ADVANCING SURGICAL STANDARDS

August 2014

Statistical methods developed for the National Hip Fracture Database annual report, 2014

A technical report

Prepared by:

Dr Carmen Tsang and Dr David Cromwell

The Clinical Effectiveness Unit, Royal College of Surgeons of England

Suggested citation. *Tsang C, Cromwell D. Statistical methods developed for the National Hip Fracture Database annual report, 2014: a technical report. London: The Royal College of Surgeons of England, 2014.*

The Royal College of Surgeons of England is dedicated to enabling surgeons achieve and maintain the highest standards of surgical practice and patient care. To achieve this, the College is committed to making information on surgical care accessible to the public, patients, health professionals, regulators and policy makers.

Registered charity number: 212808

CONTENTS

FIGURES

TABLES

APPENDICES

INTRODUCTION

The Clinical Effectiveness Unit (CEU) at the Royal College of Surgeons of England (RCS) was commissioned by the Royal College of Physicians (RCP) to produce the analyses for the National Hip Fracture Database (NHFD) 2014 annual report. The NHFD is part of the Falls and Fragility Fractures Audit Programme (FFFAP), which is commissioned by the Healthcare Quality Improvement Partnership (HQIP) and forms part of the National Clinical Audit and Patient Outcomes Programme (NCAPOP).

This report describes work to develop the statistical methods used in the analyses for the NHFD 2014 annual report. The work was undertaken between March and May 2014, and used historic NHFD audit data, and data extracts from Hospital Episode Statistics (HES) and Patient Episode Database for Wales (PEDW):

- NHFD data from 1st January 2011 to 31st March 2013
- HES data from financial year 2011/12 (1st April 2011 to 31st March 2012)
- PEDW data from financial year 2012/13 (1st April 2012 to 31st March 2013)

The main element of the work was the evaluation of statistical approaches for risk (casemix) adjusted analyses of outcome measures (mortality within 30 days and return home from home within 30 days). To ensure consistency across NHFD analyses and to assess the quality of the NHFD data, the work also included refinement of NHFD audit case ascertainment, and an investigation of the agreement between mortality within 30 days of patients in the NHFD audit and crude (unadjusted) in-hospital mortality.

This report was prepared by Dr Carmen Tsang and Dr David Cromwell at the CEU for the NHFD in August 2014. All analyses were carried out using STATA version 11.

If you have queries about this document, please contact Carmen Tsang, Lecturer at the CEU, ctsang@rcseng.ac.uk

DATA QUALITY SCREENING

Basic data quality checks were undertaken prior to this developmental work (Table 1). The same checks were performed on the data used in the final analyses for the 2014 annual report. Further checks were made in the risk adjusted outcome analyses and are described later in this document.

Table 1 Data quality screening steps

Data field	Action
Sex	Drop if missing
Age	Drop if <60 years or >110 years
Died before admitted	Drop if death date before admission date
Duplicate records (NHFD only)	Drop if duplicate record (based on Crown Informatics' "Possible Dupe" flag)

CASE ASCERTAINMENT

The original case ascertainment method for the 2013 NHFD annual report was modified to take into account changes to hospitals and NHS trusts (reorganisation and closures).[1, 2] The original method for calculating the denominator from administrative data excluded elective admissions for hip fracture. Although the majority of hip fracture cases will be emergency admissions, elective cases were included in the analyses for the 2014 annual report to ensure full capture of all cases.

As linked NHFD and HES data were not available, the refinement of case ascertainment was achieved by:

- NHFD compared to HES: numerator (NHFD audit) and denominator (standalone HES) 2011/12
- NHFD compared to PEDW: numerator (NHFD audit) and denominator (standalone PEDW) 2012/13

NHFD CASE ASCERTAINMENT FOR ENGLISH HOSPITALS

The case ascertainment for English hospitals in the NHFD was based on HES data from the financial year 2011/12. The validity of denominators derived from HES data was assessed by comparing the estimated number of hip fracture cases in HES with the number of cases in the NHFD Facilities Audit (from the 2012 NHFD annual report).[3] Analyses were based on 162 hospitals in England.

The number of cases captured in the Facilities Audit was 58,010. In the HES data, there were 58,123 cases. Additional checks were made by focusing on the case ascertainment in 27 hospitals with historically good data completion in the NHFD audit.[1] The results were explored graphically (not shown). Overall, there was reasonable agreement between NHFD and HES.

NHFD CASE ASSERTAINMENT FOR WELSH HOSPITALS

The PEDW dataset for financial year 2012/13 was used to assess the case ascertainment of 13 eligible hospitals in Wales. There were 3,665 hip fracture cases reported in the NHFD Facilities Audit, and 3,639 cases in PEDW. Some potentially ineligible cases may have been included in these analyses as PEDW age data were only available in ten year bands (e.g. 100 years or older). There were 39 records with age recorded as “100 years or older”.

Case ascertainment in NHFD and PEDW was similar for the majority of hospitals in Wales. However, for one hospital, there were marked differences in the number of hip fracture cases recorded in PEDW compared with the NHFD audit and Facilities Audit (same number of cases reported in the two latter datasets). PEDW contained 194 fewer cases for the hospital.

CRUDE RATES OF MORTALITY WITHIN 30 DAYS

To assess the representativeness of NHFD data, crude rates of mortality within 30 days of admission were calculated using HES and NHFD data from 2011/12. Agreement between HES and NHFD for rates in 162 hospitals in England was explored graphically using scatter plots and Bland-Altman plots.[4] Overall agreement in mortality rates between the datasets was reasonable, with six hospitals falling outside the 95% limits of agreement. Sub-group analyses were performed using data from the 27 hospitals with a history of consistent audit participation, with good agreement between the datasets for all hospitals except one, which was below the lower 95% limit of agreement.

CRUDE RATES OF IN-HOSPITAL MORTALITY

A further assessment of the representativeness of NHFD data was made by comparing the number of recorded deaths before discharge with figures derived from HES data. These analyses were performed for the 162 hospitals in England that were included in the NHFD 2012 report.[3] Bland Altman plots were again used to graphically assess agreement between the two data sources. Reasonable agreement was found, with nine out of the 162 hospitals (5.56%) falling outside of the 95% limits of agreement (three hospitals above the upper 95% limit and six hospitals below the lower 95% limit).

RISK ADJUSTED OUTCOME MEASURES

We conducted work to assess and refine statistical models used to produce risk adjusted outcomes in the NHDS annual report. Analyses were performed on:

1. Mortality within 30 days (2013 and 2011 to 2013)
2. Return home from home within 30 days (2013)

The methods used to develop the statistical models, select the final risk adjustment models and the results of the analyses are presented in the following report sections.

RISK ADJUSTMENT METHODS

The analyses for the NHFD annual reports in previous years were carried out by Quantics.[5] They had developed a classification tree (CART) approach for risk (casemix) adjustment. The NHFD project team decided to examine the statistical methods for the 2014 annual report, with a focus on evaluating how patient characteristics that may affect patient outcomes were taken into account in the risk adjusted analyses of patient outcomes.

The work examined the potential of using logistic regression as an alternative approach to CART for risk adjustment. The first step in the work was to compare the performances (discrimination and calibration) of the existing CART model and a model created using logistic regression that contained the same patient variables. The second step was to develop a number of logistic regression models with different combinations of the patient variables to identify the best performing model for use in the NHFD 2014 annual report.

CHANGES FROM THE NHFD 2013 REPORT ANALYSES

In the CART approach, records with missing values of patient characteristic(s) used in the risk adjusted analyses were reassigned to other categories. In contrast, the refined analyses using logistic regression did not make any assumptions about the missing values but instead retained these records in a category for missing values.

When comparing the funnel plots presented in developmental work and the 2014 annual report with those published in previous NHFD reports, it should be noted that the method used to define the control limits (95% and 99.8% control limits) was not clearly documented in the 2012 and 2013 NHFD reports. In this report, we follow the standard approach outlined by Spiegelhalter (2005).[6]

MORTALITY WITHIN 30 DAYS

This section describes the methods used to develop a risk adjustment model for mortality within 30 days (2013 data and 2011 to 2013 data).

The analysis used NHFD data from the financial year 2012/13. The first regression models (NHFD models 1 and 2) contained only the four patient variables used in the NHFD classification tree (age at admission, sex, American Society of Anesthesiologists (ASA) grade and accompanied to walk outdoors status). The performances of these two models were then compared with the performance of the classification tree used in the 2013 NHFD annual report.[1]

SELECTION OF CASEMIX VARIABLES

Two variables were aggregated to replicate the 2013 annual report analyses. First, ASA grade was aggregated to:

- ASA grades 1 to 3 and unknown
- ASA grade 4
- ASA grade 5

The second variable that was aggregated was accompanied to walk outdoors status (categories: yes, no, wheelchair or bedbound, never goes outdoors, and unknown). The “no” category was combined with “unknown” in the CART model.

In developing the new regression models, alternative methods of aggregating ASA grade were tested. In particular, due to small numbers, it was decided to use the following aggregation:

- ASA grades 1 to 2
- ASA grade 3 and unknown
- ASA grades 4 and 5

Additional variables available in the NHFD dataset were included in the CEU regression models and were selected in consultation with the NHFD Clinical Leads to ensure their inclusion was clinically sound. All variables of interest (Table 2) were statistically significantly associated with mortality within 30 days in crude analyses (x^2 test and Mann Whitney Wilcoxon test). Fracture type was also aggregated to intracapsular versus extracapsular on recommendation by the NHFD Clinical Leads.

We note that the original question on walking outdoors was superseded by “mobility” in the NHFD 2014 audit. This new “mobility” measure was developed by Quantics prior to 2014 and combines all four walking ability measures (walking ability indoors and outdoors, as well as accompanied to walk indoors and outdoors). However, because many cases have missing “mobility” status (43.1% of cases had unknown or missing mobility status in the 2012/13 dataset), the former walking ability measures were used to develop the models that are evaluated in this report. Future work to improve the risk adjustment models should consider using “mobility” as completeness of this data field improves over time.

Table 2 Strength of association between patient characteristics used in risk adjustment models and mortality within 30 days

Variable	x^2	p-value
ASA grade	1547.2	<0.001
Mobility status	967.2	<0.001
Age at fracture	747.1	<0.001
Walking ability indoors	679.5	<0.001
Source of admission	504.1	<0.001
Accompanied to walk outdoors	473.4	<0.001
Sex	323.1	<0.001
Preoperative Abbreviated Mental Test (AMT) score	112.4	<0.001
Fracture type	16.8	0.001
Social deprivation	3.5	<0.001

Table 2 highlights that ASA grade was most strongly associated with mortality within 30 days (x^2 1547.2, p<0.001). It also shows that out of the three walking ability measures evaluated in univariate analyses, mobility status demonstrated strongest explanatory power (x^2 967.2, p<0.001) with approximately double the power of accompanied to walk outdoors status to predict mortality within 30 days (x^2 473.4, p<0.001). Walking ability indoors was also strongly associated with mortality (x^2 679.5, p<0.001). Fracture type and social deprivation had the weakest explanatory power out of all the patient characteristics that were assessed.

INCLUSION AND EXCLUSION CRITERIA

The inclusion criteria for the risk adjusted mortality analyses were:

- Admitted between 1st January and 31st December 2013 (in the developmental work - between 1st March 2012 and 28th February 2013)
- Valid sex
- Aged between 60 and 110 years
- Known mortality status at 30 days
- Records in "draft status" in the NHFD system

In the developmental work, records were only included in the risk adjusted analyses if they were from hospitals that were included in the 2013 annual report and that also had at least 80% case ascertainment. In the final analyses (2013 and 2011 to 2013), all hospitals were included because sensitivity analyses indicated that the exclusion of hospitals with less than 80% case ascertainment had negligible effects on risk adjustment.

Exclusion criteria:

- Records not matched to ONS death data
- Date of death before date of admission
- Records from hospitals with no deaths within 30 days of admission
- Potentially duplicate records (based on Crown Informatics flag for duplicate record)

Hospitals in Northern Ireland were excluded from the developmental work because validated mortality data were not available for the four hospitals (ALT, CRG, RVB and NUH). Two hospitals in England were excluded from the analyses because of poor quality data (BRT and HOR).

SELECTION OF A FINAL RISK ADJUSTMENT MODEL FOR MORTALITY

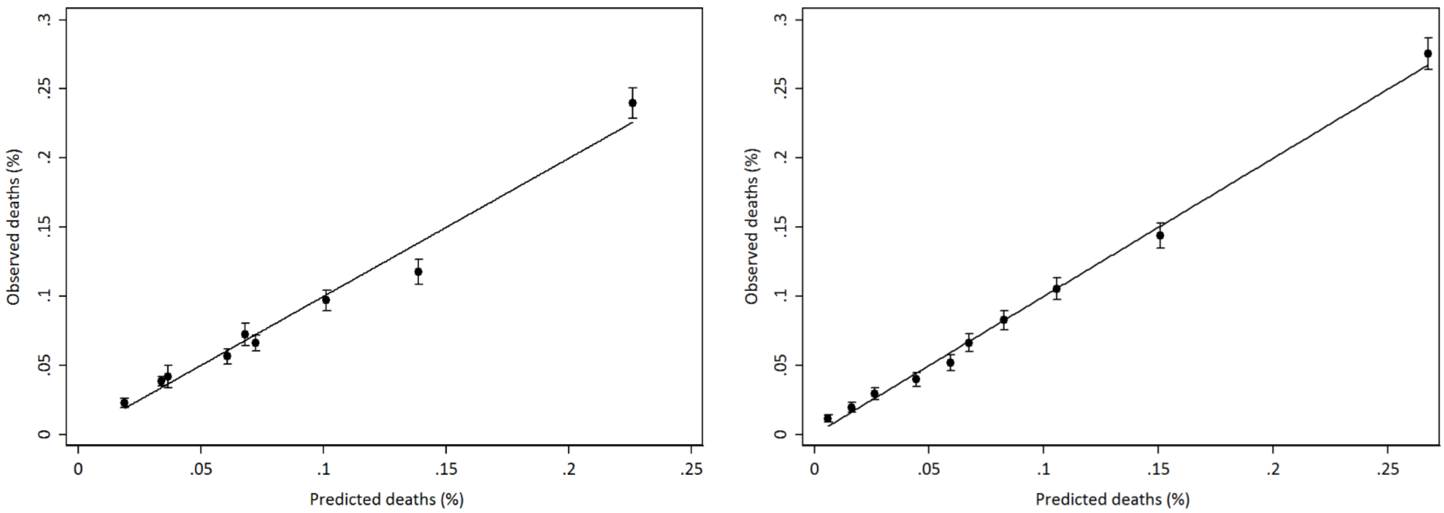
The performances of all risk adjustment models using NHFD 2012/13 data are shown in Appendix 1. Model performance was assessed graphically in calibration plots (which show the difference between the actual and predicted number of deaths within 30 days by groups of increasing risk) and Receiver Operating Characteristic (ROC) curve graphs. The area under the ROC curve (c-statistic) describes the model's discrimination between patients of high and low risk, with values closer to 1 indicating better performance. The CEU models included variables in a linear fashion by order of predictive power estimated in univariate analyses.

CEU model 12 was the best performing risk adjustment model, with good discrimination (c-statistic 0.756) and better calibration than most other models (Hosmer-Lemeshow test, x^2=34.9, 8 df, p-value<0.001). This model contained seven variables: age, ASA grade (grouped), sex, source of admission, mobility, preoperative AMT score and social deprivation. However, this model was rejected because deprivation, in the form of Index of Multiple Deprivation (IMD) scores, is derived from Lower Super Output Area (LSOA) codes and these are only valid for English households. Therefore, patients residing in Wales, Scotland and Northern Ireland were missing deprivation data. AMT scores (preoperative and postoperative) were also inconsistently recorded outside of England. While the issue of missing values for social deprivation and preoperative AMT score could be mitigated by using imputed values, alternative regression models without these two variables performed comparably well to CEU model 12 (Appendix 1).

CEU model 17 contained six variables: age, ASA grade (grouped), sex, source of admission (grouped), walking ability indoors and fracture type. Appendix 1 shows that this model performs slightly less well than the best performing models, CEU models 12 and 16, but it was a suitable alternative model given the limitations of the deprivation and preoperative AMT variables. It also performed better than the original NHFD CART models (i.e. had a higher c-statistic). While both the original NHFD and final CEU (model 17) models were reasonably well calibrated, the CEU model fitted the data better (Figure 1). This improved calibration is indicated by the decile groups being more evenly spaced and being closer to the fitted line. N.B. The NHFD model contained 9 instead of 10 groups of the predicted probability of death due to a large number of ties in the predicted values.

Figure 1 Calibration of NHFD and CEU models for mortality within 30 days (2012/13)

NHFD model 1 (n=58,314) **CEU model 17 (n=59,654)**



EFFECT OF RISK ADJUSTMENT ON MORTALITY WITHIN 30 DAYS

The 2014 NHFD report presents crude and risk adjusted analyses for mortality within 30 days using data from three years (2011 to 2013) in a funnel plot. The control limits were derived using the normal approximation to the binomial distribution, and were defined to correspond to two and three standard deviations above and below the overall national average, respectively. The control limits can be described alternatively as 95% limits (two standard deviations) and 99.8% limits (three standard deviations). These limits indicate whether the difference between the mortality rate at a hospital and the national average is greater than would be expected from random fluctuations, and by how much.

OVER-DISPERSION

For mortality rates, the amount of expected variation between hospitals is determined by the national average (due to the relationship between the mean and the variance in a binomial distribution). However, the observed data may not follow this expected pattern and, particularly in large samples, there can be evidence of over-dispersion (i.e. the variation between hospitals is greater than expected). For the mortality rates derived from three years of data (2011 to 2013), we formally tested between-hospital variation and found a statistically significant amount of over-dispersion. Consequently, an adjustment to account for this over-dispersion was incorporated into the control limits.

Adding the measure for over-dispersion changed the typical interpretation of how hospitals performed over the three year period. In the standard approach, each hospital is assumed to perform at the same level over the entire three year period. In the approach with over-dispersion, it is assumed that each hospital had its own "true" underlying mortality rate for the three-year period, which caused it to differ by a small amount from the national average and reflected the cumulated influence of characteristics that could not be captured by the variables in the risk-adjustment model.

The amount of between-hospital variation that was added to the standard control-limits was calculated using an additive random-effects model, with 10% winsorization. The variation was assumed to be normally distributed around a mean of zero.[6]

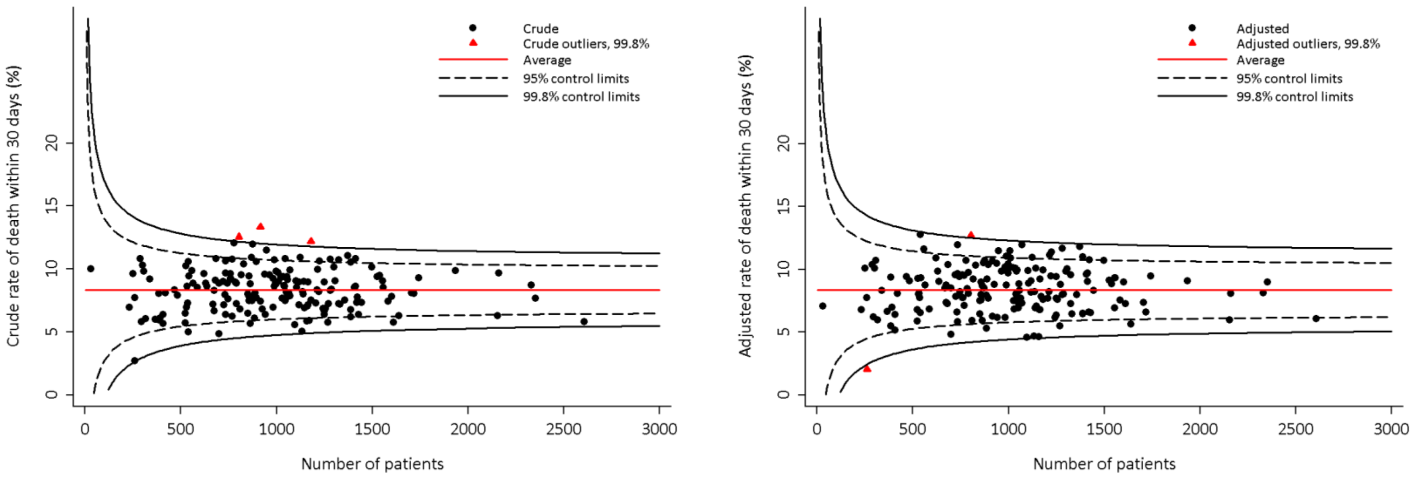
There was no evidence of over-dispersion when mortality rates were derived from one year (2013) of NHFD data.

EFFECT OF ADJUSTMENT ON MORTALITY WITHIN 30 DAYS, 2011-2013

The rate of mortality within 30 days in 178 eligible hospitals out of 185 hospitals was 8.4% (n=14,797/177,196). 16 hospitals had unadjusted mortality rates above the upper 95% limit (Figure 2). Three of these sites had crude rates above the upper 99.8% limit. 13 hospitals had mortality rates below the lower 95% limit and no hospitals had mortality rates below the lower 99.8% limit.

Figure 2 Crude and adjusted rates of mortality within 30 days with a measure for over-dispersion, 2011-2013 (n=177,196)

Crude rate **Adjusted rate**



After adjustment, only one hospital had an adjusted mortality rate above the upper 99.8% limit (Figure 2); there were 14 hospitals between the upper 95% and 99.8% limit. Ten hospitals had rates that fell between the lower 95% and 99.8% limits, and one hospital had a rate below the lower 99.8% limit.

RETURN HOME FROM HOME WITHIN 30 DAYS

This section describes the methods used to develop a risk adjustment model for the analysis of patients who returned home or were discharged to sheltered housing from home or sheltered housing within 30 days. Initial analyses used the definition of return home from home **at 30 days** from the 2013 NHFD annual report. However, subsequent analyses applied a new definition for return home from home **within 30 days** and that is why the two terms are used in this section.

DEVELOPING A RISK ADJUSTMENT MODEL FOR RETURN HOME

NHFD data from financial year 2012/13 were used to develop logistic regression models for risk adjustment. As with the models for mortality within 30 days, two NHFD models were developed for the return home analyses (NHFD models 1 and 2), with age as a categorical versus a continuous variable (all other models included age as a categorical variable only). The performances of these two models were then compared with the performance of the classification tree used in the 2013 NHFD annual report.[1]

To ensure that replication of the NHFD models was accurate, the rates of return home from home at 30 days calculated in the developmental work were compared with the rates reported in the 2013 annual report. There were minimal differences in the outcome measure (0.5% difference) and in recorded missing data on residential status at 30 days after admission (2.4%).

SELECTION OF CASEMIX VARIABLES

As with the mortality analyses, the variables ASA grade and accompanied to walk outdoors were aggregated (see Mortality within 30 days section - Selection of casemix variables). All variables of interest (Table 3) were statistically significantly associated with return home from home within 30 days in crude analyses (x^2 test and Mann Whitney Wilcoxon test).

Table 3 Strength of association between patient characteristics used in risk adjustment models and return home from home at 30 days

Variable	x^2	p-value
Age	1850.5	<0.001
Mobility	1826.0	<0.001
Walking ability indoors	1534.0	<0.001
Accompanied to walk outdoors	1330.3	<0.001
ASA grade	1196.7	<0.001
Fracture type	427.4	<0.001
Preoperative Abbreviated Mental Test (AMT) score	259.0	<0.001
Deprivation	64.0	<0.001
Sex	57.8	<0.001

INCLUSION AND EXCLUSION CRITERIA

The inclusion criteria for the risk adjusted return home analyses were:

- Admitted between 1st January and 31st December 2013 (in the developmental work - between 1st March 2012 and 28th February 2013)
- Valid sex
- Aged between 60 and 110 years
- Admitted from own home or sheltered housing
- With known residential status at 30 days (for return home from home **at 30 days** measure)
- Records in "draft status" in the NHFD system

Exclusion criteria:

- Potentially duplicate records (based on Crown Informatics flag for duplicate record)

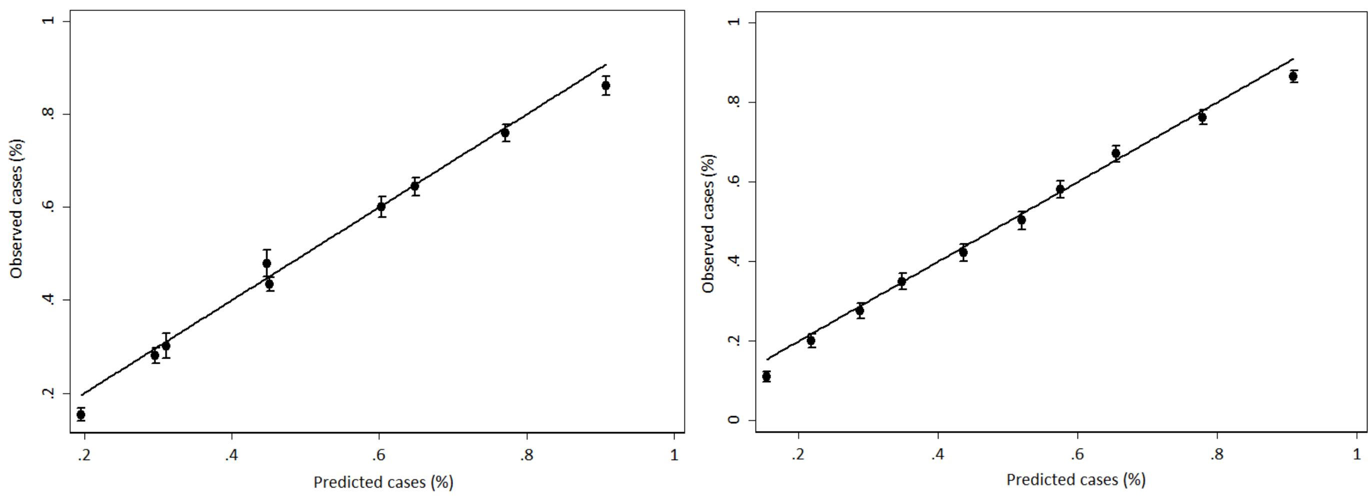
SELECTION OF THE RISK ADJUSTMENT MODEL FOR RETURN HOME

The performances of all risk adjustment models using NHFD 2012/13 data are shown in Appendix 2. Model performance was assessed graphically in calibration plots and by Receiver Operating Characteristic (ROC) curve graphs.

The best performing regression models were CEU models 7 and 8 in terms of discrimination (c-statistic 0.773 and 0.777, respectively). All the regression models were reasonably calibrated, although their poor performance on the Hosmer-Lemeshow goodness of fit tests will be due to the relatively large sample sizes. The calibration plot for model 7 is shown in Figure 3.

Figure 3 Calibration of NHFD and CEU models for return home from home at 30 days (2012/13)

NHFD model 1 (n= 19,561) **CEU model 7 (n= 20,571)**



As CEU models 7 and 8 included social deprivation and preoperative AMT score which are inconsistently recorded outside of England, further model refinement was carried out (see Mortality within 30 days section – Selection of a final risk adjustment model for mortality). CEU model 11 demonstrated good performance compared to the previously developed models, and was used in the analyses described as follows. This model included five patient characteristics: age, ASA grade, sex, walking ability indoors, and fracture type.

MODIFICATION OF THE RETURN HOME MEASURE

The original measure of return home from home was based on the "residential status at 30 days" field in the NHFD audit, but this field has several limitations:

- Some hospitals only reported on patients who are still in hospital at 30 days.
- 30 day follow up may not occur on the 30th day after admission, thus the data may not be truly representative of patients' residential status at 30 days.
- In 2013, 14.7% of cases admitted from their own home or sheltered housing had unknown (or missing) residential status at 30 days after admission (n=7,207/49,036).

Given the caveats above, a more robust measure of how many patients returned home by 30 days after admission was proposed by Andy Williams, NHFD Project Coordinator. NHFD data from 2013 were used to develop the new measure, which was derived from the number of days between the date of admission and date of discharge from the hospital trust, and the discharge destination.

TECHNICAL DETAILS OF THE NEW MEASURE

Like the original measure of return home, cases not admitted from their own home or sheltered housing were excluded (n=15,802/64,838) but cases with unknown or missing residential status at 30 days were included because the new definition did not rely on residential status at 30 days.

The original return home from home measure identified 11,620 cases admitted during 2013 as being at home or in sheltered housing at 30 days post-admission. Using the new measure, 23,639 cases were identified as being at home at 30 days. There was substantial overlap in the number of cases identified by the two measures (n=9,602) but the new measure identified more cases that were not detected using the original measure than vice versa (n=12,019).

COMPARISON OF THE TWO RETURN HOME MEASURES

Two sets of crude and adjusted figures were derived using the following return home from home definitions:

1. Return home from home at 30 days after admission, based on residential status at 30 days.
2. Return home from home within 30 days, based on date of discharge from trust and discharge from trust destination.

Each set of adjusted figures were adjusted for age, ASA grade, sex, walking ability indoors, and fracture type.

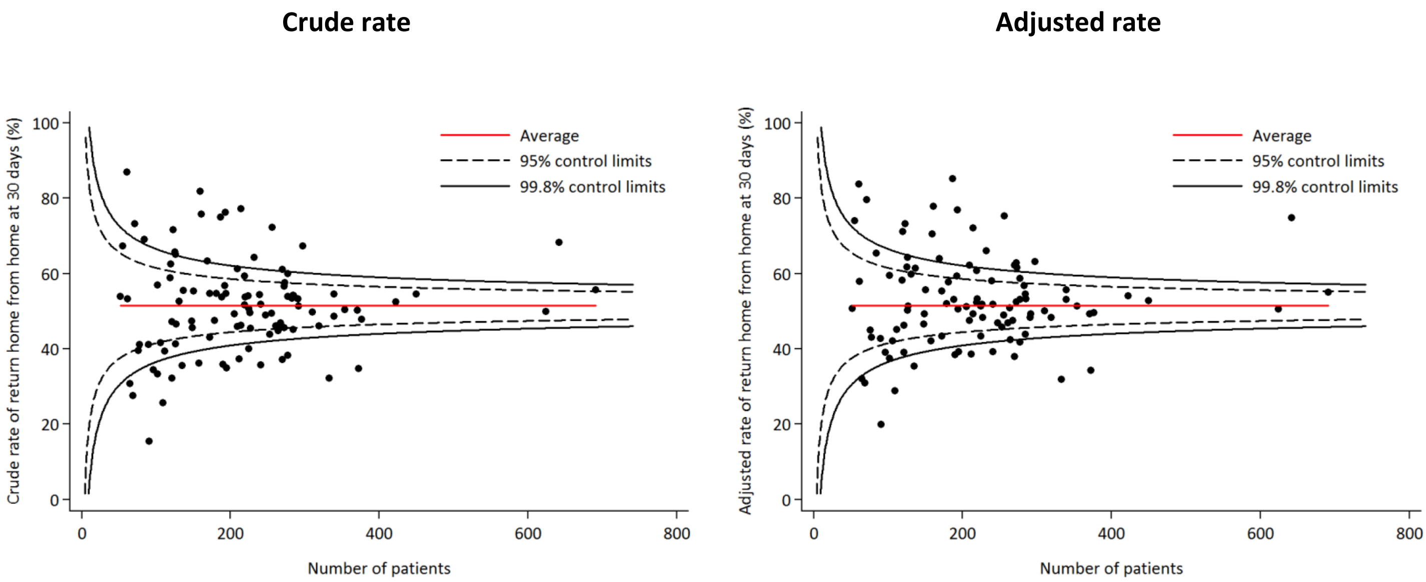
ORIGINAL MEASURE - RETURN HOME FROM HOME AT 30 DAYS

Due to issues with the quality of submitted data, it was necessary to make the following exclusions:

- Records from 73 hospitals with less than 50 eligible cases (n=1,234)
- Records from two hospitals without any cases who returned home from home at 30 days (n=214)
- Records from four hospitals where all cases had unknown or missing residential status (n=1,273).

Figure 4 shows that there was considerable scatter among the hospitals in terms of the rates of return home from home at 30 days. There were 19 hospitals with rates above the upper 99.8% limit, and 13 hospitals with rates below the lower 99.8% limit.

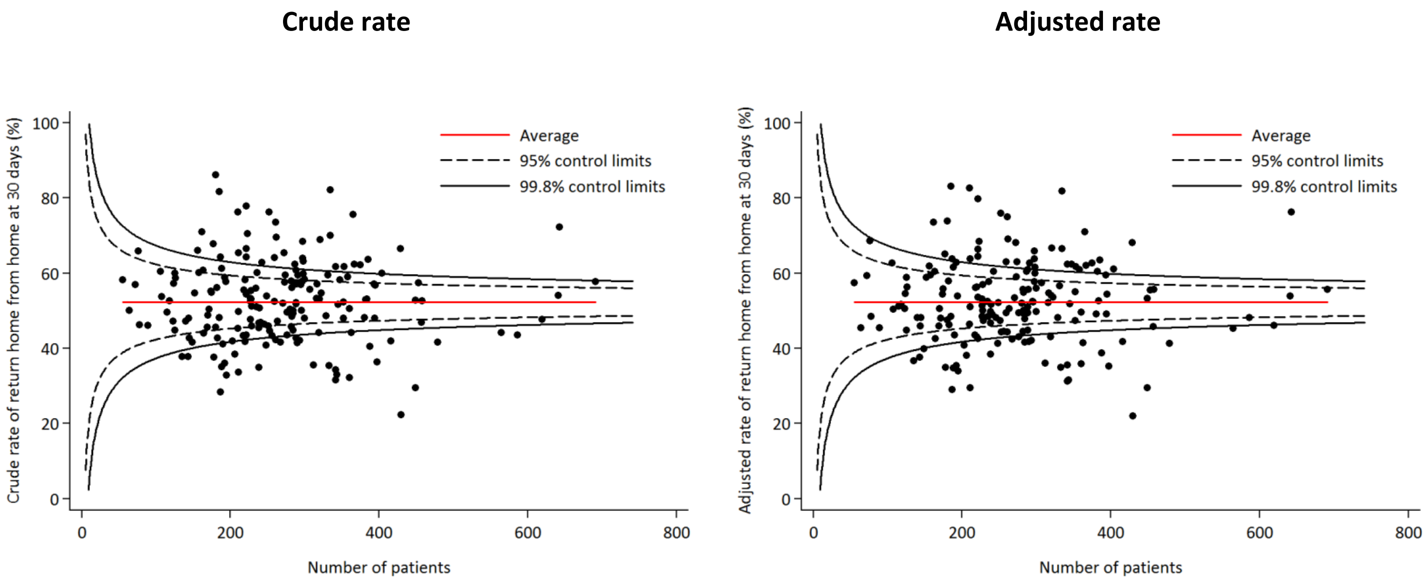
Figure 4 Crude and adjusted rates of return home from home at 30 days (original measure), 2013 (n=22,837)



NEW MEASURE - RETURN HOME FROM HOME WITHIN 30 DAYS

In the analysis using the new return home measure, it was possible to derive estimates for 182 hospitals. As before, there was wide variation in the rate of return home between hospitals in both the crude and adjusted rates (Figure 5). There is indication that there are other risk factors that influence the outcome measure which have not been accounted for in the model.

Figure 5 Crude and adjusted rates of return home from home within 30 days (new measure), 2013 (n=49,036)



COMMENT ON THE FINAL RETURN HOME MEASURE

The original measure identified a rate of return home from home at 30 days of 50.9% (n=22,837). The rate was 52.3% using the new measure of return home from home within 30 days (n=49,036).

The new measure is unable to discriminate against patients who were discharged home from hospital but subsequently died or who may no longer be home at 30 days after their index admission (e.g. readmitted to hospital). However, more patients were eligible for inclusion in these analyses compared with the original measure. Furthermore, the new measure complements the outcome measure of mortality within 30 days. Therefore, the new measure of return home from home within 30 days was used in the analyses for the 2014 annual report.

REFERENCES

1. Healthcare Quality Improvement Partnership. Falls and Fragility Fracture Audit Programme. National Hip Fracture Database. National Report 2013. London: Royal College of Physicians; 2013.
2. Neuburger J, Cromwell D. Estimating case-ascertainment and length of acute/post-acute hospital stay for patients with a hip fracture. London: The Royal College of Surgeons of England; 2012.
3. The National Hip Fracture Database. The National Hip Fracture Database National Report 2012. London: British Geriatrics Society; 2012.
4. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;327(8476):307-10
5. Quantics. *Quantics*. Available from http://www.quantics.co.uk/ [Accessed on 04 August 2014].
6. Spiegelhalter DJ. Handling over-dispersion of performance indicators. *Qual Saf Health Care*. 2005;14(5):347-51

APPENDICES

Appendix 1 Performance of regression models for mortality within 30 days, 2012/13

Model	Variables	n	Akaike Information Criterion (AIC)	Area under ROC curve (C-statistic)	Hosmer-Lemeshow test		
					x^2	df	p-value
NHFD 1	Age, ASA grade, sex, accompanied to walk outdoors	58380	28971	0.717	49.4	7	<0.001
NHFD 2	Age (continuous), ASA grade, sex, accompanied to walk outdoors	58380	28855	0.725	43.8	8	<0.001
CART	Nine groups	58380	28936	0.719	-		
CEU 1	Age, ASA grade, sex, mobility	59654	30841	0.725	76.9	8	<0.001
CEU 2	Age, ASA grade (grouped), sex, mobility	59654	30846	0.725	69.7	8	<0.001
CEU 3	Age, ASA grade (grouped), sex, deprivation	59654	30423	0.739	31.5	8	<0.001
CEU 4	Age, ASA grade (grouped), sex, source of admission	59654	30269	0.745	37.7	7	<0.001
CEU 5	Age, ASA grade (grouped), sex, preoperative AMT score	59654	30243	0.744	40.7	8	<0.001
CEU 6	Age, ASA grade (grouped), sex, fracture type	59654	30377	0.740	22.2	7	<0.001
CEU 7 (full model)	Age, ASA grade (grouped), sex, source of admission, mobility, preoperative AMT score, deprivation, fracture type	59654	29904	0.756	37.6	8	<0.001
CEU 8	Age, ASA grade (grouped)	59654	30580	0.729	40.4	8	<0.001
CEU 8	Age, ASA grade (grouped), source of admission	59654	30553	0.737	64.4	7	<0.001
CEU 9	Age, ASA grade (grouped), source of admission, mobility	59654	30293	0.746	53.3	8	<0.001
CEU 10	Age, ASA grade (grouped), source of admission, mobility, preoperative AMT score	59654	30253	0.747	60.1	8	<0.001
CEU 11	Age, ASA grade (grouped), sex, source of admission, mobility, preoperative AMT score	59654	29934	0.755	41.4	8	<0.001
CEU 12	Age, ASA grade (grouped), sex, source of admission, mobility, preoperative AMT score, deprivation	59654	29938	0.756	34.9	8	<0.001
CEU 13	Age, ASA grade (grouped), sex, source of admission (grouped), mobility (grouped)	59654	29979	0.755	56.4	8	<0.001
CEU 14	Age, ASA grade (grouped), sex, source of admission (grouped), mobility (grouped), fracture type	59654	29945	0.755	43.4	8	<0.001
CEU 15	Age, ASA grade (grouped), sex, source of admission (grouped), accompanied to walk outdoors, fracture type	60893	29981	0.760	77.0	8	<0.001
CEU 16	Age, ASA grade (grouped), sex, source of admission (grouped), accompanied to walk	60893	29886	0.762	55.3	8	<0.001

Model	Variables	n	Akaike Information Criterion (AIC)	Area under ROC curve (C-statistic)	Hosmer-Lemeshow test		
					x^2	df	p-value
	outdoors, preoperative AMT score, fracture type						
CEU 17	Age, ASA grade (grouped), sex, source of admission (grouped), walking ability indoors, fracture type	59654	30104	0.751	44.3	8	<0.001
CEU 18	Age, ASA grade (grouped), sex, source of admission (grouped), walking ability indoors, preoperative AMT score, fracture type	59654	30028	0.753	38.9	8	<0.001

Appendix 2 Performance of regression models for return home from home at 30 days, 2012/13

Model	Variables	n	Akaike Information Criterion (AIC)	Area under ROC curve (C-statistic)	Hosmer-Lemeshow test		
					x^2	df	p-value
NHFD 1	ASA grade, age, accompanied to walk outdoors	20364	24396	0.735	42.0	10	<0.001
NHFD 2	ASA grade, age (continuous), accompanied to walk outdoors	20364	24205	0.744	28.3	10	0.0004
CART	Seven groups	21083	25867	0.703	-		
CEU 1	ASA grade (grouped), age	21496	26224	0.719	38.5	7	<0.001
CEU 2	ASA grade (grouped), age, mobility (grouped)	21496	25308	0.749	45.1	9	<0.001
CEU 3	ASA grade (grouped), age, preoperative AMT score	21496	25574	0.743	62.9	9	<0.001
CEU 4	ASA grade (grouped), age, mobility (grouped), preoperative AMT score	21496	24866	0.762	64.6	10	<0.001
CEU 5	ASA grade (grouped), age, mobility (grouped), preoperative AMT score, fracture type	21490	24564	0.770	62.1	10	<0.001
CEU 6	ASA grade (grouped), age, mobility (grouped), preoperative AMT score, fracture type, deprivation	21490	24525	0.771	63.8	10	<0.001
CEU 7 (full model)	ASA grade (grouped), age, mobility (grouped), preoperative AMT score, fracture type, deprivation, sex	21490	24459	0.773	69.7	10	<0.001
CEU 8 (full model)	ASA grade (grouped), age (continuous), mobility (grouped), preoperative AMT score, fracture type, deprivation, sex	21490	24296	0.777	57.7	10	<0.001
CEU model 9	ASA grade (grouped), age, mobility (grouped), fracture type, sex	21490	24945	0.758	39.0	8	<0.001
CEU model 10	ASA grade (grouped), age, mobility (grouped), preoperative AMT score (grouped), fracture type, sex	21490	24507	0.772	66.7	8	<0.001
CEU model 11	ASA grade (grouped), age, walking ability indoors, fracture type, sex	21490	25338	0.748	48.8	8	<0.001