A Review and Methodologic Critique of the Literature Refuting Whiplash Syndrome

Michael D. Freeman, DC, PhD, MPH,* Arthur C. Croft, DC, MS,† Annette M. Rossignol, ScD,‡ David S. Weaver, DC,§ and Mark Reiser, PhD¶

Reprinted from SPINE, Vol. 24, No. 1, January 1, 1999

SPINE Volume 24, Number 1, pp 86–98

Journal Club

A Review and Methodologic Critique of the Literature Refuting Whiplash Syndrome

Michael D. Freeman, DC, PhD, MPH,* Arthur C. Croft, DC, MS,† Annette M. Rossignol, ScD,‡ David S. Weaver, DC,§ and Mark Reiser, PhD¶

The validity of whiplash syndrome has been a source of debate in the medical literature for many years. Some authors have published articles suggesting that whiplash injuries are impossible at certain collision speeds; others have stated that the problem is psychological, or is feigned as a means to obtain secondary financial gain. These articles contradict the majority of the literature, which shows that whiplash injuries and their sequelae are a highly prevalent problem that affects a significant proportion of the population. The authors of the current literature critique reviewed the biomedical and engineering literature relating to whiplash syndrome, searching for articles that refuted the validity of whiplash injuries. Twenty articles containing nine distinct statements refuting the validity of whiplash syndrome were found that fit the inclusion criteria. The methodology described in these articles was evaluated critically to determine if the authors' observations regarding the validity of whiplash syndrome were scientifically sound.

The authors of the current critique found that all of the articles contained significant methodologic flaws with regard to their respective authors' statements refuting the validity of whiplash syndrome. The most frequently found flaws were inadequate study size, nonrepresentative study sample, nonrepresentative crash conditions (for crash tests), and inappropriate study design. As a result of the current literature review, it was determined that there is no epidemiologic or scientific basis in the literature for the following statements: whiplash injuries do not lead to chronic pain, rear impact collisions that do not result in vehicle damage are unlikely to cause injury, and whiplash trauma is biomechanically comparable with common movements of daily living. Spine 1999;24:86–98

From the *Department of Public Health and Preventive Medicine, Oregon Health Sciences University School of Medicine, Portland, Oregon, the †Spine Research Institute of San Diego, San Diego, California, the ‡Department of Public Health, Oregon State University, Corvallis, Oregon, §Private Practice, Keizer, Oregon, and the ¶Department of Economics, Arizona State University, Tempe, Arizona.
Acknowledgment date: August 13, 1998.
Acceptance date: August 13, 1998.

One of the more frequently disputed conditions in the medical literature in recent decades is the constellation of symptoms comprising acute whiplash and its chronic iteration, late whiplash (collectively known as whiplash syndrome). The primary reason for the dispute stems from the fact that the validity of whiplash syndrome often is a key issue in litigation arising from the alleged etiology of the whiplash, *i.e.*, a motor vehicle crash in which the injured party is not at fault. The judge and/or jury in such cases are asked to weigh opposing medical and scientific evidence supporting both the plaintiff's position that whiplash injuries and their sequelae are real and the defense position that the injuries are manufactured or greatly exaggerated. Over $29 billion per year is spent on whiplash injuries and litigation in the United States alone.[14]

It is not surprising, considering the financial stakes, that many medical experts have dedicated their professional careers to one side or another of the whiplash controversy. These experts increasingly are relying on medical and engineering literature to support both sides of the debate over the validity of whiplash syndrome.

A recent review of the literature reported over 10,000 articles relating to whiplash injuries.[44] The majority of this literature is devoted to probing fundamental questions about whiplash injuries, such as mechanism of injury, pathogenesis, and epidemiology. More than 30 epidemiologic studies have been published that document the cumulative incidence (risk) of chronic (lasting longer than 6 months) whiplash symptoms, or "late whiplash." In a recent publication, 13 of these studies were considered sufficiently well constructed (low selection bias, sufficient study size, adequate research methodology) to be relied on for an accurate clinical projection for late whiplash.[15] A study population-weighted meta-analysis of these studies showed a 0.33 risk of late whiplash at 33 months after injury for those seeking treatment for acute whiplash injuries.[14] Therefore, the epidemiologic literature appears to support a substantial risk of chronicity after acute whiplash injury.

Federal government statistics and epidemiologic studies indicate that whiplash syndrome affects a large number of people. The National Highway Traffic Safety Administration reports that, in 1995, there were 5,500,000 Americans injured in motor vehicle crashes (MVCs).[7] A large, population-based study found that 53% of MVC

injuries include whiplash injuries, amounting to 2,900,000 acute whiplash cases in 1995,[22] or an incidence rate of 1107 per 100,000 person-years.[14] If, as is suggested by the results of the meta-analysis described earlier, 33% of acutely injured persons continue to experience symptoms at 33 months after injury, then as many as 900,000 new cases of late whiplash may have occurred in the U. S. in 1995.

A recent case–control study of 665 patients with chronic spine pain found that 45% of patients who reported having at least one intrusive episode of neck pain weekly for more than 6 months attributed the onset of their symptoms to a whiplash injury. Although it is important to keep in mind that the results of any case–control study must be interpreted carefully because of the potential effect of recall bias, if the results of this chronic neck pain study are applied to what, to the authors' knowledge, is the most conservative published estimate of the prevalence of chronic neck pain in the population (13.8%),[9] then it can be reasonably, if cautiously, estimated that 6.2%, or 15.5 million Americans, currently have late whiplash. Other authors have estimated chronic neck pain prevalence to be as high as 32.9% for women and 27.5% for men[2]; therefore, the prevalence of late whiplash could be substantially higher.

Despite the strong epidemiologic evidence supporting whiplash syndrome as a valid clinical entity that leaves many persons with permanent symptoms, numerous articles have been published, the majority since 1990, that refute the validity of some or all aspects of whiplash syndrome. And, although the entire whiplash literature base has been criticized for methodologic weakness in general,[5,8,44] the quality of the literature refuting whiplash syndrome has stood largely unchallenged.

The current study reviews, from a methodologic perspective, the literature refuting whiplash syndrome. The objective of this review is twofold. The first objective is to determine whether there are significant methodologic flaws in the individual articles that may undermine the accuracy of their conclusions regarding the biomechanics, pathogenesis, or epidemiology of whiplash syndrome. The second objective is to determine, if there are methodologic flaws in the literature, *i.e.*, whether there are categorical flaws that are common to more than one study.

Methods

The literature was searched for articles that contained statements in the abstract or conclusions that refuted the validity of part or all of whiplash syndrome. For the purpose of the current study, whiplash syndrome was defined as injuries and their sequelae resulting from indirect trauma to the spine after MVCs of low to moderate severity. Late whiplash was defined as whiplash syndrome persisting for longer than 6 months.

The literature was searched for titles or abstracts containing the term "whiplash." Literature databases searched were Medline, and those of the Society of Automotive Engineers, International Research Council on Biomechanics of Impact, and National Technical Information Service for the years 1966 through 1997, in addition to published studies the authors were aware of that contained statements refuting whiplash syndrome.

More than 2000 articles were reviewed at least cursorily to determine their relevance to the current review. Of these, more than 700 of the most relevant articles were read *in extenso*. The articles were reviewed for specific statements that were considered to be contrary to the current authors' understanding of how the majority of the current literature characterizes the biomechanics, pathogenesis, and epidemiology of whiplash syndrome. These statements were categorized and described. In addition, logical implications of the statements that may arise in a medicolegal setting were extrapolated and described. The statements and their respective implications are listed in Table 1.

The studies then were reviewed by the authors for the presence of significant methodologic flaws. A significant methodologic flaw was defined as a potential threat to the validity of the study in light of the study's conclusions regarding whiplash syndrome. In other words, although some of the study's methods and inferences regarding whiplash syndrome may have been valid, the study methods were evaluated solely in reference to its conclusion or conclusions that caused it to be included in the current critical review.

The authors of this review were asked to critique the articles individually, and if methodologic flaws were found, to describe them. The methodologic errors then were described, categorized, and put into table form (Table 2).

Results

The literature search revealed 20 articles containing statements in the abstract, conclusions, or elsewhere in the text that were interpreted as refuting whiplash syndrome. Those statements are summarized in Table 1 at the end of this results section.

The articles ranged, with respect to study type, from literature reviews to cohort studies. The articles either were designed *a priori* as a refutation of whiplash syndrome or were designed for another purpose but made extrapolative statements that refuted the validity of whiplash. The articles were divided primarily between biomedical studies and editorials, and engineering studies.

All 20 articles were found to have significant methodologic flaws relative to their proclamations regarding the validity of whiplash syndrome. These flaws were of sufficient magnitude to cast doubt on the theoretical basis for the stated link between the study results and the conclusions of the study regarding the validity of whiplash syndrome. The articles are categorized below, according to study type. A brief description of the major points of each article is given, followed by a discussion of the methodologic flaws that were found in this review. If there were flaws that were common to more than one study in a category, then all of the studies with the common flaws are listed, followed by a description of the flaws.

Table 1. Statements Refuting the Validity of Whiplash Syndrome

#	Statement	Implication
1	Acute whiplash injuries do not cause, or are unlikely to cause, chronic pain.	Those claiming chronic pain after whiplash do not actually have chronic pain that can be attributed to the whiplash trauma.
2	An identifiable threshold exists and has been documented that relates velocity at impact to injury potential. The threshold of injury is above that of vehicle damage.	Collisions in which there is no vehicle damage cannot cause occupant injury.
3	Specific actions or movements common to daily living or sports and recreational activities don't cause injury, yet involve forces similar to, or higher than, those produced in whiplash injuries.	Whiplash trauma should not cause injury because it is comparable with other, noninjury producing movements.
4	Whiplash injuries, particularly late whiplash, are less common in countries where no remuneration exists for the injuries and their long-term sequelae, or where awareness of the injury is not thought to be widespread.	Whiplash injuries, particularly late whiplash, are not real, and claims of injury, chronic pain, and disability are a result of greed or are psychogenic.
5	Symptoms commonly attributed to whiplash injuries following low-speed motor vehicle accidents, particularly chronic neck pain, are psychogenic.	Reports of chronic pain following a low-speed motor vehicle accident are not a result of a physical injury.
6	There is a direct correlation between the severity of impact forces and the probability of developing chronic symptoms.	If the impact is not severe, chronic symptoms are unlikely.
7	It is highly unlikely or impossible to injure the temporomandibular joint in a whiplash-type motor vehicle accident injury.	Claims of injury of the temporomandibular joint following whiplash trauma are false.
8	Management of acute whiplash injuries is a contributing cause of late whiplash.	No treatment is effective for whiplash injuries, and seeking and receiving treatment reinforces illness behavior in whiplash patients.
9	The prevalence of chronic neck pain in the general population is the same as the risk of late whiplash following an acute whiplash injury.	Symptoms of late whiplash are a result of a condition predating the motor vehicle accident and would be present regardless of the history of an acute whiplash injury.

Cohort Studies

Study by Shrader et al. These authors studied 202 individuals in Lithuania who had been involved in a MVC.[42] This cohort was age- and gender-matched with a control group of 202 individuals who had no history of a MVC. The two groups were surveyed for neck pain an average of 21.7 months after their crash (relative to the time of the MVC for the MVC-exposed cohort) and were found to have the same prevalence of neck pain. The authors concluded that whiplash injuries do not cause chronic symptoms, and that late whiplash exists in industrialized countries because insurance settlements are available to those claiming chronic pain.[42]

Methodologic Errors

Inadequate Sample Size. This study was criticized because only a very small proportion of the exposed cohort (15% [31 subjects]) had been injured initially and therefore exposed to the putative etiologic agent in late whiplash (an acute whiplash injury).[16] For the purposes of the current literature critique, a *post hoc* sample-size calculation was performed on the data in this study, using an alpha of 0.05 and a beta of 0.20. The smallest detectable difference between the groups was 14.6%. Thus, 94% of the acutely injured subjects (29 of 31) in this study would have had to develop chronic symptoms to enable the authors to detect a statistically significant difference between the two groups, an extremely remote possibility. A recalculation of sample size using a meta-analysis–based estimate of effect (expected proportion chronic) of 5%[14] (*i.e.*, 33% of the 15% acutely injured subjects) demonstrates that the total study cohort needed to include at least 3000 individuals to have sufficient statistical power to discern a significant difference between the two groups.

Study by Balla. Balla[4] reported on a cohort of 20 whiplash patients who were treated by an orthopedist in Singapore, with a follow-up period of more than 2 years. He reported that none of the 20 patients had symptoms of late whiplash, and concluded that late whiplash was rare in Singapore, compared with its incidence among a group of 300 Australian patients. Balla attributed the late whiplash rate difference between the two countries to cultural differences and economic factors, among other variables.

Methodologic Errors

Inappropriate Study Design. Balla compared a group of 300 patients with late whiplash with 20 patients who had been evaluated after whiplash trauma. Not only were the numbers in the two groups grossly disparate, but the patients were enrolled in two different studies using different enrollment criteria and study protocol. The 300 Australian patients were selected for study because they had late whiplash. The 20 Singaporean patients were recruited from a specialist's practice on the basis that they had sustained an acute whiplash injury. As a result of different selection criteria for the two groups and other dissimilarities, the study could not validate or invalidate the author's hypothesis that the natural history of whiplash injuries in Australia differs from that of such injuries in Singapore.

Inadequate Sample Size. Twenty patients is not a sufficient size for a prospective study of late whiplash.

Table 2. Methodologic Errors in the Critiqued Articles

Author	Statement Category Number	Methodologic Errors							
		Nonrepresentative Crash Conditions	Inadequate Sample Size	Nonrepresentative Study Sample	Inappropriate Study Design	Unsubstantiated/ Unreferenced Claims	Unsupported Conclusions	Misquoted Literature	Improper Use of Terminology	Misleading Illustration
Allen et al	2,3		X	X	X		X			X
Awerbuch	8					X				
Balla	1,4,5			X	X					
Bovim	1,9									
Castro et al	2	X	X	X			X	X		
Ferrari and Russell	2,3,5,6					X				
Heise et al	7		X		X					
Howard et al (1995)	2,7	X		X	X		X			
Howard et al (1991)	2,7				X	X				
McConnell et al (1995)	2	X	X	X						
McConnell et al (1993)	2	X	X	X						
Mertz and Patrick (1971)	2	X	X	X	X					
Mertz and Patrick (1967)	2	X	X	X	X					
Mills and Horne	4						X			
Rosenbluth and Hicks	2,3	X	X	X	X					
Schrader et al	1,4		X							
Spitzer et al	1						X		X	
Szabo et al (1994)	2	X	X	X						
Szabo and Welcher (1996)	2,7	X	X	X						

Using Balla's Singapore data, a *post hoc* power calculation was performed, assuming that the risk of late whiplash in Australia of 33% (a literature-based assumption) was an unlikely eight times greater than that in Singapore. At least 44 randomly selected subjects would be needed in Singapore for such a study. Recalculation of power using a more reasonable risk ratio of three to one results in the need for 64 randomly selected Singaporean subjects. The current authors' power calculation assumed several study factors not actually present in Balla's[4] study: identical selection criteria in both countries; random patient selection, with control for potentially confounding differences between the countries not attributable to cultural differences; and identical subject appraisal criteria.

Selection Bias. Selection bias was introduced in this study when patients in Australia were selected for study retrospectively based on their disease status (they already had late whiplash when the study was begun) and the patients from Singapore were selected prospectively based on their exposure status (an acute whiplash injury).

Study by Heise et al. Heise et al[18] reported on 155 patients treated in an emergency room after experiencing whiplash trauma. The patients were divided into two groups: 63 patients with (unspecified) radiographic evidence of cervical musculoskeletal injury and 92 patients with no radiographic evidence of injury. The two groups were examined and interviewed for temporomandibular joint (TMJ) symptoms at the time of initial presentation, then their progress was followed by conducting phone interviews 1 month and 1 year subsequently. The follow-up rate at 1 year after their initial examination was 70% of the positive radiographic findings group and 65% of the negative radiographic findings group. None of the patients who were contacted at 1 year had continued symptoms of TMJ dysfunction. The authors concluded that the incidence of TMJ injury present after whiplash trauma was "extremely low."[18]

Methodologic Errors

Inappropriate Study Design. The authors do not state their rationale for stratifying their cohort into two groups on the basis of "positive radiographic findings" of whiplash, which are unspecified. The authors of this review were unable to find any reference in the literature to a correlation between TMJ injury and radiographic findings of whiplash injury that would justify the study design used by Heise et al.

Inadequate Sample Size. Using a literature-based estimate of effect of 0.04[22] (*i.e.*, 4% of the whiplash-injured population will sustain a TMJ injury), an alpha of 0.05, and a beta of 0.20, the authors of the current review performed a *post hoc* power calculation on Heise et al's[18] data. Assuming only double the frequency of TMJ injury in the exposed group, the

authors would have needed more than 2500 subjects for their study. Assuming a highly unlikely eight times greater frequency of TMJ injury between the two groups studied, the authors still would have needed more than 650 subjects, four times greater than the number in the study.

Case Series Studies

Study by Spitzer et al. In their Quebec Task Force (QTF) on Whiplash-Associated Disorders monograph, Spitzer et al[44] conducted a retrospective case series study and a literature search, and issued a set of guidelines and recommendations based on the results. Among other findings, the QTF concluded that whiplash injuries were "short-lived," involving "temporary discomfort," that the pain resulting from whiplash was "not harmful," and that whiplash injuries have a "favorable prognosis." They also concluded that 87% and 97% of their cohort "recovered" from their whiplash injuries at 6 months and 12 months after the crash, respectively.[44]

Methodologic Errors

Improper Use of Terminology. The Results and Discussion section of the case series study contained numerous references to the percentage of the study population "recovered" at the time of cessation of compensation. However, the QTF did not gather any data regarding the symptoms, amount or type of treatment, or functional impairment of its cohort—all factors necessary to determine the level of recovery after an injury. The QTF chose to define "recovery" unconventionally as cessation of time-loss compensation. Not surprisingly, the QTF found that 87% and 97% of its cohort was "recovered" at 6 and 12 months after the crash, respectively. To refer to these individuals as recovered misrepresents the data collected.[15]

Unsupported Conclusions. In a table entitled "Prevalence of Symptoms at Follow-Up," the QTF enumerated the four studies on prognosis that were accepted for review, along with its findings, which were as follow: Norris and Watt[30] reported that 66% of their cohort had neck pain at an average of 2 years after injury; Radanov et al[35] reported that 27% of their cohort were symptomatic 6 months after the MVC, and, in a study 2 years later,[36] reported that 27% of their cohort had headaches 6 months after the MVC; and Hildingsson and Toolanen[19] reported that 44% of their cohort were symptomatic at an average of 2 years after the MVC.

Based on its literature review and its cohort study, however, the QTF concluded that "Whiplash-associated disorders are usually self-limited," and "Patients should be reassured that most WAD are benign and self-limiting," inaccurately summarizing the results of its literature review and case-series study.

Cross-sectional Study

Study by Bovim et al. In their report on chronic neck pain in the general population in Norway, Bovim et al[9] stated that 13.8% of respondents experienced "troublesome neck pain" for longer than 6 months. The authors compared this proportion with similar figures reported by previous authors regarding the risk of late whiplash following an acute whiplash injury and concluded that "chronic neck pain after whiplash injuries may be a continuation of pre-existing complaints."[9]

Methodologic Error

Misquoting Literature/Selecting Biased Literature. The basis for the primary conclusion of Bovim et al[9] is the comparison of their survey results with a literature-based estimate of the prevalence of late whiplash among the population of individuals who have sustained whiplash trauma. The authors referenced four articles that contained estimates of chronicity following whiplash. One of the articles, written in Norwegian, could not be evaluated for this critique. The remaining three articles were stated to have reported a prevalence of chronicity of 12–18%. However, the authors did not reference 27 of the 30 articles on whiplash prognosis available in indexed journals at the time of their study. A meta-analysis of the 13 highest-quality articles on whiplash chronicity showed that 33% of whiplash-injured individuals will have chronic neck pain at 33 months after the crash.[14] This more accurate appraisal of the literature-based estimate of chronicity invalidates Bovim et al's hypothesis that late whiplash is merely a continuation of pre-existing neck pain. Additionally, Bovim et al misquoted the article by Gotten,[17] claiming he had found a 12–18% chronicity rate, when in actuality, Gotten reported a prevalence of late whiplash of 46% at 12 months after the MVC.

Correlational Study

Study by Mills and Horne. Mills and Horne[29] compared the rate of whiplash injuries in Victoria, Australia, with the rate in New Zealand. They reported that the rate was substantially higher in Victoria and concluded that the difference was attributable to the fact that an injured occupant in Victoria must seek compensation through the common law system, as opposed to New Zealand, where apparently it is less difficult to gain compensation for MVC-related injuries. The authors concluded that Victorians are "more conversant with and more attuned to receiving compensation for injury, which may in itself be stimulus for claiming an injury that they would not normally have claimed for."[29]

Methodologic Error

Unsupported Conclusions. The authors do not present any evidence that supports their statement that the greater barriers to claiming compensation in

Victoria actually increase claims of whiplash injury. Indeed, the logical conclusion is quite the opposite. The difference in the whiplash rate between Victoria and New Zealand may be accounted for by any of a variety of potentially confounding factors that may exist between the two countries, including different criteria for reporting and recording whiplash injuries, different driving conditions, or different diagnostic classification systems.

Literature Reviews/Editorials

Review/Editorial by Ferrari and Russell. In their editorial/literature review, Ferrari and Russell[13] stated that more than "2000 runs of volunteer collisions have been conducted using specialized sled devices and actual vehicles (old and new, big and small), and never, ever, has the multitude of chronic symptoms of whiplash patients been reproduced."[13]

The authors stated that it is "unacceptable, however, to claim that a muscle sprain or some as yet unidentified injury is responsible for the chronic pain and the large number of symptoms of whiplash patients. Instead, the symptom complex can be explained as a whole not by an injury, but rather by a psychological disorder."

Methodologic Errors

Unsubstantiated/Unreferenced Claims. Ferrari and Russell provide no citation for their statement regarding the number or scope of crash testing. The literature review performed for the current critique revealed published accounts of fewer than 100 volunteers in crash tests, with the largest single majority (42 subjects) from one study that was published after Ferrari and Russell published their article.[43] Although the authors state that no crash test study has ever produced chronic symptoms, there is no evidence in the literature to substantiate this statement. The authors of the current critique were only able to find two studies with a total of nine volunteers that informally followed the subjects for more than a few days to determine if there were chronic symptoms after crash testing.[25,46]

The authors do not cite any references to substantiate their statement that it is "unacceptable" to claim an as-yet unidentified cause of chronic pain that follows whiplash. Although the authors state that no cause has been identified for chronic pain following whiplash, they ignore the research of Barnsley et al,[6] who have demonstrated quite convincingly the cervical zygapophyseal joints as the origin of a substantial proportion of chronic neck and head pain following whiplash trauma. Ferrari and Russell do not cite any references that substantiate their claim that late whiplash is a psychogenic illness.

Review/Editorial by Awerbuch. In his literature review/editorial, Awerbuch[3] stated that as soon as a doctor makes a diagnosis of whiplash, he or she is contributing to the patient's potential for chronicity. The author continued, "later the patient may be referred for a range of imaging (plain x-ray, computed tomography, isotope bone scan, MRI, or thermography) which can only be interpreted by the patient as being necessary to define the gravity of the 'whiplash' injury," thus, further contributing to the potential for chronicity.[3]

Methodologic Error

Unsubstantiated/Unreferenced Claim. The author does not cite any published sources to substantiate the statement that treatment and diagnosis contribute to potential chronicity of whiplash symptoms. Awerbuch[3] overlooks the alternative explanation that symptomatic patients may be more likely to need additional treatment and diagnostic testing.

Crash Test Studies

1. Study by McConnell et al (1993). McConnell et al[26] reported the results of human volunteer rear-impact crash testing of four subjects. They determined that, in reference to whiplash injuries resulting from rear-impact collisions, the threshold of a "very mild, single event musculoskeletal cervical strain injury" is a delta V (the absolute velocity change of the struck vehicle as opposed to the speed of the striking vehicle at impact) of 4–5 miles per hour (mph).[26]

2. Study by McConnell (1995). McConnell et al[25] studied the movements and acceleration forces sustained by seven human occupant volunteers subjected to repeated rear-end collisions of up to 6.8 mph delta V. They concluded that at a delta V of 5 mph "the likelihood of transient acute neck and shoulder muscle strain injury and possible mild compressive irritation of the posterior neck may increase" for the average vehicle occupant. They also concluded that any injury to the low back is "quite unlikely as a result of a low velocity rear end collision."[25]

3. Study by West et al. West et al[48] studied the acceleration forces sustained by six human volunteers in crash testing of five different vehicles. They concluded that vehicle occupants are unlikely to be injured in collisions with an equivalent barrier speed (EBS) of less than 8 mph (EBS is an estimate of impact speed based on vehicle damage, compared with a known amount of damage from a 30-mph collision with a fixed barrier). The authors also stated that they did not observe jaw opening during crash testing and that this finding rebutted claims that TMJ injury can result from whiplash trauma.[48]

4. Study by Szabo et al. Szabo et al[46] reported on human volunteer crash testing of five volunteers who were in vehicles that were struck in the rear at approximately 10 mph by another vehicle, resulting in an average delta V of 5 mph. The volunteers were evaluated by an orthopedic surgeon and underwent magnetic resonance imaging before and after the crash testing. Although four of five

volunteers reported having a headache immediately after the crashes, none had symptoms that lingered for more than 2 days, and no volunteers reported further symptoms during the subsequent year. The authors concluded that rear-end collisions with a delta V of 5 mph or less were within human tolerance levels, and that injury was unlikely after such a collision. Szabo et al concluded that the jaw does not open during whiplash trauma, and stated that their study results support an earlier author's contention that there is no potential for TMJ injury as a result of a whiplash trauma.

5. Study by Szabo and Welcher. Szabo and Welcher[47] reported on volunteer crash testing of four men and one woman. Each volunteer was exposed to two rear-end collisions with an average closing speed of 8.9 mph and an average delta V of 5.8 mph. The authors concluded that "a rear impact with a change on velocity of [5 mph] or less is within tolerance for a reasonably healthy occupant. . ."

6. Study by Mertz and Patrick (1967). Mertz and Patrick[27] studied the responses of a human volunteer, a cadaver, and anthropomorphic dummies to simulated rear-end collisions. They compared the responses of the volunteer with an index of neck injury that was developed for the study by statically loading the neck of one of the authors with tension to the point that the author believed that injury might occur. The authors concluded that a 10-mph rear-end impact for an unsuspecting occupant was within human tolerance for injury.[27]

7. Study by Mertz and Patrick (1971). Mertz and Patrick[28] used an anthropometric dummy, four cadavers, and one of the authors for sled testing simulating a rear-impact collision. The author who had volunteered for the study sustained accelerations at the head of 1.9–6.8 gravity (g) with no injury. However, a 9.8 g acceleration resulted in back and neck injury. The authors developed a guide for tolerance to injury in a whiplash trauma.[28]

8. Study by Rosenbluth and Hicks. Rosenbluth and Hicks[40] studied the acceleration forces sustained by two human crash-test volunteers who were seated in a vehicle that was struck from behind at an EBS of up to 4.8 mph. They concluded that an EBS of 4.8 mph was below the threshold of human injury tolerance. The authors also measured the acceleration forces at the head (as measured by tri-axial accelerometers affixed to a helmet) of a 7-year-old child and a 29-year-old adult skipping rope. They reported that acceleration at the head was similar to that found in the crash testing.[40]

9. Study by Howard et al. Howard et al[20] studied the acceleration forces at the TMJ that occurred during rear-impact crash testing of four human volunteers. These authors used accelerometers fitted to a bite plate to measure the acceleration forces at the approximate level of the TMJ during 5 mph delta V impacts. They concluded that the forces measured at the jaw during crash testing constitute a "minor fraction" of the normal forces experienced during mastication, and that low-velocity whiplash trauma cannot cause injury to the TMJ.[20]

10. Study by Castro et al. Castro et al[11] studied the effect of 17 rear impacts with an average delta V of 7.1 mph on 14 men and 5 women (the authors did not specify which two study participants were excluded from crash testing).[11] Of the 17 impact-exposed participants, five (29%) reported whiplash symptoms following testing, including one male participant who had objective findings of injury 10 weeks after the crash. The authors concluded that "the 'limit of harmlessness' for stresses arising from rear-end impacts with regard to the velocity changes lies between [6.2 mph] and [9.4 mph]."

Methodologic Errors

Inadequate Study Size (Studies 1–10). When attempting to study a population sample to make an inference that is applicable to a population beyond that of the study, it is essential to use inferential statistics to determine if the study results were causally related to the variables under study, or if they were caused by random variation. With crash testing, the dependent variable (the variable under study) is injury status; either an occupant is injured or not injured. Because the two outcomes are mutually exclusive, a 95% confidence interval can be established for the study results using a binomial probability distribution that is based on the study size. That is, if the study were to be repeated, the 95% confidence interval tells us how many and how few injuries are possible, based on the results of the current study. The width of a confidence interval is indirectly related to the number of participants in a study, because random error makes the interpretation of the study results less precise, *e.g.*, if a coin is tossed three times and heads is observed all three times, it is much less precise to state that the coin has heads on both sides, in comparison with 100 coin tosses resulting in heads.

Even with crash testing, with as many as 20 participants who sustain no injury in the crash test, the probability of injury in a larger population is still 0.15 (based on the confidence interval), which means that three participants could be injured the next time the same study is conducted with the same individuals, and those results would still be consistent with the results of the current study. Thus, the confidence interval for crash test studies of five or six individuals is too wide to conclude that no injury is possible under similar conditions. To describe the range of injury responses for the general population adequately, given the wide variety of human susceptibility to injury, vehicle types, crash conditions, and other such variables, many hundreds or even thousands of individuals would need to be studied in crash tests.

Nonrepresentative Study Sample (Studies 1–10). The participants in the crash test studies consisted of the

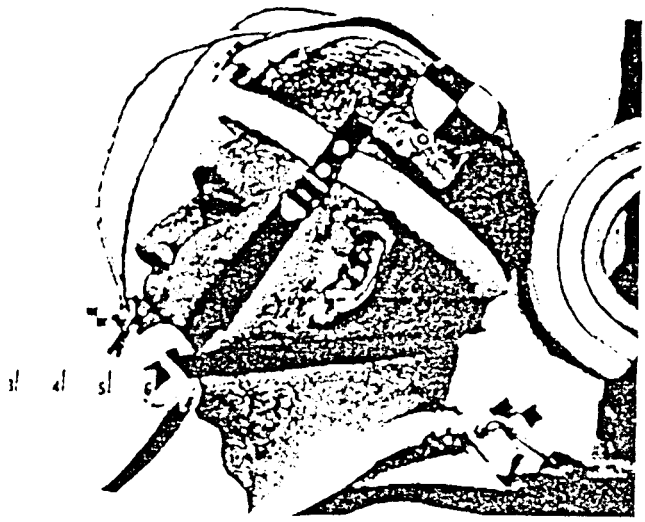

Figure 1. Reprinted with permission from SAE paper no. 710855.[28] © 1971 Society of Automotive Engineers, Inc.

authors of the studies, employees of the corporations financing the study, and other associates of the authors who may have a vested interest in the outcome of the study. In addition, almost all of the test subjects were male. To generalize the results of any study to a larger population (in this case, the general population at risk for whiplash injuries), the study population must adequately represent the larger population.

Nonrepresentative Crash Conditions (Studies 1–10). Even if the numbers of study participants were sufficient to generalize the results of the above-listed crash tests to the general population, the results would be applicable only to healthy men who were prepared for a rear impact and perfectly situated in the vehicle seat at the time of impact. Only a very small proportion of the crash-injured population fits this description.

For their crash test, Mertz and Patrick[28] used a sled seat with a specially designed head restraint that did not allow for any posterior movement of the head (Figure 1). The results of such crash testing are not generalizable to the population at risk for whiplash trauma, because car seats allow for posterior excursion of the head, which is the most significant injury-producing phase of whiplash trauma.[28]

Inappropriate Study Design (Studies 8 and 9). Howard et al[20] used a bite plate to measure forces at the TMJ, which required firm closure of the mouth on the plate. Because jaw-opening is integral to the mechanism of injury at the TMJ during whiplash,[41] having the participants keep their mandible firmly elevated during the crash testing defeated the purpose of the study, and the results are meaningless with regard to the actual forces sustained at the TMJ during *in vivo* whiplash trauma.

Rosenbluth and Hicks[40] gave no rationale for comparing whiplash trauma with rope-skipping. The maximum acceleration reported in the x vector was 3.5 g for the 7-year-old and approximately 1 g for the 29-year-old, far less than ranges of acceleration reported by other authors for low-speed, rear-impact crash testing (6–14.5 g).[43,48] The difference between the acceleration noted for the child and that for the adult may be artifactual, because the helmets were secured to the study participants with a single strap under the chin, an arrangement that may have allowed for excessive movement between the helmet and the head (Figure 2).

Unsupported Conclusions (Studies 9 and 10). Howard et al[20] compared the acceleration forces measured at the TMJ during a low-velocity rear-impact collision with those of mastication, concluding that the noninjurious forces of mastication were far greater than those of whiplash trauma. However, the authors did not study acceleration forces specifically at the TMJ; therefore they cannot compare the forces measured in their study with those found with mastication, because mastication produces a differential acceleration between the cranium and the mandible. Because the jaw was closed in this study, the mandible was accelerated at the same rate as the cranium, and no differential movement for the two osseous components of the TMJ was allowed. There was no scientific support for the conclusions of the authors regarding

Figure 2. © ASTM.[40] Reprinted with permission.

TMJ injury potential in the methods or results of this study.

Castro et al[11] noted symptoms of whiplash injury in 29% of their study subjects, yet ignored their study results when concluding that similar impacts were harmless. The authors contradicted their own study findings in their conclusions.

Biomechanical Studies

Study by Allen et al. Allen et al[1] studied the acceleration forces of common movements in eight volunteers with triaxial accelerometers affixed to a helmet. They reported peak accelerative forces, measured while subjects "plopped in a chair," that were similar to accelerative forces recorded during published accounts of volunteer crash testing. Citing the results of their study, the authors stated that "no-damage accidents," like the common movements examined in the study, were unlikely to cause injury.

Methodologic Errors

Unsupported Conclusions. Allen et al[1] concluded that whiplash trauma and ordinary daily movements were comparable, even though none of the movements studied duplicated the vector or force of whiplash trauma. The majority of acceleration in a rear-impact crash is in the *x* vector, *i.e.,* front to back. The largest single acceleration reported by these authors was 10.1 g in a diagonal vector (54.9 degrees from horizontal) during "plopping in a chair" (Figure 3) However, the *x* vector component was only 5.6 g. In "Table 2" of their article, the mean *x* vector acceleration of plopping in a chair was 3.3 g, the highest mean *x* vector acceleration of all of the movements. In actuality, Allen et al reported that 10 of the 13 movements studied had mean *x* vector accelerations of less than 2 g. In comparison, West et al[48] reported a range of peak acceleration at the head during crash testing of six volunteers of 6–14.5 g (at 9 km/hour EBS). Siegmund et al,[43] in the largest published crash test to date, reported 6.7–12 g of peak head acceleration among 41 study participants crash-tested at 8 km/hour delta V. Additionally, the duration of peak acceleration of the movements studied by Allen et al (approximately 1 millisecond) is not comparable with the duration of peak acceleration measured during whiplash trauma (70 milliseconds).[43] Taking into account both components of acceleration (magnitude and duration), whiplash trauma produces a peak accelerative force that is more than 150 times greater than that produced by plopping in a chair.

Misleading Illustration. In Allen et al's[1] illustration of the acceleration forces measured while "plopping in a chair," the authors showed a human head apparently moving into extension, with an arrow traveling rearwards through the head, and "10.1G" labeled at the arrow head (Figure 3). However, the legend of the figure parenthetically states "the apparent axis of rotation of the head in this schematic is not the true motion of the head. It is an expression of the acceleration forces." In spite of the disclaimer in the legend, it appears that the authors are attempting to convince the reader that "plopping in a chair" produces the same vector and magnitude of acceleration, as well as movement at the head, as a rear-end collision.

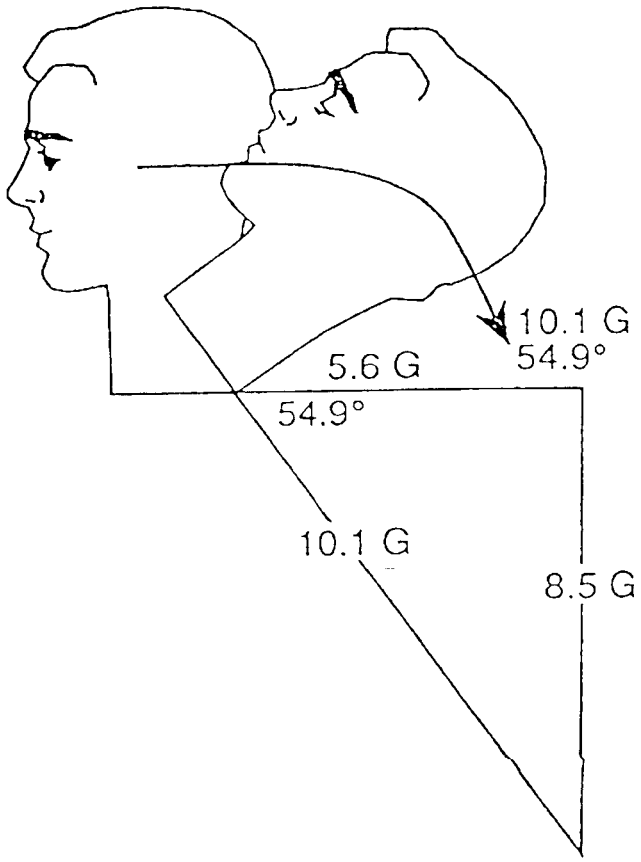

Figure 3. Vector forces. Schematic respresentation of the vector forcers and direction for the high values of "plop in chair" event, scaled to represent the angle at which the forces would be expressed. (The apparent axis of rotation of the head in this schematic is not the true motion of the head. It is an expression of the acceleration forces. (Reprinted with permission from Lippincott-Raven).[1]

Inappropriate Study Design. Allen et al[1] did not give a rationale for comparing common movements that do not usually cause injury to whiplash trauma, which results in 2.9 million injuries annually. By its design, Allen et al's study could not yield any information about whiplash injuries, because neither whiplash injuries nor the mechanism of injury in whiplash injuries was studied.

Study by Howard et al. In their article on the theoretical biomechanics of temporomandibular joint during whiplash trauma, Howard et al[21] state that "head accelerations produced by forces in the neck (extension-flexion motion). . .will generate forces in the temporomandibular joints that. . .are of substantially lower magnitude

than the forces encountered routinely with normal mastication."[21] They also state that the normal motion of chewing produced "greater potential to produce traumatic injury" than whiplash trauma.

Methodologic Error

Inappropriate Study Design. In this article,[21] the authors theorized that extension of the head with the mouth closed would not cause injury to the TMJ. Although this may be true, the most widely accepted and researched model of TMJ injury during whiplash centers around jaw opening during cervical extension, a motion that leaves the TMJ much more susceptible to posterior joint and intra-articular disc injury than when it is closed.[41] The comparison that Howard et al make between the forces acting on the TMJ during whiplash trauma and the normal forces of mastication is fundamentally unsound. The position of the joint at the point of maximum force (closed) as well as the direction (cephalad) and type (compression) of the force during mastication cannot be compared meaningfully with the position of the joint (open) and the direction (posterior) and type (shear) of force during whiplash trauma to the TMJ.

Discussion

The methodologic flaws most frequently found in the reviewed studies were the use of a nonrepresentative study sample (60% of studies), inadequate study size (60%), nonrepresentative crash conditions (50%), and inappropriate study design (45%). Other flaws found were unsupported conclusions (25% of studies), unsubstantiated/unreferenced claims (15%), misquoted literature (5%), improper use of terminology (5%), and misleading illustrations (5%) (Table 2).

All of the articles that had nonrepresentative study samples and crash conditions, inadequate sample size, and other errors resulting in poor internal validity (meaning that bias was present) also had poor external validity (lack of generalizability) as a result. In other words, if the study methods were significantly flawed, the results of the study could not be extrapolated to any population outside the study.

Although the majority of studies that were reviewed for this critique were found to be lacking in study numbers, it is doubtful that any study size or design will define a threshold for whiplash injury, because it is probable that one does not exist. This presumption is based on the confirmed existence of numerous risk factors for whiplash injury that contribute to a highly variable individual susceptibility to injury.

Variables intrinsic to the injured occupant that have been identified as risk factors for injury presence, severity, and duration following whiplash trauma are female gender,[12,33] increased age,[37] preexisting degenerative changes in the spine,[34] out-of-position occupant in the vehicle during impact,[12] rotation of the head during impact,[45] lack of preparation before impact,[23,45] and a *slender physique,*[14] *collectively referred to as* Intrinsic Injury Risk Factors (IIRF) for the purposes of this literature review. Risk factors for injury extrinsic to the occupant are direction of impact,[32,38] presence and position of a head restraint,[14,24] and presence of a shoulder restraint,[31,49] referred to as Extrinsic Injury Risk Factors (EIRF). Acceleration forces interact with the above-mentioned risk factors, as well as Unconfirmed Probable Risk Factors (UPRF), such as car seat construction and bumper dynamics, to produce injury. The number of meaningful permutations of the IIRFs, EIRFs, and UPRFs is conceivably in the thousands or tens of thousands, making volunteer crash testing a highly unlikely study design for delineating an injury threshold for an entire population.

Conclusions

The results of the current literature review and critique suggest that the methodology used by authors attempting to refute the validity of whiplash syndrome is flawed generally. With only a few exceptions, however, the studies reviewed contained other facets that involved relatively sound methods and that contributed to the knowledge base of whiplash injuries and biomechanics. Therefore, it is important to reiterate that the current critique only evaluated study methodology as it related to statements refuting whiplash syndrome.

It may be concluded, as a result of this literature critique, that there is currently no epidemiologic or scientific basis for the following statements:

1. Acute whiplash injuries do not lead to chronic pain.
2. Chronic pain resulting from whiplash injuries is usually psychogenic.
3. Whiplash injuries are unlikely to result in chronic pain in countries where there is no compensation for injury.
4. Rear-impact collisions that do not result in vehicle damage are unlikely to cause injury.
5. Whiplash trauma is biomechanically comparable with common movements of daily living.
6. There is insufficient force generated at the TMJ during whiplash trauma to cause injury.
7. TMJ injuries are not associated with whiplash trauma.
8. There is a direct correlation between vehicle damage and the probability of developing chronic pain after whiplash trauma.
9. Chronic pain following acute whiplash injury is caused or worsened by treatment and diagnostic testing.
10. The risk of chronic neck pain among acutely injured whiplash victims is the same as the prevalence of chronic neck pain in the general population.

As the body of literature about whiplash increases, reports with findings that support one side or another of the legal debate over the validity of whiplash syndrome are increasingly likely to be used in legal settings. Editors

and manuscript reviewers need to be alert for articles on whiplash with flawed methodology or that over-extrapolate their findings. The purpose of the current critique is to provide an overview of some of the weaknesses and the strengths of the literature on whiplash.

References

1. Allen ME, Weir-Jones I, Motiuk DR, et al. Accelartion perturbations of daily living: a comparison to 'whiplash.' Spine 1994;19:1285–90.
2. Andersson HI. The epidemiology of chronic pain in a Swedish rural area. Qual Life Res 1994;1(Suppl):S19–26.
3. Awerbuch MS. Whiplash in Australia: Illness or injury? Med J Aust 1992; 157:193–6.
4. Balla JI. The late whiplash syndrome: A study of an illness in Australia and Singapore. Cult Med Psychiatry 1982;6:191–210.
5. Barnsley L, Lord S, Bogduk N. Clinical review: Whiplash injury. Pain 1994; 58:283–307.
6. Barnsley L, Lord, SM, Wallis SJ, Bogduk N. The prevalence of chronic cervical zygapophysial joint pain after whiplash. Spine 1995;20:20–5.
7. Blincoe LJ. The Economic Cost of Motor Vehicle Crashes, 1994. Washington, DC: US Department of Transportation, NHTSA, 1996.
8. Bogduk N. Post whiplash syndrome. Aust Fam Physician 1994;23:2303–7.
9. Bovim G, Schrader H, Sand T. Neck pain in the general population. Spine 1994;19:1307–9.
10. Brault JR, Wheeler JB, Siegmund GP, Brault EJ. Clinical response of human subjects to rear-end automobile collisions. Arch Phys Med Rehabil 1998;79:72–80.
11. Castro WHM, Schilgen M, Meyer S, Weber M, Peuker C, Wortler K. Do "whiplash injuries" occur in low-speed rear impacts? Eur Spine J 1997;6:366–75.
12. Dolinis J. Risk factors for 'whiplash' in drivers: A cohort study of rear-end traffic crashes. Injury 1997;28:173–9.
13. Ferrari R, Russell AS. The whiplash syndrome: Common sense revisited. J Rheumatol 1997;24:618–22.
14. Freeman MD. A study of chronic neck pain following whiplash injury. Ann Arbor, MI: UMI Dissertation Services, 1998:35–9.
15. Freeman MD, Croft AC, Rossignol AM. "Whiplash Associated Disorders (WAD) - Redefining Whiplash and its Management" by the Quebec Task Force: A Critical Evaluation. Spine 1998;23:1043–9.
16. Freeman MD, Croft AC. Late Whiplash Syndrome (3rd reply). Lancet 1996; 348:125.
17. Gotten N. Survey of one hundred cases of whiplash after settlement of litigation. JAMA 1956;162:865–7.
18. Heise AP, Laskin DM, Gervin AS. Incidence of temporomandibular joint symptoms following whiplash injury. J Oral Maxillofac Surg 1992;50:825–8.
19. Hildingsson C, Toolanen G. Outcome after soft-tissue injury of the cervical spine: A prospective study of 93 car-accident victims. Acta Orthop Scand 1990; 61:357–9.
20. Howard RP, Hatsell CP, Guzman HM. Temporomandibular joint injury potential imposed by the low-velocity extension-flexion maneuver. J Oral Maxillofac Surg 1995;53:256–62.
21. Howard RP, Benedict JV, Raddin JH, Smith HL. Assessing neck extension-flexion as a basis for temporomandibular joint dysfunction. J Oral Maxillofac Surg 1991;49:1210–3.
22. Insurance Research Council. Paying for auto injuries: A consumer panel survey of auto accident victims. Oak Brook, IL: Insurance Research Council, 1994.
23. Jakobsson L, Norin H, Jernström C, et al. Analysis of different head and neck responses in rear-end car collisions using a new human-like mathematical model. Presented at the 1994 International IRCOBI Conference on the Biomechanics of Impacts, Lyon, France, September 21–23, 1994.
24. Lubin S, Sehmer J. Safety—are automobile head restraints used effectively? Survey of parked and moving cars in Vancouver. Can Fam Physician 1993;39: 1584–8.
25. McConnell WE, Howard RP, Poppel JV, et al. Human occupant kinematic response to low speed rear-end impacts. Society of Automotive Engineers Tech Paper Series 940532 1994:23–35.
26. McConnell WE, Howard RP, Guzman HM, et al. Analysis of human test subject kinematic responses to low velocity rear end impacts. Society of Automotive Engineers Tech Paper Series 930889, 1993:21–31.
27. Mertz HJ, Patrick LM. Investigation of the kinematics and kinetics of whiplash. Proceedings of the 11th Stapp Crash Conference. Society of Automotive Engineers paper 670919 1967:175–206.
28. Mertz HJ, Patrick LM. Strength and response of the human neck. Society of Automotive Engineers Transactions 710855 1971;80:2903–28.
29. Mills H, Horne G. Whiplash: Manmade disease. NZ Med J 1986;99:373–4.
30. Norris SH, Watt I. The prognosis of neck injuries resulting from rear-end vehicle collisions. J Bone Joint Surg [Br] 1983;65:608–11.
31. Nygren A. Injuries to car occupants: Some aspects of interior safety of cars. Acta Otolaryngol 1984;(Suppl #394).
32. Olsnes BT. Neurobehavioral findings in whiplash patients with long-lasting symptoms. Acta Neurol Scand 1989;80:584–8.
33. Ommaya A, Backaitis S, Fan W, Partyka S. Automotive neck injuries. Presented at the Ninth International Technical Conference on Experimental Safety Vehicles, US Department of Transportation, National Highway Traffic Safety Administration, Kyoto Japan, November 1–4, 1982:274–8.
34. Parmar HV, Paymakers R. Neck injuries from rear impact road traffic accidents: Prognosis in persons seeking compensation. Injury 1993;24:75–8.
35. Radanov BP, Di Stefano G, Schnidrig A, Sturzenegger M, Augustiny KF. Cognitive functioning after common whiplash. BMJ 1993;307:652–5.
36. Radanov BP, Sturznegger M, Di Stefano G, Schnidrig A, Aljinovic M. Factors influencing recovery from headache after common whiplash. BMJ 1993;307: 652–5.
37. Radanov BP, Sturznegger M. Predicting recovery from common whiplash. Eur Neurol 1996;36:48–51.
38. Radanov BP, Sturznegger M, Stefano GD. Long-term outcome after whiplash injury: A two-year follow-up considering features of injury mechanism and somatic, radiologic, and psychosocial factors. Medicine 1995;74:281–97.
39. Romilly DP, Thomson RW, Navin FPD, Macnabb MJ. Low speed rear impacts and the elastic properties of automobiles. Proceedings of the 12th International Conference of Experimental Safety Vehicles, Gothenburg, Nebraska, 1989;May/June:1–14.
40. Rosenbluth W, Hicks L. Evaluating low-speed rear-end impact severity and resultant occupant stress parameters. J Forensic Sci 1994;39:1393–1424.
41. Schneider K, Zernicke RF, Clark G. Modeling of jaw-head-neck dynamics during whiplash. J Dent Res 1989;68:1360–5.
42. Schrader H, Obelieniene D, Bovim G, et al. Natural evolution of late whiplash syndrome outside the medicolegal context. Lancet 1996;347:1207–11.
43. Siegmund GP, King DJ, Lawrence JM, Wheeler JP, Brault JR, Smith TA. Head/neck kinematic response of human subjects in low-speed rear-end collisions. Society of Automotive Engineers paper 973341 1997:357–85.
44. Spitzer WO, Skovron ML, Salmi LR, et al. Scientific monograph of the Quebec Task Force on Whiplash-Associated Disorders (WAD): Redefining "whiplash" and its management. Spine 1995;20(Suppl):S1–73.
45. Sturzenegger M, Di Stefano G, Radanov B, Schnidrig A. Presenting symptoms and signs after whiplash injury: The influence of accident mechanisms. Neurology 1994;44:668–93.
46. Szabo TJ, Welcher JB, Anderson RD, et al. Human head and neck kinematic response to low speed rear-end impacts. SAE Tech Paper Series 940532, 1994: 23–35.
47. Szabo TJ, Welcher JB. Human subject kinematics and electromyographic activity during low speed rear impacts. Society of Automotive Engineers Tech Paper Series 962432, 1996:295–315.
48. West DH, Goungh JP, Harper TK. Low speed collision testing using human subjects. Accident Reconstruction Journal 1993;5:22–6.
49. Yoganandan N, Haffner M, Maiman DJ, et al. Epidemiology and injury biomechanics of motor vehicle related trauma to the human spine. SAE 892438. In: Proceedings of the 33rd Stapp Car Carsh Conference. Detroit, MI: Society of Automotive Engineers, 1989:223–42.

Address reprint requests to

Dr. Michael D. Freeman
4747 River Road North
Salem, Oregon 97303
E-mail: freemn1@aol.com