Supplementary Material for "Array testing with multiplex assays"

Peijie Hou

Takeda Pharmaceutical Inc., 300 Massachusetts Avenue, Cambridge, MA 02139, USA

Joshua M. Tebbs*, Dewei Wang

Department of Statistics, University of South Carolina, Columbia, SC 29208, USA

*tebbs@stat.sc.edu

Christopher S. McMahan

Department of Mathematical Sciences, Clemson University, Clemson, SC 29634, USA

Christopher R. Bilder

Department of Statistics, University of Nebraska-Lincoln, Lincoln, NE 68583, USA

Appendix A. *Derivation of* EFF(AT) *for* $K = 2$ *diseases.* As noted in Section 3.1 in the manuscript,

$$\text{EFF(AT)} = \frac{1}{n^2}\left\{2n + n^2 \text{pr}(T_{ij1}^{(\text{AT})} + T_{ij2}^{(\text{AT})} \geq 1)\right\},$$

where $\text{pr}(T_{ij1}^{(\text{AT})} + T_{ij2}^{(\text{AT})} \geq 1) = \text{pr}(T_{ij1}^{(\text{AT})} = 1) + \text{pr}(T_{ij2}^{(\text{AT})} = 1) - \text{pr}(T_{ij1}^{(\text{AT})} = 1, T_{ij2}^{(\text{AT})} = 1)$. Calculating the marginal probabilities $\text{pr}(T_{ij1}^{(\text{AT})} = 1)$ and $\text{pr}(T_{ij2}^{(\text{AT})} = 1)$ follows from Kim et al. (2007) as described in Section 3.1. The hard part is calculating the joint probability

$$\begin{aligned}\text{pr}(T_{ij1}^{(\text{AT})} = 1, T_{ij2}^{(\text{AT})} = 1) \;=\; & \text{pr}(\mathbf{R}_i' = (1,1), \mathbf{C}_j' = (1,1)) \\ & + 2\sum_{k=1}^{2} \text{pr}\left(\mathbf{R}_i' = (1,1), C_{jk} = 1, \sum_{j'=1}^{n} C_{j'k'} = 0\right) \\ & + 2\text{pr}\left(\mathbf{R}_i' = (1,1), \sum_{j'=1}^{n} C_{j'1} = 0, \sum_{j'=1}^{n} C_{j'2} = 0\right) \\ & + 2\text{pr}\left(R_{i1} = 1, \sum_{j'=1}^{n} C_{j'1} = 0, C_{j2} = 1, \sum_{i'=1}^{n} R_{i'2} = 0\right). \quad (\text{A.1})\end{aligned}$$

We derive closed-form expressions for each probability on the RHS of Equation (A.1). Recall Assumption 3 in the manuscript which defines $S_{e:k}^{(n)}$ and $S_{p:k}^{(n)}$ as the multiplex assay sensitivity and specificity for testing row and column master pools of size n, respectively, for the kth disease ($k = 1, 2$). The following five definitions are needed.

<u>Definition 1</u>: Let $f_k(s)$ denote the probability a pool of size $s \leq n$ is diagnosed positively for the kth disease, $k = 1, 2$. Then

$$f_k(s) = S_{e:k}^{(n)} + \overline{\pi}_k^s (1 - S_{e:k}^{(n)} - S_{p:k}^{(n)}),$$

where $\overline{\pi}_k = 1 - \pi_k$ and π_k is the marginal probability of the kth disease; i.e., $\pi_1 = p_{10} + p_{11}$ and $\pi_2 = p_{01} + p_{11}$.

Definition 2: Let $f^b(s)$ denote the probability a pool of size $s \leq n$ is diagnosed positively for both diseases. Then

$$f^b(s) = \eta_0^{(s)}\overline{S}_{p:1}^{(n)}\overline{S}_{p:2}^{(n)} + \eta_1^{(s)}S_{e:1}^{(n)}\overline{S}_{p:2}^{(n)} + \eta_2^{(s)}\overline{S}_{p:1}^{(n)}S_{e:2}^{(n)} + \eta_3^{(s)}S_{e:1}^{(n)}S_{e:2}^{(n)},$$

where $\overline{S}_{e:k}^{(n)} = 1 - S_{e:k}^{(n)}$ and $\overline{S}_{p:k}^{(n)} = 1 - S_{p:k}^{(n)}$ $(k = 1, 2)$, and $\eta_0^{(s)} = p_{00}^s$, $\eta_1^{(s)} = (p_{00} + p_{10})^s - p_{00}^s$, and $\eta_2^{(s)} = (p_{00} + p_{01})^s - p_{00}^s$. The latter three quantities are the true pool status probabilities corresponding to "00," "10," and "01," respectively. The quantity $\eta_3^{(s)} = 1 - \eta_0^{(s)} - \eta_1^{(s)} - \eta_2^{(s)}$ is the probability a pool of size s is "11;" i.e., truly positive for both diseases.

Note: The expressions for $\eta_0^{(\cdot)}$, $\eta_1^{(\cdot)}$, $\eta_2^{(\cdot)}$, and $\eta_3^{(\cdot)}$ in Definition 2 are used in subsequent definitions and calculations.

Definition 3: For an array with $s_1 \leq n$ columns and $s_2 \leq n$ rows, let $g_k(c; s_1, s_2)$ denote the probability that c columns $(c = 0, 1, ..., s_1)$ are truly positive for kth disease, $k = 1, 2$. Then

$$\begin{aligned} g_1(c; s_1, s_2) &= \binom{s_1}{c}(\eta_2^{(s_2)} + \eta_3^{(s_2)})^c(1 - \eta_2^{(s_2)} - \eta_3^{(s_2)})^{s_1-c} \\ g_2(c; s_1, s_2) &= \binom{s_1}{c}(\eta_1^{(s_2)} + \eta_3^{(s_2)})^c(1 - \eta_1^{(s_2)} - \eta_3^{(s_2)})^{s_1-c}. \end{aligned}$$

Definition 4: For an array with $s \leq n$ columns and n rows, the probability c_1 columns are truly positive for the first disease and c_2 columns are truly positive for the second disease is

$$g_{c_1,c_2}(s; \boldsymbol{\eta}) = \sum_{w=0}^{\min(c_1,c_2)} \binom{s}{w}\binom{s-w}{c_1-w}\binom{s-c_1}{c_2-w}(\eta_0^{(n)})^{s-c_1-c_2+w}(\eta_1^{(n)})^{c_1-w}(\eta_2^{(n)})^{c_2-w}(\eta_3^{(n)})^{w},$$

where $\boldsymbol{\eta} = (\eta_0^{(n)}, \eta_1^{(n)}, \eta_2^{(n)}, \eta_3^{(n)})'$, for $c_1, c_2 = 0, 1, ..., s$.

Definition 5: For an array with $s_1 \leq n$ columns and $s_2 \leq n$ rows, the probability c_1 columns are truly positive for the first disease and r_2 rows are truly positive for the second disease is

$$B(s_1, c_1; s_2, r_2) = \binom{s_1}{c_1}\binom{s_2}{r_2}\overline{\pi}_2^{(s_2-r_2)c_1}\overline{\pi}_1^{(s_1-c_1)r_2}\{1 - B_1(s_1, c_1; s_2, r_2) - B_2(s_1, c_1; s_2, r_2) + B_3(s_1, c_1; s_2, r_2)\}\, p_{00}^{(s_1-c_1)(s_2-r_2)},$$

where

$$\begin{aligned} B_1(s_1, c_1; s_2, r_2) &= 1 - \{1 - \overline{\pi}_1^{r_2}(p_{00}/\overline{\pi}_2)^{s_2-r_2}\}^{c_1} \\ B_2(s_1, c_1; s_2, r_2) &= 1 - \{1 - \overline{\pi}_2^{c_2}(p_{00}/\overline{\pi}_1)^{s_1-c_1}\}^{r_2} \end{aligned}$$

and

$$B_3(s_1, c_1; s_2, r_2) = \sum_{l=1}^{c_1}\sum_{m=1}^{r_2}(-1)^{m+l}\binom{c_1}{l}\binom{r_2}{m}p_{00}^{lm}\overline{\pi}_1^{(r_2-m)l}\overline{\pi}_2^{(c_1-l)m}\left(\frac{p_{00}}{\overline{\pi}_2}\right)^{l(s_2-r_2)}\left(\frac{p_{00}}{\overline{\pi}_1}\right)^{m(s_1-c_1)}.$$

With these 5 definitions, we now provide closed-form expressions for each probability on the RHS of Equation (A.1).

The first probability $\mathrm{pr}(\mathbf{R}_i' = (1,1), \mathbf{C}_j' = (1,1))$ equals

$$\sum_{\widetilde{y}_1=0}^{1}\sum_{\widetilde{y}_2=0}^{1} \mathrm{pr}(\mathbf{R}_i' = (1,1), \mathbf{C}_j' = (1,1) | \widetilde{\mathbf{Y}}_{ij}' = (\widetilde{y}_1, \widetilde{y}_2)) \mathrm{pr}(\widetilde{\mathbf{Y}}_{ij}' = (\widetilde{y}_1, \widetilde{y}_2))$$
$$= p_{00}\{f^b(n-1)\}^2 + p_{10}(S_{e:1}^{(n)})^2\{f_2(n-1)\}^2 + p_{01}(S_{e:2}^{(n)})^2\{f_1(n-1)\}^2 + p_{11}(S_{e:1}^{(n)})^2(S_{e:2}^{(n)})^2,$$

where $f_1(\cdot)$ and $f_2(\cdot)$ are given in Definition 1 and $f^b(\cdot)$ is given in Definition 2.

The second probability $\mathrm{pr}(\mathbf{R}_i' = (1,1), C_{jk} = 1, \sum_{j'=1}^{n} C_{j'k'} = 0)$ can be written as

$$\sum_{r_k=0}^{1}\sum_{c_k=0}^{1}\sum_{r_{k'}=0}^{1}\sum_{c_{k'}=r_{k'}}^{n} \mathrm{pr}\left(\widetilde{R}_{ik} = r_k, \widetilde{C}_{jk} = c_k, \widetilde{R}_{ik'} = r_k', \sum_{j'=1}^{n} \widetilde{C}_{j'k'} = c_{k'}\right)$$
$$\times \mathrm{pr}\left(\mathbf{R}_i' = (1,1), C_{jk} = 1, \sum_{j'=1}^{n} C_{j'k'} = 0 \middle| \widetilde{R}_{ik} = r_k, \widetilde{C}_{jk} = c_k, \widetilde{R}_{ik'} = r_k', \sum_{j'=1}^{n} \widetilde{C}_{j'k'} = c_{k'}\right)$$

which equals

$$\sum_{r_k=0}^{1}\sum_{c_k=0}^{1}\sum_{r_{k'}=0}^{1}\sum_{c_{k'}=r_{k'}}^{n} \mathrm{pr}\left(\widetilde{R}_{ik} = r_k, \widetilde{C}_{jk} = c_k, \widetilde{R}_{ik'} = r_k', \sum_{j'=1}^{n} \widetilde{C}_{j'k'} = c_{k'}\right)$$
$$\times (S_{e:k}^{(n)})^{r_k+c_k}(\overline{S}_{p:k}^{(n)})^{2-r_k-c_k}(S_{e:k'}^{(n)})^{r_{k'}}(\overline{S}_{p:k'}^{(n)})^{1-r_{k'}}(\overline{S}_{e:k'}^{(n)})^{c_{k'}}(S_{p:k'}^{(n)})^{n-c_{k'}}.$$

Define

$$h_{r_k c_k r_{k'}}^{(1)}(n; k, c_{k'}) = \mathrm{pr}\left(\widetilde{R}_{ik} = r_k, \widetilde{C}_{jk} = c_k, \widetilde{R}_{ik'} = r_{k'}, \sum_{j'=1}^{n} \widetilde{C}_{j'k'} = c_{k'}\right),$$

for $r_k, c_k, r_{k'} \in \{0,1\}$, $c_{k'} \in \{r_{k'}, r_{k'}+1, \ldots n\}$, and $k = 1, 2$. By the Law of Total Probability, it follows that, for example,

$$\begin{aligned} h_{000}^{(1)}(n; k, c_{k'}) &= \sum_{c=0}^{1} \mathrm{pr}\left(\widetilde{R}_{ik} = 0, \widetilde{C}_{jk} = 0, \widetilde{R}_{ik'} = 0, \widetilde{C}_{jk'} = c, \sum_{j' \neq j} \widetilde{C}_{j'k'} = c_{k'} - c\right) \\ &= \eta_0^{(n)} \sum_{c=0}^{1} \eta_{ck'}^{(n-1)} g_{k'}(c_{k'} - c; n-1, n-1), \end{aligned}$$

where the $g_{k'}(\cdot)$ functions are given in Definition 3. The other 7 values of $h_{r_k c_k r_{k'}}^{(1)}(n; k, c_{k'})$, derived similarly, are presented below. The first two are

$$\begin{aligned} h_{001}^{(1)}(n; k, c_{k'}) &= \sum_{c=0}^{1} \eta_{ck'}^{(n)} \sum_{w=0}^{c_{k'}-c} \binom{n-1}{w} (\eta_{k'}^{(1)})^w p_{00}^{n-w-1} g_{k'}(c_{k'} - w - c; n-1, n-w-1) - h_{000}^{(1)}(n; k, c_{k'}) \\ h_{010}^{(1)}(n; k, c_{k'}) &= \eta_0^{(n)} \sum_{c=0}^{1} \eta_{(c-1)k'+3}^{(n-1)} g_{k'}(c_{k'} - c; n-1, n-1). \end{aligned}$$

The last five are

$$
\begin{aligned}
h^{(1)}_{100}(n;k,c_{k'}) &= \eta_k^{(n-1)}\sum_{c=0}^{1} p_{00}^{c}\eta_{ck'}^{(n-c)} g_{k'}(c_{k'}-c;n-1,n-1)\\
h^{(1)}_{011}(n;k,c_{k'}) &= \sum_{w=0}^{c_{k'}}\binom{n}{w}(\eta_{k'}^{(1)})^{w}p_{00}^{n-w}g_{k'}(c_{k'}-w;n-1,n-w)-h^{(1)}_{000}(n;k,c_{k'})-h^{(1)}_{001}(n;k,c_{k'})\\
&\qquad -h^{(1)}_{010}(n;k,c_{k'})\\
h^{(1)}_{101}(n;k,c_{k'}) &= \sum_{c=0}^{1}\eta_{ck'}^{(n)}g_{k'}(c_{k'}-c;n,n-1)-h^{(1)}_{000}(n;k,c_{k'})-h^{(1)}_{001}(n;k,c_{k'})-h^{(1)}_{100}(n;k,c_{k'})\\
h^{(1)}_{110}(n;k,c_{k'}) &= (\eta_0^{(n)}+\eta_k^{(n)})g_{k'}(c_{k'};n-1,n)-h^{(1)}_{000}(n;k,c_{k'})-h^{(1)}_{010}(n;k,c_{k'})-h^{(1)}_{100}(n;k,c_{k'})\\
h^{(1)}_{111}(n;k,c_{k'}) &= g_{k'}(c_{k'};n,n)-\sum_{\substack{r_k=0\\ r_k+c_k+r_{k'}<3}}^{1}\sum_{c_k=0}^{1}\sum_{r_{k'}=0}^{1}h^{(1)}_{r_kc_kr_{k'}}(n;k,c_{k'}).
\end{aligned}
$$

The third probability $\mathrm{pr}(\mathbf{R}'_i=(1,1),\sum_{j'=1}^{n}C_{j'1}=0,\sum_{j'=1}^{n}C_{j'2}=0)$ can be written as

$$
\sum_{r_1=0}^{1}\sum_{r_2=0}^{1}\sum_{c_1=r_1}^{n}\sum_{c_2=r_2}^{n}\mathrm{pr}\left(\widetilde{\mathbf{R}}'_i=(r_1,r_2),\sum_{j'=1}^{n}\widetilde{C}_{j'1}=c_1,\sum_{j'=1}^{n}\widetilde{C}_{j'2}=c_2\right)
$$
$$
\times(S_{e:1}^{(n)})^{r_1}(\overline{S}_{p:1}^{(n)})^{1-r_1}(S_{e:2}^{(n)})^{r_2}(\overline{S}_{p:2}^{(n)})^{1-r_2}(\overline{S}_{e:1}^{(n)})^{c_1}(S_{p:1}^{(n)})^{n-c_1}(\overline{S}_{e:2}^{(n)})^{c_2}(S_{p:2}^{(n)})^{n-c_2}.
$$

Define

$$
h^{(2)}_{r_1r_2}(n;c_1,c_2)=\mathrm{pr}\left(\widetilde{\mathbf{R}}'_i=(r_1,r_2),\sum_{j'=1}^{n}\widetilde{C}_{j'1}=c_1,\sum_{j'=1}^{n}\widetilde{C}_{j'2}=c_2\right),
$$

for $r_1,r_2\in\{0,1\}$, $c_1\in\{r_1,r_1+1,...,n\}$, and $c_2\in\{r_2,r_2+1,\ldots,n\}$. Using the function in Definition 4, we have

$$
\begin{aligned}
h^{(2)}_{00}(n;c_1,c_2) &= g_{c_1,c_2}(n;\eta_0^{(n)},p_{00}\eta_1^{(n-1)},p_{00}\eta_2^{(n-1)},p_{00}\eta_3^{(n-1)})\\
h^{(2)}_{01}(n;c_1,c_2) &= g_{c_1,c_2}(n;\eta_0^{(n)},p_{00}\eta_1^{(n-1)},\eta_2^{(n)},p_{00}\eta_3^{(n-1)}+p_{01}(\eta_1^{(n-1)}+\eta_3^{(n-1)}))-h^{(2)}_{00}(n;c_1,c_2)\\
h^{(2)}_{10}(n;c_1,c_2) &= g_{c_1,c_2}(n;\eta_0^{(n)},\eta_1^{(n)},p_{00}\eta_2^{(n-1)},p_{00}\eta_3^{(n-1)}+p_{10}(\eta_2^{(n-1)}+\eta_3^{(n-1)}))-h^{(2)}_{00}(n;c_1,c_2)\\
h^{(2)}_{11}(n;c_1,c_2) &= g_{c_1,c_2}(n;\eta_0^{(n)},\eta_1^{(n)},\eta_2^{(n)},\eta_3^{(n)})-\sum_{\substack{r_1=0\\ r_1+r_2<2}}^{1}\sum_{r_2=0}^{1}h^{(2)}_{r_1r_2}(n;c_1,c_2).
\end{aligned}
$$

The fourth probability $\mathrm{pr}(R_{i1}=1,\sum_{j'=1}^{n}C_{j'1}=0,C_{j2}=1,\sum_{i'=1}^{n}R_{i'2}=0)$ can be written as

$$
\sum_{r_1=0}^{1}\sum_{c_1=r_1}^{n}\sum_{c_2=0}^{1}\sum_{r_2=c_2}^{n}\mathrm{pr}\left(\widetilde{R}_{i1}=r_1,\sum_{j'=1}^{n}\widetilde{C}_{j'1}=c_1,\widetilde{C}_{j2}=c_2,\sum_{i'=1}^{n}\widetilde{R}_{i'2}=r_2\right)
$$
$$
\times(S_{e:1}^{(n)})^{r_1}(\overline{S}_{p:1}^{(n)})^{1-r_1}(\overline{S}_{e:1}^{(n)})^{c_1}(S_{p:1}^{(n)})^{n-c_1}(S_{e:2}^{(n)})^{c_2}(\overline{S}_{p:2}^{(n)})^{1-c_2}(\overline{S}_{e:2}^{(n)})^{r_2}(S_{p:2}^{(n)})^{n-r_2}.
$$

Define

$$
h^{(3)}_{r_1c_2}(n;c_1,r_2)=\mathrm{pr}\left(\widetilde{R}_{i1}=r_1,\sum_{j'=1}^{n}\widetilde{C}_{j'1}=c_1,\widetilde{C}_{j2}=c_2,\sum_{i'=1}^{n}\widetilde{R}_{i'2}=r_2\right),
$$

for $r_1, c_2 \in \{0,1\}$, $c_1 \in \{r_1, r_1+1, ..., n\}$, and $r_2 \in \{r_2, r_2+1, \ldots, n\}$. Using the function in Definition 5, we have

$$
\begin{aligned}
h_{00}^{(3)}(n; c_1, r_2) &= \eta_0^{(2n-1)} B(n-1, c_1; n-1, r_2) + \eta_0^{(n)} \eta_1^{(n-1)} B(n-1, c_1-1; n-1, r_2) \\
&\quad + \eta_0^{(n)} \eta_2^{(n-1)} B(n-1, c_1; n-1, r_2-1) \\
&\quad + p_{00} \eta_1^{(n-1)} \eta_2^{(n-1)} B(n-1, c_1-1; n-1, r_2-1) \\
h_{01}^{(3)}(n; c_1, r_2) &= \eta_0^{(n)} B(n, c_1; n-1, r_2) + \eta_2^{(n)} B(n, c_1; n-1, r_2-1) - h_{00}^{(3)}(n; c_1, r_2) \\
h_{10}^{(3)}(n; c_1, r_2) &= \eta_0^{(n)} B(n-1, c_1; n, r_2) + \eta_1^{(n)} B(n-1, c_1-1; n, r_2) - h_{00}^{(3)}(n; c_1, r_2) \\
h_{11}^{(3)}(n; c_1, r_2) &= B(n, c_1; n, r_2) - h_{00}^{(3)}(n; c_1, r_2) - h_{10}^{(3)}(n; c_1, r_2) - h_{01}^{(3)}(n; c_1, r_2).
\end{aligned}
$$

Each probability on the RHS of Equation (A.1) has been derived in closed form, thereby completing the derivation of EFF(AT) for $K = 2$ diseases.

Remark: Deriving EFF(AT) in closed form for $K > 2$ diseases would follow the same arguments as those presented herein for the $K = 2$ case; however, they would be far more tedious in the presence of potential testing error.

Appendix B. *Pooling sensitivity/specificity derivations for* AT *and* $K = 2$ *diseases.* We now derive closed-form expressions for $\text{PS}_{e:k}$ and $\text{PS}_{p:k}$ as defined in Section 3.2 in the manuscript. Recall Assumption 3 in the manuscript which defines $S_{e:k}^{(n)}$ and $S_{p:k}^{(n)}$ as the multiplex assay sensitivity and specificity for testing row and column master pools of size n, respectively, for the kth disease ($k = 1, 2$). Also in the manuscript, we let $S_{e:k}^{(1)}$ and $S_{p:k}^{(1)}$ denote the same multiplex assay accuracy probabilities for individual testing.

Recall that $\text{PS}_{e:k}$ is the probability an individual is classified as positive for the kth infection ($k = 1, 2$) given that the individual is truly positive for the kth infection; i.e.,

$$
\text{PS}_{e:k} = \text{pr}(Y_{ijk} = 1, T_{ij1}^{(\text{AT})} + T_{ij2}^{(\text{AT})} \geq 1 | \widetilde{Y}_{ijk} = 1).
$$

Using inclusion-exclusion for conditional probabilities, we can write $\text{PS}_{e:k}$ as

$$
\begin{aligned}
&\text{pr}(Y_{ijk} = 1, T_{ijk}^{(\text{AT})} = 1 | \widetilde{Y}_{ijk} = 1) + \text{pr}(Y_{ijk} = 1, T_{ijk'}^{(\text{AT})} = 1 | \widetilde{Y}_{ijk} = 1) \\
&\qquad - \text{pr}(Y_{ijk} = 1, T_{ijk}^{(\text{AT})} = 1, T_{ijk'}^{(\text{AT})} = 1 | \widetilde{Y}_{ijk} = 1). \quad \text{(B.1)}
\end{aligned}
$$

We now derive expressions for each probability on the RHS of Equation (B.1). The first probability

$$
\begin{aligned}
\text{pr}(Y_{ijk} = 1, T_{ijk}^{(\text{AT})} = 1 | \widetilde{Y}_{ijk} = 1) &= \text{pr}(Y_{ijk} = 1, R_{ik} = 1, C_{jk} = 1 | \widetilde{Y}_{ijk} = 1) \\
&\quad + 2\text{pr}\left(Y_{ijk} = 1, R_{ik} = 1, \sum_{j'=1}^{n} C_{j'k} = 0 \middle| \widetilde{Y}_{ijk} = 1 \right) \\
&= S_{e:k}^{(1)} (S_{e:k}^{(n)})^2 + 2 S_{e:k}^{(1)} S_{e:k}^{(n)} \overline{S}_{e:k}^{(n)} \left[1 - f_k(n)\right]^{n-1},
\end{aligned}
$$

where the $f_k(\cdot)$ function is given in Definition 1 (see Appendix A).

The second probability $\mathrm{pr}(Y_{ijk}=1, T_{ijk'}^{(\mathrm{AT})}=1|\widetilde{Y}_{ijk}=1)$ equals

$$
\begin{aligned}
&\frac{S_{e:k}^{(1)}}{\pi_k}\left\{\mathrm{pr}(R_{ik'}=1, C_{jk'}=1, \widetilde{Y}_{ijk}=1)+2\mathrm{pr}\left(R_{ik'}=1, \sum_{j'=1}^{n} C_{j'k'}=0, \widetilde{Y}_{ijk}=1\right)\right\} \\
&\qquad =\frac{S_{e:k}^{(1)}}{\pi_k}\left\{\eta_k^{(1)}\{f_{k'}(n-1)\}^2+p_{11}(S_{e:k'}^{(n)})^2+2p_{11}S_{e:k'}^{(n)}\overline{S}_{e:k'}^{(n)}\{1-f_{k'}(n)\}^{n-1}\right. \\
&\qquad\qquad \left.+\underbrace{\mathrm{pr}\left(R_{ik'}=1, \sum_{j'=1}^{n} C_{j'k'}=0, \widetilde{Y}_{ijk}=1, \widetilde{Y}_{ijk'}=0\right)}_{(*)}\right\},
\end{aligned}
$$

where $(*)$ can be written as

$$
\begin{aligned}
&\sum_{r_{k'}=0}^{1}\sum_{c_{k'}=r_{k'}}^{n}\mathrm{pr}\left(\widetilde{R}_{ik'}=r_{k'}, \sum_{j'=1}^{n}\widetilde{C}_{j'k'}=c_{k'}, \widetilde{Y}_{ijk}=1, \widetilde{Y}_{ijk'}=0\right) \\
&\qquad\qquad \times (S_{e:2}^{(n)})^{r_{k'}}(\overline{S}_{p:2}^{(n)})^{1-r_{k'}}(\overline{S}_{e:2}^{(n)})^{c_{k'}}(S_{p:2}^{(n)})^{n-c_{k'}}.
\end{aligned}
$$

Define

$$
h_{r_{k'}}^{(4)}(n;k,c_{k'})=\mathrm{pr}\left(\widetilde{R}_{ik'}=r_{k'}, \sum_{j'=1}^{n}\widetilde{C}_{j'k'}=c_{k'}, \widetilde{Y}_{ijk}=1, \widetilde{Y}_{ijk'}=0\right),
$$

for $r_{k'}\in\{0,1\}$, $c_{k'}\in\{r_{k'}, r_{k'}+1, ..., n\}$, and $k=1,2$. Using the functions given in Definition 3 (see Appendix A), we have

$$
\begin{aligned}
h_0^{(4)}(n;k,c_{k'}) &= \eta_k^{(1)}\{\eta_0^{(n-1)}+\eta_k^{(n-1)}\}g_{k'}(c_{k'};n-1,n) \\
h_1^{(4)}(n;k,c_{k'}) &= \eta_k^{(1)}\{\eta_0^{(n-1)}+\eta_k^{(n-1)}\}g_{k'}(c_{k'};n,n-1) \\
&\quad +\eta_k^{(1)}\{\eta_3^{(n-1)}+\eta_{k'}^{(n-1)}\}g_{k'}(c_{k'}-1;n,n-1)-h_0^{(4)}(n;k,c_{k'}).
\end{aligned}
$$

The third probability on the RHS of Equation (B.1) is the most difficult to derive. We write $\mathrm{pr}(Y_{ijk}=1, T_{ijk}^{(\mathrm{AT})}=1, T_{ijk'}^{(\mathrm{AT})}=1|\widetilde{Y}_{ijk}=1)$ as

$$
\begin{aligned}
&\mathrm{pr}(Y_{ijk}=1, R_{ik}=1, C_{jk}=1, T_{ijk'}^{(\mathrm{AT})}=1|\widetilde{Y}_{ijk}=1) \\
&\qquad +2\mathrm{pr}\left(Y_{ijk}=1, \mathbf{R}_i=(1,1)', C_{jk'}=1, \sum_{j'=1}^{n}C_{j'k}=0\middle|\widetilde{Y}_{ijk}=1\right) \\
&\qquad +2\mathrm{pr}\left(Y_{ijk}=1, \mathbf{R}_i=(1,1)', \sum_{j'=1}^{n}C_{j'1}=0, \sum_{j'=1}^{n}C_{j'2}=0\middle|\widetilde{Y}_{ijk}=1\right) \\
&\qquad +2\mathrm{pr}\left(Y_{ijk}=1, R_{i1}=1, \sum_{j'=1}^{n}C_{j'1}=0, \sum_{i'=1}^{n}R_{i'2}=0, C_{j2}=1\middle|\widetilde{Y}_{ijk}=1\right).
\end{aligned}
\tag{B.2}
$$

The first probability in (B.2) equals $(S_{e:k}^{(n)})^2\mathrm{pr}(Y_{ijk}=1, T_{ijk'}^{(\mathrm{AT})}=1|\widetilde{Y}_{ijk}=1)$, and $\mathrm{pr}(Y_{ijk}=1, T_{ijk'}^{(\mathrm{AT})}=1|\widetilde{Y}_{ijk}=1)$ was derived above; i.e., the second probability in Equation (B.1).

The second probability in (B.2) equals

$$\frac{S^{(1)}_{e:k}S^{(n)}_{e:k}\overline{S}^{(n)}_{e:k}}{\pi_k}\Big[p_{11}(S^{(n)}_{e:k'})^2\{1-f_k(n)\}^{n-1}$$
$$+f_{k'}(n-1)\text{pr}\left(R_{ik'}=1,\sum_{j'\neq j}C_{j'k}=0,\widetilde{Y}_{ijk}=1,\widetilde{Y}_{ijk'}=0\right)\Big],$$

where the $f_k(\cdot)$ function is given in Definition 1 (see Appendix A). We can write the probability $\text{pr}(R_{ik'}=1,\sum_{j'\neq j}C_{j'k}=0,\widetilde{Y}_{ijk}=1,\widetilde{Y}_{ijk'}=0)$ above as

$$\sum_{r_{k'}=0}^{1}\sum_{c_k=0}^{n-1}\text{pr}\left(\widetilde{R}_{ik'}=r_{k'},\sum_{j'\neq j}\widetilde{C}_{j'k}=c_k,\widetilde{Y}_{ijk}=1,\widetilde{Y}_{ijk'}=0\right)(S^{(n)}_{e:k'})^{r_{k'}}(\overline{S}^{(n)}_{p:k'})^{1-r_{k'}}(S^{(n)}_{p:k})^{n-c_k-1}(\overline{S}^{(n)}_{e:k})^{c_k}.$$

Define

$$h^{(5)}_{r'_k}(n;k,c_k)=\text{pr}\left(\widetilde{R}_{ik'}=r_{k'},\sum_{j'\neq j}\widetilde{C}_{j'k}=c_k,\widetilde{Y}_{ijk}=1,\widetilde{Y}_{ijk'}=0\right),$$

for $r_{k'}\in\{0,1\}$, $c_k\in\{0,1,...,n-1\}$, and $k=1,2$. Using the functions given in Definition 3 (see Appendix A), we have

$$\begin{aligned}
h^{(5)}_0(n;k,c_k) &= \eta^{(1)}_k\sum_{w=0}^{c_k}\binom{n-1}{w}(\eta^{(1)}_k)^w p_{00}^{n-1-w}g_k(c_k-w;n-1,n-w-1)\\
h^{(5)}_1(n;k,c_k) &= \eta^{(1)}_k g_k(c_k;n,n-1)-h^{(5)}_0(n;k,c_k).
\end{aligned}$$

The third probability in (B.2) equals

$$\frac{S^{(1)}_{e:k}S^{(n)}_{e:k}\overline{S}^{(n)}_{e:k}}{\pi_k}\sum_{c_k=0}^{n-1}\sum_{r_{k'}=0}^{1}\sum_{c_{k'}=r_{k'}}^{n}\text{pr}\left(\sum_{j'\neq j}\widetilde{C}_{j'k}=c_k,\widetilde{R}_{ik'}=r_{k'},\sum_{j'=1}^{n}\widetilde{C}_{j'k'}=c_{k'},\widetilde{Y}_{ijk}=1\right)$$
$$\times(\overline{S}^{(n)}_{e:k})^{c_k}(S^{(n)}_{p:k})^{n-c_k-1}(S^{(n)}_{e:k'})^{r_{k'}}(\overline{S}^{(n)}_{p:k'})^{1-r_{k'}}(\overline{S}^{(n)}_{e:k'})^{c_{k'}}(S^{(n)}_{p:k'})^{n-c_{k'}}.$$

Define

$$h^{(6)}_{r_{k'}}(n;k,c_k,c_{k'})=\text{pr}\left(\sum_{j'\neq j}\widetilde{C}_{j'k}=c_k,\widetilde{R}_{ik'}=r_{k'},\sum_{j'=1}^{n}\widetilde{C}_{j'k'}=c_{k'},\widetilde{Y}_{ijk}=1\right),$$

for $r_{k'}\in\{0,1\}$, $c_k\in\{0,1,...,n-1\}$, and $c_{k'}\in\{r_{k'},r_{k'}+1,...,n\}$.

- When $k=1$, we have

$$\begin{aligned}
h^{(6)}_0(n;1,c_1,c_2) &= \eta^{(1)}_1\{\overline{\pi}_2^{n-1}g_{c_1,c_2}(n-1;\eta^{(n)}_0,\eta^{(n)}_1,p_{00}\eta^{(n-1)}_2,\overline{\pi}_2\eta^{(n-1)}_3+\eta^{(1)}_1\eta^{(n-1)}_2)\\
&\quad+(1-\overline{\pi}_2^{n-1})g_{c_1,c_2-1}(n-1;\eta^{(n)}_0,\eta^{(n)}_1,p_{00}\eta^{(n-1)}_2,\overline{\pi}_2\eta^{(n-1)}_3+\eta^{(1)}_1\eta^{(n-1)}_2)\}\\
h^{(6)}_1(n;1,c_1,c_2) &= \{p_{11}+p_{10}(1-\overline{\pi}_2^{n-1})\}g_{c_1,c_2-1}(n-1;\eta^{(n)}_0,\eta^{(n)}_1,\eta^{(n)}_2,\eta^{(n)}_3)\\
&\quad+p_{10}\overline{\pi}_2^{n-1}g_{c_1,c_2}(n-1;\eta^{(n)}_0,\eta^{(n)}_1,\eta^{(n)}_2,\eta^{(n)}_3)-h^{(6)}_0(n;1,c_1,c_2),
\end{aligned}$$

  where the $g_{\cdot,\cdot}(\cdot;\cdot)$ function is given in Definition 4 (see Appendix A).

- When $k=2$, we have

$$\begin{aligned}
h_0^{(6)}(n;2,c_2,c_1) &= \eta_2^{(1)}\{\overline{\pi}_1^{n-1}g_{c_1,c_2}(n-1;\eta_0^{(n)},p_{00}\eta_1^{(n-1)},\eta_2^{(n)},\overline{\pi}_1\eta_3^{(n-1)}+\eta_2^{(1)}\eta_1^{(n-1)}) \\
&\quad +(1-\overline{\pi}_1^{n-1})g_{c_1-1,c_2}(n-1;\eta_0^{(n)},p_{00}\eta_1^{(n-1)},\eta_2^{(n)},\overline{\pi}_1\eta_3^{(n-1)}+\eta_2^{(1)}\eta_1^{(n-1)})\} \\
h_1^{(6)}(n;2,c_2,c_1) &= \{p_{11}+p_{01}(1-\overline{\pi}_1^{n-1})\}g_{c_1-1,c_2}(n-1;\eta_0^{(n)},\eta_1^{(n)},\eta_2^{(n)},\eta_3^{(n)}) \\
&\quad +p_{01}\overline{\pi}_1^{n-1}g_{c_1,c_2}(n-1;\eta_0^{(n)},\eta_1^{(n)},\eta_2^{(n)},\eta_3^{(n)})-h_0^{(6)}(n;2,c_2,c_1).
\end{aligned}$$

The fourth probability in (B.2) equals

$$\frac{S_{e:k}^{(1)}S_{e:k}^{(n)}\overline{S}_{e:k}^{(n)}}{\pi_k}\left\{\mathrm{pr}\left(\sum_{j'\neq j}C_{j'k}=0,\sum_{i'=1}^{n}R_{i'k'}=0,C_{jk'}=1,\widetilde{Y}_{ijk}=1,\widetilde{Y}_{ijk'}=1\right)\right.$$
$$\left.+\mathrm{pr}\left(\sum_{j'\neq j}C_{j'k}=0,\sum_{i'=1}^{n}R_{i'k'}=0,C_{jk'}=1,\widetilde{Y}_{ijk}=1,\widetilde{Y}_{ijk'}=0\right)\right\}. \tag{B.3}$$

We calculate each probability in (B.3) separately using the $B(\cdot,\cdot;\cdot,\cdot)$ function in Definition 5 (see Appendix A). The first probability in (B.3) is

$$S_{e:k'}^{(n)}\overline{S}_{e:k'}^{(n)}p_{11}\sum_{c_1=0}^{n-1}\sum_{r_2=0}^{n-1}\underbrace{\mathrm{pr}\left(\sum_{j'\neq j}\widetilde{C}_{j'1}=c_1,\sum_{i'\neq i}\widetilde{R}_{i'2}=r_2\right)}_{(**)}(\overline{S}_{e:1}^{(n)})^{c_1}(S_{p:1}^{(n)})^{n-c_1-1}(\overline{S}_{e:2}^{(n)})^{r_2}(S_{p:2}^{(n)})^{n-r_2-1},$$

where $(**)$ equals

$$\sum_{w_1=0}^{c_1}\sum_{w_2=0}^{r_2}\binom{n-1-w_1}{c_1-w_1}\binom{n-1-w_2}{r_2-w_2}\pi_1^{c_1-w_1}\overline{\pi}_1^{n-c_1-1}\pi_2^{r_2-w_2}\overline{\pi}_2^{n-r_2-1}B(n-1,w_1;n-1,w_2),$$

for $c_1,r_2\in\{0,1,...,n-1\}$. The second probability in (B.3) can be written as

$$\sum_{c_k=0}^{n-1}\sum_{c_{k'}=0}^{1}\sum_{r_{k'}=c_{k'}}^{n}\mathrm{pr}\left(\sum_{j'\neq j}\widetilde{C}_{j'k}=c_k,\sum_{i'=1}^{n}\widetilde{R}_{i'k'}=r_{k'},\widetilde{C}_{jk'}=c_{k'},\widetilde{Y}_{ijk}=1,\widetilde{Y}_{ijk'}=0\right)$$
$$\times(\overline{S}_{e:k}^{(n)})^{c_1k}(S_{p:k}^{(n)})^{n-c_k-1}(\overline{S}_{e:k'}^{(n)})^{r_{k'}}(S_{p:k'}^{(n)})^{n-r_{k'}}(S_{e:k'}^{(n)})^{c_{k'}}(\overline{S}_{p:k'}^{(n)})^{1-c_{k'}}.$$

Define

$$h_{c_{k'}}^{(7)}(n;k,c_k,r_{k'})=\mathrm{pr}\left(\sum_{j'\neq j}\widetilde{C}_{j'k}=c_k,\sum_{i'=1}^{n}\widetilde{R}_{i'k'}=r_{k'},\widetilde{C}_{jk'}=c_{k'},\widetilde{Y}_{ijk}=1,\widetilde{Y}_{ijk'}=0\right),$$

for $c_{k'}\in\{0,1\}$, $c_k\in\{0,1,...,n-1\}$, and $r_{k'}\in\{c_{k'},c_{k'}+1,...,n\}$.

- When $k=1$, we have

$$\begin{aligned}
h_0^{(7)}(n;1,c_1,r_2) &= p_{10}\overline{\pi}_2^{n-1}B(n-1,c_1;n,r_2) \\
h_1^{(7)}(n;1,c_1,r_2) &= p_{10}\sum_{w_1=0}^{c_1}\sum_{w_2=1}^{\min\{n-1,r_2\}}\binom{n-1}{w_2}\binom{n-w_1-1}{c_1-w_1}\pi_2^{w_2}\overline{\pi}_2^{n-w_2-1} \\
&\quad \times(1-\overline{\pi}_1^{w_2})^{c_1-w_1}(\overline{\pi}_1^{w_2})^{n-1-c_1}B(n-1,w_1;n-w_2,r_2-w_2).
\end{aligned}$$

- When $k=2$, we have

$$
\begin{aligned}
h_0^{(7)}(n;2,c_2,r_1) &= p_{01}\overline{\pi}_1^{n-1}B(n,r_1;n-1,c_2)\\
h_1^{(7)}(n;2,c_2,r_1) &= p_{01}\sum_{w_1=0}^{c_2}\sum_{w_2=1}^{\min\{n-1,r_1\}}\binom{n-1}{w_1}\binom{n-w_2-1}{c_2-w_2}\pi_1^{w_1}\overline{\pi}_1^{n-w_1-1}\\
&\qquad \times(1-\overline{\pi}_1^{w_1})^{c_2-w_2}\,(\overline{\pi}_2^{w_1})^{n-c_2-1}\,B(n-w_1,r_1-w_1;n-1,w_2).
\end{aligned}
$$

This completes the derivation of $\mathrm{PS}_{e:k}$ for $K=2$ diseases.

Deriving $\mathrm{PS}_{p:k}$ for $K=2$ diseases is easier because $\mathrm{PS}_{p:k}$ can be written as a function of EFF(AT) and $\mathrm{PS}_{e:k}$. Recall that $\mathrm{PS}_{p:k}$ is the probability an individual is classified as negative for the kth infection ($k=1,2$) given that the individual is truly negative for the kth infection. Therefore,

$$
1-\mathrm{PS}_{p:k}=\mathrm{pr}(Y_{ijk}=1,T_{ij1}^{(\mathrm{AT})}+T_{ij2}^{(\mathrm{AT})}\geq 1|\widetilde{Y}_{ijk}=0)=\frac{\overline{S}_{p:k}^{(1)}}{\overline{\pi}_k}\left\{\mathrm{EFF(AT)}-\frac{\pi_k\mathrm{PS}_{e:k}}{S_{e:k}^{(1)}}-\frac{2}{n}\right\}.
$$

Appendix C. *Operating characteristics of* ATM *for $K=2$ diseases.* We derive closed-form expressions for EFF(ATM) and accuracy probabilities $\mathrm{PS}_{e:k}$ and $\mathrm{PS}_{p:k}$ when ATM is used for $K=2$ diseases. The ATM algorithm was described in Section 3.3 in the manuscript. Let $S_{e:k}^{(n^2)}$ and $S_{p:k}^{(n^2)}$ denote the multiplex assay accuracy probabilities for testing the master array. The efficiency of ATM is

$$
\mathrm{EFF(ATM)}=\frac{1}{n^2}\left\{1+2n\mathrm{pr}(M_1+M_2\geq 1)+n^2\mathrm{pr}(T_{ij1}^{(\mathrm{ATM})}+T_{ij2}^{(\mathrm{ATM})}\geq 1)\right\}. \tag{C.1}
$$

We derive closed-form expressions for each probability on the RHS of Equation (C.1). The following three definitions are needed.

<u>Definition 6</u>: Let $f^o(s)$ denote the probability a pool with size $s\leq n^2$ is diagnosed positively for at least one disease. Then

$$
f^o(s)=1-\eta_0^{(s)}S_{p:1}^{(n^2)}S_{p:2}^{(n^2)}-\eta_1^{(s)}\overline{S}_{e:1}^{(n^2)}S_{p:2}^{(n^2)}-\eta_2^{(s)}S_{p:1}^{(n^2)}\overline{S}_{e:2}^{(n^2)}-\eta_3^{(s)}\overline{S}_{e:1}^{(n^2)}\overline{S}_{e:2}^{(n^2)}.
$$

<u>Definition 7</u>: Let $\widetilde{M}_k=1$ if the master array is truly positive for the kth disease; $\widetilde{M}_k=0$ otherwise. For $k=1,2$, let

$$
\begin{aligned}
\beta_k(n) &= \mathrm{pr}(R_{ik'}=1,C_{jk'}=1,\widetilde{M}_k=0)\\
&= \eta_{k'}S_{e:k'}^{(1)}S_{e:k'}^{(n)}\overline{\pi}_k^{n^2-1}+p_{00}\overline{\pi}_k^{n^2-2n+1}(\eta_0^{(n-1)}\overline{S}_{p:k'}^{(n)}+\eta_{k'}^{(n-1)}S_{e:k'}^{(n)})^2.
\end{aligned}
$$

<u>Definition 8</u>: For $k=1,2$, let

$$
\begin{aligned}
\gamma_k(n) &= \mathrm{pr}\left(R_{ik'}=1,\sum_{j=1}^{n}C_{jk'}=0,\widetilde{M}_k=0\right)\\
&= \sum_{r_{k'}=0}^{1}\sum_{c'_k=r_{k'}}^{n}\mathrm{pr}\left(\widetilde{R}_{ik'}=r_{k'},\sum_{j=1}^{n}\widetilde{C}_{jk'}=c_{k'},\widetilde{M}_k=0\right)(S_{e:k'}^{(n)})^{r_{k'}}(\overline{S}_{p:k'}^{(n)})^{1-r_{k'}}(\overline{S}_{e:k'}^{(n)})^{c_{k'}}(S_{p:k'}^{(n)})^{n-c_{k'}},
\end{aligned}
$$

for $r'_k \in \{0,1\}$. Furthermore, for $c_{k'} = \{r_{k'}, r_{k'}+1, ..., n\}$, define

$$\gamma_k^0(n; c_{k'}) \equiv \text{pr}\left(\widetilde{R}_{ik'} = 0, \sum_{j=1}^{n} \widetilde{C}_{jk'} = c_{k'}, \widetilde{M}_k = 0\right) = \eta_0^{(n)} \binom{n}{c_{k'}} (\eta_{k'}^{(n-1)})^{c_{k'}} (\eta_0^{(n-1)})^{n-c_{k'}}$$
$$\gamma_k^1(n; c_{k'}) \equiv \text{pr}\left(\widetilde{R}_{ik'} = 1, \sum_{j=1}^{n} \widetilde{C}_{jk'} = c_{k'}, \widetilde{M}_k = 0\right) = \binom{n}{c_{k'}} (\eta_{k'}^{(n)})^{c_{k'}} (\eta_0^{(n)})^{n-c_{k'}} - \gamma_k^0(n; c_{k'}).$$

With these additional definitions, we now provide closed-form expressions for each probability on the RHS of Equation (C.1). The first probability $\text{pr}(M_1 + M_2 \geq 1) = f^o(n^2)$, where $f^o(\cdot)$ is given in Definition 6. The second probability

$$\text{pr}(T_{ij1}^{(\text{ATM})} + T_{ij2}^{(\text{ATM})} \geq 1) = \sum_{m_1=0}^{1} \sum_{m_2=0}^{1} \text{pr}(M_1 + M_2 \geq 1 | \widetilde{M}_1 = m_1, \widetilde{M}_2 = m_2)$$
$$\times \text{pr}(T_{ij1}^{(\text{AT})} + T_{ij2}^{(\text{AT})} \geq 1, \widetilde{M}_1 = m_1, \widetilde{M}_2 = m_2),$$

for $m_1, m_2 \in \{0,1\}$. It is easy to show $\text{pr}(M_1 + M_2 \geq 1 | \widetilde{M}_1 = m_1, \widetilde{M}_2 = m_2) = 1 - (\overline{S}_{e:1}^{(n^2)})^{m_1} (S_{p:1}^{(n^2)})^{1-m_1} (\overline{S}_{e:2}^{(n^2)})^{m_2} (S_{p:2}^{(n^2)})^{1-m_2}$. Define

$$h_{m_1 m_2}^{(8)}(n) = \text{pr}(T_{ij1}^{(\text{AT})} + T_{ij2}^{(\text{AT})} \geq 1, \widetilde{M}_1 = m_1, \widetilde{M}_2 = m_2),$$

for $m_1, m_2 \in \{0,1\}$. Using the same conditioning arguments as in Appendix A, we have

$$h_{00}^{(8)}(n) = \eta_0^{(n^2)} \{(\overline{S}_{p:1}^{(n)})^2 + 2\overline{S}_{p:1}^{(n)} (S_{p:1}^{(n)})^n + (\overline{S}_{p:2}^{(n)})^2 + 2\overline{S}_{p:2}^{(n)} (S_{p:2}^{(n)})^n - (\overline{S}_{p:1}^{(n)})^2 (\overline{S}_{p:2}^{(n)})^2$$
$$- 2(\overline{S}_{p:1}^{(n)})^2 \overline{S}_{p:2}^{(n)} (S_{p:2}^{(n)})^n - 2(\overline{S}_{p:2}^{(n)})^2 \overline{S}_{p:1}^{(n)} (S_{p:1}^{(n)})^n - 4\overline{S}_{p:1}^{(n)} (S_{p:1}^{(n)})^n \overline{S}_{p:2}^{(n)} (S_{p:2}^{(n)})^n\}$$

$$h_{10}^{(8)}(n) = \beta_2(n) + 2\gamma_2(n) + \overline{\pi}_2^{n^2} \{(\overline{S}_{p:2}^{(n)})^2 + 2\overline{S}_{p:2}^{(n)} (S_{p:2}^{(n)})^n\} - (\overline{S}_{p:2}^{(n)})^2 \beta_2(n) - 2\overline{S}_{p:2}^{(n)} (S_{p:2}^{(n)})^n \beta_2(n)$$
$$- 2(\overline{S}_{p:2}^{(n)})^2 \gamma_2(n) - 4\overline{S}_{p:2}^{(n)} (S_{p:2}^{(n)})^n \gamma_2(n)$$

$$h_{01}^{(8)}(n) = \beta_1(n) + 2\gamma_1(n) + \overline{\pi}_1^{n^2} \{(\overline{S}_{p:1}^{(n)})^2 + 2\overline{S}_{p:1}^{(n)} (S_{p:1}^{(n)})^n\} - (\overline{S}_{p:1}^{(n)})^2 \beta_1(n) - 2\overline{S}_{p:1}^{(n)} (S_{p:1}^{(n)})^n \beta_1(n)$$
$$- 2(\overline{S}_{p:1}^{(n)})^2 \gamma_1(n) - 4\overline{S}_{p:1}^{(n)} (S_{p:1}^{(n)})^n \gamma_1(n)$$

and $h_{11}^{(8)}(n) = \text{pr}(T_{ij1}^{(\text{AT})} + T_{ij2}^{(\text{AT})} \geq 1) - h_{00}^{(8)}(n) - h_{10}^{(8)}(n) - h_{01}^{(8)}(n)$, where the functions $\beta_k(\cdot)$ and $\gamma_k(\cdot)$ are given in Definitions 7 and 8, respectively. The probability $\text{pr}(T_{ij1}^{(\text{AT})} + T_{ij2}^{(\text{AT})} \geq 1)$ was derived in Appendix A. This completes the derivation of EFF(ATM).

The pooling sensitivity for ATM is

$$\text{PS}_{e:k} = \text{pr}(Y_{ijk} = 1, T_{ij1}^{(\text{ATM})} + T_{ij2}^{(\text{ATM})} \geq 1 | \widetilde{Y}_{ijk} = 1)$$
$$= \sum_{m_{k'}=0}^{1} \text{pr}(M_1 + M_2 \geq 1 | \widetilde{M}_{k'} = m_{k'}, \widetilde{Y}_{ijk} = 1)$$
$$\times \text{pr}(Y_{ijk} = 1, T_{ij1}^{(\text{AT})} + T_{ij2}^{(\text{AT})} \geq 1, \widetilde{M}_{k'} = m_{k'} | \widetilde{Y}_{ijk} = 1). \quad \text{(C.2)}$$

The first probability on the RHS of Equation (C.3), $\text{pr}(M_1 + M_2 \geq 1|\widetilde{M}_{k'} = m_{k'}, \widetilde{Y}_{ijk} = 1) = 1 - \overline{S}_{e:k}^{(n^2)}(\overline{S}_{e:k'}^{(n^2)})^{m_{k'}}(S_{p:k'}^{(n^2)})^{1-m_{k'}}$. Define

$$h_{m_{k'}}^{(9)}(n;k) = \text{pr}(Y_{ijk} = 1, T_{ij1}^{(\text{AT})} + T_{ij2}^{(\text{AT})} \geq 1, \widetilde{M}_{k'} = m_{k'}|\widetilde{Y}_{ijk} = 1),$$

for $m_{k'} \in \{0, 1\}$. When $m_{k'} = 0$,

$$\begin{aligned} h_0^{(9)}(n;k) = \text{pr}(Y_{ijk} = 1, T_{ijk}^{(\text{AT})} = 1, \widetilde{M}_{k'} = 0|\widetilde{Y}_{ijk} = 1) \\ + \text{pr}(Y_{ijk} = 1, T_{ijk'}^{(\text{AT})} = 1, \widetilde{M}_{k'} = 0|\widetilde{Y}_{ijk} = 1) \\ - \text{pr}(Y_{ijk} = 1, T_{ij1}^{(\text{AT})} = 1, T_{ij2}^{(\text{AT})} = 1, \widetilde{M}_{k'} = 0|\widetilde{Y}_{ijk} = 1). \end{aligned} \quad (\text{C.3})$$

The first probability on the RHS of Equation (C.3) is

$$\pi_k^{-1}\{\eta_k^{(1)}\overline{\pi}_{k'}^{n^2-1}S_{e:k}^{(1)}(S_{e:k}^{(n)})^2 + 2\eta_k^{(1)}\overline{\pi}_{k'}^{n-1}(\eta_k^{(n)}\overline{S}_{e:k}^{(n)} + \eta_0^{(n)}S_{p:k}^{(n)})^{n-1}S_{e:k}^{(1)}S_{e:k}^{(n)}\overline{S}_{e:k}^{(n)}\}.$$

The second probability on the RHS of Equation (C.3) is

$$\pi_k^{-1}\eta_k^{(1)}\overline{\pi}_{k'}^{n^2-1}\{S_{e:k}^{(1)}(\overline{S}_{p:k'}^{(n)})^2 + 2S_{e:k}^{(1)}\overline{S}_{p:k'}^{(n)}(S_{p:k'}^{(n)})^n\}.$$

The third probability on the RHS of Equation (C.3) is

$$\begin{aligned} \pi_k^{-1}[\eta_k^{(1)}\overline{\pi}_{k'}^{n^2-1}S_{e:k}^{(1)}(S_{e:k}^{(n)})^2\overline{S}_{p:k'}^{(n)}\{\overline{S}_{p:k'}^{(n)} + 2(S_{p:k'}^{(n)})^n\} \\ + 2\eta_k^{(1)}\overline{\pi}_{k'}^{n-1}(\eta_k^{(n)}\overline{S}_{e:k}^{(n)} + \eta_0^{(n)}S_{p:k}^{(n)})^{n-1}S_{e:k}^{(1)}S_{e:k}^{(n)}\overline{S}_{e:k}^{(n)}\overline{S}_{p:k'}^{(n)}\{\overline{S}_{p:k'}^{(n)} + 2(S_{p:k'}^{(n)})^n\}]. \end{aligned}$$

When $m_{k'} = 1$, $h_1^{(9)}(n;k) = \text{PS}_{e:k}^{(\text{AT})} - h_0^{(9)}(n;k)$, where $\text{PS}_{e:k}^{(\text{AT})}$ was derived in Appendix B and $h_0^{(9)}(n;k)$ is given above. This completes the derivation of $\text{PS}_{e:k}$ for ATM.

The pooling specificity $\text{PS}_{p:k}$ for ATM is given by

$$\text{PS}_{p:k} = 1 - \frac{\overline{S}_{p:k}^{(1)}}{\pi_k}\left\{\text{EFF(ATM)} - \frac{\pi_k \text{PS}_{e:k}}{S_{e:k}^{(1)}} - \frac{2}{n}f^o(n^2) - \frac{1}{n^2}\right\},$$

where $f^o(\cdot)$ is given in Definition 6 and EFF(ATM) and $\text{PS}_{e:k}$ are the efficiency and pooling sensitivity derived in this web appendix.

Appendix D. *Additional results from Section 4.* Here is an outline of the material in this web appendix:

Page 12: Figure D.1. Simulation study for Cases 5-8 in Table 1. This figure complements Figure 2 in the manuscript (and is constructed in the same way as described in Section 4).

Page 13: Table D.1. Classification accuracy probabilities for Cases 1-4 in Table 1 (equal marginal disease probabilities). These are exact calculations using the derivations in Appendices B and C.

Page 14: Table D.2. Classification accuracy probabilities for Cases 5-8 in Table 1 (unequal marginal disease probabilities). These are exact calculations using the derivations in Appendices B and C.

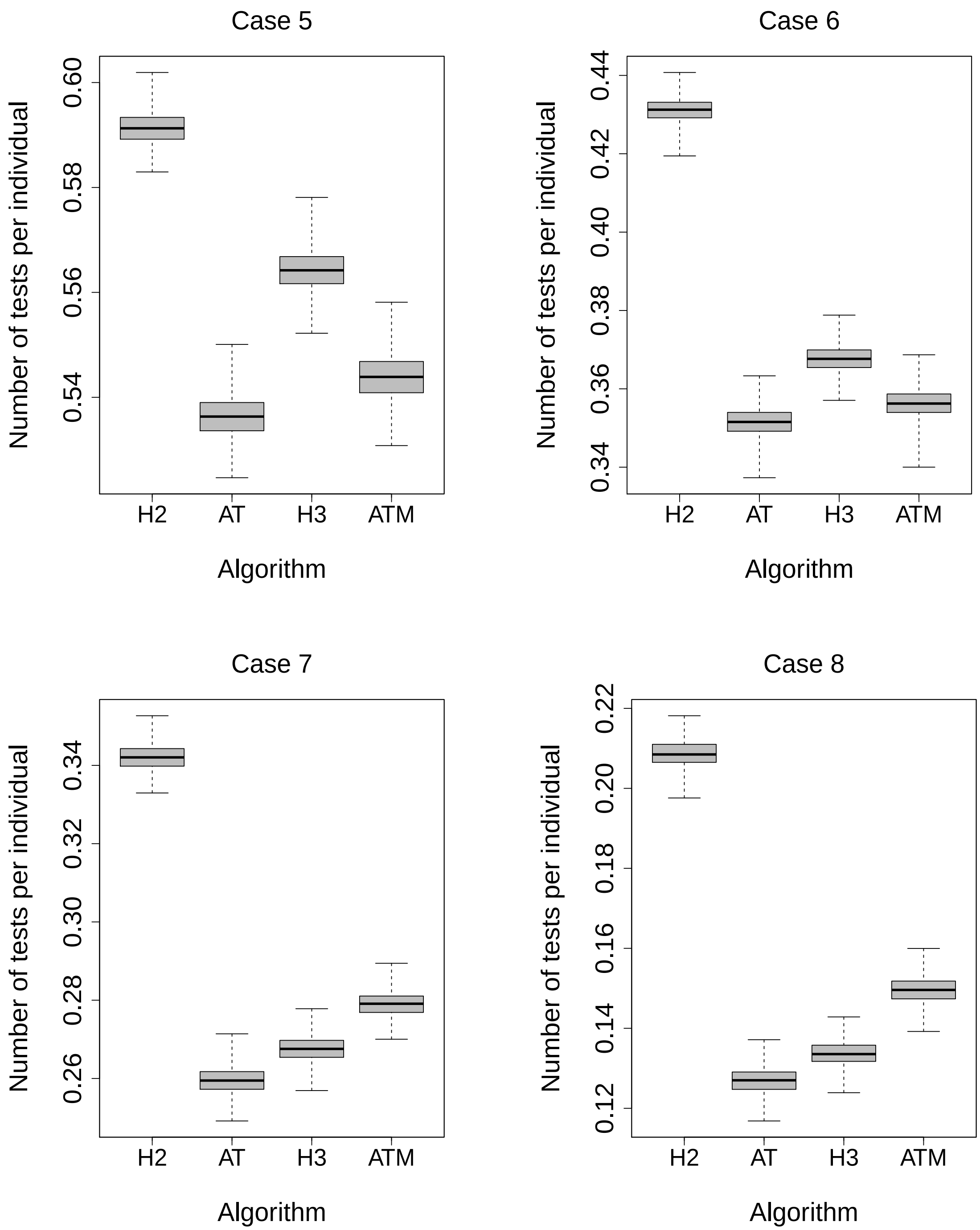

Figure D.1: Simulation study for Cases 5-8 in Table 1 in the manuscript. Boxplots of the number of tests per individual using $B = 1000$ Monte Carlo data sets. Array and hierarchical group sizes are the same as those in Table 1 in the manuscript.

Table D.1: Classification accuracy probabilities for two- and three-stage algorithms with $S_{e:k} = 0.95$ and $S_{p:k} = 0.99$ for testing all pools (regardless of size) and individuals. Cases 1-4 in this table are identical to those in Table 1 in the manuscript. Because the marginal disease probabilities $\pi_1 = p_{10} + p_{11}$ and $\pi_2 = p_{01} + p_{11}$ are equal in each case, accuracy probabilities are the same for each disease.

		Algorithm	$PS_{e:1}$	$PS_{e:2}$	$PS_{p:1}$	$PS_{p:2}$	PPV_1	PPV_2	NPV_1	NPV_2
Case 1	$p_{00} = 0.90$, $p_{10} = 0.04$, $p_{01} = 0.04$, $p_{11} = 0.02$	H2(4 : 1)	0.923	0.923	0.997	0.997	0.951	0.951	0.995	0.995
		AT(8 × 8)	0.895	0.895	0.998	0.998	0.960	0.960	0.993	0.993
		H3(9 : 3 : 1)	0.902	0.902	0.998	0.998	0.964	0.964	0.994	0.994
		ATM(8 × 8)	0.892	0.892	0.998	0.998	0.960	0.960	0.993	0.993
Case 2	$p_{00} = 0.95$, $p_{10} = 0.02$, $p_{01} = 0.02$, $p_{11} = 0.01$	H2(5 : 1)	0.921	0.921	0.998	0.998	0.931	0.931	0.998	0.998
		AT(11 × 11)	0.892	0.892	0.999	0.999	0.952	0.952	0.997	0.997
		H3(9 : 3 : 1)	0.895	0.895	0.999	0.999	0.961	0.961	0.997	0.997
		ATM(10 × 10)	0.890	0.890	0.999	0.999	0.958	0.958	0.997	0.997
Case 3	$p_{00} = 0.97$, $p_{10} = 0.01$, $p_{01} = 0.01$, $p_{11} = 0.01$	H2(7 : 1)	0.928	0.928	0.998	0.998	0.910	0.910	0.999	0.999
		AT(14 × 14)	0.903	0.903	0.999	0.999	0.949	0.949	0.998	0.998
		H3(16 : 4 : 1)	0.909	0.909	0.999	0.999	0.951	0.951	0.998	0.998
		ATM(10 × 10)	0.904	0.904	0.999	0.999	0.966	0.966	0.998	0.998
Case 4	$p_{00} = 0.99$, $p_{10} = 0.004$, $p_{01} = 0.004$, $p_{11} = 0.004$	H2(11 : 1)	0.920	0.920	0.999	0.999	0.831	0.831	1.000	1.000
		AT(29 × 29)	0.887	0.887	1.000	1.000	0.916	0.916	0.999	0.999
		H3(25 : 5 : 1)	0.892	0.892	1.000	1.000	0.925	0.925	0.999	0.999
		ATM(5 × 5)	0.912	0.912	1.000	1.000	0.977	0.977	0.999	0.999

Table D.2: Classification accuracy probabilities for two- and three-stage algorithms with $S_{e:k} = 0.95$ and $S_{p:k} = 0.99$ for testing all pools (regardless of size) and individuals. Cases 5-8 in this table are identical to those in Table 1 in the manuscript. Because the marginal disease probabilities $\pi_1 = p_{10} + p_{11}$ and $\pi_2 = p_{01} + p_{11}$ are unequal in each case, accuracy probabilities for each disease may be unequal.

		Algorithm	$PS_{e:1}$	$PS_{e:2}$	$PS_{p:1}$	$PS_{p:2}$	PPV_1	PPV_2	NPV_1	NPV_2
Case 5	$p_{00} = 0.90$	H2(4 : 1)	0.908	0.920	0.997	0.997	0.967	0.851	0.992	0.998
	$p_{10} = 0.08$	AT(8 × 8)	0.865	0.908	0.998	0.997	0.972	0.872	0.988	0.998
	$p_{01} = 0.016$	H3(9 : 3 : 1)	0.870	0.899	0.998	0.998	0.976	0.884	0.988	0.998
	$p_{11} = 0.004$	ATM(8 × 8)	0.852	0.905	0.998	0.997	0.972	0.872	0.987	0.998
Case 6	$p_{00} = 0.95$	H2(5 : 1)	0.907	0.918	0.998	0.998	0.952	0.806	0.996	0.999
	$p_{10} = 0.04$	AT(11 × 11)	0.864	0.904	0.999	0.998	0.965	0.849	0.994	0.999
	$p_{01} = 0.008$	H3(9 : 3 : 1)	0.866	0.888	0.999	0.999	0.974	0.876	0.994	0.999
	$p_{11} = 0.002$	ATM(10 × 10)	0.849	0.905	0.999	0.999	0.969	0.864	0.993	0.999
Case 7	$p_{00} = 0.97$	H2(7 : 1)	0.906	0.917	0.998	0.998	0.931	0.702	0.997	1.000
	$p_{10} = 0.025$	AT(15 × 15)	0.862	0.902	0.999	0.999	0.956	0.788	0.996	1.000
	$p_{01} = 0.004$	H3(16 : 4 : 1)	0.865	0.889	0.999	0.999	0.963	0.809	0.996	0.999
	$p_{11} = 0.001$	ATM(10 × 10)	0.845	0.915	0.999	0.999	0.973	0.851	0.996	1.000
Case 8	$p_{00} = 0.99$	H2(11 : 1)	0.906	0.915	0.999	0.999	0.874	0.613	0.999	1.000
	$p_{10} = 0.008$	AT(29 × 29)	0.862	0.889	0.999	0.999	0.934	0.757	0.999	1.000
	$p_{01} = 0.0016$	H3(25 : 5 : 1)	0.865	0.883	1.000	1.000	0.947	0.792	0.999	1.000
	$p_{11} = 0.0004$	ATM(5 × 5)	0.890	0.911	1.000	1.000	0.986	0.916	0.999	1.000

Appendix E. *Additional results from Section 5.* Here is an outline of the material in this web appendix:

Page 16: Figure E.1. Boxplots of the number of tests expended for four sex/specimen type stratum. These were constructed from the feasibility study described in Section 5 in the manuscript.

Page 17: Table E.1. Classification accuracy probabilities for H2, AT, and H3. These are exact calculations of $\text{PS}_{e:k}$, $\text{PS}_{p:k}$, PPV_k, and NPV_k using the information provided in Table 2 in the manuscript. Calculations for AT use the derivations in Appendix B. Calculations for H2 and H3 use the derivations in Tebbs et al. (2013) and Hou et al. (2017).

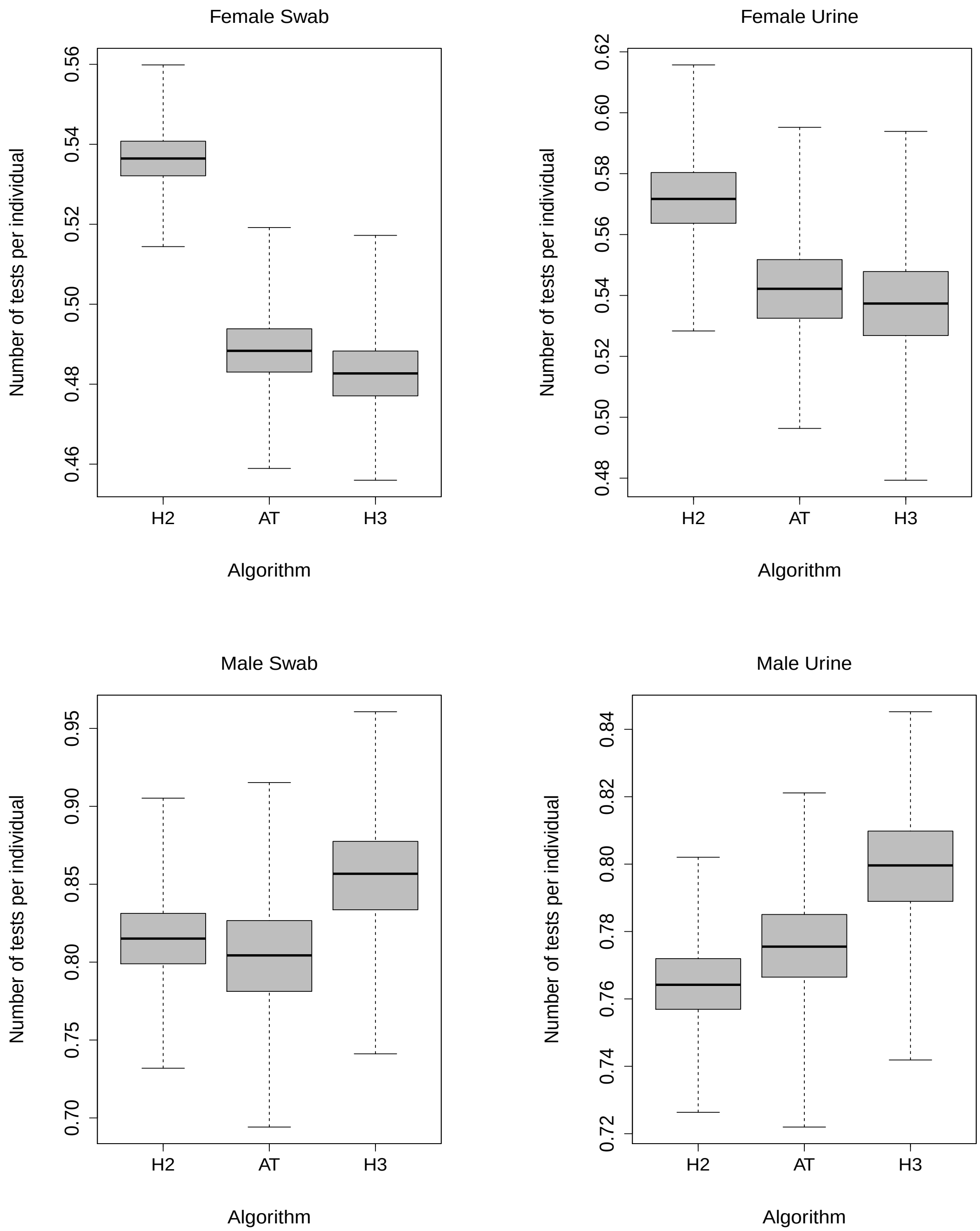

Figure E.1: Iowa SHL CT/NG feasibility study. Boxplots of the number of tests constructed from $B = 5000$ data sets as described in Section 5. The number of individuals in each sex/specimen type stratum are provided in Table 3 in the manuscript. Array and hierarchical group sizes are also provided in this table.

Table E.1: Iowa SHL example. Theoretical values of efficiency (EFF) and classification accuracy $\mathrm{PS}_{e:k}$, $\mathrm{PS}_{p:k}$, PPV_k, and NPV_k (1 = CT; 2 = NG) for each sex/specimen type stratum using the estimated probabilities $\widehat{p}_{00}$, $\widehat{p}_{10}$, $\widehat{p}_{01}$, and $\widehat{p}_{11}$ and values of $S_{e:k}$ and $S_{p:k}$ in Table 2 in the manuscript. Calculations for AT use the derivations in Appendices A and B. Calculations for H2 and H3 use the derivations in Tebbs et al. (2013) and Hou et al. (2017).

Stratum	Algorithm	EFF	$\mathrm{PS}_{e:1}$	$\mathrm{PS}_{e:2}$	$\mathrm{PS}_{p:1}$	$\mathrm{PS}_{p:2}$	PPV_1	PPV_2	NPV_1	NPV_2
Female Swab ($N = 20332$)	H2(4 : 1)	0.536	0.891	0.989	0.994	0.996	0.927	0.619	0.991	1.000
	AT(8×8)	0.487	0.844	0.989	0.996	0.997	0.938	0.664	0.988	1.000
	H3(9 : 3 : 1)	0.483	0.844	0.986	0.996	0.997	0.950	0.698	0.988	1.000
Female Urine ($N = 5998$)	H2(4 : 1)	0.571	0.900	0.877	0.997	0.998	0.969	0.736	0.990	0.999
	AT(7×7)	0.541	0.856	0.890	0.998	0.998	0.976	0.782	0.986	0.999
	H3(9 : 3 : 1)	0.537	0.856	0.851	0.998	0.998	0.978	0.787	0.986	0.999
Male Swab ($N = 1298$)	H2(3 : 1)	0.815	0.928	0.987	0.990	0.990	0.947	0.839	0.987	0.999
	AT(6×6)	0.803	0.895	0.984	0.990	0.990	0.946	0.842	0.981	0.999
	H3(9 : 3 : 1)	0.855	0.906	0.985	0.991	0.990	0.948	0.844	0.982	0.999
Male Urine ($N = 6183$)	H2(3 : 1)	0.764	0.960	0.979	0.995	0.998	0.973	0.926	0.993	1.000
	AT(5×5)	0.776	0.942	0.980	0.996	0.999	0.978	0.936	0.989	1.000
	H3(9 : 3 : 1)	0.799	0.944	0.977	0.995	0.998	0.974	0.928	0.990	0.999