This is the author's version of a work that was submitted/accepted for publication in the following source:

Mortality Following Hip Replacement – inappropriate use of National Joint Registry (NJR) data.

SL Whitehouse, PhD   Senior Research Fellow/Statistician, Orthopaedic Research Unit, Institute of Health and Biomedical Innovation, Queensland University of Technology, The Prince Charles Hospital, Rode Road, Chermside, Queensland, 4032, Australia

BJRF Bolland, MD FRCS(Tr&Orth)   Exeter Hip Fellow, Exeter Hip Unit, Princess Elizabeth Orthopaedic Centre, Royal Devon and Exeter Hospital, Barrack Road, Exeter, Devon EX2 5DW, UK

JR Howell, FRCS(Tr&Orth)   Consultant Orthopaedic Surgeon, Exeter Hip Unit, Princess Elizabeth Orthopaedic Centre, Royal Devon and Exeter Hospital, Barrack Road, Exeter, Devon EX2 5DW, UK

RW Crawford, FRACS(Orth), DPhil(Oxon)   Professor in Orthopaedic Research, Orthopaedic Research Unit, Institute of Health and Biomedical Innovation, The Prince Charles Hospital, Rode Road, Chermside, Queensland, 4032 Australia

AJ Timperley, DPhil(Oxon), FRCS(Ed)   Consultant Orthopaedic Surgeon, Exeter Hip Unit, Princess Elizabeth Orthopaedic Centre, Royal Devon and Exeter Hospital, Barrack Road, Exeter, Devon EX2 5DW, UK

Corresponding author: *Mr A John Timperley*
Consultant Orthopaedic Surgeon, Exeter Hip Unit, Princess Elizabeth Orthopaedic Centre, Royal Devon and Exeter Hospital, Barrack Road, Exeter, Devon EX2 5DW, UK
j.timperley@mac.com

Abstract

Mortality following hip replacement is affected by a large number of confounding variables each of which must be considered to enable valid interpretation.

Relevant variables available from the 2011 NJR dataset were included in the Cox model.

Mortality rates in hip replacement patients were lower than in the age-matched population across all hip types. Age at surgery, ASA grade, diagnosis, gender, provider type, hip type and lead surgeon grade all had a significant effect on mortality.

Schemper's statistic showed only 18.98% of the variation in mortality was explained by the variables available in the NJR dataset. It is inappropriate to use NJR data to study an outcome affected by a multitude of confounding variables when these cannot be adequately accounted for in the available dataset.

Introduction

Patients undergoing Total Hip Replacement (THR) have been shown to have a lower risk of mortality than the corresponding aged matched general population.[1-5] This is due to the selection of relatively fit and healthy patients for surgery rather than any protective effect of total hip replacement.[3, 6] Mortality is exponential as age increases, and is reported by the UK Office of Statistics (ONS) as death rates per 1000 population. (Fig 1)

There have been several studies published recently where Registry data has been used to study mortality rates following hip arthroplasty.[5, 7-9] There are potential problems with using observational data if used inappropriately; Registries were not set up to study predictive factors for mortality and datasets contain very few relevant confounding factors that might affect this outcome. Statistical methods to "adjust" for few available variables are likely to produce misleading results and are often misquoted.

In the 2011 report of the NJR,[10] death rates were reported per fixation type or procedure. Death rates were reported without any adjustment for confounding factors and calculated to be highest in the cemented compared to uncemented, hybrid or resurfacing procedures. A number of established confounding factors have been shown to have an effect on such rates including age, comorbidity, diagnosis (increased with inflammatory arthropathy), surgeon, hospital, socioeconomic status, etc. They can cause an effect both on an individual or combined (synergistic or inhibitory) basis.

A recent paper by McMinn et al [7] investigating mortality and revision rates following hip arthroplasty raised some controversy with regard to causes for increased mortality following cemented arthroplasty compared with uncemented (vide infra). Whilst the authors felt that the effect of other confounding factors were worthy of investigation, they concluded that the use of cement might still be influencing mortality 8 years post-surgery. This is surprising conclusion given the disparity between the groups and it seems unlikely that the presence of cement in the hip *per se* could still be having an effect on mortality years after surgery.

Cox proportional hazards regression is a well-recognised and well utilised method of analysis for survival data which allows for the analysis and adjustment for many factors, where the endpoint of interest is both censored and observed. However, the method is limited by the information available for inclusion in the analysis and it is imperative that as many confounding factors as possible are included in the analysis and that the extent to which the variables explain the variation seen is available.

The National Joint Registry (NJR) of England and Wales was established in 2002 with the explicit aim to define, improve and maintain the quality of care of individuals receiving hip and knee replacement surgery across the NHS and the independent healthcare sector [11]. The NJR records patient demographics and surgical details with a view to achieving these goals. The Registry data is updated twice yearly with mortality information.

In addition to examining data available in the NJR dataset, we also wanted to examine the effect that social deprivation might have on mortality after THR. Social deprivation can be reliably measured using the Indices of Multiple Deprivation (IMD).[12] The Indices of Multiple Deprivation [13] are constructed by the Social Disadvantage Research Centre at the University Of Oxford, [12, 14] and are on a scale of 0 (least deprived) to 100 (most deprived). The model of multiple deprivation that underpins the IMD 2010 is based on distinct dimensions of deprivation, which can be recognised and measured separately. The new IMD has seven domains including Income Deprivation (weight of 22.5%), Employment Deprivation (22.5%), Health Deprivation and Disability (13.5%), Education, Skills and Training Deprivation (13.5%), Barriers to Housing and Services (9.3%), Living Environment Deprivation (9.3%) and Crime (9.3%). Each dimension is measured independently using the best indicators available to generate a score or domain index. These domain scores are then combined with explicit weightings to generate an Index of Multiple Deprivation that is an aggregate of the component domains.

The aim of this study was to establish whether it is possible to determine a true cause and effect relationship between the risk of mortality and data that is routinely collected by the NJR and to establish the degree to which variation in the mortality rate could be explained by each variable. We carried out two analyses:

> Firstly, we conducted an analysis of data collected from the NJR dataset used in preparation of the NJR's 8th Annual Report (2011), looking for an association between the variables collected and the risk of mortality.[10]

> Secondly, as social deprivation is also known to influence mortality rates [15] but is not routinely collected as part of the NJR dataset, a further analysis was performed which included social deprivation data derived from partial postcodes.

Materials and Methods:

<u>Data Collection</u>

Following a formal request for data from ourselves, the NJR supplied anonymised data for all 384,316 primary hip procedures reported to the NJR between April 2003 and the end of December 2010 (unlinked data from the 2011 NJR 8th annual report [10]), with updated information regarding time to death where applicable. For each patient age at surgery, gender, BMI (where completed), ASA grade, [16, 17] diagnosis, hip type (cemented, uncemented, hybrid, resurfacing), surgical approach, grade of lead surgeon and provider type (NHS hospital, NHS treatment centre, independent hospital, independent treatment centre) were available. These fields constitute the data routinely collected in the NJR Minimum Dataset (MDS) that we considered likely to influence mortality. There were 2,176 simultaneous bilateral cases that were excluded from the analysis due to duplicate coding issues, leaving 382,140 cases included in the cohort. Complexity of surgery (routine or complex primary) was not collected in the third version of the MDS which meant that 185,405 cases did not have this information. This variable was therefore included in an analysis of the subset of cases with MDS 1 and 2. There were 23,686 deaths during the study period at an average of 2.60 (range 0-7.81, SD 1.83) years following surgery.

Following initial analysis of the dataset by the statistical methods described below, further application was made to the NJR for patient postcode information to better inform the model. We used the postcode to map to the Indices of Multiple Deprivation as an indicator of social deprivation, and the statistical analysis was repeated.

Indices of Multiple Deprivation can be mapped to each patient via the individual's postcode. Due to patient confidentiality issues, only the abbreviated postcodes were available from the NJR, and these abbreviated postcodes encompassed a mean of 907.3 (SD 571.35) full postcodes. Each abbreviated postcode will therefore include a wider range of social deprivation than would be found by examination of the full postcode, which unfortunately means that the impact on the model may be diluted.

The first four digits of the postcode were linked to patient information, which was in turn linked with Lower level Super Output Area (LSOA)[18] and then Indices of Multiple Deprivation (IMD) averaged over the abbreviated area.[13]

Statistical Analysis

Following confirmation of proportionality to determine that Cox proportional hazards analysis was appropriate for this dataset, statistical analysis of the NJR data for all patients within the cohort was carried out in the following two stages:

1.   Survivorship analysis was performed by Cox proportional hazards regression, using a forward conditional method of data entry with no variables forced into the model. All of the variables available in the dataset were included in the analysis. Reference categories for categorical data are indicated in each table along with hazard ratios, 95% confidence intervals and significance level (p-value) of each variable. Cases with missing values for any variable are excluded from the analysis.

2.   An examination of the degree to which variation in the mortality rate could be explained by each of the variables. We calculated the explained variation for each variable using the method described by Schemper [19] for largely censored data. Schemper's measure of predictive accuracy for Cox regression models can be calculated using the R program, [20] but is not readily available as part of the standard output using the most commonly used statistical programs such as SPSS, Stata or SAS.

The process of analysis described above was performed not only for all patients within the cohort but was also repeated for subgroups categorized by age (all patients aged >65 years, all patients aged <50 years).

The initial data analysis was performed using SPSS for Windows version 19 (SPSS Inc, IBM, Chicago, IL). Calculation of Schemper's statistic was performed using R [20, 21] (version 2.14.2).

Results

Baseline data

There were 382,140 primary, unilateral hip replacements included in this cohort – 157,479 cemented; 129,052 uncemented; 54,439 hybrid, 26,944 resurfacing hips and 14,226 with missing coding. The mean age at surgery was 68.2 (95% CI; 68.16 to 68.23). Critically, for each hip type this was 72.7 (95% CI; 72.64 to 72.73) for cemented hips; 65.4 (95% CI; 65.36 to 65.48) for uncemented hips; 68.9 (95% CI; 68.84 to 69.02) for the hybrid group and 54.9 (95% CI; 54.78 to 55.00) for the resurfacing group. It should also be noted that there were significantly more complex cases in the cemented (17.7%) and resurfacing (15.1%) than the uncemented (4.8%) and hybrid (5.6%) groups (p<0.001).

Actual mortality rates per 1000 from the NJR dataset for the four hip types (cemented, uncemented, hybrid and resurfacing) are shown in Figure 2, compared with the age matched figures from the UK Office of National Statistics death registrations, 2010 in England and Wales.[6] It can be seen that for each class of implant (cemented, uncemented, hybrid and resurfacing) the mortality rate in patients undergoing surgery (shown by dotted lines) are lower than in the age-matched population (solid lines).

Cox Regression Analysis

For the initial analysis, covariates were entered into the model in order of significance as: age at surgery, ASA grade, primary diagnosis, gender, provider type, hip type, lead surgeon grade. The surgical approach was not significant in the model and BMI was excluded due to the large number of missing responses.

Covariates in the model and their associated hazard ratios (and 95% confidence intervals) are shown in Table 1. Further analysis for those over 65s and under 50s at the time of surgery are shown in Tables 2 and 3. When complexity of surgery was included, this variable was also significant in the model and influence of the other variables was unaffected. The hazard ratio of complex compared to routine surgery was 1.06 (95% CI 1.015 to 1.105), p=0.007.

The results show that, for the cohort as a whole (Table 1), a statistically significant association exists for all variables, except surgical approach, on the outcome of mortality. As might be expected, increasing age, male sex and higher values of the ASA grade were all associated with a higher risk of mortality. Compared to patients with a diagnosis of osteoarthritis, all patients from other diagnosis groups, except for DDH, were associated with higher rates of mortality. The risk of mortality was lower in patients who had their primary operation in an independent private hospital or treatment centre and it was lower for those who had uncemented, hybrid or resurfacing procedure than those who had a cemented hip replacement. Finally, patients whose operation was performed by some grades of non-consultant surgeon had an association with lower mortality.

The pattern of results seen in the patients over 65 years of age (Table 2) was very similar to those of the overall cohort. In patients under 50 (Table 3) there were minor differences in the results of some of the variables, with the exception of Surgeon Grade. For these patients, there was no longer a statistically significant result seen in patients whose hip replacement was performed by a junior surgeon, but the number of cases in this group was small.

Schemper's Statistic

Schemper's statistic for the full dataset showed that collectively, the variables in the statistical model explained only 18.98% (20.27% in the subset with complexity of surgery) of the variation in mortality, with age accounting for 15.9% of this. Using a univariate analysis, the other variables available in the dataset explained variation in mortality as follows: ASA grade 3.3%, complexity of surgery 2.5%, provider type 1.4%, diagnosis 0.9%, hip type 0.7%, and surgeon grade 0.6%. Gender explained only a tiny amount of the variation. The sum of these individual figures does not equal the overall Schemper's statistic due to the differing numbers of valid cases for each of these variables.

For the under 50s, the variables in the statistical model explained only 4.67% of the variation in mortality whereas for the over 65s, this figure was 13.85%.

Deprivation Data

In an attempt to more accurately model the data, additional postcode data was obtained from the NJR, so that we could ascribe Indices of Multiple Deprivation to each patient. There were an average of 907.3 (95% CI: 884.03 to 930.60) full postcodes in each abbreviated postcode (Figure 3). With the addition of the postcode (and hence IMD data), the model changed accordingly (Table 4) and there was a significant association found between IMD data and the outcome of mortality. Interestingly, the order of model entry was also affected by the inclusion of this variable (age at surgery, ASA grade, primary diagnosis, gender, provider type, surgical approach, hip type, IMD, lead surgeon grade), and Surgical Approach became significant in the model, meaning that all variables are significant in this model.

Further calculation of Schemper's statistic for the whole dataset after the addition of IMD data indicated that addition of this variable actually decreased the overall amount of variation explained slightly to 17.98%. This is most likely due to the number of cases available for the model, as the valid number drops from 367,523 for the baseline model to 337,062 – a drop of 30,461 (8.3%) cases. IMD explained 2.4% of the variation when examined alone.

Discussion

The dataset collected by the NJR includes many of the variables most likely to affect joint replacement revision rates, including details of the implant, surgeon and institution at which the procedure was performed. In addition, contributory patient factors are collected including age at operation and the patients' gender. An attempt is made to capture patients' BMI, another variable

that may influence implant survival, though at the present time this data is incomplete. Patient activity level may also affect survivorship and this is not captured by the NJR.

Many joint registries have now started to publish mortality data, linked to some of the variables collected in their datasets - for example the method of implant fixation. The aim of this paper was to examine whether valid conclusions about mortality can be drawn from the data collected by the NJR.

The overall death rate within 30-90 days after total hip replacement is very low and ranges from 0%-0.45%. [1, 2, 5, 22-26] There has been a decline in these rates over the last 15 years and several studies have shown that patients who have undergone THR survive longer than do a matched control population.[1, 3-5, 22] It should be noted that mortality rates for ALL types of hip replacement are significantly lower than the national age- and gender-matched averages [3, 5, 6]. The National Joint Registry [10] has confirmed that the mortality rate following THR is significantly less than that of the general population of England and Wales (1.9%; 95% CI 1.8 to 2.0 in the 2009 report) [4, 22] using age-and gender-adjusted standardised mortality ratios. They also report that although these rates are lower than the general age- and gender-matched population, "the highest death rates were among the cemented group and the lowest were among the resurfacing group, reflecting the age distribution of these groups." This is unsurprising as the resurfacing group are on average 18 years younger than the cemented group at surgery.

The primary outcome for this analysis was the time to death. The data for patients revised during the period of the study was censored at the time of revision and these patients were treated as survivors up to that point. This was justified as revision is not a competing risk for death (unlike the reverse) and the mortality profile for revision cases is significantly different from that of primaries.

Interpretation of Results

1. Cox Regression Analysis data

The results of our regression analysis have shown that several of the variables that make up the NJR minimum dataset have a statistically significant association with mortality. Some of these results have been more surprising than others. It is, for example, perhaps to be expected that there should be an association between age and the risk of mortality, as demonstrated in our analysis. Data from the Office of National Statistics shows that mortality rates within the general population increase exponentially from the age of 75 onwards, as seen in Figure 1. In our analysis the patients' age was the most significant of the variables with respect to mortality and it was the first to be entered into

our forward conditional method of Cox regression analysis and explained the majority of the variation.

It is important to understand that although regression analysis will partially correct any of the variables with respect to their effect on mortality, the effect of such a variable on other confounding factors will not be corrected for and this is particularly important for patient age. Correction of age against mortality will make no allowances for the effect of time on other confounding factors, such as comorbidities and deprivation, be that income, health, housing, living environment or crime. This has important consequences when comparing two or more groups with respect to their mortality where the groups differ widely in respect of their age. Although their respective mortality can be corrected for the age of the groups, the effect of their different ages on other confounding variables, which may in turn have a significant effect on mortality, is not corrected for and bias is likely to persist in the analysis. It is clear that chronological age does not always correlate with physiological age and there is a large variation in life expectancy between individuals born in the same year dependent on an almost innumerable number of confounding variables.

The second most significant variable in our model was the American Society of Anaesthesiologists (ASA) grade of the patient and again this is perhaps not a surprising result. Adopted in its present form in 1963, the ASA physical status classification system is a method of pre-operative assessment of a patient's fitness to undergo an anaesthetic. [17, 27] The ASA clearly states that it does not endorse any elaboration of the simple definitions and it is well established that different anaesthetists assign different grades to the same patient [16, 28]. The ASA cannot be used to assess fully a patient's comorbid status and alternative comorbidity scores have been developed for this purpose that include the Charlson, [29] Acute Physiology and Chronic Health Evaluation (APACHE) II [30] and POSSUM. [31] From all of these, only POSSUM has been validated in orthopaedic patients but it relies on physiological data not collected by the NJR [31]. Wainwright at al [32] showed that comorbidities measured using a simple scoring system, such as the ASA grade, cannot be relied upon to differentiate between patients likely to survive and those who will not. Although ASA grade has been found to be a useful predictor of inpatient 30 day and 6 month surgical mortality, [33-35] it does not predict surgical mortality rates at 12 to 24 months [36, 37] or beyond. However, despite the limitations in its use, it is clear from our results that patients with higher ASA scores do have a higher risk of mortality than ASA grade 1 patients. Therefore it appears that the ASA grade is at least partially predictive of mortality and it seems likely that this is through some separation of patient groups with different comorbidities.

Our analysis has also produced some association between mortality and variables that may not be as obviously linked to mortality. Compared to osteoarthritis, all other diagnoses except DDH were associated with a higher mortality. This result is most likely to reflect differences between the diagnostic groups with respect to their comorbidity profiles. For instance, many of the inflammatory arthropathies are systemic diseases, affecting multiple organ systems, which may in turn adversely affect mortality.

The effect of differences in economic and social confounding variables are perhaps seen most obviously in the results that were obtained for hospital provider type, which show lower mortality for patients who received their hip replacement in an independent private hospital. While this may be due to a true difference in safety standards within different institutions, this result is far more likely to reflect the selection of healthier, fitter patients of higher socio-economic status in the private institutions, and therefore it is possibly a surrogate measure of comorbidity. Such a mechanism may also explain why operations performed by junior surgical staff are associated with a lower mortality than those performed by senior staff. It seems unlikely that this result reflects a true causative link between more junior status and better mortality. Instead, we feel it is more likely that patients selected for training are more likely to be the fitter patients, with fewer comorbidities, whilst the sicker patients are more likely to be operated on by more senior members of staff.

Of all the variables that were included in this analysis, the hip type and its relationship with mortality has caused the most controversy in the recent orthopaedic literature. McMinn et al [7] studied the relationship between implant fixation and two outcomes; mortality and revision. They found that cemented hip replacements were associated with a mortality that was initially similar to other implant types, but differences between the groups developed in the long term. In a study of 30 day mortality in over 30 000 hip replacements, Parvizi et al [38] found no statistically significant difference between cemented and uncemented implants. In addition, Lie et al [39] found that following over 180 000 lower limb arthroplasties there was an initial small increase in mortality in the surgical group, compared to the background population, but this effect was reversed within 27 days after the operation. It is therefore surprising that McMinn et al appeared to ascribe a causative relationship between prosthesis fixation and the late emergence of increased mortality.

In common with other authors, we have found an association between higher rates of mortality and cemented stem fixation, when compared to hybrid, uncemented or resurfacing designs. This includes a recent paper by Kendal et al looking at mortality alone where data was adjusted using propensity scores, a method which the authors suggest should adequately adjust for confounders.[40] The unresolved question remains whether the relationship between implant type

and mortality is causative or simply an association that reflects the influence of other, unaccounted for, confounding variables. Causation was refuted by Blom et al when they linked from the NJR with Hospital Episode Statistics (HES) data for a similar time period.[5] They were more thoroughly able to adjust for confounding variables, concluding that fixation was not a significant factor for 90 day mortality. Mortality within this time period can much more confidently be ascribed to the surgery. The most obvious source of bias in our study is the wide difference in mean age of patients within each of the hip type groups. When tracking mortality rates over time it is important to establish the average age for each group at time point zero within the study. Figure 2 demonstrates that the mortality rate for the resurfacing group 20 years after their operation (mean age at operation 54.9) is the same as the cemented group (mean age at surgery 72.7) at the time of surgery, and still before the exponential rise in mortality rate occurs. Of fundamental importance is the near 20 years difference between the two groups at the point of surgery and the observation that the exponential increase in mortality which occurs at around age 75 cannot influence the mortality rate for the resurfacing group during the first 20 years of their follow up. This means that by 10 years post surgery, of 1000 patients in the age bracket for resurfacing (average age 54.9 at surgery), approximately 63 will have died according to the ONS figures, compared with 378 in the cemented age bracket (average age of 72.7 at surgery) - a 6-fold difference that is attributable to their respective ages at the time of their operation.

In addition to the effect of age on the mortality of the different prosthesis groups, the results may also reflect the action of a number of other confounding factors. It could be that there is a difference between the groups in their patient mix that is not fully explained by age and ASA grade, as proposed by Breusch et al.[37] Correction for age and ASA in the analysis of mortality will not eradicate sources of bias from other confounding variables, as discussed above. Such confounding factors may include economic and social factors, as well as geographical or regional differences in the use of different designs of hip replacement, the effect of which would be to fix the variable of implant fixation to the underlying mortality of the host population. Additionally, surgeons may make decisions regarding implant fixation based on patient factors such as health and life expectancy, further confusing the relationship between fixation and mortality.

We have performed Cox regression analysis on the whole dataset, and then repeated that analysis for two distinct patient age groups; those over 65 and patients under 50. There were minor differences seen in the results of the two age-specific analyses, but the underlying trends identified and discussed above were generally repeated.

2. Deprivation data

Previous studies have identified numerous factors that have a significant effect on mortality, many of which are not routinely collated as part of the NJR. Examples include socio-economic status,[15, 41] details of income,[42] race,[42, 43] smoking,[44] obesity [45, 46] and medical comorbidities.[42, 43, 45, 47]

In an attempt to study the effect of social deprivation, we examined the association between a patient's postcode and their mortality. The postcode has been shown to map patients to Indices of Multiple Deprivation and can therefore be used to link mortality and deprivation. Unfortunately we were only able to utilise the abbreviated postcode information due to patient confidentiality and this is a distinct weakness of this study. This is unfortunate, as there are on average more than 900 full postcodes per abbreviated postcode, which imparts a significant dilution of information for the model. However, the variable still significantly influenced the mortality model. The inclusion of the diluted IMD data in the model changed the order the variable entered the model and made the surgical approach variable significant. We believe that this is an indication of how delicately balanced and multifactorial mortality is and further highlights the need to include as much explanatory information as possible in any statistical model – a task which unfortunately the NJR was not designed for. It also highlights that extreme care must be taken when interpreting these results and highlights the weakness in using large datasets outside their remit and specific purpose of use.

3. Schemper's statistic

Most linear regression analyses are accompanied by an assessment of the proportion of explained variation in the outcome that is attributable to the variable under examination. This statistic is given the symbol R^2, the value of which varies between 0 and 1. An R^2 value that is close to 0 indicates that little of the variation in the outcome can be explained by the tested variable and other confounding variables may be at work. Conversely, an R^2 value close to 1 indicates that much of the variation of the outcome can be attributed to the variable in question and is often expressed as a percentage of variation explained.

It is not possible to make a standard assessment of the proportion of explained variation for censored data such as the survivorship results produced by the NJR. Instead we have used the method described by Schemper to produce a measure of the predictive accuracy for our Cox regression model. Schemper's statistic is also given a value of between 0 and 1, and the value is interpreted along similar lines to the R^2.

We have found that, for the dataset as a whole, and considering the action of all the significant variables together, Schemper's statistic was low, at 0.1898. This means that only approximately 18.98% of the variation in mortality can be explained by a combination of the patients' age, sex, ASA grade and diagnosis as well as the hospital provider, the implant type, complexity of surgery and the surgeon grade. Each of these variables produced highly significant results on Cox regression analysis and these results may seem to be at odds with Schemper's statistic. However, this probably reflects the difference between a statistically significant result and one that is clinically significant and reporting of this statistic is vital to avoid the common problem of over interpretation.[7] This is a similar figure to the 15% variation explained reported by Middleton et al [9] in their mortality study of hip fractures. With the large volume of patient data entered onto the NJR, any association is likely to be highly statistically significant, even if the effect is relatively small. Of all the variables, the patients' age had the highest predictive accuracy for mortality, accounting for about 16% of the total 19% exerted by all the variables together. This result would seem to be consistent with data from the Office of National Statistics on the link between age and mortality rate for the general population.

Schemper's statistic suggests that there exists a significant effect on mortality created by confounding variables that are not part of the standard NJR dataset. We suspected that these confounding variables may include the effects of social deprivation and so we repeated the analysis with Schemper's method after the addition of the patients' abbreviated postcode data. Due to the reduction in numbers available to the analysis, the effect was to slightly decrease the value of Schemper's statistic, to 0.1798. This may indicate that social deprivation is not strongly linked to mortality but it is more likely that it reflects the heterogeneity of social deprivation encompassed by each abbreviated postcode. We think it is likely that if the full postcodes could have been used we would have seen a larger effect on Schemper's statistic exerted by the social deprivation data.

Implications of the study for interpretation of NJR results

All confounding variables need to be accounted for to enable a valid interpretation of mortality data across different groups / populations. Registry data provides information for an enormous number of patients but with limited data per episode, making it difficult to fully correct for confounding bias. Any factors not collated and adjusted for can lead to heterogeneity between groups and subsequent statistical differences on testing. The dataset collected by the NJR was not designed to explore factors contributing to mortality post surgery. If this has been an explicit aim of the Registry then collection of a different and much larger set of variables would have been necessary.

When we consider the risk of mortality in the long term we must acknowledge that revision procedures have a higher mortality rate than primary joint replacements [42, 48]. When establishing the mortality risk to a patient, therefore, we must take into account not only the index procedure but also the potential risk from subsequent revision procedures. Higher revision rates [10, 49] and reduced postoperative mobility and/or function [50, 51] for some procedures such as resurfacing and uncemented hips have been demonstrated across many national registries. Wainwright et al [32] reported that age at surgery greater than 62 years predicts that death is more likely than requiring a revision – distinctly splitting the resurfacing and cemented groups. In the NJR dataset, only one terminal event per case is recorded – revision, death or survival, which means that revision cases are effectively excluded from the mortality analysis at the time of revision. This is a potential source of bias and one that benefits the apparent mortality data for implants with higher rates of revision.

The paper published by McMinn using NJR data [7] has generated much discussion regarding case mix, [37] interpretation of Registry data [52] and more importantly access to Registry data [53] within the orthopaedic community. Macgregor [53] clarified that information on mortality available to McMinn in the dataset provided was incomplete, with no times to death given for patients who underwent revision procedures This dataset was released as part of supplier feedback to Smith and Nephew for study of implant survivorship and was not intended to be used for, nor configured for, research into mortality. The data was not linked and could not be linked to Hospital Episode Statistics (HES) (as was stated). For this reason and from the results of the current study that demonstrates the effect of confounding variables on mortality, it must be concluded that McMinn's analysis of the data was fundamentally flawed and the interpretation is, therefore, invalid.

The dangers of using NJR data for purposes that it was not designed for are proven by the results of this study, which highlights the potential that exists for the misinterpretation of results. There is political demand for transparency of data – the NHS Commissioning Board has committed to the publication of named surgeon-level data by June 30th 2013. The Profession supports transparency with appropriately interpreted, validated data. We have shown here that publication of surgeons' individual mortality rates with the data currently available would be misleading and of no benefit to healthcare commissioners, providers of healthcare and most importantly, patients.

Acknowledgements

References

1. Nunley RM, Lachiewicz PF. Mortality after total hip and knee arthroplasty in a medium-volume university practice. J Arthroplasty 2003; 18 (3): 278

2. Sharrock NE, Cazan MG, Hargett MJ, Williams-Russo P, Wilson PD, Jr. Changes in mortality after total hip and knee arthroplasty over a ten-year period. AnesthAnalg 1995; 80 (2): 242

3. Barrett J, Losina E, Baron JA, Mahomed NN, Wright J, Katz JN. Survival following total hip replacement. J Bone Joint Surg [Am] 2005; 87 (9): 1965

4. AOA. Australian Orthopaedic Association National Joint Replacement Registry Annual Report. Adelaide. 2009

5. Hunt LP, Ben-Shlomo Y, Clark EM, Dieppe P, Judge A, MacGregor AJ, Tobias JH, Vernon K, Blom AW, National Joint Registry for England W, Northern I. 90-day mortality after 409,096 total hip replacements for osteoarthritis, from the National Joint Registry for England and Wales: a retrospective analysis. Lancet 2013; 382 (9898): 1097

6. UK Office for National Statistics. Death Registrations Summary Tables, England and Wales. 2010 (accessed: 10/02/12) [Source: Office for National Statistics licensed under the Open Government Licence v.1.0]

7. McMinn DJ, Snell KI, Daniel J, Treacy RB, Pynsent PB, Riley RD. Mortality and implant revision rates of hip arthroplasty in patients with osteoarthritis: registry based cohort study. BMJ 2012; 344: e3319

8. Costain DJ, Whitehouse SL, Pratt NL, Graves SE, Ryan P, Crawford RW. Perioperative mortality after hemiarthroplasty related to fixation method. Acta Orthop 2011; 82 (3): 275

9. Middleton RG, Smith R, Young PS, Uzoigwe CE, Holt G. Hemiarthroplasty for hip fracture - does cement increase mortality? In: British Hip Society. Bristol, UK. 2013

10. The NJR Centre. National Joint Registry for England and Wales. 8th Annual Report. The NJR Centre, Hemel Hempstead. 2011

11. Horan F. The National Joint Registry for England and Wales. J Bone Joint Surg [Br] 2003; 85 (4): 618

12. Payne RA, Abel GA. UK indices of multiple deprivation - a way to make comparisons across constituent countries easier. Health Stat Q 2012; (53): 22

13. courtesy of the Department for Communities and Local Government, Indices of Deprivation 2010, http://www.nationalarchives.gov.uk/doc/open-government-licence/

14. Social Disadvantage Research Centre. [accessed 10/02/12]. Available from http://www.ons.gov.uk/ons/rel/regional-trends/atlas-of-deprivation--england/2010/atlas-of-deprivation-2010.html

15. White C, Glickman M, Johnson B, Corbin T. Social inequalities in adult male mortality by the National Statistics Socio-Economic Classification, England and Wales, 2001-03. Health Stat Q 2007; (36): 6

16. Little JP. Consistency of ASA grading. Anaesthesia 1995; 50 (7): 658

17. American Society of Anesthesiologists Inc. New classification of physical status. Anesthesiology 1963; 24 (1): 111

18. courtesy of the Office of National Statistics,

19. Schemper M, Henderson R. Predictive accuracy and explained variation in Cox regression. Biometrics 2000; 56 (1): 249

20. R Development Core Team. R: A language and environment for statistical computing. In. Vienna, Austria: R Foundation for Statistical Computing. 2012

21. Lusa L, Miceli R, Mariani L. Estimation of predictive accuracy in survival analysis using R and S-PLUS. Comput Methods Programs Biomed 2007; 87 (2): 132

22. Timperley AJ, Whitehouse SL. Mitigating surgical risk in patients undergoing hip arthroplasty for fractures of the proximal femur. J Bone Joint Surg [Br] 2009; 91 (7): 851

23. Sierra RJ, Timperley JA, Gie GA. Contemporary Cementing Technique and Mortality During and After Exeter Total Hip Arthroplasty. J Arthroplasty 2008:

24. Dearborn JT, Harris WH. Postoperative mortality after total hip arthroplasty. An analysis of deaths after two thousand seven hundred and thirty-six procedures. J Bone Joint Surg [Am] 1998; 80 (9): 1291

25. Parvizi J, Ereth MH, Lewallen DG. Thirty-day mortality following hip arthroplasty for acute fracture. J Bone Joint Surg [Am] 2004; 86-A (9): 1983

26. Parry M, Wylde V, Blom AW. Ninety-day mortality after elective total hip replacement: 1549 patients using aspirin as a thromboprophylactic agent. J Bone Joint Surg [Br] 2008; 90 (3): 306

27. Saklad M. Grading of patients for surgical procedures. Anesthesiology 1941; 2: 281

28. Haynes SR, Lawler PG. An assessment of the consistency of ASA physical status classification allocation. Anaesthesia 1995; 50 (3): 195

29. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987; 40 (5): 373

30. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med 1985; 13 (10): 818

31. Mohamed K, Copeland GP, Boot DA, Casserley HC, Shackleford IM, Sherry PG, Stewart GJ. An assessment of the POSSUM system in orthopaedic surgery. J Bone Joint Surg [Br] 2002; 84 (5): 735

32. Wainwright C, Theis JC, Garneti N, Melloh M. Age at hip or knee joint replacement surgery predicts likelihood of revision surgery. J Bone Joint Surg [Br] 2011; 93 (10): 1411

33. Glance LG, Lustik SJ, Hannan EL, Osler TM, Mukamel DB, Qian F, Dick AW. The Surgical Mortality Probability Model: derivation and validation of a simple risk prediction rule for noncardiac surgery. Ann Surg 2012; 255 (4): 696

34. Story DA, Fink M, Leslie K, Myles PS, Yap SJ, Beavis V, Kerridge RK, McNicol PL. Perioperative mortality risk score using pre- and postoperative risk factors in older patients. Anaesth Intensive Care 2009; 37 (3): 392

35. Hooper GJ, Rothwell AG, Hooper NM, Frampton C. The relationship between the American Society Of Anesthesiologists physical rating and outcome following total hip and knee arthroplasty: an analysis of theNew Zealand Joint Registry. J Bone Joint Surg [Am] 2012; 94 (12): 1065

36. Awad S, Herrod PJ, Palmer R, Carty HM, Abercrombie JF, Brooks A, de Beer T, Mole J, Lobo DN. One- and two-year outcomes and predictors of mortality following emergency laparotomy: a consecutive series from a United kingdom teaching hospital. World J Surg 2012; 36 (9): 2060

37. Breusch SJ, Howie CR. Excess mortality can be explained by case mix selection. BMJ 2012; 345: e7137

38. Parvizi J, Johnson BG, Rowland C, Ereth MH, Lewallen DG. Thirty-day mortality after elective total hip arthroplasty. J Bone Joint Surg [Am] 2001; 83-A (10): 1524

39. Lie SA, Pratt N, Ryan P, Engesaeter LB, Havelin LI, Furnes O, Graves S. Duration of the increase in early postoperative mortality after elective hip and knee replacement. J Bone Joint Surg [Am] 2010; 92 (1): 58

40. Kendal AR, Prieto-Alhambra D, Arden NK, Carr A, Judge A. Mortality rates at 10 years after metal-on-metal hip resurfacing compared with total hip replacement in England: retrospective cohort analysis of hospital episode statistics. BMJ 2013; 347: f6549

41. Clement ND, Muzammil A, Macdonald D, Howie CR, Biant LC. Socioeconomic status affects the early outcome of total hip replacement. J Bone Joint Surg [Br] 2011; 93 (4): 464

42. Mahomed NN, Barrett JA, Katz JN, Phillips CB, Losina E, Lew RA, Guadagnoli E, Harris WH, Poss R, Baron JA. Rates and outcomes of primary and revision total hip replacement in the United States medicare population. J Bone Joint Surg [Am] 2003; 85-A (1): 27

43. Memtsoudis SG, Pumberger M, Ma Y, Chiu YL, Fritsch G, Gerner P, Poultsides L, Valle AG. Epidemiology and risk factors for perioperative mortality after total hip and knee arthroplasty. J Orthop Res 2012:

44. Singh JA. Smoking and outcomes after knee and hip arthroplasty: a systematic review. J Rheumatol 2011; 38 (9): 1824

45. Bozic KJ, Lau E, Kurtz S, Ong K, Rubash H, Vail TP, Berry DJ. Patient-related risk factors for periprosthetic joint infection and postoperative mortality following total hip arthroplasty in Medicare patients. J Bone Joint Surg [Am] 2012; 94 (9): 794

46. Huddleston JI, Wang Y, Uquillas C, Herndon JH, Maloney WJ. Age and obesity are risk factors for adverse events after total hip arthroplasty. Clin Orthop Relat Res 2012; 470 (2): 490

47. Aynardi M, Jacovides CL, Huang R, Mortazavi SM, Parvizi J. Risk factors for early mortality following modern total hip arthroplasty. J Arthroplasty 2013; 28 (3): 517

48. Zhan C, Kaczmarek R, Loyo-Berrios N, Sangl J, Bright RA. Incidence and short-term outcomes of primary and revision hip replacement in the United States. J Bone Joint Surg [Am] 2007; 89 (3): 526

49. Morshed S, Bozic KJ, Ries MD, Malchau H, Colford JM, Jr. Comparison of cemented and uncemented fixation in total hip replacement: a meta-analysis. Acta Orthop 2007; 78 (3): 315

50. Parker MJ, Gurusamy K. Arthroplasties (with and without bone cement) for proximal femoral fractures in adults. Cochrane Database Syst Rev 2004; (2): CD001706

51. Weidenhielm LR, Mikhail WE, Nelissen RG, Bauer TW. Cemented collarless (Exeter-CPT) versus cementless collarless (PCA) femoral components. A 2- to 14-year follow-up evaluation. J Arthroplasty 1995; 10 (5): 592

52. Jameson SS, Baker P, Deehan D, Reed M, Mason J. Paper's conclusions will only cause confusion. BMJ 2012; 345: e7005

53. MacGregor AJ. Clarification from the National Joint Registry. BMJ 2012; 345: e7199

Figure 1: Mortality rate per 1000 population for 5 year age groups from UK Office of National Statistics (ONS).
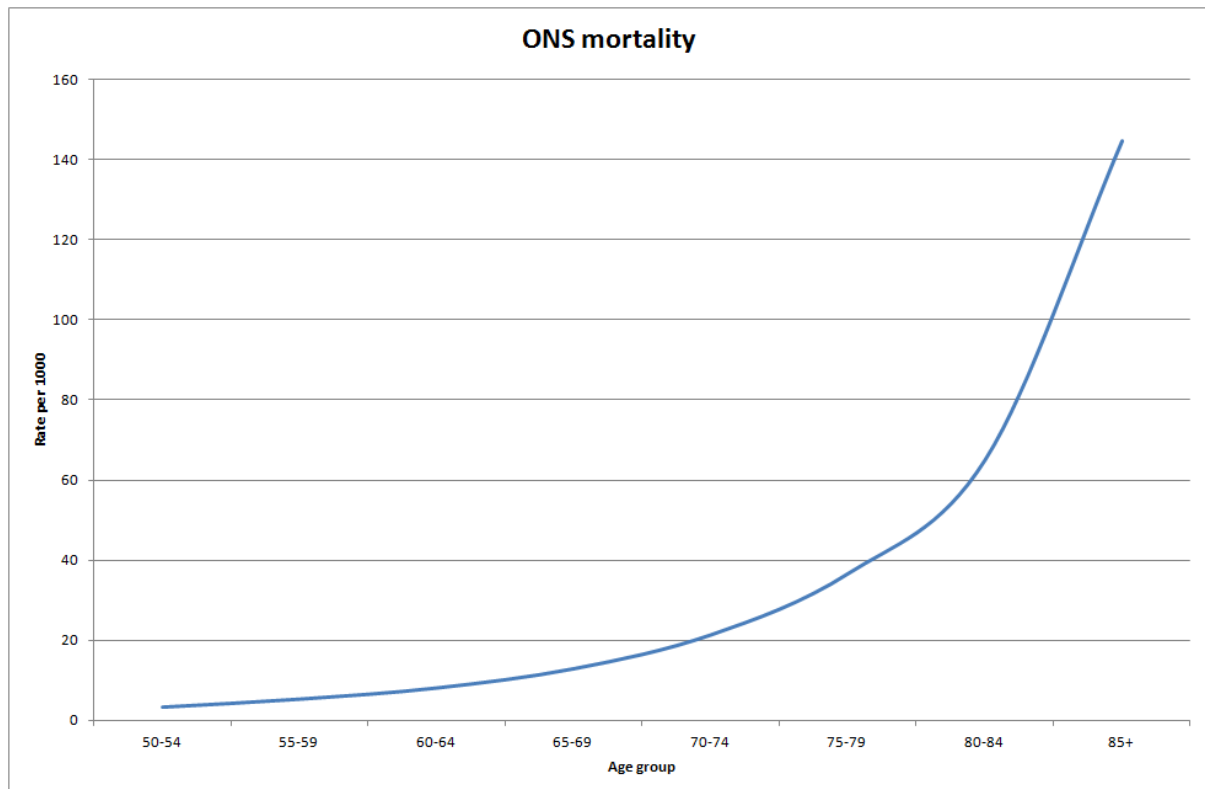
Figure 2: ONS death rate per 1000 in age bands according to mean age at surgery for each hip type – year 0 being the year of operation and representing the mortality rate for the age band of each hip type. Actual death rates from the NJR dataset are also shown as dashed lines.
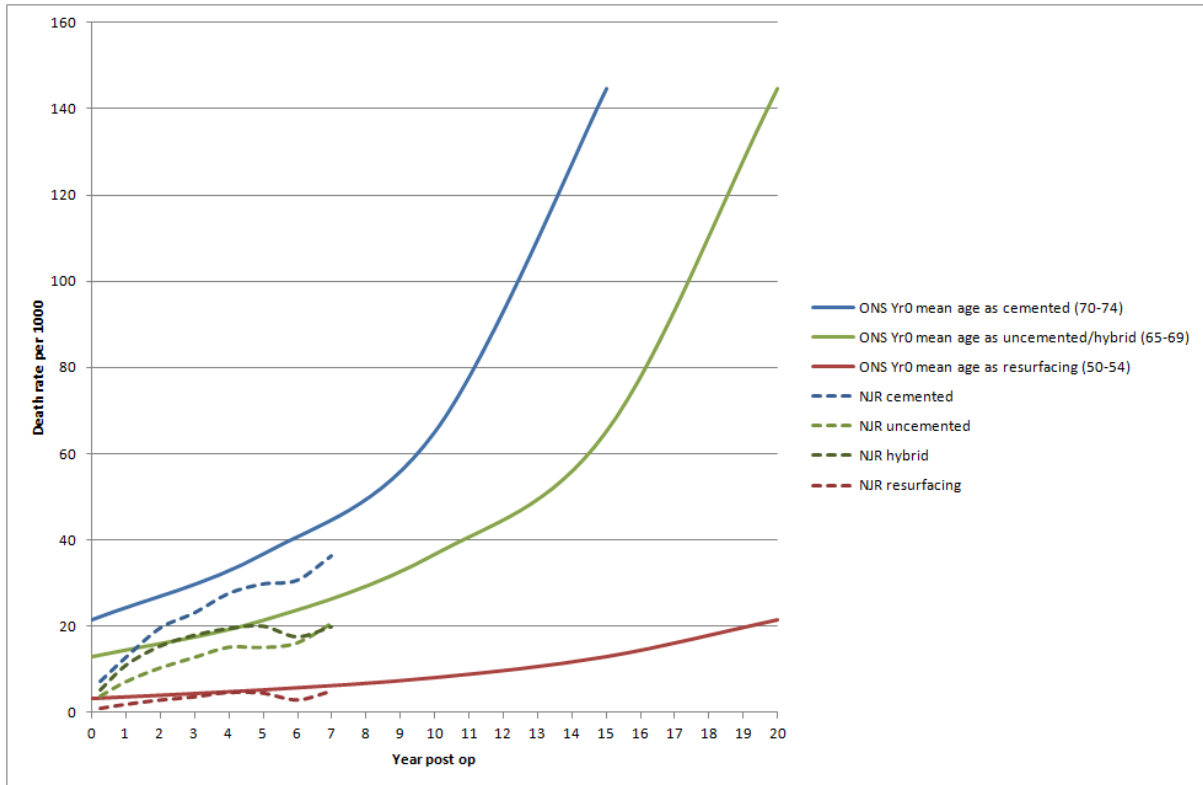
Figure 3: Histogram of number of full postcodes per abbreviated postcode.
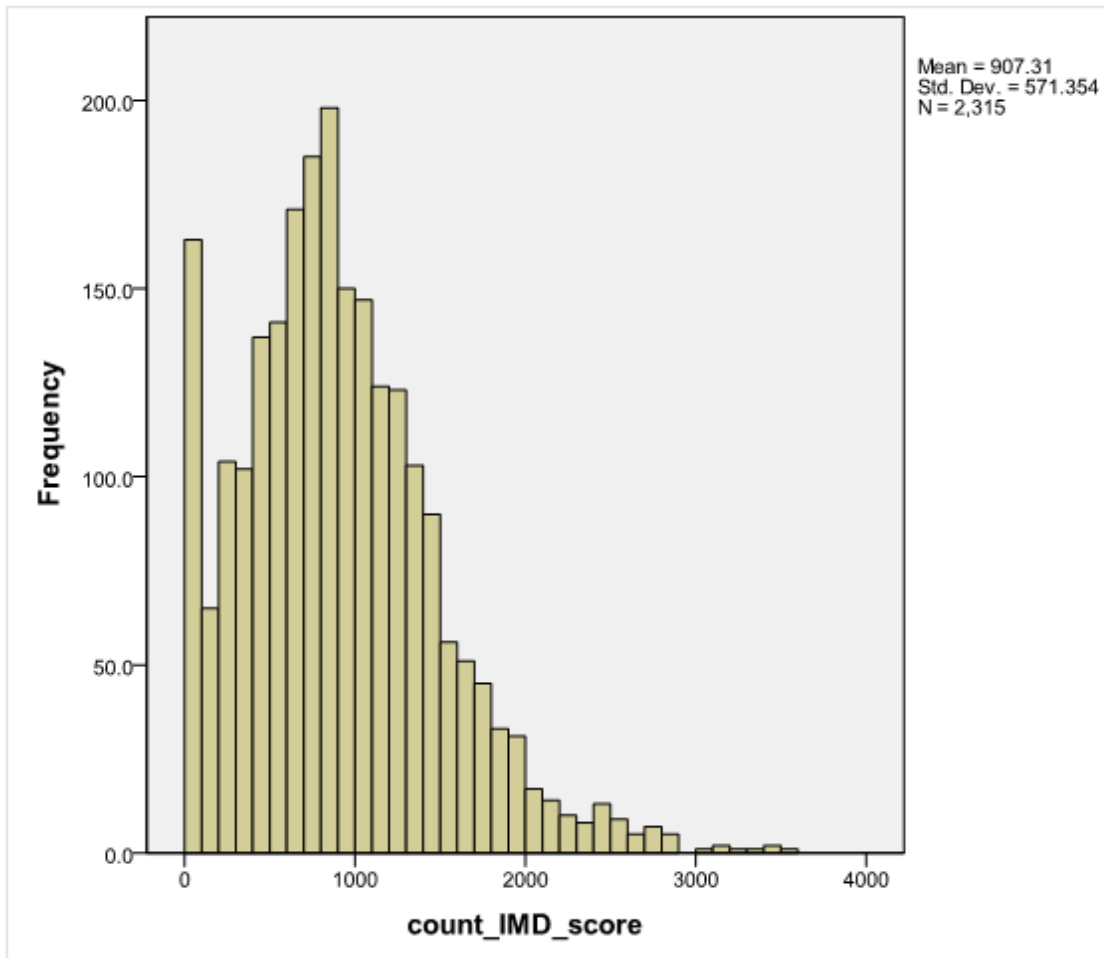
Table 1: Variables in the initial model and associated hazard ratios (95% confidence intervals) (valid n=367,523, missing 14,617)

Variable (reference variable)	n	Hazard ratio	95% CI	p-value
Age at primary (per additional year)	381,922	1.08	1.077 to 1.081	p<0.001
ASA grade (1)	74,418	*reference*		
2	242,186	1.27	1.212 to 1.324	p<0.001
3	48,872	2.37	2.259 to 2.491	p<0.001
4	1931	4.47	4.056 to 4.916	p<0.001
5	116	2.88	1.832 to 4.521	p<0.001
Diagnosis (OA)	331,611	*reference*		
Other arthritides	5070	1.71	1.544 to 1.894	p<0.001
Previous trauma	2607	1.79	1.518 to 2.100	p<0.001
Acute trauma	4815	2.64	2.441 to 2.852	p<0.001
Previous hip surgery	2062	2.14	1.928 to 2.381	p<0.001
DDH	5727	0.99	0.808 to 1.207	p=0.904 (ns)
AVN	8283	1.55	1.438 to 1.669	p<0.001
Other (inc infection)	7348	2.38	2.222 to 2.547	p<0.001
Gender (Female)	219,074	*reference*		
Male	148,449	1.49	1.450 to 1.532	p<0.001
Provider type (NHS)	241,544	*reference*		
NHS treatment centre	13,843	0.94	0.868 to 1.007	p=0.078 (ns)
Independent hospital	94,521	0.73	0.702 to 0.754	p<0.001
Independent TC	17,615	0.65	0.598 to 0.714	p<0.001
Hip type (Cemented)	157,257	*reference*		
Uncemented	128,960	0.89	0.858 to 0.919	p<0.001
Hybrid	54,391	0.94	0.904 to 0.979	p=0.003
Resurfacing	26,915	0.55	0.493 to 0.617	p<0.001
Lead surg grade (Consultant)	300,364	*reference*		
SpR/ST3-8	33,051	0.97	0.924 to 1.009	p=0.122 (ns)
F1-ST2	375	0.85	0.608 to 1.193	p=0.351 (ns)
Speciality Dr	21,019	0.95	0.903 to 0.993	p=0.024
Other	12,714	0.88	0.809 to 0.962	p=0.004
(ns) is not significant. Approach is not significant in this model, p=0.520				
Approach (Hardinge/lateral/ant-lat)	154,721			
Posterior	177,555			p=0.456 (ns)
Anterior	1464			p=0.205(ns)
Trochanteric osteotomy	1499			p=0.348 (ns)
Other	32,284			p=0.885 (ns)

Table 2: Variables in the model and associated hazard ratios (95% confidence intervals) for over 65s (valid n=237,791, missing 9500).

Variable (reference variable)	n	Hazard ratio	95% CI	p-value
Age at primary (per additional year)	247,291	1.10	1.094 to 1.098	p<0.001
ASA grade (1)	31,164	*reference*		
2	165,705	1.20	1.142 to 1.260	p<0.001
3	39,280	2.11	2.004 to 2.231	p<0.001
4	1563	3.66	3.289 to 4.062	p<0.001
5	79	2.64	1.611 to 4.312	p<0.001
Diagnosis (OA)	220,467	*reference*		
Other arthritides	2424	1.80	1.604 to 2.015	p<0.001
Previous trauma	1577	1.65	1.383 to 1.977	p<0.001
Acute trauma	3348	2.19	2.005 to 2.396	p<0.001
Previous hip surgery	1452	1.96	1.749 to 2.188	p<0.001
DDH	799	0.92	0.745 to 1.305	p=0.921 (ns)
AVN	4002	1.35	1.238 to 1.468	p<0.001
Other (inc infection)	3722	1.83	1.684 to 1.984	p<0.001
Gender (Female)	150,768	*reference*		
Male	87,023	1.55	1.504 to 1.595	p<0.001
Provider type (NHS)	157,606	*reference*		
NHS treatment centre	8996	0.91	0.837 to 0.982	p=0.016
Independent hospital	59,107	0.73	0.700 to 0.756	p<0.001
Independent TC	12,082	0.67	0.609 to 0.735	p<0.001
Hip type (Cemented)	128,349	*reference*		
Uncemented	69,720	0.93	0.895 to 0.964	p<0.001
Hybrid	36,770	0.96	0.918 to 1.000	p=0.050
Resurfacing	2952	0.62	0.504 to 0.757	p<0.001
Lead surg grade (Consultant)	186,575	*reference*		
SpR/ST3-8	24,501	0.96	0.918 to 1.008	p=0.103 (ns)
F1-ST2	291	0.76	0.527 to 1.106	p=0.154 (ns)
Speciality Dr	16,577	0.95	0.908 to 1.002	p=0.060 (ns)
Other	9847	0.89	0.817 to 0.978	p=0.015

(ns) is not significant. Approach is not significant in this model, p=0.336

Variable (reference variable)	n	Hazard ratio	95% CI	p-value
Approach (Hardinge/lateral/ant-lat)	107,824			
Posterior	106,805			p=0.252 (ns)
Anterior	956			p=0.178 (ns)
Trochanteric osteotomy	831			p=0.370 (ns)
Other	21,375			p=0.856 (ns)

Table 3: Variables in the model and associated hazard ratios (95% confidence intervals) for under 50s (valid n=23,999, missing 905).

Variable (reference variable)	n	Hazard ratio	95% CI	p-value
Age at primary (per additional year)	24,904	1.03	1.014 to 1.050	p<0.001
ASA grade (1)	11,764	*reference*		
2	10,832	2.41	1.788 to 3.235	p<0.001
3	1331	7.13	5.070 to 10.015	p<0.001
4	69	30.57	18.269 to 51.144	p<0.001
5	7	0.00	0 to 1.0×10^{269}	p=0.979 (ns)
Diagnosis (OA)	16,098	*reference*		
Other arthritides	867	0.37	0.150 to 0.914	P=0.031
Previous trauma	292	1.67	0.530 to 5.270	P=0.380 (ns)
Acute trauma	288	6.14	3.926 to 9.592	p<0.001
Previous hip surgery	205	6.38	3.420 to 11.888	p<0.001
DDH	2848	0.92	0.564 to 1.502	p=0.741 (ns)
AVN	1861	2.71	1.939 to 3.785	p<0.001
Other (inc infection)	1544	4.92	3.645 to 6.634	p<0.001
Provider type (NHS)	16,615	*reference*		
NHS treatment centre	939	1.45	0.858 to 2.454	p=0.165 (ns)
Independent hospital	5893	0.43	0.292 to 0.644	p<0.001
Independent TC	556	0.68	0.252 to 1.831	p=0.444 (ns)
Hip type (Cemented)	3120	*reference*		
Uncemented	10,851	0.42	0.323 to 0.524	p<0.001
Hybrid	2819	0.45	0.314 to 0.635	p<0.001
Resurfacing	7213	0.24	0.164 to 0.353	p<0.001

(ns) is not significant. Gender, Surgeon Grade and Approach are not significant in this model, p=0.416, p=0.388 and p=0.361 respectively

Variable (reference variable)	n	Hazard ratio	95% CI	p-value
Gender (Female)	11,381			
Male	12,622			
Lead surg grade (Consultant)	22,179			
SpR/ST3-8	1175			p=0.598 (ns)
F1-ST2	10			p=0.708 (ns)
Speciality Dr	341			p=0.119 (ns)
Other	298			p=0.293 (ns)
Approach (Hardinge/lateral/ant-lat)	7016			
Posterior	14,762			p=0.223 (ns)
Anterior	117			p=0.911 (ns)
Trochanteric osteotomy	170			p=0.118 (ns)
Other	1938			p=0.433 (ns)

Table 4: Updated model output (with IMD included) (valid n=337,062, missing 45,078).

Variable (reference variable)	n	Hazard ratio	95% CI	p-value
Age at primary (per additional year)	381,922	1.08	1.077 to1.081	p<0.001
ASA grade (1)	69,328			
2	222,512	1.21	1.149 to 1.269	p<0.001
3	43,471	2.32	2.189 to 2.447	p<0.001
4	1649	4.52	4.047 to 5.043	p<0.001
5	102	2.94	1.765 to 4.880	p<0.001
Diagnosis (OA)	304,600			
Other arthritides	4658	1.66	1.475 to 1.873	p<0.001
Previous trauma	2427	1.51	1.244 to 1.838	p<0.001
Acute trauma	4308	2.65	2.425 to 2.900	p<0.001
Previous hip surgery	1822	2.30	2.038 to 2.596	p<0.001
DDH	5291	1.03	0.818 to 1.286	p=0.826 (ns)
AVN	7402	1.56	1.426 to 1.697	p<0.001
Other (inc infection)	6554	2.43	2.243 to 2.623	p<0.001
Gender (Female)	201,396			
Male	135,666	1.53	1.482 to 1.578	p<0.001
Provider type (NHS)	216,415			
NHS treatment centre	13,318	0.92	0.845 to 1.004	p=0.061 (ns)
Independent hospital	90,474	0.74	0.711 to 0.771	p<0.001
Independent TC	16,855	0.59	0.531 to 0.653	p<0.001
Hip type (Cemented)	145,392			
Uncemented	118,742	0.84	0.802 to 0.869	p<0.001
Hybrid	48,513	0.87	0.845 to 0.930	p<0.001
Resurfacing	24,415	0.55	0.486 to 0.628	p<0.001
Lead surg grade (Consultant)	274,874			
SpR/ST3-8	30,578	0.98	0.928 to 1.026	p=0.339 (ns)
F1-ST2	358	0.97	0.683 to 1.384	p=0.876 (ns)
Speciality Dr	19,307	0.98	0.926 to 1.033	p=0.443 (ns)
Other	11,945	0.75	0.669 to 0.824	p<0.001
Approach (Hardinge/lateral/ant-lat)	141,636			
Posterior	162,193	0.93	0.899 to 0.965	p<0.001
Anterior	1417	0.82	0.630 to 1.067	p=0.139 (ns)
Trochanteric osteotomy	1445	0.51	0.309 to 0.826	p=0.006
Other	30,371	1.27	1.213 to 1.327	p<0.001
IMD (per additional unit)	350,609	1.01	1.005 to 1.008	p<0.001

(ns) is not significant