MILLIMAN CASE STUDY

Using the Milliman Advanced Risk Adjusters™ (MARA™) tool in the UK PMI market

Joanne Buckle, FIA
Tanya Hayward, FIA
Kishan Desai

Milliman

In this case study, we discuss risk categorisation techniques in the context of UK private medical insurance (PMI) and how the Milliman Advanced Risk Adjusters™ (MARA™) tool has been applied successfully in this environment.

What is population stratification?

Population stratification is the process that allows us to stratify a population by predefined characteristics (e.g., age/sex/clinical condition), where members within a particular population stratification group are considered to have similar risk profiles. For example, in the context of healthcare, we may expect patients within the same group to have similar levels of healthcare resource utilisation. Risk adjustment is the process that allows us to analyse the healthcare resource utilisation of these groups by taking their specific risk profile characteristics into account. A successful population stratification and risk adjustment methodology will allow us to:

- Understand the risk profile of a particular population/ sub-population
- Assess how risk profiles have changed over time
- Compare risk profiles among different populations

Challenges in the UK PMI market

Traditional risk adjustment methodologies rely on comprehensive and accurate member-level data in order to be effective. Developing a robust population stratification methodology in the UK PMI environment has some challenges which prevent us from being able to develop a complete clinical and claims profile for covered lives.

UK PMI plans have limited primary care coverage and limited explicit coverage for chronic conditions or emergency and maternity services. The restricted benefit coverage means that there is only partial patient medical information and, consequently, comorbidity profiles are not available. Because UK PMI mainly covers elective services, a large proportion of covered lives will have no claims experience within a year and we are not able to build any expected claims profiles for these members based on claims data alone. There are also data limitations where secondary diagnosis and procedure codes are not captured.

Can a risk adjustment algorithm do an effective job in population risk stratification, given these data and system challenges, such that we will be able to better understand member risk profiles within the context of UK PMI? To answer this question, we explore using our MARA tool to see what insights we are able to glean about members' expected relative healthcare resource utilisation risks.

We note that we have produced a similar article[1] exploring how our Chronic Condition Hierarchical Grouper™ (CCHG™) tool for population stratification has been applied to UK PMI populations and, in a later section of this article, we demonstrate how the outputs of MARA and the CCHGs can be combined to provide further powerful insights.

WHAT ARE THE CCHGS?

The CCHGs were developed by Milliman in the United States in association with Dr. Michael Chernew, a Harvard University health economist and coeditor of the *American Journal of Managed Care*. The tool assigns individuals to unique categories using a clinically relevant hierarchy based on how healthcare providers make treatment decisions. It considers the entire set of diseases that a member faces and how they interact. All members are assigned to 43 mutually exclusive categories over a 12-month rolling look-back period.

[1] Buckle, J., Hayward, T., & Singhal, N. (March 2018). A Case Study: Risk Adjustment in the UK PMI Market. Milliman White Paper. Retrieved November 28, 2018, from http://www.milliman.com/uploadedFiles/insight/2018/risk-adjustment-uk-pmi-market.pdf.

[2] More information on the CCHG tool can be found on Milliman's Medinsight website at http://www.medinsight.milliman.com/MedInsight/Products/Medinsight-Tools/?prid=71829

The four key questions

Before embarking on any population stratification process, we ask ourselves the following four key questions, defined by Lisa Iezzoni in *Risk Adjustment for Measuring Healthcare Outcomes*,[3] shown in Figure 1.

FIGURE 1: THE FOUR KEY QUESTIONS

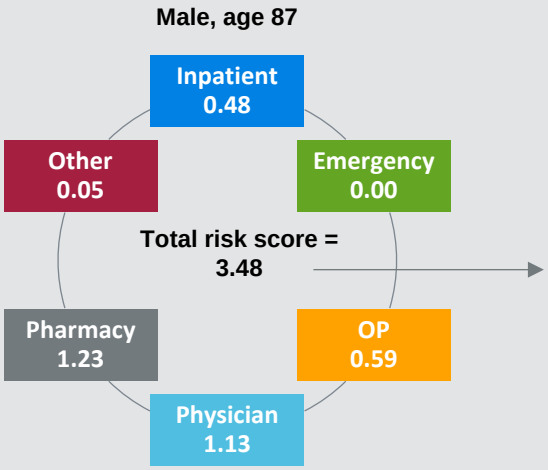
QUESTION	EXAMPLE ANSWERS
Risk of what outcome?	High claims experience, mortality, hospital admission or readmission.
Over what timeframe?	One year, hospital admission or clinical episode.
For what population?	Entire membership, clinical definition, regional stratification or member characteristic such as age group.
For what purpose?	Disease management programme, alternative reimbursement, provider profiling or clinical analysis.

MARA tool

MARA is a population stratification tool that uses longitudinal data assets and advanced statistical methods to calculate a total risk score per member in the population being considered. The total risk score represents the expected overall healthcare resource utilisation for each member relative to the average member in the population. This risk score can be further broken down into inpatient, outpatient, physician, emergency, pharmacy (Rx) and other service categories. These results can also be used to calculate each member's likelihood of hospitalisation within a 12-month period. In addition, MARA produces output summarising each member's clinical conditions as well as whether and how they contribute to the overall risk score.

MARA IN ACTION

The healthcare resource utilisation for this particular policyholder is expected to be 3.48 times as great as the average policyholder in the population. This risk score is made up of subcomponent risk scores for each major service category to allow for more detailed analysis at this level. We are able to garner further clinical insights by understanding the composition of the risk score by clinical condition and determining what contribution each clinical condition that the member has makes to the total risk score.



Conditions driving risk score of 3.48	
Condition	Contribution to risk score
Glaucoma	77.1%
Back Sprain or Strain	22.9%
Entropion	0.0%
Keratosis	0.0%

[3] Iezzoni, L. (2012). Risk Adjustment for Measuring Healthcare Outcomes, Fourth Edition.

Returning to our four questions, MARA can be applied in the following way:

FIGURE 2: MARA AND THE FOUR KEY QUESTIONS

KEY QUESTION	APPLICATION OF MARA
Risk of what outcome?	Healthcare resource utilisation and cost
Over what timeframe?	One year
For what population?	Can be applied to entire population or subpopulation
For what purpose?	Provider profiling, evaluating return on investment of care management programmes, disease management, clinical analysis, informing benefit design and product development processes, developing population budgets and alternative reimbursement arrangements.

Meaningful results

We used our UK PMI Health Cost Guidelines™ (HCGs[4]) data to conduct a feasibility study and found that the data was of sufficient quality and granularity to apply the MARA tool.

Although the MARA tool has been calibrated on a sample of US commercial data, it has produced a robust set of meaningful results when applied to UK PMI data. In the US, an independent study by the Society of Actuaries found that MARA was the best-performing risk adjuster, with the most predictive power, when compared to other risk adjustment tools available in the market.[5] A future possibility for the tool is to develop a UK-calibrated version, which would likely lead to additional performance gains.

MARA has three types of predictive models which are selected according to the purpose of the exercise and the data available. The DxAdjuster is run solely on medical claims data, the RxAdjuster is run solely on prescription drugs data and the CxAdjuster is run using a combination of medical claims and prescription drugs data.

Each of these three models can be run on a concurrent or prospective basis. The concurrent model calculates the expected resource use in the 12-month base period, based on the claims inputs, while the prospective model predicts resource use in the 12-month period following the lag period, based on the 12-month base year claims inputs. For our purposes, we have opted to use the MARA DxAdjuster Concurrent model due to the lack of prescription drugs coverage and data in the UK PMI market. Although the DxAdjuster model does not use prescription drugs data as an input, it produces an Rx risk score based on demographic and claims data.

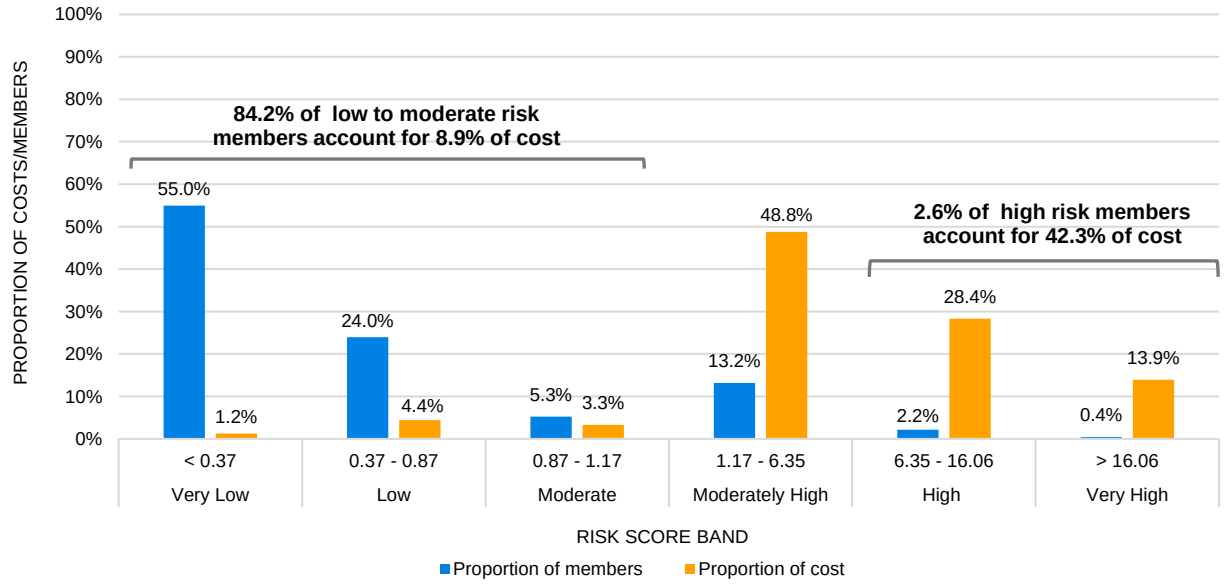
MARA calculates a risk score for each member included in the population that relates that person's expected healthcare resource utilisation and cost—expressed as per member per month (PMPM) costs—to the population average. For ease of interpretation, these risk scores are grouped into categories ranging from very low to very high.

Assessing the distribution of members and cost by risk band, shown in Figure 3, we see that MARA produces results that we would expect to see from a successful population stratification exercise. That is, we are able to identify the small proportion of high-risk members who are responsible for a large proportion of costs and to differentiate them from the large proportion of lower-risk members responsible for a substantially smaller proportion of costs.

[4] Milliman UK PMI HCGs are a tool for modelling healthcare cost and utilisation by service categories based on data we collected from PMI contributors in 2015. This data covers several million life-years for the analysis period (2012-2015).

[5] Society of Actuaries (2016). Accuracy of Claims-Based Risk Scoring Models. Retrieved November 28, 2018, from https://www.soa.org/Files/Research/research-2016-accuracy-claims-based-risk-scoring-models.pdf.

FIGURE 3: DISTRIBUTION OF MEMBERS AND COSTS BY MEMBERS' RISK SCORES, 2014



Using additional data categories, the tool can provide useful insights and results validation to help explain the relationships between a member's risk score and other relevant factors.

MARA is able to successfully consider clinical and demographic data inputs and synthesise them into a single, meaningful risk score to accurately categorise members' risk levels. For example, in the charts in Figures 4 and 5 we observe the resulting MARA risk scores by age band and location compared to PMPM costs observed in each grouping. The MARA risk score results we observe within each of these categories follow a similar pattern to what we would expect. We observe an increase in average risk score as age increases as well as higher risk scores in London and the South East of England compared to other regions.

These insights can be applied in a multitude of areas. For example, in provider and hospital profiling initiatives, we could compare the overall risk profile of patients by hospital group to draw meaningful comparisons when benchmarking hospital cost and service mix.

FIGURE 4: RELATIVE PMPM AND MARA RISK SCORES BY AGE GROUP, 2014

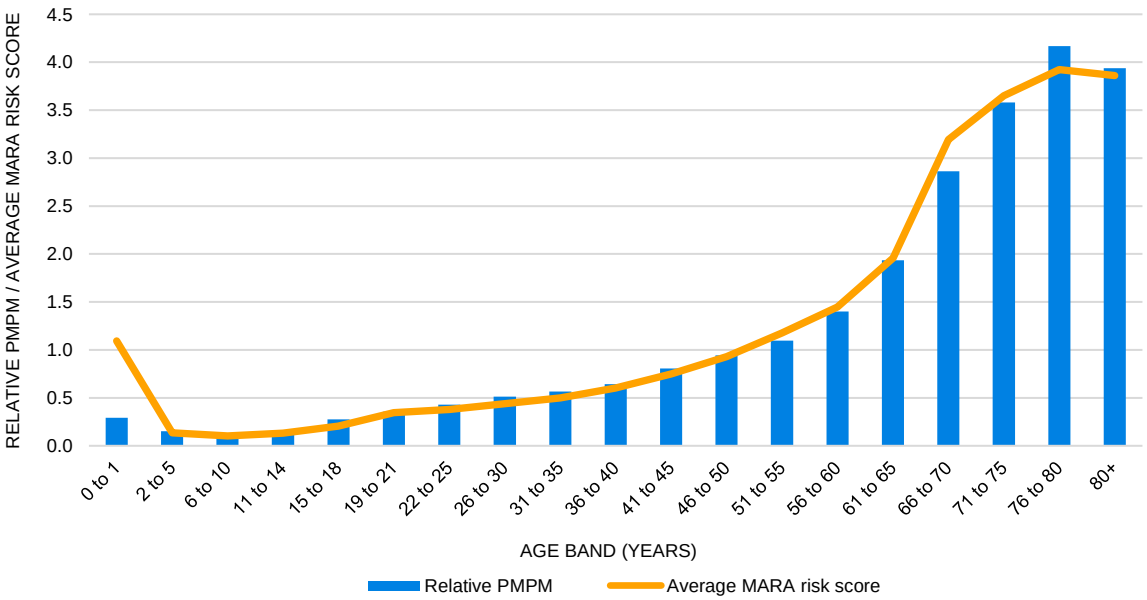
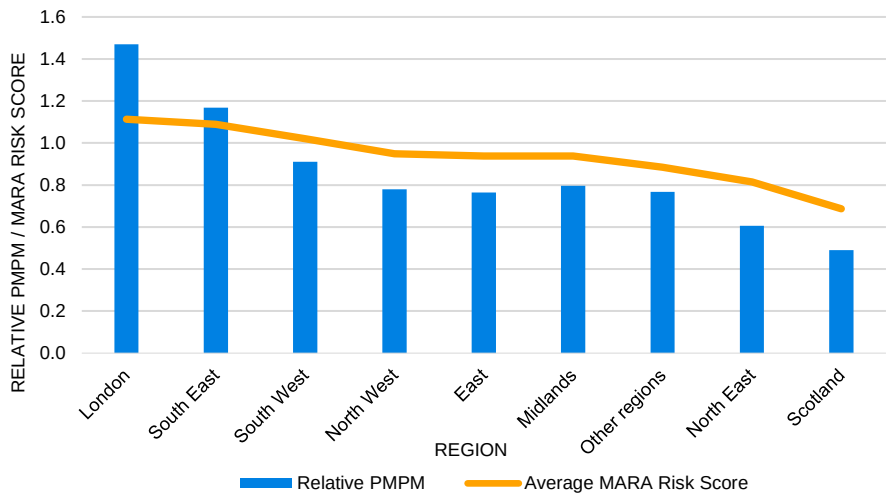


FIGURE 5: RELATIVE PMPM AND MARA RISK SCORE BY LOCATION, 2014



Combined power of MARA and CCHGs

By combining the MARA tool with the CCHGs, we amplify the power of our risk stratification approach as it allows us to investigate the following:

1. MARA risk scores for each CCHG category.
2. Variation and distribution of MARA risk scores within each CCHG category.

In Figure 6, we illustrate how we are able to calculate the average MARA risk score within CCHG categories and how they broadly follow the patterns we would expect (i.e., higher-ranking CCHG categories generally have higher risk scores, fewer members and higher proportions of costs compared to lower-ranking CCHG categories).

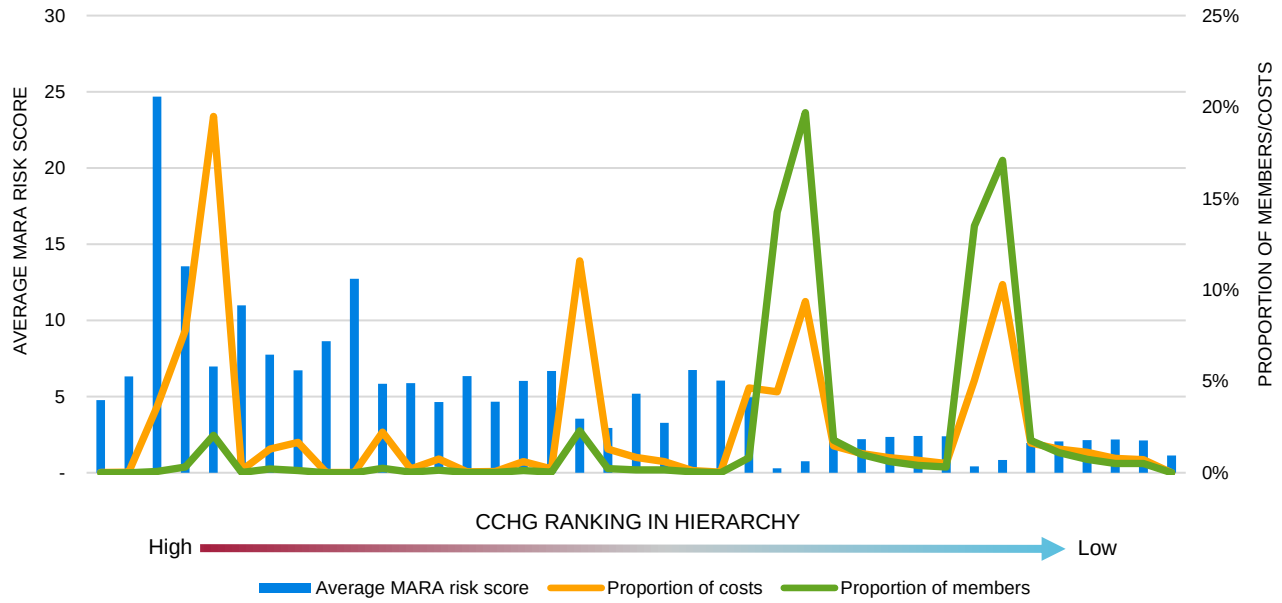
Note that this expectation would not apply strictly across all categories and we would not necessarily expect a strictly downward trend of PMPMs as CCHG categories descend in hierarchy ranking, because the CCHGs hierarchy ranking relates to severity of condition and prioritisation of treatment rather than expected cost. These insights could prove useful when designing and monitoring care management programmes, as they would be powerful aids from the early stages of case-finding right through to measuring return on investment (ROI).

FIGURE 6: COMBINED MARA AND CCHG RESULTS FOR TOP 10 CCHGS IN HIERARCHY

CCHG	AVERAGE MARA RISK SCORE	PMPM RELATIVITY FACTOR[6]	PROPORTION OF COSTS	PROPORTION OF MEMBERS
Major psychosis	4.8	8.3	0.0%	0.0%
Severe dementia	6.3	11.2	0.0%	0.0%
Active cancer, three or more cancers	24.7	56.6	3.6%	0.1%
Active cancer, two cancers	13.6	26.3	7.8%	0.3%
Active cancer, one cancer	7.0	10.2	19.5%	2.0%
Renal failure, post-transplant	11.0	8.8	0.1%	0.0%
Severe rheumatic and other connective tissue disease	7.7	6.4	1.3%	0.2%
Severe heart failure/transplant/rheumatic heart disease/non-rheumatic valvular heart disease	6.7	14.9	1.7%	0.1%
Hemophilia, sickle cell and chronic blood disorders	8.6	8.8	0.0%	0.0%
Both coronary artery disease and diabetes	12.7	9.0	0.0%	0.0%

[6] PMPM relative to average PMPM for population under consideration.

FIGURE 7: AVERAGE MARA RISK SCORE, PROPORTION OF COSTS AND MEMBERS BY CCHG CATEGORY



Clinical insights

The value of the MARA tool extends beyond calculated risk scores to meaningful clinical insights. We are able to classify members by each clinical condition they have and summarise them by condition group or chronic status.[7] We are then able to calculate the average MARA risk score within each condition group as well as the relative contribution of conditions to the total risk score.

For example, members with spinal stenosis have an average MARA risk score of 5.5 and, on average, more than half of this score is directly attributable to spinal stenosis. This shows that, on average, members with this condition have a risk score that indicates 'moderately high risk' and when this condition is present, it is a major contributor to the overall risk score. Conversely, members who have had UK PMI claims for headaches are also typically considered moderately high risks but headaches contribute a smaller proportion of their total risk scores—less than 20% of the overall average risk score of 4.8. This signals that higher-risk members assigned to the headache condition typically have other conditions contributing a greater amount to their overall risk scores. For insurers designing disease management programmes around specific conditions, it is helpful to understand how much future resource for a cohort of members with each condition will be driven by that specific condition and how much by peripheral costs.

Combining the risk scores and clinical outputs from the MARA tool has many potential applications. For example, it could be used in the design of care management or disease management programmes, for case finding or for provider profiling to help identify more efficient providers, to perform clinical analysis and in many other areas where combining relative risk scores with clinical classifications is required. Traditionally insurers have tended to focus on analysis by disease and condition area, rather than on cohorts of distinct patients. This one-dimensional view has limited the cost-effectiveness of management interventions to control costs and utilisation.

[7] Although UK PMI coverage does not typically include chronic conditions, we do sometimes see chronic conditions coded in claims data where members experience acute exacerbations of their chronic conditions.

FIGURE 8: CLINICAL OUTPUT INSIGHTS FROM MARA

CONDITION	CONDITION GROUP	CHRONIC INDICATOR	PREVALENCE RATE PER 1,000 LIVES	AVERAGE MARA RISK SCORE WITHIN GROUP	RELATIVE CONTRIBUTION OF CONDITION TO TOTAL MARA RISK SCORE
Other Derangement of Joint	Musculoskeletal System	non-chronic	18.6	5.1	35.3%
Fatigue, Asthenia Other Than Chronic Fatigue Syndrome	Symptoms, Signs, and Ill-Defined Conditions	non-chronic	17.8	5.3	0.0%
Ligament Injury – Other	Musculoskeletal System	non-chronic	17.0	3.9	20.5%
Arthralgia	Musculoskeletal System	non-chronic	16.6	3.4	20.6%
Back Sprain or Strain	Musculoskeletal System	non-chronic	13.4	3.3	12.1%
Abdominal Pain	Symptoms, Signs, and Ill-Defined Conditions	non-chronic	11.5	6.1	23.0%
Spinal Stenosis	Musculoskeletal System	non-chronic	10.7	5.5	50.9%
Other General Screenings	Services/Therapy/ Vaccines/Exams	non-chronic	8.0	6.6	no contribution
Ligament Injury – Ankle	Musculoskeletal System	non-chronic	7.0	3.4	17.6%
Headaches	Nervous System and Sense Organs/Nervous	non-chronic	6.7	4.8	16.7%

Goodness of fit

The R^2 measure was used to test the goodness of fit for the MARA tool. Each member's expected PMPM cost was calculated by multiplying each member's normalised individual risk score by the average PMPM. We censored the data using the inter-quartile range method by calculating a censor point for each MARA risk band. Members with total annual claims costs in excess of the relevant censor point had their costs adjusted downwards (censored) to the censor point.[8] Censoring the data improves the goodness of fit by removing the impact of extreme outliers.

We compared these goodness-of-fit results to a simple deterministic age/sex adjustment and found that the MARA tool adds tremendous predictive power.

When applying the CCHGs to the UK PMI data, we segmented the 'active cancer' category further by the number of distinct cancer diagnoses a member had in a year and found that this improved the goodness of fit significantly. The goodness-of-fit results for MARA are better than CCHGs alone but using CCHGs with further cancer stratification produces a marginally better goodness-of-fit result than MARA. The results produced by each tool are complementary, with each providing different insights that can be used in a meaningful way on a standalone basis or in conjunction with each other.

Although a significant proportion of costs are censored in all methods, the associated proportion of members whose costs are censored is approximately 1%. This indicates that a minor proportion of members are responsible for a high proportion of the outlier costs.

The dramatic improvement in goodness of fit from using MARA compared to age/sex alone indicates a key result:

MARA is able to synthesise each member's clinical and demographic information into a single score that has significantly greater predictive power than the usual age/sex risk adjustment techniques used by actuaries, health economists and statisticians.

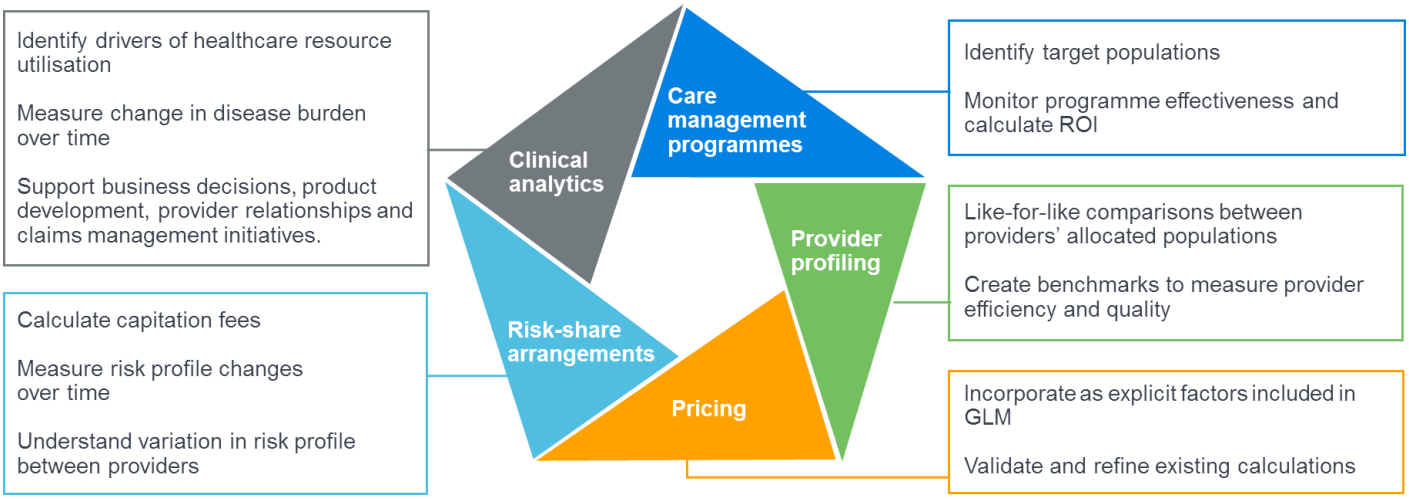
[8] Censor point = 25th percentile + K * (75th percentile – 25th percentile). We found that using a value of K = 3 provided the best balance of proportion of data censored versus goodness of fit.

FIGURE 9: GOODNESS OF FIT WITH AND WITHOUT CENSORING

POPULATION STRATIFICATION METHODOLOGY	R^2 BEFORE CENSORING	R^2 AFTER CENSORING	PROPORTION OF MEMBERS WITH RECORDS CENSORED	PROPORTION OF COSTS CENSORED
Age/sex	2.7%	8.0%	1.1%	23.7%
CCHGs	11.2%	22.3%	1.1%	19.0%
MARA tool	13.6%	25.6%	1.1%	23.0%
CCHGs with further cancer stratification	17.2%	28.8%	1.0%	13.9%

Applications of population stratification with risk adjustment in the UK PMI environment

FIGURE 10: APPLICATION OF POPULATION STRATIFICATION IN THE UK PMI ENVIRONMENT



Conclusion

Using MARA, a risk score was assigned to each member included in our data set, based on a variety of clinical and demographic input factors. The MARA tool proved to have strong predictive power and using this tool together with our Chronic Condition Hierarchical Grouper (CCHG) enhances the value of these insights further. Due to its powerful predictive capabilities, this tool would be perfectly applicable in the context of clinical and cost analytics, pricing, provider profiling, care management programmes, developing risk-share arrangements and various other predictive modelling applications for population stratification and population health management.

CONTACT

Joanne Buckle
joanna.buckle@milliman.com

Tanya Hayward
tanya.hayward@milliman.com

Kishan Desai
kishan.desai@milliman.com