Canadian Institute of Actuaries
Institut canadien des actuaires

Study

Predictive Analytic Models for Canadian Group Disability Termination Experience

Research Council –

Experience Research Committee

May 2019

Document 219060

Ce document est disponible en français

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

This analysis of termination experience under Canadian group long-term disability (LTD) policies was conducted by the Research Council of the Canadian Institute of Actuaries (CIA).

The mandate for this research assignment was to extend the results of the 2018 LTD Termination Study, which produced updated LTD termination tables using experience from the period 2009–2015.

The new disability tables (like earlier versions) vary by gender, age, duration of claim, and region (Québec and other). The objective of this project was to create factors, formulae, or methods that, in conjunction with the four tables, could be used to incorporate the effect of the following variables and pairs and groups of the variables into the estimation of termination rates:

1. Diagnosis;
2. Age;
3. Gender;
4. Residence;
5. Duration of Disability;
6. Industry Code;
7. Monthly Benefit;
8. Salary;
9. Tax Status;
10. CPP/QPP Integration – status and amount;
11. Workers' Compensation Integration;
12. Pre-LTD Benefits – yes/no and type;
13. Initial Definition of Disability;
14. Maximum Benefit Duration; and
15. Elimination Period.

The CIA retained Fraser Group and Denis Garand & Associates to act as the study managers. Their mandate was to:

- Conduct a literature review;
- Explore various analytic techniques that might be used to create the desired predictive tools;
- Create one or more practical tools that an actuary could implement based solely on a reading of the report; and
- Prepare appropriate documentation, including this report.

PROJECT GOVERNANCE

Chair of the Research Council:

Keith Walter.

The Project Oversight Group responsible for this project consisted of:

Frank Reynolds (Chair);

Jean-François Blais;

Pierre-Philippe Carle-Mossdorf;

Erin Crump;

Rhys DeGrave;

Lina Forner;

Tim Griffin;

Kateri Laneuville;

Stella-Ann Ménard; and

Keith Walter (liaison to the Research Council).

PROJECT ACTIVITY

This research project relied on the database that had been previously created for the development of the Group LTD Termination Tables (2005–2015). The data for that study came from 16 insurance companies accounting for approximately 99% of the Canadian market for group LTD insurance.

Insurance Companies Contributing Data

- Assumption Life;
- Blue Cross Life;
- Co-operators Life;
- Desjardins Financial;
- Empire Life;
- Equitable Life;
- Great-West Life;
- Industrial Alliance;
- Humania;
- La Capitale;
- Manulife;
- Pacific Blue Cross (BC Life);
- RBC Life;
- SSQ;
- Sun Life; and
- Wawanesa Life.

Project Team

As mentioned above, the CIA retained Fraser Group and Denis Garand & Associates to act as the study managers. The project leaders were Ken Fraser and Denis Garand, FCIA.

The project team also included:

- Donna Swiderek, ACIA, from Denis Garand & Associates;
- Clayton Zaluski, FCIA; Stephen Swenarchuk, ACIA; and Merv Worden, FCIA, from Worden Zaluski Consulting Actuaries; and
- Taehan Bae, PhD, ACIA, from the University of Regina.

TERMINOLOGY

This section discusses key terms used throughout this study.

A/E means *Actual to Expected* and normally refers to a ratio between the number of actual claim terminations and the number of expected terminations computed from a reference table applied to the exposure.

Any Occ and *Own Occ* refer to the definitions of disability being used in the LTD contract. *Own Occ* defines disability as the inability of the claimant to perform the essential duties of his *own* occupation while *Any Occ* defines disability as the inability of the claimant to perform the duties of *any* occupation for which the employee is qualified by training, education, or experience.

CiD, or *change in definition*, refers to the provision in most LTD contracts that shifts the definition of disability from Own Occ to the more stringent Any Occ basis after an initial period of disability (usually two years). Thus, an individual may qualify for disability benefits for a certain period and then be ineligible for benefits even though there has been no change in the medical or vocational evidence.

Designated Tables refers to the Group LTD termination tables published in the most recent (2009–2015) CIA study, the Group Long-term Disability Termination Study (Document 219012).

Exposure has its usual actuarial sense and refers to claims that are active and thus "exposed" to a contingent termination event. Exposure is quantified as the number of claims (rather than amount of benefit).

LTD means *long-term disability insurance*. In this study, it exclusively refers to coverage provided on a group basis. This is discussed in greater detail in the next section, DATA RESOURCES.

Recovery is used in this study to refer to any termination that is not due to mortality. While this includes the plain-language meaning (i.e., claimants have made a medical recovery from their injury or illness and have returned to work), *recovery* in this study also includes any situation where a claim was terminated by an insurance company other than for death. Notably, this includes CiD scenarios where the claimant no longer qualifies under a more stringent definition of disability. It would also include situations where claimants abandon a claim by not submitting required information, and where the insurance company determines that the evidence does not support the continued payment of benefits.

STD refers to Short Term Disability benefits. These are also known as Weekly Income benefits.

Tables Study refers to the CIA study on LTD termination experience for the years 2009–2015: Group Long-term Disability Termination Study (Document 219012).

Termination refers to any contingent event that terminates an otherwise active claim. Thus, a claim that ends due to the attainment of a maximum benefit period (e.g., age 65) is not a termination. In this study, *termination* is used to include both mortality and recovery (see above).

DATA RESOURCES

No data collection or validation was required for this project since we used the database developed for the Tables Study. The reader should refer to the Tables Study[1] for further details on data collection and validation and table construction.

This section summarizes the conceptual framework for the data made available to this project. The claims used in this project are characterized by the following:

- They are from group policies issued to Canadian employers, multiple-employer trusts, and union welfare trusts; and
- They have a date of disability prior to December 31, 2015, and were "in payment" for some period between January 1, 2009, and December 31, 2015.

External Data

We considered the possibility of including external data such as inflation, interest, and unemployment rates in this project. However, our research indicated that these influences have remained relatively stable over the period covered by the experience data. Incorporating these elements into the study would have been resource-intensive with little expectation of useful results. However, this may be an area for exploration in future investigations.

Designated Tables

The mandate of this project is to create adjustment factors that can be used in conjunction with published tables from the Tables Study to better predict termination experience. Thus, these Designated Tables can be considered a part of the data resources for this project.

The construction of a table from raw data requires decisions on many practical issues and often there is some tension among competing technical objectives. For this reason, it is useful to articulate the expected uses that guided the project team as documented in that report.

These uses are identified as:

- Valuation by insurance companies of Canadian LTD open claim liabilities in financial statements;

[1] The primary authors of this study also authored the Tables Study.

- Calculation by insurance companies of claim liabilities in the experience-rated accounting for specific policyholders;
- Use by insurance companies in the development of manual rates for group LTD benefits; and
- Valuation by self-funded plan sponsors of Canadian LTD open claim liabilities in financial statements.

The authors of the Tables Study set out the following description of the scenario for which the Designated Tables might be considered an optimal fit. The elements of this model include:

- Employee benefits plans;
- Written on an insured basis (i.e., not Administrative Services Only);
- Canadian employees;
- Groups of varying size;
- High levels of enrolment;
- Primarily guaranteed issue with individual underwriting of excess amounts;
- Elimination periods of four to six months;
- Benefits payable to age 65;
- High replacement ratios but less than 100%;
- Two-year Own Occ definition of disability;
- Industry-standard provisions for recurrent disability, all sources limits, rehab, etc.; and
- Industry-standard claim management practices such as early intervention.

Users are alerted that there may be a need for adjustments if they face a situation that is widely variant from the model scenario; for instance, long waiting periods or variant contract provisions.

DESIGNATED TABLES STRUCTURE

This section summarizes the structure of the Designated Tables. The factors developed in this project are intended to adjust the Base Table values.

In the Designated Tables, termination values are developed by adding factors from two component tables:

- Base Table; and
- CiD Adjustment.

Rates are provided separately for:

- Total terminations;
- Terminations due to death (Mortality); and
- Terminations for other reasons (Recovery).

The tables provided are in four sections segmented by:

- Québec versus Rest of Canada; and
- Female versus Male.

Each section contains:

- Select values for the first 120 months of disability. These provide monthly resolution for five-year age groups (age at disability) from months five to 60. Annual resolution is provided for the final five.
- Ultimate values for durations beyond 10 years. These are by gender and by attained age and are not differentiated by Québec/Rest of Canada.

Rates are shown as monthly values for the first 60 months and annual thereafter.

INITIAL INVESTIGATIONS

Several predictive analytic techniques were considered for use in this project. These included:

- Generalized linear models (GLMs): logistic regression (LR);
- Survival models; and
- Minimum bias procedure (MBP).

In each case, we attempted to construct a model within the framework of the given technique and to use the model to produce the desired tools. The models were constructed using Python and R and evaluated based on the criteria described in the following sections.

Generalized Linear Models: Logistic Regression

GLMs are widely used in property and casualty insurance due to their predictive power. GLMs have been used in the United States to develop mortality tables. By using a "logit" link function and assuming the error term follows the binomial distribution, a specific type of GLM called a logistic regression model may be constructed.

LR models predict a binary dependent variable: 0 or 1. An LR model can be constructed for an LTD termination study where:

- 0 means the claim did not terminate; and
- 1 means the claim did terminate.

The LR model produces an equation to predict claim termination probability based on any number of independent variables. Through testing and construction of LR models, we determined the following strengths and weaknesses:

Table 1 Strengths and Weaknesses of LR Models

Strengths	Weaknesses
• Predictive power is well documented • Easily produce useful statistics (p-values, confidence intervals) • Provide inherent "smoothing", eliminating the need for algebraic methods	• Complex • Objective is to develop factors, formulae, or methods to adjust Designated Table values; LR does not accomplish this • Difficult to account for the effect of CiD

Ultimately, it was decided that an LR model would not be pursued because it would not meet the study objective. An LR model would produce new base rate values that would be challenging to tie back to the Designated Tables. However, we did explore the use of an LR model as a method to develop the base termination rate tables (by age, gender, claim duration, and Québec/Non-Québec). This is described in more detail in Appendix 2.

Survival Models

Survival models estimate the distribution of the random variable T (time to termination). Both non- and semi-parametric methods can be used, and we explored each possibility.

The most common non-parametric method is the Kaplan-Meier product limit estimator. This method is based on the number of claims active just prior to each observed termination.

The Cox Proportional Hazard Model (CPHM) is a common semi-parametric model used in duration analysis due to its ability to incorporate the baseline hazard function and a linear combination of predictor variables. It allows formal hypothesis testing to be performed regarding the significance of predictor variables.

Our experience with survival models leads us to the following strengths and weakness:

Table 2 Strengths and Weaknesses of Survival Models

Strengths	Weaknesses
• Theoretically effective • Can measure predictive power of a set variables • Provide inherent "smoothing", eliminating the need for algebraic methods	• Computationally intensive • Difficult to calibrate the model fit for specific durations • Difficult to isolate the effect of CiD • CPHM proportionality assumption does not hold in this dataset

Minimum Bias Procedure

The MBP is a predictive modelling method that is commonly used for insurance rate making. A multiplicative MBP model generates factors for each predictor variable.

The process is iterative and requires the input of Actual and Expected values. The Expected value is the probability of termination as determined by the Designated Tables. The Actual value is a 1 if the claim did terminate and 0 if the claim did not terminate. The actual and expected amounts must be determined for each claim at each duration.

We consider the following to be the strengths and weaknesses of the MBP:

Table 3 Strengths and Weaknesses of MBP Models

Strengths	Weaknesses
• Familiar output • Can be directly applied to the Designated Tables termination rates • Readily fulfill study objective	• Not as sophisticated as other techniques • Do not produce statistical measures of fit • The results implicitly assume that the value of each variable is the same impact at all durations

Having evaluated the outcome of the initial investigations it was decided to focus on the MBP as being best suited to fulfill the study objectives.

MODEL CONSTRUCTION

This section outlines how we used the MBP to develop the adjustment factors recommended in this report.

Data Description

The data from the Tables Study provided the following variables that could be used in a predictive model (as indicated, we selected some and excluded others):

Table 4 Data Variables Available

Data Variable	Included in Model	Reason
Age	No	Included in Designated Tables
Gender	No	Included in Designated Tables
Duration of Disability	No	Included in Designated Tables
Initial Definition of Disability	No	Included in Designated Tables
Province	Yes	
Diagnosis	Yes	
Industry	Yes	
Monthly Benefit	Yes	
Pre-LTD Benefits	Yes	
Elimination Period	Yes	
Maximum Benefit Duration	No	Included in exposure measurement
Salary	No	Correlated with Monthly Benefit
Tax Status	No	Poor data quality

CPP/QPP Integration	No	Poor data quality
Workers' Compensation	No	Poor data quality

Variables included in the Designated Tables were not included in this project to avoid double counting. Monthly benefit amounts are typically calculated as a percentage of monthly salary. Monthly Benefit was included and Salary[2] was excluded to avoid double counting.

Maximum Benefit Duration is used to calculate exposures and was excluded from the study to avoid double counting.

Data regarding Tax Status, CPP/QPP Integration, and Workers' Compensation were not consistently provided by all companies and contained many unknowns and suspect entries. These fields were excluded to maintain the integrity of results.

In total, there were six variables that had relatively robust data and were not already built into the structure of the Designated Tables:

- Province;
- Diagnosis;
- Industry ;
- Monthly Benefit;
- Pre-LTD Benefits; and
- Elimination Period.

These six variables were included in the analysis.

[2] The Tables Report indicated that the Salary data was not consistently available for nearly 30% of the data.

Association among Variables

Since the purpose of the proposed model is to unravel the effects of correlated predictor variables, we reviewed the variables for independence or correlation.

In the first stage, using the chi-square test, the independence hypothesis between every pair of predictor variables was rejected since the p-values were all close to zero.[3]

The second stage involved assessing the degree of correlation among variables. The strength of association between variables can be measured by Cramér's V statistic.[4] Cramér's V statistic is similar to the r^2 statistic but is suitable for use with categorical variables. As with the r^2 statistic, a value of 0.00 indicates no association and a value of 1.00 indicates perfect association.

A value of 0.25 is a typical threshold used to indicate a moderate association while values below 0.15 indicate very weak association.

The following table shows that the level of association ranges from 0.055 (Diagnosis – Elimination Period) to 0.253 (Industry – Elimination Period).

Table 5 Strength of Association – Cramér's V Statistic

	Elimination Period	Pre-LTD Benefits	Monthly Benefit	Diagnosis	Province
Industry	0.25	0.18	0.18	0.06	0.09
Elimination Period		0.18	0.16	0.06	0.09
Pre-LTD Benefits			0.18	0.09	0.22
Monthly Benefit				0.07	0.14
Diagnosis					0.08

Given that there is some association between some pairs of variables, it is expected that a predictive analytical model will add value by identifying the relative contribution of each variable.

[3] Note, however, that a small p-value does not necessarily mean that there is a strong dependence between variables. It simply indicates the independence assumption is not appropriate based on the data. For a large dataset, even a very small association may result in a very low p-value.

[4] Cramér's V statistic is calculated as:

$$\sqrt{\frac{\chi 2}{N(k-1)}}$$

where
χ^2 is the Pearson chi-square statistic;
N is the sample size involved in the test; and
k is the lesser of the number of categories in either variable.

BUILDING THE MODEL

Data Manipulation

The MBP Model requires expansion of claim data to create a record for each claim month of exposure. The data table was transformed from 485,000 claim records to 10.8 million claim exposure month records. This process was completed using logic consistent with the Tables Study with one notable difference. The Tables Study calculated precise exposure to the day, while the MBP requires exposure to be calculated for each duration month. This resulted in slight differences compared to the Tables Study that are noted below.

The MBP requires that all variables be categorical. Continuous variables like Benefit Amount have been grouped into buckets or ranges to achieve this. Ranges were determined based on exposure and reasonable consistency with the Tables Study analysis.

Model Construction Algorithm

The first step was to adjust Expected terminations so that the overall A/E ratio is equal to 1.00. To do this, the Expected terminations were multiplied by a flat factor. For our study, this flat factor (the "weighting variable") was 1.02, which indicates that the Actual terminations are 2% higher than the Expected terminations derived from the Designated Tables. This difference is due to exposure being calculated to the nearest month in this project compared to the Tables Study, where it was calculated to the day.

The MBP Model is built in an iterative process with predictive variables being added one step at a time. In each step, the MBP software[5] adjusts Expected claims so that the A/E ratio is equal to 1.00 for each category within an independent variable. Eventually the A/E ratios converge to 1.00, meaning no more adjustments are required.

The final factor output is derived from the product of adjustment factors used in each MBP iteration.

The output of the MBP process is a set of multiplicative adjustment factors that can be used to alter expected termination rates for claims with given characteristics. For example, the industry category White Collar and Professional has an MBP Industry factor of 1.018 (from Table 6 below). In valuing a claim, the termination rates from the Designated Tables would be multiplied by 1.018 if the claimant was in the White Collar and Professional category. Additional factors would be applied based on each of the other predictor variables.

The factors for all variables are set out in the following table.

[5] Proprietary software built for this project.

Table 6 MBP Predictive Model Factors – Version 1 (All Durations)

INDUSTRY	MBP	EXPOSURE	
Heavy Blue Collar	1.038	1,332,695	12.3%
Manufacturing	0.994	1,522,271	14.1%
Wholesale, Retail Trade	1.019	1,265,085	11.7%
White Collar and Professional	1.018	1,616,981	14.9%
Health, Education, Social Services	1.025	1,312,515	12.1%
Other Services (Private Sector)	0.990	875,876	8.1%
Public Administration	0.920	1,802,829	16.7%
Unknown	1.017	1,088,653	10.1%

ELIMINATION PERIOD	MBP	EXPOSURE	
0 to 3 months	0.948	2,101,997	19.4%
4 months	1.018	4,772,036	44.1%
5 to 6 months	1.012	2,746,656	25.4%
Greater than 6 months	0.968	1,196,216	11.1%

PRE-LTD BENEFITS	MBP	EXPOSURE	
Our STD	1.181	1,936,321	17.9%
Other or None	0.939	8,880,584	82.1%

BENEFIT AMOUNT		MBP	EXPOSURE	
Unknown		1.082	460,056	4.3%
Less than $1,499		1.009	2,022,371	18.7%
$1,500 to $1,999		0.975	2,102,488	19.4%
$2,000 to $2,499		1.003	1,877,541	17.4%
$2,500 to $3,249		1.018	2,111,827	19.5%
Greater than $3,250[6]		0.969	2,242,622	20.7%

DIAGNOSIS		MBP	EXPOSURE	
Mental Disorders		1.026	3,072,438	28.4%
Musculo-skeletal		0.900	2,371,654	21.9%
Neoplasms (Cancers)		1.236	1,016,681	9.4%
Circulatory		0.854	877,167	8.1%
Nervous System		0.526	1,200,137	11.1%
Accidents		1.219	772,068	7.1%
All Other Identified Causes		1.049	1,336,964	12.4%
Not Stated or Unknown		1.059	169,796	1.6%

PROVINCE		MBP	EXPOSURE	
British Columbia		0.999	1,286,554	11.9%
Alberta		1.189	1,100,046	10.2%
Saskatchewan		1.242	281,222	2.6%
Manitoba		1.112	359,099	3.3%
Ontario		0.966	4,299,717	39.7%
Québec[7]		0.976	2,352,523	21.7%
Other Canada		0.906	1,137,744	10.5%

The Province variable warrants some additional discussion. The Designated Tables provide separate values for Québec and the Rest of Canada. These tables provided the expected values used in building the MBP Model. Consequently, the MBP Province factor for Québec (0.976 in the table) is not intended to represent the difference between Québec and the Rest of Canada. Since the Designated Tables already account for major differences between Québec and the Rest of Canada, the MBP Province factor is only adjusting for the residual difference. It is less than 1.00 because a small amount of the difference between Québec and the Rest of Canada is accounted for by the other variables in the model.

---

[6] Data groupings were selected to provide sufficient credibility in each category and to limit the complexity of the final model. Table 25 of the Tables Study indicates that there are only modest variations in A/E ratios until monthly amounts reach $15,000.

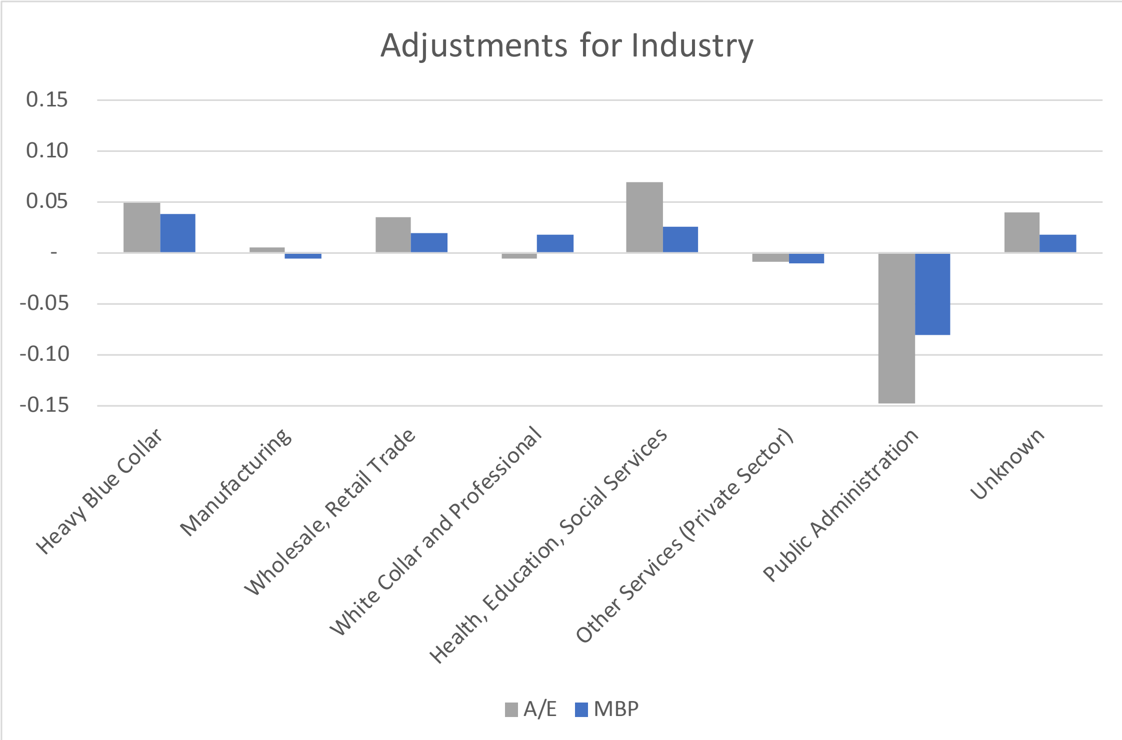
[7] Note that the indicated MBP factor is applied to the termination rates contained in the Québec portion of the Designated Tables. Although the adjustment factors are similar in Ontario and Québec, this does not indicate that termination behaviour is similar in the two provinces.

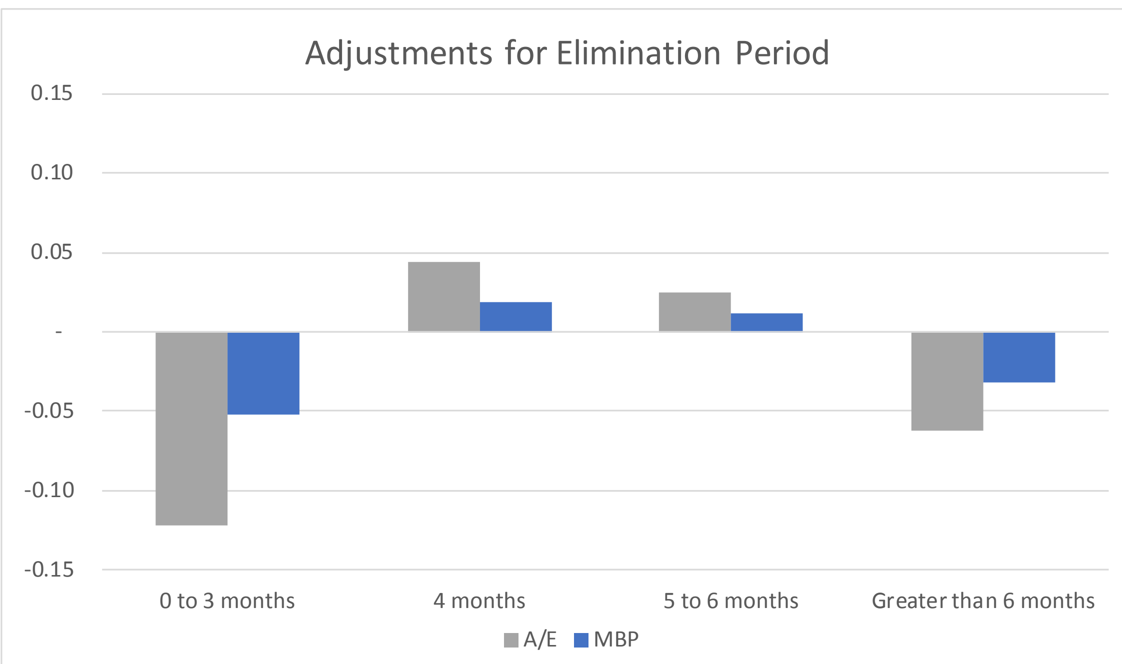
The value of the MBP Model compared with traditional one-way A/E analysis is that MBP factors have removed the effect of all other predictive variables. For example, in Table 7 below, the one-way analysis suggests that termination rates for the Health, Education, Social Services industry sector are 8% higher than for White Collar and Professional (1.07-0.99 = 0.08). The MBP Model indicates no difference (MBP adjustment factor is 1.02 for both). The apparent difference is actually due to other factors (Diagnosis, perhaps).

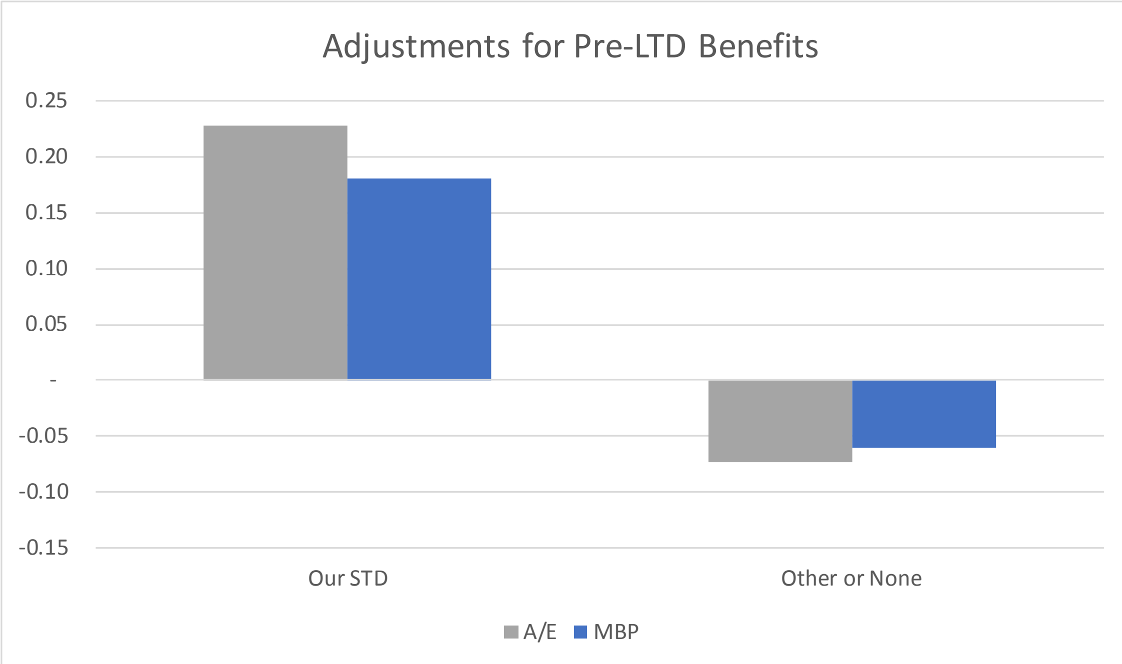
Table 7 Example – MBP versus Traditional A/E

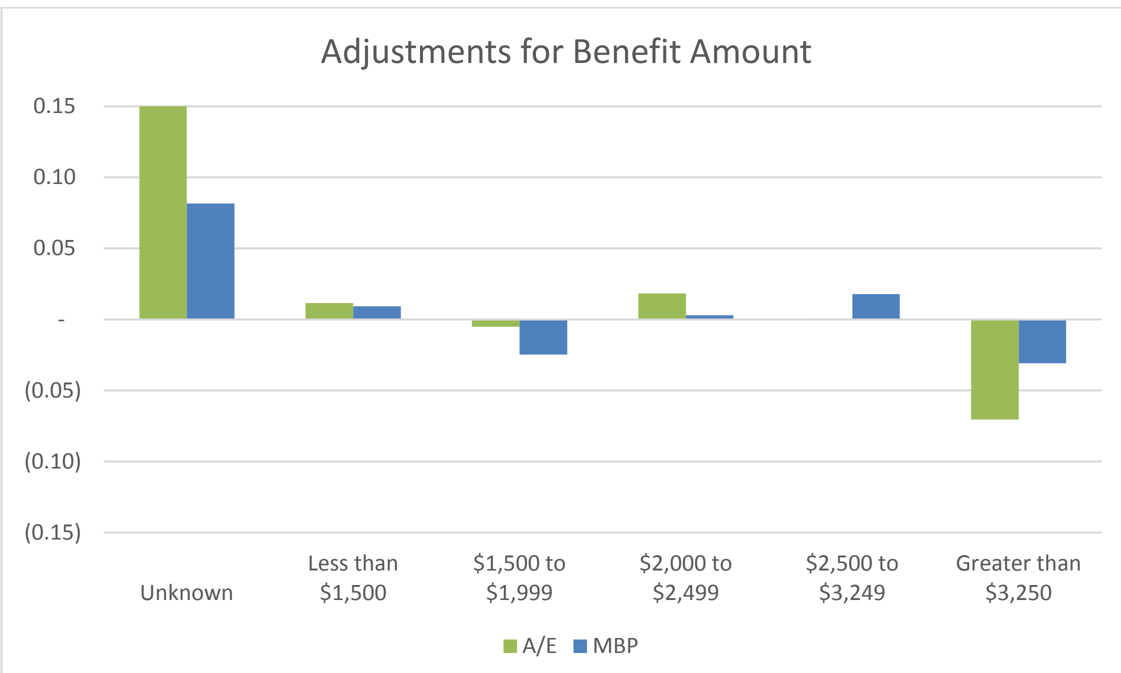
	A/E Factor	MBP Factor	Difference
Heavy Blue Collar	1.05	1.04	-0.01
Manufacturing	1.01	0.99	-0.01
Wholesale, Retail Trade	1.04	1.02	-0.02
White Collar and Professional	0.99	1.02	0.02
Health, Education, Social Services	1.07	1.02	-0.04
Other Services (Private Sector)	0.99	0.99	-0.00
Public Administration	0.85	0.92	0.07
Unknown	1.04	1.02	-0.02

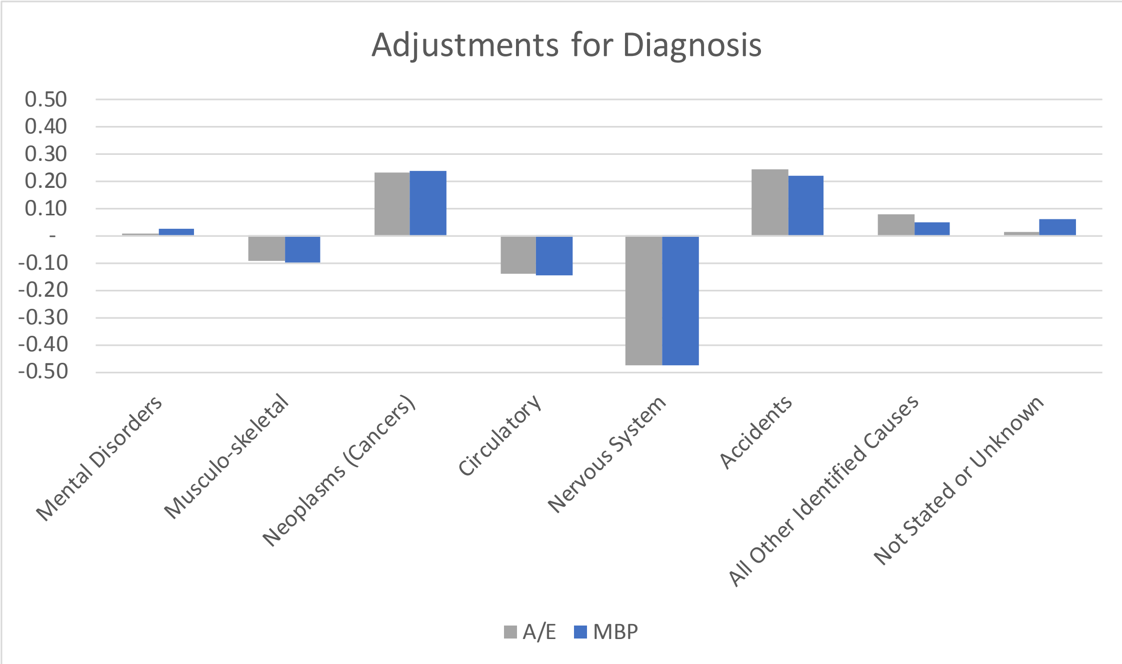
The following charts compare the A/E derived factors with the MBP factors for each variable. As a general observation, the MBP factors will show similar or less variability than the one-dimensional A/E analysis. The factors will be similar for a variable when there is very little correlation or association with other variables.

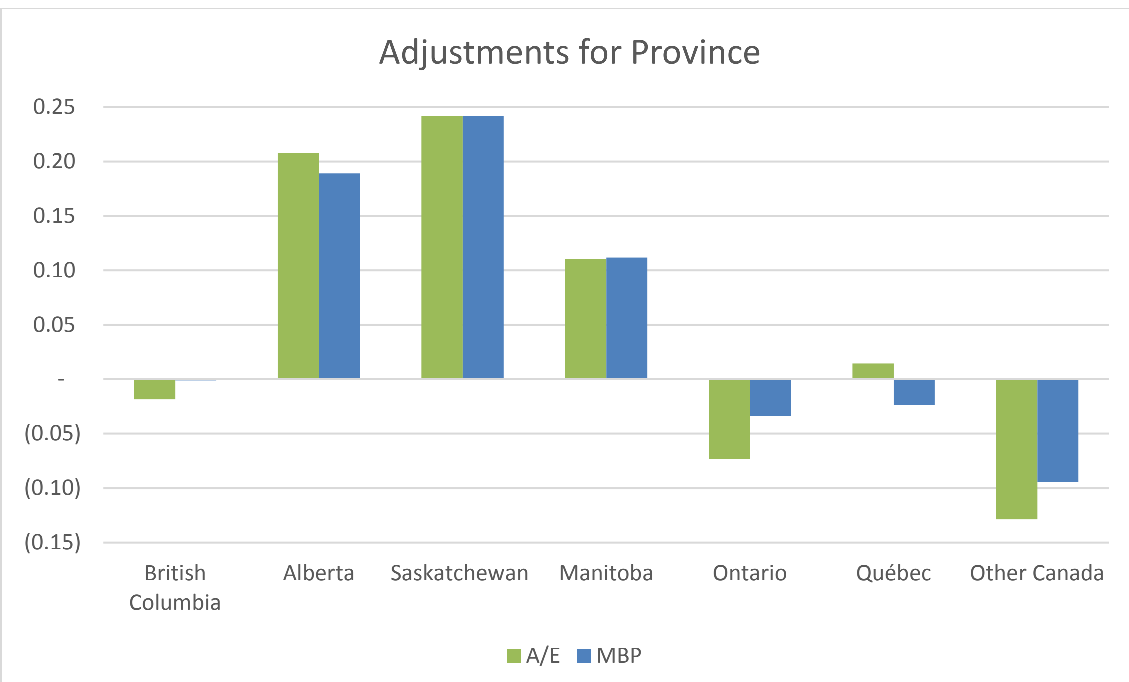

Adjustments for Industry
0.15
0.10
0.05
-
-0.05
-0.10
-0.15
Heavy Blue Collar
Manufacturing
Wholesale, Retail Trade
White Collar and Professional
Health, Education, Social Services
Other Services (Private Sector)
Public Administration
Unknown
A/E
MBP


Adjustments for Elimination Period
0.15
0.10
0.05
-
-0.05
-0.10
-0.15
0 to 3 months
4 months
5 to 6 months
Greater than 6 months
A/E
MBP


Adjustments for Pre-LTD Benefits
0.25
0.20
0.15
0.10
0.05
-
-0.05
-0.10
-0.15
Our STD
Other or None
A/E
MBP


Adjustments for Benefit Amount
0.15
0.10
0.05
-
(0.05)
(0.10)
(0.15)
Unknown
Less than $1,500
$1,500 to $1,999
$2,000 to $2,499
$2,500 to $3,249
Greater than $3,250
A/E
MBP





Adjustment Factors – Version 2

One concern about the MBP, as noted earlier, is that it implicitly assumes that the indicated adjustment is constant for all durations. It is unlikely that one multiplicative factor would be appropriate for a claim at all durations given that termination behaviour is driven by different forces along the duration curve. For instance, early-duration terminations are mostly recoveries where the claimant returns to work. After 36 months, termination rates tend to be dominated by mortality.

To account for this, we extended the Version 1 model by creating two additional sets of factors using the experience for:

- Claim durations of one month to 36 months; and
- Claim durations beyond 36 months.

This produces more precise results by generating different factors to modify termination rates at different durations.

In theory, this process could be extended to provide even greater resolution by dividing the data into additional shorter intervals. The recommendation for two intervals is based on considerations of data credibility and ease of use for the ultimate user.

The Version 2 factors for all variables are set out in Table 8 below. Appendix 5 summarizes both versions of the model and provides complete exposure data.

Table 8 MBP Predictive Model Factors – Version 2 (by Duration)

	MBP FACTORS for Months	
INDUSTRY	**1 to 36**	**36+**
Heavy Blue Collar	1.033	1.105
Manufacturing	0.997	0.941
Wholesale, Retail Trade	1.022	0.994
White Collar and Professional	1.025	0.950
Health, Education, Social Services	1.024	1.018
Other Services (Private Sector)	0.989	0.968
Public Administration	0.906	1.083
Unknown	1.025	0.928

	MBP FACTORS for Months	
ELIMINATION PERIOD	**1 to 36**	**36+**
0 to 3 months	0.945	0.961
4 months	1.021	0.984
5 to 6 months	1.011	1.008
Greater than 6 months	0.954	1.099

	MBP FACTORS for Months	
PRE-LTD BENEFITS	**1 to 36**	**36+**
Our STD	1.193	0.901
Other or None	0.933	1.019

BENEFIT AMOUNT	MBP FACTORS for Months	
	1 to 36	**36+**
Unknown	1.080	1.080
Less than $1,499	1.003	1.087
$1,500 to $19990	0.974	0.991
$2,000 to $2,499	1.002	1.012
$2,500 to $3,249	1.017	1.009
Greater than $3,250	0.976	0.893

DIAGNOSIS	MBP FACTORS for Months	
	1 to 36	**36+**
Mental Disorders	1.036	0.872
Musculo-skeletal	0.906	0.822
Neoplasms (Cancers)	1.181	2.656
Circulatory	0.854	0.877
Nervous System	0.506	0.661
Accidents	1.227	1.036
All Other Identified Causes	1.038	1.170
Not Stated or Unknown	1.086	0.811

PROVINCE	MBP FACTORS for Months	
	1 to 36	**36+**
British Columbia	1.002	0.980
Alberta	1.192	1.145
Saskatchewan	1.245	1.212
Manitoba	1.107	1.170
Ontario	0.963	1.009
Québec	0.976	0.966
Other Canada	0.913	0.842

IMPLEMENTING THE MODEL

Both Version 1 and Version 2 factors are recommended for use in pricing and valuation efforts. Version 2 provides more precision, but this may not be required for all applications. Possible uses might include:

- Valuing open claim reserves for refund-rated cases;
- Doing segmented profitability analysis when segments are defined along any of the dimensions used in the model (e.g., Province or Industry); and
- Valuation of statutory claim reserves, especially for carriers with a small block of business or where the carrier believes its portfolio is skewed relative to the overall market.

The recommended procedure for implementing the model developed here is as follows:

1. For a given claim, determine a termination rate from the Designated Tables based on age at disability, gender, region, and duration of claim.
2. Multiply the rate determined above by each of the factors for the six variables in the model. The final product is the adjusted termination rate.

See Appendix 6 for examples of this process.

If the user does not wish or is unable to use one of the variables (due to lack of available data, perhaps), the variable may be omitted (by setting all values for that variable to 1.00). This, naturally, reduces the predictive power of the model.

EVALUATION

In this section, we discuss the practical implications of the findings presented above.

The model provides adjustments for six variables in addition to Age, Gender, and Duration contained in the Designated Tables. The analysis above clearly indicates that there is potential for improved accuracy if Diagnosis and Province adjustments are incorporated into valuation calculations.

The Industry variable is problematic. Except for Public Administration, most categories are close to 1.00 over all durations. For practicality, it may be useful to collapse this variable into Public Administration and All Other.

For the Benefit Amount and Elimination Period variables, the improvement in accuracy appears minimal for individual claims and probably non-existent over a portfolio of claims. It is unlikely that adding these variables would warrant the additional complexity in processing and data validation burdens. There is, of course, value in knowing that these variables are relatively unimportant.

The Pre-LTD Benefits variable suggests that claims reserves could be 25% lower for those claimants who also received short-term disability benefits from the same carrier. The suggested explanation is that claims in this category benefit from the carrier having earlier access to data and an earlier opportunity for return to work management.

Given that this hypothesis is dependent on the internal procedures within each carrier, it would be prudent to replicate this finding on one's own block of business before adopting this adjustment.

CAVEATS

Users of this study should take note of the following comments.

1. The Designated Tables and the adjustment factors recommended in this paper are based on lives, not on benefit amount.
2. The work presented here is novel in the context of Canadian LTD termination experience and there is no body of published research supporting the use of the MBP in the development of group LTD expected termination rates.
3. The factors recommended here are designed to be used with the Designated Tables and have been validated on total industry experience. They may not be fully applicable if the Designated Tables are modified to reflect a particular case or portfolio.

4. The factors recommended here are designed to be used with the total termination rates in the Designated Tables and have not been validated for use with the mortality-only rates in the Designated Tables.
5. The recommended model employs the product of multiple adjustment factors. Although the MBP used to create the recommended model operates to minimize the bias associated with correlation among predictive variables, it does not necessarily remove all bias.
6. The model factors (e.g., for Province or Industry) should not be used directly as adjustment factors in a pricing model without due consideration for possible correlation between incidence rates and termination rates.

CONCLUSIONS AND RECOMMENDATIONS

1. Valuation actuaries should consider using the models proposed in combination with the recently released Designated Tables.
2. While a full implementation of the Version 2 model offers the maximum potential benefits, it would be appropriate to implement a more limited scheme initially. The Version 1 model may be used to simplify administrative requirements. In addition, any of the variables in either Version may be excluded.
3. When the CIA considers an update to the recently published 2008–2015 termination tables, we recommend that predictive analytical techniques be considered as an alternative or supplement to traditional methodology. Regardless of the methodology used, we recommend that the resulting table/model should incorporate, at a minimum:
    a. Age;
    b. Gender;
    c. Province (in lieu of Québec/Rest of Canada);
    d. Diagnosis; and
    e. Industry.

APPENDIX 1 – LITERATURE REVIEW

The following provides a brief summary of the review of relevant literature to date.

Industry Termination Rates Tables

The termination rate for a particular claim duration since disablement is derived as a function of factors (e.g., age and gender) corresponding to the profile of each claim. The actuarial exposure method based on a large sample is used to estimate the termination rate for each duration for a segment of claim profile. For some tables, the mortality rates and recovery rates are separated, and a graduation process is employed.

- U.S.: Commissioner's Disability Tables (64CDT); Commissioner's Individual Disability Table A (85CIDA); Society of Actuaries (SOA) 1987 Basic Group LTD Table (87GLTD); SOA Group Term Life Waiver Table (2005GLTD, 2008GLTD); 2004–2012 GLTD Database;
- Canada: CIA 1988–1994 GLTD study, CIA 1988–1997 GLTD Termination Experience;
- UK: Continuous Mortality Investigation Bureau (CMIR12 Table); and
- AU: Institute of Actuaries of Australia (IAD1989-93 Table).

Statistical Analysis

A large number of studies have been conducted based on the survival or duration analysis methods to estimate probability of termination (death or recovery). The methods range from non-parametric (Kaplan-Meier) and semi-parametric (Cox proportional hazards regression) to parametric (accelerated life time regression) models. A few references to note include:

D. J. Doudna (1977). Effect of the Economy on Group Long Term Disability Claims. *Journal of Risk and Insurance*, 44(2), 223–235.

S. M. Mulla, S. Makosso-Kallyth, N. St-Hilaire, K. Munsch, P. B. Gove, D. Heels-Ansdell, G. H. Guyatt, and J. W. Busse (2017). Factors associated with the duration of disability benefits claims among Canadian workers: a retrospective cohort study. *CMAJ Open*, 5(1): 109–115.

D. Pitt (2007). Modelling the Claim Duration of Income Protection Insurance Policyholders using Parametric Mixture Models. *Annals of Actuarial Science*, 2(1), 1–24.

Predictive Modelling

Various statistical learning and data mining methods have been proposed to better utilize the information in insurers' information system, e.g., Predictive Modeling: A Modeler's Introspection (SOA, June 2015) and Predictive Analytics in Life and Health insurance (SOA blog, February 2018). However, there are only a few studies available specifically for the prediction of disability termination rates:

Mervyn Kopinsky (2015). Predicting Group Long Term Disability Recovery and Mortality Rates using Tree Models. Society of Actuaries. Available at https://www.soa.org/experience-studies/2017/2017-gltd-recovery-mortality-tree/.

- This study applies the regression tree method on the SOA 2004–2012 Group Long Term Disability (GLTD) Database.

Q. Liu, D. Pitt, and X. Wu (2014). On the prediction of claim duration for income protection insurance policyholders. *Annals of Actuarial Science*, 8(1): 42–62.

- The prediction accuracy of several statistical learning methods such as linear regression, linear and quadratic discriminant analyses, LRs, and K-nearest-neighbour methods are compared.

APPENDIX 2 – USING LOGISTIC REGRESSION TO PRODUCE DESIGNATED TABLES

Logistic regression (LR) models predict a binary dependent variable. The independent variables may either be continuous or categorical.

Although it was outside the scope of this project, we explored using LR models to create base termination rate tables.

The dependent variable was defined as whether a claim terminated (0 = claim did not terminate, 1 = claim did terminate). The independent variables were:

- Age at Disability;
- Gender;
- Claim Duration; and
- Non-Québec/Québec.

Claim Duration may be included as either continuous or categorical. We treated Claim Duration as a categorical variable for two reasons:

1. To be consistent with the Tables Study, which presents annual rather than monthly termination rates after duration 60; and
2. To avoid "over-smoothing" as it may not be appropriate for termination rates to be always decreasing.

The analysis leads to the following conclusions regarding LR:

- An advantage of this approach is that the model inherently smooths the raw data and eliminates the need for other approaches (i.e., moving averages) that are tedious, require actuarial judgment, and are not statistical in nature.
- The results are statistically sound and provide p-values and other measures of goodness of fit.
- Final tables may be constructed using a single equation with variables that vary based on a claimant's characteristics.
- A disadvantage of this approach is that it is difficult to account for the effect of the CiD of disability for a claim. We believe it is possible to do so by isolating portions of the data, but we did not explore this as it is out of the scope of this study.

LR is a viable predictive modelling method that may be explored further in future studies.

APPENDIX 3 – INDUSTRY CODING

A variety of coding systems were used by the contributing companies.

The Tables Study used the following coding scheme, which is based on the first two digits of the Canadian version of the 2008 North American Industry Classification System (NAICS) with the adjustments noted.

Codes 96, 97, 98, and 99 are not part of NAICS.

- *96 Invalid Code* means that the submitted code was not valid within the carrier's own coding system (probably due to data entry error).
- *97 No Data* means that the data submitted did not provide any information on industry.
- *98 Unmappable* means that the data submission did provide industry information, but we were unable to map these data into our coding scheme.
- *99 Unknown* means that the data submission did provide industry information telling us that the industry was "unknown".

For each carrier, we created concordance tables that mapped submitted codes to the study scheme.

Table 9 Industry Codes

Code	Description	NAICS Adjustments
11	Agriculture, Forestry, Fishing, and Hunting	
21	Mining and Oil and Gas Extraction	
22	Utilities	
23	Construction	
31	Manufacturing	Plus NAICS 32, 33
41	Wholesale Trade	
44	Retail Trade	Plus NAICS 45
48	Transportation and Warehousing	Plus NAICS 49
51	Information and Cultural Industries	
52	Finance and Insurance	
53	Real Estate and Rental and Leasing	
54	Professional, Scientific, and Technical Services	
55	Management of Companies and Enterprises	
56	Waste Management	NAICS 562 ONLY, Excludes 561
61	Educational Services	
62	Health Care	NAICS 621, 622, 623
63	Social Services	NAICS 624
71	Arts, Entertainment, and Recreation	
72	Accommodation and Food Services	
81	Other Services (except Public Administration)	Plus NAICS 561
91	Public Administration	
96	Invalid Code	
97	No Data	
98	Unmappable	
99	Unknown	

For purposes of the model constructed in this project, the industry categories were collapsed as indicated below.

Table 10 Mapping of Industry Codes

Industry Category	Codes Included
Heavy Blue Collar	11, 21, 22, 23, 48, 56
Manufacturing	31
Wholesale, Retail Trade	41, 44
White Collar and Professional	51, 52, 53, 54, 55
Health, Education, Social Services	61, 62, 63
Other Services (Private Sector)	71, 72, 81
Public Administration	91
Unknown	96, 97, 98, 99

APPENDIX 4 – DIAGNOSIS CODING

Most companies have coding systems based on the International Classification of Diseases, 9th Revision (ICD-9). Three companies use ICD-10. A few have proprietary schemes. One company provided only free-form text descriptions and we manually coded each record into our coding scheme.

For the purposes of this study, we created the following coding scheme, which is also based on ICD-9.

Table 11 Diagnosis Codes

Code	Description
A	Infectious and Parasitic Diseases
B	Neoplasms
C	Endocrine, Nutritional, and Metabolic Diseases, and Immunity Disorders
D	Diseases of the Blood and Blood-Forming Organs
E	Mental Disorders
F	Diseases of the Nervous System and Sense Organs
G	Diseases of the Circulatory System
H	Diseases of the Respiratory System
I	Diseases of the Digestive System
J	Diseases of the Genitourinary System
K	Complications of Pregnancy, Childbirth, and the Puerperium
L	Diseases of the Skin and Subcutaneous Tissue
M	Diseases of the Musculo-skeletal System and Connective Tissue
N	Congenital Anomalies
O	Certain Conditions Originating in the Perinatal Period
P	Symptoms, Signs, and Ill-Defined Conditions
Q	Injury and Poisoning
U	Unknown
X	No Data
Y	Unmappable

As is typical of claim records, there is often some ambiguity as to the exact cause of disability or the appropriate code to be used when the claimant suffers from several conditions. Where necessary, we exercised our judgment. Normally, we did not request additional data from the actual claim file.

Table 12 Mapping of Diagnosis Codes

Diagnosis Category	Codes Included
Mental Disorders	E
Musculo-skeletal	M
Neoplasms (mostly cancers)	B
Circulatory	G
Nervous System	F
Accidents	Q
All Other Identified Causes	A, B, C, D, H, I, J, K, L, N, O, P
Not Stated or Unknown	U, X, Y, Z

APPENDIX 5 – SUMMARY OF MODEL FACTORS

Version 1

All Durations

Industry	**Exposure**	**Factors**
Heavy Blue Collar	1,332,695	1.038
Manufacturing	1,522,271	0.994
Wholesale, Retail Trade	1,265,085	1.019
White Collar and Professional	1,616,981	1.018
Health, Education, Social Services	1,312,515	1.025
Other Services (Private Sector)	875,876	0.990
Public Administration	1,802,829	0.920
Unknown	1,088,653	1.017
Total	**10,816,905**	

Elimination Period	**Exposure**	**Factors**
0 to 3 months	2,101,997	0.948
4 months	4,772,036	1.018
5 to 6 months	2,746,656	1.012
Greater than 6 months	1,196,216	0.968
Total	**10,816,905**	

Version 2

Under/Over 36 Months' Duration

Industry	**Exposure**		**Factors**	
	1 to 36 Months	**36+ Months**	**1 to 36 Months**	**36+ Months**
Heavy Blue Collar	682,892	649,803	1.033	1.105
Manufacturing	712,457	809,814	0.997	0.941
Wholesale, Retail Trade	715,760	549,325	1.022	0.994
White Collar and Professional	800,860	816,121	1.025	0.950
Health, Education, Social Services	693,151	619,364	1.024	1.018
Other Services (Private Sector)	340,455	535,421	0.989	0.968
Public Administration	745,577	1,057,252	0.906	1.083
Unknown	377,210	711,443	1.025	0.928
Total	**5,068,362**	**5,748,543**		

Elimination Period	**Exposure**		**Factors**	
	1 to 36 Months	**36+ Months**	**1 to 36 Months**	**36+ Months**
0 to 3 months	873,909	1,228,088	0.945	0.961
4 months	2,443,044	2,328,992	1.021	0.984
5 to 6 months	1,228,463	1,518,193	1.011	1.008
Greater than 6 months	522,946	673,270	0.954	1.099
Total	**5,068,362**	**5,748,543**		

Pre-LTD Benefits	Exposure	Factors
Our STD	1,936,321	1.181
Other or None	8,880,584	0.939
Total	**10,816,905**	

Exposure		Factors	
1 to 36 Months	36+ Months	1 to 36 Months	36+ Months
1,157,505	778,816	1.193	0.901
3,910,857	4,969,727	0.933	1.019
5,068,362	**5,748,543**		

Monthly Benefit	Exposure	Factors
Unknown	460,056	1.082
Less than $1,499	2,022,371	1.009
$1,500 to $1,999	2,102,488	0.975
$2,000 to $2,499	1,877,541	1.003
$2,500 to $3,249	2,111,827	1.018
Greater than $3,250	2,242,622	0.969
Total	**10,816,905**	

Exposure		Factors	
1 to 36 Months	36+ Months	1 to 36 Months	36+ Months
280,099	179,957	1.080	1.080
752,597	1,269,774	1.003	1.087
917,803	1,184,685	0.974	0.991
926,853	950,688	1.002	1.012
991,775	1,120,052	1.017	1.009
1,199,235	1,043,387	0.976	0.893
5,068,362	**5,748,543**		

Claim Diagnosis	Exposure	Factors
Mental Disorders	3,072,438	1.026
Musculo-skeletal	2,371,654	0.900
Neoplasms (Cancers)	1,016,681	1.236
Circulatory	877,167	0.854
Nervous System	1,200,137	0.526
Accidents	772,068	1.219
All Other Identified Causes	1,336,964	1.049
Not Stated or Unknown	169,796	1.059
Total	**10,816,905**	

Exposure		Factors	
1 to 36 Months	36+ Months	1 to 36 Months	36+ Months
1,430,744	1,641,694	1.036	0.872
1,141,163	1,230,491	0.906	0.822
693,484	323,197	1.181	2.656
362,680	514,487	0.854	0.877
390,049	810,088	0.506	0.661
433,256	338,812	1.227	1.036
559,035	777,929	1.038	1.170
57,951	111,845	1.086	0.811
5,068,362	**5,748,543**		

Province	Exposure	Factors
British Columbia	1,286,554	0.999
Alberta	1,100,046	1.189
Saskatchewan	281,222	1.242
Manitoba	359,099	1.112
Ontario	4,299,717	0.966
Québec	2,352,523	0.976
Other Canada	1,137,744	0.906
Total	**10,816,905**	

Exposure		Factors	
1 to 36 Months	36+ Months	1 to 36 Months	36+ Months
550,187	736,367	1.002	0.980
558,443	541,603	1.192	1.145
142,158	139,064	1.245	1.212
157,845	201,254	1.107	1.170
1,819,043	2,480,674	0.963	1.009
1,361,635	990,888	0.976	0.966
479,051	658,693	0.913	0.842
5,068,362	**5,748,543**		

APPENDIX 6 – EXAMPLES

These examples are also available in an Excel workbook.

Example 1 – Find Termination Rate for Duration (Month) 18

Claimant Information

Basic

Residence	Alberta
Gender	Male
Age at Disability	37

Additional

Industry	Education
Elimination Period	4 months
Pre-LTD Benefits	None
Benefit Amount	$2,200
Diagnosis	Musculo-skeletal
Province	Alberta

From the Designated Tables, find the appropriate termination rate.

	Rest of Canada tables, male age 35–39
Unadjusted termination rate	0.04147

From Appendix 5 of this report, find the relevant adjustment factors.

	Model 2, Under 36 months
Industry	1.024
Elimination Period	1.021
Pre-LTD Benefits	0.933
Benefit Amount	1.002
Diagnosis	0.906
Province	1.192

Multiply all model factors together.

Composite Factor	1.056

Multiply the base rate by the product of all factors from the model.

Adjusted termination rate	0.043774

Example 2 – Find Termination Rate for Duration (Month) 45

Claimant Information

Basic

Residence	Québec
Gender	Female
Age at Disability	52

Additional

Industry	Public Administration
Elimination Period	12 months
	Own
Pre-LTD Benefits	STD
Benefit Amount	$5,000
Diagnosis	Nervous System (Parkinson's Disease)
Province	Québec

From the Designated Tables, find the appropriate termination rate.

	Rest of Canada tables, male age 35–39
Unadjusted termination rate	0.00834

From Appendix 5 of this report, find the relevant adjustment factors.

	Model 2, Over 36 months
Industry	1.083
Elimination Period	1.099
Pre-LTD Benefits	0.901
Benefit Amount	0.893
Diagnosis	0.661
Province	0.966

Multiply all model factors together.

Composite Factor	0.611

Multiply the base rate by the product of all factors from the model.

Adjusted termination rate	0.0051