# THE INCIDENCE AND COSTS OF CHEMOTHERAPY SIDE EFFECTS

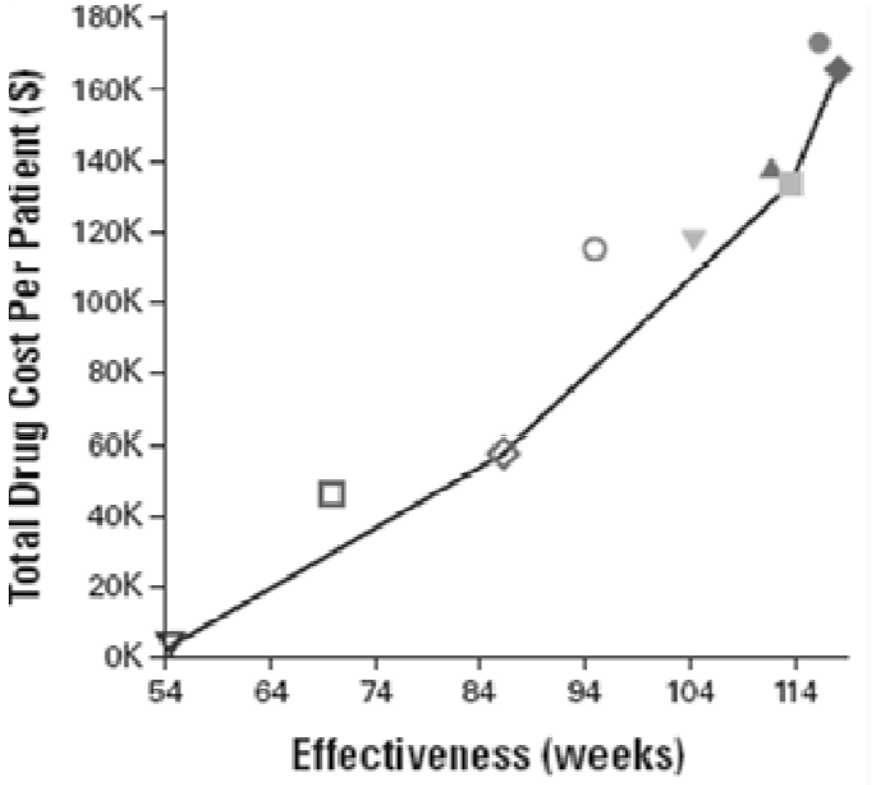
Alison Pearce - PhD candidate Centre for Health Economics Research and Evaluation, UTS

Supervisors: Marion Haas, Rosalie Viney

CAER 2013

# Chemotherapy

- Chemotherapy drugs can be life extending for people with cancer. But…
  - they contribute a small amount to survival
  - they are increasingly expensive
  - they cause side effects



# Chemotherapy side effects

- Chemotherapy side effects can:
  - Impact on patients physical wellbeing
  - Impact on patients quality of life (QoL)
  - Potentially impact on cancer survival
  - Be expensive to manage




# Economic evaluation

- In Australia, new drugs are listed for public subsidy by PBAC on the basis of economic evaluation
- Literature review examined how side effects are incorporated into economic evaluations of chemotherapy
  - Costs and outcomes of side effects are not included in any systematic way
  - Clinical trials are the primary source of probabilities
  - Resource use is often estimated with expert opinion or based on best practice
- These data sources may not reflect clinical practice
- If side effects aren't accounted for (accurately) then outcomes of economic evaluations may be biased

# Aims & Objective

- Overall objective:
  - To better inform models of chemotherapy cost effectiveness
- Aims:
  - Explore in clinical practice:
    1. the incidence of chemotherapy side effects
    2. the factors which influence the incidence of chemotherapy side effects
    3. the resource use associated with chemotherapy side effects

# Department of Veterans Affairs

- The Australian Government Department of Veterans Affairs provides services to nearly 500,000 war veterans and their families in Australia
- Clients with a 'gold card' are entitled to the full range of services at DVA's expense
- DVA has actively encouraged the use of their data to undertake pharmacoepidemiological research

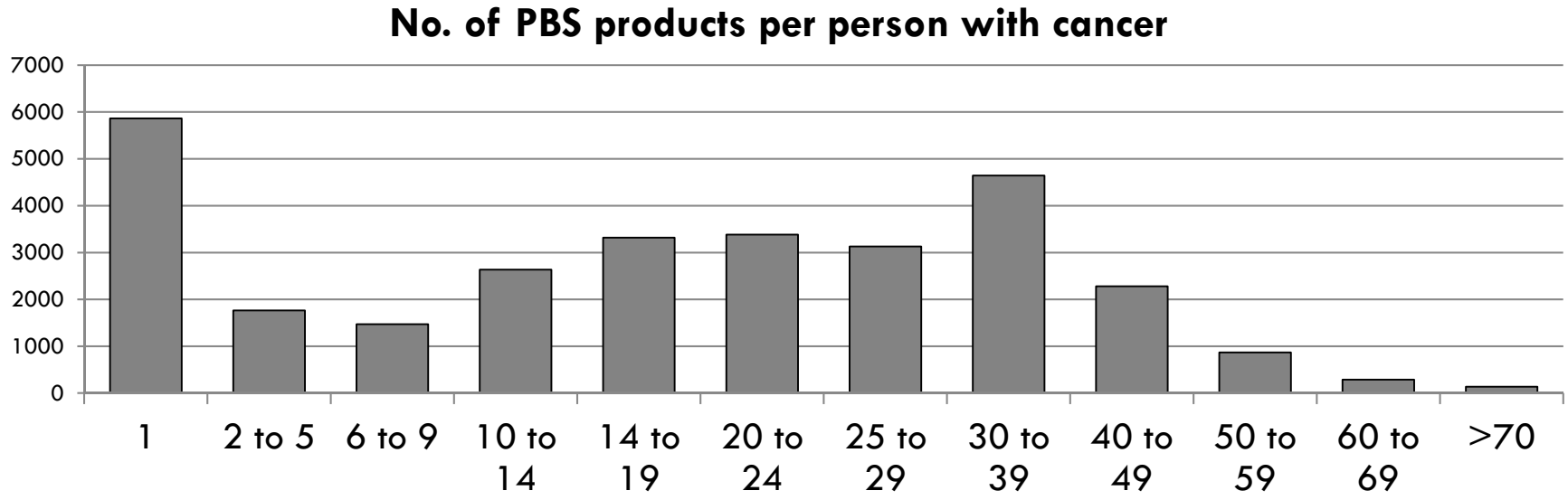
# Data linkage

- Extract from DVA client database – individuals residing in NSW 1994 – 2007
- Linked by CHeReL to NSW population data

| **Registry** | **Start Date** | **End Date** |
|---|---|---|
| **NSW Cancer Registry** | Jan 1994 | Dec 2009 |
| **Repatriation PBS** | 01 July 2004 | 31 Jan 2010 |
| **Repatriation MBS** | 01 Jan 2000 | 31 Jan 2010 |
| **Admitted Patient Data Collection** | 01 July 2000 | 30 June 2009 |
| **Emergency Department Data** | 01 Jan 2005 | 31 Dec 2009 |
| **Resource utilisation period** | **01 Jan 2005** | **30 June 2009** |

# Sample

- Individual Gold Card Holders 129,307
  - Individuals with a cancer diagnosis 29,480
  - Individuals who received chemotherapy 12,030
  - Total doses of chemotherapy 111,059



# Demographics

| **Demographic** | | **Chemo cohort** |
|---|---|---|
| Proportion males | | 72% |
| Mean age (median) in years | | 81 (83) |
| age range | | 46 - 106 |
| age group | <70 yrs | 14% |
| | 70-80 yrs | 23% |
| | >80 yrs | 63% |
| Mean Rx Risk score (weighted comorbidities) | | 8.83 |
| RxRisk score range | | 0 - 26 |

# Cancer

| Cancer site | N | % of cancer |
|---|---|---|
| **Prostate** | 3124 | 39.17 |
| **Breast** | 1059 | 13.28 |
| **Melanoma of skin** | 881 | 11.05 |
| **Colon** | 491 | 6.16 |
| **Lung** | 354 | 4.44 |
| **Non-Hodgkin's lymphoma** | 349 | 4.38 |
| **Rectum, rectosigmoid, anus** | 279 | 3.5 |
| **Bladder** | 186 | 2.33 |
| **Ill-def & unspec site** | 136 | 1.71 |
| **Head & neck** | 591 | 0.65 |

# Chemotherapy

| Drug | Frequency | % of chemo | Used to treat… |
|---|---|---|---|
| **Fluorouracil** | 2198 | 18.20 | Breast, colorectal |
| **Goserelin acetate** | 1909 | 15.80 | Prostate, breast |
| **Leuprorelin acetate** | 1307 | 10.82 | Prostate |
| **Bicalutamide** | 1005 | 8.32 | Prostate, breast |
| **Tamoxifen citrate** | 776 | 6.42 | Breast |
| **Capecitabine** | 327 | 2.71 | Breast, colorectal |
| **Rituximab** | 321 | 2.66 | Lymphoma |
| **Cyclophosphamide** | 305 | 2.53 | Breast, leukemia |
| **Anastrazole** | 280 | 2.32 | Breast |
| **Gemcitabine** | 276 | 2.28 | Breast, lung, bladder, pancreas |

# Overview of methods

- 4 common side effects examined:
  - Diarrhoea, anaemia, nausea and vomiting (N&V), and neutropenia

- Aim 1 – incidence of side effects
  - The incidence of each side effect was calculated
- Aim 2 – factors influencing incidence of side effects
  - Multiple regression analysis using generalised estimating equations identified factors which influence the incidence of each side effect
- Aim 3 – resource use associated with side effects
  - Multiple linear regression identified whether those who experienced a side effect had higher chemotherapy costs

# Overview of assumptions

- No direct data on whether someone experiences a side effect, so require a proxy
- Specific treatments are likely (based on best practice) to be given when an individual experiences a side effect
- These treatments can be related to chemotherapy administration by time
- In interpretation, need to consider:
  - "Individuals treated for a likely side effect"
  - individuals having these treatments for reasons other than side effects
  - individuals having side effects and not receiving these treatments
- Treatment of a side effect was considered related to chemotherapy when it occurred on or within three days after a chemotherapy dose

# Incidence of side effects - method

- An analysis dataset was generated for each side effect
- For each dose of chemotherapy dispensed, a search was done of any side effect treatments which were given to the same individual within 3 days
- The incidence was calculated by dose of chemotherapy, and then by individual

# Incidence of side effects - results

| | **Side effects** | **No. with chemotherapy** | **No. with side effect** | **% with side effect** |
|---|---|---|---|---|
| By doses | Diarrhoea | 89,594 | 879 | 1% |
| | Anaemia | 84,872 | 638 | <1% |
| | Nausea & vomiting | 84,378 | 5,415 | 6% |
| | Neutropenia | 84,495 | 601 | <1% |
| By person | Diarrhoea | 7,978 | 396 | 5% |
| | Anaemia | 8,158 | 330 | 4% |
| | Nausea & vomiting | 9,173 | 1,535 | 17% |
| | Neutropenia | 8,069 | 242 | 3% |

# Factors influencing side effects - methods

- Multiple regression used to identify factors which influence the incidence of each side effect
- Binary outcome, so logistic model required
- Correlated data noted
  - Can restructure data to remove correlation, using a summary measure (eg: ever had a side effect), or
  - Can use technique designed for correlated data, such as Generalised Estimating Equations (GEE)

# Generalised estimating equations

- Allow the correlation of outcomes within an individual to be estimated and taken into account in the regression coefficients and their standard errors
- The regression coefficients obtained from GEE are correctly interpreted in a population averaged manner
- Specifications of my GEE models
  - Repeated subject variable: PPN
  - Distribution: Binomial
  - Link function: Logit

# GEE Correlation structures

- Independent – simplest assumption, but usually incorrect
  - Each observation for an individual is uncorrelated with every other observation for that individual.
  - The GEE reduces to the independence (GLM) estimating equation
- Exchangeable (compound symmetry)
  - Every observation within an individual is equally correlated with every other observation from that individual.
  - Fully characterised by the intraclass correlation coefficient
- Auto-regressive
  - Derived from time series analysis
  - Two observations taken close in time within an individual tend to be more highly correlated that two observations taken far apart in time from the same individual.
- Others, inc unstructured and user fixed – more complicated and situation specific

# Factors influencing side effects - methods

$$side\ effect \sim \alpha + gender + age + RxRisk + chemo + cancer + \varepsilon$$

| **Variable** | **Levels** |
|---|---|
| **Side effect** | **Yes / No** |
| Gender | Male / Female |
| Age | Continuous, or<br><70 years<br>70 – 79 years<br>>79 years |
| RxRisk (comorbidities) | Quartiles (0-7, 8-9, 10-12, 13-26) |
| Chemo | Consolidated to 8 levels based on ATC code |
| Cancer | Consolidated to 7 levels based on ICD classification |

# Factors influencing side effects - models

- Tested correlation structures to maximise model fit with all variables at least aggregated level
  - Autoregressive consistently chosen as most appropriate
  - Indicates that there is correlation based on time as well as individuals
- Tested models with aggregated variable levels for age (continuous vs 4 levels) and chemotherapy category (2 categorisations each with 8 levels)
  - Model 1 (continuous age and standard chemo categories) most appropriate for ¾ side effects

# Summary of results

| **Variable** | **Diarrhoea** | **Nausea & vomiting** | **Anaemia** | **Neutropenia** |
|---|---|---|---|---|
| Gender (female) | ND | Increase*** | ND | ND |
| Age (younger) | Increase*** | Increase*** | ND | ND |
| RxRisk (fewer co-morbidities) | Decrease* | Decrease* | Decrease*** | Decrease** |
| | | | | * <0.05, **<0.01, ***<0.001 |

- Females are1.6 times more likely to experience N&V
- Every additional year of age decreases odds of diarrhoea by 4% and decreases odds of N&V by 3%
- Moving from highest to lowest RxRisk reduces odds of a side effect by 25% (N&V) to 60% (neutropenia)

# Summary of results

| **Variable** | **Diarrhoea** | **Nausea & vomiting** | **Anaemia** | **Neutropenia** |
|---|---|---|---|---|
| Breast cancer | ND | Decrease* | ND | Increase*** |
| Colorectal cancer | ND | ND | ND | Increase*** |
| Genital cancer | ND | ND | ND | Increase*** |
| Lung cancer | Decrease* | ND | ND | Increase*** |
| Non-solid tumours | Decrease* | Decrease*** | ND | Increase*** |
| Other | ND | ND | ND | Increase*** |

- Compared to urinary cancer: * <0.05, **<0.01, ***<0.001
  - diarrhoea odds were 70% lower in lung and 60% lower in non-solid cancers
  - N&V odds were reduced by nearly half in breast cancer and by over 60% in non-solid tumours
  - The increase of odds of neutropenia was highest for non-solid tumours (50-fold) and lung cancers (20-fold)

# Summary of results

| Variable | Diarrhoea | N&V | Anaemia | Neutropenia |
|---|---|---|---|---|
| Antineoplastic | Decrease*** | Increase*** | ND | Increase* |
| Progestogens | ND | Increase* | ND | ND |
| LHRH agnoists | Decrease*** | Increase*** | Decrease** | Increase*** |
| Anti-estrogens | Decrease* | Increase*** | ND | Increase*** |
| Anti-androgens | Decrease** | Increase*** | Decrease*** | Increase* |
| Aromatase inhibitors | Decrease* | ND | Decrease* | ND |
| Immunostimulants | ND | ND | ND | Increase*** |

* <0.05, **<0.01, ***<0.001

- Compared to immunosuppresants:
  - Antineoplastics lower odds of diarrhoea by over 70%
  - Anti-androgens increased odds of N&V by 13-fold
  - AIs decrease odds of anaemia by 84%
  - Immunostimulants increased odds of neutropenia by 700-fold

# Resource use - methods

$$Total\ cost \sim \alpha + gender + age + RxRisk + cancer + doses + any\ se + \varepsilon$$

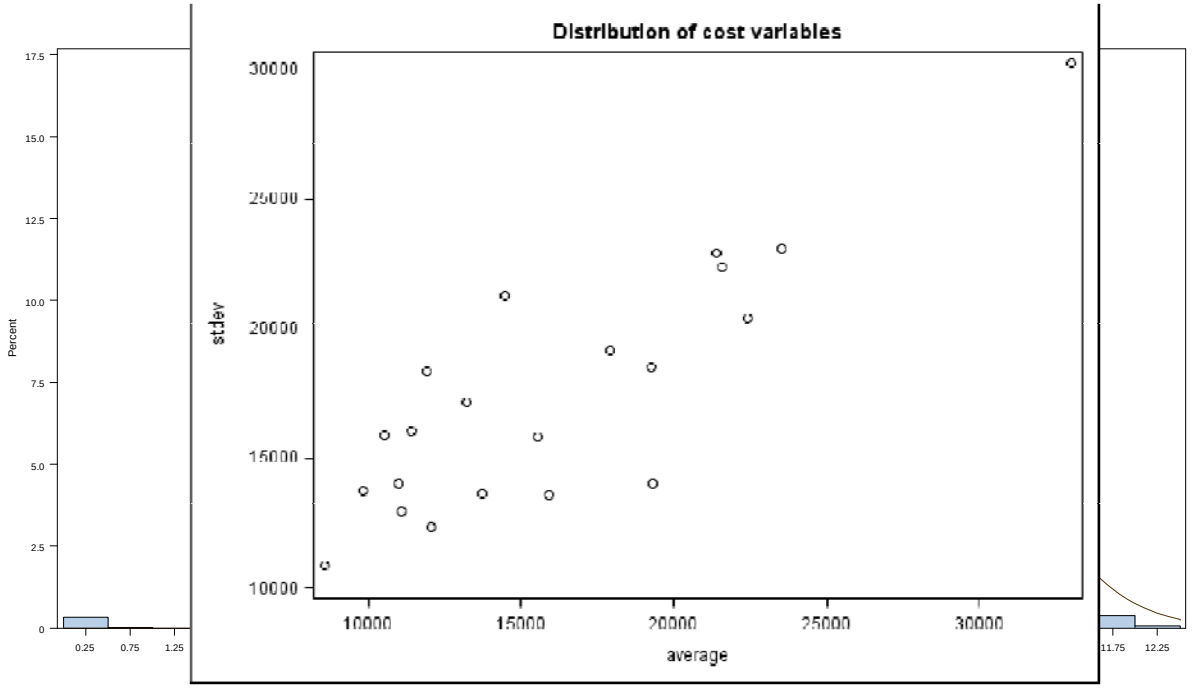
| **Variable** | **Levels** |
|---|---|
| Total cost | Total health care expenditure (medical services, hospitalisation &/or pharmaceuticals) during the 6-month period following the first dose of a new chemotherapy regimen from 1st Jan 2005 |
| Gender | Male / Female |
| Age | <70 years<br>70 – 79 years<br>>79 years |
| RxRisk | Quartiles (0-7, 8-9, 10-12, 13-26) |
| Doses | Total number of doses of chemotherapy (continuous) |
| Cancer | Consolidated to 7 levels based on ICD classification |
| Any side effect | Diarrhoea OR Anaemia OR N&V OR Neutropenia |

# Resource use- data distribution

- Cost data are typically positively skewed and truncated at zero, making parametric tests difficult
- Options include:
  - If large sample size, ignore skew (central limit theorem)
  - Non-paramatric tests – inappropriate for decision makers
  - Transform data – retransformation difficult
  - Non-parametric bootstrapping – simulation method, but doesn't model the skewness of the data
  - Generalised linear modelling – allows responses to be distributed in other ways (often gamma distribution is appropriate for cost data)

# Resource use - results

- Data highly skewed
- Log-transformed data approaches normal
- Mean vs standard deviation for raw costs shows an approximate constant coefficient of variation



# Resource use – raw cost results

| Solution for Fixed Effects - Simple linear regression of costs and each AE | | | | |
|---|---|---|---|---|
| **Effect** | **Category** | **Estimate** | **Standard Error** | **Pr > \|t\|** |
| **Intercept** | | 39705 | 3131.98 | <.0001 |
| **Sex (vs male)** | **Female** | -1418.69 | 599 | 0.0179 |
| **age** | | **-140.26** | 30.3976 | <.0001 |
| **RxRisk** | | **552.77** | 59.6786 | <.0001 |
| **Cancer site (vs urinary)** | **Breast** | **-4148.06** | 1299.15 | 0.0014 |
| | **CRC** | 616.02 | 1206.16 | 0.6096 |
| | **Genital** | **-3231.73** | 1097.67 | 0.0033 |
| | **Lung** | 237.14 | 1395.47 | 0.8651 |
| | **Non-solid** | **4655.44** | 1214.67 | 0.0001 |
| | **Other** | -2693.62 | 1150.71 | 0.0193 |
| **Any diarrhoea** | **No** | 2498.68 | 977.5 | 0.0106 |
| **Any nausea/vomit** | **No** | **-7511.1** | 543.34 | <.0001 |
| **Any anaemia** | **No** | **-4724.43** | 1042.62 | <.0001 |
| **Any neutropenia** | **No** | **-10631** | 1141.47 | <.0001 |

# Resource use – log-transformed results

| Solution for Fixed Effects - Regression of log costs – each AE | | | | |
|---|---|---|---|---|
| **Effect** | **Category** | **Estimate** | **Standard Error** | **Pr > \|t\|** |
| **Intercept** | | 10.2124 | 0.2335 | <.0001 |
| **Sex (vs male)** | **Female** | **-0.2062** | 0.04466 | <.0001 |
| **age** | | -0.00565 | 0.002266 | 0.0127 |
| **RxRisk** | | **0.06941** | 0.004449 | <.0001 |
| **Cancer site (vs urinary)** | **Breast** | **-0.3471** | 0.09686 | 0.0003 |
| | **CRC** | -0.077 | 0.08992 | 0.3919 |
| | **Genital** | -0.1911 | 0.08184 | 0.0195 |
| | **Lung** | -0.167 | 0.104 | 0.1084 |
| | **Non-solid** | 0.1749 | 0.09056 | 0.0535 |
| | **Other** | **-0.3751** | 0.08579 | <.0001 |
| **Any diarrhoea** | **No** | -0.01491 | 0.07288 | 0.8379 |
| **Any nausea/vomit** | **No** | **-0.5665** | 0.04051 | <.0001 |
| **Any anaemia** | **No** | **-0.3472** | 0.07773 | <.0001 |
| **Any neutropenia** | **No** | **-0.5458** | 0.0851 | <.0001 |

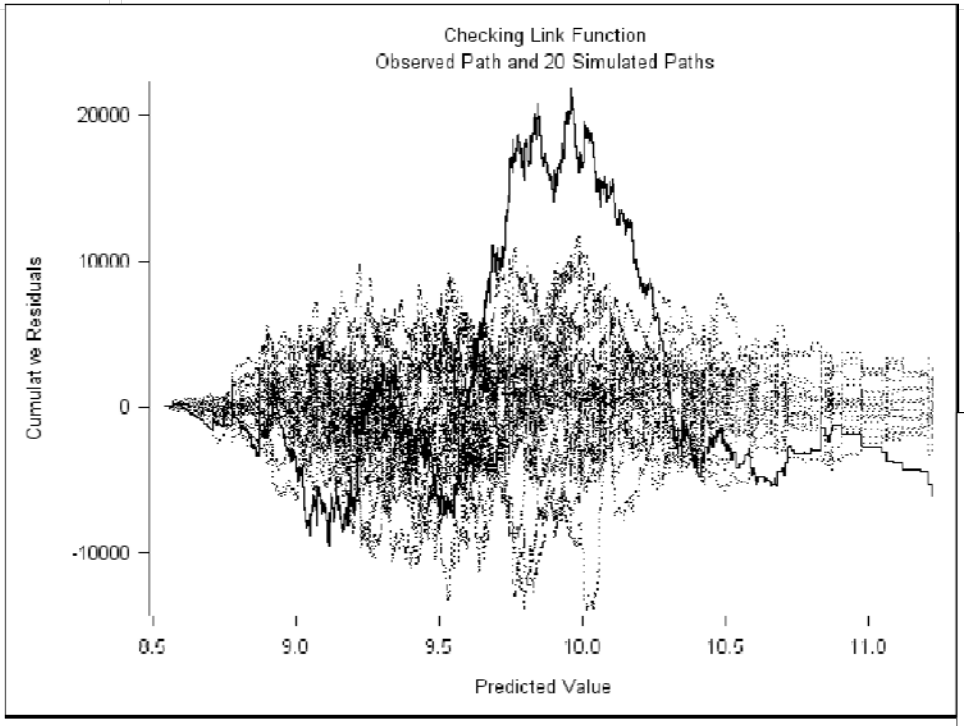
# Resource use – GLM results

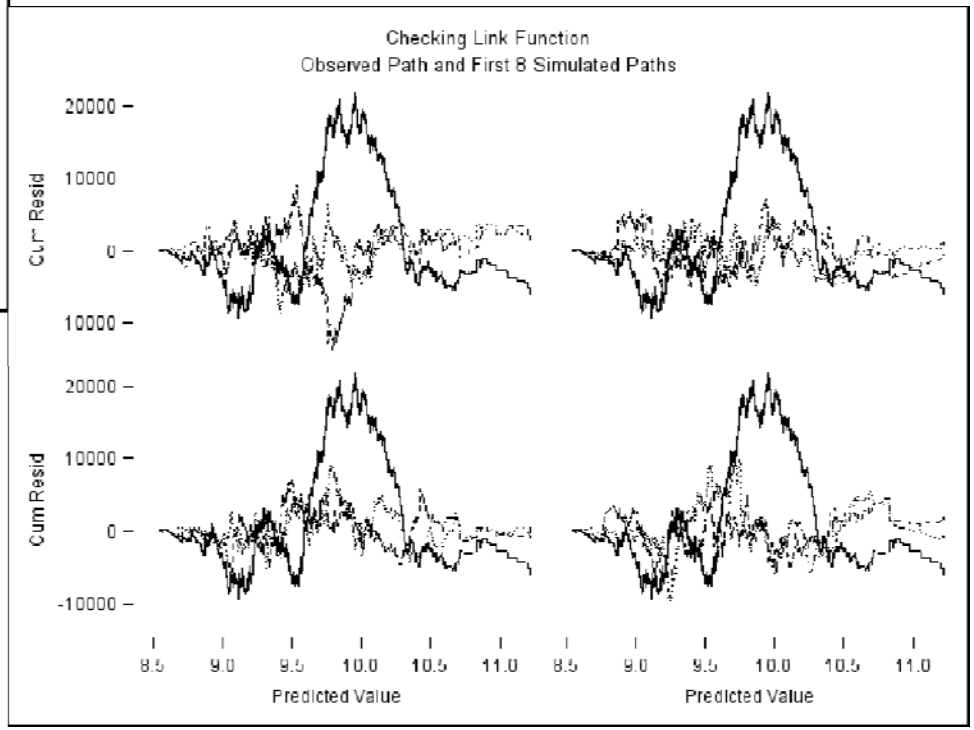
| Parameter | Category | Exp (Estimate) | Exp (Wald 95% Confidence Limits) | |
|---|---|---|---|---|
| **Intercept** | | 14237.01 | 10515.44 | 19273.76 |
| **Sex** | **F** | **0.91** | 0.84 | 0.97 |
| **age** | | **0.99** | 0.99 | 1.00 |
| **RxRisk** | | **1.05** | 1.04 | 1.06 |
| **sitecatb** | **Breast** | **0.67** | 0.59 | 0.75 |
| **sitecatb** | **Genita** | **0.76** | 0.70 | 0.83 |
| **sitecatb** | **Lung** | **1.10** | 0.96 | 1.25 |
| **sitecatb** | **Nosoli** | 1.25 | 1.12 | 1.39 |
| **sitecatb** | **Other** | **0.76** | 0.69 | 0.83 |
| **sitecatb** | **Urinar** | **1.01** | 0.87 | 1.16 |
| **anydia** | **1** | 0.89 | 0.79 | 1.00 |
| **anynausea** | **1** | **1.61** | 1.51 | 1.72 |
| **anyanaemia** | **1** | **1.33** | 1.18 | 1.51 |
| **anyneut** | **1** | **1.54** | 1.34 | 1.76 |
| **Scale** | | 2.95 | 2.85 | 3.06 |

# Resource use – GLM results

- Test model
  - Plot cumulative residuals to assess fit of covariates or appropriateness of link function
  - Assesses if the simulated residual patterns (with a log-link) that would be generated by the model under the specified assumptions are statistically different from the one actually generated

# Resource use – GLM results



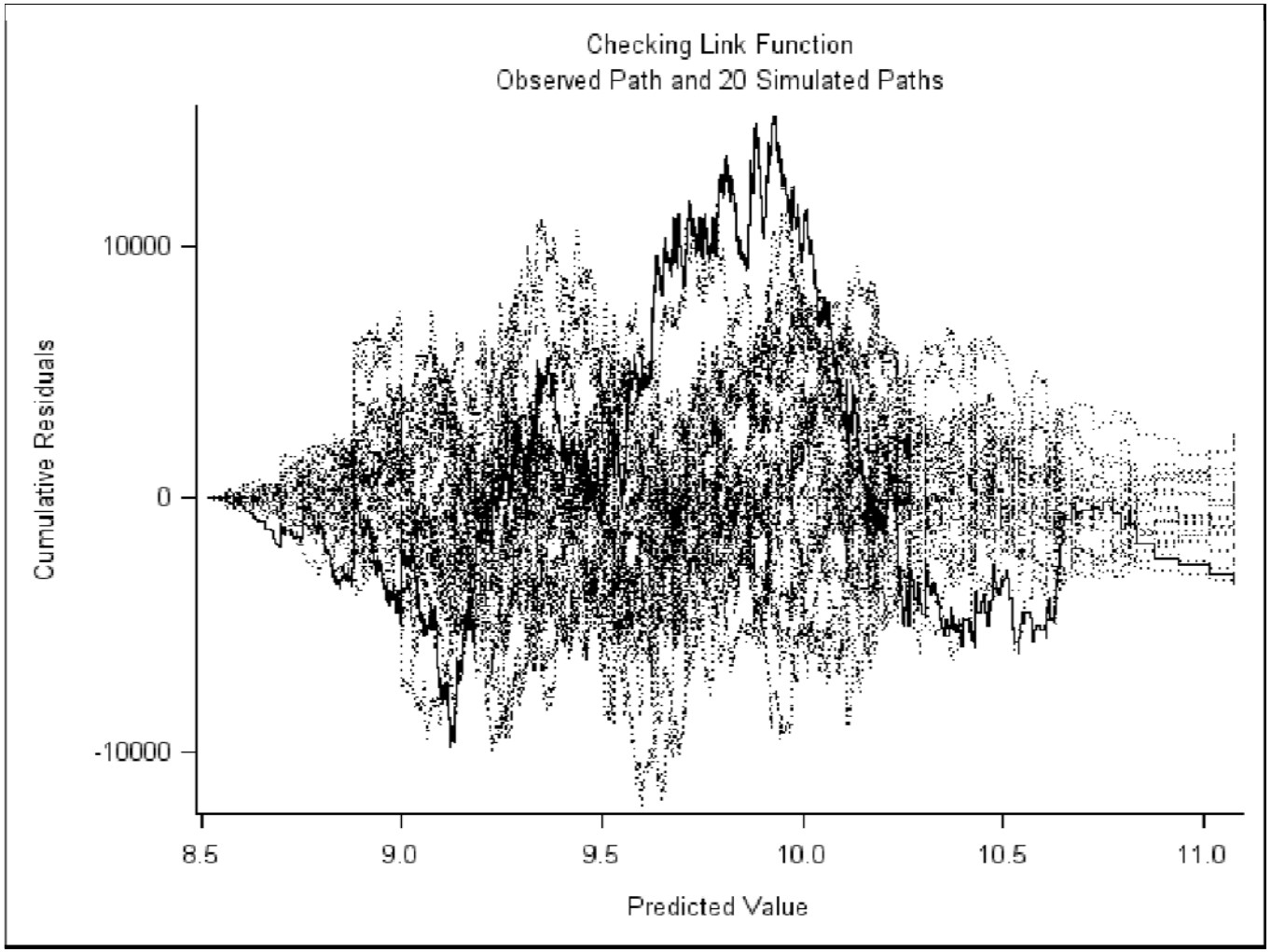
# Resource use – GLM results

- Plots indicate an artefact in the data
- Exploratory analysis of model with interaction terms
  - Between side effects
    - 2/3 anaemia interactions were significant at $p<0.05$ level
  - Between the type of cancer and the side effect
    - Nausea had the strongest association with type of cancer
  - Between age and comorbidities
    - Not significant

# Resource use – GLM results

- Final model included:
  - Main effects
  - Interaction term for anaemia and other side effects
  - Interaction term for nausea and cancer type
- Little impact on the significance of the main effects on total cost
- A number of interaction terms appear to significantly influence total cost
- Inclusion of interaction terms appears to improve model fit

# Resource use – GLM results



# Conclusions

- This large administrative dataset provides an opportunity to examine 'real life' incidence of chemotherapy side effects in older people
- Being treated for a likely side effect is more common in individuals who are older or who have more co-morbidities
- Being treated for a likely side effect may be influenced by the type of cancer and chemotherapy an individual has
- Being treated for a likely side effect significantly increases overall healthcare costs

# Acknowledgements


- This work was funded through:
  - UTS Doctoral Scholarship
  - EMCaP PhD Scholarship
- This work was supported by:
  - Department of Veterans Affairs
  - *This conference presentation has been reviewed by DVA prior to presentation and the views expressed are not necessarily those of the Australian Government*
- With special thanks to:
  - Sallie-Anne Pearson and Preeyaporn Srasuebkul
  - Marion Haas and Rosalie Viney

# THE INCIDENCE AND COSTS OF CHEMOTHERAPY SIDE EFFECTS

Alison Pearce - PhD candidate Centre for Health Economics Research and Evaluation, UTS

Supervisors: Marion Haas, Rosalie Viney

CAER 2013

# Spare slides

# Data issues – dataset size

- Large number of individuals, each with multiple observations
- 63Gb dataset for PBS data alone
- Data management and analysis techniques to manage this included:
  - Network location to reduce processing time
  - Removal of text variables to reduce data size
  - Use of SQL programming to improve efficiency

| GEE Parameter Estimates (Odds Ratios) | | | | |
|---|---|---|---|---|
| **Parameter** | | **Diarrhoea** | **N&V** | **Anaemia** |
| **Intercept** | | 0.04*** | 0.06*** | 0.03*** |
| **Gender** | Female | 1.00 | 1.41*** | 0.81 |
| | Male | . | . | . |
| **Age** | < 70 yrs | 0.74 | 0.58*** | 0.26*** |
| | 70 - 80 yrs | 1.35* | 1.08 | 0.94 |
| | > 80 Yrs | . | . | . |
| **Rx Risk** | 0 to 7 | 0.57*** | 0.81* | 0.49*** |
| | 8 to 9 | 0.70 | 1.09 | 0.77 |
| | 10 to 12 | 0.79 | 1.20 | 0.81 |
| | 13 + | . | . | . |
| **Cancer** | Breast | 0.54 | 0.83 | 0.69 |
| | Colorectal | 2.38* | 1.39 | 0.95 |
| | Genital | 0.64 | 0.99 | 1.18 |
| | Lung | 0.39* | 2.49*** | 0.86 |
| | Non-solid | 0.29*** | 0.82 | 1.96 |
| | Other | 0.77 | 0.82 | 0.69 |
| | Urinary | . | . | . |
| **Chemo** | Antineoplastic | 0.58 | 2.87*** | 0.67 |
| | Progestogens | 0.49 | 1.96 | 0.41 |
| | LHRH agonist | 0.25*** | 0.28*** | 0.29* |
| | Anti-estrogens | 0.47 | 0.40*** | 0.60 |
| | Anti-androgens | 0.33* | 0.46*** | 0.26* |
| | Aromatase inhibitors | 0.45 | 0.66 | 0.18 |
| | Immunostimulants | 0.53 | 1.20 | 1.49 |
| | Immunosuppresants | . | . | . |

# Comparing correlation structures

| **GEE Model Information** | | | |
|---|---|---|---|
| Correlation Structure | **Exchangeable** | **Independent** | **AR(1)** |
| Subject Effect | ppn (7976 levels) | ppn (7976 levels) | ppn (7976 levels) |
| Number of Clusters | 7976 | 7976 | 7976 |
| Clusters With Missing Values | 5 | 5 | 5 |
| Correlation Matrix Dimension | 151 | 151 | 151 |
| Maximum Cluster Size | 151 | 151 | 151 |
| Minimum Cluster Size | 1 | 1 | 1 |
| **Algorithm converged** | Yes | Yes | Yes |
| **GEE Fit Criteria** | | | |
| QIC | 9450.3736 | 9433.6513 | 9429.4868 |
| QICu | 9378.4911 | 9320.2626 | 9348.8765 |
| **Exchangeable Working Correlation** | | | |
| Correlation | 0.151245715 | . | . |

# Generalised Estimating Equations

- A GEE model extends generalised linear models in 2 ways:
  - Allows the correlation of outcomes within an individual to be estimated and taken into account in the regression coefficients and their standard errors
  - Permits the calculation of *robust* estimates for the standard errors (empirical standard errors) of the regression coefficients
    - ensures consistent inferences even if the correlation structure is incorrect, or the strength of the correlation varies between people
- *"Main benefit is that besides the seeking of more efficient estimators of regression parameters, the GEE produces reasonably accurate standard errors, hence confidence intervals with the correct coverage rates."*

# GEE continued

- GEE can deal with different number of observations per person (ie no need for designs to be balanced)
- GEE ideal when it is the fixed regression parameters which are of primary interest, and the correlation structure is merely a nuisance
- GEE not suitable if it is the correlation structure that you wanted to model (better to use multi-level modelling)
- Typically used in epidemiology and health, with responses that are binomially distributed or Poisson distributed.

# Generalised Estimating Equations

- Allow the correlation of outcomes within an individual to be estimated and taken into account in the regression coefficients and their standard errors
- The correlation structure is pre-specified
- Despite incorrect working correlation structure the resulting regression coefficient estimate is still consistent
- Can deal with different number of observations per person
- Produces reasonably accurate standard errors, and therefore confidence intervals with the correct coverage rates

# Steps to a GEE

1. Fit a standard regression model - assuming all observations are independent
2. Take the residuals from the regression and use these to estimate the parameters which quantify the correlation between observations in the same individual
3. Refit the regression model using a modified algorithm incorporating a matrix which reflects the magnitude of the correlation estimated in step 2
4. Keep alternating between steps 2 and 3 until the estimates all stabilize (ie model converges)

# Interpreting GEE results

- The regression coefficients obtained from GEE are correctly interpreted in a population averaged manner
  - This is same as regular logit (specified for cluster)
  - Ie: Estimates the odds of the average male getting diarrhoea compared to the odds of the average female getting diarrhoea.
  - Different to random effects logit, which calculates the individual effect, ie the odds of a person having diarrhoea if male compared to the odds of the same person getting diarrhoea if female
- In many cases the population averaged and subject specific estimates are very close, but not always

# GEE References


- Burton, Gurrin & Sly. Extending the simple linear regression model to account for correlated responses: an introduction to generalised estimating equations and multi-level mixed modelling. Statistics in Medicine. 1998. 17:1261-1291
- Sribney. What is the difference between random effects and population-averaged estimators. 1999 (revised 2007). STATA support
- Hanley, Negass, Edwards, Forrester. Statistical analysis of correlated data using generalised estimating equations: An orientation. American journal of epidemiology. 2003, 157(4):364-375
- Pan & Connett. Selecting the working correlation structure in generalised estimating equations with application to the lung health study. Statistics Sinica. 2002, 12:475-490
- Liang & Zeger. Longitudinal data analysis using generalised linear models. Biometrics. 1986, 78(1):13-22

# Change data structure

- Summary statistics can be used to address the correlation
  - Changed data structure so only one observation per patient
  - Summary statistic ‘ever diarrhoea’ (assoc with chemo)
  - Replace chemotherapy details with number of doses of chemotherapy
  - Use simple logistic regression
- Removes potential patient correlation
- Similar analysis to that used in clinical trials
- Model converges

# Logistic regression with summary statistic – model fit

| Model Fit Statistics | | |
|---|---|---|
| **Criterion** | **Intercept Only** | **Intercept and Covariates** |
| **AIC** | 3152.002 | 2862.676 |
| **SC** | 3158.986 | 2960.451 |
| **-2 Log L** | 3150.002 | 2834.676 |

| Testing Global Null Hypothesis: BETA=0 | | | |
|---|---|---|---|
| **Test** | **Chi-Square** | **DF** | **Pr > ChiSq** |
| **Likelihood Ratio** | 315.3261 | 13 | <.0001 |
| **Score** | 411.3727 | 13 | <.0001 |
| **Wald** | 323.4319 | 13 | <.0001 |

# Logistic regression with summary statistic

| Odds Ratio Estimates | | | |
|---|---|---|---|
| **Effect** | **Point Estimate** | **95% Wald Confidence Limits** | |
| **Sex Female vs Male** | 1.137 | 0.852 | 1.519 |
| **Age <70 vs 80+** | 1.271 | 0.962 | 1.678 |
| **Age 70 – 79 vs 80+** | 1.792* | 1.387 | 2.316 |
| **RxRisk 0-7 vs 13+** | 0.439* | 0.325 | 0.595 |
| **RxRisk 8-9 vs 13+** | 0.723* | 0.532 | 0.981 |
| **RxRisk 10-12 vs 13+** | 0.795 | 0.599 | 1.054 |
| **Cancer site Breast vs Urinary** | 0.719 | 0.361 | 1.433 |
| **Cancer site CRC vs Urinary** | 2.979* | 1.605 | 5.526 |
| **Cancer site Genital vs Urinary** | 0.679 | 0.365 | 1.263 |
| **Cancer site Lung vs Urinary** | 0.627 | 0.275 | 1.428 |
| **Cancer site Nosoli vs Urinary** | 0.678 | 0.34 | 1.35 |
| **Cancer site Other vs Urinary** | 0.817 | 0.426 | 1.567 |
| **Chemotherapy doses** | 1.035* | 1.03 | 1.04 |

# Logistic regression with a summary statistic

- While this method has advantages, it is not always ideal
- It requires use of a summary statistic, which may not answer the question
- It is inefficient, as it only uses some of the data in the dataset
- Generalised Estimating Equations can be used when a simple logistic regression would be appropriate, but there is correlation in the data, such as repeated measures

| GEE Parameter Estimates - Nausea & Vomiting | | | | | | |
|---|---|---|---|---|---|---|
| **Parameter** | | **Estimate (Odds Ratio)** | **Standard Error** | **95% Confidence Limits** | | **Pr > \|Z\|** |
| **Intercept** | | 0.06*** | 1.32 | 0.03 | 0.10 | <.0001 |
| **Gender** | Female | 1.41*** | 1.09 | 1.19 | 1.67 | <.0001 |
| | Male | . | . | . | . | . |
| **Age** | < 70 yrs | 0.58*** | 1.11 | 0.47 | 0.71 | <.0001 |
| | 70 - 80 yrs | 1.08 | 1.07 | 0.95 | 1.22 | 0.23 |
| | > 80 Yrs | . | . | . | . | . |
| **Rx Risk** | 0 to 7 | 0.81* | 1.11 | 0.66 | 0.99 | 0.04 |
| | 8 to 9 | 1.09 | 1.11 | 0.89 | 1.34 | 0.40 |
| | 10 to 12 | 1.20 | 1.10 | 0.99 | 1.45 | 0.07 |
| | 13 + | . | . | . | . | . |
| **Cancer** | Breast | 0.83 | 1.31 | 0.49 | 1.40 | 0.48 |
| | Colorectal | 1.39 | 1.25 | 0.90 | 2.16 | 0.14 |
| | Genital | 0.99 | 1.26 | 0.63 | 1.55 | 0.96 |
| | Lung | 2.49*** | 1.26 | 1.59 | 3.91 | <.0001 |
| | Non-solid | 0.82 | 1.25 | 0.53 | 1.28 | 0.39 |
| | Other | 0.82 | 1.26 | 0.53 | 1.29 | 0.39 |
| | Urinary | . | . | . | . | . |
| **Chemo** | Antineoplastic | 2.87*** | 1.19 | 2.04 | 4.02 | <.0001 |
| | Progestogens | 1.96 | 1.45 | 0.95 | 4.06 | 0.07 |
| | LHRH agonists | 0.28*** | 1.23 | 0.19 | 0.42 | <.0001 |
| | Anti-estrogens | 0.40*** | 1.29 | 0.24 | 0.66 | 0.00 |
| | Anti-androgens | 0.46*** | 1.26 | 0.29 | 0.72 | 0.00 |
| | Aromatase inhibitors | 0.66 | 1.30 | 0.39 | 1.11 | 0.11 |
| | Immunostimulants | 1.20 | 1.47 | 0.56 | 2.56 | 0.64 |
| | Immunosuppresants | . | . | . | . | . |

| GEE Parameter Estimates - Anaemia | | | | | | |
|---|---|---|---|---|---|---|
| **Parameter** | | **Estimate (Odds Ratio)** | **Standard Error** | **95% Confidence Limits** | | **Pr > \|Z\|** |
| **Intercept** | | 0.03*** | 1.85 | 0.01 | 0.09 | <.0001 |
| **Gender** | Female | 0.81 | 1.23 | 0.54 | 1.21 | 0.30 |
| | Male | . | . | . | . | . |
| **Age** | < 70 yrs | 0.26*** | 1.41 | 0.13 | 0.51 | <.0001 |
| | 70 - 80 yrs | 0.94 | 1.16 | 0.70 | 1.26 | 0.68 |
| | > 80 Yrs | . | . | . | . | . |
| **Rx Risk** | 0 to 7 | 0.49*** | 1.22 | 0.33 | 0.72 | 0.00 |
| | 8 to 9 | 0.77 | 1.22 | 0.53 | 1.13 | 0.18 |
| | 10 to 12 | 0.81 | 1.20 | 0.56 | 1.15 | 0.24 |
| | 13 + | . | . | . | . | . |
| **Cancer** | Breast | 0.69 | 2.10 | 0.16 | 2.94 | 0.61 |
| | Colorectal | 0.95 | 1.58 | 0.39 | 2.33 | 0.91 |
| | Genital | 1.18 | 1.55 | 0.50 | 2.80 | 0.71 |
| | Lung | 0.86 | 1.61 | 0.34 | 2.19 | 0.76 |
| | Non-solid | 1.96 | 1.52 | 0.86 | 4.45 | 0.11 |
| | Other | 0.69 | 1.58 | 0.28 | 1.69 | 0.42 |
| | Urinary | . | . | . | . | . |
| **Chemo** | Antineoplastic | 0.67 | 1.66 | 0.25 | 1.81 | 0.43 |
| | Progestogens | 0.41 | 2.36 | 0.08 | 2.19 | 0.30 |
| | LHRH agonists | 0.29* | 1.70 | 0.10 | 0.81 | 0.02 |
| | Anti-estrogens | 0.60 | 2.04 | 0.15 | 2.42 | 0.47 |
| | Anti-androgens | 0.26* | 1.72 | 0.09 | 0.75 | 0.01 |
| | Aromatase inhibitors | 0.18 | 2.71 | 0.03 | 1.27 | 0.09 |
| | Immunostimulants | 1.49 | 1.80 | 0.47 | 4.74 | 0.50 |
| | Immunosuppresants | . | . | . | . | . |

# Interpreting results – diarrhoea

- No difference in the odds of the average male getting diarrhoea compared to the odds of the average female getting diarrhoea
- Moving from the highest RxRisk category to the lowest, the average person will experience a 43% reduction in the odds of getting diarrhoea
- Compared to urinary cancer, the average person with CRC is 2.4 times more likely to experience diarrhoea.
- For the average person, the use of LHRH agonists or anti-androgens will reduce the odds of diarrhoea, compared to immunosuppressant's

# Interpreting results – Nausea & vomiting

- The average female is 1.4 times more likely to experience N&V than the average male
- The average individual in the younger age group has a 40% lower odds of experiencing N&V
- Compared to urinary cancer, the average person with lung cancer is more likely to experience N&V
- The average person on antineoplastic chemotherapy is almost 3 times more likely to experience N&V than the average person on immunosuppresants
- LHRH agonists, anti-estrogens and anti-adrogens all decrease the odds of N&V in the average person

# Interpreting results - anaemia

- No difference based on gender
- The younger age group has an 80% reduction in the odds of having anaemia
- Moving from the highest RxRisk category to the lowest, the average person will experience a 50% reduction in the odds of getting diarrhoea
- Type of cancer does not appear to have a significant impact on the experience of anaemia
- LHRH-agonists and anti-androgens reduce the odds of anaemia compared to immunosuppressant's

# What next?

- Model impact of having one AE on odds of having additional AE(s)
  - Correlation analysis
  - Two-part model
- Incorporate results into decision models and examine differences in outcomes
- Questions?