Statistical Issues in Designing an Optimal Detection System with Multiple Heterogeneous Sensors

Carol Y. Lin
Global Biometric Sciences, Bristol-Myers Squibb Company
311 Pennington-Rocky Hill Road, Pennington, NJ 08534

Lance A. Waller
Department of Biostatistics, Rollins School of Public Health
Emory University, 1518 Clifton Road NE, Atlanta, GA 30322
lwaller@sph.emory.edu

Robert H. Lyles
Department of Biostatistics, Rollins School of Public Health
Emory University, 1518 Clifton Road NE, Atlanta, GA 30322

P. Barry Ryan
Department of Environmental and Occupational Health, Rollins School of Public Health
Emory University, 1518 Clifton Road NE, Atlanta, GA 30322
Abstract

Combining individual tests or sensors in a detection system often improves overall diagnostic performance, and many optimal decision-theoretic combinations under many different criteria exist. However, when designing a detection system, cost-effectiveness is important, particularly when combining a set of sensors with heterogeneous individual cost and performance. We consider the expected cost of correct and incorrect decisions of the system, constrained by the budget available to select an optimal system with various combinations of different types of conditionally independent and dependent sensors. We illustrate the approach using a hypothetical network of two different types of air pollution monitors.

Keywords: data fusion, decision fusion, combining information, false alarm rate, decision theory
1 Introduction

Detection (diagnosis) techniques play an important role in clinical medicine, environmental monitoring, industrial quality control, as well as many other fields. The purpose of detection is to provide reliable information on the presence or absence of a underlying phenomenon (disease status, critical levels of air pollution, presence of explosives, etc.) in order to allow early intervention. Detection could also provide better understanding of the mechanisms driving an underlying phenomenon so one can take measures to prevent its occurrence.

Combination of multiple tests or sensors into a system often improves performance over that of individual sensors. For example, physicians may use a set of biomarkers or combination of biomarkers and screening tests to diagnose a disease of interest (Su and Liu 1993, Woolas et al. 1995, Zhang et al. 1999); environmental regulators may use a network of air pollution monitors to detect air pollution levels exceeding a regulatory threshold; or an epidemiologist may use a disease surveillance system monitoring patient diagnoses for disease incidence associated with an influenza outbreak.

The probability of detection and the false alarm rate are common measures used to evaluate the performance for an individual test or sensor and similar definitions apply to a detection system. The probability of detection (sensitivity) is the conditional probability that a phenomenon is detected given it is present. The false alarm rate (1-specificity) is the conditional probability of an “alarm” given the phenomenon is not present. Cost constraints influence a system’s design, in particular, various types of sensors might come with different costs and systems generally are constructed under a budget. In addition, the consequences of false alarms and false negatives bear their own costs in terms of resulting subsequent events or actions. Incorporating all of these costs in a multiple-sensor detection system design is crucial for system design and success.

To set the stage for our development, consider the following situation. Suppose we wish to design a system to detect the overall presence or absence of some phenomenon of interest. Our system will be comprised of several individual sensors, but suppose we can choose from two types of sensor: expensive, precise sensors or inexpensive, imprecise sensors. Under a given budget, we can purchase a few expensive sensors or many inexpensive sensors. Our basic design criterion is
to define how many of each type of sensor to include in our system in order to maximize system performance and minimize decision costs.

Several optimal statistical approaches for “fusing” decisions from a system of distributed sensors appear in the literature. Many approaches build upon likelihood ratio methods stemming from signal detection theory developed in the 1950s and 1960s (Green and Swet 1966, Eagen 1975, Kay 1993). Relevant theoretical studies and applications appear in the engineering, signal processing, and medical diagnostic testing literature. However, there has been little cross-referencing of results between the various subject-specific literatures and a brief review provides valuable context for our development.

In the signal processing literature, Tenney and Sandell (1981) were among the first to study the problem of detection with a system of distributed sensors. Fukushima et al. (1983) proposed a neural network approach to the problem. In development more closely related to the work below, Chair and Varshney (1986) used a Bayes decision rule to derive an optimal fusion algorithm for systems of heterogeneous sensors. Their approach has the intuitive appeal of weighting observations from each sensor relative to the individual reliability of the associated sensor. Raibman and Nolte (1987a,b) further extended the ideas of local decision optimization by Tenney and Sandell (1981) together with the fusion optimization considerations of Chair and Varshney (1986) to derive an overall optimal fusion design. Kam and Chang (1991), Kam et al. (1992), Drakopoulos and Lee (1991) and Willett et al. (2000) further extended the approach to systems of correlated individual sensors.

The medical literature on combining results from multiple diagnostic tests applied to the same patient addresses the problem with a variety of approaches, e.g., linear combinations of results (Su and Liu 1993, Baker 2000), regression trees (Kodell and Chen 1999), and, similar to the signal processing work of Fukushima et al. (1983), artificial neural networks (Zhang et al. 1999). Su and Lin (1993), Woolas et al. (1995), Zhang et al. (1999) and Chen and Kodell (1999) apply such methods to combinations of multiple biomarker results. In addition, Dudoit et al. (2000) review and compare related approaches arising from gene expression array applications. Since none of the regression trees, artificial neural network and other nearest neighbor methods necessarily yield
optimal combinations of biomarkers, McIntosh and Pepe (2002) extend Baker’s nonparametric algorithms (Baker 2000) to define an optimal risk score. This risk score maximizes the true positive rate simultaneously for each false-alarm rate through a binary regression approach in order to provide an optimal binary combined decision.

In this paper, we build on these developments to address the problem of designing an optimal detection system in terms of minimum decision costs. We derive and assess optimal combinations of binary detectors taking into account associated issues such as the number of individual sensors (tests) within a detection system and correlations between the sensors. In section 2, we use expected cost combined with Bayesian optimal decision criteria under minimum cost as a guide for a decision maker to select an optimal system design. In section 3, we derive a combined probability of detection for systems comprised of independent homogeneous or heterogenous sensors, then extend the derivation to cover systems of correlated individual sensors. In section 4, we assess how the numbers of individual sensors and inter-sensor correlation impact the performance of a detection system. An example derived from a heterogeneous system of air pollution monitors is given in section 5. Discussion and further extensions of the approach appear in section 6.

2 Optimal Decision Rule and Bayesian Minimum Cost Criteria

Suppose we have a set of parallel individual sensors (Figure 1) designed to detect the phenomenon of interest, and consider a statistical test associated with each sensor assessing the simple hypotheses $H_0$ : the phenomenon is absent versus $H_1$ : the phenomenon is present. This is the basic structure for binary “decision fusion” where individual sensors pass along binary decisions (but not data) to a central “fusion center” to make a final binary decision for the system (Desarathy 1994, Hall and Llinas 2001). To set notation, let $P(H_0)$ and $P(H_1)$ denote the a priori probabilities of the two hypotheses, reflecting the prevalence of the phenomenon in the general population of observations. Let $x_i$ denote an observation from the $i^{th}$ individual sensor and the individual binary decision, $u_i$, is a function of the individual observation, $x_i$ ($u_i = 1$ if deciding $H_1$ and $u_i = 0$ if $H_0$). Let
\( \mathbf{u} = (u_1, u_2, \ldots, u_n) \) be the vector of individual sensor (test) decisions. \( P_{D_i} \) denotes the probability of detection and \( P_{F_i} \) denotes the false alarm rate for the \( i^{th} \) individual sensor. Let \( U_f \) stand for the combined (binary) decision at the fusion center. Also, let \( P_D \) denote the system-wide probability of detection and \( P_F \) denote the associated system-wide false alarm rate. Finally, let \( C \) denote cost and \( C_{r,s} \) denote the cost of deciding \( H_r \) when \( H_s \) is true, where \( r, s = 0 \) or \( 1 \). For design purposes, we assume these costs to be known.

Given we know the probability of detection and false alarm rate of the individual sensors, our objective is to select an optimal detection system design. The optimal detection system design yields the highest performance in terms of the lowest cost, given our monetary constraints. Let the risk function be defined as the average cost of correct and incorrect decisions:

\[
R = E(C) = \sum_{r=0}^{1} \sum_{s=0}^{1} C_{r,s} P(U_f = r, H_s) = \sum_{r=0}^{1} \sum_{s=0}^{1} C_{r,s} P(U_f = r | H_s) P(H_s). \tag{1}
\]

Let us consider a comparison of two detection systems: system A and system B. Systems A and B can differ by numbers of individual sensors or systems A and B can be systems with different combinations of heterogenous types of sensors. Let \( R^A \) be the risk function of system A and \( R^B \) be the risk function of system B. Let \( (P^A_D, P^A_F) \) and \( (P^B_D, P^B_F) \) denote the probability of detection and false alarm rate for system A and B, respectively. Now,

\[
R^A - R^B = P(H_1)(C_{01} - C_{11})(P^B_D - P^A_D) + P(H_0)(C_{10} - C_{00})(P^A_F - P^B_F). \tag{2}
\]

Therefore, B is a better design if \( R^A - R^B > 0 \). Given \( P(H_1) \) and \( C_{r,s} \) are known, the magnitude of the risk differences is a function of the two bivariate measures of system performance: \( (P^A_D, P^A_F) \) and \( (P^B_D, P^B_F) \).

To obtain \( P_D \) and \( P_F \), a decision criteria is needed to make a global decision, \( U_f \), given a set of individual sensor results, \( \mathbf{u} \). A Bayesian optimal decision criterion under minimum cost combines individual sensor decisions via the likelihood ratio yielding the system decision criterion:

\[
\Lambda(\mathbf{u}) = \frac{P(\mathbf{u}|H_1)}{P(\mathbf{u}|H_0)} \begin{cases} \frac{P(H_0)(C_{00})}{P(H_1)(C_{01} - C_{11})} \\ \frac{P(H_1)(C_{10} - C_{00})}{P(H_0)(C_{11})} \end{cases} \tag{3}
\]
\[ E(C|u, U_f = 0) \geq_{H_1} E(C|u, U_f = 1) \]


This criterion allows decision makers to take into account the costs of consequences of all alternatives. That is, decisions made by such criteria minimize average costs of correct and incorrect decisions, i.e. \{phenomenon presence, alarm\}, \{phenomenon absence, no alarm\}, \{phenomenon presence, no alarm\}, and \{phenomenon absence, alarm\}.

3 Combined Probabilities of Detection and False Alarm Rates

Suppose there are \(n\) individual tests, \(1, 2, 3, \ldots, n\), and \(j\) types of individual sensors, \(1, 2, 3, \ldots, z\). The Bayesian optimal decision criteria with minimum cost in equation (3) are used to make a fused decision at the fusion center. \(P_D\) and \(P_F\) are given by:

\[
P_D = P(\Lambda(u) \geq \gamma|H_1) = \sum_{\Lambda(u) \geq \gamma} P(u|H_1),
\]

\[
P_F = P(\Lambda(u) \geq \gamma|H_0) = \sum_{\Lambda(u) \geq \gamma} P(u|H_0),
\]

where \(\Lambda(u) = \frac{P(u|H_1)}{P(u|H_0)}, \gamma = \frac{P(H_0)(C_{10} - C_{00})}{P(H_1)(C_{01} - C_{11})} \).

3.1 Independent Sensors

Let us first consider the scenario that individual sensors are conditionally independent given the true presence / absence of the phenomenon and their conditional probabilities are

\[
P(u_i|H_1) = (P_{D_i})^{U_i}(1 - P_{D_i})^{1-U_i},
\]

and

\[
P(u_i|H_0) = (P_{F_i})^{U_i}(1 - P_{F_i})^{1-U_i}.
\]

In this case, the conditional likelihood at the fusion center becomes,

\[
P(u|H_1) = \prod_{i=1}^{n} P(u_i|H_1) = \prod_{i=1}^{n} (P_{D_i})^{U_i}(1 - P_{D_i})^{1-U_i},
\]

7
\[ P(u|H_0) = \prod_{i=1}^{n} P(u_i|H_0) = \prod_{i=1}^{n}(P_{F})^{U_i}(1 - P_{F})^{1-U_i}, \]

yielding a likelihood ratio of

\[ \Lambda(u) = \frac{P(u|H_1)}{P(u|H_0)} = \prod_{i=1}^{n} \left( \frac{P_{D}}{P_{F}} \right)^{U_i} \left( \frac{1 - P_{D}}{1 - P_{F}} \right)^{1-U_i}. \] (4)

Considering the scenario where \( P_{D_i} > P_{F_i} \), \( \Lambda(u) \) becomes a monotone increasing function with the number of positive individual decisions, \( u_i = 1 \).

Let \( n \) be the total number of individual sensors and \( m \) be the total number of sensors yielding a positive decision (phenomenon present) within a detection system. Let \( n_j \) denote the number of sensors of type \( j \), and let \( \sum n_j = n \). Let \( (m_1, m_2, \ldots, m_z) \) denote the vector of the number of positive \( u_i \)s from type 1, 2, \ldots, \( z \) sensors and \( (k_1, k_2, \ldots, k_z) \) denote the vector of the number of positive sensors required from each type to minimally satisfy the condition, \( \Lambda(u) \geq \gamma \). Also, let \( P_{D_j} \) denote the individual probability of detection and \( P_{F_j} \) the individual false alarm rate for the \( j^{th} \) type of sensor. In the heterogeneous system, the system-wide probability of detection and false alarm rate becomes:

\[ P_D = P(\Lambda(u) \geq \gamma|H_1) \]

\[ = \sum_{\Lambda(u) \geq \gamma} P(u|H_1) \]

\[ = \sum_{(m_1, m_2, \ldots, m_z) \geq (k_1, k_2, \ldots, k_z)} \prod_{j=1}^{z} \left( \frac{n_j}{m_j} \right) (P_{D_j})^{m_j}(1 - P_{D_j})^{n_j-m_j}, \] (5)

and

\[ P_F = P(\Lambda(u) \geq \gamma|H_0) \]

\[ = \sum_{\Lambda(u) \geq \gamma} P(u|H_0) \]

\[ = \sum_{(m_1, m_2, \ldots, m_z) \geq (k_1, k_2, \ldots, k_z)} \prod_{j=1}^{z} \left( \frac{n_j}{m_j} \right) (P_{F_j})^{m_j}(1 - P_{F_j})^{n_j-m_j}. \] (6)

where, again, \( (k_1, k_2, \ldots, k_z) \) denotes the minimum of the possible combinations of \( (m_1, m_2, \ldots, m_z) \) which satisfy the condition, \( \Lambda(u) \geq \gamma \).
3.2 Correlated Sensors with Gaussian Noise

We next extend our calculations to the situation where sensors are no longer assumed to be conditionally independent. As noted by Willett et al. (2000), system performance is considerably complicated in this setting and few analytical results exist.

We address this problem in the following manner. First, suppose there are \( n \) identical correlated sensors, \( 1, 2, 3, \ldots, n \). Next, consider our observations defined by \( x_i = a \ast \mu_i + \epsilon_i \), where \( \mu_i \) is the signal that the phenomenon exists, \( \epsilon_i \) is the noise, and \( a \) is an indicator variable where \( a = 0 \) if the phenomenon is absent and \( a = 1 \) if it is present. Without loss of generality, we assume observations, \( x_1, x_2, \ldots, x_n \), are distributed as \( N(0, \Sigma) \) under \( H_0 \), and are distributed \( N(\mu, \Sigma) \) under \( H_1 \), where \( \mu = (\mu_1, \mu_2, \ldots, \mu_j, \ldots, \mu_z) \) represents the mean value of various types of sensors and

\[
\Sigma = \begin{pmatrix}
\sigma^2 & \rho & \rho & \cdots & \rho \\
\rho & \sigma^2 & \rho & \cdots & \rho \\
\rho & \rho & \sigma^2 & \cdots & \rho \\
& & & & & \\
\rho & \rho & \rho & \cdots & \sigma^2
\end{pmatrix}, 0 < \rho < 1.
\]

For simplicity of notation and development, the covariance matrix remains unchanged between \( H_0 \) and \( H_1 \), and all subunit decision thresholds are the same and equal to \( t \), i.e., \( u_i(x_i) = 0 \) if \( x_i < t \) and \( u_i(x_i) = 1 \) if \( x_i \geq t \).

The likelihood ratio now becomes:

\[
\Lambda(u) = \frac{P(u|H_1)^{n_1}}{P(u|H_0)^{n_0}} \frac{P_o(C_{10} - C_{00})}{P_1(C_{01} - C_{11})} \Rightarrow \frac{P(u_1, u_2, \ldots, u_n|H_1)^{n_1}}{P(u_1, u_2, \ldots, u_n|H_0)^{n_0}} \frac{P_o(C_{10} - C_{00})}{P_1(C_{01} - C_{11})} \Rightarrow \int_{u_1} \int_{u_2} \cdots \int_{u_n} P_{x_1, \ldots, x_n}(X_1, \ldots, X_n|H_1) dx_1 \cdots dx_n \\
\int_{u_1} \int_{u_2} \cdots \int_{u_n} P_{x_1, \ldots, x_n}(X_1, \ldots, X_n|H_0) dx_1 \cdots dx_n \Rightarrow \frac{n_1}{n_0} \frac{P_o(C_{10} - C_{00})}{P_1(C_{01} - C_{11})}. 
\]

Then, \( P_D \) and \( P_F \) for a system with \( j \) types of sensors can be obtained through Gupta (1963) as
follows:

\[ P_D = P(\Lambda(u) \geq \gamma | H_1) \]

\[ = \sum_{\Lambda(u) \geq \gamma} P(u | H_1) \]

\[ = \sum_{m_z \geq k_z} \binom{n_1}{m_2} \cdots \binom{n_z}{m_z} \int_{u_1} \cdots \int_{u_n} P_{X_1, \ldots, X_n}(X_1, \ldots, X_n | H_1) \, dx_1 \cdots dx_n \]

\[ = \sum_{m_z \geq k_z} \binom{n_1}{m_2} \cdots \binom{n_z}{m_z} \int_{u_1} \cdots \int_{u_m} \int_{u_{m+1}} \cdots \int_{u_n} P_{X_1, \ldots, X_n}(X_1, \ldots, X_n | H_1) \, dx_1 \cdots dx_n \]

\[ = \sum_{m_z \geq k_z} \binom{n_1}{m_2} \cdots \binom{n_z}{m_z} \int_{-\infty}^{\infty} \int_{-\infty}^{t} \cdots \int_{-\infty}^{t} P_{X_1, \ldots, X_n}(X_1, \ldots, X_n | H_1) \, dx_1 \cdots dx_n \]

\[ = \sum_{m_z \geq k_z} \binom{n_1}{m_2} \cdots \binom{n_z}{m_z} \int_{-\infty}^{t} \prod_{j=1}^{z} \Phi \left( \frac{t - u_j - \sqrt{\rho y}}{\sqrt{\sigma^2 - \rho}} \right)^{(n_j - m_j)} \left[ 1 - \Phi \left( \frac{t - u_j - \sqrt{\rho y}}{\sqrt{\sigma^2 - \rho}} \right) \right]^{m_j} \phi(y) \, dy, \tag{8} \]

and

\[ P_F = P(\Lambda(u) \geq \gamma | H_0) \]

\[ = \sum_{\Lambda(u) \geq \gamma} P(u | H_0) \]

\[ = \sum_{m_z \geq k_z} \binom{n_1}{m_2} \cdots \binom{n_z}{m_z} \int_{u_1} \cdots \int_{u_n} P_{X_1, \ldots, X_n}(X_1, \ldots, X_n | H_0) \, dx_1 \cdots dx_n \]

\[ = \sum_{m_z \geq k_z} \binom{n_1}{m_2} \cdots \binom{n_z}{m_z} \int_{u_1} \cdots \int_{u_m} \int_{u_{m+1}} \cdots \int_{u_n} P_{X_1, \ldots, X_n}(X_1, \ldots, X_n | H_0) \, dx_1 \cdots dx_n \]

\[ = \sum_{m_z \geq k_z} \binom{n_1}{m_2} \cdots \binom{n_z}{m_z} \int_{t}^{t} \int_{t}^{t} \cdots \int_{t}^{t} P_{X_1, \ldots, X_n}(X_1, \ldots, X_n | H_0) \, dx_1 \cdots dx_n \]

\[ = \sum_{m_z \geq k_z} \binom{n_1}{m_2} \cdots \binom{n_z}{m_z} \int_{-\infty}^{t} \prod_{j=1}^{z} \Phi \left( \frac{t - \sqrt{\rho y}}{\sqrt{\sigma^2 - \rho}} \right)^{(n_j - m_j)} \left[ 1 - \Phi \left( \frac{t - \sqrt{\rho y}}{\sqrt{\sigma^2 - \rho}} \right) \right]^{m_j} \phi(y) \, dy, \tag{9} \]

where \( \Phi(.) \) is the standard cumulative univariate normal distribution function, and \( \phi(.) \) is the standard normal density function. Recall \( k_z = (k_1, k_2, \ldots, k_z) \) denotes the minimum of the possible combinations of the \( m_z = (m_1, m_2, \ldots, m_z) \) which satisfy the condition, \( \Lambda(u) \geq \gamma \). Since \( P(H_1) \) and \( C_{r,s} \) are assumed to be known, the expected cost can be calculated by plugging \( P_D \) and \( P_F \) into (1).
4 Assessing the Impact of System Size and Correlations via Simulation

Given known values for $P(H_1)$ and $C_{r,s}$, the minimum expected decision cost (risk) of a system is only a function of $P_D$ and $P_F$. The magnitude of risk difference (2) between two systems can be compared by their respective values of $P_D$ and $P_F$. In this section, we apply the derivations above in a variety of settings to illustrate how the number of and correlation between individual sensors impact $P_D$ and $P_F$ for several types of detection systems. For simplicity, we assume $C_{10} - C_{00} = C_{01} - C_{11}$ and $P(H_1) = P(H_0) = 0.5$, but note that the results above provide a very general framework for assessment given any set of costs, a priori prevalence of the phenomenon, and individual detection probabilities. To compare systems with various numbers of homogeneous tests, we calculate probabilities of detection and false alarm rates for systems of independent sensors via equations (5) and (6) and for systems of correlated individual sensors via equations (8) and (9).

4.1 Homogeneous Individual Sensors

To illustrate comparisons between systems comprised of homogeneous individual sensors, we consider examples based on three different sets of individual probabilities of detection and false alarm rates, namely: probability of detection equals 1-false alarm rate (0.691, 0.309), probability of detection smaller than 1-false alarm rate (0.460, 0.135) and probability of detection greater than 1-false alarm rate (0.864, 0.540). The numbers of individual sensors range from 2 to 100. For correlated individual sensors, due to computational complexity, we limit reported results to scenarios in which the probability of detection equals 1-the false alarm rate (0.691, 0.309), while allowing the correlation $\rho$ to vary from 0, 0.2, 0.4 to 0.6. The number of correlated sensors considered ranges from 3 to 20.

4.2 Heterogeneous Individual Sensors

To illustrate application to systems comprised of heterogeneous sensors, we report results based on systems with 8 individual sensors (4 from each type), systems with 2 individual sensors (1 from
each type) and investigate whether the system with 8 individual sensors yields higher probability of
detection and lower false alarm rate than the 2-sensor system. We assume type II sensors perform
better than type I sensors. More specifically, we assume the probability of detection of type II
sensors is higher than for type I sensors but that the false alarm rates are the same. We vary the
probability of detection of type I sensors from 0.242 to 0.864 and of type II sensors from 0.618 to
0.982. The corresponding false alarm rates for both types of the sensors vary from 0.045 to 0.540.

For correlated individual sensors scenarios, the correlations between individual sensors are as-
sumed to be the same. We report results based on probabilities of detection for type I and type
II sensors of 0.691, 0.788, respectively. The false alarm rates for both types are set equal to 0.420.
Similar to the homogenous sensor system examples, we consider correlations $\rho = 0, 0.2, 0.4, 0.6$ and
then compare the corresponding combined probabilities of detection and false alarm rates. For
simplicity, we assume $P(H_1)$ is 0.5. We report results for small systems comprised of 4, 8, and 16
individual tests.

4.3 Results

4.3.1 Homogeneous Individual Sensors

The combined probability of detection and false alarm rate of the system with various numbers of
independent individual sensors are given in Table 1. As expected, in large systems, increasing the
number of sensors increases the combined probability of detection and reduces false alarm rates.
However, for systems comprised of a small number of individual sensors, the discrete nature of the
combined decision can result in a decrease in the probability of detection when adding a single
sensor to a system with an even number of sensors, since the number of positive individual sensors
required for a system alarm increases. In this case, the system probability of detection decreases
due to the increased “burden of proof” required for a system alarm (e.g., four out of seven versus
three out of six individual alarms), but the false alarm rate also decreases indicating observations
sufficient to trigger a system alarm are more likely to represent “true alarms”. When increasing
the number of sensors does not result in improvement on both probability of detection and false
alarm rate, the magnitude of improvement in probability of detection still exceeds the magnitude
of sacrifice in false alarm rate (or vice versa). This behavior illustrates the important design role for the calculations above, particularly for small systems of detectors.

Since our illustrations assume $C_{10} - C_{00} = C_{01} - C_{11}$ and $P(H_1) = P(H_0) = 0.5$, comparison of minimum cost is equivalent to comparison of magnitude of increase (reduction) of probability of detection and false alarm rates. Similar results (Table 2) are found in a system with correlated individual sensors. Furthermore, the results suggest that increasing the correlations of individual sensors decreases the combined probability of detection and increases the false alarm rate of the system. As expected, the magnitude of decline increases with the number of correlated individual sensors, and increasing correlation results in a smaller effective number of sensors.

### 4.3.2 Heterogeneous Individual Tests

The combined probabilities of detection and false alarm rates of systems with 2 and 8 individual tests are given in Table 3. As one might expect, a system with 8 individual sensors generally yields higher probability of detection and lower false alarm rate than a system with 2 individual sensors.

The results for combined probabilities of detection and false alarm rates for systems of correlated individual sensors appear in Table 4. Similar to the results of homogenous sensor systems, the simulation results also suggest that, given the number of individual sensors, the combined probability of detection decreases and false alarm rate increases with increasing correlation of the individual sensors regardless of the number of individual sensors and combination of different types of sensors. In addition, the magnitude of decline increases with the number of correlated individual sensors.

These illustrations reveal the interplay between system design and performance and provide quantitative assessment of system performance between various system configurations for large and small designs, independent and dependent observations, and homogeneous and heterogeneous sensors. The results above largely mirror intuition but reveal important complications in systems based on small numbers of sensors, in particular the “saw-tooth” pattern associated with discrete numbers of sensors.
5 Data Example

To illustrate how the results of the preceding sections might be applied in practice, consider the design of a system of air pollution monitors to detect whether the observed levels of an air pollutant exceed some specified level. This level may be a regulatory limit, or may represent a more modest level (e.g., the annual median value) to categorize days of higher or lower exposure for longitudinal studies of associated health effects.

Nitrogen dioxide, ozone, and carbon monoxide are all regulated atmospheric pollutants in the United States. Nitrogen dioxide, in particular, can cause eye, nose, and throat irritation as well as chest pain, impaired lung function and increased respiratory infections in young children. As part of an ongoing research program, a system of air pollution monitors is currently in place around Logan International Airport in Boston and surrounding communities to monitor the level of nitrogen dioxide concentrations (Ayers 2006). We focus on two types of air pollution monitors: a low-cost, but highly-variable passive monitor and an expensive, high-resolution continuous monitor. High-resolution monitors yield lower-variance measurements than the low-cost monitors, and have better probability of detection for levels exceeding a pre-specified value, especially for short durations. To compare the performance of various detection systems, we use the results from the preceding sections to calculate the combined probability of detection and false alarm rate for different hypothetical designs of a detection system in this setting.

Observations from the monitors at the normal nitrogen dioxide concentration level (noise alone) around Logan International Airport are derived by ongoing studies (Ayers 2006). The mean of the normal nitrogen dioxide concentration level is approximately 25 parts per billion (ppb), where the EPA standard is currently 53 ppb. For the purposes of illustration, the means of the observations for irregular (signal-plus-noise) nitrogen dioxide concentration levels are assumed to be 9, 20, 36, and 56 ppb higher than the mean of the normal observations. Using laboratory calibrations based on coefficients of variation for monitors, we assume that the variance of a low-cost monitor is 10% of the mean and the variance of the high resolution monitor is 5% of the mean. After a log transformation, nitrogen dioxide concentrations are assumed to be normally distributed. To illustrate our approach, we calculate the probability of detection and false alarm rate of individual
monitors using a cutoff threshold of 30 ppb representing an example of a mid-level categorization of pollution levels into “higher” and “lower” daily levels (e.g., for public health studies), rather than a regulatory alarm cutoff of 53 ppb, due to the very small number of such regulatory events.

We consider four different performance settings (probability of detection and false alarm rate pairs), namely (0.537, 0.417), (0.631,0.417), (0.702, 0.417), (0.755, 0.417) for the low cost monitors and (0.566,0.329), (0.672, 0.329), (0.747, 0.329) and (0.809, 0.329) for the expensive monitors. For simplicity, we assume the cost of a false alarm is the same as for a false negative. Again for illustration purposes, the background probability of nitrogen dioxide concentration exceeding our reference level under normal conditions is assumed to be 0.5.

First, we use equations (5) and (6) to calculate the combined probability of detection and false alarm rate for systems containing only low-cost monitors. The total numbers of monitors range from 8 to 50. The results for the homogeneous system of low-cost monitors appear in Table 5 and Figure 2. While we see general trends of improvement in performance as the number of individual monitors increases, we note that this improvement is not necessarily monotonic, particularly for the lowest values considered for the individual probability of detection. In addition, Table 5 reveals that the number of individual alarms required for a system alarm is close to, but not necessarily equal to, one-half of the (here, even) number of sensors in the system. Both settings reveal the value of the derivations above in determining system design and performance.

[Figure 2 about here.]

Next, we consider replacing 25% and 50% of the monitors in each system with the higher-cost, higher-precision monitors. Results for designs based on 8, 16, 32, and 64 monitors appear in Table 6. In general, increasing the percentage of expensive monitors improves the system performance but we again observe some initially counter-intuitive behavior in systems based on small numbers of monitors. For the individual performance values considered, Table 6 reveals a reduction in the system-level probability of detection as we replace low-performance monitors with high-performance monitors in systems of eight monitors, but these reductions are offset by reduced system-level probabilities of false alarms.
The results in Tables 5 and 6 reveal interesting patterns but fall short of a complete comparison of designs, as they do not incorporate the full design costs based on the different monitor types and associated labor costs. In the case of the Logan Airport study, the two types of monitors are very different, not only in per-unit cost and performance, but also in implementation and associated labor. For instance, the low-cost monitors cost approximately ten dollars apiece. The Logan Airport study involved 28 pairs of monitors requiring preparation time of approximately 45 minutes, field placement time of approximately four hours, and laboratory analysis time of approximately three hours (total). The continuous time monitor involves initial equipment costs of approximately $10,000 but also annual maintenance costs of approximately $30,000 per year including electricity, gases, operator time, and a trailer for storage.

It is important to note that in this example the two types of monitor are not simply exchangeable units differing only in equipment and labor costs, they also differ substantially in the amount and type of information they provide. The low-cost monitors must be placed in the field for a period of several days and provide cumulative exposure values, while the high-cost monitor provides minute-by-minute information on concentrations of NO\textsubscript{2} and other related pollutants. As a result, simply including the per-unit monitor cost into the comparisons above does not provide a comprehensive design comparison across all factors of interest in air pollution monitoring. However, the results in Tables 5 and 6, coupled with the associated costs, do provide important design information regarding the performance of a monitoring network focused on the particular purpose of detecting a specified increase in local air pollution levels. Such results could be used to quantify the value of the relatively inexpensive addition of a collection of low-cost monitors to an existing network of continuous monitoring stations in terms of increased detection performance for particular local concentrations of interest in health studies or ongoing monitoring of pollution levels.

6 Discussion

In a decision-making process, each alternative chosen leads to a possible consequence. We use minimum cost combined with Bayesian optimal decision criteria as a guide to evaluate designs of detection systems in order to identify the system yielding lowest expected decision cost as the
optimal system. Given $P(H_1)$ and $C_{r,s}$ known, the cost of each system is a function of $P_D$ and $P_F$.

In the sections above, we present a unified approach for calculating system-wide $P_D$ and $P_F$ for systems of homogeneous, heterogeneous, and correlated sensors.

In considering the impact of sample size and correlation of individual sensors (tests) on the system-level values of $P_D$ and $P_F$, we assumed that $(C_{10} - C_{00}) = (C_{01} - C_{11})$. Given the performance of individual sensors, our results quantify improvements in the combined probability of detection and false alarm rate with increasing numbers of individual sensors. These results confirm that a phenomenon, difficult to discriminate when using a single test, often may be more reliably differentiated with a multiple test system. In general, the larger the number of individual tests, the more information we gain, and the better performance of the system. However, the results above also reveal some initially counter-intuitive behavior for systems comprised of small numbers of sensors and provide an approach for calculation for detailed comparisons in such cases.

These results hold when the individual probability of detection is greater than its false alarm rate and the conditional independence assumption is satisfied. If the assumption $P_{D_i} > P_{F_i}$ is violated, the combined performance through the likelihood ratio is likely to be worse than that of some of the individual tests. This can be easily seen through equation (4). When $P_{D_i} < P_{F_i}$, increasing the number of negative decisions, $u_i = 0$, combining tests increases the likelihood ratio $\Lambda$, yielding worse combined performance than that of an individual test.

When the conditional independence assumption is violated, the performance of the detection system declines over that based on the same number of independent individual tests. Given the same increment of correlation between individual tests, the magnitude of decline increases with the number of individual tests in a manner defined by equation (8). These results suggest that when individual tests are correlated, the information gain through combining the individual tests is not as much as in the conditional independence scenario so improvement of combined performance declines. Therefore, caution is needed in designing the system to avoid the violation of conditional independence. In the situation where correlation among detectors cannot be avoided, (i.e., repeated diagnostic tests on the same subjects, multi-sensor sub-systems within a machine, individual tests geographically close to each other), a larger number of individual tests is needed to compensate for
loss in system performance due to correlations.

As a brief illustration, assume that the correlations between individual tests are the same in our simulation study and in our theoretical derivation of combined probability of detection and false alarm rate. Often in space or time related scenarios, the closer the individual tests in space or time are, the higher the correlations, e.g., consider

$$
\Sigma = \begin{pmatrix}
\sigma^2 & \rho^1 & \rho^2 & \ldots & \rho^{(n-1)} \\
\rho^1 & \sigma^2 & \rho^1 & \ldots & \rho^{(n-2)} \\
\rho^2 & \rho^1 & \sigma^2 & \ldots & \rho^{(n-3)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\rho^{(n-1)} & \rho^{(n-2)} & \rho^{(n-3)} & \ldots & \sigma^2 
\end{pmatrix}.
$$

The system performance in this scenario is better than the scenario assuming the same correlations between individual tests due to reduced correlations between distant tests (results not shown).

For simplicity of presentation, we assumed \( P(H_1) = P(H_0) = 0.5 \) and that the decision costs of false alarms and false negative responses are the same in considering the effects of system size and correlations of individual sensors on the system performance. In general, our results hold even when \( P(H_1) \neq P(H_0) \) and the costs of false alarms and false positives are different. This can be easily seen from equation (4) since the values of \( P(H_1) \) and \( C_{r,s} \) do not affect the likelihood ratio.

Bayesian optimal decision criteria with minimum cost (3) require knowledge of \( P(H_1) \), which often is not available. However, as in most Bayesian analyses, the impact of the prior \( P(H_1) \) on system performance diminishes with an increasing number of individual tests. This can be seen by rewriting the log likelihood ratio as a weighted sum of the reliability of independent sensors and of \( \log \frac{P_1}{P_0} \) (Chair and Varshney 1986, Thomopoulos et al. 1987, 1989, Viswanathan et al. 1988), i.e.,

$$
U_f = \begin{cases} 
1, & \text{if } \log \Lambda = \log \frac{P_1}{P_0} + \sum_{u_+} \log \frac{P_{D_1}}{P_{F_1}} + \sum_{u_-} \log \frac{P_{F_1}}{P_{D_1}} > \log (C_{10} - C_{00})/(C_{01} - C_{11}) \\
0, & \text{otherwise.}
\end{cases}
$$

With a two-sensor system, \( \log \frac{P_1}{P_0} \) contributes 1/3 of the magnitude of the log likelihood ratio but for an eight-sensor system \( \log \frac{P_1}{P_0} \) contributes only 1/9 of the magnitude. Also, from equation (10), we can see that changing the probability of detection or false alarm rate for any particular test changes the weight of that test in the fusion decision rule. This explains why combining the probability of

18
detection and the false alarm rate through the likelihood ratio test improves performance regardless of the individual test performance.

In our hypothetical data example, we considered designing detection systems monitoring whether ambient nitrogen dioxide concentrations exceeded a predefined level based on two types of sensors. The illustration is based on basic assumptions to highlight features of our derivations in a realistic design setting. However, the example is primarily illustrative as we did not incorporate temporal patterns or durations of concentrations, the pollutant or combinations with other pollutants in the air. Since long duration of the pollutant or/and combination effects of other pollutants might also influence the ambient pollutant level and experienced exposure levels to particular pollutants, more studies are needed to examine these effects in a more realistic setting.

Acknowledgements

This work was done as part of the first author’s dissertation in the Department of Biostatistics, Rollins School of Public Health, Emory University. The authors thank John Williamson and Lily Zhang for many helpful comments. The work of the second author was partially supported by NIEHS grant R01-ES007750, however, the opinions expressed here are those of the authors and do not necessarily represent those of NIH or NIEHS.

References


Table 1: Sample Size Calculation of Independent Subunit Sensors

<table>
<thead>
<tr>
<th>$P_{Di}$</th>
<th>$P_{Fi}$</th>
<th>Number of Sensor Detectors</th>
<th>Number of Positive Sensors Needed for Alarm</th>
<th>Combined Probability of Detection</th>
<th>Combined False Alarm Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.691</td>
<td>0.309</td>
<td>2</td>
<td>1</td>
<td>0.905</td>
<td>0.523</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.773</td>
<td>0.227</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>0.922</td>
<td>0.273</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>0.862</td>
<td>0.138</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>0.935</td>
<td>0.211</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>0.890</td>
<td>0.110</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>0.946</td>
<td>0.166</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>0.979</td>
<td>0.058</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>10</td>
<td>8</td>
<td>0.996</td>
<td>0.009</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>20</td>
<td>10</td>
<td>0.999</td>
<td>0.002</td>
</tr>
<tr>
<td>40</td>
<td>30</td>
<td>40</td>
<td>30</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td>60</td>
<td>50</td>
<td>60</td>
<td>50</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td>80</td>
<td>100</td>
<td>80</td>
<td>100</td>
<td>1.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

$P_{Di}$: probability of detection of subunit sensors.

$P_{Fi}$: false alarm rate of individual sensors.
Table 2: Combined Probabilities of Detection and False Alarm Rates of Correlated Individual Sensors

<table>
<thead>
<tr>
<th>Number of Sensor Detectors</th>
<th>Number of Positive Sensors Needed for Alarm</th>
<th>Combined Probability of Detection $\rho = 0$</th>
<th>$\rho = 0.2$</th>
<th>$\rho = 0.4$</th>
<th>$\rho = 0.6$</th>
<th>Combined False Alarm Rates $\rho = 0$</th>
<th>$\rho = 0.2$</th>
<th>$\rho = 0.4$</th>
<th>$\rho = 0.6$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$P_{D_1}$ = 0.691 and $P_{F_1}$ = 0.309</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>0.909</td>
<td>0.846</td>
<td>0.824</td>
<td>0.787</td>
<td>0.364</td>
<td>0.367</td>
<td>0.364</td>
<td>0.358</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>0.935</td>
<td>0.858</td>
<td>0.807</td>
<td>0.766</td>
<td>0.211</td>
<td>0.269</td>
<td>0.297</td>
<td>0.311</td>
</tr>
<tr>
<td>16</td>
<td>8</td>
<td>0.960</td>
<td>0.863</td>
<td>0.799</td>
<td>0.751</td>
<td>0.085</td>
<td>0.206</td>
<td>0.256</td>
<td>0.285</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>0.979</td>
<td>0.859</td>
<td>0.791</td>
<td>0.743</td>
<td>0.058</td>
<td>0.194</td>
<td>0.250</td>
<td>0.281</td>
</tr>
</tbody>
</table>

$P_{D_i}$: probability of detection of individual sensors.
$P_{F_i}$: false alarm rate of individual sensors.
$\rho$: correlations between individual sensors.

Table 3: Combined Probabilities of Detection and False Alarm Rates from Two and Eight Individual Sensors

<table>
<thead>
<tr>
<th>$P_{D_1}$</th>
<th>$P_{D_2}$</th>
<th>$P_{D(2)}$</th>
<th>$P_{D(8)}$</th>
<th>$P_{F_1}$</th>
<th>$P_{F_2}$</th>
<th>$P_{F(2)}$</th>
<th>$P_{F(8)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(H_1) = 0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.864</td>
<td>0.982</td>
<td>0.851</td>
<td>0.957</td>
<td>0.540</td>
<td>0.540</td>
<td>0.292</td>
<td>0.087</td>
</tr>
<tr>
<td>0.788</td>
<td>0.964</td>
<td>0.761</td>
<td>0.964</td>
<td>0.421</td>
<td>0.421</td>
<td>0.177</td>
<td>0.062</td>
</tr>
<tr>
<td>0.691</td>
<td>0.933</td>
<td>0.933</td>
<td>0.961</td>
<td>0.309</td>
<td>0.309</td>
<td>0.309</td>
<td>0.041</td>
</tr>
<tr>
<td>0.579</td>
<td>0.885</td>
<td>0.886</td>
<td>0.950</td>
<td>0.212</td>
<td>0.212</td>
<td>0.212</td>
<td>0.025</td>
</tr>
<tr>
<td>0.460</td>
<td>0.816</td>
<td>0.815</td>
<td>0.970</td>
<td>0.136</td>
<td>0.136</td>
<td>0.136</td>
<td>0.048</td>
</tr>
<tr>
<td>0.345</td>
<td>0.726</td>
<td>0.819</td>
<td>0.960</td>
<td>0.081</td>
<td>0.081</td>
<td>0.155</td>
<td>0.043</td>
</tr>
<tr>
<td>0.242</td>
<td>0.618</td>
<td>0.709</td>
<td>0.935</td>
<td>0.045</td>
<td>0.045</td>
<td>0.087</td>
<td>0.037</td>
</tr>
</tbody>
</table>

$P_{D_1}$: probability of detection for type I individual sensor.
$P_{F_1}$: false alarm rate for the type I individual sensor.
$P_{D_2}$: probability of detection for type II individual sensor.
$P_{F_2}$: false alarm rate for the type II individual sensor.
$P_{D(2)}$: probability of detection for system with two individual sensors.
$P_{F(2)}$: false alarm rate for a system with two individual sensors.
$P_{D(8)}$: probability of detection for system with eight individual sensors.
$P_{F(8)}$: false alarm rate for a system with eight individual sensors.
Table 4: Combined Probabilities of Detection and False Alarm Rates of Two Different Type of Sensors with Various Correlations (compound symmetric)

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$P_{D1}$</th>
<th>$P_{F1}$</th>
<th>$P_{D2}$</th>
<th>$P_{F2}$</th>
<th>$P_D$</th>
<th>$P_F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 identical sensors from each type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.691</td>
<td>0.309</td>
<td>0.788</td>
<td>0.420</td>
<td>0.931</td>
<td>0.136</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.915</td>
<td>0.175</td>
</tr>
<tr>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.905</td>
<td>0.208</td>
</tr>
<tr>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.900</td>
<td>0.238</td>
</tr>
<tr>
<td>4 identical sensors from each type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.691</td>
<td>0.309</td>
<td>0.788</td>
<td>0.420</td>
<td>0.965</td>
<td>0.062</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.934</td>
<td>0.140</td>
</tr>
<tr>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.912</td>
<td>0.203</td>
</tr>
<tr>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.897</td>
<td>0.250</td>
</tr>
<tr>
<td>8 identical sensors from each type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.691</td>
<td>0.309</td>
<td>0.788</td>
<td>0.420</td>
<td>0.996</td>
<td>0.004</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.954</td>
<td>0.082</td>
</tr>
<tr>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.915</td>
<td>0.147</td>
</tr>
<tr>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.888</td>
<td>0.199</td>
</tr>
</tbody>
</table>

$\rho$: correlations between individual sensors.
Table 5: Performance of a Detection System with Multiple, Homogenous Monitors

<table>
<thead>
<tr>
<th>Number of Monitors</th>
<th>Number of Positive Results</th>
<th>(P_D)</th>
<th>(P_F)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Needed for Alarm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>0.714</td>
<td>0.446</td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>0.708</td>
<td>0.381</td>
</tr>
<tr>
<td>40</td>
<td>20</td>
<td>0.735</td>
<td>0.183</td>
</tr>
<tr>
<td>50</td>
<td>24</td>
<td>0.829</td>
<td>0.223</td>
</tr>
</tbody>
</table>

\(P_D = 0.537, P_F = 0.417\)

| 8                  | 5                           | 0.665  | 0.201  |
| 12                 | 7                           | 0.744  | 0.190  |
| 40                 | 21                          | 0.938  | 0.111  |
| 50                 | 27                          | 0.929  | 0.053  |

\(P_D = 0.631, P_F = 0.417\)

| 8                  | 5                           | 0.810  | 0.201  |
| 12                 | 7                           | 0.885  | 0.190  |
| 40                 | 23                          | 0.970  | 0.032  |
| 50                 | 29                          | 0.977  | 0.015  |

\(P_D = 0.702, P_F = 0.417\)

| 8                  | 5                           | 0.893  | 0.201  |
| 12                 | 8                           | 0.853  | 0.073  |
| 40                 | 24                          | 0.991  | 0.015  |
| 50                 | 30                          | 0.995  | 0.007  |

\(P_D = 0.755, P_F = 0.417\)

\(P_{Di}\): probability of detection of individual test.
\(P_{Fi}\): false alarm rate of individual test.
\(P_D\): combined probability of detection of a diagnostic system.
\(P_F\): combined false alarm rate of a diagnostic system.
<table>
<thead>
<tr>
<th>Number</th>
<th>$P_{D1}$</th>
<th>$P_{F1}$</th>
<th>$P_{D2}$</th>
<th>$P_{F2}$</th>
<th>$P_D$</th>
<th>$P_F$</th>
<th>$P_D$</th>
<th>$P_F$</th>
<th>$P_D$</th>
<th>$P_F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0.537</td>
<td>0.417</td>
<td>0.566</td>
<td>0.329</td>
<td>0.714</td>
<td>0.446</td>
<td>0.700</td>
<td>0.342</td>
<td>0.645</td>
<td>0.270</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.709</td>
<td>0.334</td>
<td>0.712</td>
<td>0.230</td>
<td>0.813</td>
<td>0.260</td>
</tr>
<tr>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.726</td>
<td>0.218</td>
<td>0.822</td>
<td>0.184</td>
<td>0.878</td>
<td>0.157</td>
</tr>
<tr>
<td>64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.834</td>
<td>0.167</td>
<td>0.887</td>
<td>0.087</td>
<td>0.942</td>
<td>0.067</td>
</tr>
<tr>
<td>8</td>
<td>0.631</td>
<td>0.417</td>
<td>0.672</td>
<td>0.329</td>
<td>0.665</td>
<td>0.201</td>
<td>0.738</td>
<td>0.198</td>
<td>0.701</td>
<td>0.164</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.798</td>
<td>0.177</td>
<td>0.855</td>
<td>0.153</td>
<td>0.867</td>
<td>0.115</td>
</tr>
<tr>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.910</td>
<td>0.129</td>
<td>0.933</td>
<td>0.076</td>
<td>0.950</td>
<td>0.046</td>
</tr>
<tr>
<td>64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.961</td>
<td>0.013</td>
<td>0.981</td>
<td>0.020</td>
<td>0.991</td>
<td>0.009</td>
</tr>
<tr>
<td>8</td>
<td>0.702</td>
<td>0.417</td>
<td>0.747</td>
<td>0.329</td>
<td>0.810</td>
<td>0.201</td>
<td>0.810</td>
<td>0.145</td>
<td>0.737</td>
<td>0.120</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.829</td>
<td>0.077</td>
<td>0.895</td>
<td>0.073</td>
<td>0.907</td>
<td>0.056</td>
</tr>
<tr>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.934</td>
<td>0.033</td>
<td>0.968</td>
<td>0.024</td>
<td>0.980</td>
<td>0.015</td>
</tr>
<tr>
<td>64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.987</td>
<td>0.007</td>
<td>0.996</td>
<td>0.003</td>
<td>0.999</td>
<td>0.001</td>
</tr>
<tr>
<td>8</td>
<td>0.755</td>
<td>0.417</td>
<td>0.809</td>
<td>0.329</td>
<td>0.893</td>
<td>0.201</td>
<td>0.897</td>
<td>0.145</td>
<td>0.788</td>
<td>0.120</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.928</td>
<td>0.077</td>
<td>0.943</td>
<td>0.048</td>
<td>0.927</td>
<td>0.037</td>
</tr>
<tr>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.987</td>
<td>0.013</td>
<td>0.989</td>
<td>0.011</td>
<td>0.993</td>
<td>0.007</td>
</tr>
<tr>
<td>64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.999</td>
<td>0.003</td>
<td>0.999</td>
<td>0.001</td>
<td>0.999</td>
<td>0.000</td>
</tr>
</tbody>
</table>

$P_{D1}$: probability of detection of low-performance monitor.

$P_{F1}$: false alarm rate of low-performance monitor.

$P_{D2}$: probability of detection of high-performance monitor.

$P_{F2}$: false alarm rate of high-performance monitor.

$P_D$: combined probability of detection of a diagnostic system.

$P_F$: combined false alarm rate of a diagnostic system.
Figure 1: A schematic representing a system of $n$ parallel sensors, each testing for the presence of the same phenomenon of interest.
Figure 2: System-level performance measures for systems comprised of increasing numbers of homogeneous monitors. Open circles represent the system-level probability of detection, and filled circles denote the system-level probability of false alarm.