Management of incidentally detected Barrett's esophagus in young patients: extending the guidelines with a Markov process cost-utility analysis

Charles R. Boardman

Follow this and additional works at: http://digitalcommons.ohsu.edu/etd

Recommended Citation
http://digitalcommons.ohsu.edu/etd/869
Management of incidentally detected Barrett’s esophagus in young patients:
Extending the guidelines with a Markov process cost-utility analysis

Charles R. Boardman, MS, RN, NP

Oregon Health & Science University, School of Nursing
Student Name: Charles R. Boardman, MS, RN, NP

Degree: Doctor of Nursing Practice

Title of Clinical Inquiry Project:
Management of incidentally detected Barrett's esophagus in young patients: Extending the guidelines with a Markov process cost-utility analysis

APPROVED:

Committee Chair: Deborah C. Messecar, PHD, MPH  
(name and credentials)  
Signature:

Committee Member: Amnon Sonnenber, MD, MSc  
(name and credentials)  
Signature:

Committee Member: Gail Houck, PHD, RN, PMHNP  
(name and credentials)  
Signature:

Christine A. Tanner, R.N., Ph.D, FAAN,  
Interim Dean, School of Nursing  
Signature:

Date: 06/01/12

Submit completed original form to the Graduate Program office.

Revised 12/2011
Introduction

Esophageal adenocarcinoma (EAC) incidence has increased six-fold since 1979, faster than any other cancer in the United States (Brown, Devesa, & Chow 2008; Kubo et al., 2009). In fact, over the final quarter of the 20th century, the incidence rates of EAC increased faster than those of any other solid tumor cancer in the developed world (Claydon, 2004). EAC carries a high burden of morbidity and mortality, with over 7,500 new cases and 5,000 deaths per year; costs of treatment are significant, involving over $527M in direct costs and $1.98B in indirect costs in the United States alone (Lichtenstein et al., 2007; Everhart & Ruhl, 2009).

Diagnosed symptomatically, EAC carries a five-year survival rate of less than 20% (Corley et al., 2002; Cooper et al., 2009).

Barrett’s esophagus (BE) – the replacement of the normal squamous epithelium of the distal esophagus with columnar metaplasia – is the precursor lesion to EAC. Since the British thoracic surgeon Norman Barrett first identified the malignant potential of the lesion that bears his name over 50 years ago, a tremendous flood of research has been committed to developing effective screening and surveillance protocols to reduce deaths from EAC (Barrett, 1950; Barrett, 1957; Spechler, 2011b). However, while concomitant successes in the prevention of colorectal cancer through colonoscopic screening and surveillance have provided a promising parallel model for endoscopic prevention, such prophylactic efficacy has not been replicated in Barrett’s. Nevertheless, the three major gastroenterology societies recommend screening and surveillance for Barrett’s in higher-risk
populations based on theoretical benefit (Lichtenstein et al., 2007; Spechler et al.,
2011; Wang & Sampliner, 2008). Since incidence of both BE and EAC peaks in the
7th decade of life, all screening and surveillance guidelines devote their
recommendations to patients over the age of 50 (SEER, 2012). An exhaustive, multi-
tiered search of MEDLINE, GoogleScholar, and PubMed, finds that as of today, there
is no published guideline or paper that offers quantitatively derived, evidence-based
recommendations on the management of Barrett’s in patients under the age of 50.

Does Barrett’s exist in patients under 50? Based on numerous individual
case studies, expert experience, and abundant data from large population-wide
databases in Denmark and Northern Ireland, the answer is an unqualified ‘yes’
(Hvid-Jensen et al, 2011; Bhat et al., 2011; Coleman et al., 2011). Data from United
States and Northern Ireland show prevalence of Barrett’s in individuals under 50 to
be 0.5 - 1% (Blankenstein et al., 2005; Corley et al., 2009; Bhat et al., 2011). Since
the early 1980’s, BE – as well as Barrett’s cancers – have been diagnosed and
studied in the U.S. pediatric population, definitively demonstrating the clinically-
significant prevalence of the disease across the human lifespan (Hassall, 1993;
Hassall, Dimmick, & Magee, 1993; Jeurnink et al., 2011). The cancer risk in young
Barrett’s patients under the age of 50 appears to be lower than the cancer risk for
those above the age of 50, but remains substantial enough to warrant investigation
and analysis. Surveillance Epidemiology and End Results (SEER) data in the U.S.
shows an annual incidence of 4.5/100000 in individuals under 50; within Barrett’s
patients, this results in an annual cancer risk of 0.045% per year (SEER, 2012). Bhat
et al. (2011) found an annual cancer incidence of 0.05% per year in Barrett’s patients under 50, nearly a third the risk in patients over 50. In patients under 44 in the U.S., Everhart & Ruhl (2009) cite an annual cancer risk of 0.02%. In sum, Barrett’s appears to be prevalent in 1/100-1/200 young patients, and among those, 1/2000 – 1/5000 per year will develop a Barrett’s cancer that will leave them with only a 1/5 chance of being alive 5 years.

Young people commonly undergo upper endoscopy for a number of indications, including persistent abdominal pain, nausea and vomiting, early satiety, dysphagia, globus-sensation, iron-deficiency anemia, and malabsorption syndromes (Sonnenberg et al., 2008). During the course of these upper endoscopies, it is inevitable that the endoscope passes through the distal esophagus. Should a patient have Barrett’s esophagus, it would be detected at this point. It would be, by definition, an incidental finding as there is no evidence-based recommendation for screening patients under the age of 50 for Barrett’s. While the endoscopist pursues the work-up of the original indication – e.g. biopsying the duodenum, aspirating jejunal fluids, or imaging the pylorus – the question of what to do about the incidentally-detected Barrett’s remains unanswered. Barrett’s is indisputably the precursor to a very deadly – but also very rare – malignancy; doing nothing about its incidental detection would not seem justified. On the other hand, applying surveillance strategies based on cost and benefit calculations developed for patients in their 50’s, 60’s, and 70’s does not seem appropriate either, as it would expose both the patient and the system to a great deal of cost and risk without a meaningful
quantification of benefit. Thus, as it stands, the endoscopist must make a decision based on a nebulous admixture of intuition, emotion, and crude, off-the-cuff probability projections. In an age of evidence-based practice, this state of affairs is unacceptable.

While the ideal study with which to generate evidence-based recommendations is the randomized control trial (RCT), due to a range of factors – including cost, scope, logistics, duration, rarity, and ethical considerations – an RCT is not always the most appropriate research modality. The field of medical decision analysis has arisen, in part, to bridge gaps in the evidence landscape efficiently and soundly in such cases where an RCT is inappropriate (Pauker & Kassirer, 1987). The question of how to manage incidentally-detected BE in a young person, framed as an RCT, would necessitate decades of costly follow-up for thousands of patients in order to produce significant, adequately-powered findings. Thus, the majority of research into optimal management strategies in Barrett’s has involved modeling and decision analysis, rather than prospective trials (Barbiere & Lyratzoplous, 2009; Sonnenberg & Fennerty, 2003; Sonnenberg, Soni, & Sampliner, 2002). For reasons that will be presented in the methods section, the most appropriate model for a condition such as Barrett’s is the Markov process.

Following the vote to recommend the Doctorate of Nursing Practice (DNP) as the terminal degree for advanced-practice nurses in 2004, the American Association of Colleges of Nurses (AACN) set about defining the expectations of the role in a series of published positions papers (AACN, 2004; AACN, 2006). Common to all
proposed DNP skill-sets and capacities – in policymaking, leadership, change
management, practice, or research – was the intense and steady focus on improving
measurable practice outcomes (AACN, 2006; Fain, 2008).

Translational research – using existing fields of data in a novel conceptual
construct in order to address a particular clinical conundrum, or to patch a gap in a
guideline – is at the heart of outcomes-oriented research and practice. Medical
decision analysis – including Markov process modeling – is, in its essence, a form of
translational research. New original data is not being generated, per se: rather,
existent, disparate data are drawn together and activated within the parameters of a
context-specific model in order to produce new recommendations on the clinical
dilemma in question. Much in the way that a span bridge is held up, not by itself, but
by the solidity of the land on either side, so too is decision analysis – and all
translational research – supported by the substantial massifs of epidemiological,
biological, and clinical trial data that surround the as yet unanswered question.
Practicing the art of decision analysis – either as a consumer or a producer of such
studies – is very much at the heart of the DNP concept. Thus, this study will not only
answer an important clinical question – and thereby extend the existent evidence-
based guidelines – in a manner appropriate to the DNP formation, but it will also
furnish future DNP’s with a model for how to approach this kind of research
appropriately, rationally, and fruitfully. Indeed, any DNP wishing to impact policy,
protocol, or everyday practice must possess a working intimacy with decision
analysis. Grasping the resource-allocation recommendations of national task forces
or plenipotentiary committees depends vitally on close familiarity with the purposes, terms, and mechanisms of cost-effectiveness analysis. Moreover, once the principles are understood, small, ad hoc decision analyses may be carried out in everyday practice to address recurrent clinical dilemmas, with or without the goal of broader dissemination (e.g. publication) in mind. With decision analysis, the DNP can transcend the obscurity of guesswork and intuition by defining the parameters of the situation in answerable terms and executing a model that can rationally guide their practice.

In the clinical situation under analysis in this study, the desired outcome is to extend the existent guidelines to cover a population currently outside of the evidence-base in order to provide clinicians with a sound, quantitative, evidence-based model for management of incidentally detected BE in young patients. The purpose of this project is to create a robust and dynamic Markov model that encompasses all of the major drivers of outcome, and then to subject the model to sensitivity analyses to determine the most cost-effective strategy for prevention of cancer in young BE patients. For the cost-analysis, a health payer perspective will be adopted; while the societal perspective offers an attractive plenitude of important considerations (e.g. lost productivity, indirect costs), its incorporation into a model of this kind introduces far too much variability and scope, without conferring a corollary increase in the clarity of the findings. The hypothetical patient population in the Markov cohort will begin cycle\textsubscript{0} at age 30, in order to most broadly analyze
appropriateness of management for Barrett’s patients between the ages of 20-40. The questions that will be modeled and analyzed for this population are as follows:

1. What is the most cost-effective strategy for preventing cancer and cancer deaths in young patients with incidentally detected Barrett’s?
   a. Is one-time ablation, with no surveillance, superior to doing nothing at all?
   b. Is standard surveillance superior to doing nothing at all?
   c. Is one-time ablation superior to standard surveillance?
   d. Based on sensitivity analyses, are there circumstances in which the answers to a-c would change?

Superiority will be determined by calculating which strategy produces the lowest Incremental Cost Effectiveness Ratio (ICER) compared to the others.

**Synthesis of Evidence**

While there persists a vibrant controversy over appropriate management of BE (Spechler, 1997; Sharma & Sidorenko, 2005; Spechler, 2011b), there is little doubt that the aggressive cancer for which it is the precursor is rising vigorously in incidence in the United States, with an over 500% increase since 1975 (Pohl & Welch, 2005; SEER, 2012). Among young patients under the age of 50, there has been a >200% increase in Barrett’s cancer incidence over the past four decades (Brown et al., 2008; SEER, 2012). Caught early, Barrett’s cancer can be cured in approximately 80% of cases, employing a combined modality of neoadjuvant
chemoradiation, and surgical esophagectomy (Kuppusamy, Sylvester, & Low, 2011; Choi & Gibson, 2012). However, late stage esophageal cancer has a very poor prognosis, with less than 20% five-year survival rates despite aggressive trimodal therapy (Choi & Gibson, 2012; Cooper et al., 2009).

There are two major strategies in the endoscopic management of BE today. The first – which is the traditional, standard strategy, recommended by all three major societies – recommends surveillance endoscopies every 3 months to every 5 years depending on the findings from the most recent previous endoscopy (Lichtenstein et al., 2007; Wang & Sampliner, 2008; Spechler et al., 2011b). The traditional goal of endoscopic surveillance is not so much to prevent Barrett’s cancers, but rather, to detect them at early, curable stages. Sandick et al. (1998) and Rubenstein et al. (2008) both demonstrated that cancers detected in patients undergoing endoscopic surveillance were over twice as likely to be early stage than those that presented symptomatically from the community (e.g. with weight loss or dysphagia). Furthermore, the Sandick et al. (1998) study showed that patients with surveillance-detected cancers enjoyed a fourfold increase in 2-year survival compared with symptomatically detected patients. The goal of surveillance – as theorized, and as demonstrated in the above retrospective and prospective studies – is not, therefore, to decrease Barrett’s cancers, but is rather to decrease Barrett’s cancer-deaths through early detection at curable stages.

On the other hand, endoscopic ablation – a newer modality that has gained substantial traction in research and clinical practice over the past decade (Shaheen
& Frantz, 2010) – aims to therapeutically devitalize the Barrett’s mucosa in order to reverse metaplasia and dysplasia, restore normal neosquamous epithelium (NSE), and prevent both Barrett’s cancer and Barrett’s cancer-deaths (Shaheen & Frantz, 2010; Triadafilopoulos, 2010). The initial evidence on the efficacy, durability, and cost-effectiveness of ablation is promising, but remains shrouded in uncertainty due to lack of long-term clinical trials (several of which are underway, and data from the five-year mark is currently being studied and published), as well as to the persistently enigmatic natural history of the BE-EAC continuum itself (Spechler, 2011; Inadomi et al., 2009). Nevertheless, several leading experts in the field have already persuasively advocated that this new paradigm – performing ablation to prevent Barrett’s cancers rather than simply aiming to reduce Barrett’s cancer-deaths by finding the cancers early – should supplant the standard strategy of surveillance altogether (El Serag & Graham, 2011; Triadafilopoulos, 2010).

However, radiofrequency ablation (RFA) carries with it a great deal of cost at $5,000-$15,000 per round of treatment, which makes it between 7X – 20X more expensive than a single surveillance endoscopy (Inadomi et al., 2009). Furthermore, ablation is more risky than a routine surveillance endoscopy, with risk for iatrogenic stricture of 2.5% (1/40) and risks of perforation at 0.05% (1/2000) (Inadomi et al., 2009; Das et al., 2008). The high costs of these more-frequent complications also figure into the analysis when considered from a health payer or societal perspective; for instance, a single esophageal perforation costs, on average, $54,000 (Inadomi et al., 2009). Compared with the risk for perforation during routine surveillance
endoscopy, which is approximately 0.002%, the relative risk for this disastrous and costly complication in RFA is increased 25-fold (Inadomi et al., 2007).

Beyond the debate on ablation versus surveillance, which has focused, as is customary in the literature, on Barrett’s patients over 50, the question of what to do with Barrett’s in patients under the age of 50 remains substantially unaddressed, despite its clinical gravity and its relative inevitability in endoscopy labs evaluating patients from a general population. In the endoscopy lab at the Portland V.A. Medical Center (PVAMC), an average of 1,700 upper endoscopies are performed per year, of which approximately 10% are performed on patients under the age of 50 (Boardman & Zurfluh, 2012). Of these 170 EGD’s performed on young patients for a variety of non-BE-related indications (e.g. nausea, malabsorption), 12 yielded diagnoses of Barrett’s esophagus (5 in patients under 30), depicting a local point-prevalence of 7.1% for all patients under 50 undergoing upper endoscopy (Boardman, 2012). Despite the inevitability and relative frequency of this finding, there is no evidence-based study analyzing appropriate management strategies, or quantitatively comparing plausible options to determine optimal cost-effectiveness.

Among Barrett’s researchers in pediatric gastroenterology, there are several experts who recommend a surveillance strategy – implicitly lifelong – for pediatric patients diagnosed with BE (Hassall, Dimmick, & Magee, 1993); however, there is no explicit decision analysis or cost-effectiveness analysis supporting these recommendations. Certainly, from an ethical – and interpersonal – perspective, there is a strong desire to offer something, rather than nothing, to very young
patients who are diagnosed with a premalignant lesion in their esophagus.

However, other than non-evidence-based expert opinion, there is no guidance as to what kind of something to offer – what volume, what intensity, what frequency – in order to confer maximal benefit at minimal cost.

While a number of Barrett’s experts have advocated – and modeled – offering one-time ablation with no surveillance to Barrett’s patients over the age of 50, no such position has been taken for patients under 50 (Inadomi et al., 2009; Triadafilopoulos, 2010). Certainly, if RFA proves to reliably induce a durable remission of the Barrett’s lesion – replaced by normal neosquamous epithelium (NSE) – then the option of ablating all young BE patients once, without subsequent surveillance, appears attractive. However, if durability is less than current projections – in other words, if, after some time, the Barrett’s lesion reforms, leaving the patient essentially back where he or she started, minus the costs and risks of ablation – the strategy is less appealing. On the other hand, lifelong surveillance, beginning at age 30, could mean as many as 20-40 sedated, invasive exams over the lifespan, which would naturally engender high aggregate risks and high aggregate costs for an as of yet unquantified benefit.

Within the Portland Veterans Administration Medical Center (PVAMC) patient catchment, there has been a 14% increase in the pr enrollees under the age of 50 over the past ten years, reflecting national trends of veterans from Iraq and Afghanistan separating from the military and seeking V.A. benefits (Boardman, 2012; Golding, 2011). Consequently, over the past five years, there has been a 3%-

per-year increase in the number of upper endoscopies performed on veterans under the age of 50 (Boardman, & Zurfluh, 2012). Based on data from the local Clinical Outcomes in Research Initiative (CORI) database – which collects all endoscopic data from the PVAMC, including patient demographics and diagnoses, and stores it in remote, query-able silos – there has been a 32% increase in the finding of Barrett’s in patients under the age of 50 over the past decade (Boardman & Zurfluh, 2012). Averaged over the past decade, approximately 6% of patients under the age of 50 per year are found to have incidental Barrett’s on EGD. Within this specific population, there remains an absence of evidence-based guidance as to appropriate management.

In general the topic of how to approach incidental findings has generated increasing debate and analysis over the inevitability of the clinical conundrums incidental findings create; a MEDLINE search of the MESH term "Incidental Findings" and the keyword “incidental finding” showed that for the entire period from 1962-2002, there were 633 relevant articles. From 2002 through today, in contradistinction, the same query turns up 5,240 published articles, a nearly 1,000% increase in scholarly volume in less than a quarter of the period of time. Clearly, with the increasing sensitivity of contemporary diagnostic exams, and the increasing utilization of such exams, the question of how to manage incidental findings has accrued increasing urgency within the fields of health-care decision making, cost analysis, and ethics. Incidental findings often provoke more exams, which in turn can lead to more incidental findings, setting in motion an inflationary
chain of costs and risks without an appropriate check on the intentions and scope of the original exam. While the clinical and economic impact of incidental findings has been explicitly analyzed elsewhere (Berland et al., 2010; Ding et al. 2011; Sonnenberg & Boardman, 2012), and while it is not this project’s intention to directly engage with the topic of incidental findings, the relevance is inescapable. In this case, an exam is performed for one reason and, owing to the robust, vivid resolution of modern high-definition videoendoscopes, a host of other potentially significant findings are inevitably detected along the way, invoking a chain of clinical attention, energy, and cost that may or may not be justified by the significance of the entity detected. In this case, since Barrett’s is a premalignant lesion, there will be little dispute that even incidentally-detected BE deserves a proper, evidence-based management strategy. Indeed, it could be argued that all common incidental findings – bowel thickening or lung spots on CT scans, slightly increased white blood cell count – should be subjected to a similar evidence-based decision analysis in order to formally clarify the appropriate response.

This study was initiated after its author diagnosed a fifth case of Barrett’s in a 20-30 year old within a calendar year. Exhaustive literature searches yielded no sound, valid model for determining the most cost-effective management strategy. It is the intention of this project, therefore, to create such an evidence-based model, to provide a range of recommendations – tempered by findings of the sensitivity analyses – to local colleagues as well as to gastrointestinal endoscopists and primary care providers worldwide through publication of the model and a written
decision analysis in an international journal. The intended outcome is to offer clear, functional guidance for a difficult and not-uncommon clinical dilemma. The next step for this project involves distilling its findings into the format of a publishable decision analysis and submitting it to international gastroenterology journals. The model itself, once constructed, can be retracted or expanded to address other questions of management in Barrett’s esophagus.

**Methods**

The simplest form of decision analysis is the decision tree. A clinical question is modeled with branching paths to distinct states, each of which has an associated transition probability dictating the likelihood of transitioning to, transitioning from, and remaining within, that state. Decision trees are useful for one-, two-, or three-time probability exposures (Sonnenberg & Beck, 1993; Pauker & Kassirer, 1987). However, when there are many cycles of transition with multiple states and multiple, variable transition probabilities, decision trees become extremely “bushy” very quickly, and it becomes logistically and mathematically difficult to capture or interpret the aggregated values of the various states. In the case of Barrett’s, in which tiny transition probabilities are applied iteratively hundreds of times to distribute a very large cohort into and out of a number of states, a decision tree would become a decision forest very quickly, obscuring the cumulative costs and values of the process. Instead, for modeling Barrett’s, a Markov process would be the most productive and appropriate matrix.
In a Markov process, there is a set number of defined states (e.g. Well, Disabled, and Dead, in the simplest structure), and there are defined, evidence-based transition probabilities into and out of those states that are applied to the hypothetical cohort at each cycle until a determined horizon has been met. Utility or Disutility coefficients are applied to time spent in each state and aggregated for each strategy to summarize the total number of Quality-Adjusted-Life-Years (QALYs) that that strategy produces (Sonnenberg & Beck, 1993). Costs are also applied to time spent in each state, and those are also aggregated as the total cost of that strategy. Traditionally, both future costs and future benefits of a given strategy are discounted, in keeping with the principle of time-preference. Time preference refers to an individual’s preference for time and money now over time and money in the future. In economics, the principle of Net Present Value (NPV) is the corollary to discounting in medical decision analysis (Cohen, 2003; Claxton et al., 2010). Colloquially, time preference is evoked by the idiom: “a bird in the hand is worth two in the bush.” Thus, the QALYs generated by a particular strategy, once discounted, are referred to as dQALYs. With cost in the numerator, and dQALY’s in the denominator, the Average Cost Effectiveness Ratio (ACER) is generated for each strategy. When comparing competing strategies, the Incremental Cost Effectiveness Ratio (ICER) of each strategy, as compared with a base case strategy, is calculated. The ICER = \((\text{COSTS}_{\text{STRATEGY 2}} - \text{COSTS}_{\text{STRATEGY 1}}) / (\text{dQALY}_{\text{STRATEGY 2}} - \text{dQALY}_{\text{STRATEGY 1}})\). In short, the ICER shows how much extra it would cost to get one extra dQALY using Strategy 2 compared with Strategy 1. Classically, the accepted threshold for approving a strategy is ICER \(\leq$50,000/dQALY (Grosse, 2008).
one comparative strategy is being tested, each strategy's ICER is calculated against the base-case strategy as well as against the other comparator strategies.

There remains some controversy over the use of the $50,000-per-dQALY threshold for accepting or rejecting one strategy compared with another (Grosse, 2008). For one, the figure of $50,000 arose as an accept-reject threshold value in the early 1990's and has not been adjusted for inflation since then (Grosse, 2008). Adjusted according to inflation in the consumer price index – which has experienced far less dramatic growth than health care costs over the past decade - $50,000 in 1992 dollars would be worth approximately $112,974 today based on U.S. Medical Inflation indices (personal calculation). Clearly, the ICER threshold needs to be reframed, adjusted, and liberalized across cost-contexts. However, weaknesses notwithstanding – and there are many – a threshold of acceptability must, at some point, be determined, if only for the purposes of orienting the analysis around a figure that most clinicians and patients could reasonably understand, rather than the diffuse cumulative millions or billions of dollars generated by one strategy or another over 100,000 patients over 50 life-years. Therefore, rather than rejecting one strategy or another, this study proposes to report the competing ICERs in dCosts/dQALYs in order of ascendancy across several one-way and two-way sensitivity analyses, in order to evaluate the contexts in which each strategy flourishes or fails.

The Markov states and transition points selected for each strategy were derived from published studies, the SEER database, and personal correspondences
with lead database researchers with the Kaiser GERD/Barrett’s database (Douglas Corley, MD, PhD, MPH) and with the Northern Ireland Barrett’s Registry (Shivaram Bhat, MB, BCh, MRCP). The most difficult transition probabilities to clarify involves the annual risks for cancer within the specific age-tranches under the age of 50. Combining a matrix of observed data (SEER, 2012; Hvid-Jensen et al., 2011; Bhat et al., 2011; Brown et al., 2008), and custom incidence data from Drs. Corley and Bhat, decennial transition probabilities were assigned to the cohorts. A background, age-specific mortality rate was assigned to each living state based on data from the U.S. Vital Statistics (National Center for Health Statistics, 2008), as well as from the Centers for Disease Control and Prevention (CDC) WONDER database tables of deaths by cause and age (CDC, 2012). Costs, complication rates, durability coefficients, and stage-by-diagnosis proportions were all derived from published observational data. According to standard cost-efficiency analysis practice, the discount rate for future costs and utilities was the rounded average for the most recent 30-year Treasury Note: 3%. Naturally, such discounting is always somewhat speculative, but sensitivity analyses did not show the final outcomes – the ICERs between strategies – to change even with dramatic changes to the assumptions of the discount rate.

Table 1

Cost assumptions with sensitivity ranges

<table>
<thead>
<tr>
<th>Cost assumptions with sensitivity analysis ranges</th>
<th>Cost (USD)</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGD</td>
<td>831\footnote{1}</td>
<td>500</td>
<td>1000</td>
</tr>
<tr>
<td>EAC treatment:</td>
<td>44651\footnote{2}</td>
<td>33000</td>
<td>100000</td>
</tr>
</tbody>
</table>
Table 2

Transition probabilities between Markov states

<table>
<thead>
<tr>
<th>Transition p’s with sensitivity analysis ranges</th>
<th>Value</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>BE &gt; EAC (ages 30-39)</td>
<td>0.0002¹²³⁴⁵</td>
<td>0.0001</td>
<td>0.0005</td>
</tr>
<tr>
<td>BE&gt;EAC (Ages 40-49)</td>
<td>0.0005¹²³⁴⁵</td>
<td>0.0002</td>
<td>0.0012</td>
</tr>
<tr>
<td>BE&gt;EAC (ages 50-80)</td>
<td>0.0012¹²³⁴⁵</td>
<td>0.001</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Complication rates

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>RFA stricture</td>
<td>0.025⁶</td>
<td>0.05</td>
<td>0.005</td>
</tr>
<tr>
<td>RFA perforation</td>
<td>0.0005⁶</td>
<td>0.001</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Efficacy rates

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Durability of RFA</td>
<td>0.82⁷⁸⁹</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>Detection of early cancer by symptom</td>
<td>0.2</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Detection of early cancer by surveillance</td>
<td>0.8</td>
<td>0.5</td>
<td>1</td>
</tr>
</tbody>
</table>

Notes: BE = Barrett’s esophagus, EAC = esophageal adenocarcinoma, RFA = radiofrequency ablation.

2. Bhat et al., 2011
3. Hvid-Jensen et al., 2011
4. SEER, 2012
5. Personal correspondence with Drs. Bhat and Corley
6. Inadomi et al., 2009
Table 3

*Utility weights for one cycle spent in each Markov state*

<table>
<thead>
<tr>
<th>Utility weights for one year in each of the states</th>
<th>Utility weight</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>BE</td>
<td>$1^{1,2}$</td>
<td>0.9</td>
<td>1</td>
</tr>
<tr>
<td>Post-EAC treatment</td>
<td>$0.97^{1,2}$</td>
<td>0.8</td>
<td>1</td>
</tr>
<tr>
<td>RFA</td>
<td>$0.94^{1,2,3}$</td>
<td>0.93</td>
<td>1</td>
</tr>
<tr>
<td>Post-RFA</td>
<td>$1^{3}$</td>
<td>0.9</td>
<td>1</td>
</tr>
<tr>
<td>Early cancer</td>
<td>$0.91^{1,2,3}$</td>
<td>0.5</td>
<td>0.9</td>
</tr>
<tr>
<td>Late cancer</td>
<td>$0.34^{1,2,3}$</td>
<td>0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>Weighted symptom-detected cancer (80% late, 20% early)</td>
<td>0.452</td>
<td>0.34</td>
<td>0.9</td>
</tr>
<tr>
<td>Weighted surveillance-detected cancer (20% late, 80% early)</td>
<td>0.78</td>
<td>0.34</td>
<td>0.9</td>
</tr>
<tr>
<td>Undergoing surveillance endoscopies</td>
<td>$0.998^{2}$</td>
<td>0.9</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: BE = Barrett’s esophagus, EAC = esophageal adenocarcinoma, RFA = radiofrequency ablation.
1. Inadomi et al., 2003
2. Provenzale et al., 1994
3. Inadomi et al., 2009
The setting for the study is the gastroenterology section at PVAMC, although, in theory, the larger setting is endoscopy labs worldwide; while this analysis arose from local practice considerations, and will impact local practice decisions, its applicability is general, as the phenomenon it addresses is present across all developed countries. Within the organizational structure at PVAMC, both the Section Chief, Judy Collins, MD, and the Division Chief, David Lieberman, MD, would need to accept the necessity of the change in order for it to move forward. In light of the fact that the alternative to implementing a new, evidence-based model is to continue the discrepant, intuitional, heterogeneous strategy of 'no strategy at all,' the likelihood of leadership buy-in is high. At a recent Division meeting, the study was introduced as ‘in-progress,’ and there was warm acclaim and general interest in the outcome. There was a consensus that at the very least a functional quantitative model of the key drivers of outcome in this clinical situation needed to be constructed and analyzed. There was broad readiness to change, because, as mentioned above, the current strategy is ‘no strategy.’ The main restraining force is cost. Currently, the PVAMC is in negotiations with BARRX, the manufacturer of the radiofrequency ablation module and catheters, and, until the contract is signed and the unit is delivered, all RFA's are outsourced to the university endoscopy lab at twofold the cost of performing them in-house. However, as RFA is more and more becoming a standard tool in the Barrett's armamentarium, it is likely the contract negotiations will be finalized and ablations can be performed in-house at in-house costs. Fortunately, at current rates, there will be one new BE-under-50 diagnosis
per month, and so the marginal cost increases of the ablation strategy can be easily absorbed in the operating budget. Were the scale different – if it were ten new patients per month, or one hundred – full implementation would need to await purchase of the RFA equipment in order to remain cost-effective.

The sample chosen for study was a cohort of 100,000 hypothetical 30-year-olds with Barrett’s esophagus. Markov models classically use 100,000 person cohorts to evoke subtle discrepancies of cost and efficiency across a grander screen (Sonnenberg & Beck, 1993). The age of 30 was chosen as it is the median age of concern (between 20-40). Patients aged 40-50 are often clinically ‘rounded up’ and treated as though they were 50 and entered into standard surveillance regimens; and there are almost no patients under 20 in the PVAMC catchment. How to manage pediatric BE appears to be an interesting and worthwhile area of study, but it is not within the purview of this analysis. No recruitment was needed; the spreadsheet cells were all willing to participate in the study without objection or qualification.

While it is the intention of this project to extend the evidence-based guidelines and to thereby improve practice by illuminating a complex gray area, the project is not a practice-improvement intervention per se. It is the intention of the author to share the findings with PVAMC Section and Division colleagues, as well as to publish the model and analysis in an international gastroenterology journal in order to influence practice across settings, countries, and milieus. Thus, any evaluation of the impact of this decision analysis will have to await its wider dissemination into the practice community. As stated above, the initial
presentation of the project parameters was met with excitement in the initial locus of change, the PVAMC Division of Gastroenterology.

The main measures of data collection were deep, iterative, multi-tiered searches of MEDLINE, GoogleScholar, CINAHL, and PubMed. In MEDLINE, CINAHL, and PubMed, the MESH terms “Barrett’s esophagus,” “adenocarcinoma,” “esophagus,” “incidental findings,” “Markov chain,” “decision analysis,” “discounting,” “cost-benefit analysis,” and “epidemiology” were used in AND and OR sequences with varying year-of-publication controls. In GoogleScholar, natural language searches were performed for key phrases, concepts, and entities. Throughout this process, secondary and tertiary searches of the reference-lists of relevant articles were performed. Secondary data collection to refine model parameters was performed by contacting the lead researchers at major Barrett’s databases, namely, Dr. Corley at Kaiser, and Dr. Bhat at the Northern Ireland Barrett’s Registry (personal communication). Both were forthcoming with key data points regarding age-specific incidence of Barrett’s and Barrett’s cancers in young adults.

In the main, analysis was performed through an Excel (Microsoft, Redmond, WA) spreadsheet Markov process. The initial spreadsheet with key parameters – transition probabilities, costs, utility weights – was constructed and analyzed. Secondary one-way and two-way sensitivity analyses were performed to evaluate the deeper influence of individual parameters or parameters in relation to each other. Other than the time of the author, his committee, and correspondents
abroad, this study did not require any fiscal outlay. The data is presented in natural language, graphs, and tables, depending on which features need to be highlighted. The data is managed on the main Markov process spreadsheet, with secondary and tertiary sheets for graphing and sensitivity analyses.

No human subjects were used in this study. No protected information was analyzed, stored, or disseminated. All of the database figures from Kaiser and Northern Ireland were population-based and denuded of any identifying demographic information.

The findings of this study will be disseminated locally to Division colleagues at the next Division meeting in June. Furthermore, the results and model will be distilled into a publishable manuscript and submitted for review to an international gastroenterology journal. The author has published for three of the major international gastroenterology journals in the past, and serves as an editorial reviewer at two currently; the author is therefore confident that this study will find an appropriate international vector of dissemination

**Results**

Of the three strategies under analysis – S1) Do nothing (natural history), S2) Standard Surveillance, and S3) Radiofrequency ablation and no surveillance – the third appeared the most cost effective. The ICER between S3 and S1 was calculated at $49,511/dQALY, below the unadjusted threshold of $50,000/dQALY, and well below the inflation-adjusted $112,974/dQALY. RFA dominated surveillance,
Table 4
*Average Cost Effectiveness Ratios (ACERs) and Incremental Cost Effectiveness Ratios (ICERs) between strategies*

<table>
<thead>
<tr>
<th></th>
<th>Do nothing</th>
<th>RFA</th>
<th>Surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACER</strong></td>
<td>$58</td>
<td>$306</td>
<td>$714</td>
</tr>
<tr>
<td><strong>ICER</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do nothing</td>
<td>*</td>
<td></td>
<td>$49,511/dQALY</td>
</tr>
<tr>
<td>RFA</td>
<td>*</td>
<td></td>
<td>$244,978/dQALY</td>
</tr>
</tbody>
</table>

*Note: RFA = radiofrequency ablation*

meaning that its ICER is superior but one of the components – the numerator or the denominator – is slightly inferior to its counterpart in the other strategy. In this case, the cumulative dQALY for surveillance is *slightly* higher than the cumulative dQALY for RFA, but its costs so far exceed those of the RFA arm, that it is “dominated” in comparative analysis. Comparing surveillance to the base-case strategy, the incremental cost per dQALY gained is ~$244,977, far in excess of even inflation-adjusted thresholds of acceptability. The main driver of costs within the surveillance arm is the frequent need for expensive endoscopy.

Total cancers were similar between S1 and S3, while S2 resulted in both the lowest total cancers and lowest total cancer-deaths. Naturally, these unadjusted numbers do not reveal the true cost-effectiveness of a strategy. Absolute Risk

<table>
<thead>
<tr>
<th></th>
<th>S1 – No surveillance</th>
<th>S2 – RFA</th>
<th>S3 - Surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total cancers</strong></td>
<td>3,657</td>
<td>639</td>
<td>3644</td>
</tr>
<tr>
<td><strong>Total cancer deaths</strong></td>
<td>2,873</td>
<td>502</td>
<td>718</td>
</tr>
</tbody>
</table>

*Note: S1 = Strategy 1, S2= Strategy 2, S3 = Strategy 3.*
Reduction (ARR) was also lowest in S2, as was Relative Risk Reduction (RRR). Of course, NNT (=1/ARR) was lowest in S2 for preventing one cancer and for presenting one cancer death. Because surveillance does not substantially reduce the total number of cancers, but, instead, reduces the number of cancer deaths because of the relative early-stage advantage, S3 shows significant RRR in cancer death (~75%) when compared with the base-case S1.

Table 5
Within-cohort risk, Absolute Risk Reduction (ARR), Relative Risk Reduction (RRR) & Number Needed to Treat (NNT)

<table>
<thead>
<tr>
<th></th>
<th>S1 - Nothing</th>
<th>S2 - RFA</th>
<th>S3 - Surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAC risk</td>
<td>0.03657</td>
<td>0.00639</td>
<td>0.03644</td>
</tr>
<tr>
<td>EAC death risk</td>
<td>0.0287</td>
<td>0.00502</td>
<td>0.00718</td>
</tr>
<tr>
<td>ARR (EAC)</td>
<td>*</td>
<td>0.03018</td>
<td>0.00013</td>
</tr>
<tr>
<td>ARR (EAC death)</td>
<td>*</td>
<td>0.02368</td>
<td>0.02152</td>
</tr>
<tr>
<td>RRR (EAC)</td>
<td>*</td>
<td>0.82526</td>
<td>0.00355</td>
</tr>
<tr>
<td>(82.53%)</td>
<td></td>
<td>(0.36%)</td>
<td></td>
</tr>
<tr>
<td>RRR (EAC death)</td>
<td>*</td>
<td>0.825087</td>
<td>0.74983</td>
</tr>
<tr>
<td>(82.51%)</td>
<td></td>
<td>(74.98%)</td>
<td></td>
</tr>
<tr>
<td>NNT to prevent one EAC</td>
<td>*</td>
<td>33</td>
<td>7,692</td>
</tr>
<tr>
<td>NNT to prevent one EAC death</td>
<td>*</td>
<td>42</td>
<td>47</td>
</tr>
</tbody>
</table>

Note: S1 = Strategy 1, S2 = Strategy 2, S3 = Strategy 3, EAC = esophageal adenocarcinoma,

Ultimately, the most meaningful output of the Markov process is the ICER between strategies. In order to evaluate the deeper influence of input parameters on the respective ICERs – to see if there are clinical circumstances in which an inferior strategy would become superior – one-way and two-way sensitivity analyses were performed. These showed that if the average number of EGD’s per year for surveillance decreased from 0.455/year to <0.15/year – i.e. one EGD every
7-10 years across the cohort – surveillance could become cost-effective compared to doing nothing. However, current surveillance guidelines typically result in 0.4-0.5 EGD’s/year across BE patients. Should our endoscopic technologies improve sufficiently in the detection of precancerous dysplastic transformations, it may become safe and appropriate to space out surveillance more, in which case this strategy might become cost-effective. This one-way sensitivity analysis, and several others, are presented graphically below, with the ICER always on the abscissa. These graphs are best read by first identifying the ICER threshold of $50,000/dQALY, and then noting when the respective sensitivity lines cross that threshold to determine under which conditions a particular strategy would provide acceptable cost-benefit ratios:
On this plot, it is clear that only when EGD’s approach 1-per-decade (<0.15/year) does the strategy dip below the $50,000/dQALY threshold. Two-way sensitivity analysis below shows that by simultaneously adjusting the cost and the frequency of EGD’s, the strategy can be cost-effective. However, for the current rates of 0.4 EGDs/year, EGDs would have to cost less than $300 for the strategy to be cost-effective; the lowest recorded reimbursement for EGD’s in the literature is ~$500 (Barbiere et al., 2008).

If the efficacy of EGD’s – as measured by the proportion of curable early stage cancers it detects – were improved to >90%, and the cost of EGD’s were <$300, the strategy could prove cost-effective, as demonstrated in the two-way sensitivity analysis below.
RFA sensitivity analyses showed that the estimated durability of RFA – a measure of its lasting eradication of the Barrett’s lesion – is strongly influential over its relative cost-effectiveness, as is the estimated cost of ablation. Other factors, such as utility
weights of RFA, or risk for stricture or perforation, were only weakly influential.

For durabilities > 60%, the RFA strategy appears to be relatively cost-effective. In other words, the current measured durability of ~80% could be a 20% overestimation without altering the fundamental practice implications of the model. The two-way sensitivity analysis below shows that even at half the current average fee for RFA, durability of 20% would not be cost-effective. If durability were improved with new techniques or technologies to near 100%, RFA would be cost-effective at prices up to ~$10,000. Above $10,000, even 100% durability does not ensure cost-effectiveness of RFA.
Discussion

This Markov process analysis shows that between the three strategies analyzed, one-time RFA ablation without any subsequent follow-up is by far the most cost-effective strategy for reducing both cancers and cancer-deaths and delivering maximum quality of life for the minimum of costs. Naturally, this hinges on assumptions about its relative durability which have yet to be borne out by long-term prospective trials. Should its durability prove to be <60% over time, or should the cost of RFA rise above $9,000, it may become a less cost-effective option when compared with doing nothing or standard surveillance. Surveillance, on the other hand, may be most effective if the time between endoscopies were extended to once
every 7-10 years and the cost-per-EGD were decreased by 20-30%. Indeed, as there have been many studies from the United Kingdom demonstrating that nurses can perform routine upper endoscopies with the same safety and efficacy as their physician colleagues, if such a program were put in place in the U.S., costs might be driven down by expanding the pool of potential endoscopists and relaxing the supply side of the curve (Pathmakanthan et al., 2001).

Within the author’s practice setting, once the RFA equipment is purchased and installed, it will be most cost-effective to follow some version of the ablate-once strategy. Further modeling and analysis might be done in the future to explore the cost-utility of ablating once and then re-surveying once in 5-10 years to capture most of those with refractory or de novo Barrett’s and enter them into either a renewed round of RFA or a surveillance program. In any event, the cost of the current ‘no-strategy’ strategy – in patient uncertainty, missed opportunities for cancer prevention, and costly, ineffective overtreatment – is certainly higher than the cost of a regimented strategy such as the RFA arm in this study.

Out of necessity, owing precisely to the gap in the literature that this study attempts to address, this model is constructed on a certain amount of assumptive or inferential transition risks within age-tranches that have not been properly, explicitly studied. Furthermore, much of the RFA data is predicated on ongoing, as of yet incomplete prospective trials whose results may radically shift the model’s output one way or the other. Lastly, as is the case with every model, the major limitation of this study is that it is a model. Models are inherently abstract and
cannot possibly absorb or reflect the innumerable biopsychosocial features of a real patient in a real clinical context. However, models should not be seen as dogmatic, proscriptive templates; rather, they provide an opportunity to analyze the abstract drivers of outcome in order to explore the deeper contours of decision-making. Back in the office, with an actual living patient, a different decision than that prescribed by the model may well be the best one. However, the experience of testing the parameters of our decision-making with a rigorous quantitative abstraction enables us to participate in human-to-human clinical care more humanely and concretely in the end.
References


Centers for Disease Control and Prevention, National Center for Health Statistics. Underlying Cause of Death 1999-2009 on CDC WONDER Online Database, released 2012. Data for year 2009 are compiled from the Multiple Cause of Death File 2009, Series 20 No. 20, 2012, Data for year 2008 are compiled from the Multiple Cause of Death File 2008, Series 20 No. 2N, 2011, data for


