Computational Statistics and Data Analysis 122 (2018) 156–166

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Group testing case identification with biomarker information

Dewei Wang [a], Christopher S. McMahan [b], Joshua M. Tebbs [a,*], Christopher R. Bilder [c]

[a] Department of Statistics, University of South Carolina, Columbia, SC 29208, USA
[b] Department of Mathematical Sciences, Clemson University, Clemson, SC 29634, USA
[c] Department of Statistics, University of Nebraska-Lincoln, Lincoln, NE 68583, USA

ARTICLE INFO

Article history:
Received 9 August 2017
Received in revised form 11 January 2018
Accepted 12 January 2018
Available online 1 February 2018

Keywords:
Classification
Measurement error
Pooled testing
Screening
Sensitivity
Specificity

ABSTRACT

Screening procedures for infectious diseases, such as HIV, often involve pooling individual specimens together and testing the pools. For diseases with low prevalence, group testing (or pooled testing) can be used to classify individuals as diseased or not while providing considerable cost savings when compared to testing specimens individually. The pooling literature is replete with group testing case identification algorithms including Dorfman testing, higher-stage hierarchical procedures, and array testing. Although these algorithms are usually evaluated on the basis of the expected number of tests and classification accuracy, most evaluations in the literature do not account for the continuous nature of the testing responses and thus invoke potentially restrictive assumptions to characterize an algorithm's performance. Commonly used case identification algorithms in group testing are considered and are evaluated by taking a different approach. Instead of treating testing responses as binary random variables (i.e., diseased/not), evaluations are made by exploiting an assay's underlying continuous biomarker distributions for positive and negative individuals. In doing so, a general framework to describe the operating characteristics of group testing case identification algorithms is provided when these distributions are known. The methodology is illustrated using two HIV testing examples taken from the pooling literature.

1. Introduction

Testing individual specimens in pools, which is known as group testing (or pooled testing), is widespread in disease screening applications. Individuals in pools that test negatively are declared to be negative, and positive pools are resolved (or "decoded") to determine which individuals are positive. The origins of group testing are usually traced back to Dorfman (1943), who proposed that it be used to screen World War II soldiers for syphilis. Since this seminal work, group testing has been applied to numerous infectious disease applications. A literature review reveals recent public health and surveillance applications for HIV (Krajden et al., 2014), HBV and HCV (Page-Shafer et al., 2008; Candotti and Allain, 2009), chlamydia and gonorrhea (Lewis et al., 2012), West Nile virus (Busch et al., 2005), and influenza (Edouard et al., 2015). Group testing is also routinely used by national organizations around the world to screen blood and plasma donations for HIV/HBV/HCV and other diseases (see, e.g., Schmidt et al., 2010; O'Brien et al., 2012; Stramer et al., 2013).

The original procedure proposed by Dorfman (1943) is a two-stage hierarchical algorithm; i.e., non-overlapping pools are tested in the first stage and individuals from positive pools are tested in the second. Hierarchical algorithms using a

* Corresponding author.
E-mail address: tebbs@stat.sc.edu (J.M. Tebbs).

https://doi.org/10.1016/j.csda.2018.01.005
0167-9473/

larger number of stages can reduce the number of tests needed when the disease prevalence is small. For example, Mehta et al. (2011) describe a three-stage algorithm for HIV testing in San Diego that uses master pools of size 10 in the first stage, subpools of size 5 in the second stage, and individual testing in the third. The most common non-hierarchical algorithm is two-dimensional array testing (Phatarfod and Sudbury, 1994; Hudgens and Kim, 2011; McMahan et al., 2012b), where individuals are tested in the rows and columns of an array. A recent HIV application in New Jersey (Martin et al., 2013) illustrates how array testing can even be used in higher dimensions (Kim and Hudgens, 2009). Comprehensive summaries of group testing algorithms and their operating characteristics are found in Kim et al. (2007) and Westreich et al. (2008).

When faced with the task of choosing an appropriate case identification algorithm for screening purposes, public health officials and lab technicians are interested in cost and accuracy. Laboratories with large budgets may opt to test specimens individually as pooling can reduce an assay's sensitivity. In the group testing literature, this reduction is known as "the dilution effect" and can result in an increased number of false negative diagnoses. Group testing algorithms can be selected on the basis of minimizing the expected number of tests per individual to minimize costs (Kim et al., 2007; Westreich et al., 2008) or perhaps in a way that incorporates both the expected number of tests and classification accuracy (see, e.g., Malinovsky et al., 2016). Of course, additional practical considerations such as testing platform constraints, the time needed for testing, and the availability of individuals to pool should also be carefully considered.

When an individual or pooled specimen is tested, an assay typically elicits a binary diagnosis (positive/negative) that is derived from measuring a continuous biomarker; large values of this continuous measurement are usually evidence that the disease is present. Although it is widely known that dichotomizing a continuous outcome can lead to a loss in information, previous evaluations in group testing have largely ignored this underlying aspect and instead have relied explicitly on binary results. Doing so helps to facilitate the derivation of closed-form expressions for the expected number of tests and classification accuracy probabilities; however, this also usually requires one to make assumptions such as (a) the sensitivity and specificity are unaffected by pool size; i.e., there is no dilution effect, and (b) testing outcomes on pools containing common individuals are independent conditional on the true pool statuses. An important contribution of this article is to provide a general framework for case identification evaluation where these assumptions are not needed.

In offering this framework, our approach exploits the underlying continuous biomarker distributions associated with positive and negative individuals. In other words, we do not dichotomize testing outcomes into "positive" or "negative" categories, but instead we make our evaluations in terms of the biomarker distributions themselves. Our work is related to the methodology in Wein and Zenios (1996), who proposed using biomarker concentrations to determine an optimized Dorfman algorithm for HIV testing. However, our article takes a somewhat different perspective. We are not focused on determining optimal designs for specific group testing procedures per se; instead, our goal is to enhance previous case identification algorithm evaluations, such as those in Kim et al. (2007) and Westreich et al. (2008), in group testing applications where biomarker distributions are known. Our evaluations can be performed for any group testing procedure, including Dorfman testing, higher-stage hierarchical algorithms, and array testing. We obtain closed-form expressions for operating characteristics for normally distributed biomarkers in specific algorithms; however, even these expressions may be of limited utility for practitioners. We therefore use simulation to overcome the computational challenges when incorporating biomarker information.

2. Notation and preliminaries

We modify the notation from Wang et al. (2015), who used biomarker distributions to acknowledge the dilution effect in group testing regression. Let $T_i = 1$ if the ith individual is truly positive; $T_i = 0$ otherwise. We assume the T_i's are independent and identically distributed statuses with $\text{pr}(T_i = 1) = p$, the prevalence of the population. Generalizing our evaluation framework to allow for unequal individual disease probabilities (McMahan et al., 2012a, b) or correlated individuals (Lendle et al., 2012) is straightforward; see Section 6. Let $\widetilde{\mathcal{C}}_i$ denote the true biomarker level of the ith individual (e.g., viral load, optical density reading, antibody concentration, etc.). We assume the $\widetilde{\mathcal{C}}_i$'s are mutually independent random variables and that the conditional probability density function of $\widetilde{\mathcal{C}}_i$ given the true status $T_i = t$ is

$$f_{\widetilde{\mathcal{C}}_i|T_i=t}(u) = tf_{\widetilde{\mathcal{C}}^+}(u) + (1-t)f_{\widetilde{\mathcal{C}}^-}(u),$$

where $f_{\widetilde{\mathcal{C}}^+}$ and $f_{\widetilde{\mathcal{C}}^-}$ denote the true biomarker density functions for positive and negative individuals, respectively. In other words, positive individuals in the population have true biomarker levels described by the common density $f_{\widetilde{\mathcal{C}}^+}$; similarly, negative individuals' true biomarker levels are described by $f_{\widetilde{\mathcal{C}}^-}$.

We are interested in calculating quantities like the expected number of tests per individual and classification accuracy probabilities commonly seen in the group testing case identification literature (i.e., pooling sensitivity, pooling specificity, predictive values). To set our ideas, we assume a hierarchical group testing algorithm is used in $S \geq 2$ stages, although we later modify our notation to account for array testing in two dimensions (Phatarfod and Sudbury, 1994; Hudgens and Kim, 2011; McMahan et al., 2012b); see Section 3.3. An S-stage hierarchical algorithm begins by testing a master pool of individual specimens. If the master pool tests negatively, all individuals are declared to be disease-free and no further testing is performed. Otherwise, non-overlapping subpools are formed and are tested in the second stage. Any second-stage subpool that tests positively is split again while subpools that test negatively in the second stage are declared to be disease-free. This process continues until all subpools in a particular stage test negatively or until individual testing (in stage S) is performed.