LATENT CLASS ANALYSIS: DEPENDENT MISCLASSIFICATION ERRORS

LATENT CLASS ANALYSIS OF DIAGNOSTIC TESTS:

THE EFFECT OF DEPENDENT MISCLASSIFICATION ERRORS

By

VIRGINIA L. TORRANCE, B.Sc.

A Project

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

MASTER OF SCIENCE (1994) McMASTER UNIVERSITY

Mathematics and Statistics Hamilton, Ontario

TITLE: Latent Class Analysis of Diagnostic Tests: The Effect of Dependent Misclassification Errors

AUTHOR: Virginia L. Torrance, B.Sc. (McMaster University)

SUPERVISOR: Dr. Stephen D. Walter

NUMBER OF PAGES: v, 202

ABSTRACT

Latent class modelling is one method used in the evaluation of diagnostic tests when there is no gold standard test that is perfectly accurate. The technique determines maximum likelihood estimates of the prevalence of a disease or a condition and the error rates of diagnostic tests or observers. This study reports the effect of departures from the latent class model assumption of independent misclassifications between observers or tests conditional on the true state of the individual being tested. It is found that estimates become biased in the presence of dependence. Most commonly the prevalence of the disease is overestimated when the true prevalence is at less than 50% and the error rates of dependent observers are underestimated. If there are also independent observers in the group, their error rates are overestimated. The most dangerous scenario in which to use latent class methods in the evaluation of tests is when the true prevalence is low and the false positive rate is high. This is common to many screening situations.

ACKNOWLEDGEMENTS

Most importantly, I would like to thank my supervisor Dr. Stephen Walter for his help, advice, and enthusiasm for this work. Dr. Walter was always available for discussion of ideas and problems. His encouragement of my independence in developing this project created an opportunity to gain practical experience in independent research. In addition, I greatly appreciate the time spent reading and the useful suggestions made by the reviewers of this project, Dr. N. Balakrishnan and Dr. H. Shannon.

I would also like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for their support of my graduate studies in statistics.

Mathew Rynard served as an extremely helpful resource for the computer programming used in this project. His valuable advice and quick problem solving helped speed along the programming process.

Finally, I would like to thank my family for the examples set by their academic successes and for their loving support.

TABLE OF CONTENTS

1. INTRODUCTION

In medicine and epidemiology it is common to classify individuals according to some characteristic of interest. Often these classifications are made into two categories such as, diseased or not diseased, symptomatic or not symptomatic, or exposed to some risk factor or not exposed. Classifications can also be made into more than two categories or levels. For example, an individual's disease status could be classified as none, mild, moderate, or severe.

Classifications can be done through a variety of tests or observations. Often the methods of classification are not perfect. In some instances new tests are compared to a gold standard test that is considered error free. This provides a method of determining the error rates in the new test. If the gold standard test is truly error free then this is the ideal way of calculating test error rates.

When there is no gold standard, error rates in classification methods are still of interest but cannot be determined so easily. One common approach is to compare tests or observers to something that is not truly a gold standard. This approach leads to biased results [1]. For example, determining the error rates in diagnosis of junior clinicians by comparing them to a senior clinician will generally lead to overestimation of the error rates if the senior clinician is not always 100% accurate.

Latent class modelling techniques have been developed for dichotomous classifications without assuming a gold standard [1]. They provide maximum likelihood estimates of the error rates for each test or observer and the prevalence of the characteristic of interest. These models are called latent class models because the true classification of each individual remains unknown. Latent class methods have been used in a large variety of situations. For example, the estimation of the error rates for three colorectal cancer screening tests [2] and the evaluation of the accuracy of tests for tuberculosis [3] have made use of these methods.

Walter and Irwig [1] provide an extensive review of data analysis using latent class models. A variety of situations are considered, each characterized by the number of populations or subgroups in the data and the number of observations made on each individual. When there is only one population, it is necessary to impose constraints on the data if there are less than three observers.

As with most statistical models, latent class models rely on some assumptions. One assumption is that misclassifications are independent between individuals or subjects. A second more dubious assumption is that misclassifications are independent between observers conditional on the true state of the individual. This second assumption may not hold in many situations. Consider a classification that dichotomizes a continuous characteristic. Individuals close to the cut off are likely to be misclassified in the same way by clinicians, especially if the clinicians received similar training. As this example shows, the association of errors between observers is likely to be positive, that is, observers or tests tend to make the same errors when classifying the same individuals.

Walter and Irwig discuss some methods that have been developed to handle conditional dependence of errors between observers. However, they point out that little work has been done on the effect this dependence has on the resulting estimates if uncontrolled. Vacek [4] has done some analytic work in the case of two populations and two observers. The estimators for this situation were analytically derived by Hui and Walter [3]. Vacek showed that the error rates can be underestimated if the errors are conditionally dependent and that the prevalence can be biased in either direction.

This paper extends the previous work by reporting the effect that a violation of the assumption of independent observer errors has on estimates obtained from latent class modelling. Only the more likely case of violation by a positive association is considered. This project is also restricted to situations involving one population with three or four observers, where all observers observe all individuals. Vacek's method of parameterization of the dependence of errors between observers is adapted. Computer simulations are used to determine the results. The methods used in this project are described in Chapter 2. Chapter 3 contains the results. The conclusions, applications, and opportunities for further research are discussed in Chapter 4.

2. METHODS

2.1 Terminology and Notation

For the remainder of this report the characteristic of interest will be referred to as disease, and the methods of classification will be referred to as multiple observers. It should be noted that this is done only to maintain simple and consistent terminology and should not be considered to limit the applications of the methods in any way. For example, disease could be replaced by symptom, exposure, or condition and multiple observers could be replaced by multiple screening tests or multiple diagnostic tests. The notation developed here will closely follow that presented by Walter and Irwig [1].

The true prevalence of the disease will be represented by θ. The false negative rate for the i^{th} observer will be β_i. This is the probability that the i^{th} observer classifies an individual who is truly positive for the disease as not having the disease. The probability that the i^{th} observer classifies an individual who truly does not have the disease as having the disease is the false positive rate and will be given by α_i.

Other common terms used when assessing the accuracy of observers are sensitivity and specificity. Sensitivity is the probability of correctly classifying a truly positive individual and specificity is the probability of correctly classifying a truly negative individual. Thus, the sensitivity and specificity of the i^{th} observer are the complements of the false negative rate and the false positive rate and are given by $1-\beta_i$ and $1-\alpha_i$, respectively.

If there are r observers classifying each individual then there are 2^r possible combinations of classifications for each individual. These combinations of classifications will be referred to as outcome categories. For example, when there are three observers there are $2^3 = 8$ outcome categories which can be represented by: −−−, −−+, −+−, −++, +−−, +−+, ++−, +++. The first outcome category is classification as negative for the disease by all three observers. The second category is classification as negative by the first and second observer and as positive by the third observer. All other categories can be interpreted similarly.

More formally, the classification by the i^{th} observer will be represented by the random variable X_i where

$$X_i = 0 \text{ if the classification is negative,}$$
$$= 1 \text{ if the classification is positive.}$$

The true disease status of the individual will be represented by Y where

$$Y = 0 \text{ if the individual is negative,}$$
$$= 1 \text{ if the individual is positive.}$$

The number of outcome categories determines the number of statistical degrees of freedom, *df*, available for estimation. For a fixed number of individuals there are 2^r-1 *df*. Thus in the three observer case there are $8-1=7$ *df* and for four observers there are $16-1=15$ *df*. When there are three observers, there are seven parameters to estimate: prevalence, three false negative rates, and three false positive rates. When there are fewer than three observers, there are more parameters than degrees of freedom. It is then necessary to impose constraints on the data to permit estimation of the parameters. When there are greater than three observers, there are more degrees of freedom than parameters, leaving all parameters estimable with excess degrees of freedom available for goodness of fit testing. In the four observer case there are nine parameters to estimate: prevalence, four false negative rates, and four false positive rates. Therefore, there are 15−9=6 *df* available for goodness of fit testing.

2.2 Probabilities and Likelihoods

Probabilities and likelihoods will be shown for the three observer case. Extensions to more observers are straightforward. The probability of an individual being in a particular outcome category can be calculated conditional on the true disease status of the individual. For example, the probability that an individual who does not have the disease is classified as negative by all three observers is the product of the observers' specificities,

$$\Pr(---\mid -) = \Pr(X_1=0, X_2=0, X_3=0 \mid Y=0)$$

$$= (1-\alpha_1)(1-\alpha_2)(1-\alpha_3).$$

This formula holds under the assumption that errors in classification are independent between observers. There are 2 x 8 = 16 such independent conditional probabilities in the three observer case.

By the total law of probability [6], the unconditional probability of an individual being classified as negative by all three observers is

$$\begin{aligned}\Pr(---) &= \Pr(\text{individual is positive})\Pr(---\mid +) + \Pr(\text{individual is negative})\Pr(---\mid -)\\ &= \Pr(Y=1)\Pr(X_1=0,X_2=0,X_3=0\mid Y=1) + \Pr(Y=0)\Pr(X_1=0,X_2=0,X_3=0\mid Y=0)\\ &= \theta\beta_1\beta_2\beta_3 + (1-\theta)(1-\alpha_1)(1-\alpha_2)(1-\alpha_3).\end{aligned}$$

There are eight such unconditional probabilities in the three observer case as shown below.

Outcome Category	Probability
−−−	$\theta\beta_1\beta_2\beta_3 + (1-\theta)(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)$
−−+	$\theta\beta_1\beta_2(1-\beta_3) + (1-\theta)(1-\alpha_1)(1-\alpha_2)\alpha_3$
−+−	$\theta\beta_1(1-\beta_2)\beta_3 + (1-\theta)(1-\alpha_1)\alpha_2(1-\alpha_3)$
−++	$\theta\beta_1(1-\beta_2)(1-\beta_3) + (1-\theta)(1-\alpha_1)\alpha_2\alpha_3$
+−−	$\theta(1-\beta_1)\beta_2\beta_3 + (1-\theta)\alpha_1(1-\alpha_2)(1-\alpha_3)$
+−+	$\theta(1-\beta_1)\beta_2(1-\beta_3) + (1-\theta)\alpha_1(1-\alpha_2)\alpha_3$
++−	$\theta(1-\beta_1)(1-\beta_2)\beta_3 + (1-\theta)\alpha_1\alpha_2(1-\alpha_3)$
+++	$\theta(1-\beta_1)(1-\beta_2)(1-\beta_3) + (1-\theta)\alpha_1\alpha_2\alpha_3$

Thus, the probability of any individual being classified as $(X_1=x_1, X_2=x_2, X_3=x_3)$ is

$$Pr(x_1, x_2, x_3) = \theta \prod_{i=1}^{3} \beta_i^{1-x_i} (1-\beta_i)^{x_i} + (1-\theta) \prod_{i=1}^{3} \alpha_i^{x_i} (1-\alpha_i)^{1-x_i} .$$

In general, for r observers and N individuals where $N(\mathbf{x})$ are classified as the vector $\mathbf{x}=(x_1, x_2, ..., x_r)$ such that $\sum_{\mathbf{x}} N(\mathbf{x}) = N$, the joint probability mass function when interpreted as a function of the parameters, θ, α_1, α_2,..., α_r, β_1, β_2,..., β_r becomes the likelihood function,

$$L = \prod_{\mathbf{x}} \left[\theta \prod_{i=1}^{r} \beta_i^{1-x_i} (1-\beta_i)^{x_i} + (1-\theta) \prod_{i=1}^{r} \alpha_i^{x_i} (1-\alpha_i)^{1-x_i} \right]^{N(\mathbf{x})} .$$

Taking the logarithm of this equation for a given data set gives the log likelihood which must be numerically maximized to obtain maximum likelihood estimates of the parameters. Recall that all parameters are not estimable for $r < 3$.

A computer program created by S. Walter called Latent [7] performs the numerical maximization for three to five observers. The program allows the application of these methods to appropriate data sets in order to obtain prevalence and error rate estimates. The program also produces large sample theory estimates of standard errors for the parameter estimates by using the inverse of the expected Fisher information matrix to estimate the variance-covariance matrix for the parameters.

2.3 Development of dependence parameters

The conditional dependence of errors in classification between observers will be quantified as positive terms to be added to or subtracted from the independent conditional probabilities. This follows the approach used by Vacek [2] where the probability that two observers classify an individual as positive when the individual is truly positive is equal to the product of the observers' sensitivities plus the dependence term for a truly positive individual, δ. That is,

$$\Pr(++|+) = (1-\beta_1)(1-\beta_2) + \delta.$$

The probability of the two observers disagreeing is reduced by the dependence. For example, the probability of observer 1 classifying a positive individual correctly while observer 2 misclassifies the individual is

$$\Pr(+-|+) = (1-\beta_1)\beta_2 - \delta.$$

Since the dependence parameters represent conditional dependence there must be separate terms for true positive individuals and true negative individuals. Consider only pairwise or two-way dependence terms The dependence between observers i and j will be represented by δ_{ij} for truly positive individuals and by ε_{ij} for truly negative individuals. As mentioned earlier, only positive association between observers will be considered, therefore, $\delta_{ij} \geq 0$ and $\varepsilon_{ij} \geq 0$, for all $i=1,2,...r$ and $j=1,2,...r$, $i \neq j$. Thus, the independent unconditional probabilities shown in the last section for three observers can be rewritten to incorporate dependence between pairs of observers as shown below.

Outcome Category	Probability
$---$	$\theta[\beta_1\beta_2\beta_3+\delta_{12}+\delta_{13}+\delta_{23}] + (1-\theta)[(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)+\varepsilon_{12}+\varepsilon_{13}+\varepsilon_{23}]$
$--+$	$\theta[\beta_1\beta_2(1-\beta_3)+\delta_{12}-\delta_{13}-\delta_{23}] + (1-\theta)[(1-\alpha_1)(1-\alpha_2)\alpha_3+\varepsilon_{12}-\varepsilon_{13}-\varepsilon_{23}]$
$-+-$	$\theta[\beta_1(1-\beta_2)\beta_3-\delta_{12}+\delta_{13}-\delta_{23}] + (1-\theta)[(1-\alpha_1)\alpha_2(1-\alpha_3)-\varepsilon_{12}+\varepsilon_{13}-\varepsilon_{23}]$
$-++$	$\theta[\beta_1(1-\beta_2)(1-\beta_3)-\delta_{12}-\delta_{13}+\delta_{23}] + (1-\theta)[(1-\alpha_1)\alpha_2\alpha_3-\varepsilon_{12}-\varepsilon_{13}+\varepsilon_{23}]$
$+--$	$\theta[(1-\beta_1)\beta_2\beta_3-\delta_{12}-\delta_{13}+\delta_{23}] + (1-\theta)[\alpha_1(1-\alpha_2)(1-\alpha_3)-\varepsilon_{12}-\varepsilon_{13}+\varepsilon_{23}]$
$+-+$	$\theta[(1-\beta_1)\beta_2(1-\beta_3)-\delta_{12}+\delta_{13}-\delta_{23}] + (1-\theta)[\alpha_1(1-\alpha_2)\alpha_3-\varepsilon_{12}+\varepsilon_{13}-\varepsilon_{23}]$
$++-$	$\theta[(1-\beta_1)(1-\beta_2)\beta_3+\delta_{12}-\delta_{13}-\delta_{23}] + (1-\theta)[\alpha_1\alpha_2(1-\alpha_3)+\varepsilon_{12}-\varepsilon_{13}-\varepsilon_{23}]$
$+++$	$\theta[(1-\beta_1)(1-\beta_2)(1-\beta_3)+\delta_{12}+\delta_{13}+\delta_{23}] + (1-\theta)[\alpha_1\alpha_2\alpha_3+\varepsilon_{12}+\varepsilon_{13}+\varepsilon_{23}]$

These dependence terms add to the independent probability of a certain classification when two dependent observers agree and subtract from the probability when two dependent observers disagree. Consider the first outcome category where all three observers agree that the classification is negative. There are three terms added to the conditional probability when the individual is truly positive. The terms are added because observers 1 and 2 agree, observers 1 and 3 agree, and observers 2 and 3 agree giving $\beta_1\beta_2\beta_3+\delta_{12}+\delta_{13}+\delta_{23}$ as the dependent conditional probability. There are also three terms added to the conditional probability of a correct classification by all three observers when the individual is truly negative. Again, this is because all pairs of observers agree. The probability given the individual is truly negative becomes

$$(1-\alpha_1)(1-\alpha_2)(1-\alpha_3)+\varepsilon_{12}+\varepsilon_{13}+\varepsilon_{23}.$$

A more formal interpretation of these parameters will now be discussed. Note six restrictions on the probabilities for a true positive case:

$$\Pr(---|+) + \Pr(--+|+) + \Pr(-+-|+) + \Pr(-++|+) = \beta_1$$

$$\Pr(+--|+) + \Pr(+-+|+) + \Pr(++-|+) + \Pr(+++|+) = 1-\beta_1$$

$$\Pr(---|+) + \Pr(--+|+) + \Pr(+--|+) + \Pr(+-+|+) = \beta_2$$

$$\Pr(-+-|+) + \Pr(-++|+) + \Pr(++-|+) + \Pr(+++|+) = 1-\beta_2$$

$$\Pr(---|+) + \Pr(-+-|+) + \Pr(+--|+) + \Pr(++-|+) = \beta_3$$

$$\Pr(--+|+) + \Pr(-++|+) + \Pr(+-+|+) + \Pr(+++|+) = 1-\beta_3$$

The first restriction is arrived at by summing the probabilities over the possible outcomes for observers 2 and 3 while holding observer 1's outcome constant at negative. This is the probability that observer 1 classifies a truly positive individual as negative, that is, observer 1's false negative rate, β_1. All other restrictions can be explained similarly. Depdendence terms should be developed in such a way that these restrictions remain true in the presence of dependence.

Now recall the random variables X_1, X_2, X_3, and Y that take on the values 0 or 1 to represent an observer's classification and the true classification as negative or positive. When dependence is present,

$$
\begin{aligned}
E(X_1 \mid Y=1) &= (1)\Pr(X_1=1 \mid Y=1) + (0)\Pr(X_1=0 \mid Y=1) \\
&= \Pr(X_1=1 \mid Y=1) \\
&= \Pr(+-- \mid +) + \Pr(+-+ \mid +) + \Pr(++- \mid +) + \Pr(+++ \mid +) \\
&= (1-\beta_1)\beta_2\beta_3-\delta_{12}-\delta_{13}+\delta_{23} + (1-\beta_1)\beta_2(1-\beta_3)-\delta_{12}+\delta_{13}-\delta_{23} \\
&\qquad + (1-\beta_1)(1-\beta_2)\beta_3+\delta_{12}-\delta_{13}-\delta_{23} + (1-\beta_1)(1-\beta_2)(1-\beta_3)+\delta_{12}+\delta_{13}+\delta_{23} \\
&= (1-\beta_1)[\beta_2\beta_3 + \beta_2(1-\beta_3) + (1-\beta_2)\beta_3 + (1-\beta_2)(1-\beta_3)] \\
&= 1-\beta_1
\end{aligned}
$$

This shows that the addition of the dependence terms into the probability equations does not affect the previously mentioned restrictions. All other restrictions can be similarly verified. For example, it can be shown that,

$$E(X_2 \mid Y=1) = 1-\beta_2.$$

Now,

$$
\begin{aligned}
E(X_1X_2 \mid Y=1) &= \Pr(X_1=1,X_2=1 \mid Y=1) \\
&= \Pr(++- \mid +) + \Pr(+++ \mid +) \\
&= (1-\beta_1)(1-\beta_2)\beta_3+\delta_{12}-\delta_{13}-\delta_{23} + (1-\beta_1)(1-\beta_2)(1-\beta_3)+\delta_{12}+\delta_{13}+\delta_{23} \\
&= (1-\beta_1)(1-\beta_2) + 2\delta_{12}.
\end{aligned}
$$

So,

$$
\begin{aligned}
Cov(X_1,X_2 \mid Y=1) &= E(X_1X_2 \mid Y=1) - E(X_1 \mid Y=1)E(X_2 \mid Y=1) \\
&= (1-\beta_1)(1-\beta_2) + 2\delta_{12} - (1-\beta_1)(1-\beta_2) \\
&= 2\delta_{12}.
\end{aligned}
$$

Thus,

$$\delta_{12} = 1/2\ Cov(X_1,X_2 \mid Y=1).$$

The dependence parameter δ_{12} is one half of the conditional covariance between the classifications of observers 1 and 2 given the truth is positive. Similarly, the dependence parameter ε_{12} is one half of the conditional covariance between the classifications of observers 1 and 2 given the truth is negative. The other dependence parameters can also be shown to be equal to one half of their corresponding conditional covariances in the three observer case.

In the four observer case there are eight restrictions. Similar calculations show that $\delta_{12} = 1/4 \, Cov(X_1, X_2 \mid Y=1)$ and that the other pairwise dependence parameters can be shown to equal one quarter of their corresponding conditional covariances.

Although this project concentrates on three and four observers, only the pairwise or two-way dependence terms will be considered. These terms have an intuitive interpretation and can be expressed in terms of covariances. It is difficult to rationalize the existence of any higher order dependence terms in this context. For example, a three-way dependence term would represent how the dependence between two observers is affected by the classification given by a third observer. It is unlikely that such a dependence would occur in a practical situation where all observers are classifying individuals physically removed from and without communications with the other observers.

2.4 Bounds on Dependence Parameter Values

Dependence parameters are defined in terms of probability, therefore, they have an upper bound. All conditional probabilities must be between 0 and 1. Thus, for a given set of parameter values each dependence parameter has a maximum possible value. It is necessary to calculate these maximum values prior to simulation. The simulations will involve cases with dependence between one pair of observers and cases with dependence between all pairs of observers. Details of the simulations will be discussed in the next two sections. The maxima for the dependence parameter values will be determined for four separate cases:

i) Dependence between observers 1 and 2 for three observers,

ii) Equal dependence between all pairs of observers for three observers,

iii) Dependence between observers 1 and 2 for four observers,

iv) Equal dependence between all pairs of observers for four observers.

In general, it is necessary that the sum of the false positive rate and the false negative rate for the i^{th} observer is less than one, that is, $\alpha_i + \beta_i < 1$, for the i^{th} observer to be useful. This work will make the further restriction that each of the error rates is at most .5. Thus, only observers with reasonable error rates will be considered. This bound on the error rates will be used in the following determinations of dependence maxima.

Starting with case i), the dependence is between observers 1 and 2 with respect to false negative rate, thus, we are considering true positive individuals. The restrictions on the conditional probabilities and the resulting restrictions on the dependence parameters follow.

Outcome Category	Restrictions on Conditional Probability	Restrictions on Dependence Parameter
− − −	$\beta_1\beta_2\beta_3 + \delta_{12} \le 1$	$\delta_{12} \le 1 - \beta_1\beta_2\beta_3$
− − +	$\beta_1\beta_2(1-\beta_3) + \delta_{12} \le 1$	$\delta_{12} \le 1 - \beta_1\beta_2(1-\beta_3)$
− + −	$\beta_1(1-\beta_2)\beta_3 - \delta_{12} \ge 0$	$\delta_{12} \le \beta_1(1-\beta_2)\beta_3$
− + +	$\beta_1(1-\beta_2)(1-\beta_3) - \delta_{12} \ge 0$	$\delta_{12} \le \beta_1(1-\beta_2)(1-\beta_3)$
+ − −	$(1-\beta_1)\beta_2\beta_3 - \delta_{12} \ge 0$	$\delta_{12} \le (1-\beta_1)\beta_2\beta_3$
+ − +	$(1-\beta_1)\beta_2(1-\beta_3) - \delta_{12} \ge 0$	$\delta_{12} \le (1-\beta_1)\beta_2(1-\beta_3)$
+ + −	$(1-\beta_1)(1-\beta_2)\beta_3 + \delta_{12} \le 1$	$\delta_{12} \le 1 - (1-\beta_1)(1-\beta_2)\beta_3$
+ + +	$(1-\beta_1)(1-\beta_2)(1-\beta_3) + \delta_{12} \le 1$	$\delta_{12} \le 1 - (1-\beta_1)(1-\beta_2)(1-\beta_3)$

Recalling that all error rates are at most .5 the most strict restrictions can be determined. Note that the first two and the second to last conditions are not as strict as the last. That is,

$$1 - \beta_1\beta_2\beta_3 \ge 1 - (1-\beta_1)(1-\beta_2)(1-\beta_3),$$

$$1 - \beta_1\beta_2(1-\beta_3) \ge 1 - (1-\beta_1)(1-\beta_2)(1-\beta_3), \text{ and}$$

$$1 - (1-\beta_1)(1-\beta_2)\beta_3 \ge 1 - (1-\beta_1)(1-\beta_2)(1-\beta_3).$$

Similarly,

$$\beta_1(1-\beta_2)(1-\beta_3) \ge \beta_1(1-\beta_2)\beta_3 \text{ and}$$

$$(1-\beta_1)\beta_2(1-\beta_3) \ge (1-\beta_1)\beta_2\beta_3.$$

The three restrictions left are

$$\delta_{12} \le \beta_1(1-\beta_2)\beta_3, \qquad (1)$$

$$\delta_{12} \le (1-\beta_1)\beta_2\beta_3, \text{ and} \qquad (2)$$

$$\delta_{12} \le 1 - (1-\beta_1)(1-\beta_2)(1-\beta_3). \qquad (3)$$

It is shown in Appendix A that condition (1) and (2) are more strict than condition (3). Therefore, determining the maximum value for δ_{12} for any error rate and prevalence setting is done by restricting

$$\delta_{12} \leq \beta_1(1-\beta_2)\beta_3 \text{ and } \delta_{12} \leq (1-\beta_1)\beta_2\beta_3.$$

This means that for any given setting of error rate values, the maximum for δ_{12} is determined by the minimum value given by the above two conditions. It can similarly be shown that the maximum value of ε_{12} must be determined by restricting

$$\varepsilon_{12} \leq \alpha_1(1-\alpha_2)\alpha_3 \text{ and } \varepsilon_{12} \leq (1-\alpha_1)\alpha_2\alpha_3.$$

For case ii) there is dependence between all pairs of observers with a total of three observers. The same procedure can be followed to determine the maximum value for the dependence parameters. In these simulations the dependence values will be set equal for all pairs, that is, $\delta_{12}=\delta_{13}=\delta_{23}=\delta$ In addition, all observers' false negative rates will be equal and all observers' false positive rates will be equal. The individual parameters will be retained in the calculations that follow to preserve the general case. Only a special case of dependence between all pairs of observers is being considered in this study. The restrictions on the conditional probabilities and the resulting restrictions on the dependence parameters are shown below. Again we will consider dependence with respect to false negative rates.

Outcome Category	Restrictions on Conditional Probability	Restrictions on Dependence Parameter
$---$	$\beta_1\beta_2\beta_3 + 3\delta \leq 1$	$\delta \leq (1-\beta_1\beta_2\beta_3)/3$
$--+$	$\beta_1\beta_2(1-\beta_3) - \delta \geq 0$	$\delta \leq \beta_1\beta_2(1-\beta_3)$
$-+-$	$\beta_1(1-\beta_2)\beta_3 - \delta \geq 0$	$\delta \leq \beta_1(1-\beta_2)\beta_3$
$-++$	$\beta_1(1-\beta_2)(1-\beta_3) - \delta \geq 0$	$\delta \leq \beta_1(1-\beta_2)(1-\beta_3)$
$+--$	$(1-\beta_1)\beta_2\beta_3 - \delta \geq 0$	$\delta \leq (1-\beta_1)\beta_2\beta_3$
$+-+$	$(1-\beta_1)\beta_2(1-\beta_3) - \delta \geq 0$	$\delta \leq (1-\beta_1)\beta_2(1-\beta_3)$
$++-$	$(1-\beta_1)(1-\beta_2)\beta_3 - \delta \geq 0$	$\delta \leq (1-\beta_1)(1-\beta_2)\beta_3$
$+++$	$(1-\beta_1)(1-\beta_2)(1-\beta_3) + 3\delta \leq 1$	$\delta \leq (1-(1-\beta_1)(1-\beta_2)(1-\beta_3))/3$

It can immediately be seen that the first restriction is not as strict as the last. In addition, the fourth restriction is not as strict as the second or third, and that the sixth and seventh restrictions are not as strict as the fifth. The four remaining restrictions are

$$\delta \leq \beta_1\beta_2(1-\beta_3), \tag{4}$$

$$\delta \leq \beta_1(1-\beta_2)\beta_3, \tag{5}$$

$$\delta \leq (1-\beta_1)\beta_2\beta_3 \text{ and,} \tag{6}$$

$$\delta \leq (1-(1-\beta_1)(1-\beta_2)(1-\beta_3))/3. \tag{7}$$

Appendix A gives the proof that restrictions (4), (5), and (6) are more strict than restriction (7). Therefore, the maximum value of δ is determined by restricting

$$\delta \leq \beta_1\beta_2(1-\beta_3),\ \delta \leq \beta_1(1-\beta_2)\beta_3, \text{ and } \delta \leq (1-\beta_1)\beta_2\beta_3.$$

Similarly, the maximum value of $\varepsilon=\varepsilon_{12}=\varepsilon_{13}=\varepsilon_{23}$ is determined by restricting

$$\varepsilon \leq \alpha_1\alpha_2(1-\alpha_3),\ \varepsilon \leq \alpha_1(1-\alpha_2)\alpha_3, \text{ and } \varepsilon \leq (1-\alpha_1)\alpha_2\alpha_3.$$

The four observer cases will now be discussed. Consider dependence between observers 1 and 2, case iii). The restrictions on the probabilities and the resulting restrictions on the dependence values are given below for the case of dependence with respect to false negative rate.

Outcome Category	Restrictions on Conditional Probability	Restrictions on Dependence Parameter
----	$\beta_1\beta_2\beta_3\beta_4 + \delta_{12} \le 1$	$\delta_{12} \le 1 - \beta_1\beta_2\beta_3\beta_4$
---+	$\beta_1\beta_2\beta_3(1-\beta_4) + \delta_{12} \le 1$	$\delta_{12} \le 1 - \beta_1\beta_2\beta_3(1-\beta_4)$
--+-	$\beta_1\beta_2(1-\beta_3)\beta_4 + \delta_{12} \le 1$	$\delta_{12} \le 1 - \beta_1\beta_2(1-\beta_3)\beta_4$
--++	$\beta_1\beta_2(1-\beta_3)(1-\beta_4) + \delta_{12} \le 1$	$\delta_{12} \le 1 - \beta_1\beta_2(1-\beta_3)(1-\beta_4)$
-+--	$\beta_1(1-\beta_2)\beta_3\beta_4 - \delta_{12} \ge 0$	$\delta_{12} \le \beta_1(1-\beta_2)\beta_3\beta_4$
-+-+	$\beta_1(1-\beta_2)\beta_3(1-\beta_4) - \delta_{12} \ge 0$	$\delta_{12} \le \beta_1(1-\beta_2)\beta_3(1-\beta_4)$
-++-	$\beta_1(1-\beta_2)(1-\beta_3)\beta_4 - \delta_{12} \ge 0$	$\delta_{12} \le \beta_1(1-\beta_2)(1-\beta_3)\beta_4$
-+++	$\beta_1(1-\beta_2)(1-\beta_3)(1-\beta_4) - \delta_{12} \ge 0$	$\delta_{12} \le \beta_1(1-\beta_2)(1-\beta_3)(1-\beta_4)$
+---	$(1-\beta_1)\beta_2\beta_3\beta_4 - \delta_{12} \ge 0$	$\delta_{12} \le (1-\beta_1)\beta_2\beta_3\beta_4$
+--+	$(1-\beta_1)\beta_2\beta_3(1-\beta_4) - \delta_{12} \ge 0$	$\delta_{12} \le (1-\beta_1)\beta_2\beta_3(1-\beta_4)$
+-+-	$(1-\beta_1)\beta_2(1-\beta_3)\beta_4 - \delta_{12} \ge 0$	$\delta_{12} \le (1-\beta_1)\beta_2(1-\beta_3)\beta_4$
+-++	$(1-\beta_1)\beta_2(1-\beta_3)(1-\beta_4) - \delta_{12} \ge 0$	$\delta_{12} \le (1-\beta_1)\beta_2(1-\beta_3)(1-\beta_4)$
++--	$(1-\beta_1)(1-\beta_2)\beta_3\beta_4 + \delta_{12} \le 1$	$\delta_{12} \le 1-(1-\beta_1)(1-\beta_2)\beta_3\beta_4$
++-+	$(1-\beta_1)(1-\beta_2)\beta_3(1-\beta_4) + \delta_{12} \le 1$	$\delta_{12} \le 1-(1-\beta_1)(1-\beta_2)\beta_3(1-\beta_4)$
+++-	$(1-\beta_1)(1-\beta_2)(1-\beta_3)\beta_4 + \delta_{12} \le 1$	$\delta_{12} \le 1-(1-\beta_1)(1-\beta_2)(1-\beta_3)\beta_4$
++++	$(1-\beta_1)(1-\beta_2)(1-\beta_3)(1-\beta_4) + \delta_{12} \le 1$	$\delta_{12} \le 1-(1-\beta_1)(1-\beta_2)(1-\beta_3)(1-\beta_4)$

Of all the dependence parameter restrictions beginning with "1-" the last is the most strict since all error rates are at most 0.5. The sixth, seventh, and eighth restrictions are less strict than the fifth and the tenth, eleventh, and twelfth are less strict than the ninth. Hence, the most strict restrictions are

$$\delta_{12} \le \beta_1(1-\beta_2)\beta_3\beta_4, \tag{8}$$

$$\delta_{12} \le (1-\beta_1)\beta_2\beta_3\beta_4 \tag{9}$$

$$\delta_{12} \le 1-(1-\beta_1)(1-\beta_2)(1-\beta_3)(1-\beta_4) \tag{10}$$

It is shown in Appendix A that conditions (8) and (9) are more strict than condition (10). Thus, the restrictions of

$$\delta_{12} \le \beta_1(1-\beta_2)\beta_3\beta_4 \text{ and } \delta_{12} \le (1-\beta_1)\beta_2\beta_3\beta_4$$

are used to determine the maximum possible value of δ_{12} in the four observer case with dependence between observers 1 and 2 only. It can similarly be shown that the restrictions

$$\varepsilon_{12} \le \alpha_1(1-\alpha_2)\alpha_3\alpha_4 \text{ and } \varepsilon_{12} \le (1-\alpha_1)\alpha_2\alpha_3\alpha_4$$

determine the maximum possible value for ε_{12} with dependence between observers 1 and 2 only with respect to false positive rate.

The last case to be discussed here, case iv), involves four observers with all pairwise dependence terms present. All dependence terms will be set equal, $\delta_{12}=\delta_{13}=\delta_{14}=\delta_{23}=\delta_{24}=\delta_{34}=\delta$. Again, as in the three observer case, all false negative rates are equal and all false positive rates are equal. The probability and dependence term restrictions for this case are given below for the case of dependence with respect to false negative rate.

Outcome Category	Restrictions on Conditional Probability	Restrictions on Dependence Parameter
----	$\beta_1\beta_2\beta_3\beta_4 + 6\delta \le 1$	$\delta \le (1-\beta_1\beta_2\beta_3\beta_4)/6$
---+	$0 \le \beta_1\beta_2\beta_3(1-\beta_4) + 0\delta \le 1$	No restriction
--+-	$0 \le \beta_1\beta_2(1-\beta_3)\beta_4 + 0\delta \le 1$	No restriction
--++	$\beta_1\beta_2(1-\beta_3)(1-\beta_4) - 2\delta \ge 0$	$\delta \le (\beta_1\beta_2(1-\beta_3)(1-\beta_4))/2$
-+--	$0 \le \beta_1(1-\beta_2)\beta_3\beta_4 + 0\delta \le 1$	No restriction
-+-+	$\beta_1(1-\beta_2)\beta_3(1-\beta_4) - 2\delta \ge 0$	$\delta \le (\beta_1(1-\beta_2)\beta_3(1-\beta_4))/2$
-++-	$\beta_1(1-\beta_2)(1-\beta_3)\beta_4 - 2\delta \ge 0$	$\delta \le (\beta_1(1-\beta_2)(1-\beta_3)\beta_4)/2$
-+++	$0 \le \beta_1(1-\beta_2)(1-\beta_3)(1-\beta_4) + 0\delta \le 1$	No restriction
+---	$0 \le (1-\beta_1)\beta_2\beta_3\beta_4 + 0\delta \le 1$	No restriction
+--+	$(1-\beta_1)\beta_2\beta_3(1-\beta_4) - 2\delta \ge 0$	$\delta \le ((1-\beta_1)\beta_2\beta_3(1-\beta_4))/2$
+-+-	$(1-\beta_1)\beta_2(1-\beta_3)\beta_4 - 2\delta \ge 0$	$\delta \le ((1-\beta_1)\beta_2(1-\beta_3)\beta_4)/2$
+-++	$0 \le (1-\beta_1)\beta_2(1-\beta_3)(1-\beta_4) + 0\delta \le 1$	No restriction
++--	$(1-\beta_1)(1-\beta_2)\beta_3\beta_4 - 2\delta \ge 0$	$\delta \le ((1-\beta_1)(1-\beta_2)\beta_3\beta_4)/2$
++-+	$0 \le (1-\beta_1)(1-\beta_2)\beta_3(1-\beta_4) + 0\delta \le 1$	No restriction
+++-	$0 \le (1-\beta_1)(1-\beta_2)(1-\beta_3)\beta_4 + 0\delta \le 1$	No restriction
++++	$(1-\beta_1)(1-\beta_2)(1-\beta_3)(1-\beta_4) + 6\delta \le 1$	$\delta \le (1-(1-\beta_1)(1-\beta_2)(1-\beta_3)(1-\beta_4))/6$

The first restriction is less strict than the last restriction. All of the remaining restricitons are of equivalent form. It is shown in Appendix A that the last restriction is not as strict as these others. Thus, in this case the maximum for δ is determined by restricting

$$\delta \leq \beta_1\beta_2(1-\beta_3)(1-\beta_4)/2,$$

$$\delta \leq \beta_1(1-\beta_2)\beta_3(1-\beta_4)/2,$$

$$\delta \leq \beta_1(1-\beta_2)(1-\beta_3)\beta_4/2,$$

$$\delta \leq (1-\beta_1)\beta_2\beta_3(1-\beta_4)/2,$$

$$\delta \leq (1-\beta_1)\beta_2(1-\beta_3)\beta_4/2,$$

$$\text{and } \delta \leq (1-\beta_1)(1-\beta_2)\beta_3\beta_4/2.$$

It can be shown that the restrictions on ε are

$$\varepsilon \leq \alpha_1\alpha_2(1-\beta_3)(1-\alpha_4)/2,$$

$$\varepsilon \leq \alpha_1(1-\alpha_2)\alpha_3(1-\alpha_4)/2,$$

$$\varepsilon \leq \alpha_1(1-\alpha_2)(1-\alpha_3)\alpha_4/2,$$

$$\varepsilon \leq (1-\alpha_1)\alpha_2\alpha_3(1-\alpha_4)/2,$$

$$\varepsilon \leq (1-\alpha_1)\alpha_2(1-\alpha_3)\alpha_4/2,$$

$$\text{and } \varepsilon \leq (1-\alpha_1)(1-\alpha_2)\alpha_3\alpha_4/2.$$

In the simulations the dependence values will be expressed in terms of the proportion of the maximum possible value. This value will be determined by the most strict condition for each case. To obtain varying degrees of dependence, the parameters will be set at: no dependence, one half of maximum dependence, and full dependence.

2.5 Simulation Procedure

Simulations were run on a 486DX computer with 4 MBytes of memory and a 220 MByte hard drive using programs written in Borland Pascal, version 7.0, from Borland International, Inc. The values for prevalence and observer error rates along with dependence structure and the number of individuals in the population are input to the simulation program. The program calculates the probability of an individual being in each outcome category using the formulae previously given in Section 2.3. Individuals are assigned a random number between zero and one which is then used to classify them into the various outcome categories with the proper probabilities. This gives frequencies of individuals in each outcome category that are entered into the regular program, Latent. The Latent program was translated to Borland Pascal from Fortran for the purpose of this project. The latent estimates are obtained under the standard assumption of independence of errors between observers.

For any parameter setting many simulations are run. Each simulation creates a population distributed in the outcome categories. The mean of the estimates for each parameter over all simulations is produced. These means are then compared to the originally set parameters to determine the effect that a particular dependence structure has on the latent estimates. In addition, the true standard deviation of the estimates is produced to permit comparison to the large sample theory estimate of standard error given by the Latent program.

2.6 Outcome Summaries

The comparison of estimated parameters and the true parameters involves looking at the percent bias in the estimates. A 20% bias is used as a clinically significant bias in the results. Sample size, meaning the number of simulations, is calculated based on this clinically important amount. Details of these calculations will be discussed in the next section. Tables included in this report display the mean of the parameter estimates for each parameter setting and dependence structure used. Graphs of the means are also included to aid in identifying trends. The graphs show the mean of the estimates versus dependence for each combination of true prevalence values and error rates.

In the cases of four observers, there are enough degrees of freedom to permit goodness of fit testing. The means of the χ^2 values for the simulations are included in the tables. In addition, an estimate of power is included. This estimate is calculated as the proportion of χ^2 values that are above the critical value at the 5% significance level of $\chi^2_{.05,6} = 12.59$. As a result, an estimate of the power of the test to detect a departure from the model assumptions of dependence between observers is obtained. It is expected that at no dependence the value for the power estimate will be .05.

The true standard deviation of the simulated estimates is compared with the estimated standard error from large sample theory. The standard error was estimated from the Latent program using expected frequencies as input. The ratio of standard error to standard deviation will be used as an outcome summary for this comparison.

2.7 Selection of Parameter Values

Parameters that must be set for the simulations include: disease prevalence, θ; false negative rates for all observers, $\beta_1,\beta_2,\ldots,\beta_r$, where r is the number of observers; false positive rates for all observers, $\alpha_1,\alpha_2,\ldots,\alpha_r$; and all dependence parameters, δ_{ij} and ε_{ij}, for $i=1,2,\ldots,r$ and $j=1,2,\ldots,r$ and $i\neq j$. In addition, the number of individuals in the population, N, must be set as well as the number of simulations, n, for each parameter setting. Explanations of how these values are chosen or calculated follows.

Prevalence

In order to obtain an appropriate range of prevalence values, the prevalence is set at .05, .15, and .40. This will be considered low, medium, and high prevalence. Prevalence need not be set at values over .5 due to the symmetry involved in the problem. For example, in a given population the prevalence of condition A is .6 and the false positive rate for observer 1 is .1 and the false negative rate for observer 1 is .15. If an individual does not have condition A then they have condition A^c, A complement. Thus, the same scenario could be reworked by reporting the prevalence of condition A^c as .4 and the false positive rate for observer 1 is .15 while the false negative rate for observer 1 is .1.

Error rates

Error rates are typically set at .1 and .25. Combinations include representation of situations where all observers are equally accurate in their classifications and situations where one or two observers stand out as particularly accurate or inaccurate.

Population size

The population size is set at 2000 for simulations with three observers and at 4000 for simulations with four observers. These numbers are chosen to generally avoid outcome category frequencies of less than 5 in most simulations.

Number of simulations

The number of simulations needed for each combination of the above parameter values is determined using standard sample size calculations. For any parameter, a clinically significant amount of bias in the mean of the estimates is 20%. It is desired to show, at the 5% significance level with 80% power, a clinically significant difference between the mean of the estimates and the true parameter if the difference truly exists. Consider θ to be a general parameter of interest. θ_0 is the true value and θ_1 is the clinically significant result of interest. Large sample theory permits the use of the formula,

$$n = \left[\frac{\sigma\,(z_{\alpha/2} + z_\beta)}{\theta_0 - \theta_1}\right]^2$$

where $z_{\alpha/2} = 1.96$ and $z_\beta = 0.84$ [8]. σ is estimated by the large sample theory estimate of the standard error given by the Latent program under the expected frequencies for the given parameter setting. For any particular parameter setting the number of simulations will be determined by the single parameter that requires the largest sample size. All sample sizes will be at least 30.

2.8 Simulations Runs

Although a factorial design could be implemented here, it was decided to design the simulations using a more basic "one factor at a time" approach. This approach was adopted for ease of understanding and explaining results and has been used in other work in this general area [9]. Various situations will be simulated. All situations will be repeated at all three prevalence levels. Specific input parameter settings for the situations considered are listed below. The first six runs are for different dependence structures in the three observer case. The last six runs are the corresponding situations in the four observer case.

Run 1.

Observers: $r = 3$

Population size: N=2000

Prevalence: θ=.05, .15, .40

Error rates: $\beta = \beta_1 = \beta_2 = \beta_3 = .1, .25$ and $\alpha = \alpha_1 = \alpha_2 = \alpha_3 = .1, .25$

Required number of simulations:

	θ=.05		θ=.15		θ=.40	
	$\alpha = .1$	$\alpha = .25$	$\alpha = .1$	$\alpha = .25$	$\alpha = .1$	$\alpha = .25$
$\beta = .1$	77	469	30	54	30	30
$\beta = .25$	30	167	30	30	30	30

Dependence: Dependence is between observers 1 and 2 with respect to false positive rate, at zero, one half, and maximum. All other dependencies are 0.

	θ=.05, θ=.15, θ=.40	
	$\alpha = .1$	$\alpha = .25$
$\beta = .1$	ε_{12}= 0	ε_{12}= 0
	ε_{12}= .0045	ε_{12}= .0234375
	ε_{12}= .009	ε_{12}= .046875
$\beta = .25$	ε_{12}= 0	ε_{12}= 0
	ε_{12}= .0045	ε_{12}= .0234375
	ε_{12}= .009	ε_{12}= .046875

Run 2.

Observers: $r = 3$

Population size: N=2000

Prevalence: θ=.05, .15, .40

Error rates: $\beta = \beta_1 = \beta_2 = \beta_3 = .1, .25$ and $\alpha = \alpha_1 = \alpha_2 = \alpha_3 = .1, .25$

Required number of simulations:

	θ=.05		θ=.15		θ=.40	
	$\alpha = .1$	$\alpha = .25$	$\alpha = .1$	$\alpha = .25$	$\alpha = .1$	$\alpha = .25$
$\beta = .1$	77	469	30	54	30	30
$\beta = .25$	30	167	30	30	30	30

Dependence: Dependence is between observers 1 and 2 with respect to false negative rate, at zero, one half, and maximum. All other dependencies are 0.

	θ=.05, θ=.15, θ=.40	
	$\alpha = .1$	$\alpha = .25$
$\beta = .1$	$\delta_{12}= 0$ $\delta_{12}= .0045$ $\delta_{12}= .009$	$\delta_{12}= 0$ $\delta_{12}= .0045$ $\delta_{12}= .009$
$\beta = .25$	$\delta_{12}= 0$ $\delta_{12}= .0234375$ $\delta_{12}= .046875$	$\delta_{12}= 0$ $\delta_{12}= .0234375$ $\delta_{12}= .046875$

Run 3.

Observers: $r = 3$

Population size: N=2000

Prevalence: θ=.05, .15, .40

Error rates: $\beta = \beta_1 = \beta_2 = \beta_3 = .1, .25$ and $\alpha = \alpha_1 = \alpha_2 = \alpha_3 = .1, .25$

Required number of simulations:

	θ=.05		θ=.15		θ=.40	
	$\alpha = .1$	$\alpha = .25$	$\alpha = .1$	$\alpha = .25$	$\alpha = .1$	$\alpha = .25$
$\beta = .1$	77	469	30	54	30	30
$\beta = .25$	30	167	30	30	30	30

Dependence: Dependence is between observers 1 and 2 with respect to false positive and false negative rate, at zero, one half, and maximum. All other dependencies are 0.

	θ=.05, θ=.15, θ=.40	
	$\alpha = .1$	$\alpha = .25$
$\beta = .1$	ε_{12}= 0, δ_{12}= 0 ε_{12}= .0045, δ_{12}= .0045 ε_{12}= .009, δ_{12}= .009	ε_{12}= 0, δ_{12}= 0 ε_{12}= .0234375, δ_{12}= .0045 ε_{12}= .046875, δ_{12}= .009
$\beta = .25$	ε_{12}= 0, δ_{12}= 0 ε_{12}= .0045, δ_{12}= .0234375 ε_{12}= .009, δ_{12}= .046875	ε_{12}= 0, δ_{12}= 0 ε_{12}= .0234275, δ_{12}= .0234375 ε_{12}= .046875, δ_{12}= .046875

Run 4.

Observers: $r = 3$

Population size: N=2000

Prevalence: θ=.05, .15, .40

Error rates: $\beta_1 = \beta_2 = .1$, $\beta_3 = .25$ and $\alpha_1 = \alpha_2 = .1$, $\alpha_3 = .25$

Required number of simulations:

	θ=.05	θ=.15	θ=.40
$\beta_1 = \beta_2 = .1$, $\beta_3 = .25$ and $\alpha_1 = \alpha_2 = .1$, $\alpha_3 = .25$	317	41	30

Dependence: Dependence is between observers 1 and 2 with respect to false positive and false negative rate, at zero, one half, and maximum. All other dependencies are 0.

	θ=.05, θ=.15, θ=.40
$\beta_1 = \beta_2 = .1$, $\beta_3 = .25$ and $\alpha_1 = \alpha_2 = .1$, $\alpha_3 = .25$	ε_{12}= 0, δ_{12}= 0 ε_{12}= .01125, δ_{12}= .01125 ε_{12}= .0225, δ_{12}= .0225

Run 5.

Observers: $r = 3$

Population size: N=2000

Prevalence: θ=.05, .15, .40

Error rates: $\beta_1 = \beta_2 = .25$, $\beta_3 = .1$ and $\alpha_1 = \alpha_2 = .25$, $\alpha_3 = .1$

Required number of simulations:

	θ=.05	θ=.15	θ=.40
$\beta_1 = \beta_2 = .25$, $\beta_3 = .1$ and $\alpha_1 = \alpha_2 = .25$, $\alpha_3 = .1$	1756	174	30

Dependence: Dependence is between observers 1 and 2 with respect to false positive and false negative rate, at zero, one half, and maximum. All other dependencies are 0.

	θ=.05, θ=.15, θ=.40
$\beta_1 = \beta_2 = .25$, $\beta_3 = .1$ and $\alpha_1 = \alpha_2 = .25$, $\alpha_3 = .1$	ε_{12}= 0, δ_{12}= 0 ε_{12}= .009375, δ_{12}= .009375 ε_{12}= .01875, δ_{12}= .01875

Run 6.

Observers: $r = 3$

Population size: N=2000

Prevalence: θ=.05, .15, .40

Error rates: $\beta = \beta_1 = \beta_2 = \beta_3 = .1, .25$ and $\alpha = \alpha_1 = \alpha_2 = \alpha_3 = .1, .25$

Required number of simulations:

	θ=.05		θ=.15		θ=.40	
	α = .1	α = .25	α = .1	α = .25	α = .1	α = .25
β = .1	77	469	30	54	30	30
β = .25	30	167	30	30	30	30

Dependence: Equal dependence between all pairs of observers with respect to false positive and false negative rate, at zero, one half, and maximum. That is, $\varepsilon = \varepsilon_{12} = \varepsilon_{13} = \varepsilon_{23}$, and $\delta = \delta_{12} = \delta_{13} = \delta_{23}$.

	θ=.05, θ=.15, θ=.40	
	α = .1	α = .25
β = .1	ε= 0, δ= 0	ε= 0, δ= 0
	ε= .0045, δ= .0045	ε= .0234375, δ= .0045
	ε= .009, δ= .009	ε= .046875, δ= .009
β = .25	ε= 0, δ= 0	ε= 0, δ= 0
	ε= .0045, δ= .0234375	ε= .0234275, δ= .0234375
	ε= .009, δ= .046875	ε= .046875, δ= .046875

Run 7.

Observers: $r = 4$

Population size: N=4000

Prevalence: θ=.05, .15, .40

Error rates: $\beta = \beta_1 = \beta_2 = \beta_3 = \beta_4 = .1, .25$ and $\alpha = \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = .1, .25$

Required number of simulations:

	θ=.05		θ=.15		θ=.40	
	$\alpha = .1$	$\alpha = .25$	$\alpha = .1$	$\alpha = .25$	$\alpha = .1$	$\alpha = .25$
$\beta = .1$	30	55	30	30	30	30
$\beta = .25$	30	30	30	30	30	30

Dependence: Dependence is between observers 1 and 2 with respect to false positive rate, at zero, one half, and maximum. All other dependencies are 0.

	θ=.05, θ=.15, θ=.40	
	$\alpha = .1$	$\alpha = .25$
$\beta = .1$	ε_{12}= 0 ε_{12}= .00045 ε_{12}= .0009	ε_{12}= 0 ε_{12}= .005859375 ε_{12}= .01171875
$\beta = .25$	ε_{12}= 0 ε_{12}= .00045 ε_{12}= .0009	ε_{12}= 0 ε_{12}= .005859375 ε_{12}= .01171875

Run 8.

Observers: $r = 4$

Population size: N=4000

Prevalence: θ=.05, .15, .40

Error rates: $\beta = \beta_1 = \beta_2 = \beta_3 = \beta_4 = .1, .25$ and $\alpha = \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = .1, .25$

Required number of simulations:

	θ=.05		θ=.15		θ=.40	
	$\alpha = .1$	$\alpha = .25$	$\alpha = .1$	$\alpha = .25$	$\alpha = .1$	$\alpha = .25$
$\beta = .1$	30	55	30	30	30	30
$\beta = .25$	30	30	30	30	30	30

Dependence: Dependence is between observers 1 and 2 with respect to false negative rate, at zero, one half, and maximum. All other dependencies are 0.

	θ=.05, θ=.15, θ=.40	
	$\alpha = .1$	$\alpha = .25$
$\beta = .1$	$\delta_{12}= 0$ $\delta_{12}= .00045$ $\delta_{12}= .0009$	$\delta_{12}= 0$ $\delta_{12}= .00045$ $\delta_{12}= .0009$
$\beta = .25$	$\delta_{12}= 0$ $\delta_{12}= .005859375$ $\delta_{12}= .01171875$	$\delta_{12}= 0$ $\delta_{12}= .005859375$ $\delta_{12}= .01171875$

Run 9.

Observers: $r = 4$

Population size: $N = 4000$

Prevalence: $\theta = .05, .15, .40$

Error rates: $\beta = \beta_1 = \beta_2 = \beta_3 = \beta_4 = .1, .25$ and $\alpha = \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = .1, .25$

Required number of simulations:

	θ=.05		θ=.15		θ=.40	
	α = .1	α = .25	α = .1	α = .25	α = .1	α = .25
β = .1	30	55	30	30	30	30
β = .25	30	30	30	30	30	30

Dependence: Dependence is between observers 1 and 2 with respect to false positive and false negative rate, at zero, one half, and maximum. All other dependencies are 0.

	θ=.05, θ=.15, θ=.40	
	α = .1	α = .25
β = .1	$\varepsilon_{12}= 0, \delta_{12}= 0$ $\varepsilon_{12}= .00045, \delta_{12}= .00045$ $\varepsilon_{12}= .0009, \delta_{12}= .0009$	$\varepsilon_{12}= 0, \delta_{12}= 0$ $\varepsilon_{12}= .005859375, \delta_{12}= .0045$ $\varepsilon_{12}= .01171875, \delta_{12}= .009$
β = .25	$\varepsilon_{12}= 0, \delta_{12}= 0$ $\varepsilon_{12}= .00045, \delta_{12}= .005859375$ $\varepsilon_{12}= .0009, \delta_{12}= .01171875$	$\varepsilon_{12}= 0, \delta_{12}= 0$ $\varepsilon_{12}= .005859375, \delta_{12}= .005859375$ $\varepsilon_{12}= .01171875, \delta_{12}= .01171875$

Run 10.

Observers: $r = 4$

Population size: $N = 4000$

Prevalence: $\theta = .05, .15, .40$

Error rates: $\beta_1 = \beta_2 = .1$, $\beta_3 = \beta_4 = .25$ and $\alpha_1 = \alpha_2 = .1$, $\alpha_3 = \alpha_4 = .25$

Required number of simulations:

	θ=.05	θ=.15	θ=.40
$\beta_1 = \beta_2 = .1$, $\beta_3 = \beta_4 = .25$ and $\alpha_1 = \alpha_2 = .1$, $\alpha_3 = \alpha_3 = .25$	59	30	30

Dependence: Dependence is between observers 1 and 2 with respect to false positive and false negative rate, at zero, one half, and maximum. All other dependencies are 0.

	θ=.05, θ=.15, θ=.40
$\beta_1 = \beta_2 = .1$, $\beta_3 = \beta_4 = .25$ and $\alpha_1 = \alpha_2 = .1$, $\alpha_3 = \alpha_4 = .25$	$\varepsilon_{12}= 0$, $\delta_{12}= 0$ $\varepsilon_{12}= .0028125$, $\delta_{12}= .0028125$ $\varepsilon_{12}= .005625$, $\delta_{12}= .005625$

Run 11.

Observers: $r = 4$

Population size: $N = 4000$

Prevalence: $\theta = .05, .15, .40$

Error rates: $\beta_1 = \beta_2 = .25$, $\beta_3 = \beta_4 = .1$ and $\alpha_1 = \alpha_2 = .25$, $\alpha_3 = \alpha_4 = .1$

Required number of simulations:

	θ=.05	θ=.15	θ=.40
$\beta_1 = \beta_2 = .25$, $\beta_3 = \beta_4 = .1$ and $\alpha_1 = \alpha_2 = .25$, $\alpha_3 = \alpha_4 = .1$	59	30	30

Dependence: Dependence is between observers 1 and 2 with respect to false positive and false negative rate, at zero, one half, and maximum. All other dependencies are 0.

	θ=.05, θ=.15, θ=.40
$\beta_1 = \beta_2 = .25$, $\beta_3 = \beta_4 = .1$ and $\alpha_1 = \alpha_2 = .25$, $\alpha_3 = \alpha_4 = .1$	$\varepsilon_{12} = 0$, $\delta_{12} = 0$ $\varepsilon_{12} = .0009375$, $\delta_{12} = .0009375$ $\varepsilon_{12} = .001875$, $\delta_{12} = .001875$

Run 12.

Observers: $r = 4$

Population size: $N = 4000$

Prevalence: $\theta = .05, .15, .40$

Error rates: $\beta = \beta_1 = \beta_2 = \beta_3 = \beta_4 = .1, .25$ and $\alpha = \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = .1, .25$

Required number of simulations:

	θ=.05		θ=.15		θ=.40	
	α = .1	α = .25	α = .1	α = .25	α = .1	α = .25
β = .1	30	55	30	30	30	30
β = .25	30	30	30	30	30	30

Dependence: Equal dependence between all pairs of observers with respect to false positive and false negative rate, at zero, one half, and maximum. That is, $\varepsilon = \varepsilon_{12} = \varepsilon_{13} = \varepsilon_{14} = \varepsilon_{23} = \varepsilon_{24} = \varepsilon_{34}$, and $\delta = \delta_{12} = \delta_{13} = \delta_{14} = \delta_{23} = \delta_{24} = \delta_{34}$.

	θ=.05, θ=.15, θ=.40	
	α = .1	α = .25
β = .1	ε= 0, δ= 0 ε= .00045, δ= .00045 ε= .0009, δ= .0009	ε= 0, δ= 0 ε= .005859375, δ= .00045 ε= .01171875, δ= .0009
β = .25	ε= 0, δ= 0 ε= .00045, δ= .005859375 ε= .0009, δ= .01171875	ε= 0, δ= 0 ε= .005859375, δ= .005859375 ε= .01171875, δ= .01171875

3. RESULTS

3.1 Bias in Estimates

Dependence between two observers means that the observers classify individuals in the same way more often than in the independent case. In particular, dependence will increase the frequency of simultaneous classification errors. Through the maximum likelihood procedure the apparent probability that the misclassifications of two dependent observers are correct will increase. This will cause the error rates for the dependent observers to be underestimated.

Since the prevalence will always be less than 50% there will always be more truly negative than truly positive individuals in the simulated populations. This leads to the expectation that the false negative rate estimates will be less precise and more susceptible to bias due to the smaller sample of positive than negative individuals. These estimates will be more biased because the larger proportion of dependent errors being made on the truly negative individuals will overpower the smaller sample of truly positive individuals, thus the false negative rates will be more biased.

As discussed earlier the mean of the estimates for a parameter will be considered seriously biased if it differs from the truth by more than 20%. In the tables to be discussed in the following sections, means biased by more than 20% are printed in bold.

The situations simulated were listed and described in Section 2.8. The results from these simulation runs will now be reported. The first six runs are for the three observer case while the next 6 runs are identical situations with four observers. The discussion of the effects of the dependence on the estimates will be detailed for the three observer cases. The discussion for the four observer cases will centre mainly on the differences between the four observer and the corresponding three observer case. In simulations where all error rates are equal, FPR and FNR will be used to refer to the false positive rates ($\alpha_1 = \alpha_2 = \alpha_3$) and the false negative rates ($\beta_1 = \beta_2 = \beta_3$), respectively.

3.1.1 RUN 1 - Dependence between two observers with respect to false positive rate - Three Observer Case

This run includes all combinations of θ, FPR, FNR with dependence between observers 1 and 2 with respect to false positive rate only. Tables 1a, 1b and 1c show the results for $\theta = .05$, $\theta = .15$, and $\theta = .40$, respectively. Figures 1a to 1g are graphs which display the trends in the mean of the estimates as the dependence increases for each prevalence value.

OBSERVATIONS

Effect on prevalence estimates

From the tables and Figure 1a, it can be seen that the dependence tends to produce a positive bias in the prevalence estimates. The most serious bias occurs when the true prevalence is low at .05 or .15 and the FPR is high at .25. For example, in Table 1a when the FNR = .1 and the FPR = .25 the mean of the prevalence estimates is at 11% at one half of maximum dependence and increases to 17% at full dependence when the true prevalence is 5%. There is no substantial bias in the mean of the estimates of prevalence when the true prevalence is high.

Effect on false negative rate estimates

False negative rates for the two dependent observers are very seriously underestimated even at one half of maximum dependence. There is an inverse effect on the third observer's false negative rate which is seriously overestimated. The effect is strongest for low prevalence values and high FPR. These trends are easily seen in Figures 1b to 1d.

Effect on false positive rate estimates

A bias is seen in the two dependent observers' false positive rates when the FPR is at .25. This underestimation of the rates is seen at all prevalences but increases slightly as the prevalence decreases, see Figure 1e and 1f. Figure 1g shows an apparent inverse effect on the third observer's false positive rate. This effect appears to be strongest for high prevalence values.

EXPLANATIONS

The dependence between observers 1 and 2 with respect to the false positive classifications causes these observers to classify truly negative individuals in the same way more often than in the independent case. The more often both observer 1 and 2 make a false positive error, the higher the prevalence estimate will be, because the errors will be considered more likely to be correct since they are occurring together. Thus, when the false positive rate is high, the prevalence will be more biased. Since these false positive errors are being considered correct, the false positive rates for the two dependent observers drop. In addition the false negative rates will also drop due to the increased agreement between the dependent observers. The lower the prevalence the more serious the bias because there are more negative individuals to be classified in the same way by the dependent observers.

Since observers 1 and 2 are being considered correct a great proportion of the time, observer 3 is estimated to have a high false negative rate due to an increased lack of agreement with the two dependent observers. The dependent observers misclassify negative individuals as positive, leaving observer 3 with what appears to be a likely erroneous classification as negative. The false negative rate is more strongly affected than the false positive rate because the false negative rate is based on truly positive individuals of which there are a small number.

IMPLICATIONS

In situations where the false positive rates are high and the prevalence is low, dependence between two observers with respect to false positive classifications can lead to very serious bias in all estimates except the third observer's false positive rate. This situation commonly occurs in population screenings where the tests are often designed to have high sensitivity (a high proportion of true positives classified as positive) and lower specificity (a lower proportion of true negatives classified as negative). Sackett et. al. [5] refer to this kind of a test as SnNout, meaning a test with a sufficiently high ***Sen***sitivity so that a *N*egative result rules ***out*** the disease. Therefore, the use of latent class methods to evaluate screening tests for a rare disease should be approached with the following warning. The presence of any dependence between the tests with respect to false positive classifications, meaning any commonalities between the tests when classifying truly negative individuals that will lead to false positive classifications, will cause seriously underestimated error rates for these tests, especially the false negative rates. Equivalently, the specificity and especially the sensitivity of the two dependent tests will be overestimated leading to a false sense of confidence regarding the accuracy of these tests. Most care should be taken at very low prevalences where even a slight dependence will result in very biased results. The higher the prevalence the more accurate are results from the Latent estimation procedure in the presence of this kind of dependence. Even at a high prevalence, if the false positive rate is high then dependence will cause bias in most of the error rates.

TABLE 1a
Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false positive rate only

True Prevalence = .05

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	5.23 (0.75)	10.49 (6.51)	10.75 (5.77)	10.89 (6.19)	9.95 (0.84)	9.75 (0.74)	9.97 (0.70)
	0.5	5.34 (0.66)	**2.14 (2.87)**	**1.82 (3.31)**	**18.38 (5.44)**	9.44 (0.79)	9.31 (0.73)	10.21 (0.75)
	1	**6.05 (0.53)**	**0.02 (0.22)**	**0.00 (0.00)**	**24.19 (4.29)**	8.63 (0.73)	8.36 (0.67)	10.12 (0.82)
β =.1 α =.25	0	5.77 (1.90)	11.43 (10.87)	12.14 (10.43)	11.74 (10.51)	24.64 (1.36)	24.73 (1.29)	24.69 (1.32)
	0.5	**11.21 (0.86)**	**0.02 (0.25)**	**0.05 (0.43)**	**40.89 (3.91)**	**19.27 (1.02)**	**19.22 (1.04)**	24.35 (1.00)
	1	**17.46 (0.87)**	**0.00 (0.00)**	**0.00 (0.00)**	**47.12 (2.79)**	**13.06 (0.93)**	**13.03 (0.91)**	22.93 (1.07)
β =.25 α =.1	0	4.98 (1.21)	24.55 (8.90)	22.19 (8.75)	23.33 (10.97)	10.19 (0.94)	10.02 (0.75)	10.36 (0.75)
	0.5	4.68 (0.78)	**8.25 (7.00)**	**9.08 (6.15)**	**32.69 (7.87)**	9.50 (0.78)	9.74 (0.62)	10.66 (0.91)
	1	4.90 (0.47)	**0.50 (1.81)**	**0.80 (1.83)**	**35.97 (4.35)**	9.07 (0.64)	9.21 (0.62)	10.66 (0.88)
β =.25 α =.25	0	**6.47 (4.46)**	22.95 (17.41)	25.27 (17.45)	24.46 (18.30)	24.41 (1.82)	24.85 (1.87)	24.77 (1.73)
	0.5	**9.76 (0.83)**	**0.14 (1.24)**	**0.20 (1.33)**	**49.22 (4.41)**	**19.58 (1.03)**	**19.63 (0.98)**	24.93 (1.09)
	1	**16.10 (0.85)**	**0.00 (0.00)**	**0.00 (0.00)**	**52.16 (3.28)**	**13.56 (0.96)**	**13.64 (0.90)**	23.56 (0.99)

* Proportion of maximum dependence between observers 1 and 2 for true negative subjects.
β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.
Means with 20% or greater bias are displayed in bold.

TABLE 1b

Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false positive rate only

True Prevalence = .15

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	14.82 (0.96)	10.02 (3.15)	9.34 (2.75)	9.58 (2.62)	10.13 (1.03)	10.13 (0.96)	9.99 (0.80)
	0.5	15.10 (0.87)	**7.55 (2.24)**	**6.89 (2.10)**	**12.86 (2.08)**	9.19 (0.70)	9.38 (0.74)	10.60 (0.78)
	1	15.16 (0.81)	**3.57 (1.63)**	**4.04 (1.56)**	**15.93 (2.41)**	9.09 (0.73)	8.95 (0.70)	11.06 (0.85)
β =.1 α =.25	0	14.84 (2.37)	9.72 (5.11)	9.37 (5.26)	9.53 (5.26)	25.14 (1.29)	25.08 (1.24)	24.85 (1.60)
	0.5	**18.40 (1.06)**	**0.66 (1.62)**	**0.65 (1.58)**	**26.12 (3.07)**	20.32 (1.31)	20.26 (1.16)	25.75 (1.08)
	1	**23.75 (1.07)**	**0.00 (0.00)**	**0.00 (0.00)**	**33.45 (2.71)**	**14.25 (0.86)**	**14.46 (0.99)**	24.79 (1.16)
β =.25 α =.1	0	14.55 (1.55)	24.45 (4.15)	24.48 (4.71)	23.66 (3.28)	10.29 (1.09)	10.23 (0.99)	10.18 (1.04)
	0.5	14.75 (1.14)	20.27 (3.97)	**19.79 (3.89)**	28.42 (3.00)	9.39 (0.90)	9.17 (0.80)	10.83 (0.97)
	1	14.44 (1.21)	**16.42 (4.54)**	**16.57 (4.00)**	**30.45 (3.37)**	9.16 (0.98)	8.91 (0.87)	11.41 (0.69)
β =.25 α =.25	0	17.85 (3.77)	27.04 (6.74)	28.81 (6.18)	28.39 (6.24)	23.87 (1.53)	24.08 (1.46)	24.42 (1.97)
	0.5	15.00 (2.03)	**2.34 (3.43)**	**4.33 (5.61)**	**37.43 (3.56)**	21.18 (1.64)	21.90 (1.38)	27.18 (1.58)
	1	**19.69 (1.07)**	**0.00 (0.00)**	**0.00 (0.00)**	**43.52 (1.90)**	**16.21 (0.99)**	**15.89 (0.97)**	26.51 (1.13)

* Proportion of maximum dependence between observers 1 and 2 for true negative subjects.
β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.
Means with 20% or greater bias are displayed in bold.

TABLE 1c

Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false positive rate only

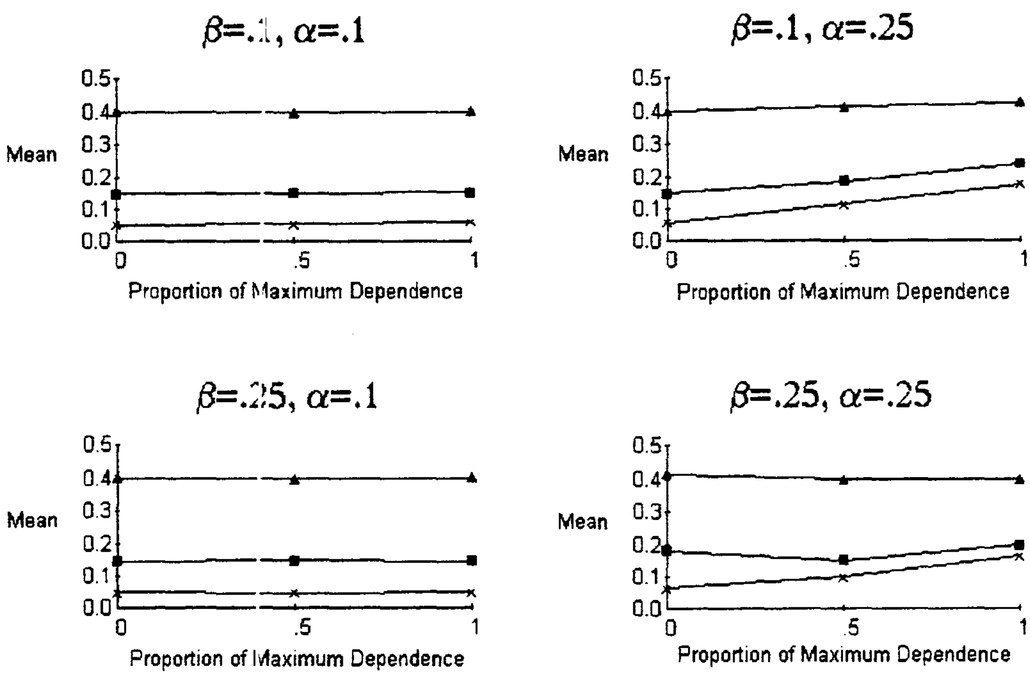
True Prevalence = .40

Truth		Mean of Estimates (%) (standard deviation (%))						
Error Rates	Dep*	θ	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	39.96 (1.32)	9.96 (1.32)	9.97 (1.43)	9.99 (1.28)	9.95 (0.84)	10.08 (1.09)	9.99 (0.95)
	0.5	39.80 (1.36)	9.24 (1.06)	9.11 (1.07)	10.71 (1.37)	9.44 (1.06)	9.38 (0.91)	10.46 (1.11)
	1	39.97 (1.64)	8.20 (1.24)	8.55 (1.53)	11.70 (1.35)	8.71 (1.06)	8.69 (1.09)	11.07 (0.99)
β =.1 α =.25	0	39.85 (1.72)	10.04 (1.93)	10.32 (1.98)	10.38 (2.04)	25.13 (1.79)	25.31 (1.75)	25.25 (1.62)
	0.5	41.23 (1.78)	**5.35 (1.56)**	**6.18 (1.59)**	**15.36 (1.86)**	20.57 (1.75)	21.19 (1.86)	27.64 (1.60)
	1	42.48 (1.33)	**3.09 (1.76)**	**2.81 (1.60)**	**18.95 (1.73)**	**16.76 (1.99)**	**17.23 (1.38)**	28.60 (1.61)
β =.25 α =.1	0	40.07 (2.02)	24.67 (2.40)	24.40 (1.99)	25.27 (2.24)	9.98 (1.07)	9.88 (1.38)	10.13 (1.59)
	0.5	39.69 (1.55)	24.16 (1.53)	23.50 (2.00)	25.80 (2.14)	9.60 (1.45)	9.61 (0.95)	10.93 (1.23)
	1	40.08 (2.02)	22.93 (2.50)	23.37 (2.28)	27.20 (2.64)	8.61 (1.15)	9.16 (1.28)	11.28 (1.07)
β =.25 α =.25	0	41.47 (3.51)	25.96 (2.54)	24.85 (3.51)	26.46 (3.51)	24.67 (2.12)	24.27 (2.24)	24.53 (2.36)
	0.5	39.76 (3.23)	**17.97 (3.29)**	**18.70 (3.09)**	**30.28 (2.24)**	20.54 (2.03)	21.02 (2.19)	28.88 (1.97)
	1	39.75 (3.04)	**12.58 (2.77)**	**12.88 (3.45)**	**33.46 (2.36)**	**16.85 (2.14)**	**17.26 (2.19)**	**30.93 (1.55)**

* Proportion of maximum dependence between observers 1 and 2 for true negative subjects.
β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.
Means with 20% or greater bias are displayed in bold.

Figure 1a

Mean of Estimates of Prevalence (θ) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

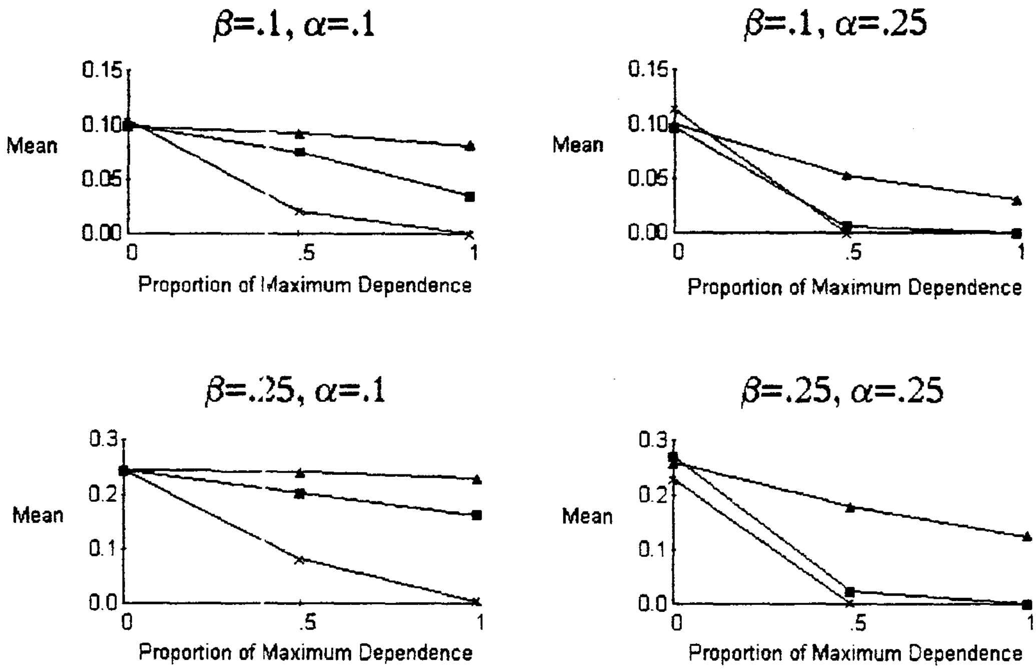
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 1b

Mean of Estimates of false negative rate for observer 1 (β_1) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

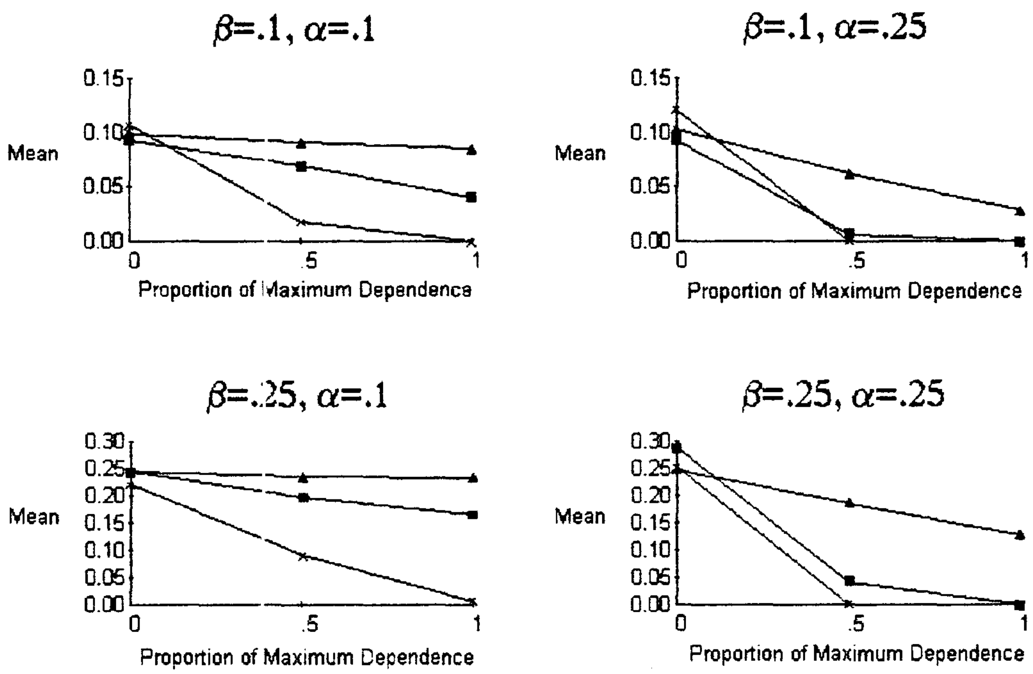
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 1c

Mean of Estimates of false negative rate for observer 2 (β_2) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

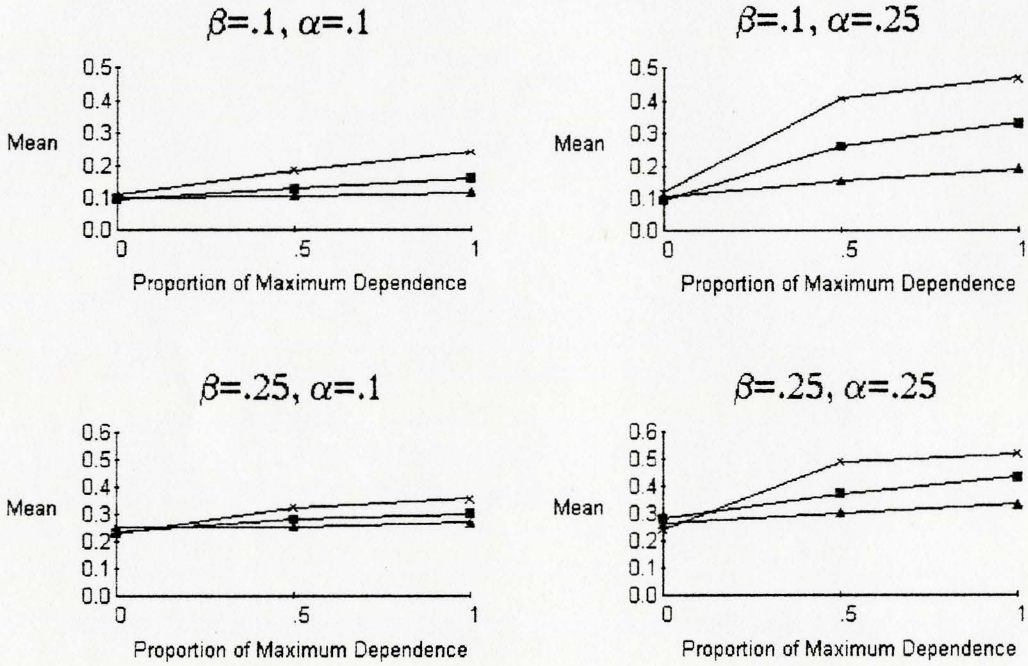
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 1d

Mean of Estimates of false negative rate for observer 3 (β_3) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

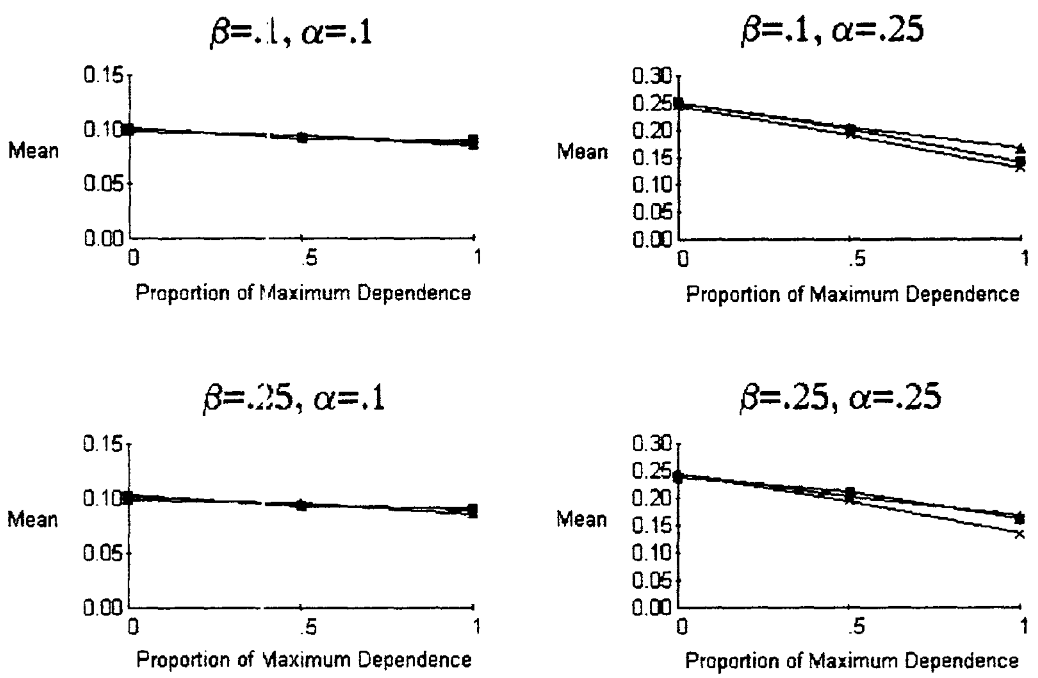
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 1e

Mean of Estimates of false positive rate for observer 1 (α_1) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

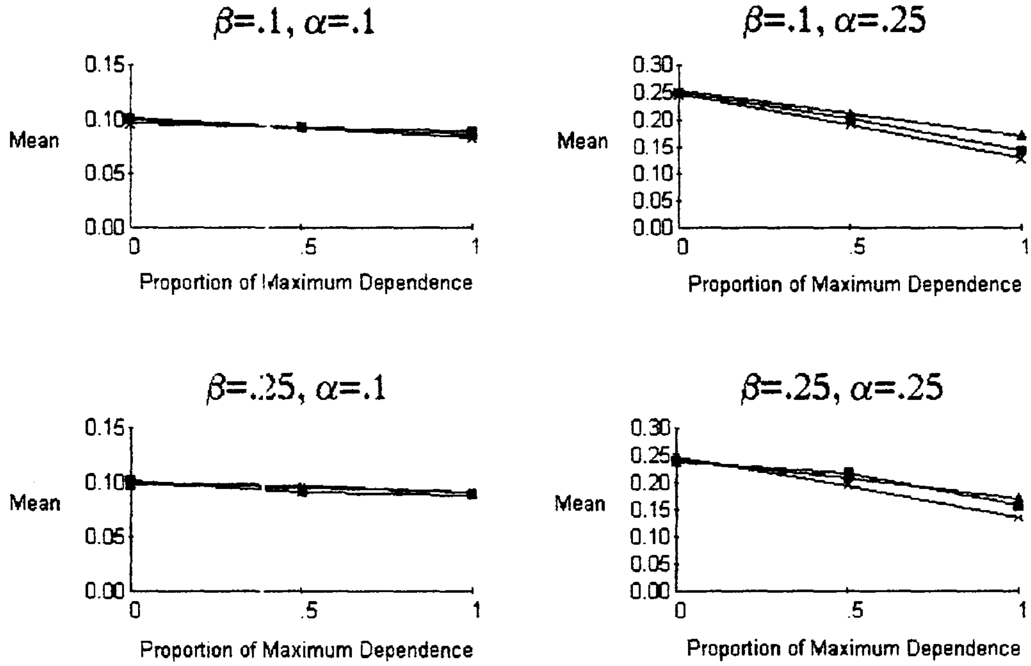
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 1f

Mean of Estimates of false positive rate for observer 2 (α_2) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

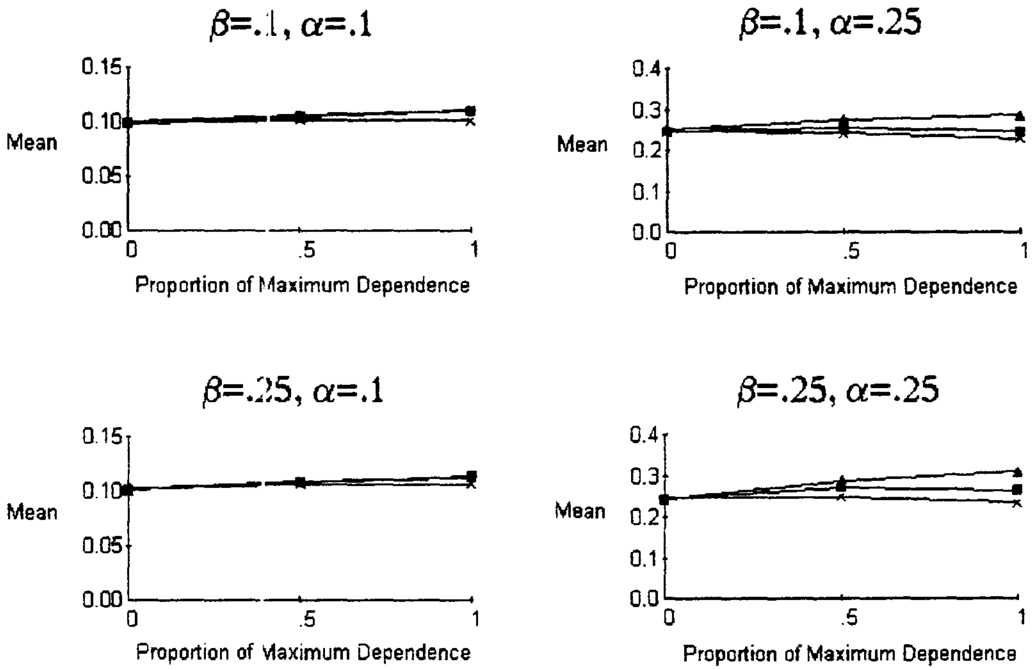
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 1g

Mean of Estimates of false positive rate for observer 3 (α_3) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

3.1.2 RUN 2 - Dependence between two observers with respect to false negative rate - Three observer case

This run includes all combinations of θ, FPR, FNR with dependence between observers 1 and 2 with respect to false negative rate only. The results for $\theta = .05$, $\theta = .15$, and $\theta = .40$ are given in tables 2a, 2b and 2c, respectively. Figures 2a to 2g display the mean of the estimates as the dependence increases for each prevalence value.

OBSERVATIONS

Effect on prevalence estimates

Most prevalence estimates show no substantial bias. Table 2a shows that the means of the prevalence estimates all overestimate the truth by more than 20% when the FNR = .25 and the FPR = .25, even the estimate from independent observers. A more detailed look at the actual estimates and the fact that the standard deviations of the estimates are quite large indicated that a few simulations found the symmetrical solution to the problem. This means that the prevalence estimates were closer to 95% than 5% and some error rates were estimated to be above .5. This instability most likely occurred in a few instances due to small frequencies in some outcome categories because of the low value for true prevalence.

Effect on false negative rate estimates

The most serious bias is seen when the FNR is high at .25 where for all prevalences the false negative rate for the two dependent observers is underestimated by more than 20% at maximum dependence. The third observer's false negative rate is slightly overestimated in the same situations, just barely reaching 20% bias at full dependence. These biases are not nearly as severe as those seen in Run 1.

Effect on false positive rate estimates

At θ = .05 and θ = .15 there is no apparent effect from the dependence on the false positive rate estimates However, at the high prevalence of .40 the rates for the dependent observers are underestimated when FNR = .25. The third observer's false positive rate is slightly overestimated when FNR = .25.

EXPLANATIONS

The dependence between observers 1 and 2 with respect to the false negative classifications causes these observers to classify truly positive individuals the same more often than if they were independent. This run can be thought of as complementary to Run 1, meaning that if Run 1 were repeated with prevalences above .5 then the results should be the same as those found here. True positive individuals in this run would be true negative individuals in Run 1, false negative classifications would be false positives, and false positive classifications would be false negatives. For example,

Run 2	is equivalent to	Run 1
$\theta = .05$	$\Rightarrow$	$\theta = .95$
$\alpha = .1$	$\Rightarrow$	$\beta = .1$
$\beta = .1$	$\Rightarrow$	$\alpha = .1$
$\delta = \max$	$\Rightarrow$	$\varepsilon = \max$

For this reason trends that were discovered in Run 1 will also be present here in a complementary direction. However, there are fewer truly positive than truly negative individuals, so the effects of this dependence will not be as strong as in Run 1.

The more often both observer 1 and 2 make a false negative error simultaneously the more often truly positive individuals will simultaneously be classified as negative. These false negative errors lead to a higher estimated probability that the individual is negative. As a result the false negative rates for the two dependent observers drop, especially when the true false negative rates are high. Since the number of truly positive individuals is small, the dependence does not effect the prevalence or the false positive rates. The false positive rates begin to be affected when the prevalence becomes high at 40%. This is complementary to the effects on the false negative rates in Run 1 that were stronger when the prevalence was low.

IMPLICATIONS

If the true prevalence is at most 50% and it is suspected that there is dependence between two observers with respect to classification of positive individuals then the danger of biased estimates is not nearly as great as if the dependence was with respect to classification of negative individuals or both, as will be discussed next. In this situation large false negative rates will cause false negative rate estimates to be low for the dependent observers and high for the independent observer. At high prevalence values the false positive rate estimates will be underestimated for the dependent observers and overestimated for the independent observer if the false negative rates are high. Hence, in this situation the dangerous parameter values are high false negative rates.

TABLE 2a
Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false negative rate only

True Prevalence = .05

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	5.15 (0.72)	9.87 (6.62)	10.30 (5.97)	11.27 (5.63)	9.94 (0.77)	9.81 (0.81)	9.84 (0.67)
	0.5	5.07 (0.66)	9.18 (6.57)	9.01 (5.65)	10.45 (5.49)	9.96 (0.87)	9.97 (0.88)	10.11 (0.74)
	1	4.91 (0.70)	8.28 (6.46)	**7.94 (5.59)**	10.23 (5.91)	9.92 (0.80)	9.93 (0.84)	10.19 (0.87)
β =.1 α =.25	0	5.87 (2.25)	**12.28 (11.16)**	11.68 (10.62)	**12.29 (11.13)**	24.76 (1.34)	24.71 (1.39)	24.73 (1.36)
	0.5	5.86 (1.99)	**12.20 (10.35)**	11.70 (10.58)	11.95 (10.76)	24.69 (1.33)	24.61 (1.30)	24.69 (1.27)
	1	5.91 (2.07)	**12.13 (10.92)**	11.15 (10.41)	**13.08 (11.06)**	24.70 (1.34)	24.68 (1.31)	24.71 (1.29)
β =.25 α =.1	0	4.90 (1.11)	24.16 (9.27)	21.93 (9.41)	25.03 (9.46)	10.34 (0.85)	9.99 (0.85)	10.18 (0.64)
	0.5	4.76 (0.71)	**19.06 (8.22)**	**19.60 (6.90)**	27.01 (8.03)	9.72 (0.71)	10.12 (0.71)	10.35 (0.72)
	1	4.52 (0.78)	**15.56 (8.44)**	**12.34 (7.78)**	**30.18 (8.06)**	9.81 (0.72)	9.58 (0.89)	10.57 (0.87)
β =.25 α =.25	0	**8.80 (12.02)**	26.71 (19.01)	26.15 (20.40)	24.81 (19.15)	24.17 (3.37)	24.13 (2.95)	24.10 (2.67)
	0.5	**6.60 (4.78)**	20.45 (17.27)	22.46 (17.12)	26.97 (16.02)	24.36 (1.87)	24.51 (1.86)	24.76 (1.83)
	1	**7.48 (7.58)**	**16.99 (16.62)**	**19.61 (17.14)**	**32.55 (15.74)**	23.91 (2.12)	23.98 (2.10)	24.88 (1.74)

* Proportion of maximum dependence between observers 1 and 2 for true positive subjects.
β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.
Means with 20% or greater bias are displayed in bold.

TABLE 2b
Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false negative rate only

True Prevalence = .15

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	15.44 (0.87)	9.91 (2.31)	9.98 (1.85)	10.53 (2.05)	10.15 (0.78)	10.16 (0.91)	9.97 (0.77)
	0.5	15.07 (1.11)	8.49 (2.05)	9.34 (2.22)	10.75 (2.50)	9.91 (0.81)	9.87 (0.75)	10.22 (0.74)
	1	15.09 (0.87)	8.98 (2.27)	8.44 (2.07)	10.67 (2.10)	9.95 (0.85)	9.65 (0.78)	10.09 (0.85)
β =.1 α =.25	0	14.63 (1.81)	8.98 (5.14)	8.47 (4.57)	9.47 (4.56)	24.92 (1.33)	24.95 (1.56)	25.12 (1.42)
	0.5	15.26 (2.23)	10.37 (5.47)	8.87 (4.91)	10.52 (5.00)	24.84 (1.57)	24.96 (1.45)	24.68 (1.42)
	1	15.52 (2.20)	8.55 (4.71)	9.52 (5.23)	11.44 (4.92)	24.43 (1.44)	24.43 (1.64)	24.95 (1.42)
β =.25 α =.1	0	14.96 (1.76)	25.66 (3.47)	25.86 (3.35)	24.74 (4.25)	9.98 (0.97)	10.15 (0.97)	9.97 (1.06)
	0.5	14.97 (1.25)	20.90 (4.00)	20.46 (4.57)	27.03 (3.52)	9.39 (1.04)	9.68 (0.85)	10.72 (0.85)
	1	13.99 (1.23)	**16.29 (3.89)**	**15.84 (3.41)**	29.95 (4.29)	9.23 (0.83)	9.37 (0.67)	11.59 (0.86)
β =.25 α =.25	0	14.58 (5.54)	23.37 (8.16)	22.26 (9.64)	22.20 (11.18)	25.33 (2.29)	25.35 (2.08)	25.37 (2.35)
	0.5	14.70 (4.17)	21.03 (9.51)	17.34 (9.35)	28.16 (6.84)	24.32 (1.81)	24.00 (1.97)	25.93 (1.64)
	1	15.15 (4.21)	**14.05 (9.55)**	**15.14 (8.04)**	**32.64 (4.94)**	23.63 (2.00)	23.56 (2.51)	26.44 (1.84)

* Proportion of maximum dependence between observers 1 and 2 for true positive subjects.
β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.
Means with 20% or greater bias are displayed in bold.

TABLE 2c
Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false negative rate only

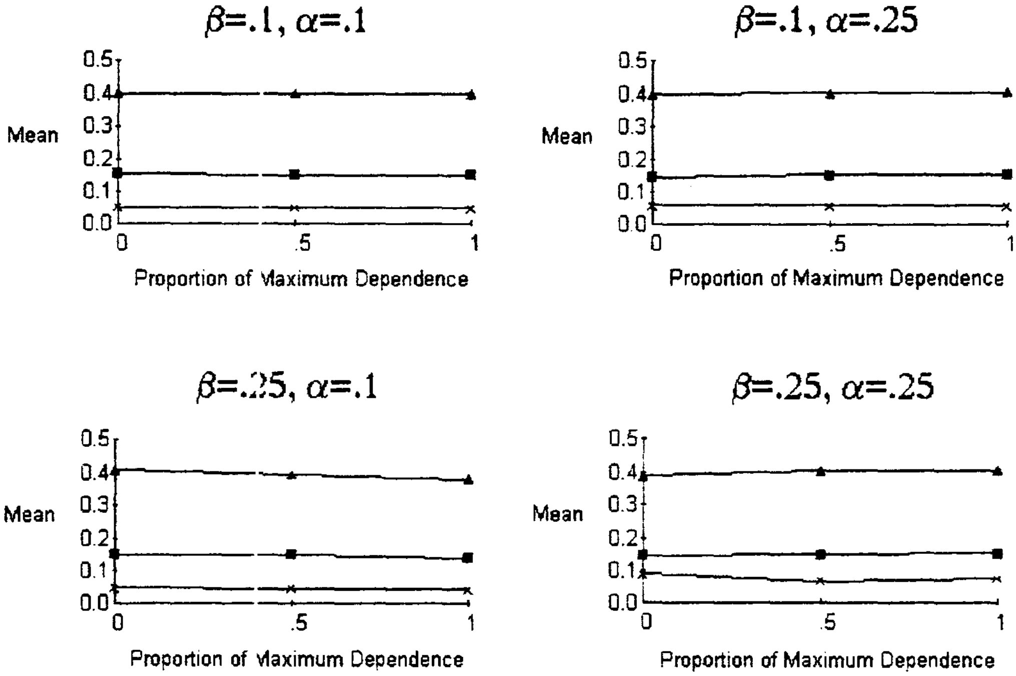
True Prevalence = .40

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	40.06 (1.04)	10.10 (1.40)	10.02 (1.34)	10.24 (1.23)	10.26 (1.01)	9.64 (1.01)	10.15 (0.90)
	0.5	39.88 (1.35)	9.53 (1.21)	9.48 (1.15)	10.46 (1.32)	9.51 (0.98)	9.67 (1.07)	10.23 (1.01)
	1	39.64 (1.27)	8.78 (1.12)	8.63 (0.96)	11.11 (1.12)	9.07 (0.80)	9.43 (0.83)	10.78 (1.05)
β =.1 α =.25	0	39.80 (2.71)	9.63 (2.39)	9.68 (2.21)	9.73 (2.34)	25.55 (1.99)	25.10 (2.32)	25.37 (2.22)
	0.5	40.17 (2.47)	9.44 (1.95)	8.82 (2.34)	10.81 (2.13)	24.47 (1.96)	24.27 (2.37)	25.50 (1.95)
	1	40.49 (1.67)	8.25 (1.84)	8.81 (2.10)	11.94 (1.63)	23.43 (1.46)	23.98 (1.93)	26.15 (1.60)
β =.25 α =.1	0	40.72 (1.84)	25.68 (2.44)	25.42 (2.79)	25.07 (2.10)	9.84 (1.30)	9.94 (1.04)	9.62 (1.22)
	0.5	39.40 (1.91)	21.43 (1.90)	20.79 (2.50)	27.57 (1.90)	8.21 (1.44)	8.36 (1.03)	**12.56 (1.46)**
	1	38.09 (1.73)	**16.92 (2.11)**	**15.74 (1.94)**	29.26 (2.02)	**6.90 (1.18)**	**6.76 (1.32)**	**14.60 (1.69)**
β =.25 α =.25	0	38.73 (3.45)	22.84 (4.42)	24.47 (2.77)	23.63 (3.19)	25.00 (2.10)	25.64 (2.43)	25.69 (1.98)
	0.5	40.28 (4.01)	**19.63 (3.30)**	20.94 (3.93)	29.18 (2.61)	21.48 (2.95)	22.45 (1.84)	27.49 (2.31)
	1	40.52 (3.47)	**16.87 (3.43)**	**15.99 (3.68)**	**31.32 (2.74)**	**19.95 (2.53)**	**18.88 (2.70)**	29.35 (1.63)

* Proportion of maximum dependence between observers 1 and 2 for true positive subjects.
β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.
Means with 20% or greater bias are displayed in bold.

Figure 2a

Mean of Estimates of Prevalence (θ)
in the case of three observers, with dependence between observers 1 and 2 with respect to false negative rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

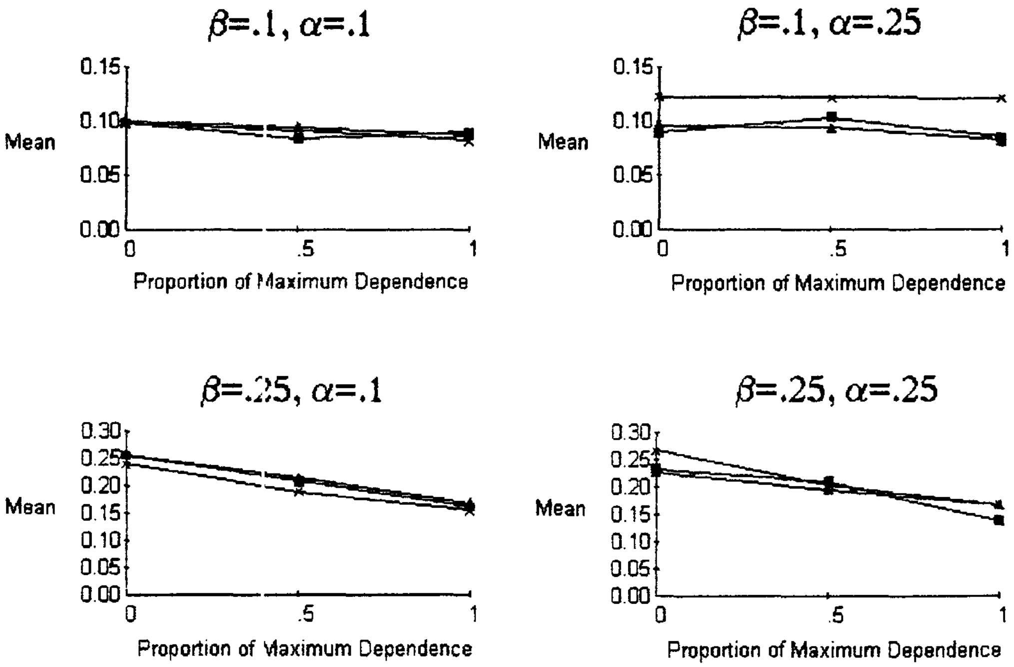
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 2b

Mean of Estimates of false negative rate for observer 1 (β_1) in the case of three observers, with dependence between observers 1 and 2 with respect to false negative rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

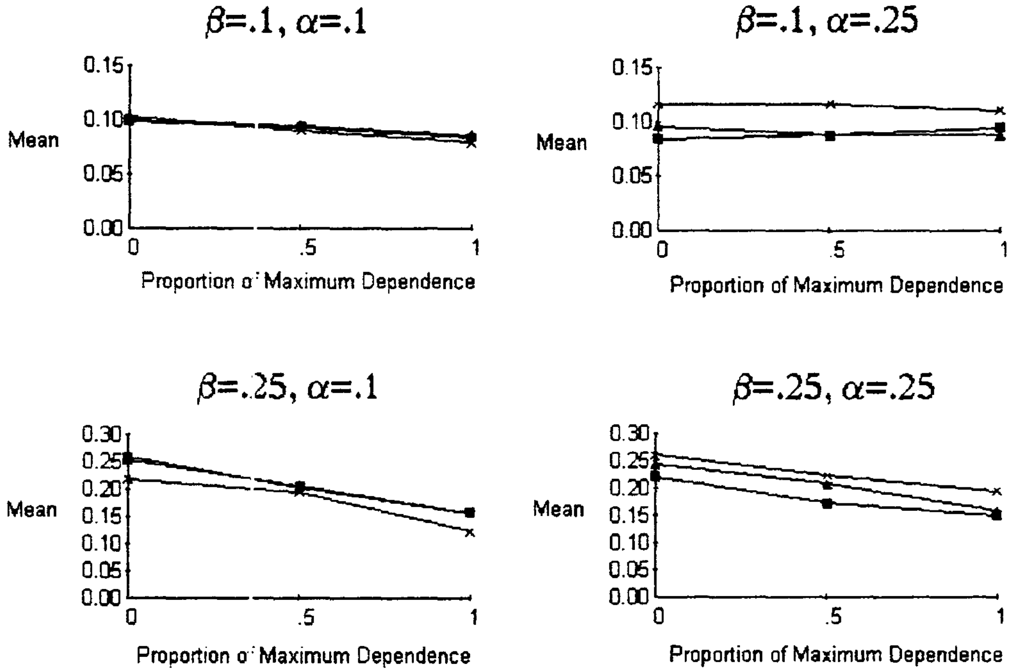
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 2c

Mean of Estimates of false negative rate for observer 2 (β_2) in the case of three observers, with dependence between observers 1 and 2 with respect to false negative rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

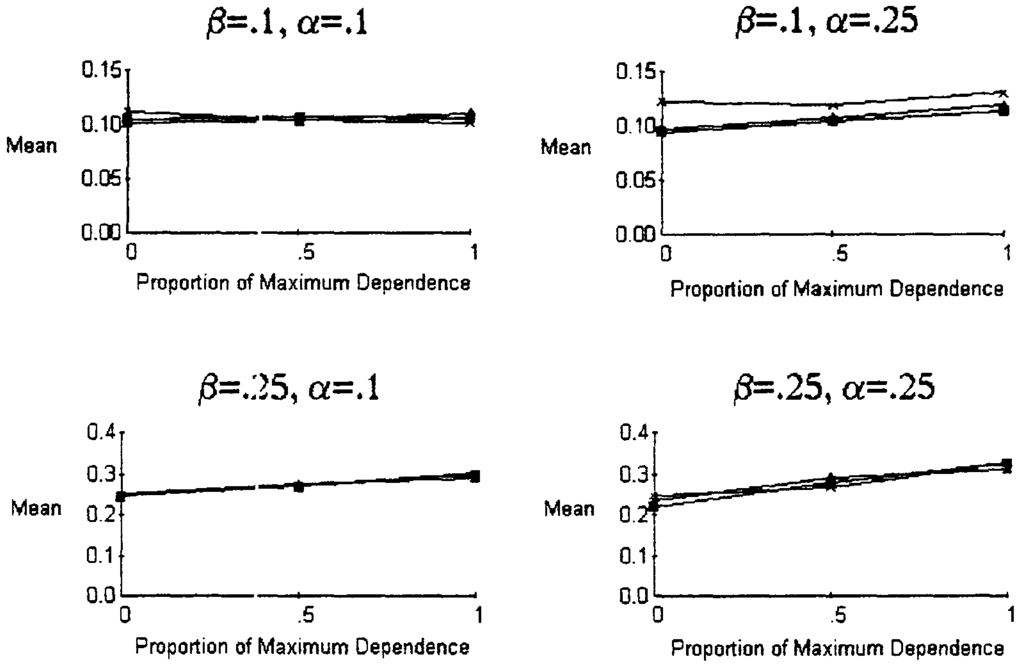
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 2d

Mean of Estimates of false negative rate for observer 3 (β_3) in the case of three observers, with dependence between observers 1 and 2 with respect to false negative rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

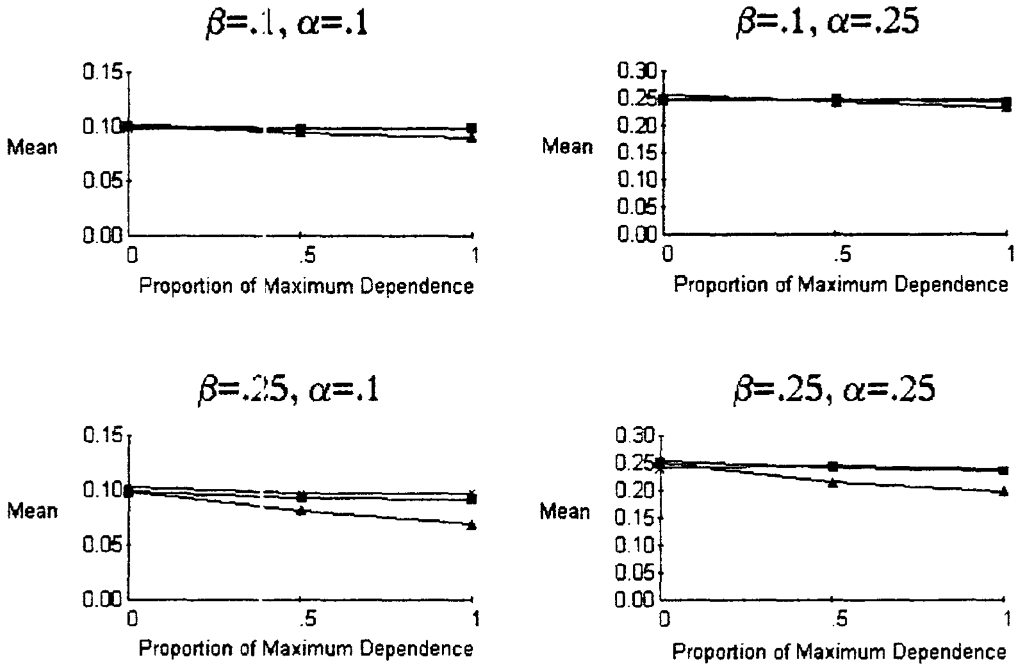
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 2e

Mean of Estimates of false positive rate for observer 1 (α_1) in the case of three observers, with dependence between observers 1 and 2 with respect to false negative rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

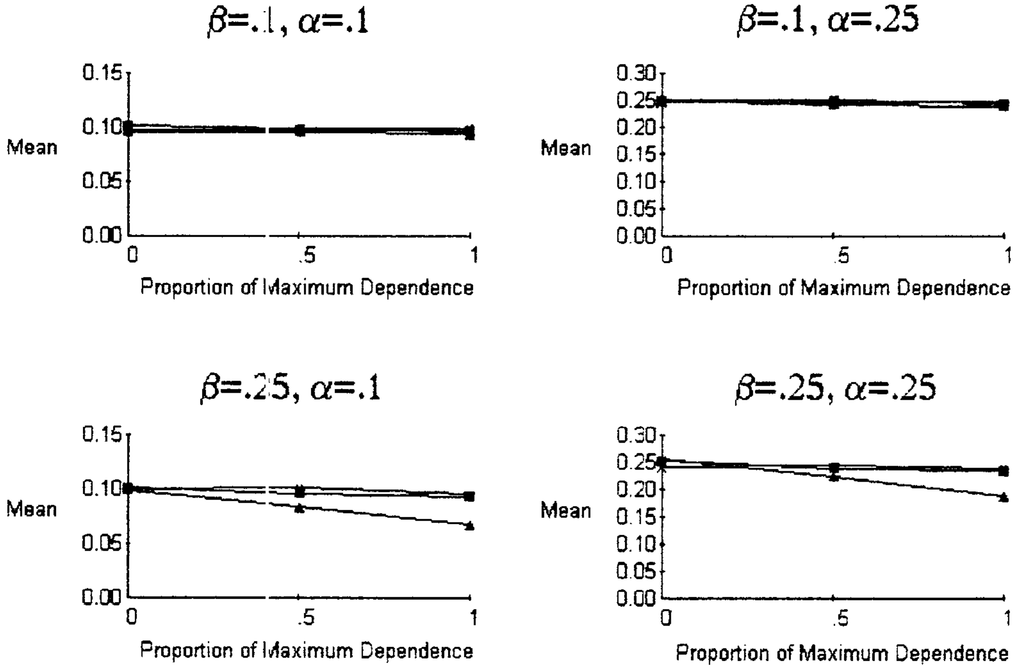
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 2f

Mean of Estimates of false positive rate for observer 2 (α_2) in the case of three observers, with dependence between observers 1 and 2 with respect to false negative rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

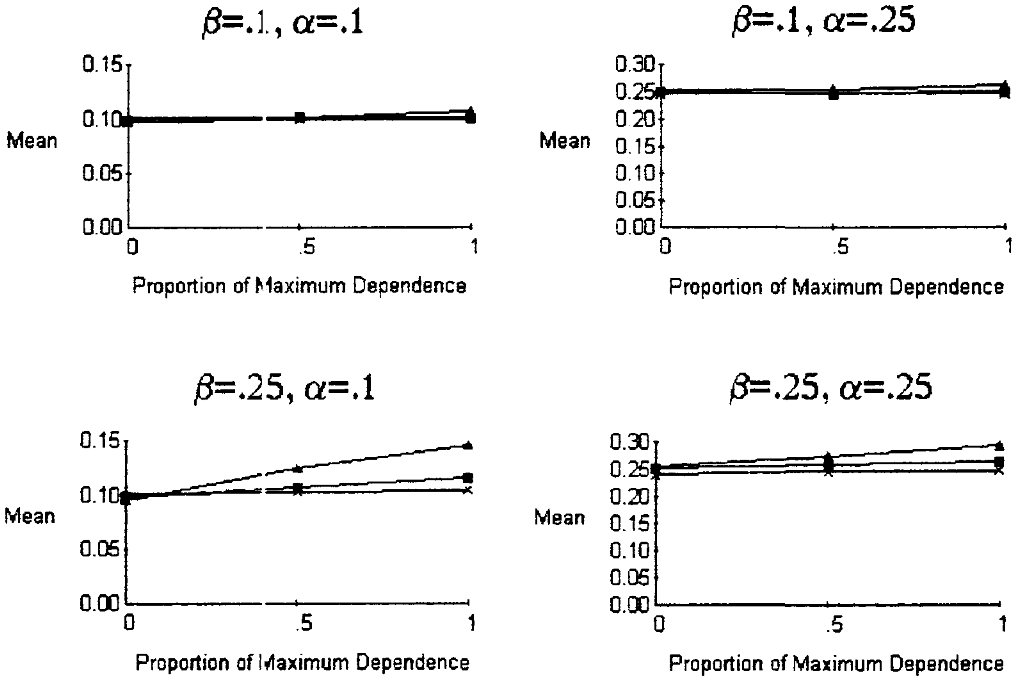
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 2g

Mean of Estimates of false positive rate for observer 3 (α_3) in the case of three observers, with dependence between observers 1 and 2 with respect to false negative rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.
× represents true prevalence (θ) = .05.
■ represents true prevalence (θ) = .15.
▲ represents true prevalence (θ) = .40.

3.1.3 RUN 3 - Dependence between two observers with respect to false positive and false negative rates - Three Observer Case

In this simulation there was a positive dependence between observers 1 and 2 with respect to errors classifying both positive and negative individuals. Tables 3a, 3b, and 3c show the results for the three different prevalence values which will be interpreted below. Graphs of the means of the parameter estimates versus dependence level can be found in Figures 3a to 3g.

OBSERVATIONS

Effect on prevalence estimates

In this simulation run it was found that when FPR is high at .25, dependence causes a low prevalence to be overestimated. For example, Table 3a shows that when FNR is .1 and FPR is .25 the estimate of the prevalence which is truly .05 increases to .11 at one half of maximum dependence and to .18 at maximum dependence. Table 3b which is for a true prevalence of .15 also demonstrates this effect. In Figure 3a, the graphs to the right, with FPR=.25, show the prevalence estimates increasing for low true prevalence values when compared with the other graphs which show no effect.

Effect on false negative rate estimates

The false negative rates for the two dependent observers are seriously underestimated. The third observer's false negative rate is overestimated. These effects become more severe with lower prevalence. The bias is strongest when the FPR is high at .25.

Effect on false positive rate estimates

The effects on the false positive rates are smaller than those on the false negative rates. The false positive rates for the two dependent observers are underestimated. At low prevalences of .05 and .15 this underestimation is only strong when the FPR is high at .25. The third observer's false positive rate is slightly overestimated with the strongest effect occurring when the prevalence is at .40.

EXPLANATIONS

The results from this run can be best understood after observing the results from Run 1 and Run 2. This run can be considered a combination of Run 1 and Run 2 and therefore shares the main effects from these runs. As in Run 1 the prevalence becomes overestimated with dependence because negative individuals are being simultaneously misclassified by observers 1 and 2 as positive. This leads to an overabundance of positive classifications, an increased prevalence estimate, and a decreased false positive rate estimate for the dependent observers. The false negative rate for the dependent observers is also underestimated as in Run 1. The influence of Run 2, or more precisely of the dependence with respect to false negative classifications, can be seen in the false positive rate estimates for the independent third observer which become overestimates at a high prevalence of .4 and at a high FPR of .25.

IMPLICATIONS

As discussed in Run 1, SnNout tests are particularly dangerous if dependence between two of three observers is suspected. The danger increases as the true prevalence decreases and the true error rates increase. The Latent procedure will produce estimates of prevalence that are too high and estimates of dependent observer error rates that are far too low if there is dependence between two tests or observers. In addition, the estimated error rates for the independent third observer will increase.

TABLE 3a

Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate

True Prevalence = .05

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	4.95 (0.71)	9.56 (6.19)	10.47 (5.70)	8.52 (5.54)	9.86 (0.76)	10.22 (0.69)	9.96 (0.77)
	0.5	5.31 (0.62)	**1.95 (3.26)**	**1.76 (2.62)**	**18.67 (5.05)**	9.39 (0.73)	9.30 (0.72)	10.45 (0.78)
	1	**6.35 (0.58)**	**0.00 (0.00)**	**0.00 (0.02)**	**24.72 (4.56)**	8.33 (0.63)	8.39 (0.57)	10.05 (0.70)
β =.1 α =.25	0	5.92 (2.24)	**12.13 (10.88)**	**12.39 (11.04)**	**12.47 (10.66)**	24.59 (1.35)	24.68 (1.29)	24.73 (1.42)
	0.5	**11.19 (0.88)**	**0.03 (0.46)**	**0.01 (0.18)**	**41.47 (3.82)**	**19.13 (1.03)**	**19.16 (1.04)**	24.46 (1.02)
	1	**17.56 (0.91)**	**0.00 (0.00)**	**0.00 (0.00)**	**46.98 (2.97)**	**12.91 (0.88)**	**12.87 (0.86)**	22.94 (1.00)
β =.25 α =.1	0	4.88 (1.07)	23.07 (7.81)	24.16 (11.56)	24.06 (8.43)	9.84 (0.96)	10.03 (0.81)	10.04 (0.82)
	0.5	4.68 (0.59)	**7.64 (7.08)**	**5.14 (5.39)**	**36.30 (6.30)**	9.41 (0.70)	9.21 (0.80)	10.52 (0.84)
	1	5.24 (0.56)	**0.00 (0.00)**	**0.00 (0.00)**	**37.94 (5.45)**	8.28 (0.59)	8.39 (0.61)	10.59 (0.65)
β =.25 α =.25	0	**8.38 (7.49)**	27.77 (18.39)	26.04 (18.85)	26.89 (18.96)	24.25 (2.15)	24.16 (2.16)	24.28 (2.28)
	0.5	**10.11 (0.87)**	**0.11 (0.72)**	**0.04 (0.53)**	**49.30 (4.24)**	**19.32 (1.00)**	**19.33 (1.00)**	24.93 (1.18)
	1	**16.79 (0.96)**	**0.00 (0.00)**	**0.00 (0.00)**	**52.45 (2.79)**	**12.93 (0.87)**	**12.88 (0.84)**	23.39 (0.94)

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.

β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

Means with 20% or greater bias are displayed in bold.

TABLE 3b

Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate

True Prevalence = .15

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	15.09 (1.05)	10.72 (2.64)	10.11 (2.97)	8.93 (2.49)	10.14 (0.92)	10.46 (0.78)	10.09 (0.86)
	0.5	14.82 (1.03)	**6.01 (2.90)**	**6.42 (2.36)**	**12.73 (2.82)**	9.40 (0.93)	9.48 (0.89)	10.72 (0.91)
	1	15.00 (0.82)	**2.32 (2.20)**	**3.10 (2.03)**	**16.06 (2.32)**	8.86 (0.77)	8.62 (0.74)	10.95 (0.65)
β =.1 α =.25	0	15.09 (2.23)	9.91 (4.94)	10.44 (4.87)	9.98 (4.65)	24.98 (1.37)	24.75 (1.34)	24.94 (1.65)
	0.5	**18.55 (1.10)**	**0.62 (1.41)**	**0.36 (1.06)**	**26.67 (2.82)**	**19.93 (1.17)**	20.08 (1.29)	25.99 (1.09)
	1	**24.07 (0.82)**	**0.00 (0.00)**	**0.00 (0.00)**	**34.45 (2.53)**	**13.74 (0.80)**	**13.89 (0.98)**	24.81 (0.88)
β =.25 α =.1	0	15.55 (1.77)	25.38 (4.05)	27.04 (4.43)	26.18 (6.05)	9.73 (0.96)	9.86 (1.05)	9.72 (1.24)
	0.5	13.90 (1.00)	**14.73 (3.51)**	**15.87 (2.94)**	29.71 (3.98)	9.18 (1.13)	9.24 (0.93)	11.47 (0.86)
	1	13.02 (0.95)	**5.81 (3.68)**	**5.33 (3.04)**	**32.41 (3.67)**	8.75 (0.85)	8.72 (0.76)	**12.62 (0.83)**
β =.25 α =.25	0	14.95 (4.56)	25.17 (7.62)	22.24 (9.78)	21.79 (9.12)	24.88 (1.82)	25.16 (1.79)	24.85 (2.09)
	0.5	15.57 (1.30)	**1.44 (2.34)**	**2.40 (3.70)**	**39.32 (3.09)**	20.68 (1.39)	20.55 (1.04)	27.28 (0.96)
	1	**21.76 (1.10)**	**0.00 (0.00)**	**0.00 (0.00)**	**44.38 (2.71)**	**14.04 (1.30)**	**14.27 (1.15)**	26.49 (1.12)

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.

β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

Means with 20% or greater bias are displayed in bold.

TABLE 3c

Means ar d standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate

True Prevalence = .40

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	40.40 (1.14)	9.80 (1.33)	10.24 (1.12)	9.96 (1.15)	9.73 (0.89)	10.00 (1.17)	10.04 (1.02)
	0.5	39.89 (1.23)	8.40 (1.34)	8.65 (1.07)	11.49 (1.34)	8.86 (0.89)	9.08 (1.08)	10.64 (1.11)
	1	40.14 (1.32)	**7.14 (1.35)**	**7.49 (1.46)**	**13.09 (1.22)**	8.16 (0.92)	**7.97 (1.03)**	**12.00 (0.96)**
β =.1 α =.25	0	40.10 (1.84)	9.77 (1.94)	10.09 (2.03)	9.51 (1.88)	25.03 (1.33)	25.07 (1.71)	25.13 (1.95)
	0.5	40.75 (2.04)	**5.66 (2.45)**	**4.89 (1.98)**	**15.71 (1.74)**	20.68 (2.01)	20.30 (1.96)	28.42 (1.79)
	1	42.11 (1.83)	**1.56 (1.25)**	**1.81 (1.44)**	**19.94 (1.52)**	**16.07 (1.75)**	**16.27 (1.17)**	29.72 (1.60)
β =.25 α =.1	0	39.93 (1.85)	24.58 (2.24)	25.16 (1.63)	24.56 (1.73)	9.88 (1.40)	10.31 (1.50)	9.92 (1.56)
	0.5	38.77 (1.56)	**19.65 (2.12)**	**18.95 (1.62)**	28.07 (2.58)	**7.99 (1.00)**	**7.83 (1.29)**	**13.18 (1.06)**
	1	38.01 (1.30)	**13.72 (1.81)**	**14.08 (1.94)**	**30.07 (2.11)**	**5.92 (1.07)**	**5.95 (1.18)**	**15.99 (0.88)**
β =.25 α =.25	0	41.05 (6.25)	25.88 (4.74)	24.85 (4.71)	25.21 (3.90)	24.99 (2.59)	24.21 (3.19)	24.44 (3.04)
	0.5	40.69 (3.20)	**15.12 (3.10)**	**15.46 (2.94)**	**32.92 (1.96)**	**17.58 (2.34)**	**18.27 (1.99)**	29.83 (2.10)
	1	39.00 (2.17)	**6.35 (2.45)**	**5.80 (2.52)**	**36.00 (1.68)**	**13.67 (1.79)**	**13.75 (1.98)**	**32.67 (1.67)**

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.

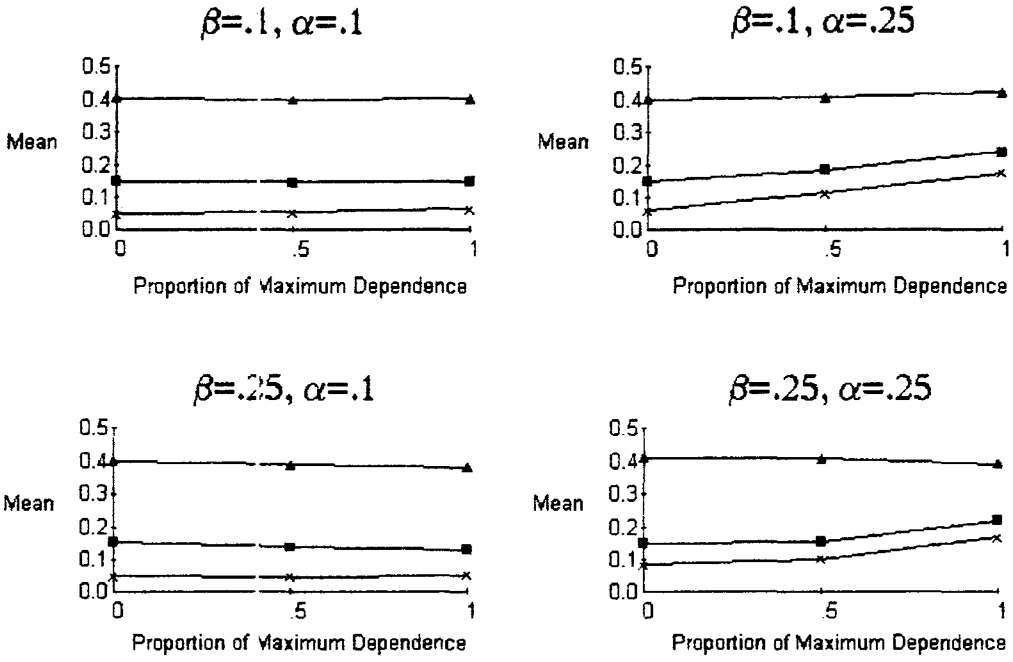
β represents the true false negat ve rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

Means with 20% or greater bias are displayed in bold.

Figure 3a

Mean of Estimates of Prevalence (θ) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

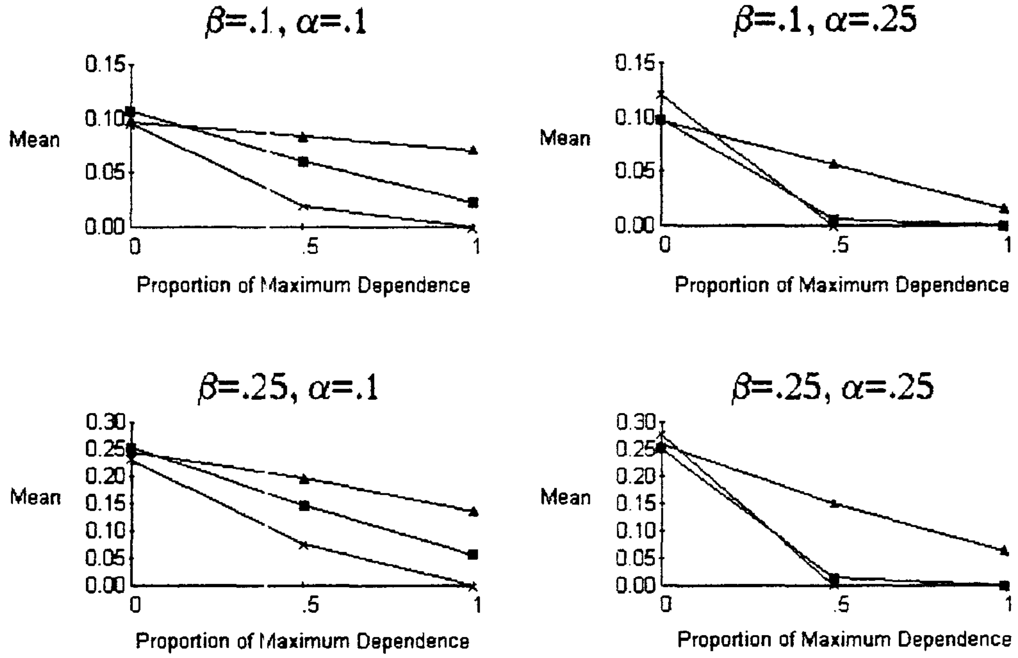
× represents true prevalence (θ) = .05

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 3b

Mean of Estimates of false negative rate for observer 1 (β_1) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

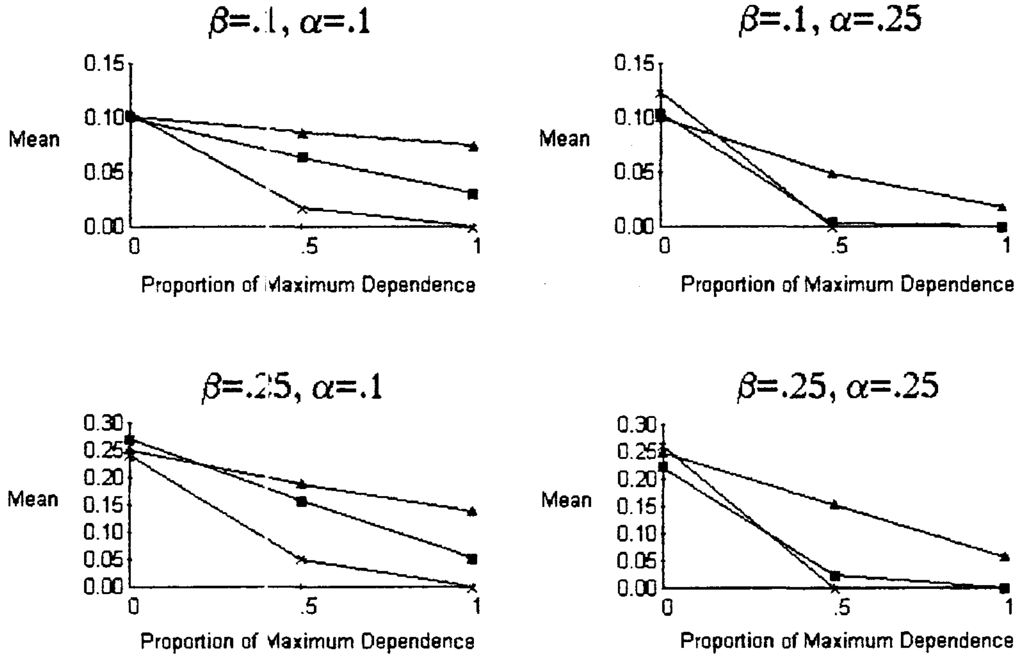
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 3c

Mean of Estimates of false negative rate for observer 2 (β_2) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

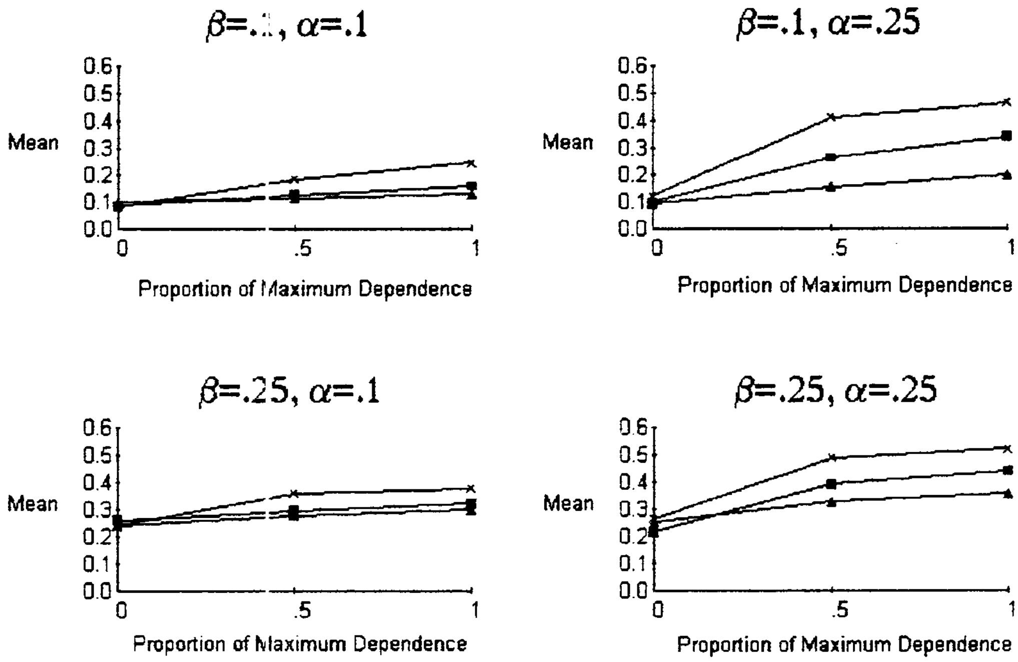
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 3d

Mean of Estimates of false negative rate for observer 3 (β_3) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

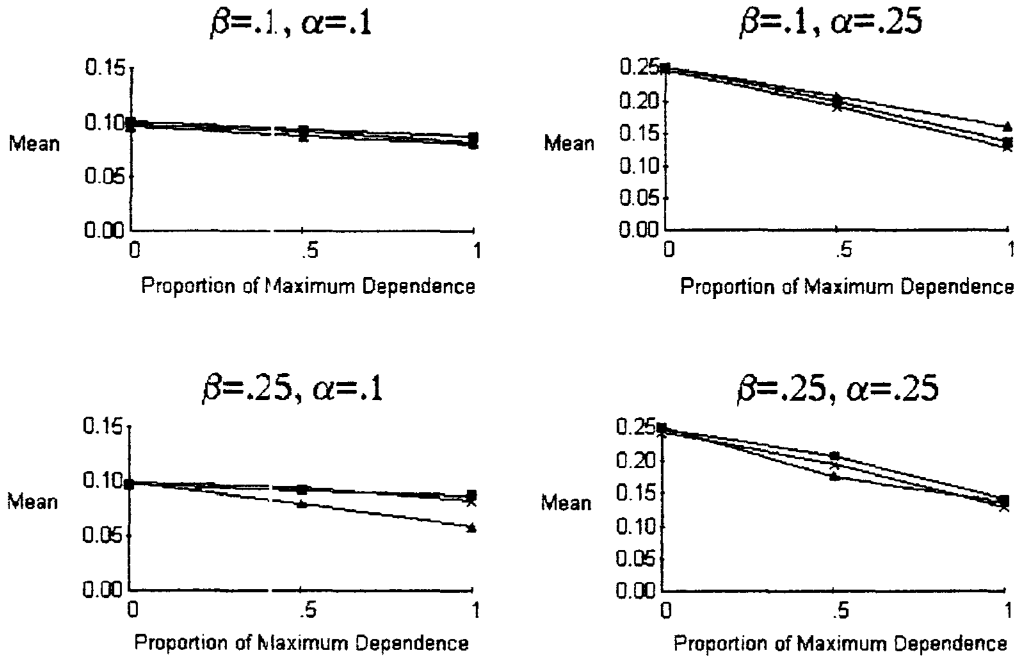
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 3e

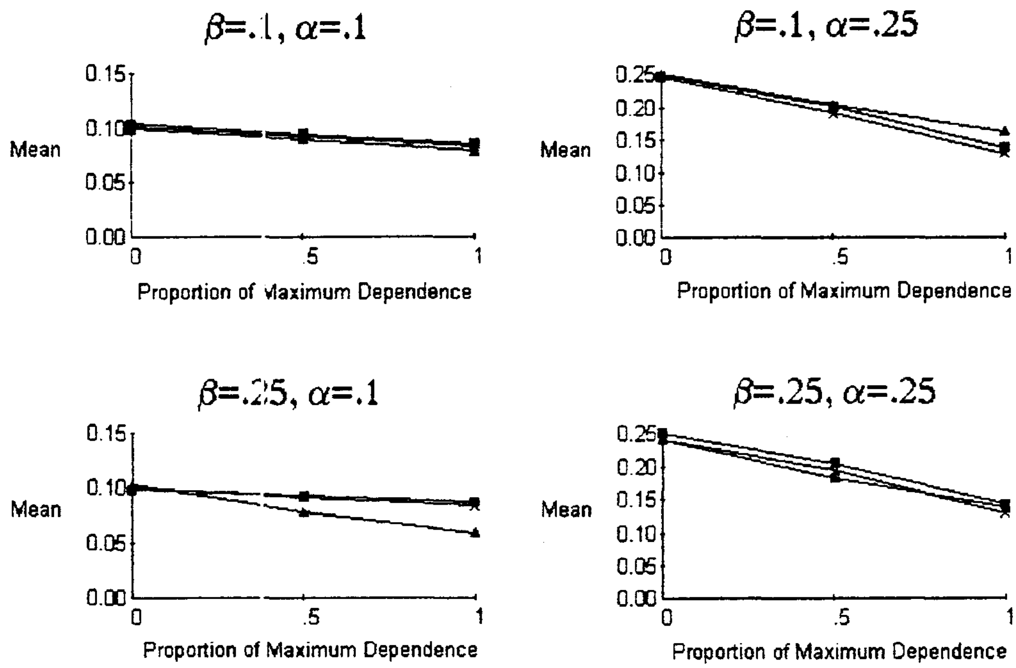
Mean of Estimates of false positive rate for observer 1 (α_1) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.
× represents true prevalence (θ) = .05.
■ represents true prevalence (θ) = .15.
▲ represents true prevalence (θ) = .40.

Figure 3f

Mean of Estimates of false positive rate for observer 2 (α_2) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

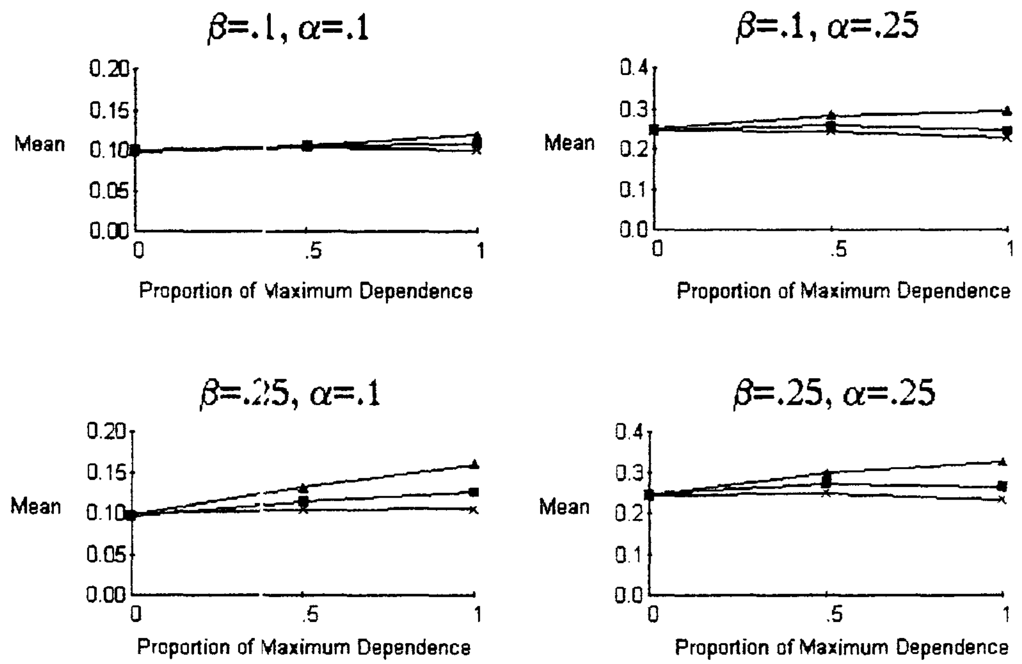
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 3g

Mean of Estimates of false positive rate for observer 3 (α_3) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

3.1.4 RUN 4 - Dependence between two accurate observers - Three observer case

This run only includes two observers who are fairly accurate with false negative and false positive error rates of .1 and a third observer who is less accurate with error rates of .25. There is dependence between the two more accurate observers when classifying all individuals. This situation is run for prevalences of .05, .15, and .40 and the results can be found in Tables 4a, 4b and 4c, respectively. Figures 4a to 4g show graphs of the mean of the estimates as the dependence increases.

OBSERVATIONS

Effect on prevalence estimates

The true prevalence of .05 is overestimated as can be seen in Table 4a. At one half of maximum dependence the mean of the estimates is .07 and at full dependence the mean is .09. All other prevalences show no effect of the dependence in this run.

Effect on false negative rate estimates

The two dependent observers have false negative rates seriously underestimated. The effect is strongest for low prevalences as was the case in previous runs. There is an inverse effect on the third observer's false negative rate. It is more severely overestimated for low prevalence values.

Effect on false positive rate estimates

There is an underestimation of the two dependent observers' false positive rates. The amount of bias is similar for all prevalences. The third observer's false positive rate is not affected by the dependence.

EXPLANATIONS

As seen in earlier runs dependence between two observers makes those two observers appear more accurate than they truly are. As a result the third observer appears less accurate. A lower prevalence will cause more serious bias in the estimates due to the dependence with respect to false positive classifications as explained in Run 1.

IMPLICATIONS

This run reflects practical situations in diagnostic test evaluation. For example, consider two fairly accurate diagnostic tests with similar components that cause their misclassifications to be dependent while a third test is less accurate and remains independent from the others. The results show that the prevalence estimate from the Latent procedure may be biased if the true prevalence is low. In addition, the estimated error rates of the dependent tests will appear too low and the third test will appear even less accurate. Again, the most dangerous situation is when the true prevalence is low.

TABLE 4a

Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative Rate

True Prevalence = .05

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
$\beta_1, \beta_2 = .1$	0	5.36 (1.18)	11.43 (9.91)	12.01 (10.30)	24.89 (6.51)	9.86 (0.95)	9.86 (0.95)	24.81 (1.12)
$\beta_3 = .25$ $\alpha_1, \alpha_2 = .1$	0.5	**6.71 (0.62)**	**0.24 (1.20)**	**0.20 (1.11)**	**36.15 (4.42)**	**7.81 (0.67)**	**7.83 (0.64)**	24.85 (1.03)
$\alpha_3 = .25$	1	**9.26 (0.63)**	**0.00 (0.00)**	**0.00 (0.00)**	**40.93 (3.87)**	**5.19 (0.54)**	**5.26 (0.51)**	24.35 (1.00)

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.
Means with 20% or greater bias are displayed in bold.

TABLE 4b

Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative Rate

True Prevalence = .15

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
$\beta_1, \beta_2 = .1$	0	15.50 (1.53)	9.65 (5.65)	11.20 (4.31)	25.30 (2.49)	9.77 (0.88)	9.94 (1.00)	24.93 (1.39)
$\beta_3 = .25$ $\alpha_1, \alpha_2 = .1$	0.5	15.29 (0.97)	**1.39 (1.85)**	**1.27 (2.02)**	**30.19 (3.17)**	8.31 (0.85)	8.42 (0.86)	25.80 (1.24)
$\alpha_3 = .25$	1	17.18 (0.85)	**0.00 (0.00)**	**0.00 (0.00)**	**32.99 (2.37)**	**5.77 (0.59)**	**5.78 (0.66)**	25.31 (1.01)

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.
Means with 20% or greater bias are displayed in bold.

TABLE 4c

Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative Rate

True Prevalence = .40

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
$\beta_1, \beta_2 = .1$	0	39.75 (1.49)	10.38 (2.04)	9.52 (1.83)	25.01 (2.07)	9.92 (1.57)	10.03 (1.23)	25.14 (1.33)
$\beta_3 = .25$ $\alpha_1, \alpha_2 = .1$	0.5	40.22 (1.35)	**5.96 (1.62)**	**7.11 (2.01)**	26.85 (2.03)	**7.40 (0.99)**	**7.82 (1.51)**	24.21 (1.70)
$\alpha_3 = .25$	1	39.72 (1.77)	**3.29 (1.45)**	**3.41 (1.37)**	28.14 (1.55)	**5.74 (1.19)**	**5.80 (0.90)**	27.11 (1.63)

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.
Means with 20% or greater bias are displayed in bold.

Figure 4a

Mean of Estimates of Prevalence (θ)
in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .1, \beta_3 = .25; \quad \alpha_1, \alpha_2 = .1, \alpha_3 = .25$

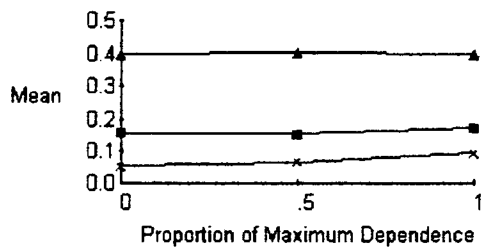
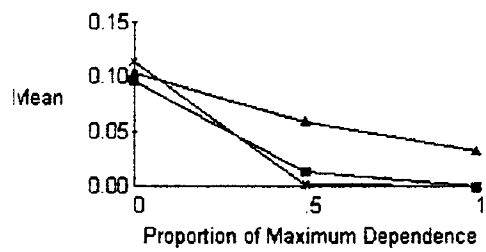


Figure 4b

Mean of Estimates of false negative rate for observer 1 (β_1)
in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .1, \beta_3 = .25; \quad \alpha_1, \alpha_2 = .1, \alpha_3 = .25$



× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 4c

Mean of Estimates of false negative rate for observer 2 (β_2) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .1, \beta_3 = .25; \quad \alpha_1, \alpha_2 = .1, \alpha_3 = .25$

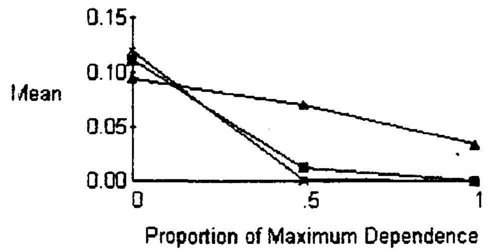
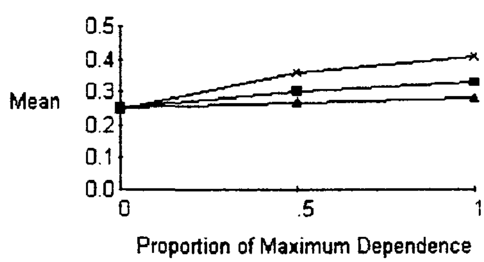


Figure 4d

Mean of Estimates of false negative rate for observer 3 (β_3) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .1, \beta_3 = .25; \quad \alpha_1, \alpha_2 = .1, \alpha_3 = .25$



× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 4e

Mean of Estimates of false positive rate for observer 1 (α_1) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .1, \beta_3 = .25; \quad \alpha_1, \alpha_2 = .1, \alpha_3 = .25$

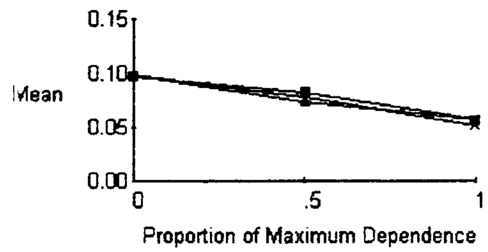
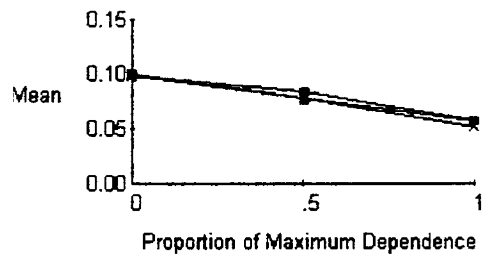


Figure 4f

Mean of Estimates of false positive rate for observer 2 (α_2) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .1, \beta_3 = .25; \quad \alpha_1, \alpha_2 = .1, \alpha_3 = .25$



× represents true prevalence (θ) = .05.

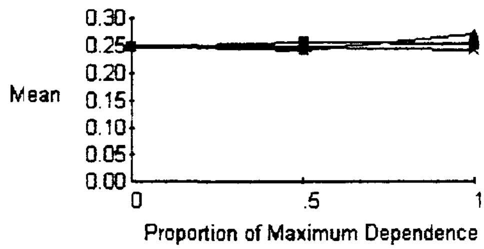
■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 4g

Mean of Estimates of false positive rate for observer 3 (α_3) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .1, \beta_3 = .25; \quad \alpha_1, \alpha_2 = .1, \alpha_3 = .25$



× represents true prevalence (θ) = .05.

▪ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

3.1.5 RUN 5 - Dependence between two inaccurate observers - Three observer case

This run is similar to Run 4, however, the two dependent observers are less accurate, with error rates of .25, than the independent observer who has error rates of .1. Tables 5a, 5b, and 5c show the results for prevalences of .05, .15, and .40, respectively. The graphs of the estimates against dependence can be found in Figures 5a to 5g.

OBSERVATIONS

Effect on prevalence estimates

A bias of more than 20% is shown only in Table 5a when the true prevalence is .05. At maximum dependence the mean of the prevalence estimates reaches .086. The dependence does not seriously bias any other prevalence estimates.

Effect on false negative rate estimates

The false negative rates for the two dependent observers are underestimated. Prevalence has a strong effect on the bias here as can be seen in Figures 5b and 5c. The lower the prevalence the more severely biased the estimates. There is an inverse trend for the false negative rate of the third observer where a lower prevalence leads to a larger overestimation of the rate.

Effect on false positive rate estimates

There is a much less severe effect of dependence on the false positive rates. A weak trend towards underestimation of the two dependent observers' rates can be seen in Figures 5e and 5f. Figure 5g shows that the third observer's false positive rate is overestimated at prevalences of .15 and .40. The higher the prevalence the more severe the bias in the false positive rates.

EXPLANATIONS

The two dependent observers appear more accurate due to the dependence while the third observer appears less accurate. A lower true prevalence causes more severe bias in the prevalence estimates and the false negative rate estimates through the dependence with respect to false positive classifications as seen in Run 1. A higher prevalence causes more severe bias in the false positive rate estimates through the false negative classification dependence as seen in Run 2.

IMPLICATIONS

This run reflects situations where there are two observers who perhaps had the same inadequate or incomplete training causing them to commit the same errors while a third observer had different, perhaps superior, training and commits error less often, more randomly, and independent of the others. This is a dangerous situation because the dependent observers have high error rates. A low prevalence will result in seriously biased false negative rate estimates while a high prevalence will lead to biased false positive rate estimates. At full dependence, the less accurate dependent observers appear more accurate when classifying true positives than the truly accurate observer In the last case the dependence amplified the differences between the observers false positive rates. In this case conclusions about which observer is the most sensitive could be completely wrong in the presence of dependence.

TABLE 5a

Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative Rate

True Prevalence = .05

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
$\beta_1, \beta_2 = .25$	0	5.76 (2.38)	24.68 (11.28)	24.81 (11.32)	13.93 (15.15)	24.72 (1.28)	24.75 (1.28)	9.83 (1.34)
$\beta_3 = .1$ $\alpha_1, \alpha_2 = .25$	0.5	5.70 (0.95)	**1.33 (3.23)**	**1.37 (3.34)**	**36.07 (9.01)**	23.17 (1.14)	23.19 (1.15)	11.01 (0.75)
$\alpha_3 = .1$	1	**8.60 (0.82)**	**0.00 (0.00)**	**0.00 (0.00)**	**46.44 (5.02)**	20.68 (1.04)	20.65 (1.04)	10.28 (0.72)

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.
Means with 20% or greater bias are displayed in bold.

TABLE 5b

Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative Rate

True Prevalence = .15

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
$\beta_1, \beta_2 = .25$	0	15.20 (2.16)	25.03 (4.22)	24.80 (4.69)	10.67 (7.74)	24.90 (1.33)	24.92 (1.39)	9.98 (1.40)
$\beta_3 = .1$ $\alpha_1, \alpha_2 = .25$	0.5	14.63 (1.39)	**13.49 (4.55)**	**12.55 (4.88)**	**23.12 (5.91)**	23.17 (1.50)	23.27 (1.33)	**12.61 (1.25)**
$\alpha_3 = .1$	1	14.27 (1.34)	**2.79 (3.16)**	**2.58 (3.05)**	**29.87 (3.86)**	21.63 (1.25)	21.72 (1.16)	**13.99 (0.91)**

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.
Means with 20% or greater bias are displayed in bold.

TABLE 5c

Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative Rate

True Prevalence = .40

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
$\beta_1, \beta_2 = .25$	0	40.00 (2.31)	24.88 (2.38)	24.90 (2.56)	9.61 (4.53)	25.14 (2.23)	25.55 (2.36)	10.11 (1.98)
$\beta_3 = .1$ $\alpha_1, \alpha_2 = .25$	0.5	40.44 (1.68)	20.68 (2.09)	20.59 (2.16)	**16.62 (2.62)**	21.55 (1.96)	21.93 (1.60)	**14.04 (1.61)**
$\alpha_3 = .1$	1	40.52 (2.02)	**17.13 (1.95)**	**16.69 (2.50)**	**21.19 (2.01)**	**19.21 (1.79)**	**19.12 (1.45)**	**17.06 (1.29)**

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.
Means with 20% or greater bias are displayed in bold.

Figure 5a

Mean of Estimates of Prevalence (θ)
in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .25, \beta_3 = .1; \quad \alpha_1, \alpha_2 = .25, \alpha_3 = .1$

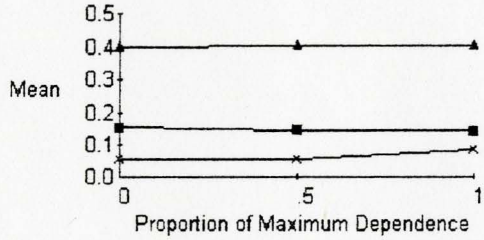
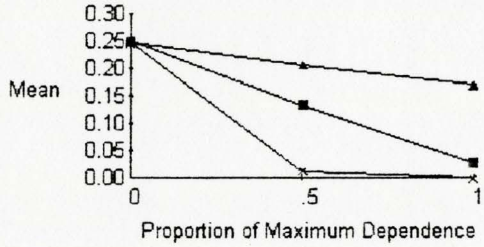


Figure 5b

Mean of Estimates of false negative rate for observer 1 (β_1)
in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .25, \beta_3 = .1; \quad \alpha_1, \alpha_2 = .25, \alpha_3 = .1$



× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 5c

Mean of Estimates of false negative rate for observer 2 (β_2) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .25, \ \beta_3 = .1; \quad \alpha_1, \alpha_2 = .25, \ \alpha_3 = .1$

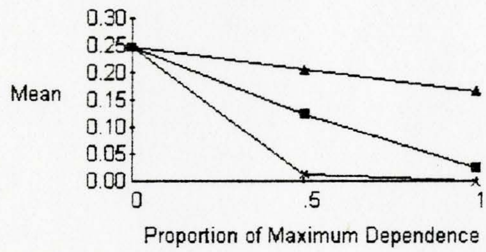
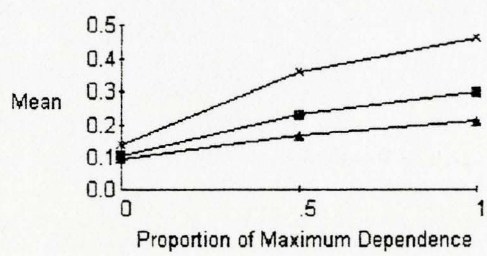


Figure 5d

Mean of Estimates of false negative rate for observer 3 (β_3) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .25, \ \beta_3 = .1; \quad \alpha_1, \alpha_2 = .25, \ \alpha_3 = .1$



× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 5e

Mean of Estimates of false positive rate for observer 1 (α_1) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .25, \ \beta_3 = .1; \quad \alpha_1, \alpha_2 = .25, \ \alpha_3 = .1$

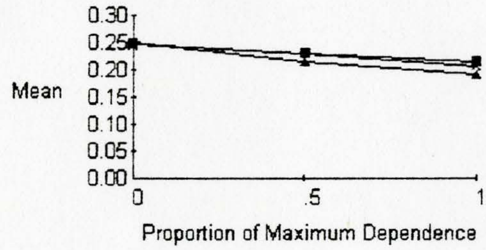
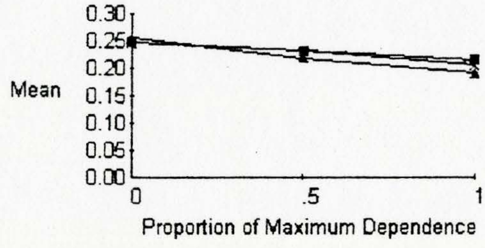


Figure 5f

Mean of Estimates of false positive rate for observer 2 (α_2) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .25, \ \beta_3 = .1; \quad \alpha_1, \alpha_2 = .25, \ \alpha_3 = .1$



× represents true prevalence (θ) = .05 .

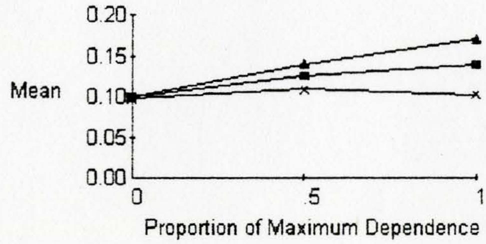
■ represents true prevalence (θ) = .15 .

▲ represents true prevalence (θ) = .40 .

Figure 5g

Mean of Estimates of false positive rate for observer 3 (α_3) in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .25, \ \beta_3 = .1; \quad \alpha_1, \alpha_2 = .25, \alpha_3 = .1$



× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

3.1.6 RUN 6 - Dependence with respect to false positive and false negative rates

In this simulation there was a positive dependence between all pairs of observers when classifying both true positive and true negative individuals. Tables 6a, 6b and 6c show the results for prevalences of .05, .15, and .40. The graphs of the means of the estimates can be found in Figure 6a to 6g.

OBSERVATIONS

Effect on prevalence estimates

Dependence causes prevalences of .05 and .15 to be overestimated when FPR = .25. For example, in Table 6a it can be seen that when FNR = .1 and FPR = .25 and the true prevalence is .05, the mean of the prevalence estimates is .12 at half maximum and .18 at maximum dependence. Figure 6a also shows these results. There is no apparent effect on the estimates of prevalence when the true prevalence is .40.

Effect on false negative rate estimates

There is symmetry here in that all three observers' rates will react to the dependence in the same way. All false negative rates are underestimated due to the pairwise dependence. The most serious effects occur when the prevalence is low and the FPR is high at .25.

Effect on false positive rate estimates

The false positive rates for all three observers are most seriously underestimated when FPR = .25. There appears to be no effect of prevalence on the amount of bias except when FNR = .25 and FPR = .1 where there is underestimation by more than 20% only when the prevalence is .40. Figures 6e to 6g clearly show the decrease in the means of the estimates as the dependence increases.

EXPLANATIONS

A positive dependence between all three observers will result in underestimation of all observers' error rates. There is more agreement between all observers than in the independent case which leads to the bias. In addition a low prevalence is overestimated when the FPR is high since there are more positive classifications occurring simultaneously due to the dependence.

IMPLICATIONS

If there is dependence between all three observers, or if there are three tests which rely on some of the same components, then the latent procedure will produce error rate estimates which may be very misleading. All error rates will be underestimated leading to a false sense of confidence for those interested in the accuracy of the observers or tests. In this run again, the most dangerous situations are when the false positive rates are high and the prevalence is low.

TABLE 6a
Means and standard deviations of parameter estimates in the case of three observers, with dependence between all pairs of observers with respect to false positive and false negative rates

True Prevalence = .05

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	5.07 (0.73)	10.11 (6.34)	9.60 (6.25)	10.22 (6.13)	9.92 (0.85)	9.86 (0.69)	10.04 (0.73)
	0.5	5.44 (0.58)	**2.37 (2.80)**	**2.11 (2.63)**	**2.24 (2.78)**	9.25 (0.70)	9.40 (0.59)	9.39 (0.62)
	1	**6.46 (0.61)**	**0.09 (0.58)**	**0.11 (0.40)**	**0.07 (0.38)**	8.14 (0.54)	8.08 (0.67)	8.08 (0.65)
β =.1 α =.25	0	5.75 (2.04)	12.26 (10.75)	11.58 (11.04)	11.43 (10.73)	24.75 (1.34)	24.76 (1.39)	24.77 (1.31)
	0.5	**11.83 (0.92)**	**1.60 (2.27)**	**1.51 (2.14)**	**1.61 (2.23)**	**18.84 (1.07)**	**18.85 (1.03)**	**18.77 (1.04)**
	1	**18.46 (0.87)**	**0.00 (0.00)**	**0.00 (0.00)**	**0.00 (0.00)**	**11.98 (0.81)**	**12.01 (0.83)**	**11.98 (0.79)**
β =.25 α =.1	0	4.99 (1.13)	24.09 (9.18)	26.87 (8.51)	22.20 (8.32)	10.08 (0.92)	10.02 (0.72)	9.69 (0.84)
	0.5	4.70 (0.57)	**5.60 (5.20)**	**6.59 (4.47)**	**4.62 (4.90)**	9.42 (0.59)	9.46 (0.79)	9.48 (0.69)
	1	5.45 (0.56)	**0.36 (1.02)**	**0.21 (0.71)**	**0.65 (1.40)**	8.28 (0.62)	8.03 (0.79)	8.20 (0.81)
β =.25 α =.25	0	**6.88 (7.33)**	24.69 (18.61)	21.66 (19.57)	23.81 (17.89)	24.76 (2.01)	24.37 (2.22)	24.64 (2.12)
	0.5	**10.79 (0.92)**	**2.75 (3.24)**	**2.66 (3.11)**	**2.60 (2.83)**	**19.16 (0.93)**	**19.09 (1.11)**	**19.05 (1.07)**
	1	**17.47 (0.79)**	**0.00 (0.00)**	**0.00 (0.00)**	**0.00 (0.00)**	**12.13 (0.84)**	**12.14 (0.85)**	**12.11 (0.80)**

* Proportion of maximum dependence between all pairs of observers for both true negative and true positive subjects.
β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.
Means with 20% or greater bias are displayed in bold.

TABLE 6b

Means and standard deviations of parameter estimates in the case of three observers, with dependence between all pairs of observers with respect to false positive and false negative rates

True Prevalence = .15

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	14.96 (1.21)	10.46 (2.49)	9.16 (3.26)	10.06 (2.97)	9.99 (0.72)	10.00 (0.83)	9.87 (0.78)
	0.5	15.12 (0.89)	**6.45 (1.88)**	**6.53 (1.92)**	**6.42 (2.15)**	9.23 (0.69)	9.36 (0.82)	9.13 (0.87)
	1	15.09 (0.89)	**2.95 (1.85)**	**3.46 (1.90)**	**3.58 (1.57)**	8.84 (0.62)	8.91 (0.72)	8.81 (0.75)
β =.1 α =.25	0	15.47 (2.11)	9.91 (4.89)	10.58 (5.12)	9.95 (4.37)	24.97 (1.42)	25.05 (1.31)	24.89 (1.32)
	0.5	**20.76 (1.27)**	**4.22 (2.28)**	**4.17 (2.16)**	**4.60 (2.08)**	**19.07 (1.27)**	**19.09 (1.32)**	**18.76 (1.24)**
	1	**24.50 (0.94)**	**0.18 (0.35)**	**0.22 (0.48)**	**0.17 (0.37)**	**13.59 (1.01)**	**13.50 (0.91)**	**13.62 (0.85)**
β =.25 α =.1	0	15.38 (1.87)	24.99 (4.24)	25.73 (4.36)	25.01 (5.45)	9.99 (0.91)	9.83 (0.99)	9.85 (0.92)
	0.5	14.01 (1.01)	**15.09 (2.56)**	**15.27 (3.01)**	**14.40 (2.61)**	9.34 (0.89)	9.52 (0.68)	8.99 (0.82)
	1	13.48 (0.86)	**7.36 (2.21)**	**7.11 (1.77)**	**7.56 (2.04)**	8.24 (0.84)	8.34 (0.65)	8.28 (0.67)
β =.25 α =.25	0	16.48 (4.48)	26.57 (9.39)	25.89 (7.43)	25.92 (7.46)	24.29 (2.16)	24.51 (1.69)	24.69 (1.96)
	0.5	**18.23 (1.33)**	**8.87 (2.83)**	**7.45 (3.34)**	**8.67 (2.91)**	**19.52 (1.21)**	**19.48 (1.19)**	**19.17 (0.96)**
	1	**22.22 (1.10)**	**0.93 (1.12)**	**0.90 (0.98)**	**1.27 (1.12)**	**13.77 (0.96)**	**13.41 (1.17)**	**13.56 (0.85)**

* Proportion of maximum dependence between all pairs of observers for both true negative and true positive subjects.
β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.
Means with 20% or greater bias are displayed in bold.

TABLE 6c
Means and standard deviations of parameter estimates in the case of three observers, with dependence between all pairs of observers with respect to false positive and false negative rates

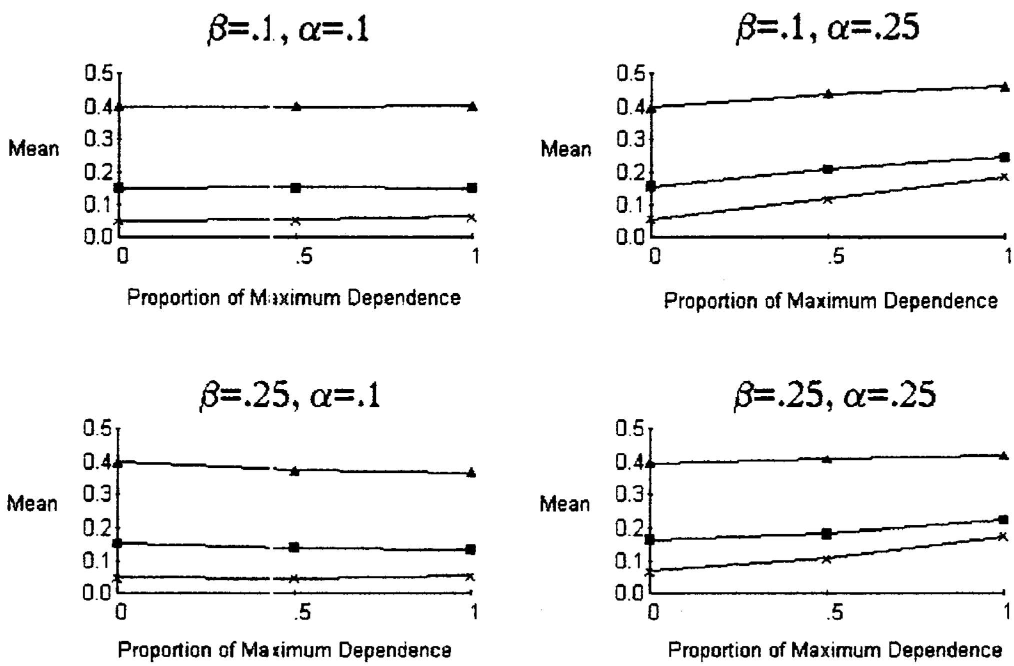
True Prevalence = .40

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	40.02 (1.36)	10.10 (1.36)	10.14 (1.28)	9.69 (1.57)	9.96 (1.24)	10.29 (1.29)	9.87 (1.00)
	0.5	40.04 (0.95)	8.67 (1.36)	8.54 (1.04)	8.79 (1.18)	9.10 (0.86)	9.36 (0.84)	9.07 (0.94)
	1	40.22 (1.40)	**7.39 (1.17)**	**7.42 (1.15)**	**7.24 (1.06)**	8.24 (0.94)	8.49 (1.01)	8.07 (0.94)
β =.1 α =.25	0	39.70 (2.02)	9.61 (2.44)	10.00 (1.90)	9.98 (2.05)	24.67 (2.04)	24.71 (1.89)	25.28 (1.87)
	0.5	43.88 (1.50)	**7.60 (1.58)**	**7.45 (1.08)**	**7.66 (0.99)**	**19.01 (1.34)**	**19.24 (1.27)**	**18.90 (1.59)**
	1	46.07 (1.32)	**5.03 (1.13)**	**5.12 (0.89)**	**4.94 (1.10)**	**13.64 (1.30)**	**13.23 (1.17)**	**13.87 (1.08)**
β =.25 α =.1	0	39.76 (1.61)	24.91 (2.45)	25.04 (2.27)	25.23 (1.95)	9.98 (1.03)	9.95 (1.28)	9.89 (1.52)
	0.5	37.23 (1.71)	**17.73 (2.19)**	**18.36 (1.96)**	**18.13 (1.80)**	8.50 (0.92)	8.57 (1.05)	8.76 (0.78)
	1	36.50 (1.03)	**12.39 (1.10)**	**12.41 (1.55)**	**12.71 (1.54)**	**6.90 (0.86)**	**6.79 (0.80)**	**6.92 (0.81)**
β =.25 α =.25	0	39.44 (4.65)	24.32 (3.68)	24.68 (2.86)	24.51 (4.21)	24.84 (2.31)	25.05 (2.75)	25.22 (2.11)
	0.5	40.82 (1.70)	**15.81 (2.01)**	**15.87 (1.68)**	**15.49 (2.34)**	**18.08 (1.43)**	**18.10 (1.54)**	**17.67 (1.44)**
	1	41.84 (1.35)	**8.50 (1.32)**	**8.81 (1.32)**	**8.64 (1.47)**	**11.44 (0.87)**	**11.77 (1.17)**	**11.86 (1.15)**

* Proportion of maximum dependence between all pairs of observers for both true negative and true positive subjects.
β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.
Means with 20% or greater bias are displayed in bold.

Figure 6a

Mean of Estimates of Prevalence (θ) in the case of three observers, with dependence between all pairs of observers with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

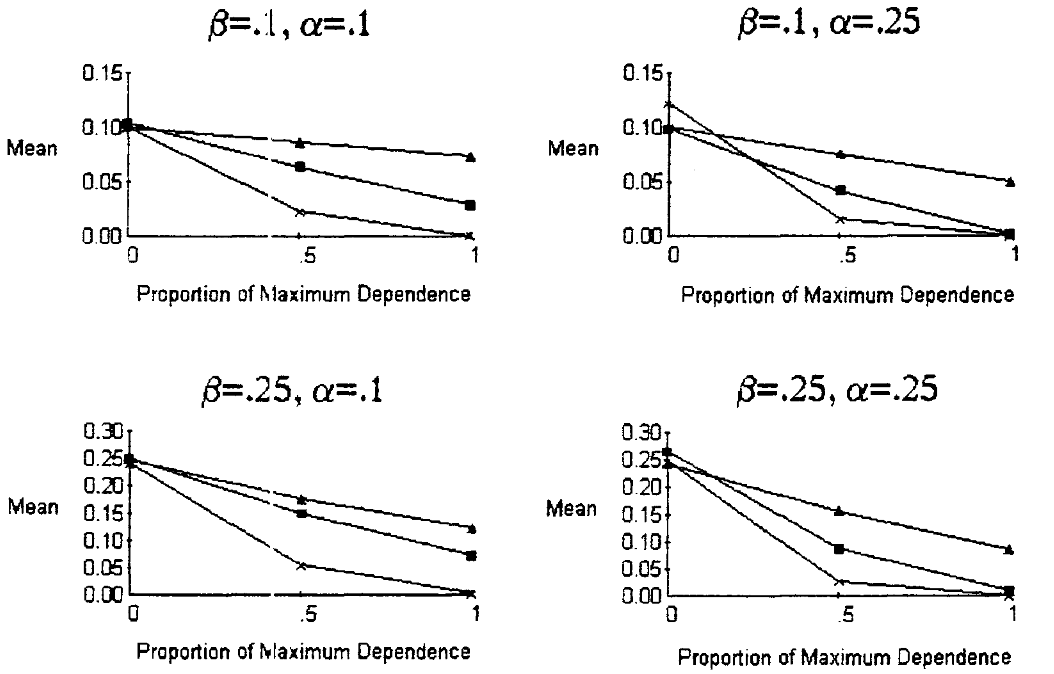
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 6b

Mean of Estimates of false negative rate for observer 1 (β_1) in the case of three observers, with dependence between all pairs of observers with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

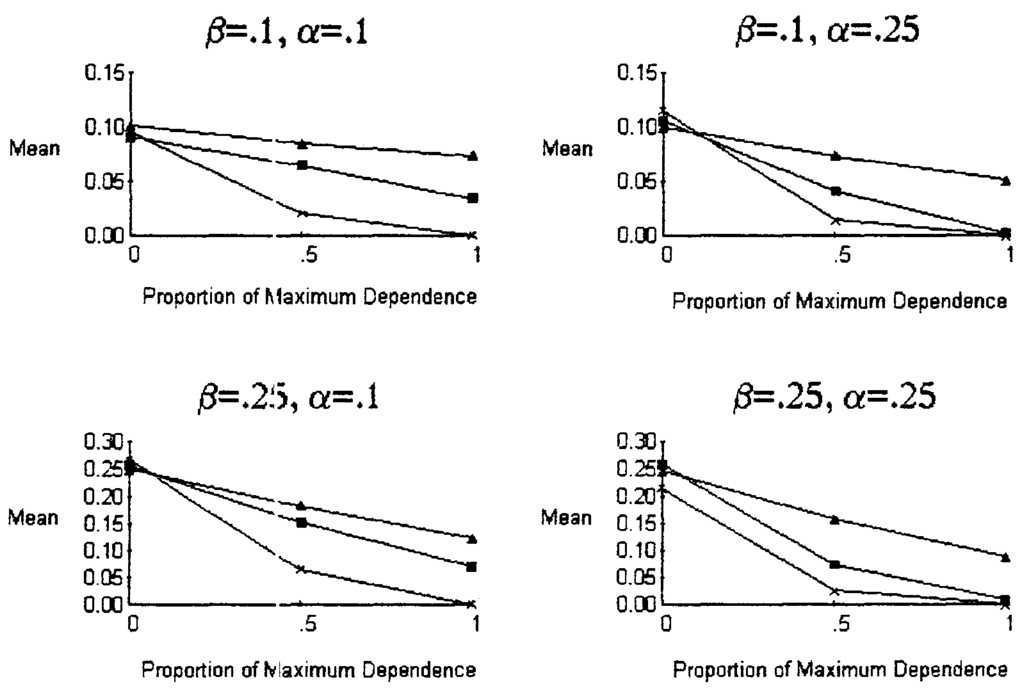
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 6c

Mean of Estimates of false negative rate for observer 2 (β_2) in the case of three observers, with dependence between all pairs of observers with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

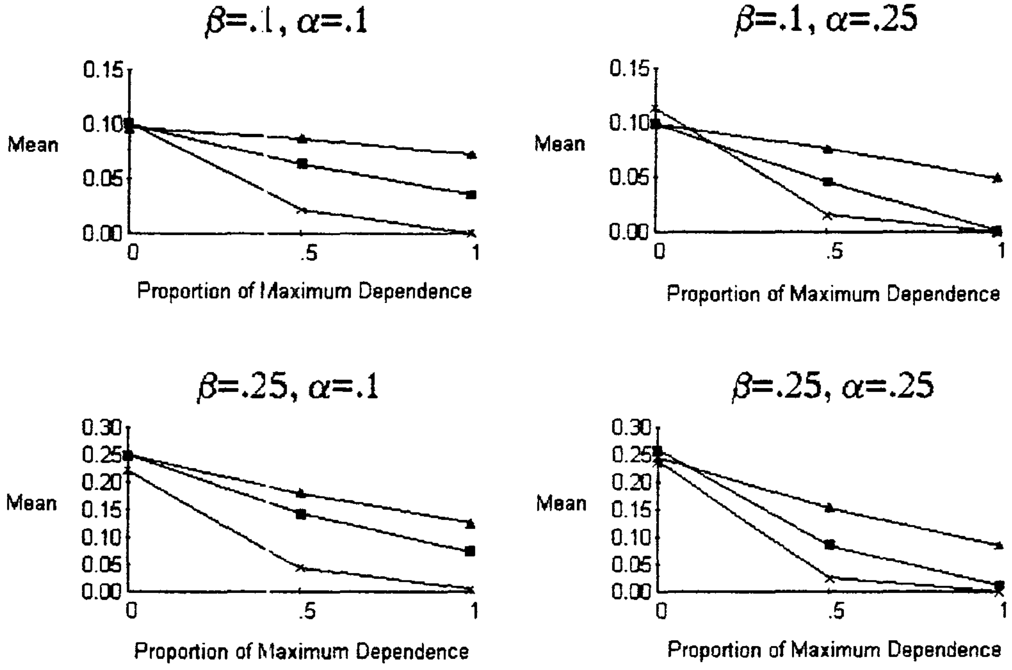
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 6d

Mean of Estimates of false negative rate for observer 3 (β_3) in the case of three observers, with dependence between all pairs of observers with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

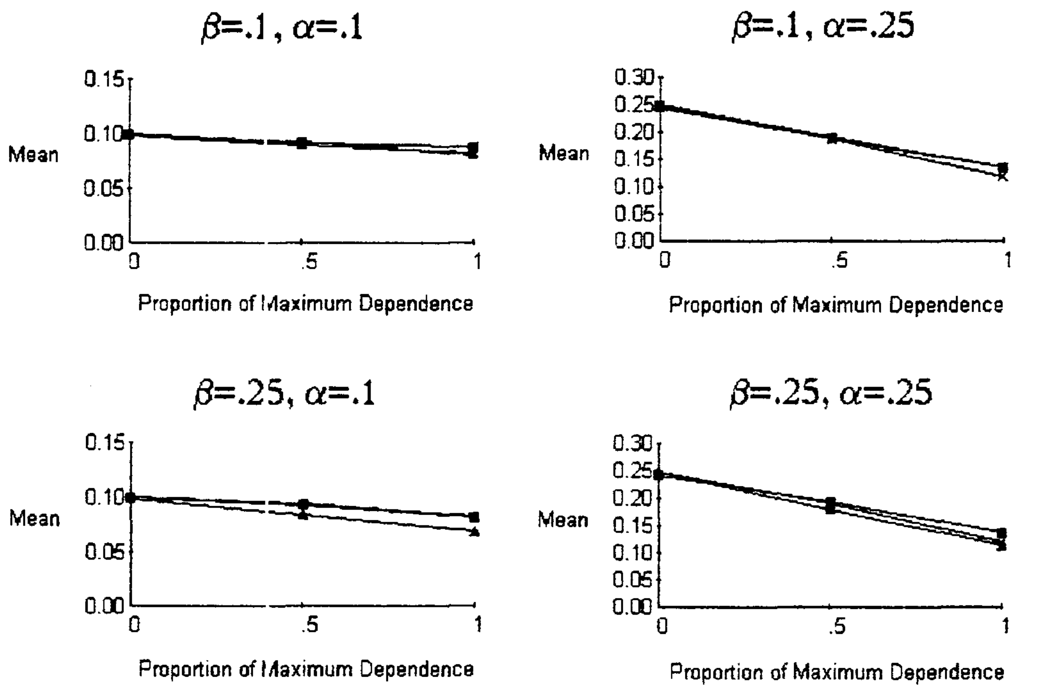
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 6e

Mean of Estimates of false positive rate for observer 1 (α_1) in the case of three observers, with dependence between all pairs of observers with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

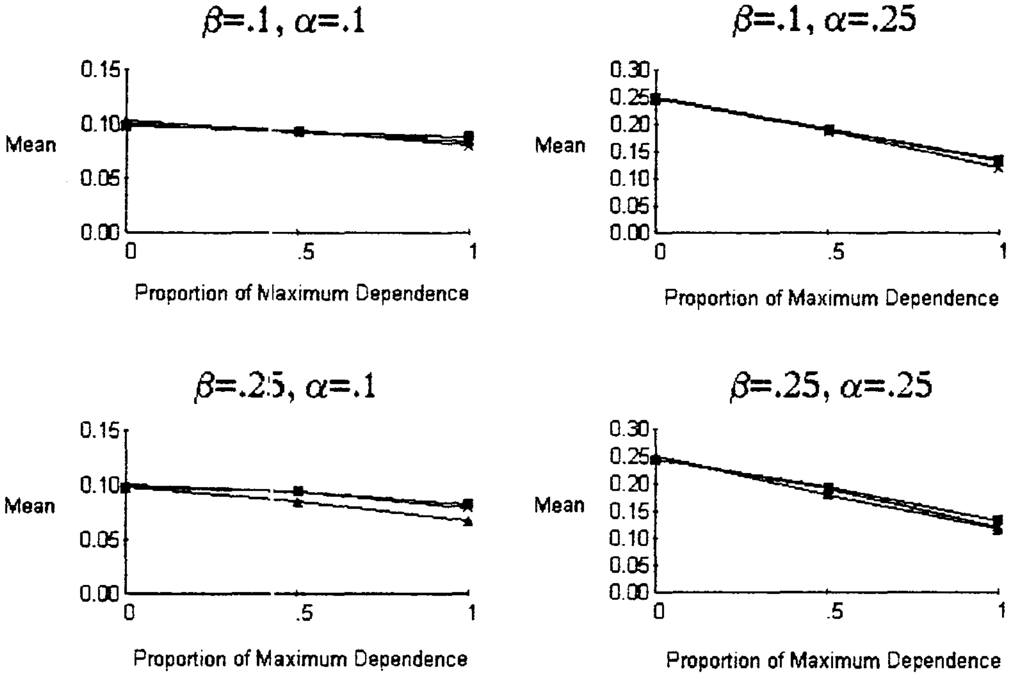
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 6f

Mean of Estimates of false positive rate for observer 2 (α_2) in the case of three observers, with dependence between all pairs of observers with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

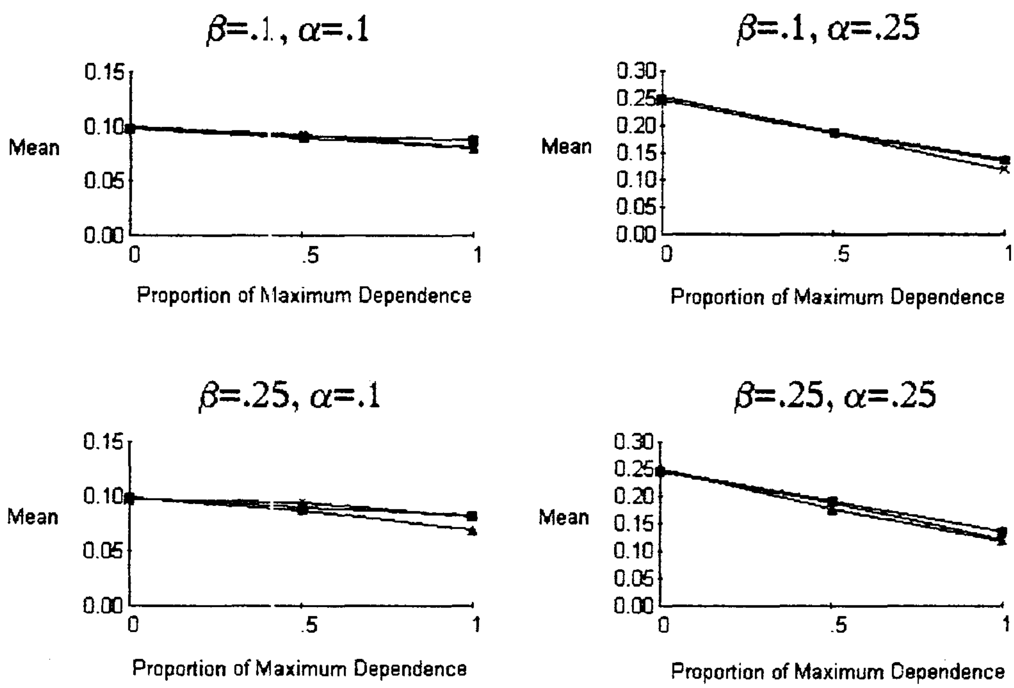
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 6g

Mean of Estimates of false positive rate for observer 3 (α_3) in the case of three observers, with dependence between all pairs of observers with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

3.1.7 RUN 7 - Dependence between two observers with respect to false positive rate - Four Observer Case

This run includes four observers with dependence between observers 1 and 2 with respect to false positive rate only. Tables 7a, 7b, and 7c give the results for θ=.05, θ=.15, and θ=.40, respectively. Figures 7a to 7i show the trends in the means of the parameter estimates as the dependence increases.

OBSERVATIONS

Effect on prevalence estimates

The dependence causes a positive bias in the prevalence estimates when the FPR is high at .25 and the true prevalence is low at .05. At other combinations of the initial parameters the prevalence estimates show no substantial bias.

Effect on the false negative rate estimates

The false negative rates for the two dependent observers are seriously underestimated when the FPR is high. This effect is strongest when the prevalence is low. At θ=.05 these false negative rates are even underestimated when the FPR is low. As was seen in the three observer case, there is an inverse effect on the false negative rates of the independent observers. These false negative rates are overestimated when the FPR is high and when the prevalence is low. These trends are most clearly seen in Figures 7b to 7e.

Effect on the false positive rate estimates

The bias in the false positive rate estimates is not as large as that in the false negative rate estimates. The false positive rate estimates for the two dependent observers tend to possess a negative bias when the FPR is high. This effect appears similar for all prevalence values. There appears to be no effect on the independent observers' false positive rates.

Goodness of fit results

The means of the χ^2 values for the simulations and the power estimates are also given in Tables 7a, 7b, and 7c. It can be seen that in general the χ^2 values increase as the dependence increases. However, as shown by the power estimates and the χ^2 values, the goodness of fit test has trouble picking up the presence of dependence when the FPR is low. This corresponds to the situation where the estimates are not seriously biased.

EXPLANATIONS

These trends can be explained in the same way as the three observer case. That is, the dependent observers, observers 1 and 2, misclassify negative individuals as positive more often than if they were independent. When the false positive rate is high there is more simultaneous misclassification. This over-abundance of positive classifications results in higher prevalence estimates. The false positive rate for the two observers drops because the simultaneous misclassifications as positive increases the apparent probability of these classifications being correct. The false negative rates also drop due to the increased agreement between these two dependent observers. Lower prevalence enhances the effect of the dependence because there are then more truly negative individuals on which the dependence is acting. The independent observers, observers 3 and 4, have their false negative rate estimates positively biased. This is due to their increased lack of agreement with the dependent observers. The false positive rates of these observers are not so much affected because they are based on truly positive individuals of which there are a small number.

This run can be compared to the equivalent run with three observers, run 1. It can be seen that all biases are larger in the three observer case. The presence of a second independent observer helps to dilute the effects of the dependence between observers 1 and 2.

IMPLICATIONS

As was stated in the three observer case, the most dangerous situation is when the prevalence is low and the false positive rate is high. However, the addition of another independent observer helps control the bias problem. Thus, if it is suspected that two of three observers or screening tests are dependent then more accurate estimates can be achieved by the addition of a fourth independent observer.

In this situation the goodness of fit test seems to have good power to identify the dependence as a departure from the model assumptions when the estimates are biased. However, in the cases where the estimates were not substantially biased the test had little power.

TABLE 7a

Means and standard deviations of parameter estimates in the case of four observers, with dependence between observers 1 and 2 with respect to false positive rate only

True Prevalence = .05

True Error Rates	Dep[*]	Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
		θ	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean $\tilde{\chi}^2$	Power[†]
β =.1 α =.1	0	5.05 (0.41)	10.67 (3.49)	9.99 (2.63)	10.98 (2.95)	9.72 (2.37)	10.05 (0.55)	9.99 (0.50)	10.06 (0.50)	10.08 (0.60)	6.31	0.03
	0.5	5.10 (0.36)	8.86 (2.50)	8.81 (2.78)	11.13 (2.30)	10.81 (2.48)	9.89 (0.51)	9.86 (0.59)	9.92 (0.48)	10.02 (0.71)	5.96	0.03
	1	4.91 (0.34)	**6.78 (2.07)**	**5.88 (2.24)**	11.38 (1.65)	11.94 (3.69)	9.95 (0.36)	9.76 (0.44)	10.13 (0.47)	10.05 (0.55)	8.19	0.17
β =.1 α =.25	0	4.98 (0.78)	10.29 (5.18)	9.19 (5.60)	9.10 (5.16)	9.03 (5.45)	25.00 (0.81)	25.04 (0.79)	24.94 (0.80)	25.04 (0.77)	6.44	0.09
	0.5	**6.88 (0.61)**	**0.35 (1.32)**	**0.24 (0.79)**	**25.34 (4.70)**	**25.66 (4.46)**	22.92 (0.77)	22.95 (0.83)	24.78 (0.79)	24.93 (0.73)	24.89	0.93
	1	**10.61 (0.59)**	**0.00 (0.00)**	**0.00 (0.00)**	**38.24 (2.99)**	**37.52 (3.10)**	**19.81 (0.66)**	**19.68 (0.74)**	24.19 (0.78)	24.06 (0.67)	81.18	1.00
β =.25 α =.1	0	4.95 (0.49)	22.70 (4.59)	24.42 (4.34)	24.43 (4.22)	24.73 (3.50)	10.03 (0.58)	10.18 (0.47)	10.03 (0.44)	9.92 (0.49)	6.07	0.00
	0.5	4.82 (0.55)	22.70 (4.76)	21.17 (3.68)	25.40 (5.33)	25.15 (4.35)	10.05 (0.42)	9.97 (0.54)	10.04 (0.49)	10.06 (0.39)	6.62	0.10
	1	4.95 (0.53)	20.40 (4.36)	20.41 (4.43)	26.60 (4.94)	25.54 (5.18)	9.75 (0.34)	9.83 (0.39)	10.15 (0.55)	10.06 (0.43)	8.41	0.10
β =.25 α =.25	0	5.19 (1.46)	24.01 (7.63)	23.94 (9.44)	24.99 (9.00)	25.62 (9.85)	24.76 (1.05)	25.11 (0.73)	24.99 (1.00)	24.83 (1.12)	6.56	0.00
	0.5	**6.07 (0.72)**	**1.71 (3.23)**	**1.73 (3.34)**	**39.48 (4.57)**	**40.14 (5.25)**	23.01 (0.88)	22.81 (0.81)	25.30 (0.70)	25.41 (0.66)	15.69	0.57
	1	**9.72 (0.65)**	**0.00 (0.00)**	**0.00 (0.00)**	**47.93 (2.54)**	**48.82 (2.83)**	**19.76 (0.87)**	**19.85 (0.82)**	24.90 (0.66)	24.99 (0.64)	48.12	1.00

[*] Proportion of maximum dependence between observers 1 and 2 for true negative subjects.

[†] Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6}$ = 12.59.

β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

Means with 20% or greater bias are displayed in bold.

TABLE 7b

Means and standard deviations of parameter estimates in the case of four observers, with dependence between observers 1 and 2 with respect to false positive rate only

True Prevalence = .15

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
		θ	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean χ^2	Power†
β =.1 α =.1	0	14.95 (0.57)	10.13 (1.45)	9.87 (1.31)	9.71 (1.52)	9.63 (1.19)	9.90 (0.40)	9.93 (0.40)	9.75 (0.49)	9.96 (0.55)	6.56	0.03
	0.5	14.85 (0.64)	9.67 (1.72)	9.61 (1.38)	9.87 (1.41)	9.90 (1.53)	9.81 (0.64)	10.14 (0.51)	9.96 (0.53)	10.08 (0.59)	5.46	0.03
	1	14.98 (0.67)	9.36 (1.16)	9.02 (1.36)	10.43 (1.60)	10.30 (1.97)	9.99 (0.63)	9.76 (0.51)	9.98 (0.55)	10.01 (0.47)	7.31	0.10
β =.1 α =.25	0	15.22 (0.78)	10.50 (1.93)	9.40 (2.71)	10.05 (1.47)	9.41 (1.83)	24.88 (0.94)	24.89 (0.83)	25.10 (0.86)	24.95 (1.07)	5.25	0.00
	0.5	15.82 (0.74)	**4.80 (1.82)**	**5.13 (1.57)**	**14.99 (1.84)**	**14.67 (2.00)**	23.36 (0.78)	23.66 (0.99)	25.37 (0.73)	25.29 (1.00)	19.63	0.90
	1	17.25 (0.89)	**0.87 (1.20)**	**1.16 (1.25)**	**20.76 (2.41)**	**21.11 (2.62)**	21.35 (0.79)	21.34 (0.96)	25.39 (0.65)	25.21 (0.66)	60.49	1.00
β =.25 α =.1	0	14.95 (0.80)	24.55 (3.31)	25.93 (1.90)	24.50 (2.43)	25.04 (2.43)	9.88 (0.52)	10.15 (0.43)	10.01 (0.56)	9.88 (0.48)	6.09	0.10
	0.5	15.08 (0.82)	24.17 (2.61)	24.47 (2.01)	25.03 (1.29)	25.10 (1.86)	9.94 (0.60)	10.05 (0.56)	10.04 (0.64)	10.14 (0.58)	7.14	0.10
	1	14.65 (0.69)	23.47 (2.35)	23.63 (2.53)	25.03 (2.34)	25.72 (2.27)	9.78 (0.55)	9.86 (0.55)	9.97 (0.51)	10.24 (0.57)	6.16	0.07
β =.25 α =.25	0	15.35 (1.92)	25.79 (3.65)	25.46 (3.61)	24.33 (3.21)	25.65 (3.49)	25.07 (0.98)	25.04 (1.08)	24.92 (0.86)	25.10 (1.28)	6.23	0.07
	0.5	15.60 (1.24)	**16.27 (2.37)**	**15.71 (2.85)**	**31.04 (3.10)**	**31.18 (3.03)**	22.91 (1.10)	22.87 (0.86)	25.80 (0.93)	25.84 (0.97)	18.16	0.70
	1	15.39 (1.28)	**5.82 (3.51)**	**6.10 (3.07)**	**36.45 (2.50)**	**36.45 (2.74)**	20.96 (1.07)	21.25 (1.12)	26.73 (1.05)	26.81 (0.83)	35.22	0.97

* Proportion of maximum dependence between observers 1 and 2 for true negative subjects.

† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6}$ = 12.59.

β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

Means with 20% or greater bias are displayed in bold.

TABLE 7c

Means and standard deviations of parameter estimates in the case of four observers, with dependence between observers 1 and 2 with respect to false positive rate only

True Prevalence = .40

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
		θ	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean χ^2	Power†
β =.1 α =.1	0	40.34 (0.69)	10.00 (0.85)	9.85 (0.72)	9.99 (0.70)	10.01 (0.91)	9.95 (0.68)	9.74 (0.78)	10.09 (0.52)	10.03 (0.60)	6.94	0.10
	0.5	39.88 (0.81)	9.98 (0.86)	9.95 (0.85)	9.83 (0.79)	9.87 (0.81)	9.68 (0.60)	10.01 (0.61)	9.97 (0.67)	9.90 (0.68)	5.90	0.03
	1	39.96 (0.87)	9.53 (0.99)	9.77 (1.05)	10.18 (0.98)	10.29 (0.89)	9.83 (0.61)	9.88 (0.58)	10.07 (0.77)	10.19 (0.83)	6.39	0.03
β =.1 α =.25	0	40.48 (1.00)	10.18 (0.88)	10.24 (1.02)	10.30 (1.15)	10.15 (0.94)	24.96 (1.17)	24.72 (1.03)	24.77 (0.96)	24.66 (1.34)	6.40	0.10
	0.5	40.87 (1.00)	8.91 (0.90)	8.86 (0.94)	11.03 (1.20)	11.39 (1.22)	23.54 (1.02)	23.39 (1.17)	25.18 (0.95)	25.21 (1.08)	17.16	0.70
	1	41.39 (0.85)	**7.46 (0.93)**	**7.40 (0.94)**	**12.71 (1.02)**	**13.02 (0.95)**	21.72 (0.97)	21.75 (1.11)	25.52 (1.08)	25.06 (0.94)	48.89	1.00
β =.25 α =.1	0	40.21 (1.05)	25.17 (1.25)	25.26 (1.50)	25.14 (1.28)	24.98 (1.16)	9.91 (0.93)	10.05 (0.92)	10.01 (0.59)	9.88 (0.83)	5.38	0.00
	0.5	40.29 (1.08)	24.88 (1.13)	24.53 (1.14)	25.11 (1.47)	24.68 (1.36)	9.81 (0.75)	9.90 (0.81)	9.97 (0.78)	9.97 (0.75)	5.82	0.13
	1	40.17 (0.80)	24.73 (1.28)	24.57 (1.39)	25.13 (1.26)	25.14 (1.17)	9.68 (0.63)	10.13 (0.55)	10.14 (0.85)	10.24 (0.63)	6.99	0.10
β =.25 α =.25	0	39.94 (2.00)	24.86 (2.30)	24.80 (1.66)	24.95 (1.41)	25.42 (2.35)	24.90 (1.68)	25.19 (1.32)	24.99 (1.57)	24.90 (1.34)	7.87	0.17
	0.5	39.86 (1.40)	21.86 (1.53)	22.34 (1.53)	26.47 (1.55)	26.27 (1.65)	23.28 (1.25)	23.40 (0.96)	26.04 (1.18)	25.94 (1.06)	11.92	0.40
	1	40.48 (1.55)	**19.60 (1.71)**	20.14 (1.89)	27.97 (1.83)	28.37 (1.54)	21.30 (1.28)	21.26 (1.41)	26.54 (1.19)	26.89 (1.31)	28.77	0.97

* Proportion of maximum dependence between observers 1 and 2 for true negative subjects.

† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6}$ = 12.59.

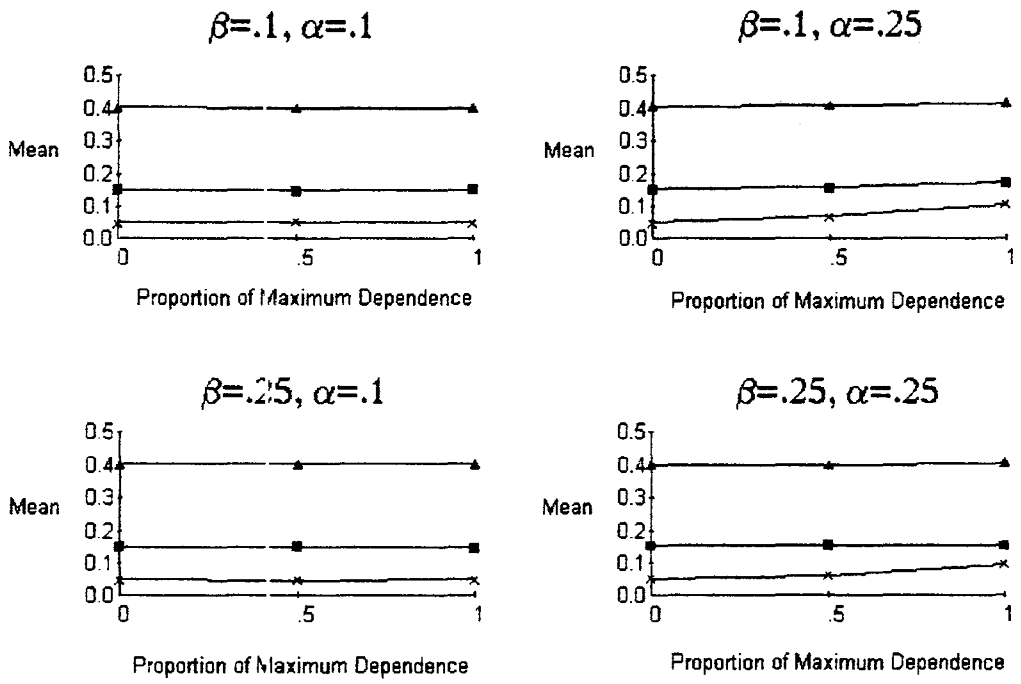
β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

Means with 20% or greater bias are displayed in bold.

Figure 7a

Mean of Estimates of Prevalence (θ) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

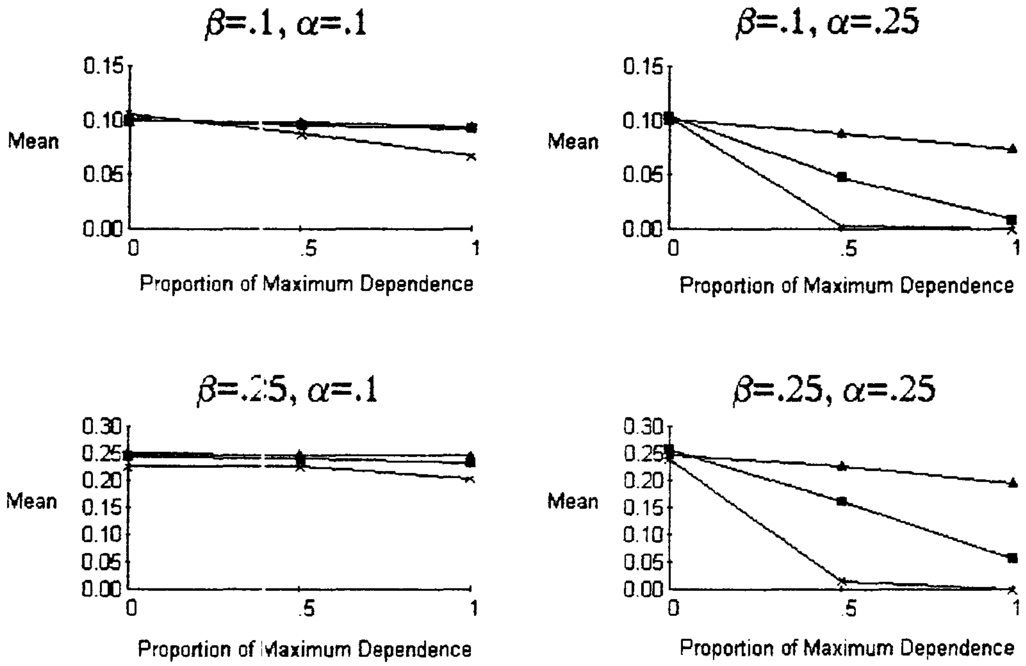
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 7b

Mean of Estimates of false negative rate for observer 1 (β_1) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

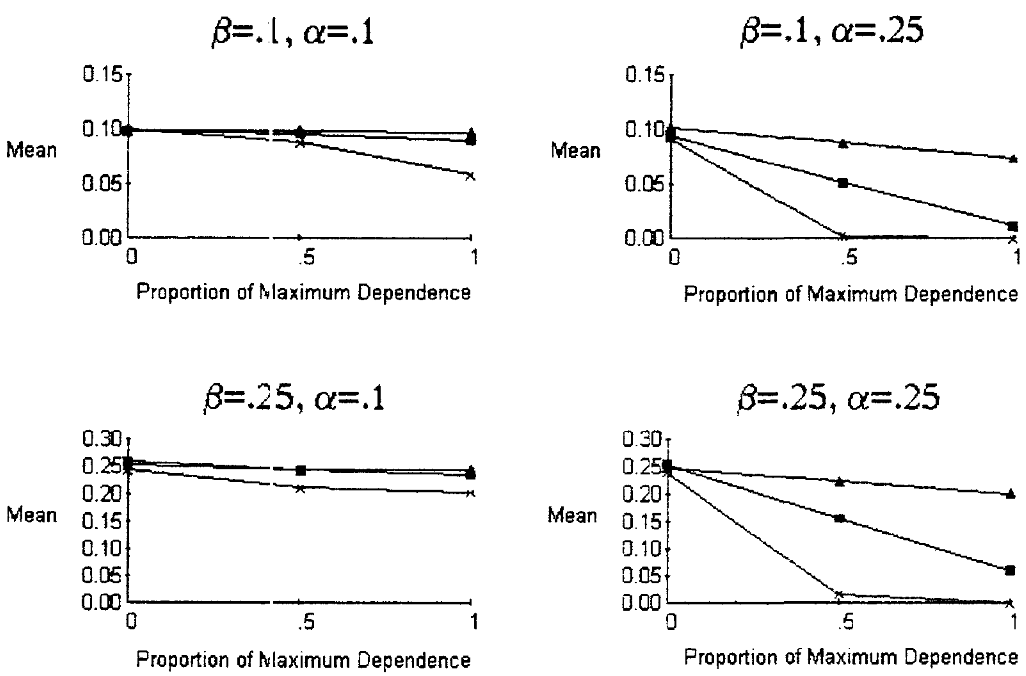
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 7c

Mean of Estimates of false negative rate for observer 2 (β_2) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

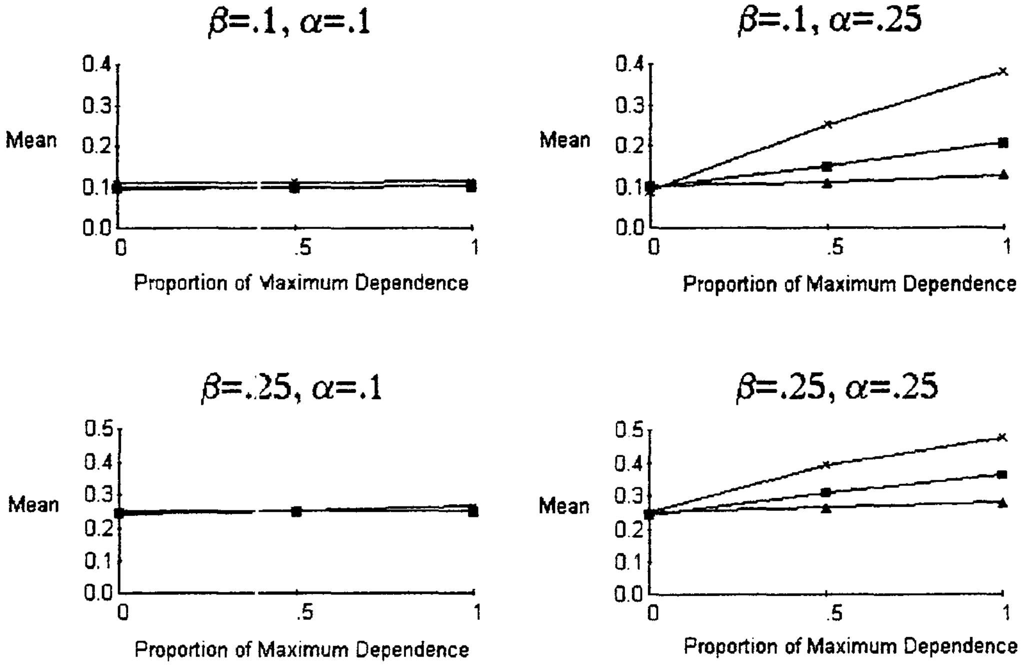
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 7d

Mean of Estimates of false negative rate for observer 3 (β_3) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

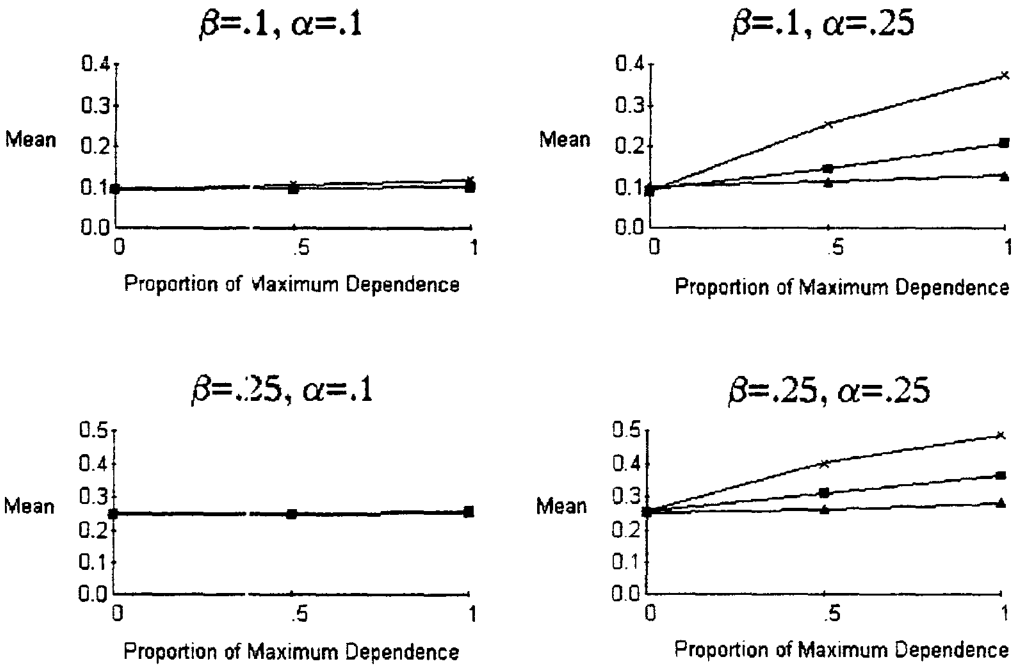
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 7e

Mean of Estimates of false negative rate for observer 4 (β_4) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

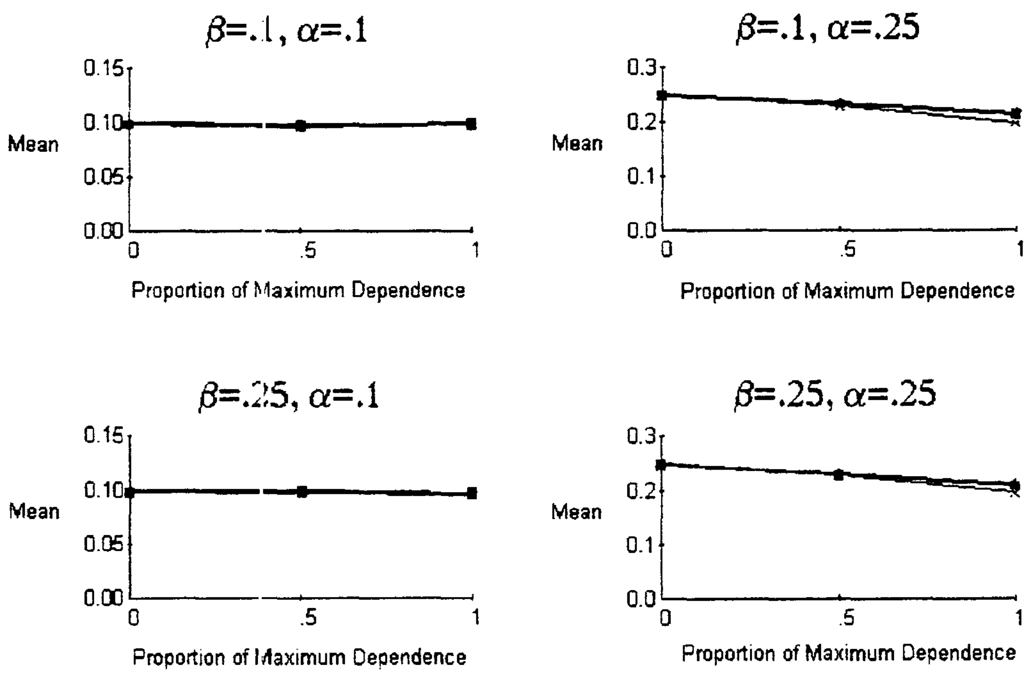
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 7f

Mean of Estimates of false positive rate for observer 1 (α_1) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

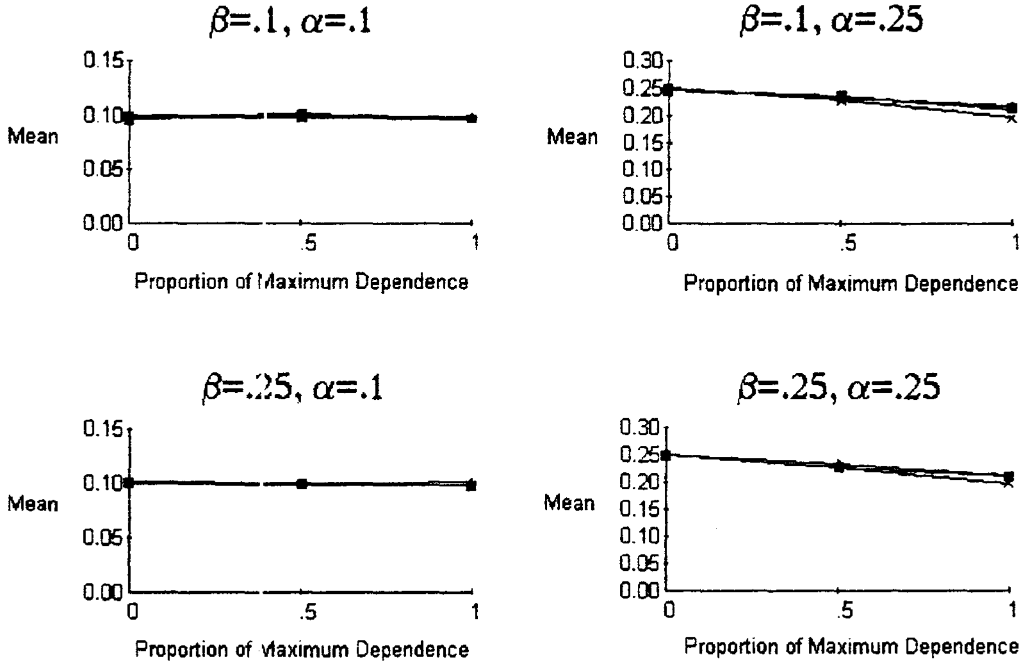
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 7g

Mean of Estimates of false positive rate for observer 2 (α_2) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

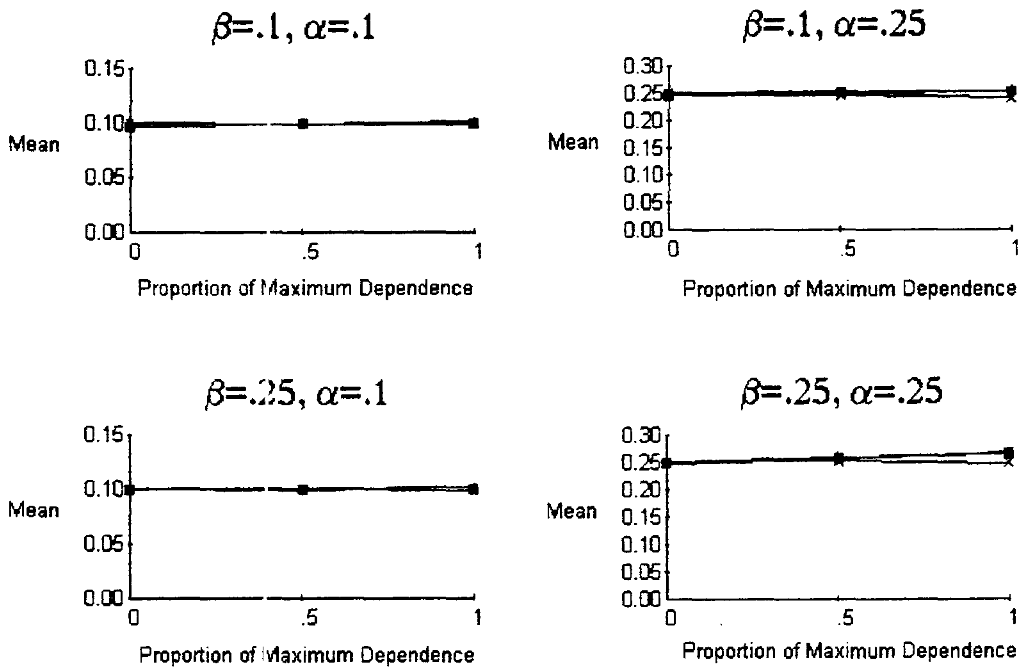
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 7h

Mean of Estimates of false positive rate for observer 3 (α_3) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

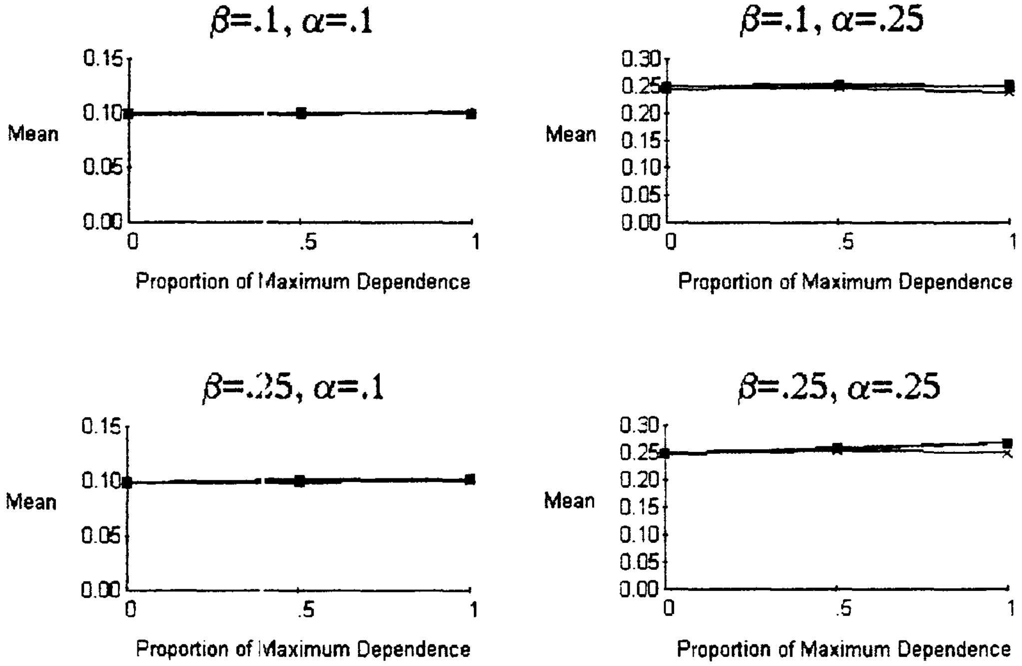
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 7i

Mean of Estimates of false positive rate for observer 4 (α_4) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

3.1.8 RUN 8 - Dependence between two observers with respect to false negative rate - Four Observer Case

The dependence in this run is between observers 1 and 2 with respect to false negative rate. Tables 8a, 8b, and 8c give the results for θ=.05, θ=.15, and θ=.40, respectively. The graphs of the means of the estimates are displayed in Figures 8a to 8i.

OBSERVATIONS

Effect on prevalence estimates

Figure 8a shows that the means of prevalence estimates remain relatively constant in the presence of this dependence. There is no detectable bias present in the estimates.

Effect on false negative rate estimates

There appears to be a very slight negative bias in the false negative rate estimates for the dependent observers when the FNR is high. For the independent observers there is no apparent effect of the dependence.

Effect on the false positive rate estimates

Any bias that may be present in the false positive rate estimates is very small, as can be seen in Figures 8f to 8i. As has been seen before, the trend is towards underestimation for the dependent observers and overestimation for the independent observers when the FNR is high

Goodness of fit results

Tables 8a to 8c present the means of the χ^2 goodness of fit values and an estimate of the power of the goodness of fit tests. In general the numbers indicate that the test has low power to detect this type of dependence. When the prevalence is .05 the test shows no power. At the prevalence of .15 there is some power shown when the FNR is high. There is fairly good power when the prevalence is .40 and the FNR is high.

EXPLANATIONS

There is not much detectable bias in this particular situation. Any bias that may be present is expected to reflect the biases seen in the three observer case, Run 2. However, it is interesting to note that the addition of the fourth observer basically alleviated the bias problems seen in the three observer case.

IMPLICATIONS

The dependency in this situation does not pose much of a threat to the latent class estimates. As was discussed in the last run, if it is suspected that two of three observers are dependent then the addition of an independent fourth observer will effectively remove any bias effects from the dependency. Again the goodness of fit test appears to only have power when the FNR is high and the estimates are showing slight trends of being biased.

TABLE 8a

Means and standard deviations of parameter estimates in the case of four observers, with dependence between observers 1 and 2 with respect to false negative rate only

True Prevalence = .05

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
		θ	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean $\tilde{\chi}^2$	Power†
β = .1, α = .1 ††	0	5.04 (0.32)	9.22 (2.49)	9.76 (2.75)	9.86 (2.55)	8.95 (2.78)	10.01 (0.42)	10.10 (0.65)	10.08 (0.55)	10.09 (0.46)	6.38	0.13
	0.5	5.13 (0.43)	9.49 (2.70)	10.23 (2.98)	9.81 (2.58)	10.35 (2.54)	9.82 (0.48)	10.00 (0.47)	10.05 (0.43)	10.07 (0.55)	5.12	0.03
	1	4.93 (0.40)	9.97 (2.44)	10.32 (3.05)	11.16 (2.53)	10.35 (2.74)	10.00 (0.56)	9.92 (0.43)	9.92 (0.41)	10.13 (0.47)	5.93	0.03
β = .1, α = .25	0	5.04 (0.80)	10.48 (5.59)	9.40 (4.95)	10.09 (5.78)	9.10 (5.67)	24.95 (0.82)	24.90 (0.97)	25.02 (0.82)	25.01 (0.80)	6.03	0.05
	0.5	5.04 (0.69)	10.51 (4.52)	10.49 (4.71)	9.46 (5.21)	10.12 (5.40)	25.00 (0.79)	25.01 (0.68)	24.96 (0.95)	24.98 (0.62)	6.60	0.04
	1	5.00 (0.72)	9.62 (4.92)	9.90 (4.85)	9.92 (5.38)	10.45 (4.90)	25.12 (0.76)	25.08 (0.81)	25.07 (0.84)	24.97 (0.69)	5.79	0.07
β = .25, α = .1	0	5.01 (0.50)	25.68 (4.15)	24.11 (4.09)	25.99 (4.13)	24.90 (4.77)	9.92 (0.60)	9.97 (0.62)	10.09 (0.53)	9.93 (0.61)	5.21	0.00
	0.5	5.03 (0.58)	23.44 (5.59)	23.04 (4.09)	26.02 (4.97)	25.38 (5.15)	9.80 (0.43)	9.89 (0.52)	10.09 (0.50)	10.02 (0.62)	7.29	0.10
	1	4.88 (0.62)	20.79 (5.42)	22.26 (3.87)	25.62 (4.76)	26.99 (4.88)	9.82 (0.53)	9.78 (0.53)	10.15 (0.64)	10.23 (0.50)	6.94	0.10
β = .25, α = .25	0	5.20 (1.53)	22.07 (10.54)	25.50 (6.56)	27.14 (6.00)	24.96 (8.38)	24.97 (0.65)	24.94 (0.88)	24.85 (0.86)	24.86 (1.11)	5.17	0.03
	0.5	5.26 (1.37)	24.93 (11.07)	23.48 (9.20)	28.23 (8.09)	26.20 (7.98)	24.99 (0.72)	24.75 (0.83)	24.87 (0.99)	24.86 (1.10)	6.20	0.03
	1	5.47 (2.15)	21.79 (9.60)	20.36 (9.99)	26.18 (9.35)	26.43 (7.74)	24.77 (1.03)	24.50 (1.09)	24.81 (0.80)	24.78 (0.81)	6.45	0.00

* Proportion of maximum dependence between observers 1 and 2 for true positive subjects.

† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6}$ = 12.59.

β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

Means with 20% or greater bias are displayed in bold.

TABLE 8b

Means and standard deviations of parameter estimates in the case of four observers, with dependence between observers 1 and 2 with respect to false negative rate only

True Prevalence = .15

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
		θ	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean χ^2	Power†
β =.1 α =.1	0	14.75 (0.57)	9.59 (1.72)	10.16 (1.50)	10.10 (1.39)	10.26 (1.52)	10.12 (0.52)	9.92 (0.53)	10.11 (0.53)	10.15 (0.47)	5.96	0.03
	0.5	14.97 (0.50)	9.88 (1.46)	10.17 (1.41)	10.08 (1.49)	10.17 (1.24)	9.89 (0.58)	9.96 (0.46)	10.00 (0.66)	9.97 (0.48)	6.59	0.00
	1	15.07 (0.60)	9.84 (1.17)	9.69 (1.48)	10.19 (1.54)	9.73 (1.38)	10.09 (0.52)	10.01 (0.46)	9.90 (0.51)	10.09 (0.54)	6.43	0.03
β =.1 α =.25	0	14.93 (0.85)	9.26 (1.97)	9.41 (1.87)	9.42 (1.90)	10.37 (2.15)	25.15 (0.76)	24.94 (0.88)	25.01 (0.99)	25.16 (0.80)	5.23	0.00
	0.5	15.01 (0.77)	9.65 (1.78)	9.95 (1.73)	10.13 (2.38)	10.15 (2.32)	24.84 (0.89)	25.02 (0.52)	24.89 (0.90)	24.98 (0.88)	5.46	0.00
	1	15.29 (0.76)	9.94 (2.59)	10.22 (1.99)	10.37 (1.64)	10.31 (1.79)	24.83 (0.97)	25.08 (0.98)	24.84 (0.85)	25.06 (0.71)	5.56	0.03
β =.25 α =.1	0	15.10 (0.75)	24.93 (2.70)	25.11 (2.56)	25.01 (2.68)	24.78 (2.37)	9.93 (0.64)	10.16 (0.61)	9.99 (0.58)	9.98 (0.55)	5.56	0.03
	0.5	14.87 (0.90)	23.44 (1.86)	23.57 (2.36)	25.40 (2.61)	25.09 (2.71)	9.80 (0.62)	9.74 (0.53)	10.06 (0.54)	10.09 (0.72)	6.35	0.07
	1	14.85 (0.72)	21.99 (2.33)	21.72 (2.50)	25.95 (1.98)	25.63 (1.95)	9.75 (0.56)	9.64 (0.55)	10.13 (0.67)	10.28 (0.47)	12.90	0.50
β =.25 α =.25	0	15.27 (1.47)	25.03 (3.94)	25.11 (3.68)	25.74 (3.06)	24.99 (3.74)	24.58 (1.13)	25.10 (0.93)	25.11 (0.99)	24.72 (0.97)	5.66	0.03
	0.5	14.62 (1.55)	22.28 (4.73)	22.42 (3.25)	26.17 (3.73)	24.56 (3.80)	24.62 (0.98)	24.53 (0.95)	25.16 (0.87)	25.22 (0.91)	7.07	0.07
	1	15.25 (1.64)	21.02 (3.76)	21.70 (3.68)	27.05 (3.08)	27.73 (3.48)	24.21 (0.87)	24.34 (0.79)	25.32 (0.96)	25.28 (0.93)	7.62	0.13

* Proportion of maximum dependence between observers 1 and 2 for true positive subjects.

† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6}$ = 12.59.

β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

Means with 20% or greater bias are displayed in bold.

TABLE 8c

Means and standard deviations of parameter estimates in the case of four observers, with dependence between observers 1 and 2 with respect to false negative rate only

True Prevalence = .40

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
		θ	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean χ^2	Power†
β =.1 α =.1	0	40.03 (0.89)	9.97 (0.83)	9.94 (0.74)	9.87 (0.85)	10.12 (0.99)	10.08 (0.68)	10.00 (0.67)	10.28 (0.54)	9.91 (0.65)	7.15	0.17
	0.5	39.96 (0.75)	9.93 (0.87)	10.25 (0.88)	10.00 (0.77)	9.93 (0.87)	9.95 (0.55)	10.01 (0.64)	10.05 (0.73)	10.03 (0.70)	5.95	0.07
	1	40.06 (0.76)	9.73 (0.99)	9.94 (0.79)	9.90 (0.64)	10.06 (0.80)	10.04 (0.54)	9.87 (0.60)	10.07 (0.64)	9.89 (0.52)	7.51	0.10
β =.1 α =.25	0	39.73 (0.86)	10.10 (0.94)	9.86 (0.87)	9.86 (0.96)	9.98 (0.93)	25.15 (0.95)	25.01 (1.00)	25.23 (0.88)	25.05 (1.00)	7.17	0.03
	0.5	39.92 (0.82)	9.66 (1.00)	9.72 (0.89)	10.07 (1.12)	10.05 (1.30)	24.91 (0.96)	25.20 (1.18)	25.22 (0.93)	24.90 (1.01)	5.49	0.00
	1	40.06 (0.97)	9.88 (0.86)	10.16 (0.72)	10.17 (0.94)	9.85 (0.78)	24.58 (0.84)	25.06 (0.94)	25.16 (1.00)	25.08 (0.82)	5.94	0.10
β =.25 α =.1	0	39.90 (1.01)	24.84 (1.11)	24.87 (1.63)	24.98 (1.42)	24.73 (1.06)	9.91 (0.94)	10.04 (0.80)	9.99 (0.81)	9.93 (0.63)	5.70	0.10
	0.5	39.87 (0.89)	23.53 (1.32)	23.46 (1.24)	25.75 (1.10)	25.63 (1.41)	9.48 (0.56)	9.26 (0.66)	10.46 (0.76)	10.59 (0.75)	13.31	0.53
	1	39.19 (0.87)	21.96 (1.64)	21.82 (1.27)	25.38 (1.31)	25.60 (1.47)	8.70 (0.58)	8.70 (0.84)	11.30 (0.78)	11.05 (0.59)	33.16	1.00
β =.25 α =.25	0	40.58 (1.92)	24.80 (1.94)	25.60 (2.08)	25.16 (1.82)	25.50 (1.58)	24.93 (1.45)	25.27 (1.02)	24.79 (1.43)	24.85 (1.60)	6.24	0.10
	0.5	39.95 (1.91)	23.30 (1.75)	22.88 (1.65)	25.60 (1.90)	25.31 (1.43)	23.60 (1.29)	23.96 (1.34)	25.73 (1.09)	25.45 (1.28)	8.93	0.13
	1	39.91 (1.80)	21.35 (2.02)	21.65 (1.93)	26.32 (1.62)	27.00 (1.63)	22.47 (1.23)	22.70 (1.23)	26.09 (1.50)	26.04 (1.20)	15.67	0.63

* Proportion of maximum dependence between observers 1 and 2 for true positive subjects.

† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6}$ = 12.59.

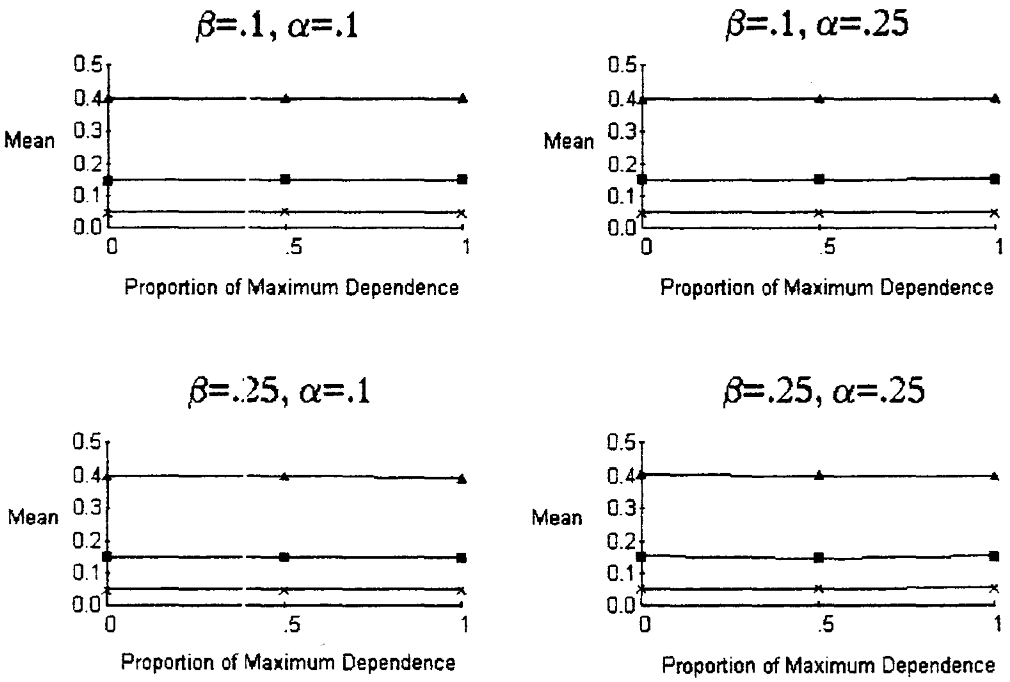
β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

Means with 20% or greater bias are displayed in bold.

Figure 8a

Mean of Estimates of Prevalence (θ)
in the case of four observers, with dependence between observers 1 and 2 with respect to false negative rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

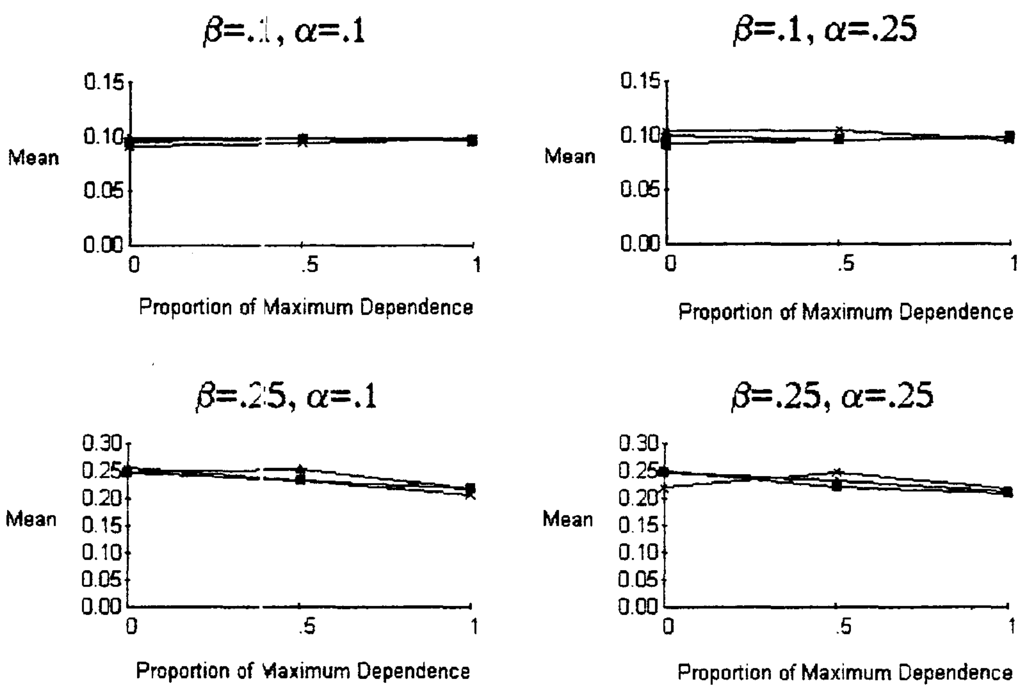
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 8b

Mean of Estimates of false negative rate for observer 1 (β_1) in the case of four observers, with dependence between observers 1 and 2 with respect to false negative rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

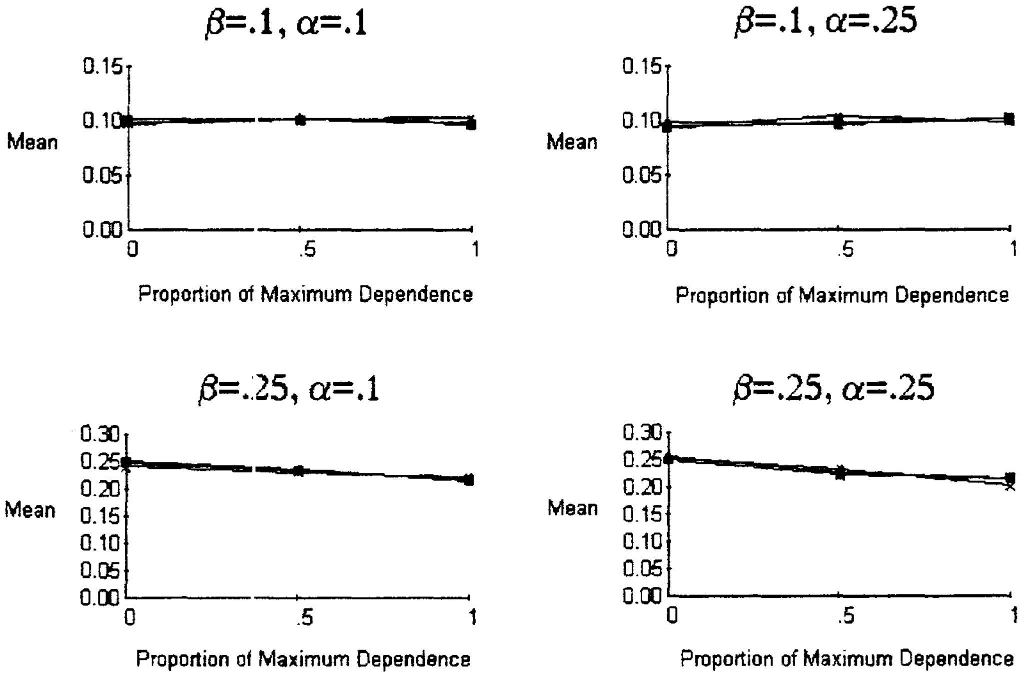
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 8c

Mean of Estimates of false negative rate for observer 2 (β_2) in the case of four observers, with dependence between observers 1 and 2 with respect to false negative rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

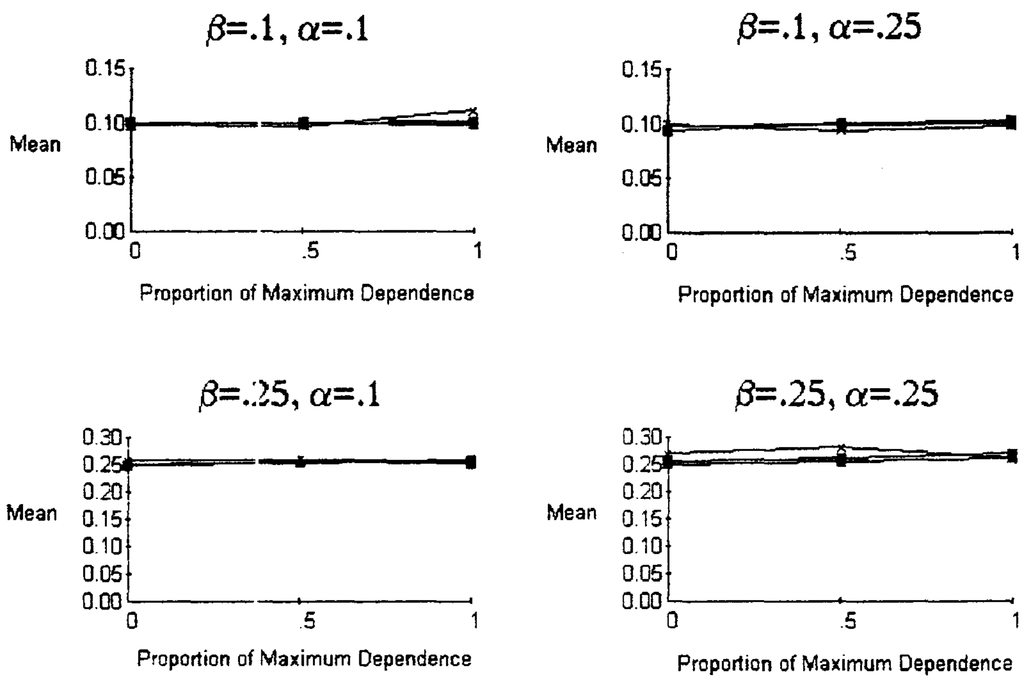
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 8d

Mean of Estimates of false negative rate for observer 3 (β_3) in the case of four observers, with dependence between observers 1 and 2 with respect to false negative rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

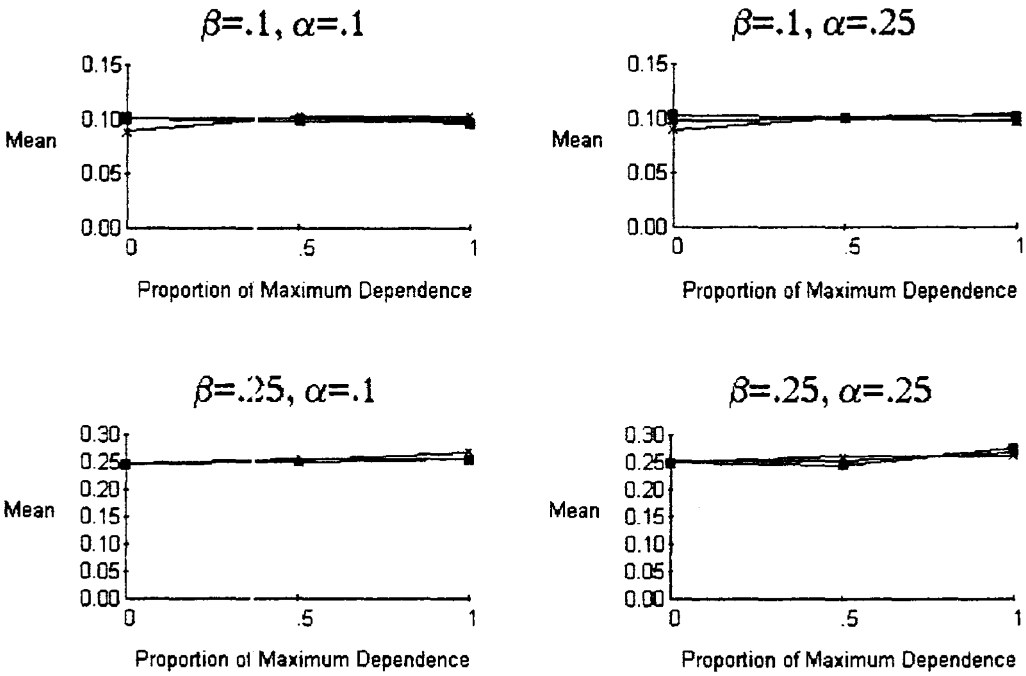
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 8e

Mean of Estimates of false negative rate for observer 4 (β_4) in the case of four observers, with dependence between observers 1 and 2 with respect to false negative rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

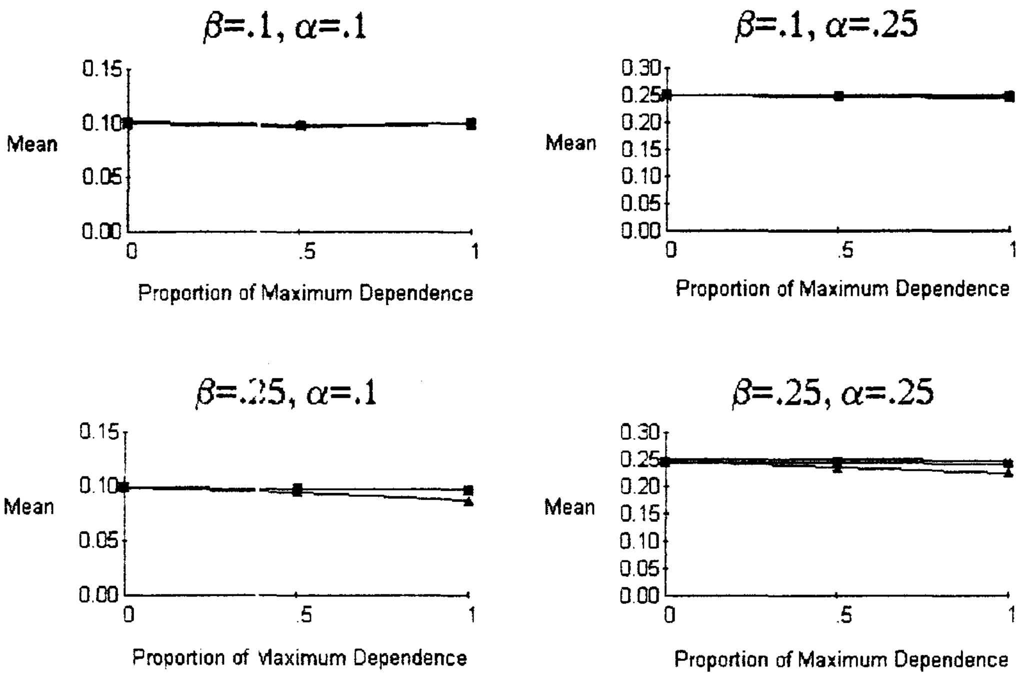
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 8f

Mean of Estimates of false positive rate for observer 1 (α_1) in the case of four observers, with dependence between observers 1 and 2 with respect to false negative rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

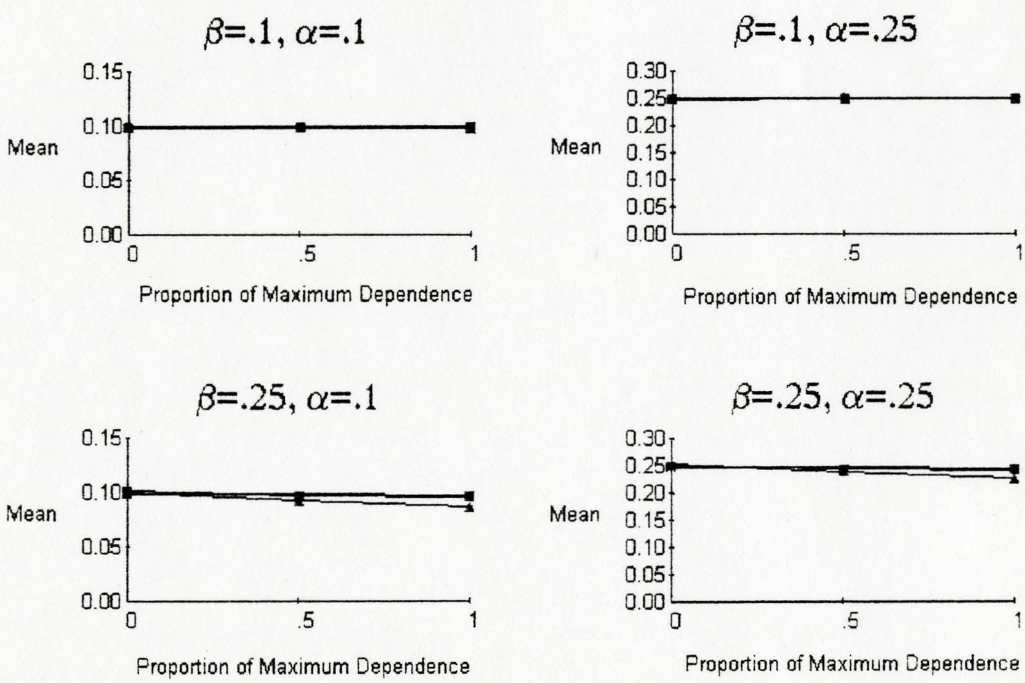
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 8g

Mean of Estimates of false positive rate for observer 2 (α_2) in the case of four observers, with dependence between observers 1 and 2 with respect to false negative rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

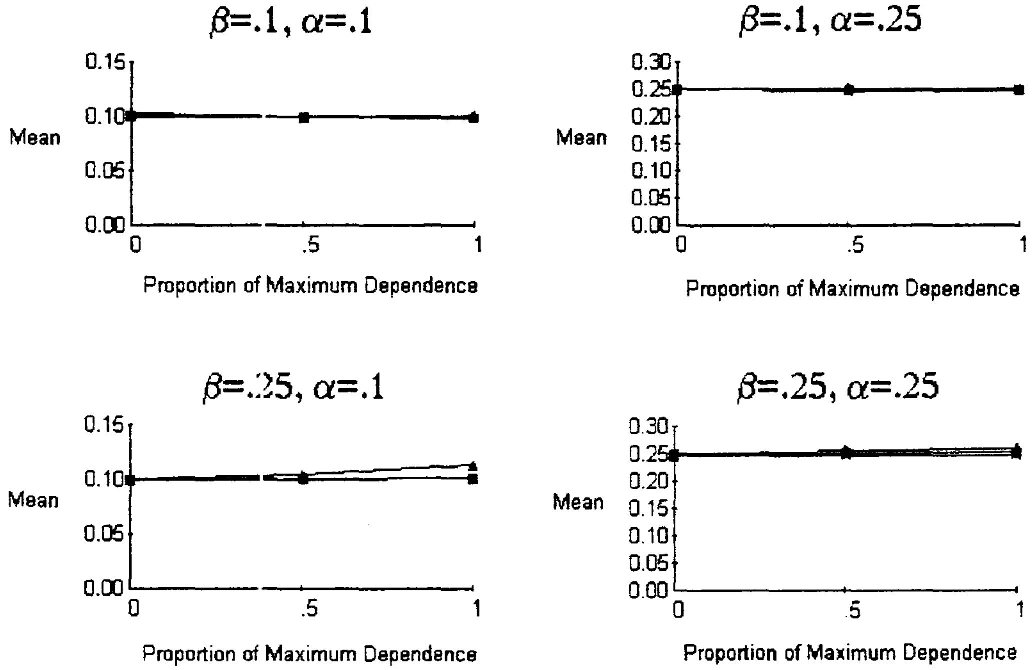
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 8h

Mean of Estimates of false positive rate for observer 3 (α_3) in the case of four observers, with dependence between observers 1 and 2 with respect to false negative rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

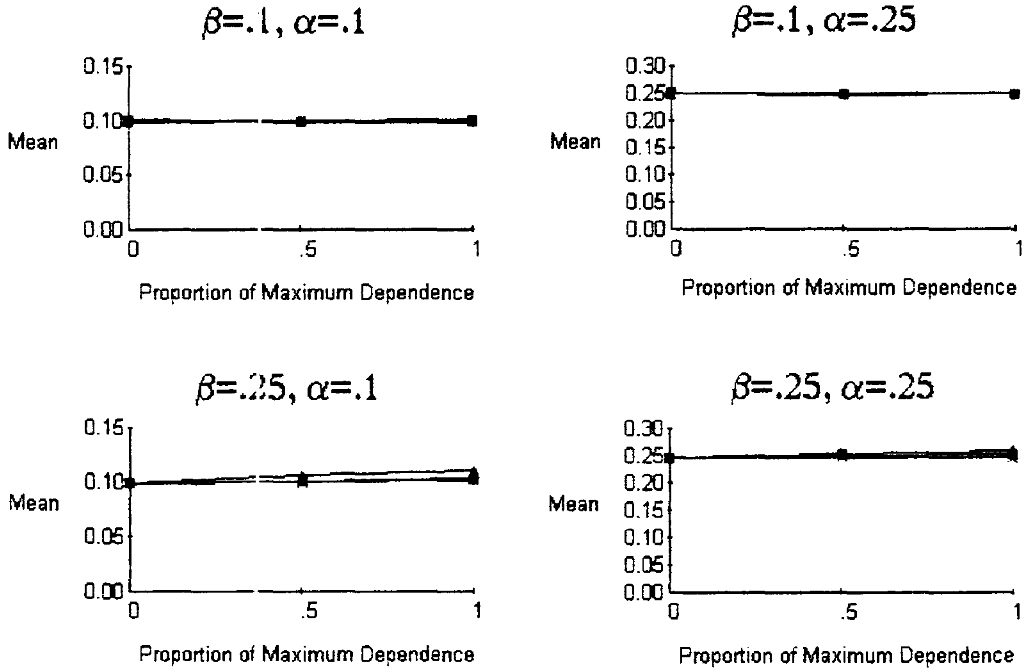
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 8i

Mean of Estimates of false positive rate for observer 4 (α_4) in the case of four observers, with dependence between observers 1 and 2 with respect to false negative rate only: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

3.1.9 RUN 9 - Dependence between two observers with respect to false positive and false negative rates - Four Observer Case

In this simulation observers 1 and 2 are dependent with respect to false positive and false negative rates. The results are shown in Tables 9a, 9b, and 9c for θ=.05, θ=.15, and θ=.40, respectively. Graphs of the means of the parameter estimates versus dependence are found in Figures 9a to 9i.

OBSERVATIONS

Effect on prevalence estimates

Figure 9a shows that the only potentially serious bias occurs when the FPR is high and the prevalence is low. In these cases the prevalence is overestimated in the presence of dependence.

Effect on the false negative rate estimates

The false negative rate estimates for the dependent observers are negatively biased. An inverse effect is seen on the two remaining independent observers, whose false negative rates are overestimated in the presence of dependence. The strength of this effect increases as the prevalence decreases and the FPR increases. Figures 9b to 9e clearly show a more serious effect for a prevalence of .05 and FPR of .25.

Effect on the false positive rate estimates

The false positive rate estimates are not affected as seriously as the false negative rate estimates. However, the trend of a negative bias for the dependent observers and a positive bias for the independent observers is still detectable. There is no apparent difference in the effect of the dependence due to prevalence.

Goodness of fit results

In this situation the χ^2 values increase with increasing dependence particularly when the FPR is high. The estimates of power as shown in Tables 9a to 9c are fairly good in most of the dependence situations. There is very little power when the prevalence is low and the FPR is low, as shown in Table 9a. There is also little power when the prevalence is higher but both the FPR and the FNR are low, see Tables 9b and 9c.

EXPLANATIONS

The effects seen here are a combination of the effects described in Run 7 and Run 8. There was not much of a bias effect in Run 8 due to the dependence with respect to false negative rate so the results in this run appear similar to those from Run 7 and are due to dependence with respect to false positive rate. This is the dependence acting on the true negative individuals. Run 3 is the corresponding three observer run. Again the addition of the fourth observer who is independent of the others decreases the strength of the effects due to the dependence.

IMPLICATIONS

The implications from Run 7 hold here. There is danger of substantially biased estimates if this form of dependence exists and there is a low prevalence and a high FPR. In addition, as stated in the earlier runs, if two of three observers are suspected to be dependent then the addition of a fourth will help to improve the estimates, although the estimates will remain biased in some situations. The goodness of fit test has better power at detecting this particular departure from the model assumptions than in Runs 7 and 8 individually.

TABLE 9a

Means and standard deviations of parameter estimates in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rates

True Prevalence = .05

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
		θ	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean χ^2	Power†
β =.1 α =.1	0	4.94 (0.35)	10.20 (3.34)	10.54 (2.86)	11.20 (2.56)	10.26 (2.68)	9.91 (0.42)	10.03 (0.53)	10.03 (0.51)	9.90 (0.54)	5.81	0.07
	0.5	5.18 (0.33)	9.40 (2.41)	8.96 (2.02)	10.82 (2.51)	11.53 (2.75)	9.96 (0.44)	9.91 (0.44)	9.93 (0.41)	10.06 (0.44)	7.24	0.07
	1	4.99 (0.38)	**7.40 (1.93)**	**7.53 (1.99)**	**12.60 (2.28)**	**12.24 (3.29)**	9.69 (0.35)	9.73 (0.47)	10.08 (0.40)	9.99 (0.61)	9.09	0.20
β =.1 α =.25	0	5.01 (0.81)	10.54 (5.49)	9.73 (5.14)	9.48 (5.42)	9.77 (5.23)	25.04 (0.86)	24.84 (0.75)	24.90 (0.79)	24.98 (0.76)	5.93	0.05
	0.5	**6.90 (0.54)**	**0.16 (0.54)**	**0.16 (0.59)**	**26.31 (4.41)**	**25.00 (4.28)**	22.85 (0.66)	22.92 (0.99)	24.95 (0.69)	24.80 (0.78)	25.51	0.91
	1	**10.57 (0.71)**	**0.00 (0.00)**	**0.00 (0.00)**	**37.93 (3.18)**	**38.33 (2.75)**	**19.73 (0.74)**	**19.68 (0.69)**	24.31 (0.89)	24.23 (0.73)	80.58	1.00
β =.25 α =.1	0	4.83 (0.41)	23.68 (3.97)	24.53 (3.77)	24.11 (4.15)	25.00 (4.20)	10.05 (0.42)	10.11 (0.44)	10.04 (0.40)	10.10 (0.42)	6.54	0.13
	0.5	4.98 (0.48)	21.38 (4.65)	21.59 (3.75)	26.05 (5.53)	26.26 (4.31)	9.70 (0.54)	9.68 (0.51)	10.11 (0.48)	10.16 (0.44)	6.61	0.07
	1	4.87 (0.44)	**18.24 (3.28)**	**16.94 (4.33)**	27.88 (4.15)	28.65 (5.06)	9.66 (0.47)	9.58 (0.56)	10.23 (0.55)	10.21 (0.54)	10.31	0.30
β =.25 α =.25	0	5.00 (1.27)	24.19 (8.58)	22.94 (7.51)	24.39 (8.06)	23.71 (9.19)	25.07 (0.69)	25.00 (0.83)	24.89 (0.89)	24.85 (0.93)	5.94	0.07
	0.5	**6.29 (0.89)**	**1.88 (4.58)**	**1.31 (3.11)**	**39.53 (4.95)**	**40.54 (3.39)**	22.91 (1.19)	22.89 (0.67)	25.38 (0.71)	25.21 (0.67)	17.25	0.73
	1	**9.84 (0.59)**	**0.00 (0.00)**	**0.00 (0.00)**	**48.45 (2.32)**	**47.55 (2.57)**	**19.61 (0.73)**	**19.49 (0.69)**	24.80 (0.85)	24.95 (0.72)	60.88	1.00

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.

† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6}$ = 12.59.

β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

Means with 20% or greater bias are displayed in bold.

TABLE 9b

Means and standard deviations of parameter estimates in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rates

True Prevalence = .15

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
		θ	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean χ^2	Power†
β =.1 α =.1	0	14.99 (0.60)	10.04 (0.88)	9.72 (1.41)	9.66 (1.59)	10.52 (1.72)	9.81 (0.47)	10.08 (0.60)	9.97 (0.50)	10.06 (0.54)	6.97	0.07
	0.5	14.90 (0.67)	9.25 (1.25)	9.41 (1.00)	10.09 (1.54)	10.18 (1.44)	10.03 (0.61)	9.77 (0.47)	10.07 (0.45)	10.09 (0.70)	6.76	0.10
	1	15.08 (0.51)	9.36 (1.74)	8.94 (1.58)	10.27 (1.46)	10.46 (1.41)	9.82 (0.55)	9.74 (0.53)	10.05 (0.48)	10.19 (0.51)	7.62	0.13
β =.1 α =.25	0	15.07 (0.87)	9.44 (2.08)	10.61 (2.27)	10.25 (2.25)	10.25 (1.91)	24.82 (0.76)	25.19 (0.93)	25.05 (0.80)	25.07 (0.82)	6.32	0.03
	0.5	16.06 (0.82)	**5.25 (1.55)**	**5.37 (1.78)**	**15.25 (1.97)**	**14.87 (1.85)**	23.26 (0.75)	23.34 (0.90)	25.17 (0.83)	25.33 (0.87)	23.60	0.93
	1	17.15 (0.79)	**0.84 (1.03)**	**0.66 (0.89)**	**20.73 (2.34)**	**20.53 (2.52)**	21.25 (0.69)	21.36 (0.67)	25.69 (0.84)	25.77 (0.78)	65.97	1.00
β =.25 α =.1	0	14.80 (0.75)	23.90 (2.29)	24.67 (1.94)	23.88 (2.52)	26.11 (2.00)	9.94 (0.66)	9.96 (0.64)	10.18 (0.51)	10.14 (0.60)	6.52	0.13
	0.5	15.11 (0.76)	22.95 (2.28)	23.32 (2.21)	25.43 (2.44)	25.25 (1.89)	9.66 (0.53)	9.77 (0.71)	10.25 (0.66)	10.15 (0.57)	10.72	0.40
	1	15.07 (0.76)	20.87 (2.44)	20.47 (2.22)	26.14 (2.93)	26.45 (2.29)	9.37 (0.61)	9.43 (0.55)	10.41 (0.56)	10.54 (0.48)	19.32	0.73
β =.25 α =.25	0	15.01 (1.84)	25.23 (4.33)	25.06 (4.26)	24.53 (4.19)	24.93 (3.59)	25.00 (0.76)	25.30 (0.89)	25.01 (0.93)	25.14 (1.05)	6.86	0.10
	0.5	15.37 (1.73)	**13.46 (3.61)**	**13.58 (3.52)**	**31.69 (3.44)**	**30.87 (2.79)**	22.64 (1.15)	22.72 (1.22)	25.84 (0.83)	25.75 (0.80)	19.86	0.87
	1	16.16 (0.95)	**3.59 (2.45)**	**4.08 (2.94)**	**38.25 (2.53)**	**38.33 (2.07)**	20.26 (0.85)	20.27 (0.85)	26.99 (0.71)	26.83 (0.67)	42.37	1.00

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.

† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6}$ = 12.59.

β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

Means with 20% or greater bias are displayed in bold.

TABLE 9c
Means and standard deviations of parameter estimates in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rates

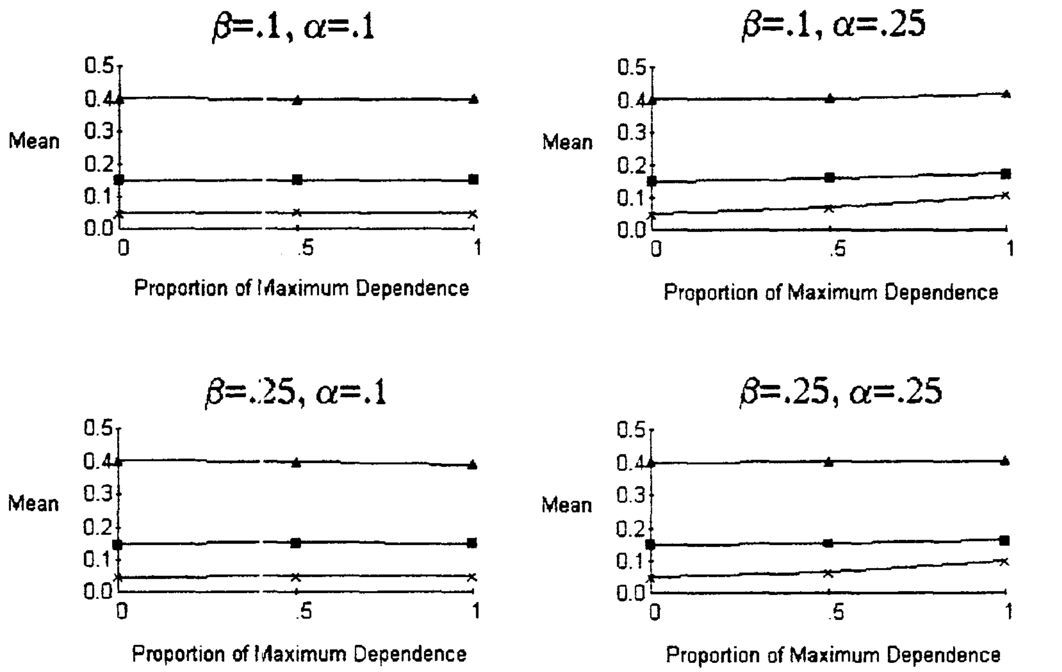
True Prevalence = .40

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
		θ	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean χ^2	Power†
$\beta=.1$ $\alpha=.1$	0	40.14 (0.90)	10.14 (0.96)	9.83 (0.71)	10.00 (0.78)	10.08 (0.95)	9.96 (0.72)	9.99 (0.64)	9.92 (0.66)	10.02 (0.68)	5.70	0.00
	0.5	39.79 (0.67)	9.66 (0.82)	9.99 (0.89)	10.05 (0.70)	9.85 (0.83)	9.90 (0.73)	10.14 (0.51)	10.14 (0.63)	10.20 (0.65)	6.37	0.10
	1	40.11 (0.89)	9.53 (0.78)	9.32 (0.85)	10.20 (0.79)	10.00 (0.83)	9.77 (0.77)	9.70 (0.51)	10.22 (0.54)	10.04 (0.58)	8.21	0.10
$\beta=.1$ $\alpha=.25$	0	40.24 (0.93)	9.99 (0.99)	10.07 (0.96)	10.44 (1.05)	9.84 (1.01)	25.05 (0.83)	25.12 (1.12)	25.18 (1.25)	24.73 (0.86)	5.48	0.00
	0.5	40.55 (0.86)	8.79 (1.07)	8.65 (0.89)	11.71 (1.10)	11.39 (1.17)	22.97 (0.89)	23.00 (1.23)	25.33 (0.86)	25.40 (0.87)	16.63	0.60
	1	41.80 (1.04)	**7.41 (0.70)**	**7.27 (0.89)**	**12.87 (1.09)**	**13.05 (1.12)**	21.20 (0.95)	21.52 (0.93)	25.54 (1.09)	25.45 (1.20)	50.76	1.00
$\beta=.25$ $\alpha=.1$	0	40.26 (0.90)	25.23 (1.46)	25.38 (1.50)	24.94 (1.40)	25.35 (1.19)	10.19 (0.85)	9.87 (0.61)	9.91 (0.66)	9.85 (0.93)	5.80	0.00
	0.5	39.71 (1.14)	23.63 (1.01)	23.54 (1.34)	25.66 (1.29)	25.16 (1.12)	9.30 (0.79)	9.11 (0.91)	10.77 (0.73)	10.66 (0.80)	12.55	0.43
	1	39.14 (1.08)	21.10 (1.11)	21.65 (1.31)	25.76 (1.30)	25.63 (1.45)	8.62 (0.70)	8.82 (0.72)	11.50 (0.76)	11.67 (0.93)	37.46	1.00
$\beta=.25$ $\alpha=.25$	0	39.98 (1.59)	24.71 (1.44)	24.87 (1.49)	25.44 (2.02)	25.20 (1.78)	25.18 (1.50)	24.88 (1.21)	24.98 (1.40)	24.72 (1.05)	5.60	0.07
	0.5	40.19 (1.69)	20.79 (1.64)	20.90 (1.59)	27.28 (1.61)	27.44 (1.61)	22.08 (1.31)	22.43 (1.29)	26.49 (0.82)	26.24 (1.41)	22.82	0.87
	1	40.47 (1.50)	**16.36 (1.69)**	**15.89 (2.01)**	**30.14 (1.50)**	**30.31 (1.68)**	**18.47 (1.29)**	**18.64 (1.23)**	28.20 (1.07)	28.36 (1.01)	57.58	1.00

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.
† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6} = 12.59$.
β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.
Means with 20% or greater bias are displayed in bold.

Figure 9a

Mean of Estimates of Prevalence (θ)
in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

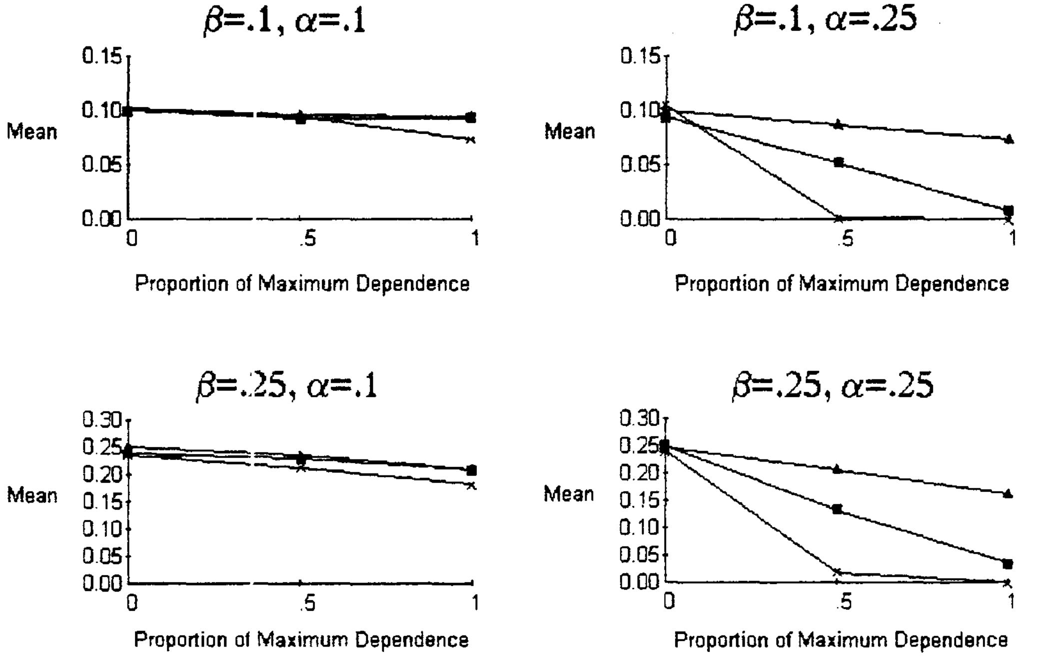
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 9b

Mean of Estimates of false negative rate for observer 1 (β_1) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

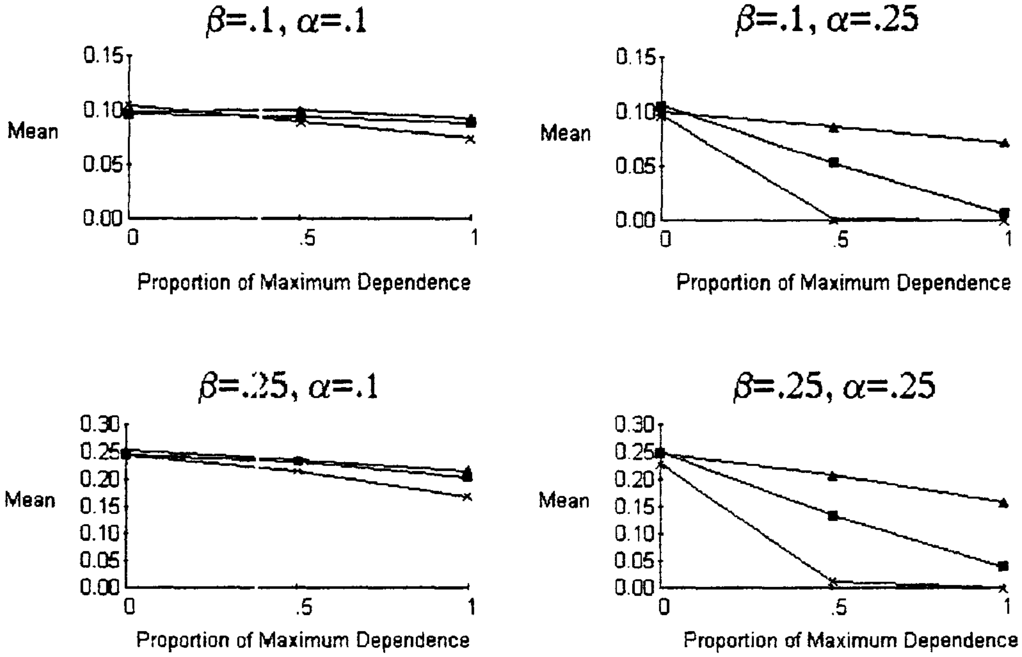
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 9c

Mean of Estimates of false negative rate for observer 2 (β_2) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

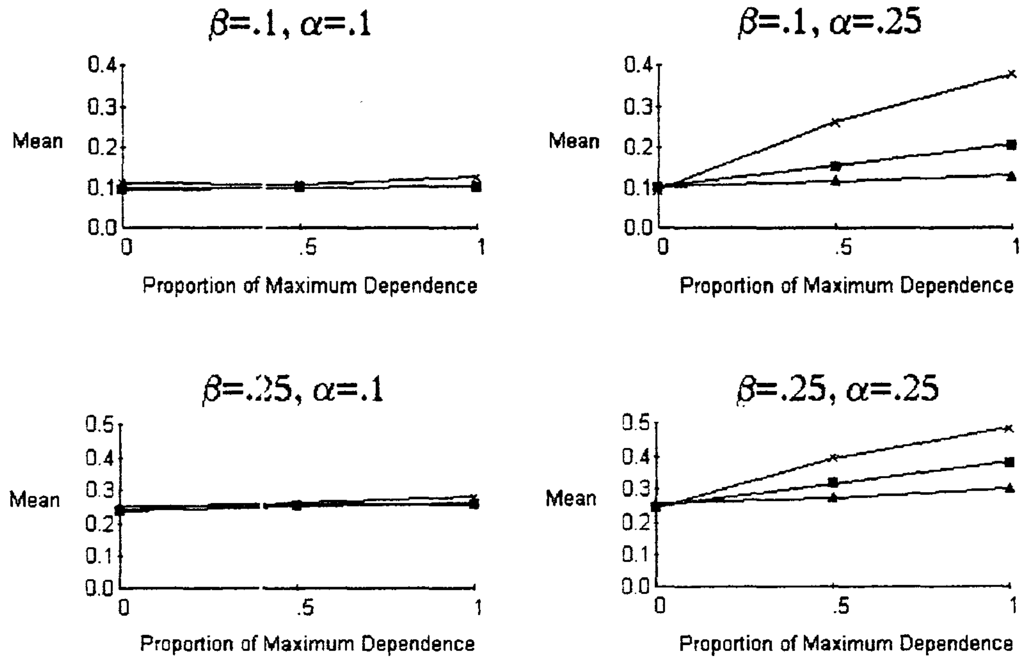
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 9d

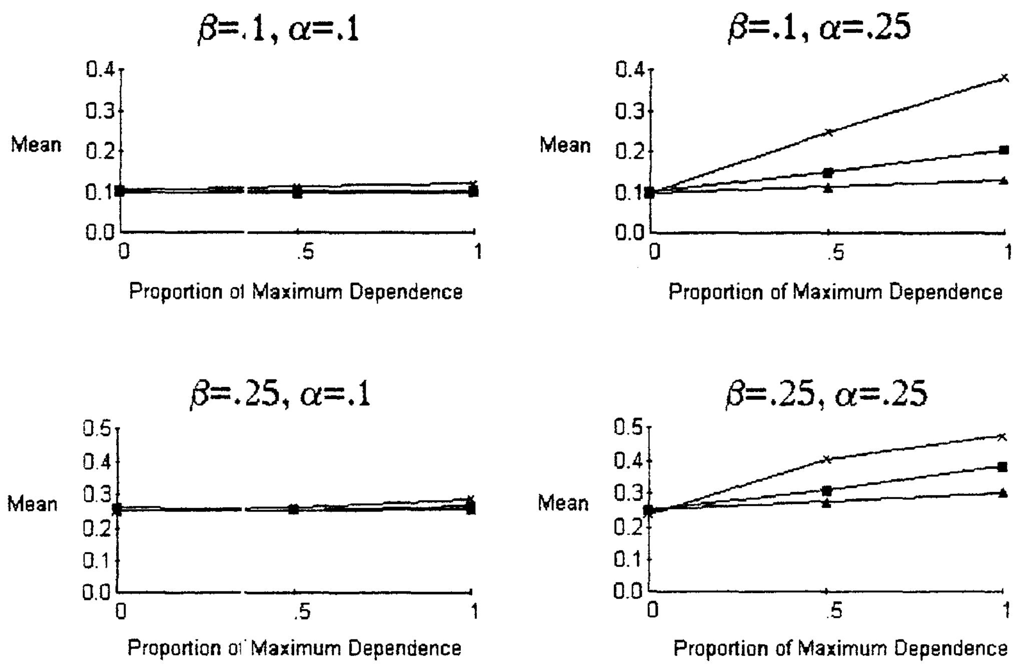
Mean of Estimates of false negative rate for observer 3 (β_3) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.
× represents true prevalence (θ) = .05.
■ represents true prevalence (θ) = .15.
▲ represents true prevalence (θ) = .40.

Figure 9e

Mean of Estimates of false negative rate for observer 4 (β_4) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

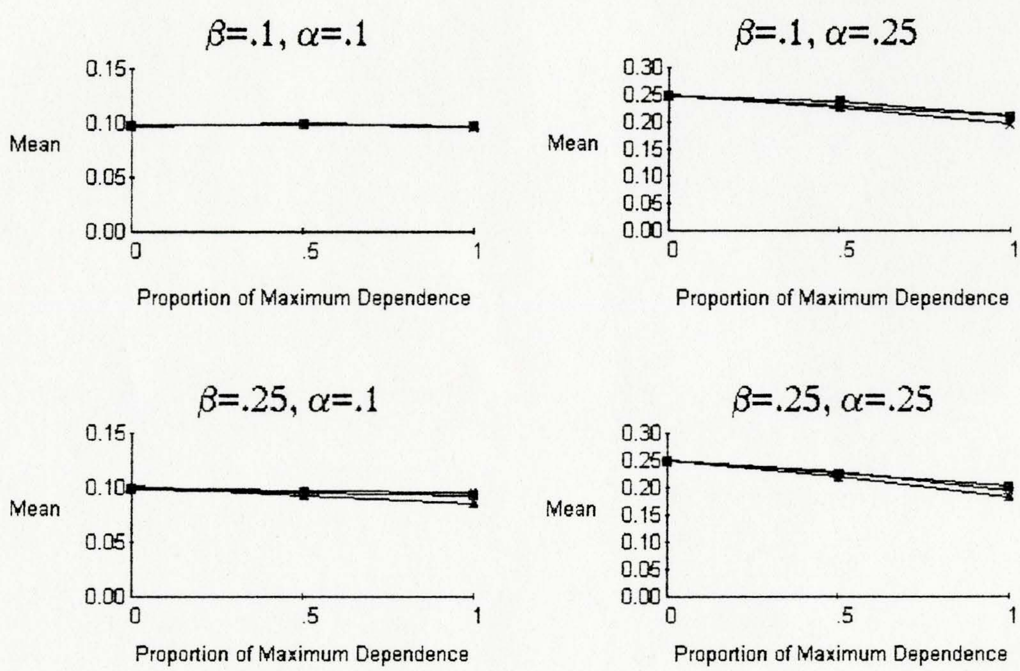
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 9f

Mean of Estimates of false positive rate for observer 1 (α_1) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

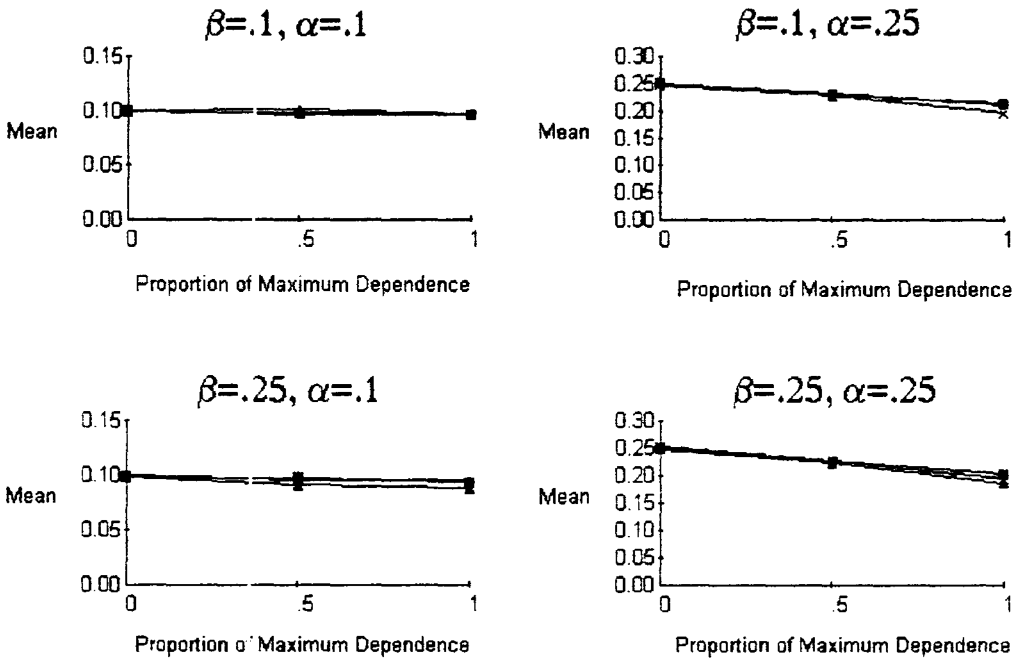
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 9g

Mean of Estimates of false positive rate for observer 2 (α_2) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

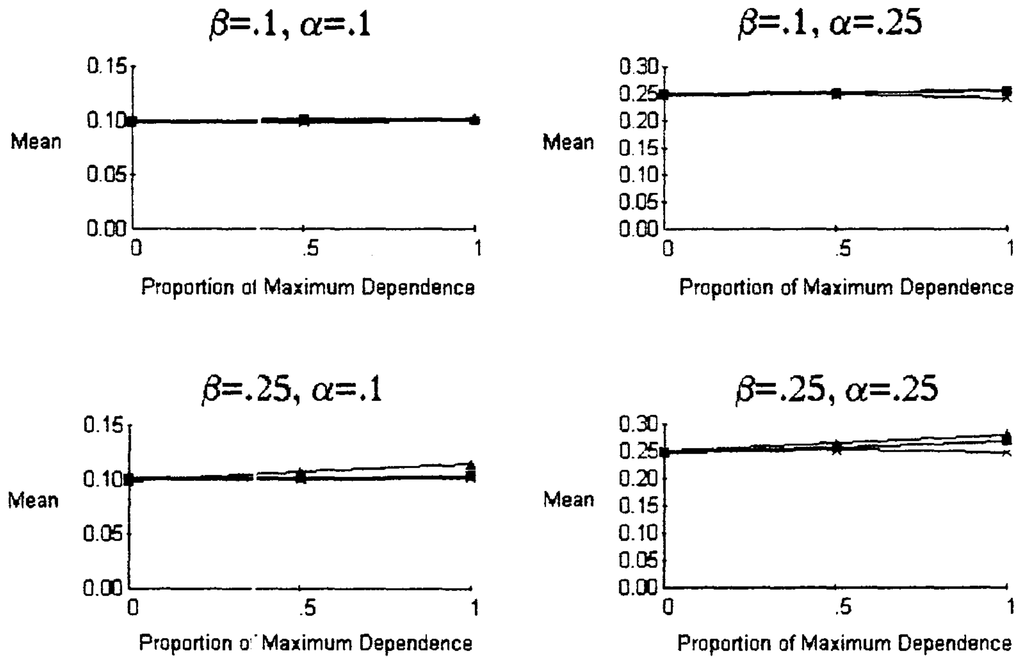
× represents true prevalence (θ) = .05.

▪ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 9h

Mean of Estimates of false positive rate for observer 3 (α_3) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

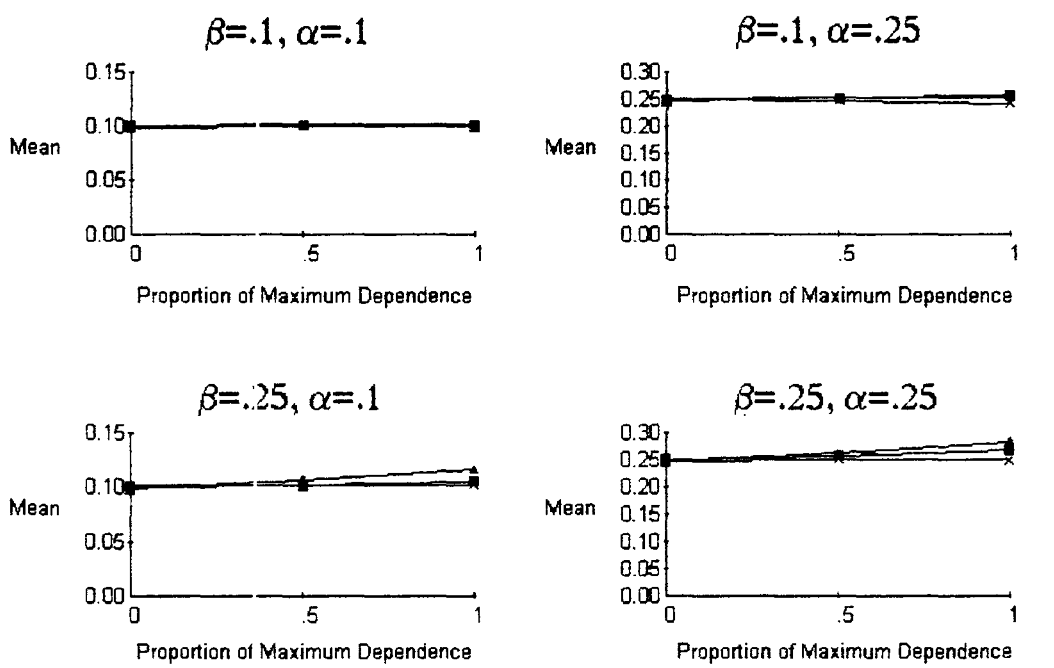
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 9i

Mean of Estimates of false positive rate for observer 4 (α_4) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

3.1.10 RUN 10 - Dependence between two accurate observers - Four Observer Case

This run includes two dependent observers who are fairly accurate with error rates of .1 and two independent observers who have higher error rates of .25. Tables 10a, 10b, and 10c display the means of the parameter estimates for prevalences of .05, .15, and .40, respectively. The graphs of the means of the parameter estimates versus dependence are in Figures 10a to 10i.

OBSERVATIONS

Effect on prevalence estimates

Estimates of the low prevalence of .05 are positively biased in the presence of dependence with a mean value of .0661 when the dependence is at maximum. No apparent bias is shown in the estimates of larger prevalences.

Effect on false negative rate estimates

The false negative rates for the two dependent observers are substantially underestimated for all prevalences. The underestimation is more severe for lower prevalences. An inverse effect is seen on the independent observers whose false negative rates are overestimated.

Effect on false positive rate estimates

There is a negative bias in the estimates of false positive rate for the dependent observers. The bias is around the 20% clinically significant level. Any effect on the false positive rates for the independent observers is not detectable.

Goodness of Fit Results

The goodness of fit results seem to indicate that the test has some power to detect the departure from model assumptions when the dependence is at maximum. The goodness of fit test is lacking power when the dependence is only at one half of maximum.

EXPLANATIONS

This run shows similar patterns to the three observer run and the bias can be explained in the same way. Again the addition of the fourth independent observer reduces the bias due to the dependence between two of the observers.

IMPLICATIONS

This situation is most dangerous when the true prevalence is low. Here again is another example of where the addition of a fourth observer will result in more accurate estimates if two of the observers are suspected to be dependent. The goodness of fit test appears to be useful if the dependence is close to maximum.

TABLE 10a

Means and standard deviations of parameter estimates in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rates

True Prevalence = .05

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
		θ	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean χ^2	Power†
$\beta_1,\beta_2=.1$ $\beta_3,\beta_4=.25$ $\alpha_1,\alpha_2=.1$ $\alpha_3,\alpha_4=.25$	0	5.12 (0.59)	10.80 (5.23)	10.81 (5.97)	25.19 (4.83)	25.47 (4.72)	10.05 (0.61)	9.93 (0.54)	24.93 (0.63)	24.97 (0.66)	6.13	0.10
	0.5	5.38 (0.49)	**1.44 (2.41)**	**1.61 (2.68)**	**31.01 (4.16)**	**30.57 (3.63)**	9.10 (0.54)	9.15 (0.42)	25.00 (0.75)	25.14 (0.73)	8.42	0.15
	1	**6.61 (0.42)**	**0.00 (0.00)**	**0.00 (0.00)**	**36.31 (2.59)**	**35.05 (3.46)**	**7.82 (0.46)**	**7.98 (0.41)**	24.94 (0.69)	25.11 (0.75)	27.24	0.98

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.

† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6}$ = 12.59.

β represents the true false negative rates, $\beta_1=\beta_2=\beta_3=\beta_4$.

α represents the true false positive rates, $\alpha_1=\alpha_2=\alpha_3=\alpha_4$.

Means with 20% or greater bias are displayed in bold.

TABLE 10b

Means and standard deviations of parameter estimates in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rates

True Prevalence = .15

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
		$\hat{\theta}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$	Mean χ^2	Power†
$\beta_1,\beta_2=.1$ $\beta_3,\beta_4=.25$	0	14.75 (0.61)	9.62 (2.01)	9.17 (2.09)	24.54 (1.87)	24.71 (2.19)	10.15 (0.65)	10.06 (0.52)	25.14 (0.80)	24.99 (0.73)	5.96	0.03
$\alpha_1,\alpha_2=.1$	0.5	15.41 (0.69)	**6.39 (2.28)**	**6.62 (1.90)**	27.94 (2.06)	27.64 (1.99)	9.05 (0.68)	9.10 (0.73)	25.26 (0.93)	25.31 (0.71)	9.13	0.17
$\alpha_3,\alpha_4=.25$	1	15.31 (0.68)	**2.25 (1.35)**	**2.47 (1.26)**	29.58 (2.13)	29.32 (1.97)	8.40 (0.60)	8.30 (0.43)	25.62 (0.70)	25.73 (0.86)	13.55	0.37

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.

† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6} = 12.59$.

β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

Means with 20% or greater bias are displayed in bold.

TABLE 10c

Means and standard deviations of parameter estimates in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rates

True Prevalence = .40

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
		θ	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean χ^2	Power†
β_1,β_2=.1 β_3,β_4=.25	0	40.11 (0.91)	9.98 (1.00)	9.87 (1.04)	25.34 (1.08)	25.13 (1.13)	9.92 (0.63)	9.72 (0.67)	24.96 (0.67)	25.25 (0.95)	6.70	0.07
α_1,α_2=.1	0.5	39.95 (0.78)	8.23 (1.05)	8.28 (1.08)	26.02 (1.11)	25.78 (1.22)	8.99 (0.68)	8.85 (0.67)	25.36 (1.07)	25.43 (0.68)	6.80	0.10
α_3,α_4=.25	1	40.23 (0.91)	**6.84 (1.11)**	**6.63 (0.79)**	26.32 (1.14)	26.78 (1.20)	**7.73 (0.63)**	**7.74 (0.61)**	26.15 (0.86)	25.90 (0.97)	13.83	0.57

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.

† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6}$ = 12.59.

β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

Means with 20% or greater bias are displayed in bold.

Figure 10a

Mean of Estimates of Prevalence (θ)
in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .1, \beta_3, \beta_4 = .25; \quad \alpha_1, \alpha_2 = .1, \alpha_3, \alpha_4 = .25$

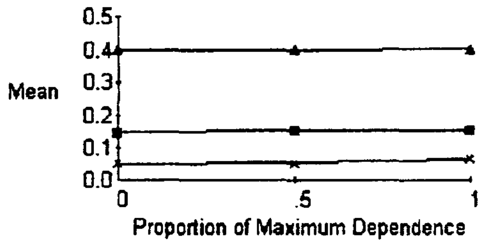
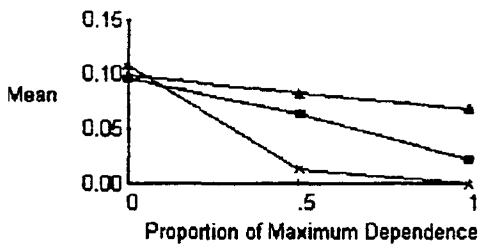


Figure 10b

Mean of Estimates of false negative rate for observer 1 (β_1)
in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .1, \beta_3, \beta_4 = .25; \quad \alpha_1, \alpha_2 = .1, \alpha_3, \alpha_4 = .25$



× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 10c

Mean of Estimates of false negative rate for observer 2 (β_2) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .1, \beta_3, \beta_4 = .25; \quad \alpha_1, \alpha_2 = .1, \alpha_3, \alpha_4 = .25$

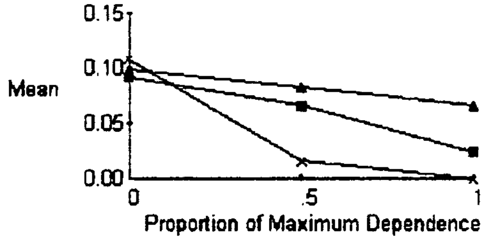
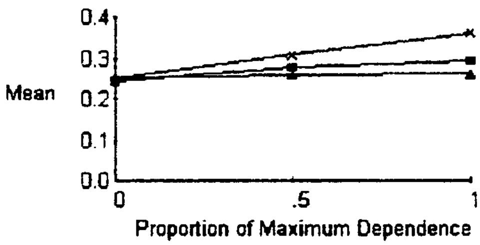


Figure 10d

Mean of Estimates of false negative rate for observer 3 (β_3) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .1, \beta_3, \beta_4 = .25; \quad \alpha_1, \alpha_2 = .1, \alpha_3, \alpha_4 = .25$



× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 10e

Mean of Estimates of false negative rate for observer 4 (β_4) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .1, \beta_3, \beta_4 = .25; \quad \alpha_1, \alpha_2 = .1, \alpha_3, \alpha_4 = .25$

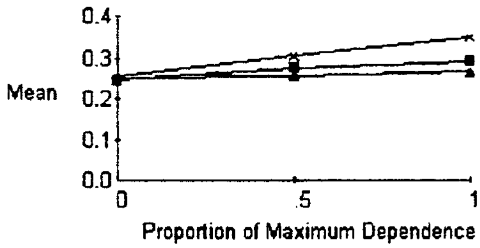
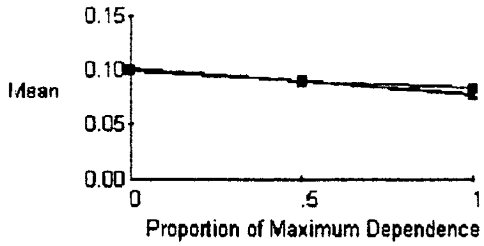


Figure 10f

Mean of Estimates of false positive rate for observer 1 (α_1) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .1, \beta_3, \beta_4 = .25; \quad \alpha_1, \alpha_2 = .1, \alpha_3, \alpha_4 = .25$



× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 10g

Mean of Estimates of false positive rate for observer 2 (α_2) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .1, \beta_3, \beta_4 = .25; \quad \alpha_1, \alpha_2 = .1, \alpha_3, \alpha_4 = .25$

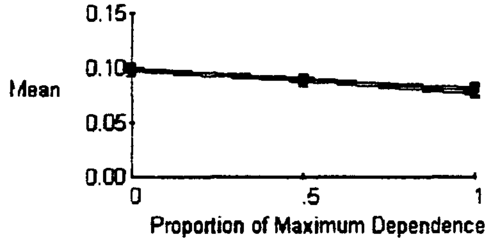
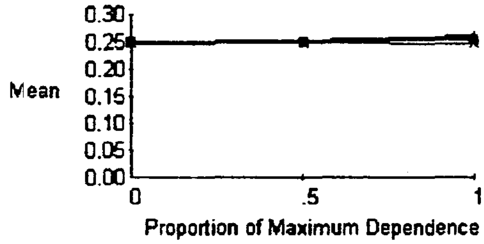


Figure 10h

Mean of Estimates of false positive rate for observer 3 (α_3) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .1, \beta_3, \beta_4 = .25; \quad \alpha_1, \alpha_2 = .1, \alpha_3, \alpha_4 = .25$



× represents true prevalence (θ) = .05.

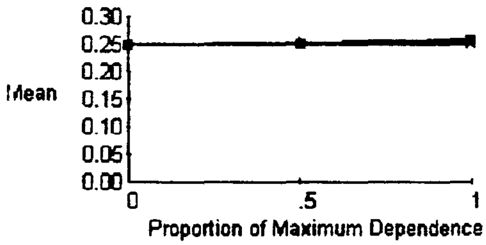
■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 10i

Mean of Estimates of false positive rate for observer 4 (α_4) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .1,\ \beta_3, \beta_4 = .25;\quad \alpha_1, \alpha_2 = .1,\ \alpha_3, \alpha_4 = .25$



× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

3.1.11 RUN 11 - Dependence between two inaccurate observers - 4 observers

This run is similar to the last with the exception that the two dependent observers have high error rates at .25 while the independent observers are more accurate with error rates of .1. The means of the parameter estimates are displayed in Tables 11a to 11c and graphed in Figures 11a to 11i.

OBSERVATIONS

Effect on prevalence estimates

Tables 11a to 11c and Figure 11a show no indications of bias in the estimates of prevalence when the dependence is present.

Effect on false negative rate estimates

The false negative rates for the dependent observers are underestimated when the prevalence is low at .05 The false negative rates for the independent observers behave in the opposite way. They are overestimated most severely when the prevalence is low.

Effect on false positive rate estimates

There is no apparent bias in the estimates of false positive rates when the dependence is present, as can be seen by the straight horizontal lines in Figures 11f to 11i.

Goodness of fit results

The power of the goodness of fit test is fairly strong when the dependence is at maximum. There is not much power when the dependence is at one half of maximum.

EXPLANATIONS

This situation can be explained in the same way as the equivalent three observer case, Run 5. The increased agreement between the dependent observers decreases their error rate estimates. It is interesting to note that the addition of the fourth observer in this case seems to improve the estimates more than in the last run. In this run the fourth observer is a more accurate observer with error rates of only .1 compared to the fourth observer added in the last run with error rates of .25.

IMPLICATIONS

In a situation where the two dependent observers are not very accurate the addition of an independent accurate observer substantially improves the estimates. The most dangerous situation remains to be when the prevalence is low. In the three observer case, maximum dependence caused the less sensitive observers to appear to be the more sensitive observers. In this four observer run the sensitivities of the observers maintain their ordinality. The results show that the goodness of fit test cannot always be relied upon to detect the dependence.

TABLE 11a

Means and standard deviations of parameter estimates in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rates

True Prevalence = .05

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
		θ	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean χ^2	Power†
β_1,β_2=.25 β_3,β_4=.1	0	5.01 (0.63)	25.80 (3.88)	24.99 (4.60)	9.65 (5.49)	10.64 (5.70)	25.11 (0.70)	25.14 (0.75)	10.09 (0.54)	9.99 (0.66)	5.42	0.02
α_1,α_2=.25	0.5	5.31 (0.60)	22.90 (4.26)	22.22 (4.42)	**13.35 (5.10)**	**13.33 (4.76)**	24.86 (0.71)	24.88 (0.81)	10.18 (0.49)	9.93 (0.60)	9.02	0.19
α_3,α_4=.1	1	5.24 (0.59)	**19.72 (4.20)**	**19.84 (3.70)**	**13.77 (4.91)**	**14.62 (4.68)**	24.55 (0.81)	24.78 (0.68)	10.02 (0.49)	10.10 (0.54)	16.02	0.61

*Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.

† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6}$ = 12.59.

β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

Means with 20% or greater bias are displayed in bold.

TABLE 11b

Means and standard deviations of parameter estimates in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rates

True Prevalence = .15

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
		θ	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean $\tilde{\chi}^2$	Power†
β_1,β_2=.25 β_3,β_4=.1 α_1,α_2=.25 α_3,α_4=.1	0	14.99 (0.75)	24.92 (2.49)	25.29 (1.87)	9.63 (2.04)	10.74 (2.11)	25.02 (0.60)	24.86 (0.79)	10.01 (0.60)	10.06 (0.49)	6.60	0.10
	0.5	15.24 (0.80)	24.56 (2.36)	24.60 (2.09)	11.37 (2.07)	10.92 (2.11)	24.58 (0.71)	24.84 (0.55)	10.09 (0.73)	10.09 (0.54)	8.38	0.13
	1	15.47 (0.78)	24.09 (2.24)	23.22 (2.42)	11.55 (2.19)	11.99 (2.23)	24.71 (0.65)	24.66 (0.68)	9.73 (0.67)	9.97 (0.64)	13.18	0.57

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.

† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6}$ = 12.59.

β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

Means with 20% or greater bias are displayed in bold.

TABLE 11c

Means and standard deviations of parameter estimates in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rates

True Prevalence = .40

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
		θ	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean χ^2	Power†
$\beta_1,\beta_2=.25$	0	40.05 (0.92)	24.91 (1.19)	25.18 (1.04)	9.90 (1.43)	10.03 (1.14)	25.12 (0.97)	25.00 (0.84)	9.85 (0.92)	9.91 (1.01)	6.40	0.03
$\beta_3,\beta_4=.1$ $\alpha_1,\alpha_2=.25$	0.5	40.06 (1.15)	24.77 (1.21)	24.95 (1.26)	9.67 (0.96)	10.36 (1.05)	24.76 (1.17)	24.72 (1.10)	9.66 (0.73)	10.17 (0.78)	8.57	0.27
$\alpha_3,\alpha_4=.1$	1	40.22 (0.85)	24.61 (1.42)	24.42 (0.99)	10.45 (0.95)	9.95 (1.14)	24.61 (0.97)	24.59 (0.90)	9.85 (0.89)	9.85 (0.65)	12.36	0.43

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.

† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6} = 12.59$.

β represents the true false negative rates, $\beta_1= \beta_2= \beta_3= \beta_4$.

α represents the true false positive rates, $\alpha_1= \alpha_2= \alpha_3= \alpha_4$.

Means with 20% or greater bias are displayed in bold.

Figure 11a

Mean of Estimates of Prevalence (θ)
in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .25, \ \beta_3, \beta_4 = .1; \quad \alpha_1, \alpha_2 = .25, \alpha_3, \alpha_4 = .1$

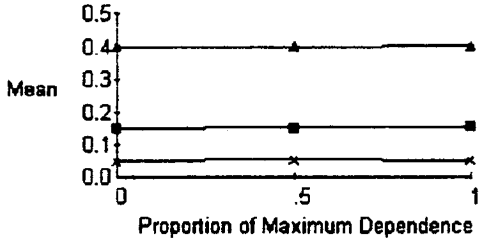
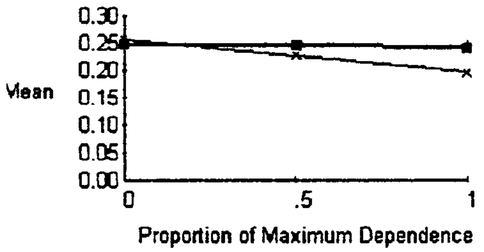


Figure 11b

Mean of Estimates of false negative rate for observer 1 (β_1)
in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .25, \ \beta_3, \beta_4 = .1; \quad \alpha_1, \alpha_2 = .25, \alpha_3, \alpha_4 = .1$



× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 11 c

Mean of Estimates of false negative rate for observer 2 (β_2) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .25, \ \beta_3, \beta_4 = .1; \quad \alpha_1, \alpha_2 = .25, \alpha_3, \alpha_4 = .1$

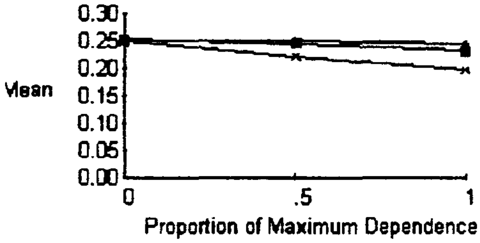
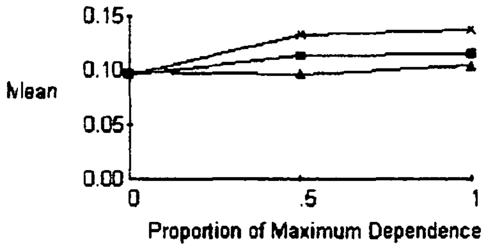


Figure 11 d

Mean of Estimates of false negative rate for observer 3 (β_3) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .25, \ \beta_3, \beta_4 = .1; \quad \alpha_1, \alpha_2 = .25, \alpha_3, \alpha_4 = .1$



× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 11 e

Mean of Estimates of false negative rate for observer 4 (β_4) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .25, \ \beta_3, \beta_4 = .1; \quad \alpha_1, \alpha_2 = .25, \ \alpha_3, \alpha_4 = .1$

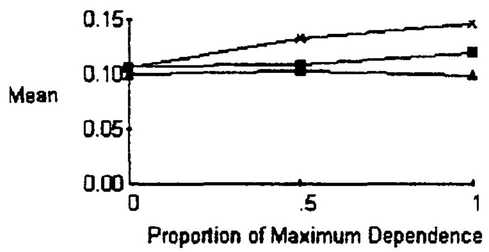
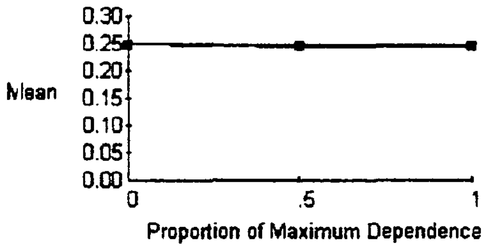


Figure 11 f

Mean of Estimates of false positive rate for observer 1 (α_1) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .25, \ \beta_3, \beta_4 = .1; \quad \alpha_1, \alpha_2 = .25, \ \alpha_3, \alpha_4 = .1$



× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 11g

Mean of Estimates of false positive rate for observer 2 (α_2) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .25, \ \beta_3, \beta_4 = .1; \quad \alpha_1, \alpha_2 = .25, \ \alpha_3, \alpha_4 = .1$

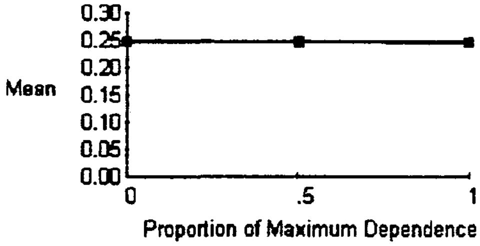
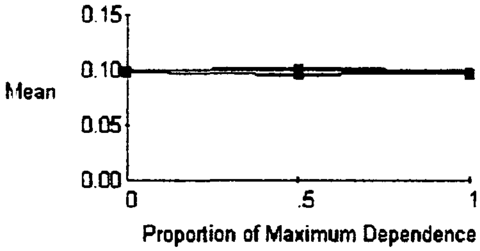


Figure 11h

Mean of Estimates of false positive rate for observer 3 (α_3) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .25, \ \beta_3, \beta_4 = .1; \quad \alpha_1, \alpha_2 = .25, \ \alpha_3, \alpha_4 = .1$



× represents true prevalence (θ) = .05.

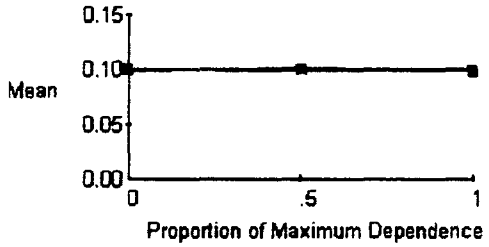
■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 11i

Mean of Estimates of false positive rate for observer 4 (α_4) in the case of four observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates

$\beta_1, \beta_2 = .25, \ \beta_3, \beta_4 = .1; \quad \alpha_1, \alpha_2 = .25, \alpha_3, \alpha_4 = .1$



× represents true prevalence (θ) = .05.

▪ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40

3.1.12 RUN 12 - Dependence between all pairs of observers - 4 observers

This run included dependence between all pairs of observers with respect to false negative and false positive rates. The results of the simulation are shown in Tables 12a, 12b, and 12c for prevalences of .05, .15, .40. Graphs of the means of the estimates are given in Figures 12a to 12i.

OBSERVATIONS

Effect on prevalence estimates

Prevalence estimates are positively biased when dependence is present. The bias is particularly severe when the true prevalence is low and when the FPR is high. Substantial decrease in the means of the prevalence estimates is seen when in Table 12c and Figure 12a where the true prevalence is .40, the FNR is .25, and the FPR is .1.

Effect on false negative rate estimates

All false negative rates are underestimated in the presence of dependence. The most severe biases are seen for low prevalence values.

Effect on false positive rate estimates

The false positive rate estimates are also underestimated when the dependence is present. The true prevalence appears to have no effect on the bias shown for these estimates.

Goodness of fit results

The goodness of fit test for this simulation has the highest power of any of the studied situations. In all cases even at one half of maximum dependence the power of the test is estimated to be 1.00.

EXPLANATIONS

Even in this situation with all pairwise dependencies present, the addition of a fourth observer appears to substantially improve the accuracy of the false negative rate estimates over the three observer case. This could be explained by the absolute values of the dependence terms added to the probabilities. In this four observer run the absolute value of the dependence term is smaller than in the three observer case. The bias in the false positive rates is not as large as the bias in the false negative rates. This shows again that the smaller number of true positive on which the false negative rate is based allows more bias into the false negative rate estimates. With the exception of the false negative rates, the addition of the fourth observer appears to have little effect.

IMPLICATIONS

Dependence between all pairs of observers will lead to substantially biased estimates. All observers appear more accurate than they truly are. Again, the most dangerous situation is when the prevalence is low. Fortunately in this situation the goodness of fit test has ample power to detect a departure from the model assumptions. A strong advantage still exists for the addition of a fourth observer, even if this observer is suspected to have pairwise dependence with the others. In addition to slightly less biased false negative rate estimates the additional observer permits the use of the goodness of fit test which in this case has very good power.

TABLE 12a

Means and standard deviations of parameter estimates in the case of four observers, with dependence between all pairs of observers with respect to false positive and false negative rates

True Prevalence = .05

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
		θ	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean χ^2	Power†
β =.1 α =.1	0	5.05 (0.29)	10.09 (2.79)	9.53 (2.58)	9.77 (2.71)	10.33 (3.06)	9.98 (0.57)	9.82 (0.39)	9.93 (0.52)	9.88 (0.52)	4.99	0.03
	0.5	**6.17 (0.42)**	**7.14 (1.48)**	**7.46 (2.23)**	8.05 (1.82)	**7.24 (2.39)**	8.88 (0.54)	8.88 (0.47)	8.74 (0.55)	8.82 (0.52)	36.89	1.00
	1	**7.29 (0.43)**	**5.64 (1.69)**	**5.88 (1.27)**	**5.90 (1.83)**	**5.43 (1.44)**	**7.62 (0.44)**	**7.72 (0.41)**	**7.69 (0.50)**	**7.57 (0.36)**	176.31	1.00
β =.1 α =.25	0	5.17 (0.79)	10.78 (5.00)	11.23 (5.16)	9.81 (5.66)	10.94 (5.71)	24.92 (0.78)	25.11 (0.65)	24.74 (0.76)	25.08 (0.88)	6.52	0.07
	0.5	**12.98 (0.66)**	9.59 (1.73)	9.30 (1.62)	9.23 (2.07)	9.53 (1.95)	**19.12 (0.68)**	**18.76 (0.64)**	**18.76 (0.69)**	**18.97 (0.68)**	63.53	1.00
	1	**18.91 (0.61)**	**6.98 (1.07)**	**6.95 (0.85)**	**6.67 (1.02)**	**6.81 (1.16)**	**13.15 (0.60)**	**13.13 (0.69)**	**12.95 (0.67)**	**13.21 (0.68)**	526.37	1.00
β =.25 α =.1	0	5.11 (0.47)	26.56 (5.96)	25.08 (5.19)	25.87 (5.38)	26.64 (4.41)	9.85 (0.52)	9.92 (0.48)	10.00 (0.63)	9.99 (0.49)	6.55	0.00
	0.5	5.42 (0.30)	**13.57 (2.80)**	**12.39 (2.18)**	**12.50 (3.03)**	**13.76 (3.00)**	8.90 (0.50)	8.92 (0.47)	8.93 (0.47)	8.98 (0.45)	31.65	1.00
	1	**6.66 (0.44)**	**8.37 (1.52)**	**8.49 (1.83)**	**9.22 (1.33)**	**9.04 (1.80)**	**7.71 (0.40)**	**7.64 (0.45)**	**7.57 (0.42)**	**7.61 (0.45)**	185.30	1.00
β =.25 α =.25	0	4.98 (1.36)	25.06 (8.12)	23.97 (7.47)	25.60 (9.63)	21.63 (7.99)	25.03 (0.78)	25.03 (0.97)	25.22 (0.86)	24.85 (0.82)	5.65	0.07
	0.5	**12.33 (0.65)**	**12.34 (2.12)**	**12.65 (2.20)**	**12.74 (2.51)**	**11.45 (1.80)**	**19.14 (0.55)**	**19.25 (0.90)**	**18.94 (0.77)**	**18.86 (0.59)**	62.45	1.00
	1	**18.23 (0.66)**	**8.21 (1.04)**	**8.15 (0.89)**	**8.08 (0.96)**	7.95 (0.93)	**13.05 (0.64)**	**13.15 (0.54)**	**13.05 (0.56)**	**13.23 (0.59)**	539.43	1.00

* Proportion of maximum dependence between observers for both true negative and true positive subjects.

† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6} = 12.59$.

β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

Means with 20% or greater bias are displayed in bold.

TABLE 12b - Dependence between all pairs of observers
with respect to False Positive and False Negative Rate

True Prevalence = .15

Truth		Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
Error Rates	Dep*	Prevalence	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean χ^2	Power†
β =.1 α =.1 ††	0	14.98 (0.64)	9.90 (1.55)	9.85 (1.69)	10.55 (1.31)	9.61 (1.61)	9.96 (0.56)	10.00 (0.52)	9.98 (0.48)	10.05 (0.56)	5.96	0.03
	0.5	15.86 (0.62)	8.20 (1.17)	8.27 (1.18)	8.04 (1.12)	7.97 (1.37)	8.81 (0.45)	8.88 (0.55)	8.83 (0.60)	8.79 (0.68)	34.57	1.00
	1	16.85 (0.59)	7.46 (0.93)	6.91 (1.10)	6.64 (0.88)	6.62 (1.14)	7.64 (0.52)	7.57 (0.42)	7.53 (0.40)	7.67 (0.53)	173.94	1.00
β =.1 α =.25	0	15.04 (0.90)	10.52 (2.46)	9.34 (2.41)	9.57 (1.60)	9.82 (1.60)	25.23 (0.87)	24.97 (1.04)	24.88 (0.64)	25.04 (0.86)	5.66	0.03
	0.5	22.18 (0.60)	9.13 (1.09)	9.22 (1.20)	9.16 (1.06)	9.44 (1.23)	19.14 (0.72)	18.94 (0.82)	18.98 (0.92)	18.92 (0.65)	57.08	1.00
	1	27.26 (0.65)	7.27 (0.68)	7.23 (0.80)	7.39 (0.96)	7.04 (0.83)	13.14 (0.59)	13.32 (0.67)	12.85 (0.49)	13.09 (0.47)	490.12	1.00
β =.25 α =.1	0	15.00 (0.67)	24.99 (2.67)	24.80 (2.38)	26.07 (2.72)	25.11 (2.37)	10.03 (0.70)	9.92 (0.58)	10.09 (0.61)	10.03 (0.66)	5.72	0.00
	0.5	14.35 (0.58)	16.97 (2.09)	16.52 (1.93)	16.41 (1.36)	16.43 (1.95)	9.01 (0.49)	9.16 (0.46)	8.95 (0.57)	8.93 (0.59)	36.39	1.00
	1	14.88 (0.50)	11.24 (1.30)	11.23 (1.28)	11.56 (1.57)	10.84 (1.38)	7.59 (0.46)	7.66 (0.45)	7.61 (0.46)	7.62 (0.39)	216.92	1.00
β =.25 α =.25	0	15.12 (1.58)	24.62 (4.33)	25.31 (3.00)	25.54 (3.63)	25.26 (3.40)	25.01 (0.99)	24.88 (0.95)	24.96 (0.80)	25.14 (0.65)	6.64	0.07
	0.5	20.71 (0.74)	15.06 (1.94)	15.57 (1.67)	14.95 (1.59)	15.04 (1.60)	18.94 (0.71)	18.78 (0.79)	19.13 (0.77)	18.91 (0.91)	60.41	1.00
	1	25.45 (0.81)	9.59 (0.97)	9.98 (0.85)	10.01 (1.06)	10.01 (1.13)	12.96 (0.53)	13.12 (0.69)	12.86 (0.66)	12.80 (0.60)	521.14	1.00

* Proportion of maximum dependence between observers for true negative and true positive subjects.

† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6}$ = 12.59

†† β and α represent β_1, β_2, and β_3 and α_1, α_2, and α_3, respectively

TABLE 12c - Dependence between all pairs of observers
with respect to False Positive and False Negative Rate

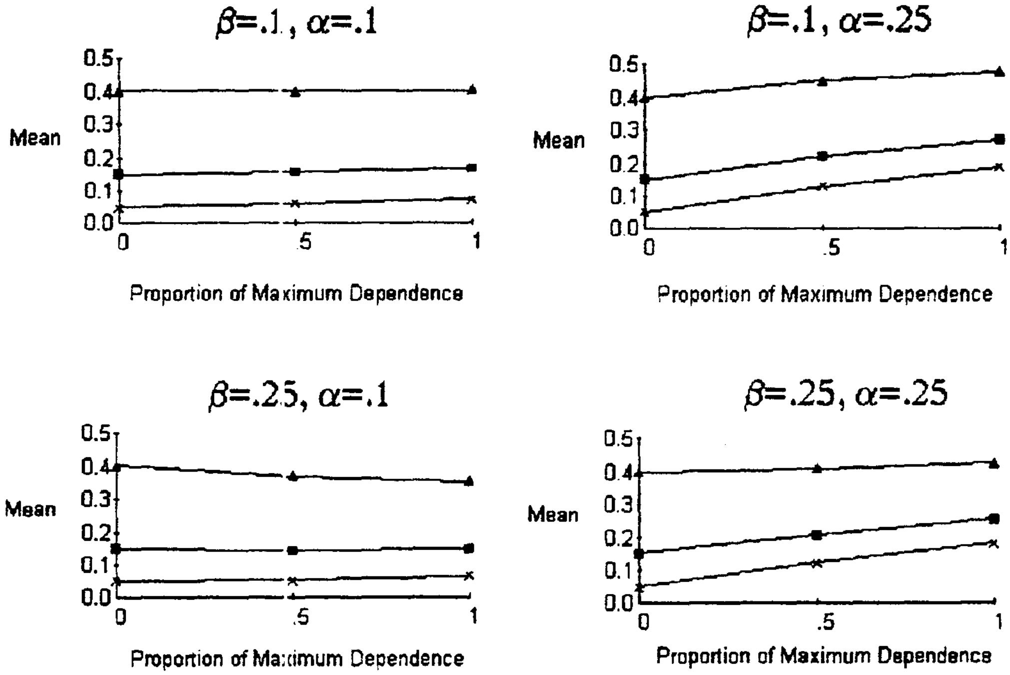
True Prevalence = .40

Truth		Mean of Estimates (%) (standard deviation (%))									Goodness of Fit	
Error Rates	Dep*	Prevalence	β_1	β_2	β_3	β_4	α_1	α_2	α_3	α_4	Mean χ^2	Power†
β =.1 α =.1 ††	0	40.14 (0.84)	9.99 (0.84)	10.03 (0.68)	9.76 (0.84)	10.05 (0.60)	9.92 (0.58)	10.12 (0.76)	9.95 (0.65)	10.07 (0.61)	5.11	0.00
	0.5	40.22 (0.84)	8.75 (0.69)	8.55 (0.66)	8.73 (0.85)	8.65 (0.88)	8.60 (0.61)	8.73 (0.58)	8.86 (0.50)	8.85 (0.58)	38.53	1.00
	1	40.42 (0.65)	7.41 (0.63)	7.30 (0.52)	7.42 (0.63)	7.42 (0.72)	7.44 (0.48)	7.67 (0.40)	7.46 (0.59)	7.41 (0.60)	171.83	1.00
β =.1 α =.25	0	39.83 (1.26)	10.24 (0.98)	9.83 (0.87)	10.23 (1.37)	9.78 (1.21)	25.23 (1.00)	24.80 (1.18)	25.48 (0.81)	24.90 (1.21)	5.72	0.00
	0.5	45.11 (0.88)	9.10 (0.72)	9.03 (0.78)	9.00 (0.73)	9.20 (0.79)	18.69 (0.86)	18.92 (1.01)	18.33 (0.70)	18.38 (0.82)	48.39	1.00
	1	47.90 (1.00)	7.42 (0.56)	7.58 (0.71)	7.34 (0.55)	7.67 (0.65)	12.92 (0.68)	12.95 (0.93)	13.01 (0.65)	13.00 (0.77)	395.70	1.00
β =.25 α =.1	0	40.30 (0.92)	25.29 (1.67)	25.40 (1.31)	25.30 (1.08)	25.04 (1.20)	10.17 (0.73)	9.74 (0.70)	9.93 (0.68)	10.02 (0.85)	6.41	0.07
	0.5	37.13 (0.94)	18.10 (1.31)	18.10 (1.08)	18.05 (1.18)	18.43 (1.11)	9.19 (0.58)	9.26 (0.77)	9.15 (0.70)	9.23 (0.47)	45.03	1.00
	1	35.49 (0.83)	12.70 (0.73)	12.73 (0.81)	12.75 (0.88)	12.45 (0.85)	7.56 (0.53)	7.54 (0.58)	7.46 (0.45)	7.65 (0.44)	314.80	1.00
β =.25 α =.25	0	39.89 (1.46)	24.21 (1.94)	24.97 (1.63)	24.80 (1.93)	24.85 (1.65)	24.80 (1.30)	24.95 (1.39)	25.29 (1.63)	25.00 (1.50)	6.33	0.00
	0.5	41.08 (1.21)	17.84 (1.05)	17.37 (1.32)	17.51 (1.28)	17.72 (1.22)	19.06 (0.98)	18.12 (0.98)	18.69 (0.89)	18.83 (0.83)	57.34	1.00
	1	42.64 (0.73)	11.73 (0.72)	11.86 (0.85)	11.93 (0.86)	11.83 (0.78)	12.56 (0.83)	12.33 (0.78)	12.33 (0.58)	12.66 (0.75)	513.30	1.00

* Proportion of maximum dependence between observers for true negative and true positive subjects.
† Proportion of χ^2 values above the critical value at the 5% significance level of $\chi^2_{.05,6}$ = 12.59
†† β and α represent β_1, β_2, and β_3 and α_1, α_2, and α_3, respectively

Figure 12a

Mean of Estimates of Prevalence (θ) in the case of four observers, with dependence between all pairs of observers with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

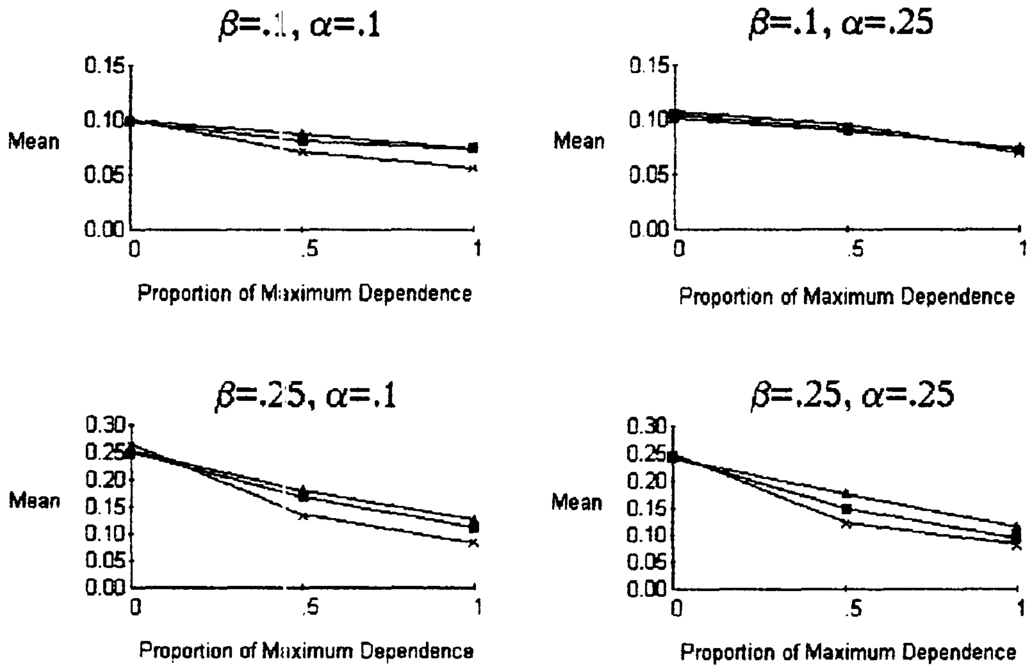
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 12b

Mean of Estimates of false negative rate for observer 1 (β_1) in the case of four observers, with dependence between all pairs of observers with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

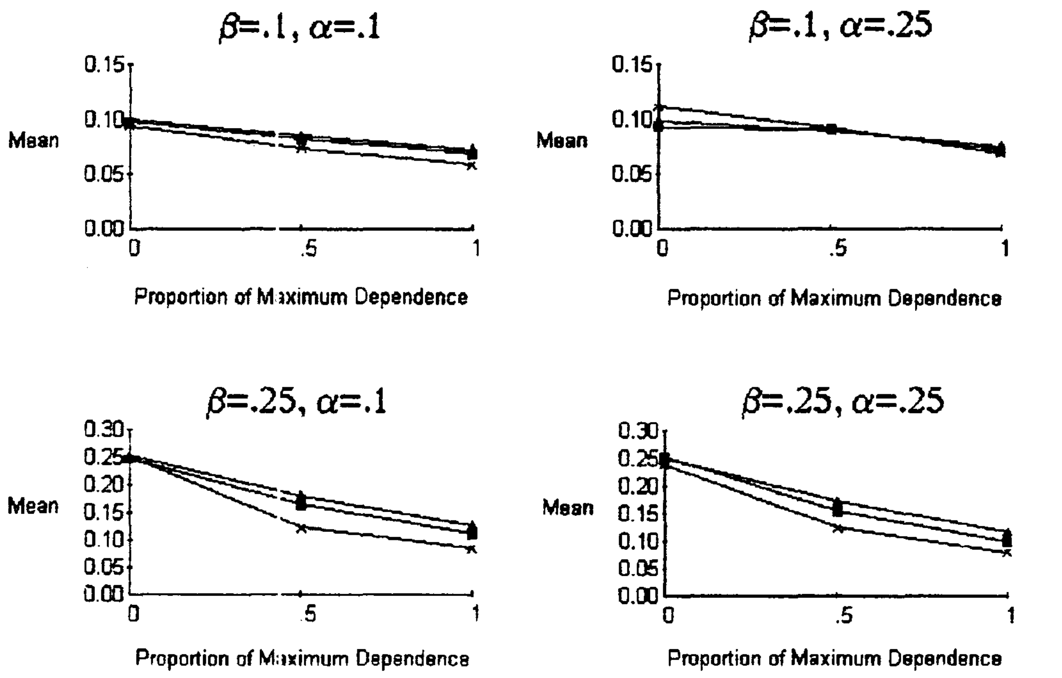
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 12c

Mean of Estimates of false negative rate for observer 2 (β_2) in the case of four observers, with dependence between all pairs of observers with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

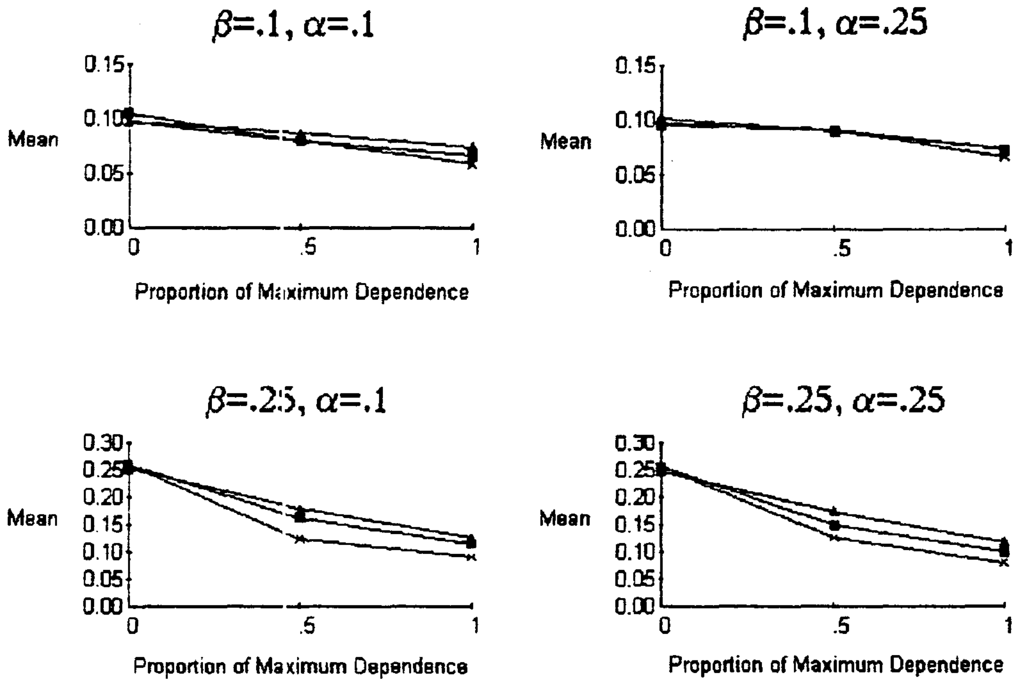
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 12 d

Mean of Estimates of false negative rate for observer 3 (β_3) in the case of four observers, with dependence between all pairs of observers with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

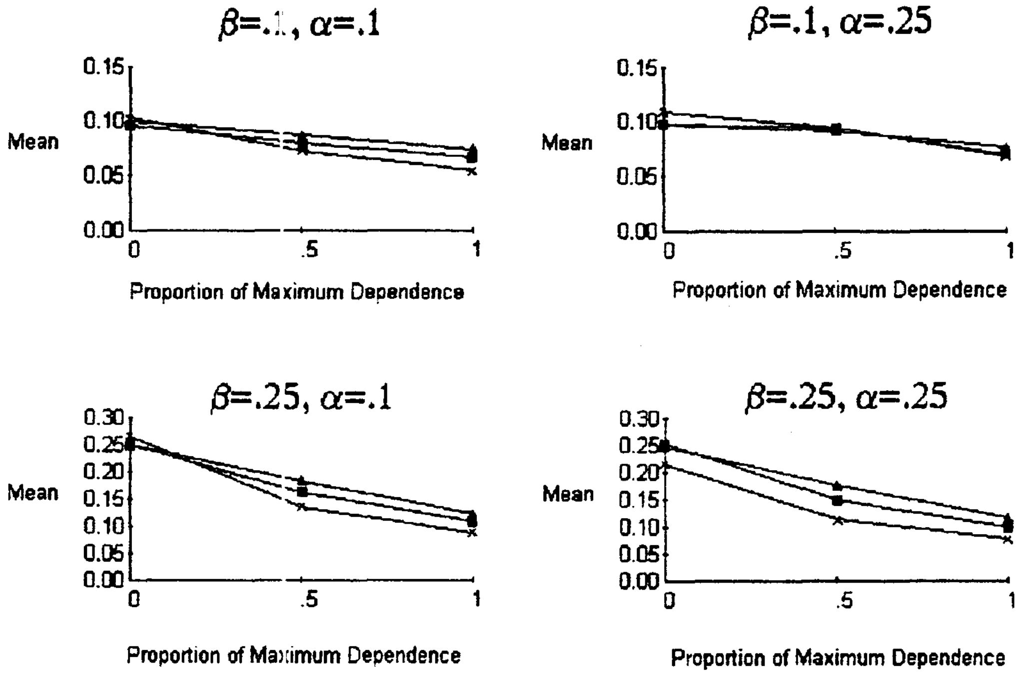
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 12e

Mean of Estimates of false negative rate for observer 4 (β_4) in the case of four observers, with dependence between all pairs of observers with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

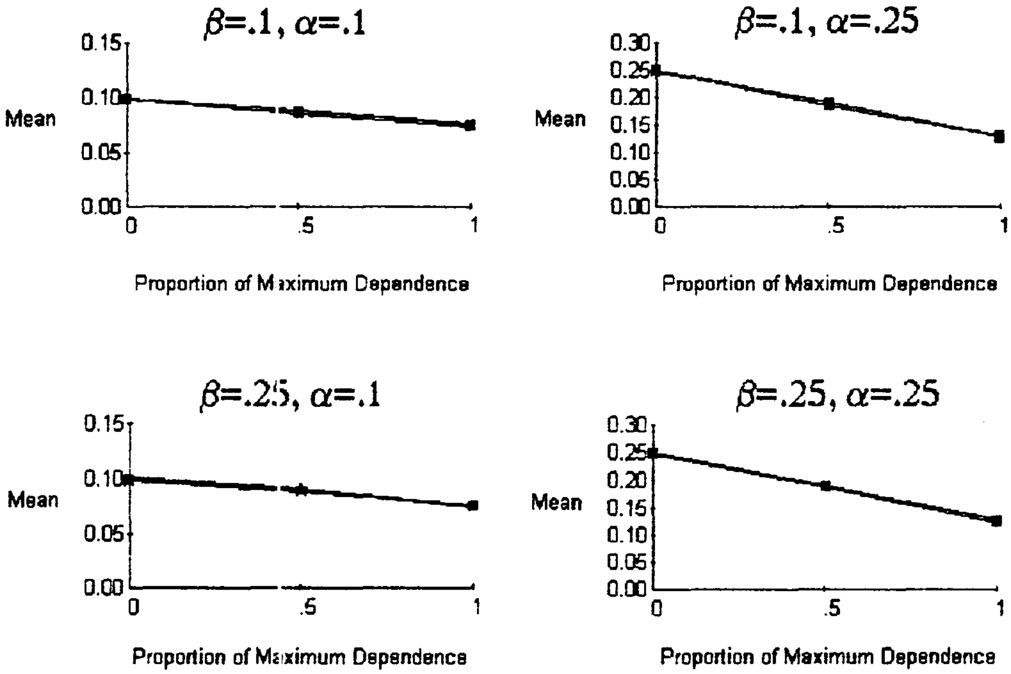
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 12f

Mean of Estimates of false positive rate for observer 1 (α_1) in the case of four observers, with dependence between all pairs of observers with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

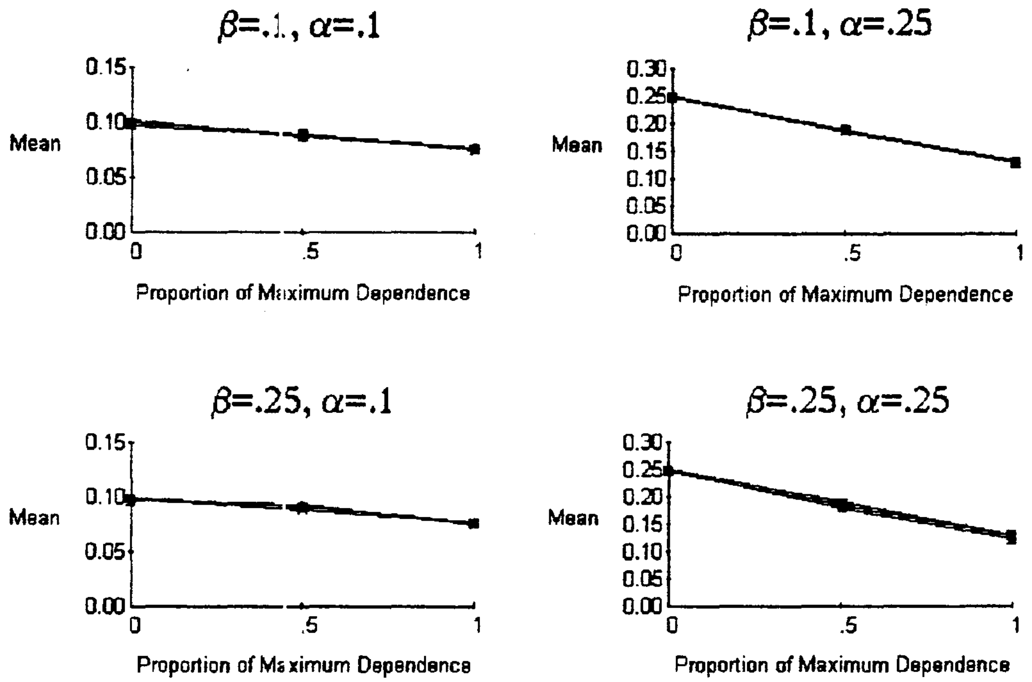
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 12g

Mean of Estimates of false positive rate for observer 2 (α_2) in the case of four observers, with dependence between all pairs of observers with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

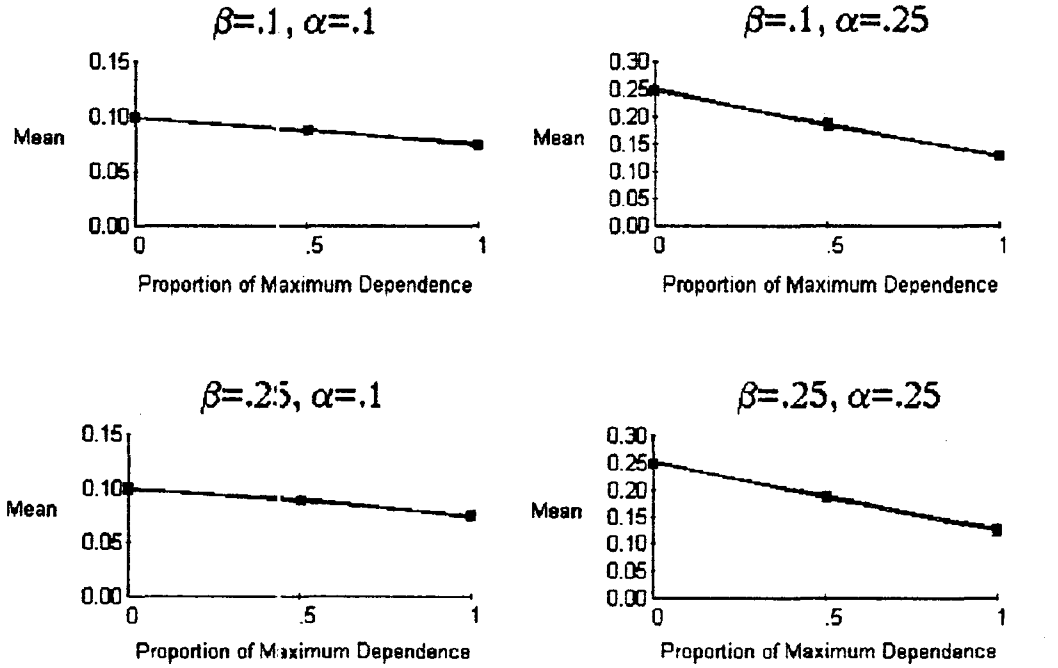
× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 12h

Mean of Estimates of false positive rate for observer 3 (α_3) in the case of four observers, with dependence between all pairs of observers with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

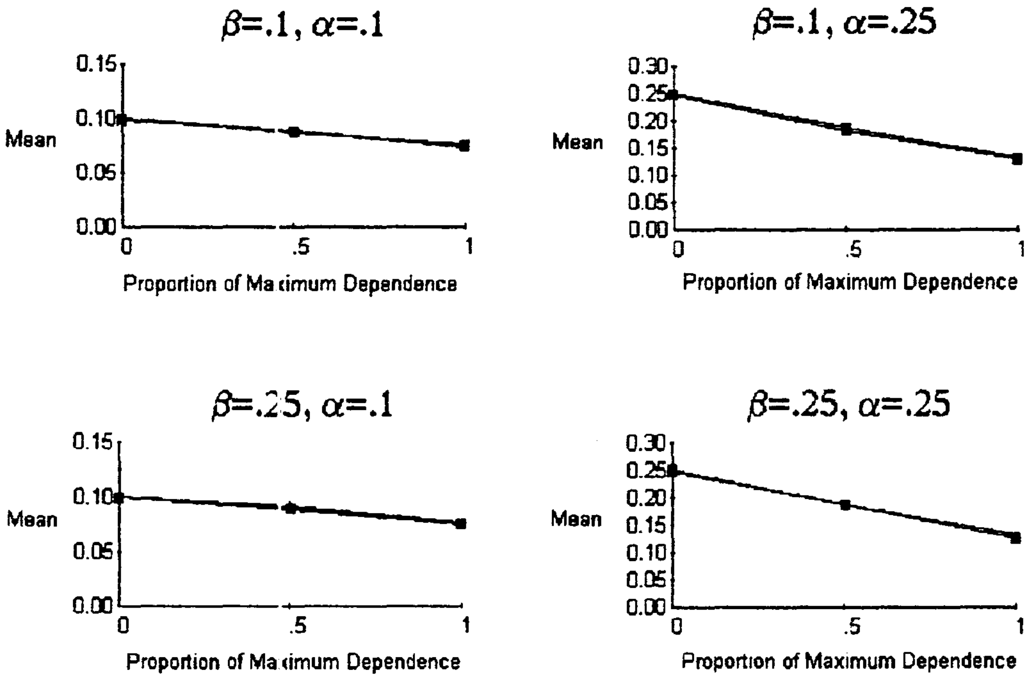
× represents true prevalence (θ) = .05.

▪ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

Figure 12 i

Mean of Estimates of false positive rate for observer 4 (α_4) in the case of four observers, with dependence between all pairs of observers with respect to false positive and false negative rate: mean vs dependence for each true prevalence by true error rates



β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3 = \beta_4$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

× represents true prevalence (θ) = .05.

■ represents true prevalence (θ) = .15.

▲ represents true prevalence (θ) = .40.

3.2 Variability of Estimates

The estimates of standard error from the Latent program were evaluated by comparison with the standard deviations of the estimates estimated empirically from the simulation runs. This was done for Run 3 only. The Latent program was used to calculate the asymptotic standard error using expected frequencies for each parameter setting. The full results from this expected run can be seen in Tables 13a to 13c for prevalences of .05, .15, and .40, respectively. The results from simulation Run 3 were displayed in Tables 3a to 3c.

Tables 14a to 14c compare the estimated standard error with the true standard deviation for each estimated parameter by reporting the ratio of standard error to standard deviation. It can be seen that most ratios are close to 1.00. The exceptions are for the false negative rates for the two dependent observers. The ratio becomes large as the dependence increases. This corresponds to the actual rate becoming more underestimated. In the simulations the estimates for the false negative rates for observers 1 and 2 bunch up at the lower bound of zero. This causes the empirical standard deviation to become smaller. Thus, the ratio becomes larger. The ratios become undefined when the empirical standard deviations become zero because all estimates lie on the lower bound of zero.

In summary, it appears that in cases where bias is not causing estimates to approach a bound, the variability estimates from the Latent program are not a problem. When the estimate is biased, concern about the standard error estimate should really be irrelevant. The standard error estimates only appear substantially different from the standard deviations in these extreme cases.

TABLE 13a

Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate

Results from expected frequencies

True Prevalence = .05

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	4.96 (0.76)	9.69 (6.25)	9.69 (6.25)	9.69 (6.25)	10.02 (0.79)	10.02 (0.79)	10.02 (0.79)
	0.5	5.06 (0.72)	**0.00 (5.70)**	**0.00 (5.70)**	**18.69 (5.19)**	9.42 (0.77)	9.42 (0.77)	10.41 (0.77)
	1	**6.14 (0.74)**	**0.00 (4.75)**	**0.00 (4.75)**	**25.37 (4.55)**	8.37 (0.74)	8.37 (0.74)	10.03 (0.76)
β =.1 α =.25	0	5.01 (2.17)	9.85 (15.45)	9.85 (15.46)	9.85 (15.46)	24.99 (1.43)	24.99 (1.43)	24.99 (1.43)
	0.5	**11.24 (2.36)**	**0.00 (10.25)**	**0.00 (10.25)**	**41.24 (4.09)**	**19.16 (1.66)**	**19.16 (1.66)**	24.39 (1.26)
	1	**17.63 (1.98)**	**0.00 (5.80)**	**0.00 (5.80)**	**47.05 (2.80)**	**12.89 (1.52)**	**12.89 (1.52)**	22.96 (1.19)
β =.25 α =.1	0	4.92 (1.13)	24.66 (9.19)	24.66 (9.19)	24.66 (9.19)	10.03 (0.85)	10.03 (0.85)	10.03 (0.85)
	0.5	4.33 (0.84)	**4.07 (9.14)**	**4.07 (9.14)**	**33.33 (6.50)**	9.51 (0.81)	9.51 (0.81)	10.78 (0.79)
	1	5.28 (0.82)	**0.00 (7.25)**	**0.00 (7.25)**	**38.83 (5.25)**	8.42 (0.78)	8.42 (0.78)	10.53 (0.77)
β =.25 α =.25	0	4.90 (4.04)	24.50 (23.05)	24.50 (23.05)	24.50 (23.05)	25.03 (1.74)	25.03 (1.74)	25.03 (1.74)
	0.5	**10.07 (3.11)**	**0.00 (15.52)**	**0.00 (15.52)**	**49.28 (4.35)**	**19.38 (2.02)**	**19.38 (2.02)**	24.90 (1.26)
	1	**16.77 (2.39)**	**0.00 (7.68)**	**0.00 (7.68)**	**52.43 (2.87)**	**12.90 (1.78)**	**12.90 (1.78)**	23.46 (1.19)

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.
β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.
Means with 20% or greater bias are displayed in bold.

TABLE 13b

Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate

Results from expected frequencies

True Prevalence = .15

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	15.08 (1.00)	10.21 (2.58)	10.21 (2.58)	10.21 (2.58)	9.96 (0.84)	9.96 (0.84)	9.96 (0.84)
	0.5	15.05 (0.97)	**6.49 (2.36)**	**6.49 (2.36)**	**13.59 (2.53)**	9.33 (0.81)	9.33 (0.81)	10.59 (0.84)
	1	14.98 (0.95)	**2.82 (2.08)**	**2.82 (2.08)**	**16.41 (2.48)**	8.75 (0.79)	8.75 (0.79)	11.15 (0.84)
β =.1 α =.25	0	14.94 (2.18)	9.93 (5.23)	9.93 (5.23)	9.93 (5.23)	25.03 (1.52)	25.03 (1.52)	25.03 (1.52)
	0.5	**18.20 (1.91)**	**0.00 (4.60)**	**0.00 (4.60)**	**26.24 (2.94)**	20.21 (1.51)	20.21 (1.51)	26.05 (1.36)
	1	**24.24 (1.66)**	**0.00 (3.12)**	**0.00 (3.12)**	**33.84 (2.28)**	**13.87 (1.38)**	**13.87 (1.38)**	24.76 (1.29)
β =.25 α =.1	0	14.83 (1.43)	24.76 (4.18)	24.76 (4.18)	24.76 (4.18)	10.04 (0.95)	10.04 (0.95)	10.04 (0.95)
	0.5	14.23 (1.23)	**15.99 (3.89)**	**15.99 (3.89)**	29.59 (3.65)	9.09 (0.90)	9.09 (0.90)	11.35 (0.90)
	1	13.43 (1.07)	**6.86 (3.45)**	**6.86 (3.45)**	**32.79 (3.29)**	8.38 (0.84)	8.38 (0.84)	**12.40 (0.88)**
β =.25 α =.25	0	15.10 (4.21)	25.17 (8.08)	25.17 (8.08)	25.17 (8.08)	24.97 (1.91)	24.97 (1.91)	24.97 (1.91)
	0.5	14.76 (2.66)	**0.00 (8.58)**	**0.00 (8.58)**	**39.14 (3.50)**	20.81 (1.85)	20.81 (1.85)	27.59 (1.37)
	1	**21.65 (2.19)**	**0.00 (5.16)**	**0.00 (5.16)**	**44.77 (2.51)**	**13.85 (1.71)**	**13.85 (1.71)**	26.22 (1.29)

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.
β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.
Means with 20% or greater bias are displayed in bold.

TABLE 13c
Means and standard deviations of parameter estimates in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate

Results from expected frequencies

True Prevalence = .40

True Error Rates	Dep*	Mean of Estimates (%) (standard deviation (%))						
		θ	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	40.08 (1.25)	10.08 (1.32)	10.08 (1.32)	10.08 (1.32)	9.95 (1.03)	9.95 (1.03)	9.95 (1.03)
	0.5	40.07 (1.23)	8.69 (1.25)	8.69 (1.25)	11.42 (1.33)	9.03 (0.99)	9.03 (0.99)	10.85 (1.03)
	1	40.06 (1.21)	**7.33 (1.18)**	**7.33 (1.18)**	**12.64 (1.34)**	8.14 (0.94)	8.14 (0.94)	11.69 (1.04)
β =.1 α =.25	0	39.98 (2.09)	9.99 (1.98)	9.99 (1.98)	9.99 (1.98)	25.01 (1.84)	25.01 (1.84)	25.01 (1.84)
	0.5	41.43 (1.85)	**5.57 (1.80)**	**5.57 (1.80)**	**15.99 (1.65)**	20.28 (1.75)	20.28 (1.75)	27.65 (1.70)
	1	42.43 (1.65)	**1.79 (1.57)**	**1.79 (1.57)**	**19.88 (1.51)**	**16.16 (1.62)**	**16.16 (1.62)**	29.49 (1.61)
β =.25 α =.1	0	39.89 (1.89)	24.91 (2.31)	24.91 (2.31)	24.91 (2.31)	10.06 (1.33)	10.06 (1.33)	10.06 (1.33)
	0.5	38.68 (1.68)	**19.41 (2.17)**	**19.41 (2.17)**	27.95 (2.13)	**7.87 (1.23)**	**7.87 (1.23)**	**13.26 (1.21)**
	1	37.73 (1.51)	**14.26 (1.98)**	**14.26 (1.98)**	**30.36 (2.01)**	**5.86 (1.10)**	**5.86 (1.10)**	**15.61 (1.16)**
β =.25 α =.25	0	40.20 (4.26)	25.12 (3.43)	25.12 (3.43)	25.12 (3.43)	24.92 (2.49)	24.92 (2.49)	24.92 (2.49)
	0.5	39.75 (3.08)	**14.72 (3.30)**	**14.72 (3.30)**	**32.42 (2.30)**	**18.43 (2.35)**	**18.43 (2.35)**	**30.10 (1.81)**
	1	38.86 (2.38)	**5.83 (2.99)**	**5.83 (2.99)**	**35.87 (1.92)**	**13.76 (2.11)**	**13.76 (2.11)**	**32.77 (1.59)**

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.
β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.
Means with 20% or greater bias are displayed in bold.

TABLE 14a

Ratios of standard error from asymptotic formula to standard deviation from empirical simulations in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate

Prevalence = .05

True Error Rates	Dep*	Ratio of standard error from asymptotic formula to standard deviation from empirical simulations						
		Prevalence	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	1.07	1.01	1.10	1.13	1.04	1.14	1.03
	0.5	1.16	1.75	2.18	1.03	1.05	1.07	0.99
	1	1.28	--	237.50	1.00	1.17	1.30	1.09
β =.1 α =.25	0	0.97	1.42	1.40	1.45	1.06	1.11	1.01
	0.5	2.68	22.28	56.94	1.07	1.61	1.60	1.24
	1	2.18	--	--	0.94	1.73	1.77	1.19
β =.25 α =.1	0	1.06	1.18	0.80	1.09	0.89	1.05	1.04
	0.5	1.42	1.29	1.70	1.03	1.16	1.01	0.94
	1	1.46	--	--	0.96	1.32	1.28	1.18
β =.25 α =.25	0	0.54	1.25	1.22	1.22	0.81	0.81	0.76
	0.5	3.57	21.56	29.28	1.03	2.02	2.02	1.07
	1	2.49	--	--	1.03	2.05	2.12	1.27

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.

β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

-- Undefined value

TABLE 14b

Ratios of standard error from asymptotic formula to standard deviation from empirical simulations in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate

Prevalence = .15

True Error Rates	Dep*	Ratio of standard error from asymptotic formula to standard deviation from empirical simulations						
		Prevalence	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	0.95	0.98	0.87	1.04	0.91	1.08	0.98
	0.5	0.94	0.81	1.00	0.90	0.87	0.91	0.92
	1	1.16	0.95	1.02	1.07	1.03	1.07	1.29
β =.1 α =.25	0	0.98	1.06	1.07	1.12	1.11	1.13	0.92
	0.5	1.74	3.26	4.34	1.04	1.29	1.17	1.25
	1	2.02	--	--	0.90	1.73	1.41	1.47
β =.25 α =.1	0	0.81	1.03	0.94	0.69	0.99	0.90	0.77
	0.5	1.23	1.11	1.32	0.92	0.80	0.97	1.05
	1	1.13	0.94	1.13	0.90	0.99	1.11	1.06
β =.25 α =.25	0	0.92	1.06	0.83	0.89	1.05	1.07	0.91
	0.5	2.05	3.67	2.32	1.13	1.33	1.78	1.43
	1	1.99	--	--	0.93	1.32	1.49	1.15

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.

β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.

α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.

-- Undefined value

TABLE 14c

Ratios of standard error from asymptotic formula to standard deviation from empirical simulations in the case of three observers, with dependence between observers 1 and 2 with respect to false positive and false negative rate

Prevalence = .40

True Error Rates	Dep*	Ratio of standard error from asymptotic formula to standard deviation from empirical simulations						
		Prevalence	β_1	β_2	β_3	α_1	α_2	α_3
β =.1 α =.1	0	1.10	0.99	1.18	1.15	1.16	0.88	1.01
	0.5	1.00	0.93	1.17	0.99	1.11	0.92	0.93
	1	0.92	0.87	0.81	1.10	1.02	0.91	1.08
β =.1 α =.25	0	1.14	1.02	0.98	1.05	1.38	1.08	0.94
	0.5	0.91	0.73	0.91	0.95	0.87	0.89	0.95
	1	0.90	1.26	1.09	0.99	0.93	1.38	1.01
β =.25 α =.1	0	1.02	1.03	1.42	1.34	0.95	0.89	0.85
	0.5	1.08	1.02	1.34	0.83	1.23	0.95	1.14
	1	1.16	1.09	1.02	0.95	1.03	0.93	1.32
β =.25 α =.25	0	0.68	0.72	0.73	0.88	0.96	0.78	0.82
	0.5	0.96	1.06	1.12	1.17	1.00	1.18	0.86
	1	1.10	1.22	1.19	1.14	1.18	1.07	0.95

* Proportion of maximum dependence between observers 1 and 2 for both true negative and true positive subjects.
β represents the true false negative rates, $\beta_1 = \beta_2 = \beta_3$.
α represents the true false positive rates, $\alpha_1 = \alpha_2 = \alpha_3$.
-- Undefined value

4. CONCLUSION

In the application of latent class models, departure from the assumption of conditional independence between the observers will often lead to serious biases in the parameter estimates. In general, dependence between observers will lead to underestimation of those observers' error rates due to the excess agreement in their classifications while the error rates of the independent observers will be overestimated. The prevalence is typically overestimated in the presence of dependence if the true prevalence is less than 50%. It has repeatedly been shown that the most severe biases occur when the prevalence is low and the false positive rates of the observers are high.

As a result, in the evaluation of diagnostic tests using latent class approaches, the most dangerous situation occurs when the specificities of the tests are low and the disease or condition is rare. This describes a typical population screening situation. Therefore, if latent class analysis is being used to evaluate several screening tests it is imperative that the tests be independent to produce accurate estimates of the parameters. If independence is not present the tests will appear deceptively accurate.

One method to help alleviate the bias problem is to add observers or tests. It was shown in this study that situations where it is suspected that two of three observers or tests are dependent benefited from the addition of an independent fourth observer. This fourth observer allowed the latent class methods to produce less biased results. Even the addition observer allowed the latent class methods to produce less biased results. Even the addition of a dependent fourth observer can be considered beneficial since with four observers the fit can be tested.

The problem with adding an additional observer is that the number of outcome categories automatically doubles when the classification is binary. For example, in the three observer case there are $2^3 = 8$ outcome categories and in the four observer case there are $2^4 = 16$ outcome categories. Hence, if there is not a large population then the analysis may become impossible with the addition of another observer because of small sample problems encountered when the existing data is spread over twice as many cells. As a result, the addition of another observer will give less biased estimates, but at the price of less precision.

A brief look at standard error estimates when dependence is present indicates that dependence does not cause a problem with the estimation of variability unless there is already a large problem with bias in the parameter estimates.

The goodness of fit test appears to possess fairly good power to detect the dependence as a departure from the model assumptions. The best power was seen when the dependence structure included dependence between all pairs of observers.

This work is just the beginning of an investigation into the behaviour of latent class estimates. Some questions about the bias of estimates in the presence of dependence have been answered here while others were brought to light. For example, the comparison of three observer cases to four observer cases may or may not reveal the same results if absolute dependence values are used as the basis of comparison in place of the proportions of maximum dependence. This question was raised by the surprising results obtained when comparing Run 6 and Run 12, where Run 12 estimates showed less bias. The problems encountered when outcome category frequencies became too small were discussed in Run 2. Thus, a question with potential for further research is, "How small is too small?" This study considered only binary classifications. The implications of more than two levels of classification is another potential area for further research into the usefulness of latent class methods. The goodness of fit statistic was simply summarized in this study. An area of further inquiry could be the interpretations of the components of the goodness of fit statistic and determination of whether the examination of the contributions to the goodness of fit statistic from the individual cells can help determine where and what kind of departure from the model assumptions is present.

REFERENCES

1. Walter SD, Irwig LM. Estimation of test error rates, disease prevalence and relative risk from misclassified data: a review. **J Clin Epidemiol** 1988; 41: 923-937.

2. Walter SD, Frommer DJ, Cook RJ. The estimation of sensitivity and specificity in colorectal cancer screening methods. **Cancer Detection & Prevention** 1991; 15: 465-470.

3. Hui SL, Walter SD. Estimating the error rates of diagnostic tests. **Biometrics** 1980; 36: 167-171.

4. Vacek PM. The effect of conditional dependence on the evaluation of diagnostic tests. **Biometrics** 1985; 41: 959-968.

5. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. **Clinical Epidemiology: A Basic Science for Clinical Medicine.** (2nd ed.) Little, Brown and Company: Boston; 1991. p 83.

6. Devore JL. **Probability and Statistics for Engineering and Sciences.** (2nd ed.) Brooks/Cole Publishing Company: Monterey, California; 1987. p 60.

7. Walter SD. LATENT Computer Software. Hamilton; 1993.

8. Devore, p 294.

9. Personal communication from Joseph L Gastwirth: Sinclair MD, Gastwirth JL. On designs for measuring prevalence rates and error rates from diagnostic testing procedures.

APPENDIX 1

DETERMINING THE MAXIMUM VALUE OF DEPENDENCE PARAMETERS

Case i) Determining the maximum value of δ_{12} in the case of three observers with dependence between observers 1 and 2

When there is dependence between observers 1 and 2 with respect to false negative rates in the three observer case, δ_{12} must be less than or equal to,

$$\beta_1(1-\beta_2)\beta_3, \tag{1}$$

$$(1-\beta_1)\beta_2\beta_3, \tag{2}$$

$$\text{and } 1-(1-\beta_1)(1-\beta_2)(1-\beta_3), \tag{3}$$

where $0 \leq \beta_1, \beta_2, \beta_3 \leq .5$. Thus the smallest of these values must be found to determine the maximum possible value of δ_{12}. Consider value (3) minus value (1).

$$\begin{aligned}(3) - (1) &= 1 - (1-\beta_1)(1-\beta_2)(1-\beta_3) - \beta_1(1-\beta_2)\beta_3 \\ &= 1 - (1-\beta_1-\beta_2-\beta_3+\beta_1\beta_2+\beta_1\beta_3+\beta_2\beta_3-\beta_1\beta_2\beta_3) - \beta_1\beta_3 + \beta_1\beta_2\beta_3 \\ &= \beta_1 + \beta_2 + \beta_3 - \beta_1\beta_2 - \beta_1\beta_3 - \beta_2\beta_3 + \beta_1\beta_2\beta_3 - \beta_1\beta_3 + \beta_1\beta_2\beta_3 \\ &= \beta_1 + \beta_2 + \beta_3 - \beta_1\beta_2 - 2\beta_1\beta_3 - \beta_2\beta_3 + 2\beta_1\beta_2\beta_3 \\ &= \beta_1(1-\beta_3) + \beta_2(1-\beta_1) + \beta_3(1-\beta_2) - \beta_1\beta_3 + 2\beta_1\beta_2\beta_3\end{aligned}$$

All terms are positive except $-\beta_1\beta_3$. However, $\beta_1(1-\beta_3) \geq \beta_1\beta_3$, so (3) - (1) is a positive value. Thus, (1) is a more strict restriction than (3). A similar argument would show that (2) is also a more strict restriction than (3). Therefore, the maximum value for δ_{12} is determined by restricting

$$\delta_{12} \leq \beta_1(1-\beta_2)\beta_3 \text{ and } \delta_{12} \leq (1-\beta_1)\beta_2\beta_3.$$

Case ii) Determining the maximum value of $\delta=\delta_{12}\delta_{13}\delta_{23}$ in the case of three observers with dependence between all pairs of observers

It has been established that the dependence term for all pairs of observers must be less than or equal to,

$$\beta_1\beta_2(1-\beta_3), \tag{4}$$

$$\beta_1(1-\beta_2)\beta_3, \tag{5}$$

$$(1-\beta_1)\beta_2\beta_3, \tag{6}$$

$$\text{and } (1-(1-\beta_1)(1-\beta_2)(1-\beta_3))/3 \tag{7}$$

where $0 \leq \beta_1, \beta_2, \beta_3 \leq .5$. The smallest of the above values must be found. The following will prove that (4) is less than or equal to (7) by showing that (7) minus (4) is greater than zero.

$$3 \times [(7) - (4)] = 1-(1-\beta_1)(1-\beta_2)(1-\beta_3) - 3\beta_1\beta_2(1-\beta_3)$$

$$= 1 - (1-\beta_1-\beta_2-\beta_3+\beta_1\beta_2+\beta_1\beta_3+\beta_2\beta_3-\beta_1\beta_2\beta_3) - 3\beta_1\beta_3 + 3\beta_1\beta_2\beta_3$$

$$= \beta_1 + \beta_2 + \beta_3 - \beta_1\beta_2 - \beta_1\beta_3 - \beta_2\beta_3 + \beta_1\beta_2\beta_3 - 3\beta_1\beta_2 + 3\beta_1\beta_2\beta_3$$

$$= \beta_1 + \beta_2 + \beta_3 - 4\beta_1\beta_2 - \beta_1\beta_3 - \beta_2\beta_3 + 4\beta_1\beta_2\beta_3$$

$$= \beta_1 - 2\beta_1\beta_2 + \beta_2 - 2\beta_1\beta_2 + \beta_3 - \beta_1\beta_3 - \beta_2\beta_3 + 4\beta_1\beta_2\beta_3$$

$$= \beta_1(1-2\beta_2)+ \beta_2(1-2\beta_1) + \beta_3(1-\beta_1-\beta_2) + 4\beta_1\beta_2\beta_3$$

Since all error rates are positive and at most 0.5 then all terms are positive. Hence, (7) - (4) is a positive value meaning that restriction (4) is more strict than restriction (7). The same procedure can be followed to show that restrictions (5) and (6) are also more strict than restriction (7). Thus, the maximum value for the dependence parameters for all pairs of observers is determined by restricting

$$\delta \leq \beta_1\beta_2(1-\beta_3),\ \delta \leq \beta_1(1-\beta_2)\beta_3, \text{ and } \delta \leq (1-\beta_1)\beta_2\beta_3.$$

Case iii) Determining the maximum value of δ_{12} in the case of four observers with dependence between observers 1 and 2

The three most strict conditions for determining the value of the dependence parameter in this case restrict δ_{12} to be less than or equal to,

$$\beta_1(1-\beta_2)\beta_3\beta_4, \tag{8}$$

$$(1-\beta_1)\beta_2\beta_3\beta_4, \tag{9}$$

$$\text{and } 1-(1-\beta_1)(1-\beta_2)(1-\beta_3)(1-\beta_4) \tag{10}$$

where $0 \leq \beta_1, \beta_2, \beta_3 \leq .5$. Taking condition (10) minus condition (8) gives,

$$\begin{aligned}
(10) - (8) &= 1-(1-\beta_1)(1-\beta_2)(1-\beta_3)(1-\beta_4) - \beta_1(1-\beta_2)\beta_3\beta_4 \\
&= 1 - (1-\beta_1-\beta_2-\beta_3-\beta_4+\beta_1\beta_2+\beta_1\beta_3+\beta_1\beta_4+\beta_2\beta_3+\beta_2\beta_4+\beta_3\beta_4-\beta_1\beta_2\beta_3-\beta_1\beta_2\beta_4 \\
&\qquad -\beta_1\beta_3\beta_4-\beta_2\beta_3\beta_4+\beta_1\beta_2\beta_3\beta_4) - \beta_1\beta_3\beta_4 + \beta_1\beta_2\beta_3\beta_4 \\
&= \beta_1 + \beta_2 + \beta_3 + \beta_4 - \beta_1\beta_2 - \beta_1\beta_3 - \beta_1\beta_4 - \beta_2\beta_3 - \beta_2\beta_4 - \beta_3\beta_4 + \beta_1\beta_2\beta_3 \\
&\qquad + \beta_1\beta_2\beta_4 + \beta_2\beta_3\beta_4 \\
&= \beta_1 - \beta_1\beta_2 - \beta_1\beta_3 + \beta_2 - \beta_2\beta_3 - \beta_2\beta_4 + \beta_3 - \beta_3\beta_4 + \beta_4 - \beta_1\beta_4 + \beta_1\beta_2\beta_3 \\
&\qquad + \beta_1\beta_2\beta_4 + \beta_2\beta_3\beta_4 \\
&= \beta_1(1 - \beta_2 - \beta_3) + \beta_2(1 - \beta_3 - \beta_4) + \beta_3(1 - \beta_4) + \beta_4(1 - \beta_1) + \beta_1\beta_2\beta_3 \\
&\qquad + \beta_1\beta_2\beta_4 + \beta_2\beta_3\beta_4
\end{aligned}$$

Note that all terms are non-negative. The value of (10) − (8) is no less than zero. Thus, restriction (8) is at least as strict as restriction (10). This argument can be repeated with restriction (9) and (10) to show that restriction (9) is also more strict than restriction (10). Therefore, restricting

$$\delta_{12} \leq \beta_1(1-\beta_2)\beta_3\beta_4 \text{ and } \delta_{12} \leq (1-\beta_1)\beta_2\beta_3\beta_4$$

will give the maximum possible value for δ_{12} in this situation.

<u>Case iv) Determining the maximum value of $\delta=\delta_{12}=\delta_{13}=\delta_{14}=\delta_{23}=\delta_{24}=\delta_{34}$ in the case of four observers with dependence between all pairs of observers</u>

In this situation, the dependence parameter δ is restricted to be less than or equal to,

$$\beta_1\beta_2(1-\beta_3)(1-\beta_4)/2, \quad (11)$$

$$\beta_1(1-\beta_2)\beta_3(1-\beta_4)/2, \quad (12)$$

$$\beta_1(1-\beta_2)(1-\beta_3)\beta_4/2, \quad (13)$$

$$(1-\beta_1)\beta_2\beta_3(1-\beta_4)/2, \quad (14)$$

$$(1-\beta_1)\beta_2(1-\beta_3)\beta_4/2, \quad (15)$$

$$(1-\beta_1)(1-\beta_2)\beta_3\beta_4/2, \quad (16)$$

$$\text{and } (1-(1-\beta_1)(1-\beta_2)(1-\beta_3)(1-\beta_4))/6, \quad (17)$$

where $0 \le \beta_1, \beta_2, \beta_3, \beta_4 \le .5$. Consider value (17) minus value (11).

$$6 \times [(17)-(11)] = 1-(1-\beta_1)(1-\beta_2)(1-\beta_3)(1-\beta_4) - 3\beta_1\beta_2(1-\beta_3)(1-\beta_4)$$

$$= 1 - (1-\beta_1-\beta_2-\beta_3-\beta_4+\beta_1\beta_2+\beta_1\beta_3+\beta_1\beta_4+\beta_2\beta_3+\beta_2\beta_4+\beta_3\beta_4-\beta_1\beta_2\beta_3-\beta_1\beta_2\beta_4-\beta_1\beta_3\beta_4$$

$$-\beta_2\beta_3\beta_4+\beta_1\beta_2\beta_3\beta_4) - 3(\beta_1\beta_2-\beta_1\beta_2\beta_3)(1-\beta_4)$$

$$= \beta_1 -- \beta_2 + \beta_3 + \beta_4 - \beta_1\beta_2 - \beta_1\beta_3 - \beta_1\beta_4 - \beta_2\beta_3 - \beta_2\beta_4 - \beta_3\beta_4 + \beta_1\beta_2\beta_3$$

$$+ \beta_1\beta_2\beta_4 + \beta_1\beta_3\beta_4 + \beta_2\beta_3\beta_4 - \beta_1\beta_2\beta_3\beta_4 - 3\beta_1\beta_2 + 3\beta_1\beta_2\beta_3 + 3\beta_1\beta_2\beta_4$$

$$- 3\beta_1\beta_2\beta_3\beta_4$$

At this point two separate cases will be considered. First for the case $0 \le \beta_2 \le .25$.

$$6 \times [(17)-(11)] = \beta_1 - 4\beta_1\beta_2 + \beta_2 - \beta_2\beta_3 - \beta_2\beta_4 + \beta_3 - \beta_1\beta_3 - \beta_3\beta_4 + \beta_4 - \beta_1\beta_4$$

$$+ 4\beta_1\beta_2\beta_3 + 4\beta_1\beta_2\beta_4 + \beta_1\beta_3\beta_4 + \beta_2\beta_3\beta_4 - 4\beta_1\beta_2\beta_3\beta_4$$

$$= \beta_1(1-4\beta_2) + \beta_2(1-\beta_3-\beta_4) + \beta_3(1-\beta_1-\beta_4) + \beta_4(1-\beta_1) + 4\beta_1\beta_2\beta_4 + \beta_1\beta_3\beta_4$$

$$+ \beta_2\beta_3\beta_4 + 4\beta_1\beta_2\beta_3(1-\beta_4)$$

All terms above are non-negative, thus, (17) minus (11) is a non-negative value. So for the case

$0 \leq \beta_2 \leq .25$, (11) is a more strict restriction. Now consider the second case, $.25 < \beta_2 \leq .5$.

$$6 \times [(17)-(11)] = \beta_1 - 2\beta_1\beta_2 + \beta_2 - 2\beta_1\beta_2 + \beta_3 - \beta_2\beta_3 - \beta_3\beta_4 + \beta_4 - \beta_1\beta_4 - \beta_2\beta_4 - \beta_1\beta_3$$
$$+ 4\beta_1\beta_2\beta_3 + 4\beta_1\beta_2\beta_4 + \beta_1\beta_3\beta_4 + \beta_2\beta_3\beta_4 - 4\beta_1\beta_2\beta_3\beta_4$$
$$= \beta_1(1 - 2\beta_2) + \beta_2(1 - 2\beta_1) + \beta_3(1 - \beta_2 - \beta_4) + \beta_4(1 - \beta_1 - \beta_2) + 4\beta_1\beta_2\beta_3 - \beta_1\beta_3$$
$$+ 4\beta_1\beta_2\beta_4 - 4\beta_1\beta_2\beta_3\beta_4 + \beta_1\beta_3\beta_4 + \beta_2\beta_3\beta_4$$
$$= \beta_1(1 - 2\beta_2) + \beta_2(1 - 2\beta_1) + \beta_3(1 - \beta_2 - \beta_4) + \beta_4(1 - \beta_1 - \beta_2) + \beta_1\beta_3(4\beta_2 - 1)$$
$$+ 4\beta_1\beta_2\beta_4(1 - \beta_3) + \beta_1\beta_3\beta_4 + \beta_2\beta_3\beta_4$$

All terms are non-negative. So, when $.25 < \beta_2 \leq .5$, (11) is a more strict restriction than (17). Because the other restrictions are of the same form, the same argument can be used to show that restritions (12), (13), (14), (15), and (16) are more strict than restriction (17). Therefore, the maximum value for the dependence parameter in this case is determined by restricting,

$$\delta \leq \beta_1\beta_2(1 - \beta_3)(1 - \beta_4)/2,$$
$$\delta \leq \beta_1(1 - \beta_2)\beta_3(1 - \beta_4)/2,$$
$$\delta \leq \beta_1(1 - \beta_2)(1 - \beta_3)\beta_4/2,$$
$$\delta \leq (1 - \beta_1)\beta_2\beta_3(1 - \beta_4)/2,$$
$$\delta \leq (1 - \beta_1)\beta_2(1 - \beta_3)\beta_4/2,$$
$$\text{and } \delta \leq (1 - \beta_1)(1 - \beta_2)\beta_3\beta_4/2.$$