Journal of Occupational Rehabilitation, Vol. 10, No. 2, 2000

Low Back Pain Prognosis: Structured Review of the Literature

Greg McIntosh,[1,5] John Frank,[2,3] Sheilah Hogg-Johnson,[2,3] Hamilton Hall,[1] and Claire Bombardier[4]

The purpose of this paper was to review and appraise pertinent articles to gain a better understanding of critical methodological issues necessary to properly design a high-quality back pain prognosis study. The review concentrated on back pain prognosis studies with epidemiologically sound designs focusing on work-disability outcomes and utilizing survival analytic methods. Nine papers were reviewed. There were few well-designed studies that achieved good scientific quality with minimal flaws. The outcomes were well defined in each paper. The age and sex characteristics of the cohorts were described in six papers and an adequate description of the study site occurred in five papers. All papers employed suitable mathematical/statistical techniques, but only one paper discussed accuracy and predictive value. No paper addressed the issue of reproducibility of the predictor variables or the final model. Most papers derived models that were clinically sensible, and the ease of use for clinicians was high. A recommended course of action for use by future patients/therapists in prognostication was rarely documented. To date, prognosis has been an inadequately studied aspect of the continuum from back injury to recovery. Researchers and clinicians interested in prognosis research need to overcome the limitations of past designs and address the methodological guidelines outlined to improve the quality of future prognosis studies.

KEY WORDS: back pain; prognosis; survival analysis; literature review.

INTRODUCTION

Although medical knowledge and understanding of disease processes continue to advance, low-back pain—despite its prevalence and high socioeconomic impact—remains enigmatic. As the cost of compensated occupational low-back pain escalates, any predictor of outcome is advantageous (1).

[1]Canadian Back Institute Research Department, Toronto, Ontario, Canada.
[2]Department of Public Health Sciences, University of Toronto, Toronto, Ontario, Canada.
[3]Institute for Work & Health, Toronto, Ontario, Canada.
[4]Clinical Epidemiology and Health Care Research Program, University of Toronto, Toronto, Ontario, Canada.
[5]Correspondence should be directed to Greg McIntosh, 153 Oliver Place, Oakville, Ontario, Canada L6H 1K9; e-mail: gmcintosh@cbi.ca.

1053-0487/00/0600-0101$18.00/0

Unfortunately, prognosis is a poorly studied area in the back pain literature. Most prognosis studies to date lack sensitive and specific predictors of clinically useful prognostic validity for the early targeting of high-risk patients in need of intensive rehabilitation (2).

The progress of low-back pain is highly variable (2,3) and it will never be possible to eliminate all the inherent uncertainty in predicting outcome, but an awareness of significant contributing physical and nonphysical factors can help to forecast its course more accurately.

The purpose of this paper was to review and appraise pertinent articles to gain a better understanding of critical methodological issues necessary to properly design a prognosis study of back pain.

The review of medical and economic literature concentrated on back pain prognosis studies with epidemiologically sound designs focusing on work-disability outcomes and utilizing survival analytic methods. Survival analysis text words were utilized in the search because these statistical methods are most appropriate for time-to-event analyses where the outcome event is return to work.

METHODS

When evaluating articles on prognosis of disease, research papers should meet certain criteria (4) as specified in Table I.

Of particular importance in an initial screen of articles is the use of an inception cohort (sampling strategy), assessment of an objective outcome, and the use of multivariable time-to-event statistical methods. An inception cohort design is essential so that subjects are all at a similar well-defined point in the course of their recovery (5). The outcome variable should help capture the episodic and recurrent nature of back pain; if return to work is used, then it must be evaluated over a time period that permits multiple spells of possible work disability (6). Survival analysis accounts for multivariable statistical modeling, time-varying effects, differential follow-up times, and censoring. Survival models are preferred over logistic models because survival analytic methods use more information—the survival times—than the logistic model does, which considers a binary-type outcome at one point in time and does not account for time to an event (7). These are critical analytical concepts because of the multifactorial nature of back pain prognosis (5,8–10).

A MEDLINE search of the relatively current health literature was conducted in April 2000. A query of the electronic database from 1985 to 2000 revealed more than 144,000 prognosis studies, more than 7,300 on back pain, and almost 14,600 involving survival analysis.

Combining these three search strategies led to the identification of 12 articles (11–22). Appendix A documents the electronic search strategy. Titles and abstracts were assessed to

Table I. Six Guidelines for Evaluating Articles About Clinical Course and Prognosis of Disease (4)

1. Was an inception cohort assembled?
2. Was the referral pattern described?
3. Was complete follow-up achieved?
4. Were objective outcome criteria developed and used?
5. Was the outcome assessment "blind"?
6. Was adjustment for extraneous prognostic factors carried out?

determine the appropriateness of each article. One of the 12 articles focused on rheumatoid arthritis (11), two on pancreatic cancer (12,13), and one on HIV infection (14) and were not reviewed.

Full texts of the remaining articles were retrieved. Review of these papers led to the exclusion of four more articles: one was not an inception cohort (15), two were primarily treatment studies (16,18), and one did not focus primarily on back pain (17). Only four papers remained (19–22), but three were added from the lead author's files that met the inclusion criteria: two did not use the specific text words in the title or abstract (23,24), and one published early in 2000 (1) that was not yet listed on MEDLINE.

The coauthors, local clinicians, and researchers proficient with current back pain literature were consulted for possible articles that may have been missed. This led to the inclusion of a paper (25) not included in MEDLINE and three from the economic literature (26–28); full text review of these four papers revealed that one of the economic studies did not include low-back pain patients (26) and one (27) was not an inception cohort, and these were not reviewed. Two well-conducted prognosis papers (9,10) were not included in this review because they are unpublished; a review of these papers is found elsewhere (29).

This led to a total of nine reviewed back pain prognosis papers (1,19–25,28) that utilized an inception cohort and survival analytic methods. Table II summarizes the methods and results of these nine papers. In true studies of survival, where the outcome is death, failure is a negative experience, but in this review, the fewer the days (time), the more "positive" is the event in terms of health; thus, hazard rate ratios (HRR) less than 1 indicate an increased risk of time on to the events and suggests possible longer time to recovery. Although this contrasts the usual terminology, the method of survival analysis and expressing HRR is consistent across all the back pain prognosis studies that were reviewed. Appendix B provides a glossary of terms common to the reviewed papers.

Methodological Standards

The nine selected articles were scrutinized more closely using more recently published methodological standards for evaluating clinical prediction papers. Wasson *et al.* (30) and Laupacis *et al.* (31) propose nine key factors when critically appraising clinical prediction papers.

1. *Outcome.* The outcome should be clearly defined and clinically important, and the presence or absence of the outcome determined without knowledge of the status of the predictor variables.
2. *Predictive variables.* The predictors should have a clear, clinically sensible, and reproducible definition. Ideally, the predictors should be collected prospectively, using a data sheet specifically developed for the study, and assessors trained in data collection.
3. *Patient characteristics.* The minimum requirement is a description of the age and sex of the subjects. A description of other patient characteristics important in the study is also useful.
4. *Study site.* The nature of the site where the study was done should be described. This will affect the type of patients and the training and experience of the clinicians as well as the treatments offered.

Table II. Summary of Relevant Health and Economic Literature, Including Location, Inclusion Criteria, Sample Size, Outcome Measure, Survival Analytic Techniques, Predictors, and Hazard Rate Ratios

Study	Location (time frame)	Main inclusion criteria	*n*	Outcome measure	Survival analysis methods	Statistically significant predictors	HRR (95% CI)
							($p < .1$ level)
McIntosh *et al.* (1)	Ontario (January to December 1994)	All WSIB lost-time claims for low-back pain and attending a specific network of physiotherapy clinics	1,752	Time receiving wage replacement benefits for first episode	Kaplan–Meier estimates Log rank tests Cox proportional hazards (semiparametric modeling)	Industry: construction	0.510 (0.388, 0.670)
						Age	0.991 (0.985, 0.997)
						Lagtime (from injury to treatment)	Dependent on time
						Leg pain	Dependent on time
						3+ nonorganic signs	Dependent on time
						Intermittent pain	1.159 (1.008, 1.333)
						Previous episode	1.538 (1.130, 2.092)
						Functional status: Modified LBOS	Dependent on time
							($p < .1$ level)
Infante-Rivard and Lortie (19)	2 Quebec rehabilitation clinics (November 1988 to May 1992)	First compensated episode of back pain	305	Duration of time off work	Kaplan–Meier estimates Cox proportional hazards (stratified on variable: pain radiating during treatment)	Age 21–30	1.43 (1.04–1.98)
						Diagnosis of strain or sprain	2.20 (1.24–3.91)
						≤30 days from accident to treatment	1.30 (0.96–1.77)
						Good flexion at start of treatment	1.52 (1.04–2.24)
						Absence of neurological signs	1.40 (0.98–2.00)
						>24-month duration of employment	1.49 (1.10–2.03)
						Industry:	
						Public work	1.63 (1.21–2.19)
						Take unscheduled breaks	1.45 (1.06–1.97)

Oleinick *et al.* (20)	Michigan (1986)	First disability episode, <8 weeks duration	8,628	Missed work time for first disability episode	Kaplan–Meier estimates Cox proportional hazards (semiparametric modeling)	Male sex Age 25–34 34–44 45–54 55–64 Dependents 2 3 4 5+ Industry: construction Occupation Production Transportation Handler Accident type Impact Falls	1.12 0.894 0.76 0.738 0.669 0.877 0.838 0.84 0.785 0.804 1.118 1.139 1.111 0.863 0.74
Cheadle *et al.* (21)	Washington State (1987 to 1989)	Employees with WCB claims involving at least 4 days of time lost from work (compensable claims)	28,473	Duration of disability	Kaplan–Meier estimates Cox proportional hazards (semiparametric modeling)	Initial hospitalization Diagnosis of CTS Age > 45 Age 30–44 Female sex Divorced Dependents At company with >50 employees Work in county with +7% unemployment Industry Construction Agriculture	0.48 (0.46–0.51) 0.55 (0.50–0.60) 0.67 (0.64–0.69) 0.78 (0.76–0.81) 0.85 (0.82–0.88) 0.90 (0.86–0.94) 0.88 (0.86–0.91) 1.06 (1.05–1.11) 0.88 (0.82–0.95) 0.83 (0.76–0.90) 0.88 (0.83–0.93)

(*Continued*)

Table II. (*Continued*)

Study	Location (time frame)	Main inclusion criteria	*n*	Outcome measure	Survival analysis methods	Statistically significant predictors	HRR (95% CI)
							($p < .1$ level)
Coste *et al.* (22)	France (June 1 to November 7, 1991)	Self-referring patients with primary complaint of back pain lasting <72 hours, without radiation below gluteal fold	103	Time to recovery	Kaplan–Meier estimates Log rank tests Cox proportional hazards (semiparametric modeling)	Not employed Previous chronic back pain Initial disability rather than initial pain intensity Pain worse on standing Pain worse on lying Compensation status	0.63 (0.38–1.05) 0.21 (0.07–0.60) 0.59 (0.31–1.12) 049 (0.30–0.77) 0.62 (0.38–1.02) 0.49 (0.24–1.05)
				Time to return to work		Male sex Previous chronic back pain Pain worse on standing Pain worse on lying Initial disability rather than initial pain intensity Compensation status Job dissatisfaction	0.62 (0.35–1.06) 0.30 (0.05–1.02) 0.52 (0.30–1.03) 0.56 (0.29–0.93) 0.65 (0.36–1.14) 0.53 (0.30–0.94) 0.57 (0.24–1.13)
							($p < .1$ level)
Sinclair *et al.* (23)	Ontario (May to November 1993)	Off work at initial contact (within 21 days)	885	Cumulative time on wage replacement benefits	Kaplan–Meier estimates Log rank tests Stratified Cox PH	Filing claim likely to affect job (males) Reporting previous claim, surgery or hospitalization for another condition	0.59 (0.45–0.78) 0.83 (0.67–1.03)
Singer *et al.* (24)	Hamilton, Freelton, Burlington, Ontario	Lumbar sacral pain with or without radiation, no back pain for at least 30 days prior to current episode	252	Time to resumption of normal activities	Kaplan–Meier estimates	Initial pain intensity	$p < .01$
				Time to recovery from pain	Cox proportional hazards (semiparametric modeling)	Initial pain intensity Pain duration Number of prior episodes	$p < .10$ $p < .05$ $p < .01$

Reid *et al.* (25)	Vermont (September 1993 to June 1994)	Reported low-back pain injury within previous 11 days	207	Time until first return to work	Kaplan–Meier estimates Log rank tests Cox PH	More pain in past few days Worst pain since injury Believed they got wrong treatment	$p < .0001$ $p = .004$ $p < .05$
Lynch (28)	Great Britain (April 6, 1979 to May 30, 1981)	All invalidity benefit claims for first time or had last claimed benefits >13 weeks prior to beginning of study	1,856	Time to termination of invalidity benefits	Cox proportional hazards (semiparametric modeling)	Male Age (log scale) Age 55–65 Benefits (log scale) Pension Pension/age 55–65 Industry: mining/oil Diseases Mental Nervous Circulatory Injury/poisoning Employment rate (log scale) Female Age (log scale) Benefits (log scale) Pension Ill defined condition Genitourinary disease Access to indoor baths (log scale)	 −0.5959 −0.315 −0.6761 0.3593 −0.3294 0.3460 −0.3858 −0.5288 −0.5455 0.2411 −0.1906 −0.7322 −0.7278 0.3128 0.4517 0.5824 −0.2135

Note. HRR: hazard rate ratio; WSIB: Workplace Safety & Insurance Board of Ontario (formerly WCB); LBOS: low back outcome score; CI = confidence interval; WCB: Workers' Compensation Board; CTS: cumulative trauma syndrome; PH = proportional hazards; log = logarithm.

5. *Mathematical techniques.* Any report on clinical prediction must adequately describe and justify the mathematical technique employed. This should include methods used to avoid multivariable analysis problems such as overfitting the data with too few outcome events (<10) per predictor.
6. *Describing the results.* Methods include sensitivity, specificity, likelihood ratios, predictive values, proportion of patients with an outcome at a particular time point, and survival curves.
7. *Reproducibility.* Predictor variables should not be included if they lack interobserver reproducibility.
8. *Sensibility.* The likelihood that a predictor will be used increases if (a) it makes clinical sense, (b) it is easy to use, and (c) it suggests a course of action. This evaluation relies more on judgement than on statistical methods.
9. *Prospective validation.* Statistical validation of the predictor variables is important in patients whom the model was developed, but prospective validation in a group different from that which was used to derive the model is essential.

RESULTS

Table III summarizes these methods as they apply to the nine reviewed papers.

The outcomes were well defined in each paper and were all deemed clinically relevant based on the rationale presented in the Introduction section. Adequate definition of the predictor variables was much less frequent (5 out of 9). Blind assessment of outcome and

Table III. Evaluation of Reviewed Papers Based on the Methodological Standards of Wasson *et al.* (30) and Laupacis *et al.* (31) for Clinical Prediction Studies

	Study									Total ✓
Methodological standard	(1)	(19)	(20)	(21)	(22)	(25)	(23)	(24)	(28)	(/9)
Outcome										
Definition	✓	✓	✓	✓	✓	✓	✓	✓	✓	9
Clinical importance	✓	✓	✓	✓	✓	✓	✓	✓	✓	9
Blind assessment	✓	?	✓	✓	?	×	?	×	✓	4
Predictive variables										
Identification & definition	✓	×	✓	✓	×	×	✓	×	✓	5
blind assessment	✓	✓	×	?	✓	✓	✓	✓	?	6
Important patient characteristics described	✓	✓	✓	✓	✓	×	×	×	✓	6
Study site described	✓	✓	✓	✓	×	×	✓	×	×	5
Mathematical techniques described	✓	✓	✓	✓	✓	✓	✓	✓	✓	9
Results described	✓	✓	✓	✓	✓	×	✓	×	?	6
Reproducibility of										
Predictive variables	×	×	×	×	×	×	×	×	×	0
The rule	×	×	×	×	×	×	×	×	×	0
Sensibility										
Clinically sensible	✓	✓	✓	✓	✓	✓	✓	✓	×	8
Easy to use	✓	✓	✓	✓	✓	✓	✓	✓	×	8
Pr (disease) described	✓	✓	✓	✓	✓	?	✓	×	?	6
Course of action	✓	×	✓	×	×	×	×	×	×	2
Prospective validation	✓	×	×	×	×	×	×	×	×	1

Note. ✓: methodological standard met; ×: not met; ?: unclear; Pr = probability.

predictor variables were both infrequent (4 and 6 out of 9, respectively). The studies that used databases for information gathering (1,20,21,28) were assumed to have a blind assessment of outcome. The age and sex characteristics of the cohorts were described in six papers. An adequate description of the study site occurred in five papers.

As expected, because of the initial screening criteria, all papers employed suitable mathematical/statistical techniques. Only one paper (1) discussed sensitivity, specificity, or predictive values but generally documented the measure of effect. One paper provided an inadequate description of the results (24) and results from two papers (25,28) were not clear, primarily because of their lack of stated HRR.

The only economic paper (28) displayed the beta coefficients from parametric regression, but no HRR. No paper addressed the issue of reproducibility of the predictor variables or the final model.

Most papers derived models that were clinically sensible (8 out of 9), and the ease of use for clinicians was high (8 out of 9). The probability of a given outcome was not stated in one article and not clear in two others. A recommended course of action for use by future patients/therapists in prognostication was rarely documented (2 out of 12). No study shared any external validation results, but one paper (1) internally validated their model using a data-splitting technique.

DISCUSSION

Pulcins *et al.* (5) noted that there is little agreement on the prognostic factors associated with acute back pain. The reviewed papers had methodological differences that make direct comparison problematic despite the careful application of inclusion criteria documented here. Part of this disagreement may be because of differences in sampling frames, patient populations, and other design features (25).

Eligibility requirements were vastly different across studies. Four papers were based primarily on Workers' Compensation (WC)-type databases (1,20,21,28). Six studies (1,19, 22–25) involved some primary data collection on patients from primary or secondary care; three of these papers (1,23,25) also included information from WC-type databases. The economic paper (28) focused on administrative datasets from insurance company records.

Some papers studied nonoccupational back pain, meaning that subjects did not necessarily have back pain originating within the context of work. It is often impossible to distinguish between back pain caused by work and pain of uncertain origin that makes job duties difficult to perform (2). This difference in case mix of injury types may contribute to interstudy disagreement because previous research has shown that compensation status is associated with recovery (32).

There was variability within the WC datasets because of interjurisdictional differences in the rules governing benefit entitlement. In Washington State, WC wage replacement starts after three days of work absence (21), whereas in Michigan, eligibility commences after seven or more consecutive workdays of work disability following an injury, not including Sundays or the day of the injury (20). In Vermont, to be eligible work-related injuries must be reported within 72 h (25). In Ontario, there is no "waiting time"; any full day of lost time due to injury or illness is compensated if the cause is "adjudicated" as occupational; eligibility for wage-replacement benefits commences as of the first completely missed shift (day) from work, but the employer must notify the WC carrier within three days of the

accident (23). Little is known about the effects of such "waiting times" on return to work for WC claimants, but they are likely complex because of inevitable compensatory behavioral responses by both injured employees and employers (8). Pulcins *et al.* (5) suggested that the role of specific characteristics of compensation systems in the genesis of disability is controversial but cannot be reasonably denied.

The outcome variable definition differed across the reviewed papers. Most studies evaluated time on some type of benefits: wage replacement (1,23), disability (21), and invalidity (28). Three studies (19,22,25) evaluated time to return to work; three (20,22,24) examined time to recovery as reported by the patient and/or clinician. Unfortunately, the time of entry into these studies was not consistent, thereby altering the temporal nature of the outcome between studies, in terms of actual time elapsed since disability or care began.

There was variability in the length of time that subjects were under study and the method of contact for follow-up. Oleinick *et al.* (20) used a 2-month follow-up and Reid *et al.* (25) followed patients for 3-months postaccident. Coste *et al.* (22) followed patients for 3 months after initial presentation to a physician. Three studies used the more common 1-year follow-up time (1,23,24). Lynch (28) and Cheadle *et al.* (21) reported 2 and 4 years of follow-up, respectively. Infante-Rivard and Lortie (19) had a maximum of 1,228 days (3.3 years) of follow-up postdischarge from treatment. Examining the type of follow-up revealed that four papers used direct patient contact (22–25), whereas five studies accessed database records (1,19–21,28). These discrepancies highlight the need for development of a theoretical base for discerning appropriate outcome measures to be used consistently across future studies (5).

Statistical Limitations

Part of the disagreement over prognostic factors is due to the lack of statistical and epidemiological sophistication among studies to date (5). Omissions in published papers made some analytic options—proportional hazards assumption, multivariable model selection, interaction, and goodness-of-fit—difficult to assess.

Proportional Hazards Assumption

Testing the proportional hazards assumption was specifically mentioned in seven papers (1,19–23,25). To verify that the proportional hazards assumption underlying the Cox model was met, McIntosh *et al.* (1) used time-varying coefficients in a piecewise proportional hazards approach. Sinclair *et al.* (23) tested the proportional hazards assumption using piecewise, linear, quadratic, and cubic interactions with duration. Infante-Rivard and Lortie (19) stratified on variables with nonproportional hazards. Reid *et al.* (25), Oleinick *et al.* (20), and Coste *et al.* (22) tested the assumption by visually checking the graphs for nonparallelism between the hazard functions against time. Cheadle *et al.* (21) tested the assumption graphically by plotting the relative hazards over time for each categorical variable.

Multivariable Model Selection

Five papers described details on the selection method of their final multivariable models (1,19,22,23,25). Infante-Rivard and Lortie (21) used a forward stepwise selection method.

McIntosh *et al.* (1) and Coste *et al.* (22) used an automated forward stepwise selection approach. Reid *et al.* (25) employed an automated backward selection procedure. Sinclair *et al.* (23) developed a final model based on 75 original variables. Although the variable reduction process was not specifically described, discussion with the authors (23) revealed that only factors thought to be related to rehabilitation clinic attendance were included in the multivariable model because the study was primarily designed to evaluate their clinics. Oleinick *et al.* (20), Cheadle *et al.* (21), Singer *et al.* (24), and Lynch (28) included all covariates in their multivariate analysis regardless of the statistical significance, thereby potentially complicating the assessment of overfitting and adequate control of confounding.

Interaction and Collinearity

Four studies mentioned testing for interaction among the various confounders in their final model (20,21,23,28). None of these papers reported testing for possible collinearity among covariates reported in their final models.

Goodness-of-Fit

Two studies (1,19) reported testing the goodness-of-fit of the multivariable models. Semiparametric (Cox regression) survival analytic models were generally used in the health literature. McIntosh *et al.* (1), and Infante-Rivard and Lortie (19) were the only health literature papers to report testing the fit of their models using −2 log likelihood statistics.

Although statistical and methodological differences lead to inconsistency between studies, Pulcins *et al.* (5) state that much of the variation in prognostic factors likely is due to lack of a common conceptual framework for the study of back pain prognosis. Back pain may be a condition where a multiplicity of complex factors—workplace, economic or social milieu—are important determinants of either recovery or chronicity (33).

Common Prognostic Factors

Pain measurements were significant predictors in three studies (22,24,25). McIntosh *et al.* (1) determined that functional status was a significant predictor. Of the demographic variables utilized, age was a significant predictor in five studies (1,19–21,28). The age effect is probably due to both the reduced physiological ability of older workers to recover from injuries and the decreased likelihood of obtaining employment once they have recovered (21).

Gender was a significant predictor in five studies (20–23,28). Two papers (20,21) showed that females were slower to recover, one paper (22) had results in the opposite direction, and one (23) stratified their results based on gender. Because women visit physicians with less provocation than men do, relatively less severe female cases may have been included in the studies showing faster recovery for females (8) especially in the primary care patient sample of Coste *et al.* (22). Conversely, perhaps the women who recovered more slowly had family care responsibilities in addition to employee duties working outside the home that may have hindered recovery. Gender has not figured prominently in some studies, perhaps because fewer women file workers' compensation claims in some jurisdictions (21), if their original jobs are not available.

Occupational and/or industry measures were significant in five studies (1,19–21,28) including the delayed recovery seen in construction workers (1,20,21). Lynch (28) found significant prognostic value for the logarithm of benefits.

Because most of the reviewed studies represent a relatively compartmentalized view of back pain, they have not contributed greatly to our practical knowledge. For example, several studies did not include any variables collected by the clinician providing care for the subject, whereas some papers included data collected from the subjects directly, presumably reflecting the authors' different intentions for the utilization of their results by various stakeholders.

Individual patient differences—age, education, duration of disability, physical impairment, perception of disability, return-to-work expectation, pain, personality style, somatic preoccupation, perceptions of the work environment—are believed to influence recovery (25). Although the biological effects of factors such as age and gender should not be expected to differ greatly across settings and populations, these variables have many indirect biopsychosocial implications likely to subtly affect the return to work, including work heaviness and labor market participation options that may vary across settings (2).

Unfortunately, it is difficult to generalize from this literature, the prognostic factors for chronic back pain, but this review has identified critical methodological issues necessary to properly design a high-quality back pain prognosis study. Of particular importance is the use of an inception cohort, assessment of an objective outcome, and the use of multivariable time-to-event statistical methods.

CONCLUSION

High-quality back pain prognosis studies will help clinicians direct specific interventions towards patients at definite risk. Systematic literature review and appraisal is a necessary initial step to gain a better understanding of critical design issues required to properly design such prognosis research.

Few studies have utilized many of these methodological standards. To date, prognosis has been an inadequately studied aspect of the continuum from back injury to recovery. There was a limited number of better-designed medical and economic studies that achieved good scientific quality with minimal flaws found in this review.

A properly conducted prognosis study can help improve the quality of future clinical trials. Researchers and clinicians interested in prognosis research need to overcome the limitations of past designs and address the methodological guidelines outlined to improve the quality of future prognosis studies.

APPENDIX A: LITERATURE SEARCH STRATEGY OF AN ELECTRONIC DATABASE

One electronic database, MEDLINE, was used to facilitate the initial search for relevant articles. MEDLINE is a powerful tool for locating studies (34) in the medical literature. Using the most current MEDLINE database at the time of the literature search (1997–2000), multiple text words were used to identify prognosis-type studies. A separate search was conducted for each text word (predict, or prognosis, or prognostic, or risk factor), which

were then combined using the Boolean "OR" operator. This resulted in one set of 40,527 prognosis papers.

A similar strategy resulted in a set of 2,106 back pain/work injury papers that contained all studies with the text words: back pain, or back problem, or back injury, or work injury, or work disability, or work-related disability.

A third strategy resulted in a set of 6,214 papers involving survival analysis that contained all studies with the text words: survival analysis, or cox, or kaplan-meier, or kaplan meier, or proportional hazard, or proportional hazards.

The three sets of search results were then combined using the Boolean "AND" operator to create one final set that contained all studies that included at least one text word from the prognosis set, one from the back pain set, and one from the survival analysis set. This entire strategy was repeated a second time for the MEDLINE database 1993–1996, a third time for the 1987–1992 database, and a fourth time for the 1985–1986 database. Because of the paucity of studies found in the earlier databases, and the relatively new methods for employing survival analysis techniques, a search of earlier databases was not performed.

APPENDIX B: GLOSSARY OF TERMS

Collinearity: very high correlation between variables (35).

Data-splitting: a method of obtaining a nearly unbiased internal assessment of accuracy. With this technique, a random sample of the full dataset (dataset split in half or some other appropriate fraction) is used for model development and the entire dataset for validation.

Goodness-of-fit: a statistical test of the hypothesis that data have been randomly sampled or generated from a population that follows a particular theoretical distribution (35).

Hazard: a synonym for risk (35).

Hazard function: can be used to characterize survival; the hazard function gives the instantaneous potential at time t for getting an event; stated another way, the hazard function is the instantaneous rate of change of failure probability, conditional on the patient surviving to time t; in contrast to the survivor function that focuses on nonoccurrence, the hazard function focuses on the event occurring. There are several types of hazard functions, three of which are as follows:

1. Exponential model: the potential for the event remains constant regardless of time.
2. Weibull model: the hazard function increases or decreases over time.
3. Log-normal model: the hazard function first increases then decreases (36).

Hazard rate ratio: a theoretical measure of the risk of occurrence of an event (35); the measure of effect obtained in survival analysis, expressed in terms of an exponential of a regression coefficient in the model (36).

Interaction: differences in the effects of one or more factors according to the level of the remaining factors (i.e., effect modification) (35).

Parametric model: a statistical model that depends upon assumption(s) about the distribution of the data; that is, the survivor function is believed to be of a certain form (e.g., exponential, Weibull, log-normal).

Proportional hazards (PH) model: often referred to as the Cox model, a statistical model in survival analysis that does not require specification of a particular probability distribution to represent survival times; thus, it is known as semiparametric (7). The PH model asserts that the effect under study does not change over time.

Proportional hazards assumption (PHA): an important feature of the Cox model that states that the hazard rate ratios do not change over time. Verifying this assumption of parellelism between hazard function curves over follow-up is essential to determine if the assumption of time-independent relationships are appropriate or if the extended Cox model is required, which considers time-dependent variables and effects.

Extended Cox model: used when the model no longer satisfies the PH assumption and time-dependent coefficients are considered; to assess whether this model is appropriate, the Cox model is extended by defining product terms involving time-independent variables multiplied by some function of time.

Piecewise proportional hazards model: assumes that the hazard is constant within a time interval, but can change from one interval to the other (37). Product terms are formed between the time levels and each variable. Statistically significant terms are regarded as evidence against the proportional hazards assumption.

Survival analysis: a collection of statistical procedures for data analysis for which the outcome variable of interest is time until an event occurs (36).

REFERENCES

1. McIntosh G, Frank JW, 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(2): 147–157.
2. Frank JW, Brooker AS, Demaio S, Kerr MS, Maetzel A, Shannon HS, Sullivan TJ, Norman RW, Wells RP. Disability resulting from occupational low-back pain. Part II: What do we know about secondary prevention? A review of the scientific evidence on prevention after disability begins. *Spine* 1996; 21(24): 2918–2929.
3. Frank JW, Pulcins IR, Kerr MS, Shannon HS, Stansfeld SA. Occupational back pain-an unhelpful polemic. *Scand J Work Environ Health* 1995; 21: 3–14.
4. How to read clinical journals: III. To learn the clinical course and prognosis of disease. *Can Med Assoc J* 1981; 124: 869–872.
5. Pulcins IR, Goel V, Cohen JE, Frank JW. A structured critical review of acute back pain prognosis studies. Working Paper No. 21-R1, Institute for Work & Health, 1995, unpublished.
6. Baldwin ML, Johnson WG, Butler RJ. The error of using returns-to-work to measure the outcomes of health care. *Am J Industr Med* 1996; 29: 632–641.
7. Allison PD. Survival analysis using the SAS system: A practical guide. SAS Institute, Inc., Cary, NC, 1995.
8. Haidar H. *Expressing effect estimates of predictors of return to work.* MS Thesis in Epidemiology. Toronto: University of Toronto Press, 1997.
9. Hogg-Johnson S, Cole DC. Early prognostic factors for duration on benefits among workers with compensated occupational soft tissue injuries. Working Paper No. 64, Institute For Work & Health, 1998, unpublished.
10. Hogg-Johnson SA, Frank JW, Rael E. Prognostic risk factor models for low-back pain: Why they have failed and a new hypothesis. Working Paper, No. 19, Institute for Work & Health, 1994, unpublished.
11. Mau W, Bornmann M, Weber H, Weidemann HF, Hecker H, Raspe HH. Prediction of permanent work disability in a follow-up study of early rheumatoid arthritis: Results of a tree structured analysis using RECPAM. *Br J Rheumatol* 1996; 35(7): 652–659.
12. Nishimura A, Sakata S, Iida K, Iwasaki Y, Takeshima T, Todoroki T, Ohara K, Hata K, Miyoshi M, Seo Y. Evaluation of intraoperative radiotherapy for carcinoma of the pancreas: Prognostic factors and survival analyses. *Radiation Med* 1988; 6(2): 85–91.
13. Ridder GJ. Klempnauer J. Back pain in patients with ductal pancreatic cancer. Its impact on resectability and prognosis after resection. *Scand J Gastroenterology* 1995; 30(12): 1216–1220.

14. Albert SM, Marder K, Dooneief G, Bell K, Sano M, Todak G, Stern Y. Neuropsychologic impairment in early HIV infection. A risk factor for work disability. *Arch Neurol* 1995; 52(5): 525–530.
15. van den Hoogen HJM, Koes BW, Deville W, van Eijk JTM, Bouter LM. The prognosis of low-back pain in general practice. *Spine* 1997; 22(13): 1515–1521.
16. Indahl A, Velund L, Reikeraas O. Good prognosis for low-back pain when left untampered. A randomized clinical trial. *Spine* 1995; 20(4): 473–477.
17. Burdorf A, Naaktgeboren B, Post W. Prognostic factors for musculoskeletal sickness absence and return to work among welders and metal workers. *Occup Environ Med* 1998; 55(7): 490–495.
18. Infante-Rivard C, Lortie M. Relapse and short sickness absence for back pain in the six months after return to work. *Occup Environ Med* 1997; 54(5): 328–334.
19. Infante-Rivard C, Lortie M. Prognostic factors for return to work after a first compensated episode of back pain. *Occup Environ Med* 1996; 53(7): 488–494.
20. Oleinick A, Gluck JV, Guire KE. Factors affecting first return to work following a compensable occupational back injury. *Am J Industr Med* 1996; 30: 540–555.
21. Cheadle A, Franklin G, Wolfhagen C, Savarino J, Liu PY, Salley C, Weaver M. Factors influencing the duration of work-related disability: A population based study of washington state workers' compensation. *Am J Public Health* 1994; 84: 190–196.
22. Coste J, Delecoeuillerie G, Cohen De Lara A, Le Parc JM, Paolaggi J. Clinical course and prognostic factors in acute low-back pain: An inception cohort study in primary care practice. *Br Med J* 1994; 308: 577–580.
23. Sinclair S, Hogg-Johnson SA, Mondloch MV, Shields SA. The effectiveness of an early active intervention program for workers with soft-tissue injuries: Early claimant cohort study. *Spine* 1997; 22(24): 2219–2231.
24. Singer J, Gilbert J, Hutton T, Taylor D. Predicting outcome in acute low-back pain. *Can Fam Physician* 1987; 33: 655–659.
25. Reid S, Haugh LD, Hazard RG, Tripathi M. Occupational low-back pain: Recovery curves and factors associated with disability. *J Occup Rehabil* 1997; 7(1): 1–14.
26. Johnson W, Ondrich J. The duration of post injury absences from work. November 1989, unpublished.
27. Johnson W, Butler R, Baldwin M. First spells of work absences among Ontario workers. March 1993, unpublished.
28. Lynch M. The duration of invalidity benefits claims: A proportional hazard model. *Applied Econom* 1991; 23: 1043–1052.
29. McIntosh G. *Back pain prognostic factors: A cohort study of 1,752 patients of a national rehabilitation clinic system.* MS Thesis in Epidemiology. Toronto: University of Toronto Press, 1999.
30. Wasson JH, Sox HC, Neff RK, Goldman L. Clinical prediction rules: Application and methodological standards. *N Engl J Med* 1985; 313: 793–799.
31. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules: A review and suggested modifications of methodological standards. *JAMA* 1997; 277(6): 488–494.
32. Greenough CG, Fraser RD. The effects of compensation on recovery from low back injury. *Spine* 1989; 14(9): 947–955.
33. Spitzer WO, Leblanc FE, Dupuis M. Scientific approach to the assessment and management of activity-related spinal disorders: A monograph for clinicians. Report of the Quebec task force on spinal disorders. *Spine* 1987; 12(7S): 1–59.
34. The Cochrane Collaboration Handbook—version 3.0.2. Institute for Work & Health 1998.
35. Last JM. *A dictionary of epidemiology.* New York: Oxford University Press, 1995.
36. Kleinbaum DG. *Survival analysis: A self learning text.* New York: Springer-Verlag, 1996.
37. Cantor AB. Extending SAS survival analysis techniques for medical research. SAS Institute, Inc., Cary, NC, 1997.