Ethics, Equity and Experiences related to Patient Safety and Artificial Intelligence

David W. Bates, MD, MSc, Professor of Medicine, Harvard

Co-Director, Center for AI and Biomedical Informatics of the Learning Healthcare System, Mass General Brigham

7th Nordic Conference on Research in Patient Safety and Quality

Overview

- SafeCare study results
  - How these fit with broader context
- Artificial intelligence and safety improvement
- Overarching conclusions

Results of the Harvard Medical Practice: published 1991, 1984 data

The NEW ENGLAND JOURNAL of MEDICINE

Vol. 324 No. 6 ADVERSE EVENTS IN HOSPITALIZED PATIENTS — LEAPE ET AL. 377

THE NATURE OF ADVERSE EVENTS IN HOSPITALIZED PATIENTS

Results of the Harvard Medical Practice Study II

Lucian L. Leape, M.D., Troyen A. Brennan, M.D., J.D., M.P.H., Nan Laird, Ph.D., Ann G. Lawthers, Sc.D., A. Russell Localio, J.D., M.P.H., Benjamin A. Barnes, M.D., Liesi Hebert, Sc.D., Joseph P. Newhouse, Ph.D., Paul C. Weiler, LL.M., and Howard Hiatt, M.D.

Abstract *Background.* In a sample of 30,195 randomly selected hospital records, we identified 1133 patients (3.7 percent) with disabling injuries caused by medical treatment. We report here an analysis of these adverse events and their relation to error, negligence, and disability.

Methods. Two physician-reviewers independently identified the adverse events and evaluated them with respect to negligence, errors in management, and extent of disability. One of the authors classified each event according to type of injury. We tested the significance of differences in rates of negligence and disability among categories with at least 30 adverse events.

Results. Drug complications were the most common type of adverse event (19 percent), followed by wound infections (14 percent) and technical complications (13 percent). Nearly half the adverse events (48 percent) were associated with an operation. Adverse events during surgery were less likely to be caused by negligence (17 percent) than nonsurgical ones (37 percent). The proportion of adverse events due to negligence was highest for diagnostic mishaps (75 percent), noninvasive therapeutic mishaps ("errors of omission") (77 percent), and events occurring in the emergency room (70 percent). Errors in management were identified for 58 percent of the adverse events, among which nearly half were attributed to negligence.

Conclusions. Although the prevention of many adverse events must await improvements in medical knowledge, the high proportion that are due to management errors suggests that many others are potentially preventable now. Reducing the incidence of these events will require identifying their causes and developing methods to prevent error or reduce its effects. (N Engl J Med 1991; 324:377-84.)

Table 1. Types of Adverse Events and Proportion of Events Involving Negligence.

Type of Event	No. of Events in Sample	Weighted Proportion of Events*		
		In Population	Due to Negligence	With Serious Disability
			percent	
Operative				
Wound infection	160	13.6	12.5†	17.9
Technical complication	157	12.9	17.6	12.0†
Late complication	137	10.6	13.6‡	35.7
Nontechnical complication	87	7.0	20.1	43.8
Surgical failure	58	3.6	36.4	17.5
All	599	47.7	17.0	24.0
Nonoperative				
Drug-related	178	19.4	17.7‡	14.1‡
Diagnostic mishap	79	8.1	75.2†	47.0‡
Therapeutic mishap	62	7.5	76.8†	35.4
Procedure-related	88	7.0	15.1	28.8
Fall	20	2.7	—	—
Fracture§	18	1.2	—	—
Postpartum¶	18	1.1	—	—
Anesthesia-related	13	1.1	—	—
Neonatal	29	0.9	—	—
System and other	29	3.3	35.9	36.0
All	534	52.3	37.2	25.3
Total	1133	100.0	27.6	24.7

*Dashes denote categories for which there were too few observations to determine a percentage.

†$P<0.001$ for the difference between this rate and all others in the same column.

‡$P<0.01$ for the difference between this rate and all others in the same column.

§Includes nonoperative fractures only.

¶Includes noncesarean deliveries only.

Take homes:
NY State rate of AE 3.2-4.2%
28% negligence
25% serious
Of the 2.6M patients in NY state 1984, 98,000 AE and
27,000 due to negligence

Call to action: develop methods to reduce errors

How safe is care today?
Results from Safe Care published January 12, 2023

The NEW ENGLAND JOURNAL of MEDICINE

SPECIAL ARTICLE

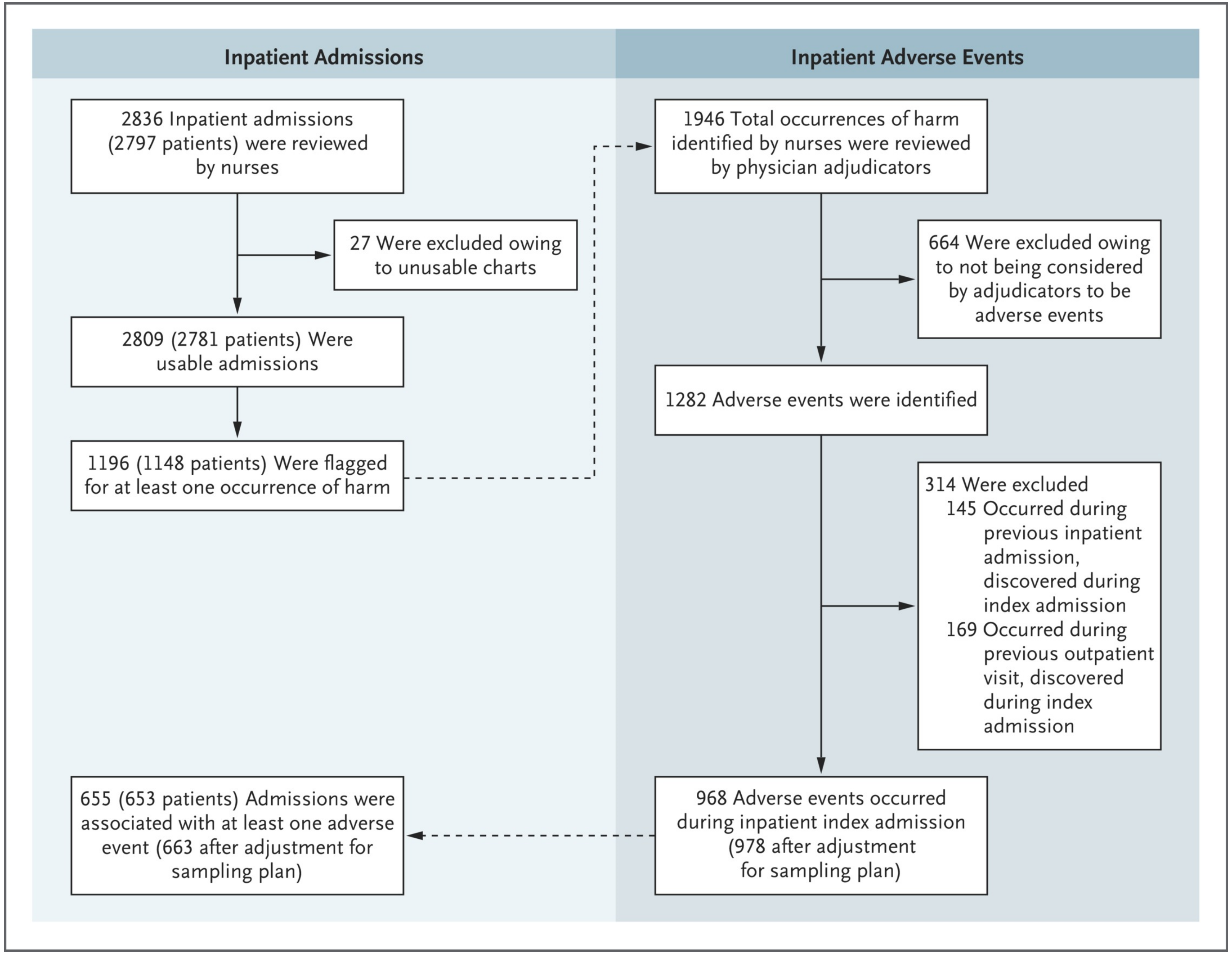
The Safety of Inpatient Health Care

David W. Bates, M.D., David M. Levine, M.D., M.P.H.,
Hojjat Salmasian, M.D., Ph.D., M.P.H., Ania Syrowatka, Ph.D., David M. Shahian, M.D.,
Stuart Lipsitz, Sc.D., Jonathan P. Zebrowski, M.D., M.H.Q.S.,
Laura C. Myers, M.D., M.P.H., Merranda S. Logan, M.D., M.P.H.,
Christopher G. Roy, M.D., M.P.H., Christine Iannaccone, M.P.H., Michelle L. Frits, B.A.,
Lynn A. Volk, M.H.S., Sevan Dulgarian, B.S., B.A., Mary G. Amato, Pharm.D., M.P.H.,
Heba H. Edrees, Pharm.D., Luke Sato, M.D., Patricia Folcarelli, Ph.D., R.N.,
Jonathan S. Einbinder, M.D., M.P.H., Mark E. Reynolds, B.A.,
and Elizabeth Mort, M.D., M.P.H.

ABSTRACT

BACKGROUND

Adverse events during hospitalization are a major cause of patient harm, as documented in the 1991 Harvard Medical Practice Study. Patient safety has changed substantially in the decades since that study was conducted, and a more current assessment of harm during hospitalization is warranted.



The NEW ENGLAND JOURNAL of MEDICINE

Differences in Adverse Events by Clinical Categories

Category	Adverse Events	Permanent Disability or Life Threatening	Death Rate	Preventable/ Probably Preventable
Medication-related	39% (380)	2% (9)	1% (2)	27% (103)
Surgical/Procedural	30% (291)	12% (36)	1% (2)	13% (39)
Patient Care*	15% (148)	3% (5)	0% (0)	38% (56)
Healthcare Acquired Infections (HAI)	12% (114)	9% (10)	3% (3)	18% (20)
Perinatal/Maternal	3% (25)	8% (2)	0% (0)	8% (2)
Blood Transfusion Reaction	1% (10)	0% (0)	0% (0)	10% (1)
Total/Overall average	**968**	**6% (62)**	**1% (7)**	**23% (221)**

* Patient care includes events such as infiltrated IV's, falls, and pressure ulcers

Inpatient Variation Exists Across Hospitals in AE Rates

# Admissions in cohort	# of AEs identified	AE rate per 100 admissions	% Admissions w/ ≥ 1 AE (n)	% Preventable or probably preventable (n)
Total: 2809	**968**	**34**	**23% (655)**	**23% (221)**
468	220	47	30% (142)	23% (50)
110	38	35	23% (25)	32% (12)
148	23	16	14% (20)	22% (5)
118	27	23	19% (22)	26% (7)
589	237	40	26% (151)	22% (53)
109	22	20	17% (18)	41% (9)
104	31	30	20% (21)	16% (5)
106	16	15	13% (14)	6% (1)
643	255	40	25% (163)	19% (49)
212	45	21	18% (39)	42% (19)
202	54	27	20% (40)	20% (11)
		range 15 - 47	**range 13% - 30%**	**range 6% - 42%**

Differences between sites for **admissions with ≥ 1 AE** statistically significant p<.0001

Differences between sites for **preventable AEs** NOT statistically significant

There is some variation in AE rate by age group

Characteristics	Distribution of admissions	Distribution of admissions w/ ≥ 1 AE*	Distribution of AEs leading to permanent disability or death*	Median length of admission (days)	
				No AE	w/ AE*
Total/Overall	n=2809	n=655	n=24	3	6
Age Group					
18-44	24% (669)	16% (104)	8% (2)	3	5
45-64	30% (835)	29% (189)	29% (7)	3	6
65-84	38% (1055)	44% (291)	50% (12)	3	6
85+	9% (250)	11% (71)	13% (3)	4	6

Difference in distributions statistically significant $p < .0001$

* At least 1 AE during admission

Males More Likely to Experience Severe AEs

Characteristics	Distribution of admissions	Distribution of admissions w/ ≥ 1 AE*	Distribution of AEs leading to permanent disability or death*
Total/Overall	n=2809	n=655	n=24
Sex			
Female	55% (1553)	49% (321)	29% (7)
Male	44% (1230)	50% (332)	71% (17)
Unknown/Other	1% (26)	<1% (2)	0% (0)

Difference in distributions statistically significant p=0.0006

Difference in distributions statistically significant p=0.009

* At least 1 AE during admission

No statistical differences by ethnicity
There is variation by race

Characteristics	Distribution of admissions	Distribution of admissions w/ ≥ 1 AE*	Distribution of AEs leading to permanent disability or death*
Total/Overall	n=2809	n=655	n=24
Ethnicity			
Hispanic	6% (158)	4% (28)	13% (3)
Non-Hispanic	80% (2254)	83% (542)	79% (19)
Unknown	14% (397)	13% (85)	8% (2)
Race			
Asian	3% (94)	2% (14)	0% (0)
Black or African American	10% (284)	10% (67)	8% (2)
White	76% (2124)	79% (515)	83% (20)
Other	6% (176)	4% (26)	4% (1)
Unknown	5% (131)	5% (33)	4% (1)

Difference in distributions not statistically significant

Difference in distributions statistically significant p=0.05

* At least 1 AE during admission

Conclusions from Inpatient Study

- Adverse events were identified in nearly one in four inpatient admissions (23.6%), and approximately one fourth (22.7%) of the events were preventable.
- Of the preventable events, 20% were serious, (caused harm that resulted in substantial intervention, or prolonged recovery), 3% were life threatening, 0.5% were fatal.
- Adverse drug events were most frequent (39%), followed by surgical/procedural (30%), patient care (15%) and hospital-acquired infections (12%).
- Patient care events were especially likely to be preventable (38%)
- These findings underscore the importance of patient safety and the need for continuing improvement.

US Dept Health and Human Services Office of Inspector General: Adverse events i hospitals

- October 2018 population sample of 770 Medicare patients
- Used trigger tool and two stage chart review, looked at POA no’s.
- **A quarter of hospitalized Medicare patients experienced harm during their hospital stays**
- **Physician-reviewers determined that 43% of adverse events and temporary harm events were preventable**
- ***Medication-related harm events*** were the most common type of harm events
- CMS’s two policies on hospital-acquired conditions create payment incentives for harm prevention but do not apply to most of harm events that patients experienced
- Nearly a quarter of patients who experienced harm events required treatment that led to additional Medicare costs
- Call to action to CMS, AHRQ, other federal programs.

Christi A. Grimm
Inspector General
May 2022, OEI-06-18-00400

US Dept Health and Human Services Office of Inspector General: Adverse events i hospitals

- October 2018 population sample of 770 Medicare patients
- Used trigger tool and two stage chart review, looked at POA no's.
- **A quarter of hospitalized Medicare patients experienced harm during their hospital stays**
- **Physician-reviewers determined that 43% of adverse events and temporary harm events were preventable**
- ***Medication-related harm events*** were the most common type of harm events
- CMS's two policies on hospital-acquired conditions create payment incentives for harm prevention but do not apply to most of harm events that patients experienced
- Nearly a quarter of patients who experienced harm events required treatment that led to additional Medicare costs
- Call to action to CMS, AHRQ, other federal programs.

Christi A. Grimm
Inspector General
May 2022, OEI-06-18-00400

Implications for Improving Safety

Detecting adverse events

Specific types of adverse events

- Adverse drug events
- Surgical events
- Patient care events
  - Falls
  - Pressure ulcers
- Hospital-acquired infections

Electronic Record Should be Used to Find Adverse Events

- Already good for hospital-acquired infections
- Reasonable too for adverse drug events though needs refinement
- Mediocre to poor for DVTs/pulmonary embolism
- Hasn't yet been added for falls, pressure ulcers
- Not yet trained for decompensation
- Works very poorly for diagnostic errors
- But mandating across institutions soon would be major forward step
- Not yet set up for outpatient setting

The Potential for Artificial Intelligence to Improve Safety

By Harm Type—Bates et al, npj Digital, 2021

Hospital-Acquired Infections

Overall: already fairly big improvement

- Especially for specific types: CLABSI, VAP, CAUTI
- But levels of implementation vary

Opportunities for improvement:

- In detection
- Also for prevention in many areas

Key use cases for AI:

- Early identification of patients with infection including but not limited to sepsis
- Assistance with triage decisions in infected patients
- Linkage between isolates from multiple patients

Detecting Adverse Drug Events

Overall: substantial recent improvement here also, mainly in reducing rates of prescribing, administration errors

Opportunities for improvement:

- Most currently implement clinical decision support not yet delivering value

Key use cases:

- Which patients may experience ADEs, leveraging:
  - Genetic/genomic data
  - Clinical information
- Which patients should have specific testing for certain SNPs
- Which patients should get specific prophylaxis

Thromboembolic Disease

Overall: has been improvement here, robust evidence about which preventive strategies work

Opportunities for improvement:

- Implementation is still uneven, many patients don't get what has been shown to work, or best prevention for them

Key use cases for AI:

- Thromboembolic risk in cancer patients
- Which patients might benefit most from specific types of prophylaxis
- Which patients with thromboembolism should have further diagnostic testing

Pressure Ulcers

Overall: these still occur far too frequently, some strategies with documented benefit

Opportunities for improvement: better sensing, especially for fluid, and when a patient is not moving

- Leveraging the data that come out of hospital beds

Key use cases for AI:

- Identifying which patients are at imminent risk using both clinical data and sensing
- Determining which patients may benefit most from expensive interventions

Falls

Overall: strategies such has FALLTIPS have proven benefit, though still not implemented in most organizations

- 25% decrease in main study—Dykes, JAMA 2010

Opportunities for improvement: reduction of rates, implementation of prevention strategies for high-risk groups

Key use cases:

- Improvement of risk stratification
  - Linkage with real-time monitoring especially from sensors
- Use in new settings such as long-term care, post-discharge patients who are high-risk

Potential Role for AI in Detecting and Improving Management of Decompensation

Overall: Has been some attention e.g. with rapid response teams, but overall hasn't been very effective especially outside of ICUs

Opportunities for improvement: detection of decompensation overall and for specific reasons such as sepsis or bleeding

Key use cases:

- Early identification of decompensation overall
- Early identification of sepsis
- Early identification of post-operative bleeding
- Decompensation in a variety of settings—post discharge high-risk, long-term care

Special Article

Automated Identification of Adults at Risk for In-Hospital Clinical Deterioration

Gabriel J. Escobar, M.D., Vincent X. Liu, M.D., Alejandro Schuler, Ph.D., Brian Lawson, Ph.D., John D. Greene, M.A., and Patricia Kipnis, Ph.D.

N Engl J Med
Volume 383(20):1951-1960

The NEW ENGLAND JOURNAL of MEDICINE

Outcomes in the Eligible Population, with Comparison between the Intervention Cohort and Comparison

Table 2. Adjusted Outcomes in the Eligible Population, with Comparison between the Intervention Cohort and Comparison Cohort.*

Variable	Study Population	Adjusted Relative Risk or Hazard Rate Ratio (95% CI)
Target population		
No. of hospitalizations	43,949	
No. of patients	35,669	
ICU admission within 30 days after alert		0.91 (0.84–0.98)
Death within 30 days after alert		0.84 (0.78–0.90)
Favorable status at 30 days after alert†		1.04 (1.02–1.06)
Hospital discharge, as assessed by proportional-hazards analysis		1.07 (1.03–1.11)
Survival, as assessed by proportional-hazards analysis		0.83 (0.78–0.89)
Nontarget population		
No. of hospitalizations	504,889	
No. of patients	313,115	
ICU admission within 30 days after admission		0.94 (0.89–0.99)
Death within 30 days after admission		0.97 (0.93–1.02)
Favorable status 30 days after admission†		1.00 (0.99–1.00)
Hospital discharge, as assessed by proportional-hazards analysis		0.98 (0.97–0.99)
Survival, as assessed by proportional-hazards analysis		0.99 (0.96–1.03)

* The analysis included 548,838 hospitalizations and 326,816 patients (a patient could be included in both the target and nontarget populations, so the numbers of patients do not sum to 326,816). For the first three analyses (ICU admission, mortality, and favorable status within 30 days after an alert), the adjusted relative risk is for whether the patient was in the intervention condition (alerts led to a clinical response), as compared with patients in the comparison condition (usual care, with no alerts). For the nontarget population, the analytic approach was the same as for the target population, except that the cohorts involved patients on the ward whose condition did not trigger an alert; since there was no alert, we used 30-day mortality. We used Cox proportional-hazard models to assess the effects of the intervention on the hospital length of stay, with censoring of a patient's data at the time of death, and long-term survival (median follow-up in the target population, 0.8 years [IQR, 0.1 to 1.8]; median follow-up in the nontarget population, 1.4 years [IQR, 0.5 to 2.4]; maximum follow-up in both populations, 3.6 years). For the hospital length of stay, the hazard rate ratio refers to the instantaneous rate of discharge from the hospital divided by the instantaneous rate of discharge from the hospital in the comparison group; a rate ratio greater than 1 indicates that the intervention shortened the time to discharge. The hazard rate ratio corresponding to long-term survival refers to the long-term mortality in the intervention group as compared with the comparison group; a ratio lower than 1 indicates lower mortality in the intervention group. Additional details are provided by Harrell,[34] Basu et al.,[35] Hosmer and Lemeshow,[36] and Mihaylova et al.[37] Confidence intervals (CIs) were calculated with the use of bootstrapping to control for within-facility and within-patient correlations; see Goldstein et al.[39]

† Favorable status at 30 days indicates that, at 30 days after an alert (in the target population) or at 30 days after admission (in the nontarget population) the patient was alive, was not in the hospital, and had not been readmitted at any time.

GJ Escobar et al. N Engl J Med 2020;383:1951-1960.

Sepsis Examples

- Hospital Corporation of America (HCA) has developed a real-time tool called SPOT (Prediction and Optimization of Therapy) with 2.5 million patients
  - Estimate 8000 lives saved as a result over 5 years
- Duke also has rolled out a system called “Sepsis Watch”—trained on 50,000 patient records, 32 million datapoints
- Many other organizations working on this area

- Spectrum.ieee.org

Original Investigation ONLINE FIRST

June 21, 2021

External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients

Andrew Wong, MD[1]; Erkin Otles, MEng[2,3]; John P. Donnelly, PhD[4]; et al

- “In this cohort study of 27 697 patients undergoing 38 455 hospitalizations, sepsis occurred in 7% of the hospitalizations. The Epic Sepsis Model predicted the onset of sepsis with an area under the curve of 0.63, which is substantially worse than the performance reported by its developer.”
- AOC reported by Epic was 76-83%
- Resulted in alerts in 1/5 patients, of whom only 12% had sepsis
- Model is widely implemented

Missed and Delayed Diagnoses

Overall: appears to represent a very large problem, especially in outpatient setting—inpatient is less clear (see NAM Report)

Opportunities for improvement—many, especially specific diagnoses like pulmonary embolism; also huge opportunity for reducing delays in diagnosis especially for common malignancies (lung, colon, breast, prostate)

Key use cases:

Identifying clinical situations in which a diagnosis may have been missed • Putting together constellations of findings	Identifying scenarios in which there has been a delay longer than an acceptable alternative

Limitations/ Reflections

Many other causes of harm

- But these are the biggest—uses pareto principle

More focus on inpatient setting than others

- Outpatient setting deserves more attention

Many areas not traditional considered safety issues like decompensation are on border but represent big opportunities

Will help a lot to have better measurement of inpatient/outpatient safety routinely

Patient engagement will play a big role—have just done a study of showing patients their own risk for safety events

Conclusions

Care should be much safer than it currently is in all countries

- Is an ethical imperative

Equity was not a big issue in this inpatient study, much more of a concern in outpatient setting, and emergency rooms

Technology exists now to allow organizations to measure all types of safety issues in real time

- Policies should be implemented to require organizations to implement one of these technologies
- Public reporting of key types of harm should be required
- Federal government should sponsor the definition of standard metrics for the key types harm

Every organization should have a concrete plan for improving safety

- Tracking is especially important—should have a dashboard
- Some areas simply involve better process
- Others involve new technology—like leveraging AI