Institute of Actuaries of Australia

Never Entirely Outsource Your Own Brain

Prepared by Len Boehm, with postscript by Julie Evans

Presented to the Institute of Actuaries of Australia
Accident Compensation Seminar
20 - 22 November 2011
Brisbane

The Institute of Actuaries of Australia
Level 7 Challis House 4 Martin Place
Sydney NSW Australia 2000
Telephone: +61 (0) 2 9233 3466 Facsimile: +61 (0) 2 9233 3446
Email: actuaries@actuaries.asn.au Website: www.actuaries.asn.au

Abstract

Why we work

Many of you will have heard me present a version of what follows at the Agent Awards, during induction sessions or simply in day to day conversations.

Even when the subject matter is largely analytical, actuarial, governance or agent management it's important that we recall why we really are all here.

For me a quick summary of my family history says it all.

Many years ago I had a break down at work as a young claims officer struggling with workloads. I lodged a stress claim and was off work for 6 weeks. I needed the help of the Workers Compensation system to recover.

Some years later my first wife was seriously injured in a motor vehicle accident and we needed the help of the (then) MAB no-fault scheme (pre TAC).

Several years later my current wife was injured when a bus wrote off her car on her journey home after picking up the kids.

And only a few years ago, my grandson's father was electrocuted at work and nearly died.

We work so that our family, our colleagues, our friends and their children are protected by a more caring, more effective safety and compensation scheme and to ensure that level of support is available for future generations.

Len Boehm

Keywords: accident compensation; scheme management; claims management

1. Introduction

For much of the last 30 years the Victorian Workers Compensation scheme has repeated a cycle characterised by blow outs in costs leading to premium increases, then major legislative change to bring things under control followed by brief periods of stability and then the cycle would all start again. This pattern of bust, legislative reform, stability and then bust again has played out time and time again in workers compensation systems across the world.

I look at 2000 as something of a watershed year. We had legislative change, with the right of injured workers to pursue common law restored, on top of an already volatile scheme and premiums were increased to 2.2% of payroll. Since that time, we've managed down the cost of the scheme - albeit with several false starts - without any legislative change (except for increases in benefits). Hundreds of millions of dollars have been removed from the cost structure of the scheme, premiums have fallen 40% and even in areas where we have not been a good as we hoped (such as worker service) a substantial improvement has at least occurred. We now have the lowest break even premium of any Australian State.

Nevertheless, there are always emerging pressures on the scheme and there are always things we could do better. We may be twice as good as we used to be but we remain only half as good as we aspire to be. In that mix of successes, opportunities and mistakes there are learnings we think are worth sharing.

This paper does not claim to be a lesson in how to run a compensation scheme, nor does it purport to represent a single unifying theory of scheme or claims management. Instead it represents an attempt to capture a mix of scheme learnings with a focus on analytical and actuarial concepts.

However, there are common threads in the examples we have chosen, in particular the value of deeper integration of actuarial style thinking into day to day claims management and the importance of all layers of management in a claims system becoming more engaged in understanding, embracing and challenging the outcome performance of their schemes. The title of this paper “Never entirely outsource your own brain” is intended to reflect these themes.

Enjoy and reach your own conclusions.

2. First some context on how the Victorian scheme works

2.1 Benefit design

The benefit design of Victoria’s workers compensation scheme can be categorised as a “longer tail” system with a higher proportion of workers receiving weekly benefits for their working life and with largely unlimited ongoing entitlements to treatment. In broad design it is closer to the South Australian and New South Wales schemes in statutory benefit coverage.

Schemes such as Western Australia and Queensland are more common law buy out systems with a much shorter statutory component and therefore a much shorter tail overall. Common law costs dominate such systems.

Victoria of course also has a substantial common law add on (for the seriously injured) but that accounts for only 20% of scheme costs, so we are still largely a no-fault statutory benefit system.

For the technically minded the mix of our outstanding claims liabilities is shown in the table below. The discounted mean term of claims is 5.4 years.

Table 1: Outstanding Claims Liabilities at 30 June 2011

Benefit Type	Estimated Liability at 30 Jun 11 $m
Weekly	2,605.2
Rehabilitation	57.2
Doctor	351.9
Hospital	296.1
Paramedical	393.7
Hearing aids	145.0
Personal & household services	126.1
Long term care	285.2
Common law	1,587.3
Common law legal	647.3
Impairment benefit	493.4
IB Medical reports	99.1
Medical reports (other)	93.9
Statutory legal	165.7
Death benefits	92.1
Investigation	25.8
Other benefits	70.7
Recoveries	-221.5
Scheme net central estimate	**7,314.3**
Claims handling expense	865.7
Scheme total net central estimate	**8,179.9**
Scheme risk margin	695.3
Scheme provision	**8,875.2**

The scheme design impacts on how schemes need to be managed and measured.

2.2 The agent model: maximising both the value of central initiatives and of agent competition

Victoria operates with day to day claims management outsourced to five external agents. This is an unusual model in world terms: liabilities are retained by the central scheme, but day-to-day management of the claims is outsourced to third parties. Only Australia appears to have tried it and even then it only applies in three states; Victoria, NSW and SA. The rest of the western world tends to go for "monopoly" or "private underwriting". That makes our scheme's experiences, challenges and opportunities somewhat unique.

Within the framework of an agency model there is a continuum of possible approaches from "set and forget" to "franchise" type approaches. A decade ago Victoria operated a largely a hands-off model setting high level incentives and then leaving it largely to agents to manage. Although the Victorian scheme relies heavily on agent autonomy and competition to drive outcomes, over time it has also become a more hands-on model (on selected issues).

In its current configuration the Victorian scheme really operates as an attempt to maximise both the value of central scheme wide strategies and the value of agent competition.

Sometimes in order to control an intractable problem it can be more effective if all agents kick down the door at the same time. A central approach can be particularly valuable where we are trying to move a trend urgently or where we are trying to influence an entire market's behaviour. For example: clinical influence approaches or strategies based on setting consistent precedents require scale and consistent messaging.

However, sometimes letting competitive forces and agent innovation drive our response is more effective. This is particularly true where an opportunity is specific to an individual agent's portfolio or where there are no obvious central strategies that could work.

The agent incentive system is key to running an effective competitive agent market. We have a fairly sophisticated approach which includes long term profit share and targeted annual performance components. The model is also very flexible, enabling annual adjustment of incentives to align with changing scheme risks and priorities.

However, financial incentives are not quite the whole competitive game. For the agents who are large insurers, the potential profit from the Victorian scheme will always represent only a small percentage of their total organisational profit. These agents are as interested in the cross sale opportunities with larger employers. So how do you raise the priority of our scheme a bit more in their eyes? For us the key lies in more transparent reputational competition.

It's taken several years for us to fully learn this lesson. A few years ago one WorkSafe Board member completed a tour of Australia and had visited the chief executives of several of the parent companies. He said *"Every boss tells me that they are either number one or two in Victoria. Given we have numerous agents that actually isn't possible is it?"* This was a clear message that the real performance story wasn't permeating through to the head office.

And so in 2004 we invented "The league table reports" which list a dozen key performance metrics and rates agents - by name - on each one. We send it to the agent head offices every quarter. This approach has attracted much more head office engagement than previous methods. Even in the most competitive forms of private enterprise the actual performance of yourself vs. your competitors is rarely known in real time. Victorian workers compensation is probably the only competitive system where detailed real time performance data is provided to all competitors about all competitors.

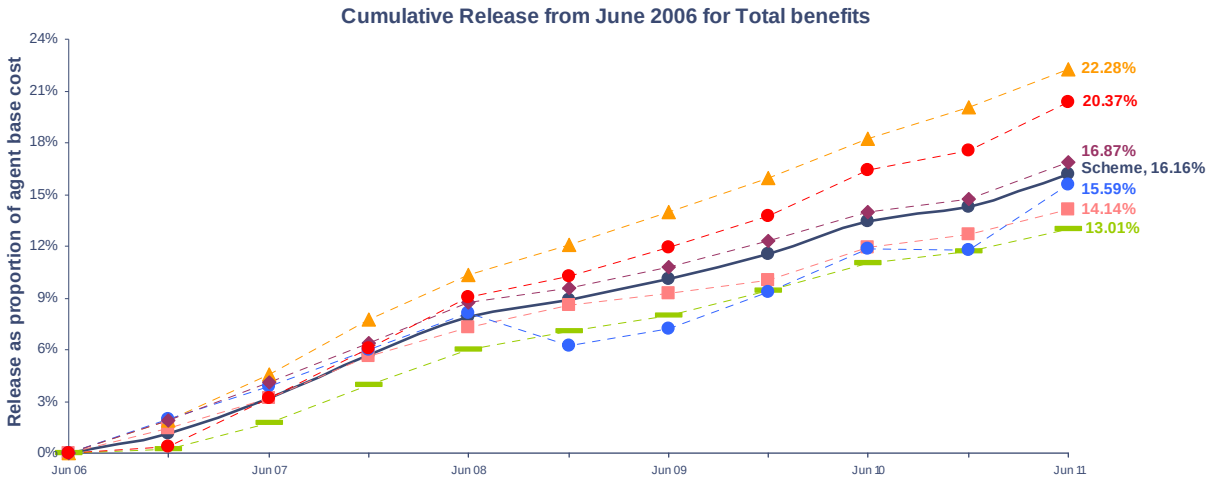
When someone says "it can't be done", don't underestimate the value of being able to show them a rival's results on the same issue with the tag line "well how come they are able to do it?"

Overall our best guess remains that "competition" and "central strategies" have contributed in about equal proportions towards the scheme turnaround of the last decade. Both competition and central strategies are integral to our model.

This is illustrated in the following chart. The chart shows the agent performance over the last contract period, on one of our measures. Broadly, the chart shows how the agents have reduced the scheme cost, as projected by the external actuary at the start of the contract period. Each line represents the actuarial release (see next section) for a different agent (there were six agents for the last contract period), with the scheme line in the middle.

The detail of how the chart is constructed, and what the numbers mean, is not in itself important for our purpose here. The point of including the chart is to show that there is difference in the performance of the different agents (i.e. central strategies are not the sole driver of performance) but also that the relative performance does change over time.

Figure 1: Agent Lump Sum Actuarial Releases



Of course attempting to manage these sometimes contradictory forces comes at a price. It makes it all the more complicated as it requires the "centre" to try to be across much more of agent day-to-day effectiveness as well as the big picture. It also creates a never ending tension between agent autonomy and central strategies and controls. It also requires an extraordinary level of transparency between WorkSafe and our agent partners. Finding the right balance with our agents day-to-day remains a constant work in progress.

It goes without saying that without the goodwill and efforts of the leadership within our agents the model could not be a viable one.

2.3 The concept of an Actuarial Release

This all started at TAC almost 15 years ago.

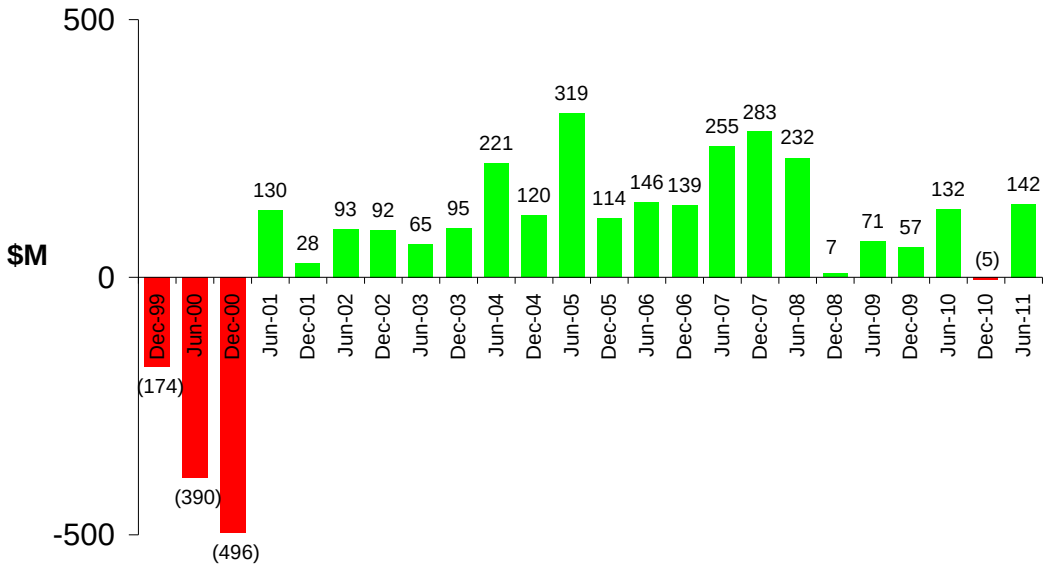
As part of the annual actuarial valuation, the actuaries project ahead one extra accident year (or more). That future prediction of outstanding claims liabilities then becomes the target that management are trying to beat (or at least equal). A year later during their formal valuation, the actuaries strip out the impact of changes in economic assumptions (AWE, bond yield etc) and restate the result based on common economic assumptions. If the resulting number is less than the projection a year before, then the scheme has achieved an "actuarial release".

Of course the actuaries methods are pretty broad and the sort of annual movements we monitor can fall within a reasonable "error margin", nevertheless taken over a number of years using this approach to measure scheme performance holds up pretty well.

The importance of the measure lies less in the maths than in how it can be used as a vehicle for cultural change. In TAC in the late 1990's, in WorkSafe from 2002, and more recently in the NZ ACC system, the approach has been used to create much greater management accountability and focus.

The chart below summarises the six monthly actuarial releases achieved by the scheme and our agent partners over the last decade.

Figure 2: Scheme Actuarial Release



It is not my intention to debate methodology. The concept and experience is included here merely to provide context to other topics in the paper.

3. Some analytical concepts, principles and lessons

3.1 Modularity

I use the term "modularity" to describe dissecting liability or service trends down to potential lever points and then managing overall scheme trends by application of targeted strategies.

Somehow, almost by accident, we began to dabble in a more modular approach to scheme management, initially driven by the need to control the latest hot spot. Over the years it has turned into a semi-systematic approach. The approach can be likened to fixing the scheme "one targeted piece at a time". It concentrates limited skills better, enables deeper strategies and - where you get the lever point right - moves scheme trends at a faster rate. When it works it also enables us to align all agents simultaneously to target the same leverage point.

In theory as the scheme matures and learns, many strategies embed into day-to-day claims management. Thus over time forming a more holistic approach one piece at a time.

The above is in no way intended to understate the importance of "end to end" claims strategies. In fact under our latest agency agreement we are working with agents to encourage greater structural and capability reform to build improved day-to-day case management and service delivery. It's just that the nature of the agency model (with culture, sub structures design, recruitment, day to day process and leadership largely delegated to agents) makes it very difficult for an outsourcer/regulator to play at that level.

Of course such a modular approach has a potential downside. Wherever we have central strategy we normally have deep real time audit, file review, agent engagement and performance monitoring. However, where we don't, the visibility over what's happening day-to-day can be weaker. We can't be everywhere and arguably deep knowledge of strategic slices of the scheme is more effective, but it means there is always a risk of gaps in our vision.

3.2 Riding the cycle

We operate under a "cycle of forced evolution" of which the six monthly actuarial valuations, the annual review of agent incentives, and the annual client satisfaction survey form the working pieces.

The six monthly actuarial valuation acts as the formal umpire's decision on how the scheme is trending in terms of claims costs and outcomes. The Executive Director must prepare an immediate and detailed management response to emerging trends at the Board meeting at which the valuation results are presented. This forces a quick and honest rethink of the state of existing strategies and a decision as to whether to adjust, abandon or add new initiatives.

The bulk of agent incentives are paid via the annual performance measures (APA) which has to be set by 30 June for the next financial year. We conduct a 4 month review from January to April of every year involving internal input and detailed debate and consultation with agents to set the new or revised measures, and weightings (the allocation of the "money on the table" between the measures). Every part of the business wants more money to create incentives focusing on their priority issues, and there isn't enough money on the table to satisfy everyone. In effect this process forces the necessary internal debate to prioritise between the numerous conflicting priorities of the scheme.

Audit processes similarly support the cycle. Agents are obliged under contract to complete an annual review of their claims and process controls. WorkSafe annually sets priority areas for such reviews based on scheme trends and complaints during the year.

Detailed monthly performance reporting and discussions with our agents further reinforces this cycle. The scale of this monitoring and the internal debate that surrounds it arguably is one of the key drivers that has created a more effective outcome culture. Every month key trends for each business theme (RTW, Treatment, Eligibility & Long Tail, Common Law and Service) are debated between internal actuaries, data analysts and business managers and agents. The inputs include not only tables and charts, but also the soft feedback from complaints, audits, file reviews and even the vibe from each agent's floor.

I admit this monthly process is somewhat of a pain. It eats a lot of time and arguably most trends don't move in a month, but it keeps us on our toes ensuring we are reasonably on the ball at any time. It also contributes to creating a culture and mindset that is more ready for change.

Pull it all together and it means that we are in a constant cycle of rethinking and informal planning whilst running the business day to day. As one of my Directors explained to me last year *"annual business planning isn't that hard when you have been refreshing where you are going throughout the year"*.

In workers compensation the speed at which trends can deteriorate, precedents change and market forces shift, makes for a volatile ride. Add to that the impacts of changing economy and employment and sometimes it feels more like a bucking bronco.

Recognising this characteristic Chris Latham, an actuarial advisor to our scheme for many years, coined a phrase about the characteristic that he thought was most important in successful scheme management: *"you need to be fleet of foot"*.

The cycle above has been crucial in our journey to create such a flexible culture.

3.3 You can't manage what you don't measure

Professor Bob Officer (former Board member) used to repeat the above mantra at every Board meeting.

A few years ago I recall puzzling over the question of why market forces hadn't really driven the evolution of a much stronger return to work (RTW) competency in insurers, regulators and schemes.

It's been over a century since all this started and economic theory says some scheme or insurer should be incredibly good at this by now.

Yet until recently no one really measured RTW. The RTW performance of insurers is not published (because it isn't measured) and so is never known to employers and so can never really be the basis of insurer selection. The choice of agents or insurer is instead largely based on cost or other intangibles. Most schemes purport to have a RTW measure but nearly always it's actually a continuance rate measure (time off benefits) which is as influenced by terminations or settlements.

Can best practice in RTW really evolve without any reliable feedback as to what did and didn't work (i.e.: real RTW outcomes)?

My comments are a little extreme but they illustrate the point. In day-to-day management whether it is service, RTW, audit compliance or liabilities the use of data to understand outcomes is vital to effective management.

3.4 The struggle to measure RTW

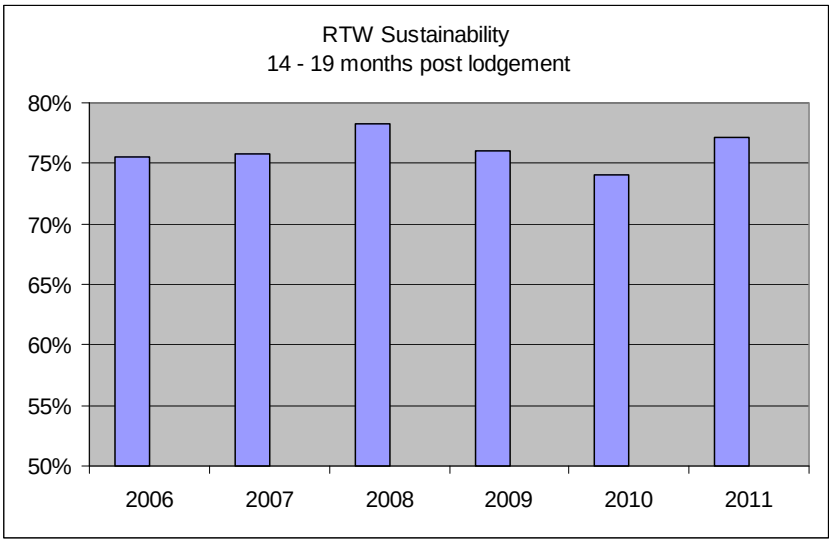
As noted above, historically most Australian workers compensation schemes didn't really measure RTW: everyone measured continuance rates (durations on weekly benefits) and just assumed it also measured RTW. It's often close, particularly for the shorter durations, but because of terminations, dispute settlements, etc it isn't actually the same. The difference may only be 5%-15% but that's more than the annual improvement targets most schemes set.

Like many workers compensation schemes, the scheme in Victoria does not pay workers directly (until they become job detached). We instead reimburse employers - often months after the absence from work. Thus we do not know the RTW status of workers at any given time.

We started trying to more accurately measure RTW about six years ago. Initially we forgot the above limitations and tried to measure the effectiveness of Occupational Rehabilitation (OR) using weekly benefit duration models. After some months of work we realised trends were mainly being driven by termination activity at 130 weeks. (So back to the drawing board.)

Next we decided to see if we could learn from the Campbell's National RTW Monitor. So several years ago we introduced a much bigger scale version surveying 4,000 workers per annum rather than 300**.** The scheme results on this measure are shown in the chart below. The approach seems to work pretty well at the "scheme" level but as an incentive measure for agents it failed.

Figure 3: RTW Sustainability (14-19 months) post lodgement



3.5 Incentive effect = $ times faith

For an agent incentive to work agents have to believe they can improve a measure. That belief is aided by greater line of sight between day-to-day activity and the outcome. Because the RTW survey is confidential (so we can't tell the agents which claims are included) and because of the randomness of non-responses, we couldn't build the direct links at a claims officer case level between interventions and an eventual RTW. Some agents attempted to build internal tracking methods linking what they knew was happening on the file with the survey outcomes, but the results were often contradictory.

We needed an additional approach to better focus agents more directly on RTW. This resulted in the creation of our *"6 months RTW index"*. The index method may be summarised as follows.

- Everything is measured at the 6 months post claim receipt date. Any worker who is in receipt of benefits for that day is assumed to be off work. Any worker who was only on benefits for up to eight weeks straight after the accident is assumed to have a RTW. This reduces the population we have to individually verify significantly.

- The agents are responsible for contacting the employer/worker to verify the status on all other claims, and to provide a RTW date. The RTW date has to be coded on the system and documented on the file. WorkSafe undertakes audits during the year to verify the accuracy level. In addition we periodically cross check the results with OR incentive outcomes.

The measure is not necessarily popular (as since the GFC we have all struggled to improve early RTW outcomes) but it at least links precisely with agent experience on each file. As far as we are aware we are the only workers compensation scheme with sizable bonuses and penalties for insurers tied to real RTW performance.

And one last experience with the measurement of RTW. For two years WorkSafe has paid OR providers on a fee for service basis with additional incentive payments for effective RTW. (The RTW has to be constant for a set period for full payment). Naturally providers make sure they have claimed their entitlements, agents check it and we periodically audit the accuracy of the results. The last audit found an error rate of less than 1%. As a result we now have fairly reliable RTW outcome measurement at the provider, agent and even injury level for rehabilitation. In addition this gives us another cross check on the validity of agent recorded RTW status.

This whole journey has taken 6 years and every step has involved claims managers, data analysts and our in house actuaries.

A continued work in progress……

3.6 Reactive is the new proactive, changing market behaviour claim by claim

There is frequent complaint in compensation schemes that we are all too reactive in what we do and that the future is more about predictive models and intervening before things become a problem. Providing such models are accurate enough to avoid us proactively intervening where we don't need to, then there is no doubt much wisdom in the ideal.

But there is also a fundamental misunderstanding of the nature of actuarial valuations. Often they are simply seen as a measure of current performance or as a reactive appraisal of how the scheme is going. That's not really true.

There is a subtlety to the accounting standard that changes the nature of the claims balance sheet. Simply put, the accounting standard requires the actuaries to book as current expenses the estimated cost of our current claims. That often means projecting the costs 10 or 20 years into the future. What the actuarial valuation actually represents therefore is "how much this thing is going to cost you if these trends continue".

Usually when we get hit with an increase in our actuarial valuation the increase in scheme payments is quite small. If we paid out $5 million more than we expected for pharmacy the actuaries would probably hit us $40 million or more because they figure that adverse trend will continue into the future for current claims. But we really only have paid out $5 million and so if we get our skates on maybe we can actually control that trend over the next few years so that the growth stops. In that case we get most of the $40 million back!

Thus each actuarial valuation is actually an attempt to project the future, but that projection counts and impacts the balance sheet now. In reacting to the valuation we are actually proactively acting to control the future cash flow of the scheme.

Of course the "future orientation" of any actuarial valuation is greater the longer the benefit durations. So this principle is most applicable to very long tail statutory benefit systems such as Victoria, NSW and SA.

Claims initiatives also can be designed with a future orientation. Our whole clinical panel allied health model is an example. At first blush it's purely reactive. We wait until a case has had a large number of treatments and then we have a phone call with the treater using a framework approved by the relevant clinical association. But in reality that program was designed as a form of direct marketing with the aim of changing the future treatment behaviour of the provider market by dialogue with a consistent clinical message claim by claim.

3.7 Impact = size times leverage

There is a tendency for people to respond to a risk or opportunity by meeting it head on. This often leads to headaches.

Following the formula that *"impact equals size times leverage"* means that targeting the biggest cost or biggest risk will only have a significant payback if you have a coherent effective strategy that can actually "lever" the gains. Leverage is a complex combination of our internal skill set, the skill set of agents, systems & controls, resources, legislative support and a practical strategy.

When we started tackling the control of treatment costs in 2002 we didn't start with the biggest treatment category (GPs) or the highest area of growth (Psychology) because we lacked an effective lever and the bandwidth to create one. Instead we focused on physiotherapy which was a medium cost moderate growth area. We knew TAC had trialled treater contact years before and so we had something to learn from. We knew this was a market with clear clinical standards supported by their clinical body. We knew that it was an area of greater evidence based clinical practice and that should lend it self to more reliable testing. We also already had some skilled practitioners contracted so the means were at hand. We also knew no one had opened the files in a decade so that the tail should be full of files where clinical practice could have got a bit weak over time (and so be open to simple clinical influence).

In plain words we picked this target area first because we had the ingredients that would make for much greater leverage in a reasonable timeframe.

3.8 The Lego toolbox lesson: recycle, recombine, evolve

Looking back over the last decade this probably is the most exciting and valuable lesson I have learnt.

What we have collected over time is a number of tested approaches for claims and agent management. Although each was originally created in isolation for specific purposes, today they form an ever growing set of tools we can re-apply in different combinations to support or create new strategies.

In 2008 we faced a sudden and alarming blow out in weekly benefit durations for larger employers. It was as though the employers were too busy worrying about the GFC and restructuring. We looked at

the mix of tools in our kit bag and put to together a new scale intervention program to address the challenge **(WorkSite Assessments)**.

- We pinched the idea of using Occupational Physicians from one of our agents (GBS).
- We took the idea of an audited/incentivized work practice from past programs where we had to embed a new approach in agents.
- We created a small team to oversee agents in real time (as we originally did with long tail weekly).
- We integrated the project with our RTW inspectorate so we could target reluctant employers for follow up if needed.
- We developed a mix of targeting outcome measurement systems. (Using all the programming tricks we have learnt from previous initiatives).

To use the Lego analogy, we frequently take a number of blocks we are familiar with and add some new elements to create a new operational approach.

3.9 Earn while you learn with scale experiments

Pilots are not always the best way to learn something if you need to change trends quickly. Normally they target only a small number of claims with an intervention and thus it is very difficult to ever prove they made an impact without an academic study several years after the event. More importantly because they are subscale they cannot actually assist the current management of scheme trends. Pilots are best used where you are seeking to find new levers but have some years to work it all out.

Our preference when dealing with major adverse claims trends is often to design a scale response even if the intervention is somewhat experimental. We can learn as we go and evolve. Thus the expression "earn while you learn".

To take the WorkSite Assessment example: we didn't run a pilot for a year first (we didn't have time as scheme trends were collapsing). Instead we set out with a target of 1200 interventions (to ensure the program was large enough to potentially move trends) but we adjusted the approach throughout the first year as we learnt from employers, workers, agents, Occupational Physicians and the data what was and wasn't working.

The result was a sharp reduction in weekly benefit durations as our scrutiny became visible to the market (with some help from the economy).

3.10 Strategies are a dime a dozen, its implementation that's a real pain

About 14 years ago when I was at the TAC I remember carefully designing an operational strategy to control home help costs. I recall following the textbook fairly closely. We clearly established the "why" and communicated it widely with all staff. We delegated design of the solution to the best line officers who enthusiastically embraced the challenge and designed what they thought was an eminently practical operational strategy. We launched fun and innovative training that scored highly on the feedback surveys. And then ten months later we were aghast when we were audited and discovered hardly any one had actually implemented the solution.

With the benefit of hindsight we had made numerous classic mistakes.

- The work practice was designed by our best and brightest and so was designed to their standard of competence (which was well beyond that of their colleagues).
- We assumed that because every one knew the why and embraced the approach they would implement it, but forgot that with heavy workloads and day to day distractions this was only one of many competing priorities.
- We assumed the local managers really knew what was going on day to day in their team. (They didn't always).

- We also neglected to build the operational measurement processes to give us feedback on what was really happening month by month.

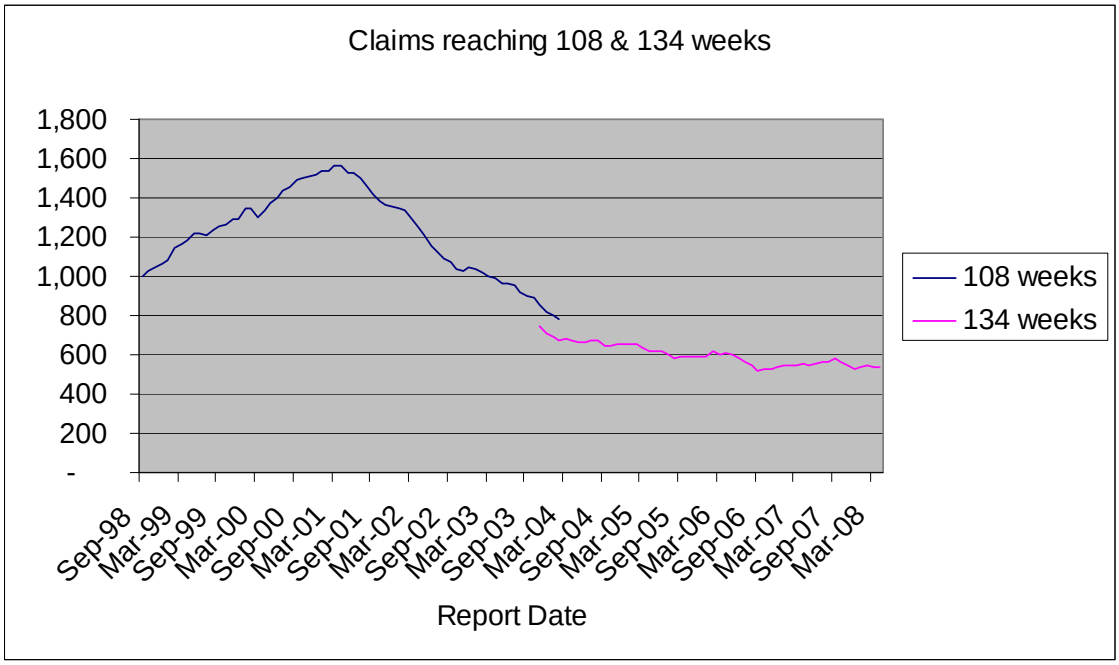
In my career this has been the most consistent and recurring weakness I have seen in all claims systems: the difficulty of converting strategic concepts into viable operational strategies through the layers of management and external providers to the actual coal face.

3.11 The challenge of cause & effect

When a trend suddenly improves the first response of an actuary often is "I don't believe it just yet". I describe this approach as "always look the gift horse in the mouth".

Often cause and effect is pretty easy to see particularly where the scale of the turnaround is sharp and matches closely with the timing of new approach which audits and monitoring show really was implemented. For example, the following graphs show the change in scheme patterns following the introduction of the long tail program.

Figure 4: Success of the Long Tail Program



However, for interventions with a more modest impact measuring whether a program is effective can be quite difficult. The following are some traps for new players.

The most common mistake in analysing scheme trends is the accidental use of undeveloped data. The tell-tail sign in monthly graphs is that the last few points always show improvements. Six months later the last few points (the latest data) still shows improvements but strangely the two months that looked good 6 months ago look worse! The problem is caused by the comparison period having a longer development time than the more recent period. The solution of course is to use "equally developed data". We rarely make this basic mistake in analysis these days.

With programs that target specific claims picked by attributes recorded in your claims database normally you can duplicate an equivalent pre-project population and compare movement trends. Of course this is less effective if there have been any substantial economic or legislative shifts in either period which may impact differently on the populations.

Projects that require considerable manual file review to target are the hardest to measure. Although the initial targeted list is based on basic attributes coded into our system most of the soft factors have to be screened manually. We may need to exclude cases that are about to RTW or where there has been

some recent action not captured by the system. As a result we can struggle to build a reliable control group because we can't duplicate the real characteristics. Adverse selection then becomes a potential risk.

Sometimes we even deliberately create the comparison problem ourselves, such as happened when we created the Worksite Assessment project in 2009. We targeted 1,200 cases per annum in larger employers with an Occupational Physician assessment and workplace visits. We could have set the project up with a control group but because of the size and urgency of the deterioration trend we decided we could not afford to leave hundreds of cases out. In that case we had to rely on monitoring the global population (continuance rates for all large population). The hope was that we were hitting enough files to move the wider population.

Another trap we still occasionally fall for is the challenge of the natural drop off rate.

The problem is most pronounced in early RTW or continuance rate measurement of programs, but it can equally be a problem when measuring any trends over time for a fixed cohort of claims. The vast majority of cases will RTW to work without any special intervention and the rate of RTW will be very high soon after accident then tapering off over time. If you are not very careful you can mistake the natural drop of rate for real improvements. For a project to be effective it has to achieve outcomes that are better than those that would occur with the normal drop off rate. And what passes for "normal", of course, varies by cohort. It's amazing how often this one gets missed.

Frankly the precision of measurement will not always be a goal we can achieve. What really matters though is creating a culture where managers learn to routinely ask the question "is this project actually making any difference" and are open to the issue being researched, argued and challenged. In that process our in house actuarial team play a key role.

3.12 The actuarial team as "thought consultants"

The accounting standard looks deceptively simple. In insurance you have to book the estimated lifetime cost of current claims as expense "now", even though the actual payments are not payable for decades. This simple concept is easy to say and at first glance easy to commit to memory. However, the impact on how you perceive risk, and leverage opportunities in a long tail scheme is profound *if you really get it!*

Most of us unaided don't really get it and need some help. I have no mathematical background (I did an arts degree) and by career am really a glorified claims officer. I cannot add, multiply or spell. It was only because of exposure to the concepts over many years working with actuaries (and Greg Tweedly and Doug Kearsley) that led me to at least partial understanding.

However, most of us as claims managers and leaders will have not fluked such similar long exposure. And so to support that analytical awareness we evolved the Actuarial Division.

In most schemes the actuarial team (if there is one) is small, is part of the finance area and does esoteric things usually associated with prudential requirements. Actuarial thinking is rarely deeply integrated into the core business.

In WorkSafe, however, we have created the actuarial team whose primary role is to support operational managers doing their job. Each Director is partnered with an in-house actuary and a performance analyst to form what I term an "analytical triumvirate". In addition the actuaries are integrated into the RTW, Treatment, Entitlement and Common Law teams to provide greater day-to-day support to help the business work out what is really going on and what they should be doing about it.

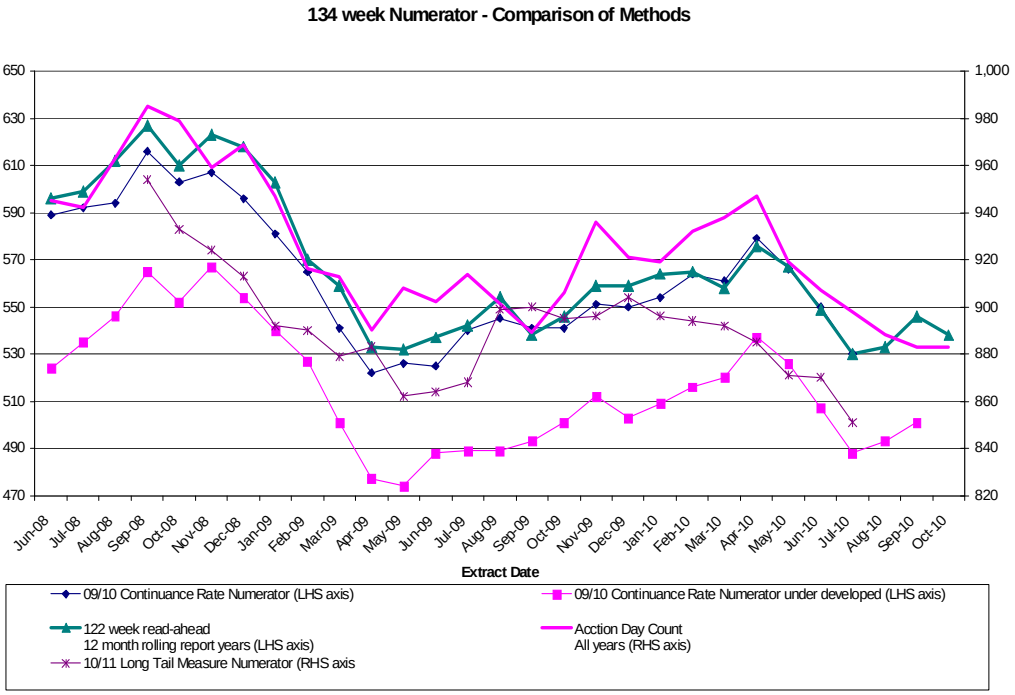
As demonstrated throughout this paper the support provided by our actuaries is often more analytical than actuarial, thus the concept of "thought consultants".

3.13 Sometimes you need multiple views

Monitoring of scheme trends sometimes has to involve multiple data sources if we are really to understand and manage the scheme. Although that makes for more complexity than is ideal, at least once a year we discover that what really is going on bears little resemblance to a headline KPI and the story has to be pieced to together from the multiple sources of evidence (qualitative and quantitative). Of course for day-to-day purposes you normally just go with the headline KPI, but periodically you have to be open to having a deeper look.

In early 2010 the main reports began to suggest that long tail weekly was deteriorating (at 130 weeks) and management became concerned that our strategies were not working. There looked to be a real risk of a blow-out in costs. However, strangely on some other cuts of the data (that we normally didn't pay much attention to) things looked more stable. In response we developed a simple chart that lined up four or five ways of cutting the data. This enabled us to more clearly see the collective trends rather than being blindsided by one. It led us to conclude we were over the worst and that the scheme was more stable than we thought.

Figure 5: Comparison of methods for comparing "long tail"



3.14 Beware the black box, or why you shouldn't entirely outsource your own brain

This is a slightly exaggerated cautionary tale.

Once upon a time, many years ago, a scheme used a massively complex statistical program to calculate the performance of agents. The agent payment component was called the TRPR. The model was developed by a number of statisticians and actuaries using the latest theories and mathematical relationships.

Exactly how it worked isn't important for this lesson. What is important is the uses to which it was put, and how it damaged agent and management culture.

In reality no user really knew how the thing worked (the maths was too complex) and so no manager could really ever challenge it. But on the surface it seemed to provide the solution every manager seeks: "I don't really have to understand because the model will work it all out and tell me the answer".

As a result, very little alternative scheme monitoring developed. Also very little analysis was happening in agents except for analysis of the TRPR model and what it was saying. Of more concern, very little management challenge of the output occurred largely because they had no alternative views of trends with which to challenge any contradictions in the output.

As with any of these sophisticated modelling exercises eventually the mathematical relationships it purported to see changed in the real world and the model became more and more unrealistic. This could have been fixed by simply re-calibrating to the main valuation and by more frequent remodelling, but those steps somehow were overlooked.

Of course management really didn't notice it because they mistook the black box for reality. But they paid agents incentive bonuses based on the black box calculations. Eventually the model got so out of sync with reality that it thought agents were performing well in a year that the official actuarial valuation reported record losses!

The TRPR in effect had become a cargo cult with everyone worshiping the technology they really didn't understand!

This cautionary tale explains my natural suspicion of "black box solutions".

Never entirely outsource your own brain.

3.15 Rationing the finite capacity for effective intervention

The classic mistake that many schemes make when trying to manage risk in claims or triage (and a mistake I have made at several points in my career) is to identify everything you see as a risk for management action. The reality, of course, is that we do not - and never will have - enough resources and skill to manage every risk on every file. In fact by trying to I believe we risk spreading our finite claims skill set so thin that we do more harm than good. I would argue that the scheme will achieve more effective outcomes for our clients if we manage 4,000 cases properly in early RTW vs. pretending to manage all 15,000 cases.

The secret is to match our finite capacity for effective intervention to risk.

In some areas we have already started to convert this to simple formulae and rules. In the case of Worksite Assessments we decided we could only afford 1,200 Occupational Physician interventions per annum. We work backwards to prioritise the 1,200 claims on which we will intervene. If the program identifies 2,000 claims then we simply tighten the criteria to match resources.

The same principle has been used for several years in our treater contact program. Our budget gives us enough clinicians to manage roughly 3,000 treater contacts a year. Each year we work out where we most need the leverage (which providers, treatment types and accident years) and set the criteria to match. No matter how many risks emerge we simply reset the priorities to match the intervention capacity.

At the global planning level, lack of accurate measures of agent workload and existence of conflicting scheme priorities means that we have been unable to systematise this "rationing" principle as much as we would like. Nevertheless in everything we do and with every new initiative we must constantly keep this philosophy it in mind and try to find the right balance.

3.16 Look for the overlooked lever if you haven't already looked

The legislative product we manage is virtually identical to that we inherited a decade ago (except for some increases in benefits). Yet the cost of the scheme is 40% less. The difference is that these days our agents manage it a bit better, or more precisely we utilise the tools within the legislation more effectively and systematically than we used to.

Most compensation legislation is full of rules and tests designed to contain the premium cost to employers. In many schemes, however, these rules are either not applied or have been applied with so little discipline that claims managers have effectively given up on them. Sometimes the real cause of the failure may be due to adverse legal precedents, but more frequently it's a management issue.

Long tail, employer reconciliations, section 99 (medical and like expenses) are all examples where we dusted off a supposedly ineffective legislative lever and made them at least partly work.

Of course the principle is equally applicable to any sort of potential lever that has been tried before. Often things fail simply because they are poorly implemented.

So keep your eyes and mind open to re-read your Act, to better understand the leverage in a new legal precedent, or even to look for the gem hidden inside a poorly implemented past project.

Sometimes the next best thing is something you tried once before.

3.17 Being accountable for ambiguity

You can have the most detailed and accurate claims trend monitoring in the universe, but sometimes what it says will be ambiguous or even contradictory. The secret of claims management is working within the ambiguity whilst still holding ourselves accountable. Although over a number of years actuarial results are very reliable, the annual actuarial release targets we have to achieve are not that much bigger than a rounding error. Sometimes a short term shift in valuation results may turn out years later to have been nothing more than a modelling error or a selection issue.

In 2000 we used to have two valuation actuaries and the Board would adopt which ever valuation was the more conservative (i.e.: the higher). That worked fine for arriving at a balance sheet figure, but made the valuation process next to useless for driving management focus and accountability. Normally the two consulting firms would come up with a similar figure for the total scheme liability, but they would often disagree on sub trends. In some years the actuaries would disagree on a particular benefit type with one saying it was improving and the other that it was deteriorating. In those circumstances the responsible manager could hide behind whichever figure looked better.

Even with day-to-day monitoring we still not infrequently find that different cuts of the data will give us conflicting views of what is really going on. As managers we eventually have to resolve this conflict sufficiently in order to be able to act.

On the one hand we need managers who will hold themselves accountable for the scheme trends and outcomes despite the ambiguity in the data. On the other hand we want managers who will not blindly accept someone else's interpretation but who will question to understand what it means without using it as an excuse not to accept the message.

3.18 A manager's role is to wrestle with the data (but then to embrace it)

There can be a tendency for us to delegate a little too much of the thinking to whomever advises us. Sometimes whatever is graphed is believed – even though the reality is that some graphs are just wrong.

An example of this is when we noted an alarming increase in files coded as TPI (Total Permanent Incapacity). We were looking at this indicator to check if long tail weekly was under any pressure and

the results looked bad. For a while we were in a panic as we interpreted this to mean that the sky was about to fall. Then someone thought to check what was really happening at the coal face. It turned out that we had changed the IT system some years previously and that that particular field was no longer used in calculations or subject to any quality control. Thus the data had become inaccurate and meaningless. Other data methods monitoring the long tail weekly trends were fine. The lesson here is that sometimes you need to ask the question: "how is the data actually being created at the coal face"?

Interpreting the data so that we know what's real requires management to take collective responsibility for understanding it. An actuary or a statistician in a corner cannot interpret the data alone without the business input and challenge to understand it.

3.19 All this is can be about much more than liabilities

In explaining these principles the examples I have given are largely about liability control strategies so naturally some will only see those principles as having relevance management of actuarial releases.

The principles themselves are quite generic and can be applied to any business problem being solved in our agent model *"as long as the outcome is measurable"*. If you can measure it, then modularity, leverage, monitoring, rationing, learning and evolving from what worked etc can be applied, as can the general project management, agent management and governance concepts.

To date we have not used these as extensively in Service - in part because our measurement systems have not been up to scratch. Another work in progress.

4. Conclusion

Many years ago I ran the data programming/ research function within TAC. In those days I thought the solution was to teach all the programmers to deeply understand the claims business. Unfortunately most programmers showed limited interest in becoming business experts, so we only succeeded with a few.

A few years later I took over a large Claims Division and I thought that maybe the answer is in getting the business to really understand the numbers. Unfortunately most of the business were too busy to deeply understand the data or did not see it as their role. We were only able to create a small number of business data champions.

With the benefit of hindsight what we have now tried at WorkSafe is a combination of the above. Building up more key data analysts and increasing the number of business data champions but supporting it all with an actuarial team deeply integrated with the claims business.

Although the anecdotes given in this paper largely relate to analytical concepts and solutions, the message overall is more about culture than maths. How to build a culture that embraces and challenges data and integrate that into day-to-day management of a compensation scheme? This paper records only some of our journey in that respect. A key message is to never entirely outsource your own brain.

5. Postscript: The view of an embedded actuary

I am approaching the end of my fourth year at WorkSafe. I came to the organisation with a reasonably long history as a consultant to the general insurance industry, and had always considered myself a long tail specialist with a particular interest in workers compensation. I was confident I knew a reasonable amount about workers compensation. I probably did, but I learn more and more every day.

One of the things that has most surprised me when I started at WorkSafe was the way in which the detail of the independent actuary's valuation of outstanding claims liabilities is used within the business. Back in the old days when I did these sorts of valuations myself, I remember that the focus was very much on the big picture, and that there was a reasonable tolerance of "swings and roundabouts" in the detail, particularly in the short term.

This is not the case at WorkSafe. Scheme experience is monitored against (amongst other things) the detailed valuation projections, and assumptions. For example, for weekly benefits the external actuary provides us with a projection of "active" weekly claims, payments and PPACs all by accident half year and projection quarter. We monitor by accident year and interpolate to get a monthly projection, and track that with detailed reporting each month. Deviations from the projection are explored and (are expected to be) explained each month.

This puts tremendous pressure on the external actuaries to understand the nuances of the data analysis and the projections, and be able to understand and explain what is different, and how and why the models are reacting as they are. In my early days at WorkSafe I felt this was a little unfair – but I got over that. It is all about accountability.

- We hold the external actuaries accountable for the reasonableness of the valuation and the movements in it. This is perfectly reasonable given:
    - we charge employers premium based on what the actuaries say
    - part of the remuneration to agents is based on what the actuaries say
    - the Board looks for a management response to what the actuaries say.
- Management is held accountable for managing the underlying experience unfolding through the valuation.

As internal actuaries, we sit somewhere in between. We too are expected to understand the nuances of the valuation actuary analysis, models and projections, and help translate those for management. We must also be on top of the claims experience and operational issues to help the rest of the business see where the opportunities and threats are. We need to be much more than a translator between the two.

I've had to reconcile myself to the idea of measuring and reporting on an actuarial release every six months. As Len says earlier in this paper, at any one valuation the movement can be at the margins from an actuarial perspective, but there is value in messages for the business, and there is certainly value in a longer term series of releases at benefit level. Part of my role is to make sure that the business understands actuarial release – both as a concept, and in practice. I take very seriously my responsibility to understand the results and the way they are used within the business.

The great advantage in being embedded in the business is being given real opportunity to contribute to how the business operates on a day to day basis. This requires deep relationships based on trust, understanding and mutual respect. I have a lot to thank my predecessors and longer-standing colleagues for: they worked hard to earn their place in the business and prove that actuaries aren't (necessarily) pedants or propeller heads. Now when my opinion is sought on a new initiative, it is not just about what it might cost (and the implications for actuarial release and breakeven premium) but what I think about the way it is going to be structured, and so on. This might be a wee tiny initiative with minor cost implications or something big and potentially ground-breaking.

I'll finish with a couple of things I have learned in my time as an embedded actuary:

- Being data rich is both a blessing and a curse. At some point you need to make a call, even if you could analyse for ever and ever! He who wrote the rest of this paper often cautions against "polishing turds". Not the most beautifully crafted of sayings, but incredibly insightful.

- It is not all about the numbers. In an environment which values numbers so highly, there is a temptation to prove that everything saves money or reduces premium. I am reminded of one initiative where those closest to it desperately wanted me to prove its financial worth. I couldn't see it in the numbers, and found that I had to work hard to convince them that this was not a problem: what they were doing was worth doing because it was the right thing to do. Nice twist for an actuary.
- Ask the question a number of different ways. Not a new lesson (I learned this one many years ago under the tutelage of Geoff Atkins), but one which has been reinforced in my years at WorkSafe. There are many dimensions to this – just a few of which follow:
    - people will answer the question they think you are asking;
    - we are all human, and we make mistakes;
    - just because it is on a chart doesn't mean it is right
    - sometimes urban myths can get in the way.
- Keep asking questions. I learn new things about our scheme every day, and the devil is in the detail.
- The value of understanding the "why". Why operational processes have been put in place, and how those decisions will play out in the experience. Why claims experience has changed - temporarily or "permanently". Why scheme participants may have changed their behaviour.
- Changes in behaviour confound everything. This is inevitable in a scheme like ours with so many different parties influencing what we see in claims data, and when we see it. I was certainly aware of this before I joined WorkSafe, but it is fair to say I didn't really understand the extent to which it plays out in every trend you look at.
- You can influence markets. The Clinical Panel initiative Len talks about in the paper is a perfect example of this.