

Verification of a Decision Analytic Model Assumption Using Real-World Practice Data: Implications for the Cost Effectiveness of Cyclo-oxygenase 2 Inhibitors (COX-2s)

Emily R. Cox, PhD; Brenda Motheral, PhD; and Doug Mager, BS

Objective: To verify the gastroprotective agent (GPA) rate assumption used in cost-effectiveness models for cyclo-oxygenase 2 inhibitors (COX-2s) and to re-estimate model outcomes using GPA rates from actual practice.

Methods: Prescription and medical claims data obtained from January 1, 1999, through May 31, 2001, from a large preferred provider organization in the Midwest, were used to estimate GPA rates within 3 groups of patients aged at least 18 years who were new to nonselective nonsteroidal anti-inflammatory drugs (NSAIDs) and COX-2 therapy: all new NSAID users, new NSAID users with a diagnosis of rheumatoid arthritis (RA) or osteoarthritis (OA), and a matched cohort of new NSAID users.

Results: Of the more than 319 000 members with at least 1 day of eligibility, 1900 met the study inclusion criteria for new NSAID users, 289 had a diagnosis of OA or RA, and 1232 were included in the matched cohort. Gastroprotective agent estimates for nonselective NSAID and COX-2 users were consistent across all 3 samples (all new NSAID users, new NSAID users with a diagnosis of OA or RA, and the matched cohort), with COX-2 GPA rates of 22%, 21%, and 20%, and nonselective NSAID GPA rates of 15%, 15%, and 18%, respectively. Re-estimation of the cost-effectiveness model increased the cost per year of life saved for COX-2s from \$18 614 to more than \$100 000.

Conclusions: Contrary to COX-2 cost-effectiveness model assumptions, the rate of GPA use is positive and marginally higher among COX-2 users than among nonselective NSAID users. These findings call into question the use of expert opinion in estimating practice pattern model inputs prior to a product's use in clinical practice. A re-evaluation of COX-2 cost-effectiveness models is warranted.

(*Am J Manag Care* 2003;9:785-794)

Decision makers have said that for pharmacoeconomic information to be useful it should, among other things, be presented in a timely manner, include head-to-head comparisons with relevant comparators, and consist of reliable data.¹ The dilemma for researchers is that meeting all of these criteria is practically impossible. If the pharmacoeconomic information is to be presented in a timely manner, typically when a product comes to market, often the only available data come from the randomized controlled trials used to establish safety and efficacy of the

product compared with placebo. At the time of product launch, studies comparing the new product with existing similar products are rarely available. Another difficulty faced by researchers is that data important to the economic evaluation of the therapy may not be available at the time of product launch, such as practice patterns reflecting real-world use, costs, resource utilization, and utilities. To provide decision makers with information in a timely manner, including use of relevant comparators, researchers utilize decision analysis techniques, relying on the best data available for model inputs. When such data are not available, expert opinion and consensus are often used to obtain estimates.² Rarely, however, are these assumptions verified once actual practice data become available.

A case in point is cost-effectiveness evaluations of cyclo-oxygenase 2 inhibitor (COX-2) therapy. Both nonselective nonsteroidal anti-inflammatory drugs (NSAIDs) and COX-2s are equally effective in treating pain. However, nonselective NSAIDs have a significantly higher rate of gastrointestinal (GI) adverse events. To reduce the risk of GI adverse events with NSAIDs, several guidelines have recommended that physicians coprescribe gastroprotective agents (GPAs), such as proton pump inhibitors (PPIs) or misoprostol to those patients at high risk of adverse GI events.³⁻⁶ Given the cost associated with GPA use and the extent of GPA coprescribing among nonselective NSAID users, estimated to be from 17% to 34%,⁷⁻⁹ failure to include these costs in economic evaluations of these products could underestimate the true cost of NSAID therapy.

From the Office of Research and Planning, Express Scripts, Inc., Maryland Heights, Mo.

This research was presented as a podium presentation at the Express Scripts Sixth Annual Outcomes Conference, St. Louis, Mo, June 3-5, 2002.

Address correspondence to: Emily R. Cox, PhD, Express Scripts, Inc, 13900 Riverport Drive, Maryland Heights, MO 63043. E-mail: ecox@express-scripts.com.

Table 1. Overview of COX-2 Cost-effectiveness Studies

COX-2 Cost-effectiveness Study	Study Perspective	Study Population	Treatment Alternatives	Nonselective NSAID GPA Rate Assumption	COX-2 GPA Rate Assumption
Rofecoxib					
Marshall et al ¹⁵	Ontario Ministry of Health	OA patients ≥65 years, not responding to acetaminophen	Nonselective NSAIDs	23%	90% less than nonselective NSAID rate
Pellissier et al ¹⁴	US third-party payer	OA	Nonselective NSAIDs	25.5%	75% less than nonselective NSAID rate*
Moore et al ¹⁶	British National Health Service	OA	Nonselective NSAIDs	20.6%	75% less than nonselective NSAID rate*
Celecoxib					
Haglund and Svarvar ¹⁰	Swedish healthcare system	OA and RA	Diclofenac/misoprostol; NSAID + PPI; NSAID + H ₂ RA; NSAID + misoprostol; rofecoxib	NA	0%
Svarvar and Aly ¹²	Norwegian healthcare system	OA and RA	Diclofenac/misoprostol; NSAID + PPI; NSAID + H ₂ RA; NSAID + misoprostol; rofecoxib	45%	0%
Zabinski et al ¹¹	Canadian Provincial Ministry of Health	OA and RA patients ≥65 years	Diclofenac/misoprostol; NSAID + PPI; NSAID + H ₂ RA; NSAID + misoprostol	NA	0%
Chancellor et al ¹³	Swiss healthcare system	OA and RA	Diclofenac/misoprostol; NSAID + PPI; NSAID + H ₂ RA; NSAID + misoprostol	NA	0%

H₂RA indicates H₂ receptor antagonist; NA, not applicable; NSAID, nonsteroidal anti-inflammatory drug; OA, osteoarthritis; PPI, proton pump inhibitor; RA, rheumatoid arthritis.

*Estimated using expert panel.

Cost-effectiveness models comparing COX-2s with nonselective NSAIDs have been published for celecoxib¹⁰⁻¹³ and rofecoxib.¹⁴⁻¹⁶ Because no differences have been noted between the efficacy of nonselective NSAIDs and that of COX-2s, these studies evaluated whether the added cost of COX-2 therapy is worth the benefit gained through reductions in GI adverse events among arthritis patients. These models have all concluded that COX-2s provide cost savings or that the benefits gained, in terms of years of life saved, is worth the added cost. Each study made use of decision analysis, but differed in study methods, inclusion of model inputs, perspectives, patient populations (ie, patients diagnosed with osteoarthritis [OA] or rheumatoid arthritis [RA]), treat-

ment alternatives, and decision pathways (Table 1). Each model assumed that a proportion of nonselective NSAID users would be prescribed a GPA. Because no empirical data were available at the time the studies were conducted to estimate the GPA rate among those on COX-2 therapy, and given clinical trial evidence supporting a more favorable GI side effect profile with COX-2s, all models assumed that the need for GPAs would be reduced or completely eliminated for COX-2 users.

In 3 cost-effectiveness studies of rofecoxib (Vioxx), based on usage patterns in the United Kingdom,¹⁶ Canada,¹⁵ and the United States,¹⁴ the authors use published estimates of GPA rates for nonselective NSAIDs of

20%, 23%, and 25.5%, respectively. Estimates of the GPA rate for patients treated with rofecoxib were derived from an expert panel in 2 of the studies.^{14,16} The expert panel, comprising gastroenterologists, rheumatologists, and general practitioners, estimated the GPA comedication rate for rofecoxib to be 75% less than the base rate for nonselective NSAIDs. In the Canadian study,¹⁵ the GPA rate was assumed to be 90% less than the nonselective NSAID rate, or 2.3%. No explanation was given as to how researchers arrived at this estimate. All rofecoxib models tested model sensitivity to variation in the base case GPA rate for nonselective NSAIDs and the percent reduction in GPA use for rofecoxib. In 2 of the 3 rofecoxib models,^{14,15} results were sensitive to the GPA comedication rate.

In the arthritis cost consequence evaluation system (ACCES) model, developed to compare celecoxib (Celebrex) with nonselective NSAIDs, the costs and GI consequences of celecoxib use were compared against 5 nonselective NSAID therapy alternatives: nonselective NSAID monotherapy, nonselective NSAID plus a PPI, nonselective NSAID plus an H₂ receptor antagonist (H₂), nonselective NSAID plus misoprostol, and Arthrotec, a combination product containing diclofenac sodium and misoprostol.¹⁷ The ACCES model assumed no GPA comedication with celecoxib. The model allows for variation in the overall GPA rate for nonselective NSAIDs depending on the sample population being modeled. For example, in an application of the model using Norwegian data, the distribution of nonselective NSAID patients included 20% taking a nonselective NSAID plus a PPI, 15% taking a nonselective NSAID plus H₂, and 10% taking Arthrotec.¹²

In another cost-effectiveness model evaluating celecoxib, the assumption was again a zero GPA rate among celecoxib users.¹⁰ No discussion was provided in any of the celecoxib models as to why a 0% GPA rate assumption was made. In addition, model sensitivity to variation in nonselective NSAID GPA rates was not evaluated.

Given evidence of model sensitivity to the GPA assumption, the use of expert opinion to estimate COX-2 GPA rates (when estimated), lack of documentation to support a 0% COX-2 GPA rate assumption and evidence contrary to these assumptions,¹⁸ and the high cost of GPA therapy relative to NSAIDs, the purpose of this study was to verify this assumption using actual practice data. We also replicated the COX-2 cost-effectiveness model developed by Pellissier and colleagues¹⁴ to re-estimate model outcomes based on observed GPA rates. We selected this model because it is the only COX-2 cost-effectiveness model to represent a US perspective, which was the perspective of our analysis.

METHODS

Pharmacy and medical claims data for a 320 000-member preferred provider organization located in the Midwest were used to identify new NSAID users. The pharmacy benefit design for plan members varied; however, more than 90% of sample members had a 3-tier benefit structure with first-tier co-pays ranging from \$8 to \$10, second-tier co-pays from \$15 to \$20, and third-tier co-pays from \$25 to \$30. Cyclo-oxygenase 2 inhibitors were on the third tier, and no other restrictions regarding their use were in place during the time of evaluation (ie, step therapy or prior approval). New NSAID users were defined as individuals with an NSAID claim within an index window from January 1, 2000, through May 31, 2000, and with no NSAID claim in the previous 12 months. Other inclusion criteria were that the individuals obtained a supply for 30 days or more for an NSAID during the 1-year follow-up; were at least 18 years of age; and were continuously eligible for the entire study period. Patients taking disease-modifying agents (ie, methotrexate, gold compounds) during the 1-year follow-up were excluded due to indications of greater disease severity.

From the first NSAID claim in the index window (ie, index prescription), patients' prescription claim activity was followed for 365 days. Use of NSAIDs was identified from the generic product identifier (GPI; First Databank [Medispan], Indianapolis, Ind) number beginning "661000" (nonselective NSAID) or "661005" (COX-2). Gastroprotective agent usage was identified using the GPI code for H₂ receptor antagonists (GPI beginning "4920"), PPIs (GPI beginning "4927"), misoprostol (GPI beginning "4925"), sucralfate (GPI beginning "4930"), and ulcer therapy combinations (eg, ulcer anti-infective with bismuth combinations or PPIs, or H₂ antagonists-antacid combinations, GPI beginning "4999"). Gastroprotective agent usage was defined as use of any of the above-listed agents on 1 or more days during the 365-day follow-up.

Three GPA rates were estimated—one for new NSAID users regardless of diagnoses, the second among new NSAID users with a diagnosis of RA or OA, and the third among a cohort of new NSAID users matched for GI risk and length of NSAID therapy. The matched cohort was included to control for differences in the GI risk across COX-2 and nonselective NSAID users, due to recent findings indicating channeling of patients at risk for GI events to COX-2s.¹⁹ Matching for GI risk was based on the Standardized Calculator Of Risk for Events (SCORE) criteria.

Although not a fully validated instrument, SCORE is a self-assessment instrument to quantify the risk of

CLINICAL

serious NSAID-related GI toxicity (eg, stomach ulcer or bleeding ulcer) and is based on 6 criteria: age, self-reported health status, diagnosis of RA, use of oral corticosteroids, hospitalization for a stomach or intestinal problem or prior history of stomach ulcer, and history of gastrointestinal side effects (heartburn, stomach pain, nausea, vomiting) when taking NSAIDs. SCORE was developed by researchers at Stanford University and based on data from the Arthritis, Rheumatism, and Aging Medical Information System (ARAMIS) database.²⁰ ARAMIS is a database that has followed more than 36 000 patients with rheumatic diseases, collecting data on health status, clinical outcomes, drug side effects, and resource utilization. We selected the SCORE criteria because it was the only algorithm available to assess NSAID-induced GI risk that combined several documented risk factors into a single score for each individual, thereby adding a level of specificity beyond a simple “yes/no” indication of risk.

Using medical and pharmacy claims data, a SCORE value was calculated for each patient using 5 of the 6 criteria, the exception being self-reported health status (Table 2). Whereas SCORE was designed to be admin-

istered by patients or physicians, the estimation of SCORE criteria using medical and pharmacy claims data has been documented.²¹ For the SCORE criteria of GI side effects when taking NSAIDs, a proxy was used. We chose this method because the study criteria excluded individuals with prior use of NSAID therapy and because prior use of NSAID therapy could have occurred earlier than 365 days before the index prescription. This proxy measure included a diagnosis of any of the following conditions: gastritis, esophagitis, ulcer-related symptoms (ie, dyspepsia), gastroesophageal reflux disease (GERD) damage, and GERD-related symptoms. Using the scoring algorithm for the 5 SCORE criteria, a summed value was generated for each member of the sample. Values could range from 0 to a maximum of 35.

Nonselective NSAID and COX-2 users were also matched for length of NSAID use, with length of therapy calculated as the summed value of the days supply field from the index prescription 365 days forward. Length of NSAID therapy was included as a match criterion due to findings of significantly greater length of therapy for individuals whose index prescription was a

Table 2. SCORE Criteria Used to Match New Nonselective NSAID and COX-2 Users

SCORE Criteria	Code	Definition
Age		Calculated as of the index prescription using member date of birth from eligibility records
Rheumatoid arthritis	Inpatient or outpatient primary, secondary, or tertiary ICD-9 codes of 7140X, 7141X, 7142X, or 7149X; GPI beginning “6629,” “6628,” “6625,” “13000020,” or “52500060”	Diagnosis or 1 or more prescription claims 365 days before or 120 days after index prescription
Oral corticosteroid use	GPI code beginning “22”	Number of 30-day equivalent prescriptions in the 12 months prior to the index prescription
Hospitalization for GI bleed or ulcer	Inpatient revenue codes 110 through 169, and 190 through 219, together with inpatient or outpatient primary, secondary, or tertiary ICD-9 codes of 531XX, 532XX, 533XX, or 534XX, or a prescription claim for ulcer therapy combination (ie, GPI 4-digit code = 4999)	Hospitalization 12 months prior to index date, with diagnoses code or GPI code within 5 days prior to or after the inpatient hospital stay
GI side effects when taking NSAIDs	Inpatient or outpatient primary, secondary, or tertiary ICD-9 codes of 5302X, 5304X, 5301X, 7871X, 7872X, 5308X, 2515X, 5303X, 5305X, 5309X, 7865X, 535XX, 7870X, 7879X, 7890X, 5369X, 5379X, 04186, 5368X	Diagnosis in the 12 months prior to the index prescription

COX-2 indicates cyclo-oxygenase 2 inhibitor; GI, gastrointestinal; GPI, generic product identifier; ICD-9, International Classification of Diseases; NSAID, nonsteroidal anti-inflammatory drug; SCORE, Standardized Calculator of Risk for Events.

Table 3. Demographics, Pattern of NSAID Use, and GPA Rate Across Study Samples

Parameter	New NSAID Users (n = 1900)	New NSAID Users With Diagnosis of OA or RA (n = 289)	New NSAID Users: Matched Cohort (n = 1232)
Average age, years (SD)	47.0 (11.7)	51.7 (10.22)	48.4 (10.3)
Female (%)	56.5	58.5	58.1
Average age by index drug, years (SD)			
COX-2	50.9 (11.27)	55.0 (9.45)	48.6 (10.52)
Nonselective NSAID	44.2 (11.12)	48.0 (9.73)	48.3 (10.15)
Percent female by index drug (%)			
COX-2	61	66	62
Nonselective NSAID	53	51	54
Index drug: COX-2, n (%)	781 (41.1)	148 (51.2)	616 (50.0)
COX-2 only	667 (85.4)	123 (83.1)	523 (84.9)
COX-2 switched to nonselective NSAID	114 (14.6)	25 (16.9)	93 (15.1)
Index drug: nonselective NSAID, n (%)	1120 (58.9)	141 (48.8)	616 (50.0)
Nonselective NSAID only	895 (79.9)	95 (67.4)	470 (76.3)
Nonselective NSAID switched to COX-2	224 (20.0)	46 (32.6)	146 (23.7)
Rate of GPA Use (%) Overall	18	18	19
Index drug: COX-2	22*	21	20
COX-2 only	23	23	20
COX-2 switched to nonselective NSAID	16	12	17
Index drug: nonselective NSAID	15	15	18
Nonselective NSAID only	13	12	16
Nonselective NSAID switched to COX-2	22	22	24

*Chi-square test for difference between COX-2 and nonselective NSAID GPA rate $P \leq .01$.

COX-2 indicates cyclo-oxygenase 2 inhibitor; GPA, gastroprotective agent; NSAID, nonsteroidal anti-inflammatory drug; OA, osteoarthritis; RA, rheumatoid arthritis.

COX-2 and the likelihood that this would inflate the GPA rate among COX-2 users. For matching purposes, days supply was categorized into 20th percentiles.

Results are presented by the pattern of switching between selective and nonselective NSAIDs during the 365 days of follow-up (ie, COX-2 users only, nonselective NSAID users only, switch from nonselective NSAID to COX-2, and switch from COX-2 to nonselective NSAID).

We replicated the cost-effectiveness model of Pellissier and colleagues¹⁴ using Microsoft Excel 97. The COX-2 and nonselective NSAID GPA rates from this study were then used to re-estimate the cost per day of therapy, the cost offset per day, cost per GI event avoided, and cost per year of life saved. The only input not provided by Pellissier was the market share of the individual GPA agents. Requests for this information were made but not obtained. Therefore, using our own data

as a reference point, we modified market share values slightly until we were able to exactly replicate Pellissier’s cost offset per day. The market share used to replicate Pellissier’s model was 74% omeprazole, 5% misoprostol, and 21% H₂s.

Finally, several sensitivity analyses were conducted to evaluate the robustness of our results to variations in other model assumptions in addition to GPA rate assumptions. The first involved using actual days supply of GPA and NSAID therapy rather than full compliance, as assumed by Pellissier. The second varied which GPA agents were included in the analysis.

.....
RESULTS

A total of 319 263 members had at least 1 day of enrollment during the index window with 17 300 having

an NSAID in the index window. A total of 9611 members were excluded for noncontinuous eligibility during the 29-month study period and 612 were excluded because they were younger than 18 years. An additional 3685 members were excluded because of prior NSAID use, and 25 were excluded due to use of disease-modifying agents. Finally, 1467 members were excluded for having less than a 30-days' supply of NSAID during the 1-year follow-up. The final sample for the analysis of new NSAID users regardless of diagnosis was 1900. From this sample, 1611 had no evidence of OA or RA in the 365 days prior to or 120 days after the index prescription, leaving a sample of 289 for the evaluation of new NSAID users with a diagnosis of OA or RA. Other diagnoses associated with the use of NSAIDs indicated many and varied musculoskeletal conditions (eg, low back pain, 20.6%; joint pain, 9.7%; knee or other joint derangement, 7.4%; sprains and strains of joints and adjacent muscles, 3.2%; dysmenorrhea, <1%; etc) and other pain-related conditions (eg, hospitalization, organ transplant, headache, etc). Approximately 26% had no documented condition that could be attributed to NSAID use.

Finally, using the sample of 1900 new NSAID users, 668 were excluded during the matching process for a final sample size of 1232 for the matched cohort analysis (616 nonselective NSAID and 616 COX-2 users).

Table 3 presents the demographics and NSAID use pattern across the 3 samples. The average age of sample subjects was 47 years (SD = 12), with 57% female. This demographic profile was not significantly different across samples with respect to percent female ($\chi^2 = 0.963$, $P = .618$). However, Bonferroni paired comparison tests indicated significant differences in age across samples at the $\alpha = 0.01$ level. The average SCORE value for the matched cohort was 8.5 (SD = 2.98).

The pattern of use and switching among those who begin on nonselective NSAIDs and COX-2s varied across the 3 samples (Table 3). For the overall sample of new NSAID users ($n = 1900$), approximately 60% had as their index prescription a nonselective NSAID. Among those with a diagnosis of OA or RA ($n = 289$), the proportion with a COX-2 as their index prescription (51%) was significantly higher compared to the overall sample ($\chi^2 = 10.49$, $P = .001$).

Across all 3 samples, from 18% to 19% of patients had evidence of GPA use during the 1-year follow-up. Across all 3 samples, the GPA rates for those who began therapy on a nonselective NSAID were lower compared with individuals who began with COX-2 therapy. However, these rates were significantly different only in the overall sample of NSAID users (15% vs 22%, for nonselective NSAID vs COX-2 users, respectively; $\chi^2 = 18.75$, $P < .001$).

To evaluate whether GPA use began during the follow-up period or represented continuation of therapy, the date GPA therapy began relative to the index prescription was captured. Among persons with GPA use during the follow-up period, 70%, 77%, and 68% of those whose first prescription was a COX-2 had evidence of GPA therapy prior to beginning their COX-2, within the sample of total new NSAID users, OA/RA patients, and the matched cohort. These same percentages for those whose first prescription was a nonselective NSAID are 53%, 43%, and 54%, respectively. These rates of prior GPA use among GPA users in the follow-up period were all significantly different at the $\alpha = 0.05$ level ($\chi^2 = 10.73$, $P = .001$; $\chi^2 = 6.45$, $P = .011$; $\chi^2 = 4.63$, $P = .031$, respectively).

Table 4 profiles the outcomes replicated using Pellissier's base case assumptions¹⁴ and re-estimated using GPA rates from our study sample. The model was exactly replicated for the outcomes of expected cost per day for nonselective NSAIDs and cost offset per day and estimated to within \$0.01 for expected cost per day for rofecoxib. The model was replicated to within 4% for cost per perforation, ulceration, and bleeding (PUB) avoided and cost per year of life saved. We believe that an error in the calculation or reporting of cost per death avoided may have occurred in Pellissier's table contributing to the 11% difference in replicated estimates.

The Pellissier model was re-estimated using GPA rates from the matched cohort, which had the smallest difference in GPA rates across nonselective NSAID and COX-2 users of 18% and 20%, respectively. Using these GPA rates, the base case expected costs per day for rofecoxib and nonselective NSAIDs were estimated to be \$3.29 and \$2.49, respectively. Separating these costs into NSAID drug cost and GI-related costs (ie, GPA and other GI medical costs), the GI-related expected cost per day for rofecoxib increased from \$0.44 to \$0.87 and decreased for nonselective NSAIDs from \$1.26 to \$1.02. The difference in the GI-related costs between nonselective NSAIDs and rofecoxib represent the expected daily cost offset for rofecoxib. Pellissier's base case cost offset per day was estimated as \$0.81, substantially higher than the \$0.15 from the re-estimated model using GPA rates from actual practice. These findings are presented in the **Figure** together with the Pellissier data, which tested model sensitivity to changes in the base case nonselective NSAID GPA rate and varied the GPA comedication reduction with rofecoxib.

Using total cost estimates (ie, NSAID drug costs and GI-related medical costs) for a cohort of 10 000 patients over 1 year, Pellissier estimated the cost per PUB avoided at \$4738 and the cost per death avoided at \$274 749.

Table 4. Replication of Pellissier Base Case-Cost-Effectiveness Model Outcomes and Re-estimated Model Outcomes Using Actual Practice Data GPA Rates

Parameter	Pellissier et al ¹⁴ Base Case Estimates	Replicated Base Case Estimates	Re-estimated Model*
Expected cost per day: rofecoxib (\$)			
Rofecoxib drug costs	2.42	2.42	2.42
Costs associated with GI prophylaxis and treatment	0.44 [†]	0.45	0.87
Total, rofecoxib	2.86	2.87	3.29
Expected cost per day: nonselective NSAIDs (\$)			
NSAID drug costs	1.47	1.47	1.47
Costs associated with GI prophylaxis and treatment	1.26 [†]	1.26	1.02
Total, nonselective NSAIDs	2.73	2.73	2.49
Costs offset per day[‡] (\$)	0.81	0.81	0.15
Events avoided with rofecoxib (n)			
PUBs	100	100	100
Deaths	2	2	2
Cost per event avoided (\$)			
PUBs	4738	4912	29 188
Deaths	274 749	245 615	1 459 414
Cost per year of life saved (\$)	18 614	17 872	106 192

*Using GPA rates of 18% for nonselective NSAIDs and 20% for rofecoxib.

[†]This value was not reported by Pellissier, but rather implied.

[‡]Calculated as the difference in costs associated with GI prophylaxis and treatment between rofecoxib and nonselective NSAIDs. GI indicates gastrointestinal; NSAID, nonsteroidal anti-inflammatory drug; PUB, perforations, ulcers, and bleeds.

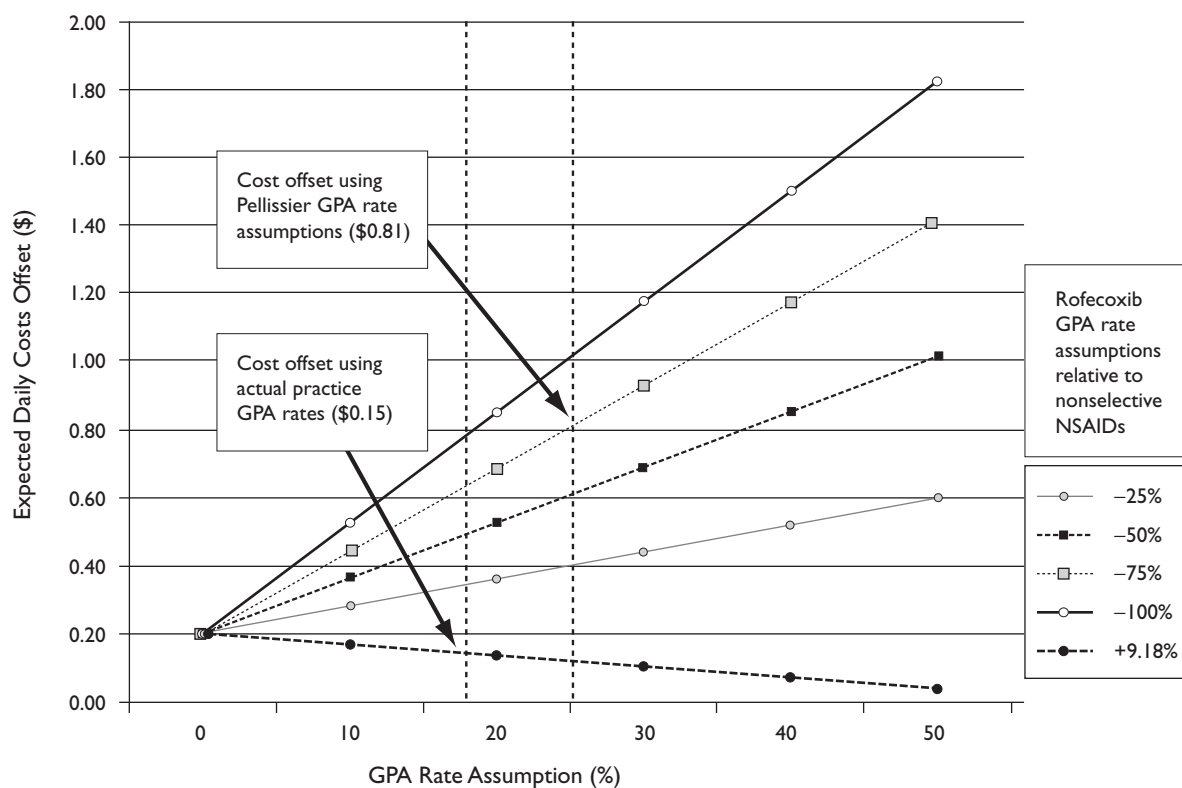
Assuming a difference of 2 deaths due to GI adverse events between rofecoxib and nonselective NSAIDs, and a loss of 18.6 years for each death discounted at 3%, Pellissier estimated a cost per year of life saved for rofecoxib of \$18 614. Re-estimating the model using GPA rates from actual practice increases the cost per PUB avoided to \$29 636 and the cost per death avoided increases to approximately \$1.5 million. Using the same assumptions regarding life years lost, the marginal cost per year of life saved for rofecoxib exceeds \$100 000.

Another assumption made by Pellissier that could not be confirmed from actual practice data was full compliance during the 1-year period. Using the matched cohort, the number of days on GPA therapy by index prescription was 35 and 28 days for COX-2 and nonselective NSAID users, respectively. The number of days on NSAID therapy by index prescription was 108 and 99 days for COX-2 and nonselective NSAID, respectively. Changing the number of days on GPA therapy in the model changed our original results

only slightly, from a cost offset per day for COX-2s of \$0.15 to \$0.19; the cost per PUB avoided decreased to \$27 850; the cost per death avoided decreased to \$1 392 508; and cost per year of life saved decreased by less than \$5000 to \$101 323. The results were more favorable for COX-2s even though the days on GPA for COX-2s are greater. This result is because GPA costs represent a larger proportion of total GI-related costs for COX-2s than for nonselective NSAIDs. When both days on nonselective NSAID and GPA therapy are modified using actual practice data, the cost per year of life saved for COX-2s increased to more than \$200 000 per year of life gained.

We also varied the inclusion of specific GPA agents. Our initial choice of GPA agents was based on the agents used by Pellissier. Re-estimating GPA rates excluding H₂s decreased the GPA rate within the matched cohort to 14% and 11%, respectively, for COX-2 and nonselective NSAID users. Using these GPA rates, the cost offset per day for COX-2s decreased slightly

Figure. Gastrointestinal Cost Offset With Rofecoxib Versus Nonselective NSAIDs by GPA Prophylactic Rate and Various Rate Reduction or Increase (Negative Value) Assumptions for Rofecoxib



GPA indicates gastroprotective agent; NSAID, nonsteroidal anti-inflammatory drug.

from our base estimates of \$0.15 to \$0.13; the cost per PUB avoided increased to \$29 783; the cost per death avoided increased to \$1 489 135; and the cost per year of life saved increased to \$108 354.

DISCUSSION

This study failed to verify the GPA comedication rate assumption made in COX-2 cost-effectiveness studies. Contrary to estimates provided by expert opinion and assumptions made by study authors, the GPA rate among COX-2 users is positive and marginally higher than the rate among nonselective NSAID users. Using the GPA estimates from our sample to re-estimate Pellissier's model, the cost per year of life saved goes beyond that considered economically acceptable using the threshold of \$50 000/year.

Recent findings would appear to provide support for this conclusion.²² In a study comparing the effect of a drug policy restricting and not restricting use of COX-2 therapy among long-term NSAID users, researchers found an incremental cost of \$31 900 and \$56 700 for

the nonrestrictive versus restrictive policy in preventing symptomatic and complicated ulcers, respectively. This incremental cost is greater than the incremental cost to prevent PUB in our replicated model (\$29 188). At this rate, the cost per life year gained would likely fall outside of the accepted threshold, as it did in our estimates.

Several study limitations may limit generalizability of the study findings. The data represent only one plan's experience, and GPA prescribing patterns could vary depending on numerous factors including physician practice style and plan design. In estimating the GPA rate, we fully recognize the limitations in using medical claims data to identify specific diagnoses, such as that used in the matching process.²³ Prior history of NSAID-related GI risk or diagnoses may have been missed because of undercoding. This can be due to lack of documentation or the fact that these diagnoses were secondary to other conditions for which the patient was being seen and not captured in the primary, secondary, or tertiary ICD-9 fields. Additionally, the time during which we evaluated individual history of GI events was limited to 365 days. A GI event could have occurred

prior to this look-back period. While we recognize these factors as limitations in the matching of patients using the SCORE method, results were consistent regardless of the sample used. In addition, because calculation of the GPA rate relied primarily on prescription claims data, which studies have shown to be a reliable and valid source of data,^{24,25} we believe that these limitations are minimized with respect to GPA estimates.

Gastroprotective agent estimates in this study did not include use of over-the-counter (OTC) GPA agents, including OTC H₂ receptor antagonists or antacids, as other estimates have.⁹ Failure to capture OTC GPA usage may underestimate the true GPA rate among NSAID users. However, OTC use typically would not be used in economic evaluations conducted from a third-party payer perspective.

These findings call into question the use of expert opinion in estimating model inputs for pharmacoeconomic evaluation, particularly when estimating practice patterns prior to the product's use in a real world setting. When compared with clinical trial data documenting GI adverse events associated with use of COX-2s and nonselective NSAIDs, these assumptions fail even the most basic test of validity, face validity. Two large trials have compared the risk of gastrointestinal complications between COX-2s and nonselective NSAIDs, the Vioxx Gastrointestinal Outcomes Research (VIGOR) trial and the Celecoxib Long-Term Arthritis Safety Study (CLASS) trial.^{26,27} In the VIGOR trial, the incidence of complicated confirmed upper GI events was 1.22% in the naproxen group and 0.52% in the rofecoxib group ($P < .005$). The CLASS trial found no statistically significant difference between celecoxib and nonselective NSAID users in the incidence of the primary endpoint of ulcer perforation, gastric-outlet obstruction, or upper GI bleeding (0.8% vs 1.5%, respectively; $P = .09$). Both rofecoxib and celecoxib have been found to significantly reduce mild-to-moderate GI adverse events (rofecoxib compared with naproxen = 23.5% vs 25.5%, $P = 0.02^{28}$; and celecoxib compared with nonselective NSAIDs = 31.4% vs 36.8%, $P \leq .05$),²⁶ a reduction of 8% to 15%, respectively. Taken together, these differences would not support the expert panel's estimate of a 75% reduction in GPA use among COX-2 users, nor the assumption that COX-2 users would not be prescribed a GPA.

We do not contend that our results are the definitive answer to the cost effectiveness of COX-2s. We do believe, however, that the cost effectiveness of COX-2s should be revisited. Sufficient data are now available to evaluate practice patterns that would affect total costs (ie, GPA use and switching). Newly developed studies should take into consideration the effect of GI risk on

the cost effectiveness of COX-2s and compare relevant treatment alternatives (eg, nonselective NSAIDs plus PPI vs COX-2s vs COX-2s plus PPI) among high-risk patients. Additionally, not all models evaluate other patterns reflecting actual practice; these include days supply of NSAID therapy and switching, either due to lack of efficacy or side effects.

.....
CONCLUSION

The error of unsubstantiated assumptions had been documented as a criticism of pharmacoeconomic evaluations.²⁹ Although Pellissier and colleagues concluded that rofecoxib is cost effective compared with nonselective NSAIDs, they also acknowledged the need for further evaluation, stating that "it is too early to tell if the economic benefits of rofecoxib versus nonselective NSAIDs will be realized in clinical practice. However, until clinical GI outcomes data becomes available from a mature marketplace, this research provides insight into the potential cost-effectiveness of rofecoxib in the management of OA."¹¹ This study demonstrates that GI clinical outcomes are not the only data that need to be verified using data from a mature marketplace in assessing the economic benefit of rofecoxib.

However, some would argue that the critics of decision analysis modeling in healthcare³⁰⁻³² are naive as to the purpose of modeling and that the cost of waiting until perfect evidence is available would paralyze the practice of medicine.³³ Weinstein and colleagues contended that we must balance the costs and consequences of obtaining and waiting for better data against the costs and consequences of permitting a synthesis of the available evidence.³³ Perhaps some believe that errors in economic modeling may not be viewed with the same degree of concern as errors in establishing clinical efficacy or safety.

The harm from incorrect assumptions, such as those found in the cost-effectiveness evaluations of COX-2s, is that they undermine the ultimate goal of pharmacoeconomics, that is, the efficient allocation of scarce healthcare resources. Whether more harm than good resulted from an invalid COX-2 cost-effectiveness model will require further examination. In the mean time, greater efforts at verifying unsubstantiated model assumptions is a critical step to undertake in order to increase the usefulness of and confidence in pharmacoeconomic information.

Acknowledgments

The authors thank Kathleen Fairman, MA, for helpful comments on the study's analysis, and Bryce Sutton, PhD, for assistance in cost-effectiveness model replication.

REFERENCES

1. Mullins CD, Wang J. Pharmacy benefit management: enhancing the applicability of pharmacoeconomics for optimal decision making. *PharmacoEconomics*. 2002;20:9-21.
2. Evans C. The use of consensus methods and expert panels in pharmacoeconomic studies: practical applications and methodological shortcomings. *PharmacoEconomics*. 1997;12:121-129.
3. Hochberg MC, Altman RD, Brandt KD, et al. Guidelines for the medical management of osteoarthritis. Part I. Osteoarthritis of the hip. American College of Rheumatology. *Arthritis Rheum*. 1995;38:1535-1540.
4. Hochberg MC, Altman RD, Brandt KD, et al. Guidelines for the medical management of osteoarthritis. Part II. Osteoarthritis of the knee. American College of Rheumatology. *Arthritis Rheum*. 1995;38:1541-1546.
5. Lanza FL. A guideline for the treatment and prevention of NSAID-induced ulcers. *Am J Gastroenterol*. 1998;93:2037-2046.
6. McCarthy DM. Prevention and treatment of gastrointestinal symptoms and complications due to NSAIDs. *Best Pract Res Clin Gastroenterol*. 2001;15:755-773.
7. Lapane KL, Spooner JJ, Pettitt D. The effect of nonsteroidal anti-inflammatory drugs on the use of gastroprotective medication in people with arthritis. *Am J Manag Care*. 2001;7:402-408.
8. Simon LS, Zhao SZ, Arguelles LM, et al. Economic and gastrointestinal safety comparisons of etodolac, nabumetone, and oxaprozin from insurance claims data from patients with arthritis. *Clin Ther*. 1998;20:1218-1235.
9. Singh G. Recent considerations in non-steroidal anti-inflammatory drug gastropathy. *Am J Med*. 1998;105:31S-38S.
10. Haglund U, Svarvar P. The Swedish ACCES model: predicting the health economic impact of celecoxib in patients with osteoarthritis or rheumatoid arthritis. *Rheumatology*. 2000;39(suppl 2):51-56.
11. Zabinski RA, Burke TA, Johnson J, et al. An economic model for determining the costs and consequences of using various treatment alternatives for the management of arthritis in Canada. *PharmacoEconomics*. 2001;19(suppl 1):49-58.
12. Svarvar P, Aly A. Use of the ACCES model to predict the health economic impact of celecoxib in patients with osteoarthritis or rheumatoid arthritis in Norway. *Rheumatology*. 2000;39(suppl 2):43-50.
13. Chancellor JVM, Hunsche E, de Cruz E, Sarasin FP. Economic evaluation of celecoxib, a new cyclo-oxygenase 2 specific inhibitor, in Switzerland. *PharmacoEconomics*. 2001;19(suppl 1):59-75.
14. Pellissier JM, Straus WL, Watson DJ, Kong SX, Harper SE. Economic evaluation of rofecoxib versus nonselective nonsteroidal anti-inflammatory drugs for the treatment of osteoarthritis. *Clin Ther*. 2001;23:1061-1079.
15. Marshall JK, Pellissier JM, Attard CL, Kong SX, Marentette MA. Incremental cost-effectiveness analysis comparing rofecoxib with nonselective NSAIDs in osteoarthritis: Ontario Ministry of Health perspective. *PharmacoEconomics*. 2001;19:1039-1049.
16. Moore RA, Phillips CJ, Pellissier JM, Kong SX. Health economic comparisons of rofecoxib versus conventional nonsteroidal anti-inflammatory drugs for osteoarthritis in the United Kingdom. *J Med Econ*. 2001;4:1-17.
17. Pettitt D, Goldstein JL, McGuire A, Schwartz JS, Burke T, Maniadas N. Overview of the arthritis cost consequence evaluation system (ACCES): a pharmacoeconomic model for celecoxib. *Rheumatology*. 2000;39(suppl 2):33-42.
18. Mamdani M, Rochon PA, Juurlink DN, et al. Observational study of upper gastrointestinal haemorrhage in elderly patients given selective cyclo-oxygenase-2 inhibitors or conventional non-steroidal anti-inflammatory drugs. *BMJ*. 2002;325:624-629.
19. Wolfe F, Flowers N, Burke TA, Arguelles LM, Pettitt DJ. Increase in lifetime adverse drug reactions, service utilization, and disease severity among patients who will start COX-2 agents: quantitative assessment of channeling bias and confounding by indication in 6689 RA and OA patients. *J Rheumatol*. 2002;29:1015-1022.
20. Singh G, Terry R, Ramey DR, Triadafilopoulos G, Brown BW. GI SCORE: a simple self-assessment instrument to quantify the risk of serious NSAID-related GI complications. *Arthritis Rheum*. 1997;40(suppl):S93.
21. Bull SA, Conell C, Campen DH. Relationship of clinical factors to the use of COX-2 selective NSAIDs within an arthritis population in a large HMO. *J Manag Care Pharm*. 2000;8:252-258.
22. Fendrick AM, Bandekar RR, Chernew ME, Scheiman JM. Role of initial NSAID choice and patient risk factors in the prevention of NSAID gastropathy: a decision analysis. *Arthritis Care Res*. 2002;47:36-43.
23. Motheral BR, Fairman KA. The use of claims databases for outcomes research: rationale, challenges, and strategies. *Clin Ther*. 1997;19:346-366.
24. Tamblin R, Lavoei G, Petrella L, Monette J. The use of prescription claims databases in pharmacoepidemiological research: the accuracy and comprehensiveness of the prescription claims database in Québec. *J Clin Epidemiol*. 1995;48:999-1009.
25. Kirking DM, Ammann MA, Harrington CA. Comparison of medical records and prescription claims files in documenting prescription medication therapy. *J Pharmacoepi*. 1996;5(1):3.
26. Silverstein FE, Faich G, Goldstein JL, et al. Gastrointestinal toxicity with celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis: the CLASS study: a randomized controlled trial. *JAMA*. 2000;284:1247-1255.
27. Bombadier C, Laine L, Reicin A, et al. Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis. *N Engl J Med*. 2000;343:1520-1528.
28. Langman MJ, Jensen DM, Watson DJ, et al. Adverse upper gastrointestinal effects of rofecoxib compared with NSAIDs. *JAMA*. 1999;282:1929-1933.
29. Hill SR, Mitchell AS, Henry DA. Problems with the interpretation and pharmacoeconomic analyses: a review of submissions to the Australian Pharmaceutical Benefits Scheme. *JAMA*. 2000;283:2116-2121.
30. Rennie D, Luft HS. Pharmacoeconomic analyses: making them transparent, making them credible. *JAMA*. 2000;283:2158-2160.
31. McCabe C, Dixon S. Testing the validity of cost-effectiveness models. *PharmacoEconomics*. 2000;17:501-513.
32. Kassirer JP, Angell M. The journal's policy on cost-effectiveness analyses [editorial]. *N Engl J Med*. 1994;331:669-670.
33. Weinstein MC, Toy EL, Sandberg EA, et al. Modeling for health care and other policy decisions: uses, roles, and validity. *Value Health*. 2001;4:348-361.