$$

and

$$\begin{aligned} &\Pr[X_0 = 1, R_i = 1, R_{i'} = 1, C_j = 1, C_{j'} = 1] \\ &\quad = (1 - S_p)^4 q^{n^2}(1 - S_p - S_e) + S_e \Pr[R_i = 1, R_{i'} = 1, C_j = 1, C_{j'} = 1]. \end{aligned}$$

The second and third parts of the right side of equation (B.7) equal

$$\Pr[X_0 = 1, R_i = 1, \sum_j C_j = 0] = (1 - S_p - S_e)\gamma_0(0)\beta_0(0) + S_e h(n)$$

and

$$\begin{aligned}
&\Pr[X_0 = 1, \sum_i R_i = 0, C_j = 1, C_{j'} = 1] \\
&\quad = (1 - S_p - S_e)\gamma_{00}(0)\beta_{00}(0) + S_e \sum_{r=0}^{n} \{\gamma_{00}(r)\beta_{00}(r) + 2\gamma_{01}(r)\beta_{01}(r) + \gamma_{11}(r)\beta_{11}(r)\}.
\end{aligned}$$

APPENDIX C

CDFs for pooling error rates

In this section, we derive the cumulative distribution function (CDF) of the observed pooling sensitivity, specificity, PPV, and NPV for any pooling algorithm $\mathcal{A}$. Suppose the pooling algorithm $\mathcal{A}$ is applied to N specimens. For $i = 1, \ldots, N$, let $Y_i = 1$ if the i^{th} specimen is positive and 0 otherwise. Let $X_i = 1$ if the i^{th} specimen tests positive by the algorithm $\mathcal{A}$, and 0 otherwise.

C.1 Pooling Sensitivity and Specificity

The observed sensitivity can be written

$$S_e^O(\mathcal{A}) \equiv \frac{\sum X_i Y_i}{\sum Y_i},$$

where the summations are from 1 to N. Note the denominator $\sum Y_i$ is Binomial(N, p) where p is the prevalence. Also note that

$$\sum X_i Y_i | \sum Y_i \sim Binomial(\sum Y_i, S_e(\mathcal{A})),$$

where $S_e(\mathcal{A}) = E(X_i|Y_i = 1)$. Therefore,

$$\begin{aligned}\Pr[S_e^O(\mathcal{A}) \leq s] &= \Pr[\sum X_i Y_i \leq s \sum Y_i] \\ &= \sum_{y=0}^{N} \Pr[\sum X_i Y_i \leq sy | \sum Y_i = y] \Pr[\sum Y_i = y] \\ &= \sum_{y=0}^{N} \sum_{x=0}^{\lfloor sy \rfloor} \binom{y}{x} S_e(\mathcal{A})^x (1 - S_e(\mathcal{A}))^{y-x} \binom{N}{y} p^y (1-p)^{N-y},\end{aligned}$$

where $\lfloor sy \rfloor$ denotes the largest integer less than or equal to sy.

Similarly, the CDF of the observed specificity equals

$$\Pr[S_p^O(\mathcal{A}) \leq s] = \sum_{y=0}^{N} \sum_{x=0}^{\lfloor sy \rfloor} \binom{y}{x} S_p(\mathcal{A})^x (1 - S_p(\mathcal{A}))^{y-x} \binom{N}{y} (1-p)^y p^{N-y}.$$

C.2 Pooling PPV and NPV

The observed PPV can be written

$$PPV^O(\mathcal{A}) \equiv \frac{\sum X_i Y_i}{\sum X_i},$$

where again the summations are from 1 to N. Note the denominator $\sum X_i$ is Binomial(N, p_1) where $p_1 = S_e(\mathcal{A})p + (1 - S_p(\mathcal{A}))q$. Also note that

$$\sum X_i Y_i | \sum X_i \sim Binomial(\sum X_i, PPV(\mathcal{A})).$$

This follows since

$$\Pr[Y_i = 1 | X_i = 1] = \frac{S_e(\mathcal{A})p}{S_e(\mathcal{A})p + (1 - S_p(\mathcal{A}))q} = PPV(\mathcal{A}).$$

Therefore,

$$\begin{aligned}\Pr[PPV^O(\mathcal{A}) \leq s] &= \Pr[\textstyle\sum X_i Y_i \leq s \sum X_i] \\ &= \textstyle\sum_{x=0}^{N} \Pr[\sum X_i Y_i \leq sx | \sum X_i = x] \Pr[\sum X_i = x] \\ &= \textstyle\sum_{x=0}^{N} \sum_{y=0}^{\lfloor sx \rfloor} \binom{x}{y} PPV(\mathcal{A})^y (1 - PPV(\mathcal{A}))^{x-y} \binom{N}{x} {p_1}^x (1-p_1)^{N-x},\end{aligned}$$

Similarly, the CDF of the observed NPV equals

$$\Pr[NPV^O(\mathcal{A}) \leq s] = \textstyle\sum_{x=0}^{N} \sum_{y=0}^{\lfloor sx \rfloor} \binom{x}{y} NPV(\mathcal{A})^y (1 - NPV(\mathcal{A}))^{x-y} \binom{N}{x} (1-p_1)^x p_1^{N-x}.$$

REFERENCES

Amemiya, C. T., Algeria-Hartman, M. J., Aslanidis, C., Chen, C., Nikolic, J., Gingrich, J. C. and de Jong, P. J. (1992). A two-dimensional YAC pooling strategy for library screening via STS and Alu-PCR methods. *Nucleic Acids Research* **25**, 2559–63.

Amos, C. I., Fazier, M. L. and Wang, W. (2000). DNA pooling mutation detection in sequence analysis. *American Journal of Human Genetics* **66**, 1689–1692.

Barillot, E., Lacroix, B. and Cohen, D. (1991). Theoretical analysis of library screening using a N-dimensional pooling strategy. *Nucleic Acids Research* **19**, 6241–7.

Berger, T., Mandell, J. W. and Subrahmanya, P. (2000). Maximally efficient two-stage screening. *Biometrics* **56**, 833–840.

Bruno, W. J., Knill, E., Balding, D. J., Bruce, D. C., Doggett, N. A., Sawhill, W. W., Stallings, R. L., Whittaker, C. C. and Torney, D. C. (1995). Efficient pooling designs for library screening. *Genomics* **26**, 21–30.

Centers for Disease Control and Prevention (2003). Detection of West Nile virus in blood donations – United States, 2003. *Morbidity and Mortality Weekly Report* **52**, 769–772.

Cook, R. J. and Dunnett, C. W. (1998). Multiple comparisons. In Armitage, P. and Colton, T., editors, *Encyclopedia of Biostatistics*, pages 2736–2746. Wiley, New York.

Daar, E., Little, S., Pitt, J., Santangelo, J., Ho, P., Harawa, N., Kerndt, P., Glorgi, J., Bai, J., Gaut, P., Richman, D., Mandel, S. and for the Los Angeles County Primary HIV Infection Recruitment Network, S. N. (2001). Diagnosis of primary HIV-1 infection. *Annals of Internal Medicine* **134**, 25–29.

Dendukuri, N. and Joseph, L. (2001). Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. *Biometrics* **57**, 158–167.

Devaney, R. L. (1992). *A first course in chaotic dynamical systems*. Perseus Press.

Devaney, R. L. (2003). *An introduction to chaotic dynamical systems*. Westview Press, second edition.

Dorfman, R. (1943). The detection of defective numbers of large populations. *Annals of Mathematical Statistics* **14**, 436–440.

Feller, W. (1957). *An Introduction to Probability Theory and its Applications*. Wiley, New York.

Finucan, H. M. (1964). The blood testing problem (Corr: V14 p210). *Applied Statistics* **13**, 43–50.

Gastwirth, J. L. (2000). The efficiency of pooling in the detection of rare mutations. *American Journal of Human Genetics* **67**, 1036–9.

Gupta, D. and Malina, R. (1999). Group testing in presence of classification errors. *Statistics in Medicine* **18**, 1049–1068.

Hecht, F., Busch, M., Rawal, B., Webb, M., Rosenberg, E., Swanson, M., Chesney, M., Anderson, J., Levy, J. and Kahn, J. (2002). Use of laboratory tests and clinical symptoms for identification of primary HIV infection. *AIDS* **16(8)**, 1119–1129.

Hochberg, Y. and Tamhane, A. C. (1987). *Multiple Comparison Procedures*. Wiley.

Hwang, F. K. (1972). A method for detecting all defective members in a population by group testing. *Journal of the American Statistical Association* **89**, 605–608.

Hwang, F. K. (1976). Group testing with a dilution effect. *Biometrika* **63**, 671–673.

Johnson, N. L., Kotz, S. and Wu, X. (1991). *Inspection errors for attributes in quality control*. Chapman and Hall Ltd.

Jones, C. M. and Zhigljavsky, A. A. (2001). Comparison of costs for multistage group testing methods in the pharmaceutical industry. *Communications in Statistics, Part A – Theory and Methods* **30**, 2189–2209.

Kacena, K. A., Quinn, S. B., Hartman, S. C., Quinn, T. C. and Gaydos, C. A. (1998). Pooling of urine samples for screening for Neisseria gonorrhea by ligase chain reaction: accuracy and application. *Journal of Clinical Microbiology* **36**, 3624–8.

Klausner, J. D. (2004). The NAAT is out of the bag. *Clinical Infectious Diseases* **38**, 820–1.

Lancaster, V. A. and Keller-McNulty, S. (1998). A review of composite sampling methods. *Journal of the American Statistical Association* **93**, 1216–1230.

Litvak, E., Tu, X. M. and Pagano, M. (1994). Screening for the presence of a disease by pooling sera samples. *Journal of the American Statistical Association* **89**, 424–434.

Mine, H., Emura, H., Miyamoto, M., Tomono, T., Minegishi, K., Murokawa, H., Yamanaka, R., Yoshikawa, A., Nishioka, K. and Group, J. R. C. N. R. (2003). High throughput screening of 16 million serologically negative blood donors for hepatitis B virus, hepatitis C virus and human immunodeficiency virus type-1 by nucleic acid

amplification testing with specific and sensitive multiplex reagent in Japan. *Journal of Virological Methods* **112**, 145–51.

NIAID Office of Communications and Public Liaison (2005). NIAID Funds Center for HIV/AIDS Vaccine Immunology (CHAVI). *NIAID News* .

Phatarfod, R. M. and Sudbury, A. (1994). The use of a square array scheme in blood testing. *Statistics in Medicine* **13**, 2337–2343.

Pilcher, C. D., Fiscus, S. A., Nguyen, T. Q., Foust, E., Wolf, L., Williams, D., Ashby, R., O'Dowd, J. O., McPherson, J. T., Stalzer, B., Hightow, L., Miller, W. C., Eron, J. J. and Cohen, M. S. (2005). Detection of acute infections during HIV testing in North Carolina. *New England Journal of Medicine* **352**, 1873–83.

Pilcher, C. D., McPherson, J. T., Leone, P. A., Smurzynski, M., Owen-O'Dowd, J., Peace-Brewer, A. L., Harris, J., Hicks, C. B., Eron, J. J. and Fiscus, S. A. (2002). Real-time, universal screening for acute HIV infection in a routine HIV counseling and testing population. *Journal of the American Medical Association* **288**, 216–21.

Pilcher, C. D., Price, M. A., Hoffman, I. F., Galvin, S., Martinson, F. E., Kazembe, P. N., Eron, J. J., Miller, W. C., Fiscus, S. A. and Cohen, M. S. (2004). Frequent detection of acute primary HIV infection in men in Malawi. *AIDS* **18**, 517–24.

Pilcher, C. D., Tien, H. C., Eron, J. J., Vernazza, P. L., Leu, S. Y., Stewart, P. W., Goh, L. E., M. S. Cohen, for the Quest Study and the Duke-UNC-Emory Acute HIV Consortium (2004). Brief but efficient: acute HIV infection and the sexual transmission of HIV. *Journal of Infectious Diseases* **189**, 1785–92.

Quinn, T. C., Brookmeyer, R., Kline, R., Shepherd, M., Paranjape, R., Mehendale, S., Gadkari, D. A. 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**, 2751–7.

Roederer, M. and Koup, R. A. (2003). Optimized determination of T cell epitope responses. *Journal of Immunological Methods* **274**, 221–228.

Samuels, S. M. (1978). The exact solution to the two-stage group-testing problem. *Technometrics* **20**, 497–500.

Schuster, E. (2002). Two-stage screening with a limited number of samples in the scheme. *Proceedings of the Conference CompStat 2002; Short Communications and Posters* .

Sobel, M. and Groll, P. (1959). Group testing to estimate efficiently all defectives in a binomial sample. *Bell System Technical Journal* **38**, 1079–1252.

Sterrett, A. (1957). On the detection of defective members of large populations. *Annals*

of Mathematical Statistics **28**, 1033–1036.

Turner, D., Tidmore, F. and Young, D. (1988). A calculus based approach to the blood testing problem. *SIAM Review* **30**, 119–122.

Venette, R. C., Moon, R. D. and Hutchinson, W. D. (2002). Strategies and statistics of sampling for rare individuals. *Annual Review of Entomology* **47**, 143–74.

Wawer, M. J., Gray, R. H., Sewankambo, N. K., Serwadda, D., Li, X., Laeyendecker, O., Kiwanuka, N., Kigozi, G., Kiddugavu, M., Lutalo, T., Nalugoda, F., Wabwire-Mangen, F., Meehan, M. P. and Quinn, T. C. (2005). Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *Journal of Infectious Diseases* **191**, 1403–9.

Wein, L. M. and Zenios, S. A. (1996). Pooled testing for HIV screening: capturing the dilution effect. *Operations Research* **44**, 543–569.

Wilks, S. S. (1962). *Mathematical Statistics*. Wiley, New York.

Wolf, J. K. (1985). Born again group testing: multiaccess communications. *IEEE Transactions and Information Theory* **31**, 185–191.

Wolfram (1998). Natural logarithm: Inequalities. *From Wolfram Function Site. http://functions.wolfram.com/01.04.29.0003.01* .

Woodbury, C. P., Fitzloff, J. F. and Vincent, S. S. (1995). Sample multiplexing for greater throughput in HPLC and related methods. *Analytical Chemistry* **67**, 885–90.

Wu, X. and Zhao, B. (1994). Optimal sampling of hierarchical screening with inspection errors. *Communications in Statistics, Part A – Theory and Methods* **23**, 803–814.

Xie, M., Tatsuoka, K., Sacks, J. and Young, S. S. (2001). Group testing with blockers and synergism. *Journal of the American Statistical Association* **96**, 92–102.

Zwaal, R. R., Broeks, A., van Meurs, J., Groenen, J. T. M. and Plasterk, R. H. A. (1993). Target-selected gene inactivation in *Caenorhabditis elegans* by using a frozen transposon insertion mutant bank. *Proceedings of the National Academy of Sciences* **1993**, 7431–7435.