Inverse decision theory with applications to screening and diagnosis of cervical intraepithelial neoplasia

Kalatu R. Davies a,⁎, Dennis D. Cox a, Richard J. Swartz b, Scott B. Cantor b, Michele Follen c,d

a Department of Statistics—MS 138, Rice University, Houston, TX 77005, USA
b Department of Biostatistics and Applied Mathematics, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030, USA
c Center for Biomedical Engineering and Department of Gynecologic Oncology, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd., Unit 193, Houston, TX 77030-4009, USA
d Department of Obstetrics, Gynecology and Reproductive Sciences, The University of Texas Health Science Center in Houston, Houston, TX 77030, USA

Abstract

Objective. Medical decision makers would like to use decision theory to determine optimal treatment strategies for patients, but this requires priors, likelihoods, and losses. It can be very difficult to specify a loss or utility function in a medical setting, especially when considering both patient health outcomes and economic costs. These issues led to the development of Inverse Decision Theory (IDT), which involves determining the set of losses under which a given decision rule is optimal.

Methods. We apply IDT to the current standard of care for the diagnosis and treatment of precancerous lesions to the cervix, using a Bayesian approach to estimate the probabilities associated with diagnostic tests and make inferences about the region of optimality. There are two ways in which Inverse Decision Theory can be useful: (i) if the decision rule of interest is optimal, then we obtain information about the losses for the optimal treatment strategy, and (ii) if the decision rule of interest is not optimal, then we characterize the losses under which it would be optimal, and assess whether or not it contains reasonable values of the losses.

Results. This paper introduces important clinical results: in particular, we find that the current standard of care for cervical precancer is probably not optimal, and a new decision rule which requires a confirmatory biopsy for all patients with a positive Pap smear test result is better.

Conclusion. We have developed a very general and flexible approach for evaluating treatment strategies that could prove useful in a variety of medical applications.

© 2007 Published by Elsevier Inc.

Keywords: Bayesian decision theory; Medical decision making; Cervical intraepithelial neoplasia; Squamous intraepithelial lesions

Introduction

The field of Decision Theory is concerned with finding optimal decision rules or procedures. It requires probabilistic modeling of the outcomes of the various actions taken in a decision rule, and costs and losses (or utilities) for those actions. Bayesian Decision Theory additionally requires probabilities for all unknowns. In principle, decision theory could have great impact on medical practice. We often can estimate the prior probability a patient has a disease by using published disease prevalences. The probabilities associated with diagnostic tests (i.e., sensitivity and specificity) are known to some degree of accuracy, and the likelihoods of various outcomes are known to some degree of accuracy (e.g., if a patient with a low-grade squamous intraepithelial lesion [SIL] is not treated, then there is a probability of progression to a high-grade lesion). Furthermore, the costs of the various procedures are known. The main difficulty in applying Bayesian Decision Theory in medical practice is our inability or unwillingness to assign a loss on the same scale as costs (i.e., in monetary units) to patient outcomes such decrease in life expectancy.

Currently, there are three main approaches in medical decision making [1]. In clinical decision analysis, one considers only patient outcomes. A second approach, cost-effectiveness analysis, involves separate consideration of monetary costs and
patient benefits. In cost–benefit analysis, all evaluations are done on a single scale by translating patient outcomes into dollars. The common outcome scale of cost–benefit analysis is what is needed to apply decision theory. Each approach has shortcomings. One cannot ignore economic costs when comparing different strategies for diagnosis and treatment. Cost-effectiveness analysis compares different treatment options and can rule out some approaches as more costly and less effective than other approaches, but it does not, in general, identify an optimal decision rule. Cost–benefit analysis presents ethical issues, since one must place a monetary value on patient health outcomes such as loss of life. Thus, a satisfactory method for quantifying losses in the medical setting still does not exist, and applying decision theory to medical decision rules becomes very difficult. To help deal with such issues, we have developed an alternative approach called Inverse Decision Theory (IDT).

The basic inverse decision theoretic problem is as follows: Given a probability model and a specific decision rule, characterize the set of costs and losses (or utilities) under which the decision rule is optimal. If we believe the given rule is optimal, then we can infer something about the actual loss function—typically inequality bounds. If we are not convinced of its optimality, then by considering the region of optimality (i.e., the set of losses under which it is optimal), we can try to assess whether there are “reasonable” losses under which it is optimal. We employ a probability density function (called a prior probability on losses) to characterize the set of reasonable losses. Higher values of the density correspond to losses that are more reasonable. This approach allows us to compute a probability that a given decision rule is optimal. We refer to IDT with the additional structure of the prior on losses as Bayesian IDT.

In this paper, we apply IDT to the detection and treatment of cervical intraepithelial neoplasia (precancerous lesions of the cervix). Specifically, we consider a single sequence of Papanicolaou (Pap) smear, possibly followed by colposcopically directed biopsy. We allow that that treatment may occur after the Pap smear, after colposcopy, and/or after biopsy. We consider three possible disease states for the patient: no disease, low-grade SIL, and a combined state of high-grade SIL and cancer. We formulate a probability density on the values of the losses in part by using an existing medical decision-making model [2]. Failure to treat low-grade and high-grade SIL has both economic consequences (e.g., cost savings from not treating now, but possible increased costs later if the disease progresses) and health outcome consequences (more radical treatments if the disease progresses to cancer and possible reduction in life expectancy) which the medical decision making model helps us to assess.

We believe that the utilization of Bayesian methodology is central to IDT. The basic medical decision problem is treated from a Bayesian perspective. Estimation of the parameters of this model (i.e., the prevalences and conditional probabilities of test outcomes given disease states) is performed in a Bayesian manner. This provides a logically consistent and simple approach to dealing with the uncertainty in these estimates. Essentially, it is the same approach as in a probabilistic sensitivity analysis. Finally, we use Bayesian methodology by specifying the prior probability density on our space of losses in order to characterize our uncertainty about the particulars of the loss function and to focus our attention on regions of the loss space which are “reasonable.” With this approach, we can compute a probability that a given decision rule is, in fact, optimal.

The results of our analysis have potentially great importance for clinical practice. We found somewhat to our surprise that the current standard of care for detecting and treating cervical neoplasia was probably not optimal. The decision rule that most likely is optimal, according to our analysis, is one that agrees with the current standard of care in all respects except that patients who have a positive Pap smear are to get one or more biopsies even if they are colposcopically negative. This, of course, leads to unnecessary biopsies for patients with a false-positive Pap smear, but detects disease in patients who have a false-negative colposcopy. Our interpretation is that the additional cost of the unnecessary biopsies is more than offset by disease that would go undetected by colposcopy.

Further investigation is needed to overcome some of the limitations of the analysis presented here, but our results indicate that Bayesian IDT has potential for application to an important medical problem: determining optimal treatment strategies.

Materials and methods

Diagnosis and treatment of cervical neoplasia

We will apply IDT to decision rules associated with the detection and treatment of the precancerous condition cervical intraepithelial neoplasia (CIN). Epidemiological and clinical evidence [3,4] shows that we can prevent the progression to cancer if abnormal cells are detected and treated in the precancerous stages, and this results in a significant reduction in the number of cancer cases and mortality. We consider several decision rules, including the standard of care (SOC), which is depicted in Fig. 1. This SOC for the detection and treatment of cervical neoplasia has been used for many years. It consists of a sequence of tests starting with a Papanicolaou (Pap) smear, which is a sample of cells from the cervix examined by a pathologist [5,6]. If this is positive, then the patient progresses to colposcopy (colpo), which is a visual examination of the cervix under magnification. Colposcopically abnormal sites are then biopsied, and the biopsy is examined by a pathologist for a final diagnosis. A negative diagnosis at any stage results in no treatment. After biopsy, a diagnosis of low-grade (LG) squamous intraepithelial lesions leads to cryosurgical treatment (Cry0) of the cervix. This involves the use of a probe cooled by nitrous oxide or other non-toxic gases to freeze the abnormal cells. A diagnosis of high-grade (HG) squamous intraepithelial lesions results in a surgical procedure called Loop Electrosurgical Excision Procedure (LEEP), where part of the cervix is removed. A diagnosis of cancer will result in more radical surgery, radiation therapy, and/or chemotherapy.

In practice, results of the tests are actually ordered categories, starting from normal and progressing through varying degrees of abnormality. To make our model more tractable, we consider three health states and dichotomize the Pap smear and colposcopy results according to the current thresholds for a “positive” result. We allow the biopsy results to be one of three disease states: normal (N), low-grade (LG) squamous intraepithelial lesions, and high-grade (HG) squamous intraepithelial lesions. The term “squamous intraepithelial lesions” (SIL) refers to a diagnostic system developed by a committee of cancer prevention experts at the National Institute of Health (NIH) in the late 1990s. The classification of histology includes normal, low grade, high grade, and cancer. In this paper, we also refer to neoplasias (cervical precancers and cancers), which is a shortened form of cervical intraepithelial neoplasia (CIN) and more refined SIL diagnostic categories. After biopsy, patients receive one of the three treatments, depending on their health status.

We do not consider cancer and its various treatments. Although this is a simplification of the screening process, the important aspects are present.
Inverse decision theory

Let $\theta$ denote the disease state of the patient, which is constant, $x$ be the observed outcome (a sequence of tests), $f(x|\theta)$ be the likelihood, the probability of observing test outcome $x$, conditional on disease state, $\theta$. Let $\pi(\theta)$ be the prior, the probability of a particular disease state before a test outcome is known. We use these functions to find a posterior, the probability of a particular disease state given specific test outcomes. Consider a loss function, $\lambda(x, a)$, where $a$ is an allowable action. Let $c_i$ be the cost for stage $i$ in the sequence of diagnostic tests and $n$ be the total number of tests. We assume a finite action space and finite parameter space. By finite, we mean the elements can be indexed by integers, $1, 2, 3, \ldots, k$ inclusive. For each decision rule, $\delta$, the Bayes risk is given by

$$B(\delta, c, \lambda) = \sum_{\theta} \sum_{x} \lambda(x, \delta(x)) f(x|\theta) \pi(\theta) + \sum_{i} c_i \Pr(\text{test } i \text{ is performed})$$

(1)

We are indicating in the notation that the Bayes risk depends on the loss function $\lambda$ and cost vector $c = (c_1, \ldots, c_k)$. A decision rule, $\delta^\star$, is optimal for a given $(c, \lambda)$, if and only if $B(\delta^\star, c, \lambda) \leq B(\delta, c, \lambda), \forall \delta \in \delta$, where $\delta$ is the set of all allowable decision rules.

Finding an optimal rule requires knowing all the quantities appearing in Eq. (1). For now we will assume that all the probability distributions are known. Then, for a given decision rule, $\delta^\star$, there will be a set of costs and losses for which $\delta^\star$ is optimal, although this set may be an empty set. We call this set the region of optimality and denote it by

$$\text{RO}(\delta^\star) = \{(c, \lambda) \in \mathcal{L} : B(\delta^\star, c, \lambda) \leq B(\delta, c, \lambda), \forall \delta\}.$$  

(2)

Here, $\mathcal{L}$ denotes the set of possible costs and losses, which we take to be nonnegative. To find the region of optimality for a decision rule of interest, $\delta^\star$, we compare the Bayes risk expression for $\delta^\star$ to the Bayes risk expressions of all other decision rules. Since the Bayes risk is linear in $(c, \lambda)$, this gives linear inequality constraints which define $\text{RO}(\delta^\star)$.

The IDT problem is to determine $\text{RO}(\delta^\star)$ for a given $\delta$. As this may be a subset of a very high dimensional space, it is a challenge to find ways of investigating it to see whether it contains values of the costs and losses which are reasonable.

One approach to this problem is to put a probability distribution on costs and losses, which can then be used to compute the probability that the specific decision rule is optimal. This is easy to do via simulation. We simply make draws from the distribution on $(c, \lambda)$ and see if they fall in $\text{RO}(\delta)$.

An IDT analysis can be interpreted in two very useful ways. First, if we believe that $\delta^\star$ is optimal, then $\text{RO}(\delta^\star)$ gives us information about the loss function. If we do not necessarily believe that $\delta^\star$ is optimal, then $\text{RO}(\delta^\star)$ characterizes the losses that would occur if $\delta^\star$ were optimal. Then these losses can be evaluated to assess whether $\delta^\star$ is an optimal or near-optimal decision rule.

Formulating the IDT problem

To apply the IDT method to any decision rule, we must first find its region of optimality. In the past, our application of the IDT method to decision rules associated with the detection and treatment of cervical neoplasias involved binary disease states, test outcomes, and treatment options [7]. We used a backward induction technique [8,9] to develop the Bayes risk expressions for the sequential decision rules. However, it becomes tedious to derive these inequalities for several disease states. Therefore, we implement a brute force algorithm to identify all possible decision rules and compare a decision rule of interest to all other decision rules.

Recall that we have simplified the problem somewhat by using only binary outcomes (positive and negative) for the Pap and colpo tests, ignoring the possibility of cancer, and assuming that all patients get the Pap smear. This latter assumption means the cost of Pap is common to all Bayes risks, and so we can ignore it when computing the Bayes risks as it would cancel out when comparing the Bayes risk of one decision rule to that of another.

Implementing the algorithm

The first part of the algorithm involves representing all possible decision rules. Fig. 2 is a diagram of the data structure that we use to represent all possible decision rules. At each node there are four possible decisions that we can make. These represent the three treatments or the decision to do further testing, either colposcopy or biopsy depending on the treatment stage. The values for each of those decisions at each node are shown below:

- treatment: 0 (NONE), 1 (Cryo), 2 (LEEP)
- continue to test: −1 (colposcopy or biopsy)
- node not reachable: NA (not applicable)

The decision rules are represented as arrays of length 6 with an entry from −1 to 2 or an NA representing the particular action as above. An entry of NA indicates that the particular node is not reached, e.g., in the SOC we stop and give a treatment of “NONE” after a negative Pap, so nodes 3 and 4 are not reached. To enumerate all the admissible decision rules, we proceed through the tree in the following manner.

1. Start at node 1 (negative Pap result) and choose one of three treatments or test (colpo).
2. If the decision is one of the three treatments, then go to step 5.
3. Else go to node 3 (negative colpo result after negative Pap) and choose one of three treatments or test (biopsy).
4. Next go to node 4 (positive colpo result after negative Pap) and choose one of three treatments or test (biopsy).
5. Next go to node 2 (positive Pap result) and choose one of three treatments or tests (colpo).
6. If the decision is one of the three treatments then the decision rule is complete.
7. Else go to node 5 (negative colpo result after positive Pap) and choose one of three treatments or test (biopsy).
8. Next go to node 6 (positive colpo result after positive Pap) and choose one of
three treatments or test (biopsy).
9. At node 6 the decision rule is complete.

We repeat this procedure until we have enumerated all decision rules. This
algorithm is further described in Table 1. We end up with 361 decision rules
using this algorithm. Note that there are $2^4 \times 4^4 = 1024$ ways of assigning a value
from $\{-1,0,1,2\}$ to nodes 1 and 2 and a value from $\{0,1,2\}$ to nodes 3 through 6
(there is no option to continue testing at these nodes). Our algorithm only lists
the unique decision rules by not assigning a value to a node that is not reached.
Table 2 gives two examples of the vectors representing decision rules: the
current standard of care and see-and-treat colposcopy. In see-and-treat
colposcopy, patients are treated based on colposcopy result, avoiding a biopsy.
This procedure has been proposed for low resource settings [10].

These decision rules are very similar. At node 1, for a negative Pap result we
decide to not treat. At node 2, a positive Pap test results in a decision for further
testing (colpo). Nodes 3 and 4 are empty because they are never reached, since
the action results in stopping at node 1. At node 5, We decided not to treat a
negative colposcopy result after a positive Pap result. At node 6, under the
current SOC we decide to do further testing, which is a biopsy. See-and-treat
colposcopy involves treating a positive colposcopy with an LEEP. Thus, at node 6,
positive colposcopy after positive Pap, we decide to LEEP the patient in
the see-and-treat colposcopy strategy. All the other decision rules are con-
structed and represented in a similar manner.

The next step of our algorithm involves finding the coefficients for the costs
and losses in the Bayes risk for each decision rule.

Finding the Bayes risk for each decision rule

Once we have enumerated all the decision rules, we need the coefficients for
the costs and losses so that we may determine the expressions for the Bayes risks
for each decision rule. First, we define our notation. Let $\text{Dis}=$ be disease state $j$
where $j \in \{0=\text{Normal}, 1=\text{LG}, 2=\text{HG}\}$. Let Pap$=$ be the Pap smear result
and Colpo$=$ be the colposcopy result where both $j$ and $m$ take values 0 or 1,
indicating a negative or positive test result, respectively. Finally, let Trt$=k$
represent treatment action where $k \in \{0=\text{None}, 1=\text{Cryo}, 2=\text{LEEP}\}$. We assume
the biopsy is perfectly accurate; once it is performed, the result equals the true
disease state, so we do not need to consider it as a separate outcome.

Let $\lambda_k$ be the loss incurred if a patient in disease state $j$ receives treatment, $k$.
In general, one would allow for negative losses since treatments are supposed to
be beneficial. However, it is easy to see that for the purpose of comparing Bayes
risks for any pair of decision rules when the losses are known, the minimal loss
for each disease state can be subtracted from all losses for that disease state
without affecting the comparison. Thus, when the losses are known, we can
assume that for each $j$, $\lambda_k=0$ for some $k$ (corresponding to the best treatment for
that disease state). We will make the assumption that $\lambda_0=0$, meaning that the
current standard treatments of care for each disease state are the best among the
three treatments for that disease state. This reduces the dimensionality of the IDT
problem because $\lambda_j$ are known, and also simplifies the problem by eliminating
negative values for the $\lambda_j$. Thus, $\lambda_k$ is the net loss for using Trt$=k$ when Dis$=j$.
There are also costs for the colposcopy and biopsy tests, denoted $c_{\text{colpo}}$ and $c_{\text{biop}}$,
respectively. It is convenient to set the cost of colposcopy equal to 1:
\[ c_{\text{colpo}} = 1. \] (3)

This pins down a proportionality constant for all losses which is indeter-
minate. Thus, all costs and losses are in “cost of colposcopy” units. This will

<table>
<thead>
<tr>
<th>Node</th>
<th>Term in Bayes risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$Pr(\text{Dis}=1, \text{Pap}=0)\lambda_{10} + Pr(\text{Dis}=2, \text{Pap}=0)\lambda_{20}$</td>
</tr>
<tr>
<td>2</td>
<td>$Pr(\text{Pap}=1)c_{\text{colpo}}$</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>$Pr(\text{Dis}=1, \text{Pap}=1, \text{Colpo}=0)\lambda_{10} + Pr(\text{Dis}=2, \text{Pap}=1, \text{Colpo}=0)\lambda_{20}$</td>
</tr>
<tr>
<td>6</td>
<td>$Pr(\text{Pap}=1, \text{Colpo}=0)c_{\text{biop}}$</td>
</tr>
</tbody>
</table>

The terms in the Bayes risk in the second column correspond to the particular
action taken at the node in the first column so that the Bayes risk of SOC is the
sum of all terms in the second column.
allow us to make inferences in terms of dollars when a monetary value is assigned to colposcopy. With these assumptions and conventions, a region of optimality \( \text{RO}(\delta) \) is a subset of a seven-dimensional space of \((\lambda_{00}, \lambda_{02}, \lambda_{12}, \lambda_{10}, \lambda_{20}, \lambda_{21}, c_{\text{colpo}})\), where all quantities are nonnegative.

Table 3 summarizes the Bayes risk expression for the current standard of care and indicates how terms in the Bayes risk are associated with nodes in our representation of the decision rule. Similarly, Table 4 gives the terms in the Bayes risk for see-and-treat colposcopy, and Table 5 presents the Bayes risk expression for an alternative decision rule in which patients are biopsied if the Pap test is positive, regardless of colposcopy outcome. We refer to this decision rule as “Confirmatory Biopsy.”

The final step of IDT involves the comparison of the Bayes risk expression for, say, the SOC to the Bayes risk expression of all other decision rules to obtain \( \text{RO}(\delta^{\text{SOC}}) \). This is done by subtracting the matrix components of all decision rules from a matrix of the same dimensions with the SOC components, forming linear constraints, where the cost \( c_{\text{colpo}} \) and losses \( \lambda_{jk} \) are unknown. Letting \( \lambda \) denote the vector of costs and losses, this will take the form \( A \lambda \leq b \) where \( A \) is \( 360 \times 7 \) matrix and \( b \) is a 360-dimensional vector (recall there are 361 decision rules). Our convention in Eq. (3) means that there will be nonzero values in \( b \) corresponding to decision rules wherein colposcopy is performed under different conditions.

### Estimation of coefficients

In this section, we present the method used to estimate the probabilities that appear in the coefficients for the unknown cost and losses in the inequality constraints defining the region of optimality. Note that these probabilities are needed for both forward and inverse decision theories. We use a Bayesian approach to estimate these probabilities. This allows us to incorporate prior information and to deal with uncertainty in our estimates by propagating it into uncertainty about \( \text{RO}(\delta) \).

The data used in our analyses are part of an ongoing program project, “Optical Technologies for Cervical Neoplasias,” funded by the National Cancer Institute. This project is further described in Ref. [11]. It includes subjects from a screening population, patients who coming in for a Pap Smear, and a diagnostic population, patients who have had a positive Pap Smear and have been referred for a colposcopy. Each data point consists of three observations (Disease State, Pap Smear, Colposcopy) for each subject. Disease state is determined by the (consensus) biopsy which has had at least two readings (a third reading if there was disagreement in the first two). Negative and positive results from the Pap smear and colposcopy will be denoted 0 and 1, respectively. Disease state for negative, low grade, and high grade are denoted as 0, 1, and 2, respectively. The screening and diagnostic populations will be denoted as \( s \) and \( d \), respectively. The data are summarized in Table 6.

Let

\[
x_p \cdot y_k = \text{number of subjects from Population p with Dis} = i, \text{ Pap} = j, \text{ and Colpo} = k,
\]

with \( \text{Dis} = i, \text{ Pap} = j, \text{ and Colpo} = k \), where \( p \in \{s, d\}, j, k \in \{0, 1\}, \text{ and } i \in \{0, 1, 2\} \). The corresponding cell probability parameters are

\[
p_{p \cdot y_k} = \text{Pr(Dis} = i, \text{ Pap} = j, \text{ Colpo} = k | \text{Population} = p).
\]

We assume independent random samples from each population so the likelihood is a product of two multinomials

\[
f(x | p) = \prod_{p \in \{s, d\}} \prod_{j,k} \frac{p_{p \cdot y_k}^{x_p \cdot y_k}}{x_p \cdot y_k!}
\]

Here, \( x \) and \( p \) denote the vectors of observations and parameters, respectively.

Because medical tests are characterized in terms of sensitivities and specificities, it is convenient to reparameterize the likelihood in terms of the conditional sensitivities, specificities, and prevalences of disease states. Define

\[
q_{p \cdot i} = \text{Pr(Dis} = i | \text{Population} = p),
\]

\[
q_{p \cdot j} = \text{Pr(Pap} = j | \text{Population} = p, \text{ Dis} = i),
\]

\[
q_{p \cdot k} = \text{Pr(Colpo} = k | \text{Population} = p, \text{ Dis} = i, \text{ Pap} = j).
\]

Note that \( q_{p \cdot 0} = 1 \) since the Pap smear is always positive in the diagnostic population. We assume that conditional on disease state and Pap smear, colposcopy is independent of the population so \( q_{0 \cdot y_k} = q_{1 \cdot y_k} = q_{2 \cdot y_k} \). With these provisions, an independent set of the \( q \)-parameters consists of

\[
q_{0 \cdot 01} \cdot q_{0 \cdot 02} \cdot q_{0 \cdot 03} \cdot q_{0 \cdot 04} \cdot q_{0 \cdot 21} \cdot q_{0 \cdot 00} \cdot q_{1 \cdot 00} \cdot q_{2 \cdot 01} \cdot q_{0 \cdot 10} \cdot q_{1 \cdot 11} \cdot q_{2 \cdot 21}.
\]

which are the prevalences in the screening and diagnostic populations, the sensitivity and specificity of the Pap smear in the screening population, and the conditional sensitivities and specificities of colposcopy given negative and positive Pap smear results, respectively. For convenience, let

\[
x_p \cdot y_k = \text{number of subjects from Population p with Dis} = i, \text{ Pap} = j, \text{ and Colpo} = k,
\]

### Tables

**Table 4**

Bayes risk for see and treat

<table>
<thead>
<tr>
<th>Node</th>
<th>Term in Bayes risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \text{Pr(Dis} = 1, \text{ Pap} = 0) \lambda_{10} + \text{Pr(Dis} = 2, \text{ Pap} = 0) \lambda_{20} )</td>
</tr>
<tr>
<td>2</td>
<td>( \text{Pr(Pap} = 1) c_{\text{colpo}} )</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>( \text{Pr(Pap} = 1, \text{ Colpo} = 0) \lambda_{01} )</td>
</tr>
<tr>
<td>6</td>
<td>( \text{Pr(Pap} = 1, \text{ Colpo} = 1) \lambda_{11} )</td>
</tr>
</tbody>
</table>

The terms in the Bayes risk for the “confirmatory biopsy” decision rule are given along with the corresponding nodes.

**Table 5**

Bayes risk for the “confirmatory biopsy”

<table>
<thead>
<tr>
<th>Node</th>
<th>Term in Bayes risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \text{Pr(Dis} = 1, \text{ Pap} = 0) \lambda_{10} + \text{Pr(Dis} = 2, \text{ Pap} = 0) \lambda_{20} )</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>( \text{Pr(Pap} = 1, \text{ Colpo} = 0) \lambda_{01} )</td>
</tr>
<tr>
<td>6</td>
<td>( \text{Pr(Pap} = 1, \text{ Colpo} = 1) \lambda_{11} )</td>
</tr>
</tbody>
</table>

The terms in the Bayes risk for the “confirmatory biopsy” decision rule are given along with the corresponding nodes.

**Table 6**

Cell counts for screening and diagnostic patients

<table>
<thead>
<tr>
<th>Population</th>
<th>Disease state</th>
<th>Pap</th>
<th>Colpo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Neg(0)</td>
<td>Pos(1)</td>
<td></td>
</tr>
<tr>
<td>Screening</td>
<td>Normal(0)</td>
<td>339</td>
<td>302</td>
</tr>
<tr>
<td></td>
<td>Neg(0)</td>
<td>302</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Pos(1)</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>LG(1)</td>
<td>Neg(0)</td>
<td>41</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Pos(1)</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>HG(2)</td>
<td>Neg(0)</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Pos(1)</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Diagnostic</td>
<td>Normal(0)</td>
<td>90</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Neg(0)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Pos(1)</td>
<td>5</td>
<td>37</td>
</tr>
<tr>
<td>LG(1)</td>
<td>Neg(0)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Pos(1)</td>
<td>5</td>
<td>73</td>
</tr>
<tr>
<td>HG(2)</td>
<td>Neg(0)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Pos(1)</td>
<td>5</td>
<td>73</td>
</tr>
</tbody>
</table>

The totals for each group are given in square brackets and index codes in parentheses.
Then re-expressing the likelihood in terms of the independent set of \( q \) parameters yields:

\[
f(x|q) \propto (q_{000}^\text{loss})^{(1 - q_{000})^{\text{loss}}} \times (q_{101}^{\text{loss}})^{q_{101}^{\text{loss}}} \times (1 - q_{101})^{q_{101}} \times (q_{111}^{\text{loss}})^{q_{111}^{\text{loss}}} \times (q_{211}^{\text{loss}})^{q_{211}^{\text{loss}}} \times (q_{201}^{\text{loss}})^{q_{201}^{\text{loss}}} \times (1 - q_{201})^{q_{201}} \\
\times (1 - q_{000})^{q_{000}^{\text{loss}}} \times (1 - q_{101})^{q_{101}^{\text{loss}}} \times (1 - q_{111})^{q_{111}^{\text{loss}}} \times (1 - q_{211})^{q_{211}^{\text{loss}}} \times (1 - q_{201})^{q_{201}^{\text{loss}}} \times (\delta^\text{loss})^{\delta^\text{loss}} \times (1 - \delta^\text{loss})^{q_{111}^{\text{loss}}},
\]

where \( q \) denotes the vector of parameters.

We will utilize independent beta priors of the form \( f(q) \propto q^a (1 - q)^b \) for all components of \( q \), with the exception of \( q_{111} \) and \( q_{201} \) which have a Dirichlet prior. The generic form of the density is \( f(q) \propto \prod_{i} q_{i}^{a_{i}} \) where \( q_{000} = 1 - q_{101} - q_{211} \). Denote the prior beta and Dirichlet parameters by \( a_{k} \), \( b_{k} \), and \( \pi \), where the subscripts follow the same convention as for the \( q \). The posterior is also a product of independent betas and a Dirichlet with parameters determined in the same way that the beta and Dirichlet parameters are updated using a single binomial or multinomial sample, e.g., the term \((q_{000}^{\text{loss}})^{r} (1 - q_{000}^{\text{loss}})^{-1} \) becomes \( q_{000}^{\text{loss}} \propto q_{000}^{a_{000}} \). Our priors are specified in terms of a prior mean, \( \mu \), and variance, \( \sigma^{2} \). These can be converted to the beta parameters via

\[
a = \mu(\mu - 1)/\sigma^{2} - 1, \quad b = (1 - \mu)/\mu.
\]

Similarly, the mean and variance of the Dirichlet parameters, \( a_{k}, \ b_{k} \in \{0,1,2\} \), are \( a_{k}/\sum_{i=0}^{k} a_{i} \) and \( b_{k}/\sum_{i=0}^{k} b_{i} \), respectively, where \( a_{0} = \Sigma a_{i} \). Because the marginals of \( q_{111} \) and \( q_{201} \) are \( \beta(a_{111}, \Sigma a_{111}) \) and \( \beta(b_{201}, \Sigma b_{201}) \), we set the value of one \( q_{i} \) to be the mean of a \( \beta(a_{i}, \Sigma a_{i}) \) and use its variance to solve for \( a_{i} \). Then we solve for the remaining \( a_{i} \) using \( a_{i} = \mu a_{i}/p_{i} \).

Each region of optimality for a decision rule of interest \( \delta^{*} \) is determined by linear inequalities that are a function of the \( q \)-parameters. Thus, a region of optimality is a random set and the posterior distribution for the \( q \)-parameters can be transformed to a posterior for \( RO(\delta^{*}) \). One way of dealing with this randomness is to simply replace the coefficients by their posterior expectation, which we call the “randomness is simply to replace the coefficients by their posterior expectation,” maximizing problem, \( \sum_{i} a_{i} q_{i} \) subject to the constraints

\[
x^{*} = \arg\max_{x} \sum_{i} a_{i} q_{i} x_{i} \text{ subject to } \sum_{i} a_{i} q_{i} x_{i} \leq r_{i} \text{ for } i = 1, \ldots, m,
\]

where \( A = \begin{bmatrix} a_{1} & a_{2} & \cdots & a_{l} \end{bmatrix} \).

To determine if the \( i \)-th constraint is active (satisfied at equality), we use LP to maximize \( a_{i} q_{i} \) subject to the constraints \( A x \leq b \). If the solution to the maximization problem, \( x^{*} \), satisfies \( a_{i} q_{i} x^{*} = b_{i} \), then \( a_{i} \) is an active constraint. Note that an active constraint for \( RO(\delta^{*}) \) corresponds to another decision rule \( \delta^{*} \) such that \( RO(\delta^{*}) \) and \( RO(\delta^{*}) \) share a common boundary region. When looking for alternative decision rules to a given one \( \delta^{*} \), it would be prudent to begin with rules which correspond to an active constraint for \( RO(\delta^{*}) \).

**Solving the IDT problem**

After we have determined the region of optimality for a decision rule, what can we infer about the losses in this subset of the loss space? Because we are dealing with a region defined by many inequality constraints with seven unknown variables, it is difficult to deal with directly. However, given a point in the loss space, we can determine if it is in the region of optimality by seeing if it satisfies the constraints. Thus, one way to deal with such a complex region is to place a prior distribution on the losses. If we believe the given decision rule is optimal, then using our constraint region, we can obtain a posterior distribution for loss by simulating values from the prior and only keeping points that fall in the region of optimality. This posterior loss distribution can be summarized by marginal distributions, posterior means, etc. Essentially, the posterior is obtained by truncating the prior to the region of optimality. Alternatively, we can simulate from the loss prior and determine for each simulated loss what is the optimal decision rule. Then we can assign a probability to each decision rule that it is optimal.

This methodology is complicated by the fact that the region of optimality is random—it is defined by linear inequalities whose coefficients are random, their distribution being given by the posterior probability parameter distribution as described in the previous section. If we believe a particular rule is optimal, obtaining a posterior loss distribution becomes somewhat more complicated. It can be accomplished by simulating from the prior on losses and the posterior on coefficients as follows.

1. We generate \( N \approx 10,000 \) data points from our prior distribution on the relative costs and losses. Each of the \( N \) points starts out with a weight of zero.
2. We make random draws from our posterior distribution of the coefficients in the constraints that define the region of optimality.
3. We plug in each of the \( N \) points simulated from the loss distribution in 1 above and see which ones satisfy the constraints.
4. After each draw in 2, the points that satisfy the constraints in 3 get their weight incremented by 1/\( m \), where \( m \) is the number of simulated cost/loss points satisfying the constraints. So all our points get reweighed so that the total weight is 1.

The result is a weighted sample of points from the posterior of cost/loss values. If we believe that the decision rule is optimal, this summarizes our information about the costs and losses, and we can explore this posterior using standard methods.

Placing a prior on the losses focuses attention on the part of \( RO(\delta) \), which we think is relevant. Also, if the priors are plausible, each random draw in 2 gives us a probability equal to \( m/N \) that the decision rule of interest is optimal, where \( m \) is the number of points satisfying the constraint, i.e., a probability (under the loss prior) that the decision rule of interest is optimal. We get an overall mean probability by taking an average of these probabilities.

Although we can be somewhat flexible with our choice of priors, we want to make sure that it is reasonable. An approach to obtain a reasonable prior mean is discussed in the next section. Once the means were obtained, we subjectively decided on a variance and created a prior from independent marginal gamma priors.

**Estimation of loss function using a model**

Researchers in the Department of Obstetrics and Gynecology at Duke University Medical Center developed a Markov model to simulate the natural history of human papillomavirus (HPV) infection in a cohort of women from ages 15 to 85 years. The HPVs are a group of more than 70 viruses, some of which are linked to cervical cancer; these are called “high-risk” HPVs. Nearly all cervical precancer and cancer patients are positive for high-risk HPV. Thus, this Markov model provides us with a tool for evaluating the natural progression of HPV infection and cervical cancer as well as the cost and effectiveness of primary and secondary prevention treatment strategies. Further details can be found in Ref. [2]. It also incorporates regular cervical screening and the treatment strategies we have outlined here (None for Normal, Cryo for LG, and LEEP for HG). We modify the model to assess the effects on cost and effectiveness if a patient is misdiagnosed and receives the incorrect treatment during a single screening year of her lifetime. First, we run the model, starting at age 12. Then, at a specific age in the cohort of patients, we interrupt the regular run of the model and select patients in one of the three precancerous health states, normal, LG, or HG, to receive one of the three treatments, no treatment, cryotherapy, or LEEP, irrespective of their true health state. All other patients not in the selected health state are sent to the
 terminal health state, so that we only observe what happens to patients starting in the health state of interest. During this intervention year, the selected patients take all three diagnostic tests so that we can establish a baseline expected monetary cost that would be the same for all patients during that year. Thus, according to the standard of care described above, patients are being overtreated, undertreated, or receiving proper treatment. The following year, the model resumes its normal operation, but with the selected set of patients who had the health state of interest in the prior year. We repeated this process for each of the three health states and the three treatments. We also performed the intervention at ages 20, 30, 40, 50, and 60 (in separate runs, so there was only one intervention in each run) to compare the costs and effects of the treatment at various ages. To get a better understanding of the process, consider the following scenario: at age 30, we interrupt the regular run of the model and select patients in the LG grade health state (by eliminating patients in all other health states). They receive the Pap smear, colposcopy, and biopsy. Then these patients receive the high-grade treatment, LEEP. We then compute the patients’ expected life-years and expected monetary cost of treatment in their lifetime. This can be compared with results of an identical run but with the LG patients receiving cryotherapy or no treatment.

Note that the Duke model [2] includes a discount fact or for cost and effectiveness, which is a common practice in the field of medical decision making. We use undiscounted costs and effectiveness values so that we can express them in terms of today’s dollars. Also, we expect that the monetary cost of diagnosing and treating patients will at least rise with inflation and we believe that there are some ethical issues with discounting life-years. Furthermore, discounting makes the effects of the intervention at different ages highly incomparable.

1. To get per-person total costs and effectiveness, divide stage costs and effectiveness at interruption age and after by the proportion of people in the health state of interest at interruption age. The cumulative sum of these corrected costs and effectiveness values are now the total costs and total effectiveness.

2. Subtract the total cost and total effectiveness of properly treating a patient, according to the standard of care, from the total cost and total effectiveness values of undertreating or overtreating a patient to obtain net cost values, ΔC and net effectiveness values, ΔE, respectively.

3. Convert the ΔE values to dollars using a conversion factor of $50,000. The $50,000 factor [12] is a willingness to pay threshold for a quality-adjusted life-year per year and combine cost and effectiveness using ΔC− (50,000 * ΔE). This is the loss associated with the misdiagnosis and mistreatment of a patient in a particular health state.

4. Divide the net losses from 3 by the cost of colposcopy to obtain losses in terms of colposcopy units.

The use of the Duke model is important because it is difficult to directly assign a “reasonable” loss to a false-negative Pap smear result or a false-positive colposcopy result. With the latter, a patient will go onto biopsy which results in unneeded economic costs, but a false-negative impacts expected life-years and economic costs later on. Thus, the model allows us to objectively quantify those costs and losses.

### Results

Using the procedure described in the previous section, we obtain the total lifetime costs shown in dollars of the various scenarios of treating patients in particular health states at various ages. Recall the health states are Normal (0), LG (1), and HG (2) and treatments are None (0), Cryo (1), and LEEP (2). Denote a “scenario” by (i,j) where i ∈ {0,1,2} denotes the health state and j ∈ {0,1,2} denotes the treatment. In our analysis, we use the age 30 results to estimate the loss function for the misdiagnosis or mistreatment of cervical precancer. We believe this is the best age because it is not too early or late in life and it will accurately reflect the rates of HPV infection in sexually active women as discussed in Ref. [13]. Thus, the effects of the incorrect treatment are correctly estimated. For the sake of brevity, we only report these results.

The cost of LEEP is virtually the same across all disease categories (but, of course, has no effectiveness for patients in Normal health state as you shall see). The highest costs occur from undertreating HG. Also, notice that if HG patients are given the Cryo treatment as opposed to an LEEP, (scenario (2,1)), the cost is higher than with no treatment (scenario (2,0)). This is perhaps due to the severity of the disease and receiving a worthless treatment as opposed to no treatment would result in a higher total cost. The total effectiveness in life-years of the same set of scenarios is also shown in Table 7. Certain trends are clear—e.g., patients with LG or HG who are undertreated suffer from reduced life expectancy, whereas overtreatment has no effect on life expectancy.

We obtain the net cost (ΔC) and net effectiveness (ΔE) values of each of the treatment scenarios by subtracting the cost and effectiveness values where patients in health state receive the wrong treatment (i ≠ j) from the cost and effectiveness values where patients receive proper treatments (i = j). The net cost and net effectiveness are zero if the patient in health state i receives the proper corresponding treatment, j, (i = j) for that health state.

Next, we apply a conversion factor of $50,000 per life-year to convert the net effectiveness values to dollars. Although this value is arbitrary, it is commonly used in the field of medical decision making [12]. Furthermore, it is only used to evaluate a mean for the prior distribution in the loss space. The spread about the mean is quite large so, in principle, we implement a

### Table 7

<table>
<thead>
<tr>
<th>Health state i</th>
<th>Treatment j</th>
<th>Cost</th>
<th>Effectiveness</th>
<th>Net cost</th>
<th>Net effectiveness</th>
<th>Losses</th>
<th>Relative losses</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>3232</td>
<td>66.2292</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>4968</td>
<td>66.2293</td>
<td>1736</td>
<td>0.000004</td>
<td>1734</td>
<td>5.78</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>6504</td>
<td>66.2293</td>
<td>3272</td>
<td>0.000004</td>
<td>3270</td>
<td>10.9</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>5004</td>
<td>66.2405</td>
<td>35.43</td>
<td>−0.0062</td>
<td>342.9</td>
<td>1.14</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>4969</td>
<td>66.2466</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>6506</td>
<td>66.2466</td>
<td>1537</td>
<td>0</td>
<td>1537</td>
<td>5.12</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>7602</td>
<td>65.7169</td>
<td>1108</td>
<td>−0.4296</td>
<td>22,589</td>
<td>75.3</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>9218</td>
<td>65.9410</td>
<td>2724</td>
<td>−0.2056</td>
<td>13,002</td>
<td>43.34</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>6494</td>
<td>65.9410</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Total cost (in dollars) and effectiveness (in life-years), net cost (ΔC), and net effectiveness (ΔE), losses (dollars), relative losses (colposcopy units) for scenario (i,j) where patients in health state i are given treatment j at age 30.
wide range of uncertainty about this value. Combining the resulting values with the net cost values, we obtained losses for each scenario in which a patient in health state, \( i \), receives the incorrect diagnosis and treatment, \( j \), i.e., \( i \neq j \). Finally, we divide each of the losses by the cost of colposcopy ($300 in the Duke model) to obtain the relative losses. The results are what we expect; there are greater losses when patients with disease receive incorrect treatment than those without disease.

The results based on the Duke model runs are used as a prior mean (approximately, we rounded them off to reasonable numbers). The standard deviations were chosen so that the priors were, in general, fairly diffuse. The prior parameters are given in Table 8. The joint prior on costs and losses was taken as independent marginal gamma distributions with these parameters.

Using the prior described above, we generate 10,000 loss points from an independent Gamma with mean and variance equal to the prior mean and variance and plugged them into the 10,000 random posterior constraint regions for the SOC. Then we checked to see how many of the points satisfy each random constraint region.

We found that 0.03% of the loss points fell in the SOC region of optimality. This raises doubts about the optimality of the current standard of care. Of course, there is uncertainty in the randomly generated probabilities and uncertainty in the losses. However, if our priors are reasonable, the current standard of care is probably not the optimal decision rule.

**Finding the optimal decision rule**

If the SOC is not optimal, then what is the optimal decision rule? Now that we have a loss function and a probability model, we can determine an optimal decision rule, e.g., the optimal decision rule under the prior mean. Using decision theory, in the typical forward fashion, we model the diagnosis and treatment and determine the Bayes risk expressions for all admissible decision rules. Then we evaluate these expressions using the prior mean from the loss function. A new decision rule—confirmatory biopsy—shown in Fig. 3, was found to be optimal. Under this rule, patients receive biopsy after a positive Pap smear irrespective of the colposcopy.

Finally, we applied IDT to the confirmatory biopsy rule and checked to see how many of the random loss points fell in its region of optimality. We found that 95.73% of the loss points satisfied the region of optimality for this new decision rule. These results indicate that this decision rule is probably the optimal decision rule. In fact, some practitioners have stated their support of this biopsy, and commonly, biopsy patients with positive Pap results, even if colposcopy is negative, because there may be lesion areas that cannot be detected by the colposcopist.

**Discussion**

The development of Inverse Decision Theory was motivated by the need for a formal mathematical approach for assessing the losses associated with medical decision rules. Similar inverse optimization approaches have been developed \([14,15]\) with applications to economics, network flow, and geophysical problems. To our knowledge, the results presented in this paper and our previous paper \([7]\) are the first developments of the inverse decision theoretic methodology and its application to medicine.

---

Table 8

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda_{01} )</td>
<td>6</td>
<td>4.5</td>
</tr>
<tr>
<td>( \lambda_{02} )</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>( \lambda_{10} )</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>( \lambda_{12} )</td>
<td>5</td>
<td>4.8</td>
</tr>
<tr>
<td>( \lambda_{20} )</td>
<td>75</td>
<td>20</td>
</tr>
<tr>
<td>( \lambda_{21} )</td>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>( \epsilon_{0n} )</td>
<td>1</td>
<td>0.5</td>
</tr>
</tbody>
</table>

\( \lambda_{ij} \) is the loss associated with giving treatment \( j \) to patient in health state \( i \). We use independent Gamma distribution with these means and standard deviations. All quantities are relative to the cost of colposcopy.

---

Fig. 3. Flow diagram of resulting optimal decision rule for diagnosing and treating cervical neoplasias after evaluating Bayes risk expressions using Duke model point.
The general method of IDT, and this study in particular, has some limitations. As with all mathematical modeling, one cannot hope to capture all of the complexity of the real world. We considered binary patient health states and treatment options [7]. In this paper, we extended the method to three patient health states and treatment options, resulting in a more realistic treatment of the decision problem. We would like to consider more refined disease states and include various stages of cancer and the various possible treatments for cervical cancer. It would also be of interest to allow different thresholds for the diagnostic tests. Furthermore, we know there is a positive probability of error in the biopsy, which is difficult to estimate since the biopsy is typically used as the “gold standard” for assessing other diagnostic tests. Introducing these complications would lead to a combinatorial explosion in the number of decision rules, and representing the decision rules algorithmically becomes a much more difficult task. We are developing approaches to solve these technical problems. Nonetheless, we believe that the results that we have obtained are clinically valid and would not change substantially if analyzed with a more complex model. We are currently developing a decision analysis model which is more elaborate than the Duke model [2], and we will perform a cost-effectiveness analysis of the confirmatory biopsy decision rule utilizing this model.

Perhaps the main advantage of IDT is that it bypasses the need to specify particular losses in seeking an optimal (or near-optimal) decision rule. The classical approach in analyzing medical decision methods is through cost-effectiveness analysis, which keeps the monetary costs and health outcomes on separate scales, but to arrive at an optimal rule, one must put them on a common scale, e.g., with the conversion factor of $50,000 per life-year. Using IDT, one can find the range of costs, effectiveness, and conversion factors optimal. Furthermore, there is no explicit perspective in IDT. A payer (e.g., health insurance company) and a patient would probably assign different values to the costs of the tests and the utility of the health outcomes, and IDT may be useful in reconciling these divergent interests. There is also the consideration of alternative settings. IDT can be adapted to evaluate the optimality of alternative care rules for situations where we cannot follow the current standard of care. For example, in the developing world, there is a shortage of trained pathologists and colposcopists to perform the diagnostic tests in the current standard of care, and these can be evaluated by focusing on different regions in the loss space where these tests are assigned higher costs. In fact, see-and-treat procedures have been recommended [16] for this scenario. We are also interested in considering alternative tests that have been recently been proposed such as HPV testing [17].

In this study, the application of Inverse Decision Theory to the problem of detecting and treating cervical neoplasia has resulted in the discovery that the current standard of care can probably be improved upon with a small increase in cost, resulting in an increased life expectancy for patients. In addition, it is applicable to other procedures and medical settings where it may be difficult to quantify treatment losses. Thus, Inverse Decision Theory becomes a useful guide for assessing the losses associated with various treatment strategies, thereby allowing us to improve the overall treatment decision-making process for patients.

Conflict of interest statement
We declare that we have no conflict of interest.

Acknowledgments
The authors would like to thank the Department of Obstetrics and Gynecology at Duke University Medical Center for use of their model [2].

References