Quality effects of operative delay on mortality in hip fracture treatment

R Sund, A Liski

Background: Most hip fracture patients undergo surgery, but there is conflicting evidence on the relation between the timing of surgery and the outcome of treatment. There is considerable variation in the length of surgical delays between hospitals, possibly reflecting the quality of care.

Aim: To examine the associations between in-hospital surgical delay and the mortality of hip fracture patients from a practical quality assessment perspective.

Methods: The effects of operative delay on mortality were estimated using various statistical methods applied to observational data from 16,881 first time hip fracture patients aged 65 or older from 47 hospitals (providers) in Finland in 1998–2001.

Results: A prolonged in-hospital operative delay was associated with a higher mortality of hip fracture patients in individual level analyses, but the instrumental variable approach indicated that the individual level effect was not caused by the operative delay but by inappropriate methodological assumptions. There was extensive variation between providers in the proportion of late surgery patients. Provider level analyses showed that the effects of the provider of operative delay on mortality are quite small, but there is a clear association between the proportion of late surgery patients and non-optimal treatment.

Conclusions: If provider level heterogeneity is not explicitly taken into account, studies of the effects of surgical delay on outcomes are prone to serious bias. The proportion of patients with prolonged waiting time for surgery at the provider level seems to work as an effective evidence-based quality indicator. Providers should reduce unnecessary delays to surgery and identify more carefully patients not suitable for early surgery.

Hip fractures are common injuries among older people and are associated with substantial morbidity and mortality. \(^1\) \(^2\) Ageing of the population has resulted in an increase in the mean age of hip fracture patients which is likely to cause additional problems in the treatment and rehabilitation of patients in the future. \(^4\) A surgical operation is performed for most patients during the acute management of hip fracture. A typical hip fracture patient is confined to hospital bed rest before surgery. A fairly short operative delay is suggested in the clinical guidelines,\(^6\) \(^7\) \(^8\) \(^9\) \(^10\) but it is known that there is considerable variation in the operative delays between providers. \(^11\) \(^12\) \(^13\) \(^14\) \(^15\) \(^16\) In fact, operative delay is a commonly used process measure in health system performance assessment. \(^18\) Minimisation of unnecessary preoperative inpatient care can be used as an economic justification for shorter operative delays. For quality measurement purposes, the operative delay should also have confirmed effects on outcomes otherwise improvements may possibly be targeted to issues which do not improve health. \(^19\) The evidence concerning the optimal timing of surgery in relation to the overall outcomes of hip fracture treatment is quite mixed. \(^20\) \(^21\) \(^22\) \(^23\) \(^24\) In short, the main reasons for variations in the operative delay are provider level system factors (availability of required resources such as operating rooms, support services, or personnel with expertise) and clinical decisions based on patient specific factors (co-morbidity, medical stabilisation). In some studies conflicting results may be due to methodological shortcomings (small numbers of observations, no adjustment for confounding factors), but more sophisticated evidence as to whether operative delay has an independent adjustment for confounding factors), but more sophisticated logical shortcomings (small numbers of observations, no adjustment for confounding factors), but more sophisticated evidence as to whether operative delay has an independent adjustment for confounding factors), but more sophisticated evidence as to whether operative delay has an independent adjustment for confounding factors), but more sophisticated evidence as to whether operative delay has an independent adjustment for confounding factors). Another open question is the definition of the time period for delayed surgery. \(^25\) The impact of provider level heterogeneity on the effects of operative delay on outcomes has also not received sufficient attention.

This study was undertaken to examine the associations between the in-hospital operative delay and the mortality of hip fracture patients in Finland in 1998–2001. Since previous studies concerning the effect of operative delay on mortality had reported conflicting results, special attention was paid to methodological issues. More specific goals for empirical analyses were (1) to identify a suitable definition for early and late surgery groups, (2) to describe the postoperative excess mortality associated with prolonged operative delay, (3) to illustrate in which respect the early and late surgery groups differ from each other and whether adjustment of these observed factors changes the association between operative delay and mortality, (4) to compare the risk adjusted proportion of late surgery patients between providers, and (5) to investigate provider level heterogeneity in terms of 1 year mortality and the proportion of patients with late surgery.

METHODS

Data set
The total population of hip fracture patients in 1998–2001 was identified in the Finnish Health Care Register. The medical histories (1987–2002) and deaths (1998–2002) of the hip fracture population were extracted from the Finnish Hospital Discharge Register, Finnish Health Care Register, and from the National Causes of Death Register using the unique personal identification numbers of the patient population. Each record in these registers includes data such as patient and provider ID numbers, age, sex, area codes, and diagnosis and operation codes, as well as dates of admission.
operation, and discharge (or death). The completeness and accuracy of the registers is known to be good.39-41

Data preprocessing
Data were preprocessed so that information concerning patients aged 65 or older with a first hip fracture could be accurately identified.42 For the purposes of this study, the data were restricted to patients operated on using an internal fixation, a prosthesis, or a total hip replacement. Patients with incomplete data were excluded from the analyses. A few providers (hospitals) had not reported operation dates for the years 1998–1999, but otherwise no systematic bias was found in the sensitivity analyses. The final study population consisted of 16 881 patients, which represents 83.3% of all first time hip fracture patients aged 65 or older. Forty six providers had at least 50 cases, while the remaining 154 cases from other providers were combined into a single figure. The existence of possible co-morbidities was extracted for each patient from his or her medical history using the diagnosis codes recorded in the data.43 The extraction method was adapted and updated from the original Charlson co-morbidity categories44 and applied to the current data set.

Statistical analyses
Cumulative probabilities of unadjusted mortality were calculated using the product limit estimators. Excess risk of death was defined as difference in cumulative probabilities of mortality between two comparable groups. Hazard ratios were obtained using the proportional hazards model allowing the adjustment of observed confounders. Since only a limited number of covariates were observed, the biasing influence of unobserved heterogeneity was statistically controlled by incorporating the gamma frailty term into the model.45

The significance of the differences between the unadjusted patient characteristics was determined using the z test (binary variables) or $\chi^2$ test (categorical variables). The logistic regression model was used to calculate odds ratios and significances of differences between adjusted patient characteristics. The same logistic model was also used for predicting the late surgery probability for each patient. These probabilities were then aggregated to the provider level, resulting in the expected number of late surgery patients for each provider.46 The provider specific ratios between observed and expected numbers of late surgery patients were further modelled using the hierarchical gamma-Poisson model to obtain the comparable risk adjusted proportions of late surgery patients for providers.46 The weekday adjusted proportion was calculated using the predictive margin technique.46 In the instrumental variables analysis49 the 1 year adjusted mortality was calculated for each admission day of the week and the null hypothesis of equal mortalities was tested using the $\chi^2$ goodness of fit test.46

RESULTS

Effect of length of operative delay on mortality
The cumulative probabilities of unadjusted mortality rates attributable to different lengths of operative delay are shown in fig 1. Waiting times of 0–2 nights had the same effect on mortality, but there was a significant increase in mortality with waiting times of 3–4 nights ($p<0.001$), and waiting times of 5 or more nights resulted in an even higher mortality rate. Early surgery was defined as a waiting time of 0–2 nights (n = 14 426, 85.5%) and late surgery as a waiting time of at least 3 nights (n = 2455, 14.5%). The unadjusted hazard ratio between late and early surgery was 1.24 (95% confidence interval (CI) 1.15 to 1.34, p<0.0001).

Excess risk of death
The unadjusted excess risk of death for late surgery compared with early surgery quickly rose to 3%, then slowly increased during follow up to about 5% at 1 year (fig 2). Complementary analyses revealed that the excess risk was quite stable and near to its maximum at 1–3 years and then began to decrease slowly. As the lower confidence limits in fig 2 indicate, the unadjusted excess risk of death became statistically significant 2 weeks after the operation.

Characteristics of early and late surgery groups
Characteristics of the patients with hip fracture classified into early and late surgery groups are shown in table 1, together with the corresponding odds ratios for late surgery adjusted for the other characteristics in the table. Early surgery was associated with being 85 years or older, of female sex, having a pterygochonteric fracture, and receiving long term inpatient care during the year preceding the admission. Admission on a Wednesday and a medical history of dementia or cataract were also associated with early surgery. The risks for late surgery were likely to be increased by admission from a nursing home, health centre or readmission shortly after...
The characteristics in table 1. Even after controlling for the
1.09 to 1.28, p
The hazard ratio for the late surgery reduced to 1.18 (95% CI
75%–79: 2955 (20.8) 491 (20.0) 0.885 0.770 to 1.017 0.086
80–84: 3483 (24.1) 616 (25.1) 0.954 0.833 to 1.092 0.404
85–89: 3379 (23.4) 556 (22.6) 0.867 0.753 to 0.998 <0.05
90+: 1848 (12.8) 288 (11.7) 0.836 0.706 to 0.989 <0.05

Sex
Male: 3443 (23.9) 696 (28.4) 1.000 0.878 to 1.050 0.827
Female: 10983 (76.1) 1759 (71.6) 0.844 0.762 to 0.935 <0.01

Fracture type
Pertrochanteric: 4529 (31.4) 694 (28.3) 1.000 0.878 to 1.050 0.827
Subtrochanteric: 835 (5.8) 165 (6.7) 1.000 0.878 to 1.050 0.827
Femoral neck: 9065 (62.8) 1596 (65.0) 0.844 0.762 to 0.935 <0.01

Medical history
Cancer: 1676 (11.6) 299 (12.2) 0.223 1.006 0.878 to 1.151 0.935
Diabetes: 1470 (10.2) 290 (11.8) <0.01 1.006 0.878 to 1.151 0.935
Dementia: 2761 (19.1) 411 (16.7) <0.01 0.863 0.762 to 0.976 <0.05
Hypertension: 2188 (15.2) 444 (18.1) <0.001 1.104 0.979 to 1.244 0.106
Cardiovascular disease: 5110 (35.4) 1108 (45.1) <0.001 1.462 1.329 to 1.609 <0.001
Cerebrovascular disease: 2807 (19.5) 562 (22.9) <0.001 1.110 0.995 to 1.238 0.062
Peripheral vascular disease: 632 (4.4) 174 (7.1) <0.001 1.447 1.207 to 1.736 <0.001
Chronic pulmonary disease: 1249 (8.7) 143 (9.9) 1.177 1.000 to 1.384 <0.05
Anemia: 1161 (8.0) 236 (9.6) <0.01 1.135 0.972 to 1.319 0.109
Diseases of the nervous system: 852 (5.9) 139 (5.7) <0.01 0.895 0.740 to 1.084 0.257
Eye disease*: 4182 (29.0) 813 (32.0) 0.038 0.903 0.816 to 1.000 0.049
Diseases of the digestive system: 1249 (8.8) 396 (16.1) 0.474 0.940 0.833 to 1.060 0.310
Other diseases**: 2429 (16.8) 432 (17.6) 0.696 0.877 0.710 to 1.084 0.184

Pre-admission residence
Home (>1 week): 9444 (65.5) 1559 (63.5) 1.000 0.878 to 1.151 0.935
Home (max 1 week): 685 (4.7) 153 (6.2) <0.001 1.006 0.878 to 1.151 0.935
Nursing home: 2508 (17.4) 416 (16.9) <0.001 0.863 0.762 to 0.976 <0.05
Health centre*: 1249 (8.7) 143 (9.9) 1.177 1.000 to 1.384 <0.05
Hospital: 540 (3.7) 84 (3.4) 0.888 0.694 to 1.135 0.341

>180 days inpatient care during year preceding fracture
2569 (17.8) 402 (16.4) <0.05 0.823 0.691 to 0.979 <0.05

Day of admission
Monday: 2479 (17.2) 307 (12.5) 1.063 1.006 0.878 to 1.151 0.935
Tuesday: 2289 (15.9) 288 (11.7) 1.026 0.864 to 1.218 0.770
Wednesday: 2248 (15.6) 220 (9.0) 0.803 0.668 to 0.964 <0.05
Thursday: 2091 (14.5) 401 (16.3) 1.567 1.334 to 1.840 <0.001
Friday: 1898 (13.2) 678 (27.6) 2.970 2.559 to 3.447 <0.001
Saturday: 1744 (12.1) 358 (14.6) 1.686 1.429 to 1.990 <0.001
Sunday: 1677 (11.6) 203 (8.3) 0.989 0.818 to 1.195 0.907

Year**
1998: 3275 (83.1) 667 (16.9) 1.000 0.878 to 1.151 0.935
1999: 3325 (84.8) 597 (15.2) 0.877 0.775 to 0.992 <0.05
2000: 3869 (86.1) 625 (13.9) 0.801 0.709 to 0.904 <0.001
2001: 3957 (87.5) 566 (12.5) 0.700 0.619 to 0.793 <0.001

*Includes Parkinson’s disease, hemiplegia or paraplegia, and epilepsy.
†Includes glaucoma and cataract.
‡Includes peptic ulcer disease, abdominal hernia, liver disease, disorders of gallbladder and biliary tract, and irritable colon.
¶Includes hypothyroidism, gout, renal disease, arthritis, and rheumatic disease.
*Health centre refers to the inpatient ward of the local primary health care unit.
**Percentages in parentheses are the proportions of the categories in the current year.

Discharge, admission between Thursday and Saturday, and co-morbidities including cardiovascular disease and peripheral vascular disease (and, prior to adjustment, diabetes, cerebrovascular disease, chronic pulmonary disease, hypertension, and anaemia). The proportion of patients having late surgery substantially decreased during the period of the study.

Adjusted effect of operative delay on mortality
The hazard ratio for the late surgery reduced to 1.18 (95% CI 1.09 to 1.28, p<0.0001) after adjusting for the provider and the characteristics in table 1. Even after controlling for the unobserved covariates, the hazard ratio remained significant, indicating an increased mortality for late surgery. However, after the pseudo randomised assignment of patients into early and late surgery groups using admission day of the week as an instrumental variable, the difference in the adjusted 1 year mortality rate was found not to be attributable to operative delay (p = 0.069).

Proportions of late surgery patients
There was extensive variation in the proportions of late surgery patients between providers (fig 3). The conservative 95% confidence interval crosses the potentially achievable
The unadjusted excess mortality for late surgery patients increased during the year following the operation. It is interesting that the short term mortality differed only slightly between early and late surgery groups, but the difference became much clearer after the perioperative period. This may be partly due to the fact that almost half the perioperative deaths are unavoidable in an unselected population. If patients are going to die shortly after the fracture—regardless of the quality of the hip fracture treatment provided—it is obvious that the treatment effect turns out to be small. The treatment effect becomes observable for long term mortality but is potentially biased because of heterogeneity between the early and late surgery groups in patient characteristics affecting the outcome.

Several patient characteristics differed between the early and late surgery groups, but the prolonged operative delay seemed to increase the (long term) mortality significantly even after adjustment for these observed patient characteristics. There were no detailed clinical data available which obviously makes the adjustments only partial. The bias attributable to the unobserved covariates was controlled by allowing extra variation in the model, but the treatment effect remained significant. However, this kind of model is still prone to bias if the relationships between the risk factors and mortality are not correctly specified in the model. The pseudo randomizing instrumental variable approach was therefore applied and indicated that the increased mortality was not due to operative delay. Since the interpretations based on the instrumental variable approach also require strong and questionable assumptions, the sizeable difference in the estimates of the effect of operative delay on mortality between methods means that there is a need for hypotheses to explain such differences.

For instance, one could try to separate the acceptable delays from the unacceptable ones or one could record the reason for the delay in the data. In the strictest sense, this kind of approach needs the assumption that the acceptable and unacceptable delays can be defined and measured uniquely. In practice, an acceptable delay corresponds to a clinical decision to postpone the operation and the assumption is approximately fulfilled if the clinical practice remains constant. Because the data used in this study did not include the reason for the delay, the hypothesis of constant clinical practice was examined indirectly using provider level analyses.
The provider specific proportion of late surgery patients, adjusted for observed patient characteristics, showed that there was extensive variation between providers, which is a common finding. A simple performance assessment interpretation is that the percentage of late surgery patients can be reduced to a potentially achievable level that can also be interpreted as the upper limit for the proportion of acceptable delayed patients. Correspondingly, the expected proportion of unacceptable delayed patients is the proportion of late surgery patients exceeding this upper limit. This leads to the hypothesis that the overall mortality of hip fracture patients should increase with an increasing proportion of late surgery patients given that the longer operative delay would increase mortality. For groups 2–4 the interpretation is more difficult than for patients in the extreme groups. Since the mortality rate in early surgery patients was lower in the providers with a small proportion of late surgery patients and there were no differences in mortality rates among early surgery patients between providers. This means that, in this group, it is essential to perform early surgery in all patients who can withstand it because the significantly prolonged surgical delay makes the patient’s condition worse and increases mortality. For patients in the most severe group mortality was also higher for the late surgery patients for all providers. However, in this group the mortality rate in the early surgery patients was lower in the providers with a small proportion of late surgery patients and there were no differences in mortality rates among early surgery patients between providers. This means that, in this group, it is essential to perform early surgery in all patients who can withstand it because the significantly prolonged surgical delay makes the patient’s condition worse and increases mortality. For groups 2–4 the interpretation is more difficult than for patients in the extreme groups. Since the mortality rate in early surgery patients increases significantly and is higher even than the mortality rate of late surgery patients for providers with a large proportion of late surgery patients, it seems that in these groups an operation performed too early may cause more harm than the prolonged waiting time.

Other differences between providers in the proportion of patients with late surgery result mainly from lack of resources such as temporary unavailability of the operating rooms or surgeons and problems with obtaining medical clearance from other specialties in off hours. The importance of these factors can be illustrated using the effect of the admission day on the in-hospital operation delay; the mean proportion of patients undergoing late surgery would be reduced from 14.5% to 9.0% (p < 0.0001) if all patients had been admitted on “the best day”.

In spite of the large nationwide database and the advanced methods, the observational study design and administrative data require a special orientation to the analyses and caution should be exercised when interpreting the results. Data were available for 1998–2001 and the practices may have disproportionately changed for the different providers during

### Table 2: Effect of proportion of late surgery patients on mortality in different severity groups of hip fracture patients

<table>
<thead>
<tr>
<th>Patient group</th>
<th>10% of patients with late surgery</th>
<th>15% of patients with late surgery</th>
<th>20% of patients with late surgery</th>
<th>30% of patients with late surgery</th>
<th>Slope* (p value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>16881</td>
<td>14426 (83)</td>
<td>2455 (15)</td>
<td>31.9 30.5 29.7 27.9</td>
<td>0.5 (&lt;0.05)</td>
</tr>
<tr>
<td>Early surgery</td>
<td>2931 (87)</td>
<td>8.8 8.6 8.4 8.1</td>
<td>15.4 13.4 12.4 11.1</td>
<td>0.2 (0.232)</td>
<td></td>
</tr>
<tr>
<td>Late surgery</td>
<td>445 (13)</td>
<td>14.0 12.4 11.4 10.1</td>
<td>17.6 16.2 14.8 13.4</td>
<td>0.2 (0.168)</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>3376 (20)</td>
<td>26.9 27.2 27.6 28.2</td>
<td>0.6 (0.101)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early surgery</td>
<td>2875 (85)</td>
<td>26.6 25.5 25.3 24.3</td>
<td>0.9 (0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late surgery</td>
<td>501 (15)</td>
<td>29.7 28.5 28.4 28.1</td>
<td>1.9 (&lt;0.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>3377 (20)</td>
<td>37.9 38.3 38.7 39.5</td>
<td>0.8 (0.126)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early surgery</td>
<td>2867 (85)</td>
<td>37.4 38.3 39.2 41.1</td>
<td>1.8 (&lt;0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late surgery</td>
<td>508 (15)</td>
<td>43.0 36.7 35.8</td>
<td>0.5 (0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>3376 (20)</td>
<td>51.4 51.9 52.5 53.5</td>
<td>1.1 (&lt;0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early surgery</td>
<td>2819 (84)</td>
<td>50.7 50.6 50.6 50.5</td>
<td>0.1 (0.455)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late surgery</td>
<td>557 (17)</td>
<td>57.9 59.3 59.9 60.6</td>
<td>0.4 (0.26)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Change in percentage of deaths when the share of patients with late surgery increased 10%.
†Corresponds to fig 4.

For groups 2–4 the interpretation is more difficult than for patients in the extreme groups. Since the mortality rate in early surgery patients increases significantly and is higher even than the mortality rate of late surgery patients for providers with a large proportion of late surgery patients, it seems that in these groups an operation performed too early may cause more harm than the prolonged waiting time.

Other differences between providers in the proportion of patients with late surgery result mainly from lack of resources such as temporary unavailability of the operating rooms or surgeons and problems with obtaining medical clearance from other specialties in off hours. The importance of these factors can be illustrated using the effect of the admission day on the in-hospital operation delay; the mean proportion of patients undergoing late surgery would be reduced from 14.5% to 9.0% (p < 0.0001) if all patients had been admitted on “the best day”.

In spite of the large nationwide database and the advanced methods, the observational study design and administrative data require a special orientation to the analyses and caution should be exercised when interpreting the results. Data were available for 1998–2001 and the practices may have disproportionately changed for the different providers during
that time. The results must therefore be interpreted as the mean situation between the years 1998–2001 for each provider. Non-operated patients were excluded from the analyses since the operative delay was not defined uniquely for these patients. Inclusion of these patients would certainly increase the mortality rate since the typical non-operated patient is in a worse condition and dies before any operation can even take place. The trend analyses are intended to summarise the systematic parts of the variation between the providers. However, there seems to be some unexplained variation, especially in the proportions of late surgery patients, because of the small number of patients for some providers, so the results of the trend analyses need to be confirmed and validated in further studies.

Study designs which explicitly take into account the heterogeneity at the provider level are a prerequisite for unbiased results since the selection bias may vary between providers. One could also include the provider level characteristics in the analyses, but easily observable characteristics such as hospital type or number of surgeons typically allow only indirect interpretations. A more direct approach would be to observe the performance of providers continuously and ask the providers themselves to clarify the reasons for their performance.

In Finland the mean proportion of late surgery patients has decreased during the years of the study, but extensive variation between providers exists. In addition to possible system problems, there also seems to be variation in the clinical practices for judging which patients are suitable for early surgery. Identification of good practices and the release of a national clinical guideline for hip fracture treatment will probably improve the situation. In this sense, the proportion of patients with a prolonged waiting time for a hip fracture operation seems to be an effective evidence based quality indicator, since it clearly indicates an area for improvement. The effects of operative delay on mortality at the patient level are quite small, but at the provider level the association can even take place. The trend analyses are intended to summarise the systematic parts of the variation between the providers, so the results of the trend analyses need to be confirmed and validated in further studies.

Authors’ affiliations
S Sund, A Liski, National Research and Development Centre for Welfare and Health (STAKES), Outcomes and Equity Research, Helsinki, Finland

This research was supported by a research grant from the Academy of Finland.

REFERENCES


