This manuscript has been accepted for publication in a future issue of Chronic Diseases in Canada (http://www.phac-aspc.gc.ca/publicat/cdic-mcc), possibly in Volume 28, Number 3, 2008. It includes all modifications from the peer review process but still hasn't been copy edited, formatted or proofed. Consequently, the journal accepts no responsibility for any errors or omissions there in.

RISK FACTORS FOR FALLING AMONG COMMUNITY-DWELLING SENIORS USING HOME-CARE SERVICES: AN EXTENDED HAZARDS MODEL WITH TIME-DEPENDENT COVARIATES AND MULTIPLE EVENTS

Bernard S. Leclerc[1,2], Claude Bégin[3], Élizabeth Cadieux[1], Lise Goulet[4], Nicole Leduc[4], Marie-Jeanne Kergoat[5], Paule Lebel[5,6]

AUTHOR REFERENCES

[1] Service de surveillance, recherche et évaluation, Direction de santé publique et d'évaluation, Agence de la santé et des services sociaux de Lanaudière, Joliette, QC, Canada.

[2] This research is a part of his Ph.D. thesis in Public Health and Epidemiology, realized under the supervision of professors Lise Goulet and Nicole Leduc, respectively from the Département de médecine sociale et préventive and the Département d'administration de la santé, Faculté de médecine, Université de Montréal, Montréal, QC, Canada.

[3] Service de prévention et de promotion, Direction de santé publique et d'évaluation, Agence de la santé et de services sociaux de Lanaudière, Joliette, QC, Canada.

[4] Groupe de recherche interdisciplinaire en santé, Université de Montréal, Montréal, QC, Canada.

[5] Research Center, Institut universitaire de gériatrie de Montréal, Montréal, QC, Canada

[6] Centre d'expertise sur la santé des personnes âgées et des aidants, Institut universitaire de gériatrie de Montréal, Montréal, QC, Canada.

Correspondence to: Bernard-Simon Leclerc, Service de surveillance, recherche et évaluation, Direction de santé publique et d'évaluation, Agence de la santé et des services sociaux de Lanaudière, 245, rue du Curé-Majeau, Joliette, J6E 8S8. Phone: 450 759-1157 extension 4324; fax: (450) 755-1969; email: Bernard-Simon_Leclerc@ssss.gouv.qc.ca

Word counts for abstract (194) and text (5364).

RISK FACTORS FOR FALLING AMONG COMMUNITY-DWELLING SENIORS USING HOME-CARE SERVICES: AN EXTENDED HAZARDS MODEL WITH TIME-DEPENDENT COVARIATES AND MULTIPLE EVENTS

ABSTRACT

The identification of risk factors for falls in longitudinal studies becomes difficult because of exposures that change during the follow-up and also because individual subjects may experience an event more than once. These issues have been neglected and improper statistical techniques have been used. The typical approaches have been to report the proportion of fallers or the time-to-first fall. Both avoid the underlying assumption of independence between events, but discard pertinent data. We review the existing methods and we propose a Cox hazards extension. We exemplify it in the study of potential risk factors associated with all falls in 959 seniors. Finally, we compare the results of the proposed Wei, Lin & Weissfeld (WLW) method with those of several techniques. Stable exposure variables measured at baseline and updated time-varying exposures include sociodemographic characteristics, BMI, nutritional risk, alcohol consumption, home hazards, gait and balance, and medications. Results demonstrate that usual methods of analyzing risk factors for falling are irrelevant, producing considerable biases with respect to the WLW model using time-dependent covariates. Results also show that modeling for first events may be inefficient given that the risk of occurrence varies between falls.

Key words: Accidental falls, Cox model, elderly, environmental hazards, negative binomial distribution, hazards model, regression analysis, survival analysis, logistic models

INTRODUCTION

Falls are common, recurrent problems with serious consequences for elderly people and the healthcare system.[1] Evidence about fall-risk factors have generally been identified by prospective observational designs. These studies may suffer from the same problems as for a cohort of other issues. These include loss to follow-up and variable follow-up time. Identification of fall-risk factors deals with additional problems such as exposure changes during the follow-up and recurrent events in the same person. These issues have been neglected and inefficient statistical techniques have been used. As a result, this may have distorted the magnitude in estimates of particular predictors or produced misleading results. Moreover, this may have eluded questions of great clinical relevance.[2-5]

More than 15 years ago, Cumming, Kelsey and Nevitt[6] advised that more attention be paid to repeated measures regarding both risk factors and rate for all falls. Despite this, few researchers have challenged the design of their studies and the analysis of their data. Unfortunately, it seems rather to have adversely affected them to circumvent the methodological complications by discarding much relevant information.

The aim of the present paper is to enhance awareness of researchers about some epidemiological and statistical considerations. We review the statistical background of methods of fall studies, we introduce the philosophical issues of time-dependent covariates and multiple events and we discuss the existing statistical techniques which deal with them. We propose an extension of the Cox proportional hazards tradition and exemplify it in the identification of potential risk factors associated with all falls in elderly people living at home. Finally, we compare the different results obtained by various statistical methods.

STATISTICAL BACKGROUND OF METHODS OF FALL STUDIES

A variety of strategies have been used to study the risk factors for recurrent falls. Their analysis is complicated by the within-subject correlation. In other words, the occurrence of an event acts on the risk of the next one. Failure to account for dependence in the data leads the usual estimator of variance to be underestimated. This produces confidence intervals that are too narrow, and test of significance too liberal (i.e. rejects the null hypothesis too often).[3,5,7]

A summary of some of the discussed methods is provided in Figure 1. A simplistic approach to such problems involves reporting the proportion of fallers (subjects who fall a least once over an arbitrarily defined period) or the time to a first fall.[8] Both possibilities avoid the underlying assumption of independent association between multiple events. However, the use of all available data for each individual could be more efficient.[4,8,9] The author of a key paper argued that the incidence rate for falls was the public health priority[6], particularly for less robust elderly people.[10] The challenge in analysing all falls arises because some people are more prone to recurrences than others, hence returning them at higher risk of fall-related injury than those who fall only once. The choice of outcome, according to whether the focus is on fallers or on the rate of falls, could also affect the conclusion with knowing if a particular exposure constitutes a risk factor. Stable over-time factors are more likely to be related to the state of “being a faller” than exposures that vary over time.[6]

Alternatives have been proposed for dealing with multiple events. Among these are the negative binomial regression, some extensions to the Cox proportional hazards model and a modified logistic regression. The dependent variable in the negative binomial regression is the individual event rate adjusted for the follow-up time, that is, the number of falls for a person divided by their specific follow-up time (Figure 1).[4,11] Since the negative binomial distribution has one more parameter than the Poisson, it naturally accommodates for over-dispersion (i.e., the variance typically exceeds the mean).[8] Therefore, this approach is robust for dependent structure data, and suitable for frequent and recurrent events.

One problem using event rates is that the likelihood of event occurrence must be assumed constant through time within one participant. A critical example could be to consider the equivalent event rates for which three participants are observed for three years and each has three falls. Consider one fell once each year, another at whom all three falls occurred in the first year, and finally the last one which would have had all three falls in the third year. The outcome variable ignores the time of occurrence for these events.[8] So, a negative binomial modeling event rates may not be the method of choice when the value of important covariates or that the likelihood of event occurrence changes with the passing of time.[3] Greater efficiency and accuracy can be obtained by modelling the lengths of inter-episode intervals via time-to-event techniques.[9] Instead of focusing on the numbers of cases, the time-to-event approach considers the time between falls. If an incidence rate is high, the intervals between events will be short, and vice versa.[3]

In addition, measured risk factors for which we want to evaluate the effects are usually only fixed variables, defined at the initial examination.[2] They refer to the intrinsic characteristics of the subjects (e.g., the sex), the past exposures (e.g., prior falls) or exposures present at baseline (e.g., consumption of drugs). Exposures that occur after the starting point or vary over time for an individual are not taken into account. Examples, which potentially cause falls through short-term exposure preceding the event, include environmental hazards, alcohol drinking and medication use. A great advantage of the time-to-event approach is its ability to handle time-dependent covariates.[3]

The hazards models include the counting process of Andersen & Gill[16] (hereafter referred to as AG), the conditional model of Prentice, Williams & Peterson[17] (PWP) and the marginal model of Wei, Lin & Weissfeld[18] (WLW). None of these approaches explicitly model the dependence structure between failure times as required. Instead, robust estimates of variance are used to account for correlated observations within subjects. That is, the so-called "variance-corrected" hazards models.[12-15]

The distinction of the hazards methods can be seen in terms of who is in the risk-set at each failure.[15,19] The AG rests on the strong assumption that the risk of an event for a given subject is unaffected by any earlier events, unless a term that captures such dependence (i.e., number of previous falls) is included as a time-dependent covariate.[3,7] That means that the data for each subject with multiple events could be described as data for multiple subjects where each has delayed entry and is followed until the next event (Figure 1). This model ignores the order of the events, that is, all falls are indistinguishable, leaving each subject to be "at risk" for an event as long as the subject is still under observation at the time the event occurs.[3,7,8,13-15,19]

The PWP is based on the idea that a subject is not technically at risk of a later event until all previous events have been experienced. This is accomplished by stratifying data by event order. Accordingly, the risk-set at time t for the kth event is limited to those subjects under study at t who have already experienced k–1 events (not exemplified in Figure 1).[13-15,19] However, Robertson[20] argued that the conditional assumption of the order of events does not hold for falls. As an illustration of her argument (personal communication), let us speculate that a person slipped on water on the floor of the kitchen without injury and another time, fell on the pavement outside and this resulted in a hip fracture. The person is at risk for both these falls from the beginning of the study period, that is the time at risk for the second fall on the pavement does not start only after the first fall in the kitchen has occurred.

The risk-set of the WLW marginal approach includes all patients under observation who have not yet experienced the *k*th event. The time for each event starts at the beginning of follow-up time for each subject. Furthermore, each subject is considered to be at risk for all events, regardless of how many events each subject actually experienced. The WLW does not impose any dependence structure among the related failure times. Thus, it ignores the ordering of events but takes into account previous events by treating each failure as belonging in an independent stratum (Figure 1).[4,7,8,13-15,19]

The logistic regression analysis is the most commonly used method in epidemiological research. D'Agostino *et al.*[21] showed that a so-called pooled logistic regression is identical to the time-dependent covariate Cox regression. This is what makes the technique attractive to evaluate the relationship of risk factors to disease development. O'Loughlin[22] applied such an approach to the study of falls. The theoretical basis for the use of this logistic regression variant is well established when the intervals between measurements of risk factors are short, the probability of an event within an interval is small, and the intercept for the pooled logistic is constant across intervals.[21] The underlying statistical requirements and data setup for the pooled logistic regression are very close to those defined for the AG. Each of the follow-up intervals for a single subject is assumed to represent intervals from different subjects. The method pools the subjects at risk and the events developed in each interval. The follow-up interview number is included as a categorical variable to test this assumption. Similarly, the dependence between multiple falls within the same individual is accounted for by considering the occurrence of previous falls as a predictor variable.[23]

However, the way that the interval observations are set up as well as the outcome variable of interest is different in both methods. The AG builds the intervals according to the precise dates of events. For example, the first interval will cover the time span from entry into the study until the time of the event, and the following interval spans the time from the first event to the next one, and so on (Figure 1).[15] In contrast, the logistic regression uses stable time points fixed by the researcher. For example, an exam could be performed at the same date each month to up-date risk factors and gather information on falls that occurred in the interval of observation (not exemplified in Figure 1).[22,23] The above analysis is, in essence, an investigation of fallers versus non-fallers in successive short intervals.[22] Even if, taken as a whole, it allows more than one outcome to occur per subject, less appreciated is that it drops all additional falls that may occur in each particular interval. It seems obvious that if three falls per month are considered equivalent to one fall for the same period, information translating the intensity of short-term phenomena is lost.

The choice of one of these models must be carried out starting from *a priori* ideas on the type of relationships which exist between the covariates and the risk of falling. In negative binomial regression, AG and pooled logistic regression, no distinction was made between the various events that succeeded one another. This restricts the baseline hazard and the regression coefficients to not vary according to the rank of recurrence. A history of previous falls is strongly recognized as a predictor of subsequent falls.[10,23] Intuitively, we would expect a first fall to differ from the aetiology of the subsequent falls. The predictors for one fall that can occur by accident might be different than those for recurrent falls that can be more associated to one's health condition.[24,25] So, researchers and practitioners are not only interested in the overall covariate effects on the risk of all failures. In fact, they may want to know the specific effects of independent variables for the first, second, or subsequent events as well. The binomial regression, AG and pooled logistic regression provide no insights to answer such questions contrary to the WLW. According to the structure of the data to be analyzed and the research question to be answered, the WLW is expected to be a naturally more relevant method for risk factor studies for falls than the other methods.

METHODS

Subjects and procedures

Subjects were volunteers recruited to form an open cohort between March 2002 and July 2005 inclusive among community-dwelling persons aged 65 years or more who receive home-care services. People who could speak neither French nor English, those not able to walk more than six meters, and those with reduced communication and cognition were excluded. All subjects gave informed consent. The study was approved by the authorities of each participating centre.

This study is a part of a research project on the evaluation of a multifaceted preventive intervention.[26] Participants were visited at home, at entry and every six months, by a trained physical rehabilitation therapist in order to ascertain information about potential risk factors. A fall was defined as an event resulting in the subject inadvertently coming to rest on the ground, floor or other lower level. Excluded were sports-related falls.[23] Subjects were asked about falls in the three months preceding the initial interview and monitored for new falls by use of a daily completed calendar and monthly phone calls.

Material and social forms of an ecological deprivation index were imputed to participants, which uses Canadian census data to match postal codes with geographic areas of residence.[27,28] Nutritional risk screening was performed on a graded 13-point scale tool.[29-32] Body weight was self-reported and height was measured using standard techniques. Gait, balance and mobility performance were assessed by the Berg scale[33-36] on a 56-point scale, and by the Timed Up & Go test[37,38] which measures the overall time, in seconds, to complete a series of functional tasks. Subjects' homes were assessed for 37 potential environmental hazards using the Gill's room-by-room assessment form[39,40] Housing types included: single-family house; apartment; row housing or other unique entrance dwelling units; private residential facilities; other housing including room in shared accommodation. Data about the use of benzodiazepines (yes/no) and number of daily consumed prescribed drugs were recorded directly from the containers. A detailed history of alcohol consumption was obtained according to a questionnaire developed by the Québec Institute of Statistics.[41,42] Responses were categorized for both drinking in the preceding week (yes/no) and usual drinking during the last 6 months (non-drinker, ⩽ 3 times a month, 1-6 times a week, every day). In a general way, higher values of the measurements denote higher risk or impairment, except for Berg scale for which the opposite is true.

Statistical analyses

Descriptive analyses were carried out using SPSS® 13.0 and regression analyses using SAS® 9.1. The adjusted effects of subject characteristics on the likelihood of falling were investigated using three survival-analysis techniques (conventional Cox regression, AG extension and WLW extension), a negative binomial regression and a logistic regression.[15,43]

The dependent variable in all survival analyses was time to fall for each participant during the follow-up, measured in days. Only cases with at least one month of follow-up fall data were included. Subjects were censored upon reaching 18 months of follow-up (optional voluntary drop out), end of study or time of withdrawal for any reason. Repeated falls were considered as occurrences of the same type of indistinguishable events. Survival analyses were performed with all covariates measured on baseline only and with updated covariates. Baseline covariates included age, sex, number of falls in the three months prior to study entry, type of residence and deprivation index. Time-varying covariates included BMI, nutritional risk, alcohol consumption, home environmental hazards, gait and balance, use of benzodiazepines and total medications, by taking the measurement closest in time preceding the fall considered.[2] Measurement of exposure to the middle of the follow-up period was used in the case of the people who did not fall. We thus tested the null hypothesis that the exposure collected during the follow-up was not associated to the risk of falling thereafter.[2] No proportional hazards assumption was required in Cox with time-dependent covariates procedure, since the hazards depend on time.[2,43]

The dependent variable in the logistic regression was the state of being a faller (subjects who fall a least once) over a 12-month period. The negative binomial and logistic regressions were performed with all covariates measured on baseline. The statistical methods are summarized in Figure 1. The linearity assumption of the relationships was checked for continuous predictor variables. All models were fit using a stepwise-like process to retain any variable in the presence of others with a *p*-value ≤ 0.05. Robust sandwich estimates of variance were used in the survival-analysis as well as the negative binomial regression techniques in order to make up for the lack of independence between multiple falls.

The WLW approach estimated both common and event-specific β for the first five falls for each subject as well as the common β for all the observed falls. The number of subjects at risk for a given stratum, after the first fall, was made up of all subjects who experienced a fall in the preceding stratum minus those who were lost in the follow-up; *n* of subjects at risk for a given pooled fall group was made up of all subjects under observation in all considered strata, as if subjects in each stratum represented a different subject. Each model was examined with and without past fall strata, because it could have masked the effects of other variables of interest.[6,23]

RESULTS

Study subjects

Of the 959 persons who met the study inclusion criteria, agreed to participate and received a home visit, 22 withdrew without completely filling the baseline assessments or before one month of follow-up. Mean and median follow-up times of the remaining 937 subjects were 488 and 458 days, respectively (range, 27 to 1330 days). Some 549 subjects (57.2%) remained in the study at 12 months and 377 (39.3%) at 18 months. Respondents were mainly women (75.7%). Mean age (standard deviation) was 79.5 (6.7), of which 76.4% were 75 years old or more. Thirty-nine percent (39.0%) experienced at least one fall in the three months prior to study entry and 14.9% have had two or more.

Comparison of statistical methods

Table 1 summarizes the differences in relative risks for falling obtained by several statistical methods. First of all, the logistic regression (1) and time-to-first fall using a standard Cox (3a), that overlook the recurrence of falls, identified less significant risk factors than negative binomial (2), AG (4a) and WLW (5a), that consider all the available information (number between parentheses refers to the concerned model in the Table 1). Logistic regression and standard Cox identified the same risk factors, whereas ignoring the time of occurrence of falling in the logistic regression led to a conclusion of higher magnitude of the related relative risks compared to standard Cox. The values by the logistic regression were 17.6% (from 1.47 to 1.25) to 39.6% (from 3.28 to 2.35) larger than those from the standard Cox.

Secondly, there were three methods which consider follow-up time, rate of all falls as well as dependence between falls by using robust estimates of variance. That is, the negative binomial regression (2), and the AG (4a) and WLW (5a) extensions of the Cox model. WLW revealed more significant fall-risk factors than both other methods and gave less importance to the history of falls in three months preceding initial interview. Notably, the negative binomial regression with respect to the WLW exhibited a difference of 48.6% (from 3.15 to 2.12) for the variable “two or more prior falls”. The different emphasis on the dependence among multiple event times made by these different approaches explains the difference in results. The negative binomial regression does not integrate at all the length of inter-fall intervals. The AG allows to explicitly model the impact of earlier falls on future events. In this regard, the incidence rate ratio (IRR, virtually equivalent to the so-called hazard ratios) of 1.10 of the time-dependent term “number of previous falls” modelled in the AG (4a) indicate that there was a 10% increase in hazard for each unit increase in number of prior falls. In contrast, WLW estimates separate relationships for each fall and computes the coefficients and the within-subject correlation more directly than does the AG, providing efficient weighted average estimates of effect (and variance).

Thirdly, results were compared for the models with and without time-dependent covariables. The number of home hazards, an exposure particularly likely to vary during the follow-up, was not significantly associated to falls in any of the models with only baseline covariates (1 to 5a). On the contrary, the variable was always statistically significant in the same models that control variation of exposure throughout time (3b to 5b). All survival models with time-varying covariables identified a greater number of fall-risk factors than the corresponding technique with only baseline covariates (3b vs 3a, 4b vs 4a, and 5b vs 5a), even if estimates were calculated from the robust variance. A more marked difference was noted for techniques that models only time-to-first fall than the ones that take into consideration time to each fall. For the marginal WLW model, inattention to time-varying covariables appeared to bias in one direction like in the other. Lastly, results from the usual methods to analyze risk factors for falling (1 and 3, in Table 1) produced considerable biases with respect to the WLW model using time-dependent covariates (5b).

Risk factors for falls

The sample of 937 subjects reported 1,270 falls during a total of 457,283 person-days of observation, given that a same person could report more than one event. Among the subjects, 495 had no falls, 192 experienced one episode, and 250 had more than one. The consideration of the first five falls gathered 90.0% of the 442 fallers and 95.3% of the 937 individuals in the sample. Of all falls for which information on consequences was available, 44.4% resulted in injuries, 25.2% in activity limitations, 17.1% in a medical consultation and 5.6% in a hospitalization. Altogether, 82.1% of falls occurred in the subjects' home.

Table 2 displays the adjusted associations between the potential risk factors and the incidence rate for specific and pooled falls. The WLW marginal risk estimates for the first fall stratum in table 2 are precisely the same as would occur if the analysis was restricted to data on time-to-first fall using a standard Cox model (column 3b, in Table 1). The only difference is that the *p*-values presented in the former were calculated from the robust rather than standard ("naïve") statistics. However, while the estimates for the first fall stratum were essentially equivalent in these two cases, results varied substantially for other strata whether or not coefficients were calculated from the robust compared to the naïve statistic, providing some indication to the degree of dependence among the events. Thus, male sex, residential facility, number of home hazards, Berg balance score and age significantly and independently predict time-to-first fall. For example, the IRR = 1.45 found for the residential facility indicated that the subjects living in such places experienced falls at a rate which was 45% higher than those living in any other kind of housing. Similarly, the IRR of 1.12 for the home hazards indicated that there was a 12% increase in hazard for each unit increase in number of items. However, since age has an IRR less than one (i.e. 0.98), increase in age by 1 year led to decrease in hazard by 2%.

Table 2 also compares the results when distinct ß were fit for each fall. Covariates as age, home hazards and Berg scale show sustained and relatively constant effects across the strata. Some others differ both in the nature and magnitude of the statistically significant variables depending on where in the sequence they occur. The greatest differences in IRR appear for the fifth episode. The entry in the last step of history of falling in the three months prior to study entry turned out to be highly significant and not alter both the magnitude and significance of the IRR for the other variables already included in all stratum models. The right-hand section of Table 2 repeats the analysis under the constraints of overall common ß (weighted average of the event-specific hazards), whether falls beyond the fifth were not applied (censored model) or all fall data were utilized (complete model). The censored model identified seven variables, three more than the time-to-first fall model (BMI, use of benzodiazepines and occasional alcohol consumption in the past six months of follow-up) and one more than the complete model (alcohol consumption). However, these additional variables are no longer significant in the context of the contribution of all others after history of falling is joined to the censored model, whereas use of benzodiazepines and alcohol consumption became insignificant in the complete set. An age-sex interaction term tested in each final model was not significant.

DISCUSSION

This article addresses the proper method of examining falls and their determinants. Even if no statistical technique reproduces human behaviour exactly, makeshift solutions to time-varying exposures and recurrence of events can lead to severe bias. To our knowledge, the first and only example where time-varying exposures and multiple falls were ascertained simultaneously was in a Ph.D. thesis deposited in 1991[22] and published later in a scientific review.[23] However, substantial statistical progress has appeared since.

Methods that may handle the aforementioned data analytical features in a statistically correct way are now available on commercial packages. They have been addressed extensively in the statistical literature, although not yet routinely applied and reported for fall studies. New advances in the statistical world are often slow to reach the clinical and public health fields.[4] We have disputed throughout our paper why the WLW approach is expected to be an appropriate choice in the context of our study. It provides a natural framework for analysing time-varying exposures and multiple events data using minimal assumptions.[2,44] Other authors reported that the WLW is robust and performs quite well for many practical situations.[14]

We illustrated the differences in the estimates obtained by several statistical methods for the analysis of risk factors for falling, according to the information they take into account. Results clearly revealed that usual methods, like are binary outcome using a logistic regression and time-to-first fall using a standard Cox, produced considerable biases with respect to the WLW model using time-dependent covariates. In addition, modeling for first events implicitly assumes that the first event is representative of all events. Our study denotes that this assumption is questionable, in the qualitative facet of IRR estimates more than in the quantitative one. Our results provide additional evidence about the convenient choice of a stratified model rather than a non-stratified, given that the risk of occurrence varies substantially between occurrences. Mahé and Chevret[45] expected such possibilities when the frequency of events per unit is “small”, such as falls among community-dwelling elderly people.

Our results are coherent with earlier findings, although we are more confident with the magnitude in estimates of predictors. A few findings merit commentaries. Number of home hazards and history of falling are strong and consistent predictors of falls, whatever their rank or pooling. Prior overall falls increase the risk of subsequent overall falls. This suggests that if the causes of past falls - for which the variable acts as a proxy - are not corrected, the chances of sustaining further falls due to the same causes are increased.[23] The people living in a residential facility are more at risk than others to fall, possibly because the variable acts as a surrogate measure of various chronic conditions and poorer functional autonomy. In the same way, younger people reveal themselves at higher risk of fall compared to older, probably because of more vigorous lifestyle activities.

Prospective design, frequent contacts, repeated measures and clinical measurements performed by a therapist limited information bias. Nonetheless, some other exposures such as nutrition screening and alcohol use were derived from self-reports. Differential misclassification could occur if the fact that people have experienced a fall or recurrent falls affected the accuracy with which they recalled relevant exposures and subsequent outcomes. It would lead to an exaggeration of the magnitude of the effect on the risk of falling.[6] Also, the length of time between a fall and the measure of follow-up exposure obviously varied according to the day the fall happened. So, accurate assessment of exactly when a change in exposure to time-dependent covariates might have happened between each six-month follow-up was not possible. It would result in nondifferential errors in the measurement of exposures, therefore diluting the observed relation. Another potential for biased results might have occurred because of dropouts, particularly when they do not have the same rate of outcome (risk of falling) as those who continue in the study. With the exception of people having refused to receive the services, who were less likely to fall than the active participants at the end of the study, none of the other reasons for loss to follow-up were associated with the falls. Male sex, ageing, residential facility, first quartile deprivation index, lower Berg score and daily alcohol drinking at baseline were associated with significant shorter duration in participation. As Campbell *et al.*[46] already noted, those who are more frail and may be more at risk of falling are the ones most difficult to involve and sustain in follow-up. This would also lead to an underestimation of the effects.

All the aforementioned reasons lead us to believe that the results observed in our study tend to be conservative. A practical drawback of the WLW is the pre-processing effort and care required in the dataset construction. The application of this method depends on the completeness of the reports of falls and knowledge of calendar dates of falls. Future researches must make the transition from risk factors for falling to community implementation of interventions.

ACKNOWLEDGEMENTS

The authors gratefully thank all older clients and health care workers from the community centres in Lanaudière who participated in the study. We also acknowledge the special contribution of Josée Payette for her effort in preparing the data files used in the analyses, Nancy Leblanc, Julie Meloche and Jean-François Allaire from the Research Centre Philippe Pinel Institute of Montreal for the statistical computations of regression analyses, and Bruce Charles Bezeau for the English revision of the manuscript. The research was sponsored by the Agence de la santé et des services sociaux de Lanaudière.

REFERENCES

1. Ministère de la Santé et des Services sociaux du Québec. La prévention des chutes dans un continuum de services pour les aînés vivant à domicile, Cadre de référence, Québec, Direction générale de la santé publique, 2004, 61 p. Available from: URL http://msssa4.msss.gouv.qc.ca/fr/document/publication.nsf/fb143c75e0c27b69852566aa0064b01c/e69780d1c112de3d85256f55007354f5?OpenDocument (last accessed July 14, 2007).

2. Desquilbet L, Meyer L. [Time-dependent covariates in the Cox proportional hazards model. Theory and practice]. Rev Épidemiol Santé Publique 2005;53(1):51-68. French.

3. Moulton LH, Dibley MJ. Multivariate time-to-event models for studies of recurrent childhood diseases. Int J Epidemiol 1997;26(6):1334-9.

4. Mahé C, Chevret S. Estimation of the treatment effect in a clinical trial when recurrent events define the endpoint. Stat Med 1999;18(14):1821-9.

5. Glynn RJ, Buring JE. Counting recurrent events in cancer research. J Natl Cancer Inst 2001;93(7):488–9.

6. Cumming RG, Kelsey JL, Nevitt MC. Methodologic issues in the study of frequent and recurrent health problems. Falls in the elderly. Ann Epidemiol 1990;1(1):49–56.

7. Goodman AC, Hankin JR, Kalist DE, Peng Y, Spurr SJ. Estimating determinants of multiple treatment episodes for substance abusers. J Ment Health Policy Econ 2001;4(2):65–77.

8. Robertson MC, Campbell AJ, Herbison P. Statistical analysis of efficacy in falls prevention trials. J Gerontol A Biol Sci Med Sci 2005;60(4):530–4.

9. Wang SJ, Winchell CJ, McCormick CG, Nevius SE, O'Neill RT. Short of complete abstinence: an analysis exploration of multiple drinking episodes in alcoholism treatment trials. Alcohol Clin Exp Res 2002;26(12):1803–9.

10. Hogan DB, MacDonald FA, Betts J, Bricker S, Ebly EM, Delarue B, Fung TS, Harbidge C, Hunter M, Maxwell CJ, Metcalf B. A randomized controlled trial of a community-based consultation service to prevent falls. CMAJ 2001;165(5):537–43.

11. Glynn RJ, Buring JE. Ways of measuring rates of recurrent events. BMJ 1996;312(7027):364–7.

12. Lin DY. Cox regression analysis of multivariate failure time data: the marginal approach. Stat Med 1994;13(21):2233–47.

13. Finkelstein DM, Schoenfeld DA, Stamenovic E. Analysis of multiple failure time data from an AIDS clinical trial. Stat Med 1997;16(8):951–61.

14. Wei LJ, Glidden DV. An overview of statistical methods for multiple failure time data in clinical trials. Stat Med 1997;16(8):833-9.

15. Cleves M. How do I analyze multiple failure-time data using Stata?, Stata FAQ, 2002, 11 p. Available from: URL http://www.stata.com/support/faqs/stat/stmfail.html (last accessed July 14, 2007).

16. Andersen PK, Gill DR. Cox's regression model for counting processes. Ann Statist 1982;10(4):1100-1120.

17. Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. Biometrika 1981;68(2):373-379.

18. Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modelling marginal distributions. J Am Stat Assoc 1989;84(408):1065-73.

19. Lancar R. [Robust analysis methods for multivariate survival times]. Rev Épidemiol SantÉ Publique 1999;47(3):287-96. French

20. Roberston MC. Development of a falls prevention programme for elderly people : evaluation of efficacy, effectiveness, and efficiency . A Ph.D. thesis submitted at the University of Otago, Department of Medical and Surgical Sciences, Dunedin, New Zealand, 2001.

21. D'Agostino RB, Lee ML, Belanger AJ, Cupples LA, Anderson K, Kannel WB. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. Stat Med 1990;9(12):1501-15.

22. O'Loughlin J. The incidence of and risk factors for falls and fall-related injury among elderly persons living in the community. A Ph.D. thesis at the McGill University, Department of Epidemiology and Biostatistics, Montreal, Québec, Canada, 1991, 270 p.

23. O'Loughlin JL, Robitaille Y, Boivin JF, Suissa S. Incidence of and risk factors for falls and injurious falls among the community-dwelling elderly. Am J Epidemiol 1993;137(3):342-54.

24. Nevitt MC, Cummings SR, Kidd S, Black D. Risk factors for recurrent nonsyncopal falls. A prospective study. JAMA 1989;261(18):2663-8.

25. Fletcher PC, Hirdes JP. Risk factors for falling among community-based seniors using home care services. J Gerontol A Biol Sci Med Sci 2002;57(8):M504-10.

26. Bégin C. Projet-pilote régional de prévention des chutes à domicile chez les personnes âgées, Devis d'implantation dans les CLSC, Saint-Charles-Borromée, Service de prévention et de promotion, Direction de santé publique, Régie régionale de la santé et des services sociaux de Lanaudière, 2002, 120 p.

27. Pampalon R, Hamel D, Raymond G. Indice de défavorisation pour l'étude de la santé et du bien-être au Québec - Mise à jour 2001. Institut national de santé publique du Québec, 2004, 11 p. Available from: URL http://www.inspq.qc.ca/pdf/publications/295-IndiceDefavorisation_2001.pdf (last accessed December 24, 2006).

28. Leclerc BS, Marquis G, Payette J. Tableau de bord lanaudois sur la défavorisation. Territoire de Lanaudière. Calibrage à l'échelle de la région de Lanaudière, Joliette. Agence de la santé et des services sociaux de Lanaudière, Direction de santé publique et d'évaluation, Service de surveillance, recherche et évaluation, 2005, 87 p.

29. Payette H. Développement, validation et évaluation d'un programme de dépistage nutritionnel pour les personnes âgées en perte d'autonomie vivant dans la communauté, Sherbrooke, Centre de recherche en gérontologie et gériatrie, Centre d'expertise en gérontologie et gériatrie inc., Institut universitaire de gériatrie de Sherbrooke, non daté, pages multiples.

30. Payette H, Guigoz Y, Vellas BJ. Study design for nutritional assessments in the elderly », in Methods in Aging Research, B.P. YU (ed), Boca Raton (Florida), CRC Press LLC, 1999, p. 301-20.

31. Laporte M, Villalon L, Payette H. Simple nutrition screening tools for healthcare facilities: development and validity assessment. Can J Diet Pract Res 2001;62(1):26-34.

32. Laporte M, Villalon L, Thibodeau J, Payette H. Validity and reliability of simple nutrition screening tools adapted to the elderly population in healthcare facilities. J Nutr Health Aging 2001;5(4):292-4.

33. Berg K. Balance and its measure in the elderly: a review. Physiother Can 1989;41(5):240-6.

34. Berg KO, Maki BE, Williams JI, Holliday PJ, Wood-Dauphinee SL. Clinical and laboratory measures of postural balance in an elderly population. Arch Phys Med Rehabil 1992;73(11):1073-80.

35. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. Can J Public Health 1992;83(suppl. 2):S7-11.

36. Berg K, Wood-Dauphinee S, Williams JI. The Balance Scale: reliability assessment with elderly residents and patients with an acute stroke. Scand J Rehabil Med 1995;27(1):27-36.

37. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991;39(2):142-8.

38. Lin MR, Hwang HF, Hu MH, Wu HD, Wang YW, Huang FC. Psychometric comparisons of the timed up and go, one-leg stand, functional reach, and Tinetti balance measures in community-dwelling older people. J Am Geriatr Soc 2004;52(8):1343-8.

39. Gill TM, Williams CS, Robison JT, Tinetti ME. A population-based study of environmental hazards in the homes of older persons. Am J Public Health 1999;89(4):553-6.

40. Gill TM, Williams CS, Tinetti ME. Environmental hazards and the risk of nonsyncopal falls in the homes of community-living older persons. Med Care 2000;38(12):1174-83.

41. Chevalier S, Lemoine O. Consommation d'alcool, in Enquête sociale et de santé 1998, Québec, Institut de la statistique du Québec, 2000, p. 117-33. (Collection La santé et le bien-être)

42. Institut de la statistique du Québec Annexe 3. Questionnaire autoadministré (QAA) VI - L'alcool, in Enquête sociale et de santé 1998, Québec, Institut de la statistique du Québec, 2000, p. 15-7. (Collection La santé et le bien-être)

43. Allison PD. Survival analysis using SAS: A practical guide. Cary NC: SAS Institute inc., 1995, 304 p.

44. Li QH, Lagakos SW.Use of the Wei-Lin-Weissfeld method for the analysis of a recurring and a terminating event. Stat Med 1997;16(8):925-40.

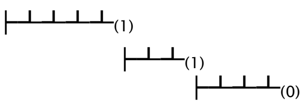
45. Mahé C, Chevret S. Analysis of recurrent failure times data: should the baseline hazard be stratified? Stat Med 2001;20(24):3807-15.

46. Campbell AJ, Robertson MC, Gardner MM, Norton RN, Buchner DM. Falls prevention over 2 years: a randomized controlled trial in women 80 years and older. Age Ageing 1999;28(6):513–8.

Standard Cox regression. One data record covers entry until the 1st fall and discards any information past that point. Total follow-up time is assigned to individual that never fell. The dependent variable is *time-to-first fall*.

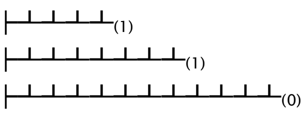
$$\log\left(\frac{\lambda_i(t)}{\lambda_0(t)}\right) = \beta' x_i(t)$$

Andersen-Gill Cox regression. Three records cover entry until the 1st fall, from the 1st until 2nd fall, from the last fall to the end of follow-up, the latter period being fall-free. The dependent variable is *time to each fall*.



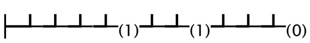
$$\log\left(\frac{\lambda_i(t \mid x_i(t))}{\lambda_0(t)}\right) = \beta' x_i(t)$$

Marginal Wei, Lin & Weissfeld regression. Three records. Each fall as well as the final fall-free period are treated in an independent stratum and time measured from entry. The dependent variable is *time to each fall*.



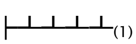
$$\log\left(\frac{\lambda_{ik}(t \mid x_{ik}(t))}{\lambda_{0k}(t)}\right) = \beta_k x_{ik}(t)$$

Negative binomial regression. One record covers entry until the end of follow-up and includes simply the total number of falls and follow-up time per subject. The dependent variable is *number of falls*.



$$\log\left(\frac{k_i}{pt_i}\right) = \beta_0 + \beta' x_i$$

Logistic regression. One data record, which does not account for follow-up time and ignores multiple falls by subject. The binary dependent variable is *status of faller*.



$$\log\left(\frac{\frac{p_e}{1-p_e}}{\frac{p_u}{1-p_u}}\right) = \beta' x_i$$

Figure 1. Schematic representation of statistical models for the study of risk factors for falls.
(Modified from a figure published by Robertson, Campbell and Herbison[8]).
Hypothetical subject with follow-up of 12 days, falls on day 5 and 8. Let (0) represent no fall and (1) a fall, x_i a risk factor of subject i measured at time t, and k_i its number of falls. Then the baseline hazard is illustrated by $\lambda_0(t)$, the hazard for a fall for the i^{th} subject by λ_i and the hazard of the k^{th} fall for this subject by λ_{ik}. Person-time, pt_i is length of time at risk for subject i, $\beta' x$ denotes the effect size of factor x, p is the probability of event in exposed, e, and unexposed, u, subjects.

Table 1. Adjusted relative risk estimates of factors for falls among the community-dwelling elderly, according to different statistical regression methods

Risk factor	*1* Logistic (fallers[c])	*2* Negative binomial (all falls)	*3a* Standard Cox/ WLW (first fall)	*4a* AG Cox[a] (all falls)	*5a* WLW (all falls)	*3b* Standard Cox/ WLW (first fall)	*4b* AG Cox[b] (all falls)	*5b* WLW (all falls)
	with baseline covariates [d]					with time-varying covariates [e]		
Home hazards (nb)	—	—	—	—	—	1.12****	1.08***	1.19****
BMI (kg/m^2)	—	—	—	0.98**	—	—	0.98**	0.99*
Berg score	—	0.98**	—	0.98****	0.98***	0.99**	0.98****	0.99****
Timed Up & Go score	—	—	—	—	0.99**	—	—	—
Male	1.47*	—	1.25*	—	1.22*	1.34**	—	1.30**
Age (yrs)	—	0.97**	—	0.97**	0.98**	0.98*	0.97**	0.98***
Residential facility housing	—	—	—	—	1.29*	1.45**	—	1.61****
One prior fall [†]	1.95***	1.41*	1.47***	1.40**	1.26*	1.45***	1.37*	1.24*
Two or more prior falls [†]	3.28****	3.15****	2.35****	2.31****	2.12****	2.07****	2.21****	1.86****

Significant (two-tailed): $*p \le 0.05$; $**p \le 0.01$; $***p \le 0.001$; $****p \le 0.0001$

Included number of previous falls during the follow-up as a time-dependent covariate to account for dependence between falls:
[a]IRR=1.10****; [b]IRR=1.09****.

[c]Subjects monitored less than 12 months who did not declare any falls (n = 221) were excluded, given that we could not define the status of faller.

[d]All covariates measured on baseline only;[e] up-dated covariates included home hazards, BMI, Berg and Timed Up & Go scales.

[†]History of falls in 3 months preceding initial interview.

Table 2. Adjusted and variance-corrected WLW incidence rate ratio by selected risk factors for falls among the community-dwelling elderly, according to the fall rank or pooled fall group

Risk factor	Fall rank number					F^{st} 5 falls	All falls
	1	2	3	4	5		
	$n^a = 937$	$n = 429$	$n = 244$	$n = 140$	$n = 93$	$n = 1843$	$n = 2169$
Falls (nb)	442	250	144	99	66	1001	1270
Home hazards (nb)	1.12****	1.19****	1.20****	1.17***	1.36****	1.16****	1.19****
BMI (kg/m^2)	—	—	0.95***	—	—	[0.99*][b]	0.99*
Berg balance score	0.99**	0.98***	0.98**	—	0.97***	0.98****	0.99****
Benzodiazepine use	—	1.37*	—	—	—	[1.22**]	[1.21*]
Alcohol use, past 6 months							
⩽2 times per month vs other categories	—	—	1.50*	—	—	[1.20*]	—
Male	1.34**	—	—	—	2.02**	1.28**	1.30**
Age (yrs)	0.98*	—	0.97*	0.96*	—	0.98**	0.98***
Residential facility housing	1.45**	1.70**	—	—	2.52*	1.51***	1.61****
Material deprivation index							
Fourth vs other quartiles	—	—	—	—	3.81****	—	—
One fall prior initial interview [c]	1.45***	—	—	—	—	1.37***	1.24*
≥ 2 falls prior initial interview [c]	2.07****	1.65**	2.15****	1.49*	—	1.95****	1.86****

Significant (two-tailed): *$p \le 0.05$; **$p \le 0.01$; ***$p \le 0.001$; ****$p \le 0.0001$

[a] n of subjects at risk for the considered fall stratum or pooled fall group.

[b] The brackets show the variables not reached statistical significance after “previous falls” were introduced.

[c] History of falls in 3 months preceding initial interview.