RECORD OF SOCIETY OF ACTUARIES
1987 VOL. 13 NO. 1

UTILIZATION REVIEW AND QUALITY OF MEDICAL CARE

Moderator: HOWARD J. BOLNICK
Panelists: MICHAEL F. ANTHONY*
MARSHA LADENBURGER**
ARNOLD MILSTEIN, M.D.***
Recorder: LEONARD KOLOMS

- Legal liability for unfavorable medical outcomes
- Effect of utilization review on the quality of care
- Review of recent care decisions affecting utilization review/quality assurance activities
- Protocol development and implementation for managed care
- Types and levels of insurance coverage necessary to protect your organization
- Operating concerns that affect the effectiveness of utilization review and patient's and provider's acceptance of these programs

MR. MICHAEL F. ANTHONY: My presentation today will focus on utilization review and more specifically the evolving theories of liability attendant to utilization review programs. After discussing the basis for liability, I will touch upon key points of the seminal case in the area, other grounds for potential liability and appropriate protections which can be effected by utilization review programs to minimize liability exposure.

* Mr. Anthony, not a member of the Society, is a Partner with McDermott, Will & Emery in Chicago, Illinois.

** Ms. Ladenburger, not a member of the Society, is Director, Quality and Utilization Management with the Daughters of Charity in St. Louis, Missouri.

*** Dr. Milstein, not a member of the Society, is President of the National Medical Audit in San Francisco, California.

The type of liability we are most concerned about in utilization review programs results from the theory of negligence in the conduct of such programs. Such negligence is not professional malpractice liability but rather the non-intentional tort of negligence which consists of four basic elements. These elements are (1) duty, (2) breach, (3) injury and (4) proximate cause. Applying these four components to a utilization review program, the duty is one of responsibility for careful payment/care decisions; a breach results from a flawed utilization review decision making process resulting from inadequate skill levels of those making decisions, amorphous appeal rights and lack of appropriate checks and balances to protect procedural due process; patient injury or aggravation of injury must be established; and the breach must be the direct cause of such injury.

There are varied organizational and administrative arrangements for utilization review programs. A common thread which runs through these programs is the prospectively communicated financial constraints on the type, amount, duration, frequency and/or setting of health services. The constraints vary in degree from the subtle to the stark. As these utilization review programs embody economic incentives to limit care, we are likely to see careless pre-admission, concurrent or referral review mechanisms in place which undoubtedly will lead to patient injury. Despite concerns of consumers and providers alike regarding the impact of utilization review on quality of care, very few cases have been documented to provide guidance about the risk of liability. It is thus not at all clear what duty of care is owed to patients by prospective utilization reviewers and whether or how negligent breach of any such duty might be linked to subsequent patient injury so as to give rise to liability.

We turn to the much discussed California decision of *Wickline vs. State of California* for such guidelines. Lois Wickline was hospitalized in 1977 for surgery to correct blocked circulation in her legs. Following Medicaid program guidelines, one of her physicians obtained prospective approval for a 10-day hospital stay for the surgery. As a result of complications during that period, the physician filed a request for an 8-day extension which was received by the designated Medicaid nurse reviewer. This nurse reviewer decided that 8 days appeared to be medically inappropriate and thus telephoned the Medicaid physician consultant on duty. This physician approved only an additional 4 days over the telephone. The patient was discharged on the order of the three physicians caring for Wickline at the end of the 14-day stay. Shortly

thereafter, Wickline began experiencing complications and had to be re-admitted to the hospital where her leg was amputated.

At the trial, two of Wickline's physicians testified that her discharge after 14 days was medically inappropriate. The third physician stated that he would have liked to have kept her in the hospital for some additional time to handle what he called "control problems." The bottom line was the jury in Los Angeles came back and awarded $500,000 for Wickline. In essence, they said that the Medicaid utilization review decisions adversely affected the outcome in terms of Wickline's treatment plan. The case was appealed and the activity at this level became much more interesting. The California Hospital Association joined with the California Medical Association (CMA) in supporting Wickline's position. These two associations wanted to demonstrate that liability can accrue when a prospective review proceeding is improperly designed or implemented. In this process of appeal, we finally see some guidelines emerging as to what utilization review programs should look like. The CMA said that the decision maker in the utilization review process should have adequate information about the patient's condition.

Think back for a moment on what took place in *Wickline*. The nurse reviewer looked at a form submitted requesting an extended stay by the attending physician. The nurse reviewed that form, flagged it for what she considered to be an excessive stay for the submitted diagnosis and she then called the consulting physician. This was standard procedure. We do not know if the physician called and acted on the basis of what was read from that form over the phone, and the case does not mention that the physician should perhaps have gone behind that form to the medical record and looked more deeply into the case. The CMA also stated that the decision maker has to be informed before making a decision about the patient's stay in the hospital.

The CMA also stated that sufficient expertise must be applied to the utilization review process. Should the nurse have made that initial screen or should it have been a physician? When looking at the physician involved in the process, should he have been a cardiovascular surgeon to opine in this case? These questions were not answered on appeal but the CMA recognized that they are important points in evaluating the efficacy of a utilization review program. Another question posed by the CMA is whether there is enough time for the

physician and nurse reviewer to consider the case? Ultimately, the CMA suggests there should be swift non-burdensome appeal mechanisms in place. These appeal mechanisms provide the attending or treating physician due process thus, in theory, avoiding unnecessary intrusion in medical judgment. Lastly, the CMA said the utilization review program should not be controlled by those with a financial interest. This poses some interesting questions, such as what does an insurer do if it wishes to implement a utilization review program? Does a physician group that takes capitation payment in an HMO arrangement and provide utilization review have a financial interest?

The result of the appeal was a reversal of the original decision. The appeal court held that the State of California was not liable as a matter of law. Thus, we still do not have a case where a utilization review program was held to be liable for the decisions made with regard to length of stay and reimbursement. The case is currently moving toward further appeal.

The dictum in this case is what has generated extreme interest. Dicta in cases are not binding on other courts, so if other courts in the State of California or outside of the State of California are faced with similar cases, the *Wickline* case would not bind these other courts. The dictum, however, is indicative of the evolving thinking of the courts on utilization review programs and attendant liability. The important dictum in *Wickline* held that a payor may be liable when "medically inappropriate decisions result from defects in the design or implementation of a cost containment program." What we haven't been left with by *Wickline* is an adequate definition of what constitutes defects in such programs. The CMA has attempted to establish guidelines for determination of such defects but there are probably other defects which could arise in these programs not mentioned by the CMA which could be the basis for liability.

The underlying theme of *Wickline* is that cost limitations imposed cannot corrupt medical judgment. A physician in this case should probably have protested the denial decision for the extension of stay. What we do have is a new liability rule emerging which has very uncertain application to the wide spectrum of utilization review programs in various parts of the country. We still do not know what the courts consider to be a defect in the utilization review process with adequate certainty to identify a breach of duty leading to a negligence determination. We also do not know when such a defect impinges so significantly

on medical judgment so as to be the proximate cause of an injury. What we do know is that the courts are still wrestling with the segregation of payment versus patient care decisions. Clearly the courts recognize payment decisions should be separate and apart from medical care decisions.

Although *Wickline* is often quoted as the seminal case in the area of potential liability resulting from utilization review programs, at least one HMO that we are aware of in California actually used *Wickline* to have the HMO removed as a defendant when so named in liability cases for its utilization review activities.

In June of 1986 the AMA, perhaps as a result of *Wickline*, through its Council of Ethical and Judicial Affairs tried to clarify the role of the physician in treating the patients involved in these utilization review programs. The two key points established by the AMA are that if a doctor disagrees with a utilization review decision, the patient should be so informed and the doctor cannot rely solely on utilization review decisions in making treatment decisions. The physician must always be guided by standards of good medical practice in deciding if care should be given to a patient.

Other legal theories creating a basis for liability generated by utilization review programs are available. Those of you who work with insurance companies are probably more familiar with bad faith breach of contract. Laws imply good faith and fair dealing in contracts by and between insurance companies and subscribers. If a utilization review decision is construed to be benefit determination and the insurer is determined to be inappropriately balancing the interests of the insurer and the insured, the insurance company may well find itself in court under the theory of bad faith breach of contract with potential for significant damages related to such an action. Such an action can produce economic damages, emotional distress and punitive damages. Because this action is founded in contract and is not a negligence action, the same level of proof is not needed to establish liability. For example, if you cannot obtain appropriate expert witnesses in a negligence case, the bad faith breach of contract might be the preferable claim. I suspect that you will see more and more of these kinds of cases arising if utilization review becomes a more popular concept with insurers. In subscriber certificates and other written documents communicating information to patients, the insurer should make it very clear that decisions

made as part of the utilization review program are not medical decisions and ultimately the physician remains responsible for medical decisions.

To finish up, let's talk briefly about protections that can be implemented. How do you adequately protect the utilization review program under the evolving theories of liability we have discussed today? First, such programs must be carefully implemented, perhaps using the CMA guidelines in designing the program. Also, the program should be adequately insured and this means coverage for bodily injury, defamation, interference with contract, invasion of privacy, business defamation and other coverages as deemed appropriate by your insurance carrier. Your insurance policy should recognize the existence of the utilization review program so that the insurer understands the activities of this program are covered by the policy. The policy should cover the organization lodging the utilization review program, its employees, committee members, the agents of the entity as well as the directors and officers. Anyone involved in the review process should be covered. Last, a clear statement should be included in subscriber certificates that outlines that the decisions of the utilization review program are not decisions that should affect medical care and that the physician is ultimately responsible for the medical care of the patient. In contracts with physicians, mention should be made of the fact that utilization review decisions relate to payment and physicians remain responsible for appropriate medical judgment.

In summary, there is little case precedent to give us clear guidelines as to the liability which is attendant to utilization review programs. Over the next two to three years, I would suspect we will see significantly more litigation evidencing the fact that the *Wickline* case is likely the tip of an iceberg. These new cases involving utilization review programs would likely produce both good and bad results. The good result would be that we will have clear guidelines to work with in designing such programs and the bad result would be new liability exposure for providers.

MR. HOWARD J. BOLNICK: I'd like to continue now with Dr. Arnold Milstein talking about the same issues from the medical perspective.

DR. ARNOLD MILSTEIN: The core idea behind National Medical Audit was to put independent physician judgment to work on behalf of people who were either

buying or sponsoring managed care systems, so they could get an independent frame of reference on how well those managed care systems were working. Of the various applications of this general concept, evaluating utilization review systems performance has probably been our greatest area of activity.

Let me start with a couple of key background concepts. The first is the concept of how much fat -- how many unnecessary services exist in current average American health care. There are a number of ways of evaluating this problem. One window that we've got on this problem is a study we did for Chevron Corporation. Chevron Corporation called us in a couple of years ago about starting a utilization review program. The finance department asked how much this program was going to cost. For their large number of employees the program is going to cost $750,000. The people in the financial department said before they spent that kind of money they wanted to know whether they had a problem. They had us pull a random sample of 120 claims paid for in the prior quarter that had not been under utilization review. We got the medical records and then used our independent physician auditor network to tell them how many days that they were currently paying for would not be approved by a topnotch UR firm. For Chevron, 32% of the days that were in that sample of 120 hospitalizations would not have been blessed by a topnotch UR firm. Clearly this magnitude of problem warranted an expenditure of $750,000.

Now we're not the only ones who have done studies like this. Last fall in the *New England Journal of Medicine* the Rand Corporation published a four-year study -- really asking the same question and randomly pulling medical records. The main difference between their study and ours is that they didn't just look at private employer cases, which we did. They also looked at Medicare, Medicaid, Workers Compensation, etc., and they came in with orders of magnitudes of unnecessary days that were roughly the same as ours. So the first key concept is the magnitude of excess hospitalization is large.

The second key concept is really an expansion of the first concept, and that is there is really so much room for improvement in an average, unreviewed health care plan that a relatively weak UR Program can look deceptively good. To illustrate this I'll take the Chevron example one step further and I'll assume that because they had lousy consultants on the case they ended up picking a mediocre UR firm, a UR firm that is only going to be successful in taking a quarter

of that 32% and reducing it. So at the end of the year Chevron's hospitalization will have gone down by 8% rather than some larger proportion of the 32%.

Still the study will show that the net of all expenses got a 6-to-1 return on that investment. But the key concept here is that this UR program would have yielded less than half of the potential savings that could have occurred had they put in a first rate UR program. There actually should have been more than double the 6 to 1. Unless they have a good perspective on what was possible from good UR they might be deceived by the 6-to-1 return investment. In fact I see some UR program sponsors proudly marketing the fact that they're yielding 2-to-1 and 3-to-1 return. I consider that a disgrace.

Let's now jump ahead to the question of how one might go about getting a handle on performance of a UR program. These are three generic basic windows on performance.

The first one can look at operational factors and there is a large list of things to look at in trying to judge a UR program. A Seattle firm recently put together a national survey of every UR program that answered a very exhausting list of operational questions, like: What is the ratio of RNs for 1,000 covered lives? How often do cases get referred to physicians for review? Typically these get done on paper. We've also been hired to actually be on site and often that is an eye opener. Often we find some very major discrepancies between how the UR program describes its functioning on paper and what you find when you actually go on site and interview the staff and observe what is happening.

Another window in performance is to analyze the impact that your program is having on utilization. This typically gets done either by examining reductions of length of stay. Another common measure is to evaluate reduction in hospital days of care per 1,000 covered lives. Of all the windows on performance that are available this is the most economical and efficient way of getting a quick handle on how well UR programs are performing. The drawback to this approach is that often the assumptions one has to make in order to drive a judgment often or cannot be founded on trustable facts.

The first of two barriers that we often run into using this mode of analysis is bad data. That is the self-insured employer or the carrier doesn't trust their data either on hospital days or on enrollment.

The other barrier is that it is often not uncommon for a health insurance program sponsor or a self-insured employer to be doing their utilization review program. They'll do other things in plan design that will also affect hospital utilization. For example add: a $250 deductible per hospitalization; a home healthcare benefit; a psychiatric half-way house benefit; put a per limit cap on mental health and chemical abuse, etc. The difficulty is then to allocate credit for that reduction over the various interventions that have been taken. Partly, as a result of some of the frustrations and drawbacks associated with these first two methods of evaluating performance, we began about 2 or 3 years ago to take a different, more clinically oriented task to evaluate UR performance. We began doing what we did for Chevron but doing it after it had been in place for a while. We began randomly sampling hospitalizations that had been subject to utilization review. Then we went to the hospitals, getting the medical records and sending the records out for independent audit by board certified physicians who were case-matched by specialty to the particular case under consideration and physicians who have a lot of experience in UR and who also were not affiliated with the UR system. How many days that we've blessed by this UR system were not medically necessary? What you are essentially doing is calculating an error rate which can be an error rate based on a percentage of days inappropriately approved or you can calculate your error rate by the percentage of cases approved. For those of you who read *Business and Health*, in February we published a summary of some of our audit findings. That article gives you a sense of a very broad range and performance that we are currently finding when we calculate these error rates for insurers, PPOs, HMOs, or self-insured employers.

Over time we have had a chance to develop this other method even further. One of the things we have started to do is to take a given error rate and divide it up over different categories of review and different medical specialty categories. Let's say 40% of the days approved by the UR system were not medically necessary -- and by the way that order of magnitude of error is unfortunately not that uncommon -- but by taking that 40% you can allocate it over different categories of review and have a sense as to exactly how the system is failing.

Let me show you some real results. Here are results from a system we evaluated about a year and a-half ago. This system had an overall error rate of about 36-38%. We then took this 36-38% and we divided up the number of possible days that they judged over different categories of review. The first subscale is pre-operative days. Of the preoperative days that were subject to judgment, 3% were inappropriately approved. Now a 3% error rate is perfection. That is, I have described to you a retrospective audit process where my auditors have perfect hindsight.

They are not the ones who are on the front line being subject to anger and threats by attending physicians and so there is a certain allowance for discrepancy that you have to make just because we have the advantage of the retrospective scope. In general any error rate of 10% or less we consider to be perfection and we give the system an A+ on that scale. They also did great on judging the necessity of surgery. We saw no situations in which surgeries whose indications were not clearly documented were approved by the system. So they got a perfect score on surgical necessity review.

The next category is necessity of admission for non-emergency admissions that they had a chance to block and for the days associated with those unnecessary admissions. Twenty-five percent of the days for this particular review system within the category of admission review were inappropriately approved. Obviously they did relatively poorly here.

The next subscale is necessity of continued stay. That is, once you have made a judgment about admission there is a certain minimum number of days that are often implied in that admission decision. After that initial number of days are over continued stay review begins.

This fourth subscale represents error rates within this subcategory of review and this was the worst area of performance for this particular UR System. Forty-one percent of the continued stay days that they reviewed and approved we thought were inappropriately approved.

The last category is avoidable delays and services. Many UR systems take it upon themselves to be on the lookout for situations in which patients are subject to hanging around the hospital just because the hospital was not efficient in

providing care. For example: A patient is supposed to have an X-ray on a given day and for one reason or another there is a foul-up in the transportation department and no one ever comes to pick up the patient. Some UR systems have it in their contract to judge and be on the lookout for these days. This particular system only had a 7% error rate on these days and that is obviously a perfect score.

You can also take this error rate and divide it over medical specialty categories. That also is often quite revealing. The first category is medical cases and 38% of the days for patients admitted for medical reasons; that is internal medicine cases, were days on which we thought a top flight UR System would not have given their blessing. That was a poor score. In terms of surgical cases the error rate was 9%; OB-GYN was 4.5%; pediatrics was 10%; in psych and substance abuse, 47% of the days.

Now when you look at that last subscale, psych and substance abuse, 47%, you might have predicted that. I know that psychiatry is an area where the symptoms are subjective; in which there is a lot of controversy in the field as to what constituted the right treatment, so I am not surprised that the score was lousy on that scale. Our feeling is that there is enough variation and opinion within psychiatry that 10% is not a reasonable standard for perfection on retrospective review. We have seen scores as low as 20% of this subscale and we think that any score of 20% or lower constitutes an A+ for psychiatry and substance abuse. We do see UR programs able to get this retrospectively judged error rate down to about the 20% range. We would have given them an A had they been down to 20%, but a 47% error rate is not an A.

Let me now jump from methods of evaluating these systems to talk about the view point of myself and the other staff and physicians at National Medical Audit who had the opportunity over the last 36 months to audit over 200 UR programs both for government and private industry, and carriers and HMOs.

The first judgment we would come to is that now we do not feel that quality of care is being substantially or very frequently jeopardized by utilization systems. As part of our audit process we ask the question, "Is there any evidence in this hospital record or on this review worksheet for a patient being pushed out of the hospital too soon?" So far, we see very little evidence of this, and this is

12 to 14 thousand medical records. There is very little evidence on the private industry side and very little evidence on the Medicare side. But again, based on that medical technical perspective both on the private industry side and the Medicare side, we just do not see evidence of major quality of care risks caused by overly aggressive UR. The endemic disease in utilization review is not overly aggressive URs; it is approvals that do not deserve to be.

The second general observation we would make is that building early alliances with patients really pays off. When we look at which UR systems are doing the best on our audits and achieving the biggest days of care reductions, are not causing patient unhappiness, and are not drawing a lot of threatened law suits, but are the people with the ideal combination of happy patients and good healthy results. One of the things we consistently see is a very savvy attempt to build an alliance with the patient early in the process. It is typical of these systems, or they will say, "Don't have the doctor call us first in your program design, have the patient call us first." Then they will use that early contact with the patients or their families to build a very supportive and very positive relationship. If the patients have any questions or worries that are on their minds they will be invited to discuss them. In American culture there is this principle of reciprocation.

People who are nice to you want you to be nice to them. This initial positive relationship really pays off later on.

The second way these alliances work is the good UR systems also use them as a basis for building patient expectations regarding short length of stay. Some of the very good and very clever UR systems we look at in the process of building this relationship will say, "Oh! Mrs. Smith, you come in for your gall bladder surgery. . . . we are going to do everything we can to help you not stay in the hospital any longer than you really need to, and if all goes well, we hope to work with your doctor to get you out of the hospital in three days." Saying that in advance has a major impact on the patient's mindset when three days later they are still having some discomfort and they don't feel totally normal. The doctor says, "Well, I think we want to consider sending you home now." Without that the patient may feel shut out early if you tap him or her on the shoulder at day three. If on the other hand you have built this positive

association between "doing well" and leaving in three days, patients are much more ready and actually happy to leave because it means they have done well.

These early alliances also serve as a basis for educating patients in the disadvantages of being in a hospital when they do not need to be. The good UR programs typically will share with patients some of the risk factors associated with being in a hospital unnecessarily and essentially pledge their help to patients in protecting them from that danger and making sure they do not stay in the hospital one day later than they have to.

The third overall judgment that we came to over the last three years is that the performance differences between UR systems that are currently being implemented in American health care are major. In some way I guess I have foreshadowed this by telling you that we see some UR systems with error rates down in the 10% range and some that are approaching 50%, which I would consider tantamount to no review. At a 50% error rate I think we are talking about URs going through the motion of putting up a process that is not influencing the course of care.

URs are obviously a complicated subject and if you try to pinpoint the key factors that affect whether or not you have a very successful or a mediocre program, its obviously going to be a long list and there are going to be a lot of inter-relationships between the variables.

Let me give you what I think are six of the most important factors in whether or not you get high impact results or mediocre results from a UR Program. The first factor is reviewer knowledge. The idea in this variable is very simple. American medicine in 1987 is just too complicated and too differentiated to expect one or two primary review physicians to be able to stand toe to toe and be critical judges of what they're being told over the phone by an attending physician.

I will give you an extreme example. When we go on-site to these review systems, we learn things that would never be learned on paper. We recently went to a UR program where we asked to see their board of physician reviewers and they gave us a list of 60 physicians in every specialty cover. When we went on-site we asked the nurses what they would do when they have a case where

they cannot agree with the doctor. Then when you go to talk to Dr. Smith, whom they say they refer these cases to, you find they virtually refer no cases out for independent review by specialists because it is "too expensive." There we have one or two doctors reviewing 99.9% of the cases and you run into trouble.

The bottom line is, the better the physician-matching the better the reviewer knowledge, and the better the decisions, then the more influential and credible the review physicians are with the attending physicians.

The second variable is called "reviewer courage" and of all the variables this is probably the most underappreciated in utilization review system performance. From a very boiled-down point of view, utilization review can be viewed as essentially telling a doctor that he is not doing it the right way. Most American physicians do not react well to that message. Often if there is not overt hostility there is an undertone of threat and anger and frustration. The personality variables of the review staff for physicians and coordinators make a big difference in whether or not that overt or threatened hostility does or does not squash the review process. Some UR program managers are very alert to how important this variable is and they will make it an explicit factor in who they select for their review staff and for the physician staff. They will select people who have that relatively unique ability to stay tied to the mask in a confrontation and above and beyond that to actually be persuasive with the attending physician. It is a very small subset in the American population that has that skill. There are people out there with that gift, and the review managers who select people on this basis end up with much better results at the end of the year than people who are blind to it.

The third variable is integrity of UR process, and this is simply a fancy way of saying any UR system has a number of essential loopholes built into it. How good have the program designers and the program implementers been at narrowing those loopholes? What is a good example? Often at a cocktail party, physicians, after a few drinks, will tell how you can just exaggerate or in some cases outright lie about the nature of the patient's symptoms, particularly over the phone, to these 800 number UR systems and "get the reviewers off your back." Well, I agree that is a vulnerability in any 800-based UR system but I have been on-site and listened to some very clever counter strategies for that. I went to

see a UR system that very subtly and diplomatically at the beginning of the UR discussion with the physician said one of the rules for this UR certification program is that all of the approvals are conditional and that there is a routine process in selected cases of getting copies of the medical record for verification. The ultimate UR decision depends on what is in the medical record.

Now, that is not going to stop every exaggeration but that is going to seriously inhibit it. I think that is a very thoughtful and clever solution to a vulnerability loophole in utilization review. There are probably half a dozen such vulnerabilities.

Next let's look at weak UR standards. I will just briefly say that if UR is basically the process of screening care against certain standards, clearly the more stringent standards, the more often care is going to be questioned, the more often will be the opportunity to influence the course of care. On the converse, if standards are relatively weak, care will be relatively less often challenged and the course of care will be relatively less often effected. If you have a chance to look at our February article in *Business and Health*, I picked four or five sample UR standards and showed how different levels of stringency could have a major difference on how many days were approved.

The fifth variable is physician hostility to UR. You might say, "Well he already said that physicians don't like it, how can that vary among UR programs?" The answer is that some UR program designers appreciate the fact that physician hostility is latent in every UR interaction and do everything they can to minimize that hostility. Others are, again, relatively oblivious to it and I will give you an example. One of the things I like to do when I go on-site is listen in on the dialogue between the review physician, the review nurses and the attending physicians. I heard the following interaction: The review coordinator called the physician on the phone because based on what the doctor's office nurse had described, she felt the case could not be approved. She introduced herself and said she was representing the XYZ UR System. At this point I would say that the physician was only mildly annoyed; he was not thrilled to be on the phone with a review coordinator but he at this point was, within reason with his reaction. She went on to say that before discussing the case there were a few minor points of information she wanted to clear up and proceeded to get the physician to verify the correct spelling of the patient's name, employee ID

number and a few other minor variables. About 90 seconds into the conversation after this part of the discussion she no longer had an audience. She was no longer dealing with a mildly irritated person, she was dealing with a mad dog. Any chance of influencing the course of care, I would say, was gone.

Make sure your underlying plan design supports a high performance program. A good example is psychiatry cases. Does your plan cover psychiatric halfway houses and/or intensive outpatient treatment programs?

Let me close by making what I think is the most important general point for you to take away from this. The difference between mediocre utilization review and excellent utilization review will translate into much greater than a 100% difference in net yield from the UR program. Very major economic implications are associated with being able to put up a high performance UR program. To illustrate this we did a cost benefit study for one of our clients. There are two things that are particularly beautiful about evaluating factors on paper and on-site. They did a very careful search of UR companies and came up with two companies they thought were terrific. They had trouble making up their minds between the two companies. Because they were not only a carrier, but because they were a worker's compensation carrier and they weren't relying on patients to notify the UR system of the admission, they had the ability to randomly assign different claims offices to each of the two UR systems. They were able to pick two systems and run a horse race. They told the vendors up front, you both look good on paper, your prices are equivalent, we are going to have a six month horse race. We are going to randomly assign claim offices to you because we are a worker's compensation carrier. You can't argue that major differences in case mix affected the result, because in workers compensation the reasons for admission to the hospital are relatively narrow.

Basically 50% are medical backs or surgical backs. The range of cause codes for the rest of the admissions is very narrow so it is a beautiful natural experiment. Three months into the horse race, after each system had reviewed just over 100 hospital patients, they had us do a cost benefits analysis. Let me tell you what that showed halfway into the race at the half mile point. Two systems, UR System A and UR System B, both looked excellent on paper. Whether you evaluate from a net saving per case point of view or return investment, there was a major difference between these two vendors. UR System B was coming out

at about triple the yield in the savings as UR System A. It doesn't make sense for me to go into it, but I will tell you that our cost benefit calculations are as conservative as you can get. We subtract out everything you can imagine including the cost of alternative care like outpatient surgery and all health care, a gigantic difference in end of the year performance. In fact, after seeing this cost benefit analysis halfway into the race the client cancelled the race. They were not willing to pay the economic consequences of running this experiment for the last half of the race and switched to UR System B, much to their economic advantage.

So, let me close by saying to you, in your advice-giving and in your program management activities, make a lot of decisions. I think that effort put into assuring high performance utilization review instead of middle-of-the-road or mediocre utilization review, really can have very major financial implications for your company's performance.

MS. MARSHA LADENBURGER: My job is two-fold. Number one is to talk about the impact of utilization review on quality and to talk about the effects of utilization review activities on hospitals, providers and patients. Coming from a provider perspective all my professional life, I have worked in hospitals or for the American Hospital Association and now for the largest non-profit health care system in the country, and I couldn't agree more with a lot of what Arnie said. There is a lot of, and has been a lot of unnecessary acute care hospitalization in this country and we are here to talk about that today.

I want to talk about it from the patient provider perspective. St. Thomas Hospital in Nashville is the Daughters of Charity Hospital. Those patients in St. Thomas Hospital are also your insureds and employees. We are not talking about two different things. We are talking about our patients or your employees.

I want to put a few things in perspective. I think it is important for us to have a historical perspective on health care in the United States, if we talk about the effects of utilization review on patient care. I also think as a health care provider that we need to be very cognizant of what our patients are seeing in the newspapers and in the lay magazines, not necessarily in the business journals or in the health care journals. A cartoon appeared in the *Chicago Tribune* last year. It is a man bringing in a lottery ticket saying to his wife,

"We won the lottery, now we can afford normal health care." The word that I zero in on there is "normal." Why is that so funny to people? Why is that so poignant? Because we have seen a lot of changes in the financing and delivery in the American health care system. I am going to take myself as an example. I was born in the late 1940s in northern Kentucky. I was born in a non-profit hospital in a small town that has never done any advertising. My mother was the last person to know the sex of the baby because she was knocked out under general anesthesia. My father had Blue Cross/Blue Shield through his employer and everything was paid for. I grew up with a family doctor, getting care in an acute care hospital. We did not know about doing things on an outpatient basis. When I went to nursing school, I was taught that in essence the patient is not responsible for his or her own health. As a student nurse I was taught that if I took a patient's blood pressure and the patient asked his or her blood pressure I was never allowed to tell them that. I was to play psychiatrist.

What is the matter? Are you worried? I gave medication to a patient and the patient would say, "What medication is that?" And I would say, "The doctor ordered it for you." They would say, "Well what medication is it?" I would say," I don't know, you will have to ask your doctor." We weren't taught that patients had any responsibility for their bodies. "Outpatient" when I grew up, and when I went to nursing school at Good Samaritan Hospital in Cincinnati, meant clinic; meant poor. You did not have money. The Visiting Nurse Association was for people who did not have money. It was a social issue, it was a poor issue, it was unprofessional for me as a nurse to even know if the patient had insurance. I was not to worry about the financial status of the patient.

I was a charge nurse in an operating recovery room in Cincinnati, and we would do the supplies and there would be a million units of penicillin missing in the morning and you would say, "Oh God, somebody didn't charge for it." We were a trauma center with a lot of surgery at night, so you would go to a patient's chart. You would find somebody in Blue Cross and you would take their address and you would charge the penicillin against them. I will never forget in the early 1970s Blue Cross in Cincinnati tried an experiment with Good Samaritan Hospital. They were going to have outpatient surgery. This was unheard of and they said it was medically safe to have certain things done on an outpatient basis. So we had this experiment going on. How many times did the patients come in with their suitcases packed; we couldn't possibly have this done on an

outpatient basis. The patients weren't attuned to this but we were beginning to try it.

So what am I saying? I am saying that in the past the payment system did not support medically necessary and appropriate care being delivered in various settings. There are many things in the past that could have been done on an outpatient basis, but it wasn't paid for. How many times do you remember a doctor saying you need to have an Upper GI and a Lower GI and a Gall Bladder. He says you are going to have it done on an outpatient basis, and you say, my insurance won't pay for it. He said, fine, I will put you in the hospital for three days. Who forced that? What I am saying is this is not one of our problems. This is a problem of the providers, the physicians, the hospitals and the health care delivery settings, the purchasers and the insured. We got to go into this thing together so that the payment didn't necessarily support medically necessary and appropriate care being delivered in various settings. What had that set up? It has set up patient expectations. These expectations deal with the issue of quality.

In the last 20 years we have had many advancements in technology. When I was first out of nursing school, if a patient was on hyperalimentation he or she was probably in the intensive care unit. Now they are probably at home. We have had many technological advances and we have also had, as we know, the rapid increase in the amount of dollars being spent for health care. After 1983 prospective pricing from Medicare answered the major purchaser's concern about cost, and we are now in an era where we are talking about what I call the cost quality interface. We are talking about the confusion that arises. What is quality and what are the resources necessary to deliver quality patient care? I want to talk about two issues. The first issue I want to talk about is quality. Then I want to talk about the effect of these utilization review activities on providers and on patients. Quality is an illusive thing. You can read all kinds of quotes on what it is and isn't. It is trying to put your arms around a cloud.

I think we can break down some of the components. I say that the components to quality are two. There is the subjective aspect of quality and there is the objective aspect of quality. The subjective aspect of quality is how the patient, your employee, perceives quality. Example: You can have a kid who was run over on a motorcycle on the expressway up here in Nashville. He has two

broken legs, bad injuries, a broken arm, ruptured spleen; he is a mess. He is taken into the Emergency Room at St. Thomas Hospital for immediate surgery and he is in the hospital for two months. Three times the orthopedic surgeon thinks he is going to have to amputate this kid's leg because he has got a bad infection. Finally, two months later, that kid walks out of the hospital -- two legs, two arms, his gut all together. The orthopedics are saying, "Wow! We thought we were going to have to amputate that kid's leg." You can bring in any orthopedic, any private review group, any physician, and they will say the care in this case was totally appropriate. Then you would say to the kid, "Did you receive quality care at St. Thomas?" He can answer many ways including, "No. They were real rude to me, they were constantly nagging me do this -- you got physical therapy . . . it hurts." Subjectively, that kid perceived that he didn't receive quality care. The same thing with people who used to have it done on an inpatient basis who have it done on an outpatient basis now, but it is required to be done on an outpatient basis. It may be medically safe and appropriate; but they are not perceiving quality. One big part of quality is the perception factor and we know that health is more than just physiological. It is also psychological and emotional so these are important factors to your employee and my patient.

The second aspect of care is the objective assessment of the quality and appropriateness of care delivered to the patient. We have a big problem in that area. We have many problems in that area. One big problem we have is lack of consensus among the medical community on what are appropriate outcomes. What structure and process leads to appropriate outcomes?

Major purchasers -- the Government and now the hospital industry, the Joint Commission on Accreditation of Hospitals, are focusing more on looking at the issue of outcome. That is the bottom line. In addition, what we are seeing and what the hospital and the medical community are acknowledging, is people want to know more about outcome. We want to know what should be done to lead to what outcome. There was a full page ad in *Time* magazine -- the last week of April 1983. That was the month that Reagan signed the Social Security Act Amendment which put the Diagnostic Related Group (DRG) prospective pricing system into effect. It is interesting -- if you read it line by line, there are several places where it is referred to as the "so be it security act," and there are some physicians who feel that this is a communist plot.

The advertisement that appeared in *Time* magazine, was from Health Insurance Association of America. It was a patient's bill, with part of somebody else's bill bandaided to it, and it says "Cost shifting adds part of somebody else's bill to yours and we have a plan to remove it." The plan from HIAA was the Prospective Pricing System. Now, we all know what has happened since 1983, where major purchasers have gone in terms of really getting a handle on the dollars being spent for health care. An advertisement appeared two years later from the same organization, one of a series of opinions on health care presented by the Health Insurance Association of America. It is the quote from Dr. Philip Kaper. Dr. Kaper and Dr. Wenberg have done a lot of work in the area of geographic differences in care. Dr. Kaper makes the statement here, that you can cut healthcare costs without cutting the quality. We can do less care without impacting on the bottom line -- the outcome. Dr. Kaper says here that optimism derives from research on variations between communities in the number and type of medical procedures performed. In Maine for example, the chances of a child being hospitalized with a diagnosis of pneumonia is 20 times higher in one community than in another. Most hospital admission shows similar variations. These variations are due more to differences among physicians beliefs about the need for and choice of treatment than the differences in patients' illnesses. Remember what Mike said -- the bottom line is that a patient is still admitted to a hospital by a physician, treated by a physician and discharged by that physician.

Physician practice patterns vary so much because good information about the outcome of many clinical procedures does not exist, so that we have got to focus more. We the delivery system need a medical community, and you are seeing it in the joint commission. We all have got to focus more on outcome. What structure and process leads to what outcome? The second thing I want to talk about is the effects of utilization review activities on patients and not hospitals. Believe it or not, hospitals have been doing a utilization review for a long time -- internal. So we believe that it is imperative for us to review the quality and appropriateness of care delivered in the hospital. The Joint Commission on accreditation of hospitals, which is the voluntary not for profit accreditation process that most hospitals opt to subscribe to, has a standard for utilization review. We talk about not just utilization review in hospitals but utilization management. Up until 1983 and the changes in the payment system, a lot of the UR in hospitals was that. We reviewed, but we didn't necessarily manage the

utilization. And there are a lot of changes that have taken place because of the malpractice situation in this country. We know that you have applied for some chance of suffering injury if you are hospitalized. That doesn't mean anybody did anything wrong. The California Medical Insurance Feasibility Study said that probably 0.8% of these injuries were caused by negligence. But the fact of life is -- it is a risk to be in a hospital and if we don't need to have you there, we're better off not to have you there and that is why we are diversifying. It's why I work for the Daughters of Charity National Health System. We don't just run hospitals; we want a magnitude of healthcare delivery systems.

I want to talk about the impact of external utilization review, which includes pre-admission certification, second surgical opinion, continued stay review, and discharge planning on patients. So with all this going on, the patient still says, when I get sick I want the best. When patients get sick, they don't care what number they are supposed to call unless it is been ingrained in them. So from the patient perspective, I would agree with what Arnold said. The better you inform the patients up front, the better it is for you and the better it is for us. You don't tell people about their insurance plan when they are in a stressful situation. It undermines the confidence that they have in the doctor and the hospital.

Do you know I have hospitals in my system that deal with 50 external private review groups? They have written computer packages to keep these things straight. It is a very difficult situation for many hospitals to deal with. When a patient comes in, he or she says "Oh, you have not heard of me, you don't know my payor." They are in a stressful situation, and it maybe undermines the confidence the patient may have in the best hospital in the country. Please let's help inform your employees, our patients, before they become our patients.

I'd like to make a couple of comments on private utilization review, and the impact on the hospital.

First, the varying requirements from the many different private review groups drives us crazy. I know that a number of years ago, American Medical Peer Review Association came to the AHA, and said, "What won't the AHA support?" That all private review activities should also go through the State Pro. Because that way a hospital would deal with the Iowa Pro for every kind of review,

Medicare review, all the private reviews, and Medicaid review for the state. In theory it makes sense, but in practice it doesn't work, and it's not ever going to work. Yes, it would be better if we could deal with one group. The second thing that has not been talked about yet is that you've got to realize that we have a right and a responsibility to protect the patient's privacy, in terms of the proper authorization you need, the private review groups need, to gain access to the records. We, the hospitals, have to protect that patient's records. And in every state there are different laws, privacy and confidentiality laws, and particularly in the area of psych, alcohol and drug abuse, AIDS patients, and sexually transmitted diseases, so we have a responsibility to protect the privacy of the patients.

The next point is, we as hospitals and we as physicians are concerned about what standards are being used to make decisions about the review process. Who are the physicians? Who is responsible for developing these standards? Who is the person on the end of the phone? You've got an orthopedic surgeon and there is a tricky question about continued stay review; if there is a pediatrician on the other end of the line, it goes down the drain.

Remember, the fact of life is it's one doctor admitting one patient to one hospital, and that patient is your employee. What we also have to remember is that medicine is an art and a science. It is that art of medicine that is goes down the drain.

One other point I want to make is that you have to remember that we as hospitals and physicians are practicing in a medical malpractice environment that has gone crazy, particularly in the area of OB-GYN. We have a lot of defensive medicine going on in this country, but one of the reasons why a lot of law suits are filed is failure to diagnose. Picture yourself as a hospital administrator, dealing with a lot of private review groups for over-utilization and you got a big admitter who is a talented young physician who gets slapped with his first medical malpractice suit. It is going to make him lose sleep for the next four years of his life and the reason why that suit was filed was failure to diagnose. We have got to recognize the medical malpractice environment in this country impacts every doctor and every physician every day of their life.

In closing I want to stress that we, all of us together, have got to put some better research activities in the area of looking at outcomes. What structure in process looks to what and achieves what outcome. I want to leave you with one of my favorite quotes. That hospitals, if they wish to be sure of improvement, must find out what the results are, must analyze their results to find their strong and weak points, must compare their results with those of other hospitals.

If you were a group of hospitals, they all compare results. Maybe we need to do it a little bit better. Hospitals must care for what cases they can care for well and avoid attempting to care for cases which they are not qualified to care for well; must assign the cases to members of the staff for treatment, for better reasons than seniority, the calendar or temporary convenience. It is very difficult to do in the small and rural hospital. Hospitals must welcome publicity not only for their successes but for their errors so that the public may give them their help when it is needed. I haven't figured out how to do that one yet. Hospitals must promote members on the staff on a basis which gives due consideration to what they can and do accomplish for their patients. We ought to call the patients a year later to see how well they are doing. There is an interesting letter in the *New England Journal of Medicine*, January 15, 1987. It is by a Dr. Steven Schrader and Dr. Schrader says, The Joint Commission's signal to explore outcome assessment is good news. At very least it will focus attention on the most important aspect of hospital care -- the state of the patient. Perhaps the first place to start might be with frequently performed elective procedures for which we know that unacceptably high complication rates exist. Outcomes are too important to ignore any further. The threat of harm to patients resulting from recent cost containment efforts makes the new Joint Commission directive especially welcome. I thank you for the opportunity to talk with you about the patient and the provider perspective.