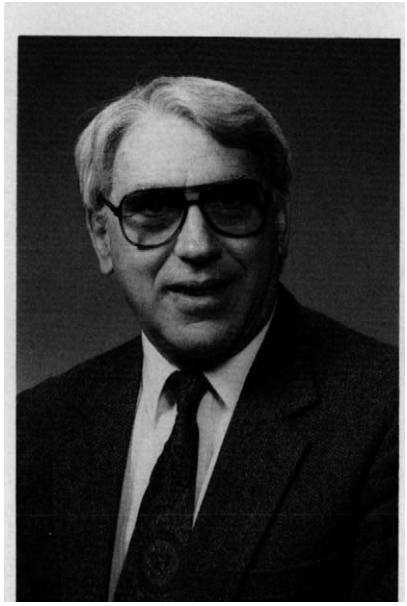
0363-5465/94/2206-0734$02.00/0
THE AMERICAN JOURNAL OF SPORTS MEDICINE, Vol. 22, No. 6



Presidential Address of the American Orthopaedic Society for Sports Medicine

Outcomes Research in the AOSSM*

Robert J. Johnson,† MD, *President, AOSSM*

The definition of outcome in Webster's dictionary is "the way something turns out; result; consequence or an effect." Outcomes research is not markedly different from clinical research as we know it. It is refined and enhanced clinical research.[4] To determine whether surgical or nonsurgical treatment has benefitted the patient, the outcome of the treatment must be assessed in an appropriate way.[6,10] The goal of outcomes research is to 1) understand use rates and the real outcomes, 2) develop accurate and comparable data bases, and 3) learn what works for individual patients. Outcomes research allows assessment of the quality and appropriateness of medical care[5]; it emphasizes patient-oriented outcomes of care rather than assessments of the process of care.[4] Process measures include items that we are used to observing and considering in the determination of the effectiveness of orthopaedic care. Examples of these process or objective measures include radiographic appearance, range of motion, laboratory results, and, in the case of knee ligament surgery, measured laxity. Patient-based outcomes measure the results of care as they are perceived by the patient.[4]

One might ask, why measure patient-oriented outcomes? Patients do not really care if there is a pivot glide or a radiographic change within their knees after treatment of an ACL disruption. They want a functionally stable, painless knee that can tolerate sports and work activities. The degree of satisfaction and the quality of life is more important to them than whether or not there is a 1° or 2° loss of range of motion, a 2- to 3-mm side-to-side difference in measured anterior-posterior laxity, or radiographic findings indicating minor degenerative changes.

"Quality of life" is the term used to establish health status. It embodies physical, social, and emotional function.[6] The quality of life is assessed by questionnaires, which are known as instruments. The quality of life assessment has been relatively rarely used in the evaluation of orthopaedic surgical procedures and, thus, is rarely encountered in our literature.[10] Perhaps the major reason for this omission has been the assumption that such quality of life information was "soft data," whereas the "objective data" that we measure as a result of surgical procedures have been considered "hard data" and are more reliable in determining the efficacy of a certain procedure. However, when we closely examine what we have considered to be objective data, there is a good deal of evidence that this information is actually softer than we would like to think.[10] For instance, our ability to measure the amount of joint laxity after an ACL reconstruction may actually be less well defined than we have been led to believe.[1] Thus, to fully appreciate and incorporate outcomes research into our literature, it is going to be necessary for orthopaedic surgeons to rethink their knowledge base and develop new information.

Relman[7] has brought to our attention that a revolution

*Presented at the 20th annual meeting of the AOSSM, Palm Desert, California, June 1994

†Address correspondence and reprint requests to: Robert J. Johnson, MD, University of Vermont, Department of Orthopaedics and Rehabilitation, Stafford Hall, Room 426A, Burlington, VT 05405-0084.

in medical care has been occurring since the 1940s. He divided the revolution into three eras. The first era, termed "the era of expansion," occurred from the 1940s through the 1960s. During that time, the number of hospitals and physicians (especially specialists), the amount of technology, and the extension of insurance coverage increased at a rapid rate. The second era, which occurred through the 1970s and 1980s, was that of cost containment. That era could be described as a revolt of the payers because of explosive, inflationary, per capita medical costs. The third era of the revolution in medical care could be termed the era of assessment and accountability. This era began in the late 1980s and is still going on. In this final era, it has become apparent that we need to know the relative cost of what we do as well as its safety and effectiveness. Armed with these facts, physicians will be in a much stronger position to advise their patients and to determine the use of medical resources, payers will be better able to decide what to pay for, and the public will have a better understanding of what is available and what they realistically want. Outcomes research can help us determine whether the present assumptions about patient demands for expensive medical care are really what we assume that they have been.[12] We do not know what the level of demand for services would be if patients were fully and realistically informed of their options. Current rates of use of invasive, high-technology medicine could be higher than what the public really wants. Certainly, patients are more adverse to accepting any risk from a new "high-tech" surgical procedure than are their physicians.

Outcomes research—its design, implementation, and dissemination—has resulted in an entirely new industry within the field of health care.[8] If we, as sports medicine orthopaedic surgeons, do not use proper outcomes research and contribute to the inevitable reforms that are coming in medicine, we will lose credibility and the ability to play a role in the issues at stake. Bernard Rineberg[8] has recently stated that orthopaedic surgery has fallen far behind in the field of outcomes research and has not taken it as seriously as some other fields of medicine.

There are three major reasons why outcomes research is needed: 1) rapidly increasing health care costs, 2) variations in the use of health care services, and 3) deficiencies in the research literature.

The percentage of our gross national product dedicated to health care has risen from 5.2% in 1960 to 14.4% in 1992.[4] This percentage is by far the highest in the world. In 1990, the percentage of gross domestic product dedicated to health care was 12.2% in the United States, 8.5% in Canada, 8.0% in Germany, and 6.3% in Japan. In spite of these tremendously high costs, our life expectancy and infant mortality rates were relatively poor compared with other developed nations. Our extra expenditures produced no obvious benefits in these areas.[4] Our increased expenditures in the field of orthopaedics may result in a higher quality of life and more effective care, but we simply cannot prove it with the information available in our present literature.[4]

Wide variations in the use of health care services occur from one area of our country to another.[4] In some areas it would appear that there is extremely high use of a particular form of orthopaedic treatment, while in others it is relatively infrequently used. There is no assurance from what we know now that the greater use rate is producing any better health care than the lower one. Obviously, both of these use rates cannot be correct. Outcomes research offers us the hope of determining the appropriate rate of use.

Deficiencies in our clinical literature result from a lack of randomized clinical trials, standard definitions, adequate outcome measures, inadequate study design, poor description of patient populations, and no evaluation of patient-oriented outcomes of care. These problems have led to rather wide disparities in the perceptions of medical care as it is now seen by various groups. For instance, health care administrators are most concerned about cost effectiveness, patients care most about the risk versus the benefits of their care, whereas clinicians have been concerned about the selection of patients who will do well and also about assessing refinements that might improve their clinical results. Some members of the health care community think that we have reached the point where health care must be rationed.[12] Patient demands are exceeding our capacity as a society to pay for all possible health care that is now available. The policy makers, patients, and payers demand that we show them that the highly variable, very expensive and complicated procedures that we now use are cost-effective.[4] Because of deficiencies in our research, we really do not have the ability to answer these challenges. If we do not devise effective answers to these questions, the policy makers, patients, and payers will simply answer us by refusing to pay for our services in the future.[4]

To understand appropriate research techniques that can lead to the answers we need to attain, we must understand the appropriate design of research studies. In an attempt to do this, I will provide a brief outline of the basics of experimental study design. The first such design is based on what are defined as observational studies. These studies observe events rather than alter them. Observational studies include case studies, case control studies, and cross-sectional studies. An example of a case study is a followup of 50 patients with a particular ACL reconstruction. From this, we get information about how these patients fare, but we cannot learn anything about which method of treatment of an ACL injury is best. This is because we only observe the results of patients who undergo a single procedure. Case studies are very frequent in the orthopaedic literature; they provide us with no control group and they result in conclusions that are difficult to interpret in terms of the relative success of the treatment rendered.

With case control studies, we are still in the category of observational research. These studies by definition are retrospective. In these studies, patients with an outcome of interest and a control group without the outcome are followed backward in time to ascertain if some earlier treatment or exposure is related to the outcome. These studies are also common in our literature, but provide biased information because one never knows how patients would have fared if they had been treated differently. Thus, the

observational studies that are so frequent in our literature are really inappropriate and inadequate to determine which type of care is best and which type is most cost-effective.

Experimental studies are potentially much stronger but are unfortunately quite rare in the orthopaedic literature. In an experimental study, a perturbation is administered to one group of individuals and the results of this treatment are compared prospectively with those of a second group who did not receive the same perturbation. This is the classic prospective randomized clinical trial. Unfortunately, this type of study is rare in the orthopaedic literature, primarily because it is difficult and expensive to perform. The randomized clinical trial is not only prospective but is also concurrently controlled, has a specified protocol for inclusion and exclusion, randomly allocates patients into two different treatment (perturbation) groups, has specific measures of the condition of the patient at both the initial state and at the time of the eventual outcome, and is the best means of avoiding susceptibility bias (Fig. 1).[9]

Biases are defined as any systematic error arising from the design or conduct of a study. There are four major biases: susceptibility, performance, detection, and transfer. Susceptibility bias is a dissimilarity in terms of the initial state. Thus, patients who have different types of injuries would not be included in a study of ACL reconstruction. Performance biases imply a dissimilarity in the treatment of the two groups studied. Detection bias means that there is dissimilarity of the measure of the initial state and the outcome of the treatment. Transfer bias implies a loss of patients at the time of followup in the two study groups. It cannot be assumed that those patients who are lost to followup have good results. Because so many of our studies have not controlled or appropriately eliminated biases appropriately, Keller et al.[5] think that, in general, the research in our literature is flawed and, if this is the case, so is the practice of our specialty.

It is hoped that the development of appropriate outcomes research methods will enable us to know the quality of our research and to understand the appropriate cost-effectiveness of treatments. Outcomes research is not a simple, single approach. It involves at least six different means of evaluating the information we are attempting to discern.[4] The outcomes methods include 1) analysis of large data bases, 2) small-area variation analysis, 3) structured literature reviews (meta analysis), 4) prospective randomized clinical trials, 5) decision analysis, and 6) guideline development.[4]

Large data base analysis is of limited value for it evaluates such presently available imperfect information as Medicaid files. However, an analysis of these files from an epidemiologic standpoint can provide us with some limited information concerning mortality, length of stay in the hospital after certain treatments, complications, and the incidence of reoperations.

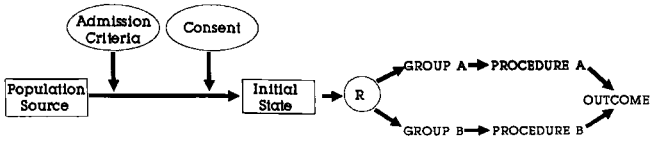


Figure 1. A randomized clinical trial.

Small-area variation analysis has demonstrated marked differences in the surgical rates per capita in various areas of the country. This type of information implies that both the high and low rates of use cannot be correct. However, we do not know the proper use rate of such procedures as disc surgery. Keller[4] thinks that outcomes research provides hope for us to establish the answer to what rate of surgery is appropriate. The fact that the use of many kinds of medical care are strikingly different between areas of the country implies that there is a difference in the beliefs among physicians about the best treatment methods. The reasons for these differences hopefully revolve around the uncertainty about what method is best rather than the financial incentives.[5]

Meta analysis demonstrates flaws in the literature. Data from many articles are pooled to form a large data base for statistical analysis. This information can give us some indication of the present status of what we know. These studies, when done appropriately, can cost between $30,000 and $50,000 but can be the first step to meaningfully determine our present state of knowledge and to identify the type of research needed to expand this knowledge.

The randomized clinical trial is perhaps the area of outcomes studies about which we are most familiar. A brief summary of outcomes studies has been presented earlier in this discussion.

Decision analysis implies that successful clinical research with meaningful outcomes has been completed. The results of this information are translated into a series of probabilities that are placed into an algorithm decision tree. This type of information can be very helpful in appropriately guiding patients toward effective proven care once these devices are established.

Valid and useful information that is necessary to develop clinical guidelines is not currently available. Almost all subjects in orthopaedics and sports medical care require improved outcomes research if we are to reach meaningful and appropriate clinical guidelines.

Patient outcomes research teams (PORT) have been developed to use all, or nearly all, of the six methods described.[4] These extremely extensive, complex means of attaining information have, to this point, been applied under the direction of the Agency for Health Care Policy and Research to only four musculoskeletal fields[8]: low back pain, arthritis of the hip, arthritis of the knee, and fractures of the hip. It is inappropriate to assume that all subjects of orthopaedic and sports medicine research could be handled by the PORT method. It is simply too cumbersome and too expensive.

In an attempt to understand outcomes research and the ways that our Society should become involved with it, an Outcomes Committee has been established and will be chaired by Jesse DeLee. There can be no doubt that orthopaedic leadership in outcomes research is a must. As

Bernard Rineberg[8] noted, "the orthopedist is more knowledgeable than federal policy makers about the clinical and personal concerns of orthopaedic patients, is best suited to design and perform the research, and is a more credible source of information than another political or scientific entity. Our national organizations are the best means to circulate the results of this information and payment policies will be based on our ability to demonstrate effectiveness of what we do." Dr. Rineberg thinks that the orthopaedic effort in outcomes research includes the delineation of priorities for topics to be studied, the development of a coalition of organizations (i.e., third-party payers, orthopaedic industries, patients, and others), collaboration with the research community, and provision of a primary role in analysis, interpretation, and dissemination of the results.

The role of the AOSSM Committee on Outcomes Research is to educate the Society to the need for this research and its methods, provide a resource center concerning this matter for the Society, develop standardized language and definitions, and work with the AAOS Committee on Outcomes Research on instrument development. We must understand that we cannot do this alone, the cost is too high, but we must participate. At the Academy level and at the level of our Committee on Outcomes Research, it has been decided that the most effective way we can contribute to the outcomes research effort is to develop outcomes assessment instruments that are of high quality, standardized, and broadly accepted. These instruments will be useful to our members not only in formal research but in their own practices. An outcomes assessment instrument is a questionnaire that can be used by any orthopaedic practice to determine the results of treatment. The goal of this type of device is to allow the individual or research group to determine how the techniques of a particular treatment measures up. The use of such an instrument would, in the future, improve the quality of all studies and make possible meaningful comparisons of various treatments that do not exist today. At present, the ideal outcome measurement for topics such as ACL reconstruction does not exist. There are no universally accepted standardized definitions and measures and, thus, the present literature is very confusing because the results of various studies cannot be compared.

Outcomes instruments can be divided into narrow- or comprehensive-focus studies. The narrow-focus instruments are specific while the comprehensive-focus instruments are generic, with a much wider spectrum. The goals of outcomes instruments can be designed for research where detailed and complex information is necessary or designed for a clinical practice setting where the goals would be quick and simple to use in the practice environment.

Outcomes instrument design can be divided into three categories: discriminitive, predictive, and evaluative.[3] The discriminitive instrument defines cross-sectional differences between groups of individuals. The predictive instrument predicts a concurrent or future gold standard of treatment or care. The evaluative instrument measures longitudinal changes over time after a treatment technique. Thus, an evaluative instrument design needs to measure the condition of an individual before the treatment and then to follow that individual to see how the treatment affects the results over a period of time. By far, the most common studies about which we are concerned in orthopaedic sports medicine would require an instrument design of the evaluative type.

The evaluative instruments must be reproducible, valid, and responsive.[3] Reproducibility means that the information attained would result in exactly the same outcome if the patient population being studied remained stable as time passes. Repeated application of the instrument would yield the same results in this case. To prove that the outcome instrument has this quality of reproducibility (repeatability), it must be tested. Thus, the same response should occur if it is administered at different times unless the patient's condition has actually changed between the two measurements.

Validity is a means of assuring the examiner that the instrument is actually measuring what it is intended to measure. Although this seems to be a rather simple concept, it is apparent from reviewing the literature that the means of establishing the validity of an outcomes instrument are far from secure. The Medical Outcome Trust Scientific Advisory Committee, which includes experts in the technical aspects of psychometrics and health status assessments instruments, is having difficulty in establishing exactly how to prove that an instrument is valid.[11] The ideal method of documentation of validity is the demonstration that the instrument results match that of some gold standard of outcomes research.[2] Since such gold standards rarely exist because there really is no standard for the quality of life, there are still arguments about the best means of establishing the validity of an instrument.

There are three types of validity.[2] The first is face validity, which implies that the instrument contains relevant points from the patient's point of view. Content validity covers all important areas of quality of life that apply to the particular outcomes method in question as agreed to by a panel of experts. The third type of validity is construct validity, which attempts to answer the question, Does the questionnaire behave in relationship to other outcome measures as one would expect if it really measures the quality of life?

The responsiveness of an outcomes instrument is whether or not the device is capable of detecting clinically important changes even if they are small.[3] This would imply that a responsive instrument is able to ascertain changes as time passes. It should be understood that a reproducible and responsive instrument may be valid for one purpose but it is not necessarily valid for another. Thus, the use of an instrument that has been standardized and validated for one purpose cannot be readily changed to another evaluation without extensive testing and proof that it will work.

The development of an outcomes instrument, which is presently being undertaken by members of the AOSSM Committee on Outcomes Research in cooperation with the efforts of the Academy, includes the following principles.

First, a task force is formed (members of COMSS, the Academy's Committee on Outcomes Research, and the categorical specialty societies) that will determine the subject that is to be covered by the instrument (in our present case, the upper or lower extremity). Second, a review of the literature is undertaken to evaluate previously developed instruments on similar subjects. Then a list of items necessary to attain the information required is assembled. The items to be selected include questions concerning the quality of life related to the specific subject, such as symptoms and physical complaints, work-related concerns, alterations in recreational activities, sports participation, and competition, life-style changes, and the emotional aspects of the loss of the ability to participate at the level the individual desires.[7] These items are carefully assessed and reduced to the minimal number necessary to adequately evaluate the results (minimal common data set). The instrument so developed is then tested for reliability, validity, and responsiveness. After all of the above steps have been completed, the instruments will become available for use by members of our Society.

The present goal is to produce generic instruments concerning the entire upper or lower extremity. It is hoped that these instruments can be used in general for almost all orthopaedic problems relative to the extremities. Our Committee on Outcomes Research will work in cooperation with the AAOS group to develop more detailed modules that can be used to study more specific problems.

We hope that in the near future instruments will be available for those of us in our Society who wish to use such instruments in their practices. Obviously, this situation is in its developmental stage and the results will soon be presented to the Society through not only the Academy group, but also through the auspices of the Outcomes Committee of our own Society.

In conclusion, it must be understood by all of us in the AOSSM that we have an obligation to understand what outcome studies are and why they are needed. If we do not adopt appropriate outcomes methods and apply them in our own practices, we will lose our credibility and ability to play a role in the issues at stake as the changes in medical care evolve in the near future. At present, it appears that our most appropriate efforts should be in cooperation with the Academy in the development of high-quality, standardized, broadly accepted, and validated outcomes instruments that can be applied in almost all of our practices and can be adopted to very specific problems by the addition of further modules as they are developed. It will be the purpose of this effort to provide not only the research scientist with appropriate methods, but also those of us in private practice with the means of evaluating how well our own patients are doing. At the very least, the application of these methods will improve the quality of our literature and research and make possible meaningful comparisons of various treatment methods that are reported in the future. I urge all members of our Society to keep abreast of the development of outcomes research options available to us as they are developed.

It has been my pleasure and honor to serve as President of the AOSSM during the past year. It is apparent to me and other members of the Board that our Society is becoming much more complex as time passes. More and more of you will be assisting in the many facets of our Society so that we can continue to improve its value to our members and to society in general. I would like to thank all of the members of the staff in Chicago, including Don Rome, Carol Rosegay, Camille Petrick, Pam Schmaranzer, and Pat Kovach, for their tireless support and help during my tenure as President. I also wish to express my appreciation to my family, including Shirley and my daughters Jennifer and Tamara, for their support and understanding during this busy year.

REFERENCES

1. Fleming BC, Beynnon BD, Johnson RJ: The use of knee laxity testers for the determination of anterior-posterior stability of the knee, in Jackson DW (ed): *The Anterior Cruciate Ligament: Current and Future Concepts.* New York, Raven Press, 1993, pp 239–250
2. Guyatt GH, Bombardier C, Tugwell PX: Measuring disease—specific quality of life in clinical trials. *Can Med Assoc J 134:* 889–895, 1986
3. Guyatt GH, Deyo RA, Charlson M, et al: Responsiveness and validity in health status measurement: A clarification. *J Clin Epidemiol 42:* 403–408, 1989
4. Keller RB: Outcomes research in orthopaedics. *J Am Acad Orthop Surg 1:* 122–129, 1993
5. Keller RB, Rudicel SA, Liang MH: Outcomes research in orthopaedics. *J Bone Joint Surg 75A:* 1562–1574, 1993
6. Mohtadi NGH: Quality of life assessment as an outcome in anterior cruciate ligament reconstructive surgery, in Jackson DW, et al (eds): *The Anterior Cruciate Ligament: Current and Future Concepts.* New York, Raven Press, 1993, pp 439–440
7. Relman AS: Assessment and accountability: The third revolution in medical care (letter). *New Engl J Med 319:* 1220–1222, 1988
8. Rineberg BA: A call to leadership (editorial). *J Bone Joint Surg 72A:* 1439–1440, 1990
9. Rudicel S: How to choose a study design. *Am J Sports Med 16:* S43-S47, 1988
10. Troidl H, Kusche J, Vestweber K-H, et al: Quality of life: An important endpoint both in surgical practice and research. *J Chron Dis 40:* 523–528, 1987
11. Ware JE: Standards for evaluating the standards. *Med Outcomes Trust Bulletin 2:* 2–3, 1994
12. Wennberg JE: Outcomes research, cost containment, and the fear of health care rationing (letter). *New Engl J Med 323:* 1202–1204, 1990