Managing Utilization Management: A Purchaser's View

Many purchasers rely on utilization management to lower health care costs. Should we be doing more to ensure its safety?

by Arnold Milstein

UTILIZATION MANAGEMENT (UM) and the reduced volume of health care services it typically fosters have struck a nerve. Over the past five years, U.S. media, legislators, and opinion polls have expressed increasing concern that reduced service levels may be jeopardizing patients' well-being.

The definition of *utilization management* varies widely in everyday use as well as in the published literature. The Institute of Medicine (IOM) in a 1989 report adopted a narrow focus on case-by-case preservice review sponsored by purchasers.[1] Eve Kerr's recent study of utilization management in capitated medical groups encompasses a broad array of methods such as physician incentives and primary care gatekeepers.[2] The definition adopted in this paper includes all interventions originating outside the physician/patient relationship with an intent to promote an economical mix of health care services.

Earlier perspectives on the value and safety of utilization management were commonly rooted in three sources of evidence. The RAND Health Insurance Experiment demonstrated that for most patients, substantially reduced utilization levels appeared to be acceptable and safe.[3] A second RAND study examined the impact of Medicare's prospective payment system (PPS).[4] It found that rapidly and substantially reduced lengths of hospital stay were not associated with detectable reductions in quality, although small adverse effects on outcome could not be ruled out. This research was germane to the issue of UM safety because intensified utilization management was a salient feature of hospitals' response to prospective payment.

The 1989 IOM study focused explicitly on utilization management.[5] It concluded that utilization management probably had reduced expenditures for some purchasers and found no evidence of quality reduction. However, both conclusions were carefully qualified with an emphasis on the inadequacy of available evidence. These and other studies of utilization management or reduced service levels did not trigger widespread initiatives to better ensure UM safety. Should such initiatives be considered today? The combined impact of six factors favors more active management of utilization management.

The Need To Better Manage Utilization Management

■ **Dramatically reduced utilization levels.** The leading edge of reduced utilization levels is far below levels found to be safe. James Robinson and Lawrence Casalino recently examined the 1994 utilization experience of a large number of nonelderly health maintenance organization (HMO) enrollees in six aggressively managed, capitated medical groups.[6] While a comparison of their findings with RAND's findings for HMO care during 1976–1980 is necessarily crude, lacking adjustments for age, income, morbidity, and outpatient surgery use, the order of magnitude of difference

Arnold Milstein is medical director of the Pacific Business Group on Health and managing director of William M. Mercer, Inc., both in San Francisco.

is stunning. Compared with a 1976–1980 fee-for-service plan, 1994 utilization levels in these advanced managed care medical groups were more than 25 percent lower for annual physician office visits and more than 80 percent lower for hospital days. Even if adjusting for uncontrolled variables were to reduce the differences by half, the result would be a sea change in service levels, without accompanying evidence of safety.

■ **Lack of evidence-based standards.** In judging the safety and patient acceptability of UM programs, we lack an evidence-based framework. While, as described below, one research group is making initial progress in studying the safety of inpatient utilization review for two diagnostic groups, what we know pales in comparison to what we would need to know to root assessment of the safety of utilization management in evidence-based standards. First, UM methods need to be examined in logical combinations. Second, major variables embedded in the design and execution of each UM method need to be specified. For example, with respect to preservice utilization review, how often should review physicians speak directly with treating physicians? Resolution of this controversial variable requires assessment of a wide range of frequencies and their near-term impact on cost, quality, and patient satisfaction. In addition, the long-term impact of such a "high-friction" UM method on physician satisfaction, demeanor, and recruitment needs to be considered. It is unlikely that researchers will make a significant dent in the implied mountain of research in the foreseeable future.

■ **Less discernable decision rationale.** As utilization management expands beyond utilization review, it is becoming less available for scrutiny and safety checks. Kerr's study of utilization management in capitated California medical groups showed that the three most widely implemented and potent methods were utilization review, physician gatekeepers, and financial incentives.[7] The latter two methods exert their effect on utilization levels via reasoning that unfolds in the distant reaches of physicians' conscious and unconscious thought processes. The majority of these thought processes are not documented in the medical record and thereby resist accountability.

■ **Inadequate detection of impaired quality.** We lack adequate sensors to detect quality impairments associated with utilization management. Although the National Committee for Quality Assurance (NCQA), the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), the Agency for Health Care Policy and Research (AHCPR), the Health Care Financing Administration (HCFA), and health services researchers are making progress in measuring quality, reliable early detection of adverse impact on quality remains largely elusive. Implementing comprehensive clinical information systems and gaining knowledge of process/outcome relationships are the most formidable underlying challenges. We are unlikely to meet them over the near term.[8]

■ **Flawed UM programs.** Although a 1994 review of evidence on differences between managed care and traditional insurance programs concluded that managed care's lower utilization levels did not appear to reduce quality and satisfaction, three more recent studies raise concern.[9] John Ware's 1996 study of four-year longitudinal change in health status showed that four HMOs examined in the late 1980s were less effective than fee-for-service care in maintaining the physical health status of older, sicker patients.[10] More specific to the issue of UM safety are findings from Kerr's study and more recent research by Tom Wickizer and Daniel Lessler at the University of Washington.[11] In Kerr's survey of capitated California medical groups, 30 percent of responding medical groups did not use any written clinical guidelines to anchor their UM programs, and an average of 9 percent of utilization review denials were made by nonphysicians. Wickizer's and Lessler's studies analyzed inpatient utilization review decisions and health insurance claims data from a private insurance company

for a group of patients with psychiatric diagnoses and a group with cardiovascular disease. Both studies found a significant relationship between limiting length-of-stay through utilization review and the risk of readmission.

■ **Few industry safeguards.** Politicians are beginning to legislate utilization standards. In the past two years, several state legislatures and Congress have passed legislation governing minimum lengths of maternity stay. In associated legislative hearings, the health care industry had little with which to reassure the public that UM methods were being carefully evaluated and managed.

Potential New Safety Features

A number of complementary paths are available to better assure UM safety.

■ **Strengthened accreditation.** We can stimulate health care accreditation organizations to expand and intensify accreditation standards for the UM component of delivery systems, health care management companies, and insurers. The Utilization Review Accreditation Commission's (URAC's) standards for utilization review organizations comprise a thoughtful basis for utilization review accreditation. Accreditation standards for other UM methods remain comparatively underdeveloped. In addition, existing UM accreditation standards could be targeted to topics of greatest concern. For example, to address concerns about underreferral to specialists, a meaningful sample of primary care medical records could be clinically audited for this problem during accreditation visits.

■ **Product labeling.** We can develop practical, easily comprehensible approaches to informing purchasers and consumers of the UM methods operating within health plans, hospitals, and medical groups. At a minimum, this should occur at the point of plan enrollment and provider selection. The NCQA's Health Plan Employer Data and Information Set (HEDIS) 3.0 encompasses this approach. However, further refinement will be required for this information to be interpretable by average Americans.

A product-labeling approach might also encompass measurable consequences of utilization management, such as service volume and patient satisfaction with service volume. To pursue this avenue, a small marketbasket of readily interpretable indices of service parsimony would need to be developed and universally reported. For example, indices of parsimony for length of maternity hospitalization might include the percentage of uncomplicated vaginal deliveries in which the length of postdelivery stay was less than twenty-four hours and/or was subsequently rated as fully satisfactory by the mother. This "consequence-oriented" approach to labeling might be more easily grasped by consumers than a listing of UM methods.

■ **Clinical ombudspersons.** We can increase the availability of clinically trained or supported ombudspersons through whom patients can (1) conveniently obtain independent assessments of the appropriateness of treatment recommended by their physician or health plan; and (2) weight neutrally presented treatment options with their personal health-related values, such as tolerance for risk, pain, disability, and uncertainty. To succeed, such programs will need to incorporate a nuanced understanding of the psychology underlying deference to physician direction, particularly by older and sicker patients.

■ **Evidence-based utilization management.** We can encourage collaboration between the health services research community and UM program managers. Subject to the limits of market competition, this will allow the results of implicit UM trials by this country's diverse managed care entities to be captured and "best UM practices" to be more rigorously derived and disseminated.

■ **Expanded use of utilization management to serve quality.** Most UM methods are also useable as vehicles for quality improvement beyond the quality gains associated with eliminating unnecessary services. However, this potential remains substantially underused. For example, utilization review of hospital care and case management of high-

cost patients are promising vantage points for identifying and correcting past or impending medical error; yet they are rarely used for this purpose, except in response to egregious quality failures. This reorientation also would make UM staff more alert to utilization management's quality risks.

■ **Revised insurance plan language.** As utilization levels decline, UM implementers need guidance on how to handle low-gain, high-cost services and how to incorporate variables such as consumer preferences and family care burden. This can be substantially advanced by reforming antiquated, narrow, and vague insurance plan language governing utilization management, such as "medically necessary," "reasonable," and "consistent with prevailing local medical practice."

■ **Improved quality sensors.** We need to accelerate the advancement of publicly available quality measures, so that the market risk to plans and providers of ill-considered UM program design and implementation is more prominent. This will reduce the need to monitor utilization management and other processes of care management. It also will require major investment in health care information systems and outcomes research.

Conclusion

No responsible stakeholder in American health care supports a wasteful service mix. Superfluous services drive higher health insurance costs and avoidable treatment risks. Accordingly, utilization management is an essential element in any delivery system that aims to satisfy value-seeking patients and purchasers. However, signs that utilization management may be cutting beneficial health care are discernible in recent research. Although the medical groups described by Robinson and Casalino are highly regarded by their peers, the prospect of less well managed delivery systems reaching for identical results is worrisome.

Utilization management is a proven savior and a potential devil wrapped in a single package. Our challenge is to maximize utilization management's net contribution to health care value while navigating in a fog of ignorance about its impact on quality. Safety enhancements may allow us to reduce the frequency of harmful collision while we map what works in medicine, reconcile it with patient and societal welfare, and transform utilization management into a better champion of these ultimate lodestars.

NOTES

1. B.H. Gray and M.J. Field, *Controlling Costs and Changing Patient Care?* (Washington: National Academy Press), 2–3.
2. E.A. Kerr et al., "Managed Care and Capitation in California: How Do Physicians at Financial Risk Control Their Own Utilization?" *Annals of Internal Medicine* (October 1995): 500–504.
3. W.G. Manning et al., "A Controlled Trial of the Effect of a Prepaid Group Practice on Use of Services," *The New England Journal of Medicine* (7 June 1984): 1505–1510.
4. K.L. Kahn et al., "Comparing Outcomes of Care before and after Implementation of the DRG-Based Prospective Payment System," *Journal of the American Medical Association* (17 October 1990): 1984–1988.
5. Gray and Field, *Controlling Costs.*
6. J.C. Robinson and L.P. Casalino, "The Growth of Medical Groups Paid through Capitation in California," *The New England Journal of Medicine* (21 December 1995): 1684–1687; and Manning et al., "A Controlled Trial of the Effect of a Prepaid Group Practice on Use of Services."
7. Kerr et al., "Managed Care and Capitation."
8. National Committee for Quality Assurance, *A Road Map for Information Systems, HEDIS 3.0* (forthcoming).
9. R.H. Miller and H.S. Luft, "Managed Care Plan Performance since 1980," *Journal of the American Medical Association* (18 May 1994): 1512–1519.
10. J.E. Ware et al., "Differences in 4-Year Health Outcomes for Elderly and Poor, Chronically Ill Patients Treated in HMO and Fee-for-Service Systems," *Journal of the American Medical Association* (2 October 1996): 1039–1047.
11. Kerr et al., "Managed Care and Capitation;" and T.M. Wickizer and D.S. Lessler, "Containing Costs while Maintaining Quality: An Unresolved Dilemma for Managed Care within Mental Health," and "The Impact of Utilization Management on Hospital Care for Patients with Cardiovascular Disease" (Presented at the Thirteenth Annual Meeting of the Association for Health Services Research, Atlanta, Georgia, June 1996).