measure what they are supposed to measure remains unclear.

Research is needed to examine which criteria for assessing the validity of cost-effectiveness studies are important determinants of study results and in what situations. For example, what is the relationship between quality scores (QHES, as an example) and treatment effect (i.e., cost-effectiveness measure)? Do lower-scoring studies tend to produce more variable estimates of cost-effectiveness? Do certain components of the checklist (e.g., sufficient time horizon) relate to the size of the treatment effect? Do quality scores vary across study type (i.e., randomized controlled trial, model, and observational study)? This type of methodological work is virtually extant in the pharmacoeconomic discipline, but with the plethora of quality checklists and the substantial resources devoted to the conduct of pharmacoeconomic studies, such a strategic approach seems viable.

Meanwhile, readers of economic evaluations should be cautious not to assume a false sense of precision in the use of summary quality scores since they generally have not been supported by empirical evidence, may actually be misleading, and are potentially more time consuming.

Brenda Motheral, RPh, MBA, PhD
Vice President of Outcomes Research
Express Scripts, Inc.

REFERENCES

1. Ofman JJ, Sullivan SD, Neumann PJ, et al. Examining the value and quality of health economic analyses: implications of utilizing the QHES. *J Managed Care Pharm.* 2003;9(1):53-61.
2. Moher D, Jadad A, Nichol G, Penman M, Tugwell T, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Control Clin Trials.* 1995;16:62-73.
3. http://p://www.cochrane.dk/cochrane/handbook/hbook67_Approaches_to_summarising_the.htm: Accessed November 2002.
4. Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA.* 1999;282:1054-60.
5. Emerson JD, Burdick E, Hoaglin DC, Mosteller F, Chalmers TC. An empirical study of the possible relation of treatment differences to quality scores in controlled randomized clinical trials. *Control Clin Trials.* 1990;11:339-52.
6. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA.* 1995;273:408-12.

Summary Quality Scores for Pharmacoeconomic Studies: Balancing Validity With Need

Once a product has received marketing approval from the U.S. Food and Drug Administration, decisions regarding insured access to these agents are immediately raised. The existence and amount of the insured benefit for specific agents requires weighing the evidence for clinical gains and their associated costs against similar measures for competing products and therapies.

Pharmacoeconomics provides a systematic, explicit, and objective basis for making and defending such drug benefit decisions. However, lack of standardization in the field¹ and the differences

in perspectives, knowledge, and interests across and within the producers and consumers of pharmacoeconomics has limited its impact on drug coverage decisions. As the methodology advances, consumers of pharmacoeconomic studies require an efficient tool to identify superior studies. In this issue of the *Journal*, Offman et al. propose such a scoring instrument, the Quality of Health Economic Studies (QHES).²

Beginning in 1973, clinical epidemiology has consistently identified large variations in the rates of performance of medical procedures and use of specific products.³ As health care costs have increased, drug costs and effectiveness analyses have become common; however, the explosion of pharmacoeconomic studies has also included some of uncertain quality, rigor, or validity. Pharmacoeconomic studies have nevertheless been subject to increasing standardization. Some are still viewed with skepticism by health plans and insurers who perceive the potential latitude in permissible assumptions as resulting in less than objective evidence. However, purchasers face constant pressure to determine the relative value of marketed pharmaceuticals and to make decisions with imperfect and disparate information. Analyses to assist these determinations come from multiple sources, with attendant variations in quality, reliability, validity, and timeliness of content. Consequently, the assessment of quality and validity of specific pharmacoeconomic results is at the center of the decision process, and uncertainty here will continue to influence the impact of pharmacoeconomic studies.

The proposed QHES instrument will be a substantial contribution if it assists end-users of pharmacoeconomic studies to discriminate among the exploding body of literature⁴ and efficiently identify the studies with superior merit. For producers of pharmacoeconomic studies, an accepted rating instrument could establish a clearer target—potentially encouraging higher quality and greater rigor. To achieve this level of acceptance and use, however, the QHES must demonstrate key validity characteristics.

A precondition for a valid rating instrument is that it be reliable. It must yield the same results on repeated trials. On this dimension, the qualitative nature of some of the QHES questions could mean lower reliability if the raters are not trained and their assessments not standardized. Otherwise, different observers may weigh the validity and reliability of health outcomes measures or scales differently. Without reliability, no instrument or measure can be valid.

Beyond being reliable, the QHES must rate studies on how well they actually answer the question posed by the research. Criterion validity, the closest aspect to what is commonly meant by validity, assesses the extent to which the measure being developed correlates with another, "gold standard" measure at the same time.⁵

Questions of external validity, or generalizability, are at the forefront of issues confronting decision makers as users of such information. Whether the original study has a societal, patient, provider, or health plan perspective will determine the relevance of results to a specific setting or decision maker. One of the biggest challenges in evaluating pharmacoeconomic studies may be the

interpretation and extension of the results to a different health care setting. Given the relative shortage of trained pharmacoeconomic analysts among management, clinicians, and other decision makers, such judgments often may be required of professionals who lack expertise in pharmacoeconomics.^{6,7}

The QHES was assessed for concurrent validity by comparison against the *British Medical Journal* checklist, the Canadian guidelines, and the *Journal of the American Medical Association* user's guide. Further, it was assessed against the global opinion of experts ("criterion validity") and validated among economists, some decision makers, and through the authors' own work. However, the method for selecting these experts, the use of convenience sampling, may present a selection bias and limit confidence in the extent of generalizability of the results.

Acceptance of an instrument as scientifically sound requires that it represent the full content of each of the attributes being measured ("content validity"). While content validity may be relatively easy to assess in established disciplines and with established tests, content validity has proven to be exceedingly difficult to establish with evolving concepts or disciplines, such as pharmacoeconomics. The QHES addresses many of the essential domains by which the soundness of an economic analysis is assessed; however, to the extent that it omits items pertinent to observational qualitative studies, its content validity might be compromised. Such studies may involve domains that are not captured by the questions in the QHES.

The value of an applied instrument is largely determined by its construct validity, a concept more appropriate to a dynamic field such as pharmacoeconomics. Construct validity is established over time by the consistency of findings across different QHES users. Such consistency was found among the experts consulted for this study; and, to that extent, the instrument was determined to have adequate construct validity. However, results from its application have yet to be demonstrated (a) across the spectrum of decision makers from health plans, managed care providers, pharmacy benefit managers, hospital Pharmacy and Therapeutics committees, or researchers, and (b) for the range of the decisions that must be made.

In general, however, it is important to note that the concept of validity is broader than just the validity of individual aspects or measurement approaches. The QHES attempts to synthesize health economic evaluations so that they are useful in decision making and, ultimately, insurance coverage determinations as well as the development of practice guidelines. Summary scores should constitute just one component of an economic evaluation. Until a gold standard for pharmacoeconomic studies is developed, more research is needed to strengthen the link between theory and practice.

Survey research supports decision makers' ability to discern differences and to balance the overall influence of socioeconomic assessments that vary in quality, availability, timeliness, comprehensiveness, and validity.⁷ Ultimately, executives and managers must make timely decisions—often with incomplete information or with information from sources from which potential conflicts of interest cannot be ruled out and must be balanced with their

results.⁸ Consequently, the QHES may initially be more useful as the first of a 2-stage screening process for decision makers under time pressure and with limited resources and less experienced analysts. Not surprisingly, the authors noted that the QHES had greater acceptance among decision makers than among health economists. For the former, the QHES might efficiently help identify those studies not to be included in a dossier or subjected to more rigorous assessment. A subsequent, in-depth second stage might then focus time and resources on critical examination of the remaining studies—possibly not being bound or influenced by the initial QHES score used to sort the studies initially.

To remain relevant for pharmacoeconomic studies, the QHES instrument must evolve and continue to improve in reliability and validity for a broad spectrum of decision makers. Performing pharmacoeconomic assessments is time-consuming and expensive; if the QHES can validly expedite these reviews, it may lower their initial cost and encourage more timely updates.

Just as the QHES will require refinement as experience with it accumulates, users will also need to address the minimal competencies required of those who use it and who make decisions based upon it. Even if the QHES validly predicts the quality of the studies being used by formulary decision makers, it does not, and probably should not, predict the extent to which these studies influence the decision-making process. A screening instrument such as the QHES should probably not be a replacement for expertise in pharmacoeconomics, only a supplement to it.

Fadia T. Shaya, PhD, MPH

Assistant Research Professor, Outcomes Research/Epidemiology,
Center on Drugs and Public Policy
University of Maryland School of Pharmacy

C. Alan Lyles, ScD, MPH

Associate Professor, Division of Government and Public Administration
University of Baltimore

REFERENCES

1. Mullins CD, Ogilvie S. Emerging standardization in pharmacoeconomics. *Clin Ther*. 1998;20(6):1194-1202.
2. Ofman JJ, Sullivan, SD, Neumann PJ, et al. Examining the value and quality of health economic analyses: Implications of utilizing the QHES. *J Managed Care Pharm*. 2003;9(1):53-61.
3. Wennberg J, Gittelsohn A. Small area variations in health care delivery. *Science*. 1973;182:1102-08.
4. Curtiss FR. Crossing the quality chasm—incremental changes through clinical practice guidelines (CPGs) [editorial]. *J Managed Care Pharm*. 2002;8(5):400-01.
5. Trochim, W. *The Research Methods Knowledge Base*. 2nd ed. Cincinnati, OH: Atomic Dog Publishing; 2000.
6. Lyles A. Decision-makers' use of pharmacoeconomics: what does the research tell us? *Expert Rev Pharmacoeconomics Outcomes Res*. 2001;1(2):133-44.
7. Lyles, A, Luce BR, Rentz A. Managed care pharmacy, socioeconomic assessments and drug adoption decisions. *Soc Sci Med*. 1997;45(4):511-21.
8. Luce, BR, Lyles CA, Rentz A. Datawatch: the view from managed care pharmacy. *Health Aff*. 1996;15(4):168-76.