School Of Health And Related Research.

An evaluation of the Summary Hospital Mortality Index

Final report

Michael J Campbell PhD CStat Professor of Medical Statistics

Richard M Jacques PhD Research Associate

James Fotheringham MbChB MRCP Research student

Tim Pearson BSc Research Associate

Ravi Maheswaran MD MRCP FFPH Reader in Public Health

Jon Nicholl DSc CStat MFPH Professor of Health Services Research

21/04/2011

ScHARR

Regent Court

30 Regent St

Sheffield S1 4DA

Executive Summary

- We have constructed a Summary Hospital Mortality Index (SHMI) based on all in-hospital deaths and deaths 30 days from discharge, using HES data from the years 2005/6 to 2009/10 in England.
- The deaths 30 days from discharge were obtained using ONS linked data.
- The linkage was very accurate with only 0.12% of in-patient deaths not correlating with ONS reported deaths. Mortality was 2.6% per spell for in-patient deaths and 3.4% including 30 day mortality.
- The SHMI is a ratio of the observed deaths in a Trust over a period of time divided by that expected, given characteristics of patients treated by the Trust.
- The denominator for the SHMI was calculated using logistic regression, with age-categories and sex as the primary predictors. The logistic regressions were carried out within grouped Clinical Classification System (CCS) diagnostic codes.
- We investigated the effect of Trust ranking by including among the predictors: method of admission (emergency or elective), Index of Multiple Deprivation score based on the post code of the patient's home, the Charlson Comorbidity score (with different coding depths and scoring methods) and number of previous admissions in the previous 12 months. We also looked at interactions between the covariates.
- We found that the SHMIs produced by models with different predictors were all highly correlated. However a model which include age, sex, method of admission and Charlson Comorbidity (full coding depth and categorised) gave a simple and stable model, and the addition of the IMD and/or the number of previous admissions in the previous 12 months on top of these terms did not change the ranking of Trusts to any great degree.
- Including interactions terms between the predictors in the model was difficult and did not improve the predictions.
- Hospitals should be evaluated using a funnel plot. The warning lines should be computed using a random effects model, to acknowledge the unmodelled variation between Trusts. However, some form of trimming should be undertaken, so that outlying Trusts do not over inflate the random effect term in a way that means they no-longer appear outlying.
- The SHMI gave stable ranks from one year to the next, and identified Trusts which had been highlighted as underperforming in other methods of assessing hospital care.
- Direct standardisation is difficult when there are large numbers of structural zeros (casemix groups which have no patients in them) and also for large numbers of variables to standardize by. Direct standardisation by age, sex and diagnosis gave comparable results to the SHMI which was indirectly standardized by the same variables (r=0.962).
- The SHMI has a strong correlation with an index based only on in-patient deaths using 2007/8 data (r=0.938) (using the same predictors).
- The Index calculated using CCS groups (259 collapsed to 137) was also strongly correlated with that using National Centre for Health Outcomes Development (NCHOD) diagnostic groupings (218 collapsed to 166 with cancer incorporated), (r=0.869) but there were a few Trusts whose ranking was markedly changed by this grouping.

- There was a strong seasonal variation in mortality, and short term SHMIs, based on 3 month periods were variable and the correlation from one quarter to the next was about 0.9.
- The SHMI was quite strongly correlated with the HSMR provided by Dr Foster for the years 2009/10 (r=0.753) implying that a significant proportion of hospitals had a different ranking by the two methods.

Contents

1. Background

In 2010 a national Steering Group was established by Sir Bruce Keogh and chaired by Ian Dalton, with the purpose of developing a consensus view of the key methodological requirements for a practical Hospital Standardised Mortality Ratio. In their report (Whalley, 2010) they suggested a new measure, the Summary Hospital-Level Mortality Indicator (SHMI). The main points about the SHMI are: i) that it covers death relating to all admitted patients that occur in all settings, i.e. including those occurring in hospital and 30 days post-discharge, ii) it adjusts as far as possible with HES data for factors outside a hospital's control that might impact on hospital mortality rates, iii) it applies to all NHS acute trust except specialist hospitals

The technical group, a subgroup of the HSMR Steering Group and chaired by Paula Whitty, proposed further work to finalize the model, and this went out to tender in December 2010. The contract was awarded to ScHARR at the University of Sheffield in Feb 2011.

1.1 The new proposals

The Steering group proposed a new ratio, based on observed and expected deaths.

For the numerator: 100% of deaths in hospital and up to 30 days after discharge from hospital (excluding death on arrival and still births). Deaths are to be counted only once, and attributed to the hospital where the patient died or from which discharged within 30 days of discharge. For 'superspells' which is defined as the time a patient enters hospital until they are discharged out of secondary care, and allows for movement between hospitals, the death is to be attributed to the last admitting hospital.

For the denominator: The expected probability of death from all patients (excluding zero length of stay, emergency cases and day cases). It will be based on spells.

Adjustment variables: age, sex admission method (elective/emergency) Diagnosis, comorbidity (Charlson), year of discharge, Deprivation score (IMD) and number of emergency admissions in previous 12 months. Notable omissions in the adjustment variables were palliative care, ethnicity, Hospital type, number and severity of prior morbidities <1 year and <5 years, length of stay and percentage of deaths in the area of the hospital that occur in the hospital.

1.2 Requirements for the Sensitivity analysis

The modelling work requested was to have three objectives i) to undertake a sensitivity analysis of the impact of or interaction between key variables, ii) to illustrate the characteristics and behaviour of the

SHMI using historic data iii) to ensure the SMHI is fully replicable by different teams using a common method and data source.

The sensitivity analysis would:

1) Contrast the SHMI and the effect on Trust ranking by including or excluding specific variables. (age, sex, IMD, comorbidity, number of emergency admissions in past 12 months and method of admission). Procedure (e.g. whether they had an operation) was 'to be discussed'
2) Compare the effect of using direct or indirect standardisation.
3) Test out the approach for i) use of last admitting hospital for the purpose of diagnosis, and ii) handling transfers.
4) Use of different depths of coding for comorbidity.
5) Using the CCS diagnostic groups or the NCHOD groupings.
6) The assessment of the potential for any confounding of including both comorbidity and deprivation in the model.
7) Assessment of interaction between variables (where relevant).
8) The assessment of multi-collinearity between variables (where relevant).
9) To note any anomalies between the HES and linked ONS data.
10) To compare the rates and relative score and rankings of hospitals using all deaths, or in-hospital death only.
11) To illustrate and assess the extent that the hospital score or ranking varies over time.
12) To examine the behaviour of the SHMI over shorter periods of time, e.g. months.
13) To compare the outputs from the SHMI with outputs from other summary level indicators, e.g. to see whether the SHMI could have triggered alarms for Mid-Staffordshire and Basildon and Thurrock.

2. Methodology

2.1 Dr Foster's HSMR

The Dr Foster technical document (Aylin et al, 2010) and Bottle et al (2011) outline the Dr Foster methodology. They describe how the data are linked and the diagnosis derivation. The methods of recoding age, admission method and other covariates are described. For calculating the risks day cases were excluded from the population and if there is more than one spell with the same diagnostic group (CCS) in a superspell, only the first occurring spell was included. They used logistic regression to derive predicted probabilities for in-hospital mortality. They used a backward elimination procedure which removes all non-significant variables (with a cut-off of P>0.1). They recategorised age, deprivation, comorbidity and number of previous admissions so that each category contained at least 20 events for standardisation. Starting with the first lowest) category, they combine it with the next lowest category if it contains less than 20 events and continue combining until at least 20 events have been reached. If the last category is left with fewer than 20 events the categories of the variable are treated as one group.

The HSMR is the ratio of the observed deaths in a Trust to the expected. The observed deaths are all deaths in the 56 diagnostic categories considered by Dr Foster, which comprise about 80% of in-hospital deaths. The expected values are the sum of the risks from the covariates describing that Trust over the 56 diagnostic groups.

They then plot a funnel plot, which is the HSMR against the expected number of deaths. A funnel plot conventionally is an outcome plotted against a measure of precision. Since the precision of an SMR can be measured by the expected value, this is what they have used. Each funnel plot has a line drawn where the process is in control. For a standard population which is the sum of the hospital populations the SMR would be 1. They also plot an upper and lower control limit line, which is defined as the three standard deviations above or below the control line (probability of being within the limits by chance is 99.8%). The control limits from Dr Foster were derived from the Poisson distribution.

2.2 ScHARR's methodology

ScHARR's methodology is essentially the same as that of Dr Foster in that to obtain the expected values we fit logistic regression models within diagnostic group for the whole data set, and then predict the probability of death for each individual in each Trust and sum over diagnostic groups to obtain the expected number of deaths in a Trust, which we can then compare with the observed number. The formulae are given in the Appendix. As a useful check the total sum of the expected deaths over the whole dataset should equal the sum of the observed numbers of deaths.

Thus both Dr Foster's methodology and ScHARR's use indirect standardisation. We are using the rates derived from the population of all admissions to predict what we would expect in a Trust, given its covariates such as the age and sex distribution of its patients. Since the national population in this case comprises simply the accumulation of the Trusts, the overall SMR for any year is defined as 1.

We did not use stepwise logistic regression, but rather fitted the same model to each diagnostic group. We also did not aggregate groups to ensure at least 20 events (deaths) in each category. We use individual case logistic regression (i.e. the dependent variable is 0/1) which does not require aggregation for a model to fit. However it will mean that some models will have too many parameters. For example, for many diseases there are no deaths at young ages and so all the young age categories will be redundant. For ovarian cancer the expected value for men will be zero. However for large datasets parsimony is not a priority, and the advantage of using the same model in every case is that a Trust could use a standard set of covariates and the weights provided by the logistic regression and predict their SHMI. Some computer programs have problems fitting values in the case of zero events but the package we used, R (GNU project, http://www.r-project.org), simply reports that a zero has been fitted. The fitted values will be not be zero, but will be a very small number (10^{-10}) and so cause no problem in computing the sum of the expected values. The regression coefficients will take very large negative values and so of themselves could not be used for prediction. This would not be a problem for a Trust wishing to predict its SHMI because the weight for certain categories and certain diseases would be zero, i.e. we would not expect any deaths in these categories. With gender specific diseases for the year 2007/8 we initially had no reported problems when a term for sex was included in the model, but cross-tabulation revealed that some 60 men were coded as having had ovarian cancer. For 2009/10 there were no men ascribed to female disease and for these diseases we had to drop the term sex.

We use two graphical methods to compare models, the funnel plot and the 'Diffsum' plot described in the Appendix 2. For the funnel plot we used two sets of warning lines. The first set, which follows Dr Foster, assumes a Poisson distribution of the observed deaths, given the expected, and the lines are 2 and 3 S.D.s (P=0.05 and P=0.001) from the assumed target SHMI=1. This assumes that all the remaining variation, given the expected value is Poisson (i.e. there are no other covariates which can account for systematic variation between Trusts). The other set of lines use a random effects model as described by Spiegelhalter (2005b). This allows an 'over-dispersion' parameter to allow for variation between trusts that has not been explained. If all the trusts were included in the estimate then truly outlying trust would inflate the over-dispersion parameter unduly, and may not appear as outliers. We adopt a 'trimming' approach as suggested by Spiegelhalter, of calculating the 'z-score' (O-E)/SE and omitting 10% of trusts at the top of z-score and 10% at the bottom, when calculating the over-dispersion parameter.

3. Data Handling

3.1 Loading

The provided data in ZIP archive form was copied from the encrypted USB hard disk onto a non-networked 64 bit computer. It was then expanded.

3.2 Importing

PASW Statistics version 18 (IBM) was used to import the delimited files into a statistical and data management environment. Loading syntax included the field name descriptors and variable format. Field names from the HES dataset were kept. Appendix 1 includes the loading syntax.

3.3 Converting to relevant formats

In order to determine the secondary diagnosis coding depth, DIAG_02 – DIAG_20 values were converted to the system missing variable if empty. Frequency of diagnosis field usage was then determined, the result of which are displayed in table 1.

Table 3.1: Frequency of diagnosis field usage

DIAG field	Percent Used	DIAG field	Percent Used	DIAG field	Percent Used	DIAG field	Percent Used
Diag_01	100	Diag_06	14.44673	Diag_11	1.620531	Diag_16	0.024613
Diag_02	70.88124	Diag_07	9.477721	Diag_12	1.048553	Diag_17	0.012452
Diag_03	45.1589	Diag_08	6.042003	Diag_13	0.602609	Diag_18	0.008646
Diag_04	31.17836	Diag_09	3.935147	Diag_14	0.166501	Diag_19	0.006037
Diag_05	21.43529	Diag_10	2.508434	Diag_15	0.035697	Diag_20	0.004275

Data containing dates were not in a format suitable for PASW and were converted into a computed field.

3.4 Calculation of comorbidity

The Charlson Index is the most widely used and validated comorbidity measure in the general population and admitted patients. The prognostic implications of the presence of co-morbidities included has changed since the measure was originally developed. As a result the Dr Foster group have re-weighted the individual diagnoses. These weights were employed for comorbidity scores in this analysis. ICD-10 codes, old weights and new weights are listed below.

Table 3.2: Charlson comorbidity groups, ICD-10 codes and weights

	Condition Name	ICD-10 Coding	New	Old
1	Acute myocardial infarction	I21, I22, I23, I252, I258	5	1
2	Cerebral vascular accident	G450, G451, G452, G454, G458, G459, G46, I60-I69	11	1
3	Congestive heart failure	I50	13	1
4	Connective tissue disorder	M05, M060, M063, M069, M32, M332, M34, M353	4	1
5	Dementia	F00, F01, F02, F03, F051	14	1
6	Diabetes	E101, E105, E106, E108, E109, E111, E115, E116, E118, E119, E131, E131, E136, E138, E139, E141, E145, E146, E148, E149	3	1
7	Liver disease	K702, K703, K717, K73, K74	8	1
8	Peptic ulcer	K25, K26, K27, K28	9	1
9	Peripheral vascular disease	I71, I739, I790, R02, Z958, Z959	6	1
10	Pulmonary disease	J40-J47, J60-J67	4	1
11	Cancer	C00-C76, C80-C97	8	2
12	Diabetes complications	E102, E103, E104, E107, E112, E113, E114, E117, E132, E133, E134, E137, E142, E143, E144, E147	-1	2
13	Paraplegia	G041, G81, G820, G821, G822	1	2
14	Renal disease	I12, I13, N01, N03, N052-N056, N072-N074, N18, N19, N25	10	2
15	Metastatic cancer	C77, C78, C79	14	3
16	Severe liver disease	K721, K729, K766, K767	18	3
17	HIV	B20, B21, B22, B23, B24	2	6

ICD-10 codes in the diagnosis fields were linked to the ICD-10 codes relating to the relevant Charlson comorbidity conditions. Charlson conditions could only be flagged once. To determine the impact of the depth of coding, Charlson scores were calculated on DIAG_02 & DIAG_03, DIAG_02 – DIAG_05 and DIAG_02 – DIAG_20. The presence of these conditions for the stated DIAG_XX depths were maintained in the dataset should there be a need to derive new weights.

3.5 Collapsing episodes to spells

The supplied dataset included a discharge date, discharge method,discharge destination and the episode_ID of the first and last episode within a spell. Spells beyond the first were not needed to determine the SHMI, so cases were filters for episodes where the episode_ID was the same as the PROVSPELL_FIRST_ID.

3.6 Removal of daycase & planned same day episodes

Cases coded as CLASSPAT = 2 and CLASSPAT = 3 were removed to exclude daycases and planned same day admissions in the 12 month admission rate and for subsequent analyses.

3.7 Removal of non-acute trusts and private providers

Providers which do not receive acute admissions, or private institutions were removed from the dataset. This was performed by selecting only providers with codes which began with the letter R.

3.8 Summary of record processing

Table 3.3: Summary of record processing

Processing Stage	**Number of records**
Full HES dataset	92,043,618
Converting episodes into spells	81,045,877
Removing daycases	54,366,309
Removing non-acute providers	53,173,352
Removing Planned Same Day admissions	47,172,029

3.9 Calculation of previous admissions in preceding 12 months

The dataset was ordered by patient_ID and increasing PROVSPELL_DISDATE. Using the LAG function in PASW, spells prior to the spell in question were queried to determine if they were within 12 months of the current spell, and counted if relevant. Spells within superspells were counted individually. This process excluded daycases and elective admissions, and counted a maximum of 50 previous admissions.

3.10 Calculation of SUPERSPELL_ID and backfilling superspell diagnosis from last admitting hospital.

With the dataset in the above order, a SUPERSPELL_ID was assigned. If the next spell had an ADMIMETH value of 81 (transferred from another hospital) and the difference in PROVSPELL_DISDATE and ADMIDATE was within +/- 1 day, and had the same patient_ID the SUPERSPELL_ID was copied forward and subsequent spells flagged as part of a superspell.

The dataset was reordered and in reverse and the DIAG_01 from the last spell of a superspell was copied forward, along with the superspell flag.

3.11 Deriving diagnosis groups

A publically available mapping of ICD-10 codes to the ICD-10 Clinical Classification System was imported and mapped to previously reported ICD10 codes used in the HES dataset for primary diagnosis. (http://www.hcup-us.ahrq.gov/toolssoftware/icd_10/ccs_icd_10.jsp) The NCHOD diagnosis grouping was mapped to the same list, using the classification sieve from the original papers appendix. These classifications mapped to all possible ICD10 codes were then joined to the dataset based on the first 4 characters of the DIAG_01 and SUPERSPELL_DIAG_01 fields for each spell.

The provided mapping of ICD10-CCS left approximately 1.3% of admissions uncoded, and these were transferred into CCS group number 259 – “other codes unclassified”. The NCHOD grouping does not include cancer, and therefore a larger proportion (6.8%) were uncoded. ICD10 codes starting with “C” were grouping into a new malignancies group with any remaining being coded as a new “other codes unclassified group”. This ensured that all deaths had a diagnostic group.

3.12 Changes in spell diagnoses using the last admitting hospital's diagnosis

Cross-tabulating the spell and superspell diagnosis for each record identified that in 33.4% of spells the reason for admission under the ICD-10 CCS grouping scheme changed. This generally reflected:

- **A better understanding of the presenting diagnosis**
  (i.e. Abdominal Pain becomes Appendicitis)
- **Reflect the outcome and rehabilitation needs of the presenting diagnosis**
  (i.e. Acute Cerebrovascular disease becomes Paralysis)

3.13 Linkage of HES dataset to the ONS dataset

The provided ONS file was expanded and imported into SPSS. Using the Add Variable command, the ONS fields were joined to the HES records using the provided common identifier field. Deaths were then assigned to the final spell within the dataset. The dataset was reordered by patient_ID and descending PROVSPELL_DISDATE, and the ONS dataset was linked on Patient_ID. If the difference between PROVSPELL_DISDATE and ONS supplied date of death was less than 30 days, the difference was recorded and a death flagged. The dataset was then scanned to determine the spell with the lowest value of this difference, and other spells had this value and the death flag set blank.

In addition, inpatient deaths determined by the date of death from the ONS data, the ADMIDATE and PROVSPELL_DISDATE were calculated to determine inaccurate coding of the PROVSPELL_DISMETH variable.

3.14 Analysis of HES-ONS linkage

A cross-tab query of Inpatient death versus DISMETH has been run to provide an indication on level of inpatient death undercoding. The output from this query is shown in table 1.

Table 3.3: Comparison of ONS coding with HES DISMETH codes

Discharge Method (PROVSPELL_DISMETH) **From Hospital Episode Statistics**			**In Patient Mortality (ONS)** Survived	Died	Total
Discharged	1	Count	51230220	16393	51246613
		%	100.0%	.0%	100.0%
Self Discharged	2	Count	613329	259	613588
		%	100.0%	.0%	100.0%
Discharged by a mental health review tribunal	3	Count	10897	68	10965
		%	99.4%	.6%	100.0%
Died	4	Count (Inpatient Death)	1042	1230629	1231671
		%	.1%	99.9%	100.0%
Baby was still born	5	Count	14022	335	14357
		%	97.7%	2.3%	100.0%
Not applicable: patient still in hospital	8	Count	29884	205	30089
		%	99.3%	.7%	100.0%
Not known: a validation error	9	Count	25868	201	26069
		%	99.2%	.8%	100.0%
Total		Count	51925262	1248090	53173352
		%	97.7%	2.3%	100.0%

Thus only 0.1% of deaths coded by DISMETH failed to appear as inpatient deaths on the ONS coding

It is understood that the HES dataset and the ONS dataset were linked using NHS Number, prior to being pseudo-anonymised by the NHS Information Centre. An NHS numbers often have to be traced or assigned when a patient is admitted there is scope for error. To explore this, dates of death from ONS were compared to HES deaths to identify any extreme anomalies.

Table 3.4: Anomalous deaths in the combined ONS-HES dataset

	ONS DoD > 1000 days before HES DoD	ONS DoD 10 - 1000 days before HES DoD	ONS DoD -10 to +30 days within HES DoD	Percentage Perceived Error
HES Year 07-08	239	209	287198	0.2%
HES Year 08-09	329	219	292326	0.2%
HES Year 09-10	496	268	280358	0.3%

It is not possible to determine whether ONS deaths occurring after discharge are anomalous or not. The proportion of erroneous linkages appears to increase as the analysis moves forward through time, but a proportion of this is will be an artefact of patients identified dead by ONS being recurrently admitted, in addition to new erroneously linked patients being admitted. The percentage is sufficiently small to not impact on a SHMI result.

4. Preliminary analysis

We carried out a preliminary analysis of the HES data for the five financial years April 2005-March 2010. We looked at the basic demographics and the factors influencing the 30 day death rate

4.1 Age, sex and death rates

Table 4.1: 30 day death rate by sex

			Men		Women	
			Died	Total	Died	Total
Age Group	<1	Count	9605	2467573	7772	2201222
		% within age group	0.4%	100%	0.4%	100%
	1-4	Count	991	896716	848	666528
		% within age group	0.1%	100%	0.1%	100%
	5-14	Count	1157	969323	1016	785484
		% within age group	0.1%	100%	0.1%	100%
	15-24	Count	3132	1216548	2036	3615454
		% within age group	0.3%	100%	0.1%	100%
	25-34	Count	5117	1279081	3900	5379284
		% within age group	0.4%	100%	0.1%	100%
	35-44	Count	13280	1699073	11401	3230349
		% within age group	0.8%	100%	0.4%	100%
	45-54	Count	31300	1896401	26146	1943750
		% within age group	1.7%	100%	1.3%	100%
	55-64	Count	79817	2472436	58953	2174243
		% within age group	3.2%	100%	2.7%	100%
	65-74	Count	156180	2839569	113961	2466798
		% within age group	5.5%	100%	4.6%	100%
	75-84	Count	269431	2700915	258229	2964397
		% within age group	10.0%	100%	8.7%	100%
	>85	Count	195447	1114559	329779	2090663
		% within age group	17.5%	100%	15.8%	100%
	Missing	Count	878	33758	701	54040
		% within age group	2.6%	100%	1.3%	100%
Total		Count	766335	19585952	814742	27572212
		% within age group	3.9%	100%	3.0%	100%

In table 4.1 we see that from the ages of 15 to 49 female patients account for more spells, explained largely by maternity admissions and accounting for low mortality rates in these age ranges. From age 80 and above there are also more female spells than male.

There were relatively few missing values for age and the death rate for data were age was missing was slightly below average, suggesting that either younger people are more likely to be to have age missing or that if age is missing, the spell is more likely to be coded as having finished alive.

For sex 0.03% was coded missing, but two thirds of these were infants under one year.

Figure 4.1: 30 Day death rate by age group

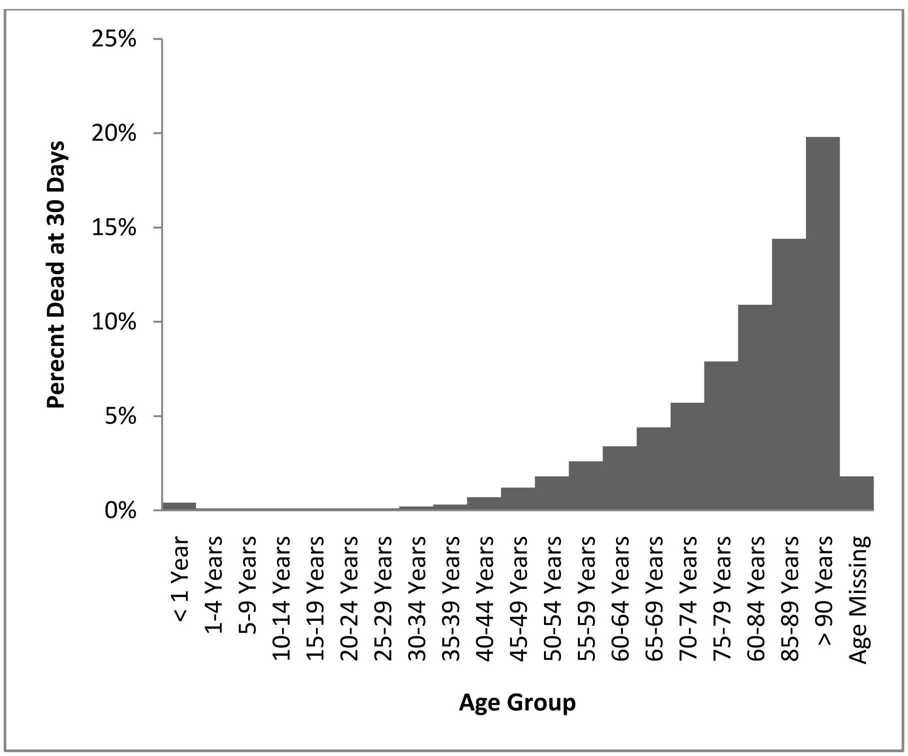


Figure 4.1 shows the 30 day death rate by age group, which shows almost exponential rise with age beyond 40-44 with a 20% mortality rate in the over 90 year olds

4.2 Index of multiple deprivation score

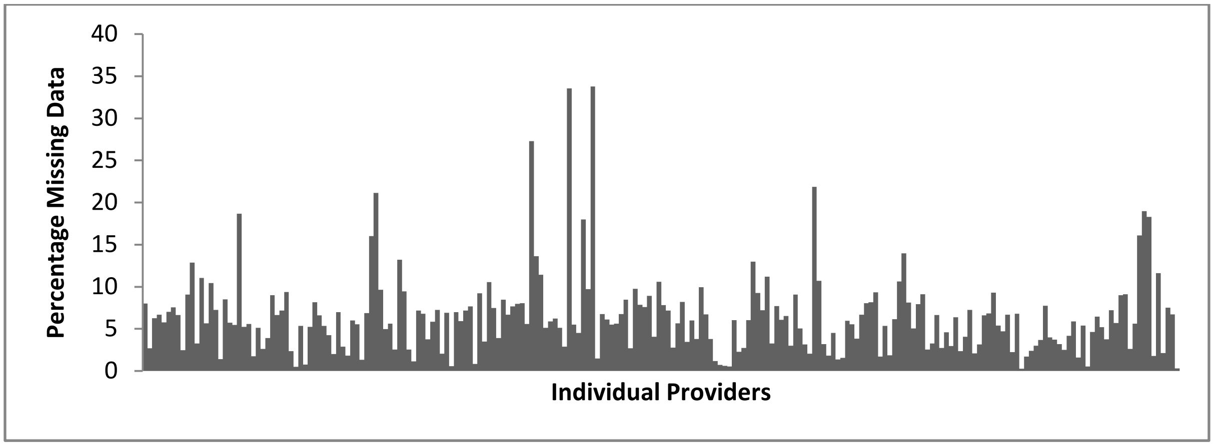
Table 4.2 shows that 30 day mortality increases with the index of multiple deprivation score, except that the most deprived have very slightly lower mortality than those in the group above with less deprivation. Those with a missing IMD also have a very low mortality, which suggests that if an individual has a missing IMD score they are more likely to be in a group with low mortality, such as babies, or specialist Trusts. However there were relatively few deaths in this category and so it was not necessary to try any special techniques such as imputation for these data.

Table 4.2: 30 day mortality by quintile of IMD (deprivation) score

	30 day deaths	Total admissions	% deaths
1st fifth (least deprived)	347413	11525763	3.01
2nd fifth	328924	9391234	3.50
3rd fifth	318454	8376898	3.80
4th fifth	301348	7688844	3.92
Top fifth (most deprived)	259822	6956512	3.73
Missing	25397	3232779	0.79
Total	1581358	47172030	3.35

Figure 4.2 shows the percentage of missing IMD data by Trust and shows that there is a wide variation in the reporting of missing values. Similar variation is observed when reviewing individual years, and missing data was not seen to improve with more recent HES returns.

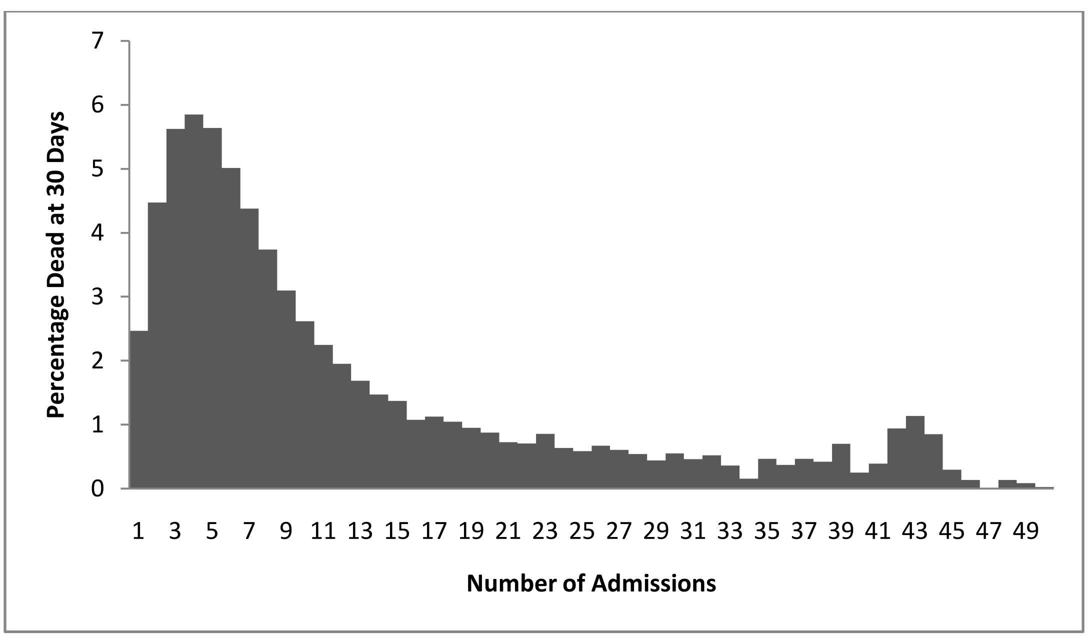
Figure 4.2: Missing data for IMD rank (all years)



4.3 Number of emergency admissions in previous 12 months

Figure 4.3 shows the death rate by number of previous acute admissions, which shows that for up to 5 previous acute admissions the death rate rises but then it steadily falls.

Figure 4.3: Number of emergency admission in previous 12 months and 30 day death rate

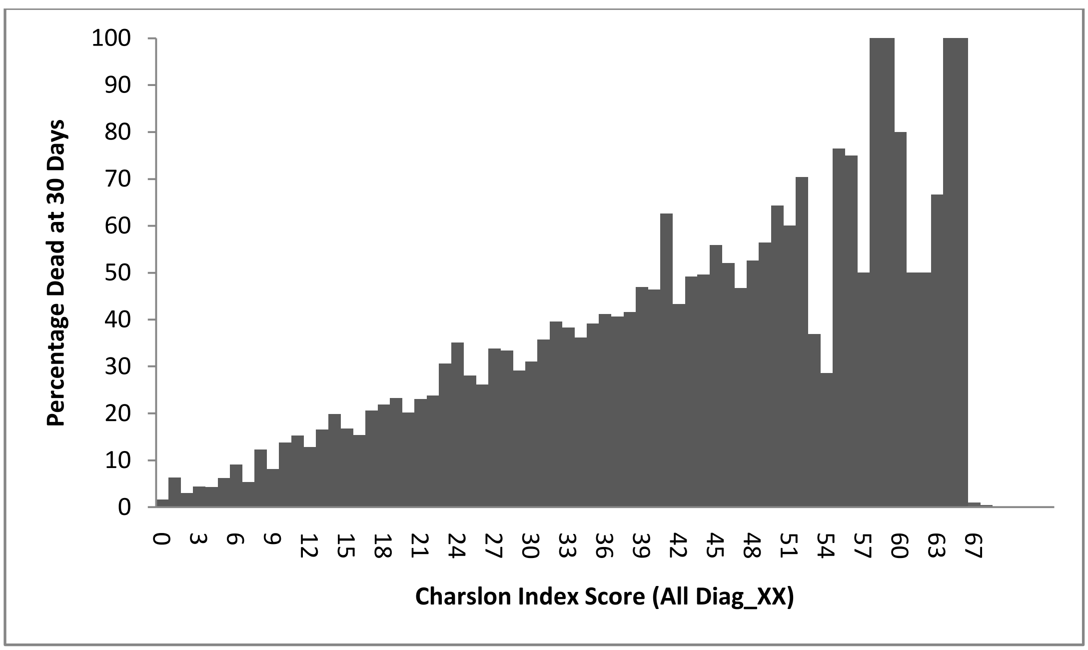


As this is a spell based rather than patient based analysis, a likely explanation is that for example a patient with 5 previous admissions is likely to appear also as a person with 4 previous admissions in an earlier spell and survived to have a further admission.

4.4 Charlson comorbidity index (All secondary HES diagnosis fields)

Figure 4.4 shows that an almost monotonic rise in the 30 day death rate by the Charlson Index of comorbidity (HES diagnosis fields DIAG_02 – DIAG_20). Observed fluctuation may be explained by weighting of individual Charlson diagnosis groups.

Figure 4.4: Charlson index and 30 day death rate



The distribution of patients' Charlson score was heavily skewed, as at least 77% of episodes scored zero at the various depths of coding explored. For this reason we were unable to convert the score into categories with equal numbers of spells. The distribution of categories for various coding depths is shown in table 4.3.

Table 4.3: Percentage of spells with categorised Charlson comorbidity scores according to depth of diagnosis coding

Charlson Score	Diag_02 & Diag_03	Diag_02 - Diag_05	All Diag_XX
0	84.6%	79.4%	77.9%
1 to 5	8.2%	10.1%	10.2%
> 5	7.2%	10.5%	11.9%

4.5 Type of admission

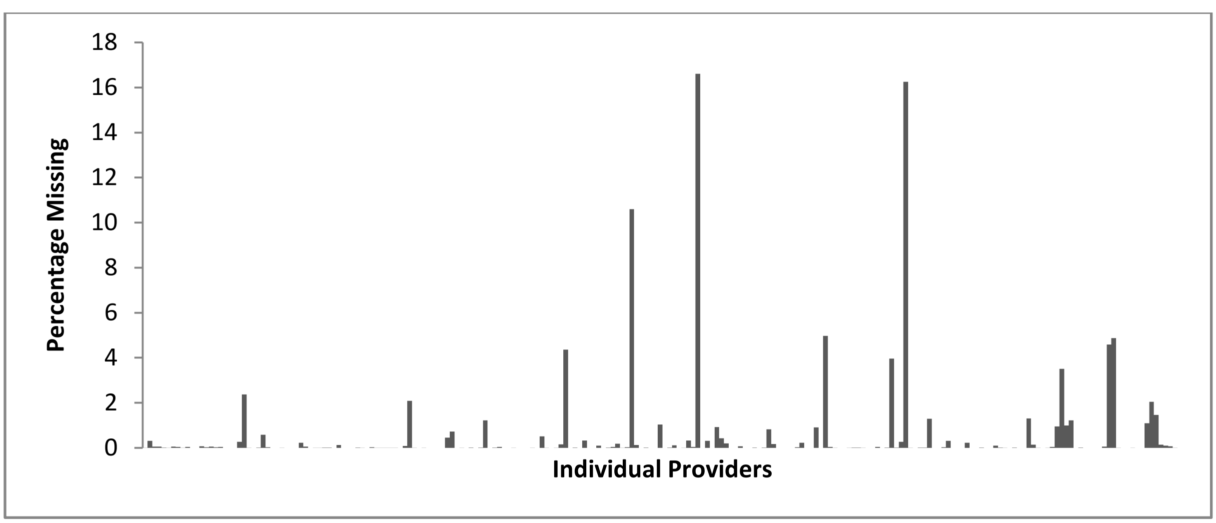
Table 4.4 shows that about one third of admissions are elective and the 30 day death rate for the elective is much smaller at 0.8% than that for emergency admissions

Table 4.4: 30 day mortality by type of admission (including 0 day length of stay)

		Died (%)	Total
Type of Admission	Emergency	1503480 (4.0%)	37443566
	Elective	77028 (0.8%)	9669896
	Missing	850 (1.5%)	58568
Total		1581358 (3.4%)	47172030

Figure 4.5 shows that there are at least 3 Trusts with relatively high missing values for type of admission. Since the numbers were small, special measures were not required to deal with missing type of admission.

Figure 4.5: Missing data for type of admission (all years)



4.6 HES in-hospital deaths by ONS 30 day mortality

Table 4.5 compares the 30 day death rate obtained from ONS with the in-patient death rate as coded by HES. In-patient deaths were 2.6% of spells and total deaths were 3.4% of spells. Of the 30 day deaths about 21% occurred out of hospital. Note there were 1490 spells which were coded as deaths by HES (0.12% of all inpatient deaths) but which were classified by ONS as having survived 30 days. Since some patient matching is done probabilistically where NHS number is not available, it is likely that these are in fact matching errors.

Table 4.5: Comparison of ONS deaths as inpatient or 30 days from discharge with in-patient deaths derived from HES

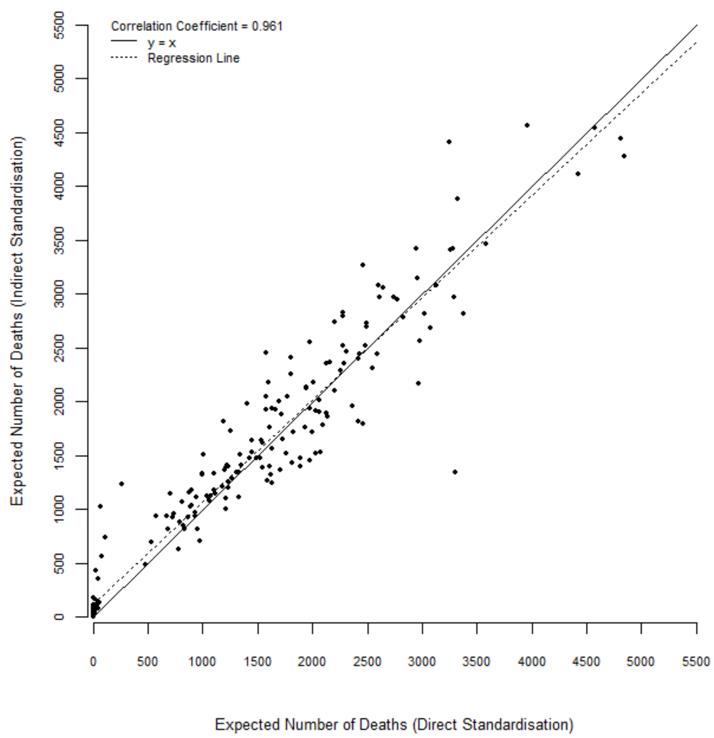
			Deaths 30 days (ONS)		
			Survived	Died	Total
In-patient Death HES	Survived	Count	45589182	335392	45924574 (97.4%)
	Died	Count	1490	1245966	1247456 (2.6%)
Total		Count	45590672	1581358	47172030
		Percentage	96.6%	3.4%	100.0%

5. Direct vs indirect standardisation

We compared standardising for age, sex and diagnosis using direct and regression based indirect methods. We restricted this analysis to these three variables because direct standardisation is difficult when there are large numbers of structural zeros (casemix groups which have no patients in them) and also for large numbers of variables to standardize by.

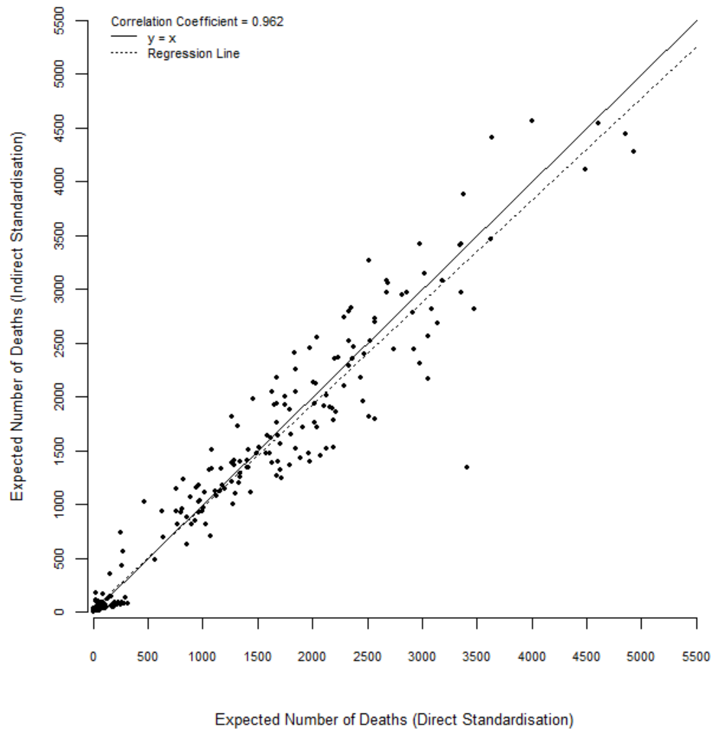
Figures 5.1 and 5.2 show scatter plots comparing indirect standardisation with two methods of direct standardisation. The first method of direct standardisation does not adjust for the structural zeros and the second method re-calculates the weights to allow for structural zeros. Details of both of these methods are given in Appendix 3.

Figure 5.1: Comparing direct and indirect standardisation (no adjustment for zeros)



Both Figures 5.1 and 5.2 show that unadjusted and adjusted direct standardisation give comparable results to indirect standardisation with Spearman correlation coefficients of 0.961 and 0.962 respectively.

Figure 5.2: Comparing direct and indirect standardisation (adjustment for zeros)



6. Sensitivity analysis

6.1 Main effects

We found it impossible to fit models to all five years of data because of the sheer quantity of data. We chose the year 2007/8 since it was mid-way through the data set therefore allowing forward and backward comparison, and investigated model fit for that year. It is also likely that annual models are to be the most useful since often institutions are monitored annually. As described in the Appendix, we fitted models within grouped diagnostic categories. We took age and sex to be the baseline model and then looked at adding covariates to the model and recalculating the SHMI. Table 2, which gives the age/sex death rates for the whole population and shows that relative death rates for men and women vary by age, the main differences occur at younger ages when death rates are low, and so we included age and sex as main effects only.

The models are defined in Table 6.1

Table 6.1: Model definitions

Model Number	Description
Model 1	Age and Sex
Model 2	Age, Sex and Admission Method
Model 3	Age, Sex and comorbidity (diag0203 – continuous variable)
Model 4	Age, Sex and comorbidity (diag02030405 – continuous variable)
Model 5	Age, Sex and comorbidity (diagALL – continuous variable)
Model 6	Age, Sex and IMD
Model 7	Age, Sex and Number of Emergency Admissions
Model 8	Age, Sex and comorbidity (diag0203 –categorical variable)
Model 9	Age, Sex and comorbidity (diag02030405 – categorical variable)
Model 10	Age, Sex and comorbidity (diagALL – categorical variable)
Model 11	Age, Sex, Admission Method and IMD
Model 12	Age, Sex, Admission Method and Number of Emergency Admissions
Model 13	Age, Sex, Admission Method and comorbidity (diag0203 – categorical variable)
Model 14	Age, Sex, Admission Method and comorbidity (diag02030405 – categorical variable)
Model 15	Age, Sex, Admission Method and comorbidity (diagALL – categorical variable)

We calculated a Spearman correlation coefficient between model 1 (allowing for age and sex) and subsequent models and these are given in Table 6.2. A model with a lower correlation would indicate more discrimination between Trusts. We can see the Model 2 (allowing for admission method) has the lowest correlation followed by the Comorbidity Charlson Score (full diagnosis depth) either as a continuous variable or a categorical one (into three groups). We used the categorical one to allow for possible non-linearities in the relationship with death rate. Table 6.3 shows the correlation coefficients of the SHMIs from model 2 with models which include admission method AND other covariates. One can see that the Comorbidity (all diagnoses) has the lowest correlation with model 2.

Table 6.2: Spearman correlation coefficients of the SHMI for various models with Model 1

	Model 2	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**	**Model 9**	**Model 10**
Correlation	0.904	0.982	0.961	0.935	0.989	0.997	0.980	0.959	0.935

Table 6.3: Spearman correlation coefficients of SHMI for various models with Model 2

	Model 11	**Model 12**	**Model 13**	**Model 14**	**Model 15**
Correlation	0.993	0.999	0.985	0.971	0.951

Tables 6.2 and 6.3 suggest that the two most influential covariates are method of admission and comorbidity (all diagnosis depth, categorical variable) which is model 15. We then explored combinations of covariates using 'diffsum' plots and funnel plots and these are shown in Appendix 3. The diffsum plots the difference in the expected values (in this case deaths) between model 1 (with just age and sex) and model 15 against the mean of the expected values. There are two guide-lines. The curved one shows the limits of random variation and the straight one shows the value where the SMHI would be expected to change by 5% if model 15 were adopted rather than model 2. Points above the upper straight line show Trusts whose SHMI would be expected to increase by at least 5% if the covariates were included in the model and points below the bottom line would show Trusts whose SHMI would be expected to decrease by 5% if the covariates were included in the model. The diffsum plots in the Appendix show that the IMD deprivation score and number of previous emergency admissions do not change the SHMI markedly for Trusts (as might be expected from the correlation coefficients). Thus we believe model 15, which includes age, sex, admission method and Charlson Comorbidity categorized into fifths is a reasonable model. Figure 6.1 07/08 shows the diffsum plot for this model relative to one with just age and sex. One can see that there are about 15 Trusts whose SHMI would increase by 5% if model 15 were adopted and a somewhat larger number whose SHMI would decrease if model 15 were adopted compared with a model which just allowed for age and sex.

Appendix 5 shows the ranking of the 30 Trusts (with >100 deaths in 2007/8) with the highest SHMI under the different models and shows how they change rank under different models. In general changes in rank are not very extreme, especially for Trusts at the top of the lists. The Trusts with the highest SHMIs are those with small expected numbers of deaths and this expected number is not unduly affected by different predictor variables.

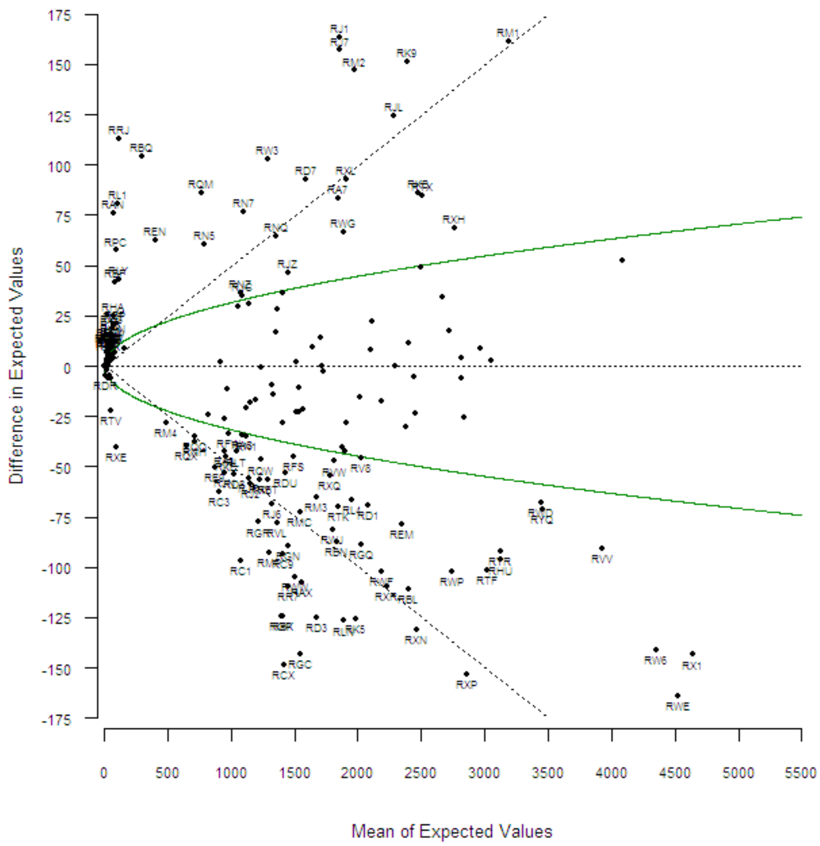
6.2 Predictive ability of the models

Logistic models are often assessed using the c-statistic. This is the area under the Receiver Operator Characteristic (ROC) Curve, but is interpreted as follows. Suppose the model produces a score, it may be a probability of death or it could be a simple linear discriminator. There are two groups: those who die and those who do not. Randomly choose a member of each group. The c-statistic is the probability that the score for the member in one group is higher than the score for the member in the other. Thus a c-statistic of 0.5 means the models have no predictive power, and a c-statistic of 1 means they have perfect predictive power. To discriminate between Trusts we would need something in between, since if we could predict deaths perfectly, the SHMI would always be 1. We computed c-statistics for each diagnostic group and calculated the mean value. The mean c-statistic for model 1 was 0.763 (range 0.515 to 0.958) The mean c-statistic for model 15 was 0.830 (range 0.534 to 0.970). This means model 15 increases the area under the ROC curve by about 0.07.

6.3 Interactions

We looked at interactions between age and admission method and age and comorbidity. We did this by including an extra term 'age x admission method' or 'age x comorbidity'. Because there are 20 age coefficients we restricted the age term in the first instance to age>65 or age<65 in the interaction term. This reduced the number of parameters and so increased the power and interpretability. It is not usual practice to examine interactions for variables when the main effect is not important, and this enabled us to cut down on the number of potential sets of interactions to consider. In both admission and comorbidity there was a high degree of collinearity and for many diagnoses the model fitting process did not converge. Thus we concluded that including interactions in the models was not a practical process and did not include any in the final model.

Figure 6.1: 07/08 Final Model – Difference in expected vales versus the sum (Model 1 v Model 15)



6.4 Funnel plots Poisson errors or random effects

Figure 6.2 shows a funnel plot of the SHMI using model 1 and Poisson error limits. One can see that the limits appear too narrow, since many points appear beyond the 99.9% limits. Figure 6.3 shows exactly the same points but with a random effect model used to derive the error limits. The rationale here is that there are a number of risk factors which contribute to the variation between trusts that have not been measured and are not included in the model, so the variation is likely to be greater than that predicted by a Poisson distribution. Figure 6.3 suggests that there are a number of Trusts with SHMIs above the 95% limit but only one (RDD Basildon and Thurrock)) above the 99.9% limit. Figure 6.4 shows the random effects funnel plot for model 15. This now shows that not only Trust RDD but also Trust RXL (Blackpool, Fylde and Wyre) are above the 99.9% limit.

Figure 6.2: 07/08 Age & Sex only – Poisson Funnel Plot

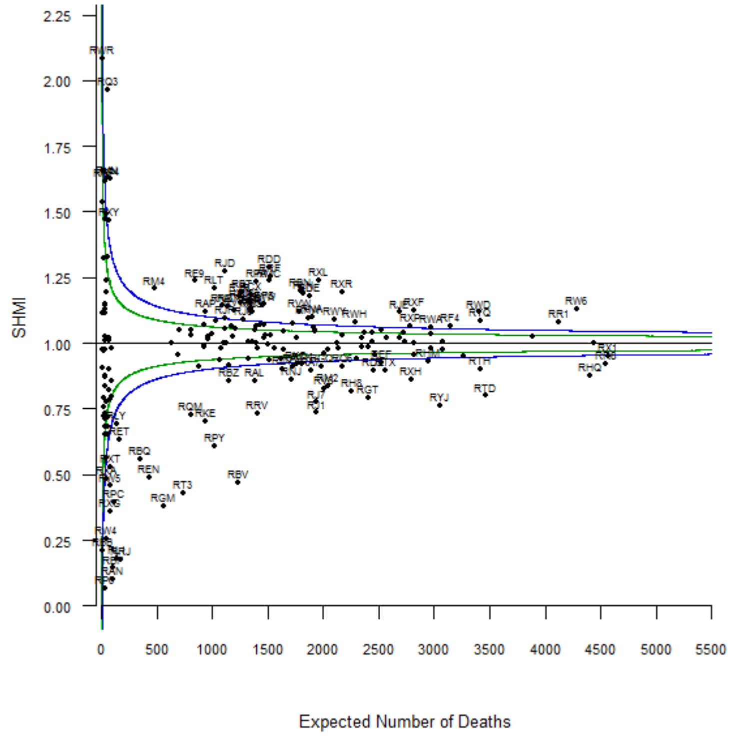


Figure 6.3: 07/08 Age & Sex only – Random Effects Funnel Plot

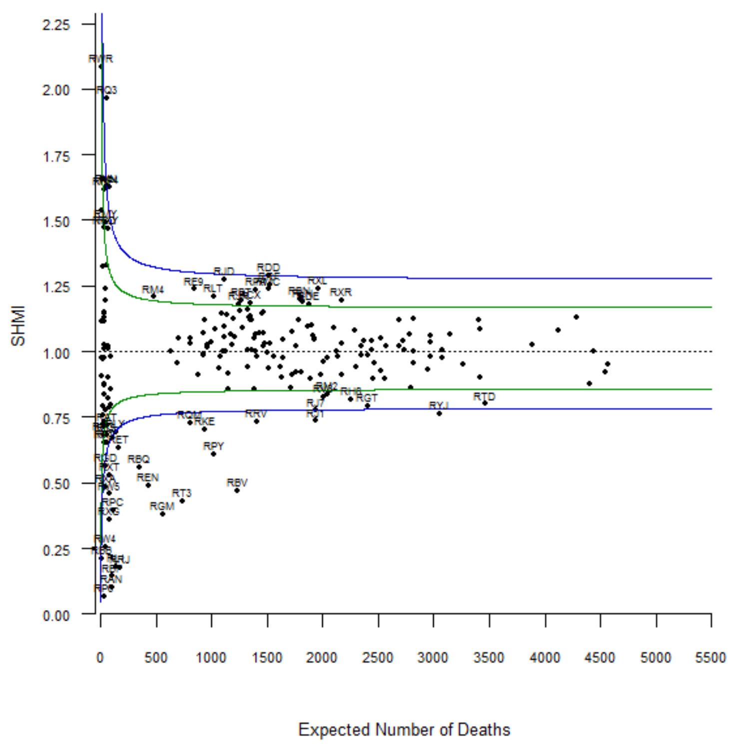
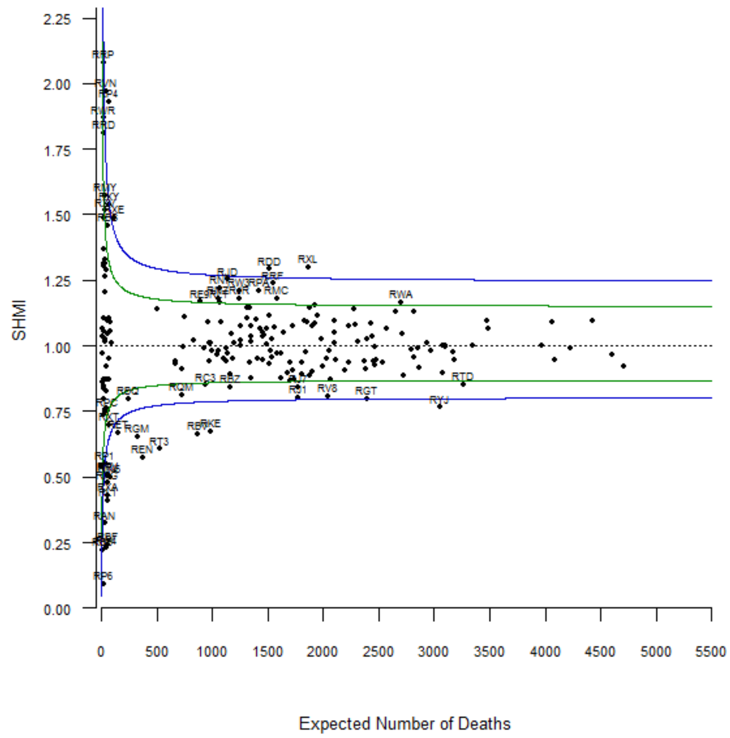


Figure 6.4 07/08 Funnel plot model 15, age, sex, type of admission and comorbidity – Random Effects



6.5 Effect by year

We then fitted model 15 to each year in turn. Table 6.4 gives the rank correlations of SHMI of Trusts by year.

As one can see, for all Trusts, the correlation from one year to the next is fairly constant at about 0.75. It drops steadily so that by 4 years the correlation is about 0.575. For larger Trusts the correlation in general is higher from year to year at about 0.85, and by 4 years has dropped to 0.625.

Table 6.4: Correlation between years – all matched trusts (216)

	SHMI 05/06	SHMI 06/07	SHMI 07/08	SHMI 08/09	SHMI 09/10
SHMI 05/06	1.000	0.778	0.637	0.640	0.575
SHMI 06/07	0.778	1.000	0.727	0.675	0.631
SHMI 07/08	0.637	0.727	1.000	0.747	0.653
SHMI 08/09	0.640	0.675	0.747	1.000	0.774
SHMI 09/10	0.575	0.631	0.653	0.774	1.000

Table 6.5 Correlation between years – matched trusts with >100 observed deaths (155)

	SHMI 05/06	SHMI 06/07	SHMI 07/08	SHMI 08/09	SHMI 09/10
SHMI 05/06	1.000	0.832	0.740	0.654	0.625
SHMI 06/07	0.832	1.000	0.863	0.799	0.741
SHMI 07/08	0.740	0.863	1.000	0.791	0.746
SHMI 08/09	0.654	0.799	0.791	1.000	0.869
SHMI 09/10	0.625	0.741	0.746	0.869	1.000

We then used funnel plots to look at Trusts which were beyond the 99.9% limit in each year, in Figures 6.5-6.9.

- In 05/06 there were 4 outliers: RJD (Mid Staffs), RLT (George Elliot), RAP (North Middlesex) and RNQ (Kettering).
- In 2006/7 there were 2 RJD (Mid Staffs) and RXL (Blackpool, Fylde and Wyre).
- In 2007/8 there were two RDD (Basildon and Thurrock) and RWA (Hull and East York).
- In 2008/9 there were three RDD, RML (Royal Bolton) and RXL, and
- In 2009/10 there were two, RXL and RWA.

Thus in general the model appears stable and Trusts are consistently ranked, in that Trusts above a warning line in one year are often Trusts above a warning line in a second year.

Figure 6.5 Funnel plot of SHMI with random effects warning lines for 2005/6

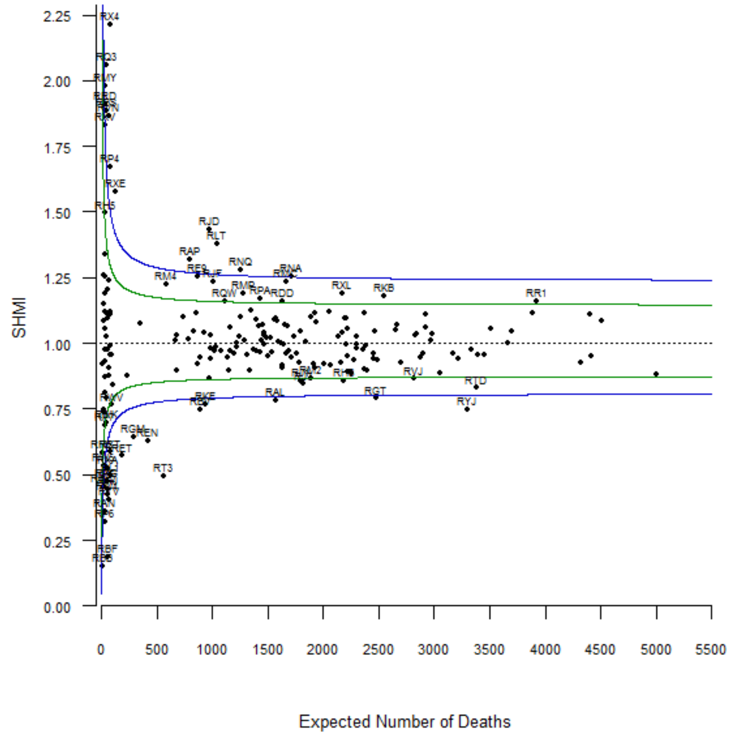


Figure 6.6 Funnel plot of SHMI with random effects warning lines for 2006/7

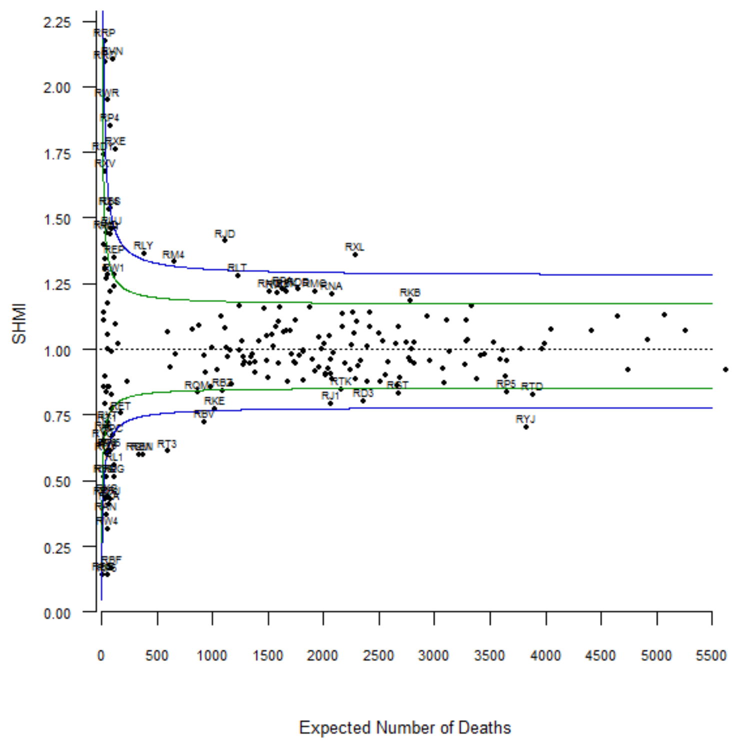


Figure 6.7 Funnel plot of SHMI with random effects warning lines for 2007/8

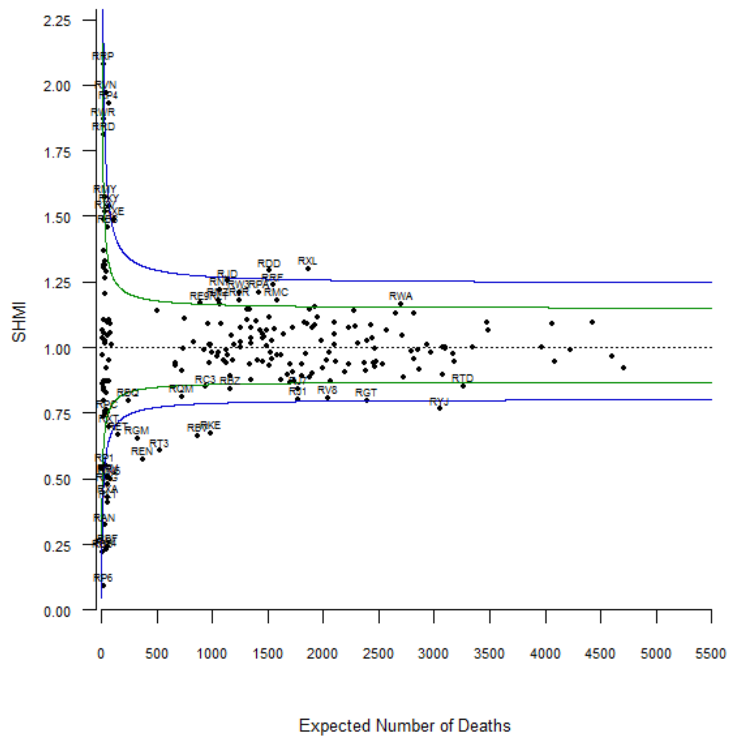


Figure 6.8 Funnel plot of SHMI with random effects warning lines for 2008/9

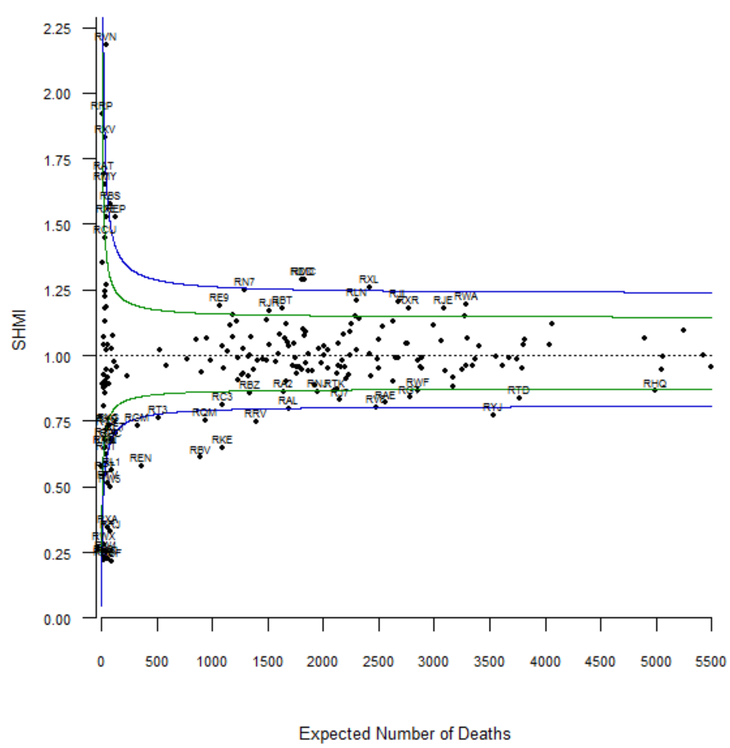
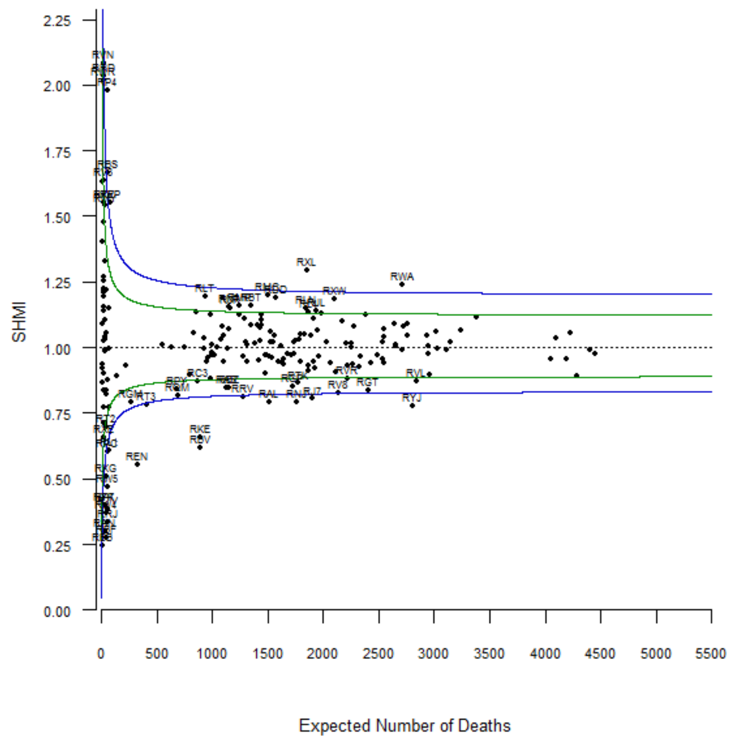


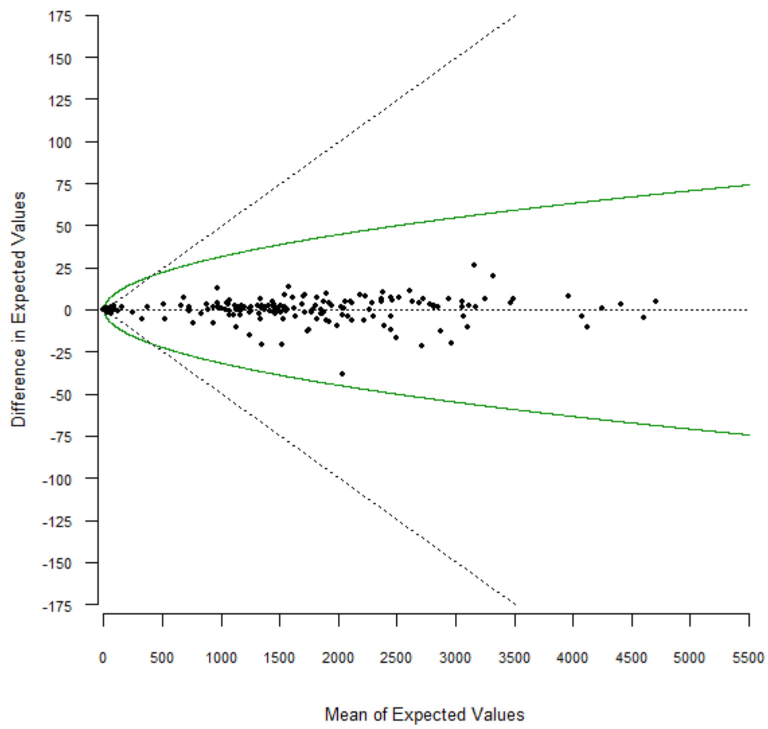
Figure 6.9 Funnel plot of SHMI with random effects warning lines for 2009/10



7. Primary diagnosis derivation from superspells rather than spells

Figure 7.1 shows the diffsum plot showing model 15 based on admission diagnosis derived from the final admitting hospital in superspell admissions. Those admissions which were not part of superspells were not altered and included in the analysis. It can be seen that there is little effect on Trust ranking of using superspell diagnosis rather than individual spell diagnosis.

Figure 7.1: Diffsum plot of difference in expected values versus the mean



8. Effect of NCHOD coding

Figure 8.1 shows the funnel plot for 07/08 for Trusts when diagnosis groups are based on the 218 NCHOD groupings which were subsequently collapsed to ensure adequate numbers of events as detailed in the appendix. Figure 8.2 shows a scatter plot of the SHMI using the CS codes and the NCHOD groups. The correlation is quite high at 0.869, but there appear to be four Trusts whose ranking would be changed to some extent if the NCHOD diagnostic grouping were to be used.

Figure 8.1: NCHOD 07/08

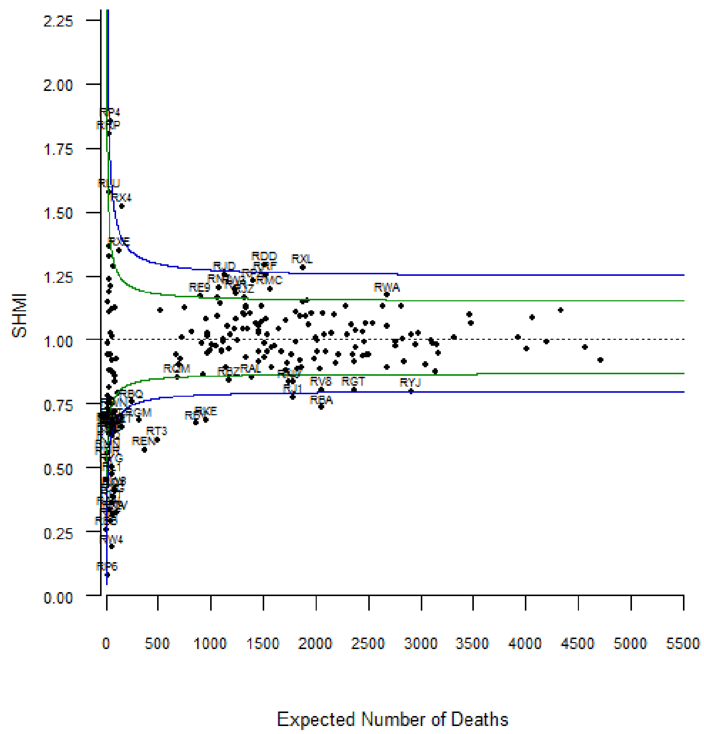
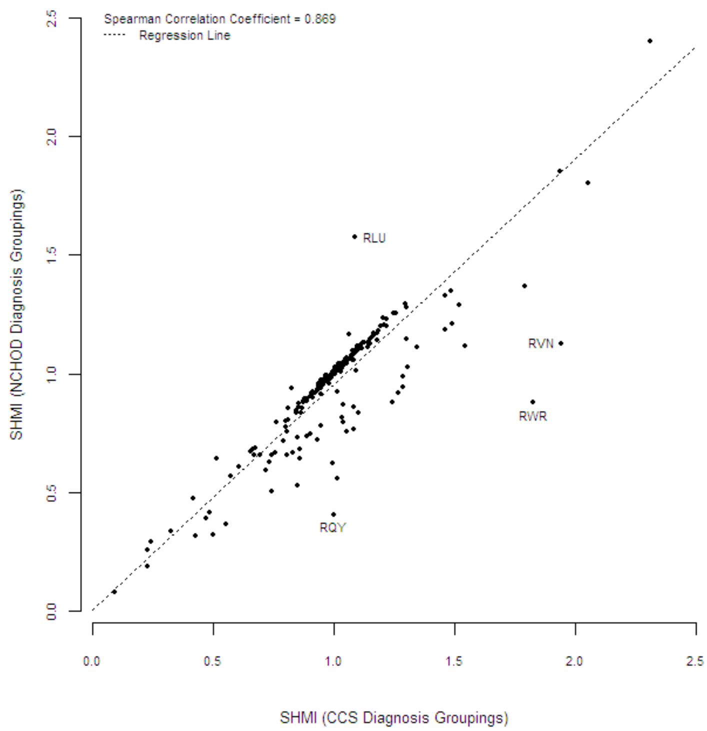


Figure 8.2: Scatterplot of SHMI based of diagnostic groups CCS or NCHOD



RLU = Birmingham Women's NHS Foundation Trust

RVN = Avon and Wiltshire Mental Health Partnership NHS Trust

RWR = Hertfordshire Partnership NHS Foundation Trust

RQY = South West London and St George's Mental Health NHS Trust

The significant alterations in SHMI based on the change in diagnostic grouping scheme may relate to how groups with small numbers of deaths were grouped in their scheme or how the NCHOD classification had to be altered to incorporate cancer deaths.

9. Effect of removing zero length of stay emergency admissions

We compared the results of calculating SHMIs for 2007/8 with and without zero length of stay emergency admissions included in the data. Table 9.1 shows that 28.3% of all emergency admissions in 2007/8 were zero length of stay and that zero length of stay emergencies accounted for 7% of all emergency deaths.

Table 9.1: Number of deaths by type of emergency admission (2007/8)

		Died	Total
Type of emergency admission	Zero length of stay	19131 7.0%	1895726 28.3%
	Other	253981 93.0%	4808356 71.7%
Total		273112 100%	6704082 100%

Figure 9.1 shows the correlation between the SHMIs is high (Spearman correlation coefficient of 0.925). This suggests that there is little change between the ranks of trusts for both methods. One trust (RLU, Birmingham Women's NHS Foundation Trust) has been identified as having a much higher SHMI when zero length of stay emergency admissions are excluded.

When calculating the SHMI with zero length of stay emergency admissions excluded in Figure 9.1 we excluded all zero length of stay emergency deaths when calculating the expected number of deaths for each trust but included them when calculating the observed number of deaths (as was suggested by the technical group). The consequence of this is that the mean SHMI for the year will be greater than 1. As an alternative comparison we calculated the SHMI again with zero length of stay emergency deaths excluded from both the calculation of the expected number of deaths and the observed number of deaths. Figure 9.2 shows that the correlation is increased when the zero length of stay emergency deaths are excluded from the observed deaths and Birmingham Women's NHS Foundation Trust is no longer outlying.

Figure 9.1: SHMI vs SHMI with zero length of stay emergencies removed (observed deaths included)

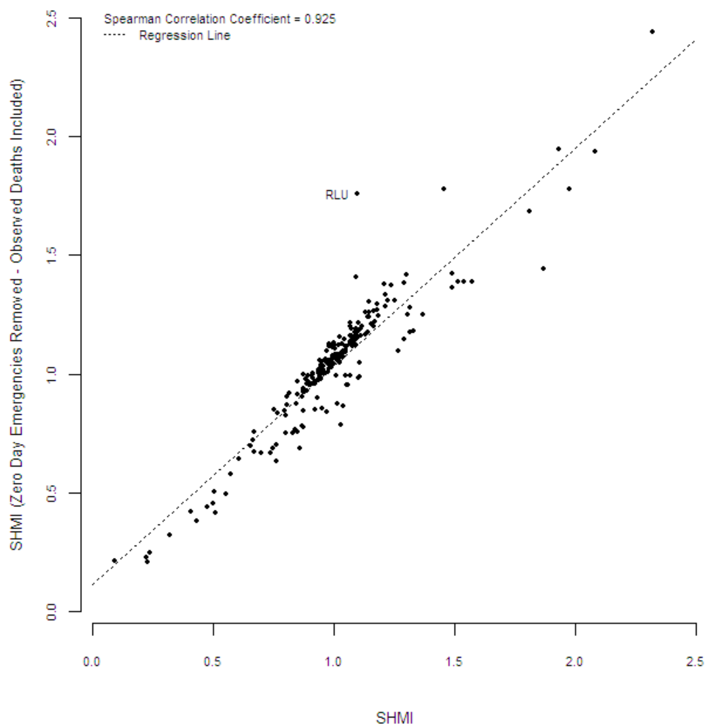
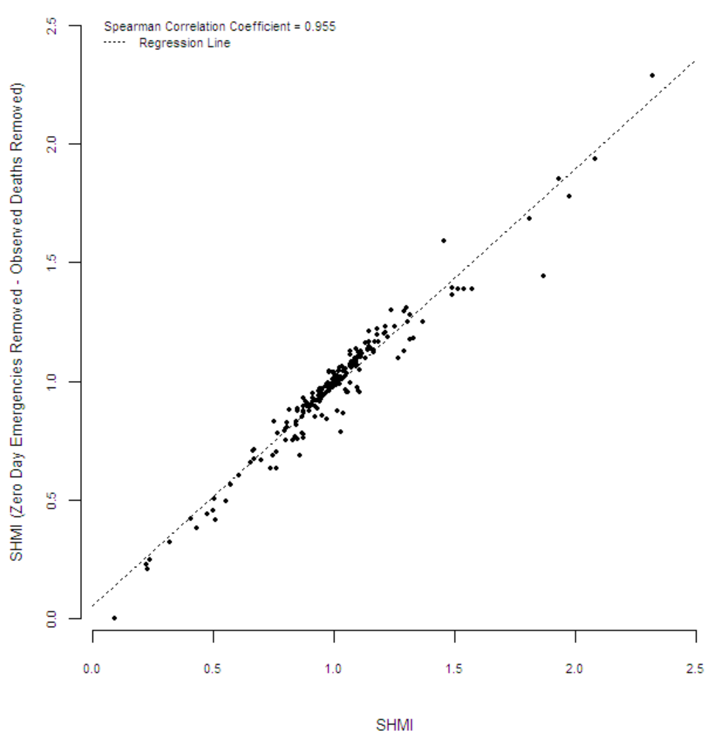


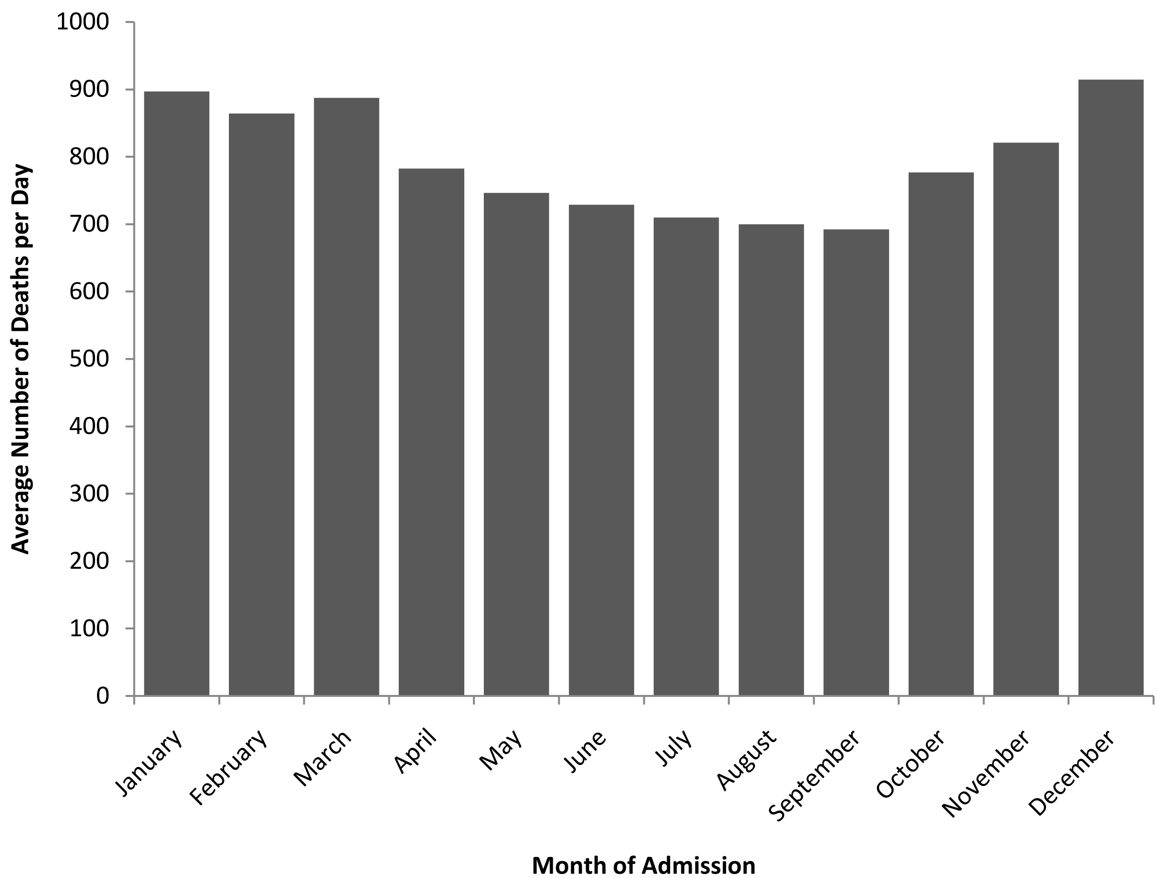
Figure 9.2: SHMI vs SHMI with zero length of stay emergencies removed (observed deaths excluded)



10. Short term SHMI

We calculated the number of deaths per day by month of admission using all available data and the results are given in Figure 10.1. As expected there is a seasonal effect on the number of deaths per day with October to March having higher deaths per day than April to September. The seasonal effect on the number of deaths will have an impact on any SHMI calculated for a period of time shorter than a year.

Figure 10.1: Average number of deaths per day by month of admission



We investigated the final SHMI model over small time periods by calculating values for the four quarters of 2009/10. These SHMIs were calculated in two ways. The first method used data from the 12 months prior to the end of the quarter to build the model and then calculated the observed and expected values for the three months contained in the quarter. Table 10.1 gives the exact dates for the data used in model building and prediction for each quarter when using this method. The second method used data from the 12 months prior to the end of the quarter to build the model and then calculated the observed and expected values for the full year. Table 10.2 gives the exact dates for the data used in model building and prediction for each quarter when using this method.

Table 10.1: Data used for calculating quarterly SHMI – method 1

Quarter	Data for Model Building	Data for Prediction
1	1st July 2008 – 30th June 2009	1st April 2009 – 30th June 2009
2	1st October 2008 – 30th September 2009	1st July 2009 – 30th September 2009
3	1st January 2009 – 31st December 2009	1st October 2009 – 31st December 2009
4	1st April 2009 – 31st March 2010	1st January 2010 – 31st March 2010

Table 10.2: Data used for calculating quarterly SHMI – method 2

Quarter	Data for Model Building	Data for Prediction
1	1st July 2008 – 30th June 2009	1st July 2008 – 30th June 2009
2	1st October 2008 – 30th September 2009	1st October 2008 – 30th September 2009
3	1st January 2009 – 31st December 2009	1st January 2009 – 31st December 2009
4	1st April 2009 – 31st March 2010	1st April 2009 – 31st March 2010

For each method we calculated the Spearman correlation coefficient between the quarters and the SHMI for the full year (as calculated previously). Tables 10.3 and 10.4 respectively show the correlations for method 1 with and without trusts with less than 100 observed deaths included. Tables 10.5 and 10.6 respectively show the correlations for method 2 with and without trusts with less than 100 observed deaths included.

Table 10.3: Spearman correlation coefficients between quarters using Method 1

	SHMI Q1	SHMI Q2	SHMI Q3	SHMI Q4	Year 09/10
SHMI Q1	1.000	0.555	0.573	0.651	0.770
SHMI Q2	0.555	1.000	0.651	0.481	0.738
SHMI Q3	0.573	0.651	1.000	0.648	0.856
SHMI Q4	0.651	0.481	0.648	1.000	0.814
Year 09/10	0.770	0.738	0.856	0.814	1.000

Table 10.4: Spearman correlation coefficients between quarters using Method 1 (Observed deaths >100)

	SHMI Q1	SHMI Q2	SHMI Q3	SHMI Q4	Year 09/10
SHMI Q1	1.000	0.726	0.712	0.731	0.883
SHMI Q2	0.726	1.000	0.710	0.701	0.863
SHMI Q3	0.712	0.710	1.000	0.743	0.868
SHMI Q4	0.731	0.701	0.743	1.000	0.905
Year 09/10	0.883	0.863	0.868	0.905	1.000

Tables 10.5 and 10.6 show that the correlation between different quarters is much higher for the second method of calculating quarterly SHMIs than the first. By calculating the observed and expected values for the SHMI over a full year the seasonal effect is removed and the quarterly results are much more consistent.

Table 10.5: Spearman correlation coefficients between quarters using Method 2

	SHMI Q1	SHMI Q2	SHMI Q3	SHMI Q4	Year 09/10
SHMI Q1	1.000	0.898	0.804	0.758	0.758
SHMI Q2	0.898	1.000	0.903	0.811	0.811
SHMI Q3	0.804	0.903	1.000	0.930	0.930
SHMI Q4	0.758	0.811	0.930	1.000	1.000
Year 09/10	0.758	0.811	0.930	1.000	1.000

Table 10.6: Spearman correlation coefficients between quarters using Method 2 (Observed deaths >100)

	SHMI Q1	SHMI Q2	SHMI Q3	SHMI Q4	Year 09/10
SHMI Q1	1.000	0.969	0.943	0.895	0.895
SHMI Q2	0.969	1.000	0.976	0.930	0.930
SHMI Q3	0.943	0.976	1.000	0.968	0.968
SHMI Q4	0.895	0.930	0.968	1.000	1.000
Year 09/10	0.895	0.930	0.968	1.000	1.000

11. In-hospital versus 30 day deaths

Figure 11.1 shows the funnel plot of the proportion of deaths that were in-hospital. The overall average is 79% but one can see there is considerable variation between Trusts. The warning lines are based on the Binomial distribution, and suggest that the variation is far from random Figure 11.2 shows the correlation between the SHMI (using age, sex, comorbidity and method of admission) for in-hospital deaths and 30 day deaths and we see a strong correlation of 0.936, but with two Trusts (both mental Health Trusts) which have rather higher SHMIs based on 30 day mortality than that based on in-hospital mortality.

Figure 11.1 In hospital vs 30 day death

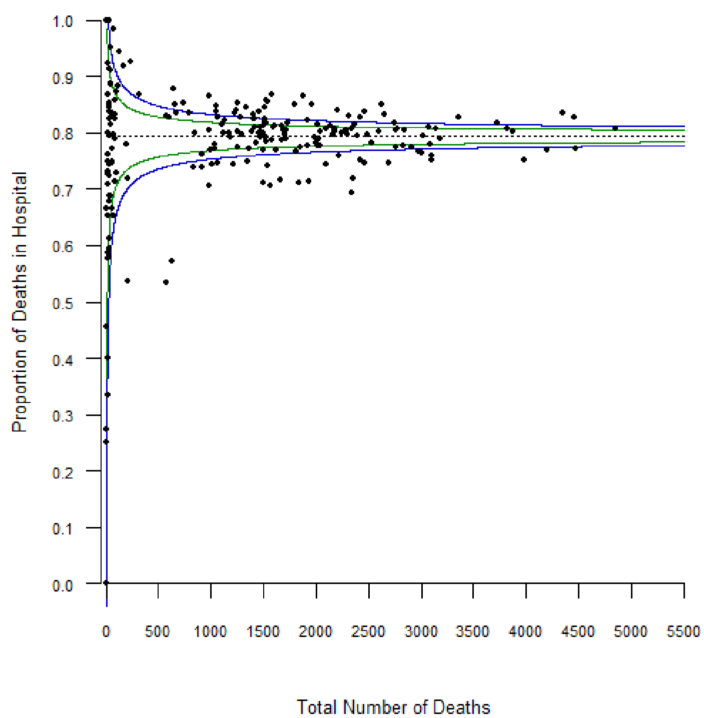
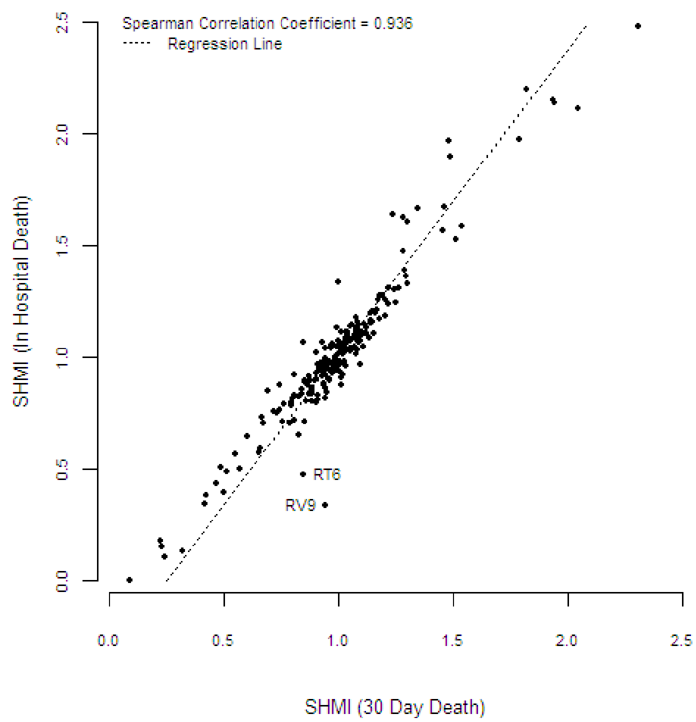


Figure 11.2 Scatter plot showing SHMI based on in-hospital deaths compared to 30 day deaths



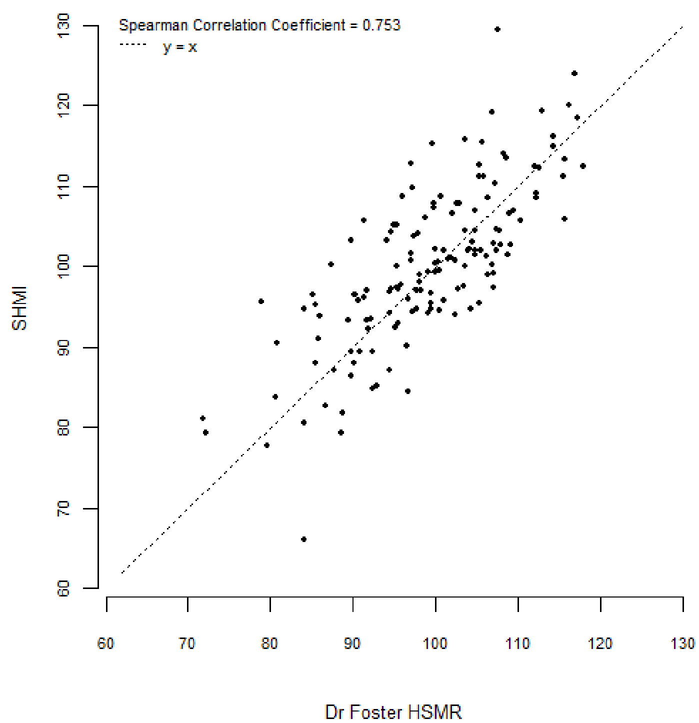
RV9 = Humber Mental Health Teaching NHS Trust

RT6 = Suffolk Mental Health Partnership NHS Trust

12. Comparison of SHMI vs Dr Foster HSMR

Dr Alex Bottle from Imperial College Dr Foster Unit kindly supplied us with data from a number of Trusts from 2009/10 giving their HSMR. There is a reasonable correlation (r=0.753) with the SHMI and no obvious outliers. We did not pursue this further as it was not part of our brief, but it of interest to note that the highest and lowest SMHI are not the highest and lowest HSMR.

Figure 12.1 Scatter plot of SMHI and Dr Foster's HSMR (selected Trusts 2009/10)



13. Discussion

The SHMI model was developed using data provided by the Department of Health to a specification agreed following a Technical Group meeting and subsequent steering group report available here:

http://www.dh.gov.uk/en/Healthcare/Qualityandproductivity/Makingqualityhappen/NationalQualityBoard/DH_102954

Wherever possible we have adhered to this document or interpreted it in a fashion favouring a robust, reproducible and timely quality measure, that includes all deaths. Variables which were potentially under a Trust's control were not included or computed.

13.1 Specialist hospitals

The HES data set we were provided with was completely unselected. One of the points in the specification of the SHMI was the exclusion of specialist hospitals. We were unable to find a clear definition of specialist hospital. In keeping with the philosophy of being inclusive, we included all hospitals. The specialist hospitals in general supply very few of the deaths by nature of their casemix (often maternity and psychiatry) and so contribute little to the weightings of the SHMI. They appear in the funnel plots near the vertical axis, since their expected deaths are low. They can appear to have high SHMIs because the denominator can be small. We did some sensitivity analysis excluding Trusts with fewer than 100 deaths per year. Small trusts may contribute to the estimation of the random effect parameter and it is possible that a 10% trim is insufficiently severe and so greater trim would be necessary.

The alterations in the SHMI for specialist hospitals based on changes in diagnostic classification and moving from 30 day to inpatient death highlights that these hospitals are sensitive to changes in methodology and should be excluded. Of note, the significant increase in SHMI moving from inpatient to 30 day death in two mental health hospitals may represent a high suicide rate following admission. ONS cause of death data and scrutiny of individual diagnostic groups would be required to explore this further.

13.2 Grouping within diagnosis classification schemes

As previously highlighted, both direct and indirect standardisation methods struggle when numbers of observations (or in this case deaths) are small. For this reason some diagnostics groups within both the CCS and NCHOD schemes were aggregated by a clinician in order to increase observations, ensuring clinical plausibility and avoid gender based anomalies. The group recognise potential limitations in performing this. Generalising coefficients for diagnostic groups which themselves have an element of case mix may decrease model fit. A similar effect may be observed when collapsing age or gender groups as performed by the Dr Foster group. Outputs from the Technical group included a need for recommendations on dealing with small numbers and the proposed collapsing of diagnostic groups is less computationally intensive than categorical methods employed by other groups. Lessons learnt from this analysis suggest that using individual ICD10 codes for diagnostic groups would not be feasible.

13.3 Using procedure as a predictor

Including treatments in a model can help predict the outcome if the treatments are (positively or negatively) effective. It is therefore natural to consider whether the treatment codes available in routine HES data, the procedure codes, should be included to improve outcome prediction and hence the comparability of performance indicators. However, we take the view that this is inappropriate in a performance model focused on monitoring the quality of care in hospitals. Quality of care is defined by whether beneficial treatments are used and harmful treatments avoided. If treatments were included in the outcome adjustment model a hospital using beneficial treatments would in effect be told that because they use beneficial treatments their outcomes are expected to be good, and so even though their observed outcomes *are* good their SMR is still 100. A hospital with poor outcomes due to poor quality care, i.e. not using beneficial treatments, would similarly be judged as having outcomes equal to those expected given their choice of treatments and would also have an SMR of 100. This is clearly not appropriate and we have not included procedure codes in any model.

13.4 Deaths from discharge or deaths from admission

SHMI proposes to use a follow-up time of 30 days post discharge. This raises two important issues that need to be resolved and agreed. First, is 30 days the right timeframe, and second should time be measured from admission or discharge?

With regard to the question of the time frame there is clearly a need to balance two contrary needs. First, the need to use a short time frame so that it is more likely that the outcome is connected to the intervention being evaluated (in this case the quality of hospital care) rather than other later interventions such as social and community care post-discharge, and second the need to use a long time frame to catch all the late impacts of care. Intensive care trials sometimes use 90 days in order to capture these late effects. On balance 30 days seems like a reasonable compromise.

The Scottish HSMR also uses 30 days but they measure it from admission not discharge. Using a fixed follow up period from date of admission would be analogous to the practice in trials of using a fixed time of follow up from randomisation (not treatment) which is designed to avoid the potential biases which result from using a variable follow up period. The advantage of using time from admission is that it defines a fixed window which is the same for all patients and hence for all hospitals whatever their discharge policies or opportunities are. Hospitals treating more deprived, socially isolated populations often find it more difficult to discharge elderly patients than hospitals treating less deprived populations. Using deaths within 30 days of discharge as a measure of performance could disadvantage the former.

The main argument for using a post discharge time frame is that if post admission is used then some patients who die in hospital will be categorised as a survivor in the SHMI analyses. These patients are still receiving hospital care and hence a SHMI calculated from admission date would not reflect the totality of hospital care, only the initial phase of care. A second reason for using post discharge is that it avoids creating an incentive to discharge patients who at 30 days are still unwell.

13.5 Direct versus indirect standardisation

Indirect standardisation can lead to contradictions, and even Simpson's Paradox, whereby two groups (A and B) can have stratum specific rates for which rates in A are always less than the rates in B but when the rates are indirectly standardised the standardised rate for A will be greater than that for B. For example Julious et al (2001) advise that direct standardisation should be used if comparisons are to be made between different groups. However, Roalfe et al (2008) state that regression based standardisation is a practical alternative to the direct method. It produces more reliable estimates than the direct method when the calculations are based on small numbers. It has greater flexibility in factor selection and allows standardisation by both continuous and categorical variables. We found direct standardisation to be difficult and when standardising simply for age and sex, and gave similar results to indirect standardisation. We plan to look at a hybrid approach whereby indirect standardisation can estimate individual risks, which can then be directly standardized, but that is work for the future.

13.6 Excluding emergency admissions with zero length of stay

We failed to exclude emergency admissions with zero length of stay, which was an omission on our part. The requirement was to exclude these spells in the predictions of the expected number of deaths but include the actual deaths. This requirement would mean that the expected number of deaths would always be less than the observed, biasing the SHMI upwards. We have conducted a sensitivity analysis for model 15 computing the SHMI with and without these spells for 2007/8. Removing them reduced the sample by 22.4%. The correlation of the SHMI with and without these spells was 0.955, and the scatter plot is given in Figure 9.2 which shows no unusual points. Thus we believe that none of our conclusions are affected by this omission.

13.7 Model selection and interpretation

With the exception of Trusts who have previously been identified using other summary quality measures and subsequently investigated and found to be underperforming, whether an individual Trust lies outside or inside a control line on a funnel plot is not a reason to prefer an individual model. Assuming variables included in the final model are outside the control of the healthcare provider and do not reflect quality of care, the best model explains the greatest variation or provides the best model fit. Indeed, during this investigation some model selection occurred when Trusts who were performing well were brought into areas in the funnel plot suggesting average performance.

By the nature of how data-derived control lines have been calculated there will inevitably be outliers even when extensive variables have been accounted for. It is not in the remit of this report to discuss whether the positioning of a Trust outside these limits based on their SHMI reflects poor quality of care.

Appendix 1

Regrouped ICD10 Clinical Classification System groups to ensure approximately 100 deaths in each group from a 10% random sample.

Summed CCS Nums	Complete Name	Total Deaths New Group
1	Tuberculosis	87
2 + 249	Septicemia (except in labor) + Shock	3550
3	Bacterial infection	75
4	Mycoses	56
5	HIV infection	105
6 + 7 + 8 + 9 + 10	Hepatitis + Viral infection + Other infections, including parasit + Sexually transmitted diseases (not + Immunization & screening for infect	91
11	Cancer of head & neck	317
12	Cancer of esophagus	722
13	Cancer of stomach	515
14	Cancer of colon	1119
15	Cancer of rectum & anus	583
16	Cancer of liver & intrahepatic bile	361
17	Cancer of pancreas	815
18	Cancer of other GI organs, peritone	264
19	Cancer of bronchus, lung	3593
20	Cancer, other respiratory & intrath	117
22 + 23	Malanomas of skin + Other non-epithelial cancer of skin	133
24	Cancer of breast	714
25	Cancer of uterus	151
26 + 28	Cancer of cervix + Cancer of other female genital organs	139
27	Cancer of ovary	395
29 + 30 + 31	Cancer of prostate + Cancer of testis + Cancer of other male genital organs	766
32	Cancer of bladder	503
33 + 34	Cancer of kidney & renal pelvis + Cancer of other urinary organs	296
35	Cancer of brain & nervous system	413
37 + 38	Hodgkin's disease + Non-Hodgkin's lymphoma	677
39	Leukemias	874
40	Multiple myeloma	384
41 + 45	Cancer, other & unspecified primary + Maintenance chemotherapy, radiother	179
42	Secondary malignancies	2058

21 + 36 + 43	Cancer of bone & connective tissue + Cancer of thyroid + Malignant neoplasm without specific	1094
44 + 167	Neoplasms of unspecified nature or + Nonmalignant breast conditions	237
46 + 47	Benign neoplasm of uterus + Other & unspec benign neoplasm	91
49	Diabetes mellitus without complicat	116
50	Diabetes mellitus with complication	374
48 + 51	Thyroid disorders + Other endocrine disorders	284
55	Fluid & electrolyte disorders	946
52 + 53 + 58	Nutritional deficiencies + Disorders of lipid metabolism + Other nutritional, endocrine & meta	359
59 + 60	Deficiency and other anemia + Acute posthemorrhagic anemia	559
63	Diseases of white blood cells	274
57 + 61 + 62 + 64	Immunity disorders + Sickle cell anemia + Coagulation & hemorrhagic disorders + Other hematologic conditions	91
65 + 68	Mental retardation + Senility & organic mental disorders	2328
66 + 67 + 69 + 72	Alcohol-related mental disorders + Substance-related mental disorders + Affective disorders + Anxiety, somatoform, dissociative,	190
71	Other psychoses	893
70 + 73 + 74 + 75	Schizophrenia & related disorders + Preadult disorders + Other mental conditions + Personal history of mental disorder	110
76 + 77 + 78	Meningitis (except that caused by T + Encephalitis (except that caused by + Meningitis (except that caused by T	173
79	Parkinson's disease	181
80 + 81	Multiple sclerosis + Other hereditary & degenerative ner	310
82 + 113	Paralysis + Late effects of cerebrovascular dis	186
83	Epilepsy, convulsions	422
85	Coma, stupor & brain damage	289
84 + 86 + 87 + 88 + 89 + 90 + 91 + 92 + 93 + 94	Headache, including migraine + Cataract + Retinal detachments, defects, vascu + Glaucoma + Blindness & vision defects + Inflammation, infection of eye (exc + Other eye disorders + Otitis media & related conditions + Conditions associated with dizzines + Other ear & sense organ disorders	202
95	Other nervous system disorders	413
96	Heart valve disorders	362
97	Peri-, endo- and myocarditis, cardi	271
98 + 99	Essential hypertension + Hypertension with complications & s	364
100	Acute myocardial infarction	4184
101	Coronary atherosclerosis & other he	1249
102	Nonspecific chest pain	620
103	Pulmonary heart disease	807
104	Other & ill-defined heart disease	75
105	Conduction disorders	202
106	Cardiac dysrhythmias	845

107	Cardiac arrest & ventricular fibril	1067
108	Congestive heart failure, nonhypert	5022
109	Acute cerebrovascular disease	9764
110 + 111 + 112	Occlusion or stenosis of precerebra + Other & ill-defined cerebrovascular + Transient cerebral ischemia	372
114	Peripheral & visceral atheroscleros	774
115	Aortic, peripheral & visceral arter	1216
116	Aortic & peripheral arterial emboli	265
117	Other circulatory disease	754
118 + 119 + 120 + 121	Phlebitis, thrombophlebitis & throm + Varicose veins of lower extremity + Hemorrhoids + Other disease of veins & lymphatics	280
122	Pneumonia (except that caused by TB	14986
125	Acute bronchitis	2547
127	Chronic obstructive pulmonary disea	3995
128	Asthma	133
129	Aspiration pneumonitis, food/vomitu	1841
130	Pleurisy, pneumothorax, pulmonary c	1094
131	Respiratory failure, insufficiency,	948
132	Lung disease due to external agents	221
56 + 133	Cystic fibrosis + Other lower respiratory disease	1077
123 + 124 + 126 + 134 + 136 + 137	Influenza + Acute & chronic tonsillitis + Other upper respiratory infections + Other upper respiratory disease + Disorders of teeth & jaw + Diseases of mouth, excluding dental	1514
135	Intestinal infection	699
138	Esophageal disorders	190
139	Gastroduodenal ulcer (except hemorr	280
140 + 141	Gastritis & duodenitis + Other disorders of stomach & duoden	124
143	Abdominal hernia	341
144	Regional enteritis & ulcerative col	125
145	Intestinal obstruction without hern	971
146 + 147	Diverticulosis & diverticulitis + Anal & rectal conditions	498
142 + 148	Appendicitis & other appendiceal co + Peritonitis & intestinal abscess	472
149	Biliary tract disease	602
150	Liver disease, alcohol-related	1130
151	Other liver diseases	1009
152	Pancreatic disorders (not diabetes)	447
153	Gastrointestinal hemorrhage	2266
154	Noninfectious gastroenteritis	1228
155	Other gastrointestinal disorders	1320
157	Acute & unspecified renal failure	2560
156 + 158	Nephritis, nephrosis, renal scleros + Chronic renal failure	463
159	Urinary tract infections	2611

160 + 161 + 162	Calculus of urinary tract + Other diseases of kidneys & ureters + Other diseases of bladder & urethra	214
163	Genitourinary symptoms & ill-define	389
164 + 165 + 166	Hyperplasia of prostate + Inflammatory conditions of male gen + Other male genital disorders	59
168 + 169 + 170 + 171 + 172 + 173 + 175	Inflammatory diseases of female pel + Endometriosis + Prolapse of female genital organs + Menstrual disorders + Ovarian cyst + Menopausal disorders + Other female genital disorders	68
174 + 176 + 177 + 178 + 179 + 180 + 181 + 182 + 183 + 184 + 185 + 186 + 187 + 188 + 189 + 190 + 191 + 192 + 193 + 194 + 195 + 196	Female infertility + Contraceptive & procreative managem + Spontaneous abortion + Induced abortion + Postabortion complications + Ectopic pregnancy + Other complications of pregnancy + Hemorrhage during pregnancy, abrupt + Hypertension complicating pregnancy + Early or threatened labor + Prolonged pregnancy + Diabetes or abnormal glucose tolera + Malposition, malpresentation + Fetopelvic disproportion, obstructi + Previous C-section + Fetal distress & abnormal forces of + Polyhydramnios & other problems of + Umbilical cord complication + Trauma to perineum & vulva + Forceps delivery + Other complications of birth, puerp + Normal pregancy and/or delivery	86
197	Skin & subcutaneous infections	659
198 + 199 + 200	Other inflammatory condition of ski + Chronic ulcer of skin + Other skin disorders	602
201	Infective arthritis & osteomyelitis	128
204	Other non-traumatic joint disorders	215
205 + 206	Spondylosis, intervertebral disc di + Osteoporosis	353
207	Pathological fracture	144
211	Other connective tissue disease	384
54 + 202 + 203 + 208 + 209 + 210 + 212	Gout and other crystal arthropathie + Rheumatoid arthritis & related dise + Osteoarthritis + Acquired foot deformities + Other acquired deformities + Systemic lupus erythematosus & conn + Other bone disease & musculoskeleta	131
213	Cardiac & circulatory congenital an	131
214 + 215 + 216 + 217	Digestive congenital anomalies + Genitourinary congenital anomalies + Nervous system congenital anomalies + Other congenital anomalies	132
219	Short gestation, low birth weight,	606
220 + 221 + 222 + 223	Intrauterine hypoxia & birth asphyx + Respiratory distress syndrome + Hemolytic jaundice & perinatal jaun + Birth trauma	90
224	Other perinatal conditions	332
226	Fracture of neck of femur (hip)	3228
229	Fracture of upper limb	353
230	Fracture of lower limb	453

225 + 227 + 228 + 231 + 232	Joint disorders & dislocations, tra + Spinal cord injury + Skull & face fractures + Other fractures + Sprains & strains	734
233	Intracranial injury	873
234	Crushing injury or internal injury	273
235	Open wounds of head, neck & trunk	234
236	Open wounds of extremities	86
237	Complication of device, implant or	562
238	Complication of surgical procedures	347
239	Superficial injury, contusion	574
240	Burns	95
241 + 242 + 243	Poisoning by psychotropic agents + Poisoning by other medications & dr + Poisoning by nonmedicinal substance	198
244	Other injuries & conditions due to	208
245	Syncope	1449
246	Fever of unknown origin	94
247 + 248	Lymphadenitis + Gangrene	142
250	Nausea & vomiting	498
251	Abdominal pain	947
252	Malaise & fatigue	540
253 + 254 + 256 + 257 + 258 + 259	Allergic reactions + Rehabilitation care, fitting of pro + Medical examination/evaluation + Other aftercare + Other screening for suspected condi + Residual codes, unclassified	2186

Appendix 2

Regrouped NCHOD Classification System groups to ensure approximately 200 deaths in each group from 07/08 HES year sample.

Summed NCHOD Nums	NCHOD Complete Name	Total Deaths New Group
1	Other bacterial intestinal infections	1875
2	Rest of Intestinal infectious diseases	276
3 + 5 + 6 + 7	Meningococcal infection + Viral infections characterized by skin & mucous membrane lesions + Chapter 1 - Remainder of chapter - Certain infectious and parasitic diseases + Chapter 1 - Remainder of chapter - Certain infectious and parasitic diseases	1996
4	Rest of Certain bacterial diseases	7283
8	Chapter 2 pt - In situ & benign neoplasms and others of uncertain/unknown behaviour	404
9	Iron deficiency anaemia	950
10	Diseases of the blood and blood-forming organs	1404
11	Chapter 3 -Remainder of chapter - Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism	592
12	Other disorders of pancreatic internal secretion	1855
13	Volume depletion	802
14	Other disorders of fluid electrolyte and acid-base balance	1843
15	Chapter 4 -Remainder of chapter - Endocrine, nutritional and metabolic diseases	1779
16 + 20	Dementia + Chapter 5 - Remainder of chapter - Mental and behavioural disorders	382
17	Schizophrenia, schizotypal and delusional disorders	332
18 + 19	Mood [affective] disorders + Neurotic behavioural & personality disorders	699
21	Inflammatory diseases of the central nervous system	357
22	Spinal muscular atrophy and related syndromes	619
23	Alzheimer’s disease	256
24 + 25	30-G32 - Rest of Other degenerative diseases (excl Alzheimer’s disease) + 30-G32 - Rest of Other degenerative diseases (excl Alzheimer’s disease)	583
26 + 33 + 34	Demyelinating diseases (incl. Multiple Sclerosis) of the central nervous system + G90-G99 - Rest of Other diseases & disorders of the nervous system + Chapter 6 - Remainder of chapter - Diseases of the nervous system	304
27 + 28	Epilepsy + Status epilepticus	256
29	Transient cerebral ischaemic attacks and related syndromes	539
30	Rest of Epilepsy migraine & other episodic disorders	404
31	Cerebral palsy & other paralytic syndromes	220
32	Other disorders of brain	692
35 + 36	Chapter 7 - Diseases of the eye and adnexa + Chapter 8 - Diseases of the ear and mastoid process	767
37 + 38	Essential (primary) hypertension + Hypertensive heart disease	1243

39	Rest of Hypertensive diseases	6889
40	Angina pectoris	1672
41 + 42 + 43 + 44	Acute myocardial infarction + Subsequent myocardial infarction + Certain current complication follow acute myocardial infarction + Other acute ischaemic heart diseases	1778
45	Chronic ischaemic heart disease	1506
46 + 47	Other pulmonary heart diseases + Rest of Pulmonary heart disease & diseases of pulmonary circulation	306
48 + 50	Other diseases of pericardium + Cardiomyopathy	224
49	Acute and subacute endocarditis	424
51	Atrioventricular and left bundle-branch block	1934
52	Other conduction disorders	213
53	Cardiac arrest	1617
54 + 56	Paroxysmal tachycardia + Other cardiac arrhythmias	842
55	Atrial fibrillation and flutter	10658
57	Heart failure	1233
58	Rest of Other forms of heart disease	3823
59	Subarachnoid haemorrhage	8040
60	Intracerebral haemorrhage	5077
61	Cerebral infarction	618
62	Stroke not specified as haemorrhage or infarction	1249
63	Other cerebrovascular diseases	402
64	Rest of Cerebrovascular diseases	2244
65 + 68	Atherosclerosis + Other peripheral vascular diseases	318
66	Aortic aneurysm and dissection	867
67	Other aneurysm	557
69 + 70	Arterial embolism and thrombosis + Other disorders of arteries and arterioles	581
71	Rest of Diseases of arteries, arterioles & capillaries	192
72 + 74	Phlebitis and thrombophlebitis + Rest of Diseases of veins & lymphatic system nec	432
73	Oesophageal varices	419
75 + 179 + 181	Hypotension + Syncope and collapse + Shock not elsewhere classified	251
76	Other and unspecified disorders of circulatory system	200
77 + 78	Rest of Other & unspecified disorders of the circulatory system + Chapter 9 - Remainder of chapter - Diseases of the circulatory system	418
79	Acute upper respiratory infections	563
80	Pneumonia due to Streptococcus pneumoniae	28241
81	Bacterial pneumonia not elsewhere classified	6481
82	Rest of Influenza & pneumonia	295
83	Other acute lower respiratory infections	8585
84	Emphysema	286
85	Other chronic obstructive pulmonary disease	341
86	Asthma	4399
87	Rest of Chronic lower respiratory diseases	466

88	Lung diseases due to external agents	1098
89	Pulmonary oedema	2366
90	Other interstitial pulmonary diseases	287
91	Pleural effusion not elsewhere classified	1633
92	Pneumothorax	244
93	Respiratory failure not elsewhere classified	374
94 + 95	Other respiratory disorders + Rest of Other diseases of the respiratory system	208
96	Chapter 10 - Remainder of chapter - Diseases of the respiratory system	309
97	Gastro-oesophageal reflux disease	337
98	Other diseases of oesophagus	721
99 + 100	Gastric ulcer + Duodenal ulcer	226
101 + 102	Peptic ulcer site unspecified + Gastritis and duodenitis	301
103 + 104	Rest of Diseases of oesophagus, stomach & duodenum + Diseases of appendix	372
105	Inguinal hernia	360
106	Femoral hernia	3233
107	Rest of Hernia	1167
108	Non-infective enteritis & colitis	2193
109	Vascular disorders of intestine	860
110	Paralytic ileus and intestinal obstruction without hernia	1031
111	Diverticular disease of intestine	786
112	Other functional intestinal disorders	991
113 + 115 + 118	Other diseases of anus and rectum + Rest of Other diseases of intestines + Rest of Diseases of peritoneum	537
114	Other diseases of intestine	829
117	Other disorders of peritoneum	2157
119	Alcoholic liver disease	323
120 + 121	Hepatic failure not elsewhere classified + Fibrosis and cirrhosis of liver	401
122 + 123	Other diseases of liver + Rest of Diseases of liver	677
124 + 125 + 126	Cholelithiasis + Cholecystitis + Other diseases of biliary tract	1587
127 + 128	Acute pancreatitis + Other diseases of pancreas	260
129	Rest of Disorders of gallbladder, biliary tract & pancreas	3665
130 + 131	Postprocedural disorders of digestive system NEC + Rest of Other diseases of the digestive system	245
132	Chapter 11 - Remainder of chapter - Diseases of the digestive system	1392
133	Cutaneous abscess furuncle and carbuncle	355
134	Cellulitis	842
135 + 136 + 137	Decubitus ulcer + Ulcer of lower limb not elsewhere classified + Chapter 12 - Remainder of chapter - Diseases of the skin and subcutaneous tissue	461
138 + 144	Pyogenic arthritis + Chapter 13 - Remainder of chapter - Diseases of the musculoskeletal system and connective tissue	1004
139	Other rheumatoid arthritis	914
140	Systemic connective tissue disorders	234
141	Soft tissue disorders	298

142	Osteoporosis with pathological fracture	2338
143	Rest of Osteopathies and chondropathies	4966
145	Acute renal failure	336
146	Chronic renal failure	7008
147	Unspecified renal failure	208
148	Other diseases of the urinary system	205
149 + 150 + 151	Diseases of male genital organs + Non-inflammatory disorders of female genital tract + Chapter 14 - Remainder of chapter - Diseases of the genitourinary system	1939
152	Chapter 15 - Pregnancy, childbirth and the puerperium	877
153	Disorders related to short gestation and low birth weight	548
154	Chapter 16 - Remainder of chapter - Certain conditions originating in the perinatal period	210
155	Chapter 17 - Congenital malformations, deformations and chromosomal abnormalities	239
156	Abnormalities of heart beat	481
157	Gangrene not elsewhere classified	243
158	Haemorrhage from respiratory passages	3257
159	Cough	1751
160	Abnormalities of breathing	178
161 + 162	Pain in throat and chest + Other symptoms and signs involving the circulatory. and respiratory systems	2325
163	Rest of Symptoms & signs inv. the circulatory/respiratory system	781
164	Abdominal and pelvic pain	431
165	Dysphagia	886
166	Unspecified jaundice	232
167	Ascites	1617
168	Other symptoms & signs involving digestive system and abdomen	189
169	Rest of Symptoms & signs involving the digestive system & abdomen	545
170	Symptoms & signs inv. the skin & subcutaneous tissue	1188
171	Symptoms & signs inv. the nervous& musculoskeletal system	370
172	Symptoms & signs involving the urinary system	2218
173 + 174	Somnolence, stupor and coma + Other symptoms & signs involve. cognitive function and awareness	308
175	Dizziness and giddiness	192
176	Rest of Symptoms & signs inv. Cognition, perception etc.	1364
177	Headache	3730
178	Malaise and fatigue	593
180	Convulsions not elsewhere classified	748
182	Symptoms and signs concerning food and fluid intake	3813
183	Other general symptoms and signs	3796
184	Rest of General symptoms & signs	748
185	Unknown & unspecified causes of morbidity	466
186	Chapter 18 - Remainder of chapter - Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	673

187 + 188	Superficial injury of head + Open wound of head	1831
189 + 190	Fracture of skull and facial bones + Intracranial injury	512
191 + 192	Other and unspecified injuries of head + Rest of Injuries to the head	370
193	Injuries to the neck	395
194 + 195	Injury of other and unspecified intrathoracic organs + Rest of Injuries to the thorax	1005
196 + 203	Injury of intra-abdominal organs + Injuries to unspecified part of trunk limb or body region	984
197	Rest of Injuries to the abdomen. lower back, lumbar spine & pelvis	391
198	Injuries to the shoulder & upper arm	179
199	Injuries to the elbow & forearm	8684
200 + 201	Other and specified injuries of hip and thigh + Rest of Injuries to the hip & thigh	874
202	Injuries to the knee & lower leg	175
204	Burns and corrosions	476
205 + 206 + 207	Poisoning by narcotics and psychodysleptics [hallucinogens] + Rest of Poisonings by drugs medicaments & biological substances + Toxic effects of substances chiefly non-medicinal as to source	339
208	Hypothermia	391
209	Rest of Other and unspecified effects of external causes	287
210	Comps of cardiac & vasc. prosthetic devices, implants & grafts	477
211	Comps of genitourinary prosthetic devices, implants & grafts	308
212	Comps of internal orthopaedic prosthetic devices implants & grafts	605
213	Comps other internal prosthetic devices implants &grafts	363
214 + 215 + 216	Rest of Complications of surgical & medical care nec. + Chapter 19 - Remainder of chapter - Injury, poisoning and certain other consequences of external causes + Chapter 19 - Remainder of chapter - Injury, poisoning and certain other consequences of external causes	184
217 + 222	Med observ and evaluation for suspected diseases and conditions + Ungrouped	278
218	Rest of Persons encountering health services for exam.	431
221	Cancer	47479

Appendix 3

Standardisation of rates for the SHMI

Notation

Suppose y_{ijkt} is the outcome for the i^{th} subject in the j^{th} Trust for clinical classification system (CCS) k (k=1…n_{CCS}) in year t. The variable y_{ijkt} is 0/1 depending on whether the subject dies or not. Let $E(y_{ijkt})=\pi_{ijkt}$. Ignore t for the time being

Indirect standardisation

For the whole population (or a random sample) for each CCS k, we fit

$$\text{logit}(\pi_{ijk}) = \beta_0 + \sum \beta_{k,age}\delta_{age} + \beta_{k,sex}\delta_{sex} + \sum_s \beta_{k,s}\, x_s \quad (1)$$

where δ_{age} is a dummy variable for each age category <1, 1-4, 5-14, …85-94), δ_{sex} is a dummy variable for sex and the x_s's are the variables in the proposal (admission method, comorbidity, year of discharge , Deprivation score, No of emergency admissions in previous 12 months.

We estimate the π_{ijk}'s from $\frac{e^{LP_{ijk}}}{1+e^{LP_{ijk}}}$ where LP_{ijk} is the linear predictor in the model.

We then find $E_j = \sum_{k=1}^{n_{ccs}} \sum_{i=1}^{n_k} \pi_{ijk}$ and $O_j = \sum_{k=1}^{n_{CCS}} \sum_{i=1}^{n_k} y_{ijk}$

We compare the E_j's and O_j's using a funnel plot, which is a plot of O_j/E_j versus E_j with warning lines as discussed by Spigelhalter (2005a,b). We used his suggestion for a ranom effects model. Assuming the variance of log(SMR)=1/O_j we used trimming whereby the scores z_i= $O_j$log(SMR) are ranked and those in the top tenth omitted and bottom tenth are omitted to find the overdispersion parameter tau^2 as described by Spiegelhalter(2005b). The warning lines for a x-axis coordinate x, are then given by $\exp(z(alpha)\surd(1/x+tau^2)$. The advantage of this method is that outlying Trusts do not inflate the estimate of the overdispersion parameter .

Model comparison

We want to see what the effect fitting a extra covariate x_h to obtain the predicted deaths would have on the ranking of a Trust.

Suppose E_i (-h) is the expected number of deaths for Trust i omitting variable x_h and $E_i(+h)$ is the expected number of deaths including variable x_h in the model.

Dr Foster plot $O_i/E_i(+h)$ vs $E_i(+h)$ and $O_i/E_i(-h)$ vs $E_i(-h)$ on the same graph and see whether or not this moves a Trust inside or outside a warning line. They also correlate $O_i/E_i(+h)$ and $O_i/E_i(-h)$.

However since O_i is common to both plots, a more sensitive analysis would be to do a 'Bland-Altman' plot of

$E_i(+h)-E_i(-h)$ versus $(E_i(+h)-E_i(-h))/2$.

This enables us to see which Trusts are affected by changes in parameters in the model. This is related to Bland-Altman plots which are used to measure agreement between different methods of measuring quantities. Bland and Altman criticise the use of correlation because one can have a high correlation and poor agreement if the slope relating the two methods is not unity.

If we think of the E_i's as being random variables then their distribution will depend on a number of factors, but as a first approximation we could suggest that they behave as Poisson variables and their variance is approximately E_i.

Then $Var\{E_i(+h)-E_i(-h)\}=Var\{E_i(+h)\}+Var\{E_i(-h)\}-2xCov\{E_i(+h),E_i(-h)\}$

If we assume that $Cov\{E_i(+h),E_i(-h)\}$ is likely to be high , say 0.90, and that

$Var\{E_i(+h)\} \approx Var\{E_i(-h)\} \approx \{E_i(+h)\}+Var\{E_i(-h)\}/2$

then we get

$Var\{E_i(+h)-E_i(-h)\} \approx 0.2x\{E_i(+h)+\{E_i(-h)\}/2$

Thus suitable guidelines might be two standard deviations either side of 0.

i.e. +/-2x√(0.2x{E_i(+h)+{E_i(-h)}/2)= +/- 0.89x√({E_i(+h)}+{E_i(-h)}/2)

Since they are just guidelines, for simplicity we chose +/-√E as the upper and lower limits rather than 0.89√E

We have also plotted guidelines y=+/- 0.05x on the graph. Points above or below this line represent Trusts which will have changed their SHMI by 5%. Let O be the observed number of deaths in a Trust and E_1 and E_2 be the expected values under models 1 and 2. Then

$$d = \frac{O}{E_1} - \frac{O}{E_2} = \frac{O}{E_2}\frac{(E_2 - E_1}{E_1} \approx \frac{O}{E_2}\frac{(E_2 - E_1}{\bar{E}}$$

.

where $\bar{E} = (E_1 + E_2)/2$.

If we put $(E_2 - E_1) = 0.05\bar{E}$ into the above formula we get

$$d = 0.05\frac{O}{E_2}$$

And thus points above the line on the graph E_1-E_2 vs $\bar{E}$ represent a change in SHMI of 5%.

We prefer to plot $\bar{E}$ rather than E_1 because then the y and x variables are uncorrelated if E_1 and E_2 are random.

Direct standardisation

Let D(a,s,d,t) be the number of deaths and N(a,s,d,t) the number of patients for each casemix group (age (a), sex (s), diagnosis (d) and trust (t)).

Let $NT(a,s,d) = \sum_t N(a,s,d,t)$ and $DT(a,s,d) = \sum_t D(a,s,d,t)$.

The death rate for each casemix group is given by $R(a,s,d,t) = \frac{D(a,s,d,t)}{N(a,s,d,t)}$.

The national proportions for each casemix group is give by $P(a,s,d,t) = \frac{NT(a,s,d,t)}{\sum_{a,s,d,t} N(a,s,d,t)}$.

Direct standardisation assumes that each casemix group has at least one patient. The directly standardized death rate for each trust is given by $R_d(t) = \sum_{a,s,d}[R(a,s,d,t) \times P(a,s,d,t)]$.

When trusts have no patients in some casemix groups direct standardisation goes wrong because the weights used to combine the casemix specific rates don't sum to 1. Furthermore, when different institutions have different zeros then they have different total weights and comparisons are unfair. An alternative method of direct standardisation to account for zeros in come casemix groups replaces the national proportions with trust specific weights. The calculations are as follows:

Let I(a,s,d,t) be an indicator function with value 1 if there are patients in the casemix group and value 0 otherwise.

Trust specific weights are given by $TW(a,s,d,t) = \frac{I(a,s,d,t) \times P(a,s,d,t)}{\sum_{a,s,d}[I(a,s,d,t) \times P(a,s,d,t)]}$.

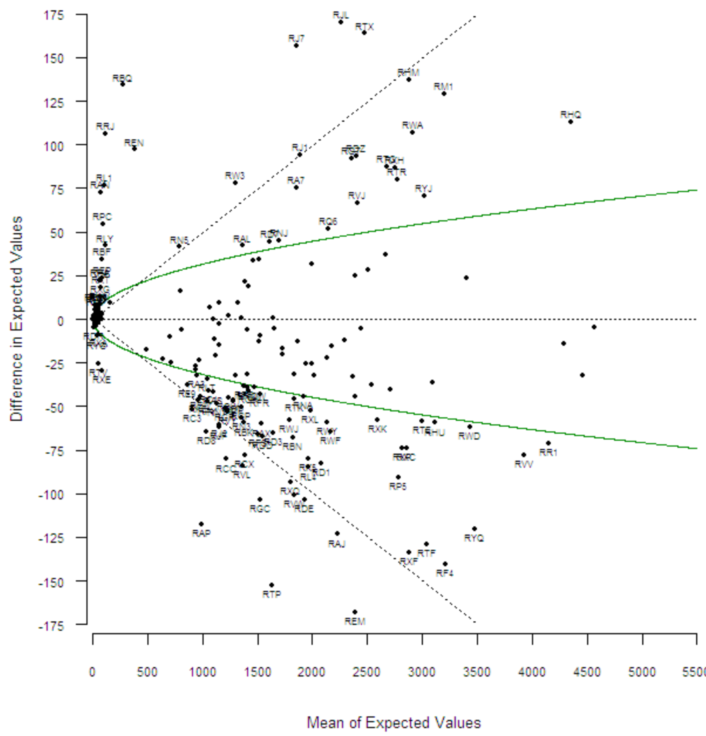
Then the directly standardized death rate (adjusted for zeroes) for each trust is given by

$R_d(t) = \sum_{a,s,d}[R(a,s,d,t) \times TW(a,s,d,t)]$.

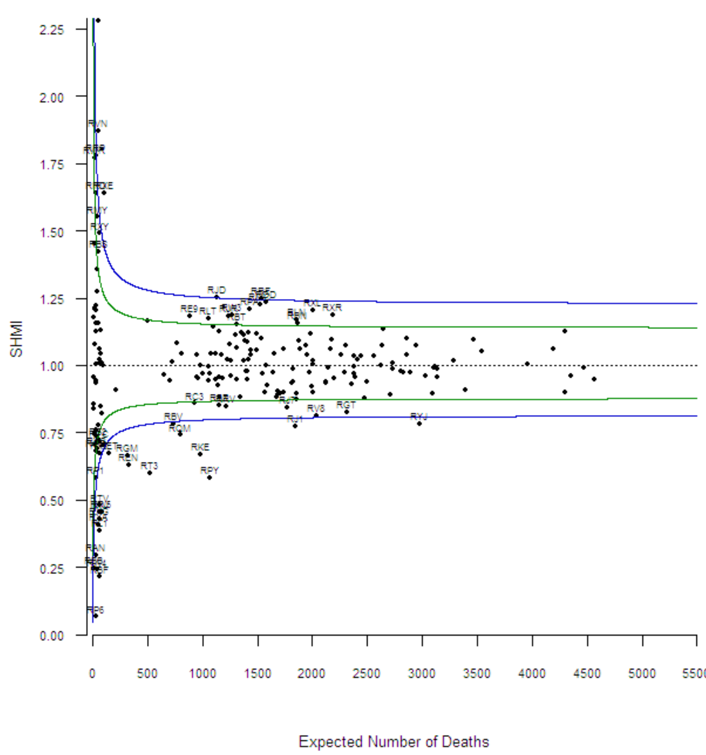
Roalfe et al (2008) state that regression based standardisation is a practical alternative to the direct method. It produces more reliable estimates than the direct or indirect method when the calculations are based on small numbers. It has greater flexibility in factor selection and allows standardisation by both continuous and categorical variables.

Appendix 4 Diffsum plots for different model contrasts

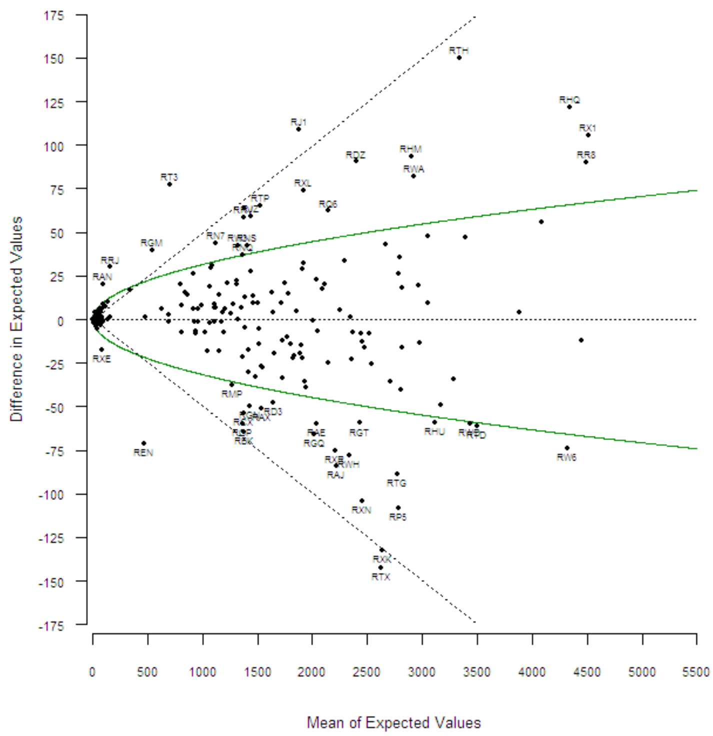
Model 1 vs Model 2



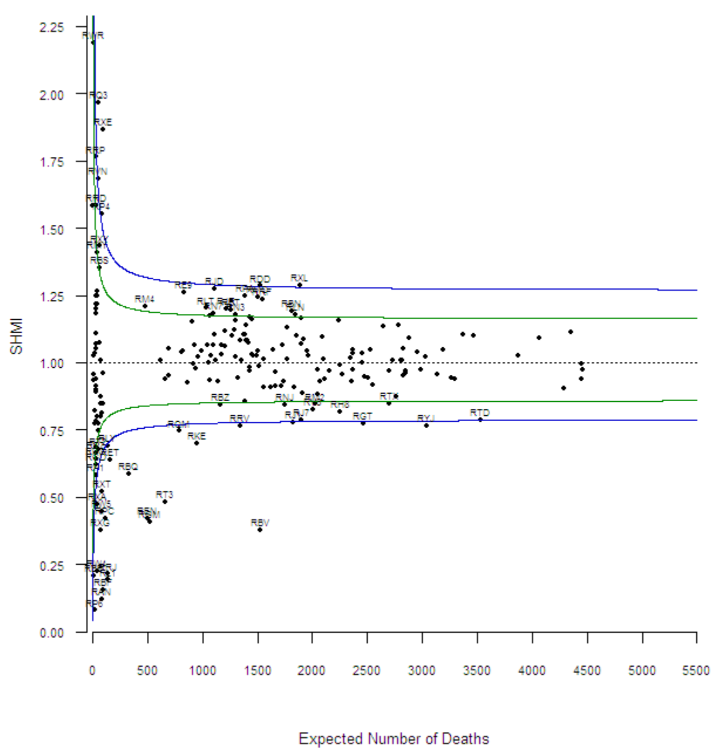
Model 2 – Random Effects



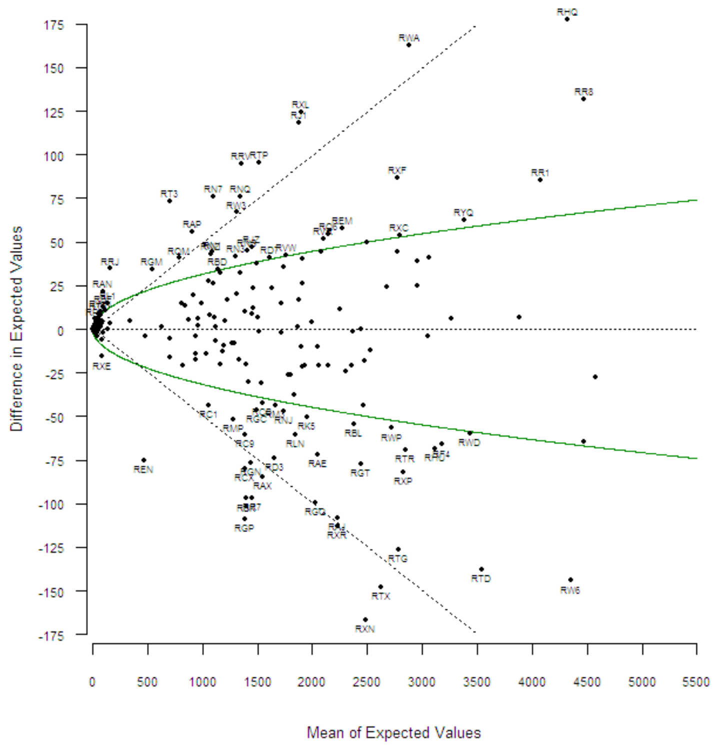
Model 1 vs Model 3



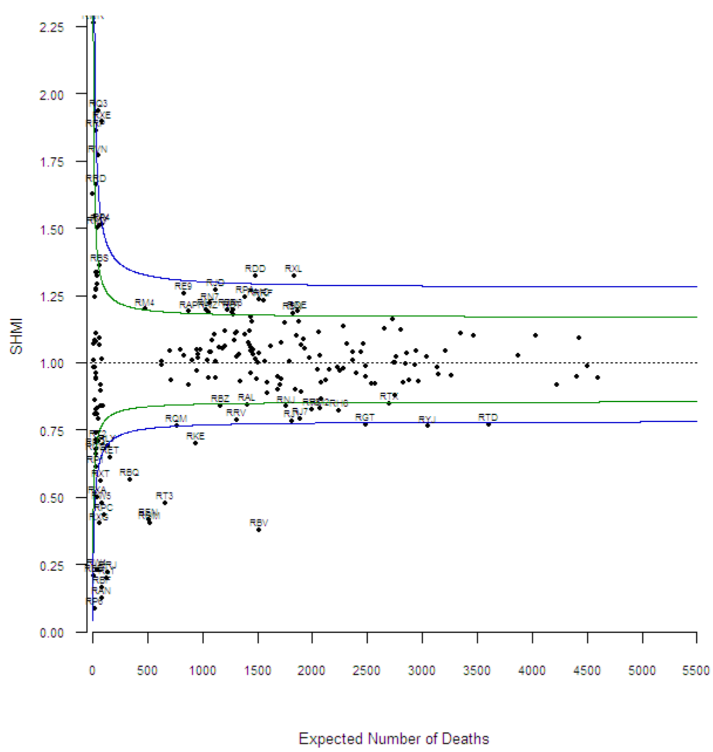
Model 3 – Random Effects



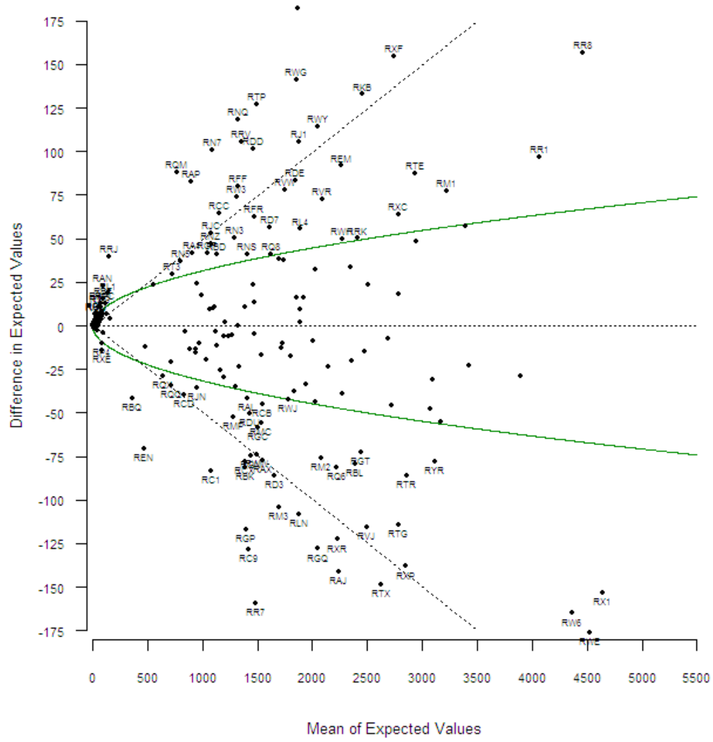
Model 1 vs Model 4



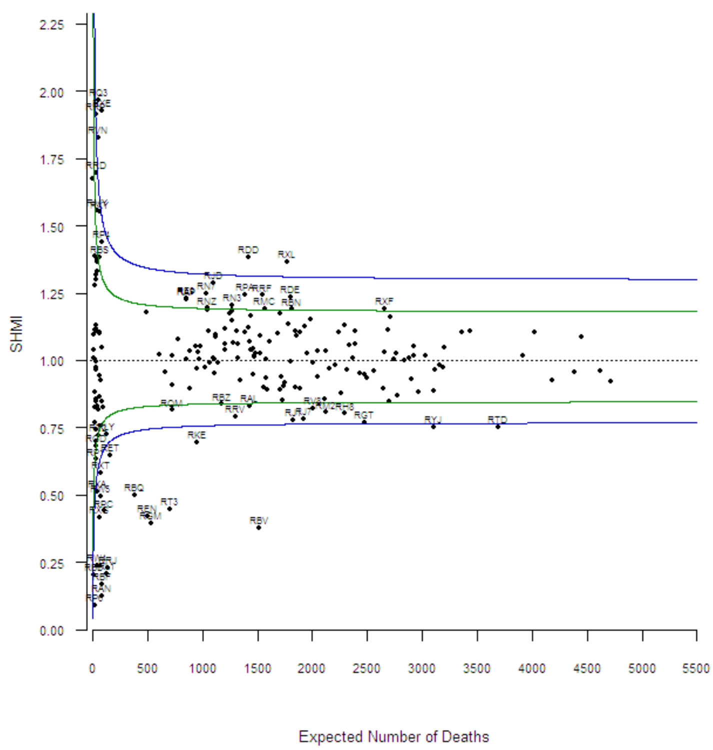
Model 4 – Random Effects



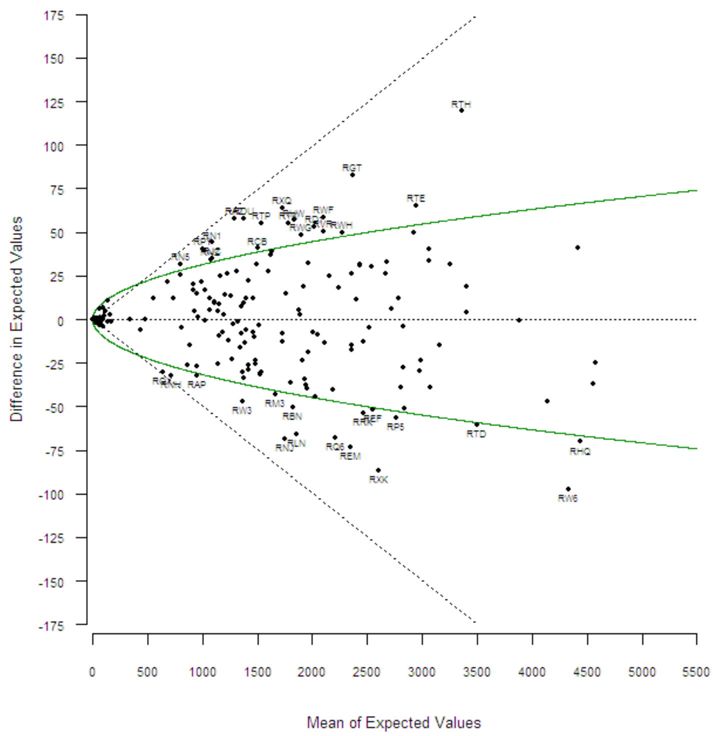
Model 1 vs Model 5



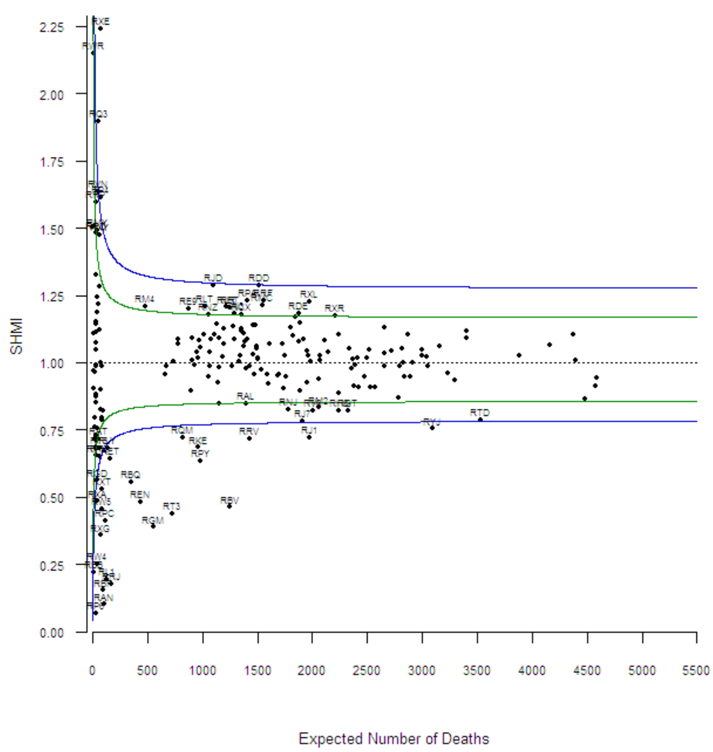
Model 5 – Random Effects



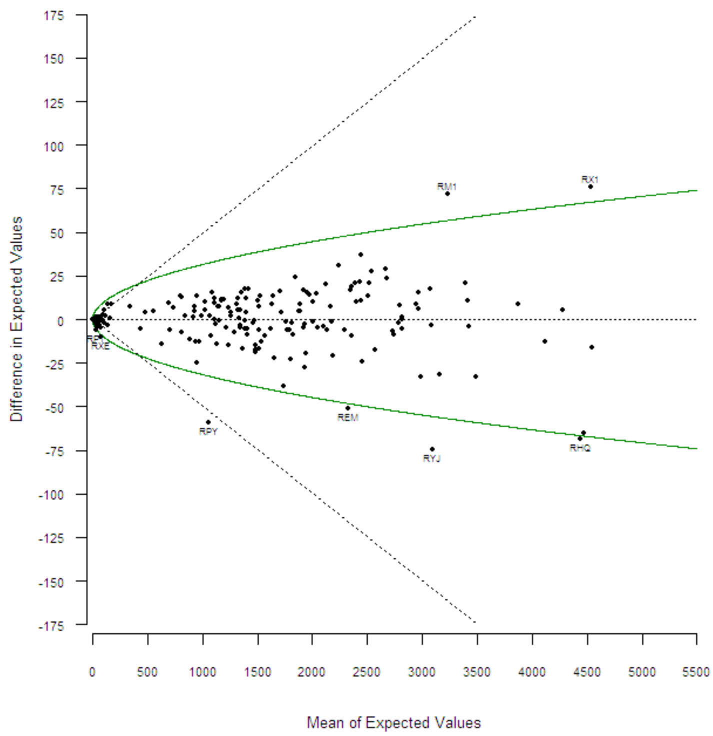
Model 1 vs Model 6



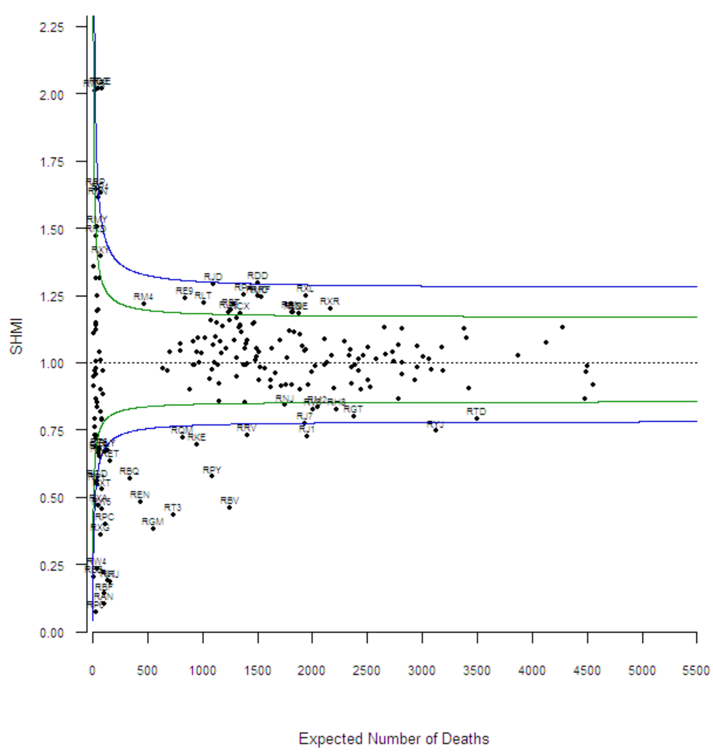
Model 6 – Random Effects



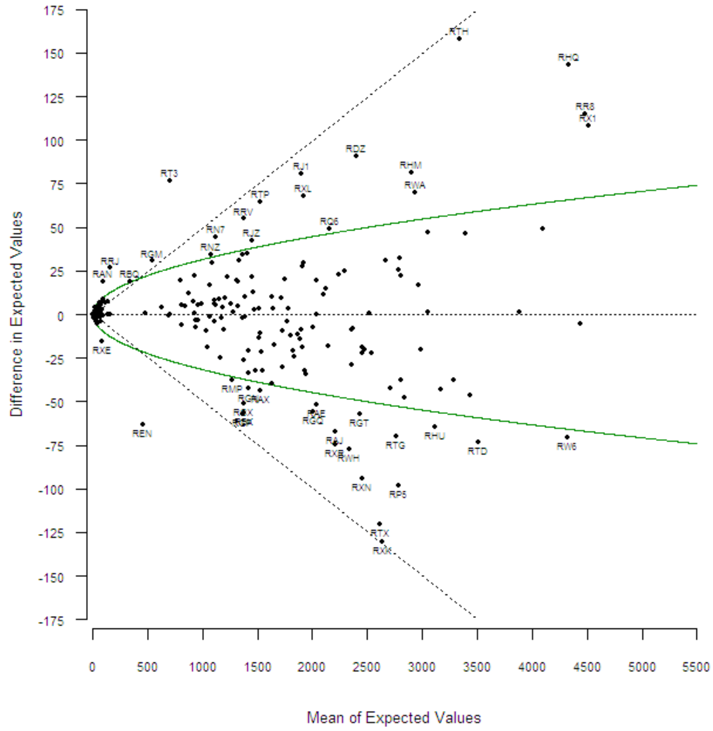
Model 1 vs Model 7



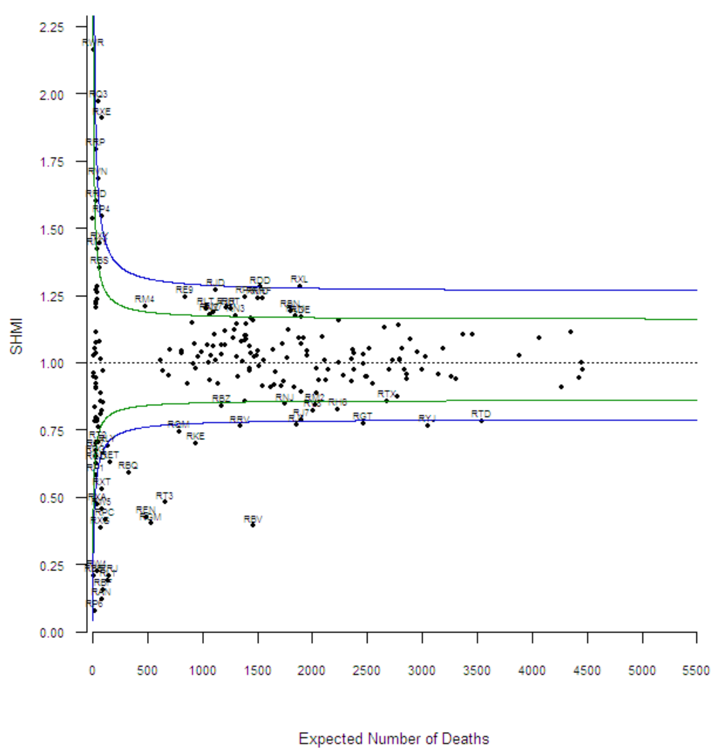
Model 7 – Random Effects



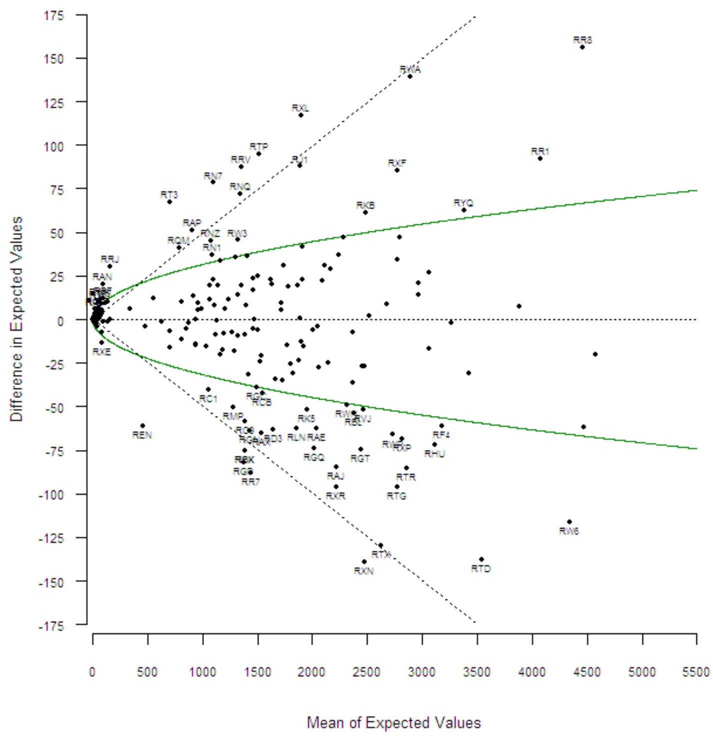
Model 1 vs Model 8



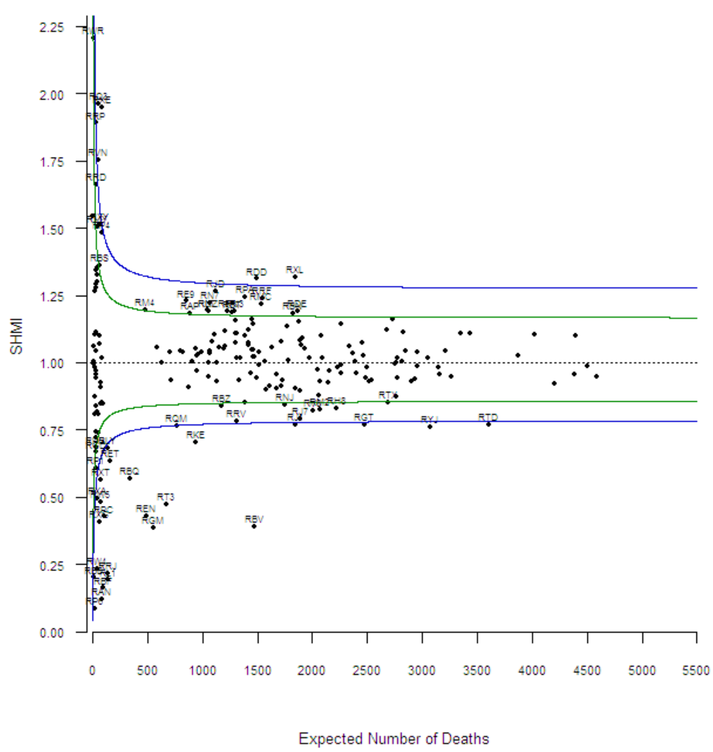
Model 8 – Random Effects



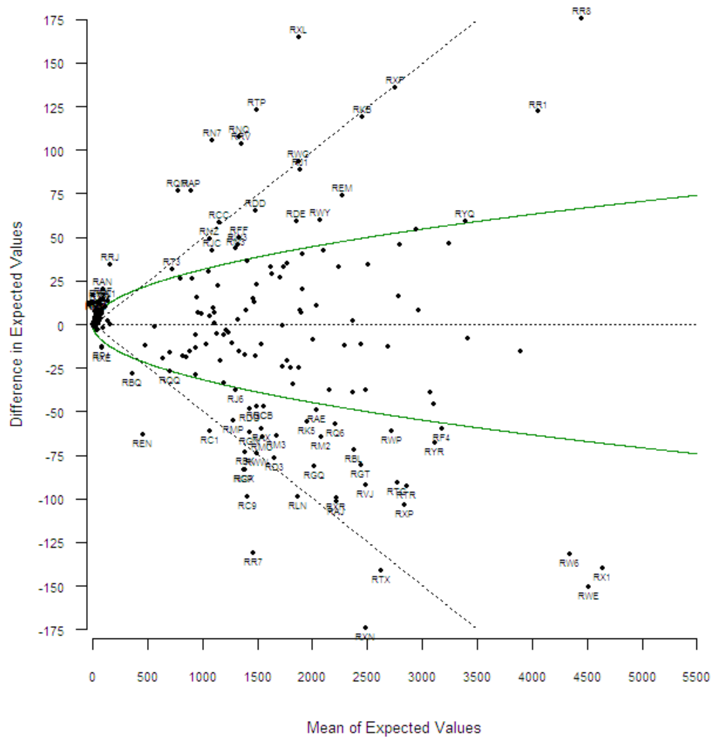
Model 1 vs Model 9



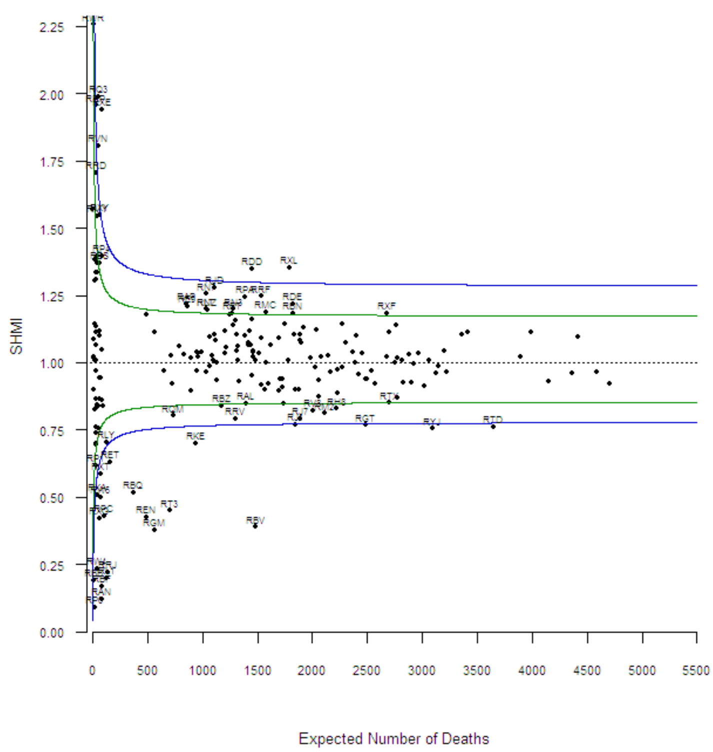
Model 9 – Random Effects



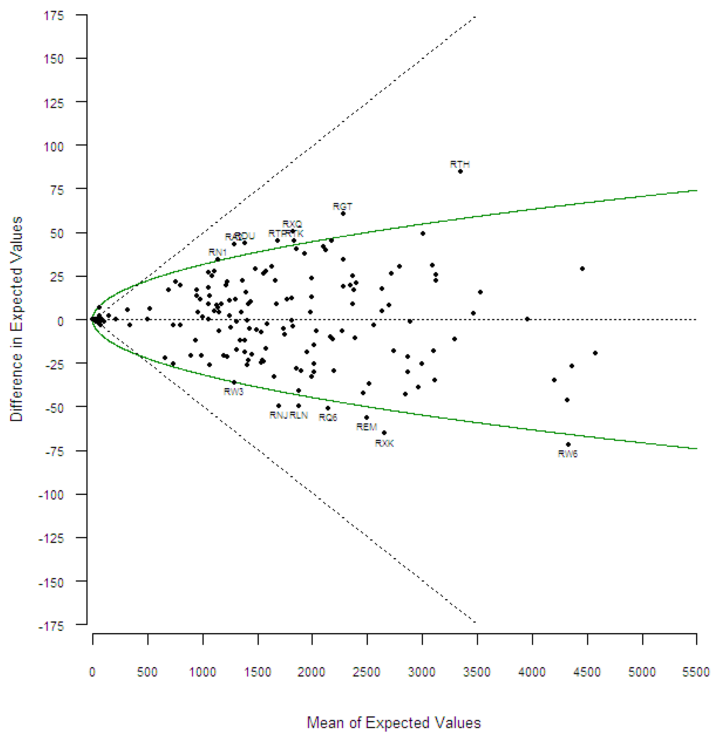
Model 1 vs Model 10



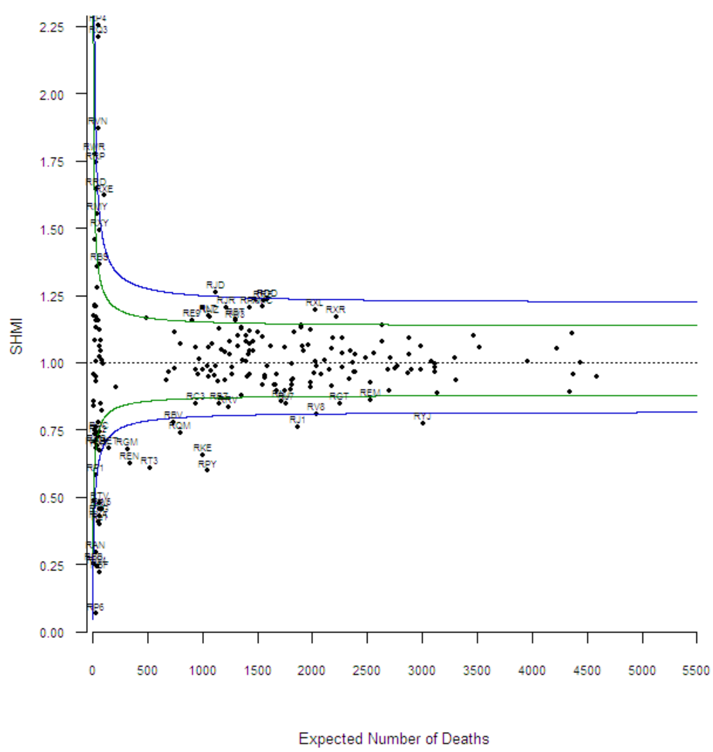
Model 10 – Random Effects



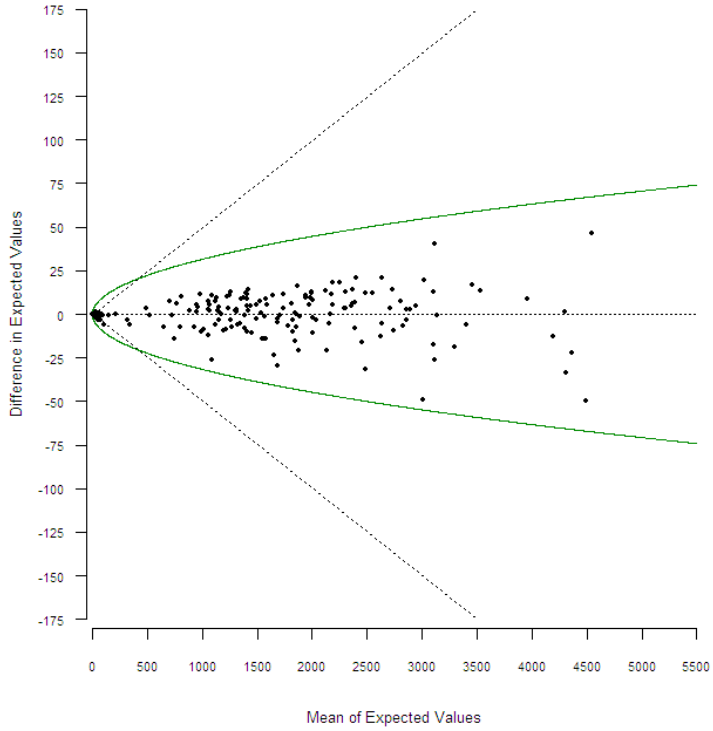
Model 2 vs Model 11



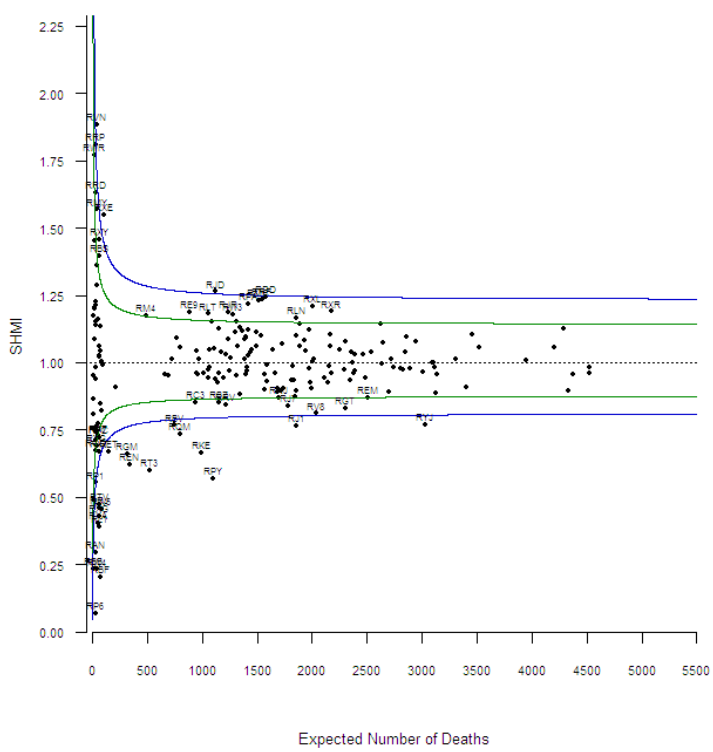
Model 11 – Random Effects



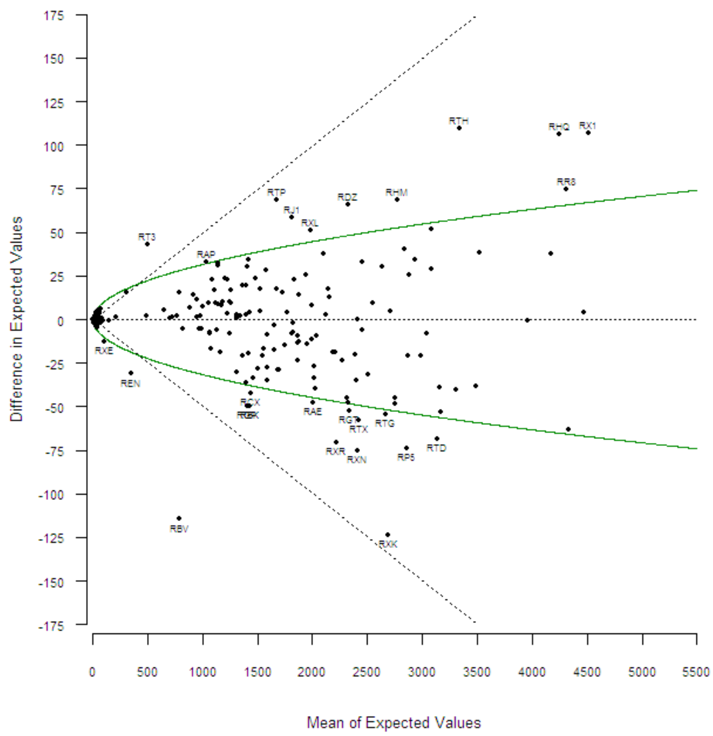
Model 2 vs Model 12



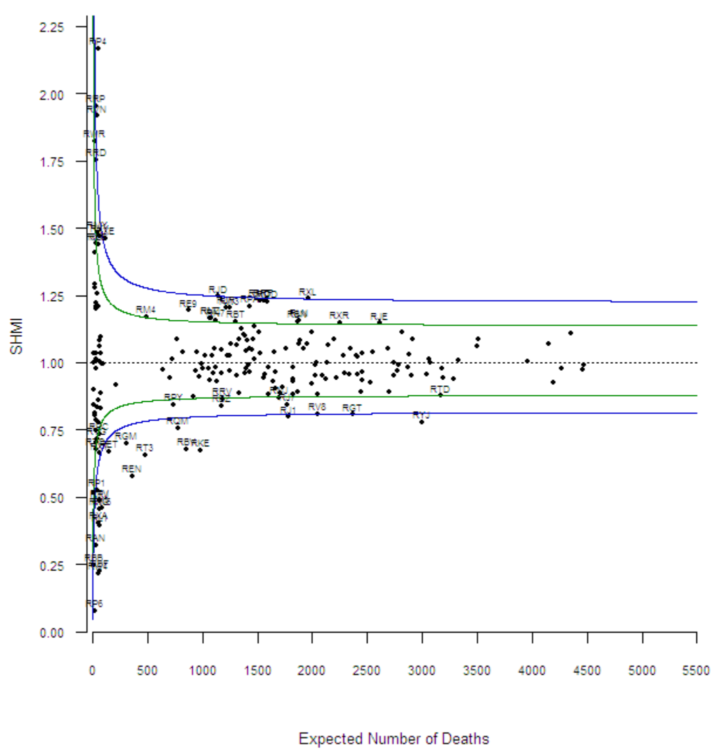
Model 12 – Random Effects



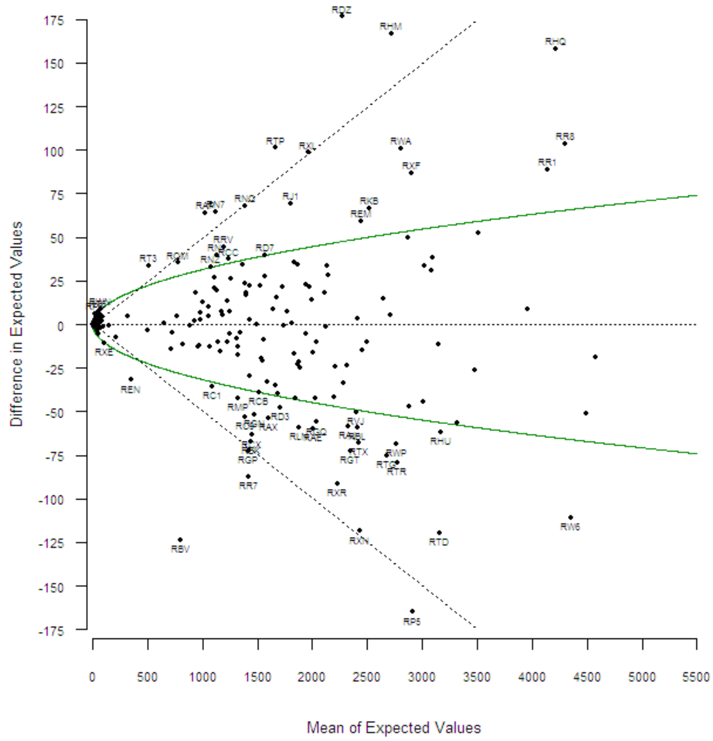
Model 2 vs Model 13



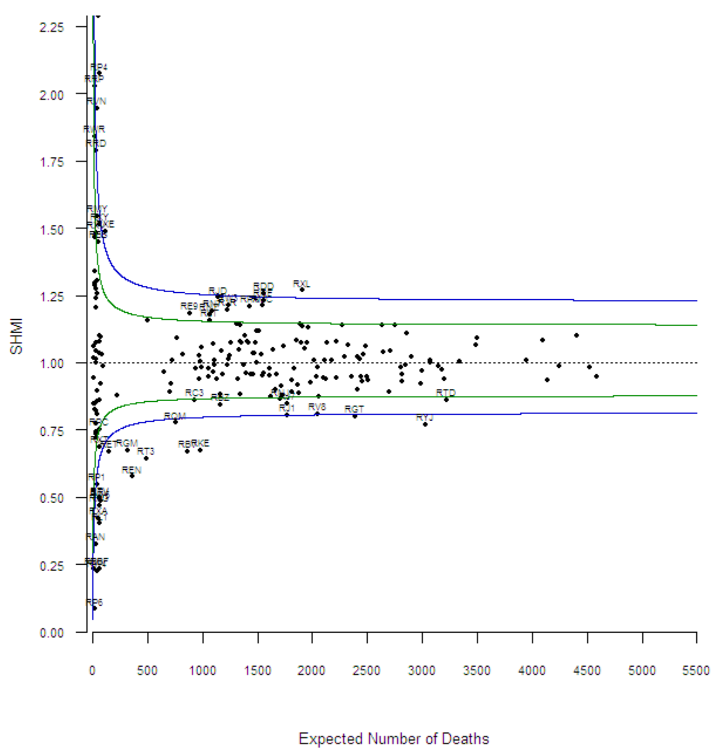
Model 13 – Random Effects



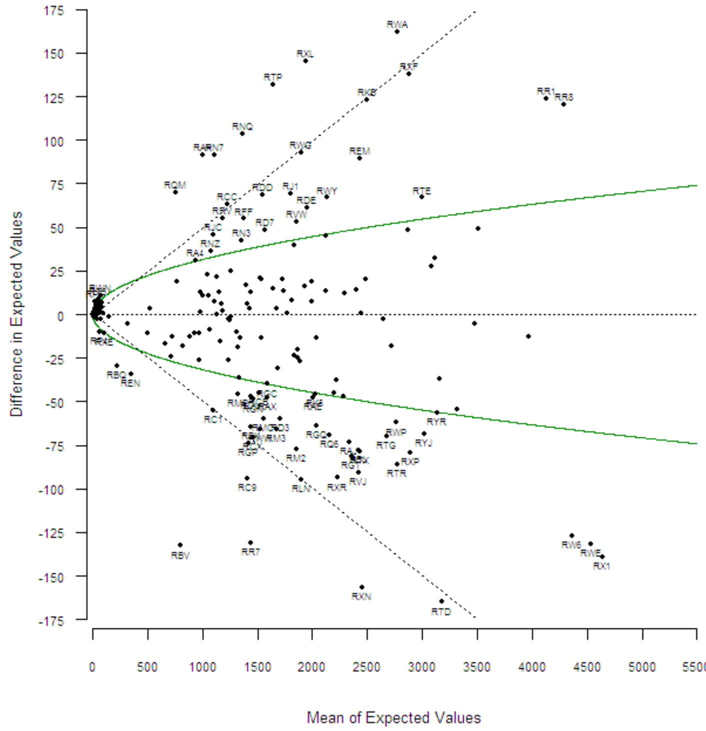
Model 2 vs Model 14



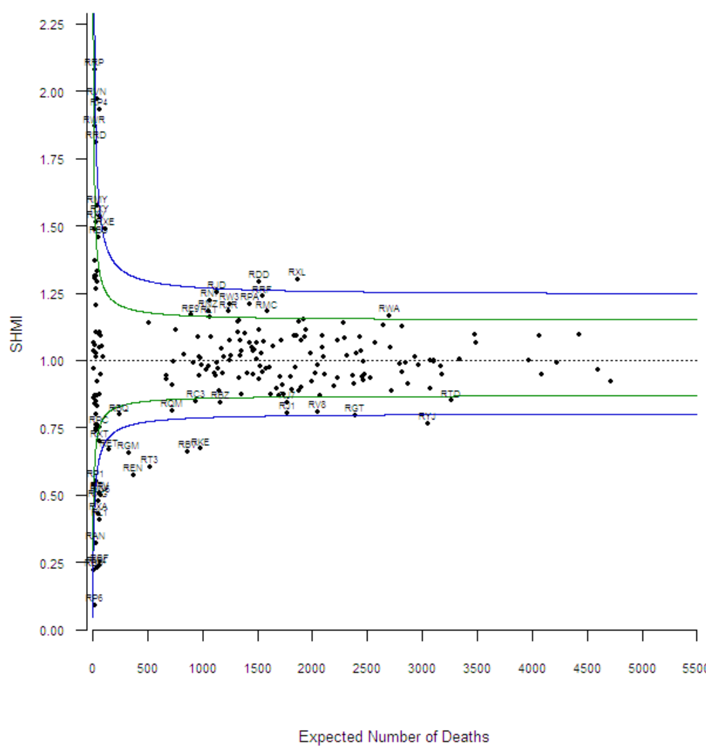
Model 14 – Random Effects



Model 2 vs Model 15



Model 15 (final) – Random Effects



Appendix 5 Tables of SHMI and (Rank) for different models

30 trusts with highest SHMI under model 1 (> 100 observed deaths) – models 1-7

Trust	SHMI 1	SHMI 2	SHMI 3	SHMI 4	SHMI 5	SHMI 6	SHMI 7
RX4	3.392 (1)	3.849 (1)	3.483 (1)	3.577 (1)	3.673 (1)	3.244 (1)	3.168 (1)
RXE	2.295 (2)	1.640 (4)	1.865 (3)	1.894 (3)	1.926 (3)	2.241 (2)	2.017 (3)
RQ3	1.963 (3)	2.291 (2)	1.965 (2)	1.934 (2)	1.966 (2)	1.896 (3)	2.020 (2)
RP4	1.625 (4)	2.28 (3)	1.553 (4)	1.512 (4)	1.440 (5)	1.612 (4)	1.629 (4)
RXY	1.467 (5)	1.490 (5)	1.433 (5)	1.510 (5)	1.550 (4)	1.476 (5)	1.394 (5)
RDD	1.290 (6)	1.235 (8)	1.286 (7)	1.323 (6)	1.383 (6)	1.287 (7)	1.296 (6)
RJD	1.276 (7)	1.253 (6)	1.275 (8)	1.268 (8)	1.289 (8)	1.288 (6)	1.290 (7)
RRF	1.255 (8)	1.248 (7)	1.234 (12)	1.231 (12)	1.242 (11)	1.231 (8)	1.245 (11)
RXL	1.238 (9)	1.206 (11)	1.286 (6)	1.322 (7)	1.365 (7)	1.226 (10)	1.248 (9)
RMC	1.237 (10)	1.227 (9)	1.245 (11)	1.236 (11)	1.193 (17)	1.212 (11)	1.246 (10)
RE9	1.237 (11)	1.184 (14)	1.260 (9)	1.257 (9)	1.232 (13)	1.200 (16)	1.240 (12)
RPA	1.235 (12)	1.208 (10)	1.246 (10)	1.244 (10)	1.244 (10)	1.229 (9)	1.251 (8)
RLT	1.211 (13)	1.173 (16)	1.203 (14)	1.195 (17)	1.189 (21)	1.211 (12)	1.223 (13)
RM4	1.207 (14)	1.165 (18)	1.211 (13)	1.198 (14)	1.178 (23)	1.208 (13)	1.218 (14)
RBN	1.201 (15)	1.157 (19)	1.192 (17)	1.184 (22)	1.189 (19)	1.168 (22)	1.186 (18)
RXR	1.196 (16)	1.187 (13)	1.155 (25)	1.137 (29)	1.132 (31)	1.174 (21)	1.200 (15)
RBT	1.194 (17)	1.152 (20)	1.197 (16)	1.186 (21)	1.189 (20)	1.206 (15)	1.198 (16)
RLN	1.191 (18)	1.170 (17)	1.177 (20)	1.152 (27)	1.124 (32)	1.149 (23)	1.185 (19)
RCX	1.185 (19)	1.120 (26)	1.139 (29)	1.119 (32)	1.120 (33)	1.178 (20)	1.182 (20)
RJR	1.181 (20)	1.183 (15)	1.202 (15)	1.198 (15)	1.175 (24)	1.207 (14)	1.188 (17)
RDE	1.179 (21)	1.117 (27)	1.167 (23)	1.190 (19)	1.234 (12)	1.182 (18)	1.182 (21)
RN3	1.159 (22)	1.111 (30)	1.178 (19)	1.197 (16)	1.205 (15)	1.183 (17)	1.167 (22)
RMP	1.154 (23)	1.113 (29)	1.121 (32)	1.108 (36)	1.108 (38)	1.133 (27)	1.156 (24)
RFS	1.149 (24)	1.099 (31)	1.159 (24)	1.167 (24)	1.167 (26)	1.141 (25)	1.148 (25)
RWW	1.146 (25)	1.117 (28)	1.168 (22)	1.153 (26)	1.091 (49)	1.140 (26)	1.144 (27)
RNZ	1.142 (26)	1.142 (21)	1.174 (21)	1.189 (20)	1.194 (16)	1.180 (19)	1.159 (23)
RN7	1.139 (27)	1.124 (24)	1.184 (18)	1.220 (13)	1.249 (9)	1.144 (24)	1.146 (26)
RBK	1.134 (28)	1.086 (37)	1.082 (47)	1.057 (55)	1.069 (55)	1.109 (34)	1.138 (28)
RGP	1.131 (29)	1.090 (36)	1.082 (46)	1.046 (63)	1.040 (67)	1.117 (33)	1.131 (32)
RW6	1.130 (30)	1.126 (23)	1.111 (34)	1.093 (43)	1.088 (50)	1.105 (38)	1.131 (31)

30 trusts with highest SHMI (> 100 observed deaths) – models 8-15

Trust	SHMI 8	SHMI 9	SHMI 10	SHMI 11	SHMI 12	SHMI 13	SHMI 14	SHMI 15
RX4	3.406 (1)	3.431 (1)	3.492 (1)	3.838 (1)	3.818 (1)	3.892 (1)	3.911 (1)	3.966 (1)
RXE	1.907 (3)	1.948 (3)	1.939 (3)	1.620 (4)	1.548 (4)	1.463 (5)	1.489 (5)	1.487 (5)
RQ3	1.969 (2)	1.962 (2)	1.989 (2)	2.209 (3)	2.325 (2)	2.298 (2)	2.289 (2)	2.318 (2)
RP4	1.543 (4)	1.483 (5)	1.394 (5)	2.255 (2)	2.313 (3)	2.167 (3)	2.073 (3)	1.930 (3)
RXY	1.445 (5)	1.513 (4)	1.547 (4)	1.493 (5)	1.459 (5)	1.470 (4)	1.514 (4)	1.536 (4)
RDD	1.281 (7)	1.311 (7)	1.348 (7)	1.233 (7)	1.242 (7)	1.228 (10)	1.257 (7)	1.291 (7)
RJD	1.271 (8)	1.266 (8)	1.277 (8)	1.262 (6)	1.263 (6)	1.246 (6)	1.241 (8)	1.252 (8)
RRF	1.238 (12)	1.239 (10)	1.246 (10)	1.229 (8)	1.237 (8)	1.231 (8)	1.231 (9)	1.237 (9)
RXL	1.283 (6)	1.317 (6)	1.352 (6)	1.197 (12)	1.211 (11)	1.237 (7)	1.268 (6)	1.300 (6)
RMC	1.240 (11)	1.218 (13)	1.186 (18)	1.207 (9)	1.233 (9)	1.231 (9)	1.211 (11)	1.18 (15)
RE9	1.244 (10)	1.229 (11)	1.210 (14)	1.157 (18)	1.187 (13)	1.194 (14)	1.183 (15)	1.168 (16)
RPA	1.244 (9)	1.242 (9)	1.242 (11)	1.203 (11)	1.219 (10)	1.211 (11)	1.210 (12)	1.211 (12)
RLT	1.200 (15)	1.194 (15)	1.198 (15)	1.173 (13)	1.181 (15)	1.164 (17)	1.159 (17)	1.163 (18)
RM4	1.210 (13)	1.197 (14)	1.178 (22)	1.165 (16)	1.174 (17)	1.170 (15)	1.158 (18)	1.140 (23)
RBN	1.193 (17)	1.184 (21)	1.185 (19)	1.133 (22)	1.144 (21)	1.152 (21)	1.144 (20)	1.145 (21)
RXR	1.156 (25)	1.145 (27)	1.143 (27)	1.171 (14)	1.19 (12)	1.150 (22)	1.139 (22)	1.138 (24)
RBT	1.200 (16)	1.187 (20)	1.184 (21)	1.162 (17)	1.153 (20)	1.154 (20)	1.145 (19)	1.143 (22)
RLN	1.175 (21)	1.151 (26)	1.130 (30)	1.14 (21)	1.165 (18)	1.156 (19)	1.134 (25)	1.113 (28)
RCX	1.142 (28)	1.122 (30)	1.116 (33)	1.116 (27)	1.122 (27)	1.088 (34)	1.072 (44)	1.067 (49)
RJR	1.202 (14)	1.193 (16)	1.177 (23)	1.204 (10)	1.187 (14)	1.206 (12)	1.196 (13)	1.181 (14)
RDE	1.170 (22)	1.191 (17)	1.217 (13)	1.12 (26)	1.123 (26)	1.111 (28)	1.130 (26)	1.153 (19)
RN3	1.176 (20)	1.191 (19)	1.198 (16)	1.129 (23)	1.119 (28)	1.127 (25)	1.140 (21)	1.147 (20)
RMP	1.121 (32)	1.109 (34)	1.106 (40)	1.098 (34)	1.114 (30)	1.087 (35)	1.078 (37)	1.075 (44)
RFS	1.158 (24)	1.162 (24)	1.16 (24)	1.094 (35)	1.100 (33)	1.112 (27)	1.116 (28)	1.115 (27)
RWW	1.163 (23)	1.142 (28)	1.091 (48)	1.112 (30)	1.115 (29)	1.135 (24)	1.117 (27)	1.069 (48)
RNZ	1.179 (19)	1.191 (18)	1.196 (17)	1.169 (15)	1.154 (19)	1.166 (16)	1.178 (16)	1.182 (13)
RN7	1.185 (18)	1.223 (12)	1.255 (9)	1.128 (24)	1.128 (24)	1.156 (18)	1.191 (14)	1.221 (10)
RBK	1.088 (45)	1.074 (50)	1.075 (52)	1.068 (45)	1.090 (38)	1.049 (53)	1.036 (58)	1.038 (58)
RGP	1.085 (47)	1.065 (51)	1.065 (55)	1.080 (39)	1.089 (39)	1.052 (52)	1.035 (59)	1.034 (61)
RW6	1.112 (34)	1.100 (40)	1.096 (46)	1.108 (31)	1.127 (25)	1.110 (29)	1.098 (31)	1.094 (33)

References

Aylin P, Bottle A, Jarman B (2009)Monitoring hospital mortality: a response to the University of Birmingham report on HSMRs. Dr Foster Unit, Imperial College.

Aylin P, Bottle A, Jen MH and Middleton S. HSMR mortality indicators. Imperial College Technical Document 26th November 2010

Bottle A, Jarman B and Aylin B. Strengths and weaknesses of hospital standardised mortality ratios. BMJ 2011: 341:c7116

Julious SA, Nicholl J, George S (2001) Why do we continue to use standardized mortality ratios for small area comparisons? Journal of Public Health Medicine 2001: 23:40-46.

Roalfe AK, Holder RL, Wilson S (2008) Standardisation of rates using logistic regression: a comparison with the direct method *BMC Health Services Research* 2008, 8:275 doi:10.1186/1472-6963-8-275

Spiegelhalter DJ (2005a) Funnel plots for comparing institutional performance. Statist. Med. 2005: 24:1185-1202.

Spiegelhalter DJ (2005b) Handling over-dispersion of performance indicators. Qual Saf Health Care 2005 Oct;14(5):347-51.

Whalley L Report from the steering group for the national review of the Hospital Standardised Mortality Ration. NHS Information Centre for Health and Social Care, July 2010