Pool Testing with Dilution Effects and Heterogeneous Priors

Gustavo Quinderé Saraiva

Latest draft: click here. This draft: June 8, 2022.
First draft: February 19th, 2021.

The Dorfman *pooled testing* scheme is a process in which individual specimens (e.g., blood, urine, swabs, etc.) are pooled and tested together; if the merged sample tests positive for infection, then each specimen from the pool is tested individually. Through this procedure, laboratories can reduce the expected number of tests required to screen the population, as individual tests are only carried out when the pooled test detects infection. Several different partitions of the population can be used to form the pools. In this study we analyze the performance of *ordered partitions*, those in which subjects with similar probability of infection are pooled together. We derive sufficient conditions under which ordered partitions outperform other types of partitions in terms of minimizing the expected number of tests, the expected number of false negatives, and the expected number of false positive classifications. These sufficient conditions can be easily verified in practical applications, once the dilution effect has been estimated. We also propose a measure of equity and present conditions under which this measure is maximized by ordered partitions.

Key words: Pool Testing, Dilution Effect, Heterogenous Priors, Dorfman Screening

Highlights

- We derive conditions under which it may be desirable to group together patients with similar probability of infection when implementing the Dorfman pooled testing procedure. We show that, depending on the dilution effect, this way of pooling subjects minimizes the expected number of tests required to screen the population, as well as the expected number of false positive and false negative classifications.
- We also propose a measure of equity and derive conditions under which this pooling method maximizes equity.
- We present a couple of case studies showing how these results can be applied to real data.

1. Introduction

For many infectious diseases it is common practice to screen the population through a *pool testing* scheme (also known as *group testing*) a process in which specimens (e.g., blood, urine, swabs) from different subjects are pooled and tested together. One of the simplest versions of pooled testing is the Dorfman procedure (due to Dorfman (1943)). In this procedure pooled samples are tested together, and

whenever a pooled test detects infection, each specimen from that group is tested individually. Compared to testing subjects individually, this procedure may potentially reduce the overall expected number of tests required to screen a population, as subjects are only tested individually in the event the pooled test detects infection. Perhaps because of its simplicity, this type of pool testing scheme is the most implemented in practical applications (e.g., McMahan, Tebbs and Bilder (2012)).

The study of pooled testing started during World War II with the seminal work of Dorfman (1943), which suggested pooling blood samples from the US military to detect syphilis in soldiers. Since then the field has evolved to produce several new applications, including the detection of other infectious diseases such as Chlamydia (e.g., McMahan, Tebbs and Bilder (2012)) and HIV (e.g., Nguyen et al. (2019)), the detection of defective parts in production lines (e.g., Sobel and Groll (1959)), the detection of data tampering using one-way hash functions (e.g., Goodrich, Atallah and Tamassia (2005)) and the allocation of transmission time slots to users in multiaccess channels (e.g., Chlebus (2001)). Recently, the field has regained new interest as pool testing techniques can be used to minimize the costs of detecting COVID-19 in a population. As of now, many countries including the US, Germany, China and Chile have experimented using pool testing to screen the population for COVID-19 (e.g., Basso et al. (2022), Grobe et al. (2020) and Fan (2020)).

For tractability reasons, most research in the literature assume that pooled testing is not subject to dilution effects, namely, that grouping non-infected with infected samples does not reduce the probability of detecting infection in the group (see Kim et al. (2007) for a thorough review). But in reality, dilution effects have proven to be non-negligible in several practical applications, including in the detection of COVID-19 using RT-PCR tests (e.g., Bateman et al. (2020)).[1] Small but non-negligible dilution effects have been reported in other applications, such as when pooling samples to test for chlamydia and gonorrhea, for pool sizes less than or equal to 10 (Morre et al. (2000) and Kacena et al. (1998*b*)).

Most research in the literature also work under the assumption that each subject has the same probability of infection (e.g., Kim et al. (2007)), when in reality the probability of infection can be highly dependent on subjects' demographics. As an example, the prevalence of chlamydia and other sexually transmitted diseases in the U.S. varies considerably with age and demographics.[2] For blood screening, the probability of HIV infection from first-time donors in the U.S. is approximately 7 times higher than that from repeat donors (Zou, Stramer and Dodd (2012)). Because some attributes of subjects can be collected by the tester, the tester could use some of this information to determine who should be matched with whom to form the pools.

---

[1] Bateman et al. (2020) estimate that the probability of detecting COVID-19 from an infected subject is 6% lower when his sample is diluted with the sample of 4 healthy subjects. The percentage reduction in the precision of the test is 8% when the infected sample is further diluted with 9 non infected samples, and 18% when the infected sample is diluted with 49 non infected specimens.

[2] See section 7 for details.

Several different partitions can be used to form the pools. In this study we analyze the performance of *ordered partitions*, those in which subjects with similar probability of infection are matched together. In the absence of dilution effects, Hwang (1975) and McMahan, Tebbs and Bilder (2012) have shown that there exists an ordered partition that minimizes the expected number of tests required to screen the population. Moreover, in the absence of dilution effects, the probability that any infected subject is incorrectly diagnosed as not infected does not depend on who the subject is matched with in the pool. So, in the absence of dilution effects, the matching criteria used to form the pools does not affect the expected number of false negatives. As to the other type of classification error, Aprahamian, Bish and Bish (2019) have shown that, in the absence of dilution effects, there exists an ordered partition that minimizes the expected number of false positives. So these results indicate that ordered partitions perform well in all of the three dimensions considered: expected number of tests, expected number of false positives and expected number of false negatives. We extend these results by showing that, when dilution effects are present but are sufficiently small, there exists an ordered partition that minimizes the expected number of tests and another ordered partition that minimizes the expected number of false positives, conditional on a given set of pool sizes. We show that, provided that an additional technical condition is met regarding the concavity of the function governing the dilution effect, pooling subjects according to an ordered partition is also optimal in terms of minimizing the expected number of false negatives. As a corollary, if all of these conditions are met and one is only considering partitions in which all pools have the same size, ordered partitions will always outperform any other partition with the same pool size configuration in all of the three attributes mentioned earlier: expected number of tests, expected number of false positives and expected number of false negatives. This result has practical applications to situations in which reconfiguring the testing machine for different pool sizes is time-consuming or impractical (e.g., Aprahamian, Bish and Bish (2020)).

We also characterize ordered partitions in terms of equity. Ideally one would want to implement a testing scheme that is fair, in the sense that it provides equitable expected payoffs to subjects. This is important because, depending on the matching criteria used to form the pools, subjects belonging to certain demographic groups can end up with a disproportionally high probability of being misclassified (either with a false negative or a false positive classification). We show that, conditional that all pools have the same size, ordered partitions do not always yield the most equitable allocation, even if they minimize both types of classification errors. But we find an instance in which ordered partitions are guaranteed to generate the most equitable allocation regardless of the concavity of the function governing the dilution effect and regardless of the distribution of priors: when all pools are comprised of only two subjects. For bigger pools, information on the dilution effect and empirical distribution of priors can be used to form sufficient conditions under which ordered partitions maximize equity.

We apply our results to the detection of chlamydia and hepatitis B through pooled testing. To do so, we first estimate the dilution effect for assays used to detect those diseases, as well as the prevalences of those diseases for different demographic groups, using data from previous studies. We then use those estimates to verify whether the hypotheses from our results are satisfied, and conduct numerical analysis comparing the performance of ordered vs random partitions (i.e., those partitions in which subjects are matched randomly). We find that, in general, the conditions that guarantee the existence of an ordered partition that minimizes the expected number of tests and the expected number of false positives are met. Though the sufficient conditions that guarantee that ordered partitions perform well in terms of minimizing the expected number of false negatives is not always met, our simulations indicate that ordered partitions tend to produce slightly less false negatives than random partitions. Our simulations also indicate that ordered partitions have a better performance than random partitions in terms of equity.

2. Related Literature

This paper is related to the work of Hwang (1975), McMahan, Tebbs and Bilder (2012) and Aprahamian, Bish and Bish (2019), which propose algorithms to determine the optimal partition used to perform the Dorfman procedure, so as to either minimize the expected number of tests (Hwang (1975) and McMahan, Tebbs and Bilder (2012)) or a convex combination of the expected number of tests and both types of classification errors (Aprahamian, Bish and Bish (2019)). But different from our setting, they follow the literature convention by assuming that pooled testing is not subject to dilution effects. In such an environment one can always find an optimal partition that is an ordered partition. But in the presence of dilution effects, an ordered partition may perform poorly in some attributes, such as in the minimization of false negative classifications. We show that if the dilution function satisfies certain properties that can be tested empirically, then one can always find an optimal partition that is an ordered partition.

This paper is also related to the work of Hwang (1976) and Wein and Zenio (1996) which vie to find optimal pool sizes for the Dorfman procedure when pooled testing is subject to dilution effects. Their work, however, follow the literature convention by assuming that the probability of infection is homogeneous across the population, whereas in our environment we allow subjects to have heterogeneous probability of infection.

Most research in the literature that simultaneously allow for dilution effects and heterogeneous priors are the ones dedicated to estimating the probability of infection conditional on the results of pooled tests, such as the work of Wang, McMahan and Gallagher (2015), Warasi et al. (2017) and Mokalled et al. (2021). To the best of our knowledge, the only theoretical research that attempt to formulate optimal matching schemes under the presence of both dilution effects and heterogeneous priors are the work of El-Amine, Bish and Bish (2017), Aprahamian, Bish and Bish (2020) and Aprahamian, Bish and Bish

(2018). El-Amine, Bish and Bish (2017) and Aprahamian, Bish and Bish (2020) propose a testing scheme in which subjects with similar probability of infection are matched together to form the pools. However, for tractability reasons they assume that the dilution effect only depends on the pool size, not on the number of subjects infected within the pool. More in line with our Aprahamian, Bish and Bish (2018) work in an environment in which the dilution effect is affected by the proportion of infected subjects within the pool. They present conditions under which an ordered partition minimizes the expected number of tests and the expected number of false negative classifications. However, different from our approach, the results they derive regarding ordered partitions do not allow samples that were detected as infected to be retested,[3] nor do they analyze the conditions under which an ordered partition minimizes the expected number of false positives and maximizes equity. They also require pool sizes to be equal, whereas we also derive results that apply to situations in which pool sizes are heterogeneous. Finally, we slightly relax the sufficient conditions derived in Aprahamian, Bish and Bish (2018) that guarantee that an ordered partition minimizes the expected number of false negatives. One practical application of this stronger result is that it implies that, for any realistic dilution effect (more precisely, for any dilution function in which the probability of detecting infection diminishes with the proportion of infected subjects within the pool), ordered partitions always minimize the expected number of false negatives conditional that all subjects are pooled into groups of size 2 (an analogous result follows for false positive classifications).

3. Environment

Suppose that there is a population $S = \{1, 2, \cdots, n\}$ of subjects to be tested. Each subject can either be infected or not infected (we use the terms *not infected* and *healthy* interchangeably). Each subject $i \in S$ is infected with probability $q_i \in [0, 1]$. Without loss of generality, we assume throughout the paper that

$$q_1 \leq q_2 \leq \cdots \leq q_n.$$

If a subject is individually tested and he is not infected, the test will classify him as healthy with probability $S_p \in [0, 1]$. An infected subject who is individually tested is classified as infected with probability $S_e \in [0, 1]$. Using the terminology from clinical trials, S_p corresponds to the *specificity* of the test, while S_e corresponds to its *sensitivity*. We assume that $S_e > 1 - S_p$, so that whenever a test detects infection, the subject is more likely to be infected than not.

In a Dorfman procedure, the subjects to be tested, S, are pooled into disjoint groups, and the samples from subjects belonging to the same group are merged and tested together. If the test detects infection for the pooled sample, then each subject within the pool is tested individually. If a group is comprised of only one subject, then this subject is only tested individually, without a followup test.

[3] Though their full specification allows for retesting, for tractability reasons, they treat the expected number of tests and false negatives from the first phase of tests separately from the results obtained in the second wave of tests.

Several criteria can be used to form the pools. One way is to group subjects randomly, without taking into account their prior probability of infection, into pools of equal size. But such pooling method does not take advantage of the information on subjects' prior probability of infection, nor does it take into account the fact that, when a subject's probability of infection is sufficiently high, it may be preferable to test that subject individually.

So next we consider a class of pooling schemes in which subjects are ordered from lowest to highest probability of infection, and those with similar probability of infection are matched together. Formally, a partition $\Omega^* = \{G_1^*, G_2^*, \cdots, G_m^*\}$ of S is said to be an *ordered partition* if, for any $g, w \in \{1, 2 \cdots, m\}$ with $g \neq w$ we have that either $q_i \geq q_j$ for all $i \in G_g^*$ and all $j \in G_w^*$ or $q_i \leq q_j$ for all $i \in G_g^*$ and all $j \in G_w^*$. In an *ordered pooling* scheme subjects are grouped according to an ordered partition.

We will compare this method to group subjects with alternative ones in terms of the expected number of tests they require to diagnose the entire population S, how many false negatives and false positives they generate (i.e., how many subjects are misclassified) and equity (i.e., how fair the final allocation is). But in order to make those comparisons, we must first make assumptions regarding how the probability of detecting infection in a pooled test is affected by the number of infected subjects within the pool.

Let $h(I,k)$ be the probability of detecting infection in a pooled sample collected from $k \in \mathbb{N}$ subjects, conditional that exactly $I \in \{0, 1, 2, \cdots, k\}$ of those subjects are infected. From our definition of sensitivity and specificity, we must have $h(1,1) = Se$ and $h(0,1) = 1 - S_p$. We will occasionally refer to h, as the *dilution function*. For each $k \in \mathbb{N}$, we assume that $h(\cdot, k)$ is (weakly) increasing, i.e., the more infected subjects there are in the pool, the more likely the pooled test will detect infection, which is arguably a very mild and reasonable assumption.

Assumption 1 *$h(I,k)$ is increasing in I (i.e., the more infected subjects there are in the group, the more likely the pooled test will detect infection for that group).*

Most of the literature assumes that pooled samples are not susceptible to dilution, so that the probability of detecting infection in a pooled sample that has at least one infected subject is the same as the probability of detecting infection from an individual test of an infected subject. This corresponds to the case in which

$$h(I,k) = \begin{cases} S_e, & \text{if } I > 0 \\ 1 - S_p, & \text{if } I = 0. \end{cases} \tag{1}$$

But in general, h could assume different formats. The faster $h(I,k)$ converges to S_e as I approaches k, the lower is the dilution effect. Figure 1 depicts different functions $h(\cdot, k)$ that satisfy assumption 1. The lower and more convex functions correspond to cases in which the dilution effect is stronger.

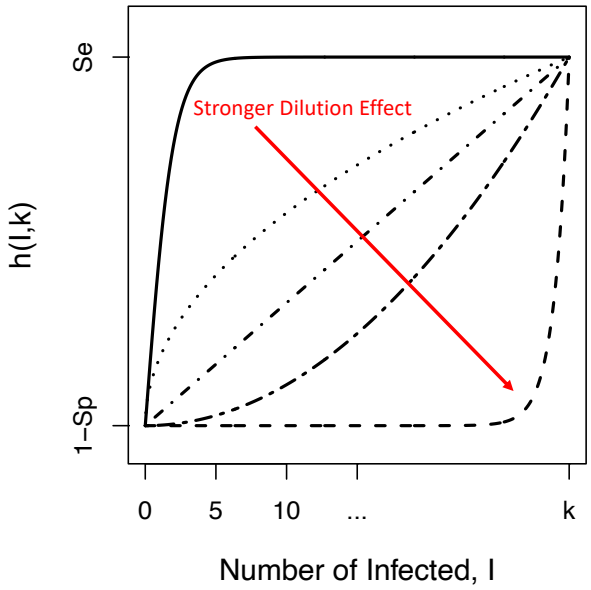


Figure 1 Different dilution functions $h(\cdot,k)$, for a given $k \in \mathbb{N}$. The dilution function is such that $h(0,k) = 1-S_p$ and $h(k,k) = S_e$.

4. Expected Number of Tests

Because testing is costly, ideally one would want to implement a partition that minimizes the expected number of tests required to screen the population. For a given partition $\Omega \equiv \{G_1, G_2, \cdots, G_{n/k}\}$ of the population S, we denote $T(\Omega)$ as the number of tests performed on the population after implementing the Dorfman procedure using this partition to determine the pools, and $\mathbb{E}[T(\Omega)]$ as its corresponding expectation.

It is straightforward to show that, for any partition $\Omega \equiv \{G_1, G_2, \cdots, G_m\}$ of S, the expected number of tests required to screen the population using the Dorfman procedure is given by

$$\mathbb{E}[T(\Omega)] \equiv \sum_{G_g \in \Omega} T_{G_g}, \tag{2}$$

where

$$T_{G_g} \equiv \begin{cases} 1, & \text{if } |G_g| = 1 \\ 1 + |G_g| \sum_{I=0}^{|G_g|} h(I,k) P_{G_g}(I), & \text{if } |G_g| > 1 \end{cases}, \tag{3}$$

$$P_{G_g}(I) \equiv \sum_{\substack{G \subseteq G_g \\ s.t. |G| = I}} \prod_{i \in G} q_i \prod_{j \in G_g \setminus G} (1 - q_j). \tag{4}$$

When $h(\cdot,k)$ is concave for all $k \in \mathbb{N}$, we show that, for any partition Ω, we can find an ordered partition Ω^* that preserves the pool sizes of Ω, and generates a weakly lower expected number of tests. Moreover, this partition is such that, whenever it individually tests a subject i with probability of infection q_i, it also individually tests all subjects with higher probability of infection.

Theorem 1 *Let* $\Omega = \{G_1, G_2, \cdots, G_m\}$ *be an arbitrary partition of* S*. Suppose that* $h(\cdot, |G_g|)$ *is concave for every* $G_g \in \Omega$*. Then, there exists an ordered partition* $\Omega^* = \{G_1^*, G_2^*, \cdots, G_m^*\}$ *such that*

a) $|G_g^*| = |G_g|$ *for all* $g \in \{1, 2, \cdots, m\}$.

b) Whenever a subject $i \in S$ *with probability of infection* q_i *is individually tested under* Ω^**, subjects with a probability of infection higher than* q_i *are also individually tested under* Ω^*.

c) $\mathbb{E}[T(\Omega^*)] \leq \mathbb{E}[T(\Omega')]$.

Intuitively, the parts a and c from theorem 1 can be explained as follows. Suppose that we were to match subjects randomly. In this case, each group would have a high probability of having at least one infected subject. A concave dilution function $h(\cdot, k)$ implies that the dilution effect is not too strong. If the dilution effect is sufficiently small, pooled tests would detect infection for most groups, resulting in many individual follow-up tests being conducted. If, on the other hand, ordered pooling was implemented, only the groups with high probability of infection would be likely to have at least one infected subject, and therefore to test positive for the disease, thus resulting in a lower number of follow-up tests.

As to part b from theorem 1, we have that, adding more infected subjects into a pool increases the probability that the pooled test detects infection, thus triggering followup tests for all the subjects within that pool. So to minimize the expected number of tests, those with highest probability of infection should be the ones, if any, allocated for individual testing.

Notice that theorem 1 requires $h(\cdot, k)$ to be concave for all different pool sizes from the original partition. This hypothesis will be satisfied if, for instance, the dilution function is assumed to depend only on the proportion of infected subjects within the pool, I/k, and such dilution function is concave. As an example, consider the following class of dilution functions previously used in the literature (e.g., Burns and Mauro (1987) and Aprahamian, Bish and Bish (2018))

$$h(I, k) = (1 - S_p) + (S_p + S_e - 1)\left(\frac{I}{k}\right)^{\delta}, \tag{5}$$

where $\delta \geq 0$. As long as $\delta \leq 1$, the function $h(\cdot, k)$ will be concave. In practice, the concavity of such function can be easily calibrated using empirical data on pooled tests (see section 7).

Theorem 1 guarantees that, for any partition Ω we can always find an ordered partition Ω^* that generates a weakly lower expected number of tests than Ω and preserves the pool sizes of Ω. So according to this theorem, if $\Omega = \{\{3,4\},\{1,2,5\}\}$, the ordered partition that generates less expected number of tests than Ω could be $\{\{1,2\},\{3,4,5\}\}$ or $\{\{1,2,3\},\{4,5\}\}$ (or both). But notice that, if all the pools of Ω have equal size, then there is essentially only one way to form an ordered partition that preserves the pool sizes of Ω, which implies, from theorem 1, that finding an ordered partition Ω^* such that $\mathbb{E}[T(\Omega^*)] \leq \mathbb{E}[T(\Omega)]$ is trivial. This has practical implications to situations in which the tester is interested in using the same pool size for all tests, say, because reconfiguring the pool sizes is too costly.

Corollary 1 *Suppose that $h(\cdot,k)$ is concave. Let Ω be a partition of $S=\{1,2,\cdots,n\}$ with $|G_g|=k\ \forall G_g\in\Omega$. Then, the following ordered partition of S*

$$\Omega^*=\{\{1,2,\cdots,k\},\{k+1,k+2,\cdots,2k\},\cdots,\{n-k+1,n-k+2,\cdots,n\}\}$$

is such that $\mathbb{E}[T(\Omega^*)]\leq\mathbb{E}[T(\Omega)]$.

5. Expected Number of False Negatives and False Positives

A false negative occurs when an infected subject is incorrectly classified as healthy, and a false positive occurs when a healthy subject is incorrectly classified as infected. For a given partition Ω of S, we denote $FN(\Omega)$ and $FP(\Omega)$ as the total number of false negatives and false positives, respectively, obtained after implementing the Dorfman procedure using the partition Ω.

It is straightforward to show that, for any partition $\Omega=\{G_1,G_2,\cdots,G_m\}$ of S, the expected number of false negatives and false positives obtained after implementing the Dorfman procedure using this partition is given by

$$\mathbb{E}[FN(\Omega)]=\sum_{G_g\in\Omega}FN_{G_g} \tag{6}$$

and

$$\mathbb{E}[FP(\Omega)]=\sum_{G_g\in\Omega}FP_{G_g}, \tag{7}$$

respectively, where, for each $G_g\in\Omega$

$$FN_{G_g}\equiv\begin{cases}(1-S_e)q_i, & \text{if } G_g=\{i\}\\ \sum_{I=0}^{|G_g|}P_{G_g}(I)I\left[1-h(I,|G_g|)S_e\right], & \text{if } |G_g|>1\end{cases},$$

$$FP_{G_g}\equiv\begin{cases}(1-S_p)(1-q_i), & \text{if } G_g=\{i\}\\ \sum_{I=0}^{|G_g|}P_{G_g}(I)h(I,|G_g|)(|G_g|-I)(1-S_p) & \text{if } |G_g|>1\end{cases}$$

and $P_{G_g}(I)$ is the probability that group $G_g\in\Omega$ has exactly I infected subjects (see equation 4).

In this section we compare the performance of ordered vs non-ordered partitions in terms of expected number of false negatives and false positives. But before we present general results, let us first look at the performance of ordered partitions for specific dilution functions.

It is a well known result that, for a given partition Ω, there always exists an ordered partition that preserves the pool sizes of Ω, and outperforms it in terms of minimizing the expected number of tests and expected number of false positives. Moreover, any other partition that preserves the pool sizes of Ω yield the same expected number of false negatives. Those results are stated formally in proposition 1.

Proposition 1 *(Aprahamian, Bish and Bish (2019)) Suppose that the dilution function is given by:*

$$h(I,k)=\begin{cases}S_e, & \text{if } I>0\\ 1-S_p, & \text{if } I=0.\end{cases}$$

Then,

a) For any partition Ω *of* S*, there exists an ordered partition* Ω^* *such that* $\mathbb{E}[T(\Omega^*)] \leq \mathbb{E}[T(\Omega)]$.

b) The probability that an infected subject is incorrectly diagnosed as healthy is the same across any partition Ω *of* S *such that* $|G_g| \geq 2$ *for all* $G_g \in \Omega$*, and is given by* $(1 - S_e^2)$.

c) For any partition Ω *of* S *such that* $|G_g| \geq 2$ *for all* $G_g \in \Omega$*, there exists an ordered partition* Ω^* *such that* $\mathbb{E}[FP(\Omega^*)] \leq \mathbb{E}[FP(\Omega)]$.

So from proposition 1, we can see that, in the absence of dilution effects, ordered partitions seem like the optimal choice, as they may potentially minimize $\mathbb{E}[T(\Omega)]$, $\mathbb{E}[FN(\Omega)]$ and $\mathbb{E}[FP(\Omega)]$.

Next, we consider the case in which $h(\cdot, k)$ is affine, i.e., when $h(I,k) = a + bI$. It can be shown that, in this case, an ordered partition also performs weakly better than any partition in which all pool sizes are equal in all of the three dimensions considered, namely, $\mathbb{E}[T(\Omega)]$, $\mathbb{E}[FN(\Omega)]$ and $\mathbb{E}[FP(\Omega)]$.

Proposition 2 *Suppose that the dilution function is given by* $h(I,k) = a + bI$*. Then,*

a) For any pair of partitions Ω *and* Ω' *of* S *such that* $|G_g| = k$ *for all* $G_g \in \Omega \cup \Omega'$*, we must have* $\mathbb{E}[T(\Omega)] \leq \mathbb{E}[T(\Omega')]$.

b) If Ω *is a partition of* S *such that* $|G_g| = k$ *for all* $G_g \in \Omega$*, and* Ω^* *is an ordered partition of* S *such that* $|G_g^*| = k$ *for all* $G_g^* \in \Omega$*, then* $\mathbb{E}[FN(\Omega^*)] \leq \mathbb{E}[FN(\Omega)]$ *and* $\mathbb{E}[FP(\Omega^*)] \leq \mathbb{E}[FP(\Omega)]$.

Finally, we consider the extreme case in which the dilution effect is so strong that the probability of detecting infection from a pool comprised of at least one healthy subject is given by $1 - S_p$, the same probability of detecting infection from a healthy subject through an individual test. In this case we show that, conditional that all pools have the same size, ordered partitions generates the *highest* expected number of tests, though it minimizes the expected number of false negatives and generates the same expected number of false positives as any other pooling criteria.

Proposition 3 *Suppose that the dilution function is given by*

$$h(l,k) = \begin{cases} 1 - S_p, & \text{if } I < k \\ S_e, & \text{if } I = k. \end{cases}$$

Then,

a) If Ω^* *is an ordered partition of* S *such that* $|G_g^*| = k$ *for all* $G_g^* \in \Omega^*$*, and* Ω *is any partition of* S *such that* $|G_g| = k$ *for all* $G_g \in \Omega$*, then* $\mathbb{E}[T(\Omega^*)] \geq \mathbb{E}[T(\Omega)]$.

b) If Ω *is a partition of* S *such that* $|G_g| = k$ *for all* $G_g \in \Omega$*, and* Ω^* *is an ordered partition of* S *such that* $|G_g^*| = k$ *for all* $G_g^* \in \Omega$*, then* $\mathbb{E}[FN(\Omega^*)] \leq \mathbb{E}[FN(\Omega)]$.

c) The probability that a healthy subject is incorrectly diagnosed as infected is the same across any partition Ω *of* S *such that* $|G_g| \geq 2$ *for all* $G_g \in \Omega$*, and is given by* $(1 - S_p)^2$.

While it is very unlikely that pooled testing schemes are ever considered as a good alternative to individual testing under such an extreme dilution effect, proposition 3 helps to illustrate that a partition used to form the pools may perform well in one dimension (e.g., minimizing false negatives), but poorly in others (e.g., does not minimize the expected number of tests).

We now derive general sufficient conditions under which ordered pooling minimizes each type of classification error.[4]

Hypothesis 1 *Suppose that the dilution function $h(\cdot,k)$ is such that, for all $I \in \{1,2,\cdots,k-1\}$,*

$$\frac{I+1}{2I}h(I+1,k)+\frac{I-1}{2I}h(I-1,k) \geq h(I,k).$$

Theorem 2 *Let $\Omega = \{G_1, G_2, \cdots, G_m\}$ be an arbitrary partition of S such that $|G_g| \geq 2$ for all $G_g \in \Omega$. Suppose that hypothesis 1 holds for every pool size $k \in \{k' \in \mathbb{N}; k' = |G_g|$, for some $G_g \in \Omega\}$. Then, there exists an ordered partition $\Omega^* = \{G_1^*, G_2^*, \cdots, G_m^*\}$ such that*

a) $|G_g^*| = |G_g|$ for all $g \in \{1,2,\cdots,m\}$,

b) $\mathbb{E}[FN(\Omega^*)] \leq \mathbb{E}[FN(\Omega)]$.

Hypothesis 1 can be interpreted as requiring that the dilution function $h(\cdot,k)$ is not "too concave". In fact, hypothesis 1 is satisfied whenever $h(\cdot,k)$ is convex, as convexity of $h(\cdot,k)$ implies that

$$\frac{I+1}{2I}h(I+1,k)+\frac{I-1}{2I}h(I-1,k) \geq h(I+1/I,k) \geq h(I,k).$$

But convexity is not a necessary condition for hypothesis 1 to hold. Indeed, for any $\delta \in (0,1)$, the dilution function introduced in equation 5 is not convex, and yet it satisfies hypothesis 1.

While Aprahamian, Bish and Bish (2019) have shown that, in the absence of dilution effects, allocating those with higher probability of infection to be individually tested minimizes false negative classifications, this is no longer the case under the presence of dilution effects. Indeed, consider the following counter-example: suppose that the dilution function is given by expression 5, with $\delta = 1/2$, $S_e = .97$ and $S_p = .95$. Then, if $S = \{1,2,3\}$ and $q_1 = .1$, $q_2 = .9$ and $q_3 = .99$ we have that

$$\mathbb{E}[FN(\{\{1\},\{2,3\}\})] = 0.143 < 0.303 = \mathbb{E}[FN(\{\{1,2\},\{3\}\})].$$

This is why theorem 2 requires all pools to have at least two subjects, as it is no longer straightforward to determine who should be individually tested to minimize false negative classifications. But notice

[4] Aprahamian, Bish and Bish (2018) had previously shown that, when all the pools have the same size $k \geq 2$, and the following condition holds

$$I\frac{\partial^2 h(I,k)}{\partial I^2}+2\frac{\partial h(I,k)}{\partial I} \geq 0 \quad \forall I \in [0,k], \tag{8}$$

then grouping subjects according to an ordered partition minimizes the expected number of false negatives. In section B from the Appendix we show that this condition implies that hypothesis 1 holds, but the converse is not necessarily true.

that theorem 2 can be relaxed to allow for individual testing in the following way: we can show that, for any partition Ω of S, there is a partition Ω^* that does not change the set of subjects who are individually tested under Ω, but implements ordered pooling for the remaining subjects, and is such that $\mathbb{E}[FN(\Omega^*)] \leq \mathbb{E}[FN(\Omega)]$. In this case, the partition Ω^* that generates less false negatives than Ω could be interpreted as a "*semi-ordered*" partition.

Corollary 2 *Let $\Omega = \{G_1, G_2, \cdots, G_m\}$ be an arbitrary partition of S. Suppose that hypothesis 1 holds for every pool size $k \in \{k' \in \mathbb{N}; k' = |G_g|$, for some $G_g \in \Omega\}$. Then, there exists a partition $\Omega^* = \{G_1^*, G_2^*, \cdots, G_m^*\}$ such that*

a) $\{i \in S; \{i\} \in \Omega\} = \{i \in S; \{i\} \in \Omega^*\}$ *(i.e., the set of subjects who are individually tested under Ω and Ω^* are the same).*

b) For every $G_g^*, G_w^* \in \Omega^*$ such that $G_g^* \neq G_w^*$, $|G_g^*| \geq 2$ and $|G_w^*| \geq 2$, we have that either $q_i \geq q_j$ for all $i \in G_g^*$ and all $j \in G_g^*$, or $q_i \leq q_j$ for all $i \in G_g^*$ and all $j \in G_g^*$.

c) $|G_g^*| = |G_g|$ *for all* $g \in \{1, 2, \cdots, m\}$,

d) $\mathbb{E}[FN(\Omega^*)] \leq \mathbb{E}[FN(\Omega)]$.

We now derive conditions under which ordered pooling minimizes false positive classifications. Those conditions will go in the opposite direction of the conditions we required for ordered pooling to minimize false negative classifications: instead of requiring the dilution function not to be "too concave", we now require it not to be "too convex".

Hypothesis 2 *Suppose that the dilution function $h(\cdot, k)$ is such that, for all $I \in \{1, 2, \cdots, k-1\}$,*

$$\frac{k-I-1}{2(k-I)} h(I+1, k) + \frac{k-I+1}{2(k-I)} h(I-1, k) \leq h(I, k).$$

Theorem 3 *Let $\Omega = \{G_1, G_2, \cdots, G_m\}$ be an arbitrary partition of S. Suppose that hypothesis 2 holds for every pool size $k \in \{k' \in \mathbb{N}; k' = |G_g|$, for some $G_g \in \Omega\}$. Then, there exists an ordered partition $\Omega^* = \{G_1^*, G_2^*, \cdots, G_m^*\}$ such that*

a) $|G_g^*| = |G_g|$ *for all* $g \in \{1, 2, \cdots, m\}$,

b) Whenever a subject $i \in S$ with probability of infection q_i is individually tested under Ω^*, subjects with a probability of infection higher than q_i are also individually tested under Ω^*.

c) $\mathbb{E}[FP(\Omega^*)] \leq \mathbb{E}[FP(\Omega)]$.

Notice that if $h(\cdot, k)$ is concave, then hypothesis 2 holds, as concavity of $h(\cdot, k)$ implies that

$$\frac{k-I-1}{2(k-I)} h(I+1, k) + \frac{k-I+1}{2(k-I)} h(I-1, k) \leq h(I - \frac{1}{k-I}, k) \leq h(I, k).$$

Corollary 3 *Let $\Omega = \{G_1, G_2, \cdots, G_m\}$ be an arbitrary partition of S such that $|G_g| \geq 2$ for all $G_g \in \Omega$. Suppose that $h(\cdot, k)$ is concave for every $k \in \{k' \in \mathbb{N}; k' = |G_g|$, for some $G_g \in \Omega\}$. Then, there exists an ordered partition $\Omega^* = \{G_1^*, G_2^*, \cdots, G_m^*\}$ such that*

a) $|G_g^*| = |G_g|$ *for all* $g \in \{1, 2, \cdots, m\}$,

b) $\mathbb{E}[FP(\Omega^*)] \leq \mathbb{E}[FP(\Omega)]$.

Intuitively, when the dilution effect is relatively small, matching subjects randomly to form the pools is likely to generate many false positives, as each group will have a high chance of having at least one infected subject who will trigger a follow-up test with high probability. If the individual tests have imperfect specificity, those follow-up tests will increase the probability that non-infected subjects are incorrectly classified as infected. Performing ordered pooling, on the other hand, results in less follow-up tests, as this matching criterion reduces the frequency with which a pooled sample is "contaminated" with an infected specimen that triggers a follow-up test.

Notice that, when we are only considering pools of size $k = 2$, we only need to check whether hypotheses 1 or 2 hold for $I = 1$. Because $h(\cdot, k)$ is increasing, both of these hypotheses are clearly satisfied at $I = 1$. So a corollary to theorems 2 and 3 is that, if all the pools from a partition have size $k = 2$, then an ordered partition in which all pools have size $k = 2$ will produce a lower expected number of false negatives and false positives.

Corollary 4 *Let* $\Omega = \{G_1, G_2, \cdots, G_m\}$ *be a partition of* S *such that* $|G_g| = 2$ *for all* $G_g \in \Omega$. *Suppose that* $h(\cdot, 2)$ *is increasing. Then, there exists an ordered partition* $\Omega^* = \{G_1^*, G_2^*, \cdots, G_m^*\}$ *such that*

a) $|G_g^*| = 2$ *for all* $G_g \in \Omega$,

b) $\mathbb{E}[FN(\Omega^*)] \leq \mathbb{E}[FN(\Omega)]$,

c) $\mathbb{E}[FP(\Omega^*)] \leq \mathbb{E}[FP(\Omega)]$.

Finally, notice that a corollary from theorems 1, 2 and 3 is that, if we are only considering partitions in which all pools are required to have the same size k that is a dividend of the population size n, finding the optimal partition when the dilution function is concave and hypothesis 2 holds is trivial: it consists in implementing the following ordered partition of S.

$$\Omega^* = \{\{1, 2, \cdots, k\}, \{k+1, k+2, \cdots, 2k\}, \cdots, \{n-k+1, n-k+2, \cdots, n\}\}.$$

Corollary 5 *Suppose that* $h(\cdot, k)$ *is concave and satisfies hypothesis 2. Let* Ω *be any partition of* $S = \{1, 2, \cdots, n\}$ *with* $|G_g| = k \ \forall G_g \in \Omega$. *Then, the following ordered partition of* S

$$\Omega^* = \{\{1, 2, \cdots, k\}, \{k+1, k+2, \cdots, 2k\}, \cdots, \{n-k+1, n-k+2, \cdots, n\}\}$$

is such that

a) $\mathbb{E}[T(\Omega^*)] \leq \mathbb{E}[T(\Omega)]$,

b) $\mathbb{E}[FN(\Omega^*)] \leq \mathbb{E}[FN(\Omega)]$,

c) $\mathbb{E}[FP(\Omega^*)] \leq \mathbb{E}[FP(\Omega)]$.

Corollary 5 has practical implications to situations in which the tester is interested in using the same pool size for all tests, say, because reconfiguring the pool sizes is too costly.

In section J from the appendix, we compare the performance of ordered vs. random pooling for different combinations of dilution functions and distribution of priors, assuming that pool sizes are homogeneous. All the simulations are consistent with the propositions proven in this section.

6. Equity

Ideally one would want to implement a testing scheme that is fair, in the sense that it provides equitable expected payoffs to subjects. This is important because, depending on the matching criteria used to form the pools, subjects belonging to certain demographic groups can end up with a disproportionally high probability of being misclassified (either with a false negative or a false positive classification). In this section we show that ordered pooling does not always implement the most equitable allocation when all pool sizes are required to be equal, even if it minimizes the overall expected number of false negative and false positive classifications. However, we show that when all pools are of size $k = 2$ and subjects either only care about false negative classifications or only care about false positive classifications, then ordered pooling is guaranteed to generate the most equitable allocation regardless of the dilution function and the distribution of priors. For pools of size $k > 2$, information on the dilution effect and distribution of priors can be used to construct sufficient conditions under which ordered pooling maximizes subjects' welfare function.

Assuming the Dorfman procedure is applied to a partition $\Omega \equiv \{G_1, G_2, \cdots, G_{n/k}\}$ of S, we denote $FN_i(\Omega)$ as an indicator variable that equals 1 whenever subject $i \in S$ is incorrectly classified as healthy, and zero otherwise. Similarly, $FP_i(\Omega)$ is an indicator variable that equals 1 whenever subject $i \in S$ is incorrectly classified as infected.

Letting $\theta \in [0,1]$ be the weight attributed to false negative classifications, and $(1-\theta)$ be the weight attributed to false positive classifications, we have that subject i's expected utility from this partition is given by

$$u_i(\Omega) \equiv 1 - \theta \mathbb{E}[FN_i(\Omega)] - (1-\theta)\mathbb{E}[FP_i(\Omega)].$$

Throughout this section we consider the *utilitarian max-min* welfare function to measure equity. As the name suggests, a utilitarian max-min welfare function is a convex combination of the utilitarian and max-min welfare functions. The utilitarian welfare function, popularized by Harsanyi (1955), is simply the sum of utilities from the agents in the economy, while the max-min welfare function, proposed by Rawls (1971), equals to the lowest utility in the economy. So a utilitarian measure of welfare values efficiency, while a max-min measure values equity. A utilitarian max-min measure puts a positive weight on

both measures. More precisely, for a given parameter $\alpha \in [0,1]$, the utilitarian max-min welfare function is given by

$$\pi_\alpha(u_1, u_2, \cdots, u_n) \equiv \alpha \min_{i \in S}(u_i) + (1-\alpha)\sum_{i \in S} u_i. \tag{9}$$

So the parameter α can be interpreted as the weight put on equity, while $1-\alpha$ is the weight put on efficiency.

The utilitarian max-min welfare function is known to satisfy desirable axioms (e.g., see Schneider and Kim (2020) and Bossert and Kamaga (2020)). But for our purposes, the main reason why we invoke this welfare function is because it satisfies the following property: if an allocation maximizes the utilitarian welfare function *and* the max-min welfare function, then it maximizes the utilitarian max-min welfare function for any parameter α. This property is not shared by some additive social welfare functions, such as the family of α*-fairness* welfare functions,[5] even though the utilitarian and the max-min welfare functions correspond to the extreme opposite instances of the α-fairness family, with the utilitarian approach putting all weight on efficiency, and the max-min putting all weight on equity (e.g., Lan et al. (2010) and Bertsimas, Farias and Trichakis (2012)).

The following example shows that, ordered pooling does not necessarily maximize social welfare, even when this pooling scheme minimizes both the expected number of false positives and expected number of false negatives.

Example 1 *Consider the dilution function* $h(I,k) = (1-S_p) + (S_p + S_e - 1)\left(\frac{I}{k}\right)^{1/10}$*, with* $S_p = .95$ *and* $S_e = .9$*. Figure 2 plots this dilution function when* $k = 3$ *(i.e., when the pool has 3 subjects).*

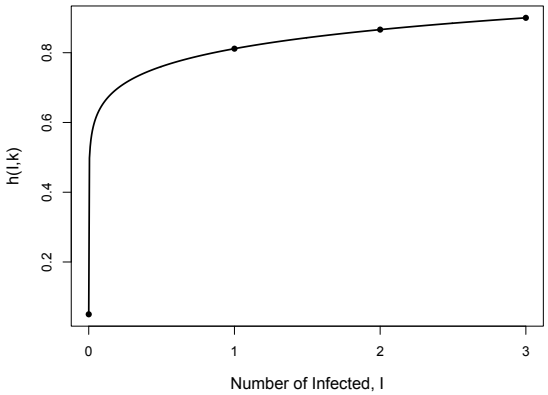


Figure 2 Dilution function $h(I,k) = (1-S_p) + (S_p + S_e - 1)\left(\frac{I}{k}\right)^{1/10}$, with $S_p = .95$, $S_e = .9$ and $k = 3$.

[5] Indeed, consider $u = (u_1, u_2, u_3, u_4) = (1,1,2,4)$ and $u = (\tilde{u_1}, \tilde{u_2}, \tilde{u_3}, \tilde{u_4}) = (36/37, 36/37, 3, 3)$. Then clearly $\min u_i > \min \tilde{u_i}$ and $\sum u_i > \sum \tilde{u_i}$. However $\sum \frac{u_i^{1-\alpha}}{1-\alpha} < \sum \frac{\tilde{u_i}^{1-\alpha}}{1-\alpha}$ for $\alpha = 2$.

Because $h(\cdot,k)$ is concave, we have from theorem 1 and corollary 3 that ordered pooling minimizes the expected number tests and expected number of false positives. Moreover one can easily show that this dilution function satisfies hypothesis 1, which, from theorem 2, implies that ordered pooling also minimizes the expected number of false negatives.

Now suppose that the utilitarian max-min welfare function 9 is such that $\alpha = 1$, so that the partition that maximizes the social welfare function is the one that maximizes the payoff from the subject with lowest expected utility. Also suppose that $\theta = 1$, so that subjects only put weight on the probability that they get a false negative result.

Now consider a population with the following vector of probability of infection:

$$q = (.01, .015, .95, .96, .98, .99).$$

Then, if we use the ordered partition, i.e., $\{\{1,2,3\},\{4,5,6\}\}$, the minimum payoff is the one from subject 3 and it is equal to 0.745. If, on the other hand, the partition $\{\{4,2,3\},\{1,5,6\}\}$ is used to group subjects, then the minimum payoff is the one from subject 6 and it is equal to 0.781. So ordered pooling does not maximize social welfare.

Similarly, if $\theta = 0$, so that subjects only care about false positive errors, and the vector of probability of infection is given by

$$q = (.001, .01, .02, .03, .98, .99),$$

then the ordered partition, i.e., $\{\{1,2,3\},\{4,5,6\}\}$, generates a minimum expected payoff of 0.956, while the partition $\{\{1,2,5\},\{3,4,6\}\}$ generates a higher minimum payoff of 0.957.

We find some instances, however, in which ordered pooling is guaranteed to maximize social welfare. Indeed, when all pools are required to be of size $k = 2$ and all the weight on individual utilities is put either on false negative or false positive classifications (i.e., $\theta = 1$ or $\theta = 0$), then ordered pooling maximizes social welfare.

Proposition 4 *Suppose that $k = 2$ and $\theta \in \{0,1\}$. Let $\Omega = \{G_1, G_2, \cdots, G_m\}$ be a partition of S such that $|G_g| = 2$ for all $G_g \in \Omega$. Suppose that $h(\cdot,2)$ is increasing. Then, there exists an ordered partition $\Omega^* = \{G_1^*, G_2^*, \cdots, G_m^*\}$ such that*
a) $|G_g^*| = 2$ for all $G_g \in \Omega$,
b) $\pi_\alpha(u_1(\Omega^*), u_2(\Omega^*), \cdots, u_n(\Omega^*)) \geq \pi_\alpha(u_1(\Omega), u_2(\Omega), \cdots, u_n(\Omega))$ for all $\alpha \in [0,1]$.

When $k > 2$, we may use information on the distribution of types to determine whether ordered pooling maximizes social welfare.

Proposition 5 *Let $\Omega^* = \{G_1^*, G_2^*, \cdots, G_m^*\}$ be an ordered partition of $S = \{1,2,\cdots,n\}$ such that $|G_g^*| = k \; \forall G_g^* \in \Omega^*$, and let $\Omega = \{G_1, G_2, \cdots, G_{n/k}\}$ be any other partition such that $|G_g| = k \; \forall G_g \in \Omega$.*

a) If $\theta = 1$, *then* $\min_{i \in G_j} u_i(\Omega) \leq u_n(\Omega^*)$.

b) If $\theta = 0$, *then* $\min_{i \in G_j} u_i(\Omega) \leq u_1(\Omega^*)$.

Notice that if the dilution function satisfies hypotheses 1 and 2, theorems 2 and 3 imply that ordered pooling minimizes both the expected number of false negatives and the expected number of false positives. In this case, a sufficient condition for ordered pooling to maximize the social welfare function is that it maximizes the utility from the subject with lowest utility. From proposition 5, this will be the case whenever $\theta = 1$ and the subject with lowest utility under ordered pooling happens to be the subject with *highest* probability of infection, or when $\theta = 0$ and the subject with lowest utility under ordered pooling happens to be the subject with *lowest* probability of infection.

Theorem 4 *Let* $\Omega^* = \{G_1^*, G_2^*, \cdots, G_m^*\}$ *be an ordered partition of* $S = \{1, 2, \cdots, n\}$ *such that* $|G_g^*| = k \ \forall G_g^* \in \Omega^*$, *and let* $\Omega = \{G_1, G_2, \cdots, G_m\}$ *be any other partition such that* $|G_g| = k \ \forall G_g \in \Omega'$. *Suppose that hypotheses 1 and 2 hold. Then:*

a) If $\theta = 1$ *and* $\min_{i \in S} u_i(\Omega^*) = u_n(\Omega^*)$, *then* $\pi_\alpha(u_1(\Omega^*), u_2(\Omega^*), \cdots, u_n(\Omega^*)) \geq \pi_\alpha(u_1(\Omega), u_2(\Omega), \cdots, u_n(\Omega))$ *for all* $\alpha \in [0, 1]$.

b) If $\theta = 0$ *and* $\min_{i \in S} u_i(\Omega^*) = u_1(\Omega^*)$, *then* $\pi_\alpha(u_1(\Omega^*), u_2(\Omega^*), \cdots, u_n(\Omega^*)) \geq \pi_\alpha(u_1(\Omega), u_2(\Omega), \cdots, u_n(\Omega))$ *for all* $\alpha \in [0, 1]$.

For situations in which the minimum utility obtained under ordered pooling does not coincide with the upper bounds derived in proposition 5, we can still conduct simulations to see how close those utilities are to those upper bounds. If they are close enough and, in addition, hypotheses 1 and 2 both hold, then ordered pooling is expected to maximize social welfare for most $\alpha \in [0, 1]$.

7. Case Study: Chlamydia Screening in the United States

Chlamydia is the most reported sexually transmitted disease in the US. It affects mostly the young and sexually active, and symptomatic cases are more frequent among women. Though in general it can be cured with antibiotics, if left untreated the disease can cause severe sequelae in women, including Pelvic Inflammatory Disease (PID), ectopic pregnancy and infertility (Centers for Disease Control and Prevention (2019*a*)).

In this section we compare the performance of ordered vs random partitions when using the Dorfman procedure to screen for Chlamydia. We also compare the performance of these two partitions with individual testing. Each of these three procedures differs in terms of their expected number of tests, expected number of false positives, and expected number of false negatives. For example, it can be shown that, for any dilution function $h(I, k)$ that is increasing in I, individual testing always generates the lowest expected number of false negatives compared to any pooling criteria used to implement the Dorfman procedure.

Proposition 6 *If the dilution function satisfies assumption 1 (i.e., if $h(\cdot,k)$ is increasing for every $k \in \mathbb{N}$), then testing all subjects individually minimizes the expected number of false negatives.*

Individual testing tends, however, to be outperformed by pooled testing in terms of the expected number of tests and expected number of false positives. So in order to compare the performance of each of these testing schemes, we need to attribute weights to each of the aforementioned attributes: i.e., weights to the expected number of tests, the expected number of false positives, and the expected number of false negatives. Following Aprahamian, Bish and Bish (2019), we use the average between the cost of sequelae for men and women, estimated by Owusu-Edusei Jr et al. (2015), to set the cost of a false negative at $2,927. Also following Aprahamian, Bish and Bish (2019), we set the screening cost per test at $55, and the cost of a false positive equal to the cost of a an additional screening test ($55). So the total expected cost of implementing a certain matching procedure is measured by

$$\mathbb{E}[C(\Omega)] \equiv \lambda_1 \mathbb{E}[FN(\Omega)] + \lambda_2 \mathbb{E}[FP(\Omega)] + \lambda_3 \mathbb{E}[T(\Omega)],$$

with $\lambda_1 = \$2,927$ and $\lambda_2 = \lambda_3 = \$55$.

The prevalence per demographic group was extracted from the Centers of Disease and Control Prevention (CDC) website for the year 2014.[6] Following Aprahamian, Bish and Bish (2019), we multiply the prevalence of each group reported by CDC by 3, in order to account for under-reporting.[7] Table 1 reports the prevalence of Chlamydia for each group in the sample.

Gender	Race/ Ethnicity	Age Group (years)	Risk (prevalence) (%)	Proportion in general population (%)
Female	Hispanic	15-24	6.54	1.41
		Other	0.65	7.01
	Black	15-24	19.19	1.07
		Other	1.22	5.67
	Other	15-24	4.38	4.29
		Other	0.25	31.31
Male	Hispanic	15-24	1.78	1.53
		Other	0.36	7.16
	Black	15-24	7.45	1.09
		Other	1.05	5.08
	Other	15-24	1.20	4.51
		Other	0.17	29.87

Table 1 Prevalence of Chlamydia and proportion in population by Gender, Age and race/Ethnicity (Aprahamian, Bish and Bish (2019) and Centers for Disease Control and Prevention 2014).

[6] https://wonder.cdc.gov/std-race-age.html

[7] In part, many end up not being screened because they exhibit no symptoms: approximately 75% of women and 50% of infected men exhibit no symptoms (Centers for Disease Control and Prevention (2000)).

To estimate the dilution effect we used the same calibration employed by Aprahamian, Bish and Bish (2018), who used data on the sensitivity of LCR pooled tests conducted by Kacena et al. (1998*a*) to obtain an approximation to the dilution function. This resulted in a dilution function of

$$\hat{h}(I,k) = (1-S_p) + (S_e + S_p - 1)(I/k)^{\delta}, \tag{10}$$

with $S_p = .98$, $S_e = 1$ and $\delta = 0.014302$. We plot this dilution function for $k = 10$ in figure 3. Details on the methodology used for this calibration is provided in section I from the appendix.

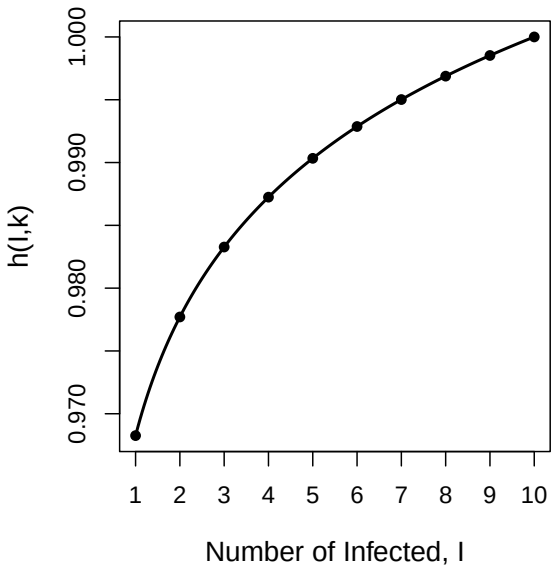


Figure 3 Calibrated dilution function for Chlamydia with parameters $S_p = .98$, $S_e = 1$ and $\delta = 0.014302$, and pool size $k = 10$.

It can be shown that the calibrated dilution function from equation (10) is concave for any $\delta \in [0,1]$ and any $S_p, S_e \in [0,1]$ such that $S_e > 1 - S_p$. Moreover, for any $\delta \geq 0$, this dilution function satisfies hypothesis 1. Therefore, if $\delta = 0.014302$, theorems 1, 2 and 3 imply that an ordered partition minimizes the expected number of tests, the expected number of false negatives and the expected number of false positives.

Conducting simulations using the data from table 1, we can confirm this qualitative result. Figures 9 to 4c display the performance of ordered pooling, random pooling and individual testing in terms of $\mathbb{E}(T(\Omega))$, $\mathbb{E}(FN(\Omega))$ and $\mathbb{E}(FP(\Omega))$, respectively, when all pools have the same size k and the total number of subjects to be tested was given by $n = 10,000$. Whenever the pool size k was not a dividend of the population size n, we grouped the subjects with highest probability of infection into the remainder of the division n/k to form the ordered partition. From these figures, we can see that ordered pooling performs better than random pooling in all of the three dimensions considered, thus resulting in a lower overall expected cost.

Consistent with proposition 6, individual testing generates less false negatives than both random pooling and ordered pooling. However, because individual testing requires more tests and produces

more false positives, it generates a higher total expected cost compared to both random and ordered pooling.

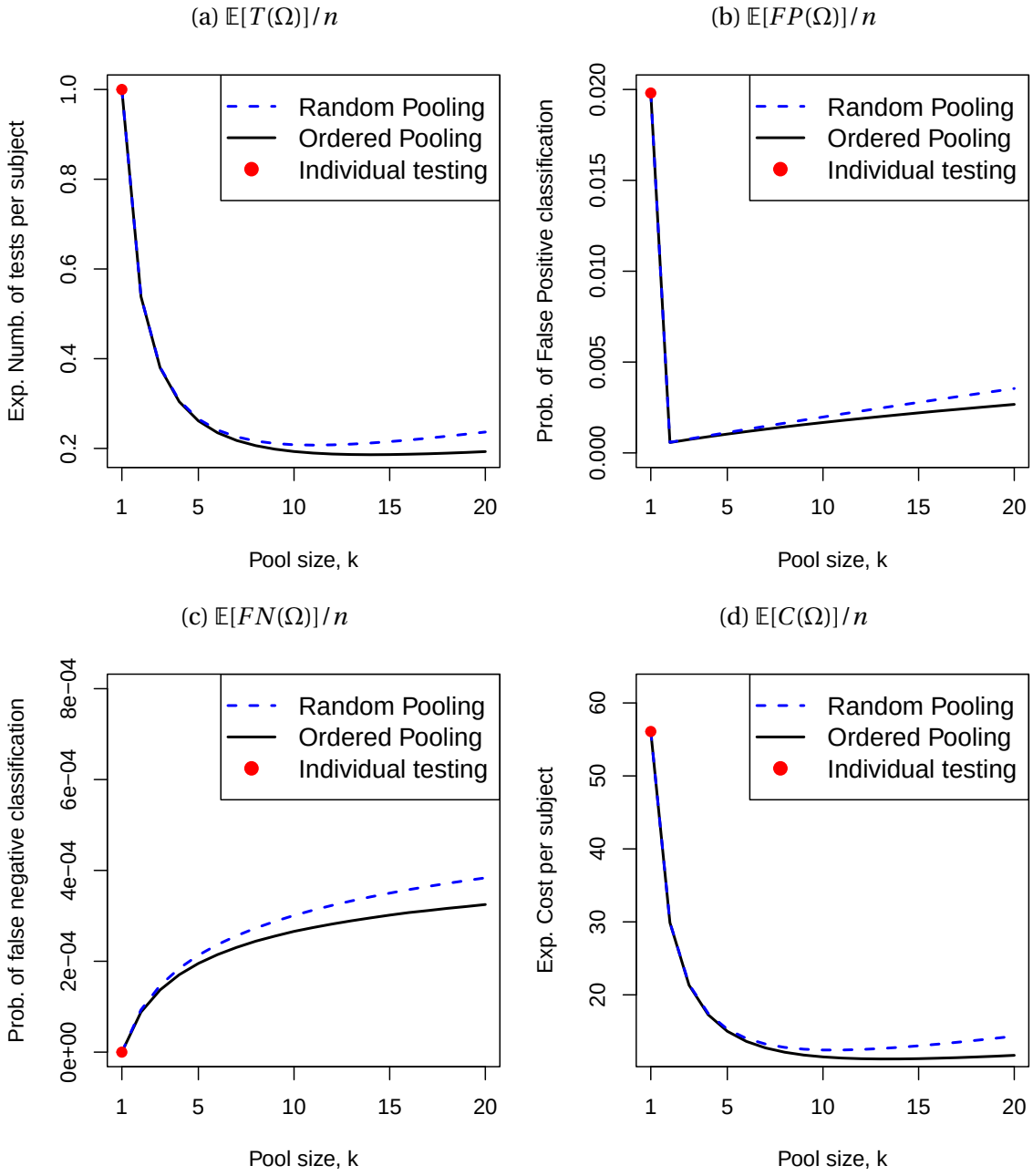


Figure 4 Expected number of tests and classification errors per subject for Chlamydia tests using ordered and random pooling. The red dot corresponds to the case in which subjects are individually tested, i.e. $k = 1$. The red dot corresponds to the case in which subjects are individually tested, i.e. $k = 1$.

Regarding equity, for each $k \geq 2$ we used proposition 5 to compute the upper bound from the minimum utility from any match when $\theta = 1$ and $\theta = 0$, and compared it with the minimum utility obtained

under ordered pooling vs random pooling. As displayed in figures 5a and 5b, for all $k \geq 2$, the average minimum utility obtained under ordered pooling was very close to the upper bound, and significantly above the minimum utility obtained under random matching. This, combined with the fact that ordered pooling maximizes the utilitarian welfare function under the dilution function $\hat{h}$, is indicative that ordered pooling should yield high welfare for any $\alpha \in [0,1]$.

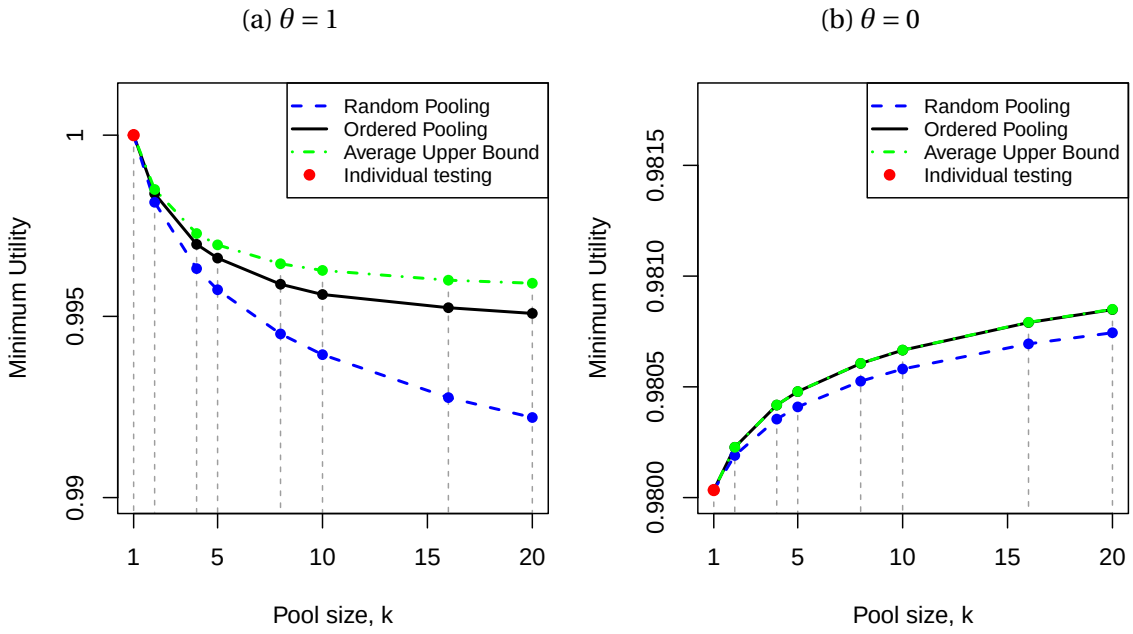


Figure 5 Minimum utility of ordered pooling vs random pooling for Chlamydia tests, for values of $k \leq 20$ that are dividends of $n = 10,000$.

8. Case Study: Hepatitis B Screening in Irish prisons

In this section we estimate the dilution effect of pooled testing for Hepatitis B using surveyed data with information on continuous biomarker readings for individual tests. We then compare the performance of random vs ordered pooling when applied to this dataset.

Hepatitis B is a liver infection disease caused by the hepatitis B virus (HBV). In 2015, 257 million people worldwide were estimated to be living with chronic hepatitis B World Health Organization (2017). Left untreated, this disease can lead to deadly sequelae, most commonly cirrhosis and hepatocellular carcinoma (primary liver cancer). In 2015 approximately 890,000 people were estimated to have died due to complications from hepatitis B World Health Organization (2017). Because hepatitis B can be transmitted by blood, the WHO recommends that all blood donations be tested for hepatitis B. Due to budget constraints, it is common for blood centers to perform pooled tests on multiple donors to detect HBV (El-Amine, Bish and Bish (2017)).

To estimate the dilution effect of pooled testing for hepatitis B, we use the dataset from a survey conducted in 1998 on 5 different prisons from Ireland, in which inmates were individually tested for Hepatitis B and other infectious diseases to determine risk factors for infection. The results of this survey have been summarized by Allwright et al. (2000). Multiple followup studies have used this dataset to estimate the dilution effect of pooled testing for Hepatitis B, including Wang, McMahan and Gallagher (2015), Warasi et al. (2017) and Mokalled et al. (2021).

The dataset contains a classification of whether an inmate was infected or not with Hepatitis B, as well as the Optical Density (OD) readings from a Murex ICE enzyme immunoassay (on oral fluid samples) used to detect the presence of the core antigen for hepatitis B (HBcAg).

In total there were 99 infected, and 1038 non infected subjects. Figure 6 displays the histogram of OD readings from both infected and non-infected subjects. While the distribution of OD readings for the non-infected inmates exhibits a unimodal format, the distribution of OD readings for infected inmates is bimodal. This is consistent with other studies in the literature that indicate a bimodal distribution of biomarkers readings for those infected with HBV (e.g., Downs et al. (2020) and Alcalde et al. (2009)).

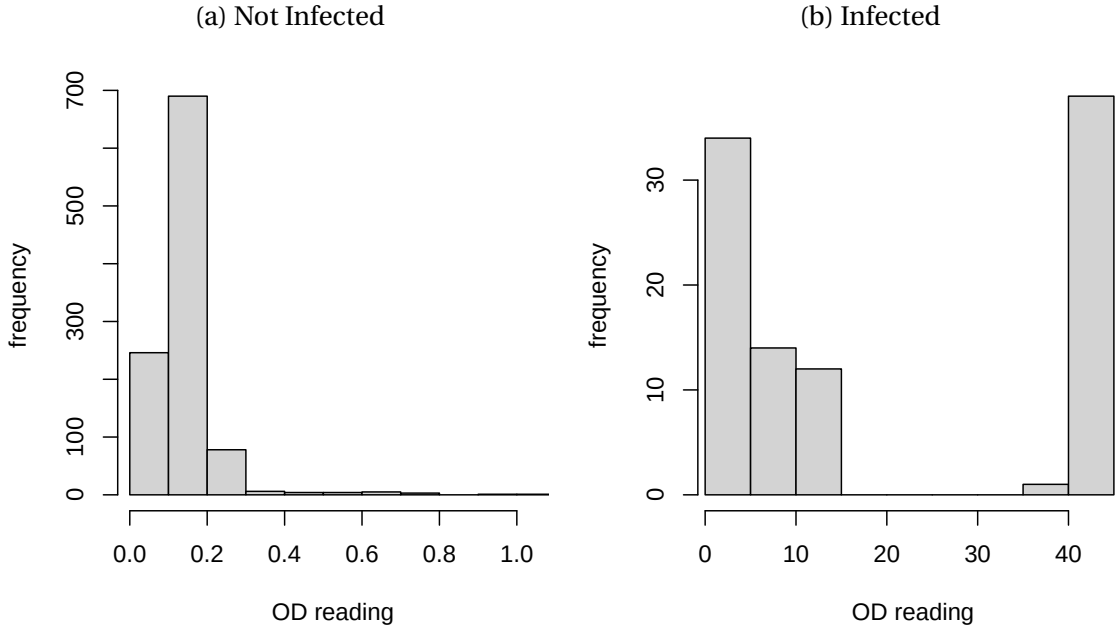


Figure 6 Histogram of OD readings for Hepatitis B of infected and non-infected inmates using the Irish Prisoner dataset collected by Allwright et al. (2000).

Let X_i be the random variable corresponding to the OD reading that subject $i \in S$ would get under individual testing. The distribution of X_i is expected to depend on whether the subject is infected or not. The previous literature has often assumed that the distribution of the OD reading of a group G_g was given by the average of the OD readings from each subject in that group, i.e., by $\sum_{i \in G} X_i / |G_g|$. A problem

with this approach, however, is that it may result in situations in which, even if I/k is much greater than I'/k', we may have $h(I,k) < h(I',k')$. For example, if we assume $X_i \sim N(0.1448, 0.007)$ when the subject is healthy and $X_i \sim N(20, 340)$ when the subject is sick (the moments of these distributions were chosen so as to match their sample counterparts), and set the cutoff values from pooled tests so as to preserve a specificity of 0.9997 for any pool size, we end up with $h(1,1) = 0.854 < 0.930 = h(2,10)$, which is arguably not very realistic.

This issue can be avoided if we adopt the following approach instead. Suppose that the *OD* reading of a healthy and infected subject are governed by the random variables X_- and X_+, respectively. We assume that, if a pool of size k has exactly $I \leq k$ infected subjects, the OD reading of the pool will be given by $X_{I,k} \equiv (I/k)X_+ + (1 - I/k)X_-$. Notice that, in this case, what ultimately affects the distribution of the OD reading is the *proportion* of healthy vs infected subjects in the pool, not the pool size.

Because the distribution of OD readings of infected subjects is bimodal, we use a non-parametric approach to estimate the densities of infected and non infected subjects, and then recover the joint distribution of OD readings using bootstrap.[8]

Because the dataset did not include the cutoff values for OD readings above which a subject is classified as infected, for individual testing we follow Mokalled et al. (2021) and Wang et al. (2018) by eliciting a threshold that maximizes the sum of specificity and sensitivity. This approach leads to an estimated specificity of $S_p = 0.996$ for individual testing. For pooled testing, we select the threshold of the OD reading so as to preserve this specificity, i.e., for each pool size $k \geq 2$, we choose the threshold of detection so that $h(0,k) = S_p = 0.996$.

Figure 7 displays the estimated dilution function using this approach for a pool of size $k = 10$. As it is clear from the figure, the dilution function is concave, which implies that ordered pooling minimizes both the expected number of tests and the expected number of false positive classifications. To assess whether ordered pooling minimizes the expected number of false positive classifications, one only needs to check whether hypothesis 1 holds. In this particular case, however, the hypothesis is not always satisfied. But if we estimate subjects' prior probability of infection, we can still conduct simulations to compare the performance of random vs. ordered pooling on this regard.

As contraction of HBV is highly dependent on age (e.g., Centers for Disease Control and Prevention (2019*b*) and World Health Organization (2017)), we estimate a logit regression of the probability of infection as a function of age, using the following quadratic specification

$$logit(Prob(Y_i = 1|Age_i)) = \beta_0 + \beta_1 \text{Age}_i + \beta_2 \text{Age}_i^{\,2} + \beta_3 \text{Age}_i^{\,3} + \varepsilon_i,$$

[8] We obtain similar results by adopting a parametric approach, and assuming that the OD readings of both infected and non infected subjects is normally distributed, though the assumption of normal distribution for infected subjects is highly inconsistent with the empirical distribution.

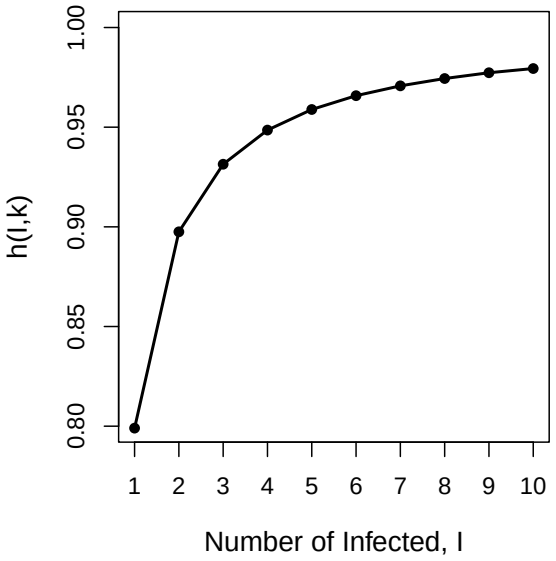


Figure 7 Calibrated dilution function for Hepatitis B using non-parametric density estimates.

where Y_i corresponds to the dummy variable that equals 1 if inmate i is infected, and zero otherwise. Age_i corresponds to the inmate's age, and $(\varepsilon_i)_{i\in S}$ are normally and independently distributed error terms. Figure 8 displays the estimated probability of infection as a function of age.

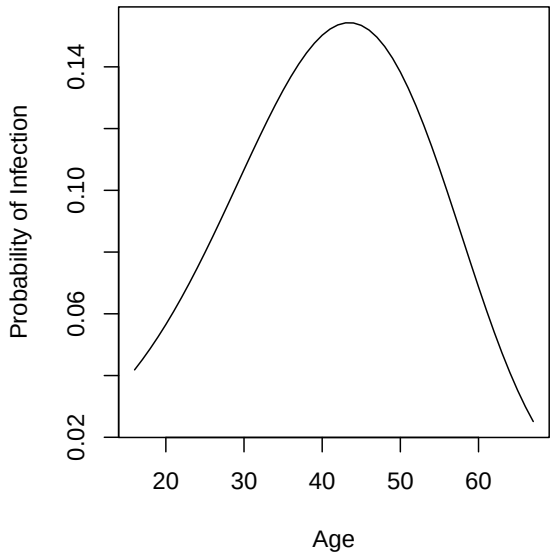
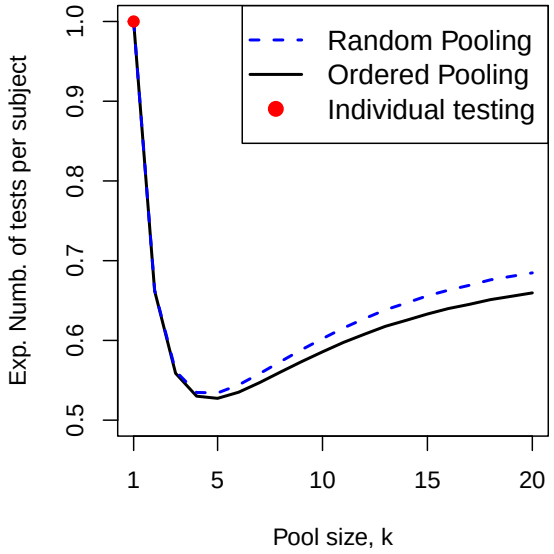


Figure 8 Estimated probability of Hepatitis B infection of inmates as a function of age.

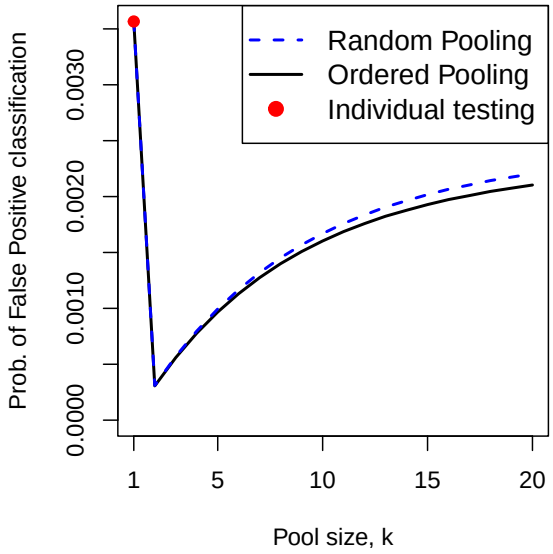
Using these estimated probabilities, we conduct simulations comparing the performance of ordered vs. random pooling in terms of expected number of tests, expected number of false negatives, expected number of false positives, and equity, when all pools are required to have the same size k. The results of these simulations are displayed in figures 9 and 10.

From figures 9a and 9b, we can see that ordered pooling outperforms random pooling both in terms of the expected number of tests and the expected number of false positives, which is consistent with theorem 1 and corollary 1. As to the expected number of false negatives, they are practically identical for both methods, with random pooling generating slightly more expected number of false negatives.

(a) $\mathbb{E}[T(\Omega)]/n$



(b) $\mathbb{E}[FP(\Omega)]/n$



(c) $\mathbb{E}[FN(\Omega)]/n$

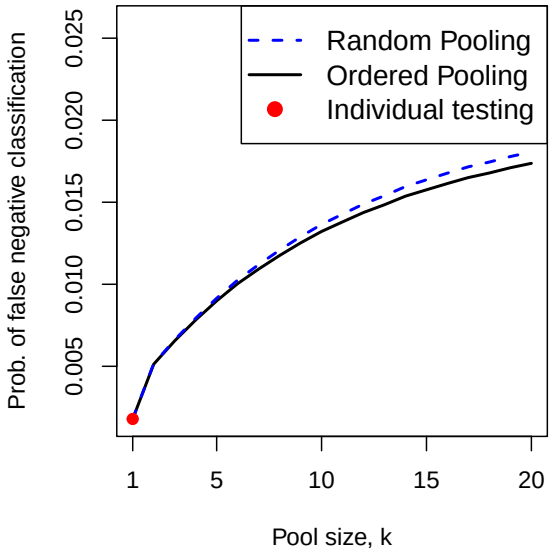


Figure 9 Expected number of tests and classification errors per subject for HBV tests using ordered and random pooling. The red dot corresponds to the case in which subjects are individually tested, i.e. $k = 1$.

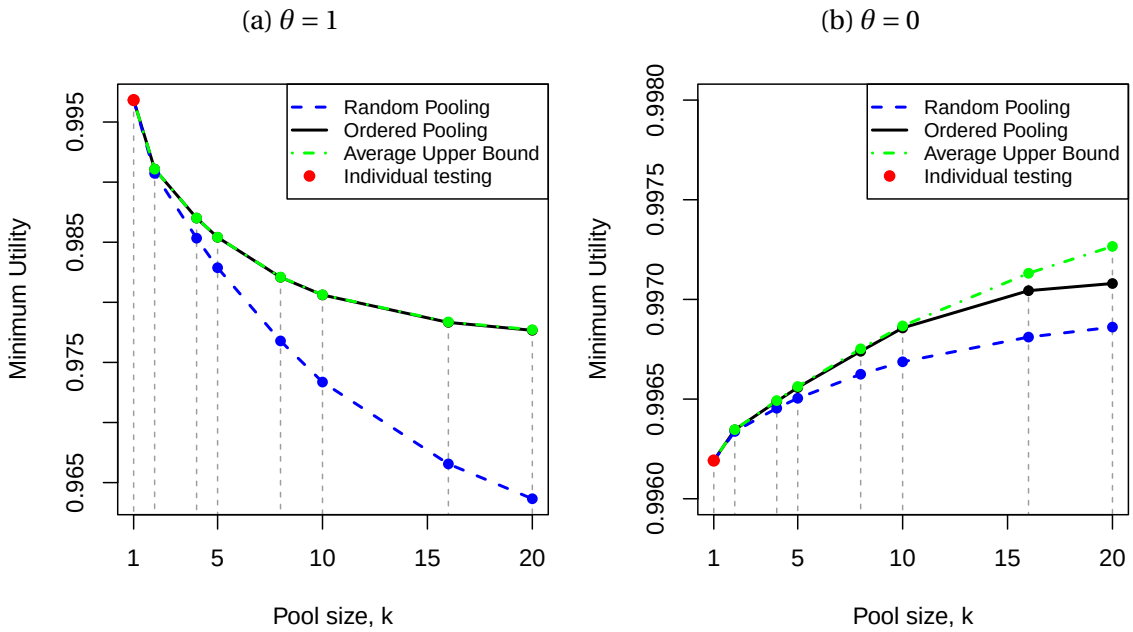


Figure 10 Minimum utility of ordered pooling vs random pooling for HBV tests, for values of $k \leq 20$ that are dividends of $n = 10,000$.

Regarding equity, similar to the previous section, ordered pooling resulted in significantly higher minimum utilities compared to random pooling, for $\theta = 1$ and $\theta = 0$. Moreover, those minimum utilities were very close to the simulated upper bounds from the minimum utility that can be achieved through a partition that preserves those pool sizes, which is indicative that the allocations from ordered pooling are highly equitable.

9. Discussion

We derived sufficient conditions under which ordered pooling minimizes the expected number of tests and both types of classification errors. In order to check whether these conditions are satisfied, one only needs to estimate the dilution function for the pool sizes being considered. Because estimating the dilution function is almost always a requirement before one even considers implementing pool testing schemes in practical applications,[9] information regarding the dilution function is usually readily available from previous studies (e.g., Bateman et al. (2020), Morre et al. (2000) and Kacena et al. (1998*b*)).

With information on subjects' prior probability of infection, one can also conduct simulations to evaluate the performance of ordered pooling in terms of equity. Our simulations suggest that, in general, ordered pooling yields more equitable allocations than random pooling.

[9] If the dilution effect is too strong, pool testing may be deemed too imprecise to be considered as a good alternative to individual testing.

Because ordered pooling is easy to implement, and does not require knowing subjects' exact probability of infection, only how those probabilities are ordered, these results may prove useful in practical applications.

References

Alcalde, Rosana, Fernando Lucas Melo, Anna Nishiya, Suzete Cleusa Ferreira, Mario Dante Langhi Júnior, Simone Sena Fernandes, Luis Augusto Marcondes, Alberto José Silva Duarte, and Jorge Casseb. 2009. "Distribution of hepatitis B virus genotypes and viral load levels in Brazilian chronically infected patients in São Paulo city." *Rev. Inst. Med. trop. S. Paulo*, 51(5): 269–272.

Allwright, Shane, Fiona Bradley, Jean Long, Joseph Barry, Lelia Thornton, and John V Parry. 2000. "Prevalence of antibodies to hepatitis B, hepatitis C, and HIV and risk factors in Irish prisoners: results of a national cross sectional survey." *BMJ*, 321(7253): 78–82.

Aprahamian, Hrayer, Douglas R. Bish, and Erub K. Bish. 2019. "Optimal Risk-Based Group Testing." *Management Science*, 65(9): 4365–4384.

Aprahamian, Hrayer, Ebru K. Bish, and Douglas R. Bish. 2018. "Adaptive risk-based pooling in public health screening." *IISE Transactions*, 50(9): 743–766.

Aprahamian, Hrayer, Ebru K. Bish, and Douglas R. Bish. 2020. "Static Risk-Based Group Testing Schemes Under Imperfectly Observable Risk." *Stochastic Systems*, 10(4): 361–390.

Basso, Leonardo, Vicente Salinas, Denis Sauré, Charles Thraves, and Natalia Yankovic. 2022. "The Effect of Correlation and False Negatives in Pool Testing Strategies for COVID-19." *Healthcare Management Science*, 25: 146–165.

Bateman, Allen C., Shanna Mueller, Kyley Guenther, and Peter Shult. 2020. "Assessing the dilution effect of specimen pooling on the sensitivity of SARS-CoV-2 PCR tests." *Journal of Medical Virology*.

Bertsimas, Dimitris, Vivek F. Farias, and Nikolaos Trichakis. 2012. "On the Efficiency-Fairness Trade-off." *Management Science*, 58(12): 2234–2250.

Bossert, Walter, and Kohei Kamaga. 2020. "An axiomatization of the mixed utilitarian–maximin social welfare orderings." *Economic Theory*, 69(2): 451–473.

Burns, Kevin C., and Carl A. Mauro. 1987. "Group testing with test error as a function of concentration." *Communications in Statistics-theory and Methods*, 16: 2821–2837.

Centers for Disease Control and Prevention. 2000. "Tracking the hidden epidemics, trends in STDs in the United States." *https://www.cdc.gov/std/trends2000/trends2000.pdf*.

Centers for Disease Control and Prevention. 2019*a*. "Sexually Transmitted Infections Treatment Guidelines, 2021." *https://www.cdc.gov/std/treatment-guidelines/chlamydia.htm*.

Centers for Disease Control and Prevention. 2019*b*. "Viral Hepatitis Surveillance Report United States 2019." *https://www.cdc.gov/hepatitis/statistics/2019surveillance/pdfs/2019HepSurveillanceRpt.pdf*.

Chlebus, Bogdan. 2001. "Randomized Communication in Radio Networks." 401–456.

Dorfman, Robert. 1943. "The Detection of Defective Members of Large Populations." *The Annals of Mathematical Statistics*, 14: 436–440.

Downs, Louise O., Sabeehah Vawda, Phillip Armand Bester, Katrina A. Lythgoe, Tingyan Wang, David A. Smith Anna L. McNaughton, Tongai Maponga, Oliver Freeman, Jim Davies Kinga A. Várnai, Kerrie Woods, Christophe Fraser, Eleanor Barnes, Dominique Goedhals, and Philippa C. Matthews. 2020. "Bimodal distribution and set point HBV DNA viral loads in chronic infection: retrospective analysis of cohorts from the UK and South Africa."

El-Amine, Hadi, Ebru K. Bish, and Douglas R. Bish. 2017. "Optimal pooling strategies for nucleic acid testing of donated blood considering viral load growth curves and donor characteristics." *IISE Transactions on Healthcare Systems Engineering*, 7(1): 15–29.

Fan, Wenxin. 2020. "Wuhan Tests Nine Million People for Coronavirus in 10 Days." *The Wall Street Journal*.

Goodrich, Michael, Mikhail Atallah, and Roberto Tamassia. 2005. "Indexing Information for Data Forensics." Vol. 3531, 206–221.

Grobe, Nadja, Alhaji Cherif, Xiaoling Wang, Zijun Dong, and Peter Kotanko. 2020. "Sample pooling: burden or solution?" *Clinical Microbiology and Infection*.

Harsanyi, John C. 1955. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility." *Journal of Political Economy*, 63(4): 309–321.

Hwang, F. K. 1975. "A Generalized Binomial Group Testing Problem." *Journal of the American Statistical Association*, 70: 923–926.

Hwang, F. K. 1976. "Group Testing with a Dilution Effect." *Biometrika*, 63(3): 671–673.

Kacena, Katherine A., Sean B. Quinn, René Howell, Guillermo E. Madico, Thomas C. Quin, and Charlotte A. Gaydos. 1998*a*. "Pooling Urine Samples for Ligase Chain Reaction Screening for Genital Chlamydia trachomatis Infection in Asymptomatic Women." *Journal of Clinical Microbiology*, 36(2): 481–485.

Kacena, Katherine A., Sean B. Quinn, Suzanne C. Hartman, Thomas C. Quinn, and Charlotte A. Gaydos. 1998*b*. "Pooling of Urine Samples for Screening for Neisseria gonorrhoeae by Ligase Chain Reaction: Accuracy and Application." *Journal of Clinical Microbiology*, 36: 3624–3628.

Kim, Hae-Young, Michael G. Hudgens, Jonathan M. Dreyfuss, Daniel J. Westreich, and Christopher D. Pilcher. 2007. "Comparison of Group Testing Algorithms for Case Identification in the Presence of Test Error." *Biometrics*, 63(4): 1152–1163.

Lan, Tian, David Kao, Mung Chiang, and Ashutosh Sabharwal. 2010. "An Axiomatic Theory of Fairness in Network Resource Allocation." 1–9.

McMahan, Christopher S., Joshua M. Tebbs, and Christopher R. Bilder. 2012. "Informative Dorfman Screening." *Biometrics*, 68(1): 287–296.

Mokalled, Stefani C., Christopher S. McMahan, Joshua M. Tebbs, Derek Andrew Brown, and Christopher R. Bilder. 2021. "Incorporating the dilution effect in group testing regression." *Statistics in Medicine*, 40(11): 2540–2555.

Morre, Servaas A., Chris J. L. M. Meijer, Christian Munk, Susanne Kruger-Kjaer, Jeanette F. Winther, Hans O. Jørgensens, and Adriaan J. C. Van Den Brule. 2000. "Pooling of Urine Specimens for Detection of Asymptomatic Chlamydia trachomatis Infections by PCR in a Low-Prevalence Population: Cost-Saving Strategy for Epidemiological Studies and Screening Programs." *Journal of Clinical Microbiology*, 38: 1679–1680.

Nguyen, Ngoc T., Hrayer Aprahamian, Ebru K. Bish, and Douglas R. Bish. 2019. "A Methodology for Deriving the Sensitivity of Pooled Testing, Based on Viral Load Progression and Pooling Dilution." *Journal of Translational Medicine*, 17(1): 1152–1163.

Owusu-Edusei Jr, Kwame, Harrell W. Chesson, Thomas L. Gift, Robert C. Brunham, and Gail Bolan. 2015. "Cost-effectiveness of Chlamydia Vaccination Programs for Young Women." *Emerging infectious diseases*, 21(6): 960–968.

Rawls, John. 1971. *A theory of justice*. The Belknap Press of Harvard University Press.

Schneider, Mark, and Byung-Cheol Kim. 2020. "The utilitarian–maximin social welfare function and anomalies in social choice." *Southern Economic Journal*, 87(2): 629–646.

Sobel, M., and P. A. Groll. 1959. "Group testing to eliminate efficiently all defectives in a binomial sample." *The Bell System Technical Journal*, 38(5): 1179–1252.

Wang, Dewei, Christopher S. McMahan, and Colin M. Gallagher. 2015. "A general regression framework for group testing data, which incorporates pool dilution effects." *Statistics in Medicine*, 34: 3606–3621.

Wang, Dewei, Christopher S. McMahan, Joshua M. Tebbs, and Christopher R. Bilder. 2018. "Group testing case identification with biomarker information." *Computational Statistics & Data Analysis*, 122: 156–166.

Warasi, Md S, Christopher McMahan, J. Tebbs, and Christopher R. Bilder. 2017. "Group Testing Regression Models with Dilution Submodels." *Statistics in Medicine*, 36 30: 4860–4872.

Wein, Lawrence M., and Stefanos A. Zenio. 1996. "Pooled testing for HIV screening: capturing the dilution effect." *Operations Research*, 44(4): 543–569.

World Health Organization. 2017. "Global Hepatitis Report 2017." *https://www.who.int/publications/i/item/global-hepatitis-report-2017*.

Zou, Shimian, Susan L. Stramer, and Roger Y. Dodd. 2012. "Donor Testing and Risk: Current Prevalence, Incidence, and Residual Risk of Transfusion-Transmissible Agents in US Allogeneic Donations." *Transfusion Medicine Reviews*, 26(2): 119–128.