Estimation through array-based group tests

Authors: João Paulo Martins
– ESTG – Polytechnic Institute of Leiria, Portugal,
CEAUL – University of Lisbon, Portugal
jpmartins@ipleiria.pt

Miguel Felgueiras
– ESTG and CIGS – Polytechnic Institute of Leiria, Portugal,
CEAUL – University of Lisbon, Portugal
mfelg@ipleiria.pt

Rui Santos
– ESTG – Polytechnic Institute of Leiria, Portugal,
CEAUL – University of Lisbon, Portugal
rui.santos@ipleiria.pt

Abstract:

- Pooling individual samples for batch testing is a common procedure for reducing costs. The recent use of multidimensional array algorithms, due to the emergence of robotic pooling, is an innovative way of pooling. We show that the two-dimensional array-based group tests can provide accurate estimates for the prevalence rate even for situations in which the traditional estimators, applied to one-dimensional arrays, are not valid. Hence, a computational script was developed to determine which prevalence rate estimate minimizes the sum of the squared deviations between the number of observed and expected rows and columns whose pooled sample had a positive test result.

Key-Words:

- *estimation; prevalence rate; pooled samples; array.*

AMS Subject Classification:

- 62F10, 62P10.

1. INTRODUCTION

Evidences of using pooled samples for batch testing date as far back as 1915 (cf. [8]). However, its use in order to reduce costs started only in 1943 with Dorfman's seminal work [4] on the detection of the syphilis antigen in U.S. soldiers during World War II. Dorfman entailed pooling together biological specimens from different individuals and testing the resulting pools of specimens rather than testing each individual.

When the aim is the detection of some binary characteristic, Dorfman's process comprehends two stages. In its first stage a pool of n individuals is homogeneously mixed and a portion of the mixed sample is analyzed. A negative result in the pooled mixture indicates that none of the n individuals has that characteristic. On the other hand, a positive result implies that at least one of the n individuals possesses the characteristic under investigation. And, in this last case, a second stage takes place in which an individual test is performed to each one of the n suspected individuals. The optimal batch size n^* is the pool size which minimizes the expected number of tests since the cost of mixing samples is, in general, negligible comparing to the cost of the experimental tests, as [13] points out.

Since Dorfman's work, the research on methodologies involving pooled sample tests has been quite active. Thence, some improvements to his work have been proposed, for instance, by [6, 22, 23]. The common idea of all these algorithm improvements was to divide each positive pool into smaller subpools until eventually all specimens are individually tested. These kind of algorithms are called hierarchical algorithms with a number of stages equal to the number of times each individual may be tested. All of them are called one-dimensional since they use non-overlapping pools. More recent works, considering the experimental errors measured by the test sensitivity and test specificity, are available in [9, 11, 19, 24]. Another branch in this area has been the application to quantitative characteristics (only reliable for underlying heavy tailed distributions). More details may be found, for instance, in [5, 15, 20, 21]. Moreover, the use of pooled samples does not refer only to the classification problem (identifying all the individuals which possess some characteristic), since it may also be useful in estimating its prevalence rate p (estimation problem), as [22] stated.

As different procedures may be applied in a problem involving pooled sample tests, the expected number of experimental tests per individual is a good measure of the savings obtained with each procedure. Hence, the *relative cost* of a procedure $\mathcal{M}$ is defined as

$$\mathrm{RC}(\mathcal{M}) = \frac{\mathrm{E}(T_N)}{N}, \tag{1.1}$$

where T_N stands for the number of experimental tests performed to screen a sample with N individuals.

When the purpose is to estimate the prevalence rate, the performing of individual tests is only optional, since the goal is no longer to identify who has the characteristic under investigation. Thus, this may lead to a lower relative cost. Furthermore, the estimators obtained by applying compound tests have, under certain conditions, better performance than the traditional estimators based on individual tests, cf. [7, 14, 22]. Hence, group testing can be more efficient as well as more accurate than individual testing.

The use of more complex schemes of mixing samples, i.e., dividing the amount of sample in two or more parts and using them in different batches had not been a reasonable choice since the complexity of the process could be itself another significant source of error. However, in the beginning of this century the emergence of the robotic and automatic pooling has turned the array-based group testing into a reliable alternative to hierarchical group testing (cf. [12]).

Our purpose is to show that in the context of a prevalence rate estimation the use of two dimensional arrays may be a reasonable alternative to the traditional one dimensional array-based procedure. We also discuss strategies to obtain an estimate from each of the possible ambiguous results of a two dimensional array. A first attempt to solve those ambiguities was performed in [17]. For this goal, an improvement of an algorithm firstly proposed in [16] is provided and a small simulation study is carried out. It is also claimed that for some low sensitivity tests, such as some enzyme immunoassays to screen for *Clostridium difficile* in fecal specimens described in [1], the performance of one dimensional arrays procedures may provide prevalence rate estimates outside the interval $[0, 1]$ whereas the two dimensional arrays procedures always provide valid estimates.

The outline of this article is as follows. Section 2 describes some general assumptions and additional notation usually used in this research field. Subsequently, Section 3 describes the use of one-dimensional and two-dimensional array-based group testing in the context of the estimation problem. The core of this work is Section 4 where it is analyzed the performance of the proposed algorithm for computing an estimate to the prevalence rate based on the results of two dimensional arrays. In particular, some possible drawbacks of the algorithm are discussed. Finally, some conclusions are provided in Section 5.

2. Framework setting

Let us consider a large population of individuals and let p stand for the probability of randomly choosing an individual infected with some disease. The value p is called the *prevalence rate* of the disease.

When dealing with an estimation problem, i.e., the problem of estimating the value of p, the most basic and common pooled sample methodology is to divide the individuals among the sample into groups with size n – one dimensional arrays. For simplicity, admit that n is a divisor of the sample size N (otherwise, one would have $\left[\frac{N}{n}\right]$ groups with n individuals and one group with $N - \left[\frac{N}{n}\right] \times n$ individuals, where $[x]$ stands for the highest integer lower than x). Then, it is required to perform $T_N = \frac{N}{n}$ tests. Let us also assume that the individual status (infected/not infected) within a pooled sample are independent. The probability of having an infected pooled sample is $\pi_n = 1 - (1-p)^n$. Hence, the total number of infected pooled samples is described by a binomial random variable $I \frown \text{Bin}(T_N, \pi_n)$, where T_N is the trials number and π_n the success probability. Thus, the maximum likelihood (ML) estimator of π_n is given by

$$\widehat{\pi_n} = \frac{I}{T_N}. \tag{2.1}$$

As p is given by a simple transformation of π_n, it is straightforward to show, applying the proprieties of the ML estimators, that the ML estimator of p is given by

$$\widehat{p} = 1 - \left(1 - \frac{I}{T_N}\right)^{1/n}. \tag{2.2}$$

For $n = 1$, $\hat{p} = 1 - \left(1 - \frac{I}{T_N}\right) = \frac{I}{T_N}$ is an unbiased estimator of p. For $n > 1$, the estimator is positively biased . Expressions for the expected value and variance of the estimator can be found in [10].

As screening errors may occur, the above binomial model is, in practice, unrealistic. Let $X_{ji} = 1$ denote an infected individual and $X_{ji} = 0$ denote a non-infected individual concerning the i-th individual of the j-th pooled sample where $i = 1, \cdots, n$ and $j = 1, \cdots, T_N$. In addition, X_{ji}^+ denotes a positive test result and X_{ji}^- a negative test result performed with a sample collected only from that individual. The probability $\varphi_s = \text{P}\left(X_{ji}^+ | X_{ji} = 1\right)$ is called the *test sensitivity* and $\varphi_e = P\left(X_{ji}^- | X_{ji} = 0\right)$ is called the *test specificity*. [19] extended the concepts of specificity and sensitivity to a specific procedure $\mathcal{M}$. These measures assess the quality of a result provided by $\mathcal{M}$. The *procedure sensitivity* is the probability of an infected individual being correctly identified by the procedure $\mathcal{M}$, that is, $\varphi_s^{\mathcal{M}} = P_{\mathcal{M}}\left(X_{ji}^+ | X_{ji} = 1\right)$. The *procedure specificity* stands for the probability of a non-infected individual being correctly classified by the procedure $\mathcal{M}$, that is, $\varphi_e^{\mathcal{M}} = P_{\mathcal{M}}\left(X_{ji}^- | X_{ji} = 0\right)$.

As some interaction between the pooled specimens may occur, some general

assumptions underlying our work must be settled (more details may be found in [18]).

- Assumption 1 – Any specimen X_{ji}, where $i = 1, \cdots, T_N$ and $i = 1, \cdots, n$, may be described by a Bernoulli random variable X_{ji} where $P(X_{ji} = 1) = p$ (infected) and $P(X_{ji} = 0) = q = 1 - p$ (non-infected).
- Assumption 2 – The methodology sensitivity equals the test sensitivity, i.e., $\varphi_s^{\mathcal{M}} = \varphi_s$. Note that this may not always be true as individually defective specimens may generate a negative result when tested in a batch. This is called an antagonism effect.
- Assumption 3 – The methodology specificity equals the test specificity, i.e., $\varphi_e^{\mathcal{M}} = \varphi_e$. As in the last assumption, in some situations the interaction between non-infected specimens may produce a positive pooled sample test result. This phenomenon is called synergism.
- Assumption 4 – Given the true status of any pool, its test result is independent of the true status and test result of any other non-overlapping pool.

3. Array-based group testing

We will briefly describe how to deal with the estimation problem concerning two different procedures. In the first one, the individuals are divided into non-overlapping groups (one-dimensional arrays). In the second one, a one-stage two-dimensional array procedure, the individuals will be tested twice.

3.1. One-dimensional arrays

One-dimensional arrays are the most common arrays used for batching individuals into groups in order to perform pooled sample tests. Each individual is allocated to one and only one group and some amount of its sample is mixed with the same amount of sample from other individual(s). This procedure will be represented by $D(n)$.

Given the j-th pooled sample of size n, the probability of it being positively classified is given by

$$\begin{aligned} &P\left(X_j^+ | \sum_{i=1}^n X_{ji} \geq 1\right)\left(1-(1-p)^n\right) + P\left(X_j^+ | \sum_{i=1}^n X_{ji} = 0\right)(1-p)^n \\ &= \varphi_s + (1 - \varphi_s - \varphi_e)(1-p)^n . \end{aligned} \tag{3.1}$$

Therefore, a ML estimator of p is

$$\widehat{p} = 1 - \left(\frac{\varphi_s - \frac{I}{T_N}}{\varphi_s + \varphi_e - 1} \right)^{1/n}. \tag{3.2}$$

The estimator only assumes a value in the interval $[0, 1]$ if

$$1 - \varphi_e \leq \frac{I}{T_N} \leq \varphi_s. \tag{3.3}$$

Whenever this condition is not fulfilled, we are not able to provide a reasonable estimate. For instance, suppose $p = 0.01$ and that a $D(3)$ procedure is performed with 50 pools where the test verifies $\varphi_e = 0.98$ and $\varphi_s = 0.6$. There is a chance of about 15% of condition (3.3) not be fulfilled. The proportion of positive samples is not an option as it may be a very biased estimator of the prevalence rate.

3.2. Two-dimensional arrays

Two-dimensional arrays are an alternative to the one dimensional arrays which uses overlapping pools. This approach is frequently employed in genetics, cf. [11], but it is rarely applied in the infectious disease setting. In its simplest two-stage version (square array), denoted by $A(n)$, a sample of size n^2 is placed in a $n \times n$ matrix in the following way. Each individual is allocated to one and only one matrix position. Then, all the individuals within the same row are gathered for batch testing, and the same procedure is applied to all the individuals within the same column. Thus, the two-stage version involves $T_{n^2} \geq 2n$ tests as subsequent individual tests can be performed to the ones lying in a row and/or column which tested positively. A variant of this methodology (a three-stage procedure) consists in performing *a priori* a pooled sample test on all the n^2 individuals (master pool). If the master pool test result is negative no further testing is needed as the individuals are all classified as negative.

The expected number of tests for all these two-dimensional array group testing procedures is derived in [18] when a classification problem is dealt. In [11] are computed the operating characteristics of these procedures (with or without a master pool). An extension to higher-dimensional arrays assuming no test errors may be found in [3]. More recently, [12] introduced the possibility of misclassification.

The performance of subsequent individual tests is required to avoid ambiguities. For instance, it is possible to have a row tested positive but all columns

tested negative (the number of infected individuals can be any integer from zero to n, since certainly there is at least one misclassification error) or to have two positive rows and columns (the number of infected individuals can be two, three or four even considering that there was no misclassification). However, it won't be required to identify who is infected or either if a row or column has infected elements when the next proposed methodology for an estimation problem is applied. Hence, this one-stage procedure is expected to decrease the relative cost.

4. Prevalence rate estimation

Two-dimensional array-based group testing allows the inclusion of each individual into two different batches. However, as previously discussed, some ambiguities may arise due to the experimental test errors described by the test sensitivity and by the test specificity.

In this case, the use of the proportion of defective individuals (without performing any individual tests) is not advised as it may lead to an underestimation of the prevalence rate or to an increase of the relative cost as [16] points out. [16] also provides some guidelines of how to compute a ML estimate of the prevalence rate using a computational script. Moreover, it computes an exact expression for the ML estimator, assuming possible test errors, for an one stage and two-dimensional array procedure $A\,(2)$.

The inputs of that script are the test sensitivity φ_s, the test specificity φ_e and the number of arrays that have $i-1$ positive rows and $j-1$ positive columns for $i = 1, 2, \cdots, r+1$ and $j = 1, 2, \cdots, c+1$. These values may be recorded in a $(r+1) \times (c+1)$ matrix $\boldsymbol{O}$ which resumes the experimental results required to compute a ML estimate.

To compute the ML function value at p_0 it is also required to compute the probability of observing $i-1$ positive rows and $j-1$ positive columns, where $i = 1, 2, \cdots, r+1$ and $j = 1, 2, \cdots, c+1$, given p_0. Let us denote this matrix of probabilities by $\boldsymbol{P}_{p_0}$. An approximation to these probabilities can be computed by the performance of a simple simulation.

The ML function value at p_0 knowing $\boldsymbol{O}$ is given by

$$\mathrm{ML}(p_0|\boldsymbol{O}) = \prod_{i,j} \boldsymbol{P}_{p_0}(i,j)^{\boldsymbol{O}(i,j)}. \tag{4.1}$$

However, it is necessary to account some special cases which can lead to inaccurate estimates. For instance, if for some value p_0 and some values i and j, $\boldsymbol{P}_{p_0}(i,j)$ is high and $\boldsymbol{O}(i,j)$ is zero then the process could converge to a value near p_0 whereas its "likelihood" is low.

To avoid having to account for this problem, we propose using the sum of the square of the differences between the values of the matrix $\boldsymbol{O}$ and the expected values for $\boldsymbol{O}$, computed using the probabilities of the matrix $\boldsymbol{P}$, as a quality measure for comparing different estimates, i.e.,

$$Dif(p_0|\boldsymbol{O}) = \sum_{i,j} \left(\boldsymbol{O}(i,j) - s \times \boldsymbol{P}_{p_0}(i,j)\right)^2, \tag{4.2}$$

where s is the total number of two-dimensional arrays (i.e., $s = \sum_{i,j} \boldsymbol{O}(i,j)$).

In the next subsections, we provide some guidelines of an algorithm to find the minimum of the function defined in (4.2) as well as a small simulation study.

4.1. A computational script

As only for a small number of rows and columns it is possible to compute an exact expression for the function Dif defined in (4.2), in general, it is just possible to find the minimum of that function using some computational method.

Next, we describe the script which was implemented and highlight the possible drawbacks due to working with very low values. Some of the issues are shared by the implementation of the well-known chi-square test of independence of two random variables.

Our script comprehends the following steps.

- Step 1 – Consider an increasing sequence of possible values for p, say $p_1 < p_2 < \cdots < p_m$.

  We used the golden section search optimization method, cf. [2], as it presents optimal properties in the numerical search of a maximum when no expression for the function of interest is available. This method uses $m = 4$ and two inner points given at each step by $p_2 = p_4 - \varphi * (p_4 - p_1)$ and $p_3 = p_1 + \varphi * (p_4 - p_1)$ where $\varphi \approx 0.61803$ is known as the golden ratio.

- Step 2 – Simulate a large number of individuals (say, equal to $r \times c \times rep$ with rep large) extracted from a population with a prevalence rate p_i for $i = 1, \cdots, m$ where r and c stand for the number of rows and columns of the array and rep is the number of arrays. These arrays will be used to obtain an estimate for the values of the matrix $\boldsymbol{P}$. We used $rep = 100$ and our simulations have shown that higher values for rep do not change significantly the final outcome of the simulation.

- Step 3 – For each p_i and for each replicate compute the (estimated) probability of observing an array corresponding to each position of the matrix

$\boldsymbol{O}$ (the matrix of the experimental results). Add that value to the position (i, j) of the matrix $\boldsymbol{P}$.

Note that regardless of the number of infected individuals in the array, it is always possible to span all possible number of positive rows and columns due to the presence of the test errors.

- Step 4 – Compute the Dif function given the matrix $\boldsymbol{O}$.
- Step 5 – Compare the Dif function at $p_1, \cdots, p_m$ and choose the two estimates which minimize the function, say p_{min} and p_{max} where $p_{min} < p_{max}$.
- Step 6 – Consider a new increasing sequence of possible values for p starting on p_{min} and finishing on p_{max}.

  In our case, $p_2 = p_4 - \varphi * (p_{max} - p_{min})$ and $p_3 = p_1 + \varphi * (p_{max} - p_{min})$.
- Step 7 – Repeat the procedure from step two up to step six until the distance between the two estimates is lower than some tolerance tol.

4.2. Possible drawbacks

Expression (4.2) for the Dif function involves very small quantities which are not a problem for the most recent software. However, for avoiding the null estimate we advise the initial choice of p_1 to be equal or higher than tol.

One possible problematic situation that must be taken into account in order to avoid underestimation occurs when the expected number of arrays with i positive rows and j positive columns, $s \times \boldsymbol{P}(i, j)$, is higher than zero and less than one. Note that, theoretically, all values of the matrix $\boldsymbol{P}$ are higher than zero but as we are not computing the exact matrix $\boldsymbol{P}$, as there are some computational restraints, it is possible to have zero in some entry of $\boldsymbol{P}$.

In that case, we suggest adding some of those low probabilities till eventually the sum of the expected number of arrays be at least one.

Hence, suppose you inspect all the values of $\boldsymbol{P}(i, j)$ in some logic sequence and you find h values of $\boldsymbol{P}$, say $\boldsymbol{P}(r_1, c_1), \cdots, \boldsymbol{P}(r_h, c_h)$, for which the sum of the expected values verifies $s \sum_{i=1}^{h-1} \boldsymbol{P}(r_i, c_i) < 1$ and $s \sum_{i=1}^{h} \boldsymbol{P}(r_i, c_i) \geq 1$. Then, add all those h probability estimates $P^* = \sum_{i=1}^{h} \boldsymbol{P}(r_i, c_i)$ and do the same for the matrix $\boldsymbol{O}$, i.e., $O^* = \sum_{i=1}^{h} \boldsymbol{O}(r_i, c_i)$. Replace the position (r_h, c_h) in the matrix $\boldsymbol{P}$ and in the matrix $\boldsymbol{O}$ by P^* and O^*, respectively. All the positions $(r_1, c_1), \cdots, (r_{h-1}, c_{h-1})$ for both matrices should be replaced by zero. This is a process with some resemblances to the one applied to contingency tables in order to improve the chi-square test performance.

In that logic sequence of inspection of all positions it is possible to get a sequence of values of $\boldsymbol{P}$ for which the sum of the expected values does not achieve

1 due to end of the inspection process. In this case, we suggest a similar process, however, the sum P^* and O^* should be added to some value of $\boldsymbol{P}$ whose expected value is at least one. Our simulations showed that the choice of this value is not relevant.

4.3. A simulation study

A small simulation was performed to assess the algorithm performance. The chosen measure to assess the accuracy of the estimates was the absolute value of the bias.

The one-stage square array-based group procedure, $A\,(2n)$, was compared to the one-dimensional alternative $D\,(n)$. Note that both procedures present the same relative cost.

We considered four different experimental tests with the sensitivity and specificity described in Table 1.

Table 1: Test sensitivity and test specificity

Test	φ_s	φ_e
A	0.99	0.99
B	0.80	0.98
C	0.60	0.98
D	0.99	0.80

It was considered four different prevalence rates: 0.01, 0.05, 0.10 and 0.25. For each one of them, 20 square arrays were simulated ($s = 20$) for applying $A(4)$ and $A(6)$ procedures. The one-dimensional alternative for each of these two procedures was $D\,(2)$, with 160 pools, for the first case and $D\,(3)$, with 240 pools, for the last one. In each pair of procedures (one and two-dimensional) the number of tests performed is the same. For all cases, the number of replicates was 100. The results are summarized in the Table 2.

Table 2: Mean of the absolute bias (multiplied by 10^2) of 100 estimates obtained by $A\,(4)\,(D\,(2))|A\,(6)\,(D\,(3))$ procedures

	p=0.01	p=0.05	p=0.10	p=0.25
A	0.48(0.34)|0.25(0.16)	0.81(0.33)|0.60(0.28)	0.96(0.49)|0.83(0.40)	1.82(0.94)|1.38(0.86)
B	0.51(—)|0.34(0.35)	0.79(0.68)|0.71(0.59)	1.12(1.03)|1.06(0.72)	2.17(1.82)|2.34(1.53)
C	0.48(—)|0.37(0.45)	0.95(1.18)|0.78(0.90)	1.24(1.41)|1.03(1.15)	2.25(2.45)|2.92(2.07)
D	0.31(0.38)|0.61(—)	1.15(1.55)|0.86(1.01)	1.46(1.61)|1.06(1.04)	1.47(1.54)|1.45(1.36)

In some cases, condition (3.3) was not fullfield leading to negative estimates.

Thus, some values are not displayed (the symbol "—" is displayed instead of a numerical value) since it was observed more than 20 (in 100) negative estimates. All estimates displayed for a prevalence rate $p = 0.01$ do not use all 100 estimates since 1 to 6 of them were negative and excluded from the calculus of the mean of the absolute bias.

Note that the mean of the absolute bias increases with the prevalence rate since for low prevalence rate values the value zero is a natural limit to the estimates. It is not surprising to observe $A(6)$ procedure performing better than $A(4)$ since it uses more individuals. However, when the prevalence rate increases the chance of having a very high number of positive tests greatly increases leading to a worse performance. Regardless the square array procedure, the one-dimensional pools generally outperform the square array procedures for moderate and high prevalence rates (note that the interval of possible prevalence rates is $]0, 0.50]$ since for values higher than 0.50 we can study $q = 1 - p$). For the most inaccurate test considered, test C, the behavior is similar.

5. Conclusion

The spreading of the possibility of robotic pooling will certainly highlight the use of arrays with dimensions higher than one as a practical alternative to the traditional one dimensional arrays for both estimation and classification problems, cf. [12].

In this work we address the problem of estimation and show that for very inaccurate tests (cheaper tests) the use of square arrays assures the experimenter a reasonable estimate (at least for low prevalence rates). However, whenever the sample size is low and φ_s and φ_e are high we have just a few square arrays and the results can be quite inaccurate. In this scenario the $D(n)$ methodology remains a more reasonable option as it could (almost certainly) provide a good estimate. Thence, $D(n)$ methodology can, in some situations, outperforms $A(2n)$ methodology whereas $A(2n)$ works well in a wider parameter support.

REFERENCES

[1] Alcalá, L.; Sánchez-Cambronero, L.; Catalán M.P.; Sánchez-Somolinos, M.; Peláez, M.T.; Marín, M. and Bouza, E. (2008). Comparison of three commercial methods for rapid detection of *Clostridium difficile* toxins A and B from fecal specimens, *J. Clin. Microbiol.*, **46**, 3833–3835.

[2] Antia, H.M. (2002). *Numerical Methods for Scientists and Engineers*, Vol. 1, 2nd ed., Springer, New York, 356–362.

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

[4] Dorfman, R. (1943). The detection of defective members in large populations, *Ann. Math. Stat.*, **14**, 436–440.

[5] Felgueiras, M.; Martins, J.P. and Santos, R. (2014). Distributions families in counting bacteria for compound sampling. *Lecture Notes in Computer Science*, **8581**, 539–551.

[6] Finucan, H.M. (1964) The blood testing problem, *App. Stat.*, **13**, 43–50.

[7] Garner, F.C.; Stapanian, M.A.; Yfantis, E.A. and Williams, L.R. (1989). Probability estimation with sample compositing techniques, *J. Off. Stat.*, **5**, 365–374.

[8] Hughes-Oliver, J.M. (2006). *Pooling Experiments for Blood Screening and Drug Discovery.* In "Screening — Methods for Experimentation in Industry, Drug Discovery, and Genetics" (A. Dean and S. Lewis, Eds.), Springer, 46–68.

[9] Hwang, F.K. (1976). Group testing with a dilution effect, *Biometrika*, **63**, 671-673.

[10] Hung, M. and Swallow, W.H. (1999). Robustness of group testing in the estimation of proportions, *Biometrics*, **55**, 231–237.

[11] Kim, H.; Hudgens, M.; Dreyfuss, J.; Westreich, D. and Pilcher, D. (2007). Comparison of group testing algorithms for case identification in the presence of testing errors, *Biometrics* 2007, **63**, 1152–1163.

[12] Kim, H. and Hudgens, M. (2009). Three-Dimensional Array-Based Group Testing Algorithms, *Biometrics*, **65**, 903–910.

[13] Liu, S.C.; Chiang, K.S.; Lin, C.H.; Chung, W.C.; Lin, S.H. and Yang, L.C. (2011). Cost analysis in choosing group size when group testing for Potato virus Y in the presence of classification errors, *Ann. Appl. Biol.*, **159**, 491–502.

[14] Loyer, M.W. (1983). Bad probability, good statistics, and group testing for binomial estimation, *Am. Stat.*, **37**, 57–59.

[15] Martins, J.P., Santos, R. and Sousa, R. (2014). *Testing the maximum by the mean in quantitative group tests.* In "New Advances in Statistical Modeling and Applications, Studies in Theoretical and Applied Statistics" (A. Pacheco, R. Santos, M.R. Oliveira and C.D. Paulino, Eds.), Springer, 55–63.

[16] Martins, J.P., Felgueiras, M. and Santos, R. (2014). Maximum Likelihood Estimation in Pooled Sample Tests. *AIP Conf. Proc.*, **1618**, 543 (online).

[17] Martins, J.P., Santos, R. and Felgueiras, M. (2015). A Maximum Likelihood Estimator for the Prevalence Rate Using Pooled Sample Tests. In " Theory and Practice of Risk Assessment" (C.P. Kitsos, T.A. Oliveira, A. Rigas and S. Gulati, Eds.), Springer, 99-110.

[18] Phatarfod, C.D. and Sudbury, A. (1994). The use of a square array scheme in blood testing, *Stat. Med.*, **13**, 2337–2343.

[19] Santos, R.; Pestana, D. and Martins, J.P. (2013). *Extensions of the Dorfman's theory.* In "Studies in Theoretical and Applied Statistics, Recent Developments in Modeling and Applications in Statistics" (P.E. Oliveira, M. Graça, C. Henriques and M. Vichi, Eds.), Springer, 179–189.

[20] Santos, R., Martins, J.P., and Felgueiras, M. (2015). Known mean, unknown maxima? Testing the maximum knowing only the mean, *Commun. Stat.-Simul. C.* , **44**, 2479–2491.

[21] Santos, R.; Martins, J.P. and Felgueiras, M. (2015). *An overview of quantitative continuous compound tests*. In "CIM Series in Mathematical Sciences, Dynamics, Games and Science III" (J.P. Bourguignon, R. Jeltsch, A. Pinto and M. Viana, Eds.), Springer, 595–607.

[22] Sobel, M. and Elashoff, R.M. (1975). Group testing with a new goal, estimation, *Biometrika*, **62**, 181–193.

[23] Sterret, A. (1957). On the detection of defective members of large populations, *Ann. Math. Stat.*, **28**, 1033–1036.

[24] Wein L.M. and Zenios S.A. (1996). Pooled testing for HIV screening: capturing the dilution effect, *Oper. Res.*, **44**, 543–569.