Interpreting Hazards: The Increasing Importance of “Antidote to Anecdote” in Managed Care

Kathleen A. Fairman, MA

The team of experts has estimated that there is a 1 in 50 chance that a crumb entering either the ear/eye or windpipe with force will cause significant irritation.¹

A September 2009 report commissioned by a British biscuit manufacturer identified the dangers of a seemingly harmless activity in which citizens of the United Kingdom (UK) engage on a daily basis—English tea time. The report was based on a “national survey” of more than 1,000 UK residents selected and queried about their tea time habits using an unreported methodology, coupled with laboratory testing of the physical properties of various biscuits. Costs from a societal perspective were based on standard pricing for items and services that are necessary, according to the project’s “team of experts,” to treat biscuit-related injuries—for example, anesthetic spray to treat burns sustained while dunking the biscuit in hot tea; heat packs to ease back strain caused by bending to pick up a dropped biscuit fragment; and visits to a National Health Service physician to treat an errant crumb lodged in an eye, ear, or windpipe. The report even included a statistical model, the Biscuit Incident Threat Evaluation (BITE), which predicted the risk associated with eating particular biscuits, accompanied by a disclaimer noting that the model “has been calculated based upon only the most obvious types of injuries and based upon common biscuit-eating behavior types. It is strongly recommended that the individual consult the biscuit manufacturer and perform a self assessment of their particular risk prior to eating biscuits.”¹

It appears unlikely that the BITE model will ever become a Medline-indexed publication. Nonetheless, its dissemination bears the markings of challenges that sometimes threaten the efforts of managed care organizations (MCOs) to identify and promote cost-effective therapies. The BITE findings were widely reported in the popular press, with headlines that included “Crumbs: Half of Britons Injured by Their Biscuits” and “Brits in Killer Biscuit Warning.”^{2,3} Consumers visiting the sponsoring company’s website were told that “there are hundreds of biscuit related injuries treated by UK doctors every year” and were encouraged to use the BITE tool lest they “risk it for a biscuit.”⁴ And, in a pattern that is familiar to most MCO decision makers, the most dangerous biscuit identified by the “study” was a particular brand of custard cream cookie, a competitor to a purportedly safer treat whose manufacturer funded and published the work.¹

MCO decision makers are routinely flooded with requests to modify coverage and benefit design policies in response to purportedly reliable evidence about clinical and economic risks.

Typically, the MCO decision maker is told that the increased acquisition cost for a given treatment or benefit design change is outweighed by clinical or economic benefits attributable to reduction in the risk of high-cost medical events, indicated by odds ratios (ORs) or hazard ratios (HRs). The challenge for a decision maker is to assess a vast amount of information in the usually limited amount of time available. The MCO’s intent, as expressed by Jackson and Barbuto in a June 2008 *JMCP* commentary on the lack of evidence for treating migraine with botulinum toxin, is to apply “antidote to anecdote.”⁵ That is, the ideal decision-making process relies on quantitative evidence reported in the peer-reviewed research literature in lieu of press reports and the ever-increasing volume of subjective and sometimes unreliable information provided on blogs. The key question for a decision maker is this: How accurate is the peer-reviewed evidence that is presented to support the proposed policy? This editorial examines some of the most common problems in claims of risk reduction and provides a checklist for hazard avoidance by MCO decision makers (Table 1).⁶

Much Ado About Nothing: The Hazard of Clinically Insignificant Effect Sizes

A December 2009 retrospective cohort analysis conducted by Simpson et al.⁷ used administrative claims data from 23 large employer groups to examine the 2-year outcomes of employed patients aged 18–64 years who were initiated for either primary or secondary prevention on simvastatin, which today is available as a generic drug for approximately \$25 per month,^{8,9} versus brand atorvastatin, which is priced at approximately \$90–\$125 per month^{10,11} when both products are obtained in 90-day supplies for their recommended starting doses.^{12,13} In a sample of 13,584 patient pairs matched on initial drug dose (low, medium, high), baseline cardiovascular events, average wage, and a propensity score for use of atorvastatin that was “generated from a logistic regression model...developed by entering all studied baseline characteristics into a stepwise selection algorithm,” the authors found what they described in the study abstract as “a reduced risk of cardiovascular events” and “reduced indirect costs” associated with atorvastatin treatment.⁷ In this study sponsored by the manufacturer of brand atorvastatin, Simpson et al. concluded that because of cost offsets for medical services and indirect costs, primarily “medically related absenteeism,” a focus on drug acquisition cost by health care payers choosing statin products is misleading. “Although it can be relatively straightforward to

TABLE 1 Guidelines for Hazard Avoidance in Adopting Policies Based on Claims of Risk Reduction

Category	Questions to Ask	Recommendations
Who?	• What were the characteristics of the group(s) studied? Are the study subjects similar to my enrollees?	• Results obtained for one group do not necessarily apply to another (e.g., high- vs. low-income, service industry vs. executives).
	• What were the reference and treatment groups included in the calculation of the hazard ratio or odds ratio? Is the proposed policy change consistent with those groups?	• Claims of risk reduction that do not specifically refer to the study reference group are thereby incorrectly stated and could be misleading.
	• Is the claim of risk reduction based on the sample overall or only on sample subgroup(s)?	• If confidence intervals for groups overlap, the differences between those groups might not be statistically significant.
What?	• Does the report contain an understandable data table of descriptive statistics on the primary study outcome(s) or only a less transparent multivariate analysis?	• View subgroup results cautiously, especially if they are internally inconsistent.
	• What are the effect sizes? Are they clinically meaningful in addition to being statistically significant?	• If multiple comparisons were used without <i>P</i> value adjustment, the results of statistical significance tests may be misleading.
	• What is the absolute amount of risk reduction—from x.x% to y.y%—not just the hazard ratio or odds ratio?	• A report that includes only multivariate analyses violates standards for presentation of statistical information and is not sufficiently useful to decision makers; view such results with caution. ^a
	• What is the number needed to treat? What is the number needed to harm? What is the cost per avoided event?	• Calculate simple measures of benefit relative to cost. An intervention that costs more than it saves could be potentially important for other reasons (e.g., quality-adjusted life years gained), but an anticipation of cost savings is not one of them.
How?	• Is the logical connection between the proposed change and the desired outcome plausible?	• View results cautiously if the claimed risk reduction is based only on association rather than a logically clear causal relationship.
	• Is the proposed intervention feasible?	

^aFairman and Curtiss (2008).⁶

implement programs designed to increase the use of generic statins over more expensive brand-name agents and to measure the impact of such programs on pharmacy costs, our findings suggest that, in the case of atorvastatin and simvastatin, cost savings predicted by this type of silo budget analysis would not materialize in bottom-line costs to employers.”⁷

In an accompanying editorial, Culler and Weintraub pointed to comparative effectiveness research, citing the analysis by Simpson et al., as a key to future health reform efforts that will require informed choices from among available treatment options.¹⁴ For example, in what the editorialists described as a “back-of-the-envelope” calculation, they concluded that with a small (10%) price reduction for atorvastatin, an employer whose average daily wage exceeds \$200 should “place atorvastatin in the tier that does not have co-pays and/or deductibles” to avoid the expense of medically related absenteeism. In contrast, Culler and Weintraub observe that an employer who pays a lower average daily wage might want to charge a higher copayment for atorvastatin.

Understandable enthusiasm for the principles underlying cost-effectiveness analysis notwithstanding, the specific findings of Simpson et al. are noteworthy both for their inconsistency and for their small magnitude. Despite the study’s extremely large sample size, of 18 comparisons on measures of cardiovascular outcomes and medical service use reported in Simpson et al.’s

primary data table, 12 were *not* statistically significant. Among the medical utilization outcomes not significantly associated with atorvastatin use during the 2-year follow-up period were number of cardiovascular events per 100 patients (mean 12.7 atorvastatin vs. 13.5 simvastatin, *P*=0.07); percentage of patients with coronary artery bypass grafting (CABG; 0.8% in both groups, *P*=0.54); percentage of patients who received angioplasty (1.7% atorvastatin vs. 1.9% simvastatin, *P*=0.24); inpatient days (mean 1.2 in both groups, *P*=0.17); emergency room (ER) days (mean 0.7 in both groups, *P*=0.77); outpatient service use days (mean 20.0 atorvastatin vs. 20.4 simvastatin, *P*=0.06); and all-cause medical service costs (mean \$9,131 atorvastatin vs. \$9,403 simvastatin, *P*=0.52).⁷

Moreover, the effect sizes for the 6 statistically significant medical utilization measures were remarkably small. Over a 2-year follow-up, comparing atorvastatin versus simvastatin, respectively, 7.5% versus 8.2% of patients experienced at least 1 inpatient cardiovascular event (*P*=0.02); 6.0% versus 6.6% had an inpatient diagnosis of ischemic heart disease (*P*=0.04); and 0.3% versus 0.5% experienced a myocardial infarction (MI, *P*=0.01). Mean cardiovascular-related nonpharmacy costs (i.e., summing medical service and “other” direct health care costs) were \$3,139 for atorvastatin-treated patients and \$3,390 for simvastatin-treated patients, a total mean difference of \$251 over a 2-year period, or just \$10.46 per month. Also notably similar

for atorvastatin- and simvastatin-treated patients were mean medically related absenteeism days, which the authors “identified from medical service claims” rather than actual days absent from work (12.2 atorvastatin vs. 12.5 simvastatin, difference of 0.0125 days per month, $P=0.02$) and estimated medically related absenteeism costs (\$2,692 atorvastatin vs. \$2,798 simvastatin, difference of \$4.42 per month, $P=0.03$).⁷

Yet, after summing total direct and indirect costs, Simpson et al. concluded that the 2-year cost offsets associated with atorvastatin (mean total costs excluding index drug, \$13,796 atorvastatin vs. \$14,213 simvastatin, \$17.38 per month, $P=0.19$) made the total treatment cost difference in choosing atorvastatin over simvastatin “minimal.” An important factor in this assessment was an estimated index drug cost difference of only \$19.04 per month, which Simpson et al. “imputed” for atorvastatin versus simvastatin “using the average reimbursed amounts in 2007 by dose and days of supply.”⁷ However, current prices, which suggest a drug cost difference of \$65-\$100 per month for a starting dose,⁸⁻¹³ make Simpson et al.’s conclusion questionable. Moreover, in the real-world environment today, generic drug discount programs in community pharmacies are common, and an average actual MCO price for generic simvastatin across all doses is less than \$6 per prescription versus approximately \$105 across all doses of Lipitor.¹⁵

Understanding and Preventing the Hazard of Misinterpretation when Sample Size Is Large

The study by Simpson et al. illustrates an unfortunately common problem with interpretation of miniscule differences as important based solely on P values. Statistical significance is strongly influenced by sample size.¹⁶ P values represent the probability that the observed results would have been obtained based on chance alone—that is, sampling error. The larger the sample size, the less likely it is that sampling error produced the study results, and the more likely it is that the results represent a real difference between study groups. However, “real” does not equate to “important.”¹⁶ As editorialists Diamond and Kaul creatively described the problem in 2004, “eventually, even the smallest difference in outcome cannot escape the pull of a ‘statistical black hole’ fueled by a sufficient mass of patients. Carried to the extreme, everything becomes ‘significant’ in a trial of infinite size.”¹⁷ In a later (February 2010) editorial in the *Journal of the American College of Cardiology*, Kaul and Diamond argued that the problem of “an erroneous tendency to equate statistical significance with clinical significance” is common in the cardiovascular disease literature.¹⁸

The question for a decision maker is how to interpret the clinical or practical significance of results obtained from research studies, especially when sample size is large. To illustrate a potential solution to the decision maker’s dilemma, Kaul and Diamond cited the results of the Treat Angina With Aggrastat and Determine Cost of Therapy With an Invasive or Conservative

Strategy—Thrombolysis In Myocardial Infarction (TACTICS-TIMI) trial, which randomized 2,220 patients with acute coronary syndrome (ACS) to an early invasive strategy of cardiac catheterization within 4 to 48 hours and revascularization if warranted versus a conservative management strategy in which catheterization was performed only upon recurrent ischemia or an abnormal stress test.¹⁹ After 6 months of follow-up, rates of the primary end point (a composite outcome of death, nonfatal MI, or rehospitalization for ACS) were 15.9% for the early invasive strategy versus 19.4% for conservative management (unadjusted OR=0.78, 95% confidence interval [CI]=0.63-0.98, $P=0.028$).

Arguing that presentation of these results alone does not permit decision making by the clinician who wants to know the probability that early invasive management produces clinically important benefit, Kaul and Diamond applied a Bayesian analysis to the TACTICS-TIMI results, producing estimates of the probability of risk reduction at any given threshold of clinical importance specified by the decision maker.^{17,20} For example, TACTICS-TIMI findings for early invasive management suggest that the probability of clinical benefit greater than 0% is 98.6%; however, the probability that the benefit exceeds 25%, the relative difference for which the TACTICS-TIMI study sample was initially powered, is only 17%. Factoring in these probabilities, as well as cost and the potential for bleeding associated with invasive treatment, Kaul and Diamond suggested that a clinical decision maker might forego invasive treatment despite the statistically significant result of TACTICS-TIMI.¹⁸

Kaul and Diamond also argued that basic measures of cost relative to risk, such as number needed to treat (NNT) and number needed to harm (NNH) are underused in the reporting of clinical trial data.¹⁸ NNT and NNH are more easily calculated than Bayesian analyses and provide valuable information. For example, using the cardiovascular event data shown in Simpson et al.’s primary data table (and ignoring for the sake of illustration the nonsignificant result on that outcome measure), the NNT to prevent 1 cardiovascular event over a 2-year time frame with atorvastatin instead of simvastatin is 125 (difference of 0.8 events per 100 patients=0.008 events per patient), at an incremental drug cost per event avoided of \$195,000-\$300,000 using current drug prices (monthly cost difference of \$65-\$100 times 24 months times 125 patients) versus \$57,125 using the drug costs reported by Simpson et al. (\$457 for 2 years times 125 patients).

Ask the Question Enough Times and You’ll Get the Desired Answer: The Hazard of Multiple Comparisons

Another common problem in the research literature is the use of multiple comparison tests, often performed on subgroups of the study sample, without statistical adjustment to the significance threshold to account for the number of comparisons being made.¹⁸ Specifically, with an a priori alpha of 0.05 and a single comparison, the probability of type 1 or “false positive” error—that is, rejecting the null hypothesis when it is actually true—is

1 in 20, or 5%. But as the number of comparisons made in a given study increases, so too does the cumulative probability of type 1 error, defined as making a false positive inference at least once in a series of comparisons. The formula for calculating the cumulative type 1 error probability is $1 - (1 - \alpha)^N$ where α is the a priori alpha value and N is the number of comparisons.¹⁸ For example, 20 subgroup analyses performed at an alpha of 0.05 yield a cumulative type 1 error probability of 64%.

As an example of the hazard of multiple comparisons, the Kaul and Diamond editorial cites the Clopidogrel for High Atherothrombotic Risk and Ischemic Stabilization, Management, and Avoidance (CHARISMA) randomized controlled trial, which compared clopidogrel+aspirin versus placebo+aspirin in the treatment of 15,603 patients with either documented cardiovascular disease (symptomatic, $n=12,153$) or cardiovascular risk factors (asymptomatic, $n=3,284$).²¹ During a median follow-up period of 28 months, clopidogrel and placebo did not significantly differ on the primary end point measure (composite of MI, stroke, or death from cardiovascular causes; 6.8% for clopidogrel vs. 7.3% for placebo, relative risk [RR]=0.93, 95% CI=0.83-1.05, $P=0.22$). However, for the subgroup of patients with active disease, statistical results favored clopidogrel (6.9% clopidogrel vs. 7.9% placebo, RR=0.88, 95% CI=0.77-0.998, $P=0.046$). The CHARISMA trial authors used the subgroup findings to conclude that “there was a suggestion of benefit with clopidogrel treatment in patients with symptomatic atherothrombosis.”²¹

In their editorial on common problems in the cardiovascular research literature, Kaul and Diamond expressed 2 concerns about this interpretation of the CHARISMA results.¹⁸ First, they argued that in a study with a negative outcome on the primary endpoint for the sample overall, subgroup analyses should not be performed because “positive subgroups within negative trials...are virtually always the result of confounding or bias.” Second, they observed that had the CHARISMA trial authors applied a Bonferroni correction for multiple comparisons—that is, dividing the a priori significance level (0.05) by the number of comparisons made (20)—the subgroup analysis would not have met the resulting significance threshold of 0.0025.¹⁸

Kaul and Diamond’s concern seems particularly appropriate in considering the Simpson et al. study of atorvastatin versus simvastatin, in which 20 comparisons of either medical service utilization or medically related absenteeism were made and the total number of statistical comparisons was 43. Of the 6 significant P values on the medical utilization outcome measures, 5 ranged from 0.01 to 0.04; thus, only 1 could possibly have maintained statistical significance after an adjustment for number of comparisons (“other” cardiovascular costs, \$250 atorvastatin vs. \$275 simvastatin, difference of \$1 per month, $P<0.001$). More importantly, in a study with 43 statistical comparisons, the cumulative probability of at least 1 false positive result using an alpha of 0.05 is 89%.

Riskier Than What? The Hazard of Forgetting the Reference Group

In January 2010, news headlines were lit up with the findings of a study of incident rates of type 2 diabetes among a sample of community residents aged 45 to 64 years who, during 9 years of follow-up, either never smoked ($n=4,090$), continued to smoke ($n=2,018$), or quit smoking ($n=380$).²²⁻²⁴ Popular press summaries of the study, which was conducted by Yeh et al., reported that “quitting smoking raises diabetes risk”²³ and that, although “smoking is linked to an increased risk of diabetes...quitting the habit, ironically, may increase diabetes risk in the short term.”²⁴

These interpretations of the smoking cessation study by Yeh et al. reflect a misunderstanding of HRs that is subtle but important and common. HRs and ORs refer to the risk for a particular group relative to a “reference group,” which serves as a point of comparison. In the analysis by Yeh et al., the reference group consisted of adults *who had never smoked*. Compared with adults *who had never smoked*, the HR of incident type 2 diabetes for those who quit smoking was 1.73 (95% CI=1.19-2.53), and the HR for those continuing to smoke was 1.31 (95% CI=1.04-1.65).²² These comparisons, although interesting and important, provide no information about whether those who stop smoking are at greater risk of diabetes than those who continue to smoke. They indicate only that *both* quitters and smokers have higher risks of incident diabetes than those who have never smoked. In addition, the overlapping 95% CIs and relatively small cohort sizes for the quitting and continued smoking groups indicate that the risks of incident type 2 diabetes for smokers versus quitters might not be significantly different.

Evidence of the Need for “Antidote to Anecdote” in this Issue of JMCP

Two articles in this issue of JMCP illustrate the potential costs, both to MCOs and to their enrollees, of using unreliable evidence in clinical and economic decision making. These articles also make clear the potential benefits of healthy skepticism about commonly accepted practices and ideas in health care.^{25,26}

Lack of Evidence for Medical Cost Offsets from Value-Based Insurance Design (VBID). An article by Melnick and Motheral examines the oft-repeated assertion that VBID, a benefit design in which copayments for efficacious chronic disease medications are lowered to encourage adherence, will improve patients’ health to such an extent that medical cost offsets will result. Melnick and Motheral present key elements of a “VBID calculator” that is available online free of charge to facilitate quantitative evaluation of these potential cost offsets.²⁷ For each of 3 therapeutic categories in which VBID programs are commonly proposed (asthma, diabetes, and statins for coronary artery disease [CAD]), users input data specific to their health plan or group: current and proposed copayments by tier level, pharmaceutical utilization,

and, if available, inpatient hospital and ER utilization and cost.²⁵ The calculator produces (a) total anticipated costs for the copayment reductions, including both the loss of copayment revenue for existing drug claims and the increase in drug claims expected from increased adherence; (b) anticipated changes in inpatient hospital and ER use, based on estimated adherence changes coupled with published efficacy data; and (c) the net economic effect of the 2 factors.

Perhaps the most valuable aspect of the calculator is its rapid computation of the cost of drug copayment reductions. This information, coupled with citations to national data on average costs for hospitalization and ER visits in commercially insured groups,²⁸ gives the user a transparent look at the basic financial feasibility of a VBID program. For example, Melnick and Motheral examine the financial implications of the drug copayment reductions that were reported by Chernew et al. in 2008 and that have been cited as evidence of the effectiveness of VBID.²⁹ In a hypothetical health plan of 10,000 members with a 3-tier copayment structure that charges \$5, \$25, and \$45 for generic, preferred brand, and nonpreferred brand drugs, respectively, the calculator estimates that the strategy employed in the Chernew et al. study—elimination of generic copayments and a 50% reduction in brand copayments—would increase the plan’s net drug costs by \$160,842 annually.²⁵ Thus, at a national average cost of approximately \$22,000 per hospitalization for CAD,^{28,30} the copayment reduction program would have to avoid more than 7 hospitalizations to break even. However, the estimated increase in statin adherence that was observed in the Chernew et al. study—a medication possession ratio change of 3.39 percentage points, or about 12 days of therapy per year—is not sufficient to generate any clinical benefit, let alone prevent 7 CAD hospitalizations in a 10,000 member plan in just 1 year’s time.

Plausibility analyses such as that provided by Melnick and Motheral, although representing the best information about VBID currently available to decision makers, are no substitute for real outcomes data. Unfortunately, the outcomes data available for VBID have not yet been reported. Although medical data were available for the unnamed employer in the 2008 Chernew et al. study,²⁹ and despite the lead author’s prediction in January 2008 that “there is considerable evidence that use of the classes of medication in this study will reduce the frequency of adverse clinical events and associated hospitalizations and ER visits,”³¹ neither economic nor clinical outcomes have been reported for that VBID intervention. A simulation analysis based on the Chernew et al. study was published as this issue of *JMCP* went to press.³² In that report, the authors indicated that they presented a simulation model instead of the project’s actual outcomes data because “preliminary statistical analysis of the spending data indicated considerable uncertainty surrounding estimates of the impact of the value-based insurance intervention on aggregate spending.”³² Similarly, the MHealthy study of a VBID program for patients with diabetes, which was implemented in July 2006, was sched-

uled for completion in January 2009 and included both drug and nondrug utilization as outcome measures; yet, no outcomes for MHealthy have been reported as of this writing.³³

While decision makers await outcomes data, Melnick and Motheral remind us that the evidence supporting the use of copayment reductions to achieve medical cost offsets is scant. Specifically, elasticity (i.e., price sensitivity) for prescription medication is generally low; rates of avoidable high-cost events are low for most patient groups; and much nonadherence is attributable to factors other than out-of-pocket cost.²⁵ Melnick and Motheral recommend more cost-effective approaches for MCOs including copayment offset policies, in which copayment reductions in some therapeutic classes are offset by copayment increases in other classes, and targeting of copayment reductions to enrollees at highest risk of cost-related nonadherence (e.g., a patient with CAD who indicates that he or she cannot afford the out-of-pocket cost for medication).

Lack of Evidence for Continuation of Stress Ulcer Prophylaxis After Hospital Discharge. Also in this issue of *JMCP* is a study by Thomas et al. of prescriptions for proton pump inhibitors (PPIs) filled within the first 30 days following a hospital discharge among patients with no PPI use in the 90 days prior to hospital admission.²⁶ Positing that these new PPI prescriptions represent a continuation of PPIs initiated during the hospital stay for prevention of stress ulcers, Thomas et al. assessed the percentage of PPI users who lacked an indication for ongoing acid-suppression therapy (AST) according to author-defined criteria. In an MCO with approximately 2.5 million members, 29,348 patients had a new post-discharge PPI prescription claim, of whom 20,197 (68.8%) had no indication for AST. The total cost for those PPI claims in the first 30 post-discharge days alone was \$3,013,069, of which \$2,148,122 was borne by the MCO.

As Thomas et al. observed, costs to the members were not only economic; chronic PPI use increases the risks of adverse events such as community-acquired pneumonia and *Clostridium difficile*-associated diarrhea.^{34,35} There may also be a duration-related increased risk of osteoporosis-related fractures, particularly hip fracture, with long-term use of PPIs.^{36,37} Moreover, although clinical guidelines suggest that stress ulcer prophylaxis is potentially appropriate only for patients in the intensive care unit (ICU) or coronary care unit (CCU),³⁸ only 37% of patients in the Thomas et al. study sample received care in either the ICU or CCU, and rates of inappropriate outpatient prescribing for ICU/CCU and general medicine patients were approximately equal.²⁶ Although Thomas et al.’s work does not address the important question of whether the inpatient use of PPIs in the study sample was appropriate, it does suggest that education of health care providers at the time of hospital discharge using evidence-based guidelines could decrease both the direct drug cost outlays and the clinical risks associated with unnecessary chronic PPI therapy.

How Should a Decision Maker View Claims About Risk Reductions?

In the context of escalating pressure on MCOs and patients to use resources more efficiently and effectively,¹⁴ an increasing focus on quantitative evidence is appropriate. Analyses that assess the costs of practices supported by questionable evidence, such as those reported by Melnick and Motheral and by Thomas et al., serve both as guideposts and as reminders that improvements are needed in the dissemination of evidence-based information to clinical and MCO decision makers. When presented with assertions that the added expense for a given policy or treatment will produce cost offsets, decision makers should ask specific questions about the validity and practical importance of research findings (Table 1). Given well-documented problems in the quality of the health care research literature,⁶ a healthy dose of “caveat emptor” is necessary for the decision maker who is asked to “buy” assertions of risk reduction.

Author

KATHLEEN A. FAIRMAN, MA, is Associate Editor and Senior Methodology Reviewer, *Journal of Managed Care Pharmacy*.

AUTHOR CORRESPONDENCE: Kathleen A. Fairman, MA, Kathleen Fairman LTD, P.O. Box 31278, Phoenix, AZ 85046. Tel.: 602.867.1343; E-mail: kfairman@amcp.org.

DISCLOSURES

The author reports no conflicts of interest related to the subjects or products discussed in this article.

REFERENCES

1. Anonymous. Biscuit incident threat evaluation (BITE). Mindlab International, LTD. September 2009. Available at: <http://www.rockybiscuit.co.uk/downloads/rocky-bite-report.pdf>. Accessed February 9, 2010.
2. Anonymous. Crumbs: half of Britons injured by their biscuits on coffee break, survey reveals. *Telegraph.co.uk*. September 8, 2009. Available at: <http://www.telegraph.co.uk/news/newstopics/howaboutthat/6153518/Crumbs-half-of-Britons-injured-by-their-biscuits-on-coffee-break-survey-reveals.html>. Accessed February 9, 2010.
3. Oliver L. Brits in killer biscuit warning. *The Sun*. September 7, 2009. Available at: <http://www.thesun.co.uk/sol/homepage/news/2625644/Brits-in-killer-biscuit-warning.html>. Accessed February 9, 2010.
4. Rocky Biscuit Company. Available at: <http://www.rockybiscuit.co.uk/bite-tool.html>. Accessed February 12, 2010.
5. Jackson M, Barbuto JP. Botox, migraine, and the American Academy of Neurology: an antidote to anecdote. *J Manag Care Pharm*. 2008;14(5):465-67. Available at: http://www.amcp.org/data/jmcp/JMCP_June08Web_465-467.pdf.
6. Fairman KA, Curtiss FR. Rethinking the “whodunnit” approach to assessing the quality of health care research—a call to focus on the evidence in evidence-based practice. *J Manag Care Pharm*. 2008;14(7):661-74. Available at: http://www.amcp.org/data/jmcp/661-674_FairmanCurtiss-Final.pdf.
7. Simpson RJ, Signorovitch J, Birnbaum H, et al. Cardiovascular and economic outcomes after initiation of lipid-lowering therapy with atorvastatin vs simvastatin in an employed population. *Mayo Clin Proc*. 2009;84(12):1065-72. Available at: <http://www.mayoclinicproceedings.com/content/84/12/1065.full.pdf+html>. Accessed February 16, 2010.
8. Drugstore.com. Price for generic simvastatin 20 mg. tablets, 90-day supply. Available at: <http://www.drugstore.com/pharmacy/prices/drugprice.asp?ndc=00093715456&trx=1Z5006>. Accessed February 12, 2010.
9. Drugstore.com. Price for generic simvastatin 40 mg. tablets, 90-day supply. Available at: <http://www.drugstore.com/pharmacy/prices/drugprice.asp?ndc=00093715556&trx=1Z5006>. Accessed February 12, 2010.
10. Drugstore.com. Price for brand Lipitor 10 mg. tablets, 90-day supply. Available at: <http://www.drugstore.com/pharmacy/prices/drugprice.asp?ndc=00071015523&trx=1Z5006>. Accessed February 12, 2010.
11. Drugstore.com. Price for brand Lipitor 20 mg. tablets, 90-day supply. Available at: <http://www.drugstore.com/pharmacy/prices/drugprice.asp?ndc=00071015694&trx=1Z5006>. Accessed February 12, 2010.
12. Zocor product label. Available at: http://www.merck.com/product/usa/pi_circulars/z/zocor/zocor_pi.pdf. Accessed February 12, 2010.
13. Lipitor product label. Available at: http://www.pfizer.com/files/products/uspi_lipitor.pdf. Accessed February 12, 2010.
14. Culler SD, Weintraub WS. Is initiation of atorvastatin for employees a good buy for employers? *Mayo Clin Proc*. 2009;84(12):1059-61.
15. RegenceRx Pharmacy Benefit Management. RxPrice guide. \$5.66 per prescription for generic simvastatin versus \$104.15 for Lipitor. Available at: <http://www.regencrx.com/learn/rxPriceGuide/index.html>. Accessed February 18, 2010.
16. Blalock HM. *Social Statistics*. New York: McGraw-Hill; 1972:162-63.
17. Diamond GA, Kaul S. Prior convictions: Bayesian approaches to the analysis and interpretation of clinical megatrials. *J Am Coll Cardiol*. 2004;43(11):1929-39. Available at: <http://content.onlinejacc.org/cgi/reprint/43/11/1929.pdf>. Accessed February 13, 2010.
18. Kaul S, Diamond GA. Trial and error: how to avoid commonly encountered limitations of published clinical trials. *J Am Coll Cardiol*. 2010;55(5):415-27.
19. Cannon CP, Weintraub WS, Demopoulos LA, et al. Comparison of early invasive and conservative strategies in patients with unstable coronary syndromes treated with the glycoprotein IIb/IIIa inhibitor tirofiban. *N Engl J Med*. 2001;344:1879-87. Available at: <http://content.nejm.org/cgi/reprint/344/25/1879.pdf>. Accessed February 13, 2010.
20. A detailed presentation of the Bayesian method is beyond the scope of this editorial. The interested reader is encouraged to view the helpful presentation in reference 17 for more information.
21. Bhatt DL, Fox KAA, Hacke W, et al. Clopidogrel and aspirin versus aspirin alone for the prevention of atherothrombotic events. *N Engl J Med*. 2006;354(16):1706-17. Available at: <http://content.nejm.org/cgi/reprint/354/16/1706.pdf>. Accessed February 13, 2010.
22. Yeh HC, Duncan BB, Schmidt MI, Wang NY, Brancati FL. Smoking, smoking cessation, and risk for type 2 diabetes mellitus. *Ann Int Med*. 2010;152(1):10-17.
23. Klein S. Study: quitting smoking raises diabetes risk. *CNN.com*. January 5, 2010. Available at: <http://www.cnn.com/2010/HEALTH/01/04/diabetes.smoking.risks/index.html>. Accessed February 13, 2010.
24. Hendrick B. Quitting smoking carries diabetes risk. *WebMD*. January 4, 2010. Available at: <http://www.webmd.com/smoking-cessation/news/20100104/quitting-smoking-carries-diabetes-risk>. Accessed February 13, 2010.

25. Melnick SJ, Motheral BR. Is ‘value-based’ value wasted? Examining value-based insurance designs through the lens of cost-effectiveness. *J Manag Care Pharm*. 2010;16(2):130-33.
26. Thomas L, Culley EJ, Gladowski P, Goff V, Fong J, Marche SM. Longitudinal analysis of the costs associated with inpatient initiation and subsequent outpatient continuation of proton pump inhibitor therapy for stress ulcer prophylaxis in a large managed care organization. *J Manag Care Pharm*. 2010;16(2):122-29.
27. Resources. VBI Calculator. CareScientific. Available free of charge at: <http://www.carescientific.com/resources.php>. Accessed February 15, 2010.
28. Andrews RA. The national hospital bill: the most expensive conditions by payer, 2006. HCUP Statistical Brief #59. Agency for Healthcare Research and Quality, September 2008. Available at: <http://www.hcup-us.ahrq.gov/reports/statbriefs/sb59.pdf>. Accessed February 16, 2009.
29. Chernew ME, Shah MR, Wegh A, et al. Impact of decreasing copayments on medication adherence within a disease management environment. *Health Aff (Millwood)*. 2008;27(1):103-12.
30. The charge to commercial payers per hospital stay for coronary artery disease was approximately \$45,000 in 2006 (reference 28).
31. Gavin K. The co-pay connection. University of Michigan press release. January 8, 2008. Available at: <http://www.med.umich.edu/opm/news-page/2008/drugcopay.htm>. Accessed February 15, 2010.
32. Chernew ME, Juster IA, Shah M, et al. Evidence that value-based insurance can be effective. *Health Aff (Millwood)*. 2010;29(3). [epub ahead of print]
33. Spaulding A, Fendrick AM, Herman WH, et al. A controlled trial of value-based insurance design—the MHealthy: Focus on Diabetes (FOD) trial. *Implement Sci*. 2009;4:19. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=19351413>. Accessed February 15, 2010.
34. Sarkar M, Hennessy S, Yang YX. Proton-pump inhibitor use and the risk for community-acquired pneumonia. *Ann Intern Med*. 2008;149(6):391-98.
35. Aseeri M, Schroeder T, Kramer J, Zackula R. Gastric acid suppression by proton pump inhibitors as a risk factor for *clostridium difficile*-associated diarrhea in hospitalized patients. *Am J Gastroenterol*. 2008;103(9):2308-13.
36. Yang Y, Lewis JD, Espstein S, Metz DC. Long-term proton pump therapy and risk of hip fracture. *JAMA*. 2006;296:2947-53. Available at: <http://jama.ama-assn.org/cgi/reprint/296/24/2947>. Accessed February 16, 2010.
37. Targownik LE, Lix LM, Metge CJ, Prior HJ, Leung S, Leslie WD. Use of proton pump inhibitors and risk of osteoporosis-related fractures. *CMAJ*. 2008;179(40):319-26. Available at: <http://www.cmaj.ca/cgi/reprint/179/4/319>. Accessed February 16, 2010.
38. American Society of Health-System Pharmacists. ASHP therapeutic guidelines on stress ulcer prophylaxis. *Am J Health Syst Pharm*. 1999;56(4):347-79.