Biostatistics Collaboration of Australia

Workplace Research Project Part C

Peter Smerdely

SID: [Redacted]

Contact: 0411168459

Total Number of Pages, including this page = 39
Table of Contents

Student’s role ................................................................................................................................. 4
Reflections on Learning .................................................................................................................. 4
Teamwork ........................................................................................................................................ 5
Ethical Considerations .................................................................................................................... 5
Project Report .................................................................................................................................. 6
1. Project Title ................................................................................................................................ 6
2. Location & Dates .......................................................................................................................... 6
3. Context ......................................................................................................................................... 6
4. Student Contribution .................................................................................................................... 6
5. Statistical Issues .......................................................................................................................... 7
6. Student Declaration ..................................................................................................................... 7
7. Supervisor’s statement: .................................................................................................................. 7

A Comparison of Hospital Outcomes in Older People ..................................................................... 8
with & without Diabetes .................................................................................................................... 8

Abstract: .......................................................................................................................................... 9

Introduction: ...................................................................................................................................... 10

Methods .......................................................................................................................................... 12
  Study design .................................................................................................................................... 12
  Setting ............................................................................................................................................. 12
  Data ................................................................................................................................................. 13

Hospital Acquired Diagnosis (HADx) ............................................................................................. 14

Unplanned Re-admission ............................................................................................................... 14

Statistical Analysis .......................................................................................................................... 14

Ethics ............................................................................................................................................... 15

Results ............................................................................................................................................ 16
  Cohort summary ............................................................................................................................ 16
  Primary Outcomes .......................................................................................................................... 21

Discussion ........................................................................................................................................ 22

Conclusion ........................................................................................................................................ 24

References ......................................................................................................................................... 25

Commentary on Statistical Methods used Work Place Project .......................................................... 28
  Data management ........................................................................................................................... 28
  Charlson Comorbidity Index (CCI) ................................................................................................. 31
  Classification of Hospital Acquired Diagnosis (CHADx) .............................................................. 32
Statistical Analysis ......................................................................................................................... 32
Choice of Dependent Variables ....................................................................................................... 32
Choice of Outcome Variables ............................................................................................................ 33
Model Development........................................................................................................................... 34
Generalised Linear Models (GLM) ..................................................................................................... 34
Negative Binomial Regression ........................................................................................................... 35
Logistic Regression ............................................................................................................................ 36
Cox Proportional Hazard .................................................................................................................... 37

Figures

Figure 1: Population by Age Group in the St George Hospital Catchment........................................... 12
Figure 2: Separations by Financial Year .............................................................................................. 17
Figure 3: Separations by Age Group .................................................................................................... 18
Figure 4: Schema for Data Management Processes ........................................................................... 29
Figure 5: Regression Coefficients & Scatterplot of Age .................................................................... 33
Figure 6: Deviance residuals V Predicted Negative Binomial Regression ......................................... 36
Figure 7: Fitted Values V Deviance Residuals .................................................................................. 37
Figure 8: Kaplan Meier Survival estimates ......................................................................................... 38
Figure 9: Survival Curves .................................................................................................................. 39

Tables

Table 1: Cohort characteristics ........................................................................................................... 17
Table 2: Outcomes by Age Group & Diabetes Status ......................................................................... 19
Table 3: Hazard Ratio for in-hospital mortality comparing Age Groups with Diabetes Diagnosis ....... 20
Preface

Student’s role

This project has its genesis in 2016 when I began exploring some data related to length of stay and diabetes in people admitted to the Department of Aged Care at St George Hospital, Kogarah, where I work. As the M. Biostatistics course progressed I was able to understand the requirements to answer the question that I was posing: “Does diabetes matter in the very elderly?”

I subsequently obtained a substantial dataset from the hospital’s casemix unit, in Excel format. I reformatted the data to make it suitable for analysis. I merged several datasets to the main dataset. With the guidance of my supervisor, I performed statistical analysis. With the analysis, I expressed the results in the form of a manuscript to be submitted for peer review.

Reflections on Learning

The project has allowed me to learn a great deal. I have applied learnings from almost every subject undertaken. I found the data management subjects particularly useful (Data Management & Statistical Computing, Introductory and Advanced Analysis of Linked Data). They gave me the skills to manage data in different formats and be able to quickly sort into different groups. I am able to link multiple complex datasets. This was required for this project.

The statistical analysis of the project required multiple attempts. Despite having achieved reasonable results in the subjects, the project provided challenges. The major type of analysis used in the study (negative binomial regression) was only mentioned in passing within the course. With the guidance of my supervisor, I was able to gain insight and learn to use this valuable technique.

In doing this project, I believe I gained skills in using Stata. I successfully converted a Stata (user-written) program to a do file for another purpose. I learned a reasonable amount of coding skills. I believe I can code efficiently.
Teamwork

While the majority of the work was self-directed, I required assistance from the casemix unit at St George Hospital. We were able to collaborate to obtain not only hospital data but extended versions containing over 400 ICD10 codes that are not routinely provided.

Ethical Considerations

Given the possibility of publication, ethics approval was sought and obtained.
Project Report

1. Project Title
A Comparison of Hospital Outcomes in Older People with & without Diabetes: A 5 year review of case-mix data at a single site.

2. Location & Dates
This project was conducted from June 2017 until January 2019. The data was provided by the St George Hospital Casemix Unit on people admitted to St George Hospital aged 65 years or more from 2012 - 2017.

3. Context
There is a philosophical conflict between geriatricians (those hospital physicians caring for the needs of older persons) and the traditional teaching for the management of diabetes. Older people often carry a burden of disease, have cognitive impairment and are physically frail. Rigid glycemic control may put unnecessary stress on people and their carers. To manage complex diabetes regimens, older patient require complex support networks or even moving to residential aged care. Anecdotally and with some emerging evidence, geriatricians are not using traditional methods for the management of diabetes often to the distress of patients and carers who are used to the more traditional methods. In this context, this study was conceived.

4. Student Contribution
Liaising with casemix unit to obtain datasets

Data management involving: Modifying datasets to be in a form that could be analysed, Merging datasets with clinical indicators Finding relevant cases to analyse

Statistical Analysis involving: Choice of methods used Ensuring methods used correctly Goodness of fit analyses

Dissemination of results: Design of tables Drawing of graphs
5. Statistical Issues

There were several statistical issues explored. The first was study design. Case-mix datasets are designed for administrative purposes. Correct formatting and removal of extraneous data was difficult at times. The next issue was ensuring the data was correct. The choice of statistical methods was another issue. The fitting of several different regression models was undertaken. The diagnostics of the analysis was another issue. My biggest issues was the distillation of a large number of outputs into simple readable and understandable tables or graphs.

6. Student Declaration

I, Peter Smerdely, declare this project is evidence of my own work, with direction and assistance provided by my project supervisor. This work has not been previously submitted for academic credit.

---------------------------------------
Peter Smerdely Date:

7. Supervisor’s statement:

I confirm that the work presented in this report has been conducted by Dr Peter Smerdely. Peter has been responsible for all aspects of this project; from the formulation of the research question, obtaining the data, conducting the analyses, through to the interpretation and writing up of the results.

---------------------------------------
Patrick Kelly Date:
A Comparison of Hospital Outcomes in Older People with & without Diabetes

A 5-year review of case-mix data at a single site.
Abstract:

Given there is little data about hospital outcome in people with diabetes aged beyond 75, this study aimed to explore the association of diabetes with hospital outcome of patients aged greater than 85 years (85+) and compare their outcomes with those aged between 65 and 85 years.

A retrospective review was conducted of the first presentation of patients age 65 years or more admitted to a Sydney teaching hospital over a 5 year period (2012-2017) exploring the outcomes of length of stay, mortality, the development of hospital acquired adverse events and unplanned readmission to hospital within 28 days of discharge. Demographic and outcome data, the presence of diabetes and co-morbidities was determined from ICD10 coding within the hospital’s electronic medical record. Logistic, negative binomial and Cox proportional hazard regression models were used to assess association of diabetes with outcome.

Of the 26673 people, 25.7% had diabetes. When compared to people without diabetes aged 85+, diabetes was not associated with increased length of stay (p=0.312), hospital acquired adverse events (p=0.374) or unplanned readmission (p=0.889). Further, people with diabetes had lower rates of mortality than those without diabetes in the 85+ years group (5.2% V 7.1%, p=0.003). Similarly, hazard ratios were significantly higher in the non-diabetic group (2.6 V 3.7, p=0.006).

Diabetes has not been shown to have a negative impact on mortality, length of stay, hospital acquired adverse events and readmission in people aged 85+ years.
Introduction:

According to the International Diabetes Federation, there are 425 million people over age 20 with diabetes worldwide in 2017 with a prevalence of 8.8% [1]. In the United States, there are 30 million people diagnosed with diabetes, with a prevalence of 9.4% in 2015 [2]. There is a prevalence of 6% with approximately 3.5 million suffering the disease in the UK [3]. Similarly, in 2012 there was a prevalence of 5.4% in Australia, or about 1 million people with diabetes [4]. It is a large problem.

Diabetes is associated with increased risk of morbidity and mortality. Diabetes UK report increased cardiovascular, renal, ophthalmic, peripheral vascular, neurological and psychiatric disease in poorly controlled diabetes with increased risk of mortality and reduced life expectancy [5]. In Australia, the Australian Institute of Health and Welfare (AIHW) report death rates being between 1.6 and 2 times as high for those with diabetes than the general population [6].

Given the increased morbidity and mortality, it is not surprising that diabetics are over represented in hospital admission statistics. Hospitalisations in people with diabetes are both more likely and more frequent as well as being longer admissions [7-9]. Similarly, the prevalence in Australia of people hospitalised for any reason who have diabetes ranges between 8.9 and 35.1% [10, 11].

Ageing is strongly associated with the development of diabetes [12]. It is likely that the interplay between genetics, environmental factors and normal ageing is the cause [13]. The changes that occur with ageing include changes in carbohydrate metabolism, metabolic alteration and glucose counter regulation [12, 13]. Thus the prevalence of diabetes increases with age with over 5 times the prevalence of diabetes in people aged 65 years or more compared to those under 45 years in the US [2]. Similar prevalence statistics have been reported in Australia with rates 3 times higher in those over 65 years of age [4]. AIHW also report that this rate has remained stable since 2007-2008 [4].

However, there is little data that explore hospitalisation or outcome in people with diabetes aged beyond 75 years with results often grouped as 65+ or 75+[14]. Anecdotally, experienced geriatricians
feel that diabetes does not confer increased morbidity or mortality in the very elderly (those aged over 85 years).

Given the paucity of literature in this regard, this study was undertaken. Its aim is to explore the effect of a diabetes diagnosis on the hospital outcome of patients aged greater than 85 years and compare their outcomes with those aged between 65 and 85 years. The hospital outcomes were defined as length of stay, mortality, the development of hospital acquired adverse events and unplanned re-admission to hospital within 28 days of discharge.
Methods

Study design

This was a retrospective study of all patients aged 65 years and older admitted for at least 24 hours to St George Hospital over a 60-month period, from 1 July 2012 to 30 June 2017. Patients under 65 years or those present for day only intervention, outpatient reviews, and ambulatory care were excluded. Additionally, subjects were excluded if they were admitted for psychiatric management or purely for renal dialysis. Subjects were analysed for their first admission only. It was expected that patients may have multiple admissions over the 5-year period.

Setting

St George Hospital is a 600-bed tertiary referral hospital in southeast Sydney with an estimated 40,951 people over the age of 65 living in the St George catchment area in 2016, see Figure 1 (Source: Australian Bureau of Statistics). The prevalence of diabetes in the local health district is 6.4% and is below the New South Wales average prevalence of 8.7% [14].

Figure 1: Population by Age Group in the St George Hospital Catchment (Source: ABS)
Data

Eligibility was based on coding within the "Patient Information Manager" (iPM), a system handling all the demographic data, discharge diagnostic codes (DRG) and separation data for all admissions. Cases were defined as those with a principal or additional diagnosis of diabetes (ICD-10-AM code: E10 Type 1 diabetes mellitus, E11 Type 2 diabetes mellitus, E12 Malnutrition-related diabetes mellitus, E13 Other specified diabetes mellitus, E14 Unspecified diabetes mellitus) as defined by International Classification of Disease 10th Revision Australian Modification (ICD-10-AM). This included new diagnosis during the admission. Demographic and separation data were obtained. Similar data was collected from a non-diabetic cohort for the same time period, for comparison.

Hospital records were supplied in three separate datasets, linkable by the patient identifier, medical record number (mrn), and unique stay number. The primary data set contained demographic data that included gender and age. Medical specialty teams and wards were provided. Separations were defined as formal discharges from the hospital where exit date and time are recorded in iPM. The separation data included length of stay and type of separation. The type of separation included death, discharge to residential aged care facility (RACF) and discharge home as well as transfer to other hospital (for the purpose of rehabilitation, repatriation or specialised care such as burns treatment). The second data set contained diagnosis and procedure codes for all diagnoses and procedures for each admission (ICD-10-AM codes). In addition, a dataset contained a flag to indicate which diagnosis code occurred during admission to classify diagnosis codes as hospital acquired. These indicate adverse events [15].

The second dataset, containing ICD-10-AM codes, was used to generate a modified Charlson Comorbidity Index (CCI) (see below) [16, 17]. The CCI was modified by removing the weights associated with diabetes. The second dataset was merged with the primary dataset.
The third dataset contained the flag for hospital acquired diagnosis (see below). Using a custom written Stata program, a hospital acquired diagnosis index (HADx) was generated, this dataset was merged into the combined dataset used for analysis.

**Hospital Acquired Diagnosis (HADx)**

The Classification of Hospital Acquired Diagnoses permits the identification of adverse events that have occurred in hospital during an admission [15, 18].

The development of the HADx stemmed from a recognition that ICD-10 codes alone were insufficient to monitor adverse events [19]. Subsequently a classification system was developed that is able to identify 17 groups of adverse events using a flag to detect the presence of ICD-10 codes occurring in hospital that has been validated across several Australian health jurisdictions [15, 18, 20]. To date, this method has been used to explore adverse drug events and effects on length of stay [21, 22]. Recently, Cromarty, et al, demonstrated the effects of hospital acquired conditions on the length of stay of people admitted to hospital with a comorbid diagnosis of diabetes [23]. Therefore, hospital acquired adverse events have been included as a possible confounder in this study.

**Unplanned Re-admission**

Unplanned re-admission was defined as being admitted for greater than 24 hours within 28 days following a previous separation. Pre-arranged or booked admissions within this time frame were excluded.

**Statistical Analysis**

Data was analysed using Stata Version 15 (StataCorp, College Station, TX, USA). Variables are described using mean and standard deviation, proportion and range. For the purpose of association, age, presence of diabetes and the interaction of these two were the variables of interest. The outcome variables of interest are length of stay, mortality within hospital (now referred to as mortality), hospital
acquired disease and readmission rates. Logistic and negative binomial regression models were used to assess association for the binary (mortality, 28 day readmission and hospital acquired disease) and count (length of stay) outcomes respectively. All models were adjusted for the potential confounding effects of gender and CCI. Effect modification (statistical interaction) was also assessed between the diabetes and the other variables in the model. A 5% two-sided significance level was used for main effects and 1% for interactions. Cox proportional hazard models were used to assess hazard ratios with respect to age and diabetes, adjusting for CCI and stratified on gender.

Sensitivity analyses and goodness of fit tests were conducted to examine the fit of the models. Deviance residuals versus fitted values were plotted and assessed for each regression. For logistic regression, goodness of fit was also assessed using the Hosmer-Lemeshow Chi Square Tests [24]. Sensitivity analysis showed the relationship of age and length of stay was not linear. Subsequent analysis was conducted using age dichotomised at age 85 years.

Ethics

The project was assessed and approved by the South Eastern Sydney Local Health District Research & Ethics Committee (HREC No: 18/007)
Results

Cohort summary

There were 69,393 separations recorded between 2012 and 2017. Of these, 52,358 had a length of stay greater than 24 hours. After excluding dialysis and psychiatry, there were 51,440 separations remaining.

There were 26,673 patients with their first separation. Presentations by financial year are shown in Figure 2 with 6,833 patients in 2012 to 4,257 patients in 2016. These 26,673 separations are shown in Figure 2 by age group, with a peak in the 80-84-year group. There were 848 separate ICD10-AM codes in the cohort with the 5 most frequent being: I21 Acute myocardial infarction, I50 Heart failure, I63 Cerebral infarction, J18 Pneumonia, unspecified organism, S72 Fracture of femur. These represent 64% of all diagnoses. The Charlson Comorbidity Index (CCI) ranged from 0 to 12 with a mean of 1.03 (SD 1.63), with 55% having a score of 0, 20% having a score of 1 and a 29% having score of 2 or more.

Table 1 summarises the cohort characteristics. Of the cohort of 26,673, 25.7% had diabetes. Those with diabetes were younger and more likely to be males, compared to those without diabetes. Those with diabetes had more separations and comorbidities.
### Table 1: Cohort characteristics

<table>
<thead>
<tr>
<th></th>
<th>Diabetes</th>
<th>No Diabetes</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number (n)</td>
<td>6843</td>
<td>19830</td>
<td></td>
</tr>
<tr>
<td>Age (Years, (SD))</td>
<td>77.8 (7.5)</td>
<td>79.5 (8.2)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age85+ (n, %)</td>
<td>1427 (19.0)</td>
<td>6088 (28.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Female (n, %)</td>
<td>3096 (45.2)</td>
<td>10615 (53.5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Separations(^c) (n, %)</td>
<td>1 3884 (56.8)</td>
<td>12171 (61.4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>2 1309 (19.1)</td>
<td>3781 (19.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3+ 1650 (24.1)</td>
<td>3878 (19.6)</td>
<td></td>
</tr>
<tr>
<td>CCI (^c) (n, %)</td>
<td>0 3361 (49.1)</td>
<td>11427 (57.6)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>1 2481 (36.3)</td>
<td>6450 (32.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2+ 1001 (14.6)</td>
<td>1953 (9.9)</td>
<td></td>
</tr>
</tbody>
</table>

\(^c\)Number of separations by each individual in the 5-year period. \(^c\)Charlson Comorbidity Index. Percentages are column percentages. P values are from chi squared assessments.

---

**Figure 2: Separations by Financial Year**
Figure 3: Separations by Age Group. Diabetes groups are overlayed
Table 2: Outcomes by Age Group & Diabetes Status. Estimates have been obtained from regression analyses

<table>
<thead>
<tr>
<th>Age Group (years)</th>
<th>Unadjusted Means (95%CI)</th>
<th>Adjusted Means (95%CI)</th>
<th>Adjusted IRR/OR*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Diabetes</td>
<td>Diabetes</td>
<td>p</td>
</tr>
<tr>
<td>LOS ( \text{^c} ) (days)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65-84</td>
<td>9.0 (8.8, 9.2)</td>
<td>10.0 (9.6, 10.4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>85+</td>
<td>9.5 (9.2, 9.8)</td>
<td>10.1 (9.6, 10.7)</td>
<td>0.049</td>
</tr>
<tr>
<td>Interaction 1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mortality (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65-84</td>
<td>2.8 (2.6, 3.1)</td>
<td>2.8 (2.3, 3.2)</td>
<td>0.825</td>
</tr>
<tr>
<td>85+</td>
<td>6.9 (6.3, 7.6)</td>
<td>5.8 (4.7, 7.2)</td>
<td>0.141</td>
</tr>
<tr>
<td>Interaction 1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>HADx ( \text{^c} ) (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65-84</td>
<td>32.2 (31.4, 33.0)</td>
<td>36.3 (35.0, 37.6)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>85+</td>
<td>36.9 (35.7, 38.1)</td>
<td>37.7 (35.2, 40.3)</td>
<td>0.583</td>
</tr>
<tr>
<td>Interaction 1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>28 Day</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65-84</td>
<td>8.6 (8.1, 9.1)</td>
<td>10.1 (9.3, 11.0)</td>
<td>0.001</td>
</tr>
<tr>
<td>Readmission (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85+</td>
<td>9.2 (8.5, 10.0)</td>
<td>9.3 (7.8, 10.9)</td>
<td>0.876</td>
</tr>
<tr>
<td>Interaction 1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

\( \text{^c} \)=Length of Stay, \( \text{^c} \)=Hospital Acquired Diagnosis. Adjusted for Age, Diabetes, Sex, comorbidity, Age/Diabetes interaction (shown), Age/Comorbidity Interaction. *IRR= incident rate ratio and applies only to LOS. OR= odds ratio and applies to Mortality, Hospital Acquired Diagnosis and 28-day readmission.
Table 3: Hazard Ratio for in-hospital mortality comparing Age Groups with Diabetes Diagnosis.

<table>
<thead>
<tr>
<th>Age Group (years)</th>
<th>No Diabetes n=19860</th>
<th>Diabetes n=6783</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>65-84 ref</td>
<td>0.8 (0.7, 1.0)</td>
<td>0.019</td>
<td></td>
</tr>
<tr>
<td>85+ Interaction</td>
<td>3.7 (2.7, 4.9)</td>
<td>2.6 (1.9, 3.8)</td>
<td>0.006</td>
</tr>
</tbody>
</table>

Adjusted for Age, Diabetes, Sex, comorbidity, Age/Diabetes interaction (shown), Age/Comorbidity Interaction. p* = comparing no diabetes with diabetes groups.
Primary Outcomes

The unadjusted comparisons are shown in Table 2. People in the younger aged cohort with diabetes have longer lengths of stay, have a higher proportion of hospital acquired disease and a higher proportion with unplanned admissions. There are no differences seen in the mortality rates between the non-diabetic and diabetic groups in the 65-84-year cohort. There are no differences between non-diabetics and diabetics in the older cohort.

Table 2 contains adjusted estimates following regression analysis (negative binomial or logistic, as appropriate). Age (as a categorical variable), diabetes diagnosis, modified Charlson Comorbidity Index and sex were highly significant in regression models (p<0.001). The interaction between age and diabetes was not significant. However, it was kept in the modelling as the interaction between age and diabetes is being explored in this study. Further, sensitivity analysis was conducted with interactions between all other variables. Only the interaction between age and CCI showed significance (p<0.001). It is also seen that people aged over 85 years with diabetes have lower estimates of mortality and odds of death than those without diabetes (Table 2).

Table 3 shows the Cox proportional hazards with those in the older cohort having higher hazard ratios. People aged 85 or more years with diabetes had significantly lower hazard ratios than non-diabetics (Table 3).
Discussion

The aim of this study was to explore the effect of a diabetes diagnosis on the hospital outcome of patients aged greater than 85 years and compare their outcomes with those aged between 65 and 85 years. This study demonstrated no effect of diabetes diagnosis in people aged 85 years or more on length of stay, hospital acquired diagnosis or 28-day unplanned readmission rates. Moreover, it showed better estimates of mortality in older diabetics than in the younger cohort. In the younger cohort, diabetes was associated with significantly worse outcomes in length of stay, hospital acquired disease and 28-day unplanned re-admission rates. There were no differences in mortality.

This study demonstrated that diabetes has an important association with length of stay in the younger cohort. This is consistent with other studies [25-28]. Ko et al showed an increased length of stay with renal abscess, Malone et al showed this with peripheral vascular disease and Nirantharakumar et al showed this with fractured proximal femur [25, 26, 28]. This may imply that length of stay related to the reason for admission rather than diabetes itself. Sajjad et al adjusted for the clinical and non-clinical measures such as glycaemia or socio-economic status and found longer length of stays[27] . On the other hand, Nirantharakumar and co-workers did not find any difference in length of stay in diabetics admitted to hospital with foot disease [29]. The present study differs from these preceding works by including only people over the age of 65 years, stratifying with age greater than 85 years and adjusting for the confounding of disease burden by using the Charlson Comorbidity Index. It remains consistent with the larger body of clinical studies. Furthermore, also consistent with these studies is the magnitude of the effect, which is small.

The diagnosis of diabetes was not associated with increased risk of mortality in this study. The increased risk was found between age groups. Interestingly, a significantly lower hazard ratio was found in diabetics who were older than 85 years compared to non-diabetics in the same age group. However, while statistically different it is likely due to immortal time bias and not to be of clinical relevance[30, 31]. Most of studies that deal with mortality and diabetes are longitudinal studies [32-
There only a few studies that explore in-hospital mortality associated with diabetes and these again are disease specific[25, 28, 29]. Two of the studies show not effect on inpatient mortality[25, 28]. The third showed an odds ratio of 1.31 (1.04-1.65) for mortality in diabetics with foot disease [29]. The study did not adjust for comorbidity. It is possible that there are disease specific subsets that are at higher risk, but this was not the aim of the present study.

This study demonstrated that the diagnosis of diabetes is associated with the development of hospital acquired adverse events in the younger cohort, but not in the older cohort. The use of a validated but not extensively used method of detection of hospital acquired diagnosis was needed to achieve this [18-23]. Using this method, Cromarty et al were able to show that 29.3% of diabetics developed hospital acquired events compared to 13% of non-diabetics[23]. The present study found a much higher proportion of hospitalised people had complication. From that high baseline, diabetics aged between 65 and 84 years had even more complications (p=0.003). However diabetes made no difference in the older cohort. The difference between the studies is likely due to the different ages of the two cohorts.

This study demonstrated that the odds of being re-admitted within 28 days of discharge was associated with the diagnosis of diabetes in the younger cohort but not in the older. Caughey et al identified older people with co-morbidities as those most likely to be re-admitted within 30 days [37]. The present study adjusted for the presence of comorbidity and found little difference, providing some support for their result. Dungan identified those with poor glycaemic control as those most likely to be re-admitted[38]. Clinical measures were not undertaken in the present work. The present study did not show why the diagnosis of diabetes did not have an impact on the very elderly. However it did confirm the anecdotal experiences of experienced geriatricians. Recently, the management of diabetes in the older person has changed with less rigid control goals [39]. There is evidence that rigid control may not have the benefits seen in younger people. A converse
viewpoint is that diabetes may not be as harmful in older people and so its control does not need to be as tight. The associations found within this study are consistent with this premise.

The current study has several limitations. It is a retrospective audit of hospital administrative data. Hospital databases have not been designed for clinical investigations. However, the large amounts of data can be used for association. Several validation studies have been conducted with sensitivities and specificities of up to 95.6% and 98.5% respectively [40]. However, studies also warn about changes in coding rules that occur over time, such as the changes that occurred in the definitions of diabetes in 2011 [41]. This has resulted in a decrease level of reporting. This study commenced in 2012 for that reason. The study is from a single site. There may be geographical, cultural and clinical practice issues that may be specific to one site. Replication of the study at other sites may be required. However, it is still reasonable to generalise.

This study did not use specific clinical measures. Glycaemic control and measures of frailty are two such measures. Several small works have examined hospital outcome based on these clinical measures [42-45]. It was beyond the scope of this study to do so. Certainly prospective studies that explore the effects of glycaemic control and frailty on hospital outcomes are needed.

The study cohort had a significantly lower proportion of people aged 85 years or more who had diabetes compared to those under the age of 85 (19.0% v 28.3%, p<0.001).

**Conclusion**

Diabetes has not been shown to have a negative impact on mortality, length of stay, hospital acquired adverse events and readmission in people aged 65 years and more, comparing the older patient to the very old patient in data derived from hospital administrative records.
References


Commentary on Statistical Methods used Work Place Project

Data management

Due to the size of the dataset, the data required was produced in 3 separate datasets linked by the patient identifier, medical record number (mrn), and unique stay number.

The primary data set contained demographic data included gender and age. Medical specialty teams and wards were provided. The separation data included length of stay and type of separation. The type of separation included death, discharge to residential aged care facility (RACF) and discharge home as well as transfer to other hospital (for the purpose of rehabilitation, repatriation or specialised care such as burns treatment).

The second data set contained diagnosis and procedure codes for all diagnoses and procedures for each admission (ICD-10-AM codes). This dataset had over 400 ICD10 codes for many patients.

In addition, an additional dataset contained a flag to indicate which diagnosis code occurred during admission to classify diagnosis codes as hospital acquired. These indicate adverse events.

The datasets were formatted and merged to allow analysis. See Figure for outline of processes used.
Figure 4: Schema for Data Management Processes. Case-mix data was converted from Excel format to Stata format prior to processing.
Each dataset was provided as an excel spreadsheet. Each excel spreadsheet was imported into Stata 15. The datasets were checked for quality and formatting. No duplicates were found. Formatting was adjusted to allow for calculations. The primary dataset was used as the base dataset. The second dataset, containing ICD-10-AM codes, was used to generate a modified Charlson Comorbidity Index (CCI) (see below). The CCI was modified by removing the weights associated with diabetes. The second dataset was merged with the primary dataset.

The third dataset contained the flag for hospital acquired diagnosis (see below). Because of the formatting provided, this dataset required significant manipulation. The flag was a series of 1’s and 2’s in “string” format. Each ICD-10 code had an individual flag. Each admission had up to 409 codes. The names of the 818 variables required renaming to allow the statistical program to manipulate the large number of variables. The two codes were combined and reformatted for the use of the program. The result was 409 variables containing only ICD-10 codes that were acquired in hospital. A separate Stata program (Stata do file) was custom written to interrogate the 409 variables for one of over 2400 ICD-10 codes grouped into 14 different domains. The result was a dataset containing 14 variables each for a different group of hospital acquired disease (see below). This dataset was merged into the combined dataset used for analysis.

Cases were excluded if they were admitted for psychiatric management or purely for renal dialysis.

It was recognised that individuals had multiple admissions during the time period. It was decided to only assess the patient’s first admission to hospital.

Prior to the deletion of subsequent admissions, the data was resorted to determine the number of admissions per patient and the length of time between these admissions. A variable was created to reflect unplanned readmissions within 28 days of discharge. Other outcomes were length of stay, death and hospital acquired diagnosis (complications of hospitalisation).
Charlson Comorbidity Index (CCI)

The information contained in the datasets has multiple diagnoses per patient with over 800 individual ICD-10 codes or diseases. These diseases are termed comorbidities. Comorbidities are additional diagnoses to the disease of interest. They are often confounders and are separately associated with mortality and morbidity. For the purposes of analysis, it is sensible to group the diagnoses into co-morbidities.

The two most commonly used summaries of comorbidities are the Charlson and Elixhauser Comorbidity Indexes [16, 46]. These have been extensively reviewed and compared. Reviews have recommended one over the other with no consensus of which should be used. Certainly, the CCI has been around longer and has had more validation studies over a wider variety of conditions. In this study, both have been implemented. After several different analyses, there was no substantial effect on outcome data using either the Charlson or Elixhauser systems within this study. Therefore, the Charlson Comorbidity Index was used in the reporting the summaries of comorbidities in this study.

Charlson, in 1987, developed a method to determine the prognostic impact of individual comorbid diseases at 1 and 10 years using two cohorts of patients. Since that time it has been used in many studies that require an estimate of burden of disease. Charlson showed that the groupings of comorbid conditions was able to predict outcome. It is a required confounder for any study of hospital based outcomes.

The Charlson Comorbidity Index has been validated for use in a wide variety of populations, ranging from different diseases, different cultures to different populations. Especially important for this study, it has been validated for use in people with diabetes and the older person.

In this study, the CCI was calculated using a module within the Stata program. The counts for diabetes and related complications were not used in the calculation of the grouped index, given it was the disease of interest. The results have been presented as the number of comorbidities, or the
more common grouping, either 0 or 1 or more comorbidity. From now the CCI will refer to this later modified grouping.

Classification of Hospital Acquired Diagnosis (CHADx)

As with CCI, the CHADx index was calculated using the Stata program. Unfortunately, there is no module available yet, so one was developed in a similar style to the CCI and Elixhauser modules. This has been described above. The results have been presented as the number of hospital acquired diagnoses or grouped as the presence or absence of a hospital acquired diagnosis.

Statistical Analysis

Choice of Dependent Variables

The question posed in this work was whether diabetes had an effect on outcome in older people. Thus, outcome variable would be regressed against the categorical variable diabetes and age. Age is a continuous variable but with respect to length of stay, the relationship is not linear. This can be seen in the results of the regression of age and the square of age. See Figure 5. Using the relationship \(-\beta_1 / 2 \times \beta_2\), the change occurred at 83.7 years. The analysis could have been conducted using fractional polynomials or cubed spline methods. However for simplicity, a dichotomous variable of age above or below 85 years was created. The age 85 years was compatible with the regression evidence as well as clinician input defining the very elderly. The relationship of age with the other outcome variables was linear.

The interaction term between age and diabetes was included in all models.
Choice of Outcome Variables

Case-mix data is limited clinical utility. Outcomes are restricted to length of stay, type of separation and diagnosis. There are 9 different types of separation. Death is recorded well. However other types of separation have varying degrees of accuracy. In general, discharge to a residential care facility is regarded as a poor outcome. However, such a separation may be recorded as a discharge to nursing home or discharge home or even transfer to another health facility. Further, the address of origin rarely accurately reflects whether they came from a residential aged care facility. Thus, new nursing home placement cannot be used as an outcome measure.

Length of stay, death, hospital acquired disease and unplanned re-admission within 28 days were chosen.
Model Development

With all models the process commenced with a minimal model which included the categorical variables of the presence of diabetes and age over 85 year and their interactions, as these were the clinical questions. The models were supplemented with confounding variables: gender, Charlson Comorbidity Index and the presence of hospital acquired diagnosis. The interactions of all the variables were then entered into a saturated model. The non-significant variables or interactions were removed in stepwise fashion until the model with the best fit statistics was achieved. When this model was achieved. Goodness of fit tests were applied. The model was adjusted when issues with goodness of fit were found. Tables were generated from the results or from linear combinations of parameters.

Generalised Linear Models (GLM)

GLM extend ordinary linear regression to encompass distributions of the response variable that are not just normal but include normal, binomial, Poisson, gamma, inverse Gaussian, geometric and negative binomial. In standard linear regression, it is assumed that (1) the response variable is normally distributed; (2) the distributions for all observations have a common variance and; (3) there is a direct relationship between the linear predictor [47]. If these assumptions were not met, using standard linear regression would not be valid. In 1972, Nelder and Wedderburn developed “a class of generalized linear models” which could be used with non-normally distributed response variables. However, GLMs have some assumptions.

The response variable must have a variance in the distribution from the exponential family. Further, there needs to be a correctly specified link function and explanatory variables are of the correct form [47]. The outcome variables used in this study are length of stay, mortality, unplanned readmission and the development of hospital acquired disease. The latter 3 are binomial. It would be reasonable to use GLM from the binomial family with a link function that is logit (Logistic Regression, see below). Length of stay are counts. The counts are not distributed in the normal
distribution, so Poisson distribution may be appropriate. But the variance of length of stay (143.8) is much greater than the mean (9.3) suggesting marked overdispersion making a Poisson regression inappropriate. Therefore, Negative Binomia Regression was undertaken for the analysis of length of stay.

**Negative Binomial Regression**

Negative binomial regression is similar to standard linear regression except the response variables observe the negative binomial distribution. It is derived from a Poisson-gamma mixture distribution.

The application of negative binomial regression is straightforward in Stata. It can be approached through the GLM family commands or through its own command. In this study, the negative binomial command was used, however, the GLM command was used to assess goodness of fit using the negative binomial family and link. The GLM method allowed for better post estimation to determine the goodness of fit. Hilbe (2011) has a demonstrated approach. Generally this is achieved firstly by examining the Pearson dispersion statistic, ideally achieving a level close to 1. The application of a robust variance estimator help inflate p values. The value of the $\alpha$ with smaller value suggesting smaller variance. Improving the Akaike Information Criterion (AIC) & Bayesian Information Criterion (BIC) statistics with modelling is important. Finally examining the analysis of the standardized deviance residuals versus the fitted values allows another dimension to goodness of fit.

The deviance/df parameter in the final model was 1.027863, which is reasonable.
In the final length of stay model, less than 2% of cases were outside the 2 SD mark, suggesting a reasonable model fit.

*Logistic Regression*

Logistic regression is the method of choice used to explore binary responses. It is the mathematical modelling approach that can be used to describe the relationship of several variables to a dichotomous dependent variable. Similar to negative binomial regression, the application of the test is straightforward. For modelling and goodness of fit, the logistic command in Stata was used. Goodness of fit was assessed by improvements in the pseudo R², improvement with deviance with each predictor or interaction and improved likelihood ratio after testing. The Hosmer-Lemeshow and Pearson's Chi Square Tests were also employed. Examination of the analysis of the standardized deviance residuals versus the fitted values was also used to assess goodness of fit.

---

*Figure 6: Deviance residuals V Predicted Negative Binomial Regression*
The final model had goodness of fit with $P>0.3$ for the Hosmer-Lemeshow and Pearson's Chi Square Tests for all outcome measures.

Figure 7: Fitted Values V Deviance Residuals

**Cox Proportional Hazard**

Cox proportional hazards regression model was used to explore survival in the study groups. The application of the test in Stata is straightforward. Importantly, testing assumptions for proportional hazards were undertaken using Schoenfeld residuals.

The first step was to generate Kaplan-Meier Plots. These suggest that diabetes did not have an effect on survival.
Model iteration occurred using the different variables and their interactions to achieve the final model. It was found that the penultimate model failed assumptions for proportional hazards. The final model was stratified on sex. This model passed assumptions for proportional hazards (stcox i.age85##i.diabetes i.CCI i.age85#i.CCI, nolog efron strata(sex)).

Subsequently survival curves were derived. They show results consistent with the overall results namely, diabetes does not confer worse survival outcomes in the elderly.

Figure 8: Kaplan Meier Survival estimates
Figure 9: Survival Curves