


Health and Care Conference 2012

Peter Banthorpe, Jennifer Loftus and Chris Reynolds



Does where you live influence your critical illness claim rates?

1st May 2012

Disclaimer

- The views expressed today are those of the presenters and do not necessarily reflect those of their employers, and thus, their employers accept no liability as a result of any reliance you may have placed or action taken based upon the information outlined in this document / presentation

Overview

- **Introduction to the Working Party**
 - High Level Aims
- **HES Episode Data**
 - HES Data requests / Issues encountered
- **HES Data Analysis – Initial Findings**
 - Mosaic, ACORN, HealthACORN
 - IMD Quintile, Government Office Region.

Geographical Variations CI Working Party

- **Working Party formed in October 2010**
- **Members**
 - Jennifer Loftus
 - Ketiwe Nhende
 - Christopher Reynolds
 - Daniel Ryan
 - Christine Fairall
 - Peter Banthorpe

High Level Aims of Working Party

- **Analysis of the impact on CI Rates of proxy rating factors:**
 - Location
 - Socio-Economic Profile
 - Interactions thereof
- **Power of using these proxies for modelling mortality in the UK has already been proven**
 - Can we use the same proxies for modelling CI incidence rates?
- **Key CI Conditions:**
 - Cancer, Heart Attack, Stroke
- **We expect our results will be of interest to:**
 - Actuaries
 - Underwriters – Basis of a simplified underwriting tool?
 - Marketing and product development specialists
 - Healthcare professionals

Last Year's Presentation

- **HES Data Request Submitted**
- **Initial Analysis**
 - Health Poverty Index
 - Cancer E-Atlas/Deprivation Scores

Demographic Profilers – CACI & Experian

- **Leading providers of socio-demographic profiles in the UK (amongst six providers in the UK market)**
- **CACI**
 - Main segmentation tool is ACORN, classifies population into 5 categories, 17 groups and 56 types
 - Other ACORN indices exist, we are also looking at HealthACORN - 4 groups, 25 types
- **Experian**
 - Main segmentation tool is MOSAIC, classifies population into 15 categories, 67 groups and 141 types



Hospital Episodes Data

What the HES data looks like

Patient Identifier	Unique identifier by patient – 47m of these
Basic Patient Information	Age, gender
Basic Episode Information	Date started, date finished, admission method, current status etc
Diagnosis Information	Up to 20 different diagnoses
Procedure Information	Up to 20 different operations, with date of operation
Geographical Information	Postal district, Lower Super Output Area, IMD Rank, Mosaic Type, ACORN Type, Health ACORN type

Example data

DIAG_01	DIAG_02	DIAG_03	DIAG_04	DIAG_05	DIAG_06	DIAG_07	DIAG_08	DIAG_09	DIAG_10	DIAG_11	DIAG_12	DIAG_13	DIAG_14
I423-	J439-	Z539-											
I499-	-												
I270-	I424-	I270-	J439-	I48X-	E039-								
S6250D	W198-A												
J181-													
J181-													
I423-	I48X-	I501-	J439-										
I423-	I48X-	I501-	J439-										
I48X-	I424-	J439-	I959-										
I270-	I509-	J449-	J90X-	J439-	I48X-								
I423-	I48X-	I501-	J439-										
R14X-	R060-	I423-	I48X-	I501-									
E876-	I500-	I48X-	E058-	Y522-									
J22X	I423	J439											
K621													
K602	E059	I424	I500										
I423	I48X	I10X	J439										
I424-	I48X-	I500-	I270-										
L918	D227												
I500-	S220-	X599-	I270-	I423-	I439-	J459-	Z950-	Z921-					
I424-	I500-	I270-	I081-	I48X-	E032-	Y522-	J439-	J459-					
K602	E059	I424	I500										
R074	I48X	J439	E039	J841	I423								
I423	I48X	I10X	J439										
I424-	I500-	I48X-	J439-	Z877-									
I425	Z048												
I424-	Z450-	I500-	I48X-	T462-	Y522-	E079-	Z867-						
J22X-	R060-	T818-	Y831-	Z950-									
I500-	I429-	I48X-	I270-	J439-	K602-	K625-							
I424-	I48X-	I500-	T462-	Y522-	E079-	Z950-	Z867-						
M1097	I10X	I48X	I500	E059	J449								
I500-	I429-	I48X-	I270-	J439-	K602-	K625-							
I500-	I424-	I48X-	I270-	Z950-									
I500-	I424-	I48X-	I270-										
E876-	I424-	I517-	I48X-	I270-	Z950-								
E876-	I424-	I517-	I48X-	I270-	Z950-								
I081-	I424-	I517-	I48X-	I270-	I959-	J439-	M8190	Z950-					
I424-	I500-	I48X-	I270-	L52X-	L270-	K148-	R061-	Y041-	R55X-	I10X-	J439-	M8199	Z950-

20 Diagnosis codes

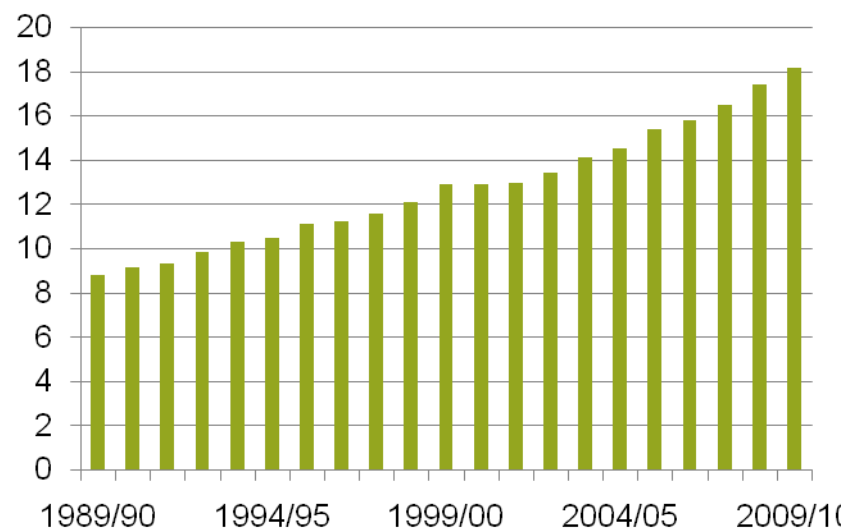
Each record is an individual episode

Each ICD code could appear multiple times as a primary diagnosis (DIAG_01) or in secondary diagnosis fields

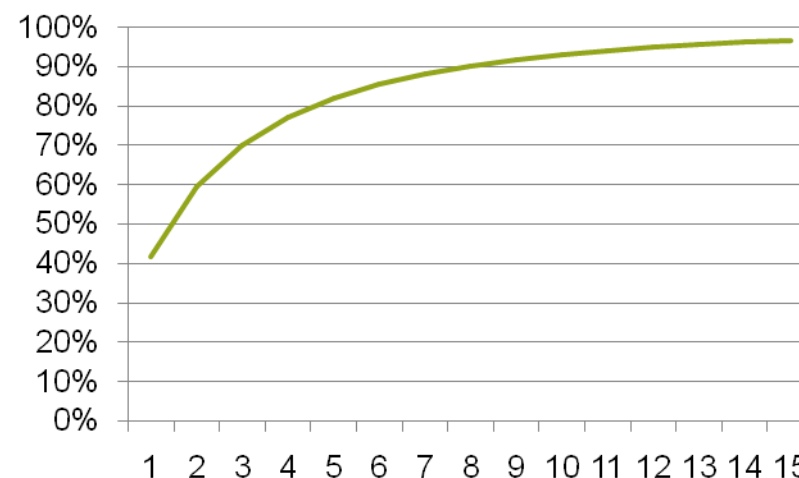
Summary of HES Data

- Data years 1989/90 to 2009/10 received.
- Only 1997/98 to 2009/10 are coded with unique patient identifiers.

Records Per Year - millions



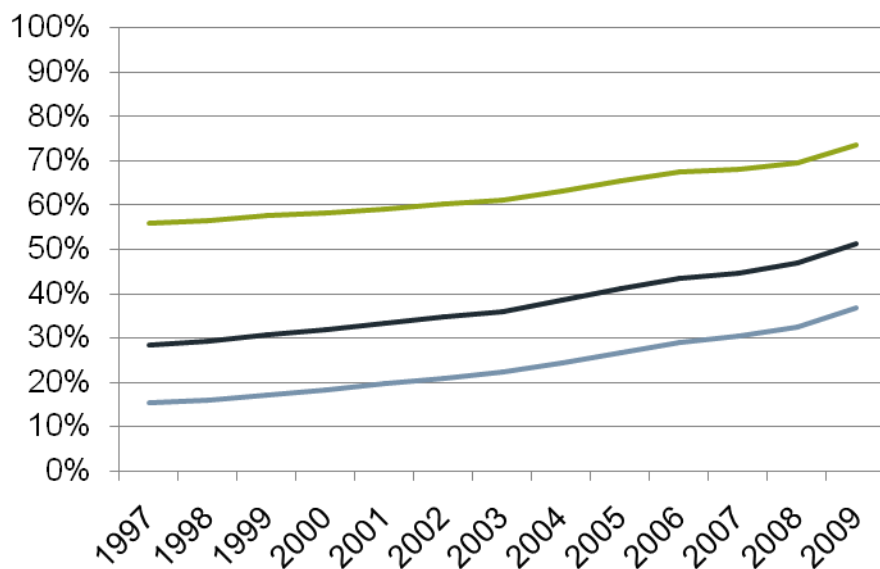
Distribution of Episodes Per HES ID



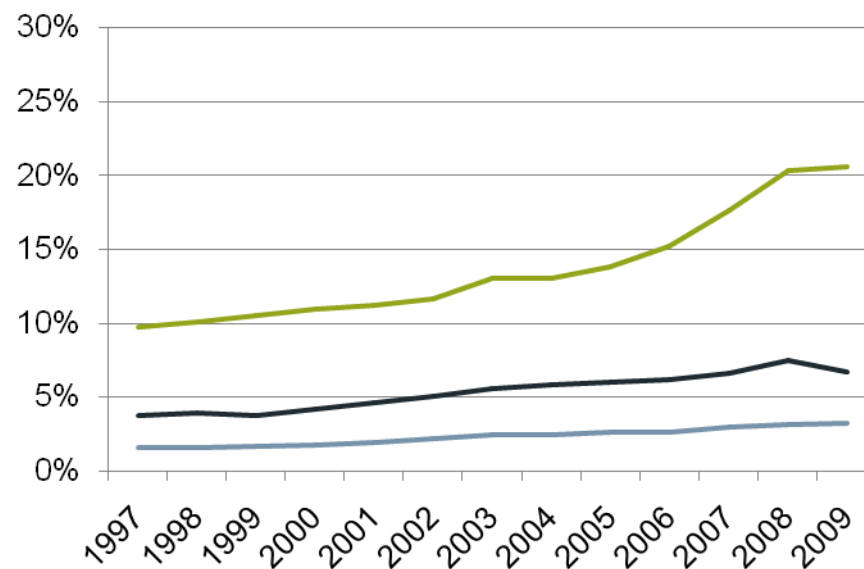
Increases in Multiple Codings

Proportion of Episodes with 2nd, 3rd, 4th Diagnosis fields being populated by year.

Any ICD Code



First Heart Attack (I21)



Data checks and Reconciliations

- Check of values in each field
 - Key fields have values corresponding to HES Data dictionary;
 - Some minor fields have unknown values;
- Reconciliations to freely available data
 - Primary Diagnosis, Finished Consultant Episodes;
 - Main procedures and interventions;
 - Total Procedures and interventions

ICD Codes Used

- Cancer, Heart Attack, Stroke reported here.
- Relevant ICD10 codes as per “Exploring the Critical Path”
 - Except I22 (Subsequent Heart Attack) codes are not used for Heart Attack.

Identifying Conditions within an Episode

- All 20 Diagnosis fields are searched for relevant codes.
- Routine works sequentially through relevant ICD codes per illness
 - Multiple relevant ICD Codes per episode results in last ICD code being allocated, for example:

Diag_01	Diag_02	Diag_03	Diag_04	Diag_05		Cancer
C01X-	C770-	C792-	Z511-	Z864-		C792-

- Not an issue for current analysis....
-but will be when looking at individual cancer sites.



HES Data - Initial Findings

Initial Analysis of HES Data

Geo-Demographic Profilers

- Processed data consists of:

Claims	Population
Cause : Cancer, HA & Stroke	
Age banded	Age banded
Gender	Gender
Geodemographic Code	Geodemographic Code
Years : 2006 - 2009	2010 (Mosaic), 2009 (CACI)

- Cancer (6.3m), Heart Attack (0.7m), Stroke (0.8m)
- Mosaic (67), ACORN (56), HealthACORN (25)
- How should we group these Geodemographic types?

Initial Analysis of HES Data

- 2 Types of Analysis

Qualitative

- Top Down Analysis. Group categories based on the geo-demographic descriptions.

Quantitative

- Bottom-up Analysis. Group categories based on homogeneity of incidence rate results.

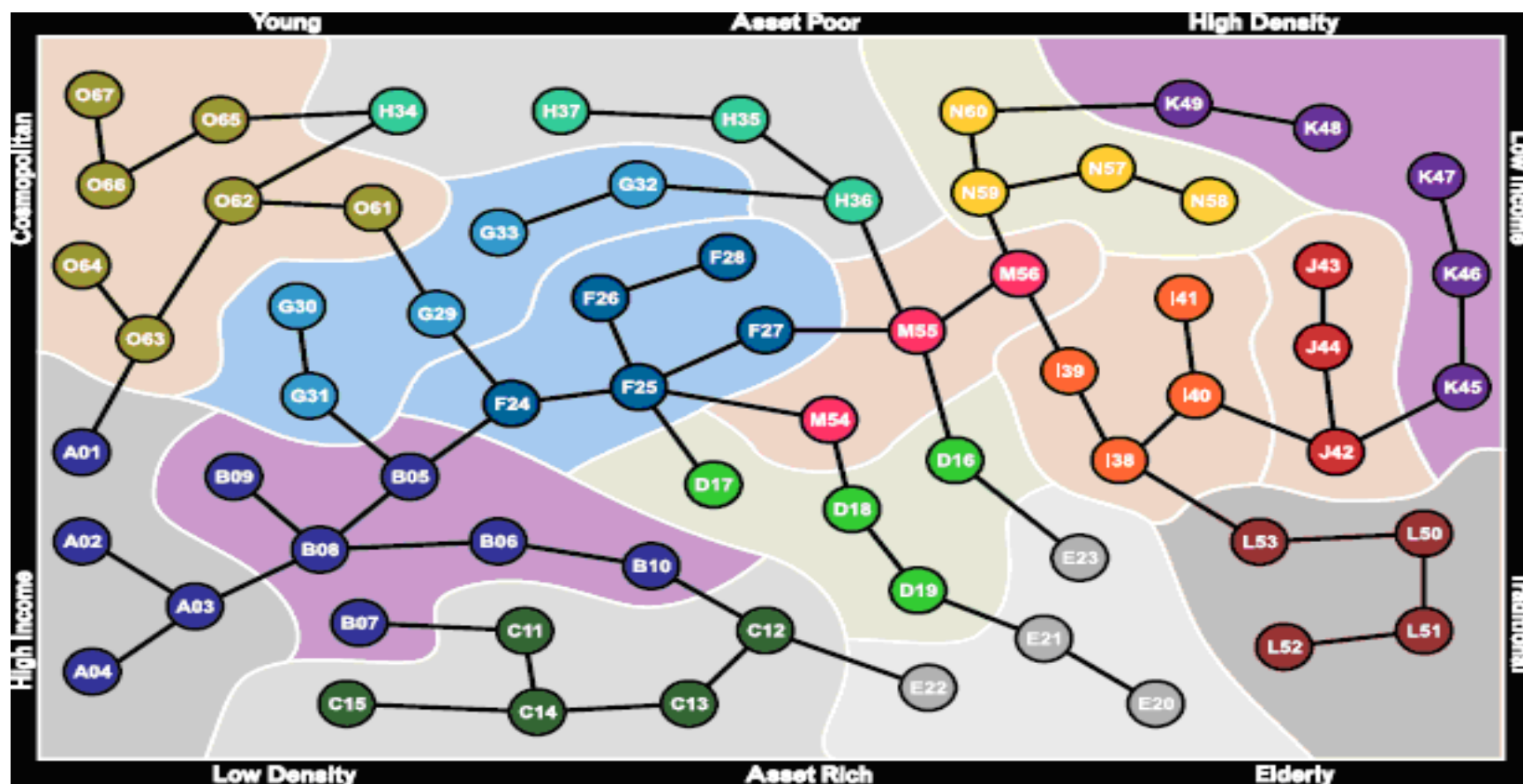
Initial Analysis of HES Data

Bottom Up Analysis

- Use HES Heart Attack Data
- Calculate “claims” and exposure split by
 - Age Banding
 - Geodemographic type
- Determine incidence rates
- Standardise using a mix provided by the CMI
- Group together homogenous results

Demographic Profilers – Experian

- MOSAIC Family Tree

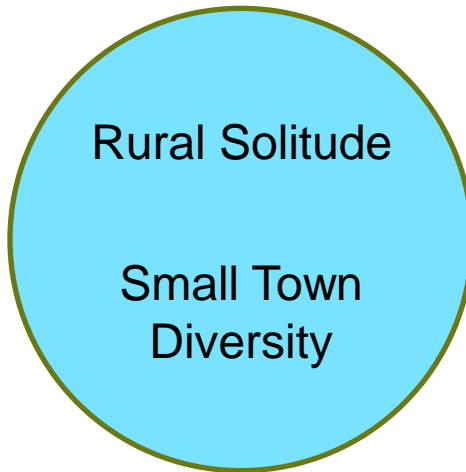


Top Down Analysis Mosaic

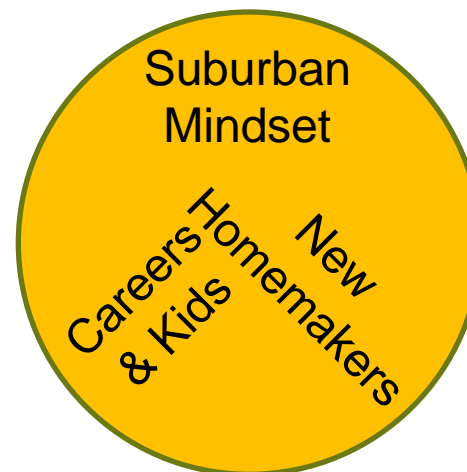
1



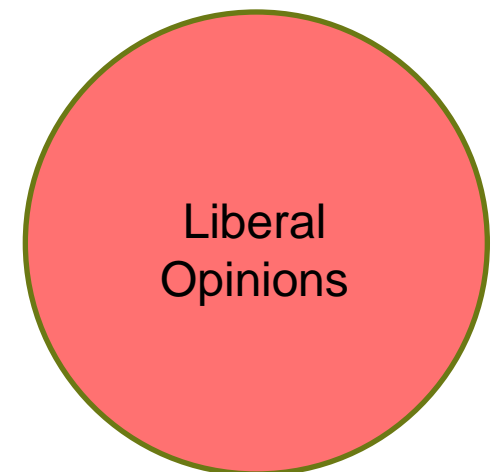
2



3



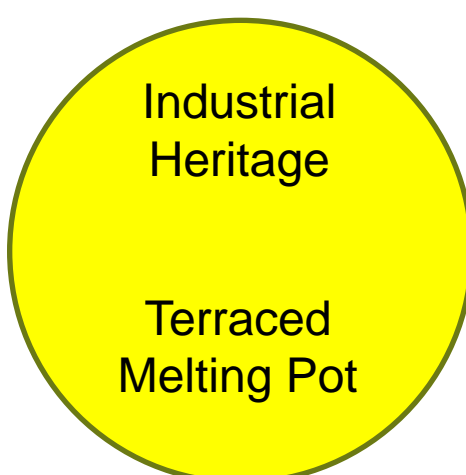
4



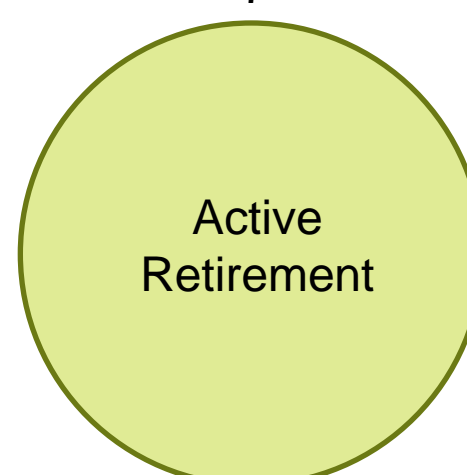
5



6



7

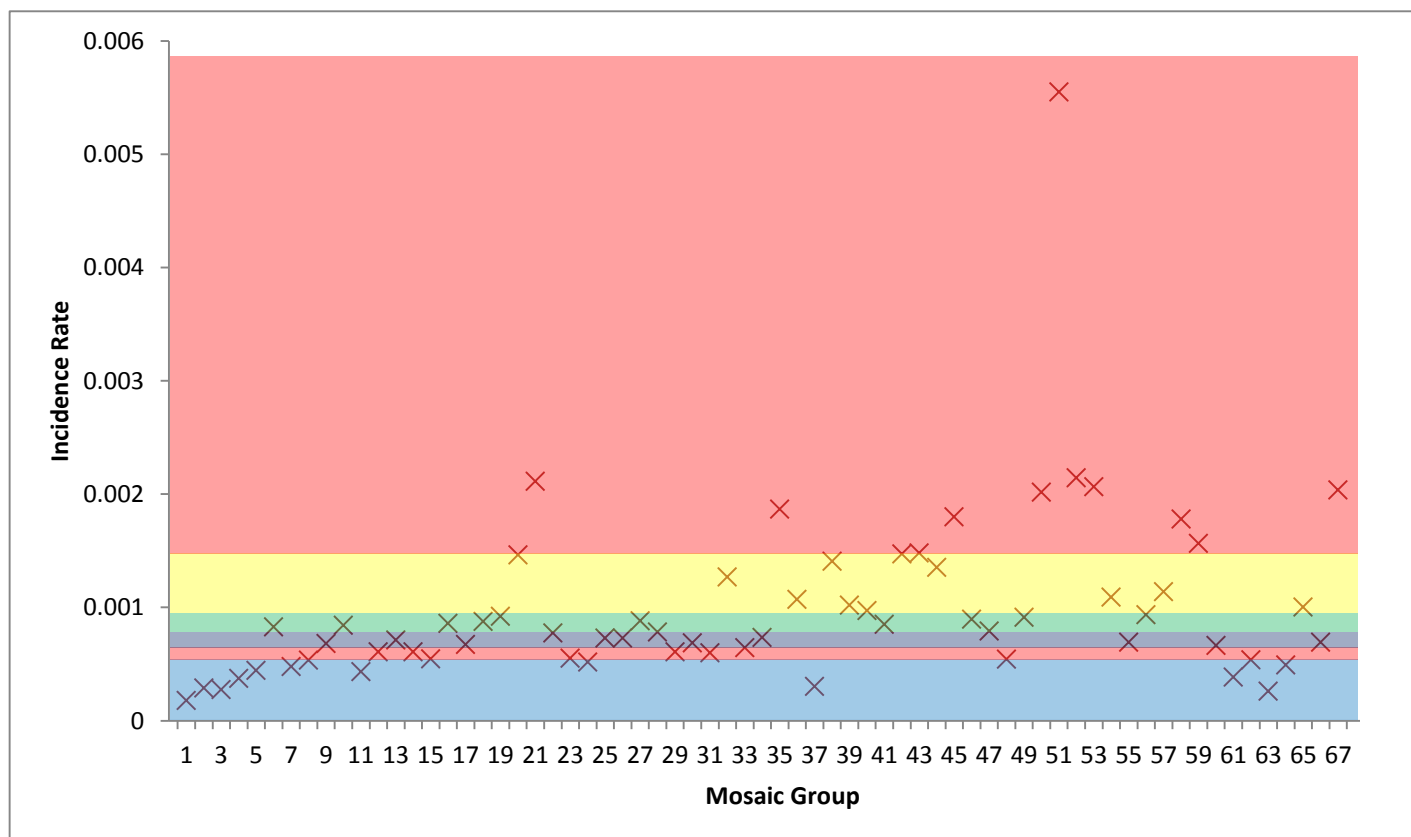


8



Bottom Up - Homogenous Groups

Mosaic

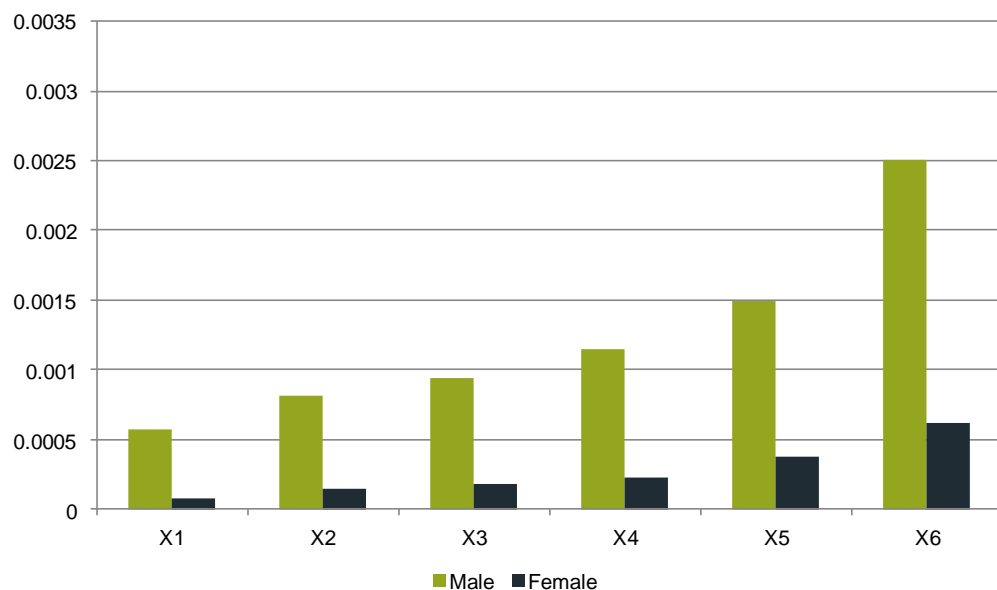


Mosaic Household	WP Grouping
1	X1
2	X1
3	X1
4	X1
5	X1
6	X4
7	X1
8	X2
!	!
59	X6
60	X3
61	X1
62	X2
63	X1
64	X1
65	X5
66	X3
67	X6

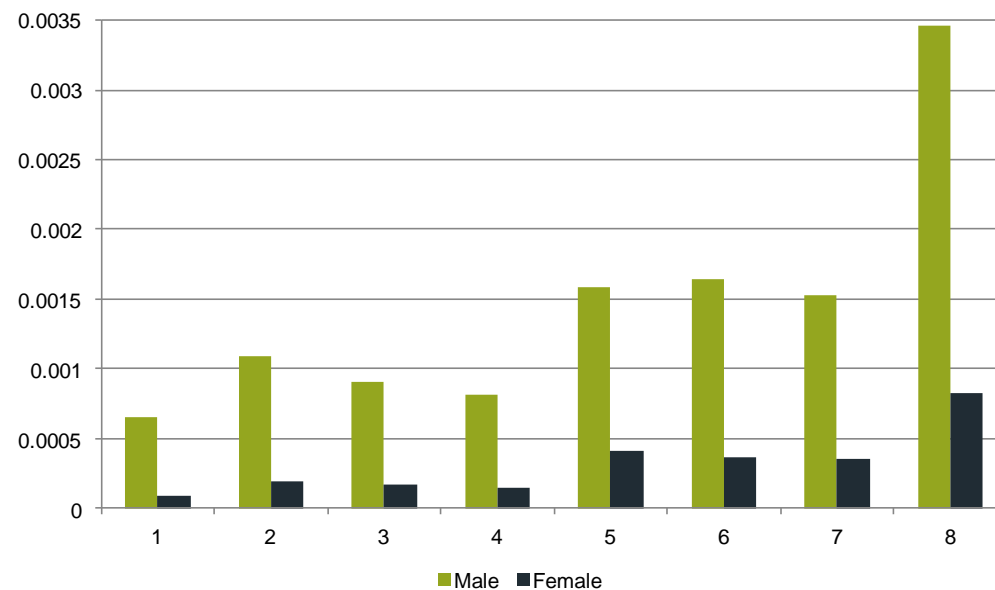
Heart Attack

Results for Males and Females

Incidence - Bottom Up



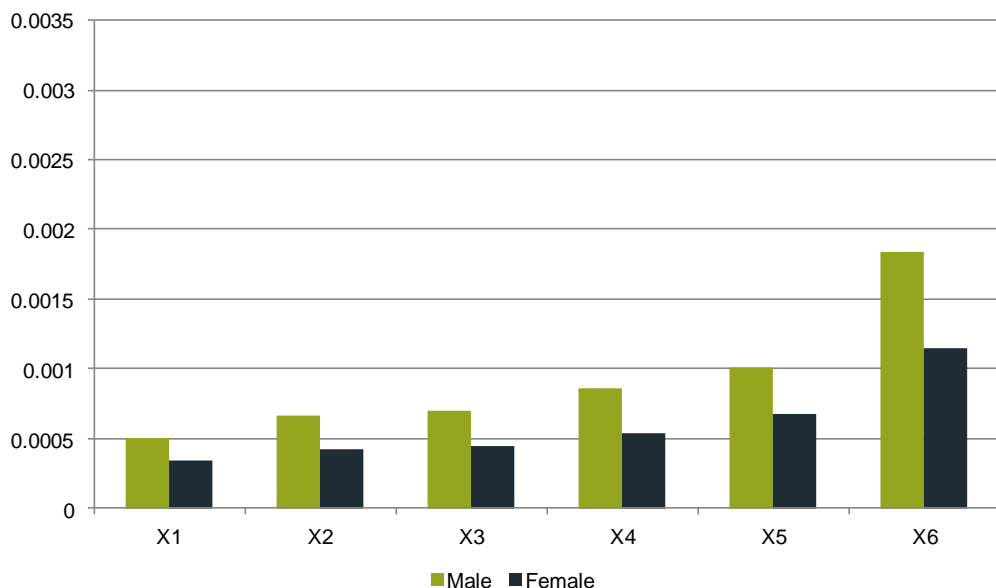
Incidence - Top Down



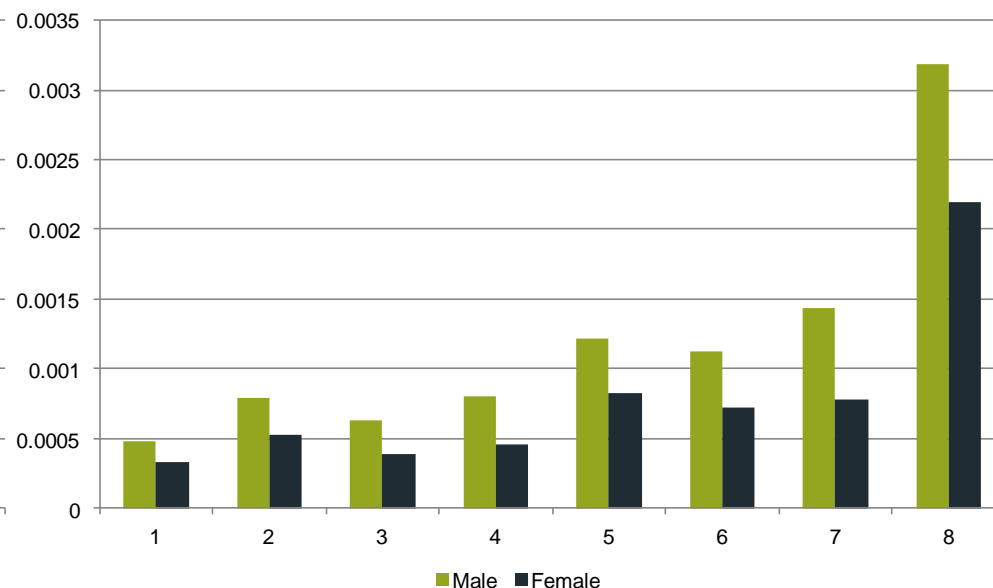
Stroke

Results for Males and Females

Incidence - Bottom Up

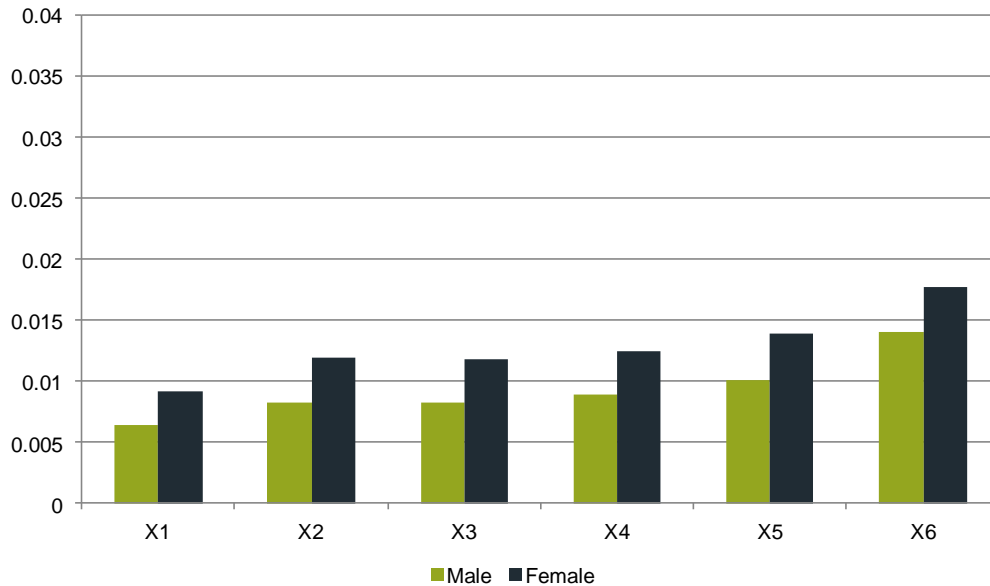


Incidence - Top Down

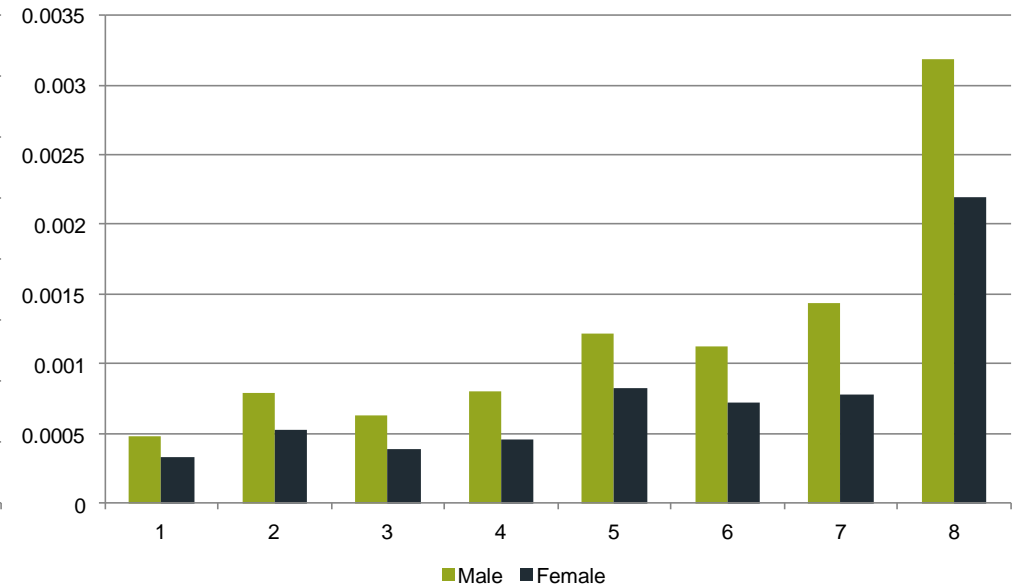


Cancer Results for Males and Females

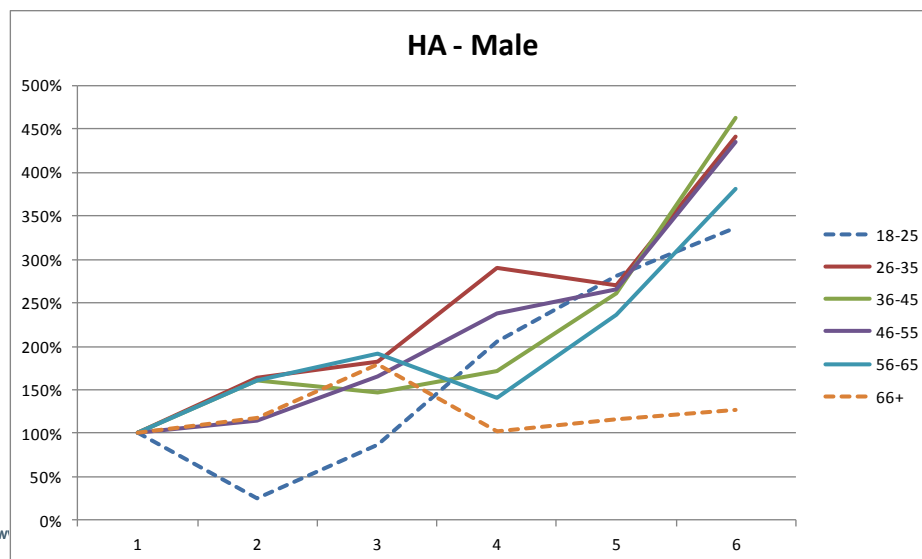
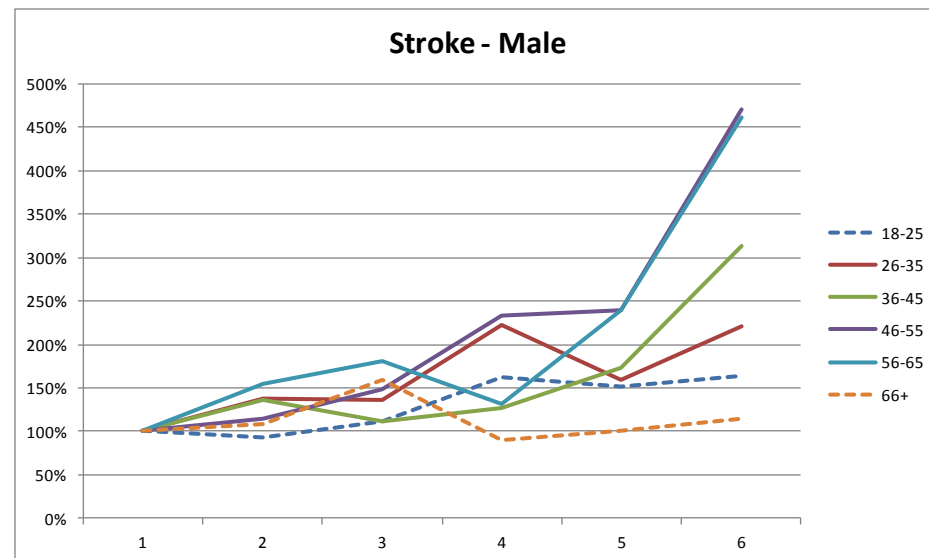
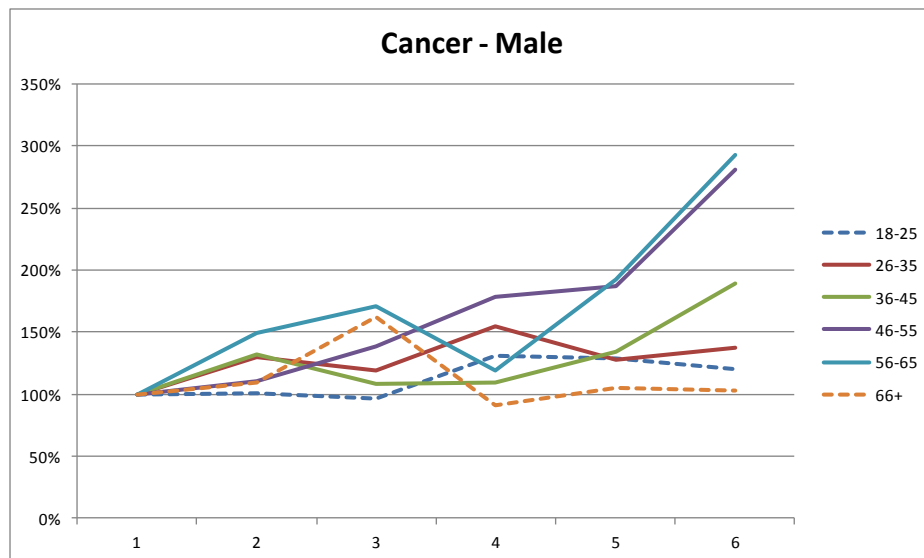
Incidence - Bottom Up



Incidence - Top Down



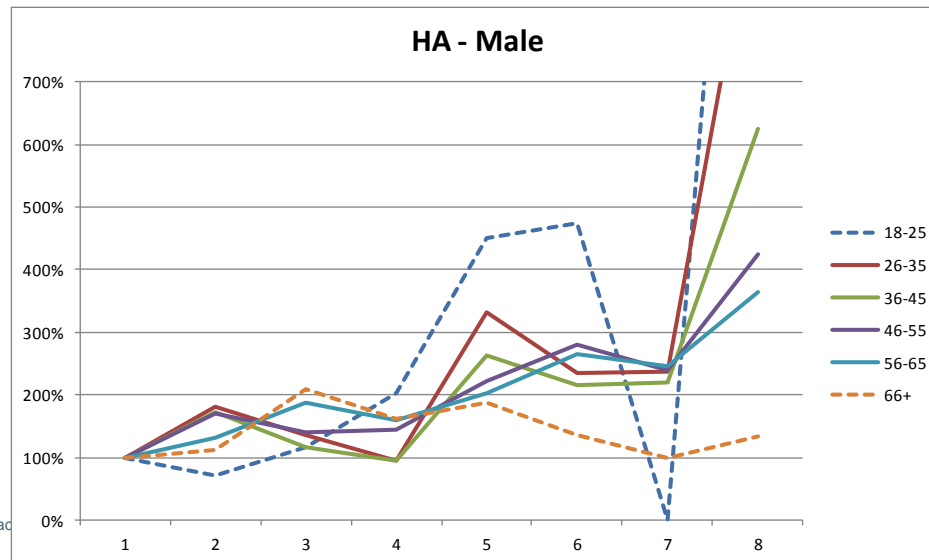
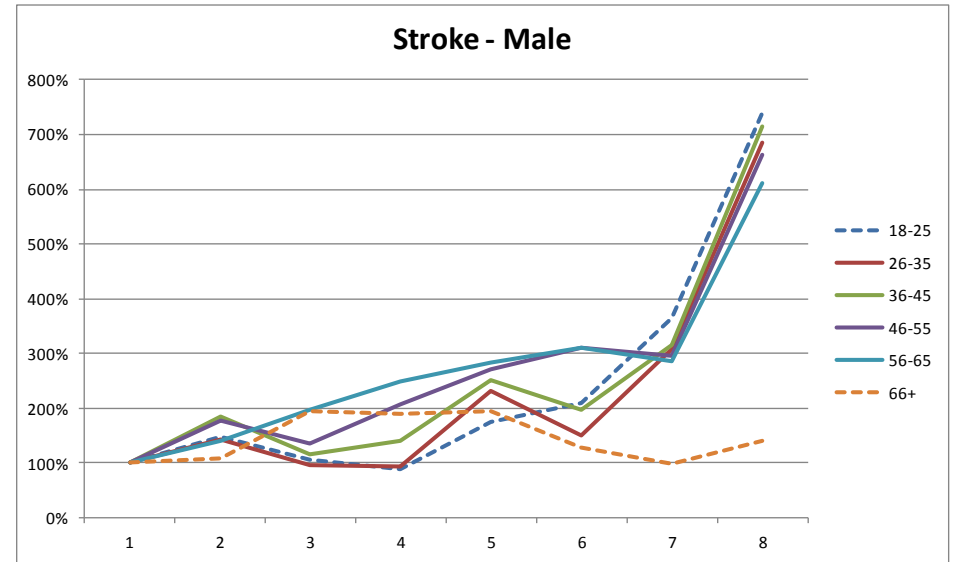
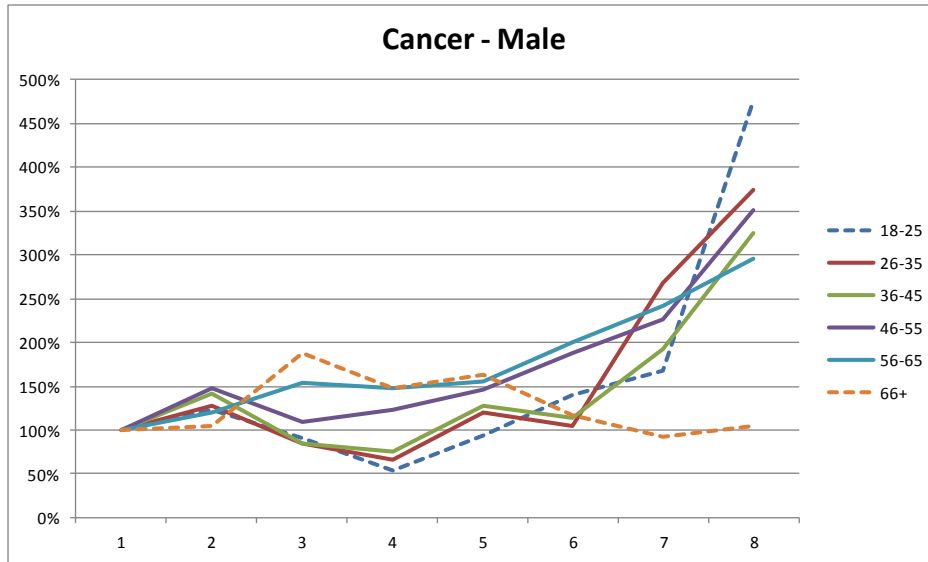
Age Bands – Bottom Up



Mosaic Classification – Copyright of Experian Limited, used with Permission. All rights Reserved.

HES Data - Copyright © 2011, Re-used with the permission of The Health and Social Care Information Centre. All rights reserved.

Age Bands – Top Down

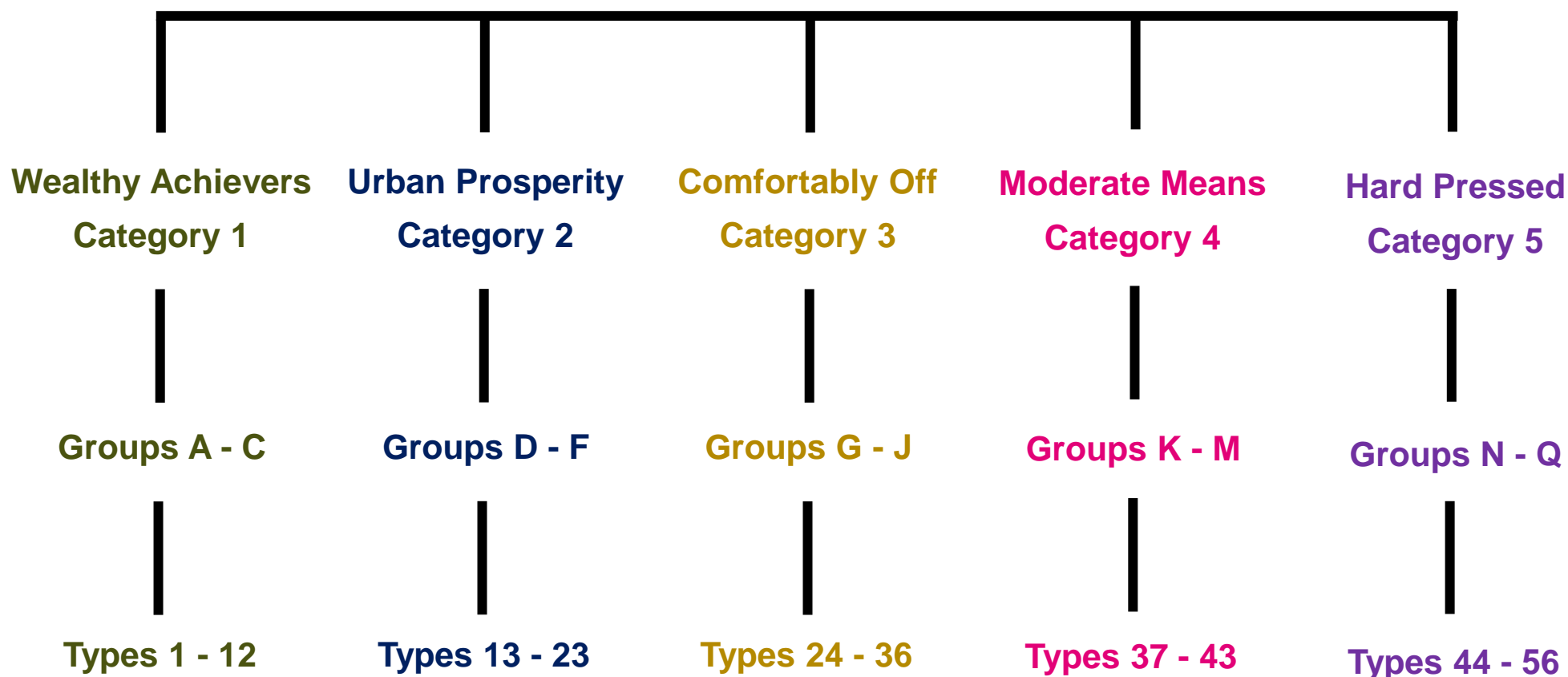


Mosaic Classification – Copyright of Experian Limited, used with Permission. All rights Reserved.

HES Data - Copyright © 2011, Re-used with the permission of The Health and Social Care Information Centre. All rights reserved.

Demographic Profilers – CACI

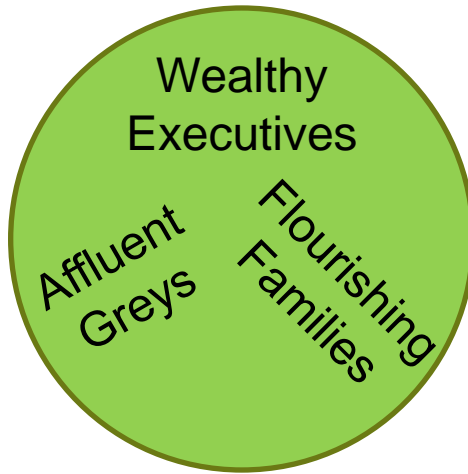
ACORN Structure:



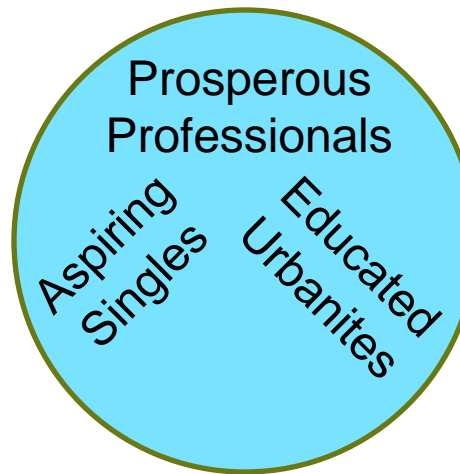
Initial Analysis of HES Data

Top Down Analysis – ACORN

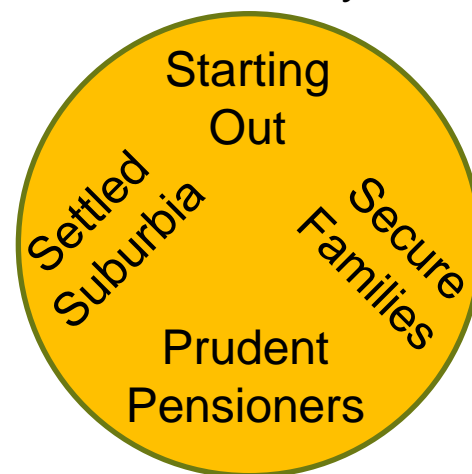
1. Wealthy Achievers



2. Urban Prosperity



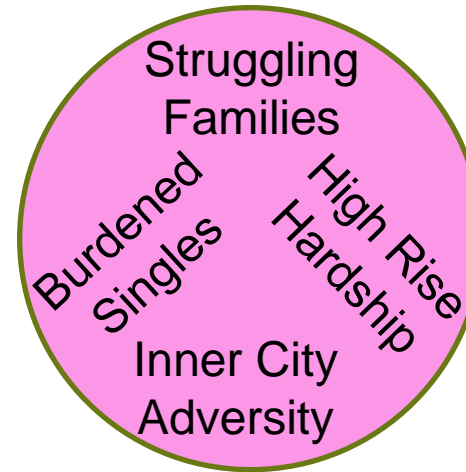
3. Comfortably Off



4. Moderate Means

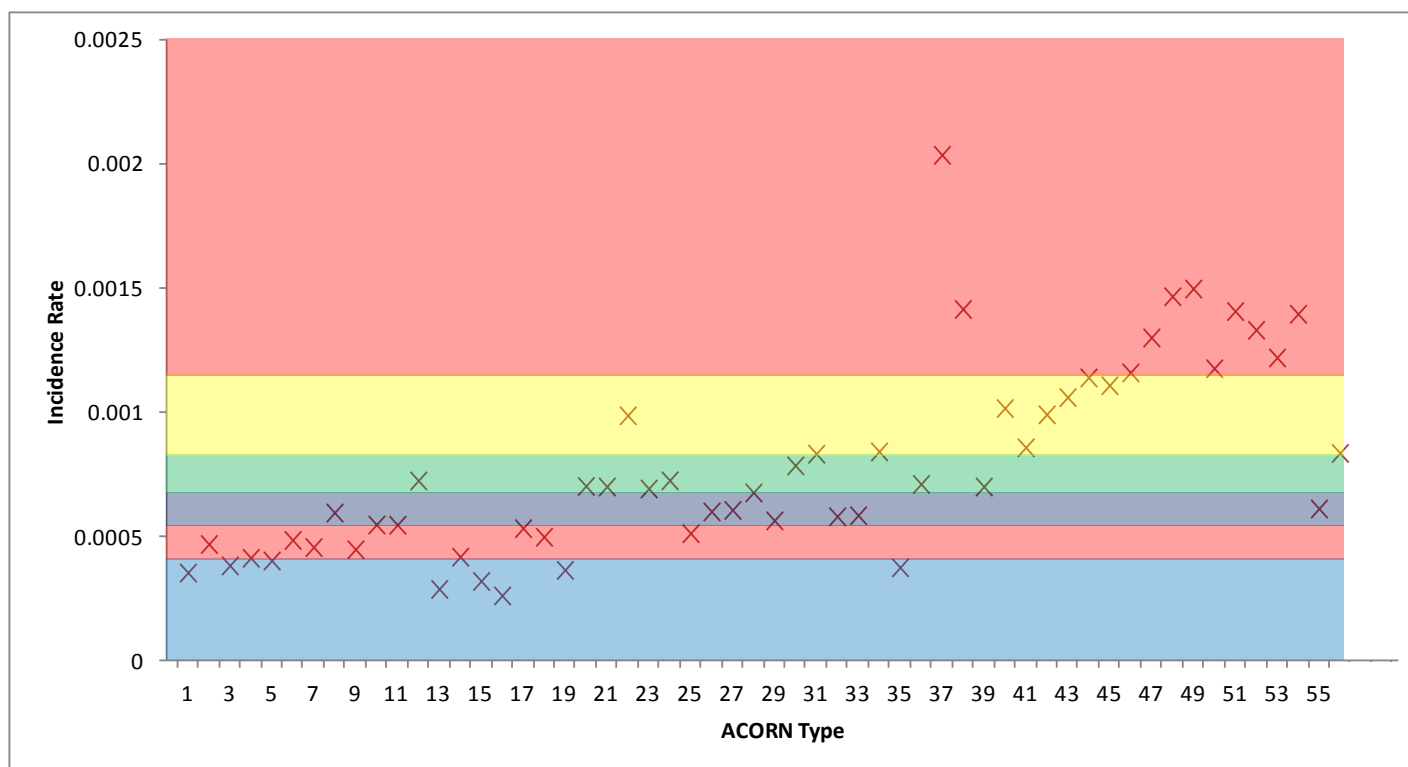


5. Hard Pressed



Initial Analysis of HES Data

Homogenous Groups

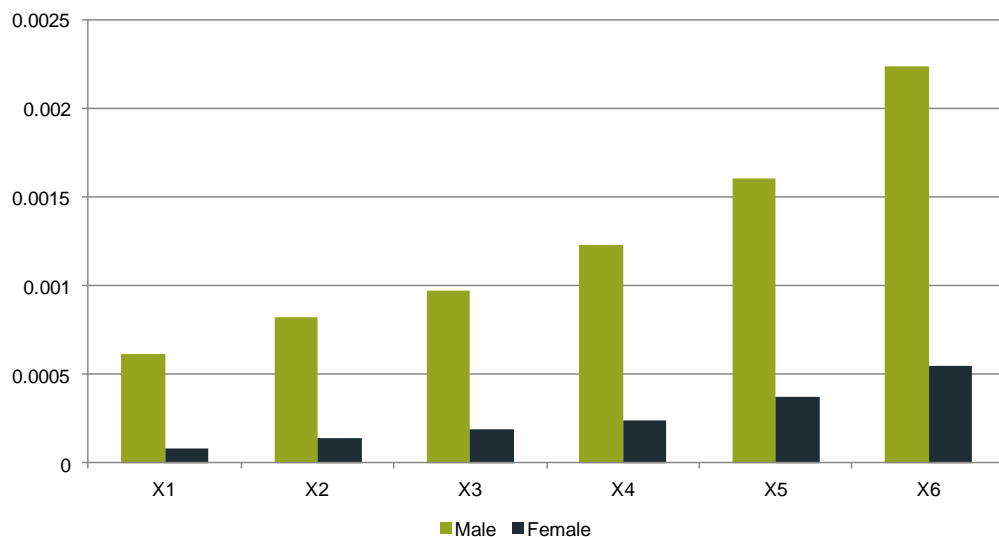


ACORN Type	WP Grouping
1	X1
2	X2
3	X1
4	X1
5	X1
6	X2
7	X2
8	X3
!	!
48	X6
49	X6
50	X6
51	X6
52	X6
53	X6
54	X6
55	X3
56	X5

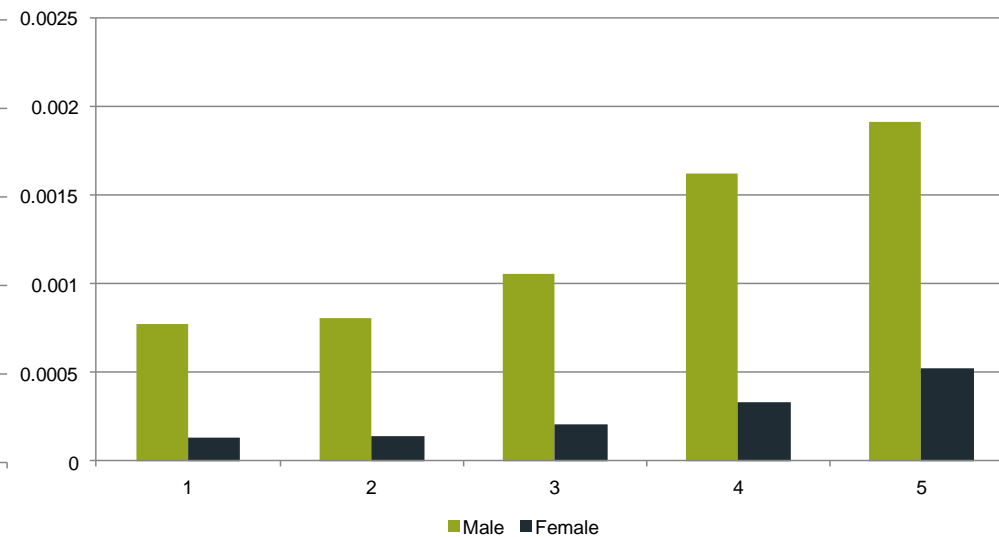
Heart Attack

Results for Males and Females

Incidence - Bottom Up



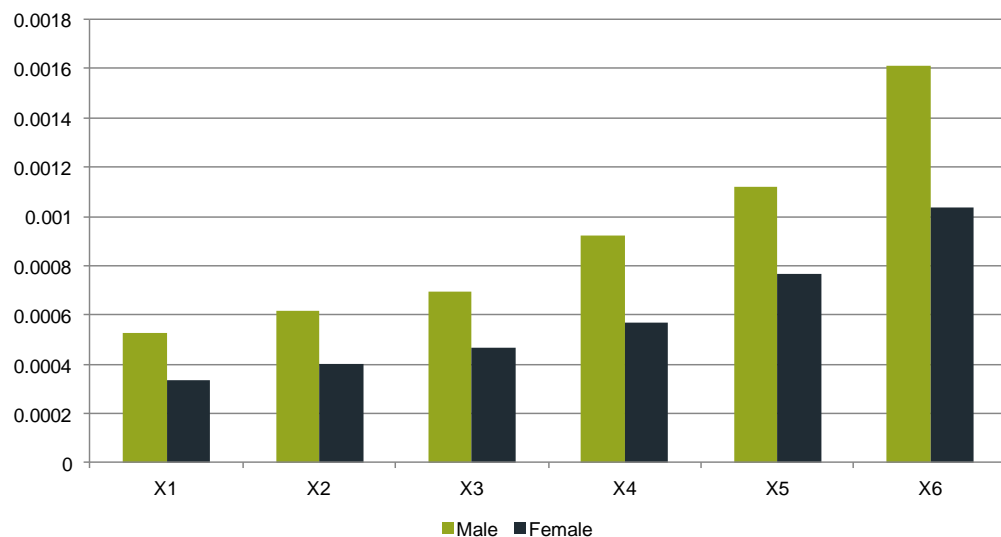
Incidence - Top Down



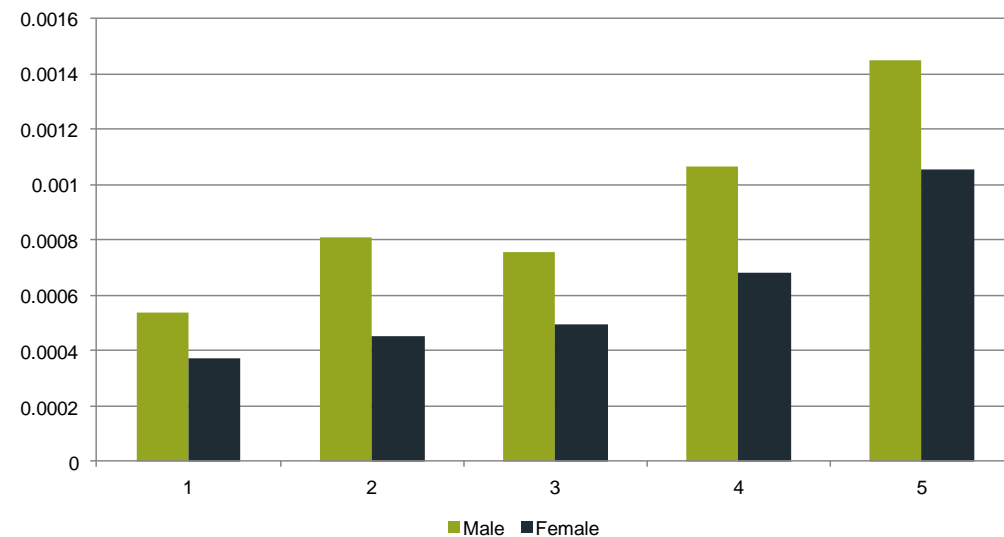
Stroke

Results for Males and Females

Incidence - Bottom Up



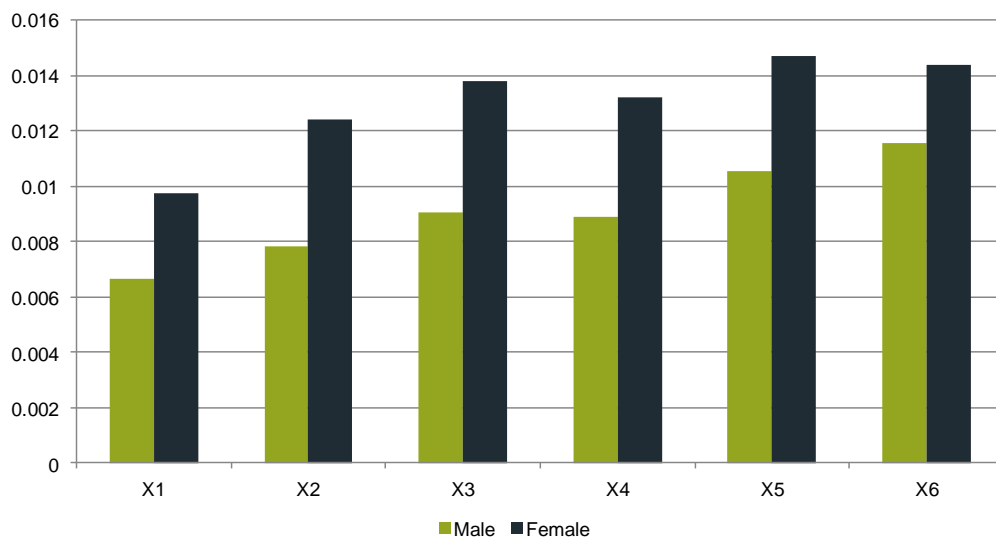
Incidence - Top Down



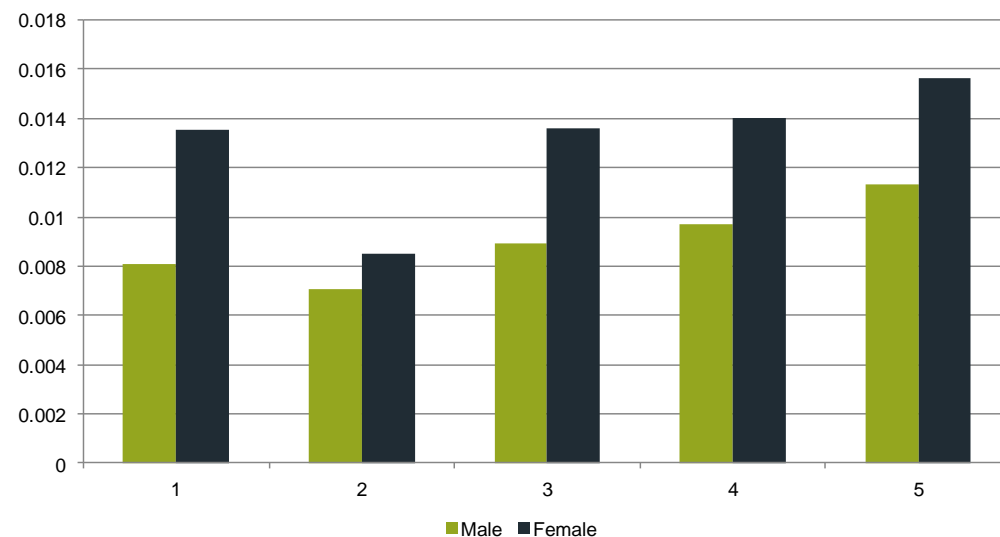
Cancer

Results for Males and Females

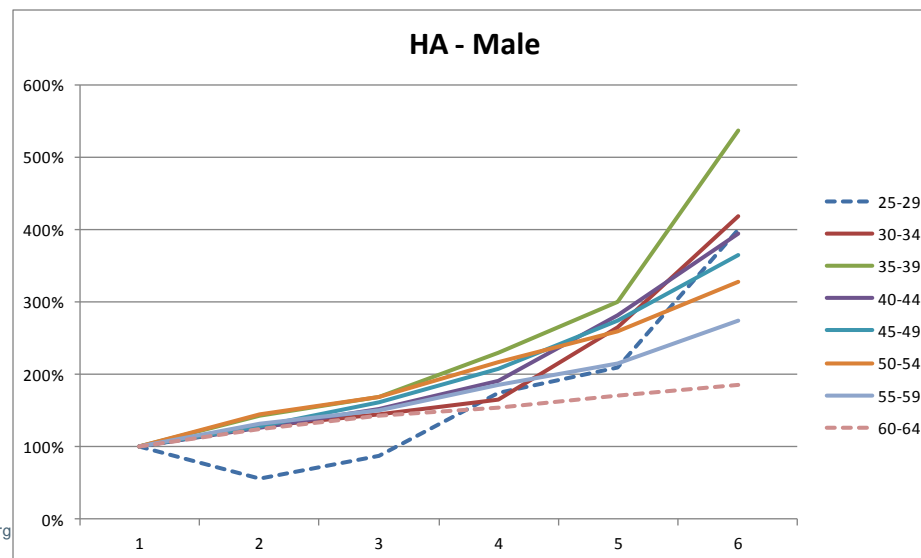
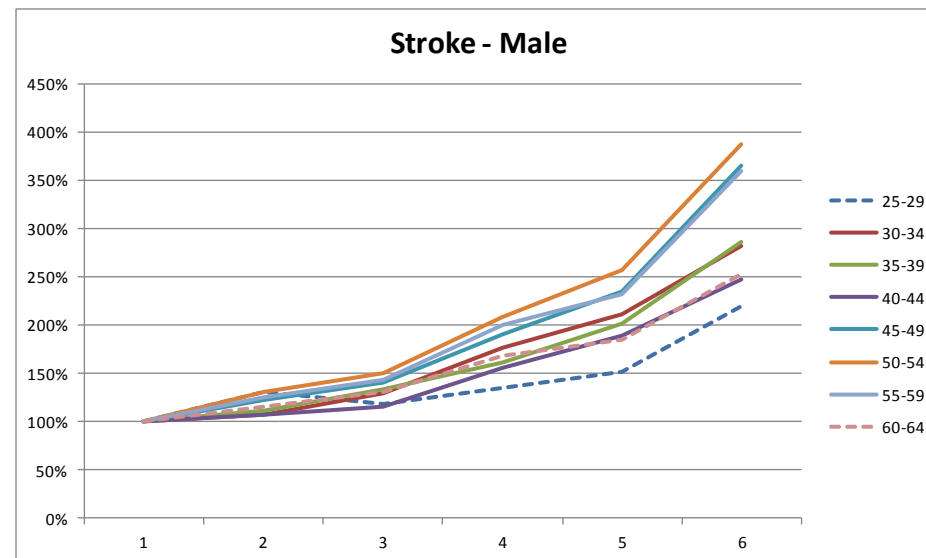
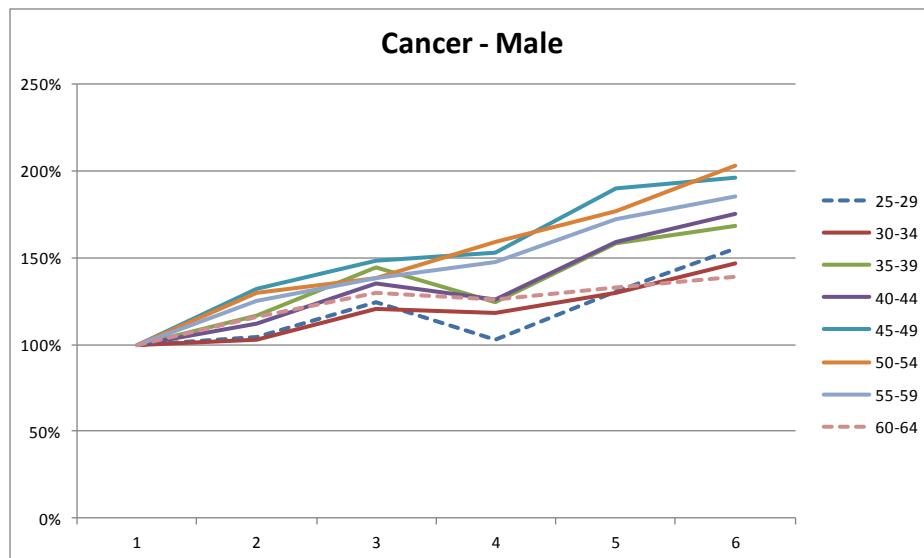
Incidence - Bottom Up



Incidence - Top Down



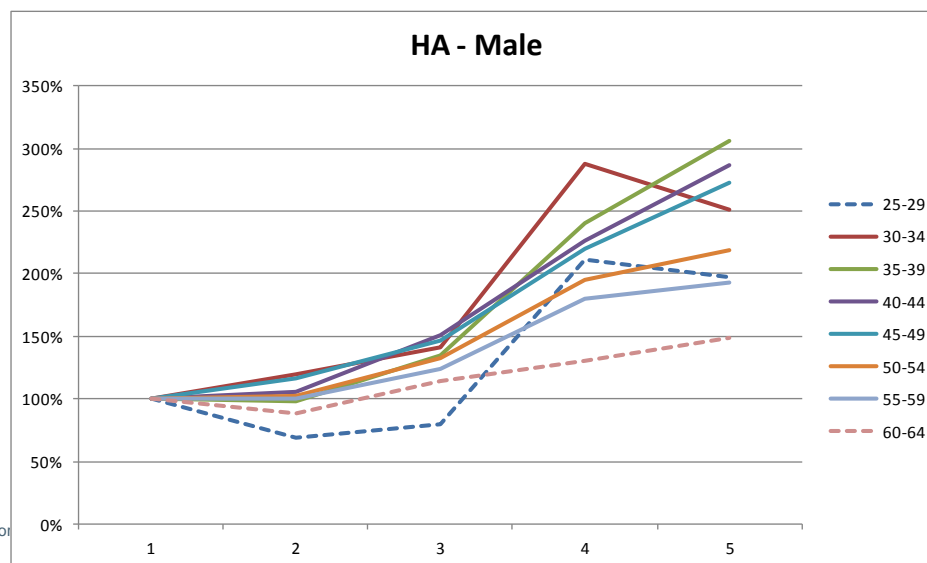
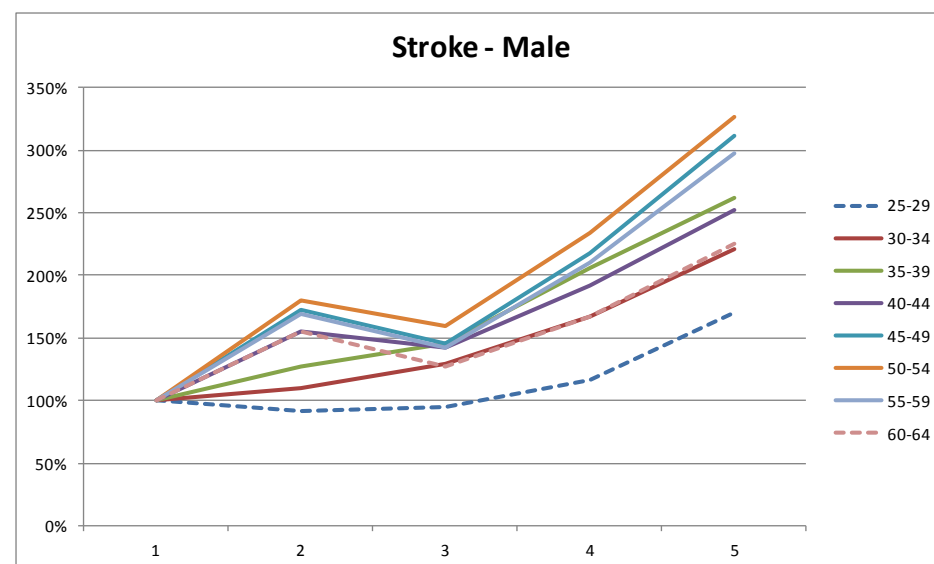
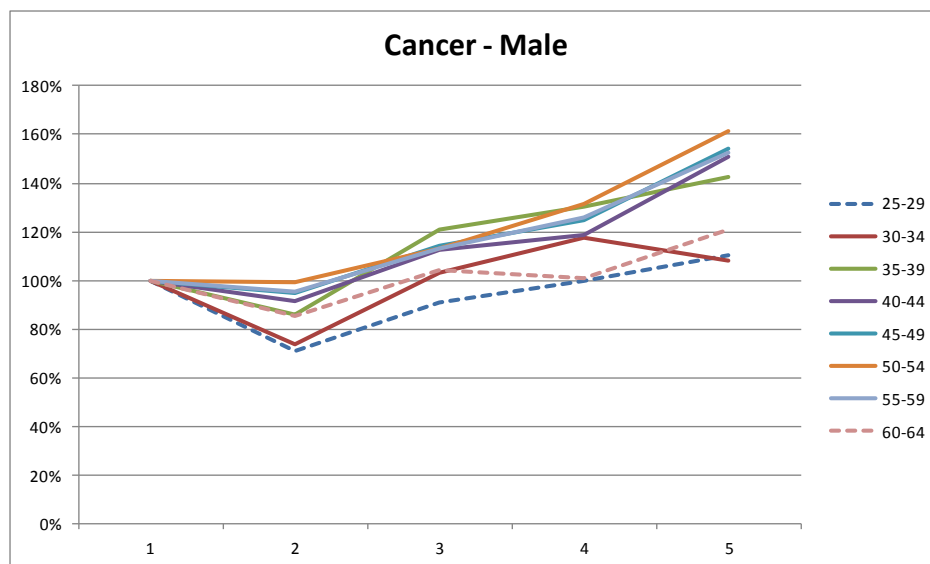
Age Bands – Bottom Up



ACORN Classification – Copyright of CACI Ltd, used with Permission. All rights Reserved.

HES Data - Copyright © 2011, Re-used with the permission of The Health and Social Care Information Centre. All rights reserved.

Age Bands – Top Down



ACORN Classification – Copyright of CACI Ltd, used with Permission. All rights Reserved.

HES Data - Copyright © 2011, Re-used with the permission of The Health and Social Care Information Centre. All rights reserved.

Demographic Profilers – CACI HealthACORN (4 Groups)

1 – Healthy



2 – Possible Future Concerns



3 – Future Problems



4 – Existing Problems

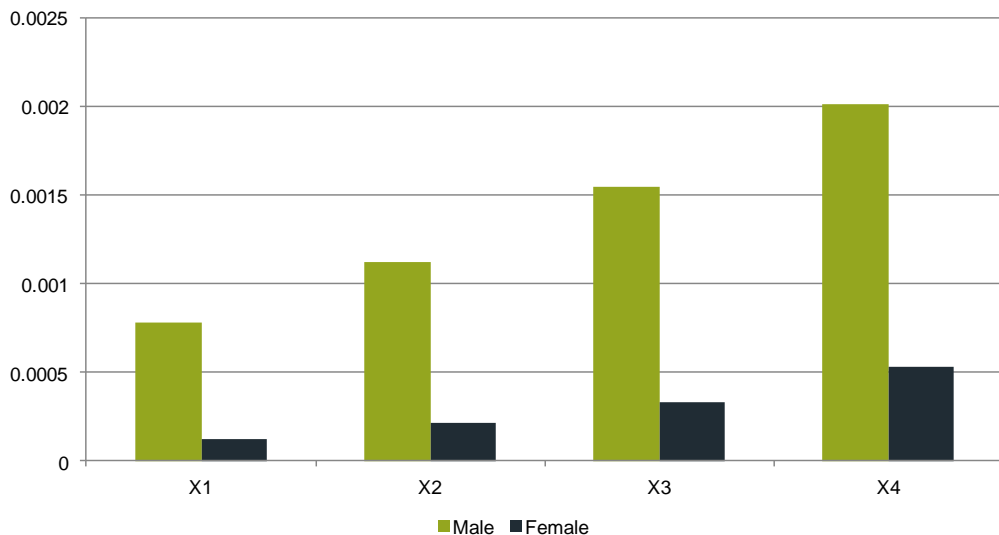


HealthACORN TYPE	Bottom Up	Top Down
1	X2	4
2	X3	4
3	X4	4
4	X3	4
5	X4	4
6	X3	4
7	X2	4
8	X4	4
9	X4	3
10	X4	3
11	X4	3
12	X4	3
13	X3	3
14	X3	2
15	X1	2
16	X1	2
17	X1	2
18	X3	2
19	X1	2
20	X2	1
21	X1	1
22	X2	1
23	X2	1
24	X2	1
25	X1	1

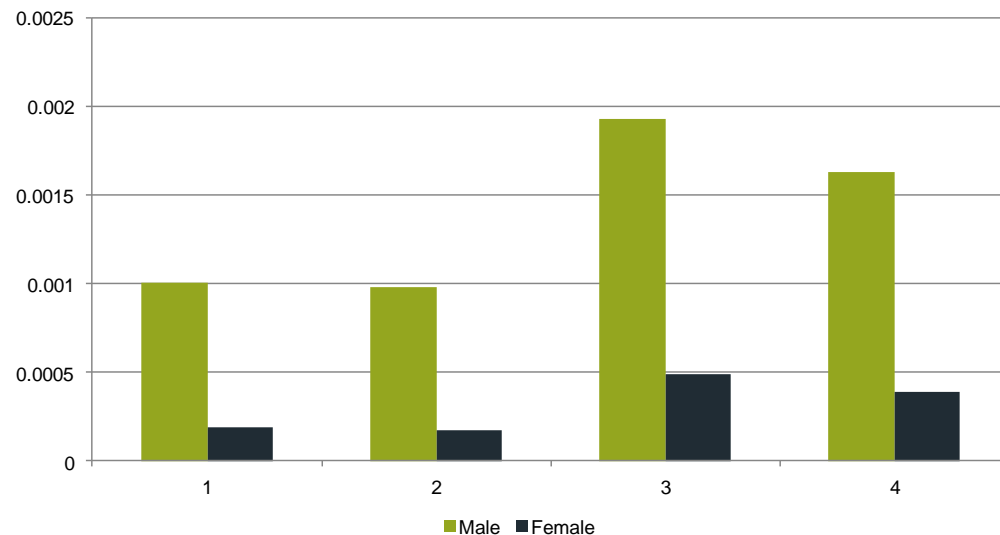
Heart Attack

Results for Males and Females

Incidence - Bottom Up



Incidence - Top Down



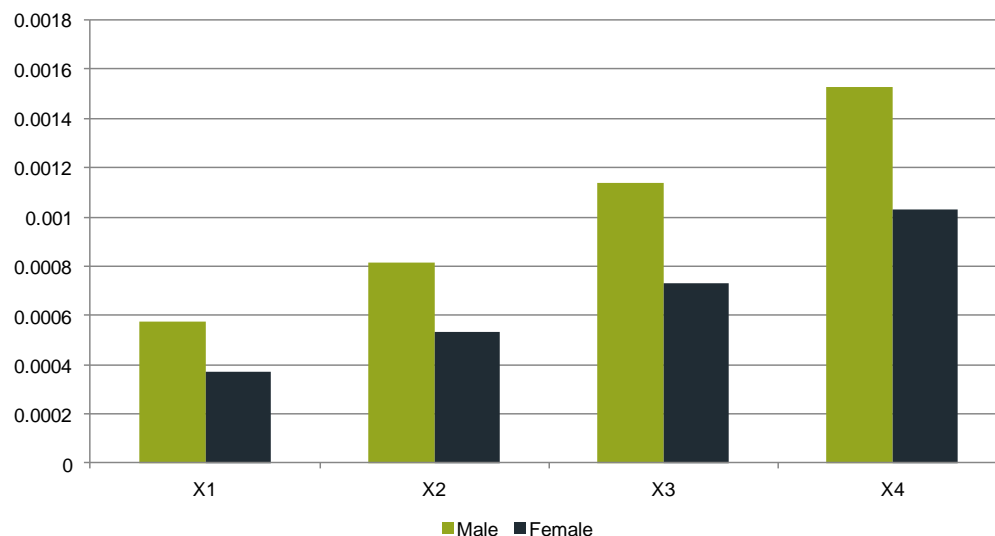
HES Data - Copyright © 2011, Re-used with the permission of The Health and Social Care Information Centre. All rights reserved.

HealthACORN Classification – Copyright of CACI Ltd, used with Permission. All rights Reserved.

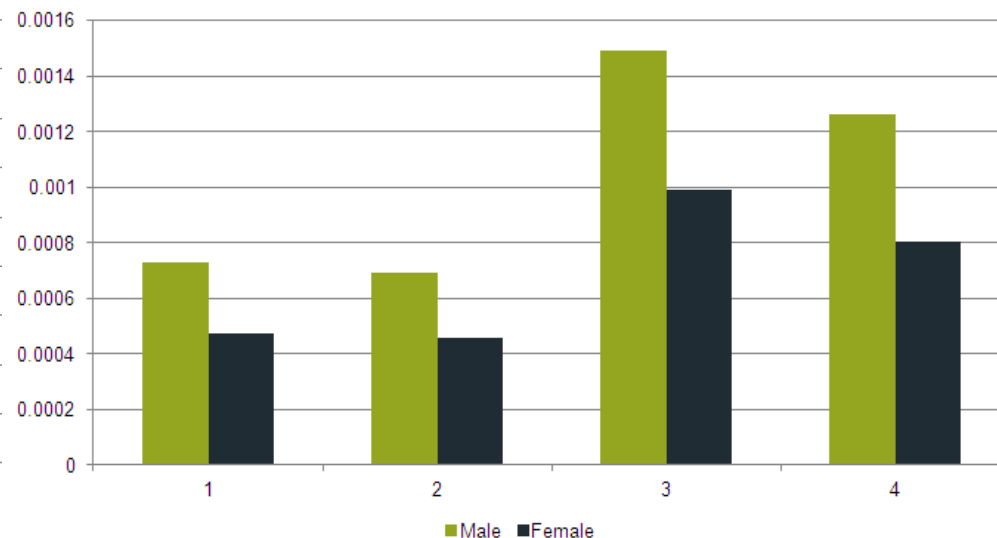
Stroke

Results for Males and Females

Incidence - Bottom Up



Incidence - Top Down

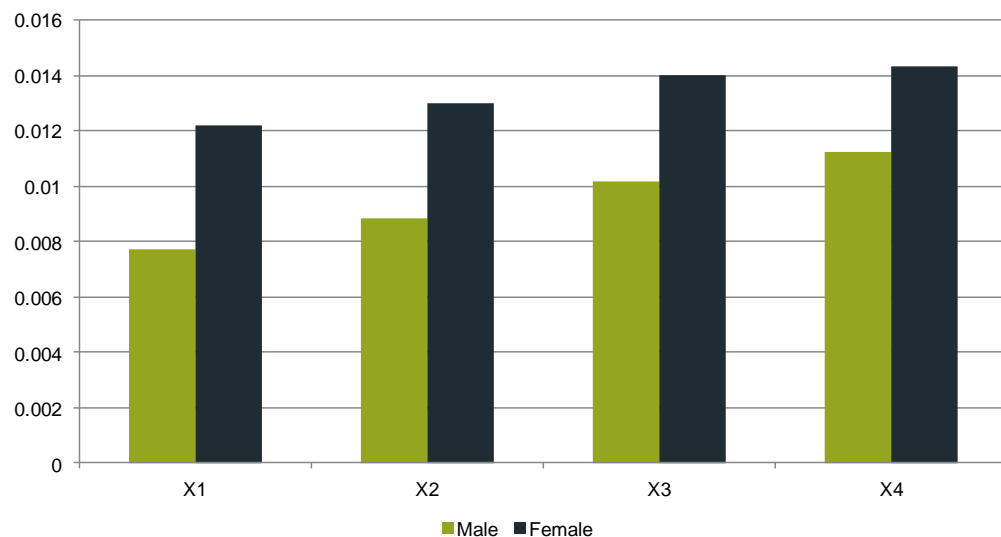


HES Data - Copyright © 2011, Re-used with the permission of The Health and Social Care Information Centre. All rights reserved.

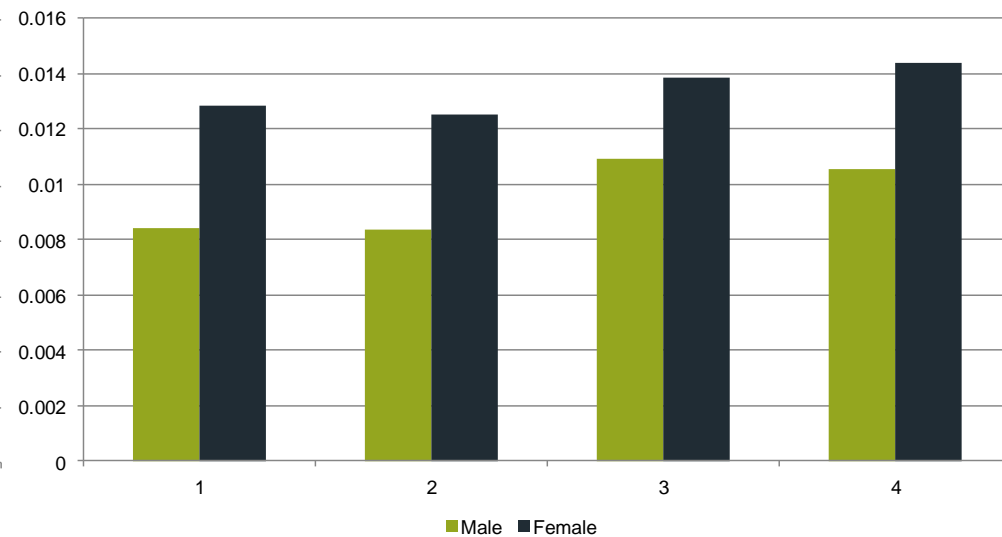
HealthACORN Classification – Copyright of CACI Ltd, used with Permission. All rights Reserved.

Cancer Results for Males and Females

Incidence - Bottom Up



Incidence - Top Down



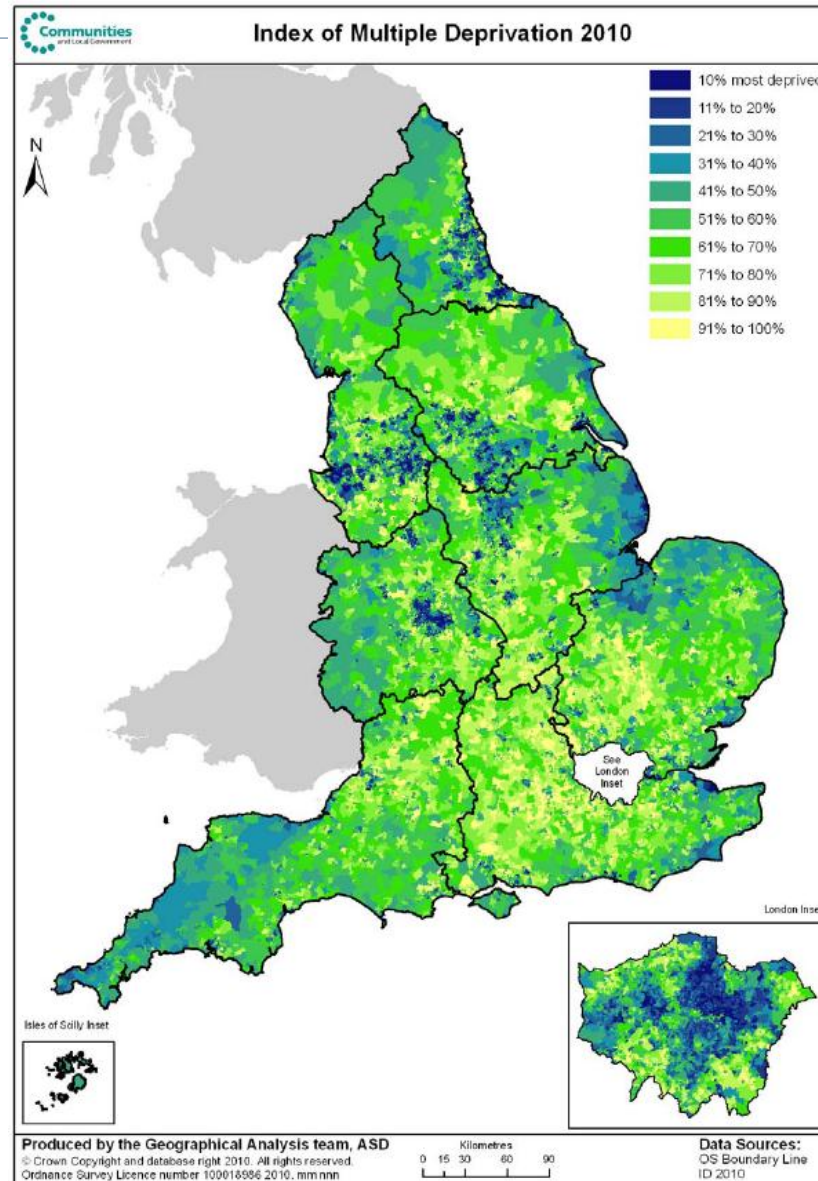
HES Data - Copyright © 2011, Re-used with the permission of The Health and Social Care Information Centre. All rights reserved.

HealthACORN Classification – Copyright of CACI Ltd, used with Permission. All rights Reserved.

Initial Analysis of HES Data by IMD Ranking

- **Index of Multiple Deprivation 2007**
 - Based on Small Area Geography – known as Lower Super Output Areas (LSOAs)
 - Average Population of an LSOA is 1,500 people
- **Brings together 37 different indicators covering:**
 - Income, Employment, Health & Disability, Education, Skills & Training, Living Environment and Crime
- **Each LSOA ranked according to level of Deprivation**
 - The higher the IMD Ranking, the higher the level of Deprivation
- **IMD Identifies Concentrations of Deprivation**
 - BUT not all deprived people live in deprived areas AND not everyone living in a deprived area is deprived!!!

Geographical Distribution of IMD Deciles



Initial Analysis of HES Data by IMD Ranking

- Processed data consists of:

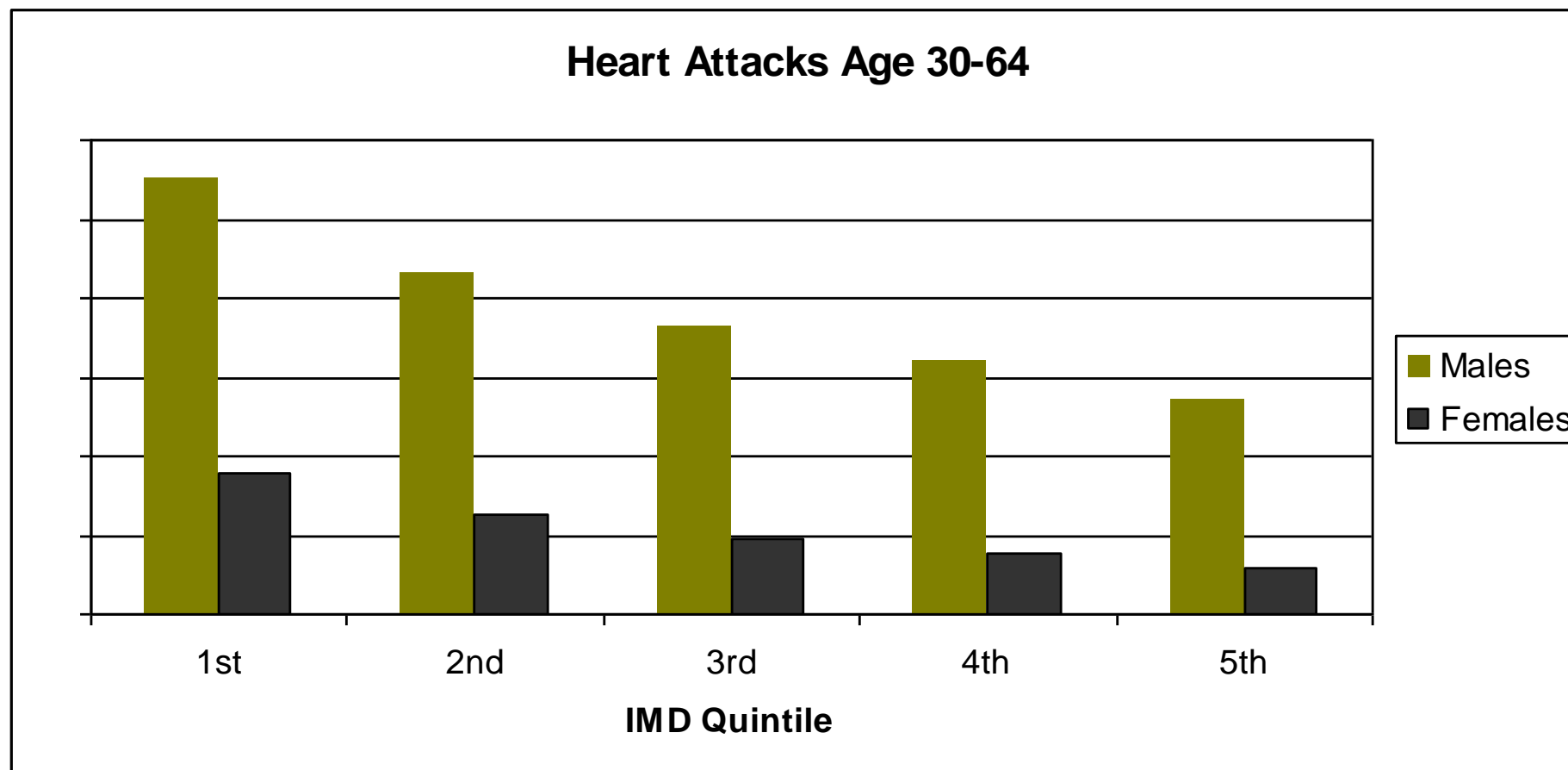
Claims	Population
Cause : Cancer, HA & Stroke	
Age banded	Age banded
Gender	Gender
LSOA	LSOA
IMD Ranking	IMD Ranking
Years : 2006 - 2009	Mid 2010

- Population Data: LSOA Population Estimates for England 2010 (ONS)**
- LSOAs for England grouped into 5 Quintiles according to IMD Ranking**

Initial Analysis of HES Data by IMD Ranking

- **Calculate “claims” and exposure for England only split by**
 - Age Banding (Ages 30 to 64)
 - IMD Quintile
- **Determine age-standardised incidence rates for each quintile**
 - 1st Quintile represents LSOAs with highest concentration of deprivation
 - 5th Quintile represents LSOAs with lowest concentration of deprivation

Heart Attack Results for Males and Females

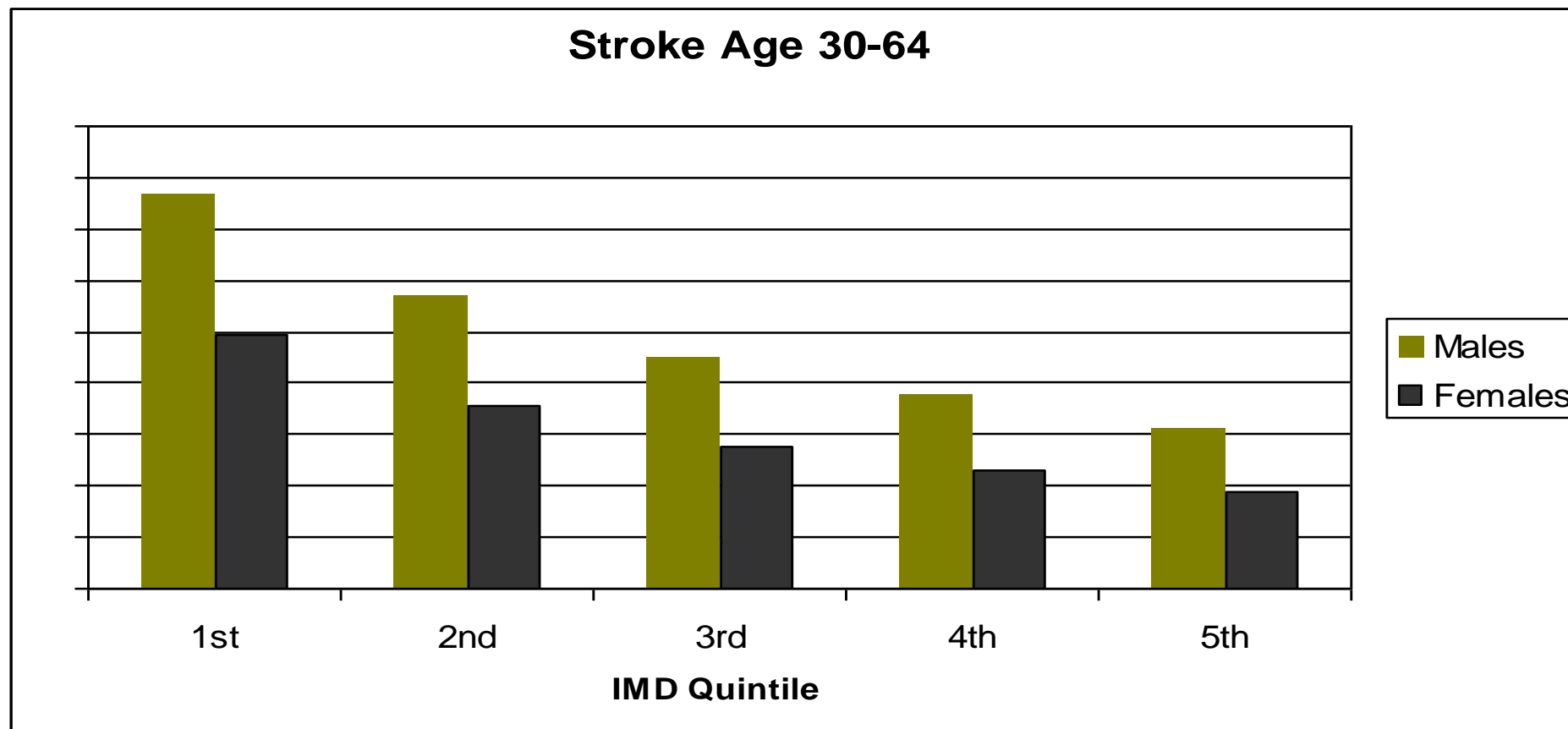


Using CMI Population Mix

HES Data - Copyright © 2011, Re-used with the permission of The Health and Social Care Information Centre. All rights reserved.

Stroke

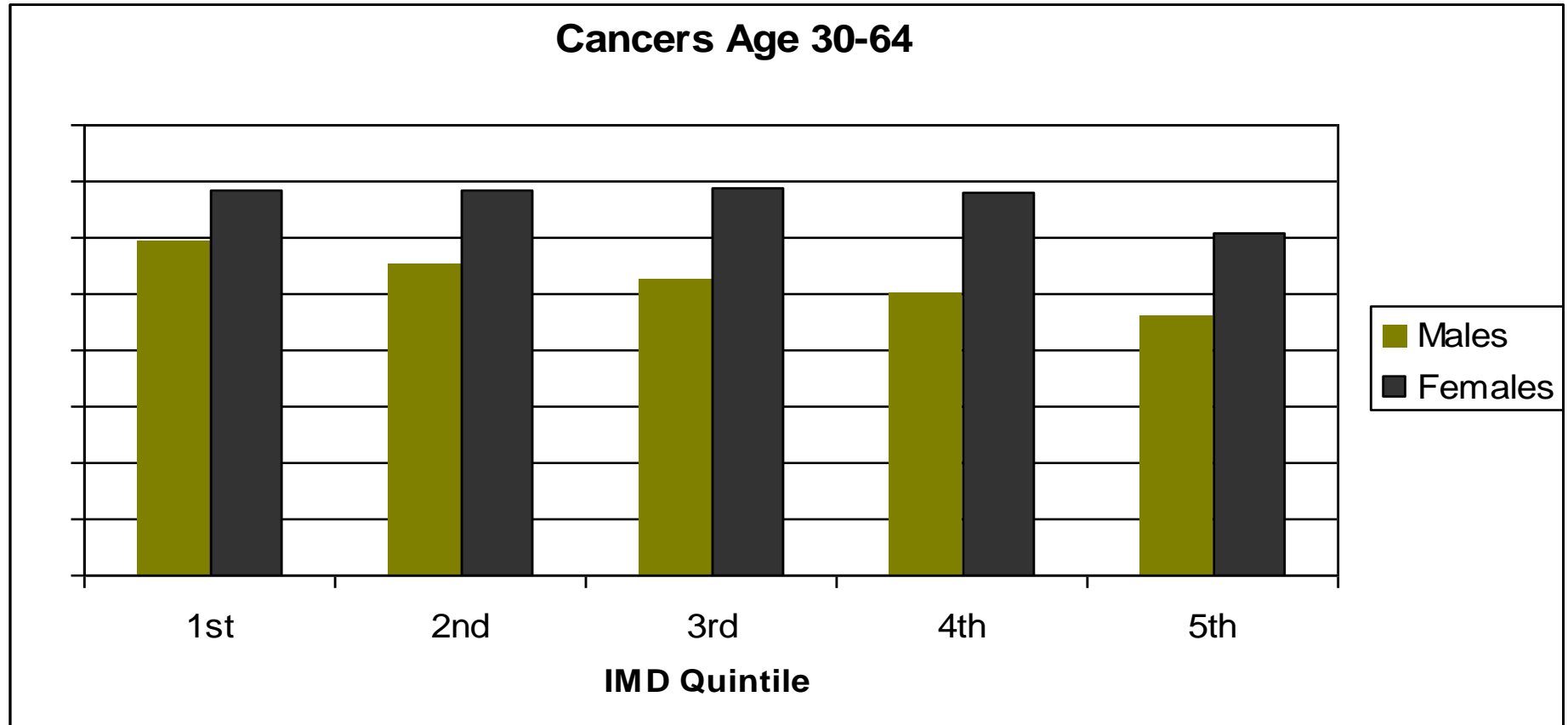
Results for Males and Females



Using CMI Population Mix

HES Data - Copyright © 2011, Re-used with the permission of The Health and Social Care Information Centre. All rights reserved.

Cancer Results for Males and Females



Using CMI Population Mix

HES Data - Copyright © 2011, Re-used with the permission of The Health and Social Care Information Centre. All rights reserved.

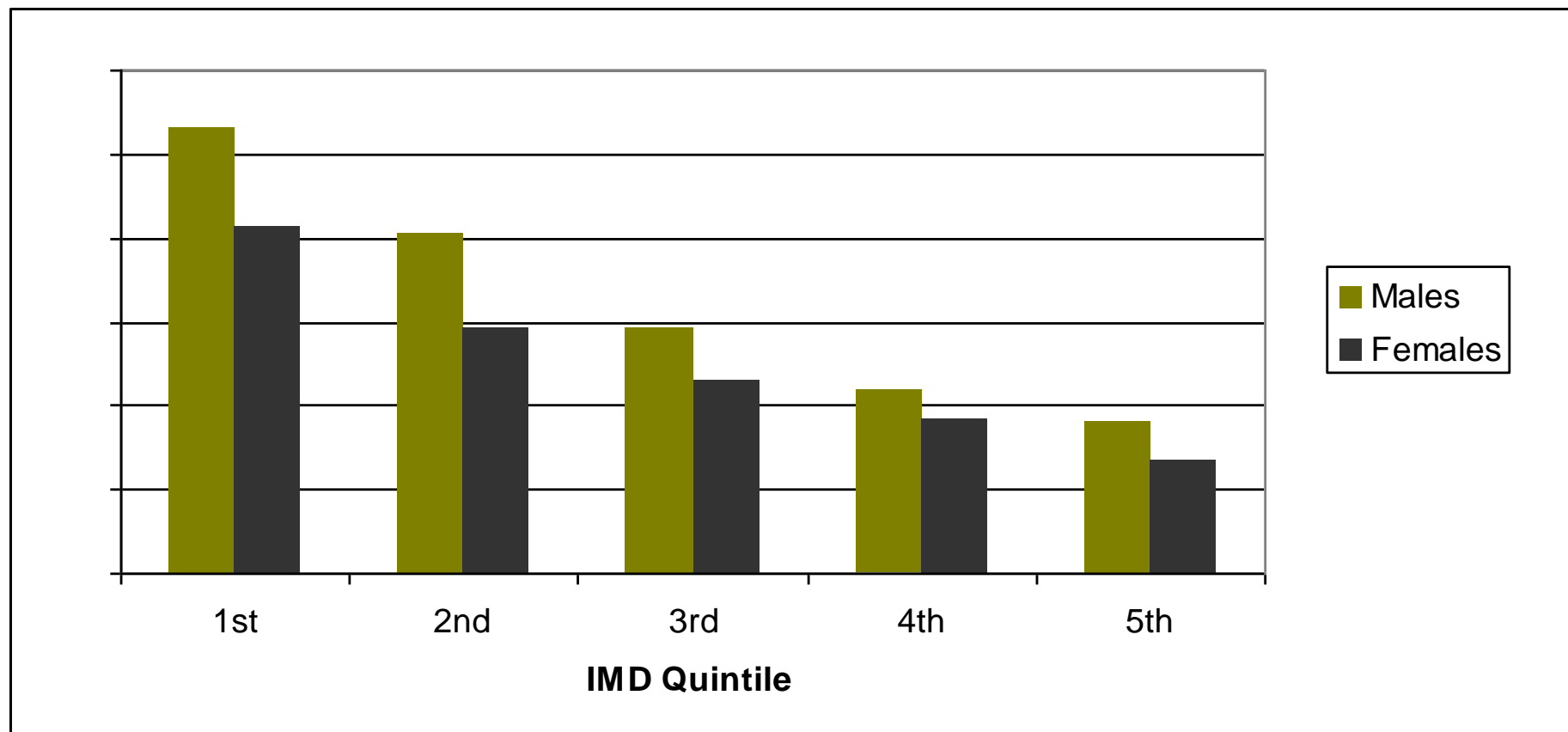
IMD Results By Cancer Site

Important Note

- In the following we perform analysis for 4 separate cancer sites;
- It should be noted that the **allocation of cancers may be incorrect** in some cases;
- This is a consequence of the current version of our algorithm which, in the case of multiple cancers being diagnosed on a single record our algorithm, labels the cancer site as the latest numerical ICD code rather than by order of diagnosis;
- It is our intention to refine this in later work.

Lung Cancer

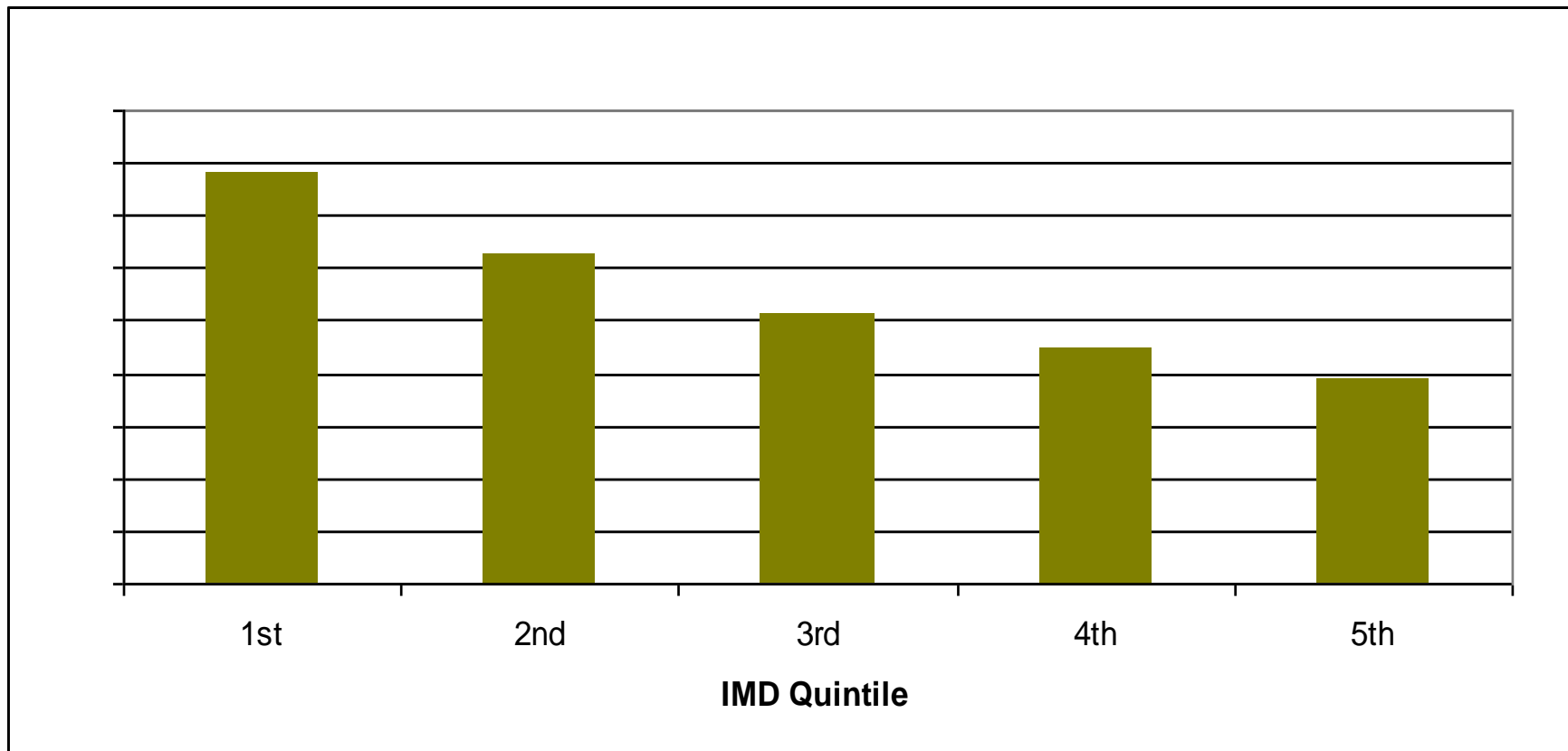
Results for Males and Females



Using CMI Population Mix

HES Data - Copyright © 2011, Re-used with the permission of The Health and Social Care Information Centre. All rights reserved.

Cervical Cancer

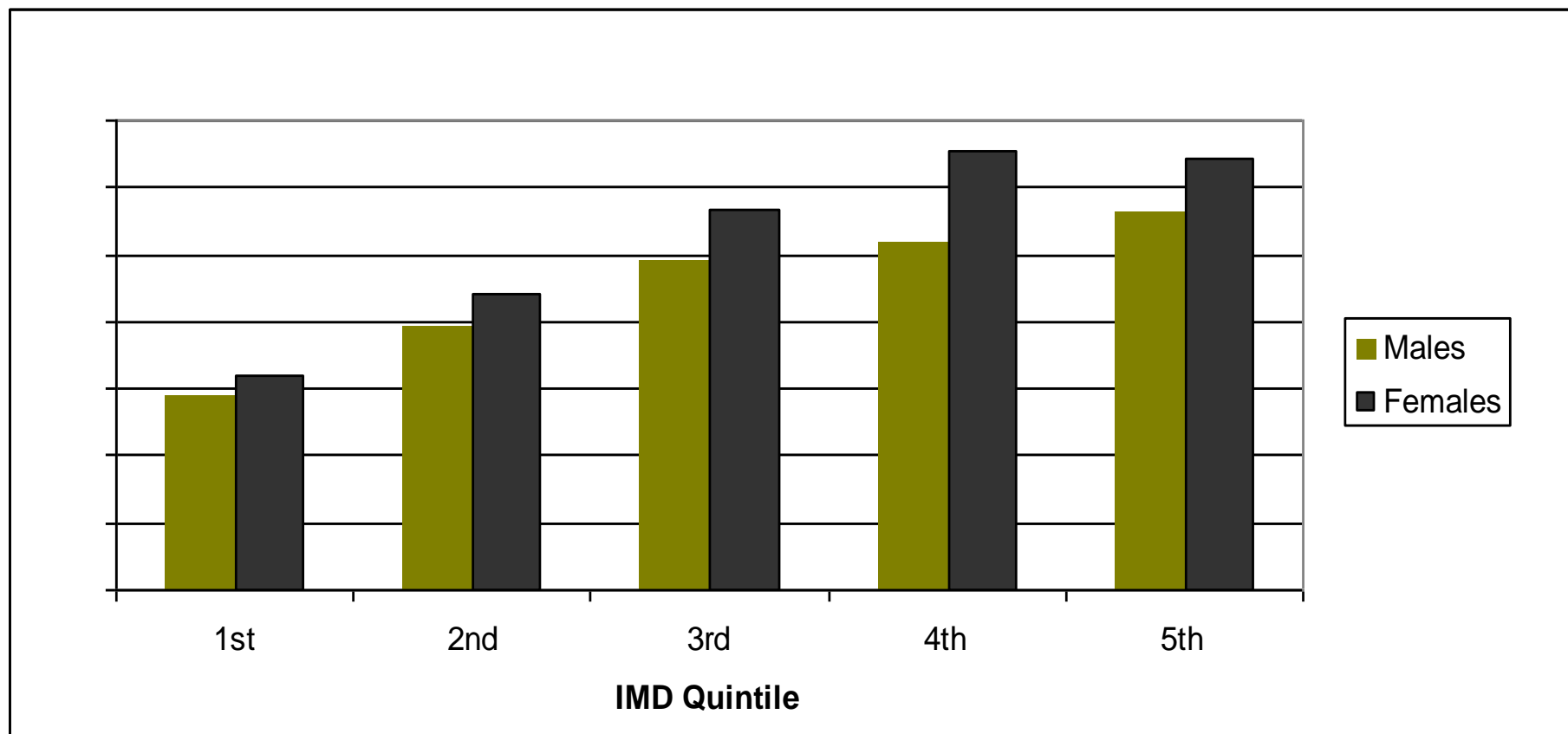


Using CMI Population Mix

HES Data - Copyright © 2011, Re-used with the permission of The Health and Social Care Information Centre. All rights reserved.

Melanoma

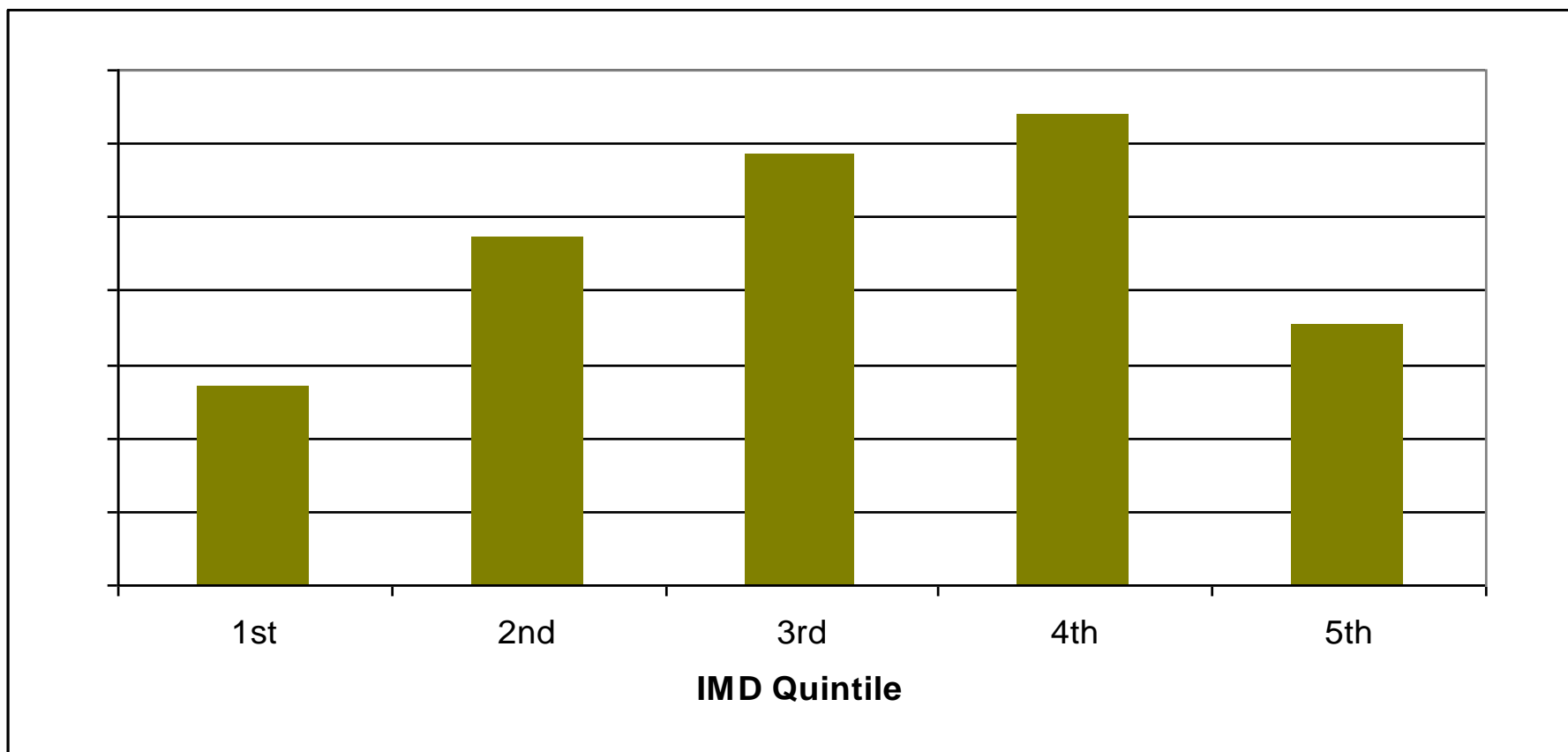
Results for Males and Females



Using CMI Population Mix

HES Data - Copyright © 2011, Re-used with the permission of The Health and Social Care Information Centre. All rights reserved.

Breast Cancer



Using CMI Population Mix

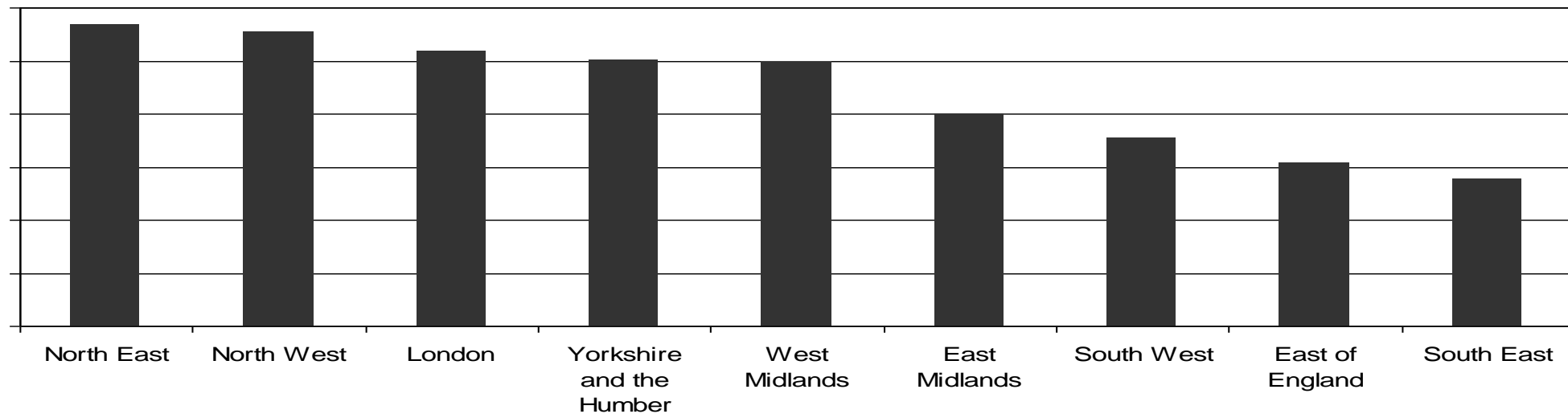
HES Data - Copyright © 2011, Re-used with the permission of The Health and Social Care Information Centre. All rights reserved.

Initial Analysis of HES Data by GOR

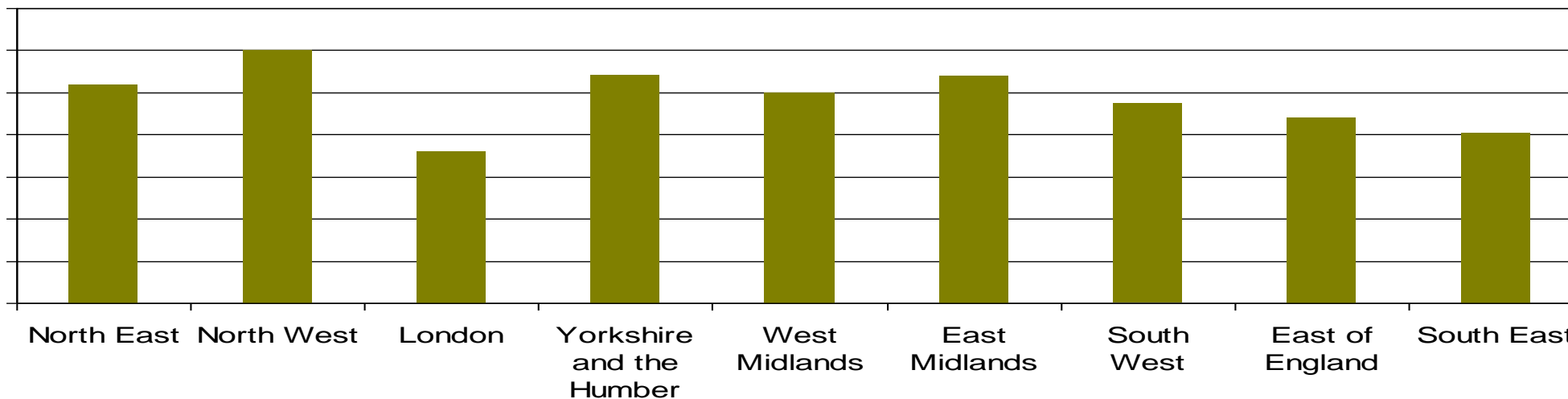
- **England divided into 9 GORs (Government Office Regions)**
 - Each LSOA mapped to a GOR
- **HES Data analysed by GORs**
 - Heart Attack
 - Stroke
 - Cancer
 - Age 30-64
- **Results compared with IMD Ranking for each GOR**

Heart Attack

IMD Scores

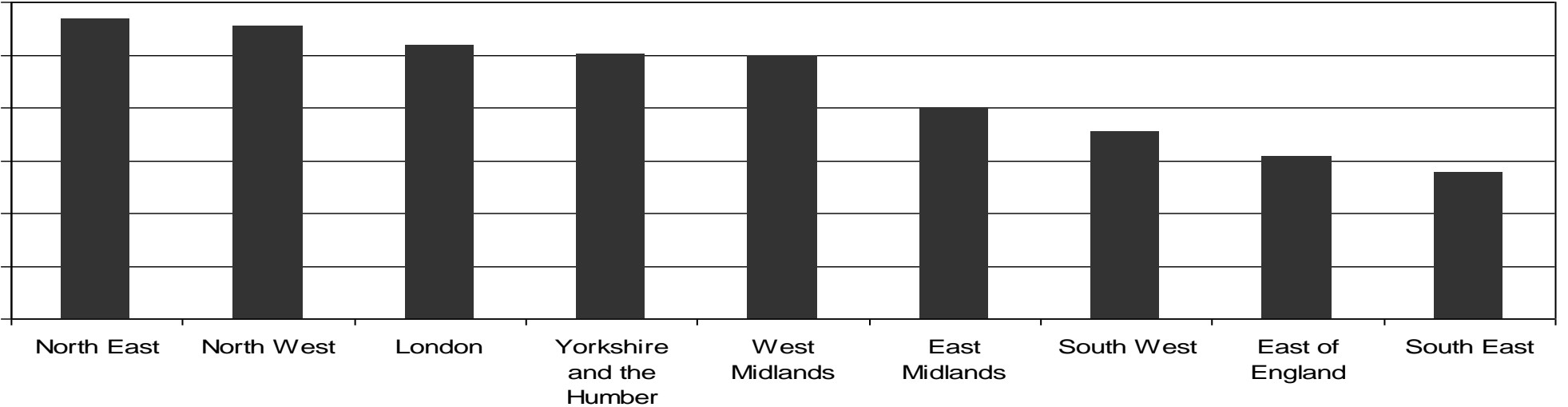


Heart Attacks Age 30-64

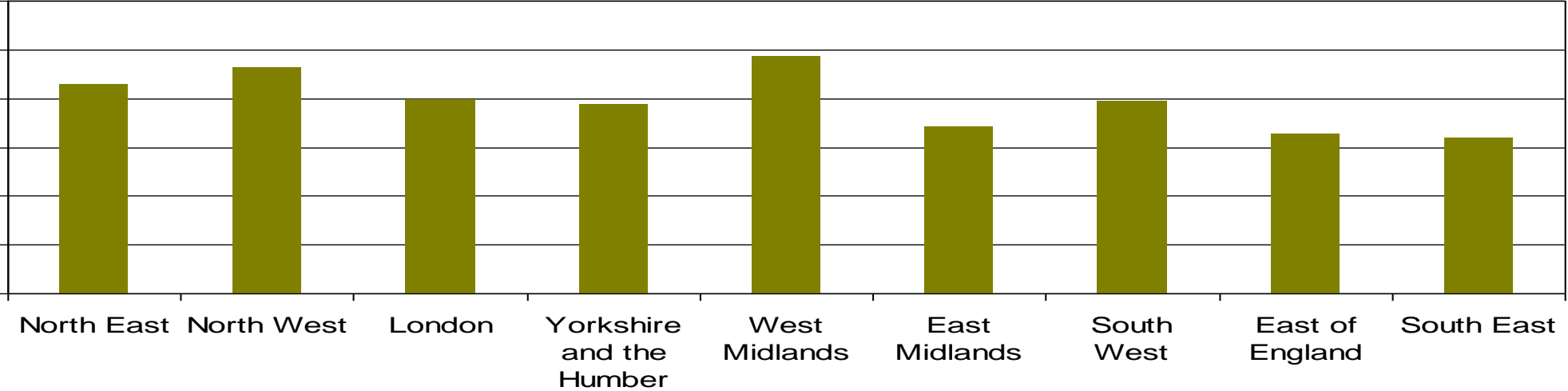


Stroke

IMD Scores

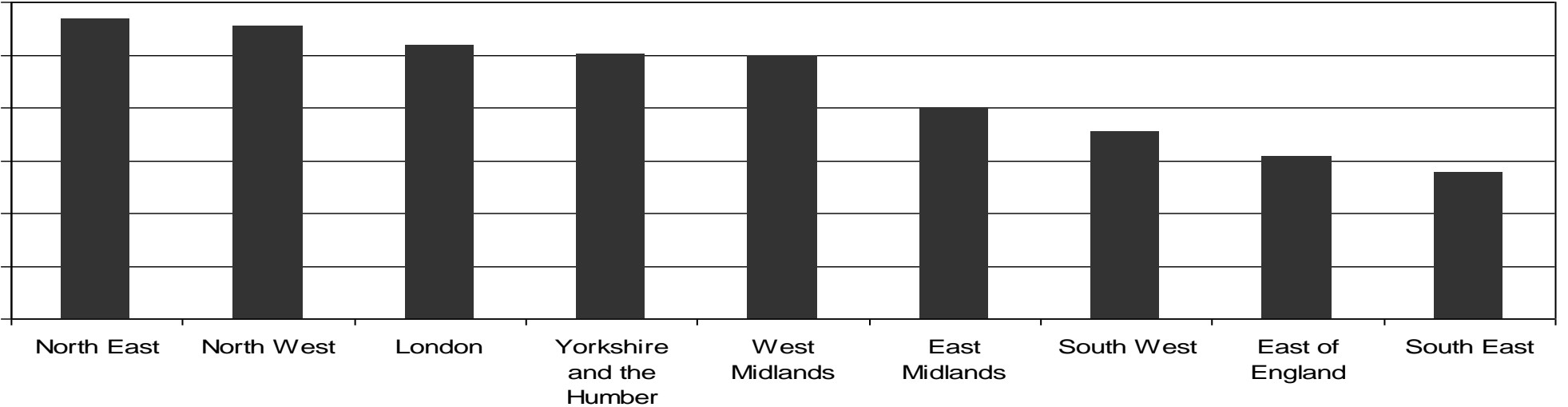


Stroke Age 30-64

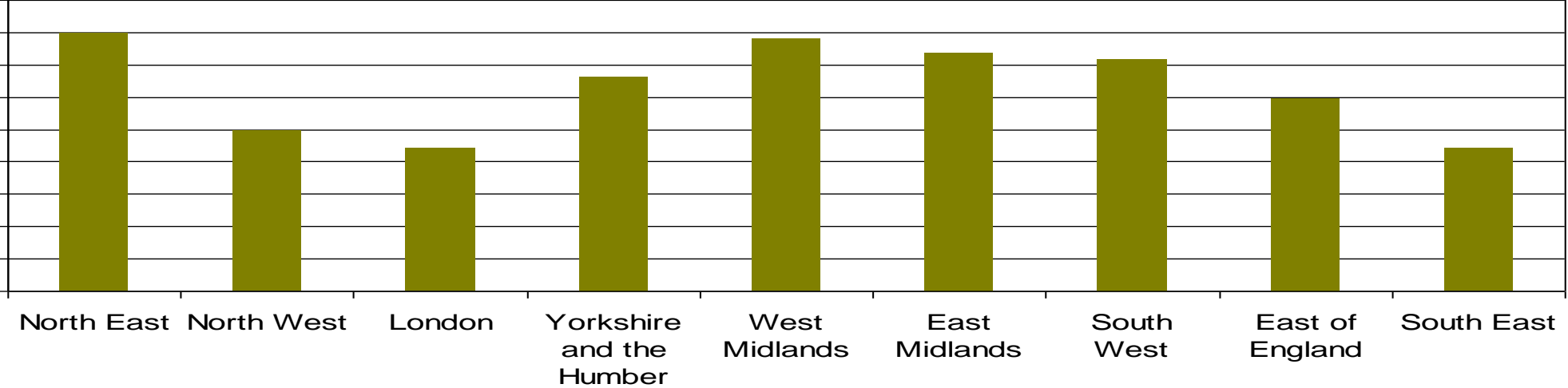


Cancer

IMD Scores



Cancer Age 30-64



Conclusion

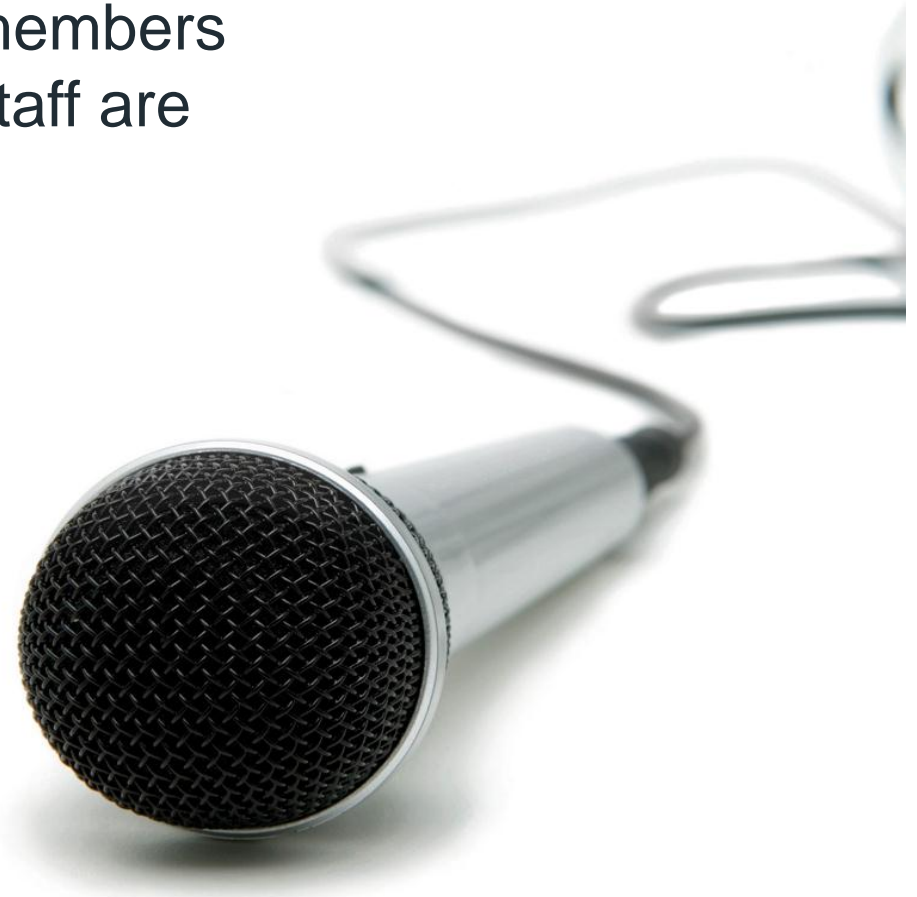
- These proxies can be used to predict Heart Attack, Stroke and Cancer incidence, indeed:
 - Top down approaches showed intuitive groupings were predictive;
 - Bottom up groupings, based on Heart Attack, transferred to other illnesses were also effective
- Results are largely as expected based on last year's work.

Next Steps

- Review cancer labelling algorithm to give primacy to primary diagnosis;
- Extend time series of analysis
 - Check models created on older data are predictive of newer data;
- Refine analysis to identify:
 - First ever events;
 - Impact of underwriting;

Questions or comments?

Expressions of individual views by members of The Actuarial Profession and its staff are encouraged





Appendix: Term of Reference

Taking the Analysis to the Next Level

Objectives and TOR

The Geographical Variations Working Party's main objectives and terms of reference:

- To analyse and understand how the incidence of certain critical illness vary according to geographical location in the UK
- To perform detailed postcode analysis of the impact of socio-economic profile and location on the critical illness (CI) incidence rates
- To build upon work already done in the SIAS papers (Exploring Critical Path, 2006; A Critical Review, 2000)
- Would be interesting to Pricing Actuaries / underwriters / actuary performing experience analysis

Terms of Reference

Regional variations in coverage and incidence rates

- The Working Party's research proposal:
 - to perform a detailed analysis of the impact of socio-economic profile and location on the CI incidence rates for the key CI conditions of Cancer, Heart Attack and Stroke
 - to analyse CI incidence rates using readily available proxies for socioeconomic categories
- The research work will involve analyses of the effectiveness of **the use of postcode in conjunction with geo-demographic profilers** as a proxy for socio-economic profile
- No such analysis has been carried out or at least known to the Working Party at present
- Already analysed and identified available data sources in the UK – HES (Hospital Episode Statistics) and the UK's cancer registries
- Taking the analysis to the next level would need to involve **effective mapping of population data into geo-demographic profilers' categorisations** by socio-economic factors, i.e. into categories and sub-groupings split by social status, wealth etc.

Terms of Reference

International variations in coverage and incidence rates

- The Working Party's research proposal:
 - a comparison of the development and coverage of CI in the established international markets. This will include the UK, South Africa, Australia, Japan, Canada and the USA
 - a high-level comparison of the incidence rates for the CI conditions of Cancer, Heart Attack and Stroke across the main international markets
 - detailed analysis on the impact of the socio-economic profile and location on CI incidence rates (as for the regional variations)
- International comparison to consider differences in the CI cover in each market, the rationale for these products, product features and characteristics
- The Working Party intends to undertake analysis to investigate product reputation at different international markets, including the use and acceptance of tiered benefits, non-disclosure and declined claims