Influence of House-Staff Experience on Teaching-Hospital Mortality: The “July Phenomenon” Revisited

Carl van Walraven, MD, MSc, FRCPC¹,²,³, Alison Jennings, MA², Jenna Wong, BSc, MSc², Alan J. Forster, MD, MSc, FRCPC¹,²,³

¹Faculty of Medicine, University of Ottawa, Ontario, Canada; ²Ottawa Health Research Institute, Ottawa, Ontario, Canada; ³Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada.

BACKGROUND: The “July Phenomenon” refers to a purported worsening of outcomes in teaching-hospital patients with the arrival of new, inexperienced house-staff. Previous quantitative studies of new house-staff and increased mortality have been limited primarily by a focused patient population and the use of limited data to adjust for severity of patient illness.

METHODS: We included all medicine, surgical, and obstetrical patients admitted to a teaching hospital in Ontario, Canada between April 15, 2004 and December 31, 2008. We calculated the ratio of observed to expected weekly number of deaths in hospital. The expected number of deaths was calculated using a validated, discriminative, and well-calibrated multivariate survival model. Collective house-staff experience was modeled from a minimum on July 1st to a maximum on June 30th using five distinct patterns.

RESULTS: We studied 259,748 encounters that included 164,318 people. The mortality rate was 3.0%. The ratio of observed to expected number of weekly deaths was not associated with collective house-staff experience, irrespective of the pattern in which it was modeled. The lack of association between risk of death in hospital and house-staff experience did not vary by admission type (urgent vs elective) or specialty (medicine vs surgery).

CONCLUSION: At our hospital, we found no association between the arrival of new house-staff and the adjusted risk of death in hospital. These data, along with the results of the vast majority of previous studies in this field, make the existence of the “July Phenomenon” for inpatient mortality extremely unlikely. Journal of Hospital Medicine 2011;6:389–394. © 2011 Society of Hospital Medicine
residents starting their first year of postgraduate training in one of 29 programs.

Inclusion Criteria
The study period extended from April 15, 2004 to December 31, 2008. We used this start time because our hospital switched to new coding systems for procedures and diagnoses in April 2002. Since these new coding systems contributed to our outcome statistic, we used a very long period (ie, two years) for coding patterns to stabilize to ensure that any changes seen were not a function of coding patterns. We ended our study in December 2008 because this was the last date of complete data at the time we started the analysis.

We included all medical, surgical, and obstetrical patients admitted to TOH during this time except those who were: younger than 15 years old; transferred to or from another acute care hospital; or obstetrical patients hospitalized for routine childbirth. These patients were excluded because they were not part of the multivariate model that we used to calculate risk of death in hospital (discussed below).24 These exclusions accounted for 25.4% of all admissions during the study period (36,820—less than 15 years old; 12,931—transferred to or from the hospital; and 44,220—uncomplicated admission for childbirth).

All data used in this study came from The Ottawa Hospital Data Warehouse (TOHDW). This is a repository of clinical, laboratory, and administrative data originating from the hospital’s major operational information systems. TOHDW contains information on patient demographics and diagnoses, as well as procedures and patient transfers between different units or hospital services during the admission.

Primary Outcome—Ratio of Observed to Expected Number of Deaths Per Week
For each study day, we measured the number of hospital deaths from the patient registration table in TOHDW. This statistic was collated for each week to ensure numeric stability, especially in our subgroup analyses.

We calculated the weekly expected number of hospital deaths using an extension of the Escobar model.24 The Escobar is a logistic regression model that estimated the probability of death in hospital that was derived and internally validated on almost 260,000 hospitalizations at 17 hospitals in the Kaiser Permanente Health Plan. It included six covariates that were measurable at admission including: patient age; patient sex; admission urgency (ie, elective or emergent) and service (ie, medical or surgical); admission diagnosis; severity of acute illness as measured by the Laboratory-based Acute Physiology Score (LAPS); and chronic comorbidities as measured by the CoMorbidity Point Score (COPS). Hospitalizations were grouped by admission diagnosis. The final model had excellent discrimination (c-statistic 0.88) and calibration (P value of Hosmer Lemeshow statistic for entire cohort 0.66). This model was externally validated in our center with a c-statistic of 0.901.25

We extended the Escobar model in several ways (Wong et al., Derivation and validation of a model to predict the daily risk of death in hospital, 2010, unpublished work). First, we modified it into a survival (rather than a logistic) model so it could estimate a daily probability of death in hospital. Second, we included the same covariates as Escobar except that we expressed LAPS as a time-dependent covariate (meaning that the model accounted for changes in its value during the hospitalization). Finally, we included other time-dependent covariates including: admission to intensive care unit; undergoing significant procedures; and awaiting long-term care. This model had excellent discrimination (concordance probability of 0.895, 95% confidence interval [CI] 0.889–0.902) and calibration.

We used this survival model to estimate the daily risk of death for all patients in the hospital each day. Summing these risks over hospital patients on each day returned the daily number of expected hospital deaths. This was collated per week.

The outcome statistic for this study was the ratio of the observed to expected weekly number of hospital deaths. Ratios exceeding 1 indicate that more deaths were observed than were expected (given the distribution of important covariates in those people during that week). This outcome statistic has several advantages. First, it accounts for the number of patients in the hospital each day. This is important because the number of hospital deaths will increase as the number of people in hospital increase. Second, it accounts for the severity of illness in each patient on each hospital day. This accounts for daily changes in risk of patient death, because calculation of the expected number of deaths per day was done using a multivariate survival model that included time-dependent covariates. Therefore, each individual’s predicted hazard of death (which was summed over the entire hospital to calculate the total expected number of deaths in hospital each day) took into account the latest values of these covariates. Previous analyses only accounted for risk of death at admission.

Expressing Physician Experience
The latent measure26 in all “July Phenomenon” studies is collective house-staff physician experience. This is quantified by a surrogate date variable in which July 1—the date that new house-staff start their training in North America—represents minimal experience and June 30 represents maximal experience. We expressed collective physician experience on a scale from “0” (minimum experience) on July 1 to “1” (maximum experience) on June 30. A similar approach has been used previously1-3 and has advantages over the other methods used to capture...
collective house-staff experience. In the “stratified, incomplete” approach, periods with inexperienced house-staff (eg, July and August) are grouped together and compared to times with experienced house-staff (eg, May and June), while ignoring all other data. The specification of cut-points for this stratification is arbitrary and the method ignores large amounts of data. In the “stratified, complete” approach, periods with inexperienced house-staff (eg, July and August) are grouped together and compared to all other times of the year. This is potentially less biased because there are no lost data. However, the cut-point for determining when house-staff transition from “inexperienced” to “experienced” is arbitrary, and the model assumes that the transition is sudden. This is suboptimal because acquisition of experience is a gradual, constant process.

The pattern by which collective physician experience changes between July 1st and June 30th is unknown. We therefore expressed this evolution using five different patterns varying from a linear change to a natural logarithmic change (see Supporting Appendix B in the online version of this article).

Analysis
We first examined for autocorrelation in our outcome variable using Ljung-Box statistics at lag 6 and 12 in PROC ARIMA (SAS 9.2, Cary, NC). If significant autocorrelation was absent in our data, linear regression modeling was used to associate the ratio of the observed to expected number of deaths (the outcome variable) with the collective first year physician experience (the predictor variable). Time-series methodology was to be used if significant autocorrelation was present.

In our baseline analysis, we included all hospitalizations together. In stratified analyses, we categorized hospitalizations by admission status (emergent vs elective) and admission service (medicine vs surgery).

Results
Between April 15, 2004 and December 31, 2008, The Ottawa Hospital had a total of 152,017 inpatient admissions and 107,731 same day surgeries (an annual rate of 32,222 and 22,835, respectively; an average daily rate of 88 and 63, respectively) that met our study’s inclusion criteria. These 259,748 encounters included 164,318 people. Table 1 provides an overall description of the study population.

Weekly Deaths: Observed, Expected, and Ratio
Figure 1A presents the observed number of deaths during the study period. There was an average of 31 deaths per week (range 15–51). Some large fluctuations in the weekly number of deaths were seen; in 2007, for example, the number of observed deaths went from 21 in week 13 up to 46 in week 15. However, no obvious seasonal trends in the observed weekly number of deaths were seen (Figure 1A, heavy line) nor were trends between years obvious.

Figure 1B presents the expected weekly number of deaths during the study period. The expected weekly number of deaths averaged 29.6 (range 22.2–38.7). The expected weekly number of deaths was notably less variable than the observed number of deaths. However, important variations in the expected number of deaths were seen; for example, in 2005, the expected number of deaths increased from 24.1 in week 41 to 29.6 in week 44. Again, we saw no obvious seasonal trends in the expected weekly number of deaths (Figure 1B, heavy line).

Figure 1C illustrates the ratio of observed to the expected weekly number of deaths. The average observed to expected ratio slightly exceeded unity (1.05) and ranged from 0.488 (week 24, in 2008) to 1.821 (week 51, in 2008). We saw no obvious seasonal trends in the ratio of the observed to expected number of weekly deaths. In addition, obvious trends in this ratio were absent over the study period.

Association Between House-Staff Experience and Death in Hospital
We found no evidence of autocorrelation in the ratio of observed to expected weekly number of deaths. The ratio of observed to expected number of hospital deaths was not significantly associated with house-staff physician experience (Table 2). This conclusion did not change regardless of which house-staff physician experience pattern was used in the linear model (Table 2). In addition, our analysis found no significant association between physician experience and patient mortality when analyses were stratified by admission service or admission status (Table 2).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n</th>
<th>Patients/hospitalizations, n</th>
<th>164,318/259,748</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deaths in-hospital, n (%)</td>
<td>7,679 (3.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length of admission in days, median (IQR)</td>
<td>2 (1–6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>124,848 (48.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at admission, median (IQR)</td>
<td>60 (46–74)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Admission type, n (%)</td>
<td>136,406 (52.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elective surgical</td>
<td>20,104 (7.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elective nonsurgical</td>
<td>71,192 (27.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergent surgical</td>
<td>7,779 (3.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergent nonsurgical</td>
<td>6,971 (2.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elixhauser score, median (IQR)</td>
<td>0 (0–4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAPIS at admission, median (IQR)</td>
<td>0 (0–15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At least one admission to intensive care unit, n (%)</td>
<td>29.6 (18.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At least one PIMR procedure, n (%)</td>
<td>−2 (−5–2)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: IQR, interquartile range; LAPIS, Laboratory-based Acute Physiology Score; PIMR, Procedural Independent Mortality Risk. The Procedural Independent Mortality Risk (PIMR) score can use administrative data to quantify the independent risk of death in hospital after procedures, 2010, unpublished work. Among admissions where at least one PIMR procedure was performed during the hospitalization.
Discussion

It is natural to suspect that physician experience influences patient outcomes. The commonly discussed “July Phenomenon” explores changes in teaching-hospital patient outcomes by time of the academic year. This serves as an ecological surrogate for the latent variable of overall house-staff experience. Our study used a detailed outcome—the ratio of observed to the expected number of weekly hospital deaths—that adjusted for patient severity of illness. We also modeled collective physician experience using a broad range of patterns. We found no significant variation in mortality rates during the academic year; therefore, the risk of death in hospital does not vary by house-staff experience at our hospital. This is no evidence of a “July Phenomenon” for mortality at our center.

We were not surprised that the arrival of inexperienced house-staff did not significantly change patient mortality for several reasons. First year residents are but one group of treating physicians in a teaching hospital.
hospital. They are surrounded by many other, more experienced physicians who also contribute to patient care and their outcomes. Given these other physicians, the influence that the relatively smaller number of first year residents have on patient outcomes will be minimized. In addition, the role that these more experienced physicians play in patient care will vary by the experience and ability of residents. The influence of new and inexperienced house-staff in July will be blunted by an increased role played by staff-people, fellows, and more experienced house-staff at that time.

Our study was a methodologically rigorous examination of the “July Phenomenon.” We used a reliable outcome statistic—the ratio of observed to expected weekly number of hospital deaths—that was created with a validated, discriminative, and well-calibrated model which predicted risk of death in hospital (Wong et al., Derivation and validation of a model to predict the daily risk of death in hospital, 2010, unpublished work). This statistic is inherently understandable and controlled for patient severity of illness. In addition, our study included a very broad and inclusive group of patients over five years at two hospitals.

Twenty-three other studies have quantitatively sought a “July Phenomenon” for patient mortality (see Supporting Appendix A in the online version of this article). The studies contained a broad assortment of research methodologies, patient populations, and analytical methodologies. Nineteen of these studies (83%) found no evidence of a “July Phenomenon” for teaching-hospital mortality. In contrast, two of these studies found notable adjusted odds ratios for death in hospital (1.41 and 1.34) in patients undergoing either general surgery or complex cardiovascular surgery, respectively. Blumberg also found an increased risk of death in surgical patients in July, but used indirect standardized mortality ratios as the outcome statistic and was based on only 139 cases at Maryland teaching hospitals in 1984. Only Jen et al. showed an increased risk of hospital death with new house-staff in a broad patient population. However, this study was restricted to two arbitrarily chosen days (one before and one after house-staff change-over) and showed an increased risk of hospital death (adjusted OR 1.05, 95% CI 1.00–1.15) whose borderline statistical significance could have been driven by the large sample size of the study (n = 299,741).

Therefore, the vast majority of data—including those presented in our analyses—show that the risk of teaching-hospital death does not significantly increase with the arrival of new house-staff. This prompts the question as to why the “July Phenomenon” is commonly presented in popular media as a proven fact. We believe this is likely because the concept of the “July Phenomenon” is understandable and has a rather morbid attraction to people, both inside and outside of the medical profession. Given the large amount of data refuting the true existence of a “July Phenomenon” for patient mortality (see Supporting Appendix A in the online version of this article), we believe that this term should only be used only as an example of an interesting idea that is refuted by a proper analysis of the data.

Several limitations of our study are notable. First, our analysis is limited to a single center, albeit with two hospitals. However, ours is one of the largest teaching centers in Canada with many new residents each year. Second, we only examined the association of physician experience on hospital mortality. While it is possible that physician experience significantly influences other patient outcomes, mortality is, obviously, an important and reliably tallied statistic that is used as the primary outcome in most “July Phenomenon” studies. Third, we excluded approximately a quarter of all hospitalizations from the study. These exclusions were necessary because the Escobar model does not apply to these people and can therefore not be used to predict their risk of death in hospital. However, the vast majority of excluded patients (those less than 15 years old, and women admitted for routine childbirth) have a very low risk of death (the former because they are almost exclusively newborns, and the latter because the risk of maternal death during childbirth is very low). Since these people will contribute very little to either the expected or observed number of deaths, their exclusion will do little to threaten the study’s validity. The remaining patients who were transferred to or from other hospitals (n = 12,931) makes a small proportion of the total sampling frame (~5% of admissions). Fourth, our study did not identify any significant association between house-staff experience and patient mortality (Table 2). However, the confidence intervals around our estimates are wide enough, especially in some subgroups such as patients admitted electively, that important changes in patient mortality with house-staff experience cannot be excluded. For example, whereas our study found that a decrease in the ratio of observed to expected number of deaths exceeding 30% is very unlikely, it is still possible that this decrease is up to 30% (the lower range of the confidence interval in Table 2). However, using this logic, it could also increase by up to 10% (Table 2). Finally, we did not directly measure individual physician experience. New residents can vary extensively in their individual experience and ability. Incorporating individual physician measures of experience and ability would more reliably let us measure the association of new residents with patient outcomes. Without this, we had to rely on an ecological measure of physician experience—namely calendar date. Again, this method is an “industry standard” since all studies quantify physician experience ecologically by date (see Supporting Appendix A in the online version of this article).
In summary, our data—similar to most studies on this topic—show that the risk of death in teaching hospitals does not change with the arrival of new house-staff.

Disclosure: This study was supported by the Department of Medicine, University of Ottawa. All authors disclose no relevant conflicts of interest.

References


