Counting the benefits - evaluation of healthcare programs

Bronwyn Hardy and Barry Leung

Abstract

There have been numerous studies that illustrate the benefits of chronic disease management programs. Many of these studies utilize different evaluation methodologies, making it difficult to compare the value offered by the initiatives.

The paper will provide guidance for actuaries who perform disease management program evaluation for private health insurers, on issues to consider when designing the measurement system and evaluation criteria including: selection of comparative groups, determination of baselines and consideration of intangible factors.

The authors would like to thank Rory Atchison and David Watson for their invaluable contribution in providing insightful peer review. Any errors or omissions remain the sole responsibility of the authors.

Keywords: Health management; Health outcomes; Health program evaluation, Broader Health Cover, Private Health Insurance

Introduction

Around the world, illness and death from chronic disease is increasingly widespread and represents a significant cost to both individuals and society. In addition to the personal and community costs, chronic diseases result in a significant economic burden due to the combined effects of health-care costs and lost productivity from illness and death. According to the AIHW, estimates based on allocated health-care expenditure indicate that the four most expensive disease groups are chronic—cardiovascular diseases, oral health[i], mental disorders, and musculoskeletal—incurring direct health-care costs of $27 billion in 2008–09. This represents 36% of all allocated health expenditure[ii].

Chronic diseases can have very broad impacts, on the individual, their friends and family, their immediate community and society generally. Some impacts include: pain and suffering, increased medical costs (to the individual or society), productive time lost and reduction in quality of life.

To help manage both the long and short term impacts of chronic diseases, disease management programs have been developed to provide patients with coordinated community support and to empower them to self-manage their diseases. Many believe that such initiatives, whilst they may not stop the disease, they may help alleviate the symptoms, slow disease progression or reduce secondary impacts.

Private health insurers in Australia have had a long running interest in disease management. Assisting policyholders to effectively manage their condition can reduce the long term claims costs, as well as improving member retention and satisfaction levels. It also aligns with the objectives of many insurers such as fostering policyholder wellbeing.

An important part of the disease management framework is evaluation and continuous improvement of programs.

The relatively short history of disease management in Australia means that the evaluation approaches adopted remain highly variable and are under constant review and development. The purpose of this paper is to provide a high level introduction to disease management program evaluation for Australian private health insurers. It does not provide an assessment of different methodologies, focussing instead on key concepts and issues to consider. Although the target audience is the evaluation teams within private health insurers, evaluators from other areas of the health sector may also find this information useful.

1. Disease management program – an overview

What is a disease management program?

There are many different definitions of disease management program. The European Observatory on Health system and Policies, in partnership with WHO, suggested that "definitions of disease management (programmes) vary substantially. Common features are: (1) an integrated approach to care/coordination of care among providers, including physicians, hospitals, laboratories and pharmacies; (2) patient education; and (3) monitoring/ collecting patient outcomes data for the early detection of potential complications"[iii].

In the USA, descriptions range from "discrete programs directed at reducing costs and improving outcomes for patients with particular conditions"[iv] to "a system of coordinated health care interventions and communications for populations with conditions in which patient self-care efforts are significant"[v]. The second definition represents a shift from single disease focus to a whole of patient focus.

In Australia, health management programs tend to target patients with chronic disease like diabetes and coronary heart disease. In recent years, the range of target audience has broadened to include persons with significant risk factors but no confirmed diagnosis of chronic diseases. The public sector in particular often supports people who have one or more of a broad range of diseases and conditions.

Why perform a program evaluation?

Program evaluation is an important stage in the lifecycle of any disease management program. Evaluation provides feedback to stakeholders on program efficacy and effectiveness, as well as opportunities to learn and improve.

There are many reasons why an evaluation may be performed, including:

- clinical efficacy assessments
- program effectiveness evaluation by funders (private or public)
- needs assessments
- feasibility studies
- clinical governance or process evaluations

The evaluations most commonly reported externally are clinical assessment and program effectiveness evaluation by funders.

Evaluation for Clinical Purposes

A large proportion of evaluations are conducted by clinicians or biomedical corporations to demonstrate the effective treatment of a disease or condition. Many of these evaluations are focused on changes in clinical indicators such as cholesterol, body mass index, blood pressure, or similar. A number of these evaluations use simple models and statistical techniques; however there are also those that utilise highly sophisticated statistical modelling. In particular, when a "matching" comparative population is being selected, the algorithms adopted can be highly technical.

Alternatively, success may be measured by patient self-assessment. There are a number of recognised approaches that are widely adopted to assess an individual's perception of their own wellness.[vi] The results of these evaluations are naturally simpler to assess. These tools have gained fairly broad acceptance, however they are still sometimes questioned due to a perceived lack of rigour,
,particularly as the consistency between individual assessments can vary significantly.

In many clinical assessments, financial factors are evaluated only once the treatment is determined to be clinically effective and is therefore being assessed to take to market.

Evaluation by Funders

Funders are those likely to pay for the programs or initiatives. They may be governments or their agencies), or health insurers.
Governments typically evaluate programs to determine:

- If the initiative is appropriate to sponsor for the general, or a specific, population
- Whether government support should be granted: for example through coordinated care trials.

Similarly health insurers wish to determine whether the expected benefits, both tangible and intangible, are sufficient to justify the costs. There are a number of measures frequently used, as discussed below.

This paper is focussed on evaluation by funders, in particular private health insurers.

2. Considerations in program evaluation

Why is this so difficult?
Like many aspects of actuarial science, there is no unique way of performing program evaluation. Despite attempts by many organisations around the world to standardise evaluation techniques, there is no "gold standard" that all countries and organisations can apply. This makes health program evaluation an interesting mix of art and science – the evaluator must take into account the surrounding circumstances to decide the best approach.

For a program to be evaluated appropriately, the evaluation methodology must align with program objectives, ethics, interests of the target audience (usually the funder), environment in which the program is delivered, and program design. There are also many constraints that evaluators must consider like data collection, timeframe, cost and practical aspects of program delivery.

Regardless of the program, there are three key pillars to any evaluation for funding purposes:

- Cost (including fees, internal costs, incentives to participate etc) versus assessed benefits
- Quality / Outcome assessment
- Experience with program

The following sections detail some of the key considerations when designing an appropriate evaluation methodology.

What is the purpose of the program?
Understanding program objectives is probably the most important part of any program evaluation. The diversity of program objectives is one of the greatest differences between programs, and can significantly impact the assessed "effectiveness" of any given initiative.

For example, if the primary objective is retention of customers, the program may be considered a success if it leads to higher retention than other similarly priced retention strategies, regardless of the outcomes in other areas. However, if retention was only a secondary consideration, and a reduction in future claims cost the primary objective, the same program may be evaluated as unsuccessful. Different evaluators, and even different parties within the funder, may rank alternate objectives differently, and therefore vary in their assessments of the initiative.

It is important to consider the balance between different objectives prior to commencing the program.

Objectives can be expressed in many ways. Those that we encountered in our research and experience range from:

- Improving clinical measures – e.g. reducing weight, improving blood pressure, cholesterol or blood sugar levels
- Improving self-assessed "wellness" - noting that this can be significantly different to the clinical results

- Reduced financial impact - fewer future admissions, reduced length of stay, absolute savings (after allowing for program costs)
- Increased customer satisfaction and retention

The aim of the program will affect all aspects of its operations. All programs should have clear and specific objectives, a definition of what success looks like, and performance indicators that track progress against targets. Data requirements, collection methodology, and assessment criteria should ideally be defined before the program starts.

Defining the measure of success can be challenging, as it can be difficult to find meaningful benchmarks. However, it remains important to have a goal, to facilitate future decision making.

In addition, while most programs focus on future claims reduction, many programs have supplementary objectives, many of which are extremely difficult to quantify. These may include:

- Improving customer satisfaction – proving a causal link between a disease management initiative and an improvement in customer satisfaction can pose a number of challenges
- Reducing direct marketing spend and
- Creating a competitive advantage

Funders, providers and independence

In Australia, funding of health management program comes from both government and non-government sources. The funder usually determines the objectives of the program. Together with the service provider, the program is designed and implemented accordingly. Evaluation can be performed by the funder, service provider, or an independent specialist (e.g. health evaluation units within universities).

Things to consider when determining the most appropriate group to perform the evaluation include:

- Is the funder independent from the service provider?
- Does the funder and/or service provider have the necessary expertise in performing the evaluation?
- How important is "evaluation independence" to both funder and service provider?
- Is there any conflict of interest, and if yes, can that be managed? For example, a service provider may have a vested interest in proving the value of its program. If the service provider performs the evaluation, strict controls, such as agreed evaluation methodologies, must be in place to avoid potential bias.

Length of program

The length of the program will affect measurement timeframes. Some programs have a finite length. If that is the case, indicators can be measured over a defined period. However some programs do not have a definite end-date i.e. the patients are being managed continuously once enrolled, possibly changing between different levels of intensity in the care or coordination received. In this case, measurement methodology will need to cater for potentially "infinite" treatment.

In addition, there is also a "performance period" in which many programs are required to show success, in whatever format has been defined, in order for the initiative to continue. This may be a short as six months, or could extend over many years.

Many programs are expected to show results over the longer term (eg weight reduction), and may be difficult to assess objectively in the period immediately following the program. In addition, some programs benefit from "top up" support at given intervals, to ensure that the individual has retained the benefits of the initiative. Further complicating this issue is that when (and if) the benefits are realised the individual may no longer be a policyholder of the sponsoring insurer. This issue is neatly resolved for government funders, who have a lifetime responsibility to their constituents.

Performance measures and data collection

The type of data collected can be classified into six basic groups:

Type	Examples	Data source
Clinical – pathological	• Cholesterol Level • HbA1c[vii] • Vitamin D level • Blood Pressure	Pathology test
Clinical – self reported	• Vaccination against flu or hepatitis • Body Mass Index (BMI) • Adherence to medical guidelines e.g. regular intake of an appropriate dosage of statins to control cholesterol • General health status using validated self-reporting tool like SF-12®[viii]	Patient Questionnaire
Lifestyle	• Level of physical activity • Diet • Smoking status • Alcohol consumption	Patient Questionnaire
Financial and utilisation	• Claims (or benefits) paid • Number of admissions and readmissions • Number of bed days utilised • Average length of stay per admission • Attendance at emergency rooms • Number of physician visits • Medicare spending or admissions	Health administration system
Customer Care / Satisfaction	• Net Promoter Scores • Number of complaints recorded	Customer Questionnaires / Feedback
Care Delivery Performance	• Enrolment rate • Retention rate • Engagement of service providers	Case management system

The type and number of indicators collected will depend on a number of factors including:

- Objective of the program

  The data collection should support the evaluation of whether the program achieves its objectives.

- Coaching channel

  The channel through which the program is delivered will determine the way data is collected and may reflect any bias within the datasets. A face to face coaching method will provide the health coach an opportunity to measure the height and weight of the patient in person objectively. If telephone coaching is utilised, the coach would usually rely on the weight and height information reported by the patient, which may contain bias.

- Cost of data collection

  While more data is usually better than less, there are both tangible and intangible costs associated with data collection. For example:
  - The more questions that the health coach asks over the phone, the longer the length of call, which in turn increases the cost of the program. There is also a risk that patients are more likely to disengage.
  - If patients are asked to fill a long questionnaire regarding their health status, their enthusiasm towards the program may reduce and also affect the data quality of self-reported measures.

- Other data limitations

  In Australia, the health system is fragmented and there is currently no central depository of all health related information across the systems. Every part of the system holds specific pieces of data and linking datasets between systems has been impossible to date. Privacy concerns also prevent data sharing in some cases. For example, although the number of visits to the emergency department will be a useful indicator, a health insurer would have low awareness of that information as emergency room visits are funded by the state government's hospital system.

Patient consent and ethics approval

All programs should receive appropriate patient consent and ethics approval for data collection and evaluation. Medical and health information is treated as sensitive personal information under the Privacy Act, and obtaining consent from patients and appropriate authorities is critical.

The big picture

It should also be remembered that the perspective of the evaluator can have a significant impact on the assessment of any given program. An initiative which moves people from privately to publically funded services may be viewed favourably by a private health insurer, but could potentially result in a higher overall cost of treatment. A number of initiatives involve the private and public sectors working together, to provide a cohesive solution. An example of this is the Victorian Carepoint program launched last year, which is jointly funded by the state government and Medibank Private.[ix]

Evaluation Methodology

While most agree that a randomised controlled trial (RCT) is scientifically the appropriate way in evaluating effectiveness of health programs, it is not possible (or necessary) to perform a RCT in all instances. A RCT requires setting up a control group and this may not be possible at all times due to ethical considerations, regulatory constraints or a limited population base.

Other methods of evaluation include:

- Pre and post treatment comparison – a comparison of the patients' performance before, during and after treatment
- Participant vs non participant comparison - a comparison of performance of those who participate in the program and those who do not enrol in the program due to incorrect contact details or opt outs
- Population trend analysis – a comparison of utilisation trend of the whole population before and after implementation of the health program.

Choice of methodology will depend on many factors, including:

- Whether the program is new or "tried and tested"
- Purpose of the program
- Whether to include all potential patients into the program
- Potential patient volume
- Ease of setting up a comparable control group

Setting comparative groups

Regardless of the methodology chosen, a "like for like" comparison between treatment and non-treatment groups should be performed. The choice of methodology will determine the comparison group utilised.

- If a "pre and post" intervention methodology is used, the comparison will be made on the same group of patients before, during, and after treatment.

- If a RCT is conducted, patients who fit the program criteria will be split into control and treatment groups. For a fair comparison, demographic and clinical characteristics of both groups should align where possible.

- If a "participant vs non participant" comparison is required, the performance of patients who enrol in the program is compared against those who do not participate. In this case, depending on the reason for non-participation, those who enrol may exhibit different characteristics to those who do not. For example, those who opt out may have less inclination to change their behaviour. In that case, if a positive outcome is demonstrated through the program, the result may not be wholly due to the program.

Time periods and measurements
The time period that measurement should take place varies between programs. Clinical indicators and patient satisfaction data is usually collected at inception, at the end of program (if program length is finite), or on a regular basis (for continuous program). The time period when financial and utilisation evaluation should occur is more contentious. It is generally agreed that it will take some time after the program for the impact on utilisation rates to crystallise. Based on our literature review, measurement of financial and utilisation performance usually occurs 12 to 24 months after treatment.

Follow up
For programs that have a finite treatment period, it may be necessary to follow up patients post treatment to ascertain positive behavioural changes are maintained. Whilst follow-up is generally considered beneficial, there is a large variance in actual experience in different programs. Follow up, when it happens, usually occurs 6 and/or 12 months after treatment. In most cases, indicators gathered at the end of program are collected again during the follow up. Adverse movement in indicators may indicate reversal of benefits generated from the program.

Regression to Mean
People selected for disease management programs have typically had a recent hospital experience. In many cases, if the operation were successful, these people would not be expected to have a readmission for a number of years, and over time may even revert to health levels close to those of the general population. This trend is referred to as Regression (or Reversion) to the Mean (RTM), and can lead to a false positive outcome for an initiative if a simple pre/post analysis is performed. However, allowing for RTM also requires judgement, and adjustments are made in many ways, such as estimating the annual improvement in claims experience or the time taken to return to base population experience. This application of judgement reduces the credibility of many evaluations.[x]

Adjustment for other factors
Many programs attempt to eliminate or adjust for the impact of other factors which may influence the outcome of the program. Such adjustments are made in an attempt to more accurately validate the results, and remove the potential impact of secondary factors. Some of the factors typically adjusted for include:

- Age
- Sex
- Geographic location (often as a proxy for broader educational and wealth status)
- Co-morbidities and other clinical factors
- Other prescriptions / treatments

Again, as there is no recognised approach for applying such adjustments, it might be argued that they provide spurious accuracy, or the ability to change the apparent outcome of the program. Others, such as Ariel Lindon, a recognised expert in the field of disease management evaluation[xi], provide pragmatic and robust methodologies for matching populations or adjusting for secondary factors.

Typically, the need for such techniques is determined on a program by program basis, depending upon the assessment criteria and available data.

Other pitfalls
Evaluation must never be an after-thought. It should be built into the program development process, with a comparison group established at inception, and data collection considered as part of service delivery. An improper evaluation will not provide the necessary insights required to ascertain value and improve program quality.

The purpose of evaluation is not just to show the program "works". While assertion of return on investment is important in any evaluation, the evaluation process should also provide inputs into the post implementation review of the program where possible, and identification of improvement opportunities.

If continuous improvement is a significant objective of the program, this should be explicitly stated at the outset, so meaningful information can be obtained from the evaluation.

Ultimately, health status is affected by many factors. It is simply impossible to isolate the impact of an individual health program on a person's health. Funders generally wish to be assured by the program design and analysis that all steps have been taken to ensure confidence that the findings can be attributed to the program. In many trials it is assumed, unless otherwise proven, that other determinants of health remain constant before and after treatment, and therefore any observed impact is attributable to the program.

3. Global Experience in program evaluation

Whilst disease management and preventative initiatives are found globally, there has been significant research in the USA and Europe regarding effective methods of evaluation, and we briefly consider the learnings from these regions.

European Experience
The European Union recently commissioned a review of the evaluation of disease management programs in Europe. The project was called DISMEVAL (for Disease Management Evaluation), and involved representatives from a number of countries. Over a two year period, they studied the evaluation of programs in many European Countries.

The working party concluded (amongst other things): *Randomised controlled trials are widely considered as the gold standard for appraising the causal relationship between a complex intervention and clinical outcomes. However they are underused and often lack methodological rigour.*[xii]

The report also noted that: *Although an evaluation may detect statistically valid associations, the interpretation of such finding, in terms of their transferability into health care decision making, remains somewhat challenging.*[xiii]

The findings did not lead to a single preferred evaluation methodology, or definitive approach, instead highlighting the differences which may be apparent between evaluations even when using a similar approach.

American Experience

In the United States, there are a number of bodies and companies that have tried to establish a standard approach, from Healthways (a major disease management service provider) to the Population Health Alliance[xiv] to the Society of Actuaries, the RAND corporation[xv] and the Congressional Budget Office.. However, to date, it does not appear that any single approach has been broadly adopted.

The Disease Management Association of America (DMAA) published "Guide to Disease Management Outcomes Evaluation" in 2004. Far from establishing once and for all methodology and principles to be followed by practitioners, it is widely-agreed… that the guide falls short of the needs of the industry in this area. Accordingly, DMAA has convened another work group in 2006 to tackle the subject again.[xvi]

Despite further research across the industry, there still appears to be no general consensus on consistent evaluation techniques.

The US Society of Actuaries developed a methodology whereby the comparative results are based on the history of the cohort under investigation. This is known as a pre-post methodology. As with all methodologies, there are significant variations in how the methodology is applied, even when groups are, in theory, basing their evaluations on similar approaches. This is illustrated in the example below, both groups are adopting a pre-post evaluation technique, comparing dollars spent on hospital benefits, however the time periods adopted could show significantly different results.

Evaluator	Pre-period assessment	Program format	Post program assessment
Group 1	12 months prior to admission to program $ spent on hospital benefits	Six month fixed duration program	Twelve months commencing with program discharge $ spent on hospital benefits
Group 2	18 month pre first approach to join program (admission process typically takes 2 months) $ spent on hospital benefits	Variable length program depending on assessment of ongoing individual needs	18 months commencing six months after successful discharge from program, or six months after 12 months continuous enrolment on the program $ spent on hospital benefits

Key Learnings
First and foremost, there is no easy answer when it comes to the best approach to evaluating the effectiveness and financial value of disease management initiatives. It is important to link evaluation to performance goals of the program.[xvii] These can differ significantly.

From a funder's perspective, a key consideration in any evaluation approach is comparability; being able to meaningfully compare alternate programs, and thus ensure that funds are appropriately directed.

Moving forward, targeting appropriate individuals is likely to become as significant as evaluation, and a major factor driving program effectiveness. In fact, identifying the right participants is critical to program design. The program is unlikely to be successful if it attracts the "worried well" rather than those most in need.

4. The Australian perspective

Overview
Disease management is currently practised by many different organisations, ranging from local hospitals to state governments and health insurers. The table below provides selected examples of disease management programs carried out by different funders:

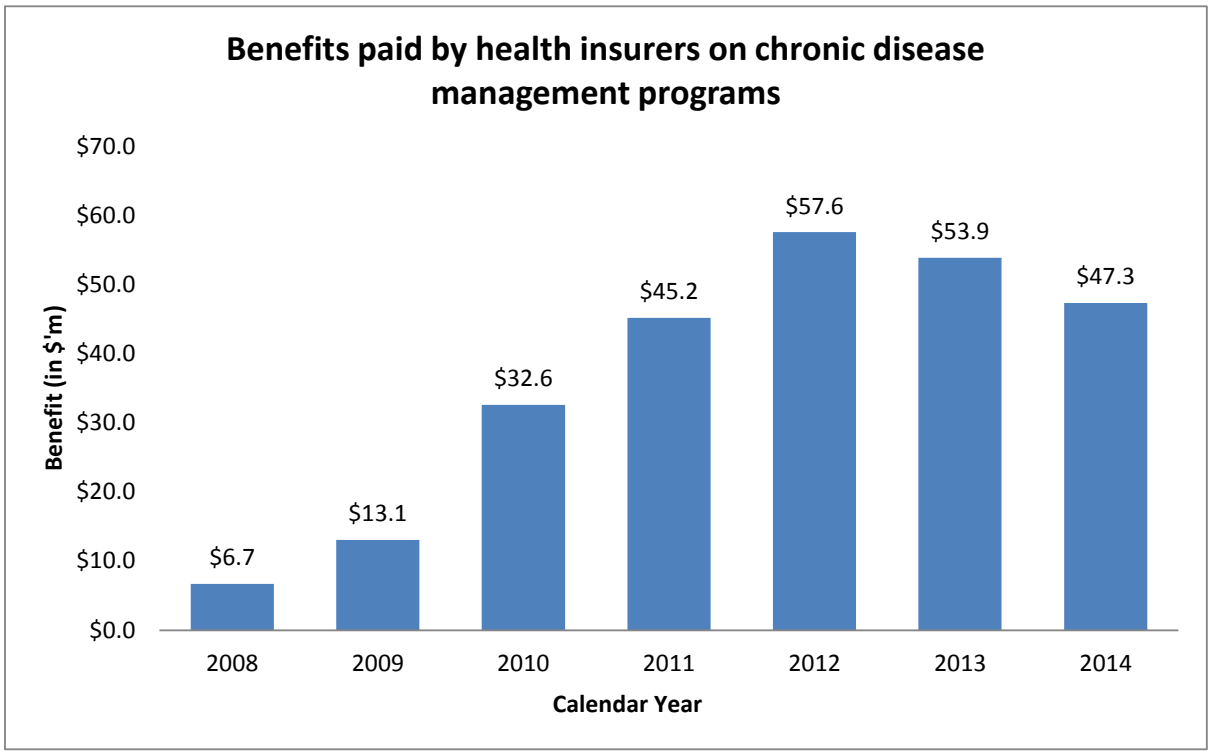
	Organisation	Program name	Program type
Hospital	Austin Health	Cardiac COACH	A telephonic program that trains patients to actively pursue coronary heart risk factors
Medicare Local	ACT Medicare Local	Heartlink	Comprehensive approach to the prevention of cardiovascular disease through systematic absolute risk identification and co-ordinated management of patients
State Government	VIC and WA state governments	Carepoint	Integrated model of care for those with chronic conditions and complex needs across service settings
	NSW, QLD, ACT, SA and TAS governments	Get healthy	Telephone service aimed at supporting adults make lifestyle changes regarding healthy eating, physical activity, and how to reach and maintain a healthy weight
Health insurers	HCF	My Health Guardian	My Health Guardian provides online and telephone-based support and guidance to help members live a healthier life and manage their chronic conditions like arthritis and diabetes.

	Organisation	Program name	Program type
Health insurers	BUPA	BUPA Health Dialogue	BUPA offers telephonic health support for members with chronic conditions – heart failure, heart disease, diabetes, back pain and lung condition. The health coaching aims to increase health literacy and management practices regarding these chronic conditions
	HBF	HBF Coach	HBF Coach is a telephone-delivered coaching program that supports members' rehabilitation by helping them understand and achieve cardiovascular risk factor targets.

The health insurer's perspective

Although private health insurers have long been involved in heath management, increased focus occurred after the enactment of the Private Health Insurance Act 2007. The concept of "Chronic Disease Management Program" was introduced as part of the legislation. Its definition as per the legislation is reproduced in Appendix 1.

The graph below shows the reported benefits paid by health insurers on chronic disease management program since 2008. It should be noted that some programs, which do not meet the regulatory definitions are paid for by health insurers as a management or marketing expense, and so do not appear in the figures below.



A sample of some of the programs funded by health insurers in Australia with published findings can be found in Appendix 2. The table provides an overview of the evaluation technique and measurement system of these programs.

There are additional considerations when a private health insurer evaluates the success and return on investment.

Risk equalisation arrangements
Private health insurance in Australia is community rated, which means in essence that everyone is entitled to buy the same product, at the same price (except for the impact of government initiatives such as Lifetime Health Cover loadings and the rebate which varies with income[xviii]), and is guaranteed the right to renew their policy.

A health fund cannot refuse to insure on the basis of a person's health or how likely you are to claim. The purpose of Risk Equalisation is to support the community rating principle. Risk Equalisation partially compensates insurers with a riskier demographic profile by re-distributing money from those insurers paying less than average hospital benefits to those paying higher than average hospital benefits.[xix]

Benefit savings should be assessed before and after risk equalisation as any reduction in benefits is shared amongst funds through the risk equalisation scheme.

Lack of access to primary care, outpatient and pharmaceutical data
Private health insurers hold a limited amount of information regarding the health of an individual. It only has visibility to the data collected as part of the program, as well as internal claims data. Overseas evidence suggests primary care and pharmaceutical usage data are very useful in the targeting process, and can also provide a more comprehensive perspective in the evaluation. However, health funds in Australia do not have access to such information. Even if access to this data is granted, it is only provided in a de-identified format, which would make it challenging for an insurer to undertake the extensive data matching with the fund's internal data before targeting or evaluation can be performed. As rapid targeting often leads to more successful programs, this process, if feasible, may not produce results in a sufficiently timely manner.

Comparing programs from an underwriting perspective
Recently in Australia, a number of private health insurers have been looking at establishing a set of standard measures for use in comparing different disease management and wellness initiatives. Evaluating all initiatives with respect to these measures will, in theory, allow the comparative value of the programs to be assessed.

Whilst there has been no formal industry consensus, a number of insurers, including the Australian Health Service Alliance[xx], have adopted the following base measures[xxi].

- Bed days per 1,000 lives
- Admission rate
- Loss ratio

It should be noted that for each of these methods, the actual "savings" achieved will depend upon the selected comparative group. Typical groups considered include:

- The same group prior to intervention
    - Absolute levels
    - Assessment of trends (noting that the period selected for determining the trend can lead to significantly different forecasts, as much of the hospital and medical data is volatile)
- The group offered the intervention who declined to participate
    - It is important to note that these people may have a naturally higher or lower outcome, which influenced their decision not to join the intervention
- A comparative group based on similar demographics to the intervention group
    - Factors considered typically include: age, sex, location, recent hospitalisations
- Those on the same product
- Insurer membership
- Industry membership
- Population statistics

There may be differing time or data constraints for each initiative, potentially leading to different approaches being adopted by the same insurer in evaluating competing programs.

In addition, it is important to ensure that all costs associated with the program are captured, including:

- Actual fees paid in respect of the program
- Back office expenses – including marketing, finance etc
- Cost of care coordination staff (if internal)
- Other costs incurred in respect of the program

Bed days per 1,000 lives (BDTL)

This is a measure of the number of days hospitalised within a given group. This should include both overnight and same day stays.

$$BDTL = \frac{\text{Total bed-days for relevant cohort over specified period}}{(Av\ Lives\ /\ 1000)}$$

Contracting arrangements can influence this factor, as some contracts offer incentives for lower bed-days, and care should be taken when interpreting the results.

A reduction in the BDTL would generally indicate an improvement in the health of the cohort, and therefore an expected reduction in costs.

The challenge in implementing this method for evaluation purposes lies in determining what the comparative BDTL would have been if there had been no intervention. Different approaches continue to be adopted across the industry, often driven by the availability of data or comparison groups, making it challenging to evaluate the relative results of different programs.

Average Lives over the period (Av Lives)

$$\text{Av Lives} = \frac{\text{lives @start + 2* (lives @ start of each intervening month) + lives @ end}}{\text{(2 * number of months in the period)}}$$

If only quarterly data is available, then use:

$$\text{Av Lives} = \frac{\text{lives @ start + 2* (lives @ start of each quarter) + lives @ end}}{\text{(2 * number of quarters in the period)}}$$

Lives at any given point in time is the number of lives counted in the relevant cohort at that time.

This approach provides for an average membership of a given cohort, allowing for the pace of growth over the period. For example, if an insurer were to grow rapidly in the first month, the average calculated using the Av Lives formula would be higher than simply taking the simple average of the start and end membership.

Admission Rate

The admission rate is a measure of the incidence of hospitalisation amongst a given cohort. This includes both overnight and same day admissions.

$$\text{Admission rate} = \frac{\text{Number of admissions for specified cohort over given time period}}{\text{(Av Lives / 1000)}}$$

Admission rates can be distorted by the treatment of readmissions as some are classified as a continuation of a previous claim, and some a new admission. This treatment may depend upon particular contracting arrangements, and can vary between insurers in respect of the same hospital group. This may be relevant for benchmarking should this evaluation method be compared between insurers.

A reduction in admission rates is generally positive, as it tends to indicate healthier members.

Loss Ratio

The loss ratio represents the ratio of claims incurred to premiums received for each cohort. It is a measure of the profitability of the group, before allowing for risk equalisation (RE) or expenses.

Prior to RE is recommended as the initial focus, as the operation of the RE pool can mask some of the program effects. However, it is also appropriate for insurers to consider the loss ratios inclusive of risk equalisation, as this can significantly alter the financial outcome of the initiative.

$$\text{Loss ratio} = \frac{\text{Claims incurred (charges) for a cohort over specific observation period.}}{\text{Premiums earned in respect of the same cohort over the same period}}$$

Charges are utilised, instead of benefits paid, as these are not impacted by any excesses or co-payments made by the policyholder. They also provide a stable reflection of incurred costs, removing many of the impacts of product design.

However, it is noted that many insurers are less stringent in their reporting of charges than they are of benefits, and care should be taken in reviewing the results obtained for reasonableness.

Generally claims are focussed on hospital/medical claims, however some insurers argue that general treatment costs should also be considered, as these can also reflect overall "wellness". When comparing the evaluation results of different programs, it is essential to understand what is included and/or excluded from the assessment.

It is important that this ratio is based on incurred claims (claims relating to a specific period, regardless of when they are paid), and earned premiums (premiums relating to a specific period, regardless of when they are actually received), as this is comparing like with like. Otherwise, variations in payment patterns could lead to a distortion of results.

This means, that claims will not be reliable until approximately three months after the end of the specified period (as most claims are paid within three months of occurrence).

Furthermore, premiums, like other commercial prices, are impacted by a range of factors including: profitability requirements, a contribution to management expenses, availability of capital and actual and/or expected claims experience.

As this assessment technique involves a comparison to premiums, it has been argued that the results cannot be effectively compared between different insurers, or potentially even those on different products within an insurer.

Alternate evaluation techniques
In addition, some funders utilise a range of alternate techniques.

Direct comparison
Using a direct comparison with a relevant peer group, the funder simply compares the experience of the control group with that who have undertaken the program. If the cost of those who have been part of the program is lower, the difference is taken to represent the program savings.

Return on investment
This can be utilised with many of the other evaluation techniques. The measurement is converted to a dollar value saving, which is expressed as a percentage of the total cost of offering the program.

Other Factors
In addition, a number of self-assessed wellness and improvement factors are frequently used. These tend to vary by insurer and program.

Secondary impacts, such as customer satisfaction, retention rates or marketing impact tend to be measured more sporadically in Australian private health insurers at this stage. We are aware of with little comparison of these impacts across programs.

5. So how do I start?

Initiating a new program evaluation can be daunting. The following checklist has been developed to ensure that critical elements of program design are captured in a timely manner, to assist with effective evaluation.

- Program objectives
    - Ensure these are clearly articulated, and that there is a common understanding of the language used
    - Test whether there are secondary, or tertiary, objectives
        - Marketing
        - Retention
        - Customer satisfaction
    - Consider the most effective ways to measure success

- Agree high level evaluation techniques
    - Clinical (if appropriate)
    - Financial
        - Return on investment (need to agree how savings and costs will be determined)
        - Loss ratio
        - Admission Rates
        - Bed Days per 1,000 lives
        - Absolute cost of program (pre or post risk equalisation)
        - Cost of capital
    - Other

- Specify timeframes for evaluation
    - Pre-program evaluation (the period selected can significantly influence the reported results)
    - Intervention period
    - Post-intervention period

- Determine whether participants will be classified into cohorts (the initial members often have a different experience to those who join later)
    - How will cohorts be determined

- Agree the comparative group – noting again that the reported results will vary depending upon the group selected
    - The control group if a RCT is performed
    - This may be the same cohort pre-intervention (with or without allowance for trends)
    - Alternatively it may be a different group (eg those who declined to participate)
    - Consider also comparing to a broader baseline of product or insurer membership

- Ensure required data will be collected and available in a usable format
    - How will the data be collected?
    - Is there anything that needs to be recorded differently

- Consider adjustments
    - Will allowance be made for reversion to the mean effects?
    - Will the comparative group be adjusted to reflect differences in characteristics with intervention cohort?

- Ensure the evaluation will assist the business in making strategic decisions
    - Would the results of this evaluation enable you to choose appropriately between competing options? Would the results help the business to improve existing programs and other practices?

6. Conclusion

Disease management programs are widely accepted as a way to improve self-management practices amongst chronic disease patients, and slow the rate of disease progression while maximising health and wellbeing. Although disease management in Australia is still quite immature compared to the US and Europe, it is gaining momentum. As the number of programs increases, the need for program evaluation is going to grow.

Actuaries have a role to play as the actuarial skillset is closely aligned with the requirements of a health program evaluator. The authors hope this paper will provide some useful insights for those who undertake program evaluation. This paper recognises that every program is different in nature and evaluators should adapt their methodologies to suit the circumstances. There are many considerations that evaluators need to take into account when performing an evaluation.

While there is no recognised "gold standard" in program evaluation, the exercise is still worthwhile for both the funder as well as the wider community. As the knowledge of disease management increases through quality evaluations, better programs can be designed, and most importantly, we can help improve future patient outcomes.

Appendix 1

Definition of Chronic Disease Management Program in the Australian private health insurance

The definition of chronic disease management program is defined in section 12 of the Private Health Insurance (Health Insurance Business) Rules 2013, which is as follows:

1) A **chronic disease management program** means a program that is intended to:

a) either:
   i) reduce complications in a person with a diagnosed chronic disease; or
   ii) prevent or delay the onset of chronic disease for a person with identified multiple risk factors for chronic disease; and

b) requires the development of a written plan that:
   i) specifies the allied health service or services, and any other goods and services to be provided; and
   ii) specifies the frequency and duration of the provision of those goods and services; and
   iii) specifies the date for review of the plan; and
   iv) has been provided to the patient for consent, and consent is given to the program, before any services under the program are provided; and

c) is coordinated by a person who has accepted responsibility for:
   i) ensuring the services are provided according to the plan; and
   ii) monitoring the patient's compliance with the agreed goals and activities specified in the plan.

2) In this rule:

allied health service means a health service provided by any of the following allied health professionals who were eligible, at the time the service was provided, to claim a Medicare rebate for a service of that type:
   a) an Aboriginal health worker;
   b) audiologist
   c) chiropodist
   d) chiropractor
   e) diabetes educator
   f) dietician
   g) exercise physiologist
   h) mental health worker
   i) occupational therapist
   j) osteopath
   k) physiotherapist
   l) psychologist
   m) speech pathologist

chronic disease is a disease that has been, or is likely to be, present for at least 6 months, including, but not limited to, asthma, cancer, cardiovascular illness, diabetes mellitus, a mental health condition, arthritis and a musculoskeletal condition.

risk factors for chronic disease include, but are not limited to:

a) lifestyle risk factors, including, but not limited to, smoking, physical inactivity, poor nutrition or alcohol misuse; and
b) biomedical risk factors, such as high cholesterol, high blood pressure, impaired glucose metabolism or excess weight; and
c) family history of a chronic disease.

Appendix 2
Examples of health management program funded by private health insurers with published findings

Funder	Program	Disease targeted	Evaluation methodology	Time period post intervention for PHI measures	PHI measures	Pathological measures	Other biometric measures	Other measures
Remedy clients	Bone Health	Osteoarthritis	Compare results of treatment and non-treatment groups	Continuous tracking post-graduation	Hospital benefit per member	Vitamin D level	Vitamin D intake Calcium intake Medication adherence Level of weight bearing exercise Smoking status Alcohol intake	Patient Satisfaction
Remedy clients	Diabetes	Diabetes	Compare results of treatment and non-treatment groups	Continuous tracking post-graduation	Claims per member	HBA1c LDL Total Cholesterol Triglycerides Blood pressure	BMI Level of Physical Activity	Patient Satisfaction
HCF	My Health Guardian	Heart Disease or Diabetes	Pre and Post intervention comparison	12 and 18 months	Hospital Admissions Hospital readmissions Average Length of stay			

Funder	Program	Disease targeted	Evaluation methodology	Time period post intervention for PHI measures	PHI measures	Pathological measures	Other biometric measures	Other measures
Medibank	better health on call	CHF COPD CAD Diabetes	RCT		Numerous including use of allied health services, hospital utilisation, ALOS and Claims per member	Disease specific measures e.g. for CAD Self-reported cholesterol Self-reported Hypertension	Disease specific measures e.g. for CAD Weight Have a written action plan Influenza vaccination Smoking status Level of physical activity Low salt diet	Patient satisfaction with program & Medibank in general
nib	COACH	Cardiac	Compare results of treatment and non-treatment groups	12 and 24 months	Drawing rate before risk equalisation			

Bibliography

AHSAEvaluation Methodology for HSS Broader Health Cover Programs, August 2012, EQHealth

Allen, K, et al (2007) Evaluation of the Medibank betterhealth On Call Program for Chronic Disease Management: Pilot cohort report. Monash University, Melbourne.

Bachler R et al, Evaluating the Results of Care Management Interventions: Comparative Analysis of Different Outcomes Measures – various papers https://www.soa.org/research/research-projects/health/research-evaluating-the-results-of-care-management-interventions-comparative-analysis-of-different-outcomes-measures-claims.aspx#sthash.xqdSoY0l.dpbs

Bull D. The COACH Program® results in reduced hospital admission costs with 2 years: Experience at nib health fund

Brent Hamar G et al. Impact of a Chronic Disease Management Program on Hospital Admissions and Readmissions in an Australian
Population with Heart Disease or Diabetes, Population Health Management 2012, 16: 125-131

Cullen M. MCKESSON Results of Three Randomised Controlled Trials of Telephone Self-management

Elissen AMJ et al, Advancing current approaches to disease management evaluation: capitalising on heterogeneity to understand what works and for whom, BMC Medical Research Methodology 2013 13:40

Grindlay G. Review of a Telephone Based Self-Management Program for Type 2 Diabetes, presented at the 2012 ADMA conference

Grindlay G. Preliminary results of the Remedy Healthcare Bone Health Program, presented at the 2011 ADMA conference

Linden A, What will it take for disease management to demonstrate a return on investment? New perspectives on an old theme, American Journal of Managed Care 12 no 4 2006

Linden A, Disease Management: Truth in Advertising

Linden A, Assessing regression to the mean effects in health care initiatives, BMC Medical Research Methodology 2013 13:119

Mattke S et al. Measuring and Reporting the Performance of Disease Management Programs. National Evidence Based Guidelines for the Prevention and Management of Type 2 Diabetes- Public Consultation Draft August 2008 Diabetes Australia http://www.diabetesaustralia.com.au

Nolte E and Hinrichs S, DISMEVAL: Developing and validating disease management evaluation methods for European healthcare systems Final Report, Technical Report 2012

Nolte E et al, DISMEVAL: Evaluating chronic disease management, recommendations for funders and users, Technical Report 2012

Optum, Moneyball Analytics, Connecting and leveraging the best data across the health care continuum, 2014

School of Public Health and Preventive Medicine Centre for Health Economics Monash University. Evaluation of the Medibank betterhealth On Call Program for Chronic Disease Management. June 2010

Vale M et al. Leading the world in coaching for prevention of chronic disease

Vrijhoef B, DISMEVAL: A study on disease management in Europe, Integrated chronic disease management forum, Department of Health, 2011
Wilson T et al, Evaluation in disease management studies, 2004-2007, Disease Management & Health Outcomes 2008 vol 16 No 5 pp 365-373

Wilson T & Linden A, Disease Management's Economic Impact: Unproven?

Wootton R. Telephone-supported care coordination in an Australian veterans population: a randomized controlled trial. IJ Telemed Telecare March 2010 16:57—62

Young D et al. Patient Engagement and Coaching for Health: The PEACH study – a cluster Randomised controlled trial using the telephone to coach people with type 2 diabetes to Engage with their GPs to improve diabetes care: a study protocol. Published: 11 April 2007 Received: 12 March 2007, BMC Family Practice 2007, 8:20

Useful websites

http://www.dismeval.eu/

http://www.lindenconsulting.org/

https://www.soa.org/

http://www.populationhealthalliance.org/

[i] Oral health may not typically be considered a chronic disease or condition; however the AIHW ranks it with other leading diseases due to the costs attributed to this area. Refer for example to http://www.aihw.gov.au/media-release-detail/?id=60129546452

[ii] Australia's Health 2014, AIHW

[iii] The European Observatory on Health Systems and Policies. Assessing chronic disease management in European health systems: Concepts and approaches. Page 10

[iv] Rothman AA, Wagner EH (2003). Chronic illness management: what is the role of primary care? Annals of Internal Medicine, 138(3):256–261.

[v] Population Health Alliance (2014). PHM glossary: D. Disease management. (http://www.populationhealthalliance.org/research/phm-glossary/d.html,accessed 4 March 2015).

[vi] Some frequently used assessment tools include: SF-12, HeiQ, EQ-5D, AQOL and HLQ.

[vii] **HbA1c** (hemoglobin A1c) is a blood test measure used to indicate diabetes

[viii] A short health evaluation survey: http://www.sf-36.org/tools/sf12.shtml

[ix] http://www.health.vic.gov.au/news/carepoint.htm

[x] The interested reader is referred to papers such as: **Linden A**. Assessing regression to the mean effects in health care initiatives. *BMC Medical Research Methodology* 2013;(119):1-7.

[xi] http://www.lindenconsulting.org/

[xii] DISMEVAL Technical Report 2012, page 63

[xiii] DISMEVAL Technical Report 2012, page 57

[xiv] The Population Health Alliance was previously the Care Continuum Alliance and before that the Disease Management Association of America http://www.populationhealthalliance.org/newsroom/care-continuum-alliance-changes-name-to-embrace-population-health-industry-expansion.html

[xv] http://www.rand.org/topics/health-and-health-care.html

[xvi] An introduction to care management interventions and their implication for actuaries, Ian Duncan, p 1 https://www.soa.org/Files/Research/Projects/Paper-9.pdf

[xvii] DISMEVAL, Technical report, page xiv

[xviii] A full listing of government incentives and surcharges can be found at the following link http://www.privatehealth.gov.au/healthinsurance/incentivessurcharges/

[xixix] Further details on Risk Equalisation: http://phiac.gov.au/wp-content/uploads/2014/12/Factsheet-Risk-Equalisation-ATT-9.pdf

[xx] The Australian Health Service Alliance (AHSA) is responsible for facilitating arrangements between hospitals, doctors and health service providers on behalf of its participating funds

[xxi] AHSAEvaluation Methodology for HSS Broader Health Cover Programs, August 2012, EQHealth