SOA PREDICTIVE ANALYTICS SEMINAR
SEOUL, SOUTH KOREA
SEP. 8, 2017

Session 1

Application of Predictive Analytics in Life Insurance

Winnie Liu, FSA, FCIA, FSAS

SOA Predictive Analytics Seminar

8, Sept. 2017

SOCIETY OF ACTUARIES

Applications of Predictive Analytics in Life Insurance

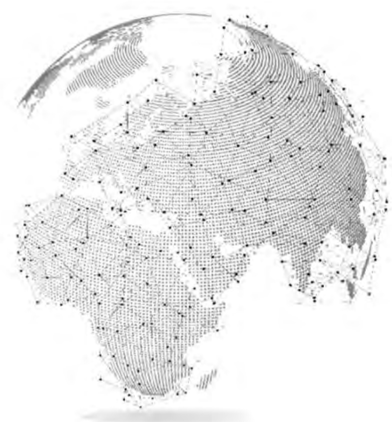
WINNIE LIU, FSA, FCIA, FSAS

Head of Actuarial Modelling & Experience Study, Munich Re – Asia Pacific, Middle East & Africa (Life)

September 8th, 2017

SOCIETY OF ACTUARIES

Let's think…



Big data, artificial intelligence and advanced analytics is transforming business models across the globe.

What does this mean for the future of insurance business?

A changing reality for insurers

Today's reality

- Ever-changing data formats require constant development and adoption
- New sources of data provide greater insights, might trigger new questions
- Smart data is connected to everything, changes customers' behaviour
- Data access and improved insights change the way we do business
- Technology, analytics methods and human skills are improving daily

What it means for insurers

- New data formats and sources create opportunities for differentiation
- Mobile phone data could show driver and driving behaviour, lead to different underwriting outcome
- Continuous increase in number of data sources imply need for constant algorithm development
- Increased need for closer collaboration across all business disciplines
- Insurance-specific know-how in a business development context needs to be developed

Predictive analytics used across the insurance value chain

Pricing and product development

- Are **best estimate** assumptions adequate?
- What **pricing basis** would be accurate for a new product?
- What are the **risk drivers** and how do they affect the current experience?
- Which **product features** appeal most to the target segment?

Sales & marketing

- What concrete **up-sell** opportunities exist in the current portfolio?
- Which clients are most likely to take up **cross-sell** offers?
- Which groups should we target for sales **campaigns**?
- Which distribution **channel** is performing best/worst?

Underwriting

- Are there any groups for which we can **simplify the underwriting process** (to improve customer experience)?
- Can we **reduce** the need for **medical exams** (and the underwriting cost)?
- How can we use underwriting **resources most effectively**?

Portfolio management

- How **profitable** is the business really?
- Which customers are at **highest risk of lapsing** and how can we retain them instead?
- How does the **portfolio composition** compare with pricing assumptions?

Claims

- How good is the **risk selection** process?
- Are we attracting poorer risks than we intended to?
- How can we **streamline the claims process**?
- Which claimants should I prioritize for fraud investigation?
- Can we **standardize** the claim cause by text mining?

Munich Re examples of Value-added through Predictive Analytics*

*For confidentiality reason, some variables/numbers in the examples may have been modified for demonstration purpose only

Big Data and Analytics Platform
Designed to serve our clients

25	physical server machines, including 18 data notes
1,75TB	SAS in-memory capacity (LASR server)
288	Core Hadoop cluster
5	Core Hadoop / Spark memory
170	HDFS storage net (510TB)

CASE STUDY 1

Up-selling
with no or limited underwriting

Problem statement:

"I want to offer existing customers an increase in cover with a very simple sales process."

"Who should be eligible and which underwriting questions should I ask in an up-selling campaign?"

Our approach:

- Developed a predictive model to estimate the underwriting decision (using Tree-boosting algorithm)
- Applied the model to the data to identify customers eligible for up-selling
- Design and execute campaigns based on risk appetite

XXX 000	lives analysed
XX	Variables collected on application form, plus underwriting decision
YY	Variables predict the underwriting result with similar accuracy to 55

The outcome:

Compared to previous campaigns, take-up rate

Doubled

CASE STUDY 1 CONTINUED

Up-selling campaign design

1 Up-selling campaign with no additional underwriting

Direct offer

No underwriting

- For customers with no medical loadings, capped age and in-force policy duration
- X 000 customers offered option to increase sum insured (up to defined limit)
- No underwriting required. Results of predictive model used to estimate likely underwriting outcome.

2 Up-selling campaign with limited underwriting

Call centre

Limited underwriting

- Eligibility extended to customers with limited medical loadings, no limitation to in-force policy duration
- XX 000 customers offered option to increase sum insured (up to defined limit)
- Customers only need to answer a few questions about their medical condition to take up offer.

TAKE-UP RATE **DOUBLED** COMPARED TO PREVIOUS CAMPAIGNS

CASE STUDY 2

Streamlined underwriting *with loyalty programme data*

Problem statement:

"I want to significantly reduce the underwriting requirements with a minimal price impact while retaining sound risk management."

"Which variables from my loyalty programme data can streamline an up- and cross-selling campaign?"

Our approach:

- Developed predictive models of the underwriting decision with machine learning methods a limited number of variables
- Applied the models to the data to identify eligible customers for streamlined underwriting

XXX 000 lives analysed

XXX Application form and loyalty programme variables analysed

YY Variables predict decision with similar accuracy to 150

The outcome:

85%

Reduction in underwriting requirements to identify standard risks

(with <5% false positive rate)

>XX%

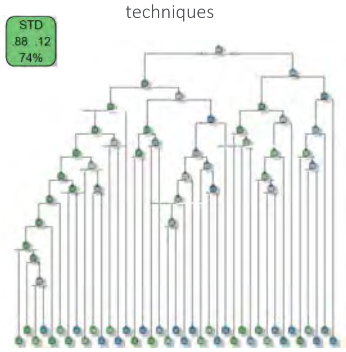
of customers qualify for streamlined up-selling

CASE STUDY 2 CONTINUED

Streamlined underwriting - *modelling consideration*

Which customers should I offer streamlined underwriting to?

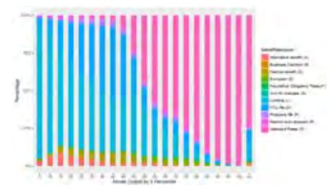
Criteria that define eligible customers are identified with machine learning techniques



Which predictive models should I use?

Advanced analytics methods are benchmarked to identify the best performing models, e.g. random forest.

How much (mortality) allowance should I make for false positives?



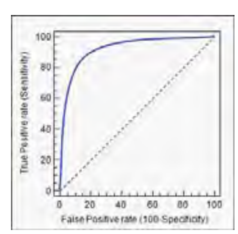
5 - 15 variables identify 45 - 60% of customers within which >90% are standard risks

CASE STUDY 2 CONTINUED

Streamlined underwriting – *the price of simplification*

- Trend towards: 1 click, elimination of fluids, walk-in bank customers
- Risk of anti-selection, non-disclosure, deteriorating experience
- Can also consider answering existing questions better vs. removing existing questions – Multiply data!
- Question of reliability of proxy data
- Quantifying the price impact of Type I
- Objective is not just minimising the false positive rate
- But also…maximising the true positive rate

	Predicted decision	
Actual decision	Standard	Non-standard
Standard		
Non-standard		



CASE STUDY 3

No fluid underwriting *with 3rd party data*

Problem statement:

"For customers under 50 with sums insured under $1 million, I want to remove all fluids in the underwriting process with a negligible price impact."

Our approach:

- Developed a predictive model to estimate underwriting outcome using 3rd party data instead of fluids
- Determined the impact on price if predictive model is used instead of traditional underwriting (with fluids)
- Deployed and monitored the model at the point of underwriting via an automated underwriting platform

X00 Variables (from application form and 3rd party data) analysed

YY Variables (excluding fluids) predict the underwriting result with similar accuracy to 300

The outcome:

Conclusion: It is possible to predict the underwriting decision using external data rather than fluids, with a minimal price impact.

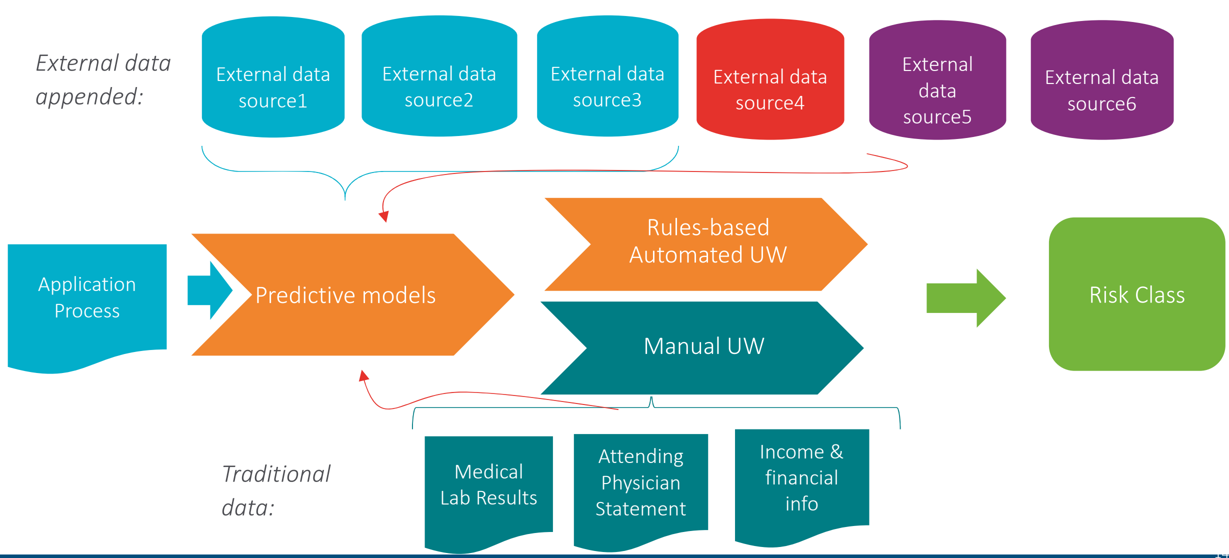
67% Less underwriting requirements

0 Invasive tests in underwriting process

<X% Net impact on mortality

CASE STUDY 3 CONTINUED

No fluid underwriting – *work flow*



CASE STUDY 3 CONTINUED

No fluid underwriting – *predictors identified*

Demographic factors:
- Age
- Gender
- …

Build:
- Height
- Weight
- …

Motor Vehicle Records:
- Total risk factors
- Number of violations in last 3 years
- …

Treatment (Rx):
- High count
- Medium count
- …

Questions and disclosures:
- Anxiety, depression and bipolar disorder
- Disorders or diseases of the blood, skin, thyroid, lymph or other gland
- …

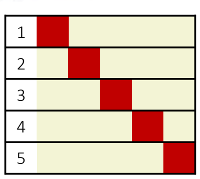
Unlike rules, predictive models use a multi-variate approach and is a weighted combination of all input parameters

CASE STUDY 4

Algorithmic pricing
taking wisdom out of the experience data

Advanced analytics models tailored for mortality/incidence rates

Advanced analytics methods combined with our actuarial expertise in identifying biometric risk

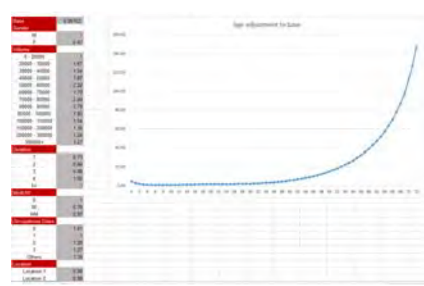


Which are the best performing predictive models?

- Which combination of rating factors and interactions should I include in my pricing basis?
- Which produce higher predictive power than my existing rates?

Easy to use pricing model that can be scaled much faster

Calculator of mortality incidence with standardised, interpretable adjustments per rating factor



CASE STUDY 4

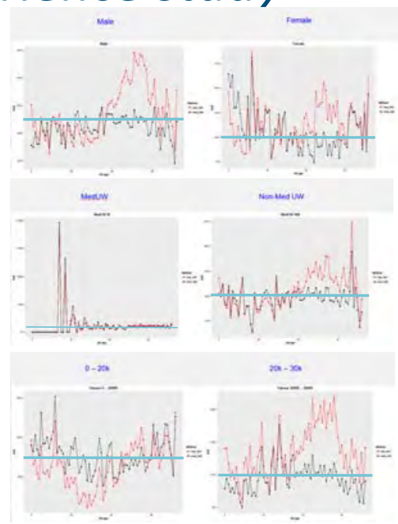
Algorithmic pricing
advantages over traditional experience study

Disadvantages of traditional experience study

- All experience data within the study period is used in the study – may lead to over-fitting
- Importance of each risk factor is not easy to identify
- Correlation of the risk factors is hidden – adjustments to risk factors may overlap
- Lack of ability to predict the "unseen" future experience

Advantages of algorithmic pricing (non-traditional experience study)

- Experience data is divided into training/validation (used to train the model) and testing (unseen future) data
- Risk factors are ranked based on their importance to help select the features to be used in the model
- Interaction of the risk factors can be easily validated by looking at the testing deviance
- Model performance is always tested on the testing (unseen future) data



CASE STUDY 5

Insuring previously uninsurable risks

Problem statement:

"How can I start offering HIV-positive applicants the same fully underwritten cover as all other applicants?"

Our approach:

- Derived an underwriting guideline using the largest private healthcare cohort of HIV positive lives globally.
- Predictive model applied to derive extra mortality (EM) loadings per sub-group of applicants.
- Guideline adopted in MIRA, available to clients

XXX 000 Patients' data analysed

10 Years of observation with longitudinal medical markers of disease progression

The outcome:

40% reduction in best estimate claims cost

X Continents launched new product / solution

>50% of HIV positive applications accepted, (from 0% for some)

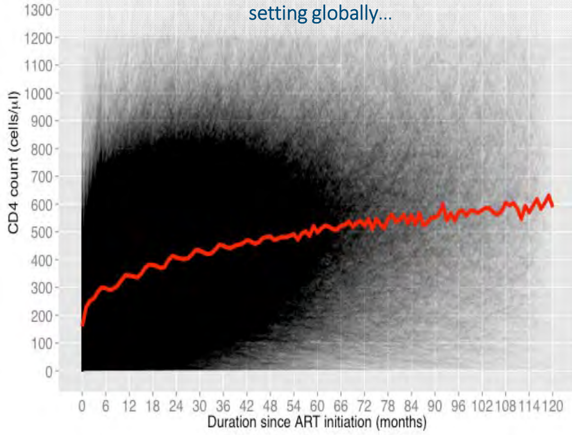
100% Reduction in cost of continuous underwriting

400% Increase in annual new business volumes since 20XX

CASE STUDY 5
CONTINUED

Insuring previously uninsurable risks
HIV positive applicants

Munich Re accesses data on the **largest cohort of HIV-infected lives receiving treatment in a private health care setting globally…**



"…**using predictive analytics methods**, Munich Re produced an underwriting guideline that enables clients to underwrite HIV like other impairments. HIV was often declined before or offered unaffordable terms."

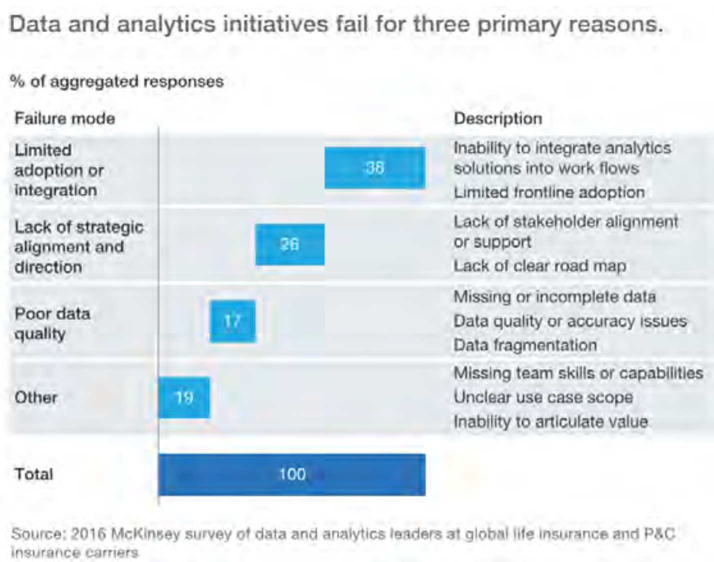
How to be successful in predictive analytics

Key groups:

- Sponsor
- Advisor
- Business/Product owner
- **Working team**
- Support
- End user

Working team:

- Business/Team lead
- Project manager
- Business analyst
- Machine learning expert
- Actuarial predictive modeler
- Data engineer
- Data scientist
- Data journalist



SOCIETY OF ACTUARIES®