A Practicable and Systematic Approach to Medicolegal Causation

Michael D. Freeman, MedDr, PhD, MPH

The kinds of questions that are commonly asked of treating orthopedic surgeons in the medicolegal context are often outside of the scope of ordinary clinical practice. Although questions regarding diagnosis and treatment are easily answered, questions regarding causal attribution, functional impairment and disability, prognosis, and future need for care are more difficult to address. The primary reason for this difficulty is the degree to which probabilistic reasoning and inference is relied on as a basis for the opinion, rather than observation and interpretation, the latter being the most common mode of medical practice.

A Systematic Approach to Medicolegal Causation

The nature of the types of injuries encountered in an orthopedic practice makes causation assessment a straightforward process in many instances. This is not the case with orthopedic injuries that do not require a high degree of energy, that can be present without producing symptoms, and that can have an insidious symptom onset. Spinal disk and joint degeneration, hip labral tears, and rotator cuff injuries are examples of conditions that commonly result from the wear and tear of the physiologic forces of daily life and that exist, to some degree, in a large portion of the asymptomatic adult population. They are also conditions that make individuals more prone to injury following a relatively minor trauma, such as a minor traffic crash or a ground-level fall. The methodology described here is directed at assessing the causation of such conditions when they have become symptomatic after minor trauma.

The scientific bases for general (population) and specific (individual) determinations of cause and effect were first formally described in the inductive canons of John Stuart Mill in 1862 and the rules proposed by the philosopher David Hume in 1739.[1] In the modern era, rules for assessing potential causal relationships have been formalized in the form of the "Hill criteria," a set of 9 viewpoints set forth by Sir Austin Bradford Hill in 1965.[2] Although Hill did not personally endorse the term "criteria" because he did not want the viewpoints to be seen as a causation checklist, this description is commonly used.

The Hill criteria can be distilled to 3 basic elements or steps for practical application in clinical assessments of causation[3-5]:

1. Plausibility. This first step addresses whether it is biologically *possible* for the injury event to have caused the condition. This step is also known as "general causation" in that it only addresses whether the injury *can* be caused by the trauma, not *whether* the injury was caused by the trauma. Epidemiologic and other scientific literature can serve as a basis for a determination of plausibility, but clinical experience is also an adequate basis for the determination. It is appropriate to use prior clinical experience with an injury mechanism, gleaned from patient histories, as a foundation for assessing whether an injury plausibly resulted from a similar injury mechanism.

Although it is common in injury litigation for the defendant to assert minimal vehicle damage or minimal force, as reconstructed by an engineer or other nonclinical expert, as a basis for disputing injury cause, the approach has virtually no relevance to plausibility, as such analyses do not have sufficiently low error rates to establish *implausibility* and at best can only be used to suggest a low frequency or risk of injury in the general population.

The author is from the Department of Psychiatry, Oregon Health & Science University School of Medicine, Portland, Oregon.

The author has no relevant financial relationships to disclose.

Correspondence should be addressed to: Michael D. Freeman, MedDr, PhD, MPH, Department of Psychiatry, Oregon Health & Science University School of Medicine, 4500 Kruse Way, Plaza I, Ste 385, Lake Oswego, OR 97035 (forensictrauma@gmail.com).

doi: 10.3928/01477447-20180227-02

2. Temporality. This second step, which examines the clinical and other evidence of the timing between the onset of the symptoms of injury and the injury event, must be satisfied to assess the specific cause of the patient's injury. First, it must be established that the sequence of the injury and the event is appropriate; the symptoms cannot be identically present prior to the event. Further, the onset of the symptoms of injury cannot be implausibly latent, relative to the injury event. For example, although the symptoms of a spinal disk injury in the neck most often will not immediately include upper extremity radiculopathy (and most disk injuries are initially diagnosed as a simple sprain or strain), a complete absence of symptoms in the neck for several months after a traffic crash, followed by the sudden insidious onset of symptoms of a cervical disk injury with radiculopathy, could not be plausibly related to the traffic crash in most cases.

Of course, there are injuries that are always immediately apparent, such as bony fractures, and some injuries are considered "distracting" from others because the pain they generate distracts attention from other, less painful injuries. And there are still other injuries that can mask or enhance pain from nearby parts of the body through neurological mechanisms that are mediated centrally at the brain and spinal cord. The assessment of the timing of the onset of symptoms reasonably attributed to an injury can sometimes be nuanced and complicated, requiring careful examination and consideration.

3. Lack of a more probable alternative explanation. This final step is akin to a differential diagnosis process (more accurately termed a differential *etiology*) in that the probability of alternative explanations is assessed and weighed relative to the injury risk of the traumatic event in question. Of course, any evidence of an intervening injury event must be evaluated as the most obvious competing explanation. Barring such evidence, however, the likelihood of the condition occurring spontaneously in the patient, given what is known about the patient's history and the incidence of the condition in the population generally, must also be assessed. Although traumatic injuries such as fractures do not occur insidiously in healthy individuals, symptomatic joint and spinal disk injuries can and do begin in many cases without antecedent overt trauma. Tempering this fact is that the competing risk of the insidious onset of symptoms is adjusted for the time interval between the injury event and the onset of symptoms attributable to the injury diagnosis, assessed and quantified in the second step of the causation analysis. The estimation of the likelihood of the competing explanation can come from clinical experience, epidemiologic study, or a combination of sources.

A recent legal decision from the 10th Circuit US Court of Appeals endorsed the 3-step causal methodology as the generally accepted approach for injury causation.[6] The decision was based on the 2009 publication that first described the application of the methodology to the analysis of spinal disk injury following a traffic crash.[4]

A brief example of the practical clinical application of the 3-step methodology is as follows: A 60-year-old female patient presents with a history of first-ever shoulder pain beginning within a day of a minimal damage rear impact collision. This ultimately leads to the need for an arthroscopic decompression and repair. A physical examination and magnetic resonance imaging reveal a partial tear of the supraspinatus tendon with degenerative changes at the labrum and a large spur at the acromion. Contained within this brief description is both a *plausible* explanation for the injury (traumatic loading of the shoulder secondary to a traffic crash) and the *temporal* relationship between the trauma and the symptom onset indicative of the diagnosed condition (less than a day following the crash). Thus, the first 2 steps of the analysis are easily satisfied.

The third step (whether there is a more likely alternative explanation) requires a more detailed analysis. A 60-year-old female patient, despite having no history of shoulder problems, is likely to have degenerative changes in the rotator cuff because of the high prevalence of such changes in the population of 60-year-old women.[7] At the same time, such a patient is unlikely to spontaneously develop symptoms requiring surgery, in part because of her asymptomatic history and in part because the annual incidence of shoulder surgery in the population of 60-year-old women is relatively low. It is reasonable to conclude that the chance that an asymptomatic 60-year-old woman would ever need shoulder surgery is small, that the chance that she would have needed surgery within a year of a traffic crash is much smaller, and that the chance that she would have developed the onset of symptoms leading to shoulder surgery on the same day as a crash is remarkably small. This probability can be roughly approximated from an estimate of the annual risk of surgery for the relevant age and gender group, which, for 2006, was less than 200 per 100,000 (1 in 500).[8] A daily risk can be calculated from the annual risk by dividing the 1 in 500 by 365 days, which yields 1 in 182,500. It is thus reasonable to estimate the competing risk that the patient would have developed symptoms and the need for surgery spontaneously on the same day as the crash (but had the crash not occurred) as less than 1 in 182,500. In comparison, even very low speed (no damage) traffic crashes carry at least a 1 in 100 risk of significant and persisting musculoskeletal injury.[9] The large difference in risk allows for quite a bit of room for error; even if the competing risk was underestimated by 100 times, it would only be 1 in 1825, which is still much lower than the risk of injury from a no damage traffic crash. Thus, even a traffic crash or other traumatic event with a very low risk of injury still carries substantially greater risk than the time-adjusted competing probability of spontaneous symptom onset in most cases with a close temporal association between the event and the symptoms indicative of injury.

Conclusion

A 3-step systematic approach to medicolegal injury causation, incorporating plausibility, temporality, and the lack of a more probable alternative explanation, provides a practicable and reliable methodology that meets both medical and legal standards of practice.

References

1. Morabia A. Hume, Mill, Hill, and the sui generis epidemiologic approach to causal inference. *Am J Epidemiol.* 2013; 178(10):1526-1532.
2. Hill AB. The environment and disease: association or causation? *Proc R Soc Med.* 1965; 58:295-300.
3. Freeman MD, Rossignol AM, Hand ML. Forensic epidemiology: a systematic approach to probabilistic determinations in disputed matters. *J Forensic Leg Med.* 2008; 15(5):281-290.
4. Freeman MD, Centeno CJ, Kohles SS. A systematic approach to clinical determinations of causation in symptomatic spinal disc injury following motor vehicle crash trauma. *PM R.* 2009; 1(10):951-956.
5. Koehler S, Freeman MD. Forensic epidemiology: a method for investigating and quantifying specific causation. *Forensic Sci Med Pathol.* 2014; 10(2):217-222.
6. Etherton v Owners Insurance Company, No. 14-1164 (10th Cir 2016).
7. Sher JS, Uribe JW, Posada A, Murphy BJ, Zlatkin MB. Abnormal findings on magnetic resonance images of asymptomatic shoulders. *J Bone Joint Surg Am.* 1995; 77(1):10-15.
8. Colvin AC, Egorova N, Harrison AK, Moskowitz A, Flatow EL. National trends in rotator cuff repair. *J Bone Joint Surg Am.* 2012; 94(3):227-233.
9. Freeman MD. Biomechanical, mechanical, and epidemiologic characteristics of low speed rear impact collisions. In: *Proceedings of the 67th Annual Meeting of the American Academy of Forensic Sciences.* Colorado Springs, CO: American Academy of Forensic Sciences; 2015.