Supplementary Material for "The Optimal Group Size Controversy for Infectious Disease Testing: Much Ado About Nothing?!"

Brianna D. Hitt
Department of Statistics, University of Nebraska-Lincoln
Lincoln, NE 68583
brianna.hitt@huskers.unl.edu

Christopher R. Bilder
Department of Statistics, University of Nebraska-Lincoln
Lincoln, NE 68583

Joshua M. Tebbs
Department of Statistics, University of South Carolina
Columbia, SC 29208

Christopher S. McMahan
Department of Mathematical Sciences, Clemson University
Clemson, SC 29634

August 24, 2018

Contents

List of Tables

S1. Derivation of $E(C)$

Because infectious disease testing is imperfect, define Y_i as the final diagnosis based on the group testing algorithm and $\tilde{Y}_i$ as the true status of individual i, for $i = 1, \ldots, I$. The probability that individual i is truly positive is denoted as $P(\tilde{Y}_i = 1) = p_i$. Also, define C as the number of correct classifications for a group of size I. The expected number of correct classifications for the group is

$$\begin{aligned} E(C) &= \sum_{i=1}^{I} P(Y_i = 0, \tilde{Y}_i = 0) + P(Y_i = 1, \tilde{Y}_i = 1) \\ &= \sum_{i=1}^{I} P(Y_i = 0|\tilde{Y}_i = 0)P(\tilde{Y}_i = 0) + P(Y_i = 1|\tilde{Y}_i = 1)P(\tilde{Y}_i = 1) \\ &= \sum_{i=1}^{I} PS_{p,i}(1 - p_i) + PS_{e,i}p_i \end{aligned}$$

where $PS_{p,i}$ and $PS_{e,i}$ are the pooling specificity and pooling sensitivity, respectively, for individual i.

S2. Additional results for Section 3.1

We provide additional results to coincide with our investigations in Section 3.1. Overall, these additional results continue to show that the OTCs have the same or very similar operating characteristics when using either objective function. We also include in our summaries the pooling positive predictive value, $PPPV$, and the pooling negative predictive value, $PNPV$, as additional accuracy measures. The pooling positive (negative) predictive value is the probability that an individual who is determined to be positive (negative) by the testing algorithm is truly positive (negative). Predictive values simply provide an alternative way of looking at accuracy in comparison to the pooling sensitivity and pooling specificity. Expressions for all accuracy measures are available in Kim et al. (2007), McMahan et al. (2012a), McMahan et al. (2012b), and Black et al. (2015).

Table 1 in the paper provides a summary of the optimal testing configurations (OTCs) and their operating characteristics when $p = 0.01$. This table is reproduced here as Table S1 with the addition of the predictive values. Similar tables for $p = 0.05$ and $p = 0.10$ are shown in Tables S2 and S3, respectively. The largest differences between operating characteristics for OTCs are shown in Table 2 of the main paper. Table S4 displays the same results with the addition of the predictive values.

S3. Additional results for Section 3.2

S3.1. Tables

We provide additional results to coincide with our investigations in Section 3.2. Once again, these additional results show that the same or very similar operating characteristics are obtained regardless of which objective function is used. Table S5 displays the same results as Table 3 in the paper but with the addition of the predictive values. Similar tables for $E(P_i) = 0.05$ and $E(P_i) = 0.10$ are provided in Tables S6 and S7, respectively. Table S8 displays the same findings as Table 4 in the paper but with the addition of the predictive values.

Because informative group testing results in potentially different accuracy measures for each individual tested,

Table S1: OTC summary for $p = 0.01$ under non-informative group testing.

Algorithm	S_e	S_p	Objective function	OTC	$E(T)/I$	PS_e	PS_p	$PPPV$	$PNPV$
2-stage hierarchical	0.99	0.99	O_{ET}	11-1	0.2035	0.9801	0.9990	0.9052	0.9998
			O_{MAR}	11-1	0.2035	0.9801	0.9990	0.9052	0.9998
	0.95	0.95	O_{ET}	11-1	0.2351	0.9025	0.9932	0.5727	0.9990
			O_{MAR}	11-1	0.2351	0.9025	0.9932	0.5727	0.9990
	0.90	0.90	O_{ET}	12-1	0.2742	0.8100	0.9816	0.3081	0.9980
			O_{MAR}	12-1	0.2742	0.8100	0.9816	0.3081	0.9980
3-stage hierarchical	0.99	0.99	O_{ET}	25-5-1	0.1354	0.9703	0.9996	0.9604	0.9997
			O_{MAR}	25-5-1	0.1354	0.9703	0.9996	0.9604	0.9997
	0.95	0.95	O_{ET}	24-6-1	0.1443	0.8574	0.9973	0.7634	0.9986
			O_{MAR}	24-6-1	0.1443	0.8574	0.9973	0.7634	0.9986
	0.90	0.90	O_{ET}	24-6-1	0.1562	0.7290	0.9938	0.5437	0.9973
			O_{MAR}	24-6-1	0.1562	0.7290	0.9938	0.5437	0.9973
Array w/o master pooling	0.99	0.99	O_{ET}	25-1	0.1378	0.9703	0.9995	0.9529	0.9997
			O_{MAR}	25-1	0.1378	0.9703	0.9995	0.9529	0.9997
	0.95	0.95	O_{ET}	25-1	0.1475	0.8575	0.9970	0.7456	0.9986
			O_{MAR}	24-1	0.1475	0.8575	0.9972	0.7566	0.9986
	0.90	0.90	O_{ET}	25-1	0.1611	0.7291	0.9926	0.4996	0.9973
			O_{MAR}	24-1	0.1611	0.7291	0.9930	0.5112	0.9973
Array w/ master pooling	0.99	0.99	O_{ET}	625-25-1	0.1364	0.9606	0.9995	0.9529	0.9996
			O_{MAR}	625-25-1	0.1364	0.9606	0.9995	0.9529	0.9996
	0.95	0.95	O_{ET}	625-25-1	0.1402	0.8146	0.9972	0.7458	0.9981
			O_{MAR}	576-24-1	0.1402	0.8146	0.9974	0.7569	0.9981
	0.90	0.90	O_{ET}	625-25-1	0.1450	0.6562	0.9934	0.4997	0.9965
			O_{MAR}	576-24-1	0.1450	0.6562	0.9937	0.5115	0.9965

NOTE: Equally sized groups are optimal at each stage; thus, an OTC of "24-6-1" means that stage 1 has a group of size 24, stage 2 has four groups of size 6, and stage 3 has twenty-four groups of size 1. Differences between O_{ET} and O_{MAR} are highlighted.

we formed weighted averages across all individuals tested to present one overall value for each accuracy measure. These weighted averages are developed from accuracy definitions given by Altman and Bland (1994a) and Altman and Bland (1994b) and were used by Black et al. (2015). The pooling sensitivity is defined as

$$PS_e^W = \frac{\sum_i p_i PS_{e,i}}{\sum_i p_i}, \tag{S3.1}$$

and the pooling specificity is defined as

$$PS_p^W = \frac{\sum_i (1-p_i) PS_{p,i}}{\sum_i (1-p_i)}. \tag{S3.2}$$

Similarly, the pooling positive predictive value is defined as

$$PPPV^W = \frac{\sum_i p_i PS_{e,i}}{\sum_i p_i PS_{e,i} + (1-p_i)(1-PS_{p,i})}, \tag{S3.3}$$

and the pooling negative predictive value is defined as

$$PNPV^W = \frac{\sum_i (1-p_i) PS_{p,i}}{\sum_i (1-p_i) PS_{p,i} + p_i(1-PS_{e,i})}. \tag{S3.4}$$

Expressions (S3.1) through (S3.4) represent weighted averages over all I individuals within the initial group for a hierarchical algorithm or all I^2 individuals within the array for an array-based algorithm.

Table S2: OTC summary for $p = 0.05$ under non-informative group testing.

Algorithm	S_e	S_p	Objective function	OTC	$E(T)/I$	PS_e	PS_p	$PPPV$	$PNPV$
2-stage hierarchical	0.99	0.99	O_{ET}	5-1	0.4317	0.9801	0.9981	0.9642	0.9990
			O_{MAR}	5-1	0.4317	0.9801	0.9981	0.9642	0.9990
	0.95	0.95	O_{ET}	5-1	0.4536	0.9025	0.9892	0.8141	0.9948
			O_{MAR}	5-1	0.4536	0.9025	0.9892	0.8141	0.9948
	0.90	0.90	O_{ET}	6-1	0.4786	0.8100	0.9719	0.6027	0.9898
			O_{MAR}	6-1	0.4786	0.8100	0.9719	0.6027	0.9898
3-stage hierarchical	0.99	0.99	O_{ET}	9-3-1	0.3773	0.9703	0.9990	0.9812	0.9984
			O_{MAR}	9-3-1	0.3773	0.9703	0.9990	0.9812	0.9984
	0.95	0.95	O_{ET}	9-3-1	0.3798	0.8574	0.9950	0.8993	0.9925
			O_{MAR}	9-3-1	0.3798	0.8574	0.9950	0.8993	0.9925
	0.90	0.90	O_{ET}	12-4-1	0.3806	0.7290	0.9853	0.7227	0.9857
			O_{MAR}	12-4-1	0.3806	0.7290	0.9853	0.7227	0.9857
Array w/o master pooling	0.99	0.99	O_{ET}	10-1	0.3809	0.9705	0.9986	0.9735	0.9984
			O_{MAR}	10-1	0.3809	0.9705	0.9986	0.9735	0.9984
	0.95	0.95	O_{ET}	10-1	0.3852	0.8581	0.9926	0.8597	0.9925
			O_{MAR}	10-1	0.3852	0.8581	0.9926	0.8597	0.9925
	0.90	0.90	O_{ET}	10-1	0.3907	0.7302	0.9842	0.7086	0.9858
			O_{MAR}	10-1	0.3907	0.7302	0.9842	0.7086	0.9858
Array w/ master pooling	0.99	0.99	O_{ET}	100-10-1	0.3772	0.9608	0.9986	0.9736	0.9979
			O_{MAR}	100-10-1	0.3772	0.9608	0.9986	0.9736	0.9979
	0.95	0.95	O_{ET}	100-10-1	0.3660	0.8152	0.9930	0.8600	0.9903
			O_{MAR}	100-10-1	0.3660	0.8152	0.9930	0.8600	0.9903
	0.90	0.90	O_{ET}	100-10-1	0.3517	0.6572	0.9858	0.7091	0.9820
			O_{MAR}	100-10-1	0.3517	0.6572	0.9858	0.7091	0.9820

NOTE: Equally sized groups are optimal at each stage; thus, a "24-6-1" means that stage 1 has a group of size 24, stage 2 has four groups of size 6, and stage 3 has twenty-four groups of size 1. There are no differences between the OTCs.

S3.2. OTCs for informative group testing

Due to the lack of available space, Tables S5, S6, and S7 display at most only the initial (stage 1) group size for the informative hierarchical algorithms. We display their full algorithms in Tables S9 - S14. As described in Section 2, I_{sj} denotes the size of group j at stage s. For two-stage hierarchical testing, individuals are assembled into blocks (McMahan et al. 2012a), where we use a block size of 50. Thus, we have $\sum_j I_{1j} = 50$ by design.

To better understand the displayed OTCs in the tables, consider the OTC given in the first row of results in Table S12. The algorithm is performed over $S = 3$ stages with an initial group size of $I_{11} = 10$ individuals. If this initial group tests positively, four new groups are formed for the second stage of testing with sizes $I_{21} = 4$, $I_{22} = 3$, $I_{23} = 2$, and $I_{24} = 1$. Informative group testing always orders individuals by their probabilities of positivity. Therefore, the first group consists of the individuals with the four smallest probabilities, and the last group consists of the individual with the largest probability. If any of these groups test positively and has a size greater than 1, individual testing is performed upon its group members. For the first group in stage 2, this means that individual tests would be performed on each of its members in stage 3 ($I_{31} = I_{32} = I_{33} = I_{34} = 1$). For the last group in stage 2, no subsequent retesting would be performed. Figure S1 provides a pictorial representation of this group testing algorithm.

Table S3: OTC summary for $p = 0.10$ under non-informative group testing.

Algorithm	S_e	S_p	Objective function	OTC	$E(T)/I$	PS_e	PS_p	$PPPV$	$PNPV$
2-stage hierarchical	0.99	0.99	O_{ET}	4-1	0.5970	0.9801	0.9972	0.9753	0.9978
			O_{MAR}	4-1	0.5970	0.9801	0.9972	0.9753	0.9978
	0.95	0.95	O_{ET}	4-1	0.6095	0.9025	0.9853	0.8722	0.9891
			O_{MAR}	4-1	0.6095	0.9025	0.9853	0.8722	0.9891
	0.90	0.90	O_{ET}	4-1	0.6251	0.8100	0.9683	0.7396	0.9787
			O_{MAR}	4-1	0.6251	0.8100	0.9683	0.7396	0.9787
3-stage hierarchical	0.99	0.99	O_{ET}	9-3-1	0.5836	0.9703	0.9981	0.9827	0.9967
			O_{MAR}	9-3-1	0.5836	0.9703	0.9981	0.9827	0.9967
	0.95	0.95	O_{ET}	9-3-1	0.5733	0.8574	0.9905	0.9091	0.9843
			O_{MAR}	9-3-1	0.5733	0.8574	0.9905	0.9091	0.9843
	0.90	0.90	O_{ET}	9-3-1	0.5619	0.7290	0.9808	0.8081	0.9702
			O_{MAR}	9-3-1	0.5619	0.7290	0.9808	0.8081	0.9702
Array w/o master pooling	0.99	0.99	O_{ET}	7-1	0.5821	0.9705	0.9978	0.9800	0.9967
			O_{MAR}	7-1	0.5821	0.9705	0.9978	0.9800	0.9967
	0.95	0.95	O_{ET}	7-1	0.5776	0.8585	0.9888	0.8950	0.9843
			O_{MAR}	7-1	0.5776	0.8585	0.9888	0.8950	0.9843
	0.90	0.90	O_{ET}	7-1	0.5722	0.7310	0.9772	0.7808	0.9703
			O_{MAR}	7-1	0.5722	0.7310	0.9772	0.7808	0.9703
Array w/ master pooling	0.99	0.99	O_{ET}	49-7-1	0.5767	0.9608	0.9978	0.9800	0.9957
			O_{MAR}	49-7-1	0.5767	0.9608	0.9978	0.9800	0.9957
	0.95	0.95	O_{ET}	49-7-1	0.5491	0.8156	0.9894	0.8952	0.9797
			O_{MAR}	49-7-1	0.5491	0.8156	0.9894	0.8952	0.9797
	0.90	0.90	O_{ET}	49-7-1	0.5154	0.6579	0.9795	0.7812	0.9626
			O_{MAR}	49-7-1	0.5154	0.6579	0.9795	0.7812	0.9626

NOTE: Equally sized groups are optimal at each stage; thus, a "24-6-1" means that stage 1 has a group of size 24, stage 2 has four groups of size 6, and stage 3 has twenty-four groups of size 1. There are no differences between the OTCs.

S4. R programs

To reproduce the research in this paper, we make available a set of R functions in the `binGroup` package that

- Calculate $E(T)$ and associated accuracy measures for several different objective functions, and
- Find the optimal testing configuration over a wide variety of settings.

All calculations for the paper were performed in version 3.4.1 of R (R Core Team 2017).

The examples provided next show how to use `binGroup` to reproduce results from Tables 1 and 3. Examples 1 and 2 use non-informative group testing with an overall disease prevalence of $p = 0.01$. Examples 3 and 4 use informative group testing with an overall disease prevalence of $E(P_i) = 0.01$. Estimated running times for each example were calculated using a computer with 16 GB of RAM and one core of an Intel i7-6500U processor.

```
# Example 1
# Finding the OTC using non-informative
# three-stage hierarchical testing, where
#   p denotes the overall prevalence of disease,
#   Se denotes the sensitivity of the diagnostic test,
#   Sp denotes the specificity of the diagnostic test,
#   group.sz denotes the range of initial pool sizes for consideration, and
#   obj.fn specifies the objective functions for which to find results.

# This example takes approximately 2.5 minutes to run.
> results1 <- OTC(algorithm="D3", p=0.01, Se=0.99, Sp=0.99, group.sz=3:40,
```

Table S4: Largest differences between operating characteristics for OTCs under non-informative group testing.

Algorithm	S_e	S_p	Frequency	Largest difference $E(T)/I$	PS_e	PS_p	$PPPV$	$PNPV$
2-stage hierarchical	0.99	0.99	0	-	-	-	-	-
	0.95	0.95	3	0.0018	0.0000	0.0049	0.0262	0.0001
	0.90	0.90	4	0.0023	0.0000	0.0054	0.0345	0.0001
3-stage hierarchical	0.99	0.99	0	-	-	-	-	-
	0.95	0.95	1	0.0014	0.0000	0.0051	0.0296	0.0001
	0.90	0.90	3	0.0015	0.0000	0.0049	0.0575	0.0001
Array w/o master pooling	0.99	0.99	0	-	-	-	-	-
	0.95	0.95	5	0.0010	0.0018	0.0026	0.0195	0.0003
	0.90	0.90	8	0.0028	0.0022	0.0054	0.0305	0.0005
Array w/ master pooling	0.99	0.99	2	0.0005	0.0006	0.0008	0.0046	0.0001
	0.95	0.95	4	0.0012	0.0017	0.0026	0.0198	0.0003
	0.90	0.90	8	0.0015	0.0018	0.0051	0.0307	0.0005

NOTE: Values of p range from 0.005 to 0.150 by 0.005. The frequency column denotes the number of times a different OTC was found for O_{ET} and O_{MAR} among these values of p. Differences between operating characteristics are rounded to four decimal places. Note that operating characteristics are always smaller for O_{ET} than for O_{MAR} when differences exist.

```
                   obj.fn=c("ET", "MAR"))
You have specified an overall probability of disease.
A probability vector will be generated based on the algorithm specified.
Algorithm: Non-informative three-stage hierarchical testing
Initial Group Size = 3
Initial Group Size = 4
Initial Group Size = 5

<OUTPUT EDITED>

Initial Group Size = 38
Initial Group Size = 39
Initial Group Size = 40
 Number of minutes running: 2.429667

# Print the results.
> data.frame("Obj.Fn"=c("O_ET", "O_MAR"),
             "OTC"=c(paste(results1$opt.ET$OTC$Stage1,
                           results1$opt.ET$OTC$Stage2[1], 1, sep="-"),
                     paste(results1$opt.MAR$OTC$Stage1,
                           results1$opt.MAR$OTC$Stage2[1], 1, sep="-")),
             "ET.I"=c(round(results1$opt.ET$ET/results1$opt.ET$OTC$Stage1, 4),
                      round(results1$opt.MAR$ET/results1$opt.MAR$OTC$Stage1, 4)),
             "PSe"=c(round(results1$opt.ET$PSe, 4),
                     round(results1$opt.MAR$PSe, 4)),
             "PSp"=c(round(results1$opt.ET$PSp, 4),
                     round(results1$opt.MAR$PSp, 4)))
  Obj.Fn    OTC   ET.I    PSe    PSp
1   O_ET 25-5-1 0.1354 0.9703 0.9996
2  O_MAR 25-5-1 0.1354 0.9703 0.9996

# Example 2
# Finding the OTC using non-informative
# array testing with master pooling.
# The OTC differs for the ET and MAR objective functions in this example.

# This example takes approximately 2 minutes to run.
> results2 <- OTC(algorithm="A2M", p=0.01, Se=0.90, Sp=0.90, group.sz=3:30,
                  obj.fn=c("ET", "MAR"))
You have specified an overall probability of disease.
A probability vector will be generated based on the algorithm specified.
Algorithm: Non-informative square array testing with master pooling
Row/Column Size = 3, Array Size = 9
Row/Column Size = 4, Array Size = 16
Row/Column Size = 5, Array Size = 25

<OUTPUT EDITED>

Row/Column Size = 28, Array Size = 784
Row/Column Size = 29, Array Size = 841
```

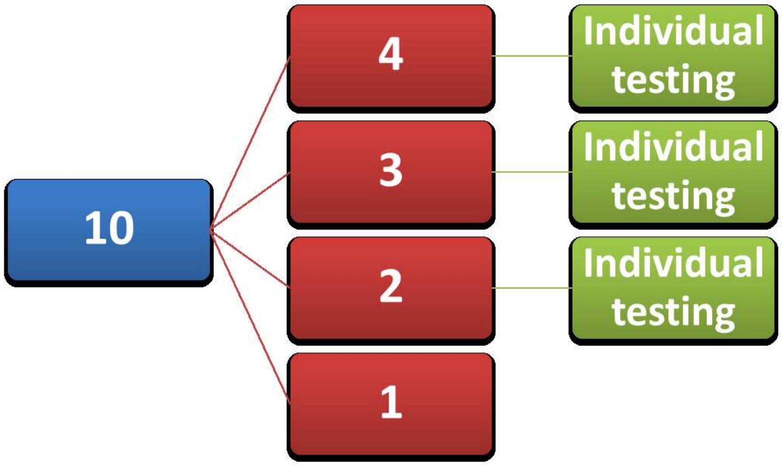

Figure S1: Diagram of the group testing algorithm described in Section S3.2. Group sizes are provided within nodes.

```
Row/Column Size = 30, Array Size = 900
 Number of minutes running: 1.745667

# Print the results.
> data.frame("Obj.Fn"=c("O_ET", "O_MAR"),
             "OTC"=c(paste(results2$opt.ET$OTC$Array.sz,
                           results2$opt.ET$OTC$Array.dim, 1, sep="-"),
                     paste(results2$opt.MAR$OTC$Array.sz,
                           results2$opt.MAR$OTC$Array.dim, 1, sep="-")),
             "ET.I"=c(round(results2$opt.ET$ET/results2$opt.ET$OTC$Array.sz, 4),
                      round(results2$opt.MAR$ET/results2$opt.MAR$OTC$Array.sz, 4)),
             "PSe"=c(round(results2$opt.ET$PSe, 4),
                     round(results2$opt.MAR$PSe, 4)),
             "PSp"=c(round(results2$opt.ET$PSp, 4),
                     round(results2$opt.MAR$PSp, 4)))
  Obj.Fn       OTC  ET.I    PSe    PSp
1   O_ET 625-25-1 0.145 0.6562 0.9934
2  O_MAR 576-24-1 0.145 0.6562 0.9937

# Example 3
# Finding the OTC using informative two-stage
# hierarchical testing, implemented via the pool-specific optimal Dorfman
# (PSOD) method described in McMahan et al. (2012), where
# alpha denotes the level of heterogeneity in the beta distribution
# used to generated the vector of individual probabilities.
# The OTC differs for the ET and MAR objective functions in this example.

# Depending on the specified probability, level of heterogeneity,
# and initial group size, simulation may be necessary in order
# to generate an ordered vector of individual probabilities. This
# is done with the beta.dist() function (see Black et al. 2015)
# using 10,000 simulated data sets.
# The user will need to set a seed in order to reproduce results.

# This examples takes approximately 2.5 minutes to run.
> set.seed(1002)
> results3 <- OTC(algorithm="ID2", p=0.01, Se=0.95, Sp=0.95, group.sz=50,
                  obj.fn=c("ET", "MAR"), alpha=2)
You have specified an overall probability of disease.
A probability vector will be generated based on the algorithm specified.
A single group size was provided. The optimal testing configuration will be found
 over all possible testing configurations for the specified group size.
NOTE: You have specified a maximum group size larger than 50.
This function may take a VERY long time to run.
Press 'ESC' if you wish to cancel the submitted statements.
Algorithm: Informative Dorfman testing
[1] "Using simulation"
Block Size = 50
[1] "Using simulation"
[1] "Using simulation"
 Number of minutes running: 2.617833

# Print the results.
> data.frame("Obj.Fn"=c("O_ET", "O_MAR"),
```

```
                 "OTC"=c(results3$opt.ET$OTC$Block.sz,
                         results3$opt.MAR$OTC$Block.sz),
                 "ET.I"=c(round(results3$opt.ET$ET/results3$opt.ET$OTC$Block.sz, 4),
                          round(results3$opt.MAR$ET/results3$opt.MAR$OTC$Block.sz, 4)),
                 "PSe"=c(round(results3$opt.ET$PSe, 4),
                         round(results3$opt.MAR$PSe, 4)),
                 "PSp"=c(round(results3$opt.ET$PSp, 4),
                         round(results3$opt.MAR$PSp, 4)))
  Obj.Fn OTC   ET.I    PSe    PSp
1   O_ET  50 0.2264 0.9025 0.9931
2  O_MAR  50 0.2264 0.9025 0.9931

# Example 4
# Finding the OTC using informative
# array testing without master pooling.

# This example takes approximately 2.5 minutes to run.
> set.seed(1002)
> results4 <- OTC(algorithm="ID3", p=0.01, Se=0.95, Sp=0.95, group.sz=3:40,
                  obj.fn=c("ET", "MAR"), alpha=0.5)
You have specified an overall probability of disease.
A probability vector will be generated based on the algorithm specified.
Algorithm: Informative three-stage hierarchical testing
Initial Group Size = 3
Initial Group Size = 4
Initial Group Size = 5

<OUTPUT EDITED>

Initial Group Size = 38
Initial Group Size = 39
Initial Group Size = 40
 Number of minutes running: 2.614333

# Print the results.
> data.frame("Obj.Fn"=c("O_ET", "O_MAR"),
             "OTC"=c(results4$opt.ET$OTC$Stage1,
                     results4$opt.MAR$OTC$Stage1),
             "ET.I"=c(round(results4$opt.ET$ET/results4$opt.ET$OTC$Stage1, 4),
                      round(results4$opt.MAR$ET/results4$opt.MAR$OTC$Stage1, 4)),
             "PSe"=c(round(results4$opt.ET$PSe, 4),
                     round(results4$opt.MAR$PSe, 4)),
             "PSp"=c(round(results4$opt.ET$PSp, 4),
                     round(results4$opt.MAR$PSp, 4)))
  Obj.Fn OTC   ET.I    PSe    PSp
1   O_ET  28 0.1291 0.8574 0.9977
2  O_MAR  28 0.1291 0.8574 0.9977
```

References

Altman, D. and Bland, J. (1994a). Diagnostic tests 1: sensitivity and specificity. *BMJ*, 308:1552.

Altman, D. and Bland, J. (1994b). Diagnostic tests 2: predictive values. *BMJ*, 309:102.

Black, M. S., Bilder, C. R., and Tebbs, J. M. (2015). Optimal retesting configurations for hierarchical group testing. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, 64(4):693–710.

Kim, H. Y., Hudgens, M. G., Dreyfuss, J. M., Westreich, D. J., and Pilcher, C. D. (2007). Comparison of group testing algorithms for case identification in the presence of test error. *Biometrics*, 63(4):1152–1163.

McMahan, C. S., Tebbs, J. M., and Bilder, C. R. (2012a). Informative Dorfman screening. *Biometrics*, 68(1):287–296.

McMahan, C. S., Tebbs, J. M., and Bilder, C. R. (2012b). Two-dimensional informative array testing. *Biometrics*, 68(3):793–804.

R Core Team (2017). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org.

Table S5: OTC summary for $E(P_i) = 0.01$ under informative group testing.

Algorithm	α	S_e	S_p	Objective function	Initial group size for OTC	$E(T)/I$	PS_e^W	PS_p^W	$PPPV^W$	$PNPV^W$
2-stage hierarchical	2	0.99	0.99	O_{ET}	-	0.1947	0.9801	0.9991	0.9204	0.9998
				O_{MAR}	-	0.1947	0.9801	0.9991	0.9204	0.9998
		0.95	0.95	O_{ET}	-	0.2264	0.9025	0.9931	0.5687	0.9990
				O_{MAR}	-	0.2264	0.9025	0.9931	0.5687	0.9990
		0.90	0.90	O_{ET}	-	0.2657	0.8100	0.9822	0.3143	0.9981
				O_{MAR}	-	0.2657	0.8100	0.9822	0.3143	0.9981
	0.5	0.99	0.99	O_{ET}	-	0.1683	0.9801	0.9992	0.9245	0.9998
				O_{MAR}	-	0.1683	0.9801	0.9992	0.9245	0.9998
		0.95	0.95	O_{ET}	-	0.2019	0.9025	0.9943	0.6115	0.9990
				O_{MAR}	-	0.2019	0.9025	0.9943	0.6115	0.9990
		0.90	0.90	O_{ET}	-	0.2439	0.8100	0.9843	0.3389	0.9981
				O_{MAR}	-	0.2439	0.8100	0.9843	0.3389	0.9981
3-stage hierarchical	2	0.99	0.99	O_{ET}	26	0.1285	0.9703	0.9996	0.9635	0.9997
				O_{MAR}	26	0.1285	0.9703	0.9996	0.9635	0.9997
		0.95	0.95	O_{ET}	26	0.1375	0.8574	0.9974	0.7655	0.9986
				O_{MAR}	26	0.1375	0.8574	0.9974	0.7655	0.9986
		0.90	0.90	O_{ET}	26	0.1497	0.7290	0.9939	0.5422	0.9973
				O_{MAR}	26	0.1497	0.7290	0.9939	0.5422	0.9973
	0.5	0.99	0.99	O_{ET}	33	0.1197	0.9703	0.9996	0.9637	0.9997
				O_{MAR}	33	0.1197	0.9703	0.9996	0.9637	0.9997
		0.95	0.95	O_{ET}	28	0.1291	0.8574	0.9977	0.7883	0.9986
				O_{MAR}	28	0.1291	0.8574	0.9977	0.7883	0.9986
		0.90	0.90	O_{ET}	29	0.1422	0.7290	0.9942	0.5558	0.9973
				O_{MAR}	29	0.1422	0.7290	0.9942	0.5558	0.9973
Array w/o master pooling	2	0.99	0.99	O_{ET}	25	0.1349	0.9703	0.9995	0.9556	0.9997
				O_{MAR}	25	0.1349	0.9703	0.9995	0.9556	0.9997
		0.95	0.95	O_{ET}	25	0.1448	0.8575	0.9972	0.7548	0.9986
				O_{MAR}	25	0.1448	0.8575	0.9972	0.7548	0.9986
		0.90	0.90	O_{ET}	25	0.1585	0.7291	0.9929	0.5089	0.9973
				O_{MAR}	25	0.1585	0.7291	0.9929	0.5089	0.9973
	0.5	0.99	0.99	O_{ET}	28	0.1277	0.9703	0.9995	0.9542	0.9997
				O_{MAR}	28	0.1277	0.9703	0.9995	0.9542	0.9997
		0.95	0.95	O_{ET}	28	0.1379	0.8574	0.9971	0.7491	0.9986
				O_{MAR}	27	0.1379	0.8574	0.9972	0.7581	0.9986
		0.90	0.90	O_{ET}	28	0.1519	0.7290	0.9927	0.5017	0.9973
				O_{MAR}	27	0.1519	0.7290	0.9930	0.5113	0.9972

NOTE: Multiple initial group sizes for 2-stage hierarchical algorithms are found within a block size of 50, so they are not displayed here. The full OTCs are given in Tables S9 and S10. Differences between O_{ET} and O_{MAR} are highlighted.

Table S6: OTC summary for $E(P_i) = 0.05$ under informative group testing.

Algorithm	α	S_e	S_p	Objective function	Initial group size for OTC	$E(T)/I$	PS_e^W	PS_p^W	$PPPV^W$	$PNPV^W$
2-stage hierarchical	2	0.99	0.99	O_{ET}	-	0.4101	0.9801	0.9981	0.9645	0.9990
				O_{MAR}	-	0.4101	0.9801	0.9981	0.9645	0.9990
		0.95	0.95	O_{ET}	-	0.4321	0.9025	0.9892	0.8152	0.9948
				O_{MAR}	-	0.4321	0.9025	0.9892	0.8152	0.9948
		0.90	0.90	O_{ET}	-	0.4586	0.8100	0.9733	0.6149	0.9898
				O_{MAR}	-	0.4586	0.8100	0.9733	0.6149	0.9898
	0.5	0.99	0.99	O_{ET}	-	0.3584	0.9801	0.9984	0.9705	0.9990
				O_{MAR}	-	0.3584	0.9801	0.9984	0.9705	0.9990
		0.95	0.95	O_{ET}	-	0.3830	0.9025	0.9908	0.8371	0.9948
				O_{MAR}	-	0.3830	0.9025	0.9908	0.8371	0.9948
		0.90	0.90	O_{ET}	-	0.4124	0.8100	0.9761	0.6403	0.9899
				O_{MAR}	-	0.4124	0.8100	0.9761	0.6403	0.9899
3-stage hierarchical	2	0.99	0.99	O_{ET}	10	0.3687	0.9725	0.9990	0.9803	0.9986
				O_{MAR}	10	0.3687	0.9725	0.9990	0.9803	0.9986
		0.95	0.95	O_{ET}	11	0.3709	0.8574	0.9940	0.8821	0.9925
				O_{MAR}	11	0.3709	0.8574	0.9940	0.8821	0.9925
		0.90	0.90	O_{ET}	12	0.3724	0.7290	0.9862	0.7357	0.9857
				O_{MAR}	12	0.3724	0.7290	0.9862	0.7357	0.9857
	0.5	0.99	0.99	O_{ET}	11	0.3365	0.9757	0.9988	0.9768	0.9987
				O_{MAR}	11	0.3365	0.9757	0.9988	0.9768	0.9987
		0.95	0.95	O_{ET}	11	0.3433	0.8822	0.9934	0.8763	0.9938
				O_{MAR}	11	0.3433	0.8822	0.9934	0.8763	0.9938
		0.90	0.90	O_{ET}	10	0.3503	0.7582	0.9871	0.7560	0.9873
				O_{MAR}	10	0.3503	0.7582	0.9871	0.7560	0.9873
Array w/o master pooling	2	0.99	0.99	O_{ET}	10	0.3677	0.9705	0.9988	0.9761	0.9984
				O_{MAR}	10	0.3677	0.9705	0.9988	0.9761	0.9984
		0.95	0.95	O_{ET}	10	0.3731	0.8582	0.9933	0.8703	0.9925
				O_{MAR}	10	0.3731	0.8582	0.9933	0.8703	0.9925
		0.90	0.90	O_{ET}	11	0.3784	0.7295	0.9836	0.7004	0.9857
				O_{MAR}	11	0.3784	0.7295	0.9836	0.7004	0.9857
	0.5	0.99	0.99	O_{ET}	13	0.3372	0.9703	0.9986	0.9727	0.9985
				O_{MAR}	13	0.3372	0.9703	0.9986	0.9727	0.9985
		0.95	0.95	O_{ET}	13	0.3421	0.8575	0.9924	0.8551	0.9926
				O_{MAR}	13	0.3421	0.8575	0.9924	0.8551	0.9926
		0.90	0.90	O_{ET}	13	0.3484	0.7292	0.9837	0.7000	0.9859
				O_{MAR}	13	0.3484	0.7292	0.9837	0.7000	0.9859

NOTE: Multiple initial group sizes for 2-stage hierarchical algorithms are found within a block size of 50, so they are not displayed here. The full OTCs are given in Tables S11 and S12. There are no differences between the OTCs.

Table S7: OTC summary for $E(P_i) = 0.10$ under informative group testing.

Algorithm	α	S_e	S_p	Objective function	Initial group size for OTC	$E(T)/I$	PS_e^W	PS_p^W	$PPPV^W$	$PNPV^W$
2-stage hierarchical	2	0.99	0.99	O_{ET}	-	0.5674	0.9801	0.9975	0.9772	0.9978
				O_{MAR}	-	0.5674	0.9801	0.9975	0.9772	0.9978
		0.95	0.95	O_{ET}	-	0.5815	0.9025	0.9863	0.8798	0.9891
				O_{MAR}	-	0.5815	0.9025	0.9863	0.8798	0.9891
		0.90	0.90	O_{ET}	-	0.5973	0.8100	0.9681	0.7382	0.9787
				O_{MAR}	-	0.5973	0.8100	0.9681	0.7382	0.9787
	0.5	0.99	0.99	O_{ET}	-	0.4868	0.9833	0.9977	0.9793	0.9981
				O_{MAR}	-	0.4868	0.9833	0.9977	0.9793	0.9981
		0.95	0.95	O_{ET}	-	0.5054	0.9148	0.9873	0.8887	0.9905
				O_{MAR}	-	0.5054	0.9148	0.9873	0.8887	0.9905
		0.90	0.90	O_{ET}	-	0.5271	0.8333	0.9695	0.7522	0.9813
				O_{MAR}	-	0.5271	0.8333	0.9695	0.7522	0.9813
3-stage hierarchical	2	0.99	0.99	O_{ET}	5	0.5567	0.9764	0.9981	0.9824	0.9974
				O_{MAR}	5	0.5567	0.9764	0.9981	0.9824	0.9974
		0.95	0.95	O_{ET}	8	0.5550	0.8691	0.9891	0.8985	0.9855
				O_{MAR}	8	0.5550	0.8691	0.9891	0.8985	0.9855
		0.90	0.90	O_{ET}	8	0.5461	0.7500	0.9781	0.7916	0.9724
				O_{MAR}	8	0.5461	0.7500	0.9781	0.7916	0.9724
	0.5	0.99	0.99	O_{ET}	40	0.5074	0.9733	0.9976	0.9786	0.9970
				O_{MAR}	40	0.5074	0.9733	0.9976	0.9786	0.9970
		0.95	0.95	O_{ET}	40	0.5050	0.8711	0.9876	0.8863	0.9857
				O_{MAR}	40	0.5050	0.8711	0.9876	0.8863	0.9857
		0.90	0.90	O_{ET}	40	0.4994	0.7469	0.9739	0.7604	0.9719
				O_{MAR}	40	0.4994	0.7469	0.9739	0.7604	0.9719
Array w/o master pooling	2	0.99	0.99	O_{ET}	7	0.5585	0.9706	0.9981	0.9823	0.9967
				O_{MAR}	7	0.5585	0.9706	0.9981	0.9823	0.9967
		0.95	0.95	O_{ET}	7	0.5563	0.8587	0.9900	0.9050	0.9844
				O_{MAR}	7	0.5563	0.8587	0.9900	0.9050	0.9844
		0.90	0.90	O_{ET}	8	0.5494	0.7297	0.9757	0.7697	0.9701
				O_{MAR}	8	0.5494	0.7297	0.9757	0.7697	0.9701
	0.5	0.99	0.99	O_{ET}	8	0.5092	0.9704	0.9982	0.9837	0.9967
				O_{MAR}	8	0.5092	0.9704	0.9982	0.9837	0.9967
		0.95	0.95	O_{ET}	9	0.5076	0.8576	0.9892	0.8978	0.9843
				O_{MAR}	9	0.5076	0.8576	0.9892	0.8978	0.9843
		0.90	0.90	O_{ET}	9	0.5040	0.7294	0.9777	0.7842	0.9702
				O_{MAR}	9	0.5040	0.7294	0.9777	0.7842	0.9702

NOTE: Multiple initial group sizes for 2-stage hierarchical algorithms are found within a block size of 50, so they are not displayed here. The full OTCs are given in Tables S13 and S14. There are no differences between the OTCs.

Table S8: Largest differences between operating characteristics for OTCs under informative group testing.

Algorithm	α	S_e	S_p	Frequency	Largest difference $E(T)/I$	PS_e^W	PS_p^W	$PPPV^W$	$PNPV^W$
2-stage hierarchical	2	0.99	0.99	0	-	-	-	-	-
		0.95	0.95	7	0.0006	(0.0023)	0.0011	0.0156	0.0004
		0.90	0.90	12	0.0010	(0.0052)	0.0023	0.0160	0.0007
	0.5	0.99	0.99	0	-	-	-	-	-
		0.95	0.95	3	0.0003	(0.0035)	0.0011	0.0102	0.0003
		0.90	0.90	15	0.0008	(0.0103)	0.0022	0.0277	0.0007
3-stage hierarchical	2	0.99	0.99	1	0.0000	(0.0019)	0.0002	0.0057	(0.0001)
		0.95	0.95	2	0.0035	0.0219	0.0033	0.0270	0.0034
		0.90	0.90	6	0.0044	0.0152	0.0062	0.0409	0.0023
	0.5	0.99	0.99	1	0.0000	0.0001	0.0001	0.0018	0.0000
		0.95	0.95	0	-	-	-	-	-
		0.90	0.90	3	0.0010	0.0250	0.0033	0.0296	0.0025
Array w/o master pooling	2	0.99	0.99	1	0.0003	0.0004	0.0005	0.0039	0.0001
		0.95	0.95	2	0.0011	0.0012	0.0027	0.0169	0.0002
		0.90	0.90	5	0.0016	0.0012	0.0040	0.0265	0.0003
	0.5	0.99	0.99	0	-	-	-	-	-
		0.95	0.95	4	0.0003	0.0004	0.0015	0.0129	0.0001
		0.90	0.90	14	0.0015	0.0004	0.0032	0.0194	0.0002

NOTE: Values of $E(P_i) = p$ range from 0.005 to 0.150 by 0.005. The frequency column denotes the number of times a different OTC was found among these values of p. Differences between operating characteristics are rounded to four decimal places. Note that the operating characteristic value for O_{ET} is always subtracted from the operating characteristic value for O_{MAR}. Thus, a negative value (indicated with parentheses) means that the value for O_{ET} was larger than the value for O_{MAR}.

Table S9: Full OTCs for $E(P_i) = 0.01$ under informative two-stage hierarchical group testing.

α	S_e	S_p	Objective function	Block size	$E(T)/I$	Group sizes I_{11}	I_{12}	I_{13}	I_{14}	I_{15}
2	0.99	0.99	O_{ET}	50	0.1947	16	11	9	8	6
			O_{MAR}	50	0.1947	16	11	9	8	6
	0.95	0.95	O_{ET}	50	0.2264	18	13	11	8	
			O_{MAR}	50	0.2264	18	13	11	8	
	0.90	0.90	O_{ET}	50	0.2657	18	13	11	8	
			O_{MAR}	50	0.2657	18	13	11	8	
0.5	0.99	0.99	O_{ET}	50	0.1683	25	12	8	5	
			O_{MAR}	50	0.1683	25	12	8	5	
	0.95	0.95	O_{ET}	50	0.2019	25	12	8	5	
			O_{MAR}	50	0.2019	25	12	8	5	
	0.90	0.90	O_{ET}	50	0.2439	25	12	8	5	
			O_{MAR}	50	0.2439	25	12	8	5	

NOTE: There are no differences in the OTCs for O_{ET} and O_{MAR}.

Table S10: Full OTCs for $E(P_i) = 0.01$ under informative three-stage hierarchical group testing.

α	S_e	S_p	Objective function	I_{11}	$E(T)/I$	Group sizes I_{21}	I_{22}	I_{23}	I_{24}	I_{25}
2	0.99	0.99	O_{ET}	26	0.1285	9	5	5	4	3
			O_{MAR}	26	0.1285	9	5	5	4	3
	0.95	0.95	O_{ET}	26	0.1375	10	7	5	4	
			O_{MAR}	26	0.1375	10	7	5	4	
	0.90	0.90	O_{ET}	26	0.1497	10	7	5	4	
			O_{MAR}	26	0.1497	10	7	5	4	
0.5	0.99	0.99	O_{ET}	33	0.1197	15	6	5	4	3
			O_{MAR}	33	0.1197	15	6	5	4	3
	0.95	0.95	O_{ET}	28	0.1291	13	7	5	3	
			O_{MAR}	28	0.1291	13	7	5	3	
	0.90	0.90	O_{ET}	29	0.1422	14	7	5	3	
			O_{MAR}	29	0.1422	14	7	5	3	

NOTE: There are no differences in the OTCs for O_{ET} and O_{MAR}.

Table S11: Full OTCs for $E(P_i) = 0.05$ under informative two-stage hierarchical group testing.

α	S_e	S_p	Objective function	Block size	$E(T)/I$	Group sizes I_{11}	I_{12}	I_{13}	I_{14}	I_{15}	I_{16}	I_{17}	I_{18}	I_{19}
2	0.99	0.99	O_{ET}	50	0.4101	9	7	6	6	5	5	4	4	4
			O_{MAR}	50	0.4101	9	7	6	6	5	5	4	4	4
	0.95	0.95	O_{ET}	50	0.4321	9	7	6	6	5	5	4	4	4
			O_{MAR}	50	0.4321	9	7	6	6	5	5	4	4	4
	0.90	0.90	O_{ET}	50	0.4586	10	8	7	6	5	5	5	4	
			O_{MAR}	50	0.4586	10	8	7	6	5	5	5	4	
0.5	0.99	0.99	O_{ET}	50	0.3584	16	9	6	5	4	4	3	3	
			O_{MAR}	50	0.3584	16	9	6	5	4	4	3	3	
	0.95	0.95	O_{ET}	50	0.3830	16	9	6	5	4	4	3	3	
			O_{MAR}	50	0.3830	16	9	6	5	4	4	3	3	
	0.90	0.90	O_{ET}	50	0.4124	17	9	7	5	5	4	3		
			O_{MAR}	50	0.4124	17	9	7	5	5	4	3		

NOTE: There are no differences in the OTCs for O_{ET} and O_{MAR}.

Table S12: Full OTCs for $E(P_i) = 0.05$ under informative three-stage hierarchical group testing.

α	S_e	S_p	Objective function	I_{11}	$E(T)/I$	Group sizes I_{21}	I_{22}	I_{23}	I_{24}
2	0.99	0.99	O_{ET}	10	0.3687	4	3	2	1
			O_{MAR}	10	0.3687	4	3	2	1
	0.95	0.95	O_{ET}	11	0.3709	5	3	3	
			O_{MAR}	11	0.3709	5	3	3	
	0.90	0.90	O_{ET}	12	0.3724	5	4	3	
			O_{MAR}	12	0.3724	5	4	3	
0.5	0.99	0.99	O_{ET}	11	0.3365	6	3	1	1
			O_{MAR}	11	0.3365	6	3	1	1
	0.95	0.95	O_{ET}	11	0.3433	6	3	1	1
			O_{MAR}	11	0.3433	6	3	1	1
	0.90	0.90	O_{ET}	10	0.3503	6	3	1	
			O_{MAR}	10	0.3503	6	3	1	

NOTE: There are no differences in the OTCs for O_{ET} and O_{MAR}.

Table S13: Full OTCs for $E(P_i) = 0.10$ under informative two-stage hierarchical group testing.

α	S_e	S_p	Objective function	Block size	$E(T)/I$	Group sizes I_{11}	I_{12}	I_{13}	I_{14}	I_{15}	I_{16}	I_{17}	I_{18}	I_{19}	$I_{1,10}$	$I_{1,11}$	$I_{1,12}$	$I_{1,13}$
2	0.99	0.99	O_{ET}	50	0.5674	7	6	5	4	4	4	4	4	3	3	3	3	
			O_{MAR}	50	0.5674	7	6	5	4	4	4	4	4	3	3	3	3	
	0.95	0.95	O_{ET}	50	0.5815	7	6	5	4	4	4	4	4	3	3	3	3	
			O_{MAR}	50	0.5815	7	6	5	4	4	4	4	4	3	3	3	3	
	0.90	0.90	O_{ET}	50	0.5973	8	6	5	5	4	4	4	4	4	3	3		
			O_{MAR}	50	0.5973	8	6	5	5	4	4	4	4	4	3	3		
0.5	0.99	0.99	O_{ET}	50	0.4868	13	7	5	4	4	4	3	3	3	1	1	1	1
			O_{MAR}	50	0.4868	13	7	5	4	4	4	3	3	3	1	1	1	1
	0.95	0.95	O_{ET}	50	0.5054	13	7	5	5	4	4	3	3	3	1	1	1	
			O_{MAR}	50	0.5054	13	7	5	5	4	4	3	3	3	1	1	1	
	0.90	0.90	O_{ET}	50	0.5271	14	7	6	5	4	4	4	3	1	1	1		
			O_{MAR}	50	0.5271	14	7	6	5	4	4	4	3	1	1	1		

NOTE: There are no differences in the OTCs for O_{ET} and O_{MAR}.

Table S14: Full OTCs for $E(P_i) = 0.10$ under informative three-stage hierarchical group testing.

α	S_e	S_p	Objective function	I_{11}	$E(T)/I$	Group sizes I_{21}	I_{22}	I_{23}	I_{24}	I_{25}	I_{26}	I_{27}	I_{28}	I_{29}	$I_{2,10}$
2	0.99	0.99	O_{ET}	5	0.5567	3	1	1							
			O_{MAR}	5	0.5567	3	1	1							
	0.95	0.95	O_{ET}	8	0.5550	4	3	1							
			O_{MAR}	8	0.5550	4	3	1							
	0.90	0.90	O_{ET}	8	0.5461	4	3	1							
			O_{MAR}	8	0.5461	4	3	1							
0.5	0.99	0.99	O_{ET}	40	0.5074	12	6	5	4	4	3	3	1	1	1
			O_{MAR}	40	0.5074	12	6	5	4	4	3	3	1	1	1
	0.95	0.95	O_{ET}	40	0.5050	12	6	5	4	4	3	3	1	1	1
			O_{MAR}	40	0.5050	12	6	5	4	4	3	3	1	1	1
	0.90	0.90	O_{ET}	40	0.4994	12	7	5	4	4	3	3	1	1	
			O_{MAR}	40	0.4994	12	7	5	4	4	3	3	1	1	

NOTE: There are no differences in the OTCs for O_{ET} and O_{MAR}.