Cost-Effectiveness Analysis for Evaluation of Screening Programs: Hereditary Hemochromatosis

Gregory J. Buffone¹,²,³ and J. Robert Beck¹

A significant body of research over the last 10–20 years supports the hypothesis that screening for hereditary hemochromatosis (HH) may be cost-effective, given the low-cost, low-risk therapeutic options available for most homozygous individuals. The factors that confound a straightforward test of this hypothesis include the fact that the disease is not fully penetrant and that, to achieve the anticipated life-year gains, therapy must be instituted before disease complications become irreversible. Recent articles and editorials, as well as practice guidelines prepared by the College of American Pathologists, recommend screening for HH with transferrin saturation and ferritin testing, and with percutaneous liver biopsy for those with positive laboratory test results. Patients at risk would be treated with phlebotomy for life and monitored with ferritin testing. We present a cost-effectiveness analysis that evaluates the efficacy of using a screening strategy to accomplish the desired healthcare goals.

Indexing Terms: genetic testing/hemosiderosis/ferritin/transferrin/population screening

The rate of development and introduction of new technology intended for the improvement of healthcare has been rapid and seems to be gaining even greater momentum. In addition to the often bewildering rate of change in technology, healthcare institutions are faced with a proliferation of rules and regulations governing healthcare delivery and reimbursement. These regulations will almost certainly place restrictions on our ability to obtain and utilize the latest diagnostic or therapeutic options. As a consequence of market forces, we are now faced with the necessity of defining relatively precise cost and benefit projections for not only new services but also existing operations. Given the aforementioned issues, it is likely that the laissez-faire attitude toward new programs, diagnostic testing, and instrumentation for use in healthcare will not continue and that new services and programs will be scrutinized by one or more bodies or organizations within a hospital, city, or even on a regional basis. Hence there is a need for laboratory-based scientists and physicians to assume a more proactive role in the assessment of laboratory services in the "new world order" as it pertains to healthcare administration.

Commensurate with the impressive technological advances that have enhanced clinical laboratory methodologies and instrumentation over the last few decades, statistical tools for the critical analysis of laboratory methodologies and testing strategies have evolved beyond the levels of analytical precision and accuracy to include predictive value, sensitivity, specificity, decision analysis, and cost–benefit analysis (1–7). A broader view of laboratory services is required that considers new methodologies and testing strategies (e.g., reflexive testing or screening programs for diagnostic efficiency or accuracy) in light of their ability to improve diagnostic accuracy, to aid in the selection of therapeutic alternatives, and to guide early therapeutic intervention. Ultimately, the relative value of test results in influencing outcomes must be measured from several perspectives. From our perspective, the clinical chemist and pathologist can and should be making contributions to our understanding of when and how laboratory services can be used most efficiently and effectively.

What follows is a brief review of the literature and a cost-effectiveness analysis of the hypothesis that population screening for hereditary hemochromatosis as a prelude to lifelong treatment can prevent life-threatening disease complications and is cost-justified given the natural history of the disease.

Cost-Effectiveness Relating to Laboratory Medicine

Cost-effectiveness analysis applied to various aspects of clinical medicine including screening programs and treatment algorithms is well accepted, and provides useful information for practicing physicians in cases where alternative diagnostic and therapeutic options have been identified and characterized. Although decision analysis is not synonymous with cost-effectiveness analysis, it is the basis for defining the choice–chance relations required to assess cost and effectiveness. A selected set of recently published examples applying cost-effectiveness analysis to clinical and laboratory problems is briefly reviewed below to illustrate the value of this approach to assessing the utility of laboratory medicine services.

Birkmeyer and coauthors (8) examined the cost of a quality-adjusted life-year saved if autologous blood donation were provided for two relatively low-risk surgical procedures. In this study the authors examine common blood bank and clinical practices instituted in response to the perceived risk of serious disease or death associated with allogenic blood-product administration. Their findings were remarkable in that one quality-adjusted life-year saved by use of autologous blood-product processing and administration cost ~$1.5 million dollars.

¹Baylor College of Medicine and ²Texas Children's Hospital, Departments of Pathology and Information Technology, Houston, TX 77030.
³Address correspondence to this author at: Department of Pathology, Texas Children's Hospital, MC 1-2262, Houston, TX 77030. Fax 713-770-1032; E-mail gbuffone@bcm.tmc.edu.

Received December 16, 1993; accepted May 2, 1994.

CLINICAL CHEMISTRY, Vol. 40, No. 8, 1994 1631
when primary unilateral knee replacement was the procedure under review. They also examined the cost-effectiveness for bilateral and revision joint replacement and primary unilateral hip replacement when performed at different tertiary-care centers and found that they differed considerably. This study highlights the need for reassessment of laboratory-based healthcare services, in this case blood processing and product utilization, on the basis of a rational consideration of risk and cost as opposed to public or professional anxiety. Healy et al. also looked at the cost-effectiveness of autologous blood procurement in total hip arthroplasty and concluded that it is cost-effective only if secondary effects of autologous blood transfusion are considered (9). This demonstrates that cost-effectiveness analysis can identify unusual “swing variables” and determine the relative impact of various clinical parameters on cost and health outcomes.

The value of leukocyte filters in the preparation of platelets for treatment of induced marrow aplasia in individuals with acute myelogenous leukemia was studied by Baldacci et al. (10). They examined several treatment alternatives that could delay or prevent HLA alloimmunization, thereby facilitating the use of this form of supportive therapy. They concluded that using filtered blood components was more cost-effective than using pooled platelets in preventing or delaying HLA alloimmunization. The uncertainty noted in the conclusion is related to the probability of alloimmunization for each strategy and the cost of single-donor units, assuming their ready availability. An alternative conclusion is that use of filtered blood products does not increase the cost of therapy and carries a lower probability of alloimmunization compared with pooled platelet transfusion. In either case the authors demonstrate the value of cost-effectiveness analysis in assessing the relative value of laboratory products and services.

An article by Eckman and colleagues provides an excellent example of the application of cost-effectiveness analysis to long-term oral anticoagulant therapy in patients with prosthetic heart valves, given the wide variability in thromboplastin preparations used in the measurement of the prothrombin time ratio (PTR) (11). Differences in the sensitivity of rabbit thromboplastin preparations can cause variability in the measured PTR that is unrelated to the true anticoagulation intensity, resulting in increased risk of bleeding or thromboembolic events. To overcome the recognized variability in PTR, the international normalized ratio (INR) has been recommended, where each lot of thromboplastin is characterized by an international sensitivity index (ISI) that calibrates the reagent against the first World Health Organization reference human brain thromboplastin, which has an assigned value of 1.0; thus INR = PTR/ISI. The increase in cost-effectiveness ratio rose rapidly as the ISI deviated either positively or negatively from the expected North American average of 2.4. The data presented clearly demonstrate the cost-effectiveness of reporting INR or ISI in addition to PTR, as opposed to PTR alone when risk of complications and quality-adjusted life-years are considered.

Cost-Effectiveness of Screening for Hemochromatosis

In view of the preceding considerations, decisions concerning the development and implementation of new laboratory services or the revision of existing services should optimally be based on a formal review of clinical criteria, such as when and how to use the test or services, and the therapeutic outcomes that may result from interpretation of test results, as well as traditional analytical issues. Edwards and Kushner recently suggested, on the basis of a heuristic model, that screening for hereditary hemochromatosis (HH) would provide for early detection of individuals who are homozygous for the hemochromatosis gene, thereby reducing their risk for developing irreversible disease (12). Their proposed strategy involved sequential testing by transferrin saturation (%SAT) with repeat testing if the results were abnormal, serum ferritin if % SAT was increased on repeat testing, and liver biopsy if an individual had an abnormal ferritin result. The College of American Pathologists also supports this program in their draft version of a practice guideline for HH. To objectively assess societal cost and clinical impact of such a screening program, we have developed a decision model and performed a cost-effectiveness analysis on the basis of the proposed program.

Decision Model

For the purposes of this model we will limit our analysis to men ages 25 years or older with no history of alcoholism and no preexisting conditions that would predispose to iron loading. Our intention is to identify patients at risk for morbidity and mortality due to hemochromatosis, not to diagnose individuals with evidence of the disease. This approach is in keeping with the recently recommended change in the definition of hemochromatosis, which makes the presence of two HLA-linked hemochromatosis alleles sufficient for diagnosis. According to the traditional definition, based on clinical symptoms and pathology, HH is an inherited autosomal recessive trait, clinically defined by cirrhosis, diabetes, changes in skin pigmentation, endocrine failure, heart failure, and arthropathy. Although the gene for HH is not yet isolated, HLA haplotypes can be used to identify putative heterozygotes and homozygotes within a pedigree once the HLA haplotype of the proband is known (13). The background data required to set up our model, including test cost and performance and information on the natural history and characteristics of HH, are provided in Table 1.

This model is implemented as a Markov cycle tree (16). The initial portion of the tree (Fig. 1) presents the choice/chance options that lead to Markov nodes shown in detail in Fig. 2. (The Markov nodes, denoted by M, are

---

4 Nonstandard abbreviations: PTR, prothrombin time ratio; INR, international normalized ratio; ISI, international sensitivity index; HH, hereditary hemochromatosis (homozygous); % SAT, percent transferrin saturation; pHH, prevalence of HH in the population; HN, heterozygous for HH; and ASR, age, sex, and race.
Table 1. Disease and method background data summary.

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Prevalence of homozygosity for hemochromatosis in the population (pHH)</td>
<td>1/300</td>
<td>14</td>
</tr>
<tr>
<td>2. Percent of male patients who will progress to clinical disease of homozygotes ((\text{pDIS}_{\text{max}}))</td>
<td>40–50</td>
<td>See Appendix</td>
</tr>
<tr>
<td>3. Increased mortality risk for homozygotes ((\mu_{\text{HH}}))</td>
<td>0.017</td>
<td>See Appendix</td>
</tr>
<tr>
<td>4. Sensitivity of (%\text{SAT} (&gt; 82%)) for detection of homozygosity in males</td>
<td>0.92</td>
<td>14</td>
</tr>
<tr>
<td>5. Specificity of (%\text{SAT} (&gt; 82%)) for detection of homozygosity in males</td>
<td>0.98</td>
<td>14</td>
</tr>
<tr>
<td>6. Sensitivity of ferritin to iron overload</td>
<td>0.95</td>
<td>15</td>
</tr>
<tr>
<td>7. Specificity of ferritin to iron overload</td>
<td>0.95</td>
<td>15</td>
</tr>
<tr>
<td>8. Cost of clinical treatment for disease (annual)</td>
<td>$4000</td>
<td></td>
</tr>
<tr>
<td>9. Cost of phlebotomy (annual)</td>
<td>$250</td>
<td></td>
</tr>
<tr>
<td>10. Cost of (%\text{SAT}) assay</td>
<td>$10.5</td>
<td></td>
</tr>
<tr>
<td>11. Cost of ferritin assay</td>
<td>$13.5</td>
<td></td>
</tr>
<tr>
<td>12. Cost of liver biopsy</td>
<td>$350</td>
<td></td>
</tr>
</tbody>
</table>

shown in greater detail in Figs. 2 and 3 and described in the Appendix.) The choices for our model include treating the population with lifelong phlebotomy (TREAT), screening as described (LABTESTS), or no intervention until development of clinical symptoms (NOTREAT). The Markov state diagram, which provides an overview of this portion of the model, is shown in Fig. 3. The initial choice node leads to TREAT, NOTREAT, and LABTEST, which in the cases of TREAT and NOTREAT branch to chance nodes segregating the population on the basis of prevalence of HH in the population (pHH). In the case of LABTEST the population is segregated on the basis of the posttest probability for HH, given the sensitivity and specificity of the applied test as well as pHH. Bayes' theorem is used to calculate posterior probability \((P_{\text{post}})\) of disease based on disease prevalence \((P_{\text{pr}})\), sensitivity \((P_{\text{sen}})\), and specificity \((P_{\text{spec}})\) of an applied test (see Eq. 1). The predictive value of a negative result is simply \(1 - P_{\text{post}}\).

\[
P_{\text{post}} = \frac{P_{\text{pHH}}P_{\text{sen}}}{P_{\text{pHH}}P_{\text{sen}} + (1 - P_{\text{pHH}})(1 - P_{\text{spec}})}
\]  

The additional chance nodes are similarly evaluated. Costs assigned for TREAT/HH/HN (phlebotomy treatment for HH or HN)→PHLEBMHLF (phlebotomy + ferritin monitoring) are those for annual phlebotomy, and for LABTESTS are \(1.3 \times \%\text{SAT}\) cost, based on a 30% repeat testing rate. Each chance node leads to a Markov cycle tree node (Figs. 2 and 3). The Markov states shown in the Markov cycle tree (defined in legend for Figs. 2 and 3) represent the possible outcomes based on several assumptions about the natural history of the disease: (a) disease onset in HH individuals occurs some time after 20–25 years of age with \(\text{pDIS}_{\text{max}}\) (probability of disease adjusted for age) increasing with age; (b) \(\text{pDIS}_{\text{max}}\) has a maximum value of \(-0.5\); and (c) the changing values for \(\text{pDIS}_{\text{max}}\) as a function of age can be represented by a logit function (see Appendix). Given these assumptions, the time spent in a given Markov state can be calculated and the cost of treating disease complications due to HH can be applied to individuals who are assigned to a given cohort (DISEASE) for only the period of time spent in that state [Markov modeling is described in (16)]. The Markov state diagram presented in Fig. 3 shows the possible states and permitted transitions between states allowed in this portion of the overall decision model: e.g., patients who are HN will have no problems related to HH and will die of age-, sex-, and race-related causes (DIEASR); patients who are HH will either be treated with phlebotomy and die of ASR-related causes or will have no early intervention and have.

![Fig. 1. Decision analysis tree for hemochromatosis model.](image-url)

CLINICAL CHEMISTRY, Vol. 40, No. 8, 1994 1833
Fig. 2. Markov cycle tree branches for possible states subsequent to positive or negative ferritin test results.

All possible Markov states for this model are represented in these two examples. DIEASR, death due to age, sex, and race factors; DISEASE, individuals with disease complications due to HH; DEATH, deceased; OKANOW, HH individual who is clinically stable but has manifest complications due to HH; NODISEZE → DIEASR (transition → DEATH), GETSICK, or OKANOW, options for HH individual in this branch who may DIEASR, GETSICK (transition → DISEASE), or remain OKANOW; DISEASE → DIEASR, GETSICK, or OKANOW, options for HH individuals in this branch who may DIEASR, develop disease complications due to HH, or remain clinically stable; NOPROB → OK → STAYOK or DIEASR (transition → DEATH), HH individuals who remain healthy until they DIEASR.

Fig. 3. Markov state transition diagram showing the allowed transitions between states.

Lines between states indicate allowed transitions. Looping arrows originating and terminating at the same state indicate that a patient may remain in that state for one or more cycles of the model as a function of the probability assigned to a transition to another allowed state. Deaths due to age- and sex-related factors or due to HH complications are terminal states.

several possible outcomes—DIEASR, develop liver disease and then DIEASR, or die of liver disease or other complications due to HH.

Cost-Effectiveness Analysis

Cost-effectiveness analysis relies on comparison of one intervention vs another; our case involved the use of laboratory resources to assess risk for iron loading and subsequent disease vs no intervention until disease occurs. The data presented in Table 2 are derived from the computation of the alternative paths (Decision Maker version 6.2; F. A. Sonnenberg, New England Medical Center, Boston, MA) presented in the decision tree shown in Fig. 1, using data from Table 1, and represent average cost and effectiveness for each choice in the scheme. The cost-effectiveness ratio and marginal cost-effectiveness ratio represent the average cost/life-year and average cost/life-year gained for each option. Comparison of tests or testing strategies is best accomplished with an incremental or marginal cost effectiveness ratio, which is

\[
\text{Marginal cost-effectiveness} = \frac{\text{Cost of option } A - \text{cost of option } B}{\text{Effect of option } A - \text{effect of option } B} \quad (2)
\]

This approach is useful for creating a rank-order list when deciding funding priorities (Eq. 2). Marginal cost-effectiveness ratio yields the cost per unit of benefit of choosing one decision strategy over another, whereas average cost effectiveness reflects the cost per benefit of the new testing strategy independent of any other. We can readily see that, on the basis of our initial assumptions, testing, early detection, and treatment are only slightly more costly than waiting to treat at onset of symptoms; on average, $605 per life-year gained (Table 2). Differences in effectiveness are relatively small when averaged across the population, producing a relatively large difference in the cost/life-year gained. [A summary of methods for expressing cost-effectiveness data can be found in a review by Detsky and Naglie (17).]

The cost-effectiveness results obtained with this model are based on certain assumptions, some of which have moderate-to-high degrees of uncertainty. One-way sensitivity analysis, which varies one parameter within a specified range while holding constant all other variables, was performed for parameters that we judged to have a moderate-to-high degree of uncertainty. When sensitivity analysis was performed on the set of cost variables, the cost of testing for ferritin, phlebotomy, and treatment of the disease once symptoms occur were identified as factors with thresholds within the ranges tested (Table 3). Both ferritin and phlebotomy costs are important factors since these costs are incurred when individuals are appropriately or inappropriately se-

---

A threshold is found or defined by the value of a given parameter at which the cost and (or) effectiveness are equal for both strategies being examined.
selected for lifelong phlebotomy treatment and ferritin testing. The cost to treat disease symptoms is important, since 30–50% of homozygotes will develop disease symptoms related to HH during their adult life. Beyond the cost variable, the specificity of both %SAT and ferritin testing are also critical parameters since, as the specificity of these tests decreases, more individuals are inappropriately assigned to phlebotomy treatment and ferritin testing (Table 4). For other parameters, including prevalence, probability of disease in HH individuals, sensitivity of the tests, and excess mortality, the derived threshold values were very close to our initial estimates (Table 4).

Given the sensitivity of our model to test specificity and cost to treat the disease, a useful analysis is provided by a two-way sensitivity analysis of test specificity and cost to treat the disease (Figs. 4 and 5). A two-way sensitivity analysis defines a continuum of threshold values as a function of two critical variables. In our examples, any point above the line favors testing, and below the line, not testing, when cost alone is considered. It is clear from these data that if the cost of disease treatment is only marginally greater than the initial estimate of $4000 and the specificity of each test considered is at least 0.96, then the cost to test and treat, if the test is positive, is equal to or less than the cost to observe and treat when symptoms occur. Unfortunately the most poorly defined factor in our model is the cost to treat hemosiderosis. We based our initial estimate of cost for disease treatment on component cost analysis on 350 patient-years of follow-up data obtained at a Northeast teaching hospital. Unfortunately, the incidence of each complication of HH is not well known. The probability of disease occurring in homozygotes and the relation to age is also poorly defined; more work is required to obtain a good estimate for this parameter.

Nevertheless, our model and the results derived from it indicate that there is a reasonable likelihood that screening would be cost effective. This finding should encourage further efforts to refine the estimates of those parameters shown to be critical factors by our model. We did not discount our cost or life-years in arriving at our final estimates. Cost-effectiveness studies will often report results on the basis of discounts applied to cost and life-years to compensate for the investment value of today’s dollar not immediately spent on healthcare [see Weinstein and Stason for more discussion of this topic (18)]. Discounting reduces the costs and effects in this model in parallel, because most events in this clinical problem occur over a relatively long period and there is no costly or risky procedure (such as surgery) that typically occurs early in the course.

**Appendix**

Hemochromatosis Model Parameters

The model proposed in this paper combines a standard decision tree and a Markov cycle tree for assessment of the cost-effectiveness of screening the population for individuals who are homozygous for hereditary hemochromatosis.

Cost. Cost for testing is based on 1.3 × cost of %SAT
plus the cost of ferritin; adjustment of %SAT cost by 1.3 is based on a 30% repeat testing rate to evaluate an initial abnormal result due to poor control over initial sampling conditions. Test costs were derived from institution-assigned direct cost, not billable cost. Cost for treatment of homozygous individuals identified by screening is based on an initial biopsy (assume no refusal) and quarterly phlebotomy treatment and ferritin testing. We have not included the cost to rescreen individuals with increased %SAT with biannual testing of ferritin.

**Probability of disease.** Estimates of disease probability can be made from anecdotal data in the literature including information summarized by Finch and Huebers (19) and Edwards and Kushner (12). On the basis of clinical experience reported by Finch, one would predict the probability to be only 0.2 for men and women combined. Autopsy data summarized by Edwards (14) indicate that the probability would be closer to 0.3–0.5. In addition, the natural history of the disease, with the accumulation of iron over time, means that the probability of manifestation of disease symptoms and the requirement for treatment increase as one ages. Therefore, we have modeled the probability of disease as a logit function, assuming that probability increases from age 25 to 50 and remains constant thereafter, with a maximum value of \( P = 0.5 \). The following equation was used to calculate the point probability as a function of age in each Markov cycle where \( A = 0.007, B = 0.020, p\text{DIS}_{\text{max}} = 0.5, p\text{DIS}_{\text{adj}} = p\text{DIS} \) adjusted for age, and \( m\text{CYLE} \) equals time increment (1 yr/cycle) in each Markov cycle (Eq. 3).

\[
p\text{DIS}_{\text{adj}} = p\text{DIS}_{\text{max}} \times \frac{Ae^{B \cdot m\text{CYLE}}}{(Ae^{B \cdot m\text{CYLE} + 1})}
\]

(3)

The excess mortality associated with hereditary hemochromatosis (pDIEHH) was estimated with the DEALE method (Eq. 4) (20, 21). The 5-, 10-, and 20-year survival after the development of cirrhosis reported by Adams et al. (22) are 87%, 81%, and 71%, respectively.

\[
p\text{DIEHH} = \frac{1}{\left( \frac{1}{\mu\text{ASR}} + \mu\text{HH} \right) \text{pDIS} + \left( \frac{1}{\mu\text{ASR}} \right) (1 - \text{pDIS})}
\]

(4)

where \( \mu\text{ASR} \) is the predicted age, sex, and race life expectancy for the general population; \( \mu\text{HH} \) is the disease-specific mortality rate; and pDIS is the probability of disease for HH individuals.

**Note added in proof:** The following publication has come to our attention since the preparation of this manuscript: Phatak PD, Guzman G, Woll JE, Robeson A, Phelps CE. Cost-effectiveness of screening for hereditary hemochromatosis. Arch Intern Med 1994;154:769–76.

**References**