The Potential Benefits of Universal Adoption

November 2000

Research Director
John D. Birkmeyer, MD

Project Manager
Christian M. Birkmeyer

Consultants
David E. Wennberg, MD, MPH
Michael Young, MD, MS
Leapfrog Safety Standards: The Potential Benefits of Universal Adoption

Research Director
John D. Birkmeyer, MD
Assistant Professor of Surgery and Community & Family Medicine
Dartmouth Medical School

Project Manager
Christian M. Birkmeyer
Data Manager, Department of Surgery
Dartmouth Medical School

Consultants
David E. Wennberg, MD, MPH
Director, Center for Outcomes Research & Evaluation
Maine Medical Center

Michael Young, MD, MS
Assistant Professor of Medicine (Critical Care)
University of Vermont Medical Center

November 2000
The Business Roundtable is an association of chief executive officers of leading corporations with a combined workforce of more than 10 million employees in the United States. The chief executives are committed to advocating public policies that foster vigorous economic growth; a dynamic global economy; and a well-trained and productive U.S. workforce essential for future competitiveness.
# Table of Contents

**Executive Summary**  
1

**Computer physician order entry (CPOE)**  
Summary of analysis  
3  
Selected bibliography  
5  
Figure 1: Calculating serious medication errors avoided  
6  
Table 1: Studies assessing the effectiveness of CPOE  
7  
Figure 2: Sensitivity analysis  
7

**Evidence-based hospital referral (EHR): High-risk surgery**  
Summary of analysis  
8  
Selected bibliography  
11  
Table 1: Number of patients currently undergoing 5 selected procedures  
13  
Table 2: Predicted mortality rates  
14  
Table 3: Number of lives potentially saved by EHR  
14

**Evidence-based hospital referral (EHR): High-risk neonatal intensive care**  
Summary of analysis  
15  
Selected bibliography  
17  
Table 1: Number of high-risk deliveries potentially affected by EHR  
18  
Table 2: Predicted mortality rates  
18  
Table 3: Number of lives potentially saved by EHR  
19  
Figure 1: Sensitivity analysis  
19

**ICU Physician Staffing (IPS)**  
Summary of analysis  
20  
Selected bibliography  
22  
Figure 1: Number of lives potentially saved by IPS  
24  
Table 1: Pre/Post studies of the effectiveness of IPS  
25  
Table 2: Cross-sectional studies of the effectiveness of IPS  
26  
Figure 3: Sensitivity analysis  
26

**Appendix**  
Section 1: Studies assessing the relationship between volume and outcome for 5 selected procedures.  
27  
Table 1: Coronary artery bypass graft surgery (CABG)  
27  
Table 2: Coronary angioplasty  
27  
Table 3: Elective abdominal aortic aneurysm (AAA) repair  
28  
Table 4: Carotid endarterectomy  
28  
Table 5: Esophageal cancer surgery  
29  
Selected bibliography  
29  
Section 2: Medline search strategy for IPS  
31  
Acknowledgments  
32
EXECUTIVE SUMMARY

In this monograph, we describe the benefits that could be achieved if every non-rural hospital in America implemented the three hospital safety initiatives selected by The Leapfrog Group: 1) computer physician order entry (CPOE); 2) evidence-based hospital referral (EHR) for high-risk surgery and neonatal intensive care; and 3) ICU physician staffing (IPS). The Leapfrog Group is a Business Roundtable-sponsored initiative of Fortune 500 companies and other large private and public sector health care purchasers working together to drive big leaps in patient safety.

We used the same basic analysis strategy for each of the 3 safety standards. We first estimated the “population at risk”—the number of patients who are currently receiving care in suboptimal conditions and thus stand to benefit from changes imposed by Leapfrog. To avoid access issues and other unintended consequences, The Leapfrog Group is exempting hospitals in rural areas. Thus, the population at risk is restricted to patients in metropolitan areas. Second, we estimated baseline risks (of medication errors or mortality) in hospitalized patients, and the potential risk reductions associated with each of the safety standards. The following Table summarizes the results of our baseline analysis:

<table>
<thead>
<tr>
<th>Safety Initiative</th>
<th>Potential benefit with full implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPOE</td>
<td>522,000 serious medication errors avoided</td>
</tr>
<tr>
<td>EHR</td>
<td></td>
</tr>
<tr>
<td>Five high-risk procedures</td>
<td>2,581 lives saved</td>
</tr>
<tr>
<td>High-risk deliveries</td>
<td>1,863 lives saved</td>
</tr>
<tr>
<td>IPS</td>
<td>53,850 lives saved</td>
</tr>
</tbody>
</table>

Although our analysis is based on the best information currently available, we acknowledge gaps in scientific knowledge in the context of our literature synthesis and consider policy challenges. However, despite an imperfect knowledge base and implementation challenges, we believe that successful adoption of The Leapfrog Group’s three safety initiatives would significantly reduce America’s large annual toll of avoidable hospital death and patient suffering.

Computer physician order entry

In our baseline analysis, we estimated that implementation of CPOE in every non-rural hospital in the U.S. would avert approximately 522,000 serious medication errors each year in the United States. Of course, our estimates were influenced strongly by assumptions about CPOE effectiveness. Because of the relatively few careful studies in this area, our analysis relied on two well-recognized trials from a single teaching hospital. Some may question the validity of generalizing these data to other hospitals nationwide. However, we chose the most conservative estimate of CPOE effectiveness (55% medication error reduction rate) for our baseline analysis.

Although a large proportion of serious medical errors is life threatening, the number that result in fatalities cannot be determined precisely from the medical literature. Accordingly, we did not calculate the number of deaths potentially avoided by CPOE. However, if only 0.1% of such errors were fatal, over 500 deaths would be avoided every year. If the fatality rate were 1%, over 5000 deaths would be avoided.
Evidence-based hospital referral: High-risk surgery

In our baseline analysis, we estimated that full implementation of EHR would save 2,581 lives each year in the US. For assumptions about EHR effectiveness for each procedure, we used estimates from the structured literature review by Dudley et al., the most rigorous synthesis currently available. The greatest number of deaths would be prevented by evidence-based hospital referrals for coronary artery bypass graft surgery (CABG) (1,486 deaths), followed by elective abdominal aortic aneurysm (AAA) repair (464 deaths), and coronary angioplasty (345 deaths). Potential lives saved with esophagectomy and carotid endarterectomy were 168 and 118, respectively.

Our analysis estimates the benefits that could be achieved with full adherence to Leapfrog volume standards in all US (metropolitan) hospitals. The other two Leapfrog safety initiatives—CPOE and IPS—involve “all-or-none” hospital interventions. Making these changes for Leapfrog employees implies their availability to all other patients at the same hospitals. In contrast, even if EHR could be increased for Leapfrog employees, there would be no mechanism for assuring the same change in referral pattern for other patients. For this reason, a very important contribution of the Leapfrog safety initiative may occur by simply increasing public awareness of the importance of volume for selected high-risk procedures.

Evidence-based hospital referral: High-risk neonatal intensive care

Full implementation of EHR for high-risk deliveries would save 1,863 babies’ lives each year in the US: 1,369 lives for deliveries involving very low birth weight babies and 494 lives for deliveries involving babies with major congenital anomalies. Compared to the extensive scientific literature supporting EHR for high-risk surgery, however, clinical and health services research describing outcomes in neonatal ICUs is scant. By necessity, our analysis relied primarily on a single, well-recognized study based on 1990 California data by Phibbs et al. Thus, it will be important to update our analysis when more information is available about two key parameters: 1) the settings in which high-risk babies are being delivered in other regions of the US, and 2) the “volume-outcome” effect in neonatal intensive care and thus the value of moving high-risk deliveries to high volume, regional centers.

ICU physician staffing

In our baseline analysis, we estimated that full implementation of IPS would save approximately 53,850 lives each year in the US. IPS is so effective because such a large number of people die in ICUs each year (approximately one half million). Thus, even small improvements in ICU mortality rates save many lives. Although all nine studies assessing IPS to date suggest mortality reductions, all have limitations (related to their observational designs) and there may be problems with extrapolating their results nationwide. For this reason, we selected from the nine studies the most conservative estimate of IPS effectiveness (15% relative mortality reduction) for our baseline analysis.

Although workforce issues have not been studied carefully, it is unlikely that there are currently enough board-certified intensivists to fully staff ICUs at all hospitals. In hospitals with small units, meeting the Leapfrog daytime intensivist staffing standard may increase net cost per stay. For these reasons, broad implementation of intensivist model ICU staffing may require a mixture of increased fellowship training slots in critical care, consolidation of small ICUs, and advances in ICU telemedicine.
Computer physician order entry

Overview

In this section, we estimate the number of serious medication errors that could be averted by implementation of computer physician order entry (CPOE) nationwide. As defined by The Leapfrog Group, hospitals meeting the CPOE standard will require physicians to enter medication orders via computers linked to prescribing error prevention software, demonstrate that their CPOE system intercepted at least 50% of common serious prescribing errors, utilizing test cases and a testing protocol specified by First Consulting Group and the Institute for Safe Medication Practices (ISMP), require documented acknowledgement by the prescribing physician of the interception prior to any override, and post the test case interception rate on a Leapfrog-designated web site.

In our baseline analysis, we estimate that implementation of CPOE would avert approximately 522,000 serious medication errors each year in the United States (Figure 1). Since the proportion of serious medical errors that result in fatality cannot be determined precisely from the medical literature, we did not credit CPOE with any avoided deaths. However, if only 0.1% of such errors were fatal, over 500 deaths would be avoided by CPOE every year. If the fatality rate were 1%, over 5000 deaths would be avoided every year. In the following sections, we describe the methods and assumptions used in our analysis.

Methods and assumptions

As summarized in Figure 1, we calculated the number of serious medication errors potentially averted by full implementation of CPOE. We started with the population of inpatients who stand to benefit by the policy. We then estimated their baseline risks of serious medication errors and reductions expected with CPOE.

Number of patients currently admitted at hospitals without CPOE. To calculate the number of errors that could be averted with CPOE, we first estimated the size of the population at risk. According to the National Center for Health Statistics,1 35,406,187 patients were admitted to non-federal, acute care hospitals in 1997. To avoid access issues in rural areas, The Leapfrog Group is restricting CPOE, along with the other safety initiatives, to metropolitan areas. According to data from the 1996 AHA file and the US census, approximately 80% were admitted to hospitals in urban areas (defined by MSAs). Only one study by Ash et al. has assessed the proportion of hospitals with CPOE.2 Although 32% of responding hospitals had CPOE completely or partially available, fewer than 5% of these (1.6% overall) required its usage. Thus for our baseline analysis, we assumed that 98% of inpatients are currently treated at hospitals not meeting the Leapfrog standard for CPOE. Mean hospital length of stay in the US was 4.9 days in 1997;4 we assumed 4.5 days for this analysis, in keeping with national trends toward shorter LOS in general.

Baseline rate of serious medication errors. A serious medication error is a non-intercepted error in the process of ordering, dispensing, or administering a medication which causes or has the potential to cause an adverse drug event.3 In two studies by Bates et al at a Boston teaching hospital,3,4 such errors occurred at a rate of 10.7 and 7.6 per 1,000 pt-days. Expressed in terms of incidence rates per admission, 5.1% and 3.4%, respectively, of hospitalized patients experienced
at least one serious medication error. Similarly, Lesar et al documented serious medication errors in 5.3% of 211,655 patients admitted over nine years in a New York teaching hospital. In our baseline analysis, we assumed an error rate of 7.6 per 1,000 pt-days. Thus, we estimate that approximately 949,339 serious medication errors occur every year in US hospitals without CPOE.

**Efficacy of CPOE.** We identified only 2 studies—both by Bates et al—assessing the effectiveness of CPOE (Table 1). In the first cohort study of over 2,000 admissions at the Brigham and Women’s hospital in Boston, serious medication errors fell from 10.7 to 4.9 per 1000 patient-days after implementation of CPOE (55% reduction). In their second study (based on different study sample and later generation software), the proportion of patients experiencing serious medication errors fell from 7.6 to 1.1 per 1000 pt-days (88% reduction). To be conservative in our baseline calculations, we used the lower, earlier version estimate of CPOE efficacy (55%).

**Results**

In our baseline analysis, we estimate that full implementation of CPOE would avert approximately 522,000 serious medication errors each year in the US. As expected, the number of errors avoided varied according to assumptions about the effectiveness of CPOE. (Figure 2) Although we were conservative in our baseline analysis (55% reduction in error rate with CPOE), assuming higher levels of effectiveness would have significantly increased our estimates of serious medication errors averted.

**Cautions and policy considerations**

As illustrated in sensitivity analysis, our estimates are influenced strongly by assumptions about CPOE effectiveness. What proportion of serious medication errors currently occurring in US hospitals would be avoided by this decision support technology? Unfortunately, there have been relatively few careful studies of this question; our analysis was restricted to two rigorous studies from a single teaching hospital.

The effectiveness of CPOE may be influenced by the characteristics of the software employed and each hospital’s implementation skill. However, the Leapfrog performance testing requirement should significantly reduce the risk of overestimating the benefit of CPOE arising from those sources of uncertainty.

On a whole, however, the quality of CPOE, like most information technology applications, is likely increasing over time. For example, in their first trial conducted 1993-95, Bates et al. reported that CPOE reduced the serious medication error rate by 55% (the value used in our baseline analysis). In a subsequent study employing updated software, the same investigators reported 88% error reductions with CPOE.

Our estimates of the impact of implementing CPOE nationwide depend on assumptions about the baseline medication error rates, as well as CPOE effectiveness. We used the lower estimate of this parameter (7.6 per 1000 pt-days) from the two studies by Bates et al, each performed at the Brigham & Women’s hospital in Boston. Some would question the validity of generalizing error rates from a single teaching institution to other hospitals nationwide. Because of the relatively complex case-mix at Brigham & Women’s, some would argue that its baseline
medication error rate (and the apparent efficacy of CPOE) may be higher than the average rate at other hospitals. On the other hand, Brigham & Women’s has a reputation for excellence in its staff and housestaff (top ten in America’s Best Hospitals) and thus may have lower than expected medication error rates.

Although CPOE would avert more than a half million serious medication errors each year in the US, the number of lives that would be saved is difficult to determine. In the two studies by Bates et al. $^{3,4}$ more than half of all serious medication errors resulted in preventable adverse drug events. Approximately 20% of preventable adverse drug events were considered “life threatening” upon clinical review, but no patient in the two studies died as a direct result of a medication error. The two studies lacked sufficient sample size to detect a small but clinically meaningful reduction in mortality rates with CPOE. Large, multi-center studies will ultimately be needed to better characterize the impact of CPOE on mortality.

**Selected Bibliography:**


**Figure 1**: Calculating the number of serious medication errors avoided each year by computer physician order entry (CPOE).

1. **Annual Admissions To Non-Federal, Acute Care Hospitals**  
   35,406,187

2. **80%**

3. **Annual Admissions To Hospitals in Urban Areas**  
   28,324,950

4. **98%**

5. **Annual Hospital Admissions at Hospitals w/out CPOE**  
   27,758,451

6. \[ \text{x mean LOS} \]  
   4.5 days

7. **Annual Pt-Days at Hospitals w/out CPOE**  
   124,913,028

8. \[ \text{x 7.6 per 1000 days} \]

9. **Annual number of serious medication errors**  
   949,339

10. **55%**

11. **Serious medication errors avoided because of CPOE**  
   522,136
Table 1: Studies assessing the effectiveness of computer physician order entry (CPOE) in reducing serious medication errors. A serious medication error is an error in the process of ordering, dispensing, or administering a medication, which causes or has the potential to cause an adverse drug event; it does not include intercepted potential ADEs. Both studies were conducted at Brigham and Women’s Hospital, Boston, MA. 3,4

<table>
<thead>
<tr>
<th>Study Participants</th>
<th>Rate of Serious Medication Error (per 1000 pt-days)</th>
<th>RISK REDUCTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEFORE CPOE</td>
<td>AFTER CPOE</td>
<td>BEFORE CPOE</td>
</tr>
<tr>
<td>6 services chosen randomly from 23 available medical, surgical and intensive care units (2491 admissions)</td>
<td>Same 6 services plus 2 chosen randomly from same 23 available units (2047 admissions)</td>
<td>10.7</td>
</tr>
<tr>
<td>3 general medical services (379 admissions)</td>
<td>Same 3 general medical services (475 admissions)</td>
<td>7.6</td>
</tr>
</tbody>
</table>

Figure 2. Sensitivity analysis demonstrating the effect of different assumptions about the effectiveness of CPOE on the number of patients avoiding medication errors each year in the US.
Evidence-Based Hospital Referral: High-Risk Surgery

Overview

The Leapfrog Group’s Evidence-based hospital referral (EHR) standard focuses, in part, on high-risk surgical procedures (see neonatal standards in next section). In the first phase of the Leapfrog initiative, EHR for surgery will be based primarily on annual volume standards for participating hospitals, selected empirically from previous clinical research. The 5 surgical procedures—coronary artery bypass surgery (CABG), coronary angioplasty, elective abdominal aortic aneurysm (AAA) repair, carotid endarterectomy, and esophageal cancer surgery—were selected because of strong evidence of volume-outcome relationships with these procedures. In our baseline analysis, we estimate that implementation of EHR for these 5 surgical procedures would save approximately 2,581 lives each year in the US.

The initial Leapfrog standards (and the calculations herein) are based on surgical volume, though Leapfrog will also consider risk-adjusted mortality rates when these data are available. Although generally not available for non-cardiac procedures, risk-adjusted mortality rates for cardiac surgery and angioplasty are already carefully tracked in several states (e.g., NY State, Pennsylvania, and California). To the extent that such information is made available to The Leapfrog Group, mortality measures would replace surgical volume entirely as the EHR standard.

Methods and assumptions

Number of patients at risk (Table 1). To calculate the number of patients that could potentially benefit from EHR, we first estimated the total number of each of the 5 surgical procedures performed each year in US. We used data from the 1997 Nationwide Inpatient Sample to estimate the numbers of CABGs, coronary angioplasties, AAA repairs, and carotid endarterectomies performed. Because NIS data were not available for esophagectomies, we estimated total US counts using data from California state, extrapolated using 1999 US census data.

Some patients requiring emergency procedures (i.e., those which must be performed within an hour) are not appropriate for transfer to high-volume hospitals. For example, unstable patients with ruptured AAAs often need surgery at the hospital to which they present. To estimate the number of elective AAA repairs subject to EHR, we multiplied the total number of AAA repairs performed annually (44,129) by the proportion performed for non-ruptured AAAs (84%). In our calculations, we did not account for the small percentage of patients requiring truly “emergent” CABG (<5%), which occur most often from acute complications of coronary angioplasty. Such cases would only occur in hospitals continuing to perform angioplasty without surgical back-up, a controversial practice. Emergency angioplasty generally implies procedures performed for acute myocardial infarction. Because most hospitals treating AMI patients do not have this capability currently, we did not account for such cases in our calculations. Carotid endarterectomy and surgery for esophageal cancer are not performed emergently.

To avoid access issues and other unintended negative consequences in rural areas, The Leapfrog Group is restricting EHR implementation to urban areas. We assumed that 80% of the 5 procedures were performed in hospitals located in metropolitan areas, based on the proportion of hospital beds located in metropolitan statistical areas (MSAs).
To calculate the proportion of patients likely to be affected by EHR, we then estimated the proportion of patients currently undergoing the 5 procedures at low-volume hospitals (LVHs), defined according to volume standards set out by The Leapfrog Group (Table 1). Population-based studies necessary for estimating these proportions come primarily from California and New York State. For CABG and coronary angioplasty, California has substantially more LVHs than New York State, which has stricter restrictions on the delivery of cardiovascular health care (enforced in part through its Certificate of Need process). Because relevant information from other parts of the US is not available, we averaged (approximately) estimates from the two states for our baseline analysis. With carotid endarterectomy and abdominal aneurysm repair, the proportion of patients undergoing surgery at LVHs is very similar in California and New York State. Because no data were available from New York State for esophagectomies, we relied exclusively on data from California.

**Efficacy of moving patients from LVHs to HVHs (Table 2).** In estimating mortality reductions likely to be achieved with EHR for the 5 surgical procedures, we relied on point estimates from Dudley et al. for our baseline analysis. In their systematic review of the volume-outcome literature, they selected the “single best study” for each procedure based on several explicit criteria, including case-mix adjustment, sample size, and currency of the data. (For reference purposes, Appendix Tables summarize a broader number of volume-outcome studies published for each of the five procedures.) We converted adjusted odds ratios reported in Dudley et al. to relative risks using the formula described by Zhang et al. We then estimated EHR efficacy—the relative mortality reduction associated with moving patients from LVHs to HVHs—as 1 minus the relative risk.

**Mortality rates at low volume hospitals.** We then applied assumptions about EHR efficacy to estimates of current mortality rates at LVHs. With many high-risk procedures, there is evidence that surgical mortality rates have declined over the last decade. For this reason, we relied on recent overall mortality estimates from the Nationwide Inpatient Sample—instead of older data from the volume-outcome studies on which EHR efficacy was derived—for CABG, coronary angioplasty, and carotid endarterectomy. We then derived estimates of current mortality rates at LVHs by “back-calculation” using three variables: current overall mortality for each procedure, the proportion of patients at LVHs and HVHs, and the relative risk of mortality at LVHs vs. HVHs. Since similar information about current overall mortality rates is not available for elective AAA repair and esophagectomy, we used rates reported in the original volume-outcome studies.

**Lives saved (Table 3).** For each procedure, we calculated the number of lives potentially saved as the difference between expected deaths with and without EHR. Expected deaths were in turn determined by the product of three variables: 1) current number of patients at LVHs in US MSAs (Table 1); 2) current mortality rates at LVHs (Table 2); and 3) the efficacy of EHR (Table 2).

**Results**

In our baseline analysis, we estimate that full implementation of EHR would save approximately 2,581 lives each year in the US (Table 3). The greatest number of deaths would be prevented with CABG (1,486 deaths annually), followed by elective AAA repair (464 deaths annually), and coronary angioplasty (345 deaths annually). Potential lives saved with esophagectomy and CEA were 168 and 118, respectively.
Cautions and policy considerations

Our estimates of the potential number of lives saved (approximately 2500 annually) are generally consistent with two previous analyses. Dudley et al. estimated that 500 lives would be saved each year in California alone with adoption of “selective referral” for 10 procedures (which include the 5 selected by Leapfrog). Birkmeyer et al. estimated that 850 - 4200 lives could be saved by regionalizing the same ten procedures in the US Medicare population.

Of course, results from all these analyses depend heavily on point estimates of procedure-specific volume outcome relationships. We used point estimates from the structured literature review by Dudley et al., the most rigorous synthesis currently available. Because the volume-outcome literature is too heterogeneous for formal meta-analysis, the authors identified the “single best” study for each procedure based on several explicit criteria for study quality. However, relying on data from individual studies makes points estimates of volume-outcome relationships considerably less precise. It also discounts some degree of variation in results across studies (although findings from volume-outcome studies are very consistent for many procedures, Appendix).

It is important to consider other potential, indirect benefits of the Leapfrog initiative in surgery. Although its initial focus is exclusively on procedural volume standards, the Leapfrog has left open the possibility of replacing volume standards with more direct quality measures (at least for procedures performed commonly enough to allow for adequately precise measures of hospital-specific performance, such as, risk-adjusted mortality for CABG). Surgical quality could be improved considerably if hospitals responded to such incentives, collected data on their own performance, and participated in collaborative quality improvement initiatives.

Other indirect effects of the Leapfrog initiative could be negative. For example, procedure volume standards create incentives for hospitals near the volume thresholds to increase the number of procedures performed. This pressure creates the potential for unnecessary surgery for patients with “discretionary” conditions, such as lifestyle-limiting coronary artery disease, small AAAs, and asymptomatic carotid stenosis. The Leapfrog Group is relying on the integrity of surgeons to resist this pressure. Another important variable to consider is patient preferences. Although the Leapfrog initiative is exempting hospitals in rural areas to avoid excessive travel burdens for patients and access issues, some patients may still prefer receiving care in hospitals with which they are familiar.

The Leapfrog initiative could affect the delivery and (potentially) quality of care for procedures other than the 5 directly targeted. For example, encouraging referral of CABGs to high volume hospitals will affect where heart valve replacement can be offered. Policies aimed at improving the quality of AAA repair and carotid endarterectomy could diminish quality with other related procedures. For example, could hospital- or surgeon-specific performance with lower extremity bypass suffer as a result of lower total volumes of major vascular procedures? Although data necessary to answer such questions are currently not available, policy makers may need to consider individual procedures within the context of all high-risk procedures within subspecialties.

Our analysis estimates the benefits that could be achieved with full adherence to Leapfrog volume standards in all US hospitals (in metropolitan areas). How likely is this scenario? The other two Leapfrog safety initiatives—computer physician order entry and intensivist model ICUs—involv
hospital wide interventions. Making these changes for Leapfrog employees implies their availability to all other patients at the same hospitals. In contrast, even if EHR could be enforced for Leapfrog employees, there would be no mechanism for assuring the same change in referral pattern for other patients. For this reason, a very important contribution of the Leapfrog safety initiative may occur by simply increasing public awareness of the importance of volume for selected high-risk procedures.

Selected Bibliography:


Table 1. Number of patients currently undergoing 5 selected operations and number potentially affected by Leapfrog volume standards.

<table>
<thead>
<tr>
<th>Procedure (volume standard)</th>
<th>Total number performed in US/yr</th>
<th>Total number performed in US MSAs (previous column x .8)</th>
<th>Proportion Currently Performed at LVHs</th>
<th>Total Number at LVHs in US MSAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>CABG (500/yr)</td>
<td>383,788¹</td>
<td>307,030</td>
<td>66%⁶</td>
<td>164,261</td>
</tr>
<tr>
<td>Coronary Angioplasty (400/yr)</td>
<td>531,981¹</td>
<td>425,585</td>
<td>36%⁶</td>
<td>121,292</td>
</tr>
<tr>
<td>AAA Repair (30/yr)</td>
<td>37,068*¹</td>
<td>29,654</td>
<td>64%⁶</td>
<td>18,534</td>
</tr>
<tr>
<td>Carotid Endarterectomy (100/yr)</td>
<td>152,859¹</td>
<td>122,287</td>
<td>64%⁶</td>
<td>82,544</td>
</tr>
<tr>
<td>Esophagectomy (7/yr)</td>
<td>2,569#²</td>
<td>2,055</td>
<td>82.5%²</td>
<td>1,696</td>
</tr>
</tbody>
</table>

* Estimating by multiplying total number of AAA repairs (44,129),¹ by the proportion of AAA repairs performed for non-ruptured (and thus elective) AAAs (84%).⁴
** Hannan¹⁴ defined LVH as ≤26 procedures year.
# Because NIS data not available for esophagectomies, annual number performed in California² extrapolated to US estimate based on 1999 US census data.³
## No data available
Table 2. Estimated in-hospital mortality rates with 5 procedures at low-volume (LVHs) and high-volume (HVHs) hospitals.

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Estimated In-Hospital Mortality (%)</th>
<th>Relative mortality reduction with EHR*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall</td>
<td>HVH</td>
</tr>
<tr>
<td>CABG</td>
<td>2.9%</td>
<td>2.4%</td>
</tr>
<tr>
<td>Coronary Angioplasty</td>
<td>1.0%</td>
<td>0.9%</td>
</tr>
<tr>
<td>Elective AAA Repair</td>
<td>4.9%</td>
<td>4.2%</td>
</tr>
<tr>
<td>Carotid Endarterectomy</td>
<td>0.6%</td>
<td>0.5%</td>
</tr>
<tr>
<td>Esophagectomy</td>
<td>14%</td>
<td>5.9%</td>
</tr>
</tbody>
</table>

* Defined by one minus the relative risk of mortality at HVHs, compared to LVHs. Relative risks were converted from adjusted odds ratios in Dudley et al.\(^6\) using the formula describing by Zhang et al.\(^7\)

** “Back-calculated” from estimates of overall mortality rate (above), relative risks of mortality at LVHs and HVHs, and the estimated distribution of patients at LVHs and HVHs (Table 1).

Table 3. Number of lives potentially saved by implementation of EHR.

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Total patients at LVHs in MSAs</th>
<th>Expected deaths without EHR</th>
<th>Expected deaths with EHR</th>
<th>Lives saved by EHR</th>
</tr>
</thead>
<tbody>
<tr>
<td>CABG</td>
<td>164,261</td>
<td>5428</td>
<td>3942</td>
<td>1486</td>
</tr>
<tr>
<td>Coronary Angioplasty</td>
<td>121,292</td>
<td>1404</td>
<td>1059</td>
<td>345</td>
</tr>
<tr>
<td>AAA Repair</td>
<td>18,534</td>
<td>1242</td>
<td>778</td>
<td>464</td>
</tr>
<tr>
<td>Carotid Endarterectomy</td>
<td>82,544</td>
<td>543</td>
<td>425</td>
<td>118</td>
</tr>
<tr>
<td>Esophagectomy</td>
<td>1,696</td>
<td>267</td>
<td>99</td>
<td>168</td>
</tr>
<tr>
<td>Total</td>
<td>388,327</td>
<td>8,884</td>
<td>6,303</td>
<td>2,581</td>
</tr>
</tbody>
</table>
Evidence-Based Hospital Referral: High-Risk Neonatal Intensive Care

**Overview**

Evidence-based hospital referral (EHR) for neonatal ICU care will focus on very low birth weight babies (500-1500g) and those with major congenital anomalies—two groups which account for approximately 90% of all neonatal deaths. Deliveries involving these high-risk groups will be concentrated in regional NICUs (“regional” as defined by states using this classification) with average daily census levels of at least 15. This policy involves ensuring that high-risk deliveries occur in the “right” places, not that babies are transferred there after delivery. Implementation of EHR nationwide for high-risk neonatal intensive care would save approximately 1,863 lives each year in the US. In following sections, we describe the methods and assumptions used in our analysis.

**Methods and assumptions**

In calculating the number of lives potentially saved by full implementation of EHR for high-risk deliveries, we started with the number of deliveries potentially affected by the policy. We then estimated baseline mortality risks for the two high-risk groups (below) and the potential mortality reductions associated with the Leapfrog EHR standards.

**Number of patients at risk (Table 1).** To calculate the number of babies that could benefit from EHR, we applied incidence rates of very low birth weight and major congenital anomalies to the total number of US births each year (3,944,046).<sup>1</sup> In 1998, 1.26% of all births were 500-1500g babies.<sup>1</sup> The incidence of congenital anomalies targeted by The Leapfrog Group (potentially correctable conditions that frequently require surgical intervention) is 1.6% (personal communication: C. Phibbs, PhD). We assumed that 80% of births in both groups occurred in hospitals located in metropolitan areas, based on the proportion of hospital beds located in metropolitan statistical areas (MSAs) (source: 1996 AHA file).

In California, approximately 74 percent and 82 percent of 1995 births involving very low birth weight babies or major congenital anomalies, respectively, occurred in non-regional (level 1, II, or II+) NICUs or in regional NICUs with average daily census rates below 15 (personal communication: C. Phibbs, PhD). Because similar population-based data from other regions of the US are not available, we relied on data from California for our baseline analysis.

Not all high-risk births currently occurring in other settings would be appropriate for transfer to large regional ICUs before delivery. Most mothers experiencing premature labor will present to the nearest hospital or facility at which they have received their prenatal care. Some with particularly advanced or precipitous labor will not be appropriate candidates for transfer for safety reasons. This proportion cannot be determined directly from the medical literature, but an informal survey of neonatologists at Dartmouth-Hitchcock Medical Center suggests that approximately 10% of such women (in metropolitan areas) would not be appropriate for transfer. Approximately 60% of deliveries involving major congenital anomalies are not detected by prenatal ultrasound.<sup>2,3</sup> Thus, we assumed that 90% and 40% of deliveries involving very low birth weight babies or congenital anomalies, respectively, currently occurring at other facilities could be moved to large regional ICUs with implementation of EHR.

**Efficacy of EHR for high-risk neonatal intensive care (Table 2).** Only one study—based on 1990 California state data—has carefully examined mortality rates for high-risk deliveries at different types of facilities.<sup>4</sup> Babies with very low birth weight or major congenital anomalies were only 62 percent
as likely to die at regional NICUs with ADC 15+, as compared to at level I NICUs (adjusted OR 0.62, 95% CI 0.48-0.84). The adjusted odds ratio of mortality at regional NICUs with ADC 15+ compared to all other facilities was approximately 0.67 (estimated from results presented in study). Because the study did not present stratified results for the two patients subgroups, we assumed the same relative benefit (odds ratio) for very low birth weight babies and those with major congenital anomalies. We converted odds ratios to relative risks for each of the two patient subgroups using the formula described by Zhang et al.\textsuperscript{5}

Mortality rates for high-risk deliveries not at regional NICUs with ADC 15+. According to 1995 California discharge data, overall mortality rates for babies with very low birth weight or congenital anomalies (as selected by Leapfrog) was 16.8% and 9.25%, respectively (personal communication: Phibbs). Mortality was defined as death within the first 28 days of life, or within the first year of life if continuously hospitalized. We then derived estimates of current mortality rates at non-regional NICUs and regional NICUs with ADC<15 (combined) by “back-calculation” using three variables: the overall mortality rate for each of the two patient groups, the proportion of deliveries currently at regional NICUs with ADC 15+ vs. elsewhere, and relative risk of mortality at regional NICUs with ADC 15+ vs. elsewhere (Table 2).

Results

For each of the two groups, we calculated the number of lives potentially saved as the difference between expected deaths with and without EHR. Overall, the Leapfrog EHR standard for NICU care would saved 1863 babies’ lives each year in the US: 1369 lives for deliveries involving babies 500-1500g, 494 lives for deliveries involving major congenital anomalies. As expected, the estimated number of lives saved was sensitive to assumptions about EHR effectiveness (Figure 1).

Cautions and policy considerations

Estimates of the benefits likely to be achieved by EHR for high-risk deliveries should be viewed cautiously. Compared to the evidence underlying other Leapfrog safety standards, clinical and health services research examining sources of better outcomes in NICUs is relatively scant. By necessity, our analysis relied primarily on a single study based on 1990 California data by Phibbs et al. Thus, it is important to consider the generalizability of this study with regards to two key parameters: 1) where high-risk babies are currently being delivered, and 2) the potential gains with moving them to high volume, regional centers.

Although reliable data are currently not available, it is possible that the proportion of high-risk deliveries occurring in non-regional (level I-II) NICUs with ADC<15 is lower outside California (approximately 75% in 1990). For example, only 14% and 22% of very low birth weight babies were born outside level III centers in Missouri and South Carolina, respectively.\textsuperscript{6,7} However, the classifications “regional” and “level III” are not applied uniformly across states. In many states, the label is designated by hospitals themselves. Even where the market share of high-risk deliveries for level III NICUs has been described, data about NICU volume or average daily census rates (including in Missouri and South Carolina) are not available.

Although the study by Phibbs et al. is the most recognized, other investigators have examined volume-outcome relationships in neonatal intensive care, with varying conclusions. Using data from
the Vermont Oxford Network registry, Horbar found no association between NICU volume (annual number of admissions) and mortality in deliveries involving very low birth weight babies. However, this study was not population-based. Instead, the VON consortium, which was formed to promote trials, outcomes research, and quality improvement, consists disproportionately of tertiary care hospitals and those with pediatrics training programs. Unlike the Phibbs study, “outborn” infants (those born at non-study facilities) were also included. In contrast, Tilford et al. did note an inverse relationship between ICU volume and risk-adjusted mortality, but this study focused on pediatric ICUs (older children), not NICUs. Estimates of the benefits likely to be achieved through EHR in neonatal ICUs will need to be updated as more information about volume-outcome relationships in this area become available.

Selected Bibliography:


Table 1. Number of neonatal ICU babies potentially affected by EHR.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Total born in US/yr* (incidence)</th>
<th>Total born in hospitals located in MSAs</th>
<th>MSA Deliveries Not at Regional ICUs with ADC 15+</th>
<th>MSA Deliveries Eligible for EHR on “Clinical” Grounds (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Baseline est. (source)</td>
<td>n</td>
</tr>
<tr>
<td>Very low birth weight (500-1500g)</td>
<td>49,695 (1.26%)</td>
<td>39,756</td>
<td>74% (California, 1995)</td>
<td>29,419</td>
</tr>
<tr>
<td>Major congenital anomalies#</td>
<td>63,105 (1.6%, California, 1995, Phibbs)</td>
<td>50,484</td>
<td>82% (California, 1995)</td>
<td>41,397</td>
</tr>
</tbody>
</table>

* Based on total 3,944,046 babies born each year in US
** A small proportion of mothers may experience such precipitous premature labor that safe transfer to an appropriate regional neonatal ICU is not possible or safe.
# ICD-9CM codes: 741.XX, 742.0X, 742.2-742.9, 745.XX, 746.00-746.85, 747.1X-747.9, 748.0, 748.2-748.8X, 750.16, 750.3, 750.4, 750.6, 751.XX, 752.7, 753.1X, 753.3, 753.6, 756.4, 756.51, 756.55, 756.59, 756.6, 756.7X, 756.89, 756.9
## More than half of all major congenital anomalies are not detected by prenatal ultrasound.

Table 2. Estimated mortality rates at regional NICUs with average daily census at least 15, versus at other facilities.

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Observed mortality*</th>
<th>Predicted mortality**</th>
<th>Relative mortality reduction with EHR#</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Regional NICU, ADC 15+</td>
<td>Elsewhere</td>
<td></td>
</tr>
<tr>
<td>Very low birth weight (500-1500g)</td>
<td>16.8%</td>
<td>13.0%</td>
<td>18.1%</td>
</tr>
<tr>
<td>Major congenital anomalies</td>
<td>9.25%</td>
<td>6.8%</td>
<td>9.8%</td>
</tr>
</tbody>
</table>

* Adjusted mortality rates (one-year or by hospital discharge) based on 1995 California data (Phibbs: personal communication)
** Derived from overall mortality rate for each of the two patient groups, the proportion of deliveries currently at regional NICUs with ADC 15+ vs. elsewhere, and relative risk of mortality at regional NICUs with ADC 15+ vs. elsewhere, based on data from Phibbs et al, 1997.
# Defined by one minus the relative risk of mortality at regional neonatal ICUs with ADC 15+, compared to other centers. Relative risks were converted from adjusted odds ratios using the formula describing by Zhang et al.
Table 3. Number of lives potentially saved by implementation of EHR.

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Total patients moved to large NICUs with ADC 15+ (with EHR)</th>
<th>Expected deaths without EHR</th>
<th>Expected deaths with EHR</th>
<th>Lives saved by EHR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very low birth weight (500-1500g)</td>
<td>26,477</td>
<td>4804</td>
<td>3435</td>
<td>1369</td>
</tr>
<tr>
<td>Major congenital anomalies</td>
<td>16,559</td>
<td>1620</td>
<td>1126</td>
<td>494</td>
</tr>
<tr>
<td>Totals</td>
<td>43,036</td>
<td>6,424</td>
<td>4,561</td>
<td>1,863</td>
</tr>
</tbody>
</table>

Figure 1. Sensitivity analysis demonstrating the effect of different assumptions about the effectiveness of EHR on the number of lives saved each year in the US. (VLBWBs, very low birth weight babies)
ICU Physician Staffing

Overview

In this section, we estimate the number of lives that could be saved by full implementation of ICU physician staffing (IPS) nationwide. As defined by The Leapfrog Group, hospitals meeting IPS requirements will have ICUs managed by physicians who are board-certified (or -eligible) in critical care medicine. Such physicians will be present in the ICU during daytime hours and provide clinical care exclusively in the ICU. At other times, they will be able to return pages within 5 minutes and rely on in-hospital FCCS certified “effectors” (physicians or physician extenders) who can reach ICU patients within 5 minutes.

In our baseline analysis, we estimate that full implementation of IPS would save approximately 53,850 lives each year in the US. IPS is so effective because such a large number of US deaths occur in the ICU (approximately 500,000). Thus, even small improvements in ICU mortality rates save many lives. Although our analysis is based on the best data currently available, many of the variables used in our calculations cannot be estimated precisely. In instances of uncertainty, we selected parameters, which biased our calculations downward. Thus, we believe our estimate of the number of lives likely to be saved by IPS is conservative. In following sections, we describe the methods and assumptions we used in our analysis.

Methods and Assumptions

As summarized in Figure 1, we calculated the number of lives potentially saved by full implementation of intensivist model ICUs. We started with the population of ICU patients potentially affected by the policy. We then estimated their baseline in-hospital mortality risks and the potential mortality reductions associated with implementing intensivist model ICUs.

Current number of ICU admissions. To estimate the number of patients that could potentially benefit from the policy initiative we determined the number of patients admitted each year to non-intensivist ICUs. We could not directly determine the overall number of patients admitted to US ICUs. Based on analysis of the 1999 MEDPAR file, approximately 2.2 million Medicare patients were admitted to medical/surgical ICUs (excluding coronary care units) (personal communication: P. Pronovost, MD, Ph.D.). Because Medicare patients represent approximately half of all adult ICU patients, (Pronovost) we assumed in our baseline analysis that 4.4 million US patients are admitted to ICUs each year in the US.

To avoid access issues in rural areas, The Leapfrog Group is restricting policy implementation to metropolitan areas. According to analysis of the 1996 AHA file and census database, 80% of all US hospital beds (53% of hospitals) are located in metropolitan statistical areas (MSAs). Assuming that 80% of ICU admissions similarly occur in MSAs, we estimate that 3.52 million patients are admitted each year to ICUs in urban hospitals.

The current proportion of US ICUs with intensivist models is unknown, but probably low. In a 1991 national survey, only 22% of hospitals indicated that ICU order writing was restricted to unit staff (i.e., a “closed unit”). In a follow-up survey, the same group reported that 17% of ICUs had closed units with respect to order writing. Neither study described the proportion of closed units in which all ICU staff were board-certified (or -eligible) in critical care.
medicine, or met other Leapfrog criteria. In our baseline analysis, we assumed that 15% of all ICU patients are currently treated in ICUs meeting the Leapfrog standard.

**Current ICU Mortality.** We estimated average in-hospital mortality rates for ICU patients from two large multi-center studies. Zimmerman et al.\(^3\) noted an overall 12.4% in-hospital mortality rate in 38,000 patients admitted to 161 hospitals between 1993 and 1996. In another study by Shortell et al.,\(^4\) in-hospital mortality for 17,000 patients at 42 randomly selected ICUs was 16.6% between 1988 and 1990. In our baseline analysis, we selected the lower (and thus more conservative) of these two estimates—12%.

**Mortality reductions with implementing the intensivist model.** Numerous studies have evaluated the effectiveness of similar (though not identical) staffing models in reducing ICU mortality. After performing a structured literature review (Figure 2), we identified 9 studies on which to base our estimates of the effectiveness of implementing intensivist model staffing.\(^5\)-\(^{12}\) Six of these studies were based on pre/post study designs at single sites—all generally large ICUs in teaching hospitals (Table 1). Three studies had “cross sectional” designs, comparing IPS hospitals (or equivalent) with non-IPS hospitals during a single period (Table 2).

In all 9 studies, intensivist model staffing was associated with reduced ICU mortality. In 5 of the studies the mortality reductions were statistically significant. Relative mortality reductions associated with intensivist model staffing ranged varied from 15% to 60% (relative risks 0.4-0.85). To be conservative in our calculations of lives saved, we selected the estimate from the study demonstrating the least effectiveness (15% relative mortality reduction). However, we tested the effect of different assumptions about the effectiveness of intensivist model ICUs in sensitivity analysis.

**Results**

In our baseline analysis, we estimate that full implementation of intensivist model staffing would save approximately 53,850 lives each year in the US. As expected, the number of lives saved varied according to assumptions about the effectiveness of intensivist model staffing. (Figure 2) For example, assuming a 35% relative mortality reduction (a mid-range estimate from the 9 studies) instead of 15% (the most conservative estimate), 126,000 lives would be saved.

**Cautions and policy considerations**

Because so many patients die in ICUs each year in the US (approximately 500,000), even small reductions in ICU mortality rates would save many lives. If the Leapfrog initiative is successful in effecting full implementation of intensivist model ICU staffing in metropolitan areas nationwide, we estimate that approximately 53,850 lives could be saved.

Of course, our estimates depend heavily on assumptions about the effectiveness of implementing intensivist model staffing. We chose to be conservative in estimating its potential effectiveness because of several limitations in the original studies and problems with extrapolating them nationwide. First, inferences from the 6-pre/post studies are limited by secular trend bias, i.e.; mortality may have fallen at those hospitals for reasons other than implementation of intensivist model staffing. The hospitals in these studies may have changed
other aspects of care not directly related to physician staffing changes. Although there is no evidence that ICU mortality rates are declining, mortality rates with many clinical conditions are improving over time with advances in science and technology.\textsuperscript{13,14} However, given the magnitude of declines in mortality seen in many of these studies, it is very unlikely that improvements can be attributed to secular trend bias alone. Second, estimates from the 3 cross-sectional studies may have suffered from imperfect risk-adjustment. Thus, their results may be partially confounded by unmeasured differences in case-mix between control and intensivist model groups. Third, caution is required in generalizing results of the 9 studies—all based at large, teaching hospitals—to other settings.

Finally, there was substantial variation in the “intervention” in the 9 studies being assessed, and the explicitness with which the intervention was described. Some interventions involved simply adding co-management by a single intensivist to a system primarily run by non-ICU based physicians; others described extensive changes in staff organization, including complete replacement of ward-based teams by intensivists and ICU-based house staff. It is important to note, however, that the Leapfrog IPS standard falls on the latter, “stricter” side of the spectrum, and thus is likely to be more efficacious.

Although the potential benefits are large, full implementation of intensivist model ICU staffing carries several challenges. Although workforce issues have not been studied carefully, it is unlikely that there are currently enough board-certified intensivists to fully staff ICUs at all hospitals.\textsuperscript{15} In hospitals with small units, meeting the Leapfrog standard may increase net cost per stay. For these reasons, broad implementation of intensivist model ICU staffing may require a blend of increased fellowship training slots in critical care, consolidation of small ICUs, and advances in telemedicine.

Many would argue that lives saved by intensivist model ICU staffing are not equivalent to lives saved by other public health interventions (e.g. seat belt laws). ICU patients often have substantial comorbidity and thus shortened life expectancies compared to the general population. For this reason, further research should consider how improvements in ICU care affect quality of life after hospital discharge and long-term survival.

**Selected Bibliography:**


5. Reynolds HN, Haupt MT, Thill-Baharozian MC, Carlson RW. Impact of critical care


**Figure 1.** Number of lives that would be saved each year by full implementation ICU physician staffing (IPS) nationwide.

- **Annual admissions to urban ICUs in the US**: 3,520,000
- **Number of patients admitted to urban ICUs without IPS**: 2,992,000
- **Number of deaths at urban ICUs without IPS**: 359,040
- **ICU deaths despite implementation of IPS**: 305,184
- **Number of survivors at ICUs without IPS**: 2,632,960
- **Number of patients admitted to ICUs with IPS**: 528,000
- **ICU deaths averted because of IPS implementation**: 53,856

85% → 12% → 15%
### Table 1. Pre/post studies of the effectiveness of IPS.

<table>
<thead>
<tr>
<th>SETTING</th>
<th>ICU MANAGEMENT MODEL (PATIENTS, N)</th>
<th>HOSPITAL MORTALITY (%)</th>
<th>RELATIVE RISK, UNADJUSTED (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PRE</td>
<td>POST</td>
<td>PRE</td>
</tr>
<tr>
<td>Detroit Receiving Hospital (1982-1984)</td>
<td>Internists† with ICU housestaff (n=100)</td>
<td>Full-time intensivists with ICU fellows and ICU housestaff (n=112)</td>
<td>74</td>
</tr>
<tr>
<td>Surgical ICU Plains Health Care Medical Center Saskatchewan (1984-1985)</td>
<td>Attending physician or surgeon with ICU housestaff (n=223)</td>
<td>Plus co-management by full-time intensivist (n=216)</td>
<td>36</td>
</tr>
<tr>
<td>Medical ICU Long Island Jewish Medical Center (NY) (1992-1993)</td>
<td>Admitting attending with ICU housestaff (n=152)</td>
<td>Intensivists with ICU housestaff (n=154)</td>
<td>45</td>
</tr>
<tr>
<td>Bridgeport Hospital (CT) (1992-1994)</td>
<td>Private physicians with housestaff (n=459)</td>
<td>Plus full-time medical director and co-management by intensivist (n=471)</td>
<td>34</td>
</tr>
<tr>
<td>University of Chicago Hospital (1993-1994)</td>
<td>Admitting attending with ward housestaff (n=124)</td>
<td>Intensivists with ICU housestaff (n=121)</td>
<td>29</td>
</tr>
<tr>
<td>Rhode Island Hospital (1995-1996)</td>
<td>Co-management by intensivist (n=125)</td>
<td>Management by intensivist (n=149)</td>
<td>14</td>
</tr>
</tbody>
</table>

* p<0.05  
** p<0.001  
†Internists at Detroit Receiving were non-intensivists who maintained ambulatory care practices while rotating through as ICU attendings  
Note: mortality rates are unadjusted except for Univ. of Chicago study where mortality percent were adjusted as follows: (combined pre & post expected mortality)*(observed mortality/expected mortality for each group)
Table 2: Cross-sectional studies of the effectiveness of IPS.

<table>
<thead>
<tr>
<th>SETTING</th>
<th>INTERVENTION</th>
<th>MORTALITY</th>
<th>RELATIVE RISK, UNADJUSTED (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgical ICU University of Penn Medical Center (1994-1995)</td>
<td>Same* (see note below)</td>
<td>Attending surgeons with ward housestaff (n=100)</td>
<td>Full-time intensivists with ICU housestaff (n=100)</td>
</tr>
<tr>
<td>5 Maryland ICUs that manage post-op AAA repairs (1994-1995)</td>
<td>36 Maryland ICUs that manage post-op AAA repairs (1994-1995)</td>
<td>No daily rounds by intensivists (n=472)</td>
<td>Daily rounds by intensivists (n=2515)</td>
</tr>
<tr>
<td>Winthrop-University Hospital, NY (1993)</td>
<td>Long Island Jewish Medical Center, NY (1993)</td>
<td>Attending physician with ICU housestaff. (n=95)</td>
<td>Full-time intensivists with ICU housestaff (n=185)</td>
</tr>
</tbody>
</table>

* patients selected for control or intervention group (within the same ICU) by attending surgeon preference.

**p<0.01

Figure 2. Sensitivity analysis demonstrating the effect of different assumptions about the effectiveness of IPS on the number of lives saved each year in the US.
Appendix

Section 1: Selected studies assessing relationship between volume and outcome for 5 selected procedures. The following tables include studies which: 1) were population-based, 2) were published since 1985, 3) provided mortality rates and confidence intervals (or provided adequate data to calculate them), and 4) assessed volume standards similar to those selected by Leapfrog.

Table 1: Coronary Artery Bypass Graft Surgery (Standard >=500)

<table>
<thead>
<tr>
<th>Author</th>
<th>Setting (n)</th>
<th>Data type</th>
<th>LVH Definition</th>
<th>Odds ratios of surgical mortality at LVHs (95% CIs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grumbach</td>
<td>NY, California, Canada</td>
<td>Administrative</td>
<td>&lt;500</td>
<td>1.37 (1.28, 1.47)</td>
</tr>
<tr>
<td>(1995)</td>
<td>(116,593)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hannan</td>
<td>NY State, 1986 (9,774)</td>
<td>Administrative</td>
<td>&lt;310</td>
<td>1.28 (1.05, 1.57)</td>
</tr>
<tr>
<td>(1989)</td>
<td></td>
<td></td>
<td>&lt;651</td>
<td>1.47 (1.18, 1.83)</td>
</tr>
<tr>
<td>Hannan</td>
<td>New York State, 1989</td>
<td>Clinical</td>
<td>&lt;500</td>
<td>1.39 (1.16, 1.67)</td>
</tr>
<tr>
<td>(1991)</td>
<td>(12,448)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Showstack</td>
<td>California, 1983 (18,986)</td>
<td>Administrative</td>
<td>&lt;=350</td>
<td>1.34 (1.15, 1.56)</td>
</tr>
<tr>
<td>(1987)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Coronary Angioplasty (Standard >=400).

<table>
<thead>
<tr>
<th>Author</th>
<th>Setting (n)</th>
<th>Data type</th>
<th>LVH Definition</th>
<th>Odds ratios of surgical mortality at LVHs (95% CIs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hannan</td>
<td>New York State, 1991-4 (62,670)</td>
<td>Administrative</td>
<td>&lt;400</td>
<td>1.33 (1.10, 1.61)</td>
</tr>
<tr>
<td>(1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jollis</td>
<td>US Medicare, 1992 (97,478)</td>
<td>Administrative</td>
<td>&lt;200*</td>
<td>1.19 (1.10, 1.29)</td>
</tr>
<tr>
<td>(1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phillips</td>
<td>California, 1989 (20,064)</td>
<td>Administrative</td>
<td>&lt;401</td>
<td>0.92 (0.67, 1.26)</td>
</tr>
<tr>
<td>(1995)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Medicare volume represents approximately 50% of surgical volume for this procedure.
**Table 3**: Elective AAA repair (Standard >= 30).

<table>
<thead>
<tr>
<th>Author</th>
<th>Setting (n)</th>
<th>Data type</th>
<th>LVH Definition</th>
<th>Odds ratios of surgical mortality at LVHs (95% CIs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hannan</td>
<td>NY State 1982-87 (3570)</td>
<td>Administrative</td>
<td>&lt;=26</td>
<td>1.51 (1.15, 1.97)</td>
</tr>
<tr>
<td>Kazmers</td>
<td>National VA population, 1991-93 (3419)</td>
<td>Administrative</td>
<td>&lt;32</td>
<td>1.64 (1.18, 2.27)</td>
</tr>
<tr>
<td>Sollano</td>
<td>New York State, 1990-5 (9847)</td>
<td>Administrative</td>
<td>&lt;17</td>
<td>1.28 (1.18, 1.39)</td>
</tr>
<tr>
<td>Wen</td>
<td>Canada 1988-92 (5492)</td>
<td>Administrative</td>
<td>&lt;21</td>
<td>1.17 (.88, 1.56)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&lt;41</td>
<td>1.17 (.87, 1.57)</td>
</tr>
</tbody>
</table>

**Table 4**: Carotid Endarterectomy (Standard >=100)

<table>
<thead>
<tr>
<th>Author</th>
<th>Setting (n)</th>
<th>Data type</th>
<th>LVH Definition</th>
<th>Odds ratios of surgical mortality at LVHs (95% CIs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edwards</td>
<td>Tennessee 1979-88 (11,199)</td>
<td>Administrative</td>
<td>&lt; 50</td>
<td>1.23 (0.95, 1.61)</td>
</tr>
<tr>
<td>Hannan</td>
<td>NY State, 1990-95 (28,207)</td>
<td>Administrative</td>
<td>&lt;=100</td>
<td>1.28 (1.13, 1.45)</td>
</tr>
<tr>
<td>Karp</td>
<td>Georgia 1990-95 (10569)</td>
<td>Administrative</td>
<td>&lt;=50*</td>
<td>1.55 (1.09, 2.22)</td>
</tr>
<tr>
<td>Perler</td>
<td>Maryland 1990-95 (9918)</td>
<td>Administrative</td>
<td>&lt;50</td>
<td>1.31 (0.87, 1.98)</td>
</tr>
</tbody>
</table>

*Medicare volume represents approximately 50% of surgical volume for this procedure.

**Analysis based on 30-day mortality**

28
Table 5: Esophagectomy (Standard >=6).

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Setting (n)</th>
<th>Data type</th>
<th>LVH Definition</th>
<th>Odds ratios of surgical mortality at LVHs (95 % CIs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Begg (1998)</td>
<td>US 1984-93 (503)</td>
<td>SEER</td>
<td>&lt;7</td>
<td>4.11 (1.77, 9.56)</td>
</tr>
</tbody>
</table>

Selected Bibliography:


Section 2: Medline search strategy for IPS(OVID)*

**Search Terms Defined**

1. **LOCATION**: Intensive care units/ or respiratory care units, intensive care
2. **ORGANIZATIONAL INNOVATIONS**: organizational innovation, organization(floating subheading), intensivist, specialties/ medical
3. **MEDICAL OUTCOMES**: cause of death, fatal outcome, hospital mortality(floating subheading), survival rate, length of stay, treatment outcome, outcome assessment (health care)
Acknowledgments

We received valuable input from experts in different areas:
- David Bates, MD, Brigham & Women’s Hospital and Harvard Medical School (computer physician order entry)
- Peter Pronovost, MD, PhD, Johns Hopkins Medical Institutions (ICU physician staffing)
- Ciaran Phibbs, PhD, Stanford University (EHR for neonatal intensive care)

This acknowledgment does not necessarily imply, however, their agreement with all aspects of our analysis or our conclusions.