Achieving best practice tariff may not reflect improved survival after hip fracture treatment

Sameer K Khan1
Mark DF Shirley2
Clare Glennie1
Paul V Fearon1
David J Deehan1
1The Newcastle upon Tyne Hospitals NHS Foundation Trust, 2School of Biology, Newcastle University, Newcastle upon Tyne, UK

Objective: The best practice tariff (BPT) incentivizes hospitals in the England and Wales National Health Service to provide multiprofessional care to patients with hip fractures. The initial six targets included: 1) admission under consultant-led joint orthopedic–geriatric care, 2) multidisciplinary assessment protocol on admission, 3) surgery within 36 hours, 4) geriatrician review within 72 hours, 5) multiprofessional rehabilitation, and 6) assessment for falls and bone protection. We aimed to examine the relationship between BPT achievement and important patient outcomes and whether the BPT could predict these independently of other validated predictors.

Materials and methods: A retrospective review was conducted on 516 patient episodes. Four outcomes were defined: 1) 30-day mortality, 2) 365-day mortality, 3) postoperative length of stay on trauma ward (LOS-T), and 4) total post-operative hospital LOS (LOS-H). The three groups did not differ significantly in baseline characteristics or outcomes. In the event analysis, the risk of 30-day mortality was related only to abnormal creatinine ($P=0.025$); mortality at 365 days was related significantly to low albumin ($P=0.023$) and weakly to abnormal creatinine ($P=0.089$). The risks of both increased LOS-T and LOS-H were related to age only ($P=0.052$, $P<0.001$, respectively).

Conclusion: Achieving BPT does not predict any outcome of interest on its own.

Keywords: hip fractures, best practice tariff, mortality

Introduction

The incidence of hip fractures in the UK is expected to rise to 100,000 in 2033.1 They remain associated with a high 90-day mortality, ranging from 12% to 18%.2,3 Recently published literature suggests that: 1) expediting surgery reduces length of stay (LOS) and reduces medical complications and mortality,4–6 and 2) investment in dedicated geriatric care for hip fracture patients improves their outcome.7,8 These and similar studies have underpinned the development of surgical and rehabilitation guidelines in the UK,9,10 with the Department of Health formally introducing a best practice tariff (BPT) for hip fractures in April 2010. The financial “incentive” to National Health Service (NHS) hospital trusts was an incremental payment of £445 (“differential”) over the base tariff earned for treating every hip fracture (BPT = differential + base tariff).11,12 This was conditional to fulfilling all of six defined criteria (Table 1). In 2011, the criteria were expanded to include two abbreviated mental test scores (before and after surgery, but within the same patient episode). The reimbursement does not
include the cost of the implants used. The BPT has stayed the same since, but the differential has increased relative to the base tariff: from £445 to £890 for 2011–2012, and then to £1,335 in 2012–2013.

NHS hospital trusts nationwide are now reporting increasingly higher compliance rates with targets set out in the BPT.\(^1\)\(^2\) Total payment to providers for hip fracture care in 2011–2012 was £532 million (US $883 million), of which the BPT element contributed £72 million (US $120 million; 14%). Most of the improvement nationwide has been in getting patients to theater within the 36-hour window mandated by the BPT, made possible by service expansion and procedural streamlining, eg, in the reorganization of orthogeriatric ward rounds, expedition of preoperative investigations, and prioritization of hip fractures on operating lists.\(^13\)\(^14\)

Performance-linked tariff modulation has been introduced in other countries.\(^15\)\(^16\) In 2004, Israel’s health ministry incorporated surgical delay into the reimbursement formula, with hospitals receiving higher remuneration for patients operated on within 48 hours.\(^15\) Within 2 years, there was a 35% increase in patients undergoing surgery within 48 hours, with a 30% reduction in inpatient mortality and a 2-day fall in the median LOS. A recent Italian study showed a similar increase in reimbursements for patients receiving surgery within 48 hours, but did not comment on outcomes.\(^16\)

The BPT introduced in England and Wales conditions higher reimbursement to not just one parameter (earlier surgery) but several other multidisciplinary facets of care. As such, it lends itself to scrutiny as a potential predictor of patient outcomes. The only reports to date show a variable effect on LOS, with no comment on survival.

We aimed to examine the relationship between achievement of this highly lucrative target and important patient outcomes (mortality and length of postoperative hospitalization) in our cohort and whether the BPT could predict these independently of other validated predictors. Our null hypothesis was that there was no difference for survival out to 365 days after surgery for those patients who did or did not achieve this metric. An a priori power analysis of a significant difference in 365-day mortality (\(\alpha=0.05\), power =0.8, assumed odds ratio =4) suggested 98 patients were needed per group, with total sample size (all groups) equal to 294. Mortality at 30 days and LOS were investigated as secondary outcomes.

**Materials and methods**

**Setting**

A retrospective observational study was designed at a major trauma center after our institution’s research ethics committee classified this study as an in-service evaluation of practice, which did not require patients’ individual consent. A total of 566 patients with hip fractures were admitted from the emergency department and treated surgically at its trauma unit from December 1, 2008 to May 31, 2011. Exclusions included: 1) age less than 60 years, 2) pathological fractures, 3) transfers to the elective unit for total-hip arthroplasty, 4) patients treated nonoperatively, and 5) episodes with incomplete biochemical data. Extracapsular fractures (Orthopedic Trauma Association [OTA] classification, types 31-A, 32-A1.1, 32-A2.1, and 32-A3.1) were stabilized with dynamic hip screws or proximal femoral nails. Intracapsular fractures (OTA type 31-B) were treated predominantly with cemented Thompson’s hemiarthroplasties, while a small proportion were fixed in situ with cannulated screws or dynamic hip screws. The day-to-day medical management of all patients was supervised by orthogeriatricians. All patients received intravenous fluids in the perioperative period. Only nonsteroidal anti-inflammatory drug-sparing analgesia was given, unless patients were already on such medication for rheumatoid disease or other related conditions. A uniform blood-transfusion policy remained in place during this study period. Urinary tract infections were treated with antibiotics as per hospital antimicrobial guidelines. The frequency of ward rounds by the orthogeriatricians and number of rehabilitation beds increased slightly during the last 3 months of the study, but this was classed as “service development” and was financed separately by the trust, not through the earning from BPT.
Data collection
Identifiers for these patients were linked to:
1) patients’ notes: age, sex, and American Society of Anesthesiologists (ASA) grade
2) online patient-management system: dates and times of accident and emergency admission, ward admission, surgery, discharge from the trauma ward and the hospital, and inpatient death
3) online pathology results system (ICET™; Sunquest Information Systems, Norwich, UK): hematological (hemoglobin), renal (serum creatinine), and nutritional (serum albumin) indices on admission; our institution’s laboratory reference-value range remained constant for these indices during the study period: hemoglobin 130–180 g/L for men, 115–160 g/L for women, creatinine 77–133 μmol/L, and albumin 35–50 g/L
4) data on BPT targets collected quarterly and submitted by orthogeriatric nurses to the National Hip Fracture Database (NHFD), since its introduction on April 1, 2010
5) Office of National Statistics (Cardiff, UK): mortality data extracted from medical certificates of cause(s) of death, censored by general practitioners for out-of-hospital deaths, censored to 1 year postoperatively.

The included patients were divided into three groups:
1) pre-BPT patients, 2) those achieving BPT (all criteria fulfilled), and 3) patient episodes that failed BPT (even if they missed out on one criterion). Four outcomes were defined: 1) mortality at 30 days, 2) mortality at 365 days, 3) postoperative LOS on trauma ward (LOS-T), and 4) postoperative LOS in hospital (LOS-H).

Statistical analysis
The unit of analysis was the patient episode, since a small number of patients were readmitted with contralateral hip fractures during the study. The three groups were compared for mortality at 30 and 365 days (binary data) with χ² tests, and for LOS-T and LOS-H (continuous data) using Kruskal–Wallis tests. Event analysis was then done for groups 2 and 3 using generalized linear modeling, with age, sex, ASA grade, hemoglobin (low or normal), albumin (low or normal), creatinine (high or normal), and BPT (achieved or failed) evaluated as predictors. All analyses were performed in R version 3.0.2. Significance was indicated by a P-value of less than or equal to 0.05, although predictors within the range of 0.1 ≤ P > 0.05 were reported for completeness.

Results
Complete data were available for 516 patients, including 233 (group 1) treated in the 16 months before the introduction of BPT, and 283 in the 13 months after. BPT was achieved in 183 patients (group 2, 64.6%), while 100 patients were classed as BPT fails (group 3). This earned the trust an extra £89,985 (US $151,450, 65%) out of a possible extra remuneration of £138,085 (US $232,400). Surgical (52%) and orthogeriatric (43%) delays were the predominant reasons for the BPT fails. There were no significant differences in the distribution of sex, ASA grades, and fracture types between the three groups (Table 2). Mortality was not significantly different between the three groups at 30 days (P=0.683) or at 1 year (P=0.166), nor was the distribution of the causes of death at 365 days (P=0.354). LOS-T was significantly shorter in BPT fails (P=0.005), but this difference was not seen for LOS-H.

In the event analysis for the 283 BPT-eligible patients, the risk of 30-day mortality was related only to abnormal creatinine at admission (z=2.241, P=0.025). Mortality at 365 days was related significantly to low albumin (z=2.272, P=0.023) and weakly to abnormal creatinine (z=1.701, P=0.089). Risks of both increased LOS-T and LOS-H were related to age only (t=1.945, P=0.052; t=4.067, P<0.001). Low albumin was an insignificant predictor (t=1.74, P=0.083) for LOS-H.

Discussion
There were four main findings: 1) renal function on admission predicts outcome, 2) albumin is a more reliable predictor of 1-year survival, 3) LOS is related to age only, and 4) achieving BPT did not predict any outcome of interest. Whereas previous studies have reported on the influence of performance-linked tariff modulation on inpatient mortality only, our study evaluated survival out to 1 year for patients achieving and failing BPT. We believe this to be the first published study highlighting the discord between individual achievement of BPT and patient outcome. The principal limitation of this work is the exclusion of some cases for missing data, but the methodology allows for assessment of the influence of BPT upon outcome independently of other predictors in a relatively large cohort comparable in size to studies in recent literature.

We found both renal function and nutritional status to be predictive for 30-day and 1-year mortality; these are both easily ascertained on admission with routine biochemical tests. This corroborates previously published reports of poorer survival out to 1 year in hypoalbuminemic patients, and
there may be an argument for “nutritional resuscitation” prior to surgery. Hip fracture patients are a heterogeneous case mix at presentation. BPT reflects the perioperative phase of their inpatient management only, without stratifying them based on their medical needs. It mandates surgery within 36 hours, though some patients (eg, poor nutrition, renal dysfunction) may clinically require a “period for resuscitation”, whereby hepatorenal function is optimized. Therefore, such tailored individual treatment, likely ordered by a senior competent physician, would necessitate a delay in surgery but a concomitant inadvertent financial penalty.

LOS is a multifactorial issue, influenced not only by the patient’s comorbidities but also by logistical factors, eg, the provision of supported discharge and the availability of step-down and rehabilitation beds in the same or different catchment areas. Increasing LOS-T and LOS-H were related only to advancing age in the event analysis. The finding of reduced LOS-T in BPT fails in the unadjusted analysis is counterintuitive. A couple of unpredictable occurrences contributed to delayed discharge in our BPT-achieving patients: 1) the outbreak of norovirus infection on a rehabilitation ward in December 2010 meant patients had to wait longer on the trauma ward, and 2) a prolonged harsh winter in 2010–2011 caused delays in ambulance transports and eligibility assessment visits by home care staff. Our experience suggests that LOS may not be an accurate measure of outcome.

Interestingly, subgroup analysis failed to show an adverse outcome for those patients who underwent surgery beyond 36 hours in this study. The study was powered for BPT, which incorporates six other domains in addition to surgery within 36 hours. Such an arbitrary metric may be viewed as naïve when in fact previous work has found that the timing of admission to the treating hospital better reflects outcome. In previous work, patients who presented to an admitting suite between midnight and 6 am compared unfavorably for outcome when compared to other admission times, irrespective of actual time from admission to theater. Furthermore, this 36-hour surgical target may not be discriminatory, as recent work confirms that “very early” surgery (within

<table>
<thead>
<tr>
<th>Table 2 Comparison of baseline characteristics and outcomes between the three groups described in the study</th>
<th>Group 1 Pre-BPT n=233</th>
<th>Group 2 BPT achievers n=183</th>
<th>Group 3 BPT fails n=100</th>
<th>Statistical test</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>82</td>
<td>85</td>
<td>85</td>
<td>Kruskal–Wallis</td>
<td>0.032</td>
</tr>
<tr>
<td>Range</td>
<td>60–99</td>
<td>(60–100)</td>
<td>(60–100)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fracture type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intracapsular:extracapsular</td>
<td>126:107</td>
<td>101:82</td>
<td>59:41</td>
<td></td>
<td>0.708</td>
</tr>
<tr>
<td><strong>ASA grade</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td></td>
<td>0.255</td>
</tr>
<tr>
<td>2</td>
<td>45</td>
<td>45</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>148</td>
<td>115</td>
<td>66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥4</td>
<td>38</td>
<td>18</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LOS-T (days)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>15</td>
<td>12</td>
<td>12</td>
<td></td>
<td>0.005</td>
</tr>
<tr>
<td>Range</td>
<td>1–141</td>
<td>2–83</td>
<td>1–46</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LOS-H (days)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>23</td>
<td>18</td>
<td>18</td>
<td></td>
<td>0.895</td>
</tr>
<tr>
<td>Range</td>
<td>3–378</td>
<td>2–134</td>
<td>2–106</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mortality, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 days</td>
<td>16 (6.9%)</td>
<td>9 (4.9%)</td>
<td>7 (6.9%)</td>
<td></td>
<td>0.683</td>
</tr>
<tr>
<td><strong>Mortality, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>365 days</td>
<td>52 (22.3%)</td>
<td>29 (15.9%)</td>
<td>19 (18.8%)</td>
<td></td>
<td>0.166</td>
</tr>
<tr>
<td><strong>Causes of 365-day mortality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respiratory</td>
<td>19</td>
<td>15</td>
<td>8</td>
<td></td>
<td>0.354</td>
</tr>
<tr>
<td>Cardiac</td>
<td>11</td>
<td>5</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Renal</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malignancy</td>
<td>9</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>12</td>
<td>6</td>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: ASA, American Society of Anesthesiologists; BPT, best practice tariff; LOS-T, length of stay on trauma ward; LOS-H, length of stay (total) in hospital.
12 hours) is associated with far fewer hospital deaths than the 36-hour watershed.24 Additionally, BPT must be validated prospectively for functional outcome and survivorship, given that recent work suggests that the risk of death from cardiorespiratory dysfunction persists out to 100 days from time of admission.25 Further work is required to quantify the impact of outpatient intervention by stratification of patient function at the time of discharge.

A 5% increase in patients achieving BPT translates into an increase of £2.3 million (US $3.8 million) in reimbursement to NHS hospitals,26 yet the estimated annual cost to NHS hospitals of implementing measures to help achieve BPT for all patients is enormous. For example, 6,084 extra theater sessions to help reduce surgical delay will amount to £36.8 million (US $62 million), while £4 million (US $6.7 million) is required to achieve perioperative orthogeriatric assessment in 90% of patients. The costs of additional physiotherapy and occupational therapy are estimated to be £235 (US $396) per patient for both therapies. These costs are offset partially by expected savings from reduced LOS. The cost of a “bed day” for hip fractures is currently £240–£245 (US $405–$412); analysts at the NHFD predict a saving of £5.8 million (US $9.8 million) per annum if the LOS were to be reduced by 2 days.26 Moreover, bone-protection medication and falls assessment may reduce the future incidence of further fragility-related fractures, but there are no data yet to quantify or substantiate this assumption. Therefore, while the reimbursements to NHS hospitals have increased annually, the annual reference costs are expected to reduce in the longer run, making BPT payments worthwhile from a societal and health-economic point of view.

BPT was introduced to reward best practice, which was defined as demonstrable adherence to defined criteria. The principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded and the “herd” effect is lost. Continued scrutiny and audit then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded and the “herd” effect is lost. Continued scrutiny and audit then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded and the “herd” effect is lost. Continued scrutiny and audit then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded and the “herd” effect is lost. Continued scrutiny and audit then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded and the “herd” effect is lost. Continued scrutiny and audit then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rewarded principle is sound, but if less important criteria are selected, then suboptimal practice may inadvertently be rendered as therapeutic interventions. The addition of abbreviated mental test-score assessment to BPT criteria in 2011 serves as a precedent for refining the BPT. We hypothesize that such an amended inpatient-management package would portend improved patient survival.

Disclosure
The authors report no conflicts of interest in this work.

References


