Calibrating CV22 Canadian VitaCurves

In this paper we provide an overview of the techniques used to build the 'CV22' edition of our Canadian VitaCurves, which were released in July 2023. These techniques mirror those well-established for analyzing UK and US pension plan longevity, which have been peer reviewed by actuarial bodies[1,2].

The paper is aimed at a technical actuarial audience interested in understanding more about the methods we use and should be read in conjunction with our paper summarizing the data used, 'Data underpinning CV22 Canadian VitaCurves'. Further detailed technical information can be provided on request.

If any of the technical terms in this paper are unfamiliar to you, you can find definitions and related terms in our online Lexicon of Longevity. Just visit www.clubvita.net/glossary and search for a term.

1 An overview of our approach

At the heart of our approach is a statistical method known as 'Generalized Linear Modeling' (GLM), which links mortality rates to the values taken by a range of factors known to be relevant to the chance of dying (for example an individual's age, income and postal code information).

In terms of conventional actuarial notation, we can think of the chance of someone dying (typically denoted q_x) as being determined not just by their age x but a whole host of potentially relevant factors:

$$logit(q_x) = a\ linear\ function\ of\ predictors\ (e.g., age, income, postal\ code)$$

Where $logit(q_x)$ is the following logistic transformation:

$$logit(q_x) = \ln\left(\frac{q_x}{1 - q_x}\right)$$

The logistic transformation is an evolution of the basic log transformation proposed by Gompertz. In the world of Generalized Linear Modeling, it is also the 'natural' scale for analyzing events like the number of people dying at each age, which are expected to follow a Binomial distribution.

Our approach then fits a series of these models to the data, with each model optimized to provide a good fit to the underlying data. In order to fit these models, we need to identify which of the potential longevity predictors in our data is relevant, and how best to include these – we summarize this in **Section 2**, with **Section 3** covering how we create our postal code based longevity groups and pension bands for use in the models. **Section 4** describes how we fit the model to the data, with **Section 5** describing our retirement health adjustment factors to create curves for disabled and non-disabled pensioners. **Section 6** provides an overview of the goodness of fit tests applied to the models and **Section 7** covers checks done regarding the internal consistency of the models. We conclude the description of the creation of our models in **Section 8**, by explaining how we extend the fitted models to provide VitaCurves which cover the full age range (i.e., to include younger and older ages outside the range that the data is fitted to). In **Section 9** we discuss the impact of COVID-19 on this edition of Canadian VitaCurves, while in **Section 10** we draw attention to some aspects of the models that users of VitaCurves should be aware of when using the model.

---

[1] For the UK see: [What longevity predictors should be allowed for when valuing pension scheme liabilities]()

[2] For Canada see: [Key Factors for Explaining Differences in Pensioner Baseline Mortality]()

2 Allowing for different longevity predictors

Our Canadian VitaCurves generally use the GLM construct[3], with each reflecting a specific combination of data fields identified as being 'predictive' of longevity (known as 'rating factors'). Potential predictors available for each plan participant include:

- Age
- Gender
- Pensioner type (first life pensioner or in receipt of a survivor pension)
- Retirement health (ill health, non-ill health or unknown, with 'all health' meaning retirement health is not considered)
- Pension form (single or joint life)
- Amount of pension
- Last known salary amount
- Lifestyle proxies: Geo-demographic indicators derived from individual postal codes
- Collar type (manual or non-manual)

Traditionally, actuaries have fitted ('graduated') life tables by first segmenting data into groups with like values of a predictor and then smoothing based on this data. This helps to reduce heterogeneity (variability) in the underlying data, but at the expense of reducing the data volumes in each group, in turn reducing the certainty in the mortality rates and/or the potential number of groups that can be formed. The application of GLM techniques enables a wide range of internally consistent tables to be fitted simultaneously across a range of variables. This makes maximum use of the available data, improving confidence in the resulting tables while creating a model that captures the diversity of the underlying population.

Wherever practical, it is preferable to use a predictor directly in the GLM (known technically as treating the variable as a '**covariate**'). Different values of the rating factor are then modeled simultaneously within the GLM, maximizing the amount of data used to infer the impact of any particular rating factor. However, for some predictors it remains necessary to segment the data into groups before fitting the GLM to each group. Treating a longevity predictor in this way (as a '**stratifier**') is generally preferred when:

- The *meaning* of other variables differs across the predictor, greatly complicating the fitting of a GLM; or
- The *shape* of mortality with age is very different depending on the value of that variable; or
- The *data* spans different age ranges for the different values of that variable

[3] Exceptions to this are disabled pensioners (ill health) and non-disabled pensioners (non-ill health). Not all plans have reliable retirement health data, and those that do may only have a small proportion of disabled pensioners. Therefore, we develop our ill health and non-ill health adjustment factors (that vary by age and gender) relative to the corresponding 'all health' mortality for those plans where retirement health is reliable.

2.1 Splitting the data into sensible groups for modeling (stratification)

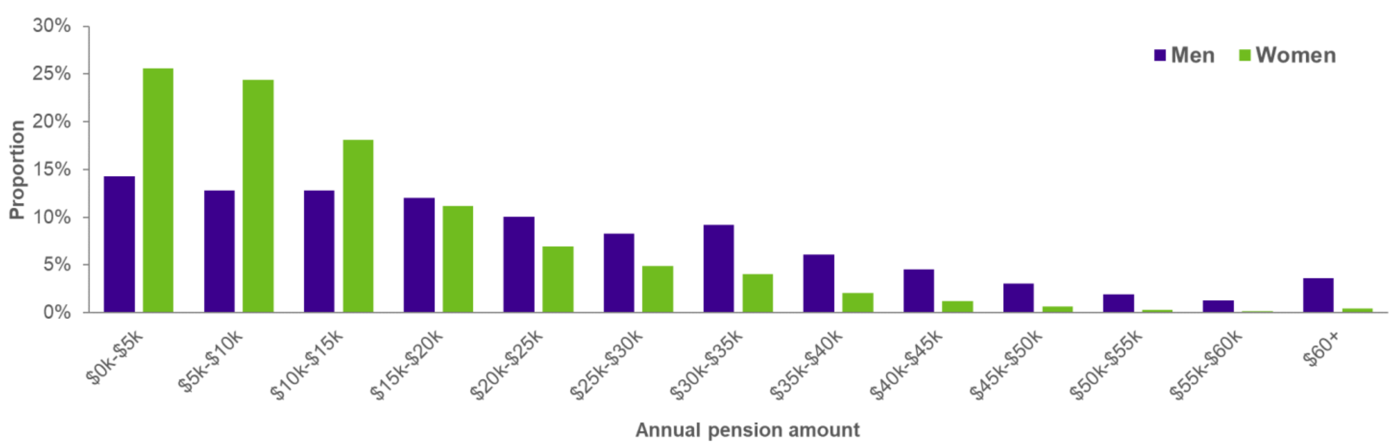
We split the data into the following eight groups (or **strata**) determined by participants' gender, retirement health and pensioner type, for the reasons set out below.

- Male 'all health' pensioners
- Female 'all health' pensioners
- Female survivors
- Male survivors
- Male ill health (or disabled) pensioners
- Female ill health (or disabled) pensioners
- Male non-ill health (or non-disabled) pensioners
- Female non-ill health (or non-disabled) pensioners

2.1.1 Why split men and women?

In Canada, like in the US and UK, the relative importance of different covariates can vary between the genders. For example, we find that the pension income of a woman is often less informative to their longevity than it is for a man. In many territories this is a legacy of both historical working patterns, and the history of pension plan participation of men and women, which has led to women typically having accrued lower pensions than men among older generations (as can be seen in Figure 1 which shows the distribution of pension amounts for male and female pensioners over age 75). Consequently, we prefer to split the data between men and women to ensure we correctly capture the importance of different rating factors for men and women.

Figure 1: Distribution of annual pension amount for male and female pensioners over age 75



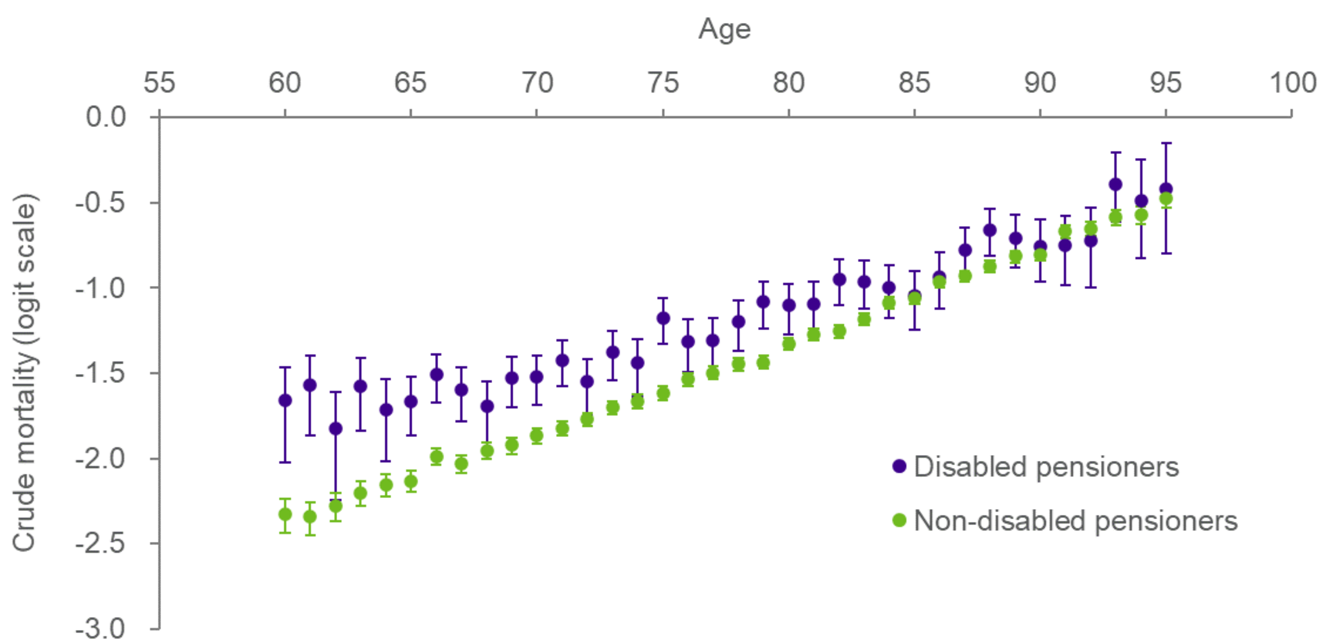
Note: Chart includes all pensioners over age 75 as of January 1, 2021 who were exposed to risk during the period 2018-2020 and so potentially were included in our modeling.

2.1.2 Disabled versus non-disabled pensioners

Disabled pensioners tend to retire younger than non-disabled pensioners and tend to have curtailed life expectancy. We split out the data for disabled pensioners since the observed proportions dying at each age ('crude mortality rates') show a different shape at younger ages, which we need to take care to ensure our models capture (see Figure 2).

As a relatively small proportion of the data relates to disabled pensioners, we do not attempt to identify the impact of other factors on the longevity of disabled pensioners.

Figure 2: Crude mortality rates (on logit scale) with 95% confidence interval for disabled and non-disabled male pensioners



2.1.3 Surviving beneficiaries of plan participants

We also split out the data relating to the surviving beneficiaries of deceased pensioners. There are fewer individuals in this group compared to pensioners and less information is available about them, limiting our ability to identify the impact of other factors on the longevity of survivors. They also tend to be older and the meaning of some of the rating factors differs (e.g., the pension amount is typically a proportion of that for the original plan participant). For this group, our VitaCurves use pension amount and postal code only[4].

The resulting rates for this group show levels of elevated mortality consistent with that seen in other studies[5], which we believe are broadly indicative of the mortality rates for these individuals.

[4] We do not allow for collar type as this relates to the original plan participant's employment and so is less relevant to the surviving beneficiary.

[5] See our publication Grieving Widows and Society of Actuaries' Exposure Draft: Pri-2012 Private Retirement Plans Mortality Tables

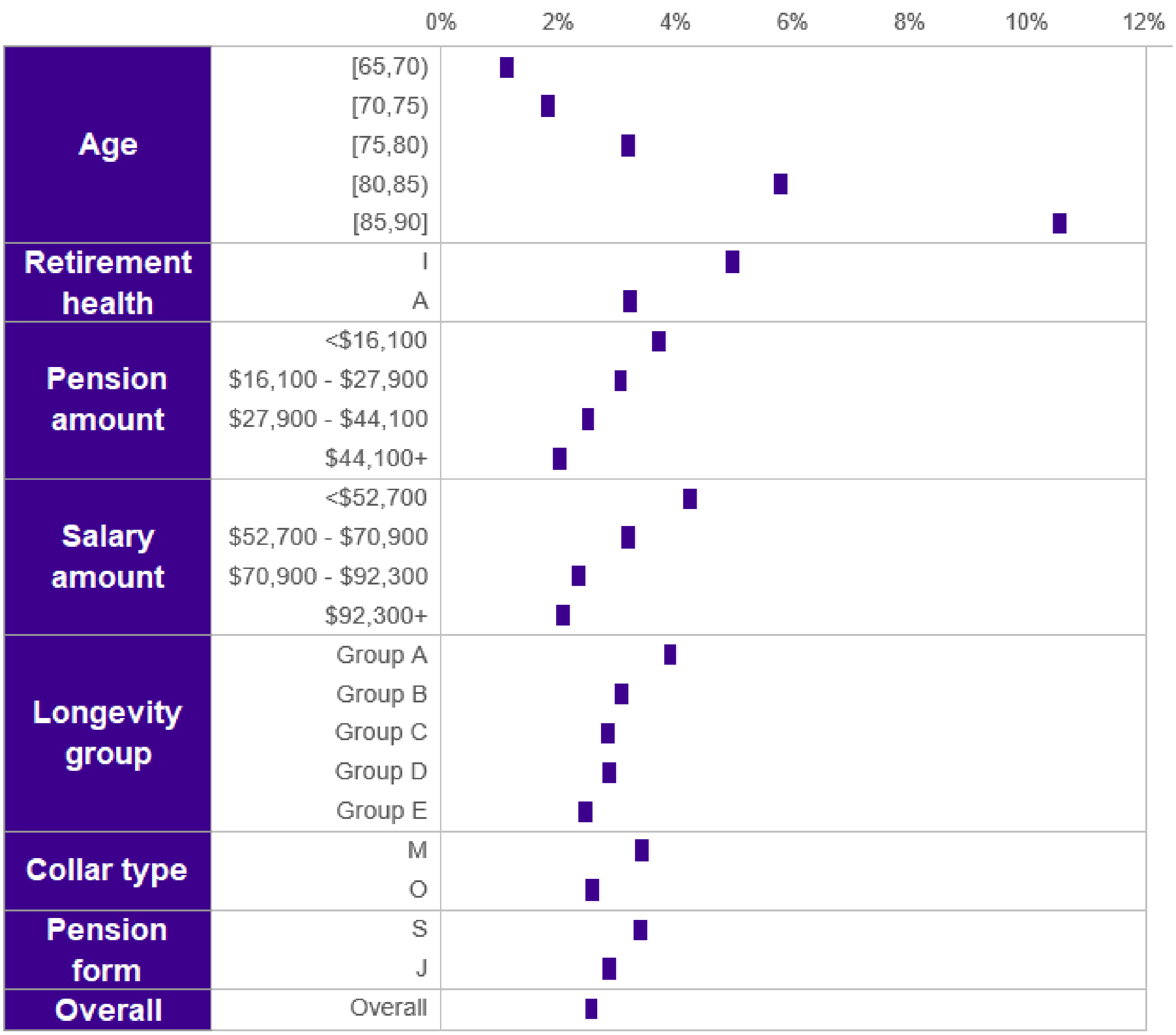
2.2 Identifying key variables which influence mortality for each group (covariates)

We've found that the following variables tend to be highly predictive of longevity:

- **Income** via salary or pension amount
- **Lifestyle** via postal code grouping (referred to as longevity groups)
- **Collar type** of former employment
- **Pension form** at retirement

For Canadian pension plan participants, all four of these variables are clearly associated with variations in mortality, and so life expectancy. This is shown in Figure 3, where the green bars represent the relative magnitude of mortality rates for different male pensioner factor groupings (the further to the right the larger the mortality rate).

Figure 3: Crude mortality rates by different rating factors for male pensioners

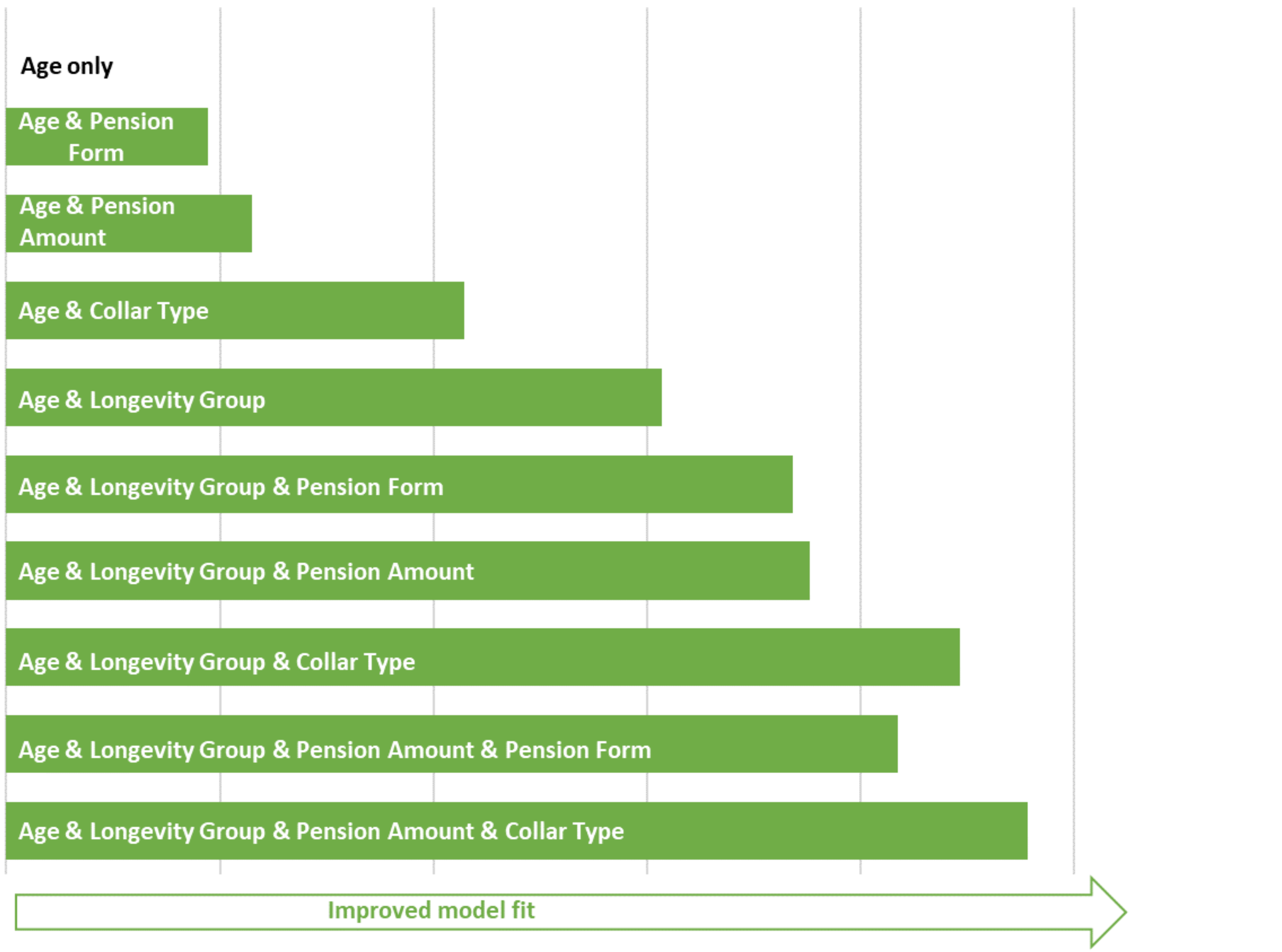


Note: Crude mortality rates here are based on male pensioners aged 65-90 and after application of quality control criteria for all variables. Mortality rates have not been age standardized. Abbreviations used: I = Ill retirement health, A = All retirement health, M = Manual occupation, O = non-manual occupation, S = Single life form of pension, J = Joint life form of pension

These predictors are not independent – often pension amounts will be higher for white collar employees, and income is closely correlated with other socio-economic factors captured by postal code. As such, we need to verify whether all these predictors are relevant to include in our modeling. We do this by sequentially adding potential rating factors to the model with the aim of improving fit. This 'stepwise method' highlights the benefit of adding specific factors to the model, and combinations thereof, but also the point where we gain limited benefit from adding further predictors.

Figure 4 illustrates the results of this for all health male pensioners where we considered models with longevity group, pension band, collar type and pension form[6]. The longer the bars extend, the better the model 'fits' to the data[7]. For the discussion below, we consider 'Age only' to be the first 'bar' in the chart.

Figure 4: Improvement in fit of model by adding different potential predictors



[6] Note that we only considered models with pension band as our affluent variable and excluded salary amounts. This is primarily done to increase the volume of the data used to fit these models, since pension amount is widely available and including salary amount can reduce the size of data.

[7] Technical Note: The bars represent the improvement in 'Akaike Information Criteria' upon adding additional variables.

We can see how:

- Allowing for any one of income, longevity group, pension form, and collar type improves the model fit, with longevity group being the single best factor *(2nd to 5th bars*).
- Using two or more factors materially improves the fit of the model *(6th to 10th bars*).
- Including longevity group in the model, the best additional variable is collar type *(8th bar*).
- A model which includes longevity group, pension amount and collar type is preferable to a model which only allows for two of these variables *(last bar compared to 6th to 8th bars*).

3 Forming ‘buckets’ of like individuals

The information we have collected contains a large number of potential covariate values and some grouping is needed before we carry out the fitting. For example:

- **Postal code information:** There are around 860,000 Canadian postal codes. Even with the use of a third-party provider to convert these postal codes into fewer than 100 different types of neighborhoods, we still have too many points to credibly model each type of neighborhood as a separate value.
- **Pension income:** Pensions can take a range of values, from a few dollars to many tens of thousands. In principle, this could be captured using a continuous variable in the GLM, but this would lead to every individual pensioner and survivor in a plan having a unique VitaCurve. This would be unwieldy for most plan valuation systems. We have elected to use an alternative approach of initially grouping pension amounts into percentage splits of the data within each of the subgroups (gender and pensioner type).

The parsimony principle

A key criterion in the development of a statistical model is parsimony;

> ***“a simpler model with few rather than many parameters is favored over comparatively complex ones, provided they fit the data about equally well.”***

The number of parameters in a model can be reduced by either:

- reducing a variable to a simpler form; or
- if possible, omitting the unnecessary variables in a model.

For example, imagine the inclusion of a 10-level discrete covariate in a regression model describing the geographical location of a pensioner based upon which province they live in. If mortality experience can be described equally as well with the inclusion of a simpler geographical variable (e.g. Western, Prairies, Central and Atlantic) then this model has materially fewer parameters (6 fewer) and is preferred over the more complex model.

For the purposes of assessing parsimony we need a measure of goodness of fit which penalizes for complexity. We generally use the Akaike Information Criterion which captures the goodness of fit via the (log-) likelihood (i.e., a measure of how likely the observed data is derived from the fitted model) with a penalty based upon the number of variables used (i.e., to avoid ‘over-fitting’ to the data).

Having reduced the list of postal codes to fewer than 100 groups (geo-demographic types), and each of salary and pension to about 30 groups, we now have a more tractable problem – namely how to further group these values to ensure we have a model which adheres to the parsimony principle (i.e., sufficient groups to capture the major variations in longevity, without having an excessively complex model). We achieve this via a two-stage process:

- Identify a potential 'optimal' grouping; and
- Verify the grouping provides a practical solution to users (and if not, restart the process).

3.1 Finding an 'optimal' grouping

Naively, we might consider an approach to determining the clustering which looks at all potential ways of grouping each of the covariates and fits the GLM for each. The combination of groupings for postal code and pension amount which results in the best fit would then be considered the most parsimonious and thus, optimal. The sheer number of possible groupings for postal codes alone makes this impractical. For example, there are over $3\text{x}10^{45}$ possibilities of reducing our neighborhood types into 5 or fewer groups.

Fortunately, statistical techniques have been developed to help overcome this challenge. These are designed to identify very good candidate clusterings, while accepting any method is unlikely to be truly optimal. The approach we have used is known as recursive partitioning[8]. Under this approach, the variable is first split into two groups, then one of these groups is split into two, and so on until there is evidence of limited benefit of further splitting in terms of goodness of fit vs complexity of the model. At each stage, splits are identified based upon a minimum required improvement in how much of the variation within the data is explained at each increase in the number of groups. This assessment of when to cease splitting into further groups is based upon cross-validation, whereby we split the entire data into several equal subsets, fit the model to all the data excluding one of these subsets, and then test the performance of the model in predicting the outcomes for the subset of the data not used for fitting. This is done for each of the subsets in turn and an average prediction error calculated. Once the prediction error falls below a certain level (typically one standard error) we have entered the realm of 'good models' and a point is usually reached where the prediction error starts to plateau (i.e., falls very little for each additional group added), indicating that there is limited benefit from creating further groups[9].

3.2 Ensuring a practical solution

It is possible that the clustering method can identify some very small groups which appear to have very different longevity, particularly for the less populous postal codes. These groups are likely to be reflecting, at least in part, data volatility, rather than genuine differences in longevity. To ensure this is avoided for the postal code model, we also:

- Require that a minimum amount of the data (5%) is included in each group during the partitioning process.
- Identify any neighborhood types with very small data volumes and/or anomalously high/low life expectancies based upon the raw data.

  These types are likely to be exhibiting artifacts of noise in the data and so we pre-group these with a neighborhood type of similar socio-economic characteristics. This is done for a very small number of neighborhood types.

[8] In our work on similar datasets in the UK, we have considered a number of possible approaches and noted they give rise to similar results.

[9] For those wishing to read more on the methods described here please see *Modern Applied Statistics with S* by Venables & Ripley (published by Springer, 2002)

It is also possible for the clustering method to identify groups which contain large volumes of data yet have very modest (but statistically significant) differences in mortality rates. These differences may not be material when converted to life expectancies and annuity values. In such instances, retaining the extra groupings is likely to be spurious, particularly once further covariates are included in the GLM providing additional differentiation between individuals. We therefore seek that the crude life expectancies at age 65 based upon the raw mortality rates (i.e., prior to the model fitting) and curtate to the maximum age used in the fitting:

- differ by at least 6 months between consecutive groups; and
- are statistically significantly different at the 95% confidence level between consecutive groups.

The first of these criteria provides confidence that we are capturing financially material differences, and the second that there will continue to be material differences in life expectancies between groups once we allow for other covariates in the GLM.

Where either of these criteria are not met, we have the choice to either revisit the partitioning to identify an alternative grouping, to combine groups, or, where the criteria are only just breached, allow it through. Judgement is applied as to which approach is adopted based upon the results from a series of diagnostics.

3.3 Resulting groups

In the case of pension amount, this approach is adopted separately for each of 3 groups: all health male and female pensioners, and female survivors. We have not incorporated pension bands for male survivors due to the small size of this group. A similar approach is taken when developing salary bands for male and female pensioners. The result is the following (annual) pension/salary bandings.

Figure 5: Pension and salary bands by gender and pensioner type

Pension Band	Male pensioners	Female pensioners	Male survivors	Female survivors
1	<$16,100	<$20,400		<$20,000
2	$16,100 - $27,900	$20,400 - $28,800		$20,000+
3	$27,900 - $44,100	$28,800+		
4	$44,100+			

Salary Band	Male pensioners	Female pensioners
1	<$52,700	$46,000
2	$52,700 - $70,900	$46,000 - $70,000
3	$70,900 - $92,300	$70,000 - $88,000
4	$92,300+	$88,000+

For postal code, the grouping is carried out separately for men and women (as socio-economic factors can influence men and women differently) using the combined all health datasets covering 3-years of experience. This results in 5 groups for men and women, labeled A (shortest lived socio-economic group) to E (longest lived socio-economic group).

4 Fitting the model

4.1 What ages and years?

In order to fit a model, we first need to decide which years we will fit to, and what age range the data can be used to reliably inform the model.

We want the results of the longevity model to be as relevant to current baseline longevity as possible. The more recent the period fitted to, the less the need to apply adjustments for the passage of time. The volatility associated with good and bad winters, flu seasons etc. means it is also preferable for any model to incorporate multiple years, rather than focusing on the most recent year alone. We have chosen to base our modeling for CV22 upon the 2018-2020 period. This means that the resulting mortality rates are applicable[10] to a pensioner or survivor of that exact age on January 1, 2019.

By including 2020 in the calibration period for CV22 we will be including for the first time the relatively high levels of mortality in that year as a result of the COVID-19 pandemic. However, while mortality rates in 2020 were slightly higher than expected based on previous trends, it is our view that the level of this 'excess' mortality in 2020 did not justify the additional complexity of publishing an 'adjusted' version of CV22 Canadian VitaCurves (unlike in the UK and US, which both saw much more material increases in mortality rates in 2020). Further details of this analysis are set out in **Section 9**.

The age ranges we fit to are determined by data volumes. The sparser the data, the lower the certainty we can have about the underlying mortality rates at each age. We therefore restrict the fitting of the Generalized Linear Model for each dataset (i.e., by pensioner type and gender) to the age ranges in the table below.

Figure 6: Fitting age range by different strata

Group	Age range fitted to
Female Pensioners, All retirement health (FPA)	60-95
Male Pensioners, All retirement health (MPA)	60-95
Female survivors (Widows), All retirement health (FWA)	60-95
Male survivors (Widowers), All retirement health (MWA)	60-95
Female Pensioners, Ill health retirement (FPI)	60-95
Male Pensioners, Ill health retirement (MPI)	60-95
Female Pensioners, Non-ill health retirement (FPN)	60-95
Male Pensioners, Non-ill health retirement (MPN)	60-95

[10] As part of the fitting we incorporate a time variable to further control for seasonality and any modest variations in exposed to risk between the individual calendar years

4.2 Including less granular curves

For some plan participants, not all data will be available – for example collar type may not be known, or salary amount not provided. Our general principle is that we want to use as much data as possible on an individual to provide as good an estimate of mortality for the individual as possible. Accordingly, we calibrate GLMs to different combinations of available information as illustrated by the table below. This means that there is always a curve available to use with the participant.

The following tables outline the functional form associated with each group of VitaCurves. Each component of a given functional form has a fitted coefficient associated with it. Rating factors are abbreviated to save space as follows: PB: pension band, SB: salary band, LG: longevity group, OC: occupation class and PF: pension form.

Figure 7: Male pensioner, all retirement health

Rating factors						Functional form
Age	PB	SB	LG	OC	PF	
✓	×	×	×	×	×	$Age^{-1} + Age^{-2} + period$
✓	×	×	✓	×	×	$Age^{-1} + Age^{-2} + Age^{-2}{:}LG + period$
✓	×	×	×	✓	×	$Age^{-1} + Age^{-2} + Age^{-2}{:}OC + period$
✓	✓	×	×	×	×	$Age^{-1} + Age^{-2} + Age^{-1}{:}PB + period$
✓	×	✓	×	×	×	$Age^{-1} + Age^{-2} + Age^{-2}{:}SB + period$
✓	×	×	×	×	✓	$Age^{-1} + Age^{-2} + Age^{-2}{:}PF + period$
✓	×	×	✓	✓	×	$Age^{-1} + Age^{-2} + Age^{-2}{:}LG + Age^{-2}{:}OC + period$
✓	✓	×	✓	×	×	$Age^{-1} + Age^{-2} + Age^{-2}{:}LG + Age^{-2}{:}PB + period$
✓	×	✓	✓	×	×	$Age^{-1} + Age^{-2} + Age^{-2}{:}LG + Age^{-2}{:}SB + period$
✓	✓	×	×	✓	×	$Age^{-1} + Age^{-2} + Age^{-2}{:}OC + Age^{-2}{:}PB + period$
✓	×	×	✓	×	✓	$Age^{-1} + Age^{-2} + Age^{-2}{:}LG + Age^{-2}{:}PF + period$
✓	✓	×	×	×	✓	$Age^{-1} + Age^{-2} + Age^{-2}{:}PB + Age^{-2}{:}PF + period$
✓	×	✓	×	✓	×	$Age^{-1} + Age^{-2} + OC + Age^{-2}{:}OC + Age^{-2}{:}SB + period$
✓	✓	×	✓	×	✓	$Age^{-1} + Age^{-2} + Age^{-2}{:}PB + Age^{-2}{:}LG + Age^{-2}{:}PF + period$
✓	✓	×	✓	✓	×	$Age^{-1} + Age^{-2} + Age^{-2}{:}OC + Age^{-2}{:}LG + Age^{-2}{:}PB + period$
✓	×	✓	✓	✓	×	$Age^{-1} + Age^{-2} + OC + SB + Age^{-1}{:}OC + Age^{-1}{:}SB + Age^{-2}{:}LG + period$
✓	×	✓	✓	×	✓	$Age^{-1} + Age^{-2} + LG + Age^{-1}{:}SB + Age^{-1}{:}PF + Age^{-1}{:}LG + period$

Figure 8: Female pensioner, all retirement health

Rating factors						Functional form
Age	PB	SB	LG	OC	PF	
✓	×	×	×	×	×	$Age^{-1} + Age^{-2} + Age^{-3} + period$
✓	×	×	✓	×	×	$Age^{-1} + Age^{-2} + Age^{-3} + LG + period$
✓	×	×	×	✓	×	$Age^{-1} + Age^{-2} + Age^{-3} + Age^{-3}: OC + period$
✓	✓	×	×	×	×	$Age^{-1} + Age^{-2} + Age^{-3} + PB + period$
✓	×	✓	×	×	×	$Age^{-1} + Age^{-2} + Age^{-3} + Age^{-2}: SB + period$
✓	×	×	×	×	✓	$Age^{-1} + Age^{-2} + Age^{-3} + Age^{-3}: PF + period$
✓	×	×	✓	✓	×	$Age^{-1} + Age^{-2} + Age^{-3} + Age^{-1}: LG + Age^{-1}: OC + period$
✓	✓	×	✓	×	×	$Age^{-1} + Age^{-2} + Age^{-3} + PB + LG + period$
✓	×	✓	✓	×	×	$Age^{-1} + Age^{-2} + Age^{-3} + LG + Age^{-1}: SB + period$
✓	✓	×	×	✓	×	$Age^{-1} + Age^{-2} + Age^{-3} + OC + Age^{-1}: PB + Age^{-2}: OC + period$
✓	×	×	✓	×	✓	$Age^{-1} + Age^{-2} + Age^{-3} + Age^{-1}: LG + Age^{-1}: PF + period$
✓	✓	×	×	×	✓	$Age^{-1} + Age^{-2} + Age^{-3} + Age^{-1}: PB + Age^{-1}: PF + period$
✓	×	✓	×	✓	×	$Age^{-1} + Age^{-2} + Age^{-3} + Age^{-2}: SB + Age^{-3}: OC + period$
✓	✓	×	✓	×	✓	$Age^{-1} + Age^{-2} + Age^{-3} + Age^{-1}: PB + Age^{-1}: LG + Age^{-1}: PF + period$
✓	✓	×	✓	✓	×	$Age^{-1} + Age^{-2} + Age^{-3} + LG + PB + Age^{-3}: OC + period$
✓	×	✓	✓	✓	×	$Age^{-1} + Age^{-2} + Age^{-3} + Age^{-3}: SB + Age^{-3}: OC + Age^{-1}: LG + period$
✓	×	✓	✓	×	✓	$Age^{-1} + Age^{-2} + Age^{-3} + Age^{-1}: SB + Age^{-1}: PF + Age^{-1}: LG + Age^{-2}: LG + Age^{-3}: LG + Age^{-2}: PF + period$

Figure 9: Female survivor

Rating factors						Functional form
Age	PB	SB	LG	OC	PF	
✓	×	×	×	×	×	$Age^{-1} + Age^{-2} + Age^{-3} + period$
✓	×	×	✓	×	×	$Age^{-1} + Age^{-2} + Age^{-3} + Age^{-1}{:}\,LG + period$
✓	✓	×	×	×	×	$Age^{-1} + Age^{-2} + Age^{-3} + Age^{-3}{:}\,PB + period$
✓	✓	×	✓	×	×	$Age^{-1} + Age^{-2} + Age^{-3} + Age^{-1}{:}\,PB + Age^{-1}{:}\,LG + period$

Figure 10: Male survivor

Rating factors						Functional form
Age	PB	SB	LG	OC	PF	
✓	×	×	×	×	×	$Age^{-1} + Age^{-2} + period$

5 Ill health and non-ill health pensioners ('I' and 'N' curves)

The users of VitaCurves may require the ability to assess mortality for a disabled pensioners (ill health) or non-disabled pensioners (non-ill health). However, Canadian VitaCurves, calibrated to our experience data, specifically relate to all health data (so make no allowance for retirement health). A convention is therefore needed as to how to predict the post retirement mortality where retirement health is either non-ill health or ill health.

We have adopted a pragmatic approach of constructing 'non-ill health' and 'ill health' curves, by adjusting the corresponding all health curves. This approach of adjusting the all health curves has been adopted to ensure consistency between all health, ill health and non-ill health curves. The adjustment factors are created, based on statistical and actuarial criteria, using only plans where complete data is available for retirement health (i.e. retirement health data is provided for the majority of both living and dead pensioners over the calibration period). Our Canadian VitaCurves technical documentation provides a more detailed description of the criteria used to determine the adjustment factors.

'Age only' curves are fitted for all health, non-ill health and ill health data, separately for male and female pensioners, based only on data from pension plans with good quality retirement health data as discussed above. We selected the quadratic model with age functional form for both male and female pensioners when fitting ill health and non-ill health models as part of this process.

The adjustment factors at each age are then derived, based on the ratio of ill health or non-ill health fitted mortality rates to all health fitted mortality rates, separately for men and women. The non-ill health adjustment factors are then applied to the set of more granular all health pensioner curves (as well as the age only all heath curves) to create corresponding non-ill health curves. For ill health, we do not publish more granular curves for male or female pensioners, and simply create the age only ill health curve by applying the adjustment factor to the all health age only curves.

6 Ensuring the model is a good fit to the data

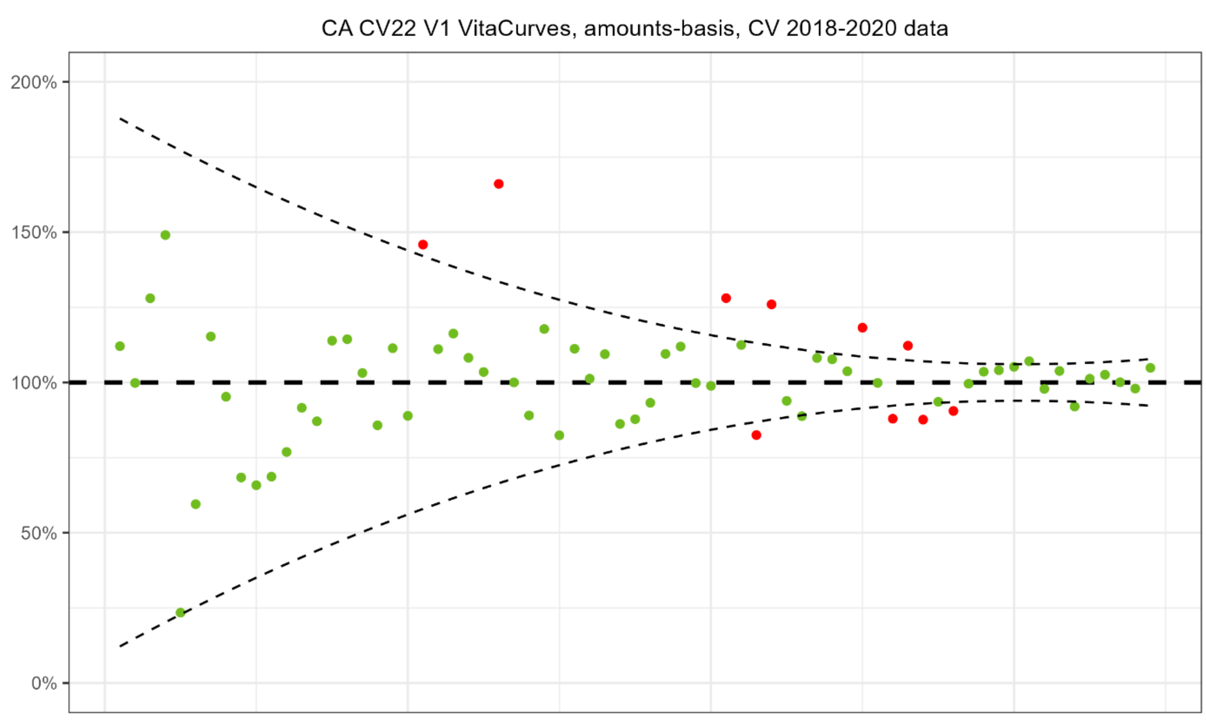
Having identified and chosen the main variables to include in the models and fitting the models, we then need to examine whether the models provide a good fit to the mortality experience of the pension plan participants observed in our dataset. There are a variety of summary test statistics and measures that are useful to assess the goodness of fit of a model. A full assessment cannot rely on one number, and instead we consider the results for multiple tests applied to each curve to identify that a model is a good fit to the underlying data. These same statistics are also used as a mechanism for comparing the merits where there may be multiple candidate models, for example for how we capture the shape of mortality with age, or how a variable changes mortality rates at different ages. More details of the statistical and actuarial tests that we apply can be found in the Canadian VitaCurves technical documentation.

6.1 Goodness of fit at the plan level

The tests described above ensure that the VitaCurves are a good fit to the dataset as a whole. However, they will be ultimately applied to value the liabilities of individual pension plans. It is important, therefore, that the curves are predictive of experience at the plan level. To assess this, we consider the ratio of actual deaths to those predicted by the VitaCurves for each plan. We do this on an amounts basis (i.e., weighting each life/death by the pension amount in payment to ensure that we capture the financial materiality of the larger pension amounts). Figure 11 shows the resulting actual-to-expected (A/E) ratio for each plan (arranged by increasing liability from left to right). Reassuringly, there is a broadly even split between the number of plans for which this ratio is above and below 100%.

The funnel dashed lines in Figure 11 represent a 95% confidence interval for the A/E ratios. The plans marked in green are those falling within the confidence interval and those marked in red fall outside. Again it is reassuring that there are only a handful of such plans.

Figure 11: Actual-to-expected (A/E) ratio by each plan with 95% confidence intervals



7 Ensuring internal consistency

We also ensure that the resulting VitaCurves are internally consistent. For example, we would expect that:

- For a man and woman who are alike in all other ways, the male mortality should exceed the female mortality;
- As we increase a covariate value (e.g. affluence), but keep all other covariate values the same, the mortality should decrease (except possibly for the oldest old);
- Former blue collar workers should have higher mortality rate than those who served as white collar workers (all other covariates alike); and
- Disabled pensioners should have higher mortality than non-disabled pensioners.

While the results within any GLM should generally be internally consistent, the use of multiple GLMs to cover different combinations of available rating factors increases the risk that we may have some internal inconsistencies. We therefore assess the number of consistency issues in VitaCurves, including 'unexpected' crossings (i.e., pairs of curves where the order of which has the highest mortality rates crosses at some point over the age spectrum) against specified criteria and tolerances. The criteria capture both the nature of consistency issue[11], and whether it is of concern[12].

With a large number of VitaCurves, it is likely that there will be a modest number of inconsistencies, and so tolerances are set on the number of inconsistencies identified. Breaching of certain thresholds is likely to lead to alternative GLMs being fitted. Where these thresholds are not breached, an experienced member of Club Vita's research team looks closely at the any inconsistencies identified before the VitaCurves are published.

8 Filling out the mortality table (extensions)

Our modeling provides mortality rates for the age range that we have fitted over. To value annuity benefits, we need to 'fill out' the VitaCurves to cover the very oldest ages. We also fill out the VitaCurves down to younger ages so that they can be used with younger participants.

8.1 Completing the tables at the highest ages

Extending the VitaCurves to the highest ages requires a balance to be struck between a range of desirable features, including continuity, monotonicity, maintaining consistency between curves, biological reasonableness and smoothness.

The approach we have used builds upon a detailed study carried out by the High Age Mortality Working Party of the Institute and Faculty of Actuaries, published over a series of Working Papers[13]. For each VitaCurve the approach requires two steps:

1. Identify a natural 'parent curve' to extend relative to; and
2. Extend the VitaCurve by exponentially decaying the gap between the VitaCurve and the parent 'curve' at a chosen rate.

[11] We require that VitaCurves should be monotonically increasing with age, similar granularity curves should be in the expected order (men above women) and less granular curves should be within the extremes of more granular curves.

[12] Some inconsistencies may be acceptable on grounds of materiality (e.g., number of ages impacted, or in size of inconsistency).

[13] See CMI Working Paper 106.

In order to implement this approach, we need to identify a suitable parent curve and an appropriate rate of decay.

8.1.1 Choice of parent curves

For male/female pensioners and survivors, each granular (i.e. not 'age only') VitaCurve is assumed to converge to a parent of the corresponding 'age only' VitaCurve (so the more granular male all health pensioner curves have the male all health pensioner age only curve as a parent, etc).

The 'age only' curves for male/female ill health pensioners and non-ill health pensioners, as well as survivors, have the corresponding male/female all heath pensioner curve as a parent. This just leaves the age only all health pensioner curves – these are extended using parent curves based on Canadian general population data as sourced from Statistics Canada. We have created this table using the published mortality rates (for 2018-2020)[14] for ages up to 109 and then converging the mortality rate at 109 to a mortality rate of 63% at age 115 using non-linear interpolation[15,16].

8.1.2 Rate of decay

Based upon the rate of convergence of mortality rates at the oldest ages in the underlying data, we have set the rate of decay for each group of VitaCurves to an appropriate level. Note the chosen rates are unchanged from CV21.

8.2 Providing mortality rates at the younger ages

The fitted VitaCurves have been extended to younger ages using a linear extension in the logit scale from the youngest age fitted (age 60) to the corresponding male/female Statistics Canada 2018-2020 Life Table rates at age 30. For ages 18 to 30 the VitaCurves mortality rates are set to be equal to the Statistics Canada 2018-2020 Life Table rates.

---

[14] Statistics Canada data available here.

[15] $q_{115} = 63\%$ has chosen consistent with studies that the force of mortality appears to flatten at around 1 at advanced ages

[16] The non-linear interpolation uses a power-law interpolation to provide curvature to the mortality rates at the oldest ages and takes the form of $\mu_x = \left(\frac{115-x}{6}\right)^C \mu_{109} + \left(1 - \left(\frac{115-x}{6}\right)^C\right)\mu_{115}$ for ages 109 to 115, where C is set to 1. Conversion between q_x and μ_x using uniform distribution of deaths ('UDD') assumption.

9 Allowing for the COVID-19 pandemic

CV22 is calibrated to Club Vita data over 2018 to 2020. It is therefore the first version of Canadian VitaCurves to include any of the COVID-19 period. We need to consider what, if any, allowance to make for the elevated mortality rates arising from the pandemic.

In the corresponding UK and US versions of CV22 we published two distinct versions:

- ‘BAU’ VitaCurves, labelled as v1, constructed using the usual calibration process; and
- ‘Adjusted’ VitaCurves, labelled as v2, where adjustments were applied to ‘strip out’ excess 2020 mortality.

However COVID-19 did not actually have a material impact on mortality rates in Canada in 2020, with mortality increasing by a much lower amount than in the UK and USA. In particular, we saw that:

- At the population level, the excess mortality in Canada (5.8%) for 2020 was significantly lower than that seen in the US (16.8%) and England & Wales (13.5%)[17].
- This 5.8% excess mortality in 2020 would be diluted over the 3 years used in calibrating CV22, and so would only decrease liabilities by of the order of 0.5%. This may not be considered sufficiently distinct from the Business as Usual (BAU) model to support the publication of two sets of tables.
- Amongst defined benefit pensioners, we would expect to see lower excess mortality than in the general population (as seen in UK and US), driven by factors such as vaccine take up and adherence to public health measures like social distancing and mask wearing, as well as the general better health of those who were healthy enough to at some point accrue defined benefits (while the general population includes those who were never fit too work).
- Indeed mortality rates in 2020 in the Canadian Club Vita data were only around 1% higher for men (and c4% higher for women) than predicted by the pre 2020 trend.

It is our view that reflecting the latest data (including 2020) in the calibration of CV22 under a ‘BAU’ approach could be justifiable, and could be viewed as a credible estimate of mortality rates, when used in combination with a ‘conventional’ mortality improvement assumption.

We therefore made a pragmatic decision not to publish ‘Adjusted’ VitaCurves which ‘strip out’ excess mortality for Canadian CV22 (unlike in the UK and US), and so only to publish one version of CV22 for Canada.

We are conscious that the approach that we have taken may not suit all circumstances. Users may therefore wish to apply their own adjustments to the BAU VitaCurves, if they have differing views on the appropriate levels of adjustment.

17 https://www.clubvita.net/ca/news-and-insights/top-charts-22-03-2021-excess-deaths

10 Some considerations for users of the VitaCurves modeling

Our VitaCurves capture a wide range of characteristics, and as such are designed to be a ‘neutral’ or ‘best estimate’ of baseline mortality. They have no deliberate prudence, nor are there areas where they knowingly overstate or understate mortality. This is demonstrated by the goodness of fit tests outlined in Section 6.

However, the curves represent an average mortality rate by age across individuals with the same longevity characteristic(s). As such, there will be some situations where plan specific considerations will lead the user to consider the most appropriate application of VitaCurves. This includes:

- **Known distortions to rating factors:** The VitaCurves capture the impact of factors such as pension amount on mortality rates. For a minority of plans, pension amounts may be particularly inflated (or indeed low) due to the history of pension accruals. Examples include where:
    - benefits are linked to length of service and the workforce has exceptionally low turnover (so potentially a bias to high pensions whereas average experience is for participants generally to have pensions spread over multiple plans); or
    - there is very high turnover / short service due to the cessation of defined benefit pension benefits.

  For such plans the user may wish to consider adjusting the pension amounts as part of using our model.

- **Disabled / Non-disabled pensioners:** We adjust our all heath curves to provide curves for disabled or non-disabled pensioners separate from all health pensioners. The disabled curves are based on records for participants who were eligible for disability provisions. The all health curves are based on the data of all pensioners regardless of retirement health. Users should be aware that:
    - If their plan does not record if a participant was eligible for disability provisions, then the pensioner curves may slightly under-state the likely mortality rates of the plan.
    - There is a variety of definitions used by plans in determining disability provisions and as such the disabled pensioner curves reflect the average level of provision across the plans in our data. For individual plans, the eligibility criteria may therefore differ to that implicit in our models.
- **Very high pensions:** If a plan has very high pensions, for example a plan focused on the executive management, then it is likely that the plan will find its liabilities highly concentrated in our top pension band. In such circumstances, the plan’s pensions may be skewed to larger pensions than seen on average in our top band and so may experience lighter mortality (higher life expectancies) than typical for that pension band. Where postal code information is available, it will help mitigate this (as it is likely that the participant will also be in the top longevity groups).
- **Grieving effect:** The mortality rates suggested for the spouse are calibrated to the mortality experience of individuals currently in receipt of a surviving spouse pension. Users should note that, in line with usual industry practice, our survivor VitaCurves include a ‘grieving effect’ and as such, may overstate mortality of the spouse prior to the original pensioner dying if applied before, as well as after, commencement of the spouse’s pension.
- **Allowing for the COVID-19 pandemic:** Different users of VitaCurves are likely to have differing views on the extent to which adjustments are required to baseline mortality in light of the pandemic, and how to combine these base tables with improvements to estimate mortality rates during the pandemic and post-pandemic periods. Users may therefore wish to apply their own adjustments to the v1 BAU CV22 curves, depending on their exact approach.

11 Want to know more?

If you have any questions on this document or would like to know additional details regarding our methods for fitting our Canadian VitaCurves, please contact any of the team below. We would be delighted to hear from you.

Saeed Ahmadi PhD
Saeed.Ahmadi@clubvita.net

Conor O'Reilly
Conor.OReilly@clubvita.net

August 2023

For and on behalf of Club Vita LLP

Reliances and Limitations

In this paper (the 'Research'), Club Vita LLP has provided an overview of the methodology used for the calibration of the seventh generation of Club Vita's Canadian VitaCurves. The Research is based upon Club Vita LLP's understanding of legislation and events as of July 2023 and therefore may be subject to change. Future actuarial measurements may differ significantly from the estimates presented in the Research due to experience differing from that anticipated by the demographic, economic or other assumptions. The Research should not be construed as advice and therefore not be considered a substitute for specific advice in relation to individual circumstances and should not be relied upon. Where the subject of the Research refers to legal matters please note that Club Vita LLP is not qualified to give legal advice, therefore we recommend that you seek legal advice if you are wishing to address any legal matters discussed in this Research. Please be advised that Club Vita LLP (not its respective licensors) does not accept any duty, liability or responsibility regarding the use of the Research, except where we have agreed to do so in writing.

When read along with the 'Data underpinning CV22 Canadian VitaCurves' paper, this paper complies with the Financial Reporting Council's Technical Actuarial Standard (TAS) 100: Principles for Technical Actuarial Work.