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

Dr George Streftaris & Dr Erengul Dodd

Work with Chunxiao Xie

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: 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
Data

CII data supplied by CMI:

- **1999-2005**
  - Details of policies inforce at the start and end of each year
  - 19,000 claims settled
- **2007-2010**
  - Grouped by various risk factors
  - 25,187 claims settled

<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:
- Claims
- Exposures
- Risk factors:
Data: 2007 - 2010

- Gender
  - Female: 52%
  - Male: 48%

- Smoker Status
  - Non-smoker: 61%
  - Smoker: 39%

- Benefit Type
  - Accelerated: 35%
  - Stand-alone: 65%

- Sum Assured
  - £0-£25,000: 23%
  - £25,001-£75,000: 26%
  - £75,001-£125,000: 17%
  - £125,000+: 32%
  - Unknown: 20%

- Distribution Channel
  - Bancassurer: 52%
  - IFA: 23%
  - Multi-tied: 14%
  - Single-tied: 17%
  - Unknown: 10%


- Distributions more “even” in 2007 – 2010
- Higher proportion of smokers in 2007 – 2010

- Higher proportion of age 16-30 in 2007 – 2010

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):
  - \( F(d; x, z) = Pr(\text{claim diagnosed age } x, \text{ risk factors } z, \text{ will be settled in } d \text{ days}) \)

\[
\begin{align*}
D_i &\sim \text{Generalised Beta2}(\alpha, \tau, \gamma, s_i) \\
\gamma_0(d_i) &= \frac{\Gamma(\alpha + \gamma)}{\Gamma(\alpha)\Gamma(\gamma)} \frac{\tau(d_i/\beta_0)^{\gamma}}{\beta_0[1 + (d_i/\beta_0)^\gamma]} \\
E(D_i) &= \exp(\eta_0) - \exp\left(\beta_0 + \sum_{j=1}^{u} \beta_j z_j + \beta_{0,k} + \beta_{0,l}\right)
\end{align*}
\]

with \( s_i \) given as function of \( \eta_0, \alpha, \tau, \gamma \).
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 $F(t)$

• 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:

![Diagram](image)

Fit Bayesian model:

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

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


Perform variable (factor) selection

Selected model includes:

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

\[ \text{Net Premium} = \text{Benefit Amount} \times \frac{\int_0^n v^t \mu_x \lambda_{x+t} dt}{\int_0^n v^t \mu_x dt} \]

where

\[ v^t \mu_x = \exp \left( - \int_{s=0}^{t} \lambda_{x+t} dt \right) \quad \text{and} \]

\[ v \] 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%

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

- Critical illness insured population distribution has some differences between 1999-2005 & 2007-2010
- 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 a 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
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, multiple sclerosis: longer delay

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