ABSTRACTS

JOINT ISOQOL PLENARY - QUALITY OF LIFE

HEALTH-STATE VALUATIONS FOR SELF VERSUS HYPOTHETICAL CONDITIONS: FINDINGS FROM A 14-COUNTRY SURVEY STUDY

Salomon J
Harvard School of Public Health, Cambridge, MA

Background: A number of previous studies have compared individuals’ valuations of their own health with societal valuations of the same states. Most studies have not examined these differences at the individual level (i.e., between a single respondent’s valuation of both self and a hypothetical state with matching description), and there has been little empirical work on valuations in developing countries. We examined differences between self- and hypothetical valuations in a multi-country sample survey study.

Methods: Population-representative surveys were conducted during 2000-2001 in fourteen countries: China, Colombia, Egypt, Georgia, India, Iran, Lebanon, Indonesia, Mexico, Nigeria, Singapore, Slovakia, Syria, and Turkey, as part of a World Health Organization study on health systems. In a health-state valuation module, respondents considered 10 hypothetical conditions with brief descriptions (e.g., “total blindness, acquired as an adult”), in addition to their own current health. For both “self” and hypothetical conditions, respondents provided “profiles” comprising categorical ratings along six dimensions (affect, cognition, mobility, pain, self-care, and usual activities), and assigned values to these states using a visual analog scale (VAS). For individuals assigning a self-matching profile to at least one hypothetical condition, we compared VAS values between the two conditions and examined determinants of the differences using multivariate regression.

Results: Out of 52,050 respondents 7498 assigned at least one hypothetical state the same 6-dimensional profile as their own health. The distribution of differences in VAS scores (self minus hypothetical) had a mean of 21, with a strong positive skew (only 4% of differences were ≤0) and considerable variation across countries (highest mean = 35 in Colombia; lowest = 16 in China). The hypothetical conditions most often assigned self-matching profiles were “mild vision problems” (44%) and “mild hearing problems” (10%). In multivariate analyses, most indicator variables for countries were significant predictors of self-hypothetical differences; sex, age, and years of education were not significant, and most variables relating to levels on the six health dimensions were not significant.

Conclusions: In this study, individuals across a diverse range of countries consistently rated their own health states higher than hypothetical states with the same profiles. These differences may have strong cultural determinants, suggested by significant cross-country variation; however, much of the variation remains unexplained even after controlling for country, sociodemographic characteristics, and health-state characteristics.

MEDICAL DECISION MAKING/JAN–FEB 2006

ABSTRACTS ORAL PRESENTATIONS

JOINT ISOQOL PLENARY - QUALITY OF LIFE

2130

PREPREFERENCE-BASED QUALITY OF LIFE IN U.S. ADULTS: ASSOCIATION WITH BODY WEIGHT

Arts-Mark B1, Lawson L1, Chen L1 and Tsevat J1
1University of Cincinnati, Cincinnati, OH; 2Cincinnati Childrens Hospital Medical Center, Cincinnati, OH

Purpose: To facilitate future economic evaluations of weight loss interventions, we examined the association of body weight with preference-based quality of life in a nationally representative sample of U.S. adults.

Methods: We analyzed data on 27,332 adults in the 2000 and 2001 Medical Expenditure Panel Surveys, two nationally representative, cross-sectional surveys of the noninstitutionalized civilian population of the United States. Height and weight were self-reported, and quality of life was assessed using the EuroQol EQ-5D, a generic, preference-based instrument that contains 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression). We used linear regression to assess the associations between National Institutes of Health body mass index (BMI) categories and utility weights on the EQ-5D. We used logistic regression to predict the probability of having at least some problems on each of the dimensions of the EQ-5D, by BMI categories. Analyses were multivariable adjusted for sociodemographic characteristics (age, gender, race, income, education, marital status, region of country, insurance type), smoking status, and 6 chronic health conditions (hypertension, diabetes, chronic heart disease, stroke, asthma, and emphysema) using STATA SE 8.2 and survey commands.

Results: Thirty-nine percent of US adults had normal body weight (BMI 18.5-24.9), 36% were overweight (BMI 25.0-29.9), 15% had class I obesity (BMI 30.0-34.9), 5% had class II obesity (BMI 35.0-39.9), 3% had class III obesity (BMI ≥ 40.0), and 2% were underweight (BMI < 18.5). When compared with normal weight adults, EQ-5D utility weights were significantly lower for overweight adults, overweight adults, and those with class I, II, and III obesity (P < 0.05). Above a BMI of 25, we observed a linear, inverse relationship between BMI and EQ-5D utility scores, where a five-point change in BMI was associated with a 0.01 change in utility (95% CI: 0.01 to 0.01; adjusted R-squared = 0.15). Obese adults were more likely than normal weight adults to report any limitations in mobility, self-care, and usual activities, and more likely to have chronic pain.

Conclusions: Body mass index is associated with lower utility in a dose-dependent fashion. Further research is needed to test the cost-effectiveness of weight loss interventions in U.S. adults using health state preferences derived from the general population.

MEDICAL DECISION MAKING/JAN–FEB 2006
2166 JOINT ISOQOL PLenary - QUALITY OF LIFE
1:00 PM - 3:30 PM
Saturday, October 22, 2005
1:30 PM Grand Ballroom

EQSD CHANGES RHEUMATOID ARTHRITIS (RA) QUALITY OF LIFE IN UNITED STATES: STUDY OF 11,289 RA PATIENTS
Micha K. and Wolfe F.
Stanford University, Stanford, CA; Arthritis Research Center Foundation, Inc., Wich-
ria, KS

Purpose: Patients with rheumatoid arthritis (RA) have reduced quality of life (QOL). Using the
EQSD health measure, we report the QOL, US EQ5D-US and European (EQSD-E) preferences and
analyse their differences by RA disability index.

Methods: During 2.5 years (2002-2004) in a US observational study, 11,289 RA pa-
tients completed 35,622 EQSD and Health Assessment Questionnaire (HAQ) measurements. The
EQSD was transformed using standard US and newly published preferences (Shaw, Share-
son, and Coons, 2005). All analyses were adjusted by sex, age, and calendar date with clustering
by patient.

Results: The study group had the following mean values (SD): 22.1% male, 62.0 (12.4 years old,
and 1.03 (0.73) HAQ. The average QOL (0.68) was similar to US EQ5D-US and European
(0.72, 0.73) EQSD-US. The table shows QOL and their difference by HAQ category. The per-
centage change in QOL over six months corresponded to a larger mean change in EQSD-E vs. EQSD-US:
0.044 > 0.04 when HAQ increased (N=4,896, mean HAQ change: 0.30 and 0.08 vs. 0.38 when
HAQ decreased (N=1,971, mean HAQ change: -0.30).

Conclusions: Even though the correlation HAQs score is the same in the US and Europe and it is
considered to be the primary outcome of disease severity and therapeutic effectiveness in RA, the
responding EQSD evaluations are significantly different and nonlinear. This impacts effectiveness
studies and could affect coverage decisions of RA treatments, as a reduction in disability corres-
ponds to a reduced improvement in QOL in the US compared to Europe even though the US QOL is
more generous.

Table. EQSD QOL By HAQ Score

<table>
<thead>
<tr>
<th>EQSD-US</th>
<th>EQSD-US</th>
<th>EQSD-E</th>
<th>EQSD-E</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - &lt;0.25</td>
<td>5801</td>
<td>0.90 (0.19)</td>
<td>0.90 (0.19)</td>
</tr>
<tr>
<td>0.25 - &lt;0.5</td>
<td>3569</td>
<td>0.84 (0.16)</td>
<td>0.84 (0.16)</td>
</tr>
<tr>
<td>0.5 - &lt;1.0</td>
<td>3498</td>
<td>0.81 (0.15)</td>
<td>0.81 (0.15)</td>
</tr>
<tr>
<td>1.0 - &lt;1.5</td>
<td>3571</td>
<td>0.78 (0.15)</td>
<td>0.78 (0.15)</td>
</tr>
<tr>
<td>1.5 - &lt;2.0</td>
<td>3908</td>
<td>0.75 (0.17)</td>
<td>0.75 (0.17)</td>
</tr>
<tr>
<td>2.0 - &lt;2.5</td>
<td>3834</td>
<td>0.71 (0.19)</td>
<td>0.71 (0.19)</td>
</tr>
<tr>
<td>2.5 - &lt;3.0</td>
<td>3683</td>
<td>0.65 (0.21)</td>
<td>0.65 (0.21)</td>
</tr>
<tr>
<td>3.0 - &lt;3.5</td>
<td>3248</td>
<td>0.59 (0.23)</td>
<td>0.59 (0.23)</td>
</tr>
<tr>
<td>3.5 - &lt;4.0</td>
<td>2287</td>
<td>0.54 (0.24)</td>
<td>0.54 (0.24)</td>
</tr>
<tr>
<td>4.0 - &lt;4.5</td>
<td>1999</td>
<td>0.49 (0.26)</td>
<td>0.49 (0.26)</td>
</tr>
<tr>
<td>4.5 - &lt;5.0</td>
<td>549</td>
<td>0.42 (0.24)</td>
<td>0.42 (0.24)</td>
</tr>
<tr>
<td>5.0 - 7.0</td>
<td>255</td>
<td>0.31 (0.32)</td>
<td>0.31 (0.32)</td>
</tr>
</tbody>
</table>

1607 JOINT ISOQOL PLenary - QUALITY OF LIFE
1:00 PM, October 22, 2005
1:00 PM Grand Ballroom

THE US AND UK VERSIONS OF THE EQ-SD PREFERENCE WEIGHTS: DO THEY MAKE DIFFERENCES?
L-Chan Huang,1 Albert W. Wu,2 and Mark J. Atkinson3
Health Policy and Management, The Johns Hopkins University, Baltimore, MD; 1La Jolla Laboratories, Worldwide Outcomes Research, San Diego, CA; 2The Johns Hopkins

Aims: Most studies using the EQ-SD in the US have applied the UK preference weights, based
on the assumption that preferences of these two populations differ minimally. We investigated the
validity of a newly developed US version of the EQ-SD preferences compared to the UK version.

Methods: During 2.5 years (2002-2004) in a US observational outcomes study, 11,289 RA pa-
tients completed 35,622 EQSD and Health Assessment Questionnaire (HAQ) measurements. We first compared the
validity of a newly developed US version of the EQ-SD preferences compared to the UK version.

Results: Results of the multi-trait multi-method (MTMM) comparing children and proxy EQSD reports. Discriminant
validity has been assessed by score comparisons between age groups, gender, countries, and social
level. Reliability of the instrument has been assessed with Cronbach's α and test-retest score
comparisons.

Results: Results of EQSA with extensive rotations and MAP analyses are satisfactory. CFA indi-
cates satisfactory fit (RMSEA=0.068, CFI=0.960). The unidimensionality of every dimension has
been confirmed (INFIT: 0.81-1.15). No DIF effect was found (Delta-R2 < 0.4% for all 4 items
across countries. Reliability was good (Cronbach's α = 0.60-0.94, ICC: 0.61-0.74). MTMM model test is
satisfactory (RMSEA<0.027, CFI=0.999). Discriminant validity was supported by significant differ-
ence between age, gender groups, and social level categories.

Conclusions: This 1st European trans-cultural instrument to assess children and adolescents
HRQOL is a new promising tool.

© 2006 Society for Medical Decision Making. All rights reserved. Not for commercial use or unauthorized distribution.
2297

ORAL CONCURRENT SESSION A - QUALITY OF LIFE AND UTILITY THEORY

5:00 PM - 6:30 PM

Saturday, October 22, 2005

5:00 PM Grand Ballroom B

DO TTO ELICITATIONS WORK OVER THE PHONE?

Laskovik A, Ubel P, Smith D and Fagerlin A

University of Michigan, Ann Arbor, MI

Purpose: In a nationally representative sample, we tested the validity of time tradeoff (TTO) elicitation conducted over the telephone. TTO elicitation is cognitively demanding, and some experts have questioned their validity in people with low numeracy. We attempted to test the validity of the phone-administered TTO by comparing it against another utility measure, the subjective health rating scale (0-100).

Methods: We surveyed a random subset of 1031 people participating in the Health and Retirement Study, a representative sample of people in the U.S. age 50 and older. The HRS includes 3 types of health measures: the SF-1, a comorbidity index, and a list of AID limitations. It also includes several measures of cognition: a 3-item numeracy measure, a serial 7 task, and a memory test. We looked at how the TTO and subjective health rating correlated with the health and AID composites. To test if the correlation between health and TTO values was a function of cognition, we divided participants by median split on each of the cognitive measures.

Results: Rating scale scores were significantly correlated with the SF-1, objective health, and AID (r's = –0.250 to –0.579, p's < 0.0001). We found no relationship between the TTO scores and either the SF-1 or objective health (p > 0.006), whereas correlation with AID was significant but low (r = 0.105, p < 0.001). Further analysis showed no relationship between TTO scores and self-reported and objective health measures for subjects who scored below median on any of the cognitive measures, a finding that extends Woloshin's research. However, TTO scores and objective health were significantly correlated for participants who scored above the median on all three cognitive tasks, although correlations were low (r's = –0.096 to –0.119, p's < 0.05). TTO scores were significantly correlated with the SF-1 and AID ratings only for subjects scoring above the median on the serial subtraction task (r = –0.089, p < 0.013 and r = –0.110, p < 0.015, respectively).

Conclusions: The validity of telephone TTO elicitation is questionable among the majority of people over the age of 50, with no correlation between health and TTO among people with below-average cognitive abilities, and low and inconsistent correlations for those with above-average abilities.

2091

ORAL CONCURRENT SESSION B - COST-EFFECTIVENESS ANALYSIS: APPLICATIONS

5:00 PM - 6:30 PM

Saturday, October 22, 2005

5:00 PM Grand Ballroom C

THE COST-EFFECTIVENESS OF ANTIRETROVIRAL THERAPY FOR INJECTION DRUG USERS IN RUSSIA

Leng D, Brandeau M, Vinichenko T, Galvin C, Tole S, Schwartz A, Sanders G and Owens D

Stanford University, Stanford, CA; Duke University, Durham, NC; VA Palo Alto Health Care System, Palo Alto, CA

The HIV epidemic in Russia grew catastrophically in the past decade, primarily because of injection drug use. Approximately 1 million people in Russia are now estimated to have HIV infection. Because of their disproportionate contribution to HIV transmission, provision of HAART to HIV-infected IDUs and non-IDUs in Russia.

Treatment of both groups relative to current practice is very cost-effective by horizon, discounted at 3% annually.

The most effective and expensive strategy is to provide HAART to all people with HIV; this result was less sensitive to assumptions related to the impact of primary angioplasty on the length of the initial hospital admission.

For a shorter delay of 30 minutes the ICER was £6,850 per QALY; increasing the time-delay to 90 minutes resulted in a mean gain of 0.29 QALYs and an additional cost of £2,680 per patient compared to thrombolysis. The associated incremental cost-effectiveness ratio (ICER) was £9,241 per QALY. At a threshold of £20,000 per QALY there was a 90% probability that primary angioplasty was cost-effective. The associated incremental cost-effectiveness ratio (ICER) was £38,237 per QALY. At a threshold of £20,000 per QALY there was a 90% probability that primary angioplasty was cost-effective.
Oral Concurrent Session A - Quality of Life and Utility Theory
5:00 PM - 6:30 PM
Saturday, October 22, 2005
5:30 PM Grand Ballroom B
Utility, Medications, and Symptoms Among Human Immunodeficiency Virus-Positive Individuals
Bavarski A
St. Michael’s Hospital, University of Toronto, Toronto, ON, Canada
Most studies of utilities among Human Immunodeficiency Virus (HIV)-positive individuals predated currently used antiretroviral therapies and accompanying toxicities. We evaluated the association between antiretroviral drugs, symptoms, and utilities. We conducted 206 interviews using a computer to illustrate tradeoffs, conduct surveys, and store data. We elicited utilities using the Rating Scale (RS), Time-Trade-Off (TTO), and Standard Gamble (SG) methods and symptoms using a modified HIV-specific scale. We retrieved drug information from respondents’ charts. The median number of cumulative antiretrovirals used was 4 (interquartile range [IQR] = 3 to 5). At the time of the interview, 69 participants (34%) were not taking antiretrovirals, 61 (30%) were taking 2 or 3 drugs, 36 (18%) were taking 4 drugs, and 65 (32%) were taking 5 or more drugs. Only 6 participants (3%) reported having no symptoms. The typical respondent had 6 (IQR=4-9) mild to 2 (0-5) severe symptoms. The most common symptom was fatigue, which was reported in 66 (34%) and severe in 56 (31%). The intraclass correlation coefficient between SG and TTO scores was 0.65. There was no association between RS, SG, or TTO utility scores and the number of drugs (p for trend = 0.18, 0.68, and 0.31, respectively). There was only one statistically significant association between any utility score and a drug class: participants using a non-nucleoside reverse transcriptase inhibitor had a higher RS score than patients not using such a drug (0.56 vs. 0.47). Several symptoms were significantly associated with SG and TTO utilities, including fatigue, fever, cognitive difficulties, diarrhea, depression, anxiety, skin problems, respiratory problems, anemia, muscle and joint pain, sexual difficulties, and weight loss. Peripheral neuropathy and headache were associated only with the SG, abdominal pain and physical changes related to the lipodystrophy syndrome only with the TTO. Not associated with the SG or the TTO were diarrhea, nausea, sleep problems, and hair loss. RS utilities were associated only with skin and respiratory problems. Each severe symptom decreased SG and TTO scores by 0.007 (0.004 to 0.008) and 0.002 (0.001 to 0.004), respectively. Utility scores were strongly associated with symptom burden in HIV but not with antiretroviral therapy use. Our study suggests that SG and TTO scores are valid measures of symptom burden but RS scores are not.
ABSTRACTS
The cost-effectiveness of HIV screening in Russia.

We evaluated the test-cost-effectiveness of HIV screening in Russia.

We used a Markov model to estimate costs, quality of life, and survival associated with a voluntary HIV screening program compared to no screening in Russia. We modeled a cohort of 15- to 49-year-olds demographically similar to Russia's population, with an average age of 32.5. HIV prevalence in the base case was 1.2%, with two-thirds of the HIV cases undiagnosed. Our base case assumed 70% of those with a positive HIV test would enter care and receive appropriate treatment, as a best case scenario. We included costs of testing, counseling, follow-up, and treatment. The annual cost of HAART was estimated at $1700 for medications plus $650 for social support services. We measured lifetime health care costs and quality-adjusted life years (QALYs) gained, discounted by 3% annually.

Once per lifetime HIV screening increased life expectancy by 1.80 QALYs for infected individuals. For the entire screened population, life expectancy increased by 3.2 quality-adjusted days at an estimated incremental cost of $66, yielding a cost-effectiveness ratio of $7660/QALY gained. When the annual cost of HAART was decreased to $1200, the cost-effectiveness ratio was $5640/QALY gained. The one-way sensitivity analysis in which only those individuals were treated, the cost-effectiveness ratio of screening increased to $8670/QALY gained. When prevalence of the screened population was lowered to 0.1%, the cost-effectiveness ratio worsened to $920/QALY gained; in a higher prevalence-sensitivity scenario, the ratio decreased modestly to $7400/QALY gained. These analyses do not include the potential public health benefits of reduced transmission due to behavior change and treatment: incorporation of these benefits would yield more favorable cost-effectiveness ratios.

Our findings suggest that once per lifetime HIV screening in Russia is cost-effective by the World Health Organization's cost-effectiveness guidelines. To improve the cost-effectiveness of screening, efforts should focus on reducing costs of HAART, ensuring that HIV-infected individuals receive appropriate treatment, and preferential screening of high-risk groups.

We empirically estimate mean and median CD4 counts for patients with history of different OIs and, for patients with CD4 count below 50 cells/L.

We give 95% confidence intervals of each estimate. Our analysis included 1359 patients' records.

Results: We base our estimation on data from the Multicenter AIDS Cohort Study (MACS). We focus our study on adult (age 18 or above), HIV-infected, and antiretroviral treatment naive patients.

We use 2-stage modified Delphi Process included: a) defining 12 classes of quality criteria; b) summarizing their theoretical and empirical links to decision quality; c) drafting and testing voting items on empirical indicators; d) tracking consensus; e) reviewing published reports; f) evaluating items on quality indicators; and g) nominated sampling of stakeholder groups (consumers, practitioners, researchers, payers). Invites were e-mailed (2 reminders) to vote on a Web site within 3 weeks. In the second round, voters were asked to vote on whether each item met their criteria.

We found 73% of 1282 invited (58%) participated on voting panels (14 countries; 77 researchers; 21 countries; 10 practitioners; 14 payers), 103/122 (85%) voted in both rounds. There was consensus on 74 of 83 quality indicators (4 equidistant scores: “9” = 41; “8” = 26; “7” = 27; “6” = 7); equidistant ideal means to categories.

We empirically estimate mean and median CD4 counts for patients with history of different OIs and, for patients with CD4 count below 50 cells/L.

We use 2-stage modified Delphi Process included: a) defining 12 classes of quality criteria; b) summarizing their theoretical and empirical links to decision quality; c) drafting and testing voting items on empirical indicators; d) tracking consensus; e) reviewing published reports; f) evaluating items on quality indicators; and g) nominated sampling of stakeholder groups (consumers, practitioners, researchers, payers). Invites were e-mailed (2 reminders) to vote on a Web site within 3 weeks. In the second round, voters were asked to vote on whether each item met their criteria.

We found 73% of 1282 invited (58%) participated on voting panels (14 countries; 77 researchers; 21 countries; 10 practitioners; 14 payers), 103/122 (85%) voted in both rounds. There was consensus on 74 of 83 quality indicators (4 equidistant scores: “9” = 41; “8” = 26; “7” = 27; “6” = 7); equidistant ideal means to categories.
RESULTS: The DCS demonstrated convergent, construct, and discriminant validity based on the total scale scores. The comparisons at the subscale score level between the intervention and control groups showed a lack of discriminating ability for the uncertainty subscale. The internal consistency reliability for the total score of the DCS was reasonably good in this sample (α = 0.81). All three uncertainty subscale items showed the weakest item-to-total correlation (0.22 < r < 0.33). When the uncertainty subscale items were eliminated, Cronbach’s alpha coefficient improved to 0.84.

Conclusions: The DCS seems to be a viable research instrument for measuring the quality of end-of-life decision making. The DCS provides good discrimination between groups and has proved reliable in the end-of-life decision-making context, especially with respect to modifiable factors contributing to uncertainty and the effectiveness of the decision-making process and the quality of decisions. However, the uncertainty inherent in such decision making may limit the applicability of the uncertainty subscale in such context.

ORAL CONCURRENT SESSION D - PATIENT AND PHYSICIAN DECISION MAKING

8:30 AM - 10:00 AM
Sunday, October 23, 2005
9:00 AM Grand Ballroom B

IMPROVING PATIENT PARTICIPATION IN DECISION MAKING FOR KNEE ARTHRITIS: RESULTS OF A PILOT TRIAL

Pravenek J, Babudini N and Wittack D
Yale University, New Haven, CT

Purpose: Recent studies have shown that patient preferences for treatment of knee osteoarthritis (OA) conflict with widespread physician practices. Improved patient participation in decision making may increase the concordance between patient treatment preferences and physician prescribing patterns. The objective of this pilot study was to test the efficacy of an interactive decision tool to improve patient participation in decision making for knee OA.

Methods: We are conducting an ongoing, single site, non-blinded pilot randomized controlled trial. Patients with knee OA and pain on most days of the preceding month (coming to see their primary care physician for a regularly scheduled visit) were randomized to receive an Arthritis Foundation pamphlet on OA or to perform an interactive computer questionnaire. The questionnaire was designed to elicit treatment preferences based on workflows for 1) types of treatment (topical, intra-articular injection, pills, or exercise), 2) benefits (decreased pain, increased strength, and endurance), and 3) side effects (risk of dyspepsia, peptic ulcer). After performing the computer task, participants were given a handout illustrating their treatment preferences. Outcomes were collected immediately after the clinic visit using validated self-administered questionnaires included 1) decisional self-efficacy, 2) preparation to participate in decision making and 3) patient satisfaction. Because the distributions of outcomes scores were skewed, we present median values. We used the Wilcoxon 2-sample test to determine whether the observed differences were statistically significant.

Results: At the time of this writing, 76 patients have been recruited, (mean age 68 ± 10 years; 100% male; 74% Caucasian; 51% high school graduates; 53% married; overall health status fair or poor in 32%). There were no significant differences in demographic or clinical characteristics across groups at baseline. The median outcome scores by treatment group are presented in the table.

Median Outcome Measures by Treatment Group

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Pamphlet (N = 36)</th>
<th>Questionnaire (N = 40)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decisional Self-Efficacy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Possible range = 0 to 32)</td>
<td></td>
<td>28</td>
<td>231</td>
</tr>
<tr>
<td>Preparedness (Possible range = 0 to 45)</td>
<td>19</td>
<td>36</td>
<td>0.0001</td>
</tr>
<tr>
<td>Satisfaction (Possible range = 8 to 32)</td>
<td>19</td>
<td>24</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Conclusions: Results of this pilot study indicate that a decision tool that explicitly elicits patient treatment preferences for OA may improve surrogate measures for improved patient participation in decision making as well as patient satisfaction. The results of this pilot study support the need for further blinded trials.
ABSTRACTS

2057

ORAL CONCURRENT SESSION E - COST-EFFECTIVENESS ANALYSIS: METHODS
8:30 AM - 10:00 AM
Sunday, October 23, 2005
9:00 AM Grand Ballroom B

EVALUATING THE PERIODICITY OF DISEASE TRANSMISSION PRIOR TO COST-EFFECTIVENESS EVALUATION OF IMMUNIZATION PROGRAMS
Bauw E1, Trico A2, Rao A3, Gilsa V1, Duval B4, Pham B2 and Krahn M5
1University of Guelph, Guelph, ON, Canada; 2GlaxoSmithKline, Canada, Oakville, ON, Canada; 3University Laval, Beaucarreau, QC, Canada; 4University Health Network, Toronto, ON, Canada

Purpose: Mass childhood vaccination against Hepatitis A virus (HAV) in areas of high or intermediate endemicity has shown remarkable efficacy. In some low endemic areas, disease transmission has fallen dramatically in the absence of vaccination programs. It is unclear whether this is due to secular disease trends, periodic, variation, or both. Lack of ongoing disease transmission might suggest that evaluating cost-effectiveness of vaccination programs is unnecessary.

Methods: Monthly incidence of HAV-reported cases were obtained from the Canadian Disease Notification Reporting for 1980-1999. A power spectral analysis (PSA) of the monthly incidence time series was conducted to decompose the series into pure component frequencies and to determine their relative strengths. A dominant peak at frequency f corresponds to periodic outbreaks every 1/ f years. Less-dominant frequencies can also be identified.

An interspecific peak can also be predicted by a Susceptible(S)-Infected(I)-Recovered(R) dynamic model. An SIR model captures the time evolution of S0, I0, and R0 using coupled equations of time-dependent rates of change. The interspecific peak in an SIR model is estimated by $\pi = \frac{\beta}{\gamma}$ where $\beta$ and $\gamma$ are the mean age of infection and duration of infectiousness, respectively. A and D were derived using age-specific incidence data and a literature search.


The power spectrum analysis revealed a dominant peak at 0.15/year, corresponding to outbreaks every 7 years. This model predicted outbreaks every 7.5 years (i.e., A=25 years, Dec. weeks). The existence of this peak suggests that much HAV in Canada comes from transmission in the general population, beyond sporadic outbreaks in high-risk groups (1990-1992, 1995-1997).

The PSA also exhibited a spectral peak every year in December (i.e., a 26% increase from 112 to 142 cases for other months), suggesting a role for seasonality in HA transmission.

Conclusions: Results from the PSA indicate ongoing transmission in the general population, and the SIR model suggests that another incidence peak in the coming years is possible. Cost-effectiveness studies of the dynamic model of different primary prevention options are indicated. Evaluation of periodicity can elucidate transmission patterns and aid in planning cost-effectiveness studies by suggesting whether they should be performed, and if so, which strategies should be considered.

2445

ORAL CONCURRENT SESSION C - PUBLIC HEALTH
8:30 AM - 10:00 AM
Sunday, October 23, 2005
9:15 AM Grand Ballroom A

THE COST-EFFECTIVENESS OF CASE FINDING FOR HEPATITIS C IN FORMER INJECTING DRUG USERS
Stein K1, Thompson Coon J2, Castelnuovo E3, Pitt M2, Price A2, Sieber U3 and Cramp M4
1Peninsular Medical School, Exeter, United Kingdom; 2University of Southampton, Southampton, United Kingdom; 3Massachusetts General Hospital, Harvard Medical School, Boston, MA; 4University of Plymouth, Plymouth, United Kingdom

Purpose: To estimate the cost utility of active case finding for hepatitis C in former injecting drug users in the UK.

Methods: A decision analytic model (combined decision tree and Markov) was developed in Microsoft Excel® to investigate the impact of case finding and treatment on progression of HCV disease in a hypothetical cohort of former IDUs. This was compared to a cohort in whom spontaneous presentation for testing may occur. Case finding was explored in a “general case” and three specific settings: family, practice, prisons, and drug services. Diagnosis is based on ELISA and PCR testing and treatment is with pegylated interferon and ribavirin.

Best available data are used to inform the model, based on systematic searches of available literature, contact with researchers in the field, and clinical expert opinion. Progression of HCV disease is modeled as transitions between discrete health states: mild, moderate, or severe hepatitis; cirrhosis; and death.

Model uncertainty was explored using extensive one-way sensitivity analyses, threshold analyses, and probabilistic sensitivity analyses. A range of scenarios were explored using stochastic analyses. Value of information analysis was carried out to determine the value of further research.

Results: Little information is available on acceptance of testing and diagnosis, and rates of spontaneous presentation are difficult to estimate. Despite limited data, case finding is likely to be considered cost-effective. Cost utility was estimated at around $16,000 per QALY. At willingness to pay of $120,000 per QALY, case finding is around 70% likely to be cost-effective. Cost-effectiveness is more favorable if case finding is targeted at people with more advanced hepatitis—in people whose infection was acquired 20 years previously, case finding yields benefits at less than $15,000 per QALY. Results are most sensitive to changes in most parameters, although discounting has important effects given the long duration of infection.

Conclusions: Case finding for hepatitis C among former drug users is cost-effective. Further research into setting up, targeting, and delivering case finding programs is justified.

2102

ORAL CONCURRENT SESSION D - PATIENT AND PHYSICIAN DECISION MAKING
8:30 AM - 10:00 AM
Sunday, October 23, 2005
9:15 AM Grand Ballroom B

PATIENT AND SURGEON CORRELATES OF SHARED DECISION MAKING FOR SURGICAL BREAST CANCER TREATMENT
Hawley S1, Lanza P2, Salem B3, Fagerlin A4, Janz N5 and Katz S6
1University of Michigan, Ann Arbor VA Health System, Ann Arbor, MI; 2University of Michigan, School of Public Health, Ann Arbor, MI; 3University of Michigan, Ann Arbor, MI; 4University of Michigan, School of Public Health. Ann Arbor, MI

Background: The choice of surgical breast cancer treatment represents an opportunity for shared decision making (SDM), since both mastectomy and breast conserving surgery are viable options. Yet women vary in their desire for involvement in this decision.

Purpose: To evaluate patient and surgeon correlates of a shared decision for surgical breast cancer treatment, and the concordance between patients’ desired and actual roles in this decision.

Methods: Study subjects were breast cancer patients of Detroit and Los Angeles SEER registries mailed a questionnaire shortly after diagnosis in 2002 (N=1,260, SR: 77%). Data were merged with a survey of all surgeons (N=456, RR: 86%) for a dataset of 1,547 patients of 318 surgeons. SDM was measured using the Control Preferences Scale and categorized into 1) surgeon-based decision (with/without patient input); 2) shared decision; 3) patient-based decision (with/without surgeon input). The concordance between a woman’s self-reported actual and desired decisional involvement was categorized as more, less, or the right amount. SDM and involvement were dependent variables. The first set of independent variables included patient clinical and demographic and surgeon demographic factors. We then included two sets of surgeon “practice style” factors potentially related to SDM, referral propensity and surgeon reports of participating in SDM. Multinomial logistic regression was done, controlling for clustering within surgeons.

Results: From the PSA indicate ongoing transmission in the general population, the SIR model suggests that another incidence peak in the coming years is possible. Cost-effectiveness studies of the dynamic model of different primary prevention options are indicated. Evaluation of periodicity can elucidate transmission patterns and aid in planning cost-effectiveness studies by suggesting whether they should be performed, and if so, which strategies should be considered.

Conclusions: Results from the PSA indicate ongoing transmission in the general population, and the SIR model suggests that another incidence peak in the coming years is possible. Cost-effectiveness studies of the dynamic model of different primary prevention options are indicated. Evaluation of periodicity can elucidate transmission patterns and aid in planning cost-effectiveness studies by suggesting whether they should be performed, and if so, which strategies should be considered.

Downloaded from http://mdm.sagepub.com by guest on June 6, 2007
© 2006 Society for Medical Decision Making. All rights reserved. Not for commercial use or unauthorized distribution.
Patients to take responsibility for treatment choices. Disparate views on the appropriate role of agency and shared decision making, and reluctance by patients in various settings have been identified. A total of 202 patients were recruited (177 male; median age 65 years; 70 taking long-term medication). All main ethnic groups and social classes were represented. We have complete data for 197 patients. When informed of their actual coronary risk, 98 participants (50%) stated that they would choose treatment that reduced their coronary risk by 30%, assuming a different number of pre-treatment five-year coronary risks. Following the first interview, participants were told they would receive treatment that reduced their coronary risk by 30%, assuming a different number of pre-treatment five-year coronary risks. Following the first interview, and on the basis of known clinical details of the patient, five-year coronary risk was calculated for each patient. The second interview took place directly from the screening program; estimated costs of complications of infection (e.g., pelvic infection) and complications of non-treatment included. The expected value of perfect information was calculated for each patient. Conclusions: Universal prenatal syphilis screening using rapid point of care tests will improve both maternal and infant outcomes and is cost-effective.

ABSTRACTS

ORAL CONCURRENT SESSION D - PATIENT AND PHYSICIAN DECISION MAKING

8:30 AM - 10:00 AM
Sunday, October 23, 2005
9:30 AM Grand Ballroom B
THE MYTH OF AGENCY AND PATIENT CHOICE IN HEALTH CARE: THE CASE OF DRUG TREATMENTS TO PREVENT CORONARY DISEASE
University of Birmingham, Birmingham, United Kingdom
Purpose: To explore the functioning of the agency relationship in primary care in the UK, using the clinical example of the decision to commence drug treatment to prevent coronary disease. Methods: A well-functioning agency relationship will see the physician providing information to the patient on alternative treatment options so that the patient can make an informed choice. This study recruited individuals likely to be at high coronary risk from 13 UK primary care practices. People were invited for risk screening, and data on preferences to receive treatment were collected through two face-to-face interviews. At the first interview, participants were asked if they would choose treatment that reduced their coronary risk by 30%, assuming a number of different pre-treatment five-year coronary risks. Following the first interview, and on the basis of known clinical details of the patient, five-year coronary risk was calculated for each patient. The second interview took place directly from the screening program; estimated costs of complications of infection (e.g., pelvic infection) and complications of non-treatment included. The expected value of perfect information was calculated for each patient. Conclusions: Universal prenatal syphilis screening using rapid point of care tests will improve both maternal and infant outcomes and is cost-effective.

ORAL CONCURRENT SESSION C - PUBLIC HEALTH

8:30 AM - 10:00 AM
Sunday, October 23, 2005
9:30 AM Grand Ballroom A
TESTING THE IMPACT OF CENSORING ON HEALTH CARE DECISION-MAKING: RATE VERSUS RHYTHM CONTROL FOR ATRIAL FIBRILLATION
Fenwick E, Levy A, Marshall D, Blackhouse G, Sloan A and Shemanski L
University of York, York, United Kingdom; University of British Columbia, Vancouver, BC, Canada; St Josephs Healthcare, Hamilton, ON, Canada; PATH, Hamilton, ON, Canada; Axiom Research Corporation, Seattle, WA
Purpose: To examine the impact of the extent of censoring for trial-based economic evaluations on the decision regarding cost-effective provision and the value of further research.
Methods: Data were obtained from a large randomized trial called the Attrial Fibrillation Follow-up Investigation of Rhythm Management (AFFIRM) study. The AFFIRM investigators compared rhythm-control to rate-control for treatment of atrial fibrillation and reported a nonsignificant mean survival gain (0.08 years) in the rate-control arm. There was considerable censoring in the patient-level data, with only 16% of patients experiencing the primary end point (death) during follow-up. We compared the results of an economic evaluation accounting for the censoring to differing extents. In the absence of censoring, for a high-risk cohort of 1000 women with 6 pregnancies each, the expected cost-effectiveness plane for RT-EIA was below the willingness-to-pay threshold.
Results: When censoring was introduced, the cost-effectiveness plane for RT-EIA was below the willingness-to-pay threshold. The uncertainty surrounding the estimates of expected cost and survival duration was largest for the RT-EIA strategy. The estimated uncertainty surrounding the decision to adopt rate-control varied between 98% (full adjustment) and 100% (partial adjustment). For a patient population of 1.5 million over a period of 5 years, the expected cost-effectiveness gain in the rate-control arm was estimated at $234,000 (partial adjustment) and $44 million (full adjustment).
Conclusions: Cost-effectiveness analyses employing patient-level data should fully account for uncertainty in both costs and effects in order to appropriately measure the extent of the uncertainty surrounding the decisions regarding cost-effective provision and the worth of further research.

1953
ORAL CONCURRENT SESSION C - PUBLIC HEALTH

8:30 AM - 10:00 AM
Sunday, October 23, 2005
9:45 AM Grand Ballroom A
HIGH SCHOOL-BASED SCREENING FOR CHLAMYDIA IN PHILADELPHIA: IDENTIFICATION OF COST-SAVINGS USING A DYNAMIC TRANSMISSION MODEL
Fisman D, Spain C, Ashol L, Goldmb M, Salmon M and Newburn E
Drexel University, School of Public Health, Philadelphia, PA; Philadelphia Department of Public Health, Philadelphia, PA
Purpose: High prevalence of infection with Chlamydia trachomatis in Philadelphia teens led to the introduction of system-wide urine-based anonymous screening in the city’s public schools in 2002-03. Prevalence in teens has declined since screening was introduced. Existing assessments of cost-effectiveness of C. trachomatis screening disregard the transmissibility of this pathogen; we sought to assess cost-effectiveness of screening male and female students using a dynamic transmission model.
Methods: A dynamic transmission model was constructed using cross-sectional prevalence data, based on the simplifying assumption that the Philadelphia Chlamydia epidemic was at equilibrium when screening was initiated. The model was calibrated to reproduce the observed changes in prevalence of Chlamydia that followed the introduction of screening. Data on test costs were derived directly from the screening program, estimated costs of complications of infection (e.g., pelvic inflammatory disease, infertility, etc.) were abstracted from the published medical literature.
Results: A model based on non-assortive mixing and 20% reduction in infection in students identified through screening reproduced observed changes in population prevalence of Chlamydia that occurred after screening was introduced. Screening both males and females resulted in a greater reduction in pelvic inflammatory disease and related sequelae than screening either gender alone. All screening strategies (including screening-only males) resulted in net societal cost-savings relative to no screening, but the greatest savings were seen when both genders were screened. Screening via kiosk testing even when initial prevalence in females was as low as 4%, but the number of years required to “break even” financially increased as prevalence declined.
Conclusions: Using a modeling approach that accounts for transmissibility, cost-savings associated with Chlamydia screening are projected to increase when screening is extended to males. Furthermore, screening is projected to be cost-saving at relatively low prevalence thresholds when transmissibility is considered.
We used a Markov model to evaluate the cost-effectiveness of genetic screening on women from four populations: the general population; the Ashkenazi Jewish population; a “high-risk” general population (women with a first degree relative with either ovarian cancer or breast cancer before age 35); and a “high-risk” Ashkenazi Jewish population (women with a first degree relative with either ovarian cancer or breast cancer before age 40). Transition probabilities, test characteristics, and costs were estimated from the literature.

Results: Screening for BRCA1/2 mutations in the general population is prohibitively expensive regardless of a woman’s preference for prophylactic surgery. Screening women from the Ashkenazi Jewish or “high-risk” general populations (BRCA1/2 prevalence = 2.5% to 4%) has an incremental cost-effectiveness ratio (ICER) <$100,000/QALY if the woman’s preferences for both surgeries are high and if BRCA1/2 women undergo prophylactic oophorectomies and BRCA2+ women undergo prophylactic mastectomy. If a woman’s preference for one surgery is high and the other surgery low, then screening for BRCA1/2 has an ICER <$100,000/QALY if women who test positive for either mutation undergo the preferred prophylactic surgery. If preferences for both prophylactic surgeries are low, screening is prohibitively expensive. Screening “high-risk” Ashkenazi Jewish women (BRCA1/2 prevalence = 10%) has an ICER <$100,000/QALY if preferences for both surgeries are high and BRCA1/2 women undergo both prophylactic surgeries and BRCA2+ women undergo prophylactic oophorectomies. If preference for one surgery is high and for the other surgery low, the results are similar to the average-risk Ashkenazi Jewish population. If preferences for both surgeries are low, screening still has an ICER <$100,000/QALY if women who test BRCA2+/2 undergo intensive surveillance.

Conclusions: The benefits and cost-effectiveness of screening for BRCA1/2 mutations vary greatly depending on a woman’s utilities for prophylactic surgeries. Current guidelines recommend screening primarily based on the prior probability of BRCA mutations and should be revised to reflect the importance of women’s utilities.
ABSTRACTS

2036
ORAL CONCURRENT SESSION C - PREFERENCE METHODS
10:30 AM - 12:00 PM
Sunday, October 23, 2005
10:30 AM Grand Ballroom B
RESPONSE SHIFT IN SELF-REPORT OF HEALTH STATUS
Sawalha Almashry M, Cox V, Comer-Spaddy B and Jones C*
1 Baylor College of Medicine/Michael E. DeBakey VA Medical Center, Houston, TX;
2 Baylor College of Medicine, Houston, TX; University of Calgary, Calgary, AB, Canada;
*University of Alberta, Edmonton, AB, Canada

Purpose: Patient-reported outcomes are increasingly being used to evaluate the effectiveness of interventions. Often, patients are asked to assess their current health in relation to their past health. Yet, if their perceptions have shifted, the assessment of improvement may be biased. Our objective was to evaluate response shift in arthritis patients’ self-report of health status 3 years after undergoing total joint replacement (TJR).

Methods: Patients who had a shift of at least 10 points, and in two-thirds of these the shift was negative (thought health was worse than had been reported). No significant associations were found between response capacity beds and surge treatment capacity in order to clarify this preparedness goal. The mathematical (game theory) tools required to capture the interaction between group behavior and individual behavior have been successfully applied in fields such as economics, ecology, and evolution but are almost entirely undeveloped in public health and vaccination policy in particular.

Conclusions: Using the latest published version of the BRAM, and a threshold ICER of £30,000/QALY ($55,000/QALY), the optimal sequence in the base case analysis was found to be sulfasalazine, methotrexate, hydroxychloroquine, leflunomide, injectable gold, penicillamine, ciclosporin, azathioprine. Different sequences were optimal at other threshold ICERs and under other modeling assumptions.

Conclusions: This article shows how an optimization algorithm from operations research can be applied to a complex model for the management of a chronic condition. It also shows that deterministic sensitivity analysis can be applied and that probabilistic sensitivity analysis can be used to estimate the expected outcome allowing for uncertainty in the model parameters.

ABSTRACTS

2179
ORAL CONCURRENT SESSION H - METHODOLOGICAL ADVANCES
10:30 AM - 12:00 PM
Sunday, October 23, 2005
10:45 AM Grand Ballroom C
A MODEL FOR PREDICTING TIME EVOLUTION OF VACCINE UPTAKE AFTER A VACCINE SCARE
Bauch C*
University of Guelph, Guelph, ON, Canada

Purpose: The literature on vaccine scares and vaccine risk perception has focused on the individual level; there is a group context to individual vaccinating behavior that should not be ignored: individuals are influenced by the opinions of other members of the population, and the vaccinating behavior of others determines the rates of vaccine uptake and hence the risk that an individual becomes infected. The mathematical (game theory) tools required to capture the interaction between group behavior and individual behavior have been successfully applied in fields such as economics, ecology, and evolution but are almost entirely undeveloped in public health and vaccination policy in particular.

Conclusions: We used the latest published version of the BRAM, and a threshold ICER of £30,000/QALY ($55,000/QALY), the optimal sequence in the base case analysis was found to be sulfasalazine, methotrexate, hydroxychloroquine, leflunomide, injectable gold, penicillamine, ciclosporin, azathioprine. Different sequences were optimal at other threshold ICERs and under other modeling assumptions.

Conclusions: This article shows how an optimization algorithm from operations research can be applied to a complex model for the management of a chronic condition. It also shows that deterministic sensitivity analysis can be applied and that probabilistic sensitivity analysis can be used to estimate the expected outcome allowing for uncertainty in the model parameters.

ABSTRACTS

© 2006 Society for Medical Decision Making. All rights reserved. Not for commercial use or unauthorized distribution.
ABSTRACTS

ORAL CONCURRENT SESSION G - PREFERENCE METHODS
10:30 AM - 12:00 PM
Sunday, October 23, 2005
10:45 AM Grand Ballroom B
GIVE ME COLOSTOMY OR GIVE ME DEATH: ATTEMPTING TO IMPROVE INTERNAL CONSISTENCY OF TREATMENT CHOICE
Kulesa J, Zikmund-Fisher B, Fagerlin A and Ubel P
University of Michigan, Ann Arbor, MI; VA Ann Arbor Healthcare System, Ann Arbor, MI
Purpose: In past research that presented treatments for colon cancer, most people rated death as worse than any possible complication after surgery (specifically: colostomy, chronic diarrhea, intermittent bowel obstruction, wound infection); however, people then go against these ratings by choosing a treatment that puts them at greater risk of death in order to avoid the chance of complications associated with other treatment options. We tested several interventions designed to counter this inconsistency by helping people to recognize that even if they got a complication, they would still be alive.
Method: We recruited subjects from a demographically balanced panel to participate in an online survey experiment. Participants were asked to imagine they have colon cancer and that there are two surgeries available to treat their illness. Surgery 1 has a lower chance of death with some chance of complications, whereas Surgery 2 has a higher chance of death with no chance of complications. In addition to a no-intervention version, we tested two versions designed to focus people’s attention on the fact that even if they got complications, they would still be alive. One version included a pictograph which used color and graphics to represent the various outcomes of treatment. The other version asked subjects to rate how good each possible outcome would be. We tested whether these interventions would reduce the proportion of individuals choosing Surgery 2.
Results: 1284 people completed the online survey. In the no-intervention version, 65.6% chose the surgery with the higher chance of death. Of the two interventions, only the pictographs had a significant effect, reducing the number who chose Surgery 2 to 55.3% (p = 0.002). Nonetheless, the majority of people in this version still chose the surgery with the higher chance of death instead of the surgery that could potentially cause complications.
Conclusions: Neither of our interventions came close to eliminating the inconsistency, suggesting this inconsistency is extremely resilient. Our interventions were both based on a common theme: helping individuals to be consistent with their value judgments (i.e., belief that complications are better than death) and their treatment choice. Because, at the very best, pictographs reduced the inconsistency by 10%, we have only limited evidence that this theme underlies the inconsistency.

ORAL CONCURRENT SESSION H - METHODOLOGICAL ADVANCES
10:30 AM - 12:00 PM
Sunday, October 23, 2005
11:00 AM Boardroom C
ROC ANALYSIS: A THEOREM REJECTED
Lukin G, Lord J and Fischer A
University of Manchester, Manchester, United Kingdom; National Institute for Health and Clinical Excellence, London, United Kingdom
Purpose: In receiver-operator curve (ROC) analysis, it has been said that “it is a theorem that if one ROC overlays another, that test improves the cost-effectiveness of treatment under all circumstances.” We sought to refute this.
Method: We employed a binomial model of diagnostic testing to a “ternaternal” model, presuming the existence of four clinically relevant subgroups having prevalences and economic characteristics as follows:

<table>
<thead>
<tr>
<th>Group</th>
<th>Prevalence</th>
<th>Treatment A, QALYs</th>
<th>Treatment B, QALYs</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>9%</td>
<td>21%</td>
<td>49%</td>
</tr>
<tr>
<td>G2</td>
<td>21%</td>
<td>49%</td>
<td>21%</td>
</tr>
<tr>
<td>G3</td>
<td>49%</td>
<td>21%</td>
<td>9%</td>
</tr>
<tr>
<td>G4</td>
<td>9%</td>
<td>21%</td>
<td>49%</td>
</tr>
</tbody>
</table>

We postulate a “gold standard” test that classifies G1 and G3 as belonging to disease X, and G2 and G4 as belonging to disease Y. This reflects the outlook with Treatment A. We now evaluate two further diagnostic tests, D1 and D2, in relation to the gold standard. The ternaternal distributions of diagnostic signal for D1 and D2 are characterized as

<table>
<thead>
<tr>
<th>ROC Curve</th>
<th>D1</th>
<th>Mean</th>
<th>SD</th>
<th>D2</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROC A</td>
<td></td>
<td>0.40</td>
<td>0.50</td>
<td></td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>ROC B</td>
<td></td>
<td>0.60</td>
<td>0.60</td>
<td></td>
<td>0.50</td>
<td>0.60</td>
</tr>
<tr>
<td>ROC C</td>
<td></td>
<td>0.50</td>
<td>0.50</td>
<td></td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>ROC D</td>
<td></td>
<td>0.60</td>
<td>0.60</td>
<td></td>
<td>0.50</td>
<td>0.60</td>
</tr>
</tbody>
</table>

As well as plotting ROC curves for D1 and D2, we plot a novel curve, “ROTS,” that tracks the economic consequences of expansion of access to Treatment B at increasingly lenient diagnostic test thresholds. ROTS is plotted in cost-effectiveness (CE) space.

Conclusions: The frame of reference for ROC analysis is determined by a preexisting diagnostic gold standard. This frame is imperceptibly calibrated for the patient decision at hand. The optimal allocation in this case is for Treatment A to G1 and G2 and Treatment B to G3 and G4. Tests D1 and D2 add additional diagnostic information independently of their ability to predict gold standard status. This is revealed by the curve ROTS, which plots the actual economic consequences of changing test thresholds. This result suggests that ROTS analysis with CE-space as the reference frame may be superior to the ROC for the evaluation of new diagnostic technology.

ORAL CONCURRENT SESSION F - SIMULATION
10:30 AM - 12:00 PM
Sunday, October 23, 2005
11:00 AM Grand Ballroom A
QUANTIFYING UNCERTAINTY OF PREDICTED HPV VACCINE EFFECTIVENESS
Briston M, Van de Velde N and Bolly M
1 Merck Frost Canada ltd, Montreal, QC, Canada; 2 Université Laval, Montreal, QC, Canada; Imperial College, London, United Kingdom
Purpose: Clinical trials have shown prophylactic HPV vaccines to be effective against persistent HPV infection, cervical intraepithelial neoplasia (CIN), and genital warts. To aid policy decisions concerning HPV vaccination, model-based estimates of population-level effectiveness and cost-effectiveness are needed. The process of model development requires choices and assumptions, which introduce uncertainty. Uncertainty should be quantified to provide policy makers with the necessary information to make appropriate decisions. The aim of this study is to examine the impact of model and parameter assumptions on the effectiveness of HPV vaccination.
Methods: A cohort model of the natural history of HPV infection (types 16, 18, and 6, and other low risk and high risk types) was developed to investigate the effectiveness of HPV vaccination at preventing infection, CIN, cervical cancer, and genital warts in Canada. Natural history parameter sets, which fit simultaneously North American longitudinal and cross-sectional data (incidence and prevalence of HPV infection, CIN, and cervical cancer), were identified as follows: 1) Uniform prior distributions were defined for each parameter using minimum and maximum values found in the literature; 2) 200,000 sets of different parameter combinations were selected from prior distributions using Latin Hypercube Sampling, and 3) GOF techniques were used to identify “best fit” parameter sets. An investigation with multivariate analyses, “Test PI,” parameter sets were used to assess the sensitivity of HPV vaccine effectiveness to plausible model and parameter assumptions. Results: The proportion of individuals protected following immunization (take) is 100%, reduction in susceptibility to HPV-16/18/6/11 infection (degree of protection) is 95% and duration of protection is lifelong, the model estimates of the type-specific vaccine effectiveness against infection, CIN and cervical cancer range from 40% to 95%. The effectiveness of vaccination is most sensitive to vaccine efficacy (take, degree, and duration), level of immunity following natural infection, and the age-specific progression and regression rates.
Conclusions: Preliminary model results suggest that vaccination against HPV 16/18/6/11 has the potential to reduce HPV infection and disease. Our modeling framework has the advantage of integrating multivariate sensitivity analysis on all the natural history and vaccine parameters when assessing the effectiveness and cost-effectiveness of vaccination. Hence, the model can provide a comprehensive, flexible, and robust framework for decisions regarding implementation of new screening and HPV vaccination programs.
ABSTRACTS

ORAL CONCURRENT SESSION H - METHODOLOGICAL ADVANCES
10:30 AM - 12:00 PM
Sunday, October 23, 2005
11:15 AM Boardroom C
DO PROPENSITY SCORE METHODS OVERCOME BIAS IN ESTIMATING AVERAGE TREATMENT EFFECTS IN OBSERVATIONAL STUDIES?

Jane Z., Gardiner J., Bradley C. and Given C.
Michigan State University, East Lansing, MI; Virginia Commonwealth University, Richmond, VA
Purpose: To compare and contrast the properties of different estimators of the average treatment effect in observational studies based on propensity score methods to that estimated in randomized controlled trials (RCT).
Background: Randomization of patients to treatment or control groups is assumed to balance the groups on both observed and unobserved patient characteristics. In non-randomized studies, patients self-select themselves to treatments. Naïve estimators of the treatment effect are subject to selection bias. Propensity score methods have been used as a means to mitigate the effect of selection bias.
Methods: We use data from two RCTs (n = 237 and n = 124) of a cognitive behavioral intervention designed to reduce the severity of symptoms in breast cancer patients. The heterogeneous treatment effects are estimated using the physical function and mental health function assessed by the SF-36. We have access to a prospective longitudinal study of a cohort of cancer patients who were undergoing chemotherapy but not cognitive behavioral intervention (n = 86). We then construct composite samples of patients from the two RCTs and comparable non-treated and non-randomized patients from the prospective study. The treatment effect is estimated using stratification, nearest neighbor matching, k-nearest neighbor matching, and bias-corrected k-nearest neighbor matching. Heteroscedasticity-robust standard errors are obtained for each estimator. Our propensity score estimators are then compared to the benchmarks from the RCTs.
Results: In comparison with the benchmark effects from the RCTs, the propensity score methods produced estimates that varied widely with the choice of comparison samples and outcomes. Agreement was closer with physical function than with mental health function. Bias was greater when the comparison sample differed from the treatment group on several patient characteristics. No simple propensity score estimator dominated the others in estimating the benchmark effects. However, the bias-corrected k-nearest neighbor matching method yielded standard errors close to those from the RCTs.
Conclusions: The choice of propensity score matching technique affected the accuracy of average treatment effect estimation. Strikingly, the effect on estimation of different composite samples we constructed was pronounced. There was closer agreement with the benchmark effect when the comparison samples were similar to the treated group. Our study suggests that propensity score methods do not always yield accurate estimates of average treatment effects. They must be used with caution in observational studies.

ORAL CONCURRENT SESSION C - PREFERENCE METHODS
10:30 AM - 12:00 PM
Sunday, October 23, 2005
11:15 AM Grand Ballroom A
A COMPUTER-GENERATED SEER DATABASE FOR THE BREAST CANCER POPULATION

Plewa R., Sigaal R., Salzman P., Gynn I. and Rosenberg J.
Stanford University, Stanford, CA; University of Rochester, Stanford, CA
Purpose: The Surveillance, Epidemiology and End Results (SEER) contains information on breast cancer patients’ demographic and tumor characteristics at the time of diagnosis and their vital status. Because it does contain information on the screening history or even mode of cancer detection and underreports the use of adjuvant therapy, the SEER limits our ability to quantify the impact of cancer control programs. Through a collaboration with Cancer Intervention and Surveillance Network (CISNET), we have developed a simulation model that brings together data from numerous national registries to develop a virtual SEER database that captures information on the screening history, model of detection, tumor characteristic, treatment, and survivorship on the individual patient level. We propose to present the underlying model that generates the virtual database and methods we used to validate it.
Methods: We developed a Monte Carlo simulation model that reproduces the life history of women in the United States who were born since 1890 and outputs population-level statistics from 1975 onward. The simulation records the following information for each breast cancer patient: her date of birth, her screening schedule, when and how she was detected, her tumor size and stage at detection, ER status at detection, what treatment she received, and her survival time from diagnosis and her cause of death. The model is informed with data from pre-screening era of SEER, the Patterns of Care Registry, and the Breast Cancer Surveillance Consortium.
Results: Our virtual breast tumor registry closely reproduces population-level trends observed in the SEER database that were not used in the model-building process. The model reproduces the tumor size distribution as a function of calendar year for individual age groups. It reproduces the proportion of tumor stage (local, regional, distant) by calendar year and age group. It demonstrates a good agreement between the model and data in terms of overall age-adjusted incidence and mortality.
Conclusion: We have generated and validated a virtual tumor registry of the US breast cancer population that merges existing national databases.

ORAL CONCURRENT SESSION G - SIMULATION
10:30 AM - 12:00 PM
Sunday, October 23, 2005
11:15 AM Grand Ballroom B
COLORRECTAL CANCER SCREENING: CONJOINT ANALYSIS OF CONSUMER PREFERENCES AND PHYSICIANS’ PERCEIVED CONSUMER PREFERENCES IN THE US AND CANADA

St. Josephs Healthcare & McMaster University, Hamilton, ON, Canada; University of California-San Francisco, SF, CA; Research Triangle Institute, Research Triangle Park, NC; McMaster University; Hamilton, ON, Canada; University of California, San Francisco, San Francisco, CA; St. Josephs Healthcare, Hamilton, ON, Canada
Purpose: To measure US and Canadian (CDN) consumer preferences for colorectal cancer (CRC) screening using conjoint analysis (CA), and compare with primary care physicians’ perceptions of those preferences.
Methods: CRC screening reduces mortality but uptake remains poor in the US and Canada. Nine attributes of CRC screening (process, pain, frequency, preparation, follow-up, specificity, sensitivity, complication rate, and cost) were identified based on focus groups and results from a previous CA survey in Canada. The fractional factorial CA survey design used to maximize D-Effectiveness included three blocks with 11 choice sets and one repeat scenario. A follow-up question asked if respondents would prefer no CRC screening at all. Survey invitations were emailed to a random sample of US and CDN consumers (aged 45-70 years) and US MDs who were online panel members. CDN MDs were posted an invitation to complete the survey online. B-coefficients from bivariate probit regression estimated the marginal utilities of the attributes including the no screening option and were scaled from 0 to 10.
Results: Response rates ranged from 7.5% (CDN MDs) to 47.9% (CDN consumers).
Order of attribute importance (rescaled b-coefficient).

US Consumers
n = 1067
CDN Consumers
n = 501
US MDs
n = 500
CDN MDs
n = 500
sensitivity (8.19) process (6.90) process (6.61) specificity (3.33) specificity (3.33) process (4.46) process (4.46)
no screening (2.09) sensitivity (5.32) no screening (6.48) frequency (1.32) frequency (1.77)
specificity (3.15) frequency (1.98) frequency (1.49) specificity (2.03) specificity (2.03)
preparation (0.98) no screening (2.09) preparation (1.63) frequency (1.44) preparation (0.42)
frequency (1.81) no screening (1.19) no screening (0.46) frequency (1.16) frequency (0.29)
follow-up (8.29) follow-up (8.29) follow-up (0.18) follow-up (0.18) follow-up (0.29)

The top two attributes were the same for US and CDN consumers (sensitivity and process) and US and CDN MDs (no screening and specificity). MDs overestimated consumer preferences for no screening and underestimated preferences for test process (e.g., stool sample vs. having an endo- copy). Both consumers and MDs assigned low preference to the need for confirmatory follow-up tests.

Conclusions: These results indicate how preferences of consumers differ from perceptions of providers about consumer preferences. This can be used to inform how CRC screening programs are presented to consumers to help increase screening uptake.

ORAL CONCURRENT SESSION F - SIMULATION
10:30 AM - 12:00 PM
Sunday, October 23, 2005
11:15 AM Grand Ballroom A
A COMPUTER-GENERATED SEER DATABASE FOR THE BREAST CANCER POPULATION

Plewa R., Sigaal R., Salzman P., Gynn I. and Rosenberg J.
Stanford University, Stanford, CA; University of Rochester, Stanford, CA
Purpose: The Surveillance, Epidemiology and End Results (SEER) contains information on breast cancer patients’ demographic and tumor characteristics at the time of diagnosis and their vital status. Because it does contain information on the screening history or even mode of cancer detection and underreports the use of adjuvant therapy, the SEER limits our ability to quantify the impact of cancer control programs. Through a collaboration with Cancer Intervention and Surveillance Network (CISNET), we have developed a simulation model that brings together data from numerous national registries to develop a virtual SEER database that captures information on the screening history, model of detection, tumor characteristic, treatment, and survivorship on the individual patient level. We propose to present the underlying model that generates the virtual database and methods we used to validate it.
Methods: We developed a Monte Carlo simulation model that reproduces the life history of women in the United States who were born since 1890 and outputs population-level statistics from 1975 onward. The simulation records the following information for each breast cancer patient: her date of birth, her screening schedule, when and how she was detected, her tumor size and stage at detection, ER status at detection, what treatment she received, and her survival time from diagnosis and her cause of death. The model is informed with data from pre-screening era of SEER, the Patterns of Care Registry, and the Breast Cancer Surveillance Consortium.
Results: Our virtual breast tumor registry closely reproduces population-level trends observed in the SEER database that were not used in the model-building process. The model reproduces the tumor size distribution as a function of calendar year for individual age groups. It reproduces the proportion of tumor stage (local, regional, distant) by calendar year and age group. It demonstrates a good agreement between the model and data in terms of overall age-adjusted incidence and mortality.
Conclusion: We have generated and validated a virtual tumor registry of the US breast cancer population that merges existing national databases.

© 2006 Society for Medical Decision Making. All rights reserved. Not for commercial use or unauthorized distribution.
The model), and external (predictive) validity.

Methods: A PubMed search was used to identify obstructive lung disease simulation models published from 2000 to 2005. Publications describing the models were reviewed to assess four types of model validation: first-order internal validity (verification/debugging), second-order internal validity (comparison of model estimates against data from studies used to construct the model), third-order internal validity (comparison of model estimates against published data not used to construct the model), and external (predictive) validity.

Results: Eight disease simulation models were identified, four for asthma and four for chronic obstructive pulmonary disease (COPD). Most were from Europe (75%) and funded by pharmaceutical companies (88%). Seven of eight were Markov models. Two models were developed for specific cost-utility analyses (CUA); these for use in future CUA, and one for exploratory analysis of future COPD burden. Seven of eight models included no mention of first-order internal validation; one did so implicitly. Five of eight models included tests of second-order internal validity. All were simulations of source study cohorts comparing modeled with observed outcomes. Disease exacerbation was the most common end point. Only one model included a statistical test of modeled versus observed outcomes. Two of eight models included tests of third-order internal validity. In both, modeled point estimates of COPD prevalence or exacerbation were compared to point estimates from published studies. No models included tests of predictive validity or plans for predictive validation. Validation was typically described as acceptable in qualitative terms, despite near-universal absence of criteria for judging adequacy of validation.

Conclusion: Most recent obstructive lung disease simulation models lack appropriate validation. For simulation modeling to be accepted as a tool for evaluating the impact of public health programs, models must be validated to the extent necessary to credibly simulate health outcomes in interest. Defining the requisite level of validation in diverse decision contexts is an important next step in promoting simulation models as practical decision tools.

ORAL CONCURRENT SESSION F - SIMULATION
10:30 AM - 12:00 PM
Sunday, October 23, 2005
11:30 AM Grand Ballroom A
VALIDATION OF OBSTRUCTIVE LUNG DISEASE SIMULATION MODELS - IS "IN-LINE WITH PUBLISHED DATA" THE STATE OF THE ART?

Ferdinands J
Centers for Disease Control and Prevention, Atlanta, GA

Background: Increasing interest in using disease simulation models to evaluate public health programs has led to closer scrutiny of model methods and rising demand for evidence that models accomplish their intended goals.

Methods: The purpose of this study was to identify obstructive lung disease simulation models from the recent literature and evaluate validation of those models.

Results: A PubMed search was used to identify obstructive lung disease simulation models published from 2000 to 2005. Publications describing the models were reviewed to assess four types of model validation: first-order internal validity (verification/debugging), second-order internal validity (comparison of model estimates against data from studies used to construct the model), third-order internal validity (comparison of model estimates against published data not used to construct the model), and external (predictive) validity.

Results: Eight disease simulation models were identified, four for asthma and four for chronic obstructive pulmonary disease (COPD). Most were from Europe (75%) and funded by pharmaceutical companies (88%). Seven of eight were Markov models. Four models were developed for specific cost-utility analyses (CUA); three for use in future CUA, and one for exploratory analysis of future COPD burden. Seven of eight models included no mention of first-order internal validation; one did so implicitly. Five of eight models included tests of second-order internal validity. All were simulations of source study cohorts comparing modeled with observed outcomes. Disease exacerbation was the most common end point. Only one model included a statistical test of modeled versus observed outcomes. Two of eight models included tests of third-order internal validity. In both, modeled point estimates of COPD prevalence or exacerbation were compared to point estimates from published studies. No models included tests of predictive validity or plans for predictive validation. Validation was typically described as acceptable in qualitative terms, despite near-universal absence of criteria for judging adequacy of validation.

Conclusion: Most recent obstructive lung disease simulation models lack appropriate validation. For simulation modeling to be accepted as a tool for evaluating the impact of public health programs, models must be validated to the extent necessary to credibly simulate health outcomes in interest. Defining the requisite level of validation in diverse decision contexts is an important next step in promoting simulation models as practical decision tools.

ORAL CONCURRENT SESSION H - METHODOLOGICAL ADVANCES
10:30 AM - 12:00 PM
Sunday, October 23, 2005
11:30 AM Boardroom C
PROBABILISTIC SENSITIVITY ANALYSES USING LOGISTIC REGRESSION INPUTS TO DECISION MODELS
Schraub D, Leonard A, Hornung R and Eckman M
University of Cincinnati, Cincinnati, OH

Purpose: Patient-level predictions from population-based predictive models are subject to uncertainty. The impact of this uncertainty can be described using probabilistic sensitivity analyses (PSA). The logistic: normal distribution is frequently used to describe probability distributions in second-order Monte Carlo simulations. However, its standard error (SE) depends upon whether the upper or lower 95% confidence limit of the probability is used in the transform, as the confidence interval (CI) is not symmetric around the mean. We describe a technique that avoids this transformation when using the output of logistic regressions as inputs to decision models.

Methods: We created a two-stage decision support tool for severe sepsis that guides the use of drotrecogin alfa (activated). First, logistic regression models were developed to calculate patient specific mortality with and without treatment. These were then used as inputs to a 75-stage Markov model. Patient-specific predictors included age, gender, and twelve readily available clinical characteristics. We calculated the logit and SE for 28-day mortality for hypothetical patients with specific risk profiles, either receiving or not receiving treatment. We also calculated the resultant probability of death with their 95% confidence intervals. For each patient, we performed PSAs in two ways: using a log normal distribution to describe the logit and its SE, and also using the approach described by Dohlet et al. creating a logistic normal distribution from the probabilities and their upper and lower 95% CIs.

Results: For a series of hypothetical patients, SIs in the distributions of the differences in expected utility between the two strategies were smallest when using the lower 95% CIs for probabilities of death less than 50% and were the largest for probabilities of death greater than 50%. Our method using log normal distributions based on the logit of the regression model yielded intermediate SIs.

Conclusions: Uncertainty in the results of PSAs may be under- or overestimated when using logistic normal distributions to describe the probabilistic inputs to decision models. Direct use of the logit and its SE through a log normal transformation simplifies calculations and eliminates these errors when performing PSAs with decision models that incorporate probabilistic logistic regressions.

ORAL CONCURRENT SESSION G - PREFERENCE METHODS
10:30 AM - 12:00 PM
Sunday, October 23, 2005
11:45 AM Boardroom C
COMPUTATIONAL IMPROVEMENT OF INTEGER PROGRAMMING MODELS FOR DETERMINING THE OPTIMAL CONFIGURATION OF REGIONS IN THE US ORGAN TRANSPLANTATION AND ALLOCATION NETWORK
Kong N, Schaefer A, Harskawd R and Roberts M
University of South Florida, Tampa, FL; University of Pittsburgh, Pittsburgh, PA

Purpose: Currently, 59 Organ Procurement Organizations (OPOs) are aggregated to 11 regions in the US organ transplantation and allocation network: allocation rules include a regional preference. Given this hierarchy, we seek to assess whether the current regional configuration optimizes transplant efficiency, that is, meets a set of criteria such as maximizing transplant cardinality and minimizing organ wastage. A prior analytic framework proposed by Stahl revealed that the current configuration is suboptimal, however, it was computationally intractable to fully optimize over all possible regional configurations. We have adapted a solution method that allows models of sufficient size and complexity to be solved, and have added additional parameters to enhance clinical realism.

Method: Integer programming (IP) is an industrial engineering technique that deals with problems making discrete choices from a potentially large set of possibilities: in this case, choosing the optimal regional configuration from a nearly infinite set of possibilities. Our analysis was conducted from the societal perspective, and the allocation system was assumed to be steady. We modified an analytic proxy of the total number of intraregional transplants to capture the effect of organ wastage. A priori we had chosen the current configuration which was suboptimal, however, it was computationally intractable to fully optimize over all possible regional configurations. We have adapted a solution method that allows models of sufficient size and complexity to be solved, and have added additional parameters to enhance clinical realism.

Conclusion: Integer programming (IP) is an industrial engineering technique that deals with problems making discrete choices from a potentially large set of possibilities: in this case, choosing the optimal regional configuration from a nearly infinite set of possibilities. Our analysis was conducted from the societal perspective, and the allocation system was assumed to be steady. We modified an analytic proxy of the total number of intraregional transplants to capture the effect of organ wastage. A priori we had chosen the current configuration which was suboptimal, however, it was computationally intractable to fully optimize over all possible regional configurations. We have adapted a solution method that allows models of sufficient size and complexity to be solved, and have added additional parameters to enhance clinical realism.

Method: Integer programming (IP) is an industrial engineering technique that deals with problems making discrete choices from a potentially large set of possibilities: in this case, choosing the optimal regional configuration from a nearly infinite set of possibilities. Our analysis was conducted from the societal perspective, and the allocation system was assumed to be steady. We modified an analytic proxy of the total number of intraregional transplants to capture the effect of organ wastage. A priori we had chosen the current configuration which was suboptimal, however, it was computationally intractable to fully optimize over all possible regional configurations. We have adapted a solution method that allows models of sufficient size and complexity to be solved, and have added additional parameters to enhance clinical realism.

Conclusion: Integer programming (IP) is an industrial engineering technique that deals with problems making discrete choices from a potentially large set of possibilities: in this case, choosing the optimal regional configuration from a nearly infinite set of possibilities. Our analysis was conducted from the societal perspective, and the allocation system was assumed to be steady. We modified an analytic proxy of the total number of intraregional transplants to capture the effect of organ wastage. A priori we had chosen the current configuration which was suboptimal, however, it was computationally intractable to fully optimize over all possible regional configurations. We have adapted a solution method that allows models of sufficient size and complexity to be solved, and have added additional parameters to enhance clinical realism.

Method: Integer programming (IP) is an industrial engineering technique that deals with problems making discrete choices from a potentially large set of possibilities: in this case, choosing the optimal regional configuration from a nearly infinite set of possibilities. Our analysis was conducted from the societal perspective, and the allocation system was assumed to be steady. We modified an analytic proxy of the total number of intraregional transplants to capture the effect of organ wastage. A priori we had chosen the current configuration which was suboptimal, however, it was computationally intractable to fully optimize over all possible regional configurations. We have adapted a solution method that allows models of sufficient size and complexity to be solved, and have added additional parameters to enhance clinical realism.
from the pre-screening era in the late 1970s, we systematically varied US all-cause mortality according to the Surveillance, Epidemiology, and End-Results Program. Using CRC incidence and mortality data, we estimated the increase in CRC cases as individuals are at risk for longer. We examined the extent to which improvements in life expectancy in the US affect CRC incidence and mortality.

Methods: We developed a Markov model to predict lifetime risk of CRC incidence and mortality. Age-specific probabilities of CRC diagnosis and cancer-specific mortality by stage were obtained from the Surveillance, Epidemiology, and End-Results Program. Using CRC incidence and mortality from the pre-screening era in the late 1970s, we systematically varied US all-cause mortality according to the Surveillance, Epidemiology, and End-Results Program. Using CRC incidence and mortality from the pre-screening era in the late 1970s, we systematically varied US all-cause mortality according to the Surveillance, Epidemiology, and End-Results Program. Using CRC incidence and mortality from the pre-screening era in the late 1970s, we systematically varied US all-cause mortality according to the Surveillance, Epidemiology, and End-Results Program.

Results: Excess life expectancy for 20-year-old US white males increased from 49.8 years in 1960 to 55.3 years in 2000. Holding all other factors constant at pre-screening levels, we predicted this increase in life expectancy over the 40-year period would result in a 44% increase in the lifetime risk of CRC incidence (from 4.5% to 6.5%) and a 50% increase in the lifetime risk of CRC mortality (from 2.5% to 3.7%). In general, for each one-year increase in remaining life expectancy for white males, the lifetime risk of CRC incidence increased by 0.004. In comparison to white males, life expectancy for black males was approximately 7 years less at both time periods. Controlling for differences in underlying cancer incidence and mortality, the lower life expectancy experienced by black males was associated with a 25% reduction in lifetime risks of CRC incidence and mortality compared with white males. In contrast, when we compared pre-screening incidence with the lower current risk levels to attributes, we found approximately 17% reductions in lifetime CRC incidence and mortality.

Conclusions: The potential effects of increased screening and improved treatment over time on observed CRC incidence and mortality are counterbalanced by increases in life expectancy. Because age-standardization of cancer statistics obfuscates the impact of increasing life expectancy, quantification of the relationship between life expectancy and colorectal cancer provides insights into the growing burden of disease.

ABSTRACTS

ORAL CONCURRENT SESSION J - MEASUREMENT OF HEALTH STATUS AND PREFERENCES

10:00 AM - 12:00 PM

Sunday, October 23, 2005

1:30 PM - 3:00 PM

Sunday, October 23, 2005

BAYesian APPROACH TO ASSESS PREFERENCES AMONG LOW LITERACY PATIENTS: AN EXAMPLE USING COLORECTAL CANCER SCREENING

Hawley S 1, Bhaia-Weiss M 1, Volk R 2 and Vernon S 3

1University of Michigan, Ann Arbor VA Health System, Ann Arbor, MI; 2Baylor College of Medicine, Houston, TX; 3University of Texas, School of Public Health, Houston, TX.

Purpose: Existing methods of preference assessment are limited by the ability of certain populations to comprehend the required techniques, such as time tradeoffs and standard gambles. Conjoint analysis (CA) facilitates preference assessment through the presentation of hypothetical scenarios that aid in the identification of the issue being studied.

Methods: This research followed the 5 stages of CA: 1) attribute identification; 2) assignment of levels to attributes; 3) creation of scenarios; 4) preference assessment using scenarios; 5) analysis. Phase I (stages 1-3) consisted of in-depth interviews with 74 patients (25 white, 27 African American, 22 Hispanic) to identify the leading attributes and associated levels relating to CRCs. Those data were used to create and pilot-test an initial CA assessment instrument: section 1) rating of individual scenarios composed of attribute/level combinations on a 1-10 scale and section 2) ranking the attributes against one another. Phase II (stages 4-5) involves preference assessment using the final instrument among 225 low literacy, minority patients.

Results: Phase I revealed that test accuracy, preparation, frequency, discomfort, and cost were the top 5 attributes relating to CRCs and that patients wanted very simple descriptions of attribute levels. Sixty-four scenarios were generated, and fractional factorial design was used to reduce the number to 14. Pilot testing found that lower literacy patients were not able to comprehend the attribute/level combinations or to complete the rating and ranking exercises. The modified CA instrument included a “story-type” scenario format, and a card-sort method for attribute ranking. Re-piloting found lower literacy patients able to perform rating and ranking and tradeoff attribute/level combinations effectively.

Conclusions: CA is a valuable and innovative method for preference elicitation research. With appropriate tailoring, CA can assess preferences among low literacy and minority populations where other preference assessment methods may fail. CA can be used in special populations to inform interventions designed to increase use of preference-based health services such as CRCs.
Approximately 30% of HIV-infected persons in the US are co-infected with hepatitis C (HCV). Co-infection has been shown to negatively impact the outcome of both HIV and HCV treatment. A prospective cohort study was conducted in HIV-infected patients on HAART therapy to evaluate the efficacy of HCV therapy in these patients. The study included 274 patients from the Kaiser Permanente of Colorado, Denver, CO. The primary outcome measure was the rate of virologic response to antiviral therapy. The study found that the rate of virologic response was comparable between HIV-HCV co-infected patients and HIV-uninfected patients. The study also found that the rate of virologic response was higher in patients with lower CD4 counts and lower viral load. The study concludes that HCV therapy is effective in HIV-HCV co-infected patients and should be considered in patients with active HCV infection and a CD4 count >200 cells/mm3. The study also highlights the importance of monitoring for resistance during antiviral therapy in these patients.
ABSTRACTS

2273

ORAL CONCURRENT SESSION I - MEASUREMENT OF HEALTH STATUS AND PREFERENCES

2:00 PM - 3:30 PM
Sunday, October 23, 2005

2:30 PM Grand Ballroom B

TOWARD AN UNDERSTANDING OF MEDICATION NON-ADHERENCE IN THE ELDERLY WITH MULTIPLE ILLNESSES

Elliott E1, Ross-Degnan D1, Adams A1, Safran D1 and Soumerai S1

1Harvard Medical School and Harvard Pilgrim Healthcare, Boston, MA; 2Tufts-New England Medical Center, Boston, MA

Purpose: Medication non-adherence is widely reported, but little is known about how multiple illnesses affect patients’ decisions around adherence. This study explored this decision-making process among older adults with multiple chronic conditions.Methods: Twenty insured community-dwelling seniors were interviewed. Interviews were semi-structured and focused on beliefs and attitudes about adherence to medications to control chronic conditions. Interviews were transcribed and analyzed using RapidQCA, a computerized procedure for qualitative content analysis. Findings: Most participants reported overtreatment, increased costs, dose changes, and avoidance of side effects as motivators for adherence. Motivators for non-adherence included concerns about burden of medications, convenience, ability to buy medications, effectiveness, and fears of adverse effects. Conclusions: The most common motivators to trade between medicines in terms of adherence were costs and side effects. In contrast, the most common motivators to trade between medicines in terms of decision-making were costs and side effects. Medication non-adherence is a complex phenomenon influenced by multiple factors. Future research should focus on understanding the underlying decision-making processes that drive adherence and non-adherence decisions.
ORAL CONCURRENT SESSION I - MEASUREMENT OF HEALTH STATUS AND PREFERENCES

2:00 PM - 3:30 PM

Sunday, October 23, 2005

2:45 PM Grand Ballroom C

CONSIDERING ADAPTATION IN PREFERENCE ELICITIONS

Damschroder L1,2, Zikmund-Fisher B3, Kulpa J4 and Ubel P1
1VA Ann Arbor Healthcare System/Univ Michigan, Ann Arbor, MI; 2Univ of Michigan & VA Ann Arbor Healthcare System, Ann Arbor, MI; 3University of Michigan, Ann Arbor, MI; 4Univ Michigan & VA Ann Arbor Healthcare System, Ann Arbor, MI

Purpose: Patients with chronic health conditions usually place significantly higher utility on their health condition as compared with other valuations. One explanation for this discrepancy is that health utility is based on the negative aspects of the onset of the condition without considering that people can emotionally adapt to the condition over time. In previous research, we found that public ratings gave significantly higher values for paraplegia after an adaptation exercise that was designed to encourage people to consider their own ability to adapt to difficult situations. This finding held true when using the rating scale for parapersonal tradeoff elicitation methods. In the current study, we tested whether this adaptation exercise would also influence people’s responses using the time tradeoff (TTO) and standard gamble (SG) utility elicitation methods.

Methods: We presented four scenarios to Internet users drawn from a demographically balanced US panel. The scenarios asked respondents to give utilities for paraplegia, below-the-knee amputation (BKA), colostomy, and severe pain. We randomly assigned subjects to one of four experimental groups in a 2x2x2 design studied by AT (with or without an adaptation exercise) study design.

Results: 1117 respondents participated in the study. Utilities derived from the TTO method were significantly higher than those obtained through the SG method (p < 0.001). Respondents placed lowest utility on paraplegia and colostomy (0.73 and 0.74, respectively). The BKA had a utility of 0.81 and severe pain a utility of 0.79, while BKA had a utility of 0.84. In contrast to our earlier research, the adaptation exercise did not influence respondents’ valuations (p = 0.26). Respondents who indicated, during the adaptation exercise, that they believed the condition would become more upsetting over time placed lowest utility on paraplegia and colostomy (0.73 and 0.74, respectively). Pain had an average utility of 0.78. Values were significantly higher than those obtained through the SG method (p < 0.001). Respondents who did an adaptation exercise before rating four chronic health conditions, using the TTO or SG elicitation method, gave ratings that were the same as those who did not do the exercise. This finding is contrary to earlier research that used a similar adaptation exercises with other valuation methods. We speculate that thinking about adaptation changes people’s evaluations of what it would feel like to live with chronic illness but does not change how much they are willing to trade off in order to avoid that chronic illness.

CONCLUSION: The use of adaptation exercises prior to rating quality of life does not lead to significantly higher utility values for paraplegia after an adaptation exercise that was designed to encourage people to consider their own ability to adapt to difficult situations. This finding held true when using the rating scale for personscale tradeoff elicitation methods. In the current study, we tested whether this adaptation exercise would also influence people’s responses using the time tradeoff (TTO) and standard gamble (SG) utility elicitation methods.

ORAL CONCURRENT SESSION J - MEASUREMENT OF HEALTH STATUS AND PREFERENCES

2:00 PM - 3:30 PM

Sunday, October 23, 2005

3:00 PM Grand Ballroom B

DOES EDUCATION REALLY MATTER? EXAMINING THE ROLE OF EDUCATION IN HEALTH PREFERENCES AMONG OLDER ADULTS

Sims T1,2, Garber A1 and Guldstein M3
1Stanford University, Stanford, CA; Center for Primary Care and Outcomes Research, Stanford, CA

Purpose: Little is known about the relationship of education level to utilities for health states of functional dependency such as the Activities of Daily Living (ADLs) among older adults.

Methods: We developed a decision-analytic Markov model to simulate outcomes for a hypothetical cohort of 60-year-old women with operable, early breast cancer (1-2 cm, estrogen receptor-positive). In the two single-test strategies examined (MRL alone, SLNB alone), positive test results were followed by axillary dissection while negative results were followed by SLNB. Markov models implementing yearly transition rates were used to estimate quality-adjusted life expectancy and lifetime costs associated with each possible final test result for axillary node status. Patients with positive test results received an anthracycline-based chemotherapy regimen (6 cycles) and tamoxifen (5 years), while patients with negative test results received tamoxifen only. Transition rates to advanced cancer states were determined by true node status, test performance, and consequent choice of chemotherapy course. Sensitivity analyses were performed to evaluate the impact of key model assumptions and input parameters on results.

Results: Values for quality-adjusted life expectancy (discounted) was similar for all diagnostic strategies: MRL alone, 14.54 life years (LY); SLNB alone, 14.57 LY; and combined MRL-SLNB, 14.58 LY. The cost (discounted) of the single-test MRL strategy was the lowest ($46,950), followed by SLNB ($65,730) and MRL-SLNB ($64,260). In an incremental cost-effectiveness analysis, the combined MRL-SLNB strategy was eliminated through extended dominance; the incremental cost-effectiveness ratio for the combined MRL-SLNB strategy compared with MRL alone was $40,730/QALY. Elimination of SLNB by extended dominance was sensitive to MRL and MRL-SLNB test performance, in particular. Strategy rankings were otherwise robust through a range of sensitivity analyses.

Conclusions: Several plausible base-case scenarios may be used to compute the cost-effectiveness of the single-test MRL strategy as compared with the single-test SLNB strategy in the auxiliary workup of patients with operable, early breast cancer. However, further studies of MRL test performance are necessary prior to routine clinical implementation of MRL-based diagnostic strategies.

2007

ORAL CONCURRENT SESSION I - HEALTH SERVICES RESEARCH

2:00 PM - 3:30 PM

Sunday, October 23, 2005

3:00 PM Grand Ballroom A

COMPARISON OF LONG-TERM CHANGES IN HEALTH CARE UTILIZATION AND COSTS OF PROSTATE CANCER PATIENTS: AN EXAMPLE OF THE VALUE OF LONGITUDINAL DATABASES AS A CLINICAL DECISION AID

Wilson L1, Sadovsky N2, Deshane J3, Greene K4, Hriover J4 and Carroll F4
1University of California San Francisco, San Francisco, CA; 2TAP Pharmaceutical Products Inc., Lake Forest, IL

Purpose: Use of longitudinal clinical databases that assess health care utilization and QOL are ideal for assessing effects of various treatments and medical decisions on outcomes and also on patient health-care burden. We compared health care utilization and costs of 4,424 newly diagnosed prostate cancer patients by type of treatment (controlling for disease risk) for up to 10 years. We also compared health care utilization by year of diagnosis, to determine if disease burden changed over time for the same risk group.

Methods: Data from CAPSURE® (Cancer of the Prostate Strategic Urological Research Endeavor), which is a national longitudinal database registry of men with all stages of biopsy-proven prostate adenocarcinoma, were used for the study. Patients are recruited from 31 academic and community-based urology practices across the United States by participating urologists who report clinical data and follow-up information on diagnostic tests and treatments. Patient-reported quality of life and utilization of services, including physician visits, long-term care, medication use, diagnostic tests, and work history. Began in 1995, data are collected every 3-6 months. Disease risk was defined using serum PSA, Gleason score, and T-stage. We compared radical prostatectomy, cryosurgery, brachytherapy, radiation, and hormonal therapies using ANOVA with Duncan Waller and GLM with Tukey’s multiple comparison tests.

Results: A number of hospitalizations and length of stay differed by type of primary treatment (F = 34.54, p < 0.05). Patients diagnosed with prostate cancer at age 65-80 (F = 33.02, p < 0.05) had higher utilizations of inpatient services than those diagnosed with prostate cancer at age 81+. Respondents who received radiation therapy had the highest hospitalization rates. Discharge rates decreased over time from 33% during the first 6 months following surgery to 15% until 14 months after surgery, and 8%-10% until 54 months. Health care utilization differed by year of diagnosis, with a significantly higher proportion of patients (2-3 times more) having a hospitalization and discharge rate over 10%.

Conclusions: The use of PSA testing and new treatments are thought to lead to earlier diagnosis and earlier more aggressive treatments, and our study demonstrates that health utilization also follows this pattern with progressively higher resource use for prostate disease diagnosed from 1995 to 2003.

© 2006 Society for Medical Decision Making. All rights reserved. Not for commercial use or unauthorized distribution.
Background: Heparin-induced thrombocytopenia (HIT) is an infrequent but serious complication of heparin therapy. HIT is a hypercoagulable state, and life-threatening thrombotic complications often occur. Diagnosis of HIT is complicated by the non-specific nature of its signs and the operating characteristics of the antibody test used to confirm its presence. Furthermore, the direct thrombin inhibitors (DTIs) available to treat HIT are expensive and may increase bleeding risks.

Objective: To evaluate the cost-effectiveness of four treatment approaches for patients receiving heparin who develop suspected HIT. These strategies are: 1) continue heparin without testing antibody levels (HEP); 2) test antibody levels, continue heparin while test is pending, and switch to a DTI for a positive result (TEST+HEP); 3) test antibody levels, switch to a DTI while test is pending, and switch back to heparin for negative result (TEST+DTI); 4) switch to a DTI without testing antibody levels (DTI).

Methods: We performed a cost-effectiveness analysis using standard decision analytic methods. Parameter estimates were obtained from literature. Costs of drug treatment, testing, and adverse events were tallied; effectiveness was captured as quality adjusted life expectancy (QALE). The effects of parameter uncertainty and variability on model results were evaluated through sensitivity analyses.

Results: HEP had the lowest cost ($1870) and resulted in the lowest QALE (11.7 years). Relative to HEP, TEST+HEP increased QALE by 0.051 quality-adjusted life years (QALY) at an incremental cost-effectiveness ratio (ICER) of $11,000 per QALY. TEST+DTI saved an additional 0.003 QALYs compared to TEST+HEP and had an ICER of $74,300 per QALY. The ICER for DTI compared to TEST+DTI was $>1 million. These results were most sensitive to assumptions about the probability of patients presenting with thrombocytopenia actually having HIT. Thrombosis rates among HIT patients can be changed to reduce the proportion of patients with blooms and mortality rates for patients with blooms can be increased to model different HIT presentations.

Conclusions: With suspected HIT, a treatment strategy of testing antibody levels, continuing heparin while test results are pending, and switching to a DTI for positive results (TEST+HEP) appears cost-effective. TEST+DTI may be the margin of acceptable cost-effectiveness. These results are relevant to clinicians facing difficult decisions surrounding the treatment of patients with suspected HIT.
Purpose: Although most public discussions of the effects of smoking on health compare smokers versus non-smokers, the health risks associated with smoking increase dramatically with the number of cigarettes smoked per day. Decreasing cigarette consumption from 30 to 20 per day (i.e., from 1.37 packs to 1 pack per day) significantly reduces health risks, more so, in fact, than decreasing from 20 to 10 cigarettes per day. But, do people understand this, and how can we frame this information to facilitate its comprehension?

Methods: 1628 people, selected from a demographically balanced national sample, completed an Internet survey on medical decision making. We randomized respondents to one of two versions of a question about smoking risks. One version asked which of two smokers had reduced her risk of lung problems the most: Sabrina, who now smokes 15 cigarettes per day instead of 20, Dawn, who now smokes 20 cigarettes per day instead of 30 cigarettes; or to indicate that both women achieved equal risk reductions. The other version compared the increased health risks of smoking more (i.e., Sabrina smokes 20 cigarettes/day instead of 15, Dawn smokes 30 cigarettes/day instead of 20). We assessed the percentage of respondents incorrectly characterizing the dose-response relationship and compared responses across question versions.

Results: When the question was framed in terms of risk reductions (i.e., smoking fewer cigarettes per day), 49% of respondents incorrectly thought that Sabrina (who cut her cigarette intake in half) had achieved greater risk reductions than Dawn (who reduced her intake by one-third), even though both women cut out the same number of cigarettes. By contrast, when the question discussed risk increases, only 4% of respondents gave the incorrect response that Sabrina’s change from 15 cigarettes/day to 10 yielded the larger risk increase, a highly significant response shift (12) = 449.54, p < 0.0001.

Conclusions: Consistent with our research on perceptions of blood pressure and weight reductions, people perceive smoking reductions as more valuable for infrequent smokers such as Sabrina. This misperception could be a health risk for some smokers (e.g., Dawn) to underestimate the benefit of significant, yet incomplete, smoking cessation efforts. Fortunately, accurately perception of the dose-response relationship was facilitated by reframing the problem in terms of risk increases, suggesting that interventions so framed might improve decision making.
2390
ORAL CONCURRENT SESSION M · TECHNOLOGY ASSESSMENT
1:00 PM - 2:30 PM
Monday, October 24, 2005
1:30 PM Grand Ballroom C
AUTOMATION BIAS IN MEDICAL DECISION MAKING: A STUDY WITH INCORRECT COMPUTER PROMPTING IN BREAST CANCER SCREENING
Alberdi E, Poryakalo A, Strigini L and Ayton P
City University, London, United Kingdom
Purpose: To investigate the effects of inappropriate computer advice on the decisions of clinicians in the investigation of symptoms of mammography. The study focused on participants in which the computer failed to detect cancers, either placing prompts elsewhere in the mamogram, or not prompting at all.
Results: Of the 183 consenting patients, 153 (83.3%) completed all questionnaires and interviews. Patients were seen by a total of 8 physicians all of whom contributed data to the study. The median patient-perceived breast cancer risk pre-visit was 0%, which declined by 11% post-visit but remained 40% higher than risk estimated by the Gail model. Physicians overestimated patients’ perceptions of risk post-visit (median risk 70%) and underestimated patients’ perceived risk pre-visit (median risk 30%). Patient preferences for the format in which they received risk information was matched by physicians 37% of the time. Physicians correctly identified patient preferences for prevention interventions in 40% of the scenarios presented.
Conclusions: Despite a wide variety of patient-physician interaction, physicians have a high level of misunderstanding about their patients’ knowledge about breast cancer and their preferences regarding decision making. Physicians’ knowledge of patient risk perception is often times inaccurate, making it difficult to identify when patients are making poorly informed decisions. The limited use of patient-preferred formats for risk communication is likely to impede physicians’ ability to guide the process of care. And physicians’ lack of understanding regarding patient’s preferences for prevention intervention points to the importance of improved physician-patient communication to guide decision making and ensure concordance between outcomes and preferences.

2380
ORAL CONCURRENT SESSION M · TECHNOLOGY ASSESSMENT
1:00 PM - 2:30 PM
Monday, October 24, 2005
1:45 PM Grand Ballroom C
A COMPARISON OF A PATIENT-LEVEL SIMULATION, COHORT SIMULATION, AND CLOSED-FORM ANALYTIC COST-EFFECTIVENESS MODELS OF THE TREATMENT OF RHEUMATOID ARTHRITIS
Hawkins N, Claxton K, Griffin S and Sculptor M
University of York, York, United Kingdom
Purpose: To investigate the feasibility of closed-form analytic solutions of decision analytic models and compare their efficiency and accuracy to cohort and patient-level simulation models. Background: Long-term cost-effectiveness models require estimates of expected costs and effects accounting for patient variation over time and discounting. This usually involves either patient-level simulation or cohort modeling. These are computationally intensive, restricting the approach to simple decision problems. Methods: Models of rheumatoid arthritis (RA) usually include components representing the initial accumulation of disease when the treatment is started, the progression of disease while on treatment, and the “rebound” in symptoms, and subsequent progression when treatment is ended. The estimates of cost-effectiveness are sensitive to the assumptions surrounding rebound and hence to the modeling delays the process of disease. A closed-form analytic solution of an RA model was derived by using MAPLE to obtain a definite integral, which provided estimates of expected costs and effects as a function of time to treatment failure for an individual. This was then integrated over the variation in treatment failure between individuals to obtain expected costs and effects. The results were compared to patient-level simulation.
Results: The results were compared to patient-level simulation. The closed-form analytic solution was faster to evaluate, the cohort model took 40 seconds to evaluate, and the closed-form analytic model took 0.0004 seconds to evaluate. The results from the closed-form and patient-level simulation models converged on the results of the closed form solution as the number of simulations and the number of time-points evaluated increased. The analytic model was the most accurate and provided an “exact” solution.
Conclusions: Closed-form analytic solutions of decision-analytic models, where feasible, have the advantage over simulation methods of being both exact and computationally efficient and should be considered when constructing models expected to be used in potential of great help in the implementation of value of information methods.

ABSTRACTS
PERCEPTIONS RELATED TO THE WEIGHTING OF TAMOXIFEN'S RISKS AND BENEFITS

Linkupa and Peters E

Duke University Medical Center, Durham, NC; Decision Research, Eugene, OR

Purpose: To assess associations among 1) interest in tamoxifen (Tam) use for breast cancer chemoprevention, 2) weighing of Tam's risks and benefits, 3) perceptions of breast cancer risk, and 4) numeracy.

Methods: 56 women recruited from gynecology clinics in North Carolina who qualified for Tam were given their five-year breast cancer risk estimate (Caile score) and a computerized decision aid that provided tailored numerical feedback as to their chances of experiencing five health risks and five health benefits of taking Tam for five years. After reviewing the information, participants were asked 1) for their perceived lifetime breast cancer risk; 2) whether Tam increased, decreased, or did not affect their chances of experiencing the 10 health events; 3) for their overall weighing of Tam's risks and benefits for self (1 = benefits outweigh the risks by a lot to 5 = risks outweigh the benefits by a lot); and 4) interest in taking Tam (1 = not at all to 5 = extremely). All women completed Likert and colleagues’ measure of numeracy.

Results: Women with greater numeracy were more likely to specify correctly how Tam affected their chances of experiencing these events (r = 0.37, p < 0.000, M = 7 out of 10 correct). Participants viewed Tam's risks as slightly outweighing the benefits (M = 3.25). Greater perceived lifetime breast cancer risk and a greater perceived risk-to-benefit ratio was related to more (r = 0.27, p < 0.000) and less interest in use (r = –0.32, p < 0.000, M = 2.3 in interest), respectively. Numeracy was not related to interest in use. Perceptions of Tam's benefit-to-risk ratio were related primarily to actual risks (r = 0.34, p < 0.01; partialling actual benefit) than actual benefits (r = –0.11, NS, partialling actual risk), despite our data showing that increasing breast cancer risk (MSG(2) = 2.44, range 1.67 to 6.71%) was related to a trend improved benefit-to-risk ratio (r = 0.34, p < 0.02) and especially actual benefit (r = 0.65, p < 0.0001).

Conclusions: These data suggest that more research is needed to understand why women disproportionate Tam's risks rather than benefits, especially when Tam's benefit-to-risk ratio improves with greater breast cancer risk. Numeracy may play an important role in how women understand numerical information on Tam. Results may have implications for how medical personnel may wish to present information on Tam’s risks and benefits.

E22  MEDICAL DECISION MAKING/JAN–FEB 2006
Purpose: Recent studies have attributed declining prostate cancer treatment costs to improvement in patients care pathways. These studies have not accounted for changes in the type of patients who are now diagnosed with prostate cancer due to widespread adoption of PSA screening. We propose a method to account for the impact of healthy screenee bias when estimating costs of cancer care.

Methods: We used SEER-Medicare claims data to estimate initial care costs among 31,643 prostate cancer patients who were treated with radiotherapy. We first employed standard methods for assessing cancer attributable costs by estimating total costs for newly diagnosed patients during a three-month initial treatment period then subtracting non-cancer costs from age-matched controls without cancer. We then performed a second analysis where we used cases as their own controls. Here, non-cancer costs were estimated based on claims among cases prior to diagnosis.

Results: Between 1991 and 1999, total costs per patient among cases declined from $10,900 to $8,700 on average, while non-cancer costs for matched controls increased from $2,280 to $5,700. Thus, cancer-attributable costs for initial treatment declined from $8,120 to $3,000 (63% decrease per year). Among cases, non-cancer costs measured prior to diagnosis actually declined slightly over the period from $3,300 to $3,200 indicating that by the end of the period cases were substantially healthier before their diagnosis than population controls without cancer. Matching on age and comorbidity in the original analysis was not sufficient to control for this healthy screenee effect. Based on our second analysis, we estimated that cancer-attributable costs of initial care changed from $8,200 in 1991 to $4,800 in 1999 (5% decrease per year).

Conclusion: Average radiotherapy costs for men newly diagnosed with prostate cancer have improved during the PSA era, although observed declines are inflated by healthy screenee bias. One third of the observed decline in initial treatment costs is because men diagnosed today include a subgroup of healthy subjects as indicated by their substantially lower non-cancer costs compared with age-matched controls without prostate cancer. Evaluating economic outcomes in screened populations should consider the impact of healthy screenee bias.

Purpose: Our objective was to determine independent predictors of pediatricians’ anticipated regret with a clinical decision.

Methods: We mailed a 4-page questionnaire to 1502 randomly selected US pediatricians. The questionnaire presented a clinical vignette of a previously healthy 1-year-old boy with a fever of 39.1°. To assess anticipated regret, we had participants choose which scenario they would regret more if they had chosen the boy with antibiotics, with subsequent anaphylaxis and complete recovery, or if not giving antibiotics, with subsequent pneumococcal sepsis and complete recovery. To assess the availability heuristic, we asked if participants would recall a potentially preventable case of sepsis that resulted in a bad outcome and/or a case of anaphylaxis caused by medication that resulted in a bad outcome. To measure fear of malpractice, we used Frank’s 6-question “Malpractice Fear Scale”.

Results: 653 pediatricians returned completed surveys (49.0% of eligible). The majority (59%) had anticipated regret from the sepsis scenario, while 41% would regret more from the anaphylaxis scenario. Among participants who recalled a bad case of sepsis (63%) than who did not (54%, p = 0.02). Similarly, having more anticipated regret from the anaphylaxis scenario was more common among participants who recalled a bad case of sepsis (63%) than who did not (54%, p = 0.04). In adjusted analysis, significant predictors of having more anticipated regret from the anaphylaxis scenario were anaphylaxis fear among participants who recalled a bad case of anaphylaxis (p = 0.04). While those patients who did not recall an anaphylaxis case had more anticipated regret from the anaphylaxis scenario compared to those who did recall an anaphylaxis case, the relative risk was not statistically significant (p = 0.7).

Conclusions: The availability heuristic and high malpractice fear independently predicted pediatricians’ anticipated regret.
ABSTRACTS

1826
ORAL CONCURRENT SESSION N - HEALTH ECONOMICS

3:00 PM - 4:30 PM
Monday, October 24, 2005

3:45 PM Grand Ballroom B
EXPECTED VALUE OF RESEARCH ON INTERVENTIONS TO INCREASE UPTAKE TO BREAST CANCER SCREENING
Wetnogen N1 and Ades A1
Bristol University, Bristol, United Kingdom; Medical Research Council, Bristol, United Kingdom
Purpose: To develop methods to assess the value of further research into interventions to increase uptake to breast cancer screening in the UK: 1) before a cluster-randomized trial reported in 2001 and 2) after that trial.
Methods: We analyzed data from a cluster-randomized 2x2 factorial trial on letter and flag interventions to increase attendance at breast cancer screening in the UK (Bankhead et al., 2001). We use a random effects logistic regression model in a Bayesian decision analytic framework, which allows reasonable estimation of the interaction term, and is embedded directly into a cost-effectiveness model incorporating costs and life-years saved by screening and early treatment. We consider how to calculate the expected value of a trial of this design 1) before the trial and 2) after the trial. Prior to the 2001 study, there was a substantial body of research into various types of intervention (mainly from Australia and the US). We use this evidence to inform a multivariate normal prior for our regression coefficients, including correlations. We then combine this prior information with the data to obtain a multivariate normal posterior that becomes the prior after the trial. We calculate, for both priors, the expected value of removing some uncertainty in intervention efficacy parameters by running a future cluster-randomized factorial design trial of given cluster size—Expected Value of Sample Information (EVSI). EVSI is calculated by simulating future data from the prior and using weighted least-squares regression to obtain multivariate normal sufficient statistics for the regression coefficients, taking correlations.
Results: Using a £100,000 valuation of a QALY, EVSI per woman invited to screening was £2.88 before and £3.18 after the trial, for the 2001 trial design study. Population EVSI over a 10-year horizon was £0.7 m before and £3.6 m after the trial.
Conclusions: By making multivariate normal approximations, we have calculated EVSI for cluster-randomized trials. There was considerable value in running the 2001 trial, and although the 2001 trial has reduced decision uncertainty, there is still value in running a similar trial in the future. We discuss methods for finding the optimal cluster sizes for a future study, allowing for the cost of the study and the opportunity cost of receiving a suboptimal intervention during the trial.

E24 • MEDICAL DECISION MAKING/JAN–FEB 2006

2397
ORAL CONCURRENT SESSION O - JUDGMENT AND DECISION MAKING: APPLICATIONS

3:00 PM - 4:30 PM
Monday, October 24, 2005

3:30 PM Grand Ballroom C
LIGHTNING MAY NOT STRIKE TWICE, BUT MY CANCER WILL: THE SPECIAL STATUS OF CANCER RECURRENCE RISKS
Zikmund-Fisher B1, Fagerlin A1 and Ubel P1
1VA Ann Arbor Healthcare System, Ann Arbor, MI; University of Michigan, Ann Arbor, MI
Purpose: Cancer survivors face risks of cancer recurrence in addition to elevated risks of new cancers. But, do people think about these risks in the same way? Concerning that the prior cancer might still “be inside me” could give it special status, making the risk of its recurrence seem more likely or more worrisome than equally likely risks of a new cancer.
Methods: 91 patients recruited from a demographically balanced sample completed an Internet survey on medical decision making. The survey asked respondents to imagine that they previously had either skin cancer or thyroid cancer (randomized), which was successfully treated into remission. Respondents were then informed of two future risks (in random order): the risk of recurrence of their prior cancer and the risk of developing an alternate cancer. The first risk seen was numerically more worrisome than equally likely new cancers, even when cancer type is randomly varied. The extra consideration given to the possibility of recurrence cancer may influence how patients assess the risk-benefit tradeoff of adjuvant therapies (likely encouraging more invasive approaches), which may in turn decrease the quality of life of cancer survivors.
ABSTRACTS

2386

ORAL CONCURRENT SESSION N - HEALTH ECONOMICS
3:00 PM - 4:30 PM
Monday, October 24, 2005
4:00 PM Grand Ballroom B
THE USE OF COST-EFFECTIVENESS ANALYSIS IN TECHNOLOGY COVERAGE DECISIONS: EVIDENCE FROM THE UK

Bryan S. Williams I and Melver S
University of Birmingham, Birmingham, United Kingdom

Purpose: To explore the use and influence of economic analyses in national technology coverage decisions and to consider how the impact of such analyses might be increased.

Methods: Over recent years there has been repeated expression of concern about the apparent lack of use of cost-effectiveness analyses (CEAs). Given that the central problem addressed by the discipline of economics is resource scarcity, this is both surprising and concerning. The National Institute for Clinical Excellence (NICE) was established in the UK with a remit of making national coverage decisions concerning the adoption of health technologies. The Institute can be viewed as an experiment in the explicit use of CEAs to inform national health policy—an independent economic analysis is commissioned for every appraisal topic. This article reports findings from a qualitative study of appraisals conducted by NICE in the first 7 technologies. Literature made available to the Committee was analyzed, formal meetings were observed, and 30 Committee members were interviewed. Data were collated and analyzed on an ongoing, iterative basis.

Results: The themes identified are grouped under three headings: political (i.e., Committee composition and role of members), conceptual and practical issues (including Committee operation and conceptual challenges), and practical issues. The article highlights the value in using CEA as the framework for discussion—a quote on this theme: “The reason why CEA is important is...the fact that it provides us with a framework to identify what we should be concerned about and where we should be pushing the argument.” In addition, important themes relating to the threshold QALY value, the shortcomings of the QALY construct itself, and equity arguments were identified. A theme of “understanding of the CEA” reveals worrying shortcomings, as shown in another quote: “A number of Committee members don’t appear to understand a word of what is going on when the health economics is discussed.”

Conclusions: The research primarily highlights the value of CEA not only in providing overall direction for coverage decisions but also in giving structure and focus to such discussions. Concerns exist in relation to CEA concepts (e.g., the QALY) and practical issues of how CEA is to be used.

2051

ORAL CONCURRENT SESSION N - HEALTH ECONOMICS
3:00 PM - 4:30 PM
Monday, October 24, 2005
4:15 PM Grand Ballroom B
USING COST-EFFECTIVENESS ANALYSIS TO DESIGN A NATIONAL CORD BLOOD BANK

Howard D. Meltzer D, Kollman C, Maires M, Horovitz M, Logan B and Setterholm M
Emory University, Atlanta, GA; University of Chicago, Chicago, IL; Jaeb Center for Health Research, Tampa, FL; National Marrow Donor Program, Minneapolis, MN; Medical College of Wisconsin, Milwaukee, WI

Purpose: Transplantation with stem cells from stored umbilical cord blood units is an alternative to living unrelated bone marrow transplantation. Both are used to treat leukemia, multiple myeloma, and a handful of rare blood diseases. Because the quality of the human leukocyte antigen (HLA) match between recipient and donor is a key determinant of transplant outcomes, it is important that transplant candidates have access to a large number of stored cord blood units from which a suitable match can be selected.

Currently, transplant candidates must conduct separate searches of each of the 22 non-profit “banks” that collect and store donated cord blood. In 2001, Congress appropriated $10 million to create a national cord blood bank in order to streamline the search process. In this study, commissioned by the Institute of Medicine, we used cost-effectiveness analysis to determine how many cord units a national bank should place in storage. The larger the inventory, the greater the likelihood that transplant candidates will match to a stored unit. However, processing new cord blood units is costly, so these benefits must be weighed against the collection and storage costs.

Methods: We estimated the likelihood that transplant candidates will match to a cord blood unit as a function of the inventory level based on the distribution of HLA types in the National Marrow Donor Program registry. We calculated lifetime gains based on these probabilities, historical survival data, and assumptions about patients’ preferences for cord blood versus marrow donors when both types of matches are available. (The time to transplant with cord blood is shorter and matching requirements are less stringent, but the volume of stem cells obtained is less than for a living donor.) Separately, we used a differential equation describing the cord inventory level to calculate total costs as a function of inventory.

Results: The cost per life year gained associated with increasing inventory from 50,000 to 100,000 to $58,000-$75,000 and from 100,000 to 150,000 units is $50,000-$42,000, depending on the assumptions about the degree to which survival rates for cord transplants vary by donor-patient match quality.

Conclusions: Assuming the monetary value of a year of life exceeds $100,000, it is cost-effective for a national bank to maintain an inventory of 100,000-150,000 units.

ABSTRACTS

2394

ORAL CONCURRENT SESSION O - JUDGMENT AND DECISION MAKING: APPLICATIONS
3:00 PM - 4:30 PM
Monday, October 24, 2005
4:00 PM Grand Ballroom C
THE PATIENT'S PERCEPTION OF CHOICE IN CANCER TREATMENT DECISION MAKING

Davies L, Grossman D and Rhodes L
White River Junction VA Medical Center and Dartmouth Medical School, White River Junction, VT; Group Health Cooperative, Seattle, WA; University of Washington, Seattle, WA

Purpose: Describe patient perceptions of their choices for, and role in, cancer treatment decision-making. Compare patient perceptions to audiotapes of the consultation.

Methods: 74 patients (32 male) were asked to complete an information sheet. After reviewing the sheet, participants were interviewed about their perceptions of choice in cancer treatment decision making. In addition, audiotapes of the patient-physician interaction were reviewed. The research was qualitative in nature. Fieldwork was conducted over 12 months and focused on 7 technologies. Literature made available to the Committee was analyzed, formal meetings were observed, and 30 Committee members were interviewed. Data were collated and analyzed on an ongoing, iterative basis.

Results: The themes identified are grouped under three headings: political (i.e., Committee composition and role of members), conceptual and practical issues (including Committee operation and conceptual challenges), and practical issues. The article highlights the value in using CEA as the framework for discussion—another quote: “The reason why CEA is important is...the fact that it provides us with a framework to identify what we should be concerned about and where we should be pushing the argument.” In addition, important themes relating to the threshold QALY value, the shortcomings of the QALY construct itself, and equity arguments were identified. A theme of “understanding of the CEA” reveals worrying shortcomings, as shown in another quote: “A number of Committee members don’t appear to understand a word of what is going on when the health economics is discussed.”

Conclusions: The research primarily highlights the value of CEA not only in providing overall direction for coverage decisions but also in giving structure and focus to such discussions. Concerns exist in relation to CEA concepts (e.g., the QALY) and practical issues of how CEA is to be used.

2413

ORAL CONCURRENT SESSION O - JUDGMENT AND DECISION MAKING: APPLICATIONS
3:00 PM - 4:30 PM
Monday, October 24, 2005
4:15 PM Grand Ballroom C
PEOPLE'S PREFERENCES FOR GRAPHICAL WAYS OF PRESENTING RISK/BENEFIT INFORMATION

Ann Arbor VA & University of Michigan, Ann Arbor, MI; VA Ann Arbor Healthcare System, Ann Arbor, MI; University of Michigan, Ann Arbor, MI; University of Michigan, Ann Arbor VA Health System, Ann Arbor, MI

Purpose: Understanding statistical information that describes the risks and benefits of treatment is difficult for many patients. Attempting to make this process easier and to improve patients’ ability to make informed treatment decisions, many decision aids and health education materials now provide graphical representations of the risks and benefits of treatment. However, little research has tested people’s perceptions of graphs (e.g., helpfulness, effectiveness, trustworthiness) or from which type of graphs people prefer to receive their medical information.

Methods: 2450 subjects from a demographically balanced panel participated in an online survey. Participants were randomized to receive risk/benefit information presented in one of five formats: a) a graph, bar graph, pie chart, sparkplg, or modified pie (clock) graph. Respondents’ perceptions of the helpfulness, effectiveness (of describing risks and benefits), scientificness, and trustworthiness were assessed through a rating exercise (0-6 Likert-type scale) consisting of 8 total questions. After reviewing all 5 formats, respondents ranked ordered them in terms of which format “they would prefer if they were trying to understand the risks and benefits of a treatment.”

Results: Bar graphs were most highly rated in terms of their perceived helpfulness (4.14) and effectiveness (4.16) of presenting risk/benefit information. However, pictographs were considered useful for determining exact numbers and were rated as the most preferred method for learning about the risks and benefits of treatment. While pictographs were rated as most trustworthy (3.62), bar graphs were rated as most scientific (3.61). People indicated that if they were reading an article, they were more likely to look at an accompanying bar graph and least likely to look at a pictograph (4.08 vs. 3.72). In all other questions, pie graphs received the lowest ratings. When patients rank ordered the 5 formats, they preferred to receive their risk/benefit information in bar graphs, the less precise nature of this format may actually result in poorer comprehension of risk/benefit information. When deciding how to present statistical information, researchers and decision aid developers should consider both the preferences of the target audience and their ability to comprehend the specific material presented.
Purpose: To capture the impact of integrated decision support for patients facing the decision of surgical options for early stage breast cancer.

Methods: A cohort of 75 newly diagnosed breast cancer patients completed a computerized in-clinic tool to help patients to make a treatment decision. This tool provided computerized, context-specific decision-making data. Following the intake, patients watched a video decision aid entitled Early Stage Breast Cancer: Choosing Your Surgery, which provided information about early stage invasive breast cancer and the surgery treatment choices. After the video, patients answered decision process questions; additional process questions were asked after the surgery consultation. Outcome measures were: 1) baseline treatment preference, actual decision following the computer-assisted decisional conflict; satisfaction with preparation for decision making, decision self-efficacy. Patients reported postvideo decisions. After watching the video and after the meeting patients were asked to "manually" confirm the accuracy of the clinical data rather than having them extracted automatically from the electronic medical record.

Results: Respondents’ knowledge scores for stroke and bleeding risk increased from 38% before using the AF-DST to 62% after (p = 0.29). The mean score on the CSUQ was 6, indicating high user satisfaction. The mean score of 6.5 on the second Likert-scale indicated that users believed that the AF-DST helped them make more informed decisions. In usability testing debriefing, clinicians favored entering the risk factor data themselves to ensure accuracy rather than having it entered automatically from the electronic medical record. Most physicians said they would likely reconsider their treatment decision when the tool recommended another strategy (mean score 5); but clinicians felt somewhat uncomfortable understanding how the “black box” calculations were made.

Conclusions: A patient-specific AF-DST can increase clinicians’ knowledge regarding risks and benefits of anticoagulation in nonvalvular atrial fibrillation. Satisfaction with the Web-based decision support tool was high, but clinicians preferred to enter the patient-specific risk factors “manually” confirm the accuracy of the clinical data rather than having them extracted automatically from the electronic medical record.

Note that all subscale scores were lower after the video except for uncertainty scores (38% of patients were unsure about a treatment preference), after meeting with their physician, these uncertainty scores dropped significantly. Total decisional conflict scores were significantly lower both after the video and after the meeting patients were asked to "manually" confirm the accuracy of the clinical data rather than having them extracted automatically from the electronic medical record.

Decisional Conflict Means (SD) (p < 0.01): *p < 0.001.
ABSTRACTS

POSTER SESSION I
7:00 AM - 8:30 AM
Saturday, October 22, 2005
5 Pacific Room EFG
A FRAMEWORK FOR MAKING DIFFICULT DECISIONS IN "UNRESECTABLE" NON-SMALL CELL LUNG CANCER CARE
Moure W1, Kee F2 and Dolan J1
1Queen’s University of Belfast, Belfast, United Kingdom; 2Queen’s University Belfast, Belfast, United Kingdom. Unity Health System, Rochester, NY
The optimal management of advanced "unresectable" non-small cell lung cancer (NSCLC) has not been established. Local clinicians identified a need to review and update existing Northern Ireland clinical guidelines.
The analytic hierarchy process (AHP) is an explicit quantitative decision methodology, which can be used to support group consensus development. This case study examined the application of an AHP framework to support the relative valuation of NSCLC treatment regimens in a local context.
A preliminary scoping review of the topic was conducted to determine appropriate parameters for the topic. Factors considered important in comparisons between lung cancer treatment regimens were identified, defined, and structured in an evaluation hierarchy. The relative importance of each evaluation criterion, in the context of different NSCLC patient populations, was derived from group pairwise judgments (Figure 1). Evaluation models were agreed for 3 different clinical approaches—radiological therapy, radical-chemoradiotherapy, and palliative chemotherapy. Relevant evidence for each regimen was identified, appraised, and summarized to support pairwise judgments. The “most appropriate” option was determined by derivation of composite (ratio-level) priority scores (Figure 2).

Figure 1: Evaluation Criteria Group Priority Scores
Figure 2: Evaluation Models - composite priority scores for regimens

Conclusion: Applying the AHP framework helped focus discussion, clarify the areas of consensus in the group, and set out a more rational basis for agreement.

POSTER SESSION I
7:00 AM - 8:30 AM
Saturday, October 22, 2005
8 Pacific Room EFG
CANCER SCREENING IN CHINA
Orhewere M1,2, Dorward S3, Forman D4 and Delaney B4
1The University of Birmingham, Birmingham, United Kingdom; 2The University of Leeds, Leeds, United Kingdom; 3University of Birmingham, Birmingham, United Kingdom
Purpose: Approximately 1% of patients taking non-steroidal anti-inflammatory (NSAI) mediation will suffer a bleeding peptic ulcer per year. As many patients with arthritis are dependent on NSAIDs to maintain mobility, a number of strategies have been studied to prevent recurrent bleeding. These include eradication of Helicobacter pylori and co-prescription of misoprostol or a proton pump inhibitor (PPI). Data are now available from Cochrane systematic reviews as to the effectiveness of these approaches. This model aimed to determine the cost-effectiveness of the strategies.
Methods: The software TreeAge Pro 2005 (TreeAge) was used. A Markov model was used to predict monthly recurrence over 5 years, and values obtained for baseline risk, effect of therapy, costs and utilities using Cochrane meta-analyses and published national cost data. Variables were specified as Beta binomial (for proportions), Lognormal (for Relative risk), or Gamma distributions (for costs). Six strategies were compared, do nothing, misoprostol, PPI, H. pylori eradication alone, H. pylori eradication with misoprostol, and H. pylori eradication with PPI after. A 1000 trial Monte Carlo simulation was carried out.
Results: H. pylori eradication alone dominated all other strategies except H. pylori eradication followed by misoprostol. H. pylori eradication alone produced a mean of 4.14 QALYs over 5 years at a cost of £78.84 per patient. Adding maintenance misoprostol after eradication added an additional 0.002 QALYs at an ICER of £653,000 per QALY. Probabilistic sensitivity analysis showed that H. pylori eradication alone dominated the other strategies with a likelihood of 90% of being the most cost-effective strategy at £1,350,000 per QALY. At £50,000 per QALY, there was only a 1% chance that adding misoprostol would be cost-effective.
Conclusion: H. pylori eradication therapy is the preferred strategy for the prevention of recurrent related peptic ulcer bleedings, it is more effective and less costly than other strategies, except adding misoprostol co-prescription. However, adding misoprostol is not cost-effective at any reasonable willingness to pay.

ABSTRACTS

POSTER SESSION I
7:00 AM - 8:30 AM
Saturday, October 22, 2005
7 Pacific Room EFG
ECO-EFFECTIVENESS OF PREVENTION STRATEGIES FOR NON-STEROIDAL INFLAMMATORY FACTOR INHIBITORS IN RHEUMATOID ARTHRITIS PATIENTS WITH INADEQUATE RESPONSE TO METHOTREXATE
Kamal K1, Miller L2, Madhavan S3, Scott V4, Kovaliokas J4 and Hornsby J4
1West Virginia University School of Pharmacy, Morgantown, WV; 2West Virginia University School of Medicine, Morgantown, WV
Purpose: To perform a cost-effectiveness analysis of the three tumor necrosis factor (TNF) inhibitors: adalimumab, etanercept, and infliximab, for patients with rheumatoid arthritis (RA) that inadequately respond to methotrexate (MTX) alone.
Methods: A Markov model was developed to estimate the health effects and costs associated with five treatment strategies for patients with RA that inadequately respond to MTX alone: 1) adalimumab plus MTX, 2) etanercept plus MTX, 3) infliximab plus MTX, 4) leflunomide plus MTX, and 5) standard therapy of MTX. A hypothetical cohort of 10,000 55-year-old women was evaluated using Monte Carlo simulation. Efficacy data and treatment withdrawal rates were based on randomised controlled trials of the various treatments conducted in patients with inadequate responses to MTX. Costs associated with joint replacement surgery were modeled in patients who did not respond to the treatments. The study was conducted from a societal perspective, and the total cost of therapy for each agent included direct costs associated with treating MTX-resistant RA combined with indirect costs incurred by the patients as a result of the disease. The main outcome measures were net gains in quality-adjusted life expectancy and incremental cost-effectiveness ratios (ICERs) in dollars per quality-adjusted life year (QALY) gained. Costs and effects were discounted at 3%. To test the robustness of the model, extensive sensitivity analyses were conducted, including a probabilistic sensitivity analysis.
Results: The combination of etanercept and MTX was the most cost-effective treatment with an ICER of $49,724/QALY when compared against traditional disease-modifying anti-rheumatic drugs (leflunomide plus MTX and standard MTX treatment) and the combination of the other two TNF inhibitors with MTX. The combination of leflunomide and MTX was the second most cost-effective option in those patients with an ICER of $53,815/QALY. One-way and probabilistic sensitivity analyses indicated that the conclusions were relatively stable to variations in model assumptions.
Conclusions: Of the three TNF inhibitors, etanercept is the most cost-effective from a societal perspective. The ICERs for both etanercept and leflunomide remained within the acceptable range of $50,000/QALY and $100,000/QALY in the simulated population under a wide range of assumptions, as compared to the other comparators.
Utilities (0.83) for prophylactic oophorectomy and mastectomy (women with a first-degree relative with either ovarian cancer or breast cancer before age 40). We evaluated women from two populations: a “high-risk” general population (women with first-degree relative with either ovarian cancer or breast cancer before age 40). We evaluated four scenarios in which women have all combinations of high utilities (0.97) or low utilities (0.83) for prophylactic oophorectomy and mastectomy.

Results: In the high-risk general population (prevalence BRCA1/2 = 8% in unaffected). Unaffected-Woman Testing resulted in slightly more quality-adjusted life years (QALYs) than did Proband Testing for women whose preference for prophylactic oophorectomy and/or mastectomy is high (range of 3.5 to 4.4 quality-adjusted days). The incremental cost-effectiveness ratio (ICER) of Unaffected-Woman Testing compared to Proband Testing is <$100,000/QALY in each scenario. In unaffected women whose utilities for both surgeries were low, no BRCA1/2 testing is the preferred strategy. For high-risk Ashkenazi Jewish women (prevalence BRCA1/2 = 20% in probands). Unaffected-Woman Testing was associated with more QALYs than Proband Testing (2.8 quality-adjusted days) for women whose preference for prophylactic mastectomy or prophylactic oophorectomy was high (ICER approximately $20,000/QALY). For women who have a high utility for prophylactic mastectomy and for prophylactic oophorectomy, Proband Testing was recommended for women whose utilities for both surgeries were low, it was cost-effective ($86,623/QALY) to first test the proband before testing the unaffected woman.

Conclusions: Although current recommendations suggest testing probands first, our analysis found that testing an unaffected relative first is a reasonable option depending on a woman’s utilities for prophylactic surgery. These findings have implication for testing guidelines.

Background: Adverse drug reactions are a common complication of chemotherapy administration. Uridine diphosphate glucuronosyltransferase 1A1 (UGT1A1) genotype can predict which patients will have severe toxicity (diabetes and neutropenia) from irinotecan-based chemotherapy. To analyze the cost-effectiveness of UGT1A1 pharmacogenomic treatment for patients with metastatic colorectal cancer.

Methods: We developed a Markov decision model using a hypothetical cohort of patients with metastatic colorectal cancer beginning chemotherapy. In the model, patients received either a pharmacogenomic treatment strategy of testing for UGT1A1 genotype or conventional treatment (no testing). All patients not tested for UGT1A1 genotype received conventional treatment with an irinotecan-based FOLFIRI regimen. Genotype-tested patients determined to be UGT1A1 deficient or heterozygous were treated with FOLFOX, a non-irinotecan-based regimen that reduces the risk of diarrhea and neutropenia for these patients without associated with a greater risk of neuropathy. Probabilities of toxicities, which included neutropenia, diarrhea, and neuropathy, were based on published literature on FOLFIRI and FOLFOX. Costs were obtained from Medicare & Medicaid Services reimbursement data to calculate costs for physician and hospital services without adjustment for geographic location, and therefore representing a national cost per procedure. Drug costs were estimated using Federal Supply Schedule. Health outcomes were measured using quality-adjusted life years (QALYs). Univariate and probabilistic sensitivity analyses were conducted to assess uncertainty in the model parameters.

Results: The pharmacogenomic strategy provided 1.183 QALYs and cost $80,535, whereas conventional treatment provided 1.057 QALYs at a cost of $62,906. The incremental cost-effectiveness ratio for pharmacogenomic treatment was $41,786/QALY in 10,000 probabilistic Monte Carlo simulations. The pharmacogenomic treatment was cost-effective in 52% of trials using a $50,000/QALY threshold. The most influential variables in the univariate sensitivity analysis were the probabilities of death associated with FOLFOX and FOLFIRI as well as the costs and years of survival associated with FOLFOX. Genotype-tested patients had a smaller benefit. Costs were estimated from the MTA clinical trial and published sources. The efficacy of treatments was taken from the MTA clinical trial.

Conclusions: The strategy of managed medication could produce an incremental cost-effectiveness ratio of less than $100,000 per QALY gained. Additional research is needed in patients with multiple comorbidities to significantly more patients than community care, while adding behavioral treatments in ADHD (MTA) clinical trial. The MTA trial compared the treatments of managed medication for all patients would require an additional investment in ADHD, depression, and conduct disorder. However, if only the group of patients with multiple illnesses such as ADHD, depression and conduct disorder, an incremental cost-effectiveness ratio of less than $100,000 per QALY is possible for multimodal treatment for ADHD to be considered cost-effective?
PID outcome costs were obtained from the literature. Multiple sensitivity analyses were performed over a 10-year time horizon using a 3% discount rate. PID outcome frequencies and utilities were differed between regimens. The baseline model examined 18-year-olds from the societal perspective were illness was absent), calculating incremental costs per QALY gained if PID complication rates exist between the least and most expensive. CDC recommends several antibiotic treatment regimens; however, there is no evidence of differences in PID complication outcomes between treatments while almost three-fold differences in cost exist between the least and most expensive.

Methods: We constructed a Markov decision model to estimate the incremental cost-effectiveness of CDC-recommended outpatient antibiotic regimens for mild to moderate PID (i.e., when severe illness was absent), calculating incremental costs per QALY gained if PID complication rates differed between regimens. The baseline model examined 18-year-olds from the societal perspective of why the differences are occurring, our results also raise the possibility that equity issues may be identified and addressed.

Conclusion: Within the cost range of CDC-recommended antibiotic regimens for outpatient PID treatment, use of more expensive antibiotics is economically reasonable if relatively small decreases in PID complication rates occur due to greater treatment effectiveness or adherence. Further work is needed to investigate differences in antibiotic effectiveness for modifying PID complication risk.

Results: Antibiotic costs vary between $64 (cefotaxim and doxycycline [CD]) and $188 (cefoxitin and doxycycline [CD]) and $188 (cefotaxim and doxycycline [CD]) and $188 (cefoxitin and doxycycline [CD]) and $188 (cefotaxim and doxycycline [CD]). The more expensive therapy decreases the relative risk of PID complications by 1% compared to the less expensive (due to greater effectiveness and/or adherence), then 1) OM compared to CD costs $38,300/QALY gained and 2) the more expensive therapy is cost saving if the difference in cost is $611, or costs $50,000/QALY if the cost difference is $585. A nearly linear relationship between greater decreases in relative PID complication risk and cost difference impact is noted (e.g., if the relative risk difference is 2%, cost savings occur if cost differences are $222).

Conclusion: The strategies “no screening” and “life-long screening of all patients under the age of 30, every N years” (N = 5, 7, 10) were compared using a Markov chain decision model. Hypothetical subgroups of patients at increased risk were defined, with relative risks (1.5, 4, 8) for the formation and rupture of aneurysms. A 5-year screening strategy due to four of recurrence screening always resulted in increased costs and a loss of quality-adjusted life years for patients with standard risk, making the “no screening” strategy dominant. Univariate sensitivity analyses indicated that the kind parameters were the incidence and rupture rate of new aneurysms. For conservative annual rates of rupture (0.4% for denovo and 2% for replacement aneurysm), screening strategies were not cost saving, not even for subgroups with a relative risk of 4.0. For annual rupture rates observed in the Netherlands (1.4% for denovo and 3.2% for replacement aneurysm), a 5-year screening strategy resulted in cost savings for patients with a relative risk exceeding 1.8. However, screening this subgroup did not result in an increase in the expected number of quality-adjusted life years. Assigning different utilities to healthy patients in the screening group and the “no screening” group had a profound impact on cost-effectiveness. A doubling exceeding 0.02, that is, accounting for being reassured or not, resulted in cost savings and increased survival for the 5-year screening strategy, already for the lower rupture rates. Cost-saving screening of these patients is not cost-effective in general, however, for subgroups of patients at increased risk, screening may be cost-effective. On patient-level, the degree to which patients are burdened by the knowledge of increased risk of recurrence predominately determines cost-effectiveness. Thus, a paradox appears: not informing patients of their increased risk leaves screening redundant, whereas informing patients may make screening highly preferred.

Conclusion: Within the cost range of CDC-recommended antibiotic regimens for outpatient PID treatment, use of more expensive antibiotics is economically reasonable if relatively small decreases in PID complication rates occur due to greater treatment effectiveness or adherence. Further work is needed to investigate differences in antibiotic effectiveness for modifying PID complication risk.
ABSTRACTS

2037
POSTER SESSION I
7:00 AM - 8:30 AM
Saturday, October 22, 2005
17 Pacific Room EFG
THE EFFECTIVENESS AND COSTS OF REDUCING LOSS TO FOLLOW-UP IN A SOUTH AFRICAN CERVICAL CANCER SCREENING STUDY
Goldhaber-Fiebert J, Deny L, De Souza M, Wright T and Goldie S
Harvard University, Cambridge, MA; University of Cape Town, Cape Town, South Africa; College of Physicians and Surgeons of Columbia University, New York, NY; Harvard School of Public Health, Cambridge, MA
Purpose: Cervical cancer disproportionately affects women in developing countries largely due to the absence of screening programs. Unless screening and treatment occur in one visit, loss to follow-up between visits reduces a program's effectiveness. This study was designed to examine the resources used in reestablishing contact with women who missed their scheduled visits and to assess the success of this effort in reducing loss to follow-up.
Methods: Women were enrolled in the screening study between 2000 and 2003, and all had scheduled 6-, 12-, or 24-month follow-up visits in 2003. The proportion of community health worker (CHW) time, vehicle use, maintenance, and depreciation spent reestablishing contact with women who had missed their appointments was estimated from weekly logs and cost accounting systems. The percentage of women who attended their scheduled visit, those who attended after CHW contact(s), and those who never returned despite CHW contact(s) were determined for each follow-up visit type. Number of CHW visits per woman by visit type was also estimated.
Results: 3,711 visits were scheduled in 2003. Of these, 2,321 (62.5%) occurred without CHW contact(s), 990 (26.4%) occurred after CHW contact(s), and 472 (12.7%) did not occur despite CHW contact(s). Loss to follow-up was reduced from 23% to 6%, 39% to 10%, and 50% to 24% for 6-, 12-, and 24-month visits, respectively. CHWs attempted 2,800 contacts in 580 trips. On average, 3 CHWs attempted to contact 6 patients over each 111-minute trip. The percent patient cost (2003 Rand) for these activities was 12.7%, 24.9%, and 40.5% for 6-, 12-, and 24-month visits respectively.
Conclusions: CHW costs for reestablishing contact were significantly increased their return rate. With more time between visits, the success of this effort decreased and the per-patient costs increased. Program costs such as these should be incorporated in cost-effectiveness analyses aimed at policy decisions about cervical cancer screening in developing countries.

2109
POSTER SESSION I
7:00 AM - 8:30 AM
Saturday, October 22, 2005
18 Pacific Room EFG
CHARACTERIZING STRUCTURAL UNCERTAINTY IN DECISION ANALYTIC MODELS: REVIEW AND APPLICATION OF AVAILABLE METHODS
Cimneliy L and Claxton K
University of York, York, United Kingdom
Purpose: To systematically review methods to characterize structural uncertainty, identify those that are applicable to decision analytic modeling, and demonstrate their applicability using case study models exhibiting different forms of structural uncertainty. We then examined how uncertainty about parameter values impacts model outcomes.
Methods: We reviewed the literature to classify types of structural uncertainty in decision analytic models and identify methods to characterize these uncertainties. A systematic review of the literature was conducted on methods to characterize structural uncertainty in decision models. Search terms included structural uncertainty, sensitivity analysis, and uncertainty characterization. Search limits included human health care. Extracted data included method type, literature source, and a model example demonstrating the method.
Results: The review identified methods for characterizing structural uncertainty; however, only a subset were applicable to decision analytic models. These are: 1) model averaging where alternative models with different specifications, are built, and their results averaged; 2) computing results for each alternative model specification and presenting these alternative results as scenario analyses; and 3) parameterizing the uncertainty directly in the model by including “uncertain” parameters. Running scenarios and model averaging were undertaken for all 4 case studies. Parameterizing the structural uncertainty was only undertaken in those models where it was technically feasible. ARMD and QIC.
In some cases, structural uncertainty had a limited impact on cost-effectiveness, for example, in the ARMD model, the ICER varied from ≤ 12,892 to ≤ 16,15% for the 2040 subgroup, and in none, it had serious impact on cost-effectiveness, for example, in the choice of the most cost-effective strategy changes (47%) or 6.7% compared to 7.7% (47% ± 5.7% 06). However, the value of additional research (ARMD) was particularly sensitive to structural uncertainty, in particular, model averaging and parameterizing uncertainty, because it introduced additional uncertainties and increased the ICER.
Neither running scenarios nor model averaging helped to inform a decision about the importance of more evident structural uncertainty with cost-effectiveness in the model in can inform this decision but unfortunately cannot be undertaken for more complex structural uncertainties such as adding additional parameterizations.
Conclusions: It is clear that alternative structural assumptions produce very different values of additional research estimates, and it is therefore essential, for decision-making purposes, to incorporate issues of structural uncertainty into the decision-modeling process.

2155
POSTER SESSION I
7:00 AM - 8:30 AM
Saturday, October 22, 2005
19 Pacific Room EFG
CONSIDERING INTERVENTION ADHERENCE AND PERSISTENCE IN COST-EFFECTIVENESS MODELS
Hoeger T, Hicks K and Sorensen S
RTI International, Research Triangle Park, NC; Centers for Disease Control and Prevention, Atlanta, GA
Background: The effectiveness of an intervention can be significantly undermined when patients do not fully participate in its essential components; at the same time, intervention costs may be reduced. For cost-effectiveness models to provide results that are meaningful and useful, they must consider patient adherence and persistence. The issue of “adherence” refers to patients only partially participating in an intervention (e.g., not taking medications at the full prescribed dosage). “Persistence” refers to how long patients continue in an intervention before stopping participation altogether.
Purpose: To outline methods for simulating intervention adherence and persistence in cost-effectiveness models and determine the impact of both on cost-effectiveness results.
Methods: We identified several relevant approaches for modeling a lack of adherence or persistence in interventions in cost-effectiveness models. The identified methods are then applied to a Markov model simulating two interventions meant to delay development of diabetes in patients with pre-diabetes—one higher effectiveness lifestyle-based and another lower effectiveness medication-based. We simulate reduced adherence or persistence for a cohort using relevant parameters for colorectal cancer screening tests. Surprisingly, we found that many patients do not prefer colorectal cancer screening testing. The costs of colorectal cancer screening tests. Surprisingly, we found that many patients do not prefer colorectal cancer screening tests.
Results: The following two approaches are relevant for modeling reduced adherence: (A1) Intervention costs and effects reduced proportionally; (A2) Full intervention costs continue to be incurred, but effects reduced proportionally. The following two approaches are relevant for modeling reduced persistence: (P1) Intervention costs and effects stop; (P2) Intervention costs remain, but effects stop. Using our diabetes development model, we found that applying method A2 to the lifestyle intervention, varying adherence between 50% and 80% resulted in ICERs from twice to six times that of the baseline ($12,14). Applying method P1 to the medication intervention, varying persistence over time between 50% and 85%, resulted in CERs all within 1 to 4 percentage points less than the baseline ($121,26).
Conclusions: The effect of considering reduced adherence and persistence in cost-effectiveness models can vary significantly, depending on the way that each applies to the intervention. Based on the potential impact of such factors, cost-effectiveness models should consider adherence in sensitivity analyses, if not baseline analyses.
REVIEW OF DECISION-ANALYTIC MODELS IN CONGESTIVE HEART FAILURE

Gohler A1, Kuntz K1, and Siebert U1

1Charité Campus Virchow Klinikum, Humboldt Medical School, Berlin, Germany; Massachusetts General Hospital, Harvard Medical School, Boston, MA

Purpose: Congestive heart failure (CHF) is a major cause of morbidity and mortality in the popula-

tion, and the societal burden of CHF is likely to escalate in the next decades. Several research groups

have developed decision-analytic models to investigate the long-term clinical effectiveness and

cost-effectiveness of several interventions. We sought to give an overview on published deci-

sion-analytic models and methodological approaches evaluating health technologies in CHF

to derive general recommendations for future comprehensive CHF decision models.

Methods: We performed a systematic literature review to identify studies that evaluated diag-

nostic, therapeutic, and disease management procedures for CHF using mathematical decision

models. Using a standardized assessment form, information on the study design, methodological

framework, and data sources were extracted from each publication and systematically reported.

Results: We identified 13 studies that used mathematical models to evaluate different pharma-

coeutical and reconditioning treatment options in CHF. Models evaluating diagnostic work-up or
disease management strategies for CHF were not identified. All identified models included clinical

economic outcomes. Modeling approaches comprised mathematical equations and Markov

models, with a time horizon ranging from one year to lifetime. Treatment effects were modeled by

lumping disease progression, which was either represented by New York Heart Association (NYHA)

stages or by the number of repeat hospital admissions. The influence of different etiologies of CHF

was not considered in the course of the disease in any model. Only one study included quality-adjusted

life years as an outcome, and no study reported external validation results.

Conclusion: Well-elaborated decision models are available for CHF treatment but are lacking for
diagnostic, patient management, and the impact of implementing disease management programs.
Future comprehensive, genetic, and flexible decision models should link diagnostic and thera-

peutic options, allow the evaluation of multiple outcomes and quality of life, integrate different endophases, and be validated with independent data.

ABSTRACTS
POTENTIAL ERRORS OF THE FRAMINGHAM RISK FUNCTION FOR CHD IN TYPE 1 DIABETES: IDENTIFYING THE GAPS
Zeller J, Ruppert K, Orchard T and Roberts M
University of Pittsburgh, Pittsburgh, PA
Purpose: Previous analyses have demonstrated that CHD risk prediction models underestimate probability of a CHD event in those with type 1 diabetes (T1D). This underestimate can alter the modeling of important outcomes in decision models of diabetes interventions. Our objective was to examine risk factors that may account for underestimated predictions.
Methods: Data were from the Pittsburgh Epidemiology of Diabetes Complications Study (EDC), a prospective cohort study of 658 subjects with childhood onset T1D, diagnosed between 1950 and 1980. Baseline exams took place 1966-1986 (mean age 28, diabetes duration 19 years). The cohort has been followed biennially since. The Framingham Risk Engine equation was applied to the EDC data to generate the probability of risk for MI or CHD death. Probabilities were then split in to deciles. Expected and observed events were compared using the Hosmer and Lemeshow goodness-of-fit statistic.
Results: The cohort experienced 123 CHD events during the follow-up period, with a mean age at onset of 40. When observed and expected probabilities were compared, there was significant lack of calibration (p < 0.001). The chi-square comparing observed to expected probabilities was highest in the top three deciles. Framingham risk factors (age, smoking, cholesterol/HDL ratio, systolic blood pressure) and other risk factors previously found to be associated with CHD in T1D were compared within the top three deciles and within the other seven deciles comparing CHD incident cases to non-cases.
Women in the top three deciles, who had an event, had significantly higher estimated glucose disposal rate (eGDR - calculated using a formula derived from euglycemic-hyperinsulinemic clamp studies), fibrinogen, white blood cell count (WBC), albumin excetration rate (AER), waist/hip ratios and Beck Depression Inventory Score at baseline than those who did not have an event. For those women in the lower seven deciles, fibrinogen and WBC were significantly different. Men from the top three risk deciles had a higher HbA1c, AER, fibrinogen, WBC, and eGDR. Those with a history of antihypertensive use also had more events. In the lower seven deciles, only AER was significantly different.
Conclusion: The Framingham risk equation does not adequately predict the probability of a CHD event in T1D. Risk factors including renal disease, WBC, and insulin resistance, not considered by the risk equation, may account for the lack of fit.
The affective and cognitive component of risk perception: decision-making in the context of prenatal screening

Van den Berg N and Timmermans D

VU University Medical Center, Amsterdam, Netherlands

This study aimed to investigate whether a cognitive and an affective component of risk perception could be identified. Furthermore, the role of risk perception in the process of making a decision to accept or decline prenatal screening was assessed. In a sample of 1658 pregnant women, risk perception, child-related anxiety, age, and the intention to have a prenatal test done were measured using postal questionnaires. Risk perception was measured both on a verbal scale (range: very small to very large) and on a numerical scale (range: 1 in 50 000 to 1 in 10). It was hypothesized that the verbal question reflects an affective component and the numerical question reflects a cognitive component. So, the verbal question should correlate with an affective variable, and the numerical risk perception question should correlate with a cognitive variable. Child-related anxiety was used as the affective variable, and age was used as the cognitive variable (age is a proxy for the knowledge of the actual risk of having a child with Down syndrome, as this risk increases with age). Furthermore, a risk perception scale (combining the verbal and numerical question) was calculated. The role of risk perception in the decision-making process was evaluated by assessing the correlation between risk perception and intention to test. For analyses of bivariate correlations, Pearson’s coefficient was used.

Numerical risk perception was more strongly correlated to age than to anxiety (0.46 vs. 0.16), and verbal risk perception was more strongly correlated to anxiety than to age (0.32 vs. 0.24). Consequently, principal component analysis revealed two components: a cognitive component (numerical risk perception and age) and an affective component (verbal risk perception and child-related anxiety). The correlation between the risk perception and intention to test was very small (0.10). All correlation coefficients were statistically significant.

From these findings it can be concluded that the perception of a risk includes both a cognitive and an affective component. This implicates that risk communication and counseling should not focus only on the correct understanding of risk figures but also on the affective reaction to those risks. Furthermore, as the correlation between risk perception and the intention to have the test done was very small, the role of risk perception in this screening decision seemed to be limited.
ABSTRACTS

E34 • MEDICAL DECISION MAKING/JAN–FEB 2006

1909
POSTER SESSION I
7:00 AM - 8:30 AM
Saturday, October 22, 2005
35 Pacific Room EFG
RECRUITING RESPONDENTS FOR A STUDY INVOLVING COMPLICATED TRADEOFFS
Verschuuren M and Van Hout B
University Medical Center Utrecht, Utrecht, Netherlands

Purpose: While piloting for a study involving complicated tradeoff questions, we were confronted with the fact that many respondents were not able to understand the task they were asked to perform. Therefore we decided to adopt our recruitment procedure to try to select those respondents with adequate cognitive abilities for performing the tradeoff task.

Methods: In our initial design, consecutive patients from an outpatient clinic for heart failure were invited to participate after they had visited their physician, and a sample of elderly persons from the general population was to be addressed by means of a short invitation letter. In our adapted recruitment procedure, all respondents were invited to participate through a detailed information letter including an example of a tradeoff question. The tradeoff comprehended a choice between two patients, for whom only 1 treatment was available. If respondents did not want to make a choice, they could “dose lots.” All respondents were asked to return the answering form, and if they were not willing to participate, they were requested to explain why.

Results: While piloting, we found that 12 out of 35 interviews failed because the respondents were not able to (completely) grasp the tradeoff task. In total, 337 people were invited according to the adapted recruitment procedure (preliminary response data). 186 returned the answering letter (55%), of whom 116 were willing to participate (62% of responders). 108 people were actually interviewed, of whom 102 (94%) successfully. 5 respondents did not understand the questions, and 1 refused to answer them. Of the 71 responding people who did not wish to participate, 24 (34%) stated that this was because they did not want to take a stand in the problem they were confronted with, while 3 (4%) declared that their inability to understand the question was the reason.

Conclusions: It is possible to efficiently select respondents who have adequate cognitive abilities for performing a complicated, abstract tradeoff task by means of an invitation letter with an example of a tradeoff. However, as a consequence of this recruitment procedure, we found people who do not want to make a tradeoff to be underrepresented in our study population.

1968
POSTER SESSION I
7:00 AM - 8:30 AM
Saturday, October 22, 2005
34 Pacific Room EFG
FROM COGNITION TO THE SYSTEM: A TAXONOMY OF PATIENT SAFETY IN FAMILY MEDICINE
Kontopoulos O
University of Birmingham, Birmingham, United Kingdom

Purpose: To develop a taxonomy of patient-safety events in family medicine based on a theoretical model of human cognition.

Methods: Observations, staff interviews, task analysis, and confidential reporting of patient-safety incidents in five clinics. Fault Trees were used to analyse the reports. Taxonomy development was an iterative process, both theory and data driven.

Results: There are three levels of classification: 1) Cognitive domain and psychological mechanisms. The information processing framework is used to classify errors according to the “cognitive domain” that failed: perception, situation assessment/response selection, memory, response execution. The psychological mechanisms aim to explain how the cognitive domain failed: expectancy, loss of attention, interference, false assumption, cognitive biases. 2) Immediate internal causes. These are affective and cognitive states that influence cognition through the psychological mechanisms: lack of knowledge, alertness/fatigue, curvilinear/hysteresis work approach, inflexible application of procedures, stress/procrastination. 3) Performance-Shaping Factors (PSFs). These are conditions that predispose to error and are divided into organizational and technical, each with a number of subcategories.

The taxonomy was used to classify 77 reported events. A tendency not to report self-made errors, generally found in reporting schemes, made it difficult to extract cognitively related information. The cognitive domain was identified in 46 reports. Most errors related to situation assessment (45%), often caused by false assumptions (64%). Response execution failures (24%) were most often due to interference/confusion (81%). Perception failures (15%) were most often due to expectations of certainty. A memory failures involved forgetting intentions (45%) and were all due to losses of attention. PSFs were almost always mentioned in the reports: task demands (37%), task fragmentation (24%), communication (12%), information presentation (11%), and information availability/delays (9%). Fragmentation and task demands were most often associated with errors in situation assessment and memory. Time pressures and information presentation were associated with perceptual errors, expectation errors happened in situations of cue similarity or proximity.

Conclusions: The taxonomy’s comprehensiveness and reliability remain to be tested. Results suggest that task fragmentation, excessive task demands, and the availability and presentation of information are important factors in safe clinical practice. Their effects on cognition can be identified and addressed across events in a database that maintains the causal links between the individual, the task, and the environment.
ABSTRACTS

2146
POSTER SESSION I
7:00 AM - 8:30 AM
Saturday, October 22, 2005
38 Pacific: Room EFG

DEVELOPMENT OF A GRAPHIC TOOL TO ASSESS THE INFLUENCE OF HISTORICAL DATA IN BAYESIAN ESTIMATION OF DIAGNOSTIC TEST PARAMETERS

Guillaumon S, Rabilfoud M and Ecchord R
Hopices Civils de Lyon, Lyon, France

Purpose: Decision making relies often partially on diagnostic tests interpretation. Unfortunately, it is well admitted that there is no perfect reference test. Bayesian approach for latent class analysis has been one of several statistical methods proposed to estimate test parameters such as sensitivity, specificity, and population prevalence as well as the sensitivities and specificities of the two tests. These two parameters are compared. The available data set, confirmed by simulations, shows that an unbalanced data set could lead to one test-parameter estimation that is very sensitive to prior information whereas others are not.

Conclusions: This original approach shows to be a useful and easy-to-handle tool to assess the weight of historical data on Bayesian estimation of diagnostic test performances. This tool is sensitive enough to highlight the parameters that would be moderately to strongly influenced by historical data. Critical informative prior distributions could then be easily identified and carefully chosen using literature review and experts' knowledge.

ABSTRACTS

2176
POSTER SESSION I
7:00 AM - 8:30 AM
Saturday, October 22, 2005
40 Pacific: Room EFG

PREFERENCES FOR COLORECTAL CANCER SCREENING AMONG LOW LITERACY AND MINORITY PRIMARY CARE PATIENTS

Hawley S, Vulk R, Jibaja-Weiss M, Vernon S and Katz S
University of Michigan, Ann Arbor VA Health System, Ann Arbor, MI; Baylor College of Medicine, Houston, TX; University of Texas, School of Public Health, Houston, TX

Background: Although colorectal cancer is preventable through regular screening, utilization rates for each of the four recommended modalities are low. A better understanding of preference variation for existing colorectal cancer screening (CRCS) tests is needed to increase patient compliance and address the first 3 stages of CA. We conducted in-depth interviews with 74 patients (25 white, 27 African American, 22 Hispanic), and found that accuracy, preparation, frequency, comfort, and cost were the top 5 attributes relating to CRCS. These data were used to create and pilot-test an initial CA assessment instrument. Phase II addresses the final 2 stages of CA and involves determination of the utilities (i.e., “part-worths”) related to specific attribute/level combinations obtained from rating and ranking of attributes by respondents. Preliminary analysis entails assessing the proportion of respondents willing to trade off between different attribute/level combinations.

Results: Phase II results indicate that patients trade off attributes of CRCS presented in CA scenario format. The leading tradeoff was a willingness to pay for a more accurate test but was modified by patient education. Among pilot interviewees with individuals with “some college or more” 86% were willing to pay more money for a more accurate test, while 55% of individuals with “high school or lower” were willing to make this tradeoff but were not willing to pay more than $25. Fifty percent of respondents were willing to trade “some discomfort” for a more accurate test, and 75% were willing to trade “some discomfort” for a less frequent test. The extent of these tradeoffs varied among racial/ethnic groups.

Conclusions: Using CA to assess preferences for CRCS provides a good understanding of how attribute/level combinations are valued among low-literacy and minority primary care patients. Further evaluation of the relationship between CRCS preferences and patient characteristics will be important in tailoring interventions designed to increase compliance with CRCS in diverse populations.
To assess the utility of a decision aid for colorectal cancer screening interest among Hispanic patients.

Purpose:
A randomized control trial was conducted among 120 Hispanic patients who were eligible for colorectal cancer screening at the Texas Tech Northeast Family Practice Clinic in El Paso. All participants provided an initial interest rating for colorectal cancer screening and completed a demographic questionnaire, a knowledge questionnaire, an accrual measure, and two patient autonomy measures. Participants read an American Cancer Society pamphlet and supplemental information on colorectal cancer screening. Participants were then randomized into one of two experimental arms. In the control condition, participants provided an interest rating about colorectal cancer screening after reading the materials. In the experimental condition, participants used a decision aid based on the Analytic Hierarchy Process prior to providing the colorectal cancer screening interest rating. Lastly, all participants completed the decisional conflict scale and the post-knowledge questionnaire.

Results:
Hierarchical multiple regressions were performed to predict interest scores, decisional conflict scores, and post-knowledge scores. After controlling for participant sex, age, location of the experimental site (home versus clinic), insurance, education, initial preference in screening, acculturation, and their ability to learn useful information as a result of using the decision aid and post-knowledge scores. Participants positively evaluated the decision aid with respect to its ease of understanding, their ability to learn useful information as a result of using the decision aid and post-knowledge scores. Participants positively evaluated the decision aid with respect to its ease of understanding, their ability to learn useful information as a result of using the decision aid.

Discussion:
Among US Hispanics, colorectal cancer is the second leading cause of cancer-related mortality. Understanding the utility of decision aids with Hispanic populations meets the goals of the Healthy People 2010 objectives, which call for the communication of health information to increase health literacy in traditionally underserved populations. These findings demonstrate that a decision aid promotes interest in screening and has favorable evaluations.

ABSTRACTS

2290
POSTER SESSION 1
7:00 AM - 8:30 AM
Saturday, October 22, 2005
42 Pacific Room EFG
DECISION AIDS AND COLORECTAL CANCER SCREENING INTEREST DECISIONS AMONG HISPANIC PATIENTS
Moraes O1, Kim L2, Fernandez N3, Urquidi U2, Gomez Y1, de la Torre M1 and Dolan J1
1University of Texas at El Paso, El Paso, TX; 2Texas Tech University Health Science Center; 3El Paso campus, El Paso, TX; 4Unity Health System, Rochester, NY
Purpose: To assess the utility of a decision aid for colorectal cancer screening interest among Hispanic patients.

Methods:
A randomized control trial was conducted among 120 Hispanic patients who were eligible for colorectal cancer screening at the Texas Tech Northeast Family Practice Clinic in El Paso. All participants provided an initial interest rating for colorectal cancer screening and completed a demographic questionnaire, a knowledge questionnaire, an accrual measure, and two patient autonomy measures. Participants read an American Cancer Society pamphlet and supplemental information on colorectal cancer screening. Participants were then randomized into one of two experimental arms. In the control condition, participants provided an interest rating about colorectal cancer screening after reading the materials. In the experimental condition, participants used a decision aid based on the Analytic Hierarchy Process prior to providing the colorectal cancer screening interest rating. Lastly, all participants completed the decisional conflict scale and the post-knowledge questionnaire.

Results:
Hierarchical multiple regressions were performed to predict interest scores, decisional conflict scores, and post-knowledge scores. After controlling for participant sex, age, location of the experimental site (home versus clinic), insurance, education, initial preference in screening, acculturation, and their ability to learn useful information as a result of using the decision aid and post-knowledge scores. Participants positively evaluated the decision aid with respect to its ease of understanding, their ability to learn useful information as a result of using the decision aid and post-knowledge scores. Participants positively evaluated the decision aid with respect to its ease of understanding, their ability to learn useful information as a result of using the decision aid.

Discussion:
Among US Hispanics, colorectal cancer is the second leading cause of cancer-related mortality. Understanding the utility of decision aids with Hispanic populations meets the goals of the Healthy People 2010 objectives, which call for the communication of health information to increase health literacy in traditionally underserved populations. These findings demonstrate that a decision aid promotes interest in screening and has favorable evaluations.
ABSTRACTS

2270

POSTER SESSION I
7:00 AM - 8:30 AM
Saturday, October 22, 2005
46 Pacific: Room EFG
DEFINING THE CLINICAL BENEFIT OF PET SCANNING IN THE MANAGEMENT OF PATIENTS WITH NON-SMALL CELL LUNG-CANCER

Kreuter, W.F.; Leatham R and Cardwell C
Queen's University Belfast, Belfast, United Kingdom

Purpose: Positron emission tomography has gained an established place in the work-up of patients with Non-Small Cell Lung Cancer (NSCLC). The Scottish Health Technology Board Report (2002) suggests its use could only be supported on cost-effectiveness grounds, for patients whose CT scans showed no hilar node spread. This conclusion was sensitive to assumptions about patient utilities for PET-detected health states.

Methods: Patients referred for the determination of surgical candidacy were recruited and interviewed by a research nurse, prior to PET. Using a series of cards that described the experiences of patients receiving specific forms of treatment and their outcomes, patients were asked to indicate, on a visual analogue scale, how they valued the avoidance of a needless thoracotomy, when compared to experiencing surgery or other treatments with or without the possibility of cure, states contingent upon the accuracy of their PET scan results. WTP performance status and quality of life were evaluated, the latter using the EDITS-Lung instrument.

Results: For our initial 25 patients, the VAS score (with 95% confidence limits) for being a “false positive” (when the PET result might deny potentially curative surgery) was 12 (4, 21), but the difference between this state and the VAS scores for the “false negative state” (that would subject them to futile thoracotomy) was not significant (24 [12, 57]). The VAS scores for being “true negative” and “true positive” were 90 (84, 97) and 72 (60, 84), respectively.

Conclusions: The only two randomized trials of PET-guided management have produced contrasting results, one showing a significant reduction in rates of futile thoracotomy and the other showing no significant difference. Our results suggest that directly eliciting patient values of possible PET outcomes would be worth incorporating into future cost-effectiveness analyses, but future work should also assess the patient’s valuations of these states after actually experiencing the outcomes.

1817

POSTER SESSION I
7:00 AM - 8:30 AM
Saturday, October 22, 2005
47 Pacific: Room EFG
INCONSISTENCIES BETWEEN VAS AND WTP FOR OWN HEALTH AND HYPOTHETICAL SCENARIOS

Byrne M; Seouchek J, Richardson M and Suarez-Almazor M
University of Miami, Miami, FL; Baylor College of Medicine/Michael E. DeBakey VA Medical Center, Houston, TX

Purpose: To examine inconsistencies between how people rate their own health and health scenarios using visual analog scales and how much they are willing to pay to improve their own health and the health scenarios.

Methods: We sampled equal numbers of Whites, Hispanics, and African-Americans from two populations: adults living in Houston, Texas (N = 193), and patients treated for knee osteoarthritis (OA) (N = 198) in Houston. Face-to-face interviews were used to collect demographic and preference data. We elicited utilities for patients’ own health and for two hypothetical scenarios depicting moderate and severe OA using visual analog scale (VAS) and willingness to pay (WTP) methodologies.

Results: Black men liked graphs more than white men for the bar graph. Stated liking (or not) for graphs was unassociated with accuracy.

Conclusion: Accuracy of interpretation is positively associated with accuracy of interpretation. Race has less impact, but Black subjects figure graph was most frequently understood, perhaps due to single concept presented. Education did better on the most complicated (bar) graph. This observed difference is not due to education.

1818

POSTER SESSION I
7:00 AM - 8:30 AM
Saturday, October 22, 2005
45 Pacific: Room EFG
ACCUACY OF GRAPH INTERPRETATION IN A DECISION-AID

Rover D; Pyle J, Willis C, Lillie J, Kelly-Blake K and Holmes-Rovner M
Michian State University, East Lansing, MI; home, E. Lansing, MI; Michian State University, E. Lansing, MI

Purpose: Informed patient decision making is the central goal of patient decision support interventions, but research is needed on optimal presentation of risk and outcome data. We examined how men 50 years or older (N = 186) interpret data from graphical presentations in a state-of-the-art videodisc decision aid about benign prostatic hypertrophy (BPH) treatment.

Method: Quasi-experimental descriptive design. Pre/post knowledge and decision survey reported previously. Audio-taped “think-aloud” semi-structured interview during and following viewing of videodisc. Following viewing, men reviewed paper copies of three graphs sampled from the video (bar, line, and stick figure graphs) and asked, “What does this graph mean to you, if anything?” and “What did you like about the graph?”

Results: There were differences in the accuracy of answers for the three types of graphs (P<0.05). Stick (N: bar graph 33.9 (177), line graph 50.3 (169), stick figure graph 75.8 (165)). College-educated men were more accurate for each graph, largest difference in bar graph. Black men were more accurate than white men for the bar graph. Stated liking (or not) for graphs was unassociated with accuracy.

Conclusion: Accuracy varied with seeming complexity of graph: bar > line > stick figure. Stick figure graph was most frequently understood, perhaps due to single concept presented. Race has less impact, but Black subjects did better on the most complicated [bar] graph. This observed difference is not due to education. Presenting risk comparisons is essential to informed decision making but appears to be compromised by available graph formats.

ABSTRACTS

E37

© 2006 Society for Medical Decision Making. All rights reserved. Not for commercial use or unauthorized distribution.

Downloaded from http://mdm.sagepub.com on guest on June 6, 2007.
Purpose: Over two-thirds of US infant homicide is attributable to inflicted neurotrauma (IN). Diagnosis of IN is challenging, and misdiagnosis leads to increased morbidity and mortality. The purpose of this study is to estimate the cost-effectiveness of cranial computed tomography (CT) to screen for IN.

Methods: We used decision analysis to compare immediate CT or discharge home from an emergency department in an asymptomatic 5-week-old with either 1) a history of an apparent life-threatening event or 2) unexplained scalp bruising. Research supports a prevalence of IN in such infants of 2.4% and 24%, respectively. We examined three clinical scenarios from three perspectives: a case-finding model to estimate costs per case found, a Markov model from a payer perspective, and a Markov model from a societal perspective. Health states modeled were no injury, undiagnosed IN, mild IN, severe IN, and death. Infants were entered into the model with no IN or mild IN according to prevalence. Probabilities of IN-related disability or death were linked to CT detection. We used available literature to estimate model parameters associated with medical, legal, and welfare costs and probabilities. Informed assumptions were made regarding quality of life in IN survivors. The model terminated at death or at 52 weeks of age to limit prediction of long-term costs and utilities of child abuse, which are poorly understood.

Results: Case finding was insensitive for both high- and low-prevalence scenarios ($813 and $1047 per case, respectively). From a payer perspective, screening under conditions of high prevalence saves money ($1088) and improves outcomes (0.009/QALY); for conditions of low prevalence, screening is expensive ($136,367/QALY) but sensitive to variation in CT cost and IN utilities. From a societal perspective however, screening is uniformly expensive ($125,630/QALY for high prevalence and $360,419/QALY for low prevalence) due to high costs of child protection following false detection.

Conclusion: From a payer perspective, CT screening for inflicted neurotrauma can be cost saving. From a societal perspective, screening appears expensive due to high costs of caring for survivors. The time-limited nature of abuse makes future benefit of future screening uncertain. Our models emphasize the challenge of selecting appropriate time horizons in pediatric cost-effectiveness analysis and the importance of improved understanding of costs and outcomes of child abuse.
VALIDATION OF THE ECONOMIC ASSESSMENT OF GLYCEMIC CONTROL AND LONG-TERM EFFECTS (EAGLE) DIABETES MODEL

Walter S, Maxion-Bergemann S, Bergemann R, Casciano R and Mueller E

Analytica International, Lorch, Germany

Purpose: To internally validate the EAGLE (Economic Assessment of Glycemic control and Long-term Effects) model version 2.0 according to current standards of validation practice. The EAGLE model simulates long-term diabetes-related complications and related costs for type 1 and type 2 diabetes. Published results from large interventional studies (DCCT, WESDR, and UKPDS) serve as the basis for EAGLE calculations.

Methods: First-order validation identified inconsistencies in results and corrected programming errors. Second-order validation included the following steps: 1) Simulation sets were created in EAGLE based on baseline data from the studies used to build the model; 2) Simulations were run. The results obtained with EAGLE were compared with published event rates. 3) Risk equations were refined if a deviation of >20% was observed between model-derived and published results. Patient numbers and iterations were systematically changed until a final run was performed with 50,000 patients and 180 iterations.

RESULTS: Initial validation was done for the analysis of the long-term diabetes-related complications and related costs in type 1 and type 2 diabetes patients. The EAGLE model is a valid and robust tool for the validation of the long-term diabetes-related complications and related costs in type 1 and type 2 diabetes.

CONCLUSIONS: The EAGLE model consistently predicts event rates reported by UKPDS, WESDR, and DCCT, for both type 1 and type 2 diabetes patients. The EAGLE model is a valid and robust tool for the validation of the long-term diabetes-related complications and related costs in type 1 and type 2 diabetes.
Method: We conducted an electronic search of five databases (Medline, Embase, Cinahl, CINAHL, and PsycINFO) from 1965 to 2004 to identify relevant studies. Two reviewers independently assessed eligibility, extracted data, and rated study quality. We used a random-effects meta-analysis to synthesize the data.

Results: Twenty-two studies were included in the meta-analysis. The estimated changes in absolute Unified Parkinson’s Disease Rating Scale scores after surgery were calculated using random-effects models. Sources of heterogeneity were explored with meta-regression models, and the possibility of publication bias was evaluated.

Conclusion: The sensitivity and specificity of calcaneal quantitative ultrasound at commonly used thresholds are too low to exclude or confirm DXA-determined osteoporosis. Until further evidence becomes available, it is premature to recommend use of calcaneal quantitative ultrasound in an evidence-based screening program for osteoporosis.
ABSTRACTS

2395

JOINT ISOOQL POSTER
3:30 PM - 5:00 PM
Saturday, October 22, 2005
1 Bayview Room

PRODUCTIVITY AND QUALITY OF LIFE ASSESSMENT: AN EMPIRICAL TEST OF A MEASUREMENT CONTROVERSY

Lavigne J1 and Sohmer M1

1University of Rochester School of Medicine and Dentistry, Rochester, NY; RAND Corporation, Pittsburgh, PA.

Purpose: Uncertainty over whether productivity effects are entirely subsumed in quality of life (QOL) measures has led to different standards for measurement in Europe and the US. We use empirical data to explain differences in QOL results when productivity and income are included versus explicitly included in QOL assessments.

Methods: The time tradeoff method (TTO) and rating scales were used to assess QOL through a series of scenarios (1) without mention of productivity or income effects, (2) explicitly stating expected productivity effects associated with the QOL scenario, and (3) explicitly stating both the productivity and income effects likely to be associated with the health state described. Respondents answered all three scenarios for each of two disease states: type II diabetes and migraine headache. Baseline scenarios specifying symptom severity and associated productivity and income estimates were drawn from the published literature. To control for any systematic bias resulting from the order of the scenarios, respondents were randomly assigned to one of three survey versions, each with a different ordering of the scenarios. Analyses included paired t-tests for differences between means and descriptive statistics.

Sample: A convenience sample of 88 faculty, staff, and students at the University of North Carolina at Chapel Hill and the University of Rochester completed the surveys.

Results: TTO results were significantly lower when productivity effects were explicitly stated in the TTO scenarios (0.80 vs 0.84, p < 0.05). When income productivity effects were included in the scenario, adding information about likely effects of symptoms on earned income did not change the TTO results significantly. In the case of migraine, however, the TTO did not change when expected productivity and income effects were explicitly stated. 26% of respondents reported identical TTO values, regardless of the explicit statement of productivity goals.

Conclusions: Productivity effects of diabetes do not appear to be subsumed within QOL assessments using the TTO, suggesting that the European approach may be more accurate than the methodological differences in how productivity and income are treated in the US. This may explain differences in QOL results when productivity and income are included versus explicitly included in QOL assessments.

2226

JOINT ISOOQL POSTER
3:30 PM - 5:00 PM
Saturday, October 22, 2005
3 Bayview Room

USING LONGITUDINAL QUALITY OF LIFE PROFILES TO EVALUATE PROGNOSIS FOR SURGICAL SUCCESS

Kuykendall J1, Kallen M1 and Wray N2

1VAMC, Houston, TX; 2Baylor College of Medicine, Houston, TX.

Purpose: To understand prognosis for alleviation of pain measured 2 years post-op using quality of life data collected 2 weeks following surgery for knee osteoarthritis.

Methods: 180 patients with knee osteoarthrosis underwent 1 of 3 surgical procedures for relief of pain and to improve functional status. Of these, 133 completed all self-reports of pain intensity during the course of 2 years of follow-up. Growth mixture modeling revealed 4 classes of patients. For 2 of these classes, a “low” pain group and a “high” pain group, 2-week pain values (average scores of 11.0 and 37.6, respectively; 0 = low pain; 50 = high pain) were predictive of their 2-year pain values (13.7 and 32.8, respectively).

However, most patients undergoing surgery (73%) experienced moderate levels of knee pain measured 2 weeks post-op. For these patients, the prognosis over the next 2 years was characterized by 3 of 2 distinct profiles. Those who reported pain that became progressively worse versus those who reported pain that progressively improved. Although similar at 2 weeks, the groups had mean 2-year pain ratings of 35.9 and 11.0, respectively. Logistic regression revealed that patients’ 2-week post-op evaluations of non-pain symptoms and operation success were highly predictive of whether pain became progressively worse or improved over the subsequent 2 years. On a 0-10 scale, where 0 is indicative of being bothered by non-pain symptoms and perceptions of operation failure, a 1-point increase was associated with a 25% increase in the odds of being classified into the group experiencing pain relief at 2 years (p < 0.05).

Conclusions: Understanding prognosis helps guide medical decisions with regard to the need for future additional intervention. By using data collected at 2-weeks post-op, it is possible to identify a subset of patients for whom additional intervention might be considered. Understanding the physical and psychological determinants for judgments of non-pain symptoms and operation success at 2 weeks holds promise for explaining why measures taken at 2 weeks are predictive of outcomes 2 years later and for informing decisions about appropriateness of surgery and additional interventions.
MEASURING THE EFFECT OF ACUTE CYSTITIS AND ANTIBIOTIC TREATMENTS ON WOMEN’S QUALITY OF LIFE

Bepu G, Ernst E, Ernst M and Hoehns J
University of Iowa, Iowa City, IA

Purpose: Although women commonly experience acute cystitis, the impact of this infection and its treatment on quality of life (QOL) has not been described. The purpose of this study was to measure the effects of acute cystitis and antibiotic treatments on QOL.

Methods: One hundred and seven women with cystitis and symptoms of acute uncomplicated cystitis were recruited at two family medicine outpatient clinics. Patients were randomized to receive TMP/SMX for 3 days, ciprofloxacin for 3 days, or nitrofurantoin for 7 days as treatment for their infections. QOL was measured using the Quality of Well-Being scale, a validated multi-attribute single-item instrument, in which 1 represents perfect health and zero, death. Follow-up telephone contacts were completed at 1, 7, 14, and 28 days after the initial visit. QOL, clinical outcomes, and adverse events were assessed during these interviews.

Results: The mean QOL score improved for all patients, from a mean (SD) of 0.68 ± 0.03 at baseline to 0.81 ± 0.11 at day 28. Patients experiencing a clinical cure with the initial course of antibiotics had significantly better QOL at Day 3 (p < 0.04), Day 7 (p < 0.01), and Day 14 (p = 0.01) compared to patients who failed their initial antibiotic treatments. Clinical cure rates for the three different antibiotic regimens were similar (p = 0.7), and there was no difference in QOL, by treatment assignment at any of the follow-up times (p > 0.05 for each). However, patients experiencing any adverse event had significantly lower QOL on Day 3 (p < 0.001) and Day 7 (p < 0.001) than subjects who did not report adverse events.

Conclusions: Women experiencing acute cystitis have measurable gain in QOL in conjunction with treatment of these infections. However, experiencing any adverse event from treatment has a negative impact on this gain during the first week of treatment as it does treatment failure. As measured by the Quality of Well-Being scale, we did not find a significant difference between the 3 antibiotic treatments.

ABSTRACTS
E43

ABSTRACTS

2184

JOINT ISOQOL POSTER

3:30 PM - 5:00 PM

Saturday, October 22, 2005

10 Bayview Room

THE IMPACT OF RECALL TIME ON THE MEASUREMENT OF HEALTH UTILITIES FOR ACUTE EVENTS

Barnack N,2 Sun H,3 Li X,1 Goh D,1 Noary B,1 Schechter M,1 Singer J,1 Anta A1 and OPTIMA Team 2

1St. Paul’s Hospital, Vancouver, BC, Canada; 2University of British Columbia, Vancouver, BC, Canada

Purpose: We examined whether the length of recall time in generic health state utility instruments determined different valuations of health state utility in patients with advanced HIV/AIDS where acute adverse events are common.

Method: The OPTIMA study (an ongoing RCT of treatment strategies for patients with advanced HIV/AIDS with whom standard therapies have failed) measures the Health Utilities Index (HUI3) and the EQ-5D at 6 monthly intervals. While both instruments were administered at the same time, the HU3 used recall time of the past week, whereas the EQ-5D asked about health status on the day of administration. To identify whether an acute adverse event was captured by the utility instrument, we classified assessments into three groups: (i) when a non-AIDS related serious adverse event (SAE) was ongoing and unresolved (so the recall period was applicable for both the HUS and EQ-5D); (ii) when SAE had resolved within the past 7 days (recall period applicable only for HUI3); and (iii) when no SAE had occurred within the past 7 days. We fitted a Generalized Linear Model (using GEE) including the instrument type, the assessment time, and their interaction terms. The effect of recall time was estimated by examining the difference between the HUS and EQ-5D in assessing events resolved in the past 7 days, after negating the effect of instrument difference estimated by assessing events which were unresolved.

Results: Among 329 patients enrolled as of March 2005, there were 295 assessments identified as type (i) and 62 as type (ii). The adjusted effect of recall time was found to be statistically significant (HU3 vs. EQ-5D = -0.152, p = 0.048) and HUS vs. EQ-5D = -0.215, p = 0.001). As anticipated, the comparison between HUI2 and HUS, where recall time is identical, showed no statistically significant recall effect (p = 0.281).

Conclusions: This study demonstrates that patients accurately adhered to the length of recall requested. The instrument shorter recall (EQ-5D) did not capture the impact of all adverse events. Whether a longer recall period than 1 week would be more appropriate in this disease group needs further research since patients might reflect on their recovery rather than the event. Validation of these findings is warranted in a new study, which should use the same instrument but include multiple recall times of different lengths.

1964

JOINT ISOQOL POSTER

3:30 PM - 5:00 PM

Saturday, October 22, 2005

12 Bayview Room

TREATMENT OF DEMENTIA WITH STANDARDIZED GINKGO BILoba EXTRACT EGb 761 IN OUTPATIENT CARE: HOW MUCH DOES AN INCREASE IN CAREGIVERS’ AND PATIENTS’ QUALITY OF LIFE COST?

Koller M,2 Heinen-Kummerer T,1 Lorenz W,1 Von den Driesch V,1 Hahs M3 and Rychlik R

1Philippus-University Marburg, Marburg, Germany; 2IFSG, Burscheid, Germany; 3Tumor Center, Regensburg, Germany; 4Schwarz Pharmaeuticals, Karlsruhe, Germany

Purpose: Dementia is a progressive disease that involves tremendous burdens for patients, their families, and society as a whole. One treatment option with proven efficacy (Birks and Evans, 2004; Cochrane Review summarizing 36 RCTs) is the Ginkgo biloba extract EGb 746. The purpose of this study was to investigate the effectiveness (caregivers’ and patients’ quality of life [QLI] and efficiency (health care costs) of EGb 761 in the routine care of non-hospitalized dementia patients.

Methods: The trial design specified a non-randomized, open label, two-arm cohort study that followed up each patient for a period of 12 months. One cohort of patients received EGb 761 from the beginning of the study, the other cohort received any form of antideementia treatment except Ginkgo biloba. Patients were eligible for this trial when they suffered form mild to moderate dementia (MMSE-score 12 to 24 points) and received home care by one primary family member. Patients were recruited in 133 family doctor practices all over Germany. It was the doctors’ choice whether or not to treat the patient with EGb 761. Primary end point was the caring person’s QLI (PLC; Siegrist et al., 1996) 12 months after start of treatment. Patients’ cognitive (MMST) and functional (Barthel index) status were assessed. Total health care costs per patient per year were calculated including all direct costs. Results: A total of N = 683 patients were included in the study, n = 281 received EGb 761 (1 - 2 x 120 mg) and n = 402 did not. At the 12-month assessment point, the caring persons’ QLI was significantly higher in the Ginkgo biloba cohort than in the control cohort (all six PLC-scores, p < 0.001). In the same manner, patients’ cognitive (MMST) and functional (Barthel index) functioning were higher (p < 0.001). Average total cost per patient amounted to 3,031.78 € in the Ginkgo biloba cohort, and 3,614.75 € in the control cohort (p = 0.16).

Conclusions: This study demonstrated in the reality of family doctor patient care, Ginkgo biloba and EGb 761 positively affect the “system” of dementia patients. Better cognitive and functional status of the patients and a better QL of their caring partners were achieved without increasing the total costs for health care. © 2006 Society for Medical Decision Making. All rights reserved. Not for commercial use or unauthorized distribution.
J O I N T I S O Q O L P O S T E R
3:30 PM - 5:00 PM
Saturday, October 22, 2005
13 Bayview Room
EQUATING SCORES ON THE HEALTH AND ACTIVITY LIMITATION INDEX AND QUALITY OF WELL-BEING SCALE IN A POPULATION OF UNITED STATES ADULTS
Bundhoo E, Anderson E, Gillespie K and Rocke D
Washington University School of Medicine, St. Louis, MO; University of Florida College of Public Health and Health Professions, Gainesville, FL; St. Louis University School of Public Health, St. Louis, MO; Saint Louis University School of Public Health, St. Louis, MO
Purpose: The Health & Activity Limitation Index (HALEX) was introduced in 1995 to measure health-related quality of life (HRQoL) for Healthy People 2000 and has been used for population-tracking activities and in a few research studies. However, little is known about how HALEX scores relate to other measures of HRQoL. The purpose of this study was to equate scores on the HALEX with scores on an accepted measure of HRQoL, the Quality of Well-Being (QWB) scale.
Methods: We administered the HALEX and QWB questions to a sample of 401 adults in St. Louis, Missouri. We used random-digit-dial techniques to select a random sample of 302 adults from the St. Louis Metropolitan area and enriched this sample by recruiting 99 participants with mobility impairments. Inclusion of the mobility-impaired participants provided representation of a broader range of disability than would typically be found in a randomly selected population. We used multiple regression analysis to examine the relationship between HALEX and QWB scores and to see how demographic and health status factors influenced this relationship.
Results: In this sample, scores on the QWB and HALEX were similar on mid-range values and discrepant at the extremes, that is, persons with extreme HALEX scores tended to have more moderate QWB scores. For instance, participants with a score of 1.0 on the HALEX had a mean score of 0.76 on the QWB, and participants with a mean score of 6.0 on the HALEX had a mean score of 4.2 on the QWB. QWB scores alone accounted for 10% of the variance in HALEX scores, the addition of five demographic and health factors increased the variance explained to 63%. HALEX scores were higher for white adults than predicted by their QWB score, and HALEX scores for people with chronic disease and disabilities were lower than predicted.
Conclusion: HRQoL scores assigned by the HALEX and QWB to the same case can vary markedly, particularly if the individual is extremely healthy or unhealthy. Including information about a few demographic and health-related characteristics greatly improves our ability to predict HALEX scores from QWB scores, suggesting that there are systematic differences in how the two scales evaluate HRQoL that requires further investigation.

J O I N T I S O Q O L P O S T E R
3:30 PM - 5:00 PM
Saturday, October 22, 2005
14 Bayview Room
DOES CANCER HINDER THE WORK OF EMPLOYED SURVIVORS?
Bunkhaut K* and Bradley C
*Michigan State University, E, Lansing, MI; Virginia Commonwealth University, Richmond, VA
Purpose: To determine if cancer interferes with the work of employed patients following a cancer diagnosis.
Methods: An inception cohort of 417 working women and 252 men newly diagnosed with breast and prostate cancer, respectively, were identified and interviewed for a period corresponding to 3 months prior to, and 12 months following, diagnosis. Four hundred eight of the women and 233 of the men were also interviewed at 18 months following diagnosis. Patients were asked if their cancer interfered with their ability to perform various job tasks. Chi-square tests and logistic regression were used to determine if there was an association between job task interference and the ability of a patient to work at 12 and 18 months.
Results: At 12 months, 32% to 52% of breast cancer patients reported interference with various physical job tasks and 21% to 32% reported cancer interference with various cognitive job tasks. There were lower rates of reported cancer interference at 18 months. At 12 months, 22% to 32% of prostate cancer patients reported interference with various physical job tasks and 10% to 26% reported cancer interference with various cognitive job tasks. There were lower reported cancer interference rates for prostate cancer patients at 18 months. Women at 12 months were less likely to be working if they reported cancer interference with job tasks requiring physical effort (p < 0.001), stooping (p = 0.001), keeping up with others (p < 0.001), concentration (p < 0.04), data analysis (p < 0.02), and learning new things (p < 0.001). Only cancer interference with having to keep up with others was statistically significantly associated with not working at 12 months for prostate cancer patients (p = 0.03). In a multivariate logistic regression analysis, controlling for demographic factors, health status, and cancer treatment, cancer interference with having to keep up with others was the only statistically significant job task interference factor associated with a breast cancer patient not working at 12 months (95% CI 0.86-0.93). There were no statistically significant job task interference factors at 18 months for breast cancer patients.
Discussion: Cancer patients are less likely to work for employed breast and prostate cancer patients. It appears that this hindrance is transient. Employers should consider making adjustments for employees with cancer who are required to keep up with others on the job.

J O I N T I S O Q O L P O S T E R
3:30 PM - 5:00 PM
Saturday, October 22, 2005
15 Bayview Room
CONCURRENT VALIDITY OF THE HEALTH STATUS CLASSIFICATION SYSTEM FOR PRE-SCHOOL CHILDREN IN VERY-LOW-BIRTH-WEIGHT CHILDREN AT AGE 24 MONTHS
van Niekerkkoek K, Raat H* and Weisgag-Kuporov N
Erasmus MC-Sophia Children's Hospital, Rotterdam, Rotterdam, Netherlands; Erasmus MC, University Medical Center Rotterdam, Rotterdam, Netherlands
Introduction: Siaqal et al proposed a Health Status Classification System for Pre-School Children (HSCS-P.S) as a multi-dimensional measure to describe quality of life of, as well as to permit evaluation of preferences for health states of children as young as 2-5 years (Qual Life Res 2005;14:243-257). We evaluated construct validity of the HSCS-P.S in a population of very-low-birth-weight (VLBW) children at the corrected age of 24 months.
Methods: Parents of 67 VLBW children (gestational age <32 weeks) visiting the neonatal follow-up program at the Erasmus MC-Sophia Children's Hospital between June 2004 and March 2005 completed the HSCS-P.S that consists of 10 dimensions with 3-5 levels each. To evaluate construct validity of the HSCS-P.S, child development was assessed with the Bayley Scale of Infant Development (BSID-II), mental (MDI) and motor (PDI) development), the Child Behavior Checklist (CBCL) and the LEXI (based on the Language Development Survey).
Results: Mean corrected age was 2.6 years (SD = 0.06), mean gestational age was 29.7 weeks (SD = 1.5), mean birth weight was 1242 g (SD = 356), mean MDI = 96 (SD = 15), mean PDI = 91 (SD = 15), and mean LEXI quotient was 97 (SD = 14). High scores (level 3, 4, or 5) on the HSCS-P.S, indicating moderate or severe impairments, were found for BSID-P.S-dimensions Self-care (9.9%), Speech (9.3%), General-health (9.5%), Dexterity (3.5%), Mobility (3.4%), and Pain & Discomfort (2.4%); Thinking & Problem solving (2.0): Social & communication (1.2%); and Behavior (1.1%). Gestational age correlated with HSCS-P.S-dimensions Pain & discomfort (r = –0.33), MDI with Seeing (r = –0.27), Speech (r = –0.36), Mobility (r = –0.39). Dexterity (r = 0.44), Self-care (r = 0.54), Emotion (r = 0.27), Learning & Remembering (r = 0.43), and Thinking & Problem solving (r = –0.35). PDI with Mobility (r = 0.34) and Social (r = 0.37). LEXI-score with Speech (r = 0.40) (p < 0.01). CBCL-scores (internalizing and externalizing behavior and total score) correlated with Emotion, Learning & Remembering and Thinking & Problem solving (r = 0.28-0.40) (p < 0.01).
Discussion: In this population of well-documented VLBW children at age 24 months, the observed correlations between parental ratings of the 10 HSCS-P.S dimensions with gestational age as indicator of pregnancy-outcome as well as with measures of specific impairments support the construct validity of the HSCS-P.S. This study showed that in a clinical setting, even at the very young age, this classification system adequately describes health and neurodevelopment of VLBW-children.
Purpose: The aim of the study was to elicit societal utility values for different stages of advanced non small cell lung cancer including responding disease, stable disease contrasting intravenous (IV) and oral treatment, progressive disease without treatment, and an "end of life" health state.

Methods: Metastatic disease-based health state descriptions were adapted to describe metastatic lung cancer health states. Content validation was performed through interviews with oncologists (n = 5). Thematic saturation was achieved, and changes were made. Revised health states were then re-reviewed by the same oncologists and a psychometric expert, which resulted in further minor changes. Revised health states were then piloted in a chained standard gamble (SG) interview followed by cognitive debrief. In the next phase of the study, each health state was valued by members of the general public (n = 78) who completed an SG interview and a visual analog scale (VAS) rating. The health state “dead” was the lowest possible value fixed at 0. Quality of life (EQ-5D) and sociodemographic data were also collected. Data were summarized and investigated using analysis of variance.

Results: The study sample was relatively well-matched to the general public in English and Wales (FINS, 2001). Mean SG utility scores were 0.70 (responding disease), 0.63 (stable disease, oral treatment), 0.58 (stable disease, IV treatment), 0.42 (progressive disease), and 0.33 (end of life). SG values declined significantly from responding disease to end of life. (F = 52.14, p < 0.0001). Mean VAS scores were 84.8 (responding disease), 41.1 (stable disease, oral treatment), 54.3 (stable disease, IV treatment), 20.4 (progressive disease), and 11.8 (end of life). (F = 98.08, p < 0.0001).

Conclusions: This study captures the utility decrement associated with deteriorating metastatic lung cancer. These values could be used in a cost-utility analysis based on societal preferences. The study also identified a preference for oral therapy over IV therapy directly in each survey. The Health Utilities Index Mark 3 (HUI3) was asked directly in Coons et al. (2001) and the 2001 Medical Expenditures Panel Survey (MEPS). The EQ-5D was asked directly in each survey. The two surveys, 3482 completed the EQ-5D and 3536 completed the HUI3. Participants, 3482 completed the EQ-5D and 3536 completed the HUI3. Of 22,523 MEPS participants, 19,629 completed the EQ-5D and 19,853 completed the HUI3. Of 22,523 MEPS participants, 19,629 completed the EQ-5D and 19,853 completed the HUI3. The absence of utility values has prevented cost-utility analysis of treatments for opiate dependence. We assessed utilities in individuals with active opiate dependence disorder.

Method: The modified Addiction Severity Index (ASI) and six quality-of-life measures were used to assess quality of life of residents of the Tenderloin neighborhood in San Francisco who provided a history and objectives of current and sustained illicit opiate use. Individuals were excluded for inability to speak coherently, aggressiveness, objective evidence of being of actively intoxicated with alcohol or heroin, or suffering severe withdrawal.

Results: Participants (N = 449) reported that in the prior 10 days, they had used heroin on average of 26.2 days and cocaine on 15.8 days. Respondents reported a mean utility of 0.518 by Quality of Well Being (QWB) scale, 0.545 by Health Utilities Index (HUI), 0.760 by EuroQol (EQ-5D), 0.768 by Standard Gamble (SG), and 0.773 on the Time Tradeoff (TTO). Their mean value on the visual analog scale was 0.632. Values of indirect utility measures (the VAS, HUI, QWB, and EQ-5D) were significantly correlated. Only 3 of the 9 correlations that involved the direct measures, the TTO and SG, were significant. There were significant relationships between drug and mental health problems and utility. We regressed utility scores on the ASI problem drug score, the ASI psychiatric problem score, and an indicator of HIV-positive status. A higher psychiatric problem score was associated with significantly lower utility by 0.6 of the 6 measures. A higher drug problem score was associated with satisfyingly lower utility by EQ-5D and QWB and significantly lower utility by 0.9 of the 6 measures. A higher drug problem score was associated with significantly lower utility by the 6 measures. A higher drug problem score was associated with significantly lower utility by 0.9 of the 6 measures.

Conclusions: Economic evaluations of drug treatment have been largely limited to cost-benefit analysis, considering reduction in criminal justice costs but not the effect of treatment on improved quality of life or mortality. We found that utilities can be assessed in drug users and that opiate dependence has a severe impact on quality of life. It thus appears feasible to use cost-utility analysis to evaluate innovation, enhancement, or expansion of substance abuse treatment. This type of analysis will allow decision makers to compare the incremental cost per QALY of changes in substance abuse treatment in parity with the economic evaluations of medical care interventions.

Methods: A comprehensive search for stuttering instruments was conducted using 1) a literature review of PsychInfo, Medline, Social Science Citations Index, Current Contents, and Health Star; 2) Web search engines for the American Speech-Language Hearing Association and Google; 3) referenced stuttering instruments listed in articles; and 4) personal files. Instruments were included in the study for evaluation if there was at least one publication using the instrument including psychometric data, in addition to instrument availability. The eight domains considered in the instrument evaluation were item information, versatility, practicality, breadth (social, role, physical, mental, and communication) and depth (floor and ceiling effects), reliability (internal consistency and test-retest), validity (discriminative and convergent), and responsiveness.

Results: Of the 18 QOL instruments identified, adequate data were available to evaluate eight based on study criteria. The eight stuttering-specific QOL metrics were Stutterer's Self-Ratings of Reactions to Speech Situations (ASRS), Erickson's scale of communication attitudes (5-scale), shortened 5-scale (S-5), Communication Attitude Test-English and revised [CAT and CAT-S], Perceptions of Stuttering Inventory (PSI) Speech Situations Checklist (SSC), and its shortened version [Shortened-SSC]. Item information was only available for S-5-scale and S-24. All instruments evaluated were designed for use in adults except for A-19 and CAT series. PSI can be used in adults and adolescents. Only S-24 and the 5 scale satisfied breadth criteria. Depth information was lacking for all instruments evaluated. Of the eight instruments, S-24 fared the best being the most widely used and associated with a significant utility decrement. The 8 instruments (omitting depth and reliability), however, due to inadequate reliability (coefficients < 0.9), authors do not recommend S-24 for individual decision making.

Conclusions: None of the eight metrics satisfied all the study criteria, and as a result none can be recommended for use in individual decision making. However, if used for group-level decision making, the authors recommended the use of S-24, which would then meet most of the study criteria reporting adequate or better performance on 7 of 8 criteria assessed (omitting instrument depth), in conjunction with a generic QOL instrument.
DIFFERENT SHADES OF “WELL”: MODELING QUALITY OF LIFE (QOL) AS A FUNCTION OF PHYSICAL ACTIVITY LEVEL

Rous U, Pratt M, Yanagawa T, Yori M, Kaplan R and Teng T

Centers for Disease Control and Prevention, Vancouver, BC; Canada; 1Centers for Disease Control and Prevention, Atlanta, GA; Centers for Disease Control and Prevention, Orlando, FL; UCSF School of Medicine, 9500 Gilman Drive, La Jolla, CA; Milliman USA, Palm Desert, CA

Purpose: The impact of physical activity (PA) on Quality of Well-being (QWB) scores, and the effects of PA on Markov models considering specific diseases or all-cause mortality, were studied.

Methods: We combined national PA and quality of life (QOL) data to derive PA-related QWB scores. Well subjects (defined as free of five inactivity-related diseases modeled) were stratified across four PA levels according to public health recommendations. Three multiplicative regressions were fit. First, using QWB data from persons with disease, QWB was fit as a function of age, gender, and type of disease. Next, using data from those who were well, QOL was fit as a function of age, gender, and PA level. Finally, using data for all persons, QOL was again fit as a function of age, gender, and exercise level. The final two regressions were used to estimate QOL, in our disease-specific model, which simulated the effect of PA on disease, and the subsequent effect of disease on QWB, mortality, and medical costs. The third regression was used in an all-cause mortality model, which simulated the direct effect of PA on QWB, mortality, and medical costs. Both models were then used to estimate the gain in QALYs from select interventions that improved PA and compared them.

Results: Among the well, there was a strong relationship between QOL and PA level. For example, 45-year-old women who were highly active had a mean QOL of 0.816, whereas those who met PA recommendations, were irregularly active, or inactive, reported mean QOLs of 0.807, 0.796, and 0.791, respectively. In the disease-specific model, the average gain in QALYs from an evaluated community-wide campaign strategy was 14.816, while in the all-cause mortality model, the same intervention had an average gain in QALYs of 15.550.

Discussion: Well adults were found to have PA-associated differences in health and QOL. Models that take into account the positive effect of healthy behavior among the well may underestimate the effect of behavior change on QALYs. When QWB, mortality, and cost data are available by behavior level, designing “all-case” models that take into account the direct effect of health behaviors on outcomes, without simulating the intervening effect of disease, may yield similar and equally accurate results.

Results: We found 35 studies, reporting SF-36 scores of 27,853 patients. Mean age was 56.9 years for HD patients and 53.6 years for PD patients, which was significantly higher than the age of RTx patients (46.3 years). Prevalence of diabetes was 21% among HD, 16% among PD, and 7% among RTx patients. In general, unadjusted SF-36 scores were higher for RTx compared to dialysis patients. HD and PD patients’ SF-36 scores were not statistically significantly different. After adjustment, the significance of the difference between dialysis groups compared to RTx recipients dissipated for the Physical Functioning scale (unadjusted: HD 52.6, PD 53.2, RTx 74.8 vs. adjusted HD 63.7, PD 56.3, RTx 66.7). In addition, the significance of the difference between HD and RTx patients disappeared on these scales: the Role Physical (unadjusted HD 41.8, PD 52.5, RTx 68.3 vs. adjusted HD 43.0, PD 52.5, RTx 67.2), the Bodily Pain (unadjusted HD 66.9, PD 65.1, RTx 74.0 vs. adjusted HD 67.6, PD 65.1, RTx 73.6), and the Vitality scales (unadjusted: HD 42.1, PD 44.8, RTx 56.6 vs. adjusted: HD 48.1, PD 45.2, RTx 52.5). The non-significantly lower score of PD as compared to HD patients on the Role Physical scale became statistically significant.

Conclusions: Differences in age and diabetes as a comorbidity among hemodialysis, peritoneal dialysis, and renal transplantation patients partly explained the differences in SF-36 scores among these treatment groups. Differences in quality of life were not statistically different among these groups after adjustment for age and diabetes.
METHODS: We surveyed 456 adult patients with diabetes from clinics affiliated with a major academic medical center and a community hospital. We elicited patients’ utilities using time tradeoff questions in face-to-face interviews. We assessed utilities for intermediate and end-stage complication states and nine treatment states. Our treatment states included a comprehensive diabetes treatment (polypharmacy), a hypothetical daily pill that combines multiple agents, the polypill, a hypothetical daily pill that combines multiple agents, conventional glucose control, conventional BP control, intensive glucose control, and intensive BP control. Certainty equivalents for each lottery were elicited; subjects were paid their earnings in cash.

RESULTS: The health risk preferences instrument was framed in terms of drugs that could potentially treat the symptoms of interest. Angina or shortness of breath. The monetary risk preferences instrument consisted of twelve $5 lottery trials. Subjects were asked to respond to the choice scenarios supposing they were experiencing angina attacks twice weekly. Each health scenario involved a choice between Drug A, with an x% probability of eradicating their symptom and a (1-x)% probability of having no effect, and Drug B, which would reduce the frequency of the symptom’s occurrence by x%.

CONCLUSIONS: Seventy-nine patients who presented for stress testing participated in this preliminary study. Of these, thirty had experienced shortness of breath and/or angina. We found a significant difference between the risk preferences of subjects who answered the hypothetical scenarios and subjects who were responding to scenarios based on the frequency of their own symptoms. Subjects with angina or shortness of breath made more risk-averse choices, choosing the risky drug in 24.2% of the scenarios, while individuals who had experienced neither of these symptoms chose the risky drug in 46.99% of the scenarios (p < 0.01). As expected, there was no significant difference in the monetary risk preferences of the two groups of subjects (p = 0.44).

Conclusion: While risk preferences differ over health and monetary choices, they also depend upon whether the health choices subjects are asked to make are real or hypothetical symptoms.

ABSTRACTS
ABSTRACTS

2117
JOINT ISOQOL POSTER
3:30 PM - 5:00 PM
Saturday, October 22, 2005
29 Bayview Room
THE EFFECTS OF DIFFICULTY WITH PREFERENCE ASSESSMENT AND COGNITIVE IMPAIRMENT ON MEASURING PREFERENCES FOR CURRENT HEALTH
King I1, Swart J and Roberts M1
1VA Connecticut Healthcare System, West Haven, CT; 1University of Cincinnati, Cincinnati, OH; 1University of Pittsburgh, Pittsburgh, PA
Purpose: Patient preferences for health states can be measured with the standard gamble (SG), time tradeoff (TTO), visual analog scale (VAS), and willingness to pay (WTP). Preference testing can be cognitively demanding and may be affected by cognitive functioning. We measured preferences in patients with cerebral aneurysms, a population vulnerable to cognitive deficits related to aneurysm rupture or treatment.
Method: A cross-sectional cohort of neurosurgery clinic patients with cerebral aneurysms completed the SG, TTO, VAS, WTP and the Mini Mental State Exam (MMSE) during a face-to-face interview. We recorded instances when patients had difficulty understanding or completing preference assessment. We examined the relationship between preferences and cognitive impairment or difficulty with testing using the Mann-Whitney U test. Multivariate linear regression models examined the relationship between preferences and patient characteristics, aneurysm history, cognitive impairment, and difficulty with preference assessment.
Results: One hundred sixty-five patients completed all 5 instruments; their mean age was 54.2 years, 72% were women, 5.2% had a history of aneurysm rupture, 46.9% had previous aneurysm treatment, and 7% had MMSE scores <24, consistent with cognitive impairment. Twenty patients (12%) had difficulty with preference assessment according to the examiner, although those patients still completed all preference instruments. In the univariate analyses, patients with cognitive impairment had lower scores on the VAS (mean = 0.55 vs. 0.68, p = 0.043), and patients with assessment difficulties had lower values on the SG (mean = 0.55 vs. 0.68, p = 0.001), and WTP ($199,100 vs. $304,700, p = 0.087). Regression models showed that cognitive impairment was associated with lower preferences measured with the SG (beta = –0.23, p = 0.084), TTO, (beta = –0.16, p = 0.067), and VAS (beta = –0.19, p = 0.032), and that difficulty with preference assessment was independently associated with lower preferences measured with the SG (beta = –0.17, p = 0.094), TTO, (beta = –0.17, p = 0.097), and WTP (beta = $136,400, p = 0.012). WTP preferences were also independently associated with income (beta = 2.1, p = 0.001).
Conclusions: Cognitive impairment and difficulty with preference assessment are independently associated with lower preference values in patients with cerebral aneurysms. Cognitive changes may affect preferences directly and/or may induce downward bias during preference assessment. Difficulty understanding or completing the preference elicitation procedures may also produce a downward bias in preference values. Utility assessments may want to incorporate both formal cognitive testing and interviewer judgments regarding subjects’ comprehension.

2133
JOINT ISOQOL POSTER
3:30 PM - 5:00 PM
Saturday, October 22, 2005
30 Bayview Room
CLOSE TO 1.0: OUR EXPERIENCE MEASURING UTILITY IN UPPER EXTREMITY PATHOLOGY
McCabe S1, Lafay A2, Myers J1, Tegoukis A1 and Goodwin A1
1University of Louisville, Louisville, KY; 2School of Public Health and Information Sciences, University of Louisville, Louisville, KY
Purpose: Clinical problems of the upper extremity are common and cause significant disability. In spite of this, there is very little published literature with utility measurement in patients with upper extremity pathology or injury. Most of these problems will have utilities close to 1.0, which may present difficulties with utility measurement. We present our overall experience measuring utilities in upper extremity problems.
Methods: We measured the utility using the visual analogue scale (VAS) and a variety of forms of the standard gamble (SG) in surrogates and patients with upper extremity problems including carpal tunnel syndrome (CTS), palmar tunnel syndrome, and fracture of the distal radius. We have measured the relationship of disease severity to utility by comparing scores with health status measures such as the carpal tunnel disease specific instrument and the Disabilities of the Arm Shoulder and Hand (DASH) where appropriate.
Results: Using the SG, we found the utilities associated with upper extremity pathology to lie between 0.9 and 1.0 on a scale of 0 to 1. For example, the average utility of CTS was found to be 0.97 and the most severe form of median nerve dysfunction 0.93. The VAS consistently yielded lower utility ratings than the SG. Patients most often provided higher utility ratings than surrogates. While the SG and VAS utilities both correlated with symptom severity measures in CTS patients, only the VAS correlated with a disability score in basal joint arthritis patients. Most patients who had experienced a distal radius fracture were not willing to gamble, providing utility estimates of 1.0 with the SG but 0.808 with the VAS. The chained version of the paper SG was found to be a useful method to expand the upper end of the utility scale to discriminate within a pathologic entity.
Conclusion: Through our experiences at measuring utility close to 1.0, we recommend the chained version of the paper form of the SG an efficient method to perform these measurements in conjunction with the VAS in a clinical population or in surrogates. Our results and experience will help the clinical researcher to embark on utility measurement for conditions with utilities close to 1.0 and are directly applicable to decision analysis and economic analysis for upper extremity pathology.

2403
JOINT ISOQOL POSTER
3:30 PM - 5:00 PM
Saturday, October 22, 2005
31 Bayview Room
STABILITY OF ADL UTILITIES AS HEALTH STATUS FLUCTUATES OVER TIME
Sims T1, Garber A1 and Goldstein M1
1Stanford University, Stanford, CA; 2Center for Primary Care and Outcomes Research, Stanford, CA
Older adults may have substantial changes in health over time. We examine the stability of older adults’ utilities for hypothetical health states of dependency in activities of daily living (ADL) over a year as a function of change in health status.
Methods: One hundred sixty-five adults (mean age = 75, 65% female, 35% non-white) reported utilities at Time 1 and one year later (Time 2) for current health (CH) and for hypothetical health states of dependency in 6 ADLs, a combination of all 6 ADLs (All6), a combination of 3 ADLs while living in a nursing home (NH) or their own home (OH), and walking. Total number of medical conditions was added at Times 1 and 2. Difference scores for baseline and follow-up utilities were computed.
Results: Participants reported more medical conditions (worse health), 44 reported fewer medical conditions (improved health), and 31 reported no change. Time 2 CH utilities were 0.84 for no change, 0.81 for improved health, and 0.77 for worse health. Time 2 mean utilities for single ADL states ranged from 0.81 (Bathing) to 0.79 (Toileting) and were 0.52 for All6, 0.53 for NH, 0.66 for OH, and 0.73 for Walking. Mean differences ranged from 0.005 (Bathing) to 0.111 (Walking) and were 0.045 for All6, 0.015 for NH, 0.043 for OH, and 0.027 for Walking.
Conclusion: Analyses of variance revealed significant effects for the lowest rated ADLs, Toileting, F(2,122) = 2.93, p < 0.05, and Walking, F(2,121) = 3.43, p < 0.05. Significant effects were also found for All6, F(2,121) = 4.125, p < 0.05; NH, F(2,120) = 5.40, p < 0.01; and OH, F(2,121) = 4.13, p < 0.05. Taken together these results revealed no difference between no change and improved health. Worse health showed a significantly greater difference over time than for All6, NH, and OH, p < 0.05 and a trend toward greater difference for Toileting, p = 0.064. Improved health differences were significantly greater than worse health for Walking, p < 0.05.
Conclusion: Change in health over time appears to affect older adults’ health preferences when rating even hypothetical health states described identically over time. Worseened health status may color an individual’s valuation of hypothetical health states. Policy analyses drawing on utilities should ensure that a broad range of health status is included in the population when deriving the utilities. In individual medical decision-making tasks, health professionals should ensure that patients have a clear grasp of the outcome scenarios that is not unduly influenced by their own recent changes in health.
Breast cancer patients' preferences for local and systemic therapy

Objective: To determine the predictors of breast cancer patients' (BCPs) willingness to accept local and systemic therapy.

Methods: A cross-sectional survey of female BCPs ages 36-80 at the University of Maryland Greenebaum Cancer Center, Baltimore, MD. Since "treatment" is considered a "short-term" health state, the chained procedure for the time tradeoff was used to assess tradeoff values. Willingness to accept therapy was determined using "minimum years cancer-free to accept therapy" as the dependent variable. The number of years was calculated based upon BCPs' time tradeoff responses for mastectomy (MM), breast-conserving therapy (BCT), chemotherapy (CTX), and tamoxifen (TAM). Independent variables included demographic factors and clinical data that were abstracted from the BCPs' medical records. Tobit regression models were used to determine the association between the dependent and independent variables because the dependent variable, minimum cancer-free years (CFYs) to accept therapy, was right-skewed for all therapies.

Results: The mean age for the 77 respondents was 56 ± 10 years (SD 18.43). 58.2% identified themselves as white; 75.6% had early-stage cancer. BCPs required more cancer-free years (median of 4) to accept MRM than to accept BCT, CTX, or TAM (median 1 year for each). For all forms of therapy, the mode was 0, suggesting that BCPs were willing to accept therapy even if it provided no guaranteed survival benefit. Late-stage patients require more CFYs to accept therapy (Median 63 ± 1, p = 0.0496; similarly, late-stage patients in good physical health required more CFYs to have MRM (β = 0.59, p = 0.0322). BCPs who had experience with a particular therapy were more willing to accept that type of therapy than those who were treatment-naive. Younger patients versus those aged 64+ required fewer CFYs to accept chemotherapy (age group 56-59, β = 4.77; p = 0.0943; age group 55-59, β = 7.23; p = 0.0192). Being white and having less education were associated with fewer CFYs to accept CTX (β = 3.86; p = 0.0087; β = 5.30; p = 0.0193, respectively).

Conclusions: BCPs require relatively few years of additional CFYs to accept treatment. Willingness to accept fewer CFYs to accept chemotherapy among those with less education and of younger age is consistent with treatment patterns previously published. The fact that non-whites (primarily African American) required fewer CFYs to accept CTX appears in stark contrast to observed undertreatment of chemotherapy among African American breast cancer patients.

ABSTRACTS
ABSTRACTS

POSTER SESSION II
7:00 AM - 8:30 AM
Sunday, October 23, 2005
2 Pacific Room EFG
TRADEOFFS BETWEEN HIP FRACTURE MORTALITY AND ORAL CORTICOSTEROID USAGES AMONG WHITE MEN AND WOMEN
Bar S, Paihelt A, Fuhlbrigge A and Kunz T
Harvard University, Boston, MA; Yale School of Medicine, New Haven, CT; Brigham and Women's Hospital, Harvard Medical School, Boston, MA; Harvard School of Public Health, Boston, MA
Purpose: Oral corticosteroid therapy increases a patient's subsequent risk for hip fracture. Although the absolute risk among white women is higher than white men, hip fracture mortality is higher among white men. We sought to quantify the tradeoffs between hip fracture incidence and mortality in both white men and white women being treated with oral corticosteroids.
Methods: We developed a Markov model to compare hip fracture outcomes over a lifetime for 50-year-old patients undergoing oral corticosteroid therapy vs. comparable patients not on therapy. Annual hip fracture rates for untreated patients were derived from the literature and were a function of age and gender. The annual risk of hip fractures among men and women receiving oral corticosteroids (1.5-10 mg/day) was estimated from the literature. Hip fracture-specific mortality rates depended on the age of fracture and the time since fracture. The first-year fracture-specific mortality rate among males/females was 0.15/0.03 for hip fractures occurring before age 75, 0.24/0.12 for hip fractures occurring at ages 75-84, and 0.41/0.23 for hip fractures occurring at ages 85 or older. Hip fracture-specific mortality rates beyond the first year were similar for men and women, although generally higher for women. We used the US life tables for other-cause mortality.
Results: We projected that out of 1,000 50-year-old persons at average risk for hip fracture, 97 men and 212 women will experience a hip fracture during their lifetime, and 53 men and 86 women will die of their hip fracture. The increased number of hip fractures resulting from taking high-dose oral corticosteroids was 54 for men and 78 for women. While the induced hip fracture incidence burden is 82% greater for white women compared with white men, the induced hip fracture mortality burden is only 46% greater.
Conclusion: Focus on hip fracture incidence as the primary side effect associated with oral corticosteroids may obscure the relative burden on white men caused by fracture-related mortality. Formal incorporation of the downstream risks shows that the differences between white men and white women in terms of hip fracture outcomes are diminished for hip fracture death.

POSTER SESSION II
7:00 AM - 8:30 AM
Sunday, October 23, 2005
3 Pacific Room EFG
EVALUATING AN INDIVIDUALIZED INFORMATION DECISION SUPPORT INTERVENTION IN URO-ONCOLOGY
Davison B
University of British Columbia, Vancouver, BC, Canada
Purpose: This study was conducted to determine if an individualized information intervention was more effective than a generic approach to provide decisional support to men newly diagnosed with prostate cancer.
Methods: A total of 324 men newly diagnosed with localized prostate cancer were recruited from community urology clinics. Interviews were conducted within two weeks following the medical consultation. The following scales were completed prior to randomization and after a treatment decision was made: Control Preferences Scale (computerized version), Satisfaction with Information and Decision Making, and Decisional Conflict. Men in the generic information group watched a video on early-stage prostate cancer, while the intervention group had a search nurse to provide disease-specific information. This session also included a discussion of the printouts of each information need rated more than 50% in importance was utilized by a registered nurse to provide disease-specific information. This session also included a discussion of the

Results: The difference in satisfaction between the two groups was not considered clinically significant. Therefore, provision of a video and written information package is suggested as an appropriate and cost-effective method of providing information decision support to these men and their families at the time of diagnosis.

Conclusion: The difference in satisfaction between the two groups was not considered clinically significant. Provision of a video and written information package is suggested as an appropriate and cost-effective method of providing information decision support to these men and their families at the time of diagnosis.

POSTER SESSION II
7:00 AM - 8:30 AM
Sunday, October 23, 2005
4 Pacific Room EFG
DECISION ANALYSIS OF THE EFFECTIVENESS OF CONTRACEPTION BY IMMEDIATE VERSUS DELAYED INSERTION OF INTRAUTERINE DEVICES AFTER ABDOMINION
Reaves M, Smith K and Creinin M
University of Pittsburgh School of Medicine, Pittsburgh, PA
Purpose: Immediate post-abortal intrauterine device (IUD) insertion has been shown to be safe and effective, yet it is not commonly available. The rate of IUD expulsion after immediate insertion has been shown not to be substantially higher than with delayed insertion.
Methods: We performed a decision analysis to examine pregnancy rates at one year following abortion with immediate or delayed IUD insertion. The base case assumed an 80% 1-year IUD continuation rate for both groups, an additional 10% expulsion risk with immediate IUD insertion, a 30% risk for not returning for delayed insertion, and pregnancy rates of 8% without an IUD and 6.5% with an IUD. Pregnancy between abortion and delayed IUD insertion was not considered. One-way, two-way, and Monte Carlo sensitivity analyses were performed.
Results: In the decision model, the 1-month pregnancy rate was 21 per 1000 women in the immediate-insertion group and 35 per 1000 women in the delayed-insertion group. Sensitivity analyses show the model to be dependent on the expulsion rate in the immediate-insertion group (varied up to 50%) and the rate of not returning in the delayed-insertion group. In one-way analysis, the overall pregnancy rate was higher in the delayed-insertion group along as the pregnancy rate in the women without IUDs was above 8.5%. In two-way analysis, immediate insertion was preferred if the rate of IUD expulsion was lower than the rate of not returning for delayed IUD insertion. Monte Carlo analysis showed that immediate insertion results in fewer pregnancies in 99.4% of scenarios, but it has an absolute mean difference of 16 pregnancies per 1000 women.
Discussion: Despite assumptions in the model favoring delayed insertion, immediate post-abortal IUD insertion results in 14 to 16 fewer pregnancies per 1000 women than delayed insertion in the following 12 months. Future research should examine cost-effectiveness and non-pregnancy outcomes.

POSTER SESSION II
7:00 AM - 8:30 AM
Sunday, October 23, 2005
5 Pacific Room EFG
A COST-EFFECTIVENESS ANALYSES BASED ON THE ASCUS AND LSIL TRIAGE STUDY (ALTS)
Harvard School of Public Health, Cambridge, MA; Duke University, Durham, NC; Agency for Healthcare Research and Quality, Rockville, MD; Georgetown University Medical Center and Cancer Clinical and Economic Outcomes Core, Washington, DC; National Cancer Institute, NIH DBHS, Rockville, MD
Purpose: Findings from the ASCUS and LSIL Triage Study (ALTS) suggest that for women with ASCUS Pap test results, human papillomavirus (HPV) DNA testing is effective at detecting underlying high-grade precancerous lesions or cancer (CIN3+). While policy analyses have been conducted to explore the long-term cost-effectiveness of such a recommendation, a cost-effectiveness analysis of the trial based on the primary data has not yet been performed. In this study, we used the ALTS trial data to determine the short-term cost-effectiveness of alternative strategies for the management of ASCUS.
Methods: Data from the ALTS trial were used in conjunction with medical care costs in a short-term decision model. The model compared the incremental costs per case of detected CIN3+ for five ASCUS triage strategies employed in the three arms of the trial: 1) immediate colposcopy (KC); 2) HPV DNA testing; and 3) conservative management (CM), involving a program of up to 1, 2, or 3 repeat cytology visits.
Results: The observed sensitivity of colposcopy and biopsy for detecting CIN3+ ranged from 46% to 91% in the trial. Using these data in the base case model, the lowest strategy cost was CM with 3 repeat cytology for referral to colposcopy. Triage to colposcopy based on a positive HPV DNA test ran an incremental cost-effectiveness ratio of $3,514 per CIN3+ detected compared to CM with 1 repeat cytology visit. Strategies of KC and CM with up to 3 repeat cytology visits were dominated by HPV DNA testing. HPV DNA triage had a lower cost-effectiveness ratio in women 30 years and older than in younger women ($2,917 vs. $3,806 per CIN3+ detected, compared to CM). Immediate colposcopy only became non-dominated when colposcopy and biopsy were assumed to be perfectly sensitive and specific, and cost $20,370 per CIN3+ detected, compared to HPV DNA triage.
Discussion: These results support that HPV DNA testing is an economically viable strategy for treating ASCUS cytology, consistent with findings from other policy analyses. The analysis also underscores the need to account for the suboptimal performance of colposcopy and biopsy in future clinical practice guidelines and policy analyses.
ABSTRACTS

POSTER SESSION II
7:00 AM - 8:30 AM
Sunday, October 23, 2005
6 Pacific Room EFG
EXTENDING THE ACCESSIBILITY OF MARKOV MODELS IN HEALTH TECHNOLOGY ASSESSMENT USING VISUALIZATION TOOLS
Pill M
Peninsula Medical School, Exeter EX2 5DQ, United Kingdom
Introduction: State-transition Markov models have become the normative approach to health technology assessment (HTA) in the UK. However, the growing sophistication of these methods has promoted the need to increase accessibility and to extend understanding of outputs beyond expert decision analysts. In this context, there are obvious benefits in exploiting the increased graphical and interactive capabilities of computers as well as the growing research literature in information visualization. The success of visual interactive modeling in discrete event simulation is testimony to the impact that graphical interfaces can bring to model accessibility.

Methods: A case study of HTA modeling in dual-chamber heart pacemakers was used to identify areas where visualization could be usefully implemented. This study used a Markov model implemented in Microsoft Excel to assess cost utility for a range of scenarios of care. One key requirement was to explore sensitivity to parameters: for ranges of model inputs and parameters for a number of different patient subgroups. A “sensitivity mixing desk™” was implemented enabling interactive control for an array of key model inputs using graphical slider controls. These were varied independently and the effects on the main model outputs observed visually in real time. In addition, a prototype animation was developed that provides a graphical depiction of the dynamics of the Markov model and allows users to control and interactively browse patient flows within the state-transition framework.

Results: Early evaluations of the prototype visualization tools in this study are promising. Specifically, the direct feedback of outputs based on interactive manipulation of model inputs given by the “sensitivity mixing desk™” is useful in promoting understanding. Users report a strong impression gained of the relative influence of input variables through the use of this application. The Markov animation tool has also proved useful in demonstrating model structure and dynamics to non-expert users in a direct and accessible graphical format.

Conclusions: These results are a considerable step for graphical tools to improve accessibility and promote understanding of Markov models in health technology assessment. The requirement for visualizing tools is increasing as the sophistication of model outputs grows and the requirement for understanding extends beyond communities of modeling and decision analytic experts. Sensitivity analysis is a key area where graphical outputs can be of specific importance.

POSTER SESSION II
7:00 AM - 8:30 AM
Sunday, October 23, 2005
7 Pacific Room EFG
PHARMACOGENOMICS AS A MECHANISM TO IMPROVE CARDIOVASCULAR OUTCOMES: SCREENING FOR THE ALPHA-ADDUCIN VARIANT
Meckley L1 and Veenstra D
University of Washington, Seattle, WA
Purpose: Studies have shown that many hypertensive patients are not taking diuretics despite being a recommended first-line therapy. Hypertensive patients with the Gly460Trp variant in the alpha-adducin gene may be more likely to benefit from diuretic therapy. The purpose of this study was to evaluate whether screening for the Gly460Trp alpha-adducin gene variant could be a cost-effective method of preventing stroke and myocardial infarction.

Methods: A Markov decision analytic model was developed for three strategies in a hypothetical cohort of non-diuretic treated hypertensive: 1) testing for the Gly460Trp variant to identify patients for addition of a diuretic, 2) no testing for the variant and no addition of diuretic (usual care), and 3) addition of a diuretic regardless of genotype. We used a lifetime horizon and payer perspective. Cost, utility, and clinical data were obtained from the literature. One-way, probabilistic, and sensitivity analyses were conducted to evaluate the uncertainty in the results. Assumptions: 70% of patients who screen positive for the alpha-adducin variant will have a diuretic added to their therapy; no patient would be switched from their current antihypertensive medication(s) to a diuretic, but instead a diuretic would be added to their current therapy to bias the results against screening; and the addition of a diuretic to antihypertensive therapy for patients with the wild-type alpha-adducin variant has a neutral effect.

Results: The testing strategy increased quality-adjusted life years (QALYs) by 0.07 (95% confidence interval range [CI]: 0.00, 0.17) and saved $5,172 (95% CI: $2,109, $8,921) compared to usual care. The most influential inputs were the strength of the association between the alpha-adducin gene variant and diuretic effect, the cost of the screening test, and the probability of adding a diuretic to hypertensive medication regimen. The strategy of adding diuretics regardless of genotype resulted in equivalent QALYs and saved $54 (95% CI: $116, $305) compared to testing.

Conclusions: Our results suggest that screening for the alpha-adducin gene variant may be a useful mechanism to identify patients most likely to benefit from diuretic therapy and improve compliance with current treatment guidelines. Additional studies documenting the association between the alpha-adducin gene variant and diuretic effect, the cost of the screening test, and the probability of adding a diuretic to hypertensive medication regimen are needed. The results of this study have implications for the future of pharmacogenomic testing.

POSTER SESSION II
7:00 AM - 8:30 AM
Sunday, October 23, 2005
9 Pacific Room EFG
ANASTRAZOLE FOR BREAST CANCER CHEMOPREVENTION: AN ECONOMIC ANALYSIS
Cukier S1 and Phifer N2
University of North Carolina, Greensboro, NC; 1 Moses Cone Hospital, Greensboro, NC
Purpose: The Breast Cancer Prevention Trial demonstrated that tamoxifen (tam) effectively reduces breast cancer risk in women. However, tam’s side-effect profile has severely constrained its use. For tam, contralateral breast cancer rates have accurately reflected its primary prevention effect. A recent breast cancer report (N = 6186) demonstrated that anastrozole reduces cancer occurrence in the contralateral breast 44%, more than tam and has fewer adverse effects. This report from the Arimidex, Tamoxifen Alone or in Combination Trials’ Group (ATAC), involved 16 months of follow-up and suggests that anastrozole is a good option for primary breast cancer prevention. However, anastrozole is a potent aromatase inhibitor and might be limited by its high price. Therefore, the purpose of our study is to evaluate the probable cost-effectiveness of anastrozole for primary, chemoprevention.

Methods: We performed decision analysis using a Markov model. Cost-effectiveness is reported from the payer’s perspective. Patient characteristics and the annual rates of breast cancer, stroke, venous thromboembolic disease, endometrial cancer, and fractures were derived from the Breast Cancer Prevention Trial of the National Surgical Adjunctive Breast and Bowel P-1 Project (NSABBP) and the ATAC group. Utility scores were measured by questioning 106 women. Sensitivity analysis was performed.

Results: Using the 3.4%, 5-year average breast cancer risk in the NSABBP combined with contralateral breast cancer data from ATAC, we determined that anastrozole increases quality-adjusted life (QALY) compared to no treatment for high-risk women aged 50 to 65 years. The cost per QALY gained ranged from $131,110 for the women aged 50-55 to $73,900 at age 65. Sensitivity analysis revealed that the utility score for breast cancer, the probability of stroke, and the cost of anastrozole influence cost-effectiveness. Incremental QALY gains of anastrozole over tam were obtained at the cost of $18,620 per QALY at age 50 and $41,700 at age 65. This model was only sensitive to the price of anastrozole, utility of breast cancer, and breast cancer risk reduction of anastrozole relative to tam. If the anastrozole price (current cost $225/month) is lowered to $128/month for 50-year-olds, the anastrozole chemoprevention strategy would be dominant over tam.
POSTER SESSION II
7:00 AM - 8:30 AM
Sunday, October 23, 2005
10 Pacific Room EFG
PHARMACOECONOMIC ANALYSIS: EXPLORATORY ESTIMATION OF THE COST-EFFECTIVENESS OF RITUXIMAB IN THE TREATMENT OF RHEUMATOID ARTHRITIS IN CANADA
Kanulee
McMaster University, Hamilton, ON, Canada
Purpose: The aim of this explorative analysis was to estimate the incremental cost-effectiveness of a treatment strategy that involves adding rituximab as another option to the sequence of disease-modifying anti-rheumatic drugs (DMARDs) for 10,000 (hypothetical) patients with rheumatoid arthritis (RA), over a 10-year time horizon. Rituximab is not currently licensed for treatment of RA in Canada but is already marketed for other indications.
Methods: Two DMARD strategies were compared—with and without rituximab. A Markov model using patient cohorts was used to simulate patients through different treatment regimens for both strategies using Microsoft Excel, reflecting the course of RA disease progression. Clinical data (ACR20 response rates not adjusted for placebo responses, drug withdrawal rates, and mortality data) were obtained from literature searches. Response data on rituximab + methotrexate as observed during the first 6 months of a phase II trial (Edwards et al., NEMD 2004) was used for the model. Drug costs were obtained from published list prices. Monitoring costs were calculated from recommendations provided in product monographs for each drug, in conjunction with procedure costs provided by the Ontario Schedule of Benefits for the year 2004. Costs and outcomes were discounted by 5% annually in accordance with provincial guidelines, and univariate sensitivity analyses were conducted to validate the results.
Results: In all scenarios analyzed, the incremental cost-effectiveness ratio (ICER) was in the range of $5,000-$17,000 per additional person-year of ACR20 response. Sensitivity analyses showed that the model was most sensitive to the rituximab re-treatment rate.
Conclusion: Based on this explorative model, the additional use of using rituximab is expected to be in a range that would be considered cost-effective in the Canadian health care setting. This analysis assumed that the indirect effects of rituximab extend clinical benefit (a reasonable cost) when incorporated into a therapeutic management plan for patients with RA. This initial pharmacoeconomic model should be expanded to include long-term data when such data become available in the future.
The effectiveness of initial endoscopy versus empirical treatment for dyspepsia was carried out. This cal-
tulated patient-level net benefits as a linear function of the patients' cost, clinical benefit, and a
treatment arm as a dummy variable. We aimed to test the normality of net benefits and to determine if
remaining as comparators. However, it may be impractical to directly compare all sequences; with the
skewed density of net benefits was better modeled by a mixture of normal densities or a single
t test showed that the net benefits were modeled by a mixture of normal densities. All analyses were done under
threshold willingness-to-pay per unit of clinical benefit. It has been proposed that net benefits
the optimum ordering is identified by estimating the expected net-benefit per unit time for each treatment—accounting for the proportion of patients who do not respond. As there will be attrition as we proceed along a treatment sequence, we maximize the total expected net benefit by
the optimum sequence identified by the proposed methodology were found to be identical to the optimum sequences identified by "a conventional" incremental analysis including all
treatment sequences as comparators. The analysis of treatments for chronic pain showed that
oxycodone CR, morphine CR, and Fentanyl would all be included in a treatment sequence at a
threshold above $10,000.

Conclusions: The proposed methodology is convenient for the identification of optimum treat-
ment sequences. The probability that a given treatment will be cost-effective at any point in the se-
quence can also be estimated. The methodology is valid for treatments that only provide benefit
rather than increasing benefit with increasing cost-effective thresholds as would be ex-
pected. Unmet OLS regression assumptions along with differences in net benefit estimates depend-
ing on robust regression methods suggest additional investigation of the use of net benefit as an
alternative to the cost-effectiveness ratio is needed.
RESULTS: The bias that was associated with using fewer than 15 categories never exceeded an absolute value of 0.85%, and always approached zero as the number of categories increased. For most parameter variations, the absolute bias never exceeded 0.45% when at least 3 categories were used.

Conclusions: Categorizing continuously valued risk factors in decision-analytic models has a negligible effect on model outcomes, remaining at less than 1% absolute bias even when only 2 categories are used.

ELEPHANT NODES SAVE THE DAY: MORE EFFICIENT COST-EFFECTIVENESS EVALUATION OF STRATEGIES INVOLVING EMBEDDED DECISIONS

Hagerty C, Sanders C, and Sonnenberg F
University of Medicine and Dentistry of New Jersey - Robert Wood Johnson Medical School, Highland Park, NJ; Duke, Durham, NC; University of Medicine and Dentistry of New Jersey - Robert Wood Johnson Medical School, New Brunswick, NJ

Background: Cost-effectiveness analyses of test/treatment strategies with sequential decisions typically require repeated analyses of sub-models that estimate utility in effectively identical contexts. We developed a technique for improving the efficiency of such analyses by storing intermediate results of a decision tree at the specific positions where they can be reused. We refer to these nodes as “Elephant Nodes,” since they both remember the utility for each context in which they appear, and require the storage of large amounts of additional storage. When subtrees involve time-consuming utility calculations such as Markov processes, the tradeoff between space and time make this approach unattractive and require potentially large amounts of additional storage. When subtrees involve time-consuming utility calculations such as Markov processes, the tradeoff between space and time make this approach unattractive and require potentially large amounts of additional storage.

Methods: We implemented Elephant Nodes in the Decision Maker software package (www.SimPal.org) so that the modeler needs to provide only a binding to a special variable, “REMEMBER,” on a branch that may benefit from memory. The value of this variable indicates an index for the context in which the associated utility values should be stored. If an identical series of events occurs via another path, the utility does not need to be recomputed. This index value can be encoded into binding expressions at design time through automatic means. Similarly, embedded decision nodes can be automatically transformed into Boolean nodes determined by a top-level strategy decision.

The strategies can be enumerated taking into account all possible combinations of embedded decisions in all event contexts. To enumerate the contexts for which subtree utilities are identical, we develop a generic decision-analytic Markov cohort model in which subjects are distributed into disease-free states defined by a categorized risk factor. From each disease-free state, we assign a unique probability of developing disease based on the average risk factor value and assume a logistic function of disease risk. Persons may die either from disease or from other causes. The outcome of interest is life expectancy with versus without a hypothetical intervention, which changes the average risk factor value by a percentage. We ran the model while varying the number of risk factor categories from 2 to 15, using 15 as the gold standard against which to compare results. We evaluated model results for changes in four key model parameters: overall disease risk, risk factor effect, disease-specific mortality, and intervention effect. Bias was defined as [LEGainN - LEGain]/ LEGain, where LEGain is the life expectancy gain from the hypothetical intervention, N is the number of categories used, and NMax is 15.

RESULTS: The bias that was associated with using fewer than 15 categories never exceeded an absolute value of 0.85%, and always approached zero as the number of categories increased. For most parameter variations, the absolute bias never exceeded 0.45% when at least 3 categories were used.

Conclusions: Categorizing continuously valued risk factors in decision-analytic models has a negligible effect on model outcomes, remaining at less than 1% absolute bias even when only 2 categories are used.

PRESENTING COST-EFFECTIVENESS INFORMATION FOR A HETEROGENEOUS POPULATION: AN ILLUSTRATION BASED ON A MODEL OF ACE-INHIBITION IN CARDIOVASCULAR DISEASE

Briggs A, Mihaylova B', Sculpher M, Hall A*, Deckers J, Simoons M, and Fox K
1University of Glasgow, Glasgow, United Kingdom; 2University of Oxford, Oxford, United Kingdom; 3University of York, York, United Kingdom; 4University of Leeds, Leeds, United Kingdom; 5Emmasius University, Rotterdam, Netherlands; 6Royal Brompton Hospital, London, United Kingdom

Objectives: To explore methods of presenting cost-effectiveness information by baseline risk of disease.

Methods: An economic model of the ACE inhibitor perindopril based on data from the EUROPA trial on Reduction Of cardiac events with Perindopril in stable coronary Artery disease (EUROPA) was developed. Cost-effectiveness is presented as a function of the baseline risk of a primary event (non-fatal myocardial infarction, cardiac arrest, or cardiovascular death) in order to identify individuals for whom treatment offers greatest value for money. While the presentation of cost-effectiveness solely on the basis of a clinical risk equation can provide useful information, it is important to also account for the impact of age and sex in the cost-per-QALY calculations. We therefore propose a variation on standard cardiovascular risk tables for presenting cost-effectiveness information by risk while adjusting for age and sex.

RESULTS: The median incremental cost per QALY gained from perindopril across the heterogeneous population of EUROPA was estimated at £9,500 (95% CI: £4,800-£27,600) with a 25th percentile of £6,500 (£2,000-£12,300) and a 75th percentile of £14,400 (£6,800-£43,600). Cost-effectiveness was strongly related to higher risk of a primary event under standard care, but was also related to age and sex, as shown in the risk tables.

Conclusions: Standard clinical risk equations are an important tool for clinical decision making. However, although cost-effectiveness results are strongly correlated with clinical risk, age and sex can also have an important influence on the cost-per-QALY calculations. Presenting cost-effectiveness results by risk, age, and sex has the potential to disentangle these important effects and lead to improved decision making.
The aim of this study was to determine in which specific locations it would be cost-effective to install a defibrillator for the treatment of out-of-hospital cardiac arrests.

A cost-effectiveness analysis compared the total costs and life expectancy of treating patients with a cardiac arrest with and without a defibrillator on site by location. Eighteen location types were considered: casinos, non-acute hospitals, shopping centers, nursing homes, penal institutions, hotels, golf courses, restaurants, airports/railway stations, marinas, schools, single stores, medical offices, office buildings, stadium, community facilities, factories, sports fields. Data on the incidence and age, sex and outcome of cardiac arrests by location were available for the province of Ontario for 1995 to 1999. A decision analysis model was created to estimate age and sex specific costs and life expectancies for individuals suffering cardiac arrests. Survival estimates and disease progression probabilities were derived from the Weinstein model, and costs were based on observed health resource utilization of incident coronary heart disease for a representative sample of the Canadian population. The benefit of on-site defibrillators was incorporated through data from a multicenter trial (RER of survival = 2). The Canadian costs of on-site defibrillators were included. Both future life expectancy and costs were discounted at a rate of 5%. Monte Carlo simulation estimated the uncertainty surrounding the estimated incremental cost-effectiveness ratios.

For each of the 18 sites, the incremental cost-effectiveness ratio associated with on-site availability of a defibrillator was calculated. There were two situations in which a defibrillator would be considered cost-effective based on a willingness to pay threshold of $50,000 per life year gained, namely, casinos and non-acute hospitals. The ICER for casinos was $5911 per life year ($closely 306 vs. $409 in the ICS+MON group and at

This study demonstrates that it would be cost-effective to install on-site defibrillators in both casinos and non-acute hospitals in Canada.
THE EFFECT OF ANTIRETROVIRAL THERAPY ADHERENCE PROGRAMS FOR HIV-INFECTED INDIVIDUALS ON HEALTH CARE UTILIZATION


Center for Disease Control and Prevention, Atlanta, GA; Los Angeles County Department of Health Services, Los Angeles, CA; Harbor-UCLA Medical Center, Torrance, CA

Purpose: To determine the impact on health care utilization of programs to promote adherence to antiretrovirals among HIV-infected individuals.

Methods: Clients in Los Angeles County were randomized to participate for six months in one of three antiretroviral adherence program arms: 1) Directly Administered Antiretroviral Therapy (DAART) in which a community worker delivered antiretrovirals to their home each weekday, 2) Impact screening use is not known. We describe patient-physician CRC-screening discussions and the association of these discussions with screening use.

El Khoury A, Shaya F, Mullins C, Weir M, Fatou H and Garber H

University of Maryland School of Pharmacy, Baltimore, MD; University of Maryland School of Medicine, Baltimore, MD; Johns Hopkins HealthCare LLC, Glen Burnie, MD

Purpose: The goal of this study is to compare the effects of combination therapies of aspirin with clopidogrel, and warfarin with aspirin or clopidogrel to monotherapies of aspirin alone, and warfarin alone, respectively, on stroke recurrence in Medicare patients who have survived a stroke.

Methods: A retrospective analysis of all medical and pharmacy claims of Medicare stroke patients (ICD-9-Codes 430-438) for the period of 1/1/01-12/31/02 was performed. Inclusion: at least one month of follow-up. (Exclusion: to obtain incident cohort): stroke diagnosis between 1/1/01 and 6/30/01. Outcome measure: non-recurrence of stroke. Using Cox proportional hazard models, we compare non-recurrence rates of combined therapies of aspirin with clopidogrel versus aspirin alone, and warfarin with aspirin or clopidogrel versus warfarin alone, while adjusting for age, gender, race, urban residence, heart disease, hypertension, diabetes, blood pressure, or lipid lowering drugs and persistence on the first stroke preventive drug taken post-stroke.

Results: Total of 925 patients, 38% older than 66, 58% African American, and 35% males. About 70% of those who started on aspirin and added clopidogrel had a recurrence, and 66% of those who started on warfarin and added either clopidogrel or aspirin had a recurrence. Patients who start on aspirin and add clopidogrel versus aspirin alone (HR = 2.43, p < 0.000, CI 1.50-3.73) and those who start on warfarin and add either clopidogrel or aspirin (HR = 1.68, p = 0.02, CI 1.08-2.62) are more likely not to have a recurrence. Patients who start on aspirin and add clopidogrel and have hypertension post-stroke (HR = 0.63, p = 0.05, CI 0.40-0.99) are less likely to avoid a recurrence. The same is true for patients who start on warfarin and add aspirin or clopidogrel (HR = 0.60, p = 0.01, CI 0.41-0.88).

Conclusion: Combination therapies of aspirin with clopidogrel and warfarin with aspirin or clopidogrel decreases the risk of recurrent stroke, when compared to monotherapies of aspirin or warfarin alone. Health care plans may consider these therapies in managing patients with stroke.
VARIABLES ASSOCIATED WITH PHYSICIANS’ ANXIETY AND RELUCTANCE TO SHARE UNCERTAINTY WITH PATIENTS IN THE CONTEXT OF SHARED DECISION MAKING

Levente L, Barrot L, Marrous H, Graham F and O’Connor A1
Laval University, Quebec, QC, Canada; Ottawa Health Research Institute, Ottawa, ON, Canada

Purpose: In the context of shared decision making, providers need to explicitly acknowledge uncertainty with patients. However, physicians’ reactions to uncertainty may influence their ability to do so. The aim of this study was to identify the sociodemographic variables that are associated with physicians’ reactions to uncertainty in a French-speaking sample of physicians. Methods: A cross-sectional survey was performed in a group of private practice as well as in a group of family practice teaching units in Québec, Canada. Two subscales from the Physicians’ Reactions to Uncertainty (PRU) that had been translated into French were used. Analyses were performed using the SAS GLM procedure. Results: 138 physicians (69 clinical teachers and 53 residents in family medicine and 16 physicians in private practice) enrolled in this study. Univariate analyses, younger age (ß = 0.0001), fewer hours worked per week (ß = 0.01), fewer years seen per patient (ß = 0.01), residency status (ß < 0.0001) were found to be positively associated with anxiety due to uncertainty but not with reluctance to disclosing uncertainty to patients. Having another diploma (ß = 0.01) was found to be negatively associated with reluctance to disclosing uncertainty to patients but not with anxiety due to uncertainty. In multivariate analyses, a female (ß = 0.483), a 1st year resident (ß = 0.597) or a 2nd year resident (ß = 1.179), and fewer hours worked per week (ß = 0.012) explained 30% of the variance in the anxiety due to uncertainty score. Having another diploma (ß = 0.510) explained 22% of the variance in the reluctance to disclosing uncertainty to patients score. Conclusions: Sociodemographic variables that are associated with anxiety due to uncertainty differed from those associated with reluctance to disclosing uncertainty to patients. In line with previous research, this study showed that female physicians experienced more anxiety due to uncertainty than their male counterparts. Number of hours worked per week and a residency status were also found to influence this type of anxiety. Having another diploma appeared to improve disclosure of uncertainty to patients. Therefore, those interested in implementing shared decision making should tailor their interventions accordingly.

ABSTRACTS

POSTER SESSION II

7:00 AM - 8:30 AM
Sunday, October 23, 2005
29 Pacific Room EFG

TRANSLATING RESEARCH INTO PRACTICE? UNFORESEEN HURDLES IN THE IMPLEMENTATION OF A DECISION AID RESEARCH PROJECT IN A HEALTH CARE SETTING

Conklin SE, McPhee J, Johnson K, Zikmund-Fisher B, Derry H, Stark A, Claus B and Fagerlin A
Group Health Cooperative, Seattle, WA; University of Michigan, Ann Arbor, MI; VA Ann Arbor Healthcare System, Ann Arbor, MI; Henry Ford Health System, Detroit, MI

Purpose: This report illuminates real-world challenges and opportunities encountered while implementing a project testing an Internet-based risk communication decision aid in two healthcare systems. Approaches to communicating tailored risk information over the Internet have not been fully optimized, though this modality has the potential to reach large populations in a replicable and sustainable way. Moreover, Web-based communication is an increasingly important adjunct in patient care.

Methods: This project is designing a tailored decision aid to communicate with women about breast cancer risk and inform them about tamoxifen as an option for prophylaxis. Collaborators include two health care systems (group model HMOs) and an academic coordinating site. Study participants will be recruited from enrolled populations of the HMOs. The academic site led development of the decision aid, which includes personalized breast cancer risk information, different methods for presenting risk information, and detailed description of the risks and benefits of tamoxifen. The health plan leaders and institutional review board of the two HMOs were presented with all materials.

Results: Our HMO expressed numerous concerns about the decision aid content, including the fact that women would be receiving medical information from a source other than their physician. Another concern was that demand for tamoxifen (cost of medication and increase in physician visits) would spiral as a result of this project. Following protracted negotiations and modifications to the original design, the study was approved by the IRB and leadership of the concerned HMO. The strong reaction to this risk communication tool was not anticipated at the inception of the study. The reason for this was that we were often caught by surprise regarding the levels of the decision aid and accompanying data collection tools. Concerns stemmed primarily from perceived inappropriate use of the decision aid by participants.

Conclusions: Translational research on clinical decision aids—particularly risk communication tools—is imperative to help both patients and doctors reach the best possible conclusion in situations of medical uncertainty. However, practitioners in delivery systems and researchers may have legitimate but divergent viewpoints about approach, salience, and impact of certain topics. Involving practicing clinicians and patients in the development of risk communication strategies and tools may be warranted.

2271 POSTER SESSION II

7:00 AM - 8:30 AM
Sunday, October 23, 2005
30 Pacific Room EPG

THE EFFECT OF OUTCOME INFORMATION ON DOCTORS’ SPONTANEOUS LEARNING

Mazzucco K1, Romiti K1 and Cherubini P1
1University of Trento, Rovereto, Italy; 2University of Padova, Padova, Italy; 3University of Milano-Città, Milano, Italy

Background: Outcome bias is a psychological tendency to use outcome information to evaluate the quality of a previous decision, where decisions with bad outcomes are undervalued and decisions with good outcomes are overvalued. Previous studies have shown that the bias can be present in some evaluations of diagnostic and/or treatment medical decisions (e.g., Baron and Hershey, 1988; Grupen, Margolin, Wisdom, and Grum, 1994). Jacobi and Cherubini (2004) recently showed that outcome bias can affect doctors evaluating their own decisions, and argued that in some circumstances, this tendency might have consequences on spontaneous learning: bad diagnostic decisions resulting, by chance, in good outcomes might be improperly repeated in the future, while good diagnostic decisions with bad outcomes might be improperly abandoned.

Methods: A total of 49 practicing Italian doctors and nurses were contacted at three times. On the first meeting, the nurses were given a clinical case that required to send out an ambulance and were asked to decide whether to send a doctor with the ambulance or not, and the doctors were given a clinical case and were asked to choose between two possible diagnoses. Participants had to make their choice and to report their confidence toward its quality. At the second meeting (2 or 3 weeks later), participants were presented with a randomly selected positive or negative outcome of their previous choice. At the third meeting (3 or 5 weeks later), participants were given a clinical case superficially different but structurally and clinically equal to the previous one and were asked to choose between two alternatives and to report their confidence toward their current choice.

Results: At the second meeting, participants in the positive outcome group showed an increase in confidence, whereas participants in the negative outcome group showed a decrease in confidence, replicating the typical outcome effect. More important, reliably more participants made a different choice in the second clinical case in the negative outcome group than in the positive outcome group.

Conclusions: The results show that outcome bias can affect spontaneous learning. After a negative outcome, clinicians changed their decisions on a structurally equal case more than after a positive outcome.
Purpose: Reimbursement authorities are faced with significant escalating health care costs, pressured budgets, and numerous launches of biotech and pharmaceutical technologies with uncertain "value for money." At the same time, biotechnology/pharmaceutical companies are assessing their future potential through probabilistic Monte Carlo sensitivity analysis was performed for an array of parameters. Results: The model estimated that the most cost-effective subpopulations included the F3 and F4 Metavir patients, compared to the broadest population. At a payer's willingness to pay threshold of $500,000/QALY, the highest incremental efficacy was achieved in F3 and F4 patients for Drug X to be cost-effective over 20 years were predicted at 16% and 45%, for annual price points ranging from US1,500 to US8,000, respectively. Sensitivity analysis determined the critical drivers included incremental efficacy/safety rates, drug price, patient subpopulation, and utility values.

Conclusion: Econometric valuation modeling has been influential in guiding the company on anticipated patient dynamics; to predict cost-effective positioning and treatment populations; and to project cost-effective pricing for new product opportunities. This approach may also help payers by identifying the patient groups that will gain clinical benefit at a cost-effective drug price. Furthermore, manufacturers are increasingly recognizing that this valuation approach may reduce product development risk by establishing more realistic estimates of market potential, pricing, and societal needs.

Methods: A deterministic Markov Excel model is illustrated comparing chronic liver disease populations treated with a new product candidate, Drug X, to established standard care. The model was tested over a three-year period using successive cohorts of experienced nurses (N = 95) registered on a specialized postgraduate module in clinical judgment and decision making. The majority of course participants were working in mental health nursing. A review of the research literature identified a set of cognitive and social/organizational factors relevant to judgment and decision making. These factors were integrated in a conceptual model and operationalized as individual items for a clinician to consider with respect to a clinical case self-selected from practice. When used in an educational context, verbal or written feedback is provided by an assessor to some or all responses, with the aim of stimulating critical reflection.

Results: The cases presented by respondents varied with respect to the clinical challenges posed (e.g., an unusual constellation of symptoms, miscommunication between health professionals, problem with several plausible causes) and were generally conceptualized using a biopsychosocial model. Several difficulties were experienced by respondents in articulating the judgment and decision-making process (e.g., categorizing information at different levels of abstraction, mapping subjective judgments onto quantitative scales, discussing patient preferences). Respondents appraised the JM-PiP tool positively in qualitative and quantitative evaluations as a useful means for understanding clinical, interpersonal, and organizational factors, it provides insights on clinical knowledge and the organization of care. Further applications of the JM-PiP may be explored, including its integration in professional development programs and use as a research tool. Considering the genetic nature of many aspects of judgment and decision making, the JM-PiP may also be used with other health care professionals.

Purpose: To explore the potential of JDM-PiP to provide insight into the decision-making process for experienced nurses. The questionnaire was then reduced and scored, based on item contents, distributions, and correlations.

Results: The broad concepts retained from literature and interviews were symptom severity, focus status, satisfaction, and adherence. The pilot questionnaire comprised 52 items in 10 sections (symptoms, allergy in daily life, motivation for desensitization, advantages, concerns, intake, outcomes, satisfaction, continuation of SIT, information). Eight items were specific to SIT-naïve patients, 21 to patients having a SIT. The mean percentage of defining was data 2.16 for treated patients and 1.46 for naïve patients. The clinicians reported high patient acceptability and major interest in using the questionnaire routinely. The prioritization given to the sections was symptoms, satisfaction, outcomes, continuation, allergy in daily life, intake, information, convenience, benefit/association, features that matter, and costs. The prioritization given was motivation, symptoms, allergy in daily life, information, advantage, convenience.

Conclusions: This self-reported questionnaire is a flexible, pragmatic tool allowing clinicians to better manage their patients in clinical practice. It is well accepted by patients and perceived as useful by clinicians. It is now available for inclusion in a prospective longitudinal study that will show its ability to predict persistence problems.
ABSTRACTS

1951
POSTER SESSION II
7:00 AM - 8:30 AM
Sunday, October 23, 2005
38 Pacific Room EFG
BENEFIT HARM TRADEOFFS ESTIMATE SIGNIFICANTLY IMPORTANT DIFFERENCE FOR COMMON COLD TREATMENTS
Barett B and Brown D
University of Wisconsin - Madison, Madison, WI
Purpose: To estimate sufficiently important difference (SID) for common cold. SID is the amount of benefit sufficient to justify treatment costs and risks.
Methods: Hypothetical benefits of reduction in duration and severity were assessed. We presented treatment scenarios based on best evidence regarding vitamin C, the herbal echinacea, zinc lozenges, and placebo, an unapproved antivirus. Costs and risks were held constant. Hypothetical benefits were varied until the cold sufferer indicated the treatment was desirable.
Results: In Cohort A, we assessed duration with 144 telephone and 149 in-person interviewees. For the vitamin, mean SID was 26.1 hours (95% CI = 23.2, 29.3), with 31% accepting without any duration benefit, and 1% rejecting even with a 4-day reduction in duration benefit. For the herbal, SID was 36.8 hours (CI = 33.4, 40.2), with 23% accepting, and 9% rejecting. For the lozenges, SID was 64.8 hours (CI = 61.0, 67.9), with 9% accepting and 24% rejecting. For the antiviral, SID was 82.6 hours (CI = 78.7, 86.7), with 6% accepting and 48% rejecting. An overall SID of 52.6 hours (CI = 49.3, 55.4) was determined by averaging across the four scenarios.
In Cohort B, we assessed severity SID with 144 telephone and 49 in-person interviewees. For the vitamin, mean SID was 24.0% (CI = 20.8, 27.2%), with 13% saying they would take it regardless of severity benefit, and 1% saying they would not, even if severity benefit exceeded 75%. For the herbal, SID was 31.8% (CI = 28.3%, 33.5%), with 6% accepting and 7% rejecting. For the lozenges, SID was 45.0% (CI = 41.5, 48.5%), with 2% accepting and 4% rejecting. For the antiviral, SID was 56.4% (CI = 52.2%, 60.6%), with 2% accepting and 36% rejecting. Overall SID was estimated at 39.3% (CI = 36.9, 41.7%) by averaging across scenarios.
Differences in SID among treatment scenarios were not due to chance. Age, sex, and severity of illness did not significantly influence response patterns. Qualitative interviews suggested that side effects, treatment type (tablet vs. lozenge vs. herbal extract), monetary costs, and opportunity costs may be important additional influences.
Conclusions: On average, people want the duration of their colds to be reduced by 26 to 65 hours in order to justify costs and risks of popular cold treatments. A prescription antiviral would require more benefit (45 hours). Alternative common cold treatments would need to reduce global severity by 24% to 45% to justify costs and risks. A prescription antiviral would require a 56% severity reduction-benefit.

2084
POSTER SESSION II
7:00 AM - 8:30 AM
Sunday, October 23, 2005
40 Pacific Room EFG
THREE-CLASS ROC ANALYSIS—DECISION THEORETIC APPROACHES
H Idol, Link J, Tout B and Frey E
Johns Hopkins Medical Institutions, Baltimore, MD
Purpose: ROC analysis is well established in the evaluation of binary task performances. However, the evaluation and optimization of some medical systems involve distinguishing among three or more classes. We have developed a practical, though not fully general, three-class ROC analysis method based on decision theory.
Methods: Four decision-theoretic optimization strategies were investigated, including the maximum utility criterion, the maximum-correctness criterion, the maximum likelihood criterion, and the Neyman-Pearson criterion.
Results: All optimization strategies lead to the use of the ideal observer as the optimal observer. Ideal observer uses the likelihood ratios as the optimal decision variables. Both Neyman-Pearson and maximum utility criteria lead to a decision model that has a 2-D decision plane and a decision structure with 5 degrees of freedom. However, this model does not provide a tractable way of determining the decision structure on the decision plane. We made the assumption that wrong decisions have equal Lagrange multipliers under the same hypothesis for the maximum utility criterion, and the assumption that wrong decisions have equal Lagrange multipliers under the same hypothesis for the Neyman-Pearson criterion. Then, using a log transformation, both criteria lead to a decision model that has a 2-D decision plane and a decision structure with 2 degrees of freedom. Moving the decision structure over the 2-D decision plane and computing the true class 1 fraction (T1F), true class 2 fraction (T2F), and true class 3 fraction (T3F) defines a 3-class ROC surface. The volume under the surface (VUS) serves as a figure of merit for the three-class task performance. The maximum-correctness criterion also leads to the same 2-D decision plane and the same ROC surface without any assumptions, while the maximum likelihood criterion results in an operating point on this ROC surface.
Conclusions: All four optimization strategies, when combined with a single assumption, lead to the same optimal observer and the same 3-class ROC surface. Thus, the proposed 3-class ROC analysis method is meaningful according to all four decision strategies.

ABSTRACTS E59
null
Purpose: While an individual's risk preferences are assumed to influence decisions that involve uncertainty, risk perceptions may also play a role. Belief in personal luck is one factor that may influence how individuals perceive risk. In this preliminary work, we examine the relationship between belief in luck and two kinds of decisions involving uncertainty, in a sample of cardiovascular patients.

Methods: In the first decision task, subjects make real choices over monetary lotteries and are paid their earnings. The second decision task involves hypothetical choices over two medications: one that will reduce the frequency of these symptoms with certainty, the other represents a gamble between complete relief of symptoms and no effect. Subjects also completed a survey measuring five concepts related to belief in luck: personal good luck, personal bad luck, general belief in luck, belief in the luck of others, and the influence of luck on decision making.

Results: 79 patients who presented for stress testing were evaluated. We find that belief in luck varies across subjects: a significant number of subjects believe in luck. While 58% of subjects strongly or somewhat agreed with the statement “It is a mistake to base any decisions on how lucky you feel,” this is not consistent with the behavior we observed. We found significant positive correlations between belief in luck and the number of risky choices made in the lottery game and the health choice game. In the lottery game, there was a significant positive correlation between belief in others’ luck and the number of times the subject’s certainty equivalent was greater than the expected value of the lottery (p = 0.22, p = 0.052) Similarly, with the health choice game, there was a significant positive correlation between both the number of times the subject chose the risky drug and belief in others’ luck (p = 0.22, p = 0.052) and general belief in luck (p = 0.22, p = 0.05). Conclusions: These results suggest that measurements of patients’ risk preferences may also reflect patients’ risk perceptions, which may in turn be influenced by belief in luck. By measuring patients’ belief in luck, we enhanced our ability to distinguish between individuals who truly have risk-seeking preferences and those who are either risk averse or risk neutral but whose perception of risk is mitigated by their belief in luck.

WILLINGNESS TO PAY FOR ANTIEMETIC REGIMENS FOR PREVENTION OF CHEMOTHERAPY-INDUCED NAUSEA AND VOMITING

Miley R and Miller1

West Virginia University, Morgantown, WV; West Virginia University School of Pharmacy, Morgantown, WV

Purpose: Patients with cancer perceive chemotherapy-induced nausea and vomiting (CINV) among the most distressing side effects of chemotherapy. Recently, new antiemetic agents such as aprepitant and palonosetron have been approved that provide better control of CINV as compared to the standard regimen of 5-HT3 receptor antagonist and corticosteroid. The objective of this study was to measure the value that cancer patients place on improved emesis control provided by new antiemetic regimens.

Methods: The contingent valuation approach was used to measure patients’ willingness to pay for CINV control. Patients were recruited from two local hospitals. Seventy subjects were approached and 55 agreed to participate. Participants were between 18 and 75 years old and were licensed drivers with a valid license for three years or more.

Results: The mean WTP was $90 (SD = $101.9). The mean WTP for a single injection of regimen B for a single month in an outpatient setting. After background information had been presented, participants were asked to state the maximum amount that they would pay out of pocket for the additional benefit provided by regimen A and B using the payment card method.

Conclusion: Almost 82% of participants were either receiving their second or later cycles of chemotherapy or had received it in the past three months. On a scale of 0 to 10, the perceived levels of importance for benefits provided by regimen A for the acute and delayed phases were 8.8 (SD = 1.7) and 9.2 (SD = 1.5), respectively, and by regimen B for the delayed phase was 8.6 (SD = 1.6). Approximately 91% of the participants were willing to pay for the new antiemetic regimens. For a 3-day course of regimen A, providing 13% improvement in acute and 11% improvement in delayed emesis, the mean WTP was $90 (SD = $101.9). The mean WTP for a single injection of regimen B for a 12% improvement in delayed emesis was $84 (SD = $84.5). Multivariate analyses conducted to identify relationships between WTP and sociodemographic and clinical variables showed that only annual household income was significantly associated with WTP.

Conclusion: This study among the first to evaluate perceived level of importance and WTP estimates for improved emesis control due to new antiemetic regimens among cancer patients in the United States. The study results indicate that cancer patients perceive improvements in the control of nausea and vomiting as important and are willing to pay out of pocket to receive the additional antiemetic benefits.
SITUATION-SPECIFIC MODEL OF REPRODUCTIVE DECISION MAKING

Howard E

Brown Medical School; Women and Infant’s Hospital, Providence, RI

Purpose: To examine the phenomenon of reproductive decision making in the context of women’s health care. Advocacy for informed choice, shared decision making, and the right to self-determination is a hallmark of women’s health care. Self-determination and empowerment are valued components in the reproductive health care of women, which leads to improved health status overall.

Reproductive decision making is viewed as a complex social process with multiple dimensions, moderated by personal and social contexts. The goal of this concept analysis is to identify the dimensions of the process of reproductive decision making.

Design: Dimensional analysis, a methodological approach to grounded theory, is based upon the work of Schatzman (1991). Through a review of the extant literature, using qualitative, grounded theory methodology, themes were identified, and dimensions of choice emerged.

Methods: 367 articles related to reproductive choices and decision making were selected. Relevant findings, recurring phrases, and themes were identified and categorized, with the aid of matrices, utilizing the process of dimensional analysis, metasynthesis techniques, and grounded theory methodology. Data sources: The medical, nursing, social science, legal, religious, and ethics literature were reviewed using computer-based and secondary bibliographic searches. Electronic databases included GINAH, MEDLINE, Cochran Review, Ethics Consult, EBC, and HAP. Search terms included fertility, pregnancy diagnosis, fetal impairment, quality of life and disability, maternal serum testing, reproductive choice, reproductive technologies, therapeutic abortions, reproductive decision making, and prenatal genetic counseling. Dictionaries, encyclopedic references, and the classical literature were consulted.

Results: Initial analysis revealed 25 different dimensions of the process of reproductive decision making. A final, integrative analysis reduced the initial 25 dimensions to 4 core dimensions of reproductive decision making: personal appraisal, the structure of health services, informed consent, and self-determination.

Conclusion: Reproductive decision making is situation specific, shaped by the interaction of personal and situational factors. It is a multidimensional social process that is based on a relational world view. Support and facilitation of an individual’s control over her reproductive choices is the ultimate goal of the women’s health care provider.

DIFFERENCES IN CANCER RISK PERCEPTION AMONG DIVERSE WOMEN

Kim S, Wong S, Pérez-Stable E, Kaplan C, Walsh J and Sawaya G

1University of California, San Francisco, San Francisco, CA; 2University of British Co-

lumbia, Vancouver, BC, Canada

Purpose: Inaccurate perceptions of risk compromise informed decision making and may have important behavioral and psychological consequences. We examined the perception of breast, cervical, and colorectal cancer risk and compared to the actual risk and the screening behavior of a sample of women from four ethnic groups.

Methods: Women, aged 50 to 80, who had made at least one visit to a primary care physician in the past two years, were recruited from ambulatory practices. Trained interviewers administered an initial screening telephone survey and a follow-up-in person interview in English, Spanish, or Chi-

nese. Multivariate regression models examined ethnic differences in risk perception and screening, controlling for education, age, income, previous cancer history, and family history of cancer.

Results: To date, 760 women completed the survey, 316 non-Latino White (41.6%), 121 African American (15.9%), 121 Latinos (15.9%), 178 Asians (24.4%), and 24 other (3.2%). Multivariate re-

sults showed that compared to White women, Latina women perceived 12.4% higher risk of breast cancer (p < 0.001), 17.8% higher risk of cervical cancer (p < 0.001), and 35.2% higher risk of colorectal cancer (p < 0.001). African American perceived 6.5% higher risk of breast cancer (p < 0.05), 8.6% higher risk of cervical cancer (p < 0.05), and 8.2% higher risk of colorectal cancer (p < 0.05); compared to non-Latino White women. Asian women perceived 5.5% lower risk of breast can-

cer (p < 0.001) and no significant difference in cervical and colorectal cancer risk, compared to non-

Latino White women. For any cancer, having a college degree was related to a lower cancer risk per-

ception (p < 0.05). Women who have had a colonoscopy reported about 9.8% lower colorectal cancer risk perception (p < 0.001). Women who had a colonoscopy reported about 8.5% higher colorectal cancer risk perception (p < 0.001). There was no significant relationship between cervical or breast cancer screening and risk perception.

Conclusion: All women did not have an accurate understanding of their cancer risk. Yet ethnic differences persisted in cancer risk perception even after controlling for education, age, income, and cancer history (self and family). Only colonoscopy screening was positively related to having a higher colorectal cancer risk perception. To help clinicians communicate with patients about risk in-

formation, we will examine whether using traditional percentage and numeracy information is effective in all diverse populations.
Our objective was to develop a natural history model of GC that would be capable of evaluating cost-benefit. A series of calibration exercises were conducted to elucidate the unobserved parameters for the disease. Modeling includes antibiotic treatment, screening for precancerous lesions, and an eventual Hp vaccine.

Primetime television programming (7–11 pm) on four network channels (ABC, CBS, Fox, NBC) was recorded for a period of one month, resulting in 38 unique prescription drug ads. Ads were coded by two independent coders for factual information about the target condition provided and narrative themes related to portrayals of medication use and non-pharmacological alternatives.

More research is needed to determine how DTCA themes influence consumer expectations of pharmaceuticals and health care services, more generally.

ABSTRACTS
Purpose: Previous models have evaluated the effectiveness and cost-effectiveness of HPV vaccination and cervical cancer screening based on few natural history parameter sets, despite considerable evidence that surrounding these parameters. More research is needed to better characterize and understand the natural history of HPV. The objective of this study is to develop a framework to identify all plausible natural history parameters of a detailed cohort model of HPV infection, consistent with empirical epidemiological trends.

Methods: We developed a 3-stages process: 1) sampling, 2) fitting, and 3) sensitivity analysis. Prior uniform probability distributions were defined for each natural history parameter using minimum and maximum parameter values derived from the literature. Numerous combinations of parameters were drawn from prior distributions using Latin Hypercube Sampling (LHS). Model predictions on age-specific incidence and prevalence of HPV infection, CIN, and cervical cancer as well as genotype distributions were compared to empirical North American longitudinal and cross-sectional data using different GOF algorithms. Sensitivity analyses of the quality of fit (GOF) to the different parameter values were performed by univariate and multivariate analysis.

Results: Preliminary analysis suggests that GOF variability due to natural history parameters is important. Given model assumptions, the most influential parameters are clearance and natural immunity. The LHS-GOF parameterization process effectively and objectively finds input parameter sets that best fit North American HPV data. Identification of parameter sets with equal GOF selects different natural history scenarios that are equally plausible.

Conclusions: This work is original because we focus on understanding the variation and sensitivity of the GOF results to parameter assumptions to answer important parameterization questions such as: 1) What is the overall GOF variability due to uncertainty in natural history parameters and GOF distributions? 2) Which input parameters explain the most of the fitting imprecision (i.e., GOF variability)? 3) Which parameter sets best fit current empirical North American HPV epidemiological trends? Have sensitivity analysis 1-3 to GOF measures used? 5) What important data are needed? The parameterization process (LHS + GOF + SA) permits to effectively explore a wide range of parameter values and to interpret the results within a statistical framework necessary to explore natural history assumptions. Such research is needed due to the structural complexity of HPV models and high degree of uncertainty in the natural history of HPV.

Methods: To develop a disease progression model of the clinical course of disease in patients with pneumonia-related sepsis that represents hospitalized patients as they progress through varying states of severity until death or hospital discharge.

Results: This work is original because we focus on understanding the variation and sensitivity of the GOF results to parameter assumptions to answer important parameterization questions such as: 1) What is the overall GOF variability due to uncertainty in natural history parameters and GOF distributions? 2) Which input parameters explain the most of the fitting imprecision (i.e., GOF variability)? 3) Which parameter sets best fit current empirical North American HPV epidemiological trends? Have sensitivity analysis 1-3 to GOF measures used? 5) What important data are needed? The parameterization process (LHS + GOF + SA) permits to effectively explore a wide range of parameter values and to interpret the results within a statistical framework necessary to explore natural history assumptions. Such research is needed due to the structural complexity of HPV models and high degree of uncertainty in the natural history of HPV.

Purpose: To develop a disease progression model of the clinical course of disease in patients with pneumonia-related sepsis that represents hospitalized patients as they progress through varying states of severity until death or hospital discharge.

Methods: We describe an empirically based, Monte Carlo microsimulation model used to represent the clinical course of disease in patients with pneumonia-related sepsis and their clinical condition over time, encompassing realism beyond that practical in Markovian, state-based models. Patients are characterized individually patients with pneumonia-related sepsis and their clinical condition over time, enabling realism beyond that practical in Markovian, state-based models. Patients are characterized.
ABSTRACTS

POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
1 Pacific Room EFG

1. POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005

1 Pacific Room EFG

PANEL DISCUSSION OFFICIAL IN-PATIENT GUIDELINE FOR POSTOPERATIVE CARE

Purpose: To develop evidence-based colorectal cancer (CRC) screening guidelines for use in average-risk members of the Kaiser Permanente Medical Care Program in the United States.

Methods: The guideline development team formulated the research questions for the literature review, with particular emphasis on the preferred choice of screening tests and the optimal interval to conduct screening. A systematic literature review was conducted examining the evidence underlying each question. Each recommendation was classified as "evidence based" or "consensus based," and the quality of underlying evidence was graded as A, B, C, or I (for insufficient evidence).

Results: We found grade "A" evidence supporting CRC screening in average-risk adults. There is insufficient evidence to determine which CRC screening modality is most effective in terms of the balance of benefits and potential harms. The evidence is strongest for fecal occult blood tests supported by randomized clinical trials (Grade: "A"), and for flexible sigmoidoscopy (FS), supported by case-control and cohort studies (Grade: "B"). A consensus-based recommendation to combine FS and FOBT can be made based on grade C evidence. There is insufficient evidence to recommend for or against the use of colonoscopy or air contrast barium enema for CRC screening. The evidence base is strongest for performing FOBT at 2-year intervals, rather than 1 year, and for performing FS at 10-year intervals, rather than 5 years.

Conclusions: Our evidence review and the guideline product are consistent with the US Preventive Services Task Force. In distinction with other published guidelines, we found insufficient evidence to justify the use of colonoscopy for screening average-risk adults for CRC. We also found that currently recommended CRC screening intervals are more frequent than what is supported by the evidence. Full adoption of our guideline would be associated with fewer colonoscopy-related complications and reduced colonoscopy resource requirements. Less reliance on constrained colonoscopy resources will allow more patients to be screened for CRC, with minimal loss of screening effectiveness.

POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005

3 Pacific Room EFG

AN EVIDENCE-BASED GUIDELINE FOR COLORECTAL CANCER SCREENING IN AVERAGE-RISK ADULTS

Levin T, Koster M, Chan W and Kempe K

Kaiser Permanente, Northern California, Oakland, CA; Kaiser Permanente, Southern California, Pasadena, CA; Kaiser Permanente, Northwest, Portland, OR; Kaiser Permanente, Colorado, Westminster, CO

Purpose: To develop evidence-based colorectal cancer (CRC) screening guidelines for use in average-risk members of the Kaiser Permanente Medical Care Program in the United States.

Methods: The guideline development team formulated the research questions for the literature review, with particular emphasis on the preferred choice of screening tests and the optimal interval to conduct screening. A systematic literature review was conducted examining the evidence underlying each question. Each recommendation was classified as "evidence based" or "consensus based," and the quality of underlying evidence was graded as A, B, C, or I (for insufficient evidence).

Results: We found grade "A" evidence supporting CRC screening in average-risk adults. There is insufficient evidence to determine which CRC screening modality is most effective in terms of the balance of benefits and potential harms. The evidence is strongest for fecal occult blood tests supported by randomized clinical trials (Grade: "A"), and for flexible sigmoidoscopy (FS), supported by case-control and cohort studies (Grade: "B"). A consensus-based recommendation to combine FS and FOBT can be made based on grade C evidence. There is insufficient evidence to recommend for or against the use of colonoscopy or air contrast barium enema for CRC screening. The evidence base is strongest for performing FOBT at 2-year intervals, rather than 1 year, and for performing FS at 10-year intervals, rather than 5 years.

Conclusions: Our evidence review and the guideline product are consistent with the US Preventive Services Task Force. In distinction with other published guidelines, we found insufficient evidence to justify the use of colonoscopy for screening average-risk adults for CRC. We also found that currently recommended CRC screening intervals are more frequent than what is supported by the evidence. Full adoption of our guideline would be associated with fewer colonoscopy-related complications and reduced colonoscopy resource requirements. Less reliance on constrained colonoscopy resources will allow more patients to be screened for CRC, with minimal loss of screening effectiveness.

POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005

3 Pacific Room EFG

AN EVIDENCE-BASED GUIDELINE FOR COLORECTAL CANCER SCREENING IN AVERAGE-RISK ADULTS

Levin T, Koster M, Chan W and Kempe K

Kaiser Permanente, Northern California, Oakland, CA; Kaiser Permanente, Southern California, Pasadena, CA; Kaiser Permanente, Northwest, Portland, OR; Kaiser Permanente, Colorado, Westminster, CO

Purpose: To develop evidence-based colorectal cancer (CRC) screening guidelines for use in average-risk members of the Kaiser Permanente Medical Care Program in the United States.

Methods: The guideline development team formulated the research questions for the literature review, with particular emphasis on the preferred choice of screening tests and the optimal interval to conduct screening. A systematic literature review was conducted examining the evidence underlying each question. Each recommendation was classified as "evidence based" or "consensus based," and the quality of underlying evidence was graded as A, B, C, or I (for insufficient evidence).

Results: We found grade "A" evidence supporting CRC screening in average-risk adults. There is insufficient evidence to determine which CRC screening modality is most effective in terms of the balance of benefits and potential harms. The evidence is strongest for fecal occult blood tests supported by randomized clinical trials (Grade: "A"), and for flexible sigmoidoscopy (FS), supported by case-control and cohort studies (Grade: "B"). A consensus-based recommendation to combine FS and FOBT can be made based on grade C evidence. There is insufficient evidence to recommend for or against the use of colonoscopy or air contrast barium enema for CRC screening. The evidence base is strongest for performing FOBT at 2-year intervals, rather than 1 year, and for performing FS at 10-year intervals, rather than 5 years.

Conclusions: Our evidence review and the guideline product are consistent with the US Preventive Services Task Force. In distinction with other published guidelines, we found insufficient evidence to justify the use of colonoscopy for screening average-risk adults for CRC. We also found that currently recommended CRC screening intervals are more frequent than what is supported by the evidence. Full adoption of our guideline would be associated with fewer colonoscopy-related complications and reduced colonoscopy resource requirements. Less reliance on constrained colonoscopy resources will allow more patients to be screened for CRC, with minimal loss of screening effectiveness.

POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005

3 Pacific Room EFG

AN EVIDENCE-BASED GUIDELINE FOR COLORECTAL CANCER SCREENING IN AVERAGE-RISK ADULTS

Levin T, Koster M, Chan W and Kempe K

Kaiser Permanente, Northern California, Oakland, CA; Kaiser Permanente, Southern California, Pasadena, CA; Kaiser Permanente, Northwest, Portland, OR; Kaiser Permanente, Colorado, Westminster, CO

Purpose: To develop evidence-based colorectal cancer (CRC) screening guidelines for use in average-risk members of the Kaiser Permanente Medical Care Program in the United States.

Methods: The guideline development team formulated the research questions for the literature review, with particular emphasis on the preferred choice of screening tests and the optimal interval to conduct screening. A systematic literature review was conducted examining the evidence underlying each question. Each recommendation was classified as "evidence based" or "consensus based," and the quality of underlying evidence was graded as A, B, C, or I (for insufficient evidence).

Results: We found grade "A" evidence supporting CRC screening in average-risk adults. There is insufficient evidence to determine which CRC screening modality is most effective in terms of the balance of benefits and potential harms. The evidence is strongest for fecal occult blood tests supported by randomized clinical trials (Grade: "A"), and for flexible sigmoidoscopy (FS), supported by case-control and cohort studies (Grade: "B"). A consensus-based recommendation to combine FS and FOBT can be made based on grade C evidence. There is insufficient evidence to recommend for or against the use of colonoscopy or air contrast barium enema for CRC screening. The evidence base is strongest for performing FOBT at 2-year intervals, rather than 1 year, and for performing FS at 10-year intervals, rather than 5 years.

Conclusions: Our evidence review and the guideline product are consistent with the US Preventive Services Task Force. In distinction with other published guidelines, we found insufficient evidence to justify the use of colonoscopy for screening average-risk adults for CRC. We also found that currently recommended CRC screening intervals are more frequent than what is supported by the evidence. Full adoption of our guideline would be associated with fewer colonoscopy-related complications and reduced colonoscopy resource requirements. Less reliance on constrained colonoscopy resources will allow more patients to be screened for CRC, with minimal loss of screening effectiveness.
POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
5 Pacific Room EFG

COST-EFFECTIVENESS OF DRUG-ELUTING STENTS

Wagner T, Engelstad L, McPhee S and Pasick R
VA Palo Alto and Stanford University, Menlo Park, CA; 1Alameda County Medical Center, Oakland CA; 2UCSF, San Francisco, CA

Purpose: Follow-up among women who have had an abnormal Papanicolaou (Pap) smear is often poor in public hospitals that serve women at increased risk for cervical cancer. In this study, we evaluated the cost-effectiveness of a tailored outreach intervention.

Method: We randomly assigned 348 women with abnormal Pap smear results who were served at intervention or usual care. In the intervention, outreach workers provided culturally tailored information and helped patients navigate the system as a supplement to the clinic’s usual care. We also evaluated the cost-effectiveness of a tailored outreach intervention.

Conclusion: In a two-year perspective, the DES strategy resulted in greater costs than the RCT strategy with any realistic combination of model parameters. Cost-effectiveness of DES depended heavily on purchasing price of the stents, risk of restenosis, and rate of reintervention in routine practice.

POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
6 Pacific Room EFG

THE COST-EFFECTIVENESS OF AN OUTREACH INTERVENTION FOR WOMEN WITH ABNORMAL PAP SMEARS

Wagner T, Engelstad L, McPhee S and Pasick R
VA Palo Alto and Stanford University, Menlo Park, CA; 1Alameda County Medical Center, Oakland CA; 2UCSF, San Francisco, CA

Purpose: Follow-up among women who have had an abnormal Papanicolaou (Pap) smear is often poor in public hospitals that serve women at increased risk for cervical cancer. In this study, we evaluated the cost-effectiveness of a tailored outreach intervention.

Method: We randomly assigned 348 women with abnormal Pap smear results who were served at intervention or usual care. In the intervention, outreach workers provided culturally tailored information and helped patients navigate the system as a supplement to the clinic’s usual care. The comparison group was usual care alone (typically a letter mailed to the woman’s home). All women were seen at a Highland Hospital, a county hospital in Oakland, CA, serving a disproportionate number of low-income and multi-ethnic women. We calculated the incremental cost per follow-up within six months of the abnormal test result. Bootstrapping was done to create confidence regions. Sensitivity analyses were done to test the robustness of the results. Long-term follow-up was not feasible because women in the control arm were assigned to the intervention group (i.e., rescued) after six months.

Conclusion: The intervention cost a total of $84,507 for 178 women served. The intervention was associated with a 28% increase in the rate of smock-follow-up. The incremental cost per follow-up was $805 (95% CI $410-$959). The intervention dramatically improved follow-up for women with a high-grade squamous intraepithelial lesion. These women have precancerous cells and face an increased risk that the abnormality will progress to invasive cancer. There were 15 such women in the intervention group and 14 in the control group. After the intervention, 14 (93%) and 6 (43%) in the intervention group and control group, respectively, had documented follow-up at six months.

Conclusions: The effect of an outreach program on the rate of follow-up has been well studied. However, the effect of an outreach program on the rate of follow-up has not been well studied. The results of this study suggest that an outreach program can improve the rate of follow-up for women with abnormal Pap smears. The results also suggest that the program can be cost-effective for women with high-grade squamous intraepithelial lesions, achieving the benefits of the program.

POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
7 Pacific Room EFG

COST-EFFECTIVENESS OF TREATING THE METABOLIC SYNDROME IN THE AFRICAN AMERICAN AND GENERAL POPULATION

Tunosa M, Alsenarney J and Schuster R
Wright State University School of Medicine, Dayton, OH

Purpose: To assess the cost-effectiveness of early treatment of metabolic syndrome (MS) risk factors in the African American and general population.

Method: A cost-effectiveness analysis was carried out using a Markov decision model to compare early (prevention) treatment of MS risk factors (hyperlipidemia, diabetes, and hypertension) with late treatment (usual care) in African Americans and the general population. The main outcome measures were the incremental cost per quality-adjusted life year (QALY). The setting was the adult medical care sector.

Conclusion: Assuming equal or better effectiveness for endovascular repair compared to open surgical repair (n = 32) or open surgical repair (n = 34) of an acute infrarenal AAA from January 1, 2001, until December 31, 2004, were retrieved. Direct hospital costs of all patients were assessed using the resource utilization approach. The in-hospital costs and costs during one-year follow-up were assessed from a health care care perspective. Patients who did not undergo a CT scan before the procedure were excluded from the analysis. Costs of one-year follow-up were completed for 26 patients who underwent endovascular repair and 25 patients who underwent open surgical repair. The mean total in-hospital costs and the mean costs during one-year follow-up were calculated for each treatment group; we used the bootstrap resample method to estimate the 95% confidence interval. Costs between the treatment groups were compared with Mann Whitney U test. In multi- and univariable analyses, we investigated the influence of clinical variables on the total in-hospital costs. Costs were expressed in 2003 Euros.

Result: Sex, age, and comorbidity did not differ significantly between the treatment groups (p > 0.05). In patients treated with open surgery, more ruptured AAA were present. Mortality and postoperative complications were lower for endovascular repair, although not significantly so. The mean total in-hospital costs were significantly lower for endovascular repair compared to open surgical repair ($21,707 versus $33,877, p = 0.03). The total costs including one-year follow-up were $21,433 versus $35,249 for patients who underwent endovascular repair and open surgical repair, respectively (p = 0.12). In multivariable analysis, complications had a significant influence on total in-hospital costs ($22,867; higher costs).

Conclusion: Assuming equal or better effectiveness for endovascular repair compared to open surgery, our results suggest that from an economic perspective, patients with an acute AAA who are eligible for endovascular repair should be treated with this procedure.

POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
8 Pacific Room EFG

COST-EFFECTIVENESS OF DRUG-ELUTING STENTS

Nordhaug 1, Rotevatn S, Sandvik A, Nordhaug I, Rotevatn L, Stoltenberg V, Essendrop J and Nordhaug A
Norwegian Knowledge Centre for Health Services, Oslo, Norway; 1Norwegian Health Services Research Centre, Oslo, Norway; 2Umeå University Hospital, Oslo, Norway; 3Haukeland University Hospital, Bergen, Norway; 4University of Southern Denmark, Odense, N-8031 Oslo, Norway

Objective: To explore cost and health consequences of using drug-eluting stents (DES) instead of bare metal stents (BMS) for patients with stable angina.

Background: Use of DES when performing percutaneous coronary intervention (PCI) for stable angina reduces the risk of reintervention but increases health care costs. Previous economic analyses may indicate that DES offer value for money, but the analyses have been based on a limited number of clinical trials and short follow-up periods.

Methods: We developed a Markov model that captures costs and outcomes the first two years after PCI with stent for stable angina. After each PCI, patients can either become well, have a new intervention (PCI or CABG), or die. The model was based on meta-analyses of trials comparing DES with BMS. The relative risk of reintervention was 86%, 49%, 36%, and 36% lower at 1, 6, 12, and 24 months follow-up when using DES instead of BMS. We assumed that DES will reduce mortality because of fewer intervention-related deaths, but also explored a potential increased mortality because the meta-analysis indicates a nonsignificant trend toward increased mortality of DES compared to BMS. One-way and Monte Carlo sensitivity analyses were applied.

Results: The estimated cost per avoided reintervention was $5,000 when BMS was replaced by DES, ranging from $200 to $16,000 in one-way sensitivity analyses. The price of a drug-eluting stent would have to be reduced from currently $2,000 to $1,400 to make the use of DES cost saving compared to BMS (current purchasing price $560).

Conclusion: The estimated cost per life year gained and quality adjusted life year gained were $122,000 and $46,000, respectively, when increased mortality was disregarded in the model. Probabilistic sensitivity analysis indicated a 64% probability that drug-eluting stents were cost-effective if society is willing to pay $55,000 per quality-adjusted life year.

When the increased mortality was included, BMS was the dominant strategy, with both lower costs and greater life expectancy.

Conclusion: In a two-year perspective, the DES strategy resulted in greater costs than the RCT strategy with any realistic combination of model parameters. Cost-effectiveness of DES depended heavily on purchasing price of the stents, risk of restenosis, and rate of reintervention in routine practice.

POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
ABSTRACTS
ABSTRACTS

2098
POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
9 Pacific Room EFG

COST-EFFECTIVENESS OF SURGICAL RESSECTION COMPARED WITH LOCAL ABLATION FOR INITIAL TREATMENT OF SOLITARY SMALL HEPATOCELLULAR CARCINOMA WITH HCV HEPATITIS C VIRUS CARRIERS

Abubakar J, Hiro K, Kurokawa F, Nishina S, Okita K and Inoue Y
Yamaguchi University School of Medicine, Ube, Japan

Purpose: The main therapy for small hepatocellular carcinoma (HCC) is surgical resection, but other treatments including percutaneous ethanol injection (PEI) and radiofrequency ablation (RFA) may have a risk of future occurrence. We evaluated the cost-effectiveness of surgical resection and local ablation in the initial treatment of solitary small (≤3 cm) HCC in HCV-related hepatitis and cirrhosis.

Methods: To simulate the process from first treatment to death from liver disease or other causes, we created a state-transition Markov model that represented the principal health states of chronic hepatitis, cirrhosis, and decompensated cirrhosis with or without HCC. We assumed that the treatment choice corresponded to HCC state features such as maximal size, number, and location; that treatment would start with resection or local ablation, including PEI and RFA; and that, as HCC progressed, local ablation plus transarterial chemoembolization (TACE), TACE, hepatic arterial infusion chemotherapy (HTAC), and conservative treatment would be applied. We retrospectively reviewed 276 HCC cases admitted to our institution, including 93 of small solitary HCC, and, for each treatment strategy, estimated the rate of complete remission (CR) and the rates among CR cases of HCC-free survival and treatment for recurrent HCC. We also estimated the rates of transition between treatments in non-CR cases. We applied data on health insurance reimbursement to estimate the annual cost of each health state and the admission cost of each treatment.

Results: Our model was validated and consistent with the 5-year survival rate from Kaplan-Meier’s analysis of a sample population at our institution and of a nationwide survey of Japan. For patients hypothesized identical with these of the sample population regarding mean age (65 years old), gender ratio (73% male), and cirrhosis (93%), surgical resection increased lifetime treat- ment cost by about $14,000 and extended average survival by about 1.4 years, yielding an incremental cost-effectiveness rate of $10,200 per life-year compared with local ablation therapy. The sensi- tivity analysis showed that only if the CR rate after resection was less than 0.57 or the death rate following resection was more than 0.1 did incremental cost-effectiveness exceed the threshold of cost-effectiveness ($50,000/LY).

Conclusions: Our analysis suggests that surgical resection is more cost-effective than local ablation for small HCC. However, this may reflect its preferential selection in cases with better functional hepatic reserve.

2340
POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
10 Pacific Room EFG

ASSESSING COST-EFFECTIVENESS AND VALUE CREATION FOR TARGETED THERAPY IN NON-SMALL CELL LUNG CANCER

Vosstraës D1, Hemmink B1, Garrison L2 and Ramsey S2
1University of Washington, Seattle, WA; 2Fred Hutchinson Cancer Research Center/University of Washington, Seattle, WA

Purpose: Recent clinical studies indicate that patients with advanced non-small cell lung cancer (NSCLC) whose tumors overexpress EGFR may respond better to EGFR tyrosine kinase inhibitors such as gefitinib and erlotinib. Targeted therapy based on EGFR status may thus influence clinical practice. The purpose of this study was to estimate the clinical and economic outcomes associated with a targeted therapy approach and to assess the value created via utilization of a diagnostic test.

Methods: We developed a disease simulation model using decision analytic techniques. Data reported in the package insert based on a randomized controlled trial of gefitinib versus best supportive care (BSC) in 711 patients with locally advanced or metastatic NSCLC after failure of at least one chemotherapy regimen were used to inform the model. Adverse event and disease progression costs were estimated from the literature. We estimated life expectancy and direct medical costs for a targeted versus a non-targeted strategy, and calculated incremental cost-effectiveness ratios (ICERs). Transition probabilities were derived from survival curves observed in the trial when available; otherwise, exponential functions were used based on median time to progression. Sensitivity analyses were conducted to assess the uncertainty in our results.

Results: For all patients in the clinical trial, treatment with gefitinib versus BSC increased life expectancy by 0.13 years and an additional cost of $12,492, thus yielding an ICER of $83,000/LY. In the subgroup of patients with known EGFR status (N = 238, 33%), targeted therapy (erlotinib and gefitinib patients) versus non-targeted therapy (erlotinib for all patients) resulted in decreased costs of $5,096 and a negligible difference in life expectancy; targeted therapy vs. BSC resulted in an ICER of $54,000/LY. The prevalence of EGFR(+)-tumors was one of the most influential parameters in the analysis.

Conclusions: Our analysis indicates utilizing an approximately 80% EGFR expression assay result in a cost savings of $5,096 per patient. Because the initial prices for the diagnostic and thera- peutic were based on on-market therapies, patients and insurers may be able to capture a substantial share of the cost savings. For patients with known EGFR status, targeted therapy based on EGFR status plus any one of the follow-up therapies (surgical resection, neoadjuvant chemotherapy, and incorporating traditional 2nd-line chemotherapy for EGFR(+) patients) will be important.

2089
POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
12 Pacific Room EFG

COST-EFFECTIVENESS ANALYSIS OF RISPERIDONE LONG-ACTING INJECTION (RISPERDAL CONSTA®) FOR THE TREATMENT OF PATIENTS WITH SCHIZOPHRENIA WHO ARE PARTIALLY MEDICATION ADHERENT IN AUSTRALIA

Schouwen B1, Crowley S2 and Neville M2
1Janssen-Cilag Australia Pty Ltd, North Ryde, Australia; 2Pretium, Sydney, Australia

Objectives: Risperidone long-acting injection (risperidone LAI) is the first long-acting atypical injectable that leads to improved medication adherence that is associated with lower recurrences of symptoms than current management. The purpose of this study was to compare the cost-effectiveness of risperidone LAI with oral risperidone, oral olanzapine, and typical depot injections in pa- tients with schizophrenia that are partially adherent to their medication. The analysis is based on the cost and outcomes incurred in the specialized public psychiatry setting. The perspective of the analysis was the Australian health care system.

Methods: A 1-year decision analytic model was developed using probabilistic sensitivity analy- sis to explore uncertainty. The outcomes used in the cost-effectiveness analysis (CEA) were relapses avoided, life years gained, deaths due to suicide, and quality-adjusted life years (QALY) gained. Local resource utilization focused on both cost of hospitalization and other community-care costs such as pharmaceuticals and outpatient consultations. Clinical trial data and other epidemi- ological assumptions based on the published literature were used to elicit the event probabilities in the decision analytic model. Utilities for each of the possible health states in the model were based on symptom severity and drug-related adverse events. These data were derived from the Australian general population using the Assessment Quality of Life (AQLQ) utility instrument.

Results: Against the weighted comparator, risperidone LAI was dominant. That is, the in- creased cost of risperidone LAI was more than offset by the reduction in hospitalizations and other costs associated with symptom recurrence. Analysis versus oral risperidone and oral olanzapine also showed that risperidone LAI was dominant. The incremental cost-effectiveness ratio (ICER) against typical depot produced a less favorable result due to the low acquisition cost of these agents. Probabilistic sensitivity analyses showed a 100% likelihood of an ICER of less than $50,000 per QALY.

Conclusions: Risperidone LAI represents a cost-effective intervention against current manage- ment for patients who are partially adherent to their medication. The introduction of this product in Australian clinical practice will result in significant clinical and economic benefits to the commu- nity because partial adherence to medication is a major reason for relapse of symptoms in patients with schizophrenia.
IMPLICATIONS FOR INTENSIVE THERAPY FOR DIABETES

EFFECTS OF PATIENT SELF-SELECTION ON COST-EFFECTIVENESS: IMPLICATIONS FOR INTENSIVE THERAPY FOR DIABETES

THE COST-EFFECTIVENESS OF ALTERNATIVE STRATEGIES FOR THE MANAGEMENT OF CHRONIC NON-MALIGNANT PAIN: BAYESIAN EVIDENCE SYNTHESIS AND PROBABLISTIC DECISION MODEL

METHODOLOGY

RESULTS

CONCLUSIONS

MEDICAL DECISION MAKING/JAN–FEB 2006

POSTER SESSION III

POSTER SESSION III

POSTER SESSION III
Purpose: Whether the expectation of compensation for morbidity costs has an influence on assessments of quality of life (QOL) has not been established. If this expectation affects QOL, then QOL assessments that ignore any expected compensation, when compensation exists, will be biased.

Most studies assume that compensation has no influence on assessments of QOL. This analysis provides empirical evidence and suggests that the expectation of compensation for morbidity costs influences assessments of QOL. The results show the importance of linking compensation for morbidity costs to QOL assessments.

Methods: 126 undergraduate students were randomly assigned to one of two groups: 1) participants who were informed they would receive compensation for time not at work or 2) participants who were not informed. They were further randomly assigned a level of illness severity (mild, moderate, or severe). Participants were asked to read a description of a health state and to assign an assessment of the QOL for that health state using the paper standard gamble. Two-factor analysis of variance was performed on the completed instruments.

Results: The overall mean QOL of those participants who were informed they would receive compensation is significantly higher than those participants who were not informed who would receive compensation (p = 0.0005, F = 6.37, df = 3, 122).

Conclusion: An individual’s assessment of quality of life is influenced by whether compensation for morbidity costs is expected. Therefore, to accurately estimate the quality of life of illnesses or health interventions with linked morbidity costs, expectations of compensation should be addressed.

Cost-effectiveness studies of HA vaccine were identified (MEDLINE, EMBASE, HSTAR, and SSIC; MSNH “cost-effectiveness” AND “hepatitis A”) and included if they were cost- effectiveness/utility studies of HA vaccine. Citations and full-text articles were reviewed independently by two reviewers. Back referencing, author searches, and expert consultation ensured literature saturation.

Quality of reporting was assessed using a 21-item quality tool (Neumann, 2000). Methodological issues specific to vaccine evaluations were appraised by Beutels’ 2002 guidelines. Key modeling issues were examined using Sculpher’s 2000 framework.

Results: Thirty-four cost-effectiveness studies were included from full-text-article review (n = 95) and citation-sampling (n = 796). Nine conducted a cost-utility analysis and six a cost-effectiveness analysis. All were model-based studies, thirteen utilizing Markov models. One used a dynamic model, capturing effects of herd immunity. Strategies included mass childhood/adult vaccination (n = 3), selective childhood/adult vaccination (n = 2), and selective childhood vaccination (n = 240) = 8). Populations assessed were healthy participants (n = 31) and patients with chronic liver disease (n = 3). The median quality using Neumann’s tool was 62% (13/21 range [14%-90%]), similar to the median quality of 58% reported for CEA studies in other clinical areas. Cost-utility studies attained higher quality (n = 9, median 76% [57%-90%]) than cost-effectiveness (n = 25, median 52% [14%- 75%]).

Cost-utility studies assessed using Beutels’ guidelines clearly stated model assumptions (79%), utilized proper time span (79%), and captured relevant costs (78%). Many did not discuss alternative modeling approaches (79%).

Using Sculpher’s framework, cost-utility studies specified model assumptions (9/9), and most justified model parameters (7/9). Some did not discuss implications of relaxing model assumptions (3/9). None adjusted for herd immunity, though four acknowledged this as a study limitation. One study adjusted for underreporting, though five studies discussed it.

Conclusions: Methods used to evaluate the cost-effectiveness of HA vaccination were inconsistent, resulting in variable cost-effectiveness ratios across studies. Failing to adjust for herd immunity among underreporting weakens scientific evidence and may result in biased estimates of the economic attractiveness of HA vaccination policies.
IMPACT OF OBESITY ON WORKFORCE PARTICIPATION AMONG US ADULTS

Bradley C, O’Hear K and Schenck M
Virginia Commonwealth University, Richmond, VA; Michigan State University, East Lansing, MI; Wayne State University, Detroit, MI

Purpose: We examined the number of days of breast or prostate cancer patients were absent from their jobs while undergoing cancer treatment.

Methods: Patients (n = 283 breast patients and n = 207 prostate patients) who were employed by large employers, and 82% had paid sick leave. Only 43% had jobs with physical demands. Eighty-two percent reported missing 1 or more days, mean days missed from work was 27. Men with prostate cancer who were treated surgically missed the greatest number of days relative to men who received radiation. In addition, a college education and being employed by small to medium employers were associated with fewer days missed from work.

Conclusions: Women with later stage disease missed more days from work than women with in situ cancer, lending support to an economic argument for early stage detection. Men with prostate cancer who were treated surgically missed the greatest number of days relative to men who received radiation.

POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
22 Pacific Room EFG

EVALUATION OF ALGORITHMS TO IDENTIFY BREAST CANCER CASES IN MEDICARE CLAIMS DATA

Gold H and Do H
Weill Cornell Medical College, New York, NY

Purpose: To test the generalizability of published algorithms designed to identify incident breast cancer cases in Medicare claims data for conducting analyses of health care service utilization.

Methods/Date: We use the Surveillance, Epidemiology, and End Results (SEER) registry data linked with Medicare physician, hospital, and outpatient claims data for breast cancer cases in 1998 and 1999. A 5% control sample of Medicare beneficiaries in SEER areas (n = 70,000; 13% cases) SEER (1982-1997) gold standard for case identification. Each algorithm uses different combinations of diagnosis and procedure codes to classify cases. We apply three algorithms to our data and evaluate the sensitivity and specificity of each compared to reported results from earlier data used for algorithm development. We compare algorithm performance by age, stage, race, and SEER region, and test via logistic regression whether adding demographic variables to the algorithm improves area under the receiver operating characteristic (ROC) curve.

Results: Sensitivity of two of these algorithms applied to the data is significantly lower than that obtained by the algorithm developers. However, sensitivity decreases as age increases (range for 80+yo: 51.4%-74.4%). Sensitivity is lower for cases with in situ and metastatic disease compared to Stage 1 or 2 disease. There also is substantial variation by SEER registry, but differences by race are insignificant. Specificity overall is about the same. Adding age, region, and race variables to the algorithm significantly improves the ROC area (p < 0.001).

Conclusions: Algorithm sensitivity is lower for the 1998 data, indicating that published algorithms may need to be updated due to changing patient characteristics or patterns of care. Differential sensitivity by SEER region likely reflects geographic variation in practice patterns. Depending on the algorithm, 3%-5% of subjects are misclassified in 1998, with false negatives highest in Freeman’s algorithm and lowest using Nottinger’s method. Misclassification disproportionately affects older women and those diagnosed with in situ and metastatic disease. Algorithms should be applied cautiously to insurance claims databases to assess health care utilization and costs of breast cancer care outside SEER-Medicare populations because of misclassification bias.

ABSTRACTS
ABSTRACTS

2314
POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
25 Pacific: Room EFG

HOW COVERAGE DECISIONS INFLUENCE HEALTH CARE


University of California-San Francisco, San Francisco, CA; Harvard, Boston, MA.

Purpose: We used the example of colonoscopy to explore how coverage decisions influence health care. The Medicare program did not cover screening colonoscopy for individuals at average risk for colorectal cancer until 2001, and many private insurers do not cover such screening. Previous studies showed a dramatic increase in colonoscopy use after the Medicare coverage, although these studies did not use nationally representative data, did not control for temporal trends, and did not examine the resultant costs and benefits.

The purpose of this study was to 1) Examine trends in colonoscopy use and insurer coverage between 2000 and 2003, during which time Medicare began covering colonoscopy for screening purposes; 2) Examine the impact of these trends on health outcomes and costs; 3) Explore the role of insurer reimbursement and health policies—particularly Medicare coverage—in facilitating use of preventive health services.

Methods: We used nationally representative data from the 2000 and 2003 National Health Interview Surveys and a previously developed cost-effectiveness model (1 Liadamou et al).

Results: We found that colonoscopy use increased from 23% in 2000 to 33% in 2003 in the Medicare population and from 17% to 27% in the insured non-Medicare population (p = 0.01 for both). Overall screening utilization and adherence to guidelines also increased. We found that colonoscopy use was significantly higher for Medicare enrollees than non-Medicare, insured individuals in 2003 (p < 0.001) but not in 2000 (p = 0.11). The observed use of screening in 2003 was relatively more cost-effective than in 2001, as it improved health outcomes, although at a higher cost.

Conclusions: We found that insurer coverage may increase the use of preventive health care services. Our findings raise important issues for other health care services about how such coverage decisions are made. Our presentation will also discuss the intersection of science and politics in coverage decisions.

1991
POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
26 Pacific: Room EFG

USE OF AGGRESSIVE MEDICAL TREATMENTS NEAR THE END OF LIFE: DIFFERENCES BETWEEN PATIENTS WITH AND WITHOUT DEMENTIA

Richardson S, Sullivan C, Hill A and Yu W

VA Palo Alto Health Care System, Menlo Park, CA; University of Arkansas for Medical Science, N. Little Rock, AR

Purpose: We analyzed whether acute care patients with dementia were more or less likely than patients without dementia to receive aggressive medical services near the end of life.

Methods: We identified all 169,036 VA patients over age 67 who died in fiscal year 2000 or 2001, and studied trends of all their VA and Medicare inpatient and outpatient utilization. To identify patients with dementia, we adapted the dementia classification method developed by the VA Dementia Registry Task Force. We measured aggressiveness of medical services using ICD admission and four surgical procedures identified by a panel of physicians for a previous study: ventilation, pulmonary artery monitor, cardiac catheterization, and dialysis. We compared the use of each aggressive treatment between acute hospital patients with dementia and those without dementia using both bivariate analyses and Probit regressions controlling for age, sex, race, principal diagnosis, and Charlson Comorbidity Score. Our primary analysis investigated care received in the final 30 days of life, but we also analyzed data for care in the final 90 days and final year of life.

Results: We identified 31.8% of decedents as having dementia, with higher prevalence among older patients, women, and blacks. In bivariate analyses stratified by age, we found that across age groups, patients with dementia were less likely to receive each aggressive treatment during a given medical/surgical hospital stay. Probit regressions found that patients with dementia were substantially less likely to receive each aggressive treatment, controlling for other factors. The Probit model predicted patients with dementia to be 20% less likely to be admitted to the ICU; 32% less likely to be placed on ventilator support; 26% less likely to receive cardiac catheterization; 53% less likely to receive pulmonary artery monitoring, and 14% less likely to receive dialysis, compared to patients without dementia. The results were similar when including care in the final 90 days and final year of life, and when limiting analyses to only VA or only Medicare.

Conclusion: Our study demonstrated a substantial disparity between the aggressiveness of in-hospital end-of-life care for patients with dementia compared to those without dementia. These differences may be related to patient preferences, or they may reflect inappropriate care. Such striking differences warrant further research into whether the observed disparity represents appropriate care.

1865
POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
28 Pacific: Room EFG

ANALYSIS OF CORE CONCEPTS WITHIN THE SOCIETY OF MEDICAL DECISION MAKING

Ozanne E and Stahl J

Massachusetts General Hospital, Boston, MA.

Purpose: This study aims to describe core concepts within the Society for Medical Decision Making (SMDM) as perceived by its members. Specifically, we look at the skills required by trainees, the concepts characterizing member's work, and concepts characterizing SMDM as a whole.

Methods: We conducted a Web-based survey. Eligible participants were members of SMDM, whose total was 813 when the survey was initiated, March 2005. Members were contacted by email and asked to complete an anonymous online survey with 20 open-ended questions. Multiple responses were made, so each member contributed up to 20 responses.

Results: A total of 8,009 RA, 7,200 IBD, 47,837 COPD, 6,093 CKD, 40,857 cancer, and 13,041 CHF patients were identified. Within each disease, patients with moderate/severe conditions incurred substantially higher average annual health care costs than those with mild condition identified by the severity staging, with an exception of CHF population. Moderate/severe RA patients (8%) incurred $25,292 compared to $10,678 by patients with mild condition. Moderate/severe IBD patients (10.9%) incurred $37,925 compared to $10,687 by patients with mild condition. Moderate/severe COPD patients (8.8%) incurred $47,223 compared to $16,658 by patients with mild condition. Moderate/severe CKD patients (41.2%) incurred $85,143 compared to $24,202 by patients with mild condition. Moderate/severe CHF patients incurred $82,622 compared to $40,850 by patients with mild condition. The differences in costs were statistically significantly different between severity levels (p < 0.05).

Conclusions: Disease-specific severity staging that identifies the level of a patient's condition can be applied as an adjustment and/or a predictive parameter in future retrospective or predictive modeling studies. Severity staging could be tailored to differentiate diseases so that different criteria are selected for each disease. The current study developed sets of criteria that would differentiate severity within the six selected chronic diseases. Although patients of moderate/severe level were a smaller proportion of the disease population, the costs incurred were significantly higher with an exception of CHF population in this study.
ABSTRACTS

POSTER SESSION III
February 24, 2005

1:00 PM - 3:00 PM
2039

COST-EFFECTIVENESS AND COST-BENEFIT OF SCREENING NEWBORNS FOR CONGENITAL ADRENAL HYPERPLASIA: DIFFERENT METHODS, DIFFERENT ANSWERS?

Yoon H and Groose S

Centers for Disease Control and Prevention, Atlanta, GA

Purpose: We conducted both benefit-cost analysis (BCA) and cost-effectiveness analysis (CEA) of screening infants for classic congenital adrenal hyperplasia (CAH), in line with an Office of Management and Budget (MBM) recommendation. Currently, newborn screening for CAH has been implemented uniformly across all states in the United States.

Methods: The primary justification for newborn screening for CAH is the prevention of acute salt-wasting crises that carry an elevated risk of death. Since CAH is a rare disorder (1-1 in 20,000) and documented CAH mortality is uncommon (<1%), the number of deaths prevented by screening newborns for CAH is not known. For a range of estimates of mortality, we calculated incremental cost-effectiveness ratios (ICERs) and threshold levels of value of statistical life (VSL) required to make net benefit equal to zero. Because of uncertainty regarding the mortality rate and treatment effects, probabilistic CEAs were conducted. All estimates are reported in SUS 2000 values using a 3% discount rate.

Results: Our results indicate that the ICER of CAH screening varies with prevalence as well as mortality rate and treatment effectiveness. In a probabilistic analysis in which mortality in untreated salt-wasting CAH is assumed to vary between 2% and 9%, the ICER is $120,000 per discounted life-year saved and the threshold VSL is $3.6 million. In a base case scenario, assuming 9% mortality and 95% treatment effectiveness, the ICER is $49,000 per life-year and the threshold VSL is $8.1 million.

Conclusions: We conclude that screening for CAH may not meet a cost-effectiveness threshold of between $50,000 and $100,000 per discounted life-year saved, except under favorable assumptions. On the other hand, for VSL of $4.0 million or greater, screening for CAH is likely to yield net economic benefit. Thus, different methods of economic evaluation, associated with differing cutoffs, can yield divergent conclusions. A complete economic evaluation of newborn screening for CAH would require data on preferences relative to other health states. In particular, early detection of virilization through the ability to prevent short stature is important screening outcomes, and their inclusion could improve estimates of value for money from CAH newborn screening, whether through willingness-to-pay in BCA or health-related quality of life in CEA.

1998

POSTER SESSION III
February 24, 2005

1:00 PM - 3:00 PM
2170

DEVELOPMENT AND PILOT TESTING OF A COMPOSITE SCALE IN STRESS URINARY INCONTINENCE TO ASSESS THERAPEUTIC RESPONSE IN CLINICAL PRACTICE AND CLINICAL RESEARCH

Arnould D1, Rosio-Le Gou G2, Cadilhac-Sous N2, Droumaguet C2, Heutebelin N2 and Costa P2

Mapgi Values, Lyon, France; Lilly France, Suresnes, France; Hospital Garemau, Nimes, France

Purpose: Clinicians in charge of Stress Urinary Incontinence (SUI) patients miss specific tools supporting pragmatic treatment decision taking into account the complexity of perceived handicap and patients’ expectations. Our aim was to develop a standardized global medical judgment scale to assess therapeutic response and support medical decision.

Methods: The concepts useful to assess SUI treatment outcomes were identified from two independent sources: literature review and clinicians’ experience (30 open interviews). A test version of the Pragmatic Assessment Form (PAF) was developed. Eight clinicians comprehensively covering various perspectives of professional medical careers in SUI were interviewed, following a semistructured interview guide assessing the relevance, comprehensiveness, and validity of the concepts covered; accessibility to source information; clarity of item wording and response choice; easiness to use in clinical practice; applicability to the various situations in real life; and appropriate level of guidance to form global medical judgment. The PAF was updated according to the comments received. Content validity was further assessed by 10 SUI patients, to quote the concepts covered by the PAF for relevance, reliability, importance, and applicability in routine clinical practice.

Results: Eight broad concepts covering symptom and functional status, impact on activities of daily life, well-being, coping, satisfaction, and expectations were identified as relevant independent outcomes and included in the test PAF. Out of 8 clinicians interviewed, 7 found the PAF complete, relevant, and specific to SUI, clear and valid to support clinical judgment. The clinicians were happy to have an assessment guide to support judgment but did not wish to have a formal scoring procedure to support decision. The balance between treatment benefit, undesirable effects, and constraints was made more explicit. An additional item describing patients’ intentions was developed. After final validation by 10 SUI patients, the pilot version of the PAF was issued in a double A4 size, including instructions for use.

Conclusions: This new instrument is a short, simple, pragmatic composite tool that aims at helping the clinicians to easily manage complexity of patients’ perspective. It provides a comprehensive coverage of relevant outcomes in SUI to assess therapeutics and support medical decision making. This instrument allows two levels of concept elicitation, one pragmatic, suitable for clinical practice, and one standardized, suitable for clinical research. Its properties will be assessed in a specific validation study.

E72 • MEDICAL DECISION MAKING/JAN–FEB 2006
dence from a relatively small clinical trial did not show improved effectiveness of CAS compared to CEA. In contrast, more recent published results obtained from surveys and registries indicate substantial clinical benefits for CAS. To enable estimation of the value of additional clinical trials, a value of information analysis was performed using the outcome of an expert opinion assessment, which provided a quantitative estimate by specialists in the field of currently available evidence.

Methods: An expert opinion survey was sent to members of the European endovascular network requesting point estimates and uncertainty margins for expected procedural complication rates. The expert opinion responses were combined into one (unweighted) overall expert opinion and used in a Markov model that describes patient health following carotid revascularization. Model parameters are based on actual trial data, published complication rates for CAS, and collected cost data. The model enables evaluation of the cost-effectiveness of CAS compared to CEA and was used to evaluate the value of information of procedural complication rates.

Results: The survey response was low (only 11%, representing 12 experts), and the estimates provided by the experts varied considerably, reflecting the uncertainty with respect to the expected procedural complication rates. As a result, the combined expert opinion shows a large uncertainty range that encloses all complication rates found in the literature. Whereas the EVPI of procedural complication rates for the most favorable and largest published series of complication rates for CAS is no more than $30 per procedure, the EVPI based on the expert opinion estimates is $2890.

Conclusions: Although performing an expert judgment assessment involves a significant amount of work and still provides a subjective rather than an objective interpretation of the available evidence, its use and subsequent value of information analysis is feasible and provides an elegant expert-based quantification of the value of information.

ABSTRACTS

2288
POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
33 Pacific Room EFG

EXPERT OPINION IN VALUE OF INFORMATION ASSESSMENT

Janssen M and Buskens E

University Medical Center Utrecht, Utrecht, Netherlands

Purpose: Stenting of the artery carotid interna (CAS) for treatment of carotid stenosis is increasingly applied. Costs of stents are high, and conclusive trials to provide evidence on the (cost-) effectiveness of CAS as compared to carotid endarterectomy (CEA) have not yet finalized. The only evidence from a relatively small clinical trial did not show improved effectiveness of CAS compared to CEA. In contrast, more recent published results obtained from surveys and registries indicate substantial clinical benefits for CAS. To enable estimation of the value of additional clinical trials, a value of information analysis was performed using the outcome of an expert opinion assessment, which provided a quantitative estimate by specialists in the field of currently available evidence.

Methods: An expert opinion survey was sent to members of the European endovascular network requesting point estimates and uncertainty margins for expected procedural complication rates. The expert opinion responses were combined into one (unweighted) overall expert opinion and used in a Markov model that describes patient health following carotid revascularization. Model parameters are based on actual trial data, published complication rates for CAS, and collected cost data. The model enables evaluation of the cost-effectiveness of CAS compared to CEA and was used to evaluate the value of information of procedural complication rates.

Results: The survey response was low (only 11%, representing 12 experts), and the estimates provided by the experts varied considerably, reflecting the uncertainty with respect to the expected procedural complication rates. As a result, the combined expert opinion shows a large uncertainty range that encloses all complication rates found in the literature. Whereas the EVPI of procedural complication rates for the most favorable and largest published series of complication rates for CAS is no more than $30 per procedure, the EVPI based on the expert opinion estimates is $2890.

Conclusions: Although performing an expert judgment assessment involves a significant amount of work and still provides a subjective rather than an objective interpretation of the available evidence, its use and subsequent value of information analysis is feasible and provides an elegant expert-based quantification of the value of information.

ABSTRACTS

2288
POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
33 Pacific Room EFG

EXPERT OPINION IN VALUE OF INFORMATION ASSESSMENT

Janssen M and Buskens E

University Medical Center Utrecht, Utrecht, Netherlands

Purpose: Stenting of the artery carotid interna (CAS) for treatment of carotid stenosis is increasingly applied. Costs of stents are high, and conclusive trials to provide evidence on the (cost-) effectiveness of CAS as compared to carotid endarterectomy (CEA) have not yet finalized. The only evidence from a relatively small clinical trial did not show improved effectiveness of CAS compared to CEA. In contrast, more recent published results obtained from surveys and registries indicate substantial clinical benefits for CAS. To enable estimation of the value of additional clinical trials, a value of information analysis was performed using the outcome of an expert opinion assessment, which provided a quantitative estimate by specialists in the field of currently available evidence.

Methods: An expert opinion survey was sent to members of the European endovascular network requesting point estimates and uncertainty margins for expected procedural complication rates. The expert opinion responses were combined into one (unweighted) overall expert opinion and used in a Markov model that describes patient health following carotid revascularization. Model parameters are based on actual trial data, published complication rates for CAS, and collected cost data. The model enables evaluation of the cost-effectiveness of CAS compared to CEA and was used to evaluate the value of information of procedural complication rates.

Results: The survey response was low (only 11%, representing 12 experts), and the estimates provided by the experts varied considerably, reflecting the uncertainty with respect to the expected procedural complication rates. As a result, the combined expert opinion shows a large uncertainty range that encloses all complication rates found in the literature. Whereas the EVPI of procedural complication rates for the most favorable and largest published series of complication rates for CAS is no more than $30 per procedure, the EVPI based on the expert opinion estimates is $2890.

Conclusions: Although performing an expert judgment assessment involves a significant amount of work and still provides a subjective rather than an objective interpretation of the available evidence, its use and subsequent value of information analysis is feasible and provides an elegant expert-based quantification of the value of information.

ABSTRACTS

2288
POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
33 Pacific Room EFG

EXPERT OPINION IN VALUE OF INFORMATION ASSESSMENT

Janssen M and Buskens E

University Medical Center Utrecht, Utrecht, Netherlands

Purpose: Stenting of the artery carotid interna (CAS) for treatment of carotid stenosis is increasingly applied. Costs of stents are high, and conclusive trials to provide evidence on the (cost-) effectiveness of CAS as compared to carotid endarterectomy (CEA) have not yet finalized. The only evidence from a relatively small clinical trial did not show improved effectiveness of CAS compared to CEA. In contrast, more recent published results obtained from surveys and registries indicate substantial clinical benefits for CAS. To enable estimation of the value of additional clinical trials, a value of information analysis was performed using the outcome of an expert opinion assessment, which provided a quantitative estimate by specialists in the field of currently available evidence.

Methods: An expert opinion survey was sent to members of the European endovascular network requesting point estimates and uncertainty margins for expected procedural complication rates. The expert opinion responses were combined into one (unweighted) overall expert opinion and used in a Markov model that describes patient health following carotid revascularization. Model parameters are based on actual trial data, published complication rates for CAS, and collected cost data. The model enables evaluation of the cost-effectiveness of CAS compared to CEA and was used to evaluate the value of information of procedural complication rates.

Results: The survey response was low (only 11%, representing 12 experts), and the estimates provided by the experts varied considerably, reflecting the uncertainty with respect to the expected procedural complication rates. As a result, the combined expert opinion shows a large uncertainty range that encloses all complication rates found in the literature. Whereas the EVPI of procedural complication rates for the most favorable and largest published series of complication rates for CAS is no more than $30 per procedure, the EVPI based on the expert opinion estimates is $2890.

Conclusions: Although performing an expert judgment assessment involves a significant amount of work and still provides a subjective rather than an objective interpretation of the available evidence, its use and subsequent value of information analysis is feasible and provides an elegant expert-based quantification of the value of information.

ABSTRACTS
Purpose: To develop a decision-making model based upon users’ judgments for identifying high-priority dimensions concerning the information and communication technology (ICT) support in quality of health service and is impossible to introduce them all. A multi-criteria decision model was constructed using the Analytic Hierarchy Process (AHP) to reflect ICT deployment in Chilean hospitals. A comparison process is pursued based on decisive factors and users’ preferences to rank the ICT support for health-related activities. Results: The AHP application allows integrating diverse judgments and preferences to achieve an overall result. The prioritization indicates ICT support has superior impact over supplying clinical care service (41%), and fixed network system (37%). From the patient perspective, the results indicated that patient care service (54%), a 28% over medical research, and least impact over the patients. Today, the priority utilization is for the support from computer-related system (Internet, e-mails, Web, etc.) (41%) and fix network system (37%). From the patient perspective, the results indicated that patient priority are mainly involved with urgent service requirement (64%). In this concern, the utilization of fix network system has the highest priority (54.8%). Concerning the gender sensitivity for patients in urgency requirement, wireless technology tends to increment (20.9%). From administrative perspective, the activities such as delivering test and exams results, within the institution or external, has the highest priority. A strong usage of fix network system (phone, fax, extensions, etc.) is detected. From medical research perspective, a strong interaction with database applications implies an increasing demand for computer-assisted support is visualized (55%). Conclusions: The AHP outcomes reveal the main ICT agents/users, the main activities involved in, and the critical needs for ICT support in a hospital requirement. This pilot study concludes that combinations of wireless networks give relevant and timely information for better decisions. The most important factors affecting quality in hospital requirement are availability of services and the need for ubiquitous access to integrated information. The methodology helps decision makers to elaborate course of action for resource distribution proportional to users requirement.
ABSTRACTS

2429
POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
41 Pacific Room EFG

MODELING RESOURCES NEEDED TO IMPROVE COLORECTAL CANCER SCREENING RATES

Bachman S1, Levin T1 and Santamaria B1
Kaiser Permanente, Oakland, CA; Kaiser Permanente, Northern California, Oakland, CA

Purpose: To develop a model to determine colonoscopy and sigmoidoscopy capacity needs for a 6-year period beginning in January 2005, with particular attention paid to quality and access goals.

Method: Kaiser Permanente Northern California adopted use of colorectal cancer screening rates as a priority in 2005. To quantify the additional resources required to improve access to colonoscopy and sigmoidoscopy, we developed a model estimating the demand associated with a target rate of colorectal cancer screening of 70%. This demand was used to estimate the number of MD and procedural assistant full-time equivalent (FTEs), procedure rooms, and recovery beds needed each year from 2005-2010. The model is based on expert assumptions regarding productivity and the proportion of the population requiring or choosing colonoscopy.

The tool is designed to apply at the regional level and at the medical center level, based on population projections specific to each area. An important feature is that several assumptions in the model can be varied in real-time by end-users. This enables local quantification of the impact of variation in productivity from the regional norm or changes in demand patterns—for example, increased preference for colonoscopy. Coupled with ongoing monitoring of productivity, positivity rates, and patient choice, the tool can be used to quickly update projections for ongoing planning. Another feature that could be added to the model is the inclusion of FOBt screening and its effect on the demand for FTEs.

Results: For a health system of 3.2 million people, including 844,000 people over the age of 50, we estimate that improving colorectal cancer screening rates will require an investment of 18 MD FTEs and 26 non-MD FTEs for colonoscopy, 11 MD/RN FTEs and 16 non-MD FTEs for sigmoidoscopy, 41 recovery beds, and 15 procedure rooms, in the first year of the program. These projections were used to determine the amount of funding to be made available to improve capacity to meet screening goals. The model is also being used by the Endoscopy Department chiefs to support the development of plans to effectively apply this funding.

Conclusion: Providing a specific, but flexible tool to quantify demand and capacity for colorectal cancer screening provides a rational approach to resource allocation to support access and quality improvement.

2434
POSTER SESSION III
7:00 AM - 8:30 AM
Monday, October 24, 2005
43 Pacific Room EFG

PERCEIVED EVIDENCE BASE AND LEADERSHIP ATTITUDES TOWARD SCREENING COLONOSCOPY

Kocherl L1 and Khurana V1
Minneapolis VAMC, Minneapolis, MN; Shreveport VAMC, Shreveport, LA

Purpose: This study seeks to understand leading GI specialists’ opinions of screening colonoscopy (SC) as a function of their perception of the evidence base and perceived barriers to implementation.

Method: Thirty-six GI specialists attending a VHA GI leadership meeting at the Digestive Disease Week 2004 conference participated in this study. 38.9% were GI staff physicians, 53.8% were chiefs of GI service, and 1 participant was a chief of medicine. Participants were given remote response keypads and viewed questions project on the auditorium screen. We asked 56 questions covering screening practices, perceived benefits, and consequences of SC; attitudes toward screening modalities, the evidence base, and perceived barriers to implementation of SC. Questions used a ten-point response scale.

Results: Three items assessed attitude toward the SC evidence base: perceived adequacy of evidence for effectiveness, safety, and cost-effectiveness. The scale had high internal consistence (alpha = .91). 42% of participants had attitude scores of 6 or more, indicating perceived insufficiency of SC evidence. Perceived insufficiency is significantly correlated with belief that screening modality must be individualized to the patient (rho = .59); less favorable attitude toward offering screening colonoscopy in the VA (rho = .40); perceived cost-effectiveness (rho = .70); anticipated impact on screening rates, colorectal cancer prevention, early detection, and mortality (rho = .41, .44, .49, .49), concerns over safety (rho = .49), perceived benefit of a 10-year rescreening interval (rho = .44), and perceived benefit of the ability to conduct polypectomy during colonoscopy (rho = .43). Regardless of attitude toward the evidence base, participants agreed that implementation of SC within the VA, under current resource constraints, would have negative consequences for adverse outcomes; for all procedures, the ability to conduct surveillance colonoscopy and screen high-risk individuals, and the overall cost of colorectal cancer screening.

Conclusions: Using colonoscopy to screen asymptomatic, average risk persons for colorectal cancer is becoming the standard of care in many communities, despite weaknesses in the evidence base. Social and economic pressures to improve colorectal cancer screening rates are compelling, yet resource constraints, the evidence base, and perceived barriers to implementation threaten the success of public screening programs. Future research should focus on development of strategies to improve awareness, knowledge, and acceptance of colorectal cancer screening among the public.

ABSTRACTS E75

Downloaded from http://mdm.asapaub.com by guest on June 6, 2007
© 2006 Society for Medical Decision Making. All rights reserved. Not for commercial use or unauthorized distribution.
PAMMEXAMINATION OF WHAT HEALTHCARE PROVIDERS SAY AND WHAT PATIENTS HEAR ABOUT ALCOHOL CONSUMPTION AND HEPATITIS C (HCV)


The Cleveland Clinic Foundation, Cleveland, OH; Case Western Reserve University, Cleveland, OH; MetroHealth Medical Center, Cleveland, OH; Case Western Reserve University, C, OH

Purpose: Clear communication is essential for effective shared decision making. Reducing con-
sumption of alcohol is recommended for patients with HCV, yet we know little about what is said and heard regarding this topic in the provider/patient encounter.

Methods: We conducted focus groups and interviews about alcohol use and HCV with both non-
dependent drinkers (as determined by the AUDIT) with HCV (N = 62) and health care providers (N = 14) who care for HCV patients at a metropolitan teaching hospital. All focus groups and inter-
views were audio-taped, transcribed, and analyzed using NVivo, a qualitative data management and analysis program.

Results: The majority of the HCV-diagnosed respondents indicated a health care provider told them to stop drinking “completely.” A smaller number said they were told an occasional drink was “fine” or “cut back.” When asked if providers explained why it was important to change this behavior, most respondents reported because of the “effect on the liver” and on disease progression. Other respondents mentioned the need to stop drinking for treatment efficacy, overall “well-being,” or a “chance to live longer.” Only a few respondents indicated they were not told to curtail their drinking or received an explanation about the effects of alcohol on their HCV. Health care provider advice about alcohol and HCV was generally concordant with what patients heard, with some strongly tell-
ing the patient “no alcohol” and others saying an “occasional drink is acceptable.” However, reports varied by provider type on how often they discussed alcohol with HCV patients. Nurses and gastroenterologists raised the issue with most patients while internists focused on alcohol use only with problematic drinkers. Despite these variations, providers were fairly consistent in telling pat-
tients that alcohol makes HCV progress faster and causes more damage to the liver than it would for someone without HCV. Gastroenterologists consistently described this graphically as “throwing fuel on a fire.”

Conclusion: The messages relayed to HCV patients about alcohol use by health care providers are generally concordant with what patients report hearing. However, the mixed messages different types of providers are sending about appropriate drinking levels is troubling. In order to make the best health-related decisions about their disease, HCV patients need to receive consistent informa-
tion about alcohol consumption from all health care providers they may encounter.

ABSTRACTS

PAMMEXAMINATION OF WHAT HEALTHCARE PROVIDERS SAY AND WHAT PATIENTS HEAR ABOUT ALCOHOL CONSUMPTION AND HEPATITIS C (HCV)


The Cleveland Clinic Foundation, Cleveland, OH; Case Western Reserve University, Cleveland, OH; MetroHealth Medical Center, Cleveland, OH; Case Western Reserve University, C, OH

Purpose: Clear communication is essential for effective shared decision making. Reducing con-
sumption of alcohol is recommended for patients with HCV, yet we know little about what is said and heard regarding this topic in the provider/patient encounter.

Methods: We conducted focus groups and interviews about alcohol use and HCV with both non-
dependent drinkers (as determined by the AUDIT) with HCV (N = 62) and health care providers (N = 14) who care for HCV patients at a metropolitan teaching hospital. All focus groups and inter-
views were audio-taped, transcribed, and analyzed using NVivo, a qualitative data management and analysis program.

Results: The majority of the HCV-diagnosed respondents indicated a health care provider told them to stop drinking “completely.” A smaller number said they were told an occasional drink was “fine” or “cut back.” When asked if providers explained why it was important to change this behavior, most respondents reported because of the “effect on the liver” and on disease progression. Other respondents mentioned the need to stop drinking for treatment efficacy, overall “well-being,” or a “chance to live longer.” Only a few respondents indicated they were not told to curtail their drinking or received an explanation about the effects of alcohol on their HCV. Health care provider advice about alcohol and HCV was generally concordant with what patients heard, with some strongly tell-
ing the patient “no alcohol” and others saying an “occasional drink is acceptable.” However, reports varied by provider type on how often they discussed alcohol with HCV patients. Nurses and gastroenterologists raised the issue with most patients while internists focused on alcohol use only with problematic drinkers. Despite these variations, providers were fairly consistent in telling pat-
tients that alcohol makes HCV progress faster and causes more damage to the liver than it would for someone without HCV. Gastroenterologists consistently described this graphically as “throwing fuel on a fire.”

Conclusion: The messages relayed to HCV patients about alcohol use by health care providers are generally concordant with what patients report hearing. However, the mixed messages different types of providers are sending about appropriate drinking levels is troubling. In order to make the best health-related decisions about their disease, HCV patients need to receive consistent informa-
tion about alcohol consumption from all health care providers they may encounter.
and the technique that a patient favors could confound attempts to help him or her form an authentic clarification of their own values. However, some patients may favor using more hands-on “explicit” values assessing their attitudes toward the pros and cons of different therapeutic options. The purpose is that this relatively non-interactive illustration helps viewers, by social matching, to clarify their own values. However, some patients may favor using more hands-on “explicit” values assessing their attitudes toward the pros and cons of different therapeutic options. To compare patients’ opinions about using implicit and explicit conceptions in the design of a PDA.

Methods: Because this was early developmental work, we asked healthy women to consider the hypothetical situation of a recent diagnosis with early-stage breast cancer. Power Point slides were used to present objective information about the surgical choice between mastectomy (with or without reconstruction) and lumpectomy with radiation. A participant was randomly allocated to either implicit Social Matching (SM) or an explicit Balance Technique (BT); at this point, she rated the technique in terms of its helpfulness, clarity, and appropriateness for other patients. Next, she viewed and rated the alternative technique. Finally, the participant reported which technique she personally favored.

Results: Eleven women participated. Recruitment was stopped when a clear preference picture emerged, as follows.

a) Implicit vs. Explicit Groups’ “Extremely/Yes” Ratings: Helpful = 1 (9%) vs. 6 (55%); Clear = 6 (55%) vs. 7 (64%); Interesting = 3 (28%) vs. 7 (64%); Appropriate = 7 (64%) vs. 10 (91%).

b) Overall Preferred Technique: 9 (82%) favored the BT; 1 (9%) the SM, and 1 (10%) was indifferent between the two techniques.

Conclusions: At least in the context of early-stage breast cancer, the effort to incorporate explicit conceptions into future PDAs would be worthwhile. Similar comparative studies could be carried out in other contexts. Then, subsequently, RTT designs could be used to test for the effects of matches/mismatches between the VC technique that a patient favors and the VC technique he or she actually encounters.

Purpose: Measures of preferences for prostate cancer treatment have incorporated attributes such as age, family, sexual, and bowel symptoms, and anxiety about mortality. While recent work has identified a broad range of prostate cancer treatment outcome domains, including self-concept and masculinity, it is not clear that men consider these impacts in making their treatment decisions. In this qualitative study, we sought to 1) determine whether African American (AA) and Caucasian American (CA) men describe an expanded range of attributes in narratives about prostate cancer treatment decisions and 2) to examine the extent of overlap in the attributes and processes described by AA and CA men.

Methods: 16 AA and 45 CA men diagnosed with localized prostate cancer were recruited from VA urology clinics and consented to participate in focused group sessions. Each of the 13 groups was homogeneous in ethnicity (i.e., AA, CA) and type of primary treatment (i.e., radical prostatectomy, brachytherapy, external beam radiation therapy, and watchful waiting). Two experienced moderators were trained in the specific study methods and conducted the groups. Moderators and groups were matched by gender and ethnicity. Content coding was conducted to identify attributes relevant to prostate cancer treatment decisions; NVivo software was used to manage the analysis.

Results: In addition to symptoms such as sexual and urinary function, both AA and CA men described a wide range of attributes considered in making prostate cancer treatment decisions. The men mentioned concerns about how treatment would influence marital, family, and social relationships apart from sexuality, impacts to self-concept and masculinity, and considerations related to family, work, and community responsibilities. Also, AA men described prostate cancer as an African American burden, and mentioned stigma and trust in health professionals as additional concerns in making prostate cancer treatment decisions.

Conclusions: These data suggest that an expanded range of attributes is likely to be important in prostate cancer treatment decisions for both AA and CA men. While there was considerable overlap in the mentions, AA men described unique influences on the decision-making process. This work suggests that an expanded conceptual framework may be important for understanding and measuring patient values and goals relevant to prostate cancer treatment decisions.

Purpose: To use the results of economic evaluations to rank subgroups in order of priority for receiving vaccination in years during which a shortage of influenza vaccine is anticipated. Methods: A systematic review of studies that evaluated the economic impact of influenza vacci- nation was conducted. Studies were identified by searching Medline, Embase, and Econlit databases and also included some works-in-progress (n = 39). Each study was assigned to 2 reviewers and un- derwent structured data abstraction. After the process, the structured review included study charac- teristics such as appropriateness of model design, model inputs (including epidemiological and clinical inputs, cost inputs, and quality adjustments), method of analysis, and sensitivity analysis. Reviewers presented summaries of each study to the entire group of reviewers who then decided on inclusion or exclusion based study characteristics. Subgroups for prioritization were stratified by age and the presence or absence of medical conditions associated with a high risk for influenza complications.

Results: Studies were excluded if they were identified as review papers (n = 8) or as having signif- icant limitations (n = 17). The primary reasons for exclusions were inappropriate assumptions regarding key inputs, primarily influenza attack rates, vaccine effectiveness, and vaccination costs, paired with insufficient sensitivity analysis. The final ranking was based on 14 studies. Within age groups, vaccinating high-risk persons was more cost-effective than vaccinating non-high-risk indivi- duals. Vaccinating among age groups was difficult due to the lack of studies that included persons of all ages, lacking of transparency in methods, and the use of different end points and methods. The groups that consistently demonstrated the highest economic returns to vaccination were high-risk elderly and non-meningitis prone infected with influenza.

Conclusions: Vaccinating the elderly was consistently cost-saving across a variety of studies using different methods, while younger age and non-high-risk groups were likely to require an invest- ment for health benefits. Ranking by economic data yielded results similar to rankings by hospitalization and death rates. Lack of transparency in methods limited the ability to use results from a number of economic studies.

Purpose: To compare patients’ opinions about using implicit and explicit VC techniques in the design of a PDA.

Methods: Because this was early developmental work, we asked healthy women to consider the hypothetical situation of a recent diagnosis with early-stage breast cancer. Power Point slides were used to present objective information about the surgical choice between mastectomy (with or without reconstruction) and lumpectomy with radiation. A participant was randomly allocated to either implicit Social Matching (SM) or an explicit Balance Technique (BT); at this point, she rated the technique in terms of its helpfulness, clarity, and appropriateness for other patients. Next, she viewed and rated the alternative technique. Finally, the participant reported which technique she personally favored.

Results: Eleven women participated. Recruitment was stopped when a clear preference picture emerged, as follows.

a) Implicit vs. Explicit Groups’ “Extremely/Yes” Ratings: Helpful = 1 (9%) vs. 6 (55%); Clear = 6 (55%) vs. 7 (64%); Interesting = 3 (28%) vs. 7 (64%); Appropriate = 7 (64%) vs. 10 (91%).

b) Overall Preferred Technique: 9 (82%) favored the BT; 1 (9%) the SM, and 1 (10%) was indifferent between the two techniques.

Conclusions: At least in the context of early-stage breast cancer, the effort to incorporate explicit conceptions into future PDAs would be worthwhile. Similar comparative studies could be carried out in other contexts. Then, subsequently, RTT designs could be used to test for the effects of matches/mismatches between the VC technique that a patient favors and the VC technique he or she actually encounters.

Purpose: Measures of preferences for prostate cancer treatment have incorporated attributes such as age, family, sexual, and bowel symptoms, and anxiety about mortality. While recent work has identified a broad range of prostate cancer treatment outcome domains, including self-concept and masculinity, it is not clear that men consider these impacts in making their treatment decisions. In this qualitative study, we sought to 1) determine whether African American (AA) and Caucasian American (CA) men describe an expanded range of attributes in narratives about prostate cancer treatment decisions and 2) to examine the extent of overlap in the attributes and processes described by AA and CA men.

Methods: 16 AA and 45 CA men diagnosed with localized prostate cancer were recruited from VA urology clinics and consented to participate in focused group sessions. Each of the 13 groups was homogeneous in ethnicity (i.e., AA, CA) and type of primary treatment (i.e., radical prostatectomy, brachytherapy, external beam radiation therapy, and watchful waiting). Two experienced moderators were trained in the specific study methods and conducted the groups. Moderators and groups were matched by gender and ethnicity. Content coding was conducted to identify attributes relevant to prostate cancer treatment decisions; NVivo software was used to manage the analysis.

Results: In addition to symptoms such as sexual and urinary function, both AA and CA men described a wide range of attributes considered in making prostate cancer treatment decisions. The men mentioned concerns about how treatment would influence marital, family, and social relationships apart from sexuality, impacts to self-concept and masculinity, and considerations related to family, work, and community responsibilities. Also, AA men described prostate cancer as an African American burden, and mentioned stigma and trust in health professionals as additional concerns in making prostate cancer treatment decisions.

Conclusions: These data suggest that an expanded range of attributes is likely to be important in prostate cancer treatment decisions for both AA and CA men. While there was considerable overlap in the mentions, AA men described unique influences on the decision-making process. This work suggests that an expanded conceptual framework may be important for understanding and measuring patient values and goals relevant to prostate cancer treatment decisions.
programs may be necessary components of any successful HIV screening intervention. While better estimates are needed, these findings suggest that effective risk-reduction counseling expanded HIV screening hinges on the assumed transmission of infection with available therapy.

Monday, October 24, 2005

54 Pacific Room EFG

A NOVEL APPROACH TO CAPTURE BOTH THE INDIVIDUAL- AND POPULATION-LEVEL EFFECTS OF HIV SCREENING

Paltiel A1, Waleronsky R1, Schackman B2, Snape III G2, Weinstein M1 and Freedberg K1

Yale School of Medicine, New Haven, CT; Massachusetts General Hospital, Boston, MA; Well Medical College of Cornell University, New York, NY; Harvard School of Public Health, Boston, MA

Purpose: To propose a unified framework for assessing both individual- and population-level outcomes of HIV screening.

Methods: We used previously developed simulations of HIV screening (Paltiel, NEJM, 2005) and HIV treatment (Freedberg, NEJM, 2002) to project individual-level QALYs and costs for one-time and repeated rapid HIV screening. We linked those projections to published data on HIV transmission risks, with and without antiretroviral therapy (ART). We incorporated model-based estimates of discounted lifetime costs (284,600) and QALY losses (16.8) attributable to new HIV infections. To describe the overall effect of ART on HIV transmission, we estimated intervention-specific values of R0, the basic reproductive number (baseline value = 1.47 to 1.50 secondary transmissions per infected individual), a widely recognized summary measure of the persistence of an epidemic.

Results: Across a range of screening frequencies in a high-risk population, the cost-effectiveness of HIV screening depends critically on the assumed effects of ART on secondary HIV transmission (see Table 1), a parameter for which current evidence is mixed. In lower-risk populations, any repeated screening strategies are dominated unless the ART impact on transmission is favorable.

Conclusions: When population-level effects are taken into account, the attractiveness of expanded HIV screening hinges on the assumed transmission of infection with available therapy. While better estimates are needed, these findings suggest that effective risk-reduction counseling programs may be necessary components of any successful HIV screening intervention.


POSTER SESSION III

1812

7:00 AM - 8:30 AM

Monday, October 24, 2005

53 Pacific Room EFG

MARKOV MODELING OF MAJOR DEPRESSION DYNAMICS

Lee S and Patton S

University of Calgary, Calgary, AB, Canada

Purpose: Most epidemiological studies of major depression report period prevalence estimates, which are of limited utility in characterizing incidence and understanding the dynamics of this condition. The objective of this study was to integrate data from three recent Canadian studies using a Markov model into a more comprehensive description of major depression epidemiology.

Methods: Data from national surveys conducted by the Canadian national statistical agency (Statistics Canada) were used. These data were analyzed using a Markov model designed to elucidate the dynamics of depression (previous version: Medical Decision Making 2004 Jul-Aug 24(4):351-8). The model adopts the format of an incidence-prevalence model, modeling the “prevalence pool” of major depression as a function of inflow to the pool (incidence) and outflow through recovery. Changes in health state (depressed and not-depressed) were evaluated over a series of one-week stages. To address declining probabilities of recovery with mounting episode length, a Markov funnel was used to represent the process of recovery. The basic model was expanded to include these recurrence categories. Monte Carlo simulations were used to calibrate model parameters to the observed data. Tracker variables were defined so that the model output could be tracked using definitions comparable to those used in the database. Tracker variables were also programmed to represent the annual number of episodes over the entire observation period.

Results: The population fell into three categories. A low recurrence group (88%) had a weekly episode incidence (transmission probability) of 0.00628. Five percent of the population occupied a moderate recurrence category in which the weekly episode incidence was 0.00198. Seven percent of the population was in a high recurrence category with a weekly episode incidence of 0.00575. The transition probabilities reflecting recovery were high in the initial weeks of the episodes, and declined by an approximately fixed proportion with each passing week.

Conclusions: Markov models can provide a mechanism for integration and interpretation of psychiatric epidemiological data. Ultimately, such models may be useful for surveillance and economic analyses of this condition. The methods employed here may also be useful for capturing the dynamics and natural history of other medical conditions.

E78 • MEDICAL DECISION MAKING/JAN–FEB 2006

2362

POSTER SESSION III

7:00 AM - 8:30 AM

Monday, October 24, 2005

54 Pacific Room EFG

CERVICAL CANCER SCREENING: POTENTIAL COST-EFFECTIVENESS OF TECHNOLOGY TO IMPROVE COLPOSCOPY AND BIOPSY PERFORMANCE

Sweet S1, Kim J2, Wright T1 and Goldie S1

Harvard School of Public Health, Cambridge, MA; Harvard University, Cambridge, MA; College of Physicians and Surgeons of Columbia University, NY, NY

Purpose: A substantial number of women with equivocal cervical cytology results are likely to harbor high-grade lesions. Recent clinical trial data have shown that only 64%-67% of lesions will be identified with colposcopy, a procedure used to visualize the cervix for abnormalities and guide the location of a biopsy. We assess the potential cost-effectiveness (CE) of a new optical detection technology (ODS) designed to improve the performance of colposcopy.

Methods: We modified a cervical cancer model to reflect recent data on false positive and false negative colposcopy errors at the level of 1) initial visualization of the cervix, 2) accuracy of biopsy location, and 3) misclassification of lesions during pathologic interpretation. We incorporated trial data on ODS technology showing a 26% improvement in detection of high-grade lesions among subjects with equivocal results. We used previously developed simulations of HIV screening (Paltiel, NEJM, 2005) and HIV treatment (Freedberg, NEJM, 2002) to project individual-level QALYs and costs for one-time and repeated rapid HIV screening. We linked those projections to published data on HIV transmission risks, with and without antiretroviral therapy (ART). We incorporated model-based estimates of discounted lifetime costs (284,600) and QALY losses (16.8) attributable to new HIV infections. To describe the overall effect of ART on HIV transmission, we estimated intervention-specific values of R0, the basic reproductive number (baseline value = 1.47 to 1.50 secondary transmissions per infected individual), a widely recognized summary measure of the persistence of an epidemic.

Results: We used previously developed simulations of HIV screening (Paltiel, NEJM, 2005) and HIV treatment (Freedberg, NEJM, 2002) to project individual-level QALYs and costs for one-time and repeated rapid HIV screening. We linked those projections to published data on HIV transmission risks, with and without antiretroviral therapy (ART). We incorporated model-based estimates of discounted lifetime costs (284,600) and QALY losses (16.8) attributable to new HIV infections. To describe the overall effect of ART on HIV transmission, we estimated intervention-specific values of R0, the basic reproductive number (baseline value = 1.47 to 1.50 secondary transmissions per infected individual), a widely recognized summary measure of the persistence of an epidemic.

Conclusions: The attractive of expanded HIV screening hinges on the assumed transmission of infection with available therapy. While better estimates are needed, these findings suggest that effective risk-reduction counseling programs may be necessary components of any successful HIV screening intervention.


POSTER SESSION III

1748

7:00 AM - 8:30 AM

Monday, October 24, 2005

55 Pacific Room EFG

THE EFFECTIVENESS ANALYSIS OF A BEHAVIORAL HIV RISK REDUCTION INTERVENTION: A STUDY FROM AN EXCEL-BASED MONTE CARLO SIMULATION MODEL

Soosapanth S1, Sansom S2 and Chick S1

Centers for Disease Control and Prevention, Atlanta, GA; INSEAD, Fontainebleau, France

Purpose: To assess the effectiveness of a behavioral HIV intervention in reducing the risk of transmission of HIV-infected persons to their sexual partners, using a Monte Carlo simulation implemented in a commercially available spreadsheet.

Methods: We simulated sexual partnership formation as a Poisson process and varied partnership duration according to an exponential distribution. During each partnership, each sexual contact and the probability of HIV transmission to an uninfected partner was modeled as a Bernoulli trial. Sexual partnerships, number of protected and unprotected sexual contacts, and the risk of HIV transmission were simulated repeatedly over time. Parameters such as the number of sexual contacts per partner and the per-contrast probability of transmission can vary with HIV disease progression and be randomly distributed to represent individual variability. We applied data from a Prevention Care Management (PCM) program to changes in transmission risk behaviors (proportion of unprotected sexual contacts and number of sex partners) and estimated the reduction of HIV transmission to partners. In the sensitivity analysis, we explored the impact of partnership duration and of behavioral changes on the transmission reduction estimates.

Results: A hypothetical cohort of 15,000 HIV-infected individuals was simulated over 5 years. Without PCM, each HIV-infected individual would infect, on average, 0.243, 0.212, and 0.2 uninfected partners, assuming the average partnership duration of 6, 12, and 24 months accordingly. The table below shows the number of transmissions to partners (%) reduction) when all PCM clients reduced their sexual risk behavior (30.8% and 20.2% reduction in unprotected sexual contacts and numbers of sex partners) for 3, 6, and 12 months.

Duration That Behavioral Changes Sustained

Average Partnership Duration 3 Months 6 Months 12 Months

6 months 0.234 [3.35%] 0.231 [5.17%] 0.218 [11.15%]
12 months 0.212 [5.61%] 0.194 [9.96%]
24 months 0.199 [9.68%] 0.191 [9.34%] 0.175 [11.08%]

Conclusions: A behavioral intervention such as PCM could potentially reduce the risk of HIV transmission, particularly if the behavioral changes are sustained over time. Information regarding partnership duration is important to better estimate transmission risk and intervention effectiveness. Spreadsheet models have the benefit of being accessible to many users while still incorporating complex parameters such as random variations in individual characteristics, including partnership formation and duration.

© 2006 Society for Medical Decision Making. All rights reserved. Not for commercial use or unauthorized distribution.
E79

ABSTRACTS

2344

POSTER SESSION III

7:00 AM - 8:30 AM
Monday, October 24, 2005
57 Pacific; Room EFG

EVALUATING THE PHARMACOGENOMICS BIAS IN DECISION-ANALYTIC MODELING


Massachusetts General Hospital, Harvard Medical School, Boston, MA; Harvard School of Public Health, Boston, MA

Purpose: Decision analyses of drug treatments in chronic diseases require modeling the progression of disease and treatment response beyond the time horizon of clinical or epidemiological studies. In such models, progression and drug effect have been applied uniformly to all patients; heterogeneity in progression, including pharmacogenomic effects, has been ignored. We sought to evaluate the direction and relative magnitude of a pharmacogenomics bias resulting from failure to adjust for genetic heterogeneity in both treatment response (HT) and heterogeneity in progression of disease (HP).

Methods: We constructed two simple Markov models with three basic health states (early-stage disease, late-stage disease, dead), one adjusting and the other not adjusting for genetic heterogeneity. Adjustment was done by creating different disease states for patients (G1) and absence (G2) of a dichotomous genetic factor. We compared the life expectancy gains attributable to treatment resulting from both models and defined pharmacogenomics bias as percent deviation of treatment-related life expectancy gains in the unadjusted model (1/AM) from those in the adjusted model (AM). We calculated the bias as a function of underlying model parameters to create generic results. We then applied our model to lipiodol-containing therapy vs prasugrel in patients with coronary artery atherosclerosis, incorporating the influence of the two TagI polymorphism variants (B1 and B2) on progression and drug efficacy as reported in the DNA subility of the REGRESS trial.

Results: For an initial volume of 1000 ORT/year, significant reduction in the WL is seen between years 3 and 4 after their introduction (p < 0.05). For initial volumes > 2500 ORT/year, significant reduction in the WL is seen in 3 years after their introduction. In case of joint HP and HT, the overall bias is likely triggered by the HP component and a purely HT-related bias is negative (conservative) and a purely HP-related bias is positive (liberal). For many typical scenarios, the absolute bias is smaller than 10%. In case of joint HP and HT, the overall bias is likely triggered by the HP component and reaches positive values >100% if functions of “fast progressors” and “strong treatment respond- ers” are low. In the pharmacogenomics example, the UAM underestimated the true life-years gained (LTYG) by 5.5%. (1.07 LTYG vs. 0.99 LTYG for 56-year-old men).

Conclusions: We have been able to predict the pharmacogenomics bias jointly caused by heterogeneity in progression of disease and heterogeneity in treatment response as a function of characteristics of patients, disease, and treatment. In case of joint presence of both types of heterogeneity, models ignoring this heterogeneity may generate results that overestimate the treatment benefit.

2150

POSTER SESSION III

7:00 AM - 8:30 AM
Monday, October 24, 2005
59 Pacific; Room EFG

VARIATION IN RISK FACTORS INCLUDED IN CANCER RISK CALCULATIONS ON THE WEB

Garmashkin A, Sonnad S, Kurichi J and Armstrong K

Dana-Farber Cancer Institute, Boston, MA; University of Pennsylvania, Philadelphia, PA

Background: Many members of the general public use the Internet as a source of health information, including using Web site calculators to assess their risk of developing serious illnesses. The risk information provided by Web risk calculators (WRC) may impact users’ worry about their risk, their health behavior, and medical decisions. The widespread use of these sites and the potential impact of the information they provide substantiates the importance of sites providing accurate risk estimates.

Purpose: We reviewed WRCs for 3 common cancers—breast, colon, and prostate—to examine what risk factors were used by each to calculate how those risk factors were quantified, and whether references for the calculations were provided.

Methods: We conducted a search of “cancer risk” in 5 popular search engines and visited the first 1,000 hits from each engine to create a list of breast, colon, and prostate WRCs. For each WRC, we reviewed source information and the risk factors incorporated in the calculation of an individual’s risk for the cancer.

Results: Web sites with risk calculators numbered 114 for breast cancer, 3 for colon cancer, and 4 for prostate cancer. No sources for the calculator appeared on the Web sites for 73% of breast cancer calculators, 67% of colon calculators, and 75% of prostate calculators. The calculators varied widely in the factors collected for risk assessment. For instance, 82% of the breast cancer calculators included age, 55% included race, 45% included alcohol consumption and menstrual status, 36% included gender, height, weight, hormone use, 27% included personal cancer history and fruit/vegeta-

tile consumption, 18% included Jewish ancestry, and 9% include tobacco use history and current use and amount of exercise.

Conclusions: Few WRCs provide source information to substantiate the risk estimates that they provide. The wide variation in the risk factors that WRCs collect to calculate cancer risk raises additional concerns about the accuracy of the risk estimates that they provide. Inaccurate risk estimates can have significant consequences for recipients, including provoking unwarranted distress and encourag-}

E79

ABSTRACTS

2086

POSTER SESSION III

7:00 AM - 8:30 AM
Monday, October 24, 2005
58 Pacific; Room EFG

HOW GOOD DO ORGAN REPLACEMENT TECHNOLOGIES NEED TO BE: VOLUME AND DURABILITY

Stahl J and Vacanti J

Massachusetts General Hospital, Boston, MA

Purpose: Determine the needed volume and durability of organ replacement technologies.

Methods: We constructed a decision analysis of the future of organ replacement with Biocor® bioreactive tissue used as a model. We considered alternative scenarios that accounted for differences in durability and volume of usage by modeling the number of organs transplanted as a function of the number of times this may occur. The WL length was then compared to the projected WL as-sociated with the number of operations. Sensitivity analysis was performed on the durability from 1 month to 5 years and projected the bias as a function of underlying model parameters to create generic results. We then applied our model to lipiodol-containing therapy vs prasugrel in patients with coronary artery atherosclerosis, incorporating the influence of the two TaqIB polymorphism variants (B1 and B2) on progression and drug efficacy as reported in the DNA subility of the REGRESS trial.

Results: Volume for an initial volume of 1000 ORT/year, significant reduction in the WL is seen between years 3 and 4 after their introduction (p < 0.05). For initial volumes > 2500 ORT/year, significant reduction in the WL is seen in 3 years after their introduction. In case of joint HP and HT, the overall bias is likely triggered by the HP component and a purely HT-related bias is negative (conservative) and a purely HP-related bias is positive (liberal). For many typical scenarios, the absolute bias is smaller than 10%. In case of joint HP and HT, the overall bias is likely triggered by the HP component and reaches positive values >100% if functions of “fast progressors” and “strong treatment respond- ers” are low. In the pharmacogenomics example, the UAM underestimated the true life-years gained (LTYG) by 5.5%. (1.07 LTYG vs. 0.99 LTYG for 56-year-old men).

Conclusions: We have been able to predict the pharmacogenomics bias jointly caused by heterogeneity in progression of disease and heterogeneity in treatment response as a function of characteristics of patients, disease, and treatment. In case of joint presence of both types of heterogeneity, models ignoring this heterogeneity may generate results that overestimate the treatment benefit.

2003

POSTER SESSION III

7:00 AM - 8:30 AM
Monday, October 24, 2005
60 Pacific; Room EFG

DIAGNOSTIC PERFORMANCE OF DUPLEX ULTRASOUND IN PATIENTS SUSPECTED OF CAROTID ARTERY DISEASE: THE IPSILATERAL VERSUS CONTRALATERAL ARTERY

Heijnenrek-Kal MJ, Nederkoorn PJ, Buskens E, Van der Graaf Y and Hunink M

Erasmus MC-University Medical Center Rotterdam, Rotterdam, Netherlands; Academic Medical Center Amsterdam, Amsterdam, Netherlands; University Medical Center Utrecht, Utrecht, Netherlands; Erasmus Medical Center, Rotterdam, Netherlands

Purpose: To evaluate duplex ultrasonographic thresholds for the determination of 70%-90% stenosis of the ipsilateral and contralateral internal carotid artery in patients with symptoms of amuscular aorta, transient ischemic attack (TIA), or minor stroke based on two criteria—maximizing accuracy and optimizing cost-effectiveness—and to compare these with current recommendations.

Methods: From January 1997 to January 2000, a prospective multicenter study was conducted including 358 consecutive patients with symptoms of amaurosis fugax, TIA, or minor stroke, who underwent bilateral duplex ultrasonography and digital subtraction angiography. A linear regression analysis was performed to estimate the degree of angiographic stenosis as a function of the peak systolic velocity (PSV). PSV thresholds were calculated for both the ipsilateral and contralateral carotid arteries based on maximizing accuracy and optimizing cost-effectiveness.

Results: The PSV measurements significantly overestimated the angiographic stenosis in the contralateral artery (9.5%; 95% CI 6.3%-12.7%) compared with the ipsilateral carotid artery. The recommended PSV threshold for the diagnosis of 70%-90% stenosis of the ipsilateral carotid artery in patients with symptoms of amaurosis fugax, transient ischemic attack (TIA), or minor stroke based on two criteria—maximizing accuracy and optimizing cost-effectiveness—and to compare these with current recommendations.

Conclusions: Few WRCs provide source information to substantiate the risk estimates that they provide. The wide variation in the risk factors that WRCs collect to calculate cancer risk raises additional concerns about the accuracy of the risk estimates that they provide. Inaccurate risk estimates can have significant consequences for recipients, including provoking unwarranted distress and encourag-