Change • Challenge • Opportunity

Injury & Disability Schemes Seminar

12 - 14 November 2017 • Sofitel • Brisbane

Actuaries Institute

Predictors of Return to Work: Experience from Workers Compensation Insurance in NSW

Dr Sriyani Dias
Catherine Day
SIRA-(State Insurance Regulatory Authority)

Purpose

To present:

- Return to Work experience of NSW
- Research findings on predictors of RTW
- Analysis results based on 3 robust data cohorts

Analysis:

- Analysis 1: Pre 2012 Claims data: Nominal Insurer
- Analysis 2: Post 2012 Claims data: NSW System
- Analysis 3: 2016 RTW Survey data : NSW System

Analysis Methods:

- Regression Analysis
- Machine Learning Model : Random Forest

SIRA as a new regulator

- SIRA regulates **workers compensation** insurance delivered by:
  - Nominal Insurer (icare)
  - SiCorp Treasury Managed Fund (icare)
  - Other 59 Self-insurers
  - 6 Specialised insurers
- SIRA regulates **motor accidents compensation** insurance delivered by private CTP insurers
- SIRA also regulates the **Home Building Compensation** Fund which is administered by SiCorp in icare
- SIRA provides assurances to government and community:
  - that the system is fair and equitable
  - achieves the aim of getting injured workers back to work

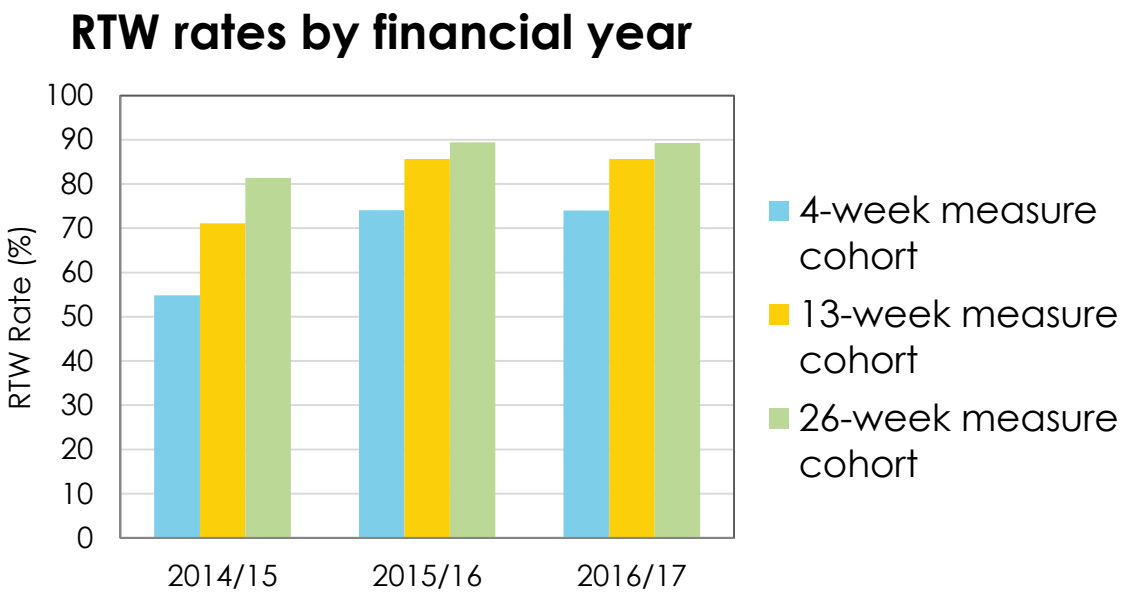
SIRA as a new regulator

Perceived conflict of interest	Impartial
Prescriptive – influence via rules	Performance based – influence via data and leadership
Leveraging legislation	Leveraging licensing conditions – insurer/provider supervision
Compliance focus	Outcome focus
Reactive, opaque methods	Proactive, transparent, and accountable methods
Data rich but analytics poor	Rapid data analytics tools & culture – evidence and risk
Perceived 'Defender' mindset	Open to improvement mindset
Poor stakeholder engagement	Excellent stakeholder engagement
Distant from customer	Strong customer understanding and impact
Culture not enabling innovation	Encouraging innovation

Why is RTW important

- For most individuals, working improves general health and wellbeing and reduces psychological distress
- An unnecessary delay in returning to work after injury is often associated with delayed recovery
- Long term work absence has a negative impact on:
  - an individual's health and wellbeing
  - an employer's workplace culture and bottom line
  - the cost of the workers compensation system

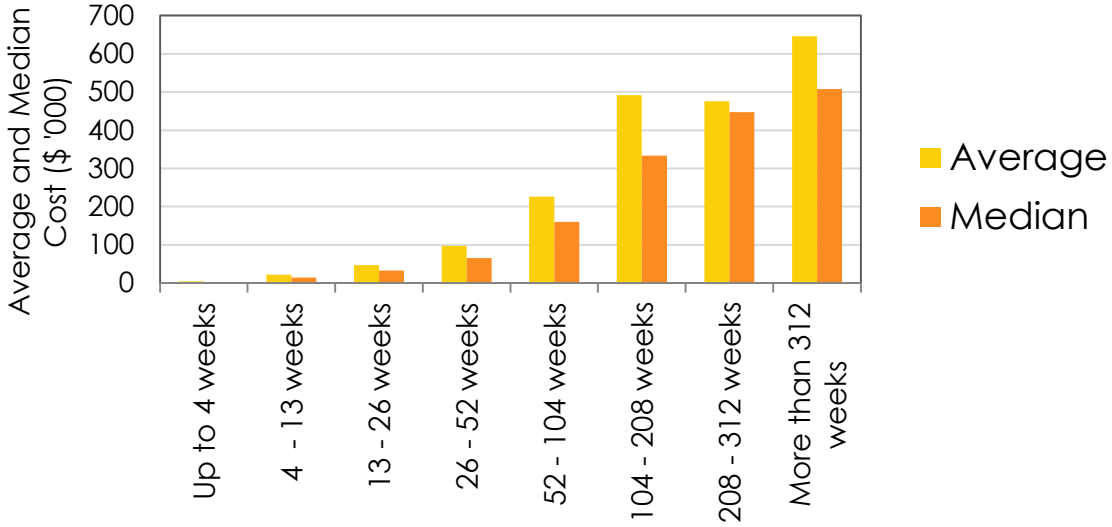
Sources:
Steenstra et. al., J, Occup. Rehab, 2017, 27,369-81
Steenstra et. al., Occup. Env. Med, 2005, 62, 851-60
Feuerstein et.al., Pain, 2003, 102,51-61



RTW Rates – NSW Workers compensation System

There is a continuing improvement in return to work rates in NSW over the past three years.

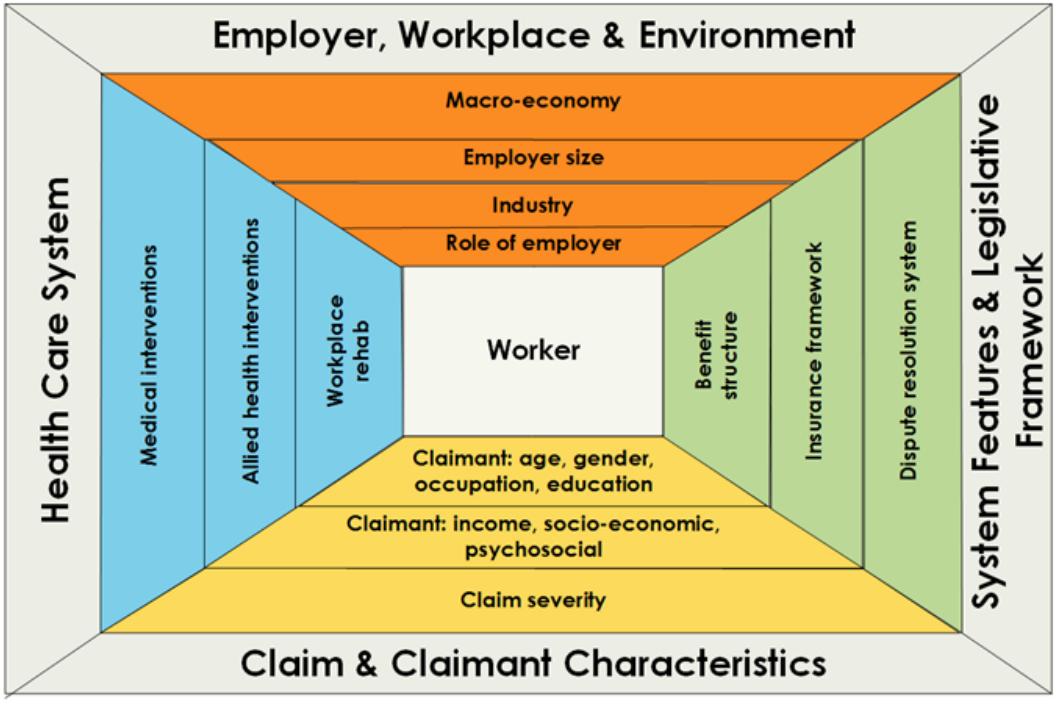
Average and Median Cost by Time lost



Cost and time lost – NSW Workers compensation System

There is an exponential increase in average and median claim cost with the increase of time away recovering from an injury

Multi-Dimensional Aspects of RTW



Source:
Loisel et. al.,J. Occup.Rehab. 2005,15(4) 507-524

Factors associated with RTW outcomes

- **Employer, Workplace & Environment:**
  - Employer support
  - Systems in place (policies and procedures)
- **Claim & Claimant Characteristics:**
  - Health status and severity of the injury
  - Claimant demographics
  - Psychosocial - job satisfaction, job demand, job control
- **System and legislative framework:**
  - Dispute resolution
  - Customer experience
- **Health care system and providers:**
  - Effective medical care and Rehab

Areas for Analysis – NSW Workers Compensation System

Item	Pre 2012 Claims data	Post 2012 Claims data	2016 RTW Survey data
Model	**Claimant model** Macro-economic Model Claimant model ANOVA , Chi Square and Multiple regression Models	**Claimant model** Chi Square and Logistic regression Model and Machine Learning Model_ Random Forest	**Injured Worker model** Chi Square and Logistic regression Model
Measure	RTW based on duration on benefits – 26 week measure Number of weeks on weekly benefits	RTW based on Work Status code – 26 week measure Work Status code - 1,2,3,4/1 or more day time lost claims	RTW based on Work Status Injured worker reported as RTW at the time of the survey
Source data	Number of claims: 38,231 Claims reported from June 2010 to November 2011 **Nominal Insurer Only**	Number of claims: 209,128 Claims reported from July 2012 to June 2016 **Workers Compensation System Data**	Number of claims: 812 RTW Survey data (2016) National RTW Survey: Injured worker interviews

Analysis 1: Pre 2012 Analysis: Claimant & macro economic models

- **Dependent variable:** Proxy measure of RTW (duration)
- **Predictive variables:** limited number available from claims data
  - Claimant demographics, e.g. age, gender, occupation, income, NESB (interpreter required)
  - Employer characteristics: size, rehab provider involvement
  - Macro economic: unemployment rate, participation rate
- Interdependence and normality analysis conducted
- **Method of analysis:** ANOVA, Chi square & Multiple regression

Analysis 1: Pre 2012 Analysis:
Multiple Regression Model: Overall findings

- 18% of RTW outcomes is influenced by claimant characteristics
- 39% of duration results is influenced by macro-economic factors
- No clear relationship between predictor variables and duration
- Following variables had strong relationship with duration:
  - **Non-English Speaking Background**
  - **Workplace rehabilitation program usage**
  - Percentage of **medical payments**

Pre 2012 Analysis:
Claimant model: Chi square (χ2) results

Claim/Workplace Characteristics	Degrees of freedom	χ^2 obtained (test statistic)	χ^2 (critical)	P value	Cramer's V
26 Week Outcome					
Employer Size	10	931.9867	18.307	<.0001	0.1104
Gender	5	139.6885	11.07	<.0001	0.0605
Required interpreter flag	5	470.5903	11.07	<.0001	0.1110
Received rehabilitation flag	5	4739.5109	11.07	<.0001	0.3522
Bodily location	40	437.5755	> 43.773	<.0001	0.0479
Industry Division	80	567.4529	> 43.773	<.0001	0.0545
Nature of Injury	5	1265.7116	11.07	<.0001	0.1820
Occupation	40	520.6599	> 43.773	<.0001	0.0522
Region	70	158.5855	> 43.773	<.0001	0.0288

Pre 2012 Analysis:

Claimant & macro economic model: ANOVA results

Variable	Number of categories	Degrees of freedom	Kruskal-Wallis χ^2	P Value
26 Week Outcome				
Claimant award rate	9	8	291.75	<.0001
Current age of claimant	12	11	273.62	<.0001
Hours worked per week	2	1	79.19	<.0001
% Medical to Total Payments	8	7	1362.43	<.0001

Variable	Number of categories	Degrees of freedom	Welch F value	Welch P Value
26 Week Outcome				
Unemployment rate - Male	2	1	19.51	<.0001
Unemployment rate - Female	2	1	54.59	<.0001
Unemployment rate - Person	2	1	51.03	<.0001
Participation rate - Male	2	1	29.11	<.0001
Participation rate - Female	4	3	12.12	<.0001

Pre 2012 Analysis:

Multiple regression model results – 26 week measure

Label	Parameter estimate (β)	Standard error	t Value	Pr > |t|
Intercept	4.69456	1.91272	2.45	0.0141
Weekly wage rate	0.00281	0.000621	4.52	<.0001
Hours work per week	0.70421	0.03471	20.29	<.0001
Legal GIC ($)	0.00524	0.000257	20.35	<.0001
percent Medical to Total Payments	-24.39321	0.94961	-25.69	<.0001
Received rehabilitation flag	27.10412	0.44735	60.59	<.0001
Required interpreter flag	19.06685	1.52285	12.52	<.0001
Current age of claimant	0.19986	0.01519	13.16	<.0001
Nature of injury/diseases	0.01754	0.0011	15.93	<.0001
Bodily location of injury/diseases	0.00277	0.00137	2.02	0.0432
Occupation types (ASCO 1)	0.44595	0.08437	5.29	<.0001
Gender (Male/Female)	-1.25518	0.45794	-2.74	0.0061
Employer size	-5.04262	0.30881	-16.33	<.0001
Region (used defined)	0.14882	0.04803	3.1	0.0019

Analysis 2: Post 2012 Analysis

Claimant model

Aim: To see any variation between pre/post periods

- Used claimant model to see the relationship between factors having affecting RTW outcomes
- **Dependent variable:** RTW Measure based on Work Status Code
- **Predictive variables:** limited number available from source
  - Claimant demographics , e.g. age, gender, occupation, income, NESB (Non-English Speaking Background : interpreter required)
  - Claim Characteristics: Psychological claim, Surgical claim, Delay to reporting
  - Employer characteristics: size
  - Provider involvement: workplace rehabilitation provider
- **Method of Analysis:** Chi square, Logistic regression & Random Forest Model

Post 2012 Analysis : Findings

Logistic regression and Random Forest model

- 14% of RTW outcomes is influenced by characteristics of injured worker together with workplace environment.
- Similar to pre 2012 model that claim characteristics explain small variance.
- Model fit statistics showed that the Regression model performed well with an overall performance of 74% as shown by the ROC curve.
- To both validate and augment the results from the logistic regression Machine learning model-Random Forest was also developed for the 26 week RTW measure.
- Variables highly impacted on RTW are:
  - Psychological claim
  - Nature of Injury
  - Body location
  - Workplace Rehabilitation

Post 2012 Analysis

Claimant model results: Logistic regression

Statistically significant results: Likelihood ratio shows significance at Pr < .0001

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	134976.20	120286.21
SC	134986.42	120603.20
-2 Log L	134974.20	120224.21

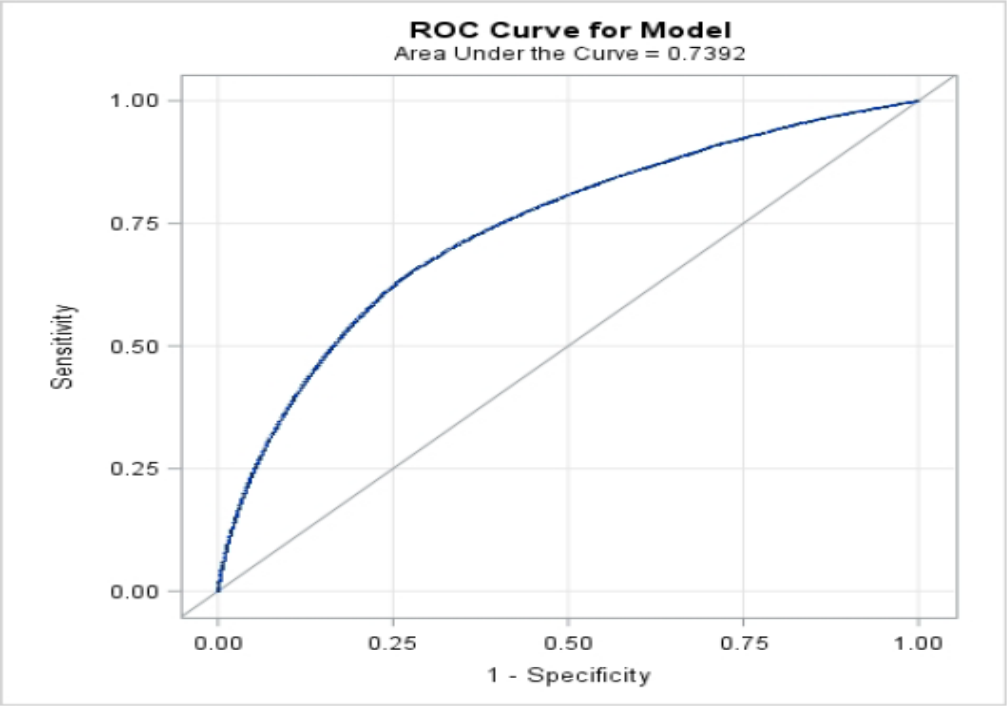
R-Square	0.0698	Max-rescaled R-Square	0.1441

Testing Global Null Hypothesis			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	14749.9844	30	<.0001

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
Gender	1	50.2887	<.0001
Age	5	601.9139	<.0001
Nature	3	184.9605	<.0001
Work size	2	160.3554	<.0001
Body location	6	92.1586	<.0001
Industry	2	101.0163	<.0001
Occupation	1	225.1555	<.0001
Pre-injury earnings	2	34.889	<.0001
Delay in reporting	5	1728.4193	<.0001
Interpreter payments	1	411.2473	<.0001
Legal payments	1	607.1722	<.0001
Rehab payments	1	4246.6209	<.0001

Post 2012 Analysis

Claimant model results: Logistic regression



Post 2012 Analysis

Regression vs Random Forest

- **Logistic Regression**
  - Powerful parametric model
  - Assumes linearity between the predictors and the logit transformed target
  - Easy to interpret
  - Limitation: Highly correlated predictor variables confound parameter interpretation

- **Random Forest Model**
  - Flexible and non-parametric in nature and therefore more able to capture non-linear structure.
  - Could be difficult to interpret
  - Able to include highly correlated variables (surgery and psychological claims)

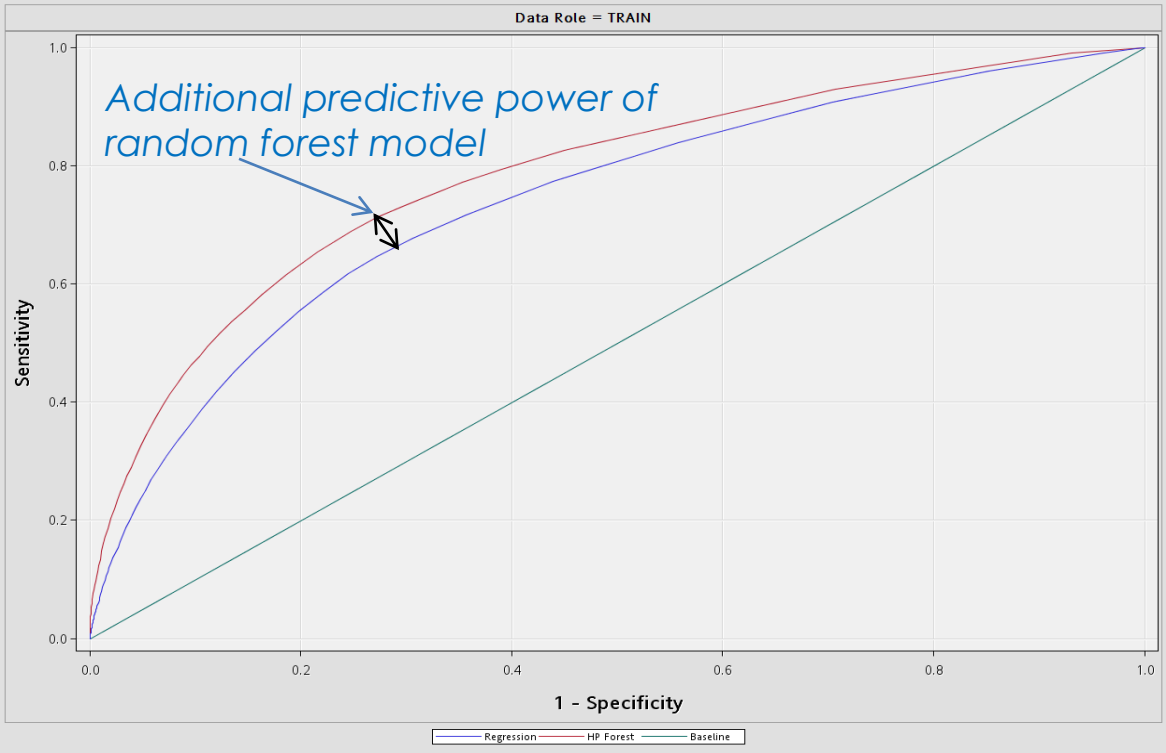
Post 2012 Analysis

Random Forest Model – The Method

- Popular choice for classification problems
- Large number of decision trees are developed on bootstrapped resampled versions of the original dataset
- Only uses a subset of predictors at any time
- The results are averaged to get an ensemble learner. This reduces over-fitting – a potential problem with machine learning models.
- However, importantly, since the growth of ROC is relatively modest from the random forest model we retain confidence in our regression results

Post 2012 Analysis: Results

This result is an improved predictive performance and shown the ROC chart

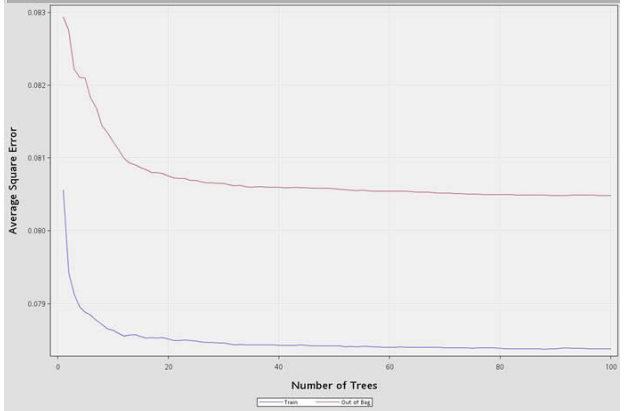


Post 2012 Analysis: Random Forest Model

Estimate of the error rate and loss reduction

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation
RTW_026w_1		_ASE_	Average Squared Error	0.078383	
RTW_026w_1		_DIV_	Divisor for ASE	407904	
RTW_026w_1		_MAX_	Maximum Absolute Error	0.978113	
RTW_026w_1		_NOBS_	Sum of Frequencies	203952	
RTW_026w_1		_RASE_	Root Average Squared Error	0.279969	
RTW_026w_1		_SSE_	Sum of Squared Errors	31972.59	
RTW_026w_1		_DISF_	Frequency of Classified Cases	203952	
RTW_026w_1		_MISC_	Misclassification Rate	0.098342	
RTW_026w_1		_WRONG_	Number of Wrong Classifications	20057	



Loss Reduction Variable Importance

Variable	Number of Rules	Gini	OOB Gini	Margin	OOB Margin
Psychological Payment	1436	0.003590	0.00465	0.007179	0.00929
Nature	2908	0.003415	0.00415	0.006830	0.01147
Body location	4681	0.002580	0.00236	0.005160	0.00679
Rehab payment	1777	0.004570	0.00234	0.009140	-0.00100
Delay in reporting	6830	0.002704	0.00191	0.005408	0.00673
Legal payment	1759	0.001411	0.00168	0.002821	0.00466
Interpreter payment	1052	0.000703	0.00100	0.001406	0.00419
Surgery payment	761	0.003358	0.00071	0.006716	-0.00284
Occupation	3483	0.000533	0.00028	0.001066	0.00083
Age	7770	0.001003	0.00012	0.002006	0.00264
Industry	4443	0.000496	-0.00003	0.000992	0.00049
Gender	4110	0.000313	-0.00014	0.000626	0.00016
Pre-injury weekly earnings	5580	0.000562	-0.00020	0.001125	0.00037
Work size	5717	0.000559	-0.00022	0.001118	0.00074

Estimate of the error rate 7.8%

Analysis 3: RTW Survey 2016

- Results of Logistic regression and Random Forest Models show limited influence of claim characteristics in explaining RTW
- Key characteristics of RTW are not all captured in traditional workers compensation data
- Supplement with RTW Survey

RTW Survey 2016

- National RTW Survey 2016 for NSW- Co-ordinated by Safe Work Australia conducted by Social Research Centre April 2016
- 2 cohorts: 'Historic Return to Work Cohort' (10 or more days compensated) and 'Balance Cohort'
- Eligible for the current survey_2016 Balance cohort
- The survey sample selection is based on the injured workers in NSW who:

- had at least one day away from work
- submitted a claim in the two years (March 2014-January 2016) prior to the interview period
- had or did not have payment-related activity within 6 months prior to the sample being drawn, and
- worked in either premium paying or self insurance (including the government sector)

RTW Survey: Sample selection

- 812 injured workers (66% male) selected for the sample
- Survey includes questions on:
    - claimant characteristics (Average age 45 years)
    - employer support and involvement
    - claim experience
    - medical care
    - claimant psycho-social factors
- 3 measures of RTW:
    - Any time RTW
    - RTW at the time of the Survey
    - RTW and at work for a period of 3 months

2016 RTW Survey Analysis:

Predictors not in other models but used in the Survey Model

Worker Health
Injured worker's perspective of their recovery

Medical Care
Access to a GP and medical treatment
Communication with GP

Psychosocial
Job satisfaction
Job control
Job demand

42 questions grouped into 5 major categories

Claims Experience
Open and honest processes and communication
Good system

Employer support
Policies and procedures
Employer commitment

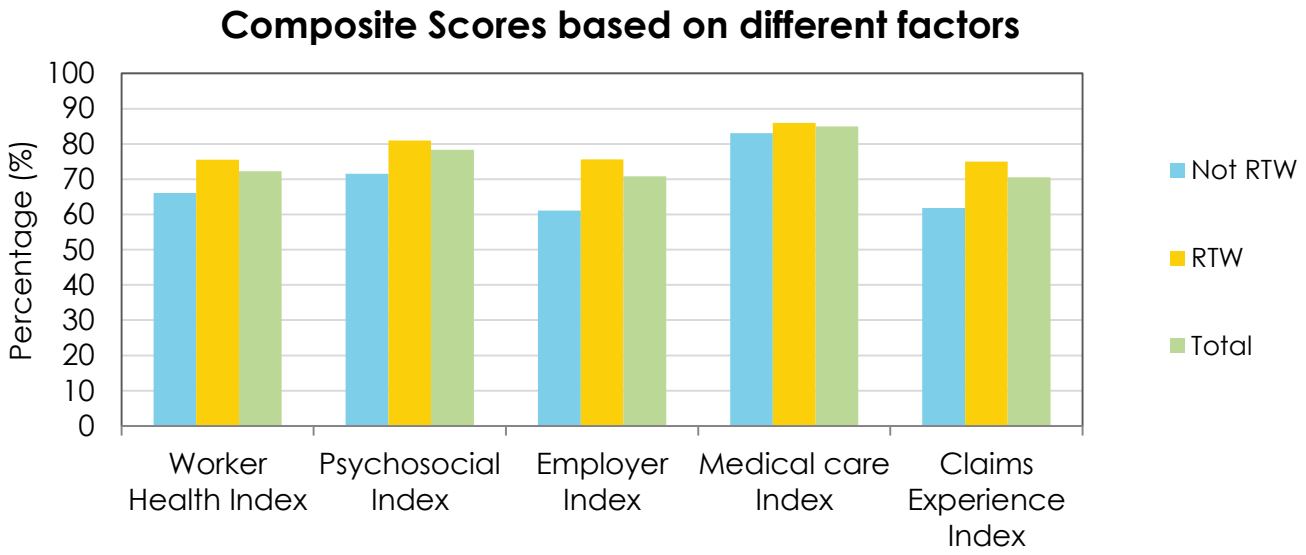
Psychosocial factors impacting RTW: Research Findings

- Hartvigsen etal., (2004)* defined a psycho-social factor as a 'measurement that potentially relates psychological'.
- Broader categories:
  - Perception of work: job security, time pressure, Job control e.g. job demands, pace, work content, work control, work tempo, quantitative/qualitative
  - Social support at work (including recognition and respect, social support, co-worker support,social relations, work relations, external relations, supervisor support)
  - Stress at work (including stress, over strain, job strain, level of distress, total mental exertion)

Occup Environ Med 2004;61:e2 (http://www.occenvmed.com/cgi/content/full/61/1/e2)

Composite Indicator Scores by category

Survey based data revealed that medical care and worker psycho-social well being as foremost contributors to return to work following an injury



Binary Logistic Regression: Results

Results revealed that medical care, psycho-social factors and workplace environment can explain about 63% of factors conducive to return to work outcomes in NSW sample

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	739.052	453.328
SC	743.751	688.303
-2 Log L	737.052	353.328

R-Square	0.3766	Max-rescaled R-Square	0.6313

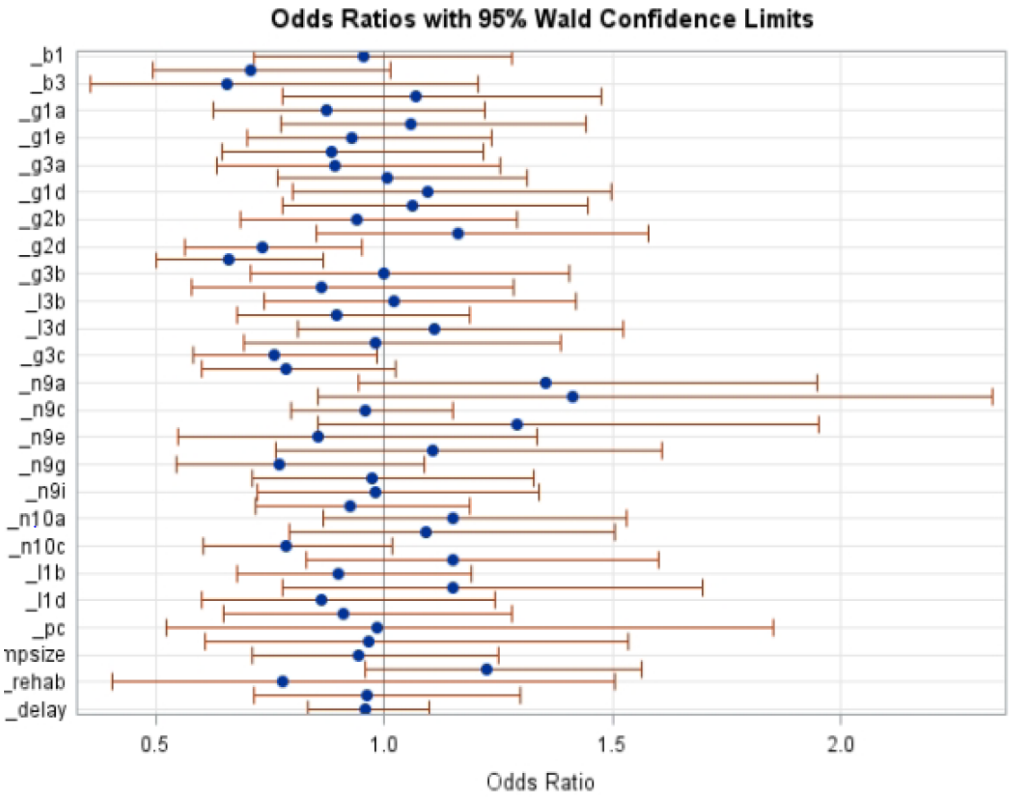
Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	383.7241	49	<.0001
Score	428.2220	49	<.0001
Wald	143.6594	49	<.0001

Model : Results and Validity

- Statistically significant results: Likelihood ratio, Score and Wald tests (Pr <.0001)
- Variables in the model improve the model fit: Chi-square test statistics and p-values (Pr <.0001)
- 63% of variance is explained by predictor variables
- Predictor variables influencing RTW outcomes:
  - Medical Care: Access to GP and medical care
  - Psychosocial: Job satisfaction, Job control
  - Employer Support: Fair treatment, communication, and access to information
  - Claims experience: Open and honest process, good communication

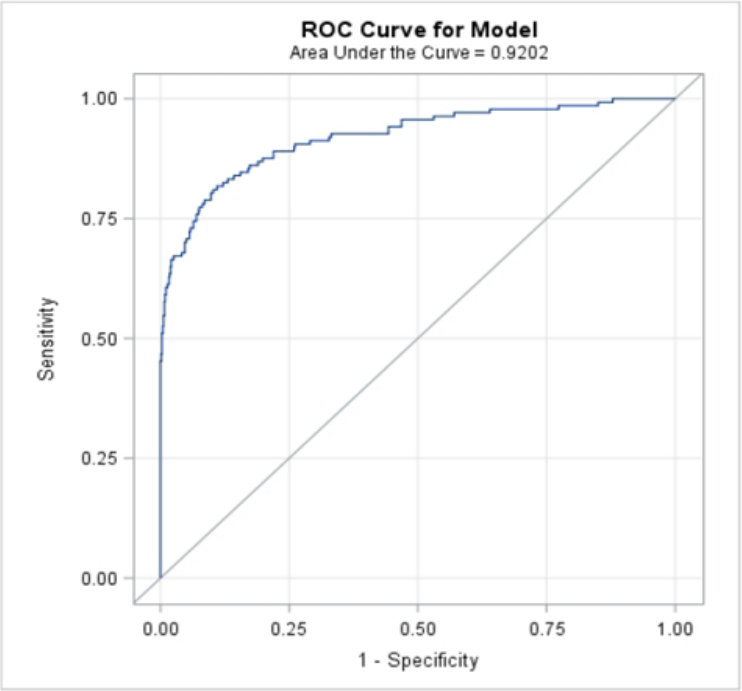
Model Results: Odds Ratios

_b1: Overall health rating
_b3: Are you recovered as much as you are going to?
_g1a: The work you are doing is important to you
_g1e: The work you are doing is valued by others at work
_g3a: Feel you are part of a community at work
_g1d: Your opinions and suggestions are considered at work
_g2b: Your skills and abilities are used appropriately
_g2d: You are physically capable of doing your job
_g3b: Employees and management are generally supportive
_l3b: Employer provided enough information
_l3d: Employer helped you with your recovery
_g3c: Your supervisor is committed to workplace safety
_n9a: You were able to easily get an appointment with a GP
_n9c: Did the GP contact your employer about you?
_n9e: You had confidence in the GP you were speaking with
_n9g: GP issued medical certificates that stated when you could return to work
_n9i: GP explained to you the physical benefits of RTW ASAP
_n10a: Able to easily access medical treatment
_n10c: Medical treatment helped you recover
_l1b: Good communication between various people I dealt with
_l1d: I believe the system treated me fairly
_pc: Has a previous claim or not?
mpsize: Employer size
_rehab: Rehab payments
_delay: Delay in reporting



RTW Survey Data: Model Performance

Claimant Survey Logistic regression; Significant performance (92%) of the data model



Summary of Findings

Post 2012 claimant model and 2016 Survey Model

Post 2012 Claimant Model	RTW 2016 Survey Model
Large dataset: Claimant details from Claims Database 209,128 Claims Ability to merge payment transactional data to RTW outcome variable	Sample Survey: Injured worker 812 Injured workers Ability to look into data not available through claims data
Static Variables: • Age, Nature, body location, Gender, Occupation, NESB, Wage rate, hours worked, employer size, Psych injury, • Dynamic variables: • Rehab involvement • Legal cost • Medical payments • Surgical involvement	Survey questions grouped under 5 broader categories: • IW Health • Medical care • Psychosocial • Claims experience • Employer involvement and support
Findings: Variables influencing RTW outcomes: • Psych injury • Nature of injury • Body location • Workplace rehabilitation	Findings: Variables influencing RTW outcomes: • Medical care- access to GP, medical treatment • Psychosocial- Job satisfaction, job demand • Employer support-commitment, support • Claims experience- Open/honest process, Good communication

Conclusions

Multifactorial nature of predictors of RTW

- Claim and claimant characteristics available in traditional workers compensation statistics explain only small proportion of variance in the data.
- Psychological injury and workplace rehabilitation explain variance from claimant models.
- Medical Care and Psychosocial factors explain a large variance from survey data.
- Employer support and claim experience also significantly associated with RTW outcomes
- Data for some other factors are not currently available
- Data items required to measure RTW outcomes need to be reviewed and multiple methods of collections need to be explored
- Need more in-depth analysis using advanced analytical methods

Interventions

- **Medical Care - Access to GP:** GP capability building strategy
- **Psychosocial - Job Satisfaction:** Working with Safework NSW on the Mentally Healthy Workplace Strategy
- **Employer support:** Building employer capability by identifying employers (small and large) that require assistance and tailoring interventions to suit their needs, developing online training modules for RTW Coordinators
- **Psychological injury:** Developing a Psychological injury strategy
- **Workplace rehabilitation:** Reviewing the regulatory approach to achieve improved RTW outcomes without a significant cost burden
- **Claims Handling:** Audit program, developing new Claims Handling Guidelines, undertaking co-design strategies with insurers to address claims hot spots

Data, Digital and Analytics

- **New BI tools and Dashboards on RTW to monitor and act on Performance**
- **Project on "Measure Well":**
  - Develop Systematic measurement of RTW outcomes, health outcomes and claimant experience in NSW to understand the impacts of policy levers, of reforms and regulatory activity
  - Robust multi-layered measures of RTW, health and experience
- **Deploy Machine Learning Models to supplement statistical models**
- **Analytics on relative performance of providers for benchmarking**

Presenters

Dr. Sriyani Dias
Manager, Insurer and Policy Analytics, State Insurance Regulatory Authority NSW
02 4321 5333 | 0436658196 | sriyani.dias@sira.nsw.gov.au

Catherine Day
Director, Claimant Outcomes, State Insurance Regulatory Authority NSW
02 4321 5914 | 0413459126 | catherine.day@sira.nsw.gov.au