CHAPTER 3

ARBITRATION OF MEDICAL AND HEALTH ISSUES

FREDRIC GERR*

Introduction

Differences of medical opinion are often at the heart of disputes requiring arbitration. Arbitrators and advocates alike need basic tools to understand and evaluate expert medical testimony. Effective cross-examination also requires a working understanding of how physicians approach issues—such as assessing the work-relatedness of disease, determining disability status, and evaluating medical tests for preplacement examination. It is axiomatic that expert medical testimony should be based on sound medical knowledge. In this paper I will discuss two broad topics that are of relevance to the arbitration process when expert medical testimony is introduced into evidence: (1) assessment of the quality of expert medical testimony, and (2) evaluation of screening tools used in preplacement evaluations. The first issue is one of long-standing importance in any setting in which important decisions are based on expert medical opinion, and the second is especially relevant in light of restrictions placed on preplacement hiring decisions by the Americans with Disabilities Act.

Assessment of the Quality of Medical Testimony

Medical Fact and Medical Opinion

An understanding of the difference between medical fact and medical opinion is necessary in assessing the quality of medical

*Assistant Professor, Emory University School of Public Health, Atlanta, Georgia.

testimony. Reviews of this topic are provided by Miller[1] and Hill and Sinicropi.[2] A brief review follows.

Medical facts are those variables or data relating to a patient which are obtained by some objective and verifiable method. Examples of medical facts include a patient's blood pressure, serum glucose level, and muscle strength. Alternatively, medical opinion is a judgment made by a physician or medical expert which requires synthesis and interpretation of the medical facts about a patient in light of the current state of relevant medical knowledge. Examples of medical opinions include assessment of disability status, determination of work-relatedness (causation) of an illness, and determination of suitability for employment on the basis of preplacement or return-to-work medical examination results.

Disputes over medical fact are often amenable to resolution by involving a third medical expert.[3] When differences of medical opinion are at the heart of a dispute, resolution is more difficult. Under these circumstances, how can both arbitrator and advocate best deal with expert medical testimony? One method involves assessment of the experts themselves, and another involves assessment of the soundness of the medical arguments put forward by the experts.

Assessment of the Experts

Often a dispute will involve the company or plant physician and a non-company physician. It has been argued that the company physician is in the better position to assess how a worker's medical condition affects employment (and vice versa).[4] After all, the company physician is most familiar and experienced with the physiological and psychological requirements and effects of the specific workplace. Alternatively, it is suggested that the outside physician, by virtue of specialty training, may be more familiar with the medical condition of the worker or have greater expertise with its treatment.

Another issue that has been discussed elsewhere is the expert's impartiality. For example, does the expert appear frequently in

---

[1]Miller, *The Use of Experts in Arbitration—I. Expert Medical Evidence: A View From the End of the Table*, in Arbitration and Social Change, Proceedings of the 22nd Annual Meeting, National Academy of Arbitrators, ed. Somers (BNA Books, 1969), 137.

[2]Hill & Sinicropi, Evidence in Arbitration, 2d ed. (BNA Books, 1980), at 166–71.

[3]Miller, *supra* note 1, at 138.

[4]*Id.* at 141.

arbitration on behalf of a certain party? It has been suggested that, if so, this fact "supports the inference that the physician is biased."[5] However, given that the company physician appears solely on behalf of a certain party, does the same reasoning not suggest potential bias in this case as well?

Although substantial disparity in the qualifications of medical experts may lend some difference in the weight given their opinions, often all experts involved are adequately credentialed. Advocates and arbitrators must have other means of evaluating more directly the quality of medical testimony. In this situation it is recommended that they adopt a method for evaluating the quality or soundness of the medical evidence. This method may also be useful in cross-examination of experts. It is a requirement that physicians rendering opinions be available for cross-examination in order to allow evaluation of the basis for their opinions.

Assessment of the Technical Quality of Medical Opinion

Given the tremendous proliferation of medical literature and the unfortunate variability in its quality, it is often possible to find published medical studies that support both sides of a particular dispute. This problem is compounded in occupational and industrial medicine by the tendency for experts to belong to either the labor or management camp. Rather than throw their hands into the air in frustration, those required to make decisions based on medical testimony or who may cross-examine medical experts should use a system for assessing the quality of medical opinion. Although nonmedical professionals may never be able to assess the quality of the medical literature to the extent that a medical scientist can, some understanding of the process by which physicians and medical scientists use the medical literature to render an expert opinion may be useful.

The assessment requires that two questions be applied to the expert opinion. First, what is the quality of the medical knowledge that supports the inference on which the opinion is based? Examples of such inferences are (1) asbestos exposure causes lung cancer, and (2) patients with unstable heart disease are at increased risk of death when performing heavy labor. Second, to what extent does that body of knowledge apply to the particular

---

[5]Hill & Sinicropi, *supra* note 2, at 176.

person or situation involved in the current dispute? Examples of this question are (1) was this worker with lung cancer actually exposed to asbestos, and (2) did this worker actually have unstable heart disease. A related medico-legal question is whether the medical evidence supports an association that meets the burden of proof required by the particular setting.

Role of the Epidemiologic Study in Determining Medical Opinion

The primary tool of medical science is the epidemiologic study. Broadly speaking, epidemiologic studies utilize either observational or experimental techniques.

Observational methods involve comparing the proportion of "exposed" individuals who have or develop a disorder to the proportion of "unexposed" individuals who have or develop a disorder. "Exposure" means employment in a particular job or significant contact with some hazard, such as noise, airborne asbestos dust fibers, or cigarette smoke. The greater the proportion of the exposed who have developed the disorder in comparison to the proportion of unexposed, the greater the risk of disease, given exposure. The results of the observational method are dependent upon accurate assessment by the investigators of both exposure and health effect, as well as assurance of the comparability or "sameness" of the exposed and unexposed groups with respect to other factors that could influence the occurrence of the disease.

Experimental methods involve deliberate allocation of "exposure" and nonexposure, typically in a random fashion, to a group of suitable subjects. This methodology is usually reserved for "exposures" that are likely to be beneficial to subjects, such as a new treatment for a disease. Studies involving experimental methods are rarely used to address questions that arise in industrial disputes.

Criteria for determining whether associations observed in epidemiologic studies are "causal" in nature have been proposed by many authors and are universally based on those proposed by the great medical epidemiologist, A.M. Hill. These criteria are useful for addressing the first of the two questions raised in the previous section, (i.e., what is the quality of the medical knowledge that supports the inference on which the opinion is based). They consist of

(1) the strength of the association,

(2) the presence of biologic gradient,
(3) consistency of results across multiple studies,
(4) biologic plausibility, and
(5) similarity to known causes.[6]

Strength of association refers to the magnitude of the increase in risk that occurs among those who are exposed. An exposure that increases the likelihood of a medical problem by tenfold is considered more important (and more likely to be causal in nature) than one that increases the likelihood by one and one-half. Biologic gradient refers to the situation in which increasing levels of exposure are associated with increasing risk of disease. This phenomenon adds weight to the belief that an association is causal in nature. Consistency across studies indicates that the association has been observed in other studies, preferably performed by different investigators using a variety of methods. Biologic plausibility refers to how well the observed association conforms to current biomedical thought. Similarity of the cause under investigation to known causes is similar to biologic plausibility. The previous criteria are used to determine whether the medical literature supports a generic association.

Regarding the second question raised above as to whether the attribution of a disease to a workplace factor is appropriate for a particular patient, a weighting scheme for certainty is proposed. The strongest evidence that a patient's condition is related to an exposure is the availability of epidemiologic studies of subjects that resemble the particular patient to the greatest extent possible and that demonstrate that the effect or disorder of interest occurs with substantially increased frequency among the exposed subjects. For example, medical investigators have found that male insulation workers exposed to asbestos for more than 20 years were at substantially higher risk of lung cancer than similar, but not asbestos exposed, workers.[7] These studies are extremely convincing when determining whether lung cancer in a male patient who happens to have been an asbestos insulation worker for more than 20 years was, in fact, caused by his occupation. To a lesser but not insubstantial degree, epidemiologic studies of subjects who do not resemble the patient fully, but who have essential features of exposure and effect in

---

[6]Kramer, Clinical Epidemiology and Biostatistics (Springer-Verlag, 1988), 264.
[7]Hammond, Selikoff, & Seidman, *Asbestos Exposure, Cigarette Smoking and Death Rates*, 330 Annals N.Y. Acad. Sci. 473 (1979).

common with the patient, can contribute substantially to the formation of an opinion. Continuing with the previous example, lung cancer in a female carpenter exposed to asbestos for 20 years would likely be attributable to work, but the association is slightly less certain than in the previous example because of differences in the characteristics of the subject (gender) and the occupation (carpenter versus insulation worker).

Epidemiologic studies in which health problems are similar or related to the condition of interest (but not identical to it) require a greater logical inference about causality and work-relatedness than directly relevant studies. For example, many studies use the presence of symptoms rather than a comprehensive diagnostic evaluation as the measure of the effect a particular exposure or job may have on health. This approach is used because symptoms are more easily, quickly, and cheaply obtained than are complete diagnostic evaluations. Such symptoms may be suggestive of disease but are not necessarily diagnostic of it. For example, studies that use the symptom of hand pain may be relevant to the evaluation of a patient with documented carpal tunnel syndrome (CTS) because hand pain is a cardinal symptom of that disorder. However, hand pain is not synonymous with CTS, and consequently such studies do not carry as much weight as those that use more formal diagnostic methods for detecting CTS among the participants.

Studies using symptoms as the measure of disease or exposure-effect are best used to corroborate the results of studies in which more definite assessment of the disease was made. For example, consider a worker employed as a data entry operator who, after many years of full-time work at a video display terminal (VDT), develops CTS. Is the disease work-related? Review of the medical literature identifies several large studies of VDT users, all of which use symptoms (i.e., hand pain, numbness, and tingling) as the sole measure of effect on health. The review also identifies two studies of CTS in assembly line workers (in which formal CTS diagnoses were made) which suggest that repetitive use of the hands significantly increases workers' risk for developing CTS. None of the studies includes subjects that match the patient at the center of the dispute. However, the existence of (1) studies that document consistent symptoms (in this example, hand pain, numbness, and tingling) in workers with the same job as the patient (data entry) and (2) studies that document the actual disease in workers who have

some critical exposure in common with the patient (repetitive hand use is the critical exposure in this case) both support the assertion that this patient's CTS is causally related to the workplace. The existence of only one of the above types of study would have been suggestive that the condition was work-related but not as strongly as both.

Case studies (i.e., reports of exposed workers with the disease or condition of interest but without comparison to an unexposed group) lend relatively little scientific support to an association. Consider the following hypothetical statement: "Three fourths of our patients with bladder cancer work in factory X, therefore working in factory X causes bladder cancer." The logic in this statement is flawed. If three fourths of all people in that town also worked in factory X, then the results do not support the conclusion. To properly address the issue of whether working in factory X increases the risk of bladder cancer, the key question that must also be asked is what fraction of individuals without bladder cancer work in factory X? Only then can the presence of a meaningful association between disease and exposure be determined. In general, case studies are suggestive but not conclusive evidence of an association.

Finally, physicians often refer to their own personal experience to support their testimony. However, the physician's personal experience is often a function of local referral patterns or medical customs or other idiosyncratic factors that may affect the opinion. As a result, the physician's personal experience is not truly objective and is subject to a number of biases that may render it untrustworthy.

Burden of Proof

When providing a medical opinion, the physician must often relate a probabilistic concept in "yes-no" terms of causality. Arbitrator, advocate, and medical expert must be clear about the level of medical certainty that constitutes sufficient burden of proof in the particular setting. Does the burden of proof require that it is more likely than not that the exposure caused the outcome? Or that the event would not have occurred in the absence of the exposure? Or does the burden of proof require only that the exposure made a substantial contribution to the occurrence of disease? The medical expert needs to know what the burden of proof is for the setting in which testimony is given

and must be prepared to explain how the facts justify the level of certainty.

Summary

Conflicts of medical testimony can present substantial difficulty for arbitrator and advocate. Considerable weight should be given to opinions that are based on epidemiologic studies of workers with exposures and diseases similar to the situation under dispute. Because a physician's personal experience may be prone to substantial bias, it may be overrated by arbitrators and advocates in terms of its usefulness.

Preplacement Medical Examinations

Role of the Americans with Disabilities Act (ADA)

The administration of preplacement medical examinations to applicants after a conditional job offer is made is currently the focus of considerable concern and controversy. The goal is to identify applicants who either cannot perform essential job functions or for whom certain job placements will result in unacceptable risk to health. Under the ADA, preplacement examinations are allowed, provided they are offered to all applicants in a particular job category and the results are not used to discriminate against those with disabilities unless they render the applicant not qualified for a particular job category.

According to the House Education and Labor Committee, employers are allowed to perform these examinations to "discover possible disabilities that do, in fact, limit the person's ability to do the job. . . ." In a discussion about the ADA, the Committee has been clear that medical testing can be used to deny employment only when the results show a "high probability of substantial harm if the candidate performed the particular functions of the job in question," and reasonable accommodations could not be made. The Committee has further stated that such assessments must be based on "valid medical analyses" and that "employers should be diligent in assuring that their examining physicians make assessments based on testing measures that actually and reliably predict the substantial, imminent degree of harm required."[8]

---

[8]The Americans With Disabilities Act: A Practical and Legal Guide to Impact, Enforcement, and Compliance (A BNA Special Report) (1990), at 246.

Assessment of Screening Tests

Two separate questions regarding risk to health arise from the use of preplacement examinations:

(1) Can the employee do the job now without experiencing an adverse health consequence (given reasonable accommodations)?
(2) Is the employee at risk for development of an adverse health consequence in the future?

For an example of a medical evaluation regarding the first question, consider work that requires the use of a respirator for control of exposure to an airborne hazard. Because respirators place an increased load on lung function, some workers with underlying lung disease are unable to use them. Physicians are often called upon to certify workers fit to use respirators. The best test for fitness is to place the worker in the work setting and observe job performance while using the respirator. Any worker who develops intolerable shortness of breath is not medically suited to the work. Reasonable accommodation might include in this case use of a "powered" respirator that reduces the work of breathing but costs considerably more than a simple non-powered respirator. This example illustrates an acceptable screening test because it is not likely to exclude from employment those able to perform the job, nor is it likely to allow employment of those unable to perform the job.

An example of an unacceptable medical test would be use of chest x-rays to assess a worker's suitability to use a respirator. This test is unacceptable because some with negative chest x-rays may have lung disease that renders them unable to use a respirator, and many with abnormal chest x-rays may have no difficulty tolerating the added physiologic burden imposed by use of the respirator.

Another example of legitimate denial of placement would be that of a worker with a known severe allergic reaction to a compound that is present as a contaminant in the work environment. Repeated exposure in this context places the worker at substantial risk of a life-threatening health event and therefore meets the ADA requirement for denial of employment.

Unfortunately, many medical situations are considerably less clear than these. For example, consider the worker who has a history of a previous heart attack (myocardial infarction) as well

as several risk factors for heart disease. Some individuals in this category may be at increased risk of a second heart attack when performing strenuous work. Tests to evaluate these workers are expensive, time consuming, and of limited value for predicting the likelihood of a serious health event. Any decision in this setting is subject to considerable error. A scientifically defensible method for making some of these decisions is currently lacking.

Review of the Technical Issues

The technical issues required for the evaluation of preplacement medical tests are those of sensitivity and specificity. Sensitivity refers to the proportion of those who actually have the condition of interest who test positively. If 90 of 100 people with the condition have a positive test result, then the sensitivity of the test is 90 percent. Specificity refers to the proportion of those who do not have the condition who test negatively. If 98 of 100 people who do not have the condition have a negative test result, then the specificity of the test is 98 percent.

Both the sensitivity and specificity of a test must be considered in order to evaluate its usefulness. Poor specificity, however, results in discrimination because many with positive tests will, in fact, not be at risk and will be incorrectly excluded from placement. For example, the following statement is often heard in defense of a test: "This test was positive in 99 percent of all workers who subsequently lost time due to illness X." This indicates that the test is "sensitive." However, the key piece of missing information is the proportion of workers who did not lose time due to illness X who also had a positive test. If the test is positive in many workers destined never to develop the condition (i.e., it has poor specificity), then its use will restrict many who are not at risk.

The previous example introduces the related issue of the use of medical tests to identify those at future risk of developing a disorder. In the area of common and costly illnesses, such as back pain and upper-extremity problems (also known as cumulative trauma disorders), considerable effort has been devoted to the development of screening tests to identify, at the time of hire, those ultimately destined to develop a work-related disorder. Scientific studies to assess the utility of such screening tests require long-term follow-up of individuals who have undergone testing, in order to determine the proportion of

those with both positive and negative test results, at the time of hire, who will eventually develop into clinical "cases" of the disorder of interest. Only after such studies have been performed can the ability of a test to predict the later development of a disorder be established. At the current time, there are virtually no medical tests available that predict future occurrence of work-related disease in otherwise healthy workers with sufficient certainty to meet the requirements established in the ADA legislation.

Summary

In summary, the use of preplacement medical examinations has been severely limited by the ADA. Under the ADA examinations can be performed only if they are offered to all applicants and can be used to deny employment only if the results indicate that the worker is incapable of performing the job (even with reasonable accommodations) or if the worker's health is at imminent, substantial risk of harm. Tests that reliably predict future occurrences of work-related disease in otherwise healthy workers are not currently available.

MANAGEMENT PERSPECTIVE

JOHN R. PHILLIPS*

Introduction

Medical and health issues highlight the arbitration agenda of the 1990s. Discipline or discharge due to an employee's physical or mental disability can be grieved. Employers may order psychiatric or psychological testing of workers to determine fitness for their jobs. No-smoking policies and AIDS issues are increasingly arbitrated. Health insurance benefits and cost containment measures continue as major items at the bargaining table and in arbitration hearings. This paper discusses these topics by reviewing selected recent arbitration decisions.

---

*Nyemaster, Goode, McLaughlin, Voigts, West, Hansell & O'Brien, P.C., Des Moines, Iowa. The author wishes to thank Penny R. Heaberlin, Drake Law School Class of 1993, for her contribution to this article.