Critical illness insurance rates: are they changing over time and how?

Dr George Streftaris & Dr Erengul Dodd

Work with Chunxiao Xie

21 May 2018

www.actuaries.org.uk/arc
Modelling, Measurement and Management of Longevity and Morbidity Risk

• Major research programme funded by the Actuarial Research Centre of the Institute and Faculty of Actuaries running from 2016 to 2020

• Significant supporting funding from the Society of Actuaries and the Canadian Institute of Actuaries

• Themes
  • Development of new single and multi-population models for mortality and new sub-population mortality datasets
  • Drivers of mortality and cause of death analysis
  • Longevity risk management
  • Stochastic models for critical illness insurance
Outline

• Critical illness insurance
• Data
• Stochastic modelling
  – Delay time distribution (diagnosis to settlement)
  – Claim rates
• Claim rates comparison
• Pricing rates
Critical illness insurance
Critical illness: Policy description

- Fixed term policy, usually ceasing at age 65
- A fixed sum insured payable on the diagnosis of one of a specified list of critical illnesses
- Covers: Cancer; *Death*; Heart attack; Stroke; Multiple Sclerosis; Total & permanent disability; Coronary artery bypass graft; Kidney failure; Major organ transplant etc.
- Policies are often sold together with term or endowment insurance
- Benefit type: Full Accelerated (FA) or Stand Alone (SA)
Data

Provided by the CMI Assurances Committee
CII data supplied by CMI:

• 1999-2005
  – Details of policies in force at the start and end of each year
  – 19,127 claims settled

• 2007-2010
  – Grouped by various risk factors
  – 20,487 claims settled
Data:

- Claims
- Exposures
- Risk factors:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (last birthday)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Gender</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Smoker</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Policy duration</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Office</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Distribution channel</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Benefit type (accelerated, standalone)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Benefit amount</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Policy type (single, joint)</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Settlement year</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Cause</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Product category</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Date of diagnosis</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>
Data: 2007 - 2010

Gender
- Female: 54%
- Male: 46%

Smoker Status
- Non-smoker: 76%
- Smoker: 24%

Benefit Type
- Accelerated: 11%
- Stand-alone: 89%

Sum Assured
- £0-£25,000: 29%
- £25,001-£75,000: 31%
- £75,001-£125,000: 9%
- £125,000+: 16%
- Unknown: 15%

Distribution Channel
- Bancassurer: 21.0%
- IFA: 29.2%
- Multi-tied: 18.3%
- Single-tied: 31.4%
- Unknown: 0.1%

- Distributions very similar between 2007 – 2010 & 1999 – 2005
- Slightly higher proportion of F and NS in 2007 – 2010

- Lower proportion of age 16-30 in 2007 – 2010
- Stand-alone only ~ 11% of claims data
Modelling

Mostly Bayesian stochastic
Stochastic modelling

- Estimation & smoothing of CI diagnosis rates
  - how do these depend on risk factors?

- Diagnosis is the insured event and there is a delay between diagnosis and settlement

- The exposure corresponds to claims settled, not to claims diagnosed

- This can lead to biased rate estimates; need to adjust it

- Also take into account uncertainty
Stochastic modelling
Delay time distribution (1999-2005)

• Diagnosis date not always recorded or available
  – 18% diagnosis dates missing

• Observed data: mean delay 185 days; sd 263 days

• Fit a delay distribution (GB2 in Bayesian GLM-type setting):
  – \( F(d; x, z) = Pr(\text{claim diagnosed age } x, \text{ risk factors } z, \text{ will be settled in } d \text{ days}) \)

\( D_i \sim \text{Generalised Beta2}(\alpha, \tau, \gamma, s_i) \)

\[
f_d(d_i) = \frac{\Gamma(\alpha + \gamma)}{\Gamma(\alpha)\Gamma(\gamma)} \frac{\tau(d_i/s_i)^{\tau-1}}{d_i \left[ 1 + (d_i/s_i)^{\tau-1} \right]^{\alpha+\gamma}}
\]

\[
E(D_i) = \exp(\eta_i) = \exp \left( \beta_0 + \sum_{j=1}^{8} \beta_j z_{ij} + \beta_{9,k} + \beta_{10,l} \right)
\]

with \( s_i \) given as function of \( \eta_i, \alpha, \tau, \gamma \).
Stochastic modelling
Delay time distribution (1999-2005)

- Most factors significant:
  - Policy duration, amount, death: shorter delay
  - Single life, stroke, multiple sclerosis: longer delay

- Non-recorded diagnosis dates estimated through delay distribution $F()$

- Data (exposures) adjusted to allow for non-settled claims

\[ E^*(u; x) = E(u; x) \times F(t-u; x) \]
Stochastic modelling

- Diagnosis date **not available**
- Assume similar delay distribution
- Match claims with common characteristics (age, policy duration etc)
- Adjust exposures as in earlier data
Stochastic modelling: Claim rates

Model:

Fit Bayesian model:

\[ N^{(j)}(x; \theta) \sim \text{Poisson} \left( \lambda_{x; \theta}^{(j)} \int_{u=0}^{4} E(u : x; \theta) F^{(j)}(4 - u : x; \theta) \, du \right) \]

- \( \lambda_{x; \theta}^{(j)} \): diagnosis (claim) rate for cause \( j \) at age \( x \) with risk factors \( \theta \)
Stochastic modelling:

Perform variable (factor) selection

Selected model includes:

✓ age (older ↑)
✓ smoker status (S ↑)
✓ distribution channel
✓ benefit type (stand-alone ↓)
✓ age x smoker
Stochastic modelling:

Selected model includes:

- ✓ policy duration (longer ↑)
- ✓ benefit amount (mid ↑)
Claim rates

Smoothed estimates, intervals
Claim rates

- Model fits crude rates (2007 – 2010) well
- 2007 – 2010 rates significantly higher
- Gap widens at younger ages
Claim rates

- Again, 2007 – 2010 rates significantly higher
- Rates higher than for Pol Duration 1
Claim rates

- Accelerated 2007 -2010 (black) higher than stand-alone (green)
- Both significantly higher than 1999 – 2005
Claim rates
Smokers & non-smokers (Accelerated, Pol Duration 1)

- 2007 – 2010 rates significantly higher, both S & NS
Claim rates
Different benefit amount (Accelerated, Smokers)

- 2007 – 2010 rates significantly higher, also for different amount
Pricing

Annual premium, paid at constant rate, n-year term:

\[
\text{Net Premium} = \text{Benefit Amount} \times \frac{\int_{t=0}^{n} v^t t p_x \lambda_{x+t} \, dt}{\int_{t=0}^{n} v^t t p_x \, dt}
\]

where

\[ t p_x = \exp \left( - \int_{s=0}^{t} \lambda_{x+t} \, dt \right) \]

and

\( v \) is the discount factor.

Then bootstrap distribution of \( \lambda \)s used to derive CIs for premiums.
Pricing

All causes, Smoker, Age 40, Policy duration 0, Benefit amount £100k, i=3%

- Since 2007 – 2010 FA rates are higher than 1999-2005 combined rates, the net premium rates are also higher.
All Cancers Excluding Non-melanoma Skin Cancer (UK)

In 1999-2005 dataset

- 49% of the claims were caused by cancer
- Death 17.6%
- Heart attack 11.6%
- CABG 2.1%

Source: cruk.org/cancerstats
Future trends of CII claims

- Cancer forms almost half of the CII claims.
  - Availability of screening (e.g. colonoscopy, mammography)
  - Social/behavioural changes (e.g. obesity, alcohol consumption)
  - New treatments (e.g. targeted immunotherapy)
  - Statistical advances (e.g. use of big data, AI methods)
Conclusions
Conclusions

• CII claimants distribution similar between 1999-2005 & 2007-2010
  (but not necessarily true for insured population)
• Time between diagnosis and settlement of a claim is important
• Claim rates (2007-2010) depend on a number of risk factors including:
  – age, smoker status, distribution channel, policy duration, benefit amount and benefit type
• Analysis suggests increase of CII claim and premium rates over time (1999-2005 v 2007-2010)
  – especially at younger ages
Continuing work

- Fit more sophisticated Bayesian model to allow for more variation in rates (e.g. hierarchical, negative binomial)
- Use of population morbidity statistics
- Liaise with CMI for knowledge exchange on data, modelling
- Compare with CMI rates
The views expressed in this presentation are those of the presenter.
Stochastic modelling: Delay time distribution
1999 – 2005 (cont.)

• Generalised Beta 2 distribution in Bayesian GLM-type setting

Most factors significant:

• Policy duration, amount, death, CABG: shorter delay

• Single life, stroke, mult sclerosis: longer delay

Figure: Posterior means (dots) and 95% credible intervals (bars) of β’s.