

# Development of a Brief Questionnaire to Predict Long-Term Disability

Deborah Fulton-Kehoe, PhD, MPH

Bert D. Stover, PhD

Judith A. Turner, PhD

Lianne Sheppard, PhD

Jeremy V. Gluck, PhD, MPH

Thomas M. Wickizer, PhD

Gary M. Franklin, MD, MPH

**Objective:** To develop a brief worker-completed questionnaire for use soon after a work-related back injury to assess risk of long-term work disability. **Methods:** In this prospective cohort study, 1885 workers provided information about pain, function, psychosocial, and work characteristics about 3 weeks after filing claims for new back injuries. Classification and regression tree analysis was used to identify the best predictive model of work disability status 1 year after claim submission. **Results:** Pain interference with ability to work (>5, 0 to 10 scale), not currently working, and radiating leg pain comprised the best model which accurately classified 77% of the workers. **Conclusions:** A brief questionnaire assessing pain interference with work, current work status, and radiating leg pain might be useful in assessing risk of long-term work disability due to back injury. (J Occup Environ Med. 2008;50:1042–1052)

Back injuries account for a significant proportion of workers' compensation claims and costs. Although most workers recover quickly, a small fraction progress to long-term disability, and these individuals account for a large percentage of medical and disability costs.<sup>1,2</sup> A brief screening instrument that could accurately identify, soon after an injury, those workers at high risk for chronic disability would be quite valuable. For example, workers identified as high risk could be offered early interventions, with the goal of facilitating a successful return to work.

Studies examining predictors of long-term back disability have found that factors within demographic, psychosocial, employment, administrative, clinical, pain, and functional status domains are associated with longer work disability.<sup>3–11</sup> In studies of subjects followed for at least 1 year, the measures most consistently associated with long-term disability are pain and functional status<sup>4,5,7–9</sup> and worker recovery expectations.<sup>3,4,7,8</sup> Nevertheless, predictors have varied across studies,<sup>6,11</sup> likely due at least in part to differences in study populations, duration of follow-up, duration of disability at the time of assessment, outcomes evaluated, and specific domains and factors assessed. Limitations of previous disability screening models include focus on only one domain<sup>12,13</sup> or failure to assess one or more important domains,<sup>9,14,15</sup> lack of worker-reported information,<sup>14</sup> mixture of acute and chronic cases,<sup>10,15</sup> small sample size,<sup>5,10,15,16</sup> short

---

From the Department of Environmental and Occupational Health Sciences (Drs Fulton-Kehoe, Stover, Sheppard, Gluck, and Franklin), University of Washington, Seattle, WA; the Department of Psychiatry and Behavioral Health Sciences and the Department of Rehabilitation Medicine (Dr Turner), University of Washington, Seattle, WA; the Department of Health Services (Dr Wickizer), University of Washington, Seattle, WA; and the Washington State Department of Labor and Industries (Dr Franklin), Olympia, WA.

Address correspondence to: Deborah Fulton-Kehoe, MPH, PhD, School of Public Health and Community Medicine, Department of Environmental and Occupational Health Sciences, Occupational Epidemiology and Health Outcomes Program, 1914 N. 34th Street, #101 Seattle, WA 98103; E-mail: debfk@u.washington.edu

Copyright © 2008 by American College of Occupational and Environmental Medicine

DOI: 10.1097/JOM.0b013e31817d361e

follow-up,<sup>10,13,16–19</sup> and statistical methods that are difficult to interpret.<sup>7,9,20</sup>

Furthermore, some studies have included long functional status or psychosocial questionnaires.<sup>5,7,10,15,16,20</sup> For example, we recently reported that a combination of factors from various domains, assessed soon after a work-related back injury, predicted work disability status 1 year after submission of a lost work-time claim.<sup>21</sup> Nevertheless, a questionnaire with all of these measures, which included the 24-item Roland-Morris Disability Questionnaire (RDQ),<sup>22</sup> would be too long to implement on a routine basis for purposes of screening. A brief (eg, less than 10-item) instrument that could be administered in different formats (eg, paper and pencil, computerized, or interview, either in-person or by telephone) would be most useful. Most measures of functional status such as the RDQ that have been found previously to predict long-term outcomes may be too long to include in a screening instrument. It is possible that individual items from such questionnaires may be good predictors of long-term disability, but few studies have examined this.<sup>4</sup>

The objective of this study was to develop a questionnaire with fewer than 10 items that could be administered 2 to 4 weeks after a work-related back injury to assess risk for long-term work disability. For this purpose, we examined a large number of potential single-item predictors obtained early after submission of a workers' compensation claim for a work-related back injury to determine the optimal combination for predicting work disability status 1 year after claim submission. Potential predictors represented factors from sociodemographic, clinical, pain and function, psychosocial, tobacco and alcohol use, employment, injury severity, and administrative domains.

## Methods

### Study Participants and Procedures

The Washington Workers' Compensation Disability Risk Identification study cohort is a prospective, population-based study designed to identify risk factors for long-term disability due to work-related injuries. Study methods have been previously described.<sup>23–25</sup> In brief, we identified all workers with accepted or provisional Washington State Department of Labor and Industries (DLI) workers' compensation back sprain claims for wage replacement benefits (work disability) between July 2002 and April 2004. The DLI State Fund insures about 70% of non-Federal workers in Washington State. Workers for about 380 large self-insured employers were not included because of insufficient administrative data.

Potential study participants were screened by telephone for study inclusion criteria. Workers were ineligible if they denied having a back injury or were unable to complete the interview in English or Spanish due to language barriers, hearing problems, or illness. Study participants completed a telephone interview that included questions in sociodemographic, clinical, pain and function, psychosocial, tobacco and alcohol use, and employment domains. The Spanish version of the interview included scales from the US Spanish Short Form-36 version 2 (SF-36v2)<sup>26</sup> and other questions were translated by an accredited translator. Medical records of each participant were reviewed by a nurse to confirm that they had a back injury and to assess the severity of the injury.

### Measures

For this analysis, we focused on potential predictive factors available by self-report by the worker. These measures were chosen based on previous research suggesting their importance.<sup>11,24,27</sup> The interview measures and variables are listed in

Table 1 and briefly described below by domain. We have previously described many of these in detail.<sup>23–25</sup>

For use in sensitivity analyses, additional information was obtained from DLI administrative data and from a review of medical records to determine if these factors added to the predictive ability of worker-reported measures.

### Sociodemographic

In the baseline interview, workers reported educational level, race, ethnicity, marital status, household income, and health insurance status. For this study, age, gender, and number of dependents were obtained from DLI.

### Clinical

Clinical questions asked about similar injuries in the past, previous injuries that resulted in more than a month missed from work, number of sick leave days in the past year, comorbidities, height and weight, current health, and health in the year before injury.

### Pain and Function

Workers were administered questions from the Graded Chronic Pain Scale<sup>28,29</sup> assessing pain intensity, pain interference with daily activities, and pain interference with work in the past week. Workers were also asked about location of pain (back, shoulder, etc); whether they had pain, numbness, or tingling that travels down their leg; and if so, whether their low back pain traveled as far as the thigh, stopped at the knee, or traveled below the knee. Participants completed the RDQ<sup>22</sup> and the Physical Functioning (PF) and Role-Physical (RP) scales from the SF-36v2.<sup>26</sup> The SF-36 and RDQ have multiple items and typically summary scores are created. Because our goal was to develop a brief screening questionnaire, we examined individual items rather than summary scores from these scales in most of the analyses.

TABLE 1

Bivariate Associations of Baseline Measures With Work Disability at 1 Yr

Sociodemographic	Accomplished less*	Downhearted and depressed*
Age*	Limited in work or regular activities*	Happy*
Gender	Difficulty performing work or regular activities*	Tobacco and alcohol
Education*	Roland	Tobacco use
Marital status	Stay at home*	AUDIT-C
No. dependents	Change position*	How often do you have a drink containing alcohol?
Race/ethnicity	Walk more slowly*	How many drinks on a typical day?
Income	Not doing jobs around house*	How often six or more on one occasion?
Health insurance*	Use handrail to get upstairs*	Employment
Clinical	Lie down to rest more often*	How often lift 10 pounds
Similar problem in past	Hold on to get out of chair*	Heavy lifting*
Previous injury >1 mo off work*	Get people to do things for me*	Vibration*
No. prior claims*	Get dressed more slowly*	Physical demands
No. sick days in prior yr	Stand up for short periods*	Working very fast*
No. days missed due to back*	Try not to bend or kneel*	Excessive amount of work*
Comorbidities	Difficult to get out of chair*	Enough time to get job done
Body mass index (BMI)	Back is painful most of time*	Job very hectic*
Current health status (other than injury)	Difficult to turn over in bed*	Can take breaks*
Health status in year before injury	Appetite is not very good*	Supervisor listens*
Health care worker talked about RTW*	Trouble putting on socks*	Job satisfaction
Pain and function	Only walk short distances*	Get along with coworkers
Pain intensity in the past week*	Sleep less well*	Full-time work
Pain interference with daily activity*	Help getting dressed*	Seasonal
Pain interference with work*	Sit down most of the day*	Temporary job
Pain in most of your body*	Avoid heavy jobs around house*	Tenure
Pain in leg*	Irritable and bad tempered*	Employer expressed concern
Pain below the knee*	Go upstairs more slowly*	Employer called about RTW* <sup>a</sup>
Change in pain since injury*	Stay in bed most of time*	Employer able to offer modifications*
SF-36 physical function	Psychosocial	Employer offered accommodation*
Limitations in:	Certainty of working in 6 mo*	Work status at time of interview*
Vigorous activity*	Blame for injury	
Moderate activity*	Catastrophizing	
Carrying groceries*	I feel I can't stand it*	
Several flights of stairs*	It is awful, it overwhelms me*	
One flight of stairs*	Want pain to stop*	
Bending, kneeling, stooping*	Fear avoidance	
Walking more than a mile*	Work might harm*	
Walking several hundred yards*	Work makes pain worse*	
Walking 100 yards*	SF-36 Mental Health	
Bathing or dressing yourself*	Very nervous*	
SF-36 Role-physical	Down in the dumps*	
Cut down time spent at work*	Calm and peaceful*	

\* $P < 0.05$ .

RTW, return to work.

## Psychosocial

Worker psychosocial characteristics were assessed using the SF-36v2 Mental Health scale<sup>26</sup> and items from the Pain Catastrophizing Scale<sup>30</sup> and the Fear-Avoidance Beliefs Questionnaire (FABQ).<sup>31</sup> The SF-36 Mental Health scale has demonstrated validity and reliability<sup>26,32</sup>; low scores indicate more psychological distress. Three questions from the Pain Catastrophizing Scale<sup>30</sup> were administered (“I feel I can’t stand it anymore,” “It is awful

and I feel that it overwhelms me,” and “I keep thinking about how badly I want it to stop”); 0 = “not at all” to 4 = “greatly”). The Pain Catastrophizing Scale has been shown to be valid, reliable, and associated with measures of disability.<sup>30,33,34</sup> Two questions from the FABQ work scale were used to assess fear-avoidance beliefs about work (belief that work might increase pain or cause harm, and that work activity should therefore be avoided) (“my work might harm

my back,” “my work makes, or might make, my pain worse”); 0 = “completely disagree” to 6 = “completely agree”).<sup>31</sup> The FABQ has been shown to be reliable and is associated with work loss.<sup>31</sup> The interview also included questions about certainty of working in 6 months (recovery expectations) and perceptions of blame for the injury (self, work factor, someone/something else, no one/nothing) from the Vermont Disability Prediction Questionnaire.<sup>17</sup>

## Health Behaviors

Participants were also asked about tobacco use and were administered the three-item Alcohol Use Disorders Identification Test-Consumption, a measure with demonstrated validity in screening for problematic alcohol use.<sup>35</sup>

## Employment

Workers were asked about job physical demands, job satisfaction, getting along with coworkers, current work status, length of time employed, and whether work at the time of injury was full-time or part-time, seasonal, or temporary.

## Administrative and Medical Records

We also obtained information about the employer size, industry, occupation, and specialty of the first provider seen for the injury from DLI administrative data. Injury severity was assessed from a review of medical records of injury-related visits in the first 8 weeks after the injury. Severity of injury was categorized as minor sprain/strain, major sprain/strain with substantial immobility, evidence of radiculopathy, and reflex/sensory/motor abnormalities.<sup>36</sup> The latter two categories were combined.

## Outcome: Work Disability

Washington State workers insured by the state workers' compensation fund receive wage replacement for temporary total disability ("work disability") after missing at least 4 days from work. Payments for temporary total disability end when a worker has returned to work or has been determined to be able to work. For this analysis, we defined long-term disability as the receipt of work disability payments 1 year after claim submission.

## Statistical Analysis

To determine the combination of items that best predict long-term disability, we conducted a binary

recursive partitioning analysis. Recursive partitioning creates a decision tree by optimally splitting the data into high- and low-risk groups. Using recursive partitioning for prediction models has a number of statistical advantages compared to logistic regression or survival analysis models. Recursive partitioning creates classification trees that can be easier to understand than predictive models based on logistic regression or survival analysis. The probability of the outcome of interest is presented in the tree and does not require additional calculations. In addition, recursive partitioning analyses can include a large number of potential predictors including correlated variables, does not assume a linear relationship, does not require proportional hazards, and can identify important interactions.

We used Classification and Regression Trees (CART) software to create classification trees.<sup>37,38</sup> CART creates trees starting with a "parent" node that is split into "child" nodes by choosing the factor that most accurately separates subjects into high- and low-risk groups. The child nodes may then be split and this process continues until the node has too few cases or further splitting does not improve classification. The trees can then be pruned to obtain a smaller number of terminal nodes. The proportion of cases with the outcome of interest (in this case, long-term disability) is shown in the terminal nodes of the trees.

We set the minimum node size limit at 25 cases. Because the goal was to create a brief screening questionnaire, trees were pruned (if necessary) to the top three to five predictive factors. The sensitivity, specificity, positive predictive value (PPV), and the area under the receiver operating characteristic curve (AUC) were computed for all models. To provide an estimate of the accuracy of the models that would be found in other samples, we used a 10-fold cross-validation procedure to replicate the trees and to estimate the error rates. Cross-

validation compares a 10% sample to a model fit to the remaining 90% of the data 10 times. Error rates across the 10 comparisons are combined to produce an overall error rate that is a reliable estimate of the predictive accuracy of the tree.

To reduce the number of items entered in the initial CART analysis, all potential predictors (shown in Table 1) were screened using logistic regression to determine if they were bivariately associated with long-term disability ( $P < 0.05$ ). The first model includes all items that were significant bivariately, with the exception of the RDQ and SF-36v2 items, which were not included. Next, we included individual items from the RDQ and the SF-36v2 to determine if they added to or replaced any variables in the first CART model. In the third model, current work status was forced to be the first variable. Our goal was to determine whether any additional questions added to the predictive ability of work status alone.

The first three models included only variables available by worker self-report. We then considered models with additional factors available from medical records and administrative files to determine whether such information (which would be more difficult to obtain for the purposes of brief screening) added substantially to the predictive ability of worker-reported measures. From the administrative records, we included worker industry, employer size, specialty of first medical provider seen for injury, number of prior claims in the past 5 years, and unemployment rate in the worker county of residence. We also included the measure of injury severity obtained from medical record reviews.

To assess the stability of the models, we repeated the analyses, varying the minimum child node size (1, 25, 50, 100). We also conducted the CART analyses with the inclusion of all variables in Table 1 to determine if the results differed from those obtained using

only variables that were significant bivariate. In addition, we repeated the CART analyses using the SF-36v2 and RDQ summary scores to determine if a longer screening questionnaire with the full scales would result in better prediction.

## Results

### Sample Characteristics

As described previously,<sup>21</sup> among 4354 claimants identified from DLI administrative records, 2147 (49.3%) enrolled in the study and completed the baseline interview, 1178 (27.1%) could not be contacted, 120 (2.8%) were ineligible, and 909 (20.9%) declined to participate. For the analyses for this report, we excluded workers who did not receive any work disability compensation within the year after injury ( $n = 240$ ), who were missing information on age ( $n = 3$ ), who were not confirmed by medical records review to have a back injury ( $n = 3$ ), and who were hospitalized for their injury ( $n = 16$ ), leaving a sample of 1885 workers. There were small but statistically significant differences between study participants and workers who either declined to participate or who could not be contacted. As compared with non-participants, participants were slightly older (mean age 39.4 versus 38.2 years), were more likely to be female participants (32.0% versus 25.6%), had a longer mean duration of work disability in the year following injury (83 versus 68 days), and were more likely to be receiving work disability compensation at 1 year (13.8% versus 11.3%). Baseline interviews were completed a median of 18 days after filing a claim [mean (standard deviation), 22 (10) days].

Study participants ranged in age from 18 to 77 years, with a mean of 39 years. Most were non-Hispanic White (70%), 68% were male participants, 52% were married, and 40% reported household incomes of less than \$30,000 per year. Over half of the participants worked in construc-

tion, manufacturing, or transportation industries, and 27% had been employed at the job of injury for less than 6 months. Twenty-one percent worked for firms with more than 200 employees, whereas 37% worked in companies with fewer than 25 employees. Among subjects in this study (workers with back injuries with some wage replacement benefits), 14% were receiving work disability benefits 1 year after claim submission.

### Bivariate Associations

The worker-reported variables bivariate ( $P < 0.05$ ) associated with disability status at 1 year are shown in Table 1. In the sociodemographic domain, these were age, education, and health insurance status. Among clinical variables, the number of prior workers' compensation claims and previous injuries that resulted in at least a month missed from work were associated with 1-year disability. Long-term disability was associated with all baseline interview measures of pain. In addition, all individual questions from the SF-36v2 PF and RP scales and the RDQ were significantly associated with long-term disability. Recovery expectations and all individual items from psychosocial scales were associated with disability status. Neither tobacco nor alcohol use was associated with disability at 1 year. Many of the items measuring physical demands of the job were associated with long-term disability, but job satisfaction and getting along with coworkers were not.

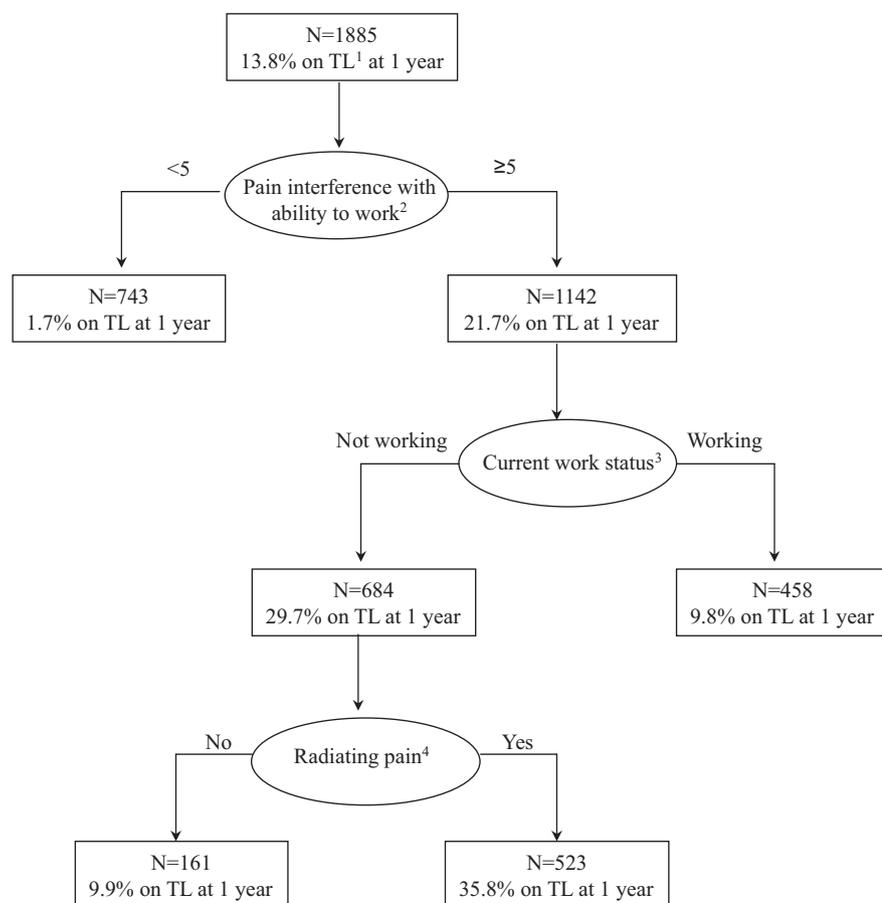
### CART Results: Primary Model

The primary CART tree (from the analysis using only variables bivariate associated with disability, not including the SF-36 and RDQ items) is shown in Fig. 1. Baseline pain interference with ability to work; current work status; and pain, numbness, or tingling in the leg were chosen as the top three variables. Subjects who rated pain interference with the ability to work as 4 or lower

on a scale of 0 (no interference) to 10 (unable to carry on any activities) had a very low risk of long-term disability (<2%). Among subjects who rated interference with work as 5 or higher, about 22% developed long-term disability. Among subjects who rated pain interference with work as 5 or higher, were not working at the time of interview, and reported symptoms radiating in the leg, 36% were receiving work disability payments 1 year later. This model accurately classified the 1-year disability status of 77% of the workers. The sensitivity was 72%, specificity was 78%, and the PPV was 35%. The AUC was 0.80 in the learning samples and 0.79 in the test samples.

### CART Results: Including SF-36 and RDQ

The next CART analysis included individual items from the SF-36v2 and RDQ scales. The default tree produced by CART included 15 terminal nodes. The model pruned to the top four variables is shown in Fig. 2. In this classification tree, "being limited in the kind of work or regular daily activities you do" (an item from the SF-36v2 RP scale) was chosen as the first node. Similar to the first tree, current work status and symptoms radiating to the leg were the second and third items chosen. Among subjects who reported being limited in work/daily activities all or most of the time, were not working at the time of the interview, and had radiating leg symptoms, 38% were receiving disability payments 1 year later. In this model, there was a fourth split for those who reported no radiating symptoms. Thirty-six percent of workers who were limited in work/daily activities all or most of the time, were not working at the time of interview, and did not have radiating pain but indicated a low level of certainty about working in 6 months were disabled at 1 year. However, this last node was based on a small number of cases. For the



**Fig. 1.** Prediction of work disability 1 year after injury, excluding SF-36 and RDQ items. <sup>1</sup>On TL = on time loss (receiving wage replacement benefits), <sup>2</sup>0 = no interference to 10 = unable to carry on any activities, <sup>3</sup>approximately 3 weeks after injury, and <sup>4</sup>pain, numbness, or tingling that travels down your leg.

pruned model, the AUC was 0.79 and 0.75 in the learning and test samples, respectively. This model accurately classified 71% of the subjects, with a sensitivity of 75%, a specificity of 71%, and a PPV of 29%.

### CART Results: Work Status Forced in First

Because injured workers returning for follow-up care are often asked about work status, we considered a model that forced current work status to be the first variable (Fig. 3). In this model, radiating symptoms was the only other variable selected. Fewer than 5% of subjects who were working at the time of the baseline interview were receiving wage replacement benefits a year later. Among those who were

not working at interview, 26% were receiving such benefits. When those not working were further divided on the basis of radiating symptoms, 33% of those with radiating symptoms and 8% without were receiving work disability compensation at 1 year. The AUC was 0.76 in the learning sample and 0.74 in the test sample. The sensitivity was 72%, specificity was 75%, PPV was 32%, and 75% were correctly classified.

### CART Results: Additional Items

Models including items obtained from administrative files and medical records were virtually identical to the models that included only self-reported items. Industry, occupation, specialty of initial treating clinician, unemployment rate, and history of prior claims were not chosen as

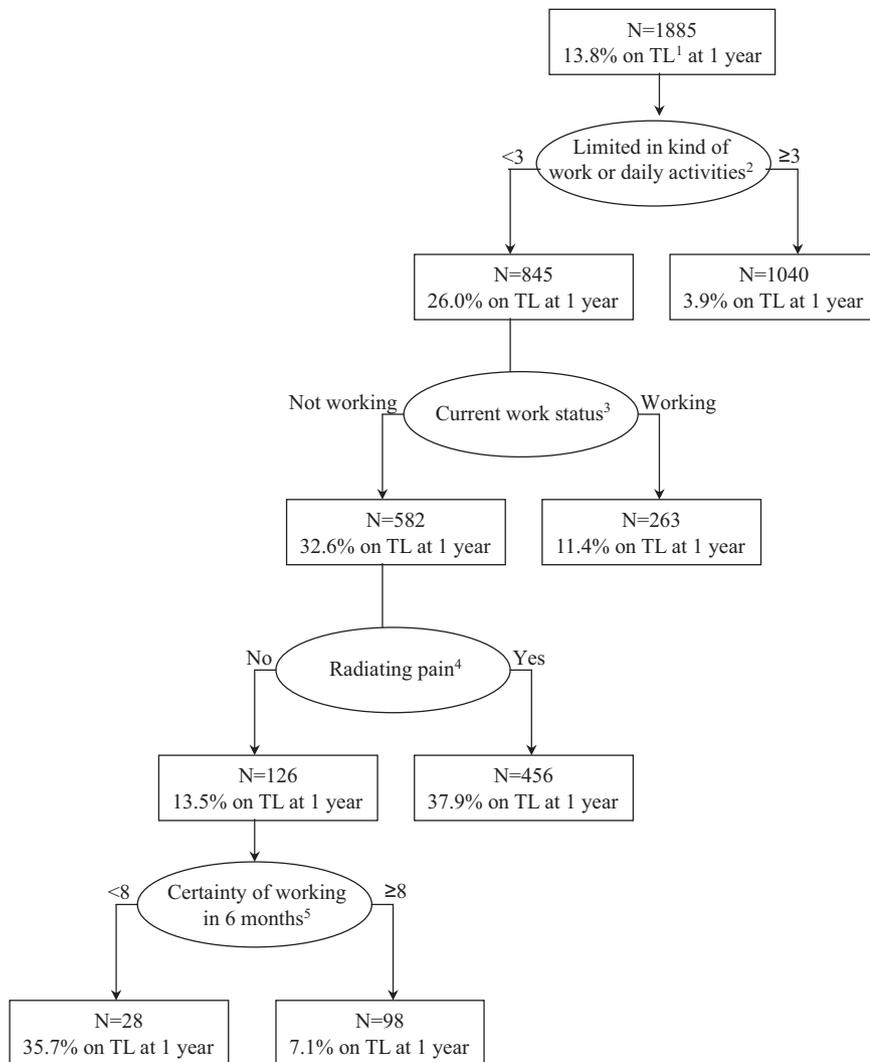
strong predictors of long-term disability. In addition to the self-report of symptoms traveling down the leg, we also examined the predictive ability of the injury severity rating, which included a category for evidence of radiculopathy. Self-report of any radiating leg symptoms was consistently chosen by the CART analyses to predict long-term disability, rather than self-report of pain specifically radiating below the knee or the injury severity rating of clinical evidence of radiculopathy based on review of medical records.

### CART Results: With SF-36 and RDQ Scale Scores

When functional status summary scores were included in the CART models, the RDQ and SF-36v2 RP and PF scales were the top three variables chosen (data not shown). The ability of this model to predict long-term disability was not improved as compared to the model with only three questions (shown in Fig. 1). Among subjects with high (worse) RDQ scores (>15.5) and low (worse) SF-36v2 PF scores (<29.1), 34% were receiving work disability payments at 1 year. The percent correctly classified (69%) and the AUC (0.81 in the learning sample and 0.75 in the test samples) were the same or slightly lower in this model than in the model shown in Fig. 1.

### Sensitivity Analyses

As a sensitivity analysis to examine the consistency of variables selected, we varied the minimum size of the child nodes (1, 25, 50, 100). In almost all cases, the top three variables chosen were the same, but there was substantial variation in the additional variables chosen (data not shown). With a minimum node size of 100, simple trees with only 2 to 4 variables were chosen. More complex trees were created with smaller node sizes. In addition, we conducted CART analyses including all variables in Table 1 regardless of bivariate as-



**Fig. 2.** Prediction of work disability 1 year after injury, including SF-36 and RDQ items. <sup>1</sup>On TL = on time loss (receiving wage replacement benefits), <sup>2</sup>1 = all of the time to 5 = none of the time, and <sup>3</sup>approximately 3 weeks after injury, <sup>4</sup>pain, numbness, or tingling that travels down your leg, <sup>5</sup>0 = not at all certain to 10 = extremely certain.

sociations with 1-year disability. For the model in Fig. 1, the tree was identical. For the model that included questions from the SF-36v2 and RDQ, the regression tree was identical to the pruned tree presented in Fig. 2.

### A Brief Questionnaire Scored by Summing Positive Responses to Screening Items

The tree shown in Fig. 1 was slightly more accurate than the other models we examined in predicting 1-year work disability. A brief screening questionnaire using the items in Fig. 1 is shown in Table 2.

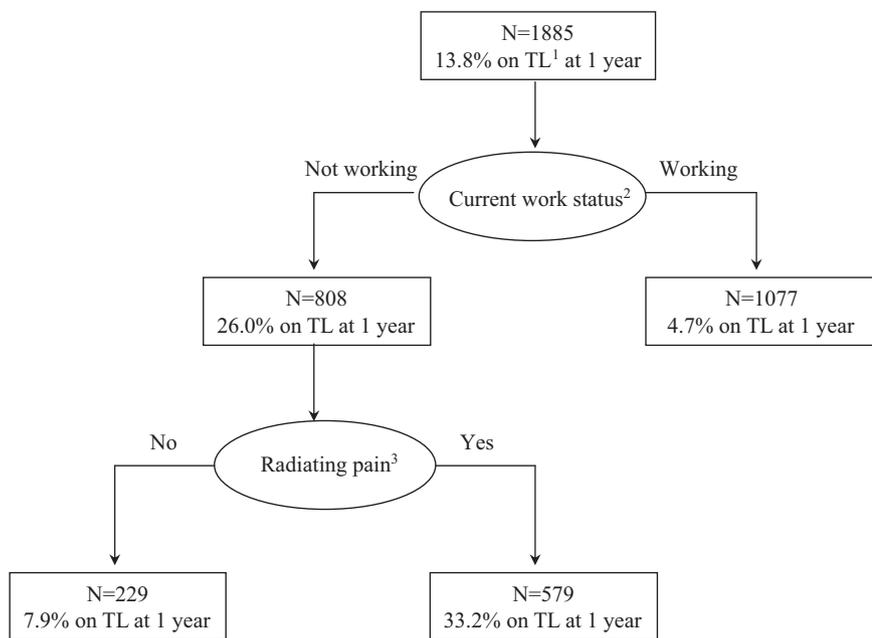
Because the order of items chosen in the CART models may differ in other study populations, another method of calculating the risk is to sum the number of positive responses to these three questions. In Table 3, we present the observed disability outcomes by the number of positive responses to Fig. 1 questions. Workers who were positive on all three risk factors shown in Fig. 1 were much more likely to be disabled at 1 year. Their mean duration of disability was 200 days of wage replacement compensation in the first year after claim submission (median 184 days) and 36% were

receiving wage replacement compensation at 1 year. Among workers with two of these risk factors, the mean number of days of wage replacement was 72 (median 23 days) and 11% were receiving compensation at 1 year. Workers who had no positive responses to these questions were at very low risk of long-term disability, with a mean of 9 days of wage replacement (median 5) and only 0.2% receiving compensation at 1 year.

### Discussion

Our goal was to develop a brief screening questionnaire that could be used a few weeks after a work-related back injury to identify workers at high risk of developing long-term disability. We found that the combination of self-reported pain interference with work, current work status, and radiating symptoms, when obtained about 3 weeks after filing a claim, was consistently predictive of long-term disability. Workers who reported a low level of pain interference with work (less than or equal to 4 on a 0 to 10 scale) had a very low risk (<2%) of being on work disability 1 year after injury. In contrast, over 35% of workers who reported the combination of pain interference with work ( $\geq 5$ ), not working at the time of the baseline interview, and radiating symptoms were still receiving work disability compensation 1 year after injury.

Of all the questions that we examined, pain interference with work had the best ability to discriminate between those at higher versus lower risk of long-term disability. This question may have been most predictive because it is strongly associated with a number of other factors known to be associated with long-term disability such as self-reported pain, pain radiating to the leg, recovery expectations, and fear-avoidance, as well as individual questions and summary scores from the RDQ and SF-36 measures of functional status. In this



**Fig. 3.** Prediction of work disability 1 year after injury, work status at baseline interview entered as first variable. <sup>1</sup>On TL = on time loss (receiving wage replacement benefits), <sup>2</sup>approximately 3 weeks after injury, and <sup>3</sup>pain, numbness, or tingling that travels down your leg.

**TABLE 2**  
Brief Screening Questionnaire

Questions	Scoring
In the past week, how much has pain interfered with your ability to work, including housework, rated on a scale from 0 to 10 when 0 is “no interference” and 10 is “unable to carry on any activities?”*	0 if rating is <5 and 1 if GE 5
During the past week have you worked for pay?	0 if yes and 1 if no
Do you have any pain, numbness, or tingling that travels down your leg?	0 if no and 1 if yes

\*Question is from the graded chronic pain scale.<sup>19,20</sup>

sample of workers with back injuries, pain interference with ability to work appears to be an efficient composite measure of both the degree of pain and the ability to function with pain.

Similar to other studies of work-related back pain,<sup>4,9,10</sup> we also found that radiating pain in the leg is an important predictor of long-term disability. However, radiating pain was not in the final model in one large study.<sup>7</sup> This difference could be because of using a different measure of radiating pain (pain below the knee instead of any radiating pain in the leg), different sample and statistical

methods, or the inclusion of a different set of potential predictor variables, including the functional status scales, which may capture the impact of radiating pain. In contrast to previous studies,<sup>4,7,23</sup> low expectation of return to work was not chosen as one of the top predictors in our analysis. Nevertheless, the worker recovery expectation rating was associated with the items that were chosen, including pain interference with work, radiating pain, and work status at the time of interview. It appears likely that because of these associations, the recovery expectation rating did not add significantly to the pre-

dictive ability of these important baseline variables.

Functional status measures assessed early in a claim are predictive of long-term disability.<sup>7,9-11</sup> The RDQ and SF-36 are both fairly long and given the constraints in the number of items to include in a brief questionnaire, summary scores from these measures were not included in the primary models. We did, however, examine individual items from these scores. When questions from the SF-36 and RDQ were included in the CART analysis, being limited in the kind of work or daily activity was the first item chosen. As compared with the rating of pain interference with work, rating of limitations in kind of work or other daily activities had a similar, although slightly less accurate, ability to classify workers according to their risk of long-term disability. We also examined predictive models that included the SF-36 and RDQ summary scores. The RDQ, SF-36 Role-Physical, and SF-36 Physical Function scales were chosen as the best predictors. Nevertheless, this model, which would require asking 35 more questions, did not improve predictive ability as compared with the model in Fig. 1. In fact, based on the three questions in Fig. 1, the percent correctly classified and the cross-validated AUC were slightly higher than for the model that included the functional status summary scores.

Our goal was to identify fewer than 10 questions that could accurately identify individuals with back injuries at high risk of developing long-term work disability. Our primary model (Fig. 1) accurately classified 77% of the recently injured workers in our sample. The models had a fairly high sensitivity (72%), but also a large proportion of false positives (PPV of 35%). CART trees with a larger number of variables improved the AUC only slightly. The accuracy of our models is comparable to that of other studies that have

**TABLE 3**  
Disability Outcomes by Number of Risk Factors

No. Positive Responses	No. Cases	Mean Duration of Wage Replacement in First Year (d)	Median Duration of Wage Replacement in First Year (d)	Interquartile Range	Percent Receiving Wage Replacement at 1 yr	Interpretation of Risk of Chronic Disability
0	432	9.3	5.0	2.5–10.0	0.2	Very low
1	429	30.3	8.0	3.0–19.0	4.4	Very low
2	505	71.9	23.0	6.0–83.5	11.3	Moderate
3	519	200.2	184.0	59.0–365.0	35.5	High

reported sensitivity and specificity or classification error rates predicting long-term disability.<sup>4,9,10,16</sup>

The current study has several strengths. This is the largest prospective study conducted to date examining predictors of long-term disability among persons with back injuries. Few previous studies have had over 1000 subjects.<sup>4,9</sup> Other strengths include examination of a comprehensive list of potential predictive factors from multiple domains including clinical, socio-demographic, pain and function, psychosocial, and employment domains, and complete 1-year follow-up information from objective administrative data on work disability status. Although we did not conduct physical examinations, all injuries were confirmed by interview and medical record review, and injury severity (including evidence of radiculopathy) was rated based on medical records of physical examinations.

The use of a regression tree analysis for prediction models is another advantage of this study over previous research in this area. The results are intuitive and easy to convert to a brief questionnaire. Regression trees can find the best cut-points on continuous variables and can identify important interactions. Despite these strengths we also offer some cautions. Some of the default trees presented by CART can be fairly complex. We chose to present pruned trees when this occurred because there was little difference in the accuracy of predicting long-term disability with

additional variables and the top three or four variables appeared to be more stable than variables entered subsequently. We suggest that CART analyses be conducted varying the node size criteria. If models vary considerably, only variables that are consistently chosen should be presented. Results are more likely to be replicated by setting a fairly high minimum node size and pruning complex trees to the top predictors.

We note a few study limitations. About 50% of identified cases completed the baseline interview. Although this is a relatively high response rate to a telephone survey, the nonresponse rate raises the question of possible sample bias. We should also emphasize that the study population was Washington State workers who had filed claims for work loss involving at least 4 days. The results may not be generalizable to workers who have not missed this amount of work or to workers covered by compensation systems dissimilar to the Washington State system. Finally, the results may not be generalizable to workers assessed either immediately after injury (for example, at the first medical visit) or workers assessed more than a month after claim filing. Nevertheless, a brief screening questionnaire might be most useful a few weeks after claim filing, when many back pain problems will have resolved (so the number of cases to be screened and the number of unnecessary screenings will be reduced) but in time for an intervention to be effective,

before pain and disability become more refractory. The questions have not yet been examined prospectively for their ability to predict disability in a new sample; we plan to conduct such a study in the future.

## Conclusions

After identifying workers who are at high risk for long-term disability, health care providers will likely need to use a variety of strategies to improve return to work outcomes. It is of interest that pain interference with work, regardless of how asked (either by a rating of degree of pain interference with work, by a rating of the degree of limitations in work or daily activities, or by simply asking whether the worker was currently working), was a more important predictor of long-term disability than was the worker's rating of pain intensity. This suggests the potential importance of interventions targeted specifically at improving the worker's ability to accomplish his or her work roles and tasks, rather than targeted solely on pain reduction. Such interventions might include telephone calls to employers to encourage job modifications such as light duty or reduced hours to enable early return to work and graded exercise programs focused on the performance of specific job tasks.

We identified a brief list of single-item questions that accurately discriminated workers on versus off work disability compensation 1 year later. The questionnaire can be

quickly administered in a primary care provider's office and does not require any complex scoring. Because such an instrument could potentially be viewed as negatively labeling an injured worker, actual implementation in clinical settings needs to occur with careful planning and in conjunction with interventions to address the identified problems.

## Acknowledgments

The authors thank Kathleen Egan for project coordination and Rae Wu for statistical analysis.

Sponsored by Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, Grant number 1 R01 OH04069.

## References

- Hashemi L, Webster BS, Clancy EA. Trends in disability duration and cost of workers' compensation low back pain claims (1988–1996). *J Occup Environ Med*. 1998;40:1110–1119.
- Hashemi L, Webster BS, Clancy EA, Volinn E. Length of disability and cost of workers' compensation low back pain claims. *J Occup Environ Med*. 1997;39:937–945.
- Cole DC, Mondloch MV, Hogg-Johnson S. Listening to injured workers: how recovery expectations predict outcomes—a prospective study. *Can Med Assoc J*. 2002;166:749–754.
- Dionne CE, Bourbonnais R, Fremont P, Rossignol M, Stock SR, Larocque I. A clinical return-to-work rule for patients with back pain. *Can Med Assoc J*. 2005;172:1559–1567.
- Enthoven P, Skargren E, Carstensen J, Oberg B. Predictive factors for 1-year and 5-year outcome for disability in a working population of patients with low back pain treated in primary care. *Pain*. 2006;122:137–144.
- Hilfiker R, Bachmann LM, Heitz CA-M, Lorenz T, Joronen H, Klipstein A. Value of predictive instruments to determine persisting restriction of function in patients with subacute non-specific low back pain. Systematic review. *Eur Spine J*. 2007;16:1755–1775.
- Hogg-Johnson S, Cole DC. Early prognostic factors for duration on temporary total benefits in the first year among workers with compensated occupational soft tissue injuries. *Occup Environ Med*. 2003;60:244–253.
- Lotter F, Burdorf A. Prognostic factors for duration of sickness absence due to musculoskeletal disorders. *Clin J Pain*. 2006;22:212–221.
- McIntosh G, Frank J, Hogg-Johnson S, Bombardier C, Hall H. Prognostic factors for time receiving workers' compensation benefits in a cohort of patients with low back pain. *Spine*. 2000;25:147–157.
- Schultz IZ, Crook JM, Berkowitz J, et al. Biopsychosocial multivariate predictive model of occupational low back disability. *Spine*. 2002;27:2720–2725.
- Turner JA, Franklin G, Turk DC. Predictors of chronic disability in injured workers: a systematic literature synthesis. *Am J Ind Med*. 2000;38:707–722.
- Boersma K, Linton SJ. Screening to identify patients at risk: profiles of psychological risk factors for early intervention. *Clin J Pain*. 2005;21:38–43.
- Linton SJ, Boersma K. Early identification of patients at risk of developing a persistent back problem: the predictive validity of the Orebro Musculoskeletal Pain Questionnaire. *Clin J Pain*. 2003;19:80–86.
- Pransky GS, Verma SK, Okurowski L, Webster B. Length of disability prognosis in acute occupational low back pain: development and testing of a practical approach. *Spine*. 2006;31:690–697.
- Schultz IZ, Crook J, Meloche GR, et al. Psychosocial factors predictive of occupational low back disability: towards development of a return-to-work model. *Pain*. 2004;107:77–85.
- Schultz IZ, Crook J, Berkowitz J, Milner R, Meloche GR. Predicting return to work after low back injury using the psychosocial risk for occupational disability instrument: a validation study. *J Occup Rehabil*. 2005;15:365–376.
- Hazard RG, Haugh LD, Reid S, Preble JB, MacDonald L. Early prediction of chronic disability after occupational low back injury. *Spine*. 1996;21:945–951.
- Margison DA, French DJ. Predicting treatment failure in the subacute injury phase using the Orebro Musculoskeletal Pain Questionnaire: an observational prospective study in a workers' compensation system. *J Occup Environ Med*. 2007;49:59–67.
- Shaw WS, Pransky G, Patterson W, Winters T. Early disability risk factors for low back pain assessed at outpatient occupational health clinics. *Spine*. 2005;30:572–580.
- Baldwin ML, Butler RJ, Johnson WG, Cote P. Self-reported severity measures as predictors of return-to-work outcomes in occupational back pain. *J Occup Rehabil*. 2007;17:685–700.
- Turner JA, Franklin G, Fulton-Kehoe D, et al. Early predictors of chronic work disability: a prospective, population-based study of workers with back injuries. *Spine*. 2008;33:199–204.
- Roland M, Morris R. A study of the natural history of back pain. Part I: Development of a reliable and sensitive measure of disability in low-back pain. *Spine*. 1983;8:141–144.
- Turner JA, Franklin G, Fulton-Kehoe D, et al. Worker recovery expectations and fear-avoidance predict work disability in a population-based workers' compensation back pain sample. *Spine*. 2006;31:682–689.
- Turner JA, Franklin G, Fulton-Kehoe D, et al. Prediction of chronic disability in work-related musculoskeletal disorders: a prospective, population-based study. *BMC Musculoskelet Disord*. 2004;5:14.
- Turner JA, Franklin G, Fulton-Kehoe D, et al. Early predictors of chronic work disability associated with carpal tunnel syndrome: a longitudinal workers' compensation cohort study. *Am J Ind Med*. 2007;50:489–500.
- Ware JE, Kosinski M, Dewey JE. *How to Score Version 2 of the SF-36 Health Survey*. Lincoln, RI: QualityMetric; 2000.
- Krause N, Frank JW, Dasinger LK, Sullivan TJ, Sinclair SJ. Determinants of duration of disability and return-to-work after work-related injury and illness: challenges for future research. *Am J Ind Med*. 2001;40:464–484.
- Von Korff M. Epidemiological and survey methods: assessment of chronic pain. In: Turk DC, Melzack R, eds. *Handbook of Pain Assessment*. 2nd ed. New York: The Guilford Press; 2001:603–618.
- Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. *Pain*. 1992;50:133–149.
- Sullivan MJL, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. *Psychol Assess*. 1995;7:524–532.
- Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A fear-avoidance beliefs questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain*. 1993;52:157–168.
- McHorney CA, Ware JE Jr, Rachel Lu JF, Donald Sherbourne C. The MOS 36-item short-form health survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care*. 1994;32:40–66.

33. Osman A, Barrios FX, Gutierrez PM, Kopper BA, Merrifield T, Grittmann L. The pain catastrophizing scale: further psychometric evaluation with adult samples. *J Behav Med.* 2000;23:351–365.
34. Sullivan MJL, Thorn B, Haythornthwaite JA, et al. Theoretical perspectives on the relation between catastrophizing and pain. *Clin J Pain.* 2001;17:52–64.
35. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol use disorders identification test. *Arch Intern Med.* 1998;158:1789–1795.
36. Stover B, Wickizer TM, Zimmerman F, Fulton-Kehoe D, Franklin G. Prognostic factors of long-term disability in a workers' compensation system. *J Occup Environ Med.* 2007;49:31–40.
37. Breiman L, Friedman J, Olshen R, Stone C. *Classification and Regression Trees.* Pacific Grove: Wadsworth; 1984.
38. Steinberg D, Colla P. *CART: Tree-Structured Non-Parametric Data Analysis.* San Diego, CA: Salford Systems; 1995.