Derivation and Validation of a Decision Rule for Predicting Seat Belt Utilization

Michael J. Lichtenstein, MD, Allen Bolton, BSc, and George Wade, MA
Nashville, Tennessee

Information from 3,108 health risk appraisals completed by Tennessee residents in 1986 was used to develop a decision rule for predicting seat belt utilization. The data set was randomly divided into derivation and validation sets. The dependent variable was self-reported seat belt use (percentage). Using multiple linear regression, the following rule was derived: score = [age (years) × 0.24] + [mood-affecting drug use × 4.09] + [miles driven per year × 5.08] + [education level × 11.18] − [race × 18.31] − [cigarette use × 2.73] − [satisfaction with life × 3.50] − [body mass (kg/m²) × 0.83] − [urban/rural residence × 4.08].

Likelihood ratios for persons stating 0 to 25 percent seat belt use were compared with those for persons stating 76 to 100 percent use. The prevalence of 0 to 25 percent seat belt use was 31 percent in the derivation set and 33 percent in the validation set. At the lowest quintile of score (−1 or less), the likelihood ratios were 4.18 and 3.31 in the derivation and validation sets, respectively. At the highest quintile of score (26 or more) the likelihood ratios were 0.29 and 0.38, respectively. At score levels less than 10 the decision rule had a sensitivity of 59 percent and 55 percent and a specificity of 80 percent and 81 percent in the derivation and validation sets, respectively. This decision rule may be used by primary care physicians to identify persons likely not to use seat belts and target them for health promotion efforts.

Injury is the leading cause of death in persons younger than 45 years old in the United States, with the largest proportion of serious injuries arising from motor vehicle accidents.[1] Prevention of motor vehicle injuries depends in part on the universal utilization of seat belts by drivers and passengers.

One strategy for improving seat belt utilization involves persuasion of persons at risk to alter their behavior to increase self-protection.[1] Investigators recently demonstrated that health risk appraisal programs utilized in both worksite and medical settings are clearly able to convince individuals to increase seat belt use.[2] Promoting healthful behavior is one of the missions of primary care. Just as physicians are expected to give cigarette smokers clear messages to stop smoking,[3] so should they be giving advice about seat belt use and traffic injury prevention.

Physicians could target and tailor their injury prevention messages if they could effectively identify persons unlikely to use seat belts. Previous studies have shown that education, race, age, alcohol use, urban or rural differences, socioeconomic status, physical activity, body mass, and driving conditions are all determinants of seat belt use.[4-6] A decision rule could combine a number of these variables and serve as a clinical diagnostic aid. In this study, data from 3,108 health risk appraisals were used to derive and validate a decision rule for predicting seat belt utilization.

Submitted, revised, December 12, 1988.

From the Division of General Internal Medicine, Vanderbilt University Medical Center, the Health Promotion Section, Tennessee Department of Health and Environment, and the Center for Health Statistics, Tennessee Department of Health and Environment, Nashville, Tennessee. Dr. Lichtenstein is a Milbank Memorial Fund Scholar in Clinical Epidemiology. Requests for reprints should be addressed to Dr. Michael J. Lichtenstein, Division of General Internal Medicine, Room 2553, The Vanderbilt Clinic, Vanderbilt University Medical Center, Nashville, TN 37232.

METHODS

Data Collection

The health risk appraisal (HRA) is a health promotion tool developed by the Centers for Disease Control to estimate a person's risk of mortality. First, risk factors are

TABLE 1. HEALTH RISK APPRAISAL QUESTIONS AND SCORING USED IN THE DECISION RULE

Question	Response and Scoring
Age at last birthday (years old)	Enter response
How often do you use drugs or medication which affect your mood or help you to relax?	Almost every day = 1 Sometimes = 2 Rarely or never = 3
Miles per year as a driver of a motor vehicle and/or passenger of an automobile (10,000 = average).	0–10,000 = 1 10,001–20,000 = 2 20,000– = 3
Education—schooling completed (one choice only)	Did *not* graduate from high school = 1 Completed high school = 2 Some college = 3 College or professional degree = 4
Race/origin	White (non-Hispanic) = 1 Black = 2
Tobacco—enter average number of cigarettes smoked per day in the last five years (ex-smokers should use the last five years before quitting)	None = 0 1–10 = 1 11–20 = 2 21– = 3
In general how satisfied are you with your life?	Mostly satisfied = 1 Partly satisfied = 2 Mostly disappointed = 3 Not sure = 4
Height (without shoes) —round up to nearest inch Weight (without shoes) —round up to nearest pound	Height and weight combined and entered as body mass index in kg/m^2
Urban or rural residence	Urban = 1 Rural = 2

assessed by self-report and measurement. Second, the risk factor profile is compared with mortality statistics to arrive at an estimate of risk-age.[7,8] The HRA version used by the Tennessee Department of Health and Environment is a 37-item multiple-choice questionnaire marketed by Planetree Medical Systems. The information from the HRA was used to conduct cross-sectional analyses of the associations between the health information and self-reported seat belt use.

Body mass index (BMI) was calculated as kilogram per meter squared (kg/m^2) from the height and weight reports.[9] Self-reported seat belt use (percentage), age (years), and body mass index were treated as continuous variables. The other variables were considered categorical.

The site of HRA completion was used to determine urban-rural differences in self-reported seat belt use. Urban areas were defined as the Standard Metropolitan Statistical Areas for Memphis, Nashville, Chattanooga, and Knoxville.

TABLE 2. DECISION RULE FOR PREDICTING SEAT BELT USE

SCORE = [Age (years) × 0.24]
+ [Drug use × 4.09]
+ [Miles driven per year × 5.08]
+ [Education level × 11.18]
− [Race × 18.31]
− [Cigarette use × 2.73]
− [Satisfaction with life × 3.50]
− [Body mass index (kg/m^2) × 0.83]
− [Urban/Rural residence × 4.08]

Derivation and Validation of Decision Rule

The data set was randomly divided into derivation and validation sets.[10] In the derivation set stepwise multiple linear regression was used to combine and adjust the variables univariately associated with seat belt use into an equation.[11] Self-reported seat belt utilization (percentage of time used) was the dependent variable. The regression coefficients from the equation were used as the weights for each of the independent variables. This equation was then applied to the derivation and validation sets, and a score for each individual was calculated. Quintiles of scores were then stratified by level of self-reported seat belt use (0 to 25 percent, 26 to 75 percent, 76 to 100 percent). From these distributions, likelihood ratios for each quintile of score were determined.[12] Likelihood ratios express the odds that a given level of multivariate score would be expected in a person who does not wear seat belts compared with a person who wears seat belts all the time.[12]

RESULTS

From January 1, 1986 to December 31, 1986, 3,140 HRAs were administered to individuals in Tennessee by the Tennessee Department of Health and Environment. More than one half of the participants (58.3 percent) completed the HRA as an optional component of a multiphasic screening clinic or a voluntary module of work site wellness programs located in Nashville and offered only to state employees. The remaining participants (41.7 percent) completed HRAs through health promotion programs throughout the state—these respondents consisted of state, public, and private employees. A small percentage (3.2 percent) of respondents were unemployed.

The racial distribution included 2,756 whites and 352 blacks. There were 32 persons of other race/ethnic groups (Hispanic, Asian, Native American) whose responses were excluded from the analysis. The remaining 3,108 participants were randomized into the derivation (1,554 subjects) and validation (1,554 subjects) data sets.

Stepwise multiple linear regression applied to the derivation set identified nine variables that remained asso-

TABLE 3. LIKELIHOOD RATIOS (LR) FOR DECISION RULE SCORE AND PERCENTAGE SEAT BELT UTILIZATION, TENNESSEE EMPLOYEE HEALTH SERVICE

Decision Rule Score	SELF-REPORT SEAT BELT UTILIZATION				
	76–100%	26–75%		0–25%	
	No.	No.	LR (95% CI)	No.	LR (95% CI)
Derivation set					
≤−1	53	96	2.87 (2.01–4.10)	159	4.18 (3.16–5.53)
0–9	81	118	2.29 (1.68–3.12)	123	2.10 (1.55–2.85)
10–18	135	80	0.93 (0.69–1.26)	82	0.84 (0.62–1.13)
19–25	189	64	0.53 (0.39–0.72)	70	0.51 (0.38–0.69)
26–	202	60	0.47 (0.34–0.64)	42	0.29 (0.20–0.41)
Total	660	418		476	
Validation set					
≤−1	57	83	2.32 (1.62–3.32)	146	3.31 (2.39–4.59)
0–9	100	95	1.50 (1.10–2.04)	128	1.65 (1.24–2.20)
10–18	123	89	1.15 (0.85–1.55)	93	0.98 (0.73–1.31)
19–25	170	77	0.72 (0.54–0.97)	77	0.59 (0.44–0.79)
26–	196	63	0.51 (0.37–0.70)	57	0.38 (0.28–0.52)
Total	646	407		501	

CI—confidence interval

ciated with self-reported seat belt use. These HRA items and the scoring for the responses are given in Table 1. The scoring for the responses are the scaled values for each variable taken directly from the HRA and then entered into the equation. The equation resulting from the regression is presented in Table 2. The constants for each variable are the regression coefficients from the equation as applied to the scaled values for each variable.

Table 3 presents likelihood ratios for quintiles of the decision rule scores stratified by level of self-reported seat belt use in both the derivation and validation data sets. The subjects who reported using seat belts 76 to 100 percent of the time were the referent group.

The prevalence of 0 to 25 percent seatbelt use was 31 percent in the derivation set. Across the quintiles of decision rule scores there is a steady gradient in likelihood ratios ranging from 0.29 for scores of 26 or more to 4.18 for scores of −1 or less. For someone with a score of −1 or less, the probability that they utilize seat belts 0 to 25 percent of the time is 65 percent; for scores of 26 or more, the probability is reduced to 11 percent. These results are confirmed in the validation set with the prevalence of 0 to 25 percent seat belt utilization being 33 percent and the likelihood ratio gradient ranging from 0.38 to 3.31. In this set the probabilities that subjects use seat belts 0 to 25 percent of the time is 62 percent and 16 percent for scores of −1 or less and 26 or more, respectively. The sensitivity of a score less than 10 is 59 percent and 55 percent and the specificity is 80 percent and 81 percent in the derivation and validation sets, respectively.

There is also a likelihood ratio gradient for identifying persons who state they use seat belts 26 to 75 percent of the time ranging from 0.47 to 2.87 in the derivation set and 0.51 to 2.32 in the validation set. These gradients are not so steep as those for identifying the 0 to 25 percent seat belt users, reflecting a less extreme mix of the independent variables, but still discriminating this group from the 76 to 100 percent referent seat belt group.

DISCUSSION

These analyses demonstrate that factors associated with seat belt use may be effectively combined into a decision rule for identifying persons who are unlikely to consistently use restraint devices while driving. For each level of the decision rule score, a probability may be calculated that the subject utilizes seat belts 0 to 25 percent or 26 to 75 percent of the time. Based on these estimates, physicians may tailor or target their health promotion messages to persons displaying certain characteristics.

Why not ask patients how often they use seat belts rather than use a mathematical equation that combines nine variables? Self-reported seat belt use, just as self-reported alcohol use, is often misreported. In addition to the decision rule, this paper identifies clinical clues that may alert the physician to identify those who do not use seat belts. For example, the high-school educated, overweight, cigarette-smoking, rural-living patient is less likely to use a seat belt than the college-educated, nonsmoking, average build, urban-living patient. For those wishing to use the rule, it does discriminate reasonably well; comparing scores of −1 or less with scores of 26 or more, the relative risk of using seatbelts 0 to 25 percent of the time is 14.4 (95 percent confidence interval [CI] = 9.1 to 22.7) in the derivation set and 8.7 (95 percent CI = 5.7 to 13.3) in the validation set.

The effectiveness and validity of the decision rule must be interpreted within the limitations of the study. First, the study sample consisted of volunteers who may not be representative of the general Tennessee population. The predicted outcome is a sociologic-behavioral outcome, and the decision rule may not be so robust when applied to a different population of subjects.[10] Even though the decision rule was validated by testing through the split-sample technique, this method of cross-validation will not eliminate effects of biases in subject selection or data collection.[10]

Second, the data were all self-reported. As examples, self-reported seat belt use figures tend to be twice as high as estimates obtained by direct observation.[2,4] Self-reported heights and weights are reliable estimates of measured heights and weights; using self-reports introduces errors that are 1 to 2 percent off the measured values.[13,14] For assessing cigarette use, 15 to 20 percent of claimed nonsmokers may in fact use cigarettes.[15] It was not possible to validate independently measures of height, weight, seat belt use, or the other reported data to assess the effect of these potential biases.

Despite these limitations, the univariate cross-sectional findings were generally consistent with previous work on determinants of seat belt use. Before combining the variables into a decision rule, the analyses were able to confirm prior reports of associations between seat belt use and education, race, age, miles driven per year, cigarette use, body mass, and urban or rural status.[4-6]

The effect of health promotion messages by physicians regarding seat belt use is likely to be small when compared with mandating seat belt use by law. Seat belt use nearly doubles after such a law is passed,[4,16] and though the law's effects may diminish after 12 months, reminder programs may temporarily boost utilization to 80 percent levels periodically.[17] In the absence of such campaigns, seat belt use remains about 40 percent in states with mandatory laws,[17] so there is a need to provide persuasive messages in other settings.

Messages to buckle up delivered by a physician are analogous to smoking cessation advice. Stop-smoking advice by physicians may induce 2 to 10 percent of smokers to quit[18]; though small in relative terms, if all physicians could convince even 2 to 5 percent of their smoking patients to quit, it would produce a large societal benefit. Using HRAs can increase seat belt utilization 5 to 10 percent[2]; if by using HRA-derived data physicians would encourage patients to use restraint devices, this injury prevention message could save a substantial number of lives. It has been estimated that a 1 percent increase in national seat belt use rates would result in a $100 million savings in death and injury costs yearly.[4]

The present study provides a diagnostic aid to help physicians identify and counsel persons who are unlikely to use seat belts. Each variable by itself provides a clue as to whether a person is more or less likely to use a seatbelt, information that the physician may follow up with further questions and counseling. To use the full decision rule efficiently would require using a programmable calculator in the practice setting.

References

1. National Research Council Committee on Trauma Research and the Institute of Medicine: Injury: Magnitude and characteristics of the problem. In Injury in America: A Continuing Public Health Problem. Washington, DC, National Academy Press, 1985, chap 1, pp 18–24
2. Perkins DD, Dunton SM: Health Risk Appraisal and Safety Belt Use. Department of Transportation publication No. HS-807-139. Washington, DC, National Highway Traffic Safety Administration, 1987
3. Health and Public Policy Committee, American College of Physicians: Methods for stopping smoking. Ann Intern Med 1986; 105:281–291
4. Mayas JM, Boyd NK, Collins MA, Harris BI: A Study of Demographic, Situational, and Motivational Factors Affecting Restraint Usage in Automobiles. Department of Transportation publication HS-806-402. Washington, DC, National Highway Traffic Safety Administration, 1983
5. Goldbaum GM, Remington PL, Powell KE, et al: Failure to use seat belts in the United States: The 1981–83 behavioral risk factor surveys. JAMA 1986; 255:2459–2462
6. Merrill BE, Sleet DA: Safety belt use and related health variables in a worksite health promotion program. Health Educ Q 1984; 11: 171–179
7. Wagner EH, Beery WL, Schoenbach VJ, Graham RM: An assessment of health hazard/health risk appraisal. Am J Public Health 1982; 72:347–352
8. Foxman B, Edington DW: The accuracy of health risk appraisal in predicting mortality. Am J Public Health 1987; 77:971–974
9. Revicki DA, Israel RG: Relationship between body mass indices and measures of body adiposity. Am J Public Health 1986; 76: 992–994
10. Wasson JH, Sox HC, Neff RK, Goldman L: Clinical prediction rules: Applications and methodologic standards. N Engl J Med 1985; 313:793–799
11. Armitage P: Statistical Methods in Medical Research. Oxford, Blackwell Scientific Publications, 1971
12. Sackett DL, Haynes RB, Tugwell P: The interpretation of diagnostic data. In Clinical Epidemiology: A Basic Science for Clinical Medicine. Boston/Toronto, Little, Brown, 1985, pp 59–138
13. Stewart AL: The reliability and validity of self-reported weight and height. J Chronic Dis 1982; 35:295–309
14. Palta M, Prineas RJ, Berman R, Hannan P: Comparison of self-reported and measured height and weight. Am J Epidemiol 1982; 115:223–230
15. Jarvis MJ, Tunstall-Pedoe H, Feyerabend C, et al: Comparison of tests used to distinguish smokers from nonsmokers. Am J Public Health 1987; 77:1435–1438
16. Hakkert AS, Zaidel DM, Sarelle E: Patterns of safety belt usage following introduction of a safety belt wearing law. Accident Anal Prev 1981; 13:65–82
17. Williams AF, Preussner DF, Blomberg RD, Lund AK: Seat belt use law enforcement and publicity in Elmira, New York: A reminder campaign. Am J Public Health 1987; 77:1450–1451
18. Folsom AR, Grimm RH: Stop smoking advice by physicians: A feasible approach? Am J Public Health 1987; 77:849–850