

Changes in Rates of Autopsy-Detected Diagnostic Errors Over Time

A Systematic Review

Kaveh G. Shojania, MD

Elizabeth C. Burton, MD

Kathryn M. McDonald, MM

Lee Goldman, MD, MPH

BEGINNING IN 1912¹ AND CONTINUING TO 2002,²⁻⁶ researchers have documented substantial discrepancies between clinical diagnoses and findings at autopsy. Time series from single institutions have examined trends in these diagnostic discrepancies and, with only 1 exception,⁷ have found no significant decreases over time.⁸⁻¹²

A possible explanation for the stability of these rates is increased selection by clinicians. In 1994, the last year for which national data exist, the autopsy rate for all nonforensic deaths decreased to less than 6%¹³ compared with average rates of 30% to 40% in the 1960s.¹⁴ With progressively fewer autopsies performed over time, clinical selection for diagnostically challenging cases might offset true gains in diagnostic accuracy. However, several prospective studies have shown clinicians to have little ability to identify cases that will yield “diagnostic surprises,”¹⁵⁻¹⁸ so clinical selection might exert little effect on rates of autopsy-detected diagnostic errors.

As part of a broader report on the autopsy as a tool for quality measurement and improvement,¹⁹ we systematically reviewed the literature to estimate the fre-

Context Substantial discrepancies exist between clinical diagnoses and findings at autopsy. Autopsy may be used as a tool for quality management to analyze diagnostic discrepancies.

Objective To determine the rate at which autopsies detect important, clinically missed diagnoses, and the extent to which this rate has changed over time.

Data Sources A systematic literature search for English-language articles available on MEDLINE from 1966 to April 2002, using the search terms *autopsy*, *postmortem changes*, *post-mortem*, *postmortem*, *necropsy*, and *posthumous*, identified 45 studies reporting 53 distinct autopsy series meeting prospectively defined criteria. Reference lists were reviewed to identify additional studies, and the final bibliography was distributed to experts in the field to identify missing or unpublished studies.

Study Selection Included studies reported clinically missed diagnoses involving a primary cause of death (major errors), with the most serious being those likely to have affected patient outcome (class I errors).

Data Extraction Logistic regression was performed using data from 53 distinct autopsy series over a 40-year period and adjusting for the effects of changes in autopsy rates, country, case mix (general autopsies; adult medical; adult intensive care; adult or pediatric surgery; general pediatrics or pediatric inpatients; neonatal or pediatric intensive care; and other autopsy), and important methodological features of the primary studies.

Data Synthesis Of 53 autopsy series identified, 42 reported major errors and 37 reported class I errors. Twenty-six autopsy series reported both major and class I error rates. The median error rate was 23.5% (range, 4.1%-49.8%) for major errors and 9.0% (range, 0%-20.7%) for class I errors. Analyses of diagnostic error rates adjusting for the effects of case mix, country, and autopsy rate yielded relative decreases per decade of 19.4% (95% confidence interval [CI], 1.8%-33.8%) for major errors and 33.4% (95% [CI], 8.4%-51.6%) for class I errors. Despite these decreases, we estimated that a contemporary US institution (based on autopsy rates ranging from 100% [the extrapolated extreme at which clinical selection is eliminated] to 5% [roughly the national average]), could observe a major error rate from 8.4% to 24.4% and a class I error rate from 4.1% to 6.7%.

Conclusion The possibility that a given autopsy will reveal important unsuspected diagnoses has decreased over time, but remains sufficiently high that encouraging ongoing use of the autopsy appears warranted.

JAMA. 2003;289:2849-2856

www.jama.com

Author Affiliations: Department of Medicine, University of California, San Francisco (Drs Shojania and Goldman); Department of Pathology and Laboratory Medicine, Baylor Health Care System, Dallas-Fort Worth, Tex (Dr Burton); and Center for Primary Care and Outcomes Research, Stanford University,

Stanford, Calif (Ms McDonald).

Corresponding Author and Reprints: Kaveh G. Shojania, MD, Department of Medicine, University of California, 533 Parnassus Ave, Room U137, Box 0131, San Francisco, CA 94143 (e-mail: shojania@medicine.ucsf.edu).

See also Patient Page.

quency with which autopsy reveals important, clinically missed diagnoses. We sought to assess the degree to which this frequency has changed over time, and the extent to which clinical selection for diagnostically challenging cases accounts for the substantial error rates that continue to be reported in autopsy studies.

METHODS

Search Strategy

We searched the MEDLINE database for English-language articles (1966–April 2002) using Medical Subject Heading terms *autopsy* and *postmortem changes*, and the title words, *autopsy*, *postmortem*, *postmortem*, *necropsy*, and *posthumous*. We then applied terms capturing aspects of study design (eg, *epidemiologic studies*, *clinical trials*) and topics relating to diagnosis (eg, *diagnostic errors*, *diagnostic techniques and procedures*) or error (eg, *medical error*, *iatrogenic disease*, *safety*). Reference lists from all relevant articles were reviewed to identify additional studies, and the final bibliography was distributed to experts in the field to identify missing or unpublished studies.

Study Selection

Included studies met the following criteria:

1. Consecutive autopsies with well-defined selection criteria (eg, all adults dying after hospital arrival and undergoing autopsy during a specified period) or random samples from such series; “convenience samples” and consecutive series missing more than 20% of eligible cases were excluded.

2. Clinical diagnoses derived from autopsy request forms submitted by clinicians or chart review performed by study investigators; assessments of clinical diagnoses based primarily on death certificates were excluded.

3. Classification of autopsy-detected errors in clinical diagnoses according to generally accepted classification schemes^{8,20}; major errors defined as clinically missed diagnoses involving a principal underlying disease or primary cause of death; and class I errors, major errors that, had they been

detected during life, “would,” “could,” “possibly,” or “might” have affected patient prognosis or outcome (at a minimum, discharge from the hospital alive). Studies that made no distinction between changes in management and changes in outcome were deemed to be reporting major errors only.

Studies reporting autopsy data from multiple institutions or observation periods were analyzed as separate series whenever possible. We did not restrict our review to studies of in-hospital deaths, although an overwhelming majority of studies involved inpatient autopsies predominantly or, in many cases, exclusively.

Quantitative Analysis

Diagnostic error rates were modeled using logistic regression analyses with country, study period, case mix, and autopsy rate as predictors and including a random study effect.²¹ Hospital teaching status was not included as a predictor, because too few studies involved nonteaching hospitals and because the nature of the teaching status was often unclear.

Autopsy rates and time were modeled as continuous variables, with the value for time designated as the midpoint of the study period. Country was simplified to United States or non-United States, but case mix was modeled as a nonordinal variable with the categories of (1) general autopsies, (2) adult medical, (3) adult intensive care, (4) adult or pediatric surgery, (5) general pediatrics or pediatric inpatients, (6) neonatal or pediatric intensive care, and (7) other. The first category, which constituted the base case mix in the regression analysis, included series reporting general autopsies (all ages, specialties, and settings), general inpatients (all ages and specialties), and general adult inpatients. We combined these 3 populations because many studies provided insufficient detail to permit reliable distinctions between these 3 and because the contribution of adult inpatients dominated samples of all 3 types.

In anticipation of methodological heterogeneity among the studies, we ab-

stracted each article for key study features plausibly related to observed error rates. These study features included (1) cohort design (prospective vs retrospective); (2) clarity of error definition (whether class I and major errors were defined using illustrative examples or if the results included a complete listing of all clinical-autopsy discrepancies designated as class I or major errors); (3) source of clinical diagnoses (chart review vs autopsy request forms); (4) involvement of clinicians in classifying errors. The regression models for major and class I error rates incorporated each of these methodological characteristics as categorical variables.

RESULTS

Included Autopsy Series

We identified 45 studies^{2–8,12,17,18,20,22–55} reporting a total of 53 distinct autopsy series meeting our inclusion criteria (TABLE 1). More studies reported major errors than reported class I errors (42 and 37 series, respectively). Twenty-six series reported both types of errors and just over half of the series (27) involved US institutions.

Although numerous studies met our inclusion criteria and offered a wide range of predictor variables for the regression model (TABLE 2), many of the studies exhibited methodological limitations. The vast majority of autopsy series were assembled retrospectively, and only half performed chart review to obtain clinical diagnoses. Clinicians played a primary role in classifying diagnostic errors in two thirds of series (TABLE 3).

Diagnostic Error Rates

The median major error rate was 23.5%, although rates ranged from 4.1% to 49.8%, with the upper bound reflecting the only series focused on postoperative deaths.⁴⁴ The median class I error rate was 9.0%, but rates ranged from 0% to 20.7%, with the upper bound again corresponding to the series focused on postoperative deaths.⁴⁴ The study reporting zero class I errors involved pediatric deaths from an emergency department.⁴⁹ The authors attributed the absence of class I errors to

Table 1. Major or Class I Error Rates of Autopsy Series Between 1959 and 1999

Year	Autopsy Series	Location	Autopsy Rate, %	Case Mix	No. of Autopsies	Major Error Rate, % (95% CI)	Class I Error Rate, % (95% CI)
General Autopsies, General Inpatients, and General Adult Inpatients (Base Case Mix)							
1959	Christian-Albrechts University Hospital ¹²	Germany	88	General inpatients	100	NA	7 (3-14)
1960	Peter Bent Brigham Hospital ⁸	Boston	75	Adult inpatients	100	22 (15-32)	8 (4-16)
1969	Christian-Albrechts University Hospital ¹²	Germany	82	General inpatients	100	NA	12 (7-20)
1970	Peter Bent Brigham Hospital ⁸	Boston	71	Adult inpatients	100	23 (15-33)	12 (7-20)
1975-1977	South Lothian District hospitals ^{17,22}	Scotland	25	General inpatients	1152	38 (36-41)	NA
1978	University of Edinburgh ²³	Scotland	64	General inpatients	154	15 (10-22)	NA
1978-1979	Winnipeg Health Sciences Center ²⁴	Canada	26	Adult inpatients	200	NA	1.5 (0.4-4.7)
1979	Royal Adelaide Hospital ²⁵	Australia	45	Adult inpatients	99	NA	2.0 (0.4-7.8)
1979	Christian-Albrechts University Hospital ¹²	Germany	58	General inpatients	100	NA	12 (7-20)
1980	Peter Bent Brigham Hospital ⁸	Boston	38	Adult inpatients	100	21 (14-30)	11 (6-19)
1980	39 Institutions performing autopsies ²⁶	Connecticut	14	General autopsies	272	NA	4.0 (2.1-7.3)
1981-1984	Belgrade University School of Medicine ²⁷	Yugoslavia	12	General inpatients	2145	29 (26.9-30.8)	NA
1983	VA Medical Center ²⁸	White River Junction	60	General inpatients	111	12.6 (7.3-20.6)	NA
1983-1988	Chandigarh Postgraduate Institute ²⁹	India	25	Adult inpatients	1000	31.7 (28.8-34.7)	NA
1984	32 University and community hospitals ²⁰	United States	30	General inpatients	2067	34 (32-36)	13 (12-15)
1984-1985	Brigham and Women's Hospital ⁸	Boston	37	Adult inpatients	175	23 (17-30)	11 (7-17)
1984-1985	Emerson Hospital ¹⁸	Concord, Mass	26	Adult inpatients	58	33 (21-46)	12 (5-24)
1985	Northwestern Memorial Hospital ³¹	Chicago	36	General inpatients	142	23 (17-31)	NA
1988	St Vincent's Hospital ³²	Australia	22	General inpatients	139	NA	6 (3-11)
1989	Christian-Albrechts University Hospital ¹²	Germany	36	General inpatients	100	NA	11 (6-19)
1990	Peterborough District Health Authority ³³	United Kingdom	13	General inpatients	63	19 (11-31)	NA
1994	University of Pittsburgh Medical Center ³⁴	Pittsburgh	19	General inpatients	172	34 (27-42)	NA
1997	Prince of Wales Hospital ³⁵	Hong Kong	17.7	General autopsies	332	23.5 (19.1-28.5)	NA
Adult Medical Patients							
1972	Zurich University Hospital ⁷	Switzerland	94	Medical patients	100	30 (21.4-40.1)	16 (10-25)
1982	Zurich University Hospital ⁷	Switzerland	89	Medical patients	100	18 (11.3-27.2)	9 (4-17)
1984	Leiden University Hospital ³⁶	Netherlands	47	Medical patients	133	41 (33-50)	NA
1992	Zurich University Hospital ⁷	Switzerland	89	Medical patients	100	14 (8.1-22.7)	7 (3-14)
1992-1993	Ben Taub General Hospital ³⁷	Houston	16	Medical patients	110	23.6 (16.3-32.9)	10.9 (6.3-16.8)
Adult Intensive Care Patients							
1983-1985	Hospital Central de la Cruz Roja ³⁸	Spain	51	Adult ICU	100	22 (14.6-31.6)	7 (3.1-14.4)
1986-1987	6 VA hospitals ³⁹	United States	43	Medical ICU	172	27.9 (21.5-35.3)	12 (8-18)
1986-1992	Hartford Hospital ⁴⁰	New Haven, Conn	29	Surgical ICU	149	23 (16.5-30.6)	11 (6.5-17.1)
1994-1995	Cleveland Clinic ²	Cleveland, Ohio	23	Medical ICU	91	19.8 (12.4-29.7)	NA
1994-1995	Hershey Medical Center ⁴¹	Hershey, Pa	31	Medical-coronary ICU	41	27 (15-43)	NA
1996-1999	Gloucestershire Royal Hospital ³	United Kingdom	40	Adult ICU	97	23.7 (15.9-33.6)	4.1 (1.3-10.8)
1998-1999	General Hospital Celje ⁵	Slovenia	47	Medical ICU	126	NA	9.5 (5.2-16.4)

(continued)

Table 1. Major or Class I Error Rates of Autopsy Series Between 1959 and 1999 (cont)

Year	Autopsy Series	Location	Autopsy Rate, %	Case Mix	No. of Autopsies	Major Error Rate, % (95% CI)	Class I Error Rate, % (95% CI)
Adult or Pediatric Surgical Patients							
1974-1975	San Francisco General Hospital and 31 other hospitals ⁴⁵	San Francisco and Orange County, Calif	100	Motor vehicle fatalities	182	NA	12.6 (8.3-18.6)
1984-1988	University of Texas Medical Branch ⁴²	Galveston	73	Adult and pediatric surgery	409	30.3 (26-35.1)	7.8 (5.5-11)
1985-1995	Johns Hopkins Hospital ⁴³	Baltimore	24	Cardiac surgery	147	38.8 (31.0-47.2)	NA
1986-1988	Royal Victoria Hospital ⁴⁴	Northern Ireland	23	Deaths within 30 days of surgery	213	49.8 (42.9-56.6)	20.7 (15.6-26.8)
1997-1998	Ryder Trauma Center ⁶	Miami	97	Adult and pediatric trauma or burn patients dying in ICU	153	15.7 (10.5-22.6)	2.6 (0.8-7.0)
General Pediatrics or Pediatric Inpatients							
1984-1993	Lutheran General Children's Hospital ⁴⁷	Park Ridge, Ill	36	General pediatrics	107	13.1 (7.6-21.3)	6.5 (2.9-13.5)
1989-1994	University of Rochester Medical Center ⁵⁰	Rochester, NY	74	General pediatrics	157	6.4 (3.3-11.7)	NA
1992	Children's Hospital of New Jersey ⁴⁶	Newark	29	General pediatrics	23	13.0 (3.4-34.7)	4.3 (0.2-24.0)
Neonatal or Pediatric Intensive Care Patients							
1984-1993	Lutheran General Children's Hospital ⁴⁷	Park Ridge, Ill	61	Neonatal ICU	296	0.3 (0.02-2.2)	11.8 (8.5-16.2)
1985-1990	Toronto Hospital for Sick Children ⁵²	Canada	62	Neonatal ICU	338	18.9 (15.0-23.6)	2.1 (0.9-4.4)
1985-1992	North Shore University Hospital ⁵¹	Manhasset, NY	26	Pediatric ICU	50	28.0 (16.7-42.7)	10 (3.7-22.6)
1991-1997	Royal Alexandra Hospital for Children ⁵³	Australia	40	Neonatal ICU	91	NA	5.5 (2.0-12.9)
1995-1996	King Edward Memorial Hospital ⁴	India	82	Neonatal ICU	197	26.9 (21.0-33.8)	12.2 (8.1-17.8)
Other							
1981	Salford Health Authority ⁵⁴	United Kingdom	35	Inpatients >85 years	129	31 (23.3-39.8)	NA
1981-1983	Beth Israel Hospital ⁵⁵	Boston	27	Adult inpatients post-CPR	130	13.8 (8.6-21.3)	NA
1984-1988	University of Massachusetts Medical Center ³⁰	Worcester	32	Deaths in ED	244	4.1 (2.1-7.6)	1.6 (0.5-4.4)
1985-1989	Children's Hospital of Western Ontario ⁴⁹	Canada	75	Pediatric deaths in ED	52	15.4 (7.3-28.6)	0 (0-8.6)
1992-1993	Ben Taub General Hospital ⁵⁷	Houston	16	Adult inpatients with AIDS	42	33.3 (20.0-49.6)	9.5 (3.1-23.6)

Abbreviations: CI, confidence interval; CPR, cardiopulmonary resuscitation; ED, emergency department; ICU, intensive care unit; NA, data not available.

the high proportion of deaths following cardiac arrest, in which survival depends predominantly on the adequacy of resuscitation rather than the accuracy of clinical diagnosis.

Effects of Country, Autopsy Rates, and Case Mix

Compared to US studies, autopsy series from outside the United States exhibited a slight, but statistically significant trend toward higher major error rates (odds ratio [OR], 1.15; 95% con-

fidence interval [CI], 1.01-1.31; $P = .03$). For class I errors, the effect was of comparable magnitude and bordered on statistical significance (OR, 1.26; 95% CI, 0.99-1.59; $P = .06$) (TABLE 4).

Autopsy rates ranged from 12% to 100% (median, 37.0%). Relative to the error rate in 1980 (the midpoint of the 40-year period spanned by the included studies), major errors decreased at a rate of 12.4% (95% CI, 7.0%-17.6%) for every 10% increase in autopsy rates. Class I errors decreased at a

rate of 17.4% (95% CI, 6.6%-27.1%) for every 10% increase in autopsies.

Autopsy series restricted to surgical patients reported significantly higher rates of both major errors (OR, 2.16; 95% CI, 1.53-3.06) and class I errors (OR, 3.01; 95% CI, 1.66-5.43). Series limited to adult medical patients reported higher class I error rates (OR, 1.84; 95% CI, 1.06-3.20); US series from adult intensive care units also had higher class I error rates (OR, 2.12; 95% CI, 1.42-3.16). Conversely, pediatric series reported sig-

nificantly lower rates of major errors, and series involving pediatric or neonatal intensive care autopsies reported significantly lower class I error rates (OR, 0.56; 95% CI, 0.32-0.98) (Table 4).

Impact of Methodological Features

None of the 4 methodological features shown in Table 3 significantly affected major error rates, but 2 methodological features significantly affected class I error rates (Table 4). Studies conducted prospectively reported higher class I error rates (OR, 1.63; 95% CI, 1.19-2.23), as did studies in which clinicians played active roles in classifying errors (OR, 2.09; 95% CI, 1.31-3.34).

Trends in Diagnostic Errors Over Time

Adjusting for the effects of country, case mix, and autopsy rates, major errors significantly decreased over time, with a relative reduction of 19.4% per decade (95% CI, 1.8%-33.8%). Adjusting for these same factors as well as the 2 significant methodological features (prospective study design and clinicians participation in classifying errors), class I error rates also decreased significantly over time, with a relative reduction of 33.4% per decade (95% CI, 8.4%-51.6%).

Despite these decreases, we estimated that a contemporary US institution with an autopsy rate of 5% (roughly the national average¹³), could observe a major error rate of 24.4% (95% CI, 18.8%-31.1%) and a class I error rate of 6.7% (95% CI, 3.8%-11.4%) (FIGURE 1 and FIGURE 2, respectively). With an autopsy rate of 37% (the median rate in the included studies), major and class I error rates in the same institution would be estimated as 17.4% (Figure 1) and 5.8% (Figure 2). This autopsy rate is much higher than the rates of 15% to 20% typically achieved in contemporary teaching hospitals⁵⁶ and, therefore, has less clinical selection. Even with extrapolation to an autopsy rate of 100% (to eliminate the effect of clinical selection completely), a US institution in 2000 would be estimated to report a major error rate of 8.4% (95% CI, 5.2%-13.1%) and a class

I error rate of 4.1% (95% CI, 1.6%-9.9%) (Figure 1 and Figure 2, respectively).

COMMENT

By analyzing the results of 53 distinct autopsy series over a 40-year period, we have shown statistically significant decreases over time for major and class I diagnostic errors detected at autopsy. By contrast, individual studies comparing rates of autopsy-detected diagnostic errors from different periods have found strikingly unchanged error rates.^{8,12,57,58} These previous results almost certainly reflect inadequate power, as well as the competing effects of improvements over time and increased clinical selection as autopsy rates decrease. In fact, the only study with high and nearly equal autopsy rates in all periods examined showed a significant decrease in major errors over time.⁷

The present data suggest that, among the approximately 850 000 individuals dying in US hospitals each year,^{59,60} a major diagnosis remains clinically undetected in at least 8.4% of cases (71 400 deaths). The data also suggest that approximately 34 850 of these patients might have survived to discharge had misdiagnosis not occurred, but this estimate depends on the accuracy of the designator of class I error. Although, this second number is more speculative, given the dependence of class I error estimates on methodological features of the primary studies, it can be considered in the context of the Institute of Medicine's estimates of 44 000 to 98 000 preventable deaths per year

due to medical error.⁶¹ These latter estimates have been debated,⁶²⁻⁶⁴ but the studies from which they were derived may not have detected many of the errors reported in our analysis.

A major limitation of any systematic review is the possibility of publication bias. Problems with existing methods of assessing publication bias^{65,66} are compounded by the opposing directions in which publication bias might operate. Lack of interest might result in fewer published reports of low error rates,

Table 2. Demographic Features of Included Autopsy Series*

Characteristic	No. of Studies Reporting	
	Major Errors (n = 42)	Class I Errors (n = 37)
Study period, y		
1960-1969†	1	3
1970-1979	4	6
1980-1989	23	19
1990-1999	14	9
Autopsy rate, %		
10.1-20	6	3
20.1-40	18	15
40.1-60	5	5
60.1-80	8	6
>80%	5	8
Case-mix		
General autopsy	15	14
Adult medical	5	4
Adult ICU	6	5
Surgery (adult or pediatric)	3	4
General pediatric	3	2
Pediatric or neonatal ICU	3	4
Other	7	4
Country		
United States	25	19
Other	17	18

Abbreviation: ICU, intensive care unit.
 *Twenty-six series reported both error types.
 †Only 1 study¹² included a series older than 1960. We listed this series in the 1960-1969 group, but retained the study period as 1959 for the regression analysis.

Table 3. Methodological Features of Included Autopsy Series

Characteristic	No. (%) of Autopsy Series		
	Major (n = 42)	Class I (n = 37)	All (n = 53)
Prospective instead of retrospective design*	11 (26)	5 (14)	11 (21)
Clinical diagnoses obtained from chart review†	20 (48)	18 (49)	27 (51)
Error classification was defined with illustrative examples, or a complete list of discrepant diagnoses was provided	30 (71)	30 (81)	39 (74)
Clinicians involved in error judgments	30 (71)	29 (78)	37 (70)

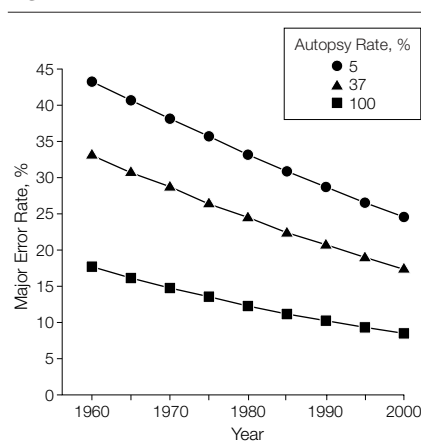
*Eleven of the retrospective series used random samples of all autopsies rather than all consecutive autopsies in the study period.
 †Indicates diagnoses from the forms or protocols used by clinicians when requesting autopsy. Series obtaining clinical diagnoses from conventional death certificates were excluded.

Table 4. Effects of Various Study Characteristics on Diagnostic Error Rates

Characteristic	Odds Ratio (95% CI)*			
	Major Errors		Class I Errors	
	All Studies	US Only	All Studies	US Only
	General†			
Study year	0.98 (0.96-0.99)	0.98 (0.96-1.00)	0.96 (0.92-1.00)	0.95 (0.88-1.02)
Autopsy rate	0.26 (0.14-0.49)	0.15 (0.06-0.34)	0.16 (0.04-0.61)	0.24 (0.02-2.75)
Other country	1.15 (1.01-1.31)	NA	1.26 (0.99-1.59)	NA
	Type of Patient‡			
Adult medical	1.33 (0.92-1.91)	0.91 (0.50-1.65)	1.84 (1.06-3.20)	2.14 (1.09-4.20)
Adult ICU	1.09 (0.79-1.51)	1.23 (0.88-1.72)	1.21 (0.74-1.97)	2.12 (1.42-3.16)
Surgery	2.16 (1.53-3.06)	2.75 (1.86-4.06)	3.01 (1.66-5.43)	1.87 (0.94-3.71)
Pediatric	0.50 (0.29-0.84)	0.59 (0.36-0.95)	1.03 (0.40-2.67)	1.17 (0.58-2.34)
Pediatric and neonatal intensive care	1.12 (0.78-1.63)	0.99 (0.63-1.54)	0.56 (0.32-0.98)	0.36 (0.16-0.82)
	Methodological Feature			
Prospective study design	1.04 (0.89-1.21)	1.00 (0.89-1.13)	1.63 (1.19-2.23)	1.17 (0.78-1.76)
Clinical diagnoses obtained from chart review	0.88 (0.70-1.12)	0.93 (0.88-1.72)	0.93 (0.75-1.14)	1.90 (0.52-6.87)
Error definition illustrated or all discrepancies listed	1.00 (0.86-1.17)	1.20 (1.03-1.40)	0.99 (0.70-1.41)	2.08 (1.18-3.66)
Clinicians involved in classifying errors	1.12 (0.86-1.45)	0.97 (0.78-1.21)	2.09 (1.31-3.34)	0.51 (0.14-1.85)

Abbreviations: CI, confidence interval; ICU, intensive care unit; NA, not applicable; US, United States.
 *Adjusting for all other characteristics in this table, except that major errors were analyzed without adjustment for any of the methodologic features and class I errors were analyzed using only the 2 significant methodologic predictors (prospective study design and clinicians involved in classifying errors).
 †Odds ratios for listed categories of case mix are relative to studies involving case mix (general autopsies), general inpatient autopsies, and general adult inpatients.
 ‡Changes are all relative to the values obtained for the base year (1980), general autopsies, country (United States), and mean autopsy rates (46.3% for major errors and 51.3% for class I errors).

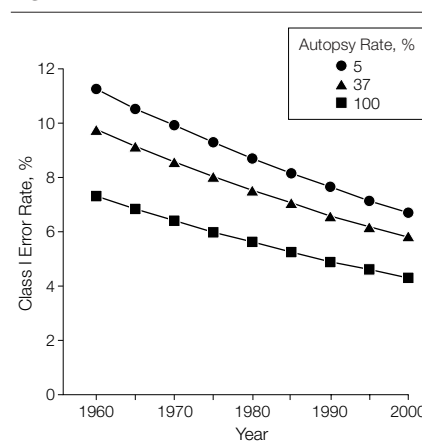
Figure 1. Major Error Rates



Major error rates were generated from the regression model, setting country to United States and using 3 different autopsy rates: 5% (roughly the national average¹³), 37% (the median rate in the included studies), and 100%. Extrapolation to an autopsy rate of 100% minimizes the effect of clinical selection and thus provides a lower bound for the major error rate.

whereas self-censorship might reduce reports of high error rates. Regardless of the true net effect of publication bias on autopsy studies, if this effect has remained reasonably stable, the observed decrease in published error rates over time would still be meaningful.

Figure 2. Class I Error Rates



Class I error rates were generated from the regression model, setting country to United States and using 3 different autopsy rates: 5% (roughly the national average¹³), 37% (the median rate in the included studies) and 100%. Extrapolation to an autopsy rate of 100% minimizes the effect of clinical selection and thus provides a lower bound for the class I error rate. Adjusting for the effects of study design (prospective vs retrospective) and clinician involvement in classifying errors significantly impacts estimates from earlier periods, but the adjusted and unadjusted estimates converge over time. For 2000, estimated class I error rates are virtually identical regardless of primary study design and clinician involvement in classifying errors.

Only 5 studies^{7,31,39,41,52} addressed the issue of reproducibility for the classification of autopsy-detected diagnos-

tic errors, and none provided sufficient detail to permit calculation of formal measures of agreement. The issue of reproducibility is particularly important for class I errors, as no study used validated criteria to guide reviewers' judgments about affects on prognosis, which are known to exhibit substantial variability.⁶²

Even more fundamental than reproducibility of the error classifications is the question of the autopsy's characteristics as a diagnostic test. Determining the sensitivity of any criterion standard, including the autopsy, presents difficulties. As reviewed in greater detail elsewhere,¹⁹ technically adequate autopsies fail to establish the cause of death in 1% to 5% of cases, although some studies have reported substantially higher rates of persistent diagnostic uncertainty after autopsy,⁴³ especially in perinatal deaths.⁶⁷⁻⁶⁹

Only 1 study⁷⁰ has assessed agreement among pathologists in determining principal underlying diseases and causes of death. Four pathologists independently reviewing 35 autopsies reported excellent to near perfect agreement for determining the principal disease (ie, underlying cause of death),

with κ values between 0.83 and 0.97 for the different pathologist pairs. For assignments of the immediate cause of death, however, the pathologists exhibited only moderate to substantial agreement (κ values ranging from 0.43-0.75).

We used the term error throughout our analysis because of its ubiquitous presence in the autopsy literature. However, it remains unclear to what extent clinically missed diagnoses represent errors per se, rather than acceptable limits of antemortem diagnosis in the face of atypical clinical presentations. In fact, because the vast majority of autopsy studies come from teaching hospitals, published autopsy series may be enriched for atypical cases. Nonetheless, the autopsy has historically helped define how cases that previously appeared atypical could more commonly be recognized antemortem. Repeated detection of certain missed diagnoses may result in the recognition that some patterns of presentation are more typical than previously appreciated.

For many physicians, interest in the autopsy as a means of detecting clinically missed diagnoses is undoubtedly offset by concerns over litigation. Only 1 study³⁴ explicitly addressed the question of whether autopsy findings influence malpractice claims. In this series of 176 autopsies from the University of Pittsburgh Medical Center (Pittsburgh, Pa) in 1994, follow-up of all cases after the statute of limitations on malpractice suits had expired identified only 1 malpractice suit. Review of the hospital record indicated that the intent to proceed to litigation in that case had become clear prior to the patient's death.

In addition to their intrinsic clinical interest, missed diagnoses detected at autopsy may have important implications for research. Health services researchers are accustomed to the problem that administrative databases contain systematic errors and biases compared with the medical record.⁷¹⁻⁷³ The data presented here indicate that the medical record itself contains substantial inaccuracies regarding the principal diagnoses causing or contributing to death. Since principal diagnoses and causes of death are de-

termined without autopsy in the vast majority of cases, vital statistics, clinical registries, and even randomized trials capture incorrect causes of death at rates comparable with the major error rates in our analysis. These inaccuracies have important policy implications, as major funding and policy decisions derive in part from vital statistics and other estimates of disease burden.⁷⁴⁻⁷⁶

Correcting such inaccuracies would not require substantial increases in autopsies at all hospitals. Perhaps a small group of hospitals funded to perform autopsies in a high percentage of deaths and according to a uniform protocol could generate accurate error rates appropriate for correcting the information contained in routinely generated death certificates and other epidemiological databases. Data from such a program would also provide the opportunity to develop an approach to enhancing the selection of autopsies likely to reveal important unsuspected diagnoses. Explicit selection of autopsy cases on the basis of diagnostic uncertainty would represent an advance over current autopsy selection, which is likely determined in large part by patients' demographic characteristics (especially age⁷⁷⁻⁷⁹) and by clinicians' comfort in requesting autopsy.^{80,81} Most importantly, further research conducted in centers with high autopsy rates would permit development of strategies for using autopsy findings to improve subsequent clinical performance.

Author Contributions: Study concept and design: Shojania, Burton, McDonald, Goldman.

Acquisition of data: Shojania.

Analysis and interpretation of data: Shojania, Burton, McDonald, Goldman.

Drafting of the manuscript: Shojania.

Critical revision of the manuscript for important intellectual content: Shojania, Burton, McDonald, Goldman.

Statistical expertise: Shojania.

Obtained funding: Shojania, McDonald, Goldman.

Administrative, technical, or material support: McDonald.

Study supervision: Goldman.

Funding/Support: The research for this article was funded by Agency for Healthcare Research and Quality contract No. 290-970013 to the University of California San Francisco-Stanford Evidence-based Practice Center. Dr Goldman also receives support from an unrestricted gift from Clarke W. Swanson and family.

Acknowledgment: We thank Preetha A. Basaviah, MD, Thomas E. Baudendistel, MD, and Chad Roach, BA, for assistance with article review and abstraction;

Alan Bostrom, PhD, and Charles E. McCulloch, PhD, for statistical expertise; and Susan B. Nguyen, BA, for technical assistance in manuscript preparation. We also gratefully acknowledge the individuals who advised us in preparing the technical report from which this work is in part derived: Lisa A. Bero, PhD, Kevin Bove, MD, Stephen A. Geller, MD, Gordon Guyatt, MD, Randy L. Hanzlick, MD, Richard J. Hausner, MD, Lee H. Hilborne, MD, MPH, Richard E. Horowitz, MD, Grover M. Hutchins, MD, George D. Lundberg, MD, Joanne Lynn, MD, Peter Nemetz, PhD, Jean Olson, MD, Eliseo Perez-Stable, MD, Thomas Russell, MD, Paul A. Raslavicus, MD, Ellen Shaffer, PhD, Patricia Styer, PhD, Victor Weedn, MD, JD.

Disclaimer: The authors are responsible for the content of this article, including any clinical or treatment recommendations. No statement in this article should be construed as an official position of the Agency for Healthcare Research and Quality or of the US Department of Health and Human Services.

REFERENCES

1. Cabot RC. Diagnostic pitfalls identified during a study of three thousand autopsies. *JAMA*. 1912;59:2295-2298.
2. Tai DY, El-Bilbeisi H, Tewari S, Mascha EJ, Wiedemann HP, Arroliga AC. A study of consecutive autopsies in a medical ICU: a comparison of clinical cause of death and autopsy diagnosis. *Chest*. 2001;119:530-536.
3. Twigg SJ, McCrerrick A, Sanderson PM. A comparison of post mortem findings with post hoc estimated clinical diagnoses of patients who die in a United Kingdom intensive care unit. *Intensive Care Med*. 2001;27:706-710.
4. Kabra NS, Udani RH. Correlation between clinical diagnoses at the time of death and autopsy findings in critically sick neonates at a regional neonatal intensive care unit in India. *J Trop Pediatr*. 2001;47:295-300.
5. Podbregar M, Voga G, Krivec B, Skale R, Pareznik R, Gabrsek L. Should we confirm our clinical diagnostic certainty by autopsies? *Intensive Care Med*. 2001;27:1750-1755.
6. Ong AW, Cohn SM, Cohn KA, et al. Unexpected findings in trauma patients dying in the intensive care unit: results of 153 consecutive autopsies. *J Am Coll Surg*. 2002;194:401-406.
7. Sonderegger-Iseli K, Burger S, Muntwyler J, Salomon F. Diagnostic errors in three medical eras: a necropsy study. *Lancet*. 2000;355:2027-2031.
8. Goldman L, Sayson R, Robbins S, Cohn LH, Bettmann M, Weisberg M. The value of the autopsy in three medical eras. *N Engl J Med*. 1983;308:1000-1005.
9. Carvalho FM, Widmer MR, Cruz M, Palomo V, Cruz C. Clinical diagnosis versus autopsy. *Bull Pan Am Health Organ*. 1991;25:41-46.
10. Poli L, Pich A, Zanolchi M, Fonte G, Bo M, Fabris F. Autopsy and multiple pathology in the elderly. *Gerontology*. 1993;39:55-63.
11. Veress B, Alafuzoff I. A retrospective analysis of clinical diagnoses and autopsy findings in 3,042 cases during two different time periods. *Hum Pathol*. 1994;25:140-145.
12. Kirch W, Schafii C. Misdiagnosis at a university hospital in 4 medical eras. *Medicine (Baltimore)*. 1996;75:29-40.
13. Burton EC, Nemetz PN. Medical error and outcomes measures: where have all the autopsies gone? *Medscape General Medicine*. 2000;2000:E8. Available at: <http://www.medscape.com/medscape/GeneralMedicine/journal/2000/v2002.n2002/mgm0428.burt/mgm0428.burt.html>.
14. Walsh CJ. Autopsy rates in US hospitals. *PAS Rep*. 1974;12:1-4.
15. Heasman MA, Lipworth L. *Accuracy of Certification of Cause of Death*. London, England: Her Majesty's Stationery Office; 1966.

16. Britton M. Diagnostic errors discovered at autopsy. *Acta Med Scand*. 1974;196:203-210.
17. Cameron HM, McGoogan E. A prospective study of 1152 hospital autopsies, I: inaccuracies in death certification. *J Pathol*. 1981;133:273-283.
18. Landefeld CS, Chren MM, Myers A, Geller R, Robbins S, Goldman L. Diagnostic yield of the autopsy in a university hospital and a community hospital. *N Engl J Med*. 1988;318:1249-1254.
19. Shojania KG, Burton EC, McDonald KM, Goldman L. *The Autopsy as an Outcome and Performance Measure (Evidence Report/Technology Assessment No. 58)*. Rockville, Md: Agency for Healthcare Research and Quality; 2002. Publication No. 02-E020. Available at: <http://www.ahrq.gov/clinic/evrptfiles.htm#autopsy>.
20. Battle RM, Pathak D, Humble CG, et al. Factors influencing discrepancies between premortem and postmortem diagnoses. *JAMA*. 1987;258:339-344.
21. SAS software, Version 8.2. SAS Institute: Cary, NC.
22. Cameron HM, McGoogan E. A prospective study of 1152 hospital autopsies, II: analysis of inaccuracies in clinical diagnoses and their significance. *J Pathol*. 1981;133:285-300.
23. Cameron HM, McGoogan E, Watson H. Necropsy: a yardstick for clinical diagnoses. *BMJ*. 1980;281:985-988.
24. Thurlbeck WM. Accuracy of clinical diagnosis in a Canadian teaching hospital. *CMAJ*. 1981;125:443-447.
25. Pounder DJ, Horowitz M, Rowland R, Reid DP. The value of the autopsy in medical audit—a combined clinical and pathological assessment of 100 cases. *Aust N Z J Med*. 1983;13:478-482.
26. Kircher T, Nelson J, Burdo H. The autopsy as a measure of accuracy of the death certificate. *N Engl J Med*. 1985;313:1263-1269.
27. Stevanovic G, Tucakovic G, Dotlic R, Kanjuh V. Correlation of clinical diagnoses with autopsy findings: a retrospective study of 2,145 consecutive autopsies. *Hum Pathol*. 1986;17:1225-1230.
28. Schned AR, Mogielnicki RP, Stauffer ME. A comprehensive quality assessment program on the autopsy service. *Am J Clin Pathol*. 1986;86:133-138.
29. Sarode VR, Datta BN, Banerjee AK, et al. Autopsy findings and clinical diagnoses: a review of 1,000 cases. *Hum Pathol*. 1993;24:194-198.
30. Burke MC, Aghababian RV, Blackbourne B. Use of autopsy results in the emergency department quality assurance plan. *Ann Emerg Med*. 1990;19:363-366.
31. Middleton K, Clarke E, Homann S, et al. An autopsy-based study of diagnostic errors in geriatric and nongeriatric adult patients. *Arch Intern Med*. 1989;149:1809-1812.
32. McKelvie PA, Rode J. Autopsy rate and a clinicopathological audit in an Australian metropolitan hospital—cause for concern? *Med J Aust*. 1992;156:456-462.
33. Harris MD, Blundell JW. Audit of necropsies in a British district general hospital. *J Clin Pathol*. 1991;44:862-865.
34. Nichols L, Aronica P, Babe C. Are autopsies obsolete? *Am J Clin Pathol*. 1998;110:210-218.
35. Tse GM, Lee JC. A 12-month review of autopsies performed at a university-affiliated teaching hospital in Hong Kong. *Hong Kong Med J*. 2000;6:190-194.
36. Boers M, Nieuwenhuyzen Kruseman AC, Eulderink F, Hermans J, Thompson J. Value of autopsy in internal medicine: a 1-year prospective study of hospital deaths. *Eur J Clin Invest*. 1988;18:314-320.
37. Bernicker EH, Atmar RL, Schaffner DL, Greenberg SB. Unanticipated diagnoses found at autopsy in an urban public teaching hospital. *Am J Med Sci*. 1996;311:215-220.
38. Fernandez-Segoviano P, Lazaro A, Esteban A, Rubio JM, Iruretagoyena JR. Autopsy as quality assurance in the intensive care unit. *Crit Care Med*. 1988;16:683-685.
39. Papadakis MA, Mangione CM, Lee KK, Kristof M. Treatable abdominal pathologic conditions and unsuspected malignant neoplasms at autopsy in veterans who received mechanical ventilation. *JAMA*. 1991;265:885-887.
40. Mort TC, Yeston NS. The relationship of pre-mortem diagnoses and post-mortem findings in a surgical intensive care unit. *Crit Care Med*. 1999;27:299-303.
41. Blosser SA, Zimmerman HE, Stauffer JL. Do autopsies of critically ill patients reveal important findings that were clinically undetected? *Crit Care Med*. 1998;26:1332-1336.
42. Stothert JC Jr, Gbaanador G. Autopsy in general surgery practice. *Am J Surg*. 1991;162:585-588.
43. Zehr KJ, Liddicoat JR, Salazar JD, et al. The autopsy: still important in cardiac surgery. *Ann Thorac Surg*. 1997;64:380-383.
44. Shanks JH, Anderson NH, McCluggage G, Toner PG. Use of the autopsy in Northern Ireland and its value in perioperative deaths. *IARC Sci Publ*. 1991;112:115-124.
45. West JG, Trunkey DD, Lim RC. Systems of trauma care: a study of two counties. *Arch Surg*. 1979;114:455-460.
46. Stolman CJ, Castello F, Holland BK, et al. Lessons learned from autopsies at a pediatric teaching hospital. *N J Med*. 1995;92:794-797.
47. Kumar P, Taxy J, Angst DB, Mangurten HH. Autopsies in children: are they still useful? *Arch Pediatr Adolesc Med*. 1998;152:558-563.
48. Kumar P, Angst D, Taxy J, Mangurten H. Neonatal autopsies: a 10-year experience. *Arch Pediatr Adolesc Med*. 2000;154:38-42.
49. Whitehouse SR, Kisson N, Singh N, Warren D. The utility of autopsies in a pediatric emergency department. *Pediatr Emerg Care*. 1994;10:72-75.
50. Goldstein B, Metlay L, Cox C, Rubenstein JS. Association of pre-mortem diagnosis and autopsy findings in pediatric intensive care unit versus emergency department versus ward patients. *Crit Care Med*. 1996;24:683-686.
51. Stambouly JJ, Kahn E, Boxer RA. Correlation between clinical diagnoses and autopsy findings in critically ill children. *Pediatrics*. 1993;92:248-251.
52. Dhar V, Perlman M, Vilela MI, Haque KN, Kirpalani H, Cutz E. Autopsy in a neonatal intensive care unit: utilization patterns and associations of clinicopathologic discordances. *J Pediatr*. 1998;132:75-79.
53. Barr P, Hunt R. An evaluation of the autopsy following death in a level IV neonatal intensive care unit. *J Paediatr Child Health*. 1999;35:185-189.
54. Puxty JA, Horan MA, Fox RA. Necropsies in the elderly. *Lancet*. 1983;1:1262-1264.
55. Bedell SE, Fulton EJ. Unexpected findings and complications at autopsy after cardiopulmonary resuscitation (CPR). *Arch Intern Med*. 1986;146:1725-1728.
56. Baker PB, Zarbo RJ, Howanitz PJ. Quality assurance of autopsy face sheet reporting, final autopsy report turnaround time, and autopsy rates: a College of American Pathologists Q-Probes study of 10003 autopsies from 418 institutions. *Arch Pathol Lab Med*. 1996;120:1003-1008.
57. Anderson RE, Hill RB, Key CR. The sensitivity and specificity of clinical diagnostics during five decades: toward an understanding of necessary fallibility. *JAMA*. 1989;261:1610-1617.
58. Veress B, Alafuzoff I. Clinical diagnostic accuracy audited by autopsy in a university hospital in two eras. *Qual Assur Health Care*. 1993;5:281-286.
59. Agency for Healthcare Research and Quality. 1999 national statistics. Available at: <http://hcup.ahrq.gov/Hcupnet.asp>. Accessed April 23, 2002.
60. American Hospital Association. *Hospital Statistics*. Chicago, Ill: Health Forum; 2002.
61. Kohn L, Corrigan J, Donaldson M, eds. *To Err Is Human: Building a Safer Health System*. Washington, DC: National Academy Press; 2000.
62. Hayward RA, Hofer TP. Estimating hospital deaths due to medical errors: preventability is in the eye of the reviewer. *JAMA*. 2001;286:415-420.
63. McDonald CJ, Weiner M, Hui SL. Deaths due to medical errors are exaggerated in Institute of Medicine report. *JAMA*. 2000;284:93-95.
64. Sox HCJ, Woloshin S. How many deaths are due to medical error? getting the number right. *Eff Clin Pract*. 2000;3:277-283.
65. Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. *J Clin Epidemiol*. 2000;53:207-216.
66. Pham B, Platt R, McAuley L, Klassen TP, Moher D. Is there a "best" way to detect and minimize publication bias? an empirical evaluation. *Eval Health Prof*. 2001;24:109-125.
67. Meier PR, Manchester DK, Shikes RH, Clewell WH, Stewart M. Perinatal autopsy: its clinical value. *Obstet Gynecol*. 1986;67:349-351.
68. Hagerstrand I, Lundberg LM. The importance of post-mortem examinations of abortions and perinatal deaths. *Qual Assur Health Care*. 1993;5:295-297.
69. Saller DN Jr, Lesser KB, Harrel U, Rogers BB, Oyer CE. The clinical utility of the perinatal autopsy. *JAMA*. 1995;273:663-665.
70. Veress B, Gadaleanu V, Nennesmo I, Wikstrom BM. The reliability of autopsy diagnostics: inter-observer variation between pathologists, a preliminary report. *Qual Assur Health Care*. 1993;5:333-337.
71. Green J, Wintfeld N. How accurate are hospital discharge data for evaluating effectiveness of care? *Med Care*. 1993;31:719-731.
72. Malenka DJ, McLerran D, Roos N, Fisher ES, Wenberg JE. Using administrative data to describe casemix: a comparison with the medical record. *J Clin Epidemiol*. 1994;47:1027-1032.
73. Romano PS. Can administrative data be used to compare the quality of health care? *Med Care Rev*. 1993;50:451-477.
74. Gross CP, Anderson GF, Powe NR. The relation between funding by the National Institutes of Health and the burden of disease. *N Engl J Med*. 1999;340:1881-1887.
75. Varmus H. Evaluating the burden of disease and spending the research dollars of the National Institutes of Health. *N Engl J Med*. 1999;340:1914-1915.
76. Lamarre-Cliche M, Castilloux AM, LeLorier J. Association between the burden of disease and research funding by the Medical Research Council of Canada and the National Institutes of Health: a cross-sectional study. *Clin Invest Med*. 2001;24:83-89.
77. Davis CJ. Autopsy rates: the decline with age. *PAS Rep*. 1979;16:31-41.
78. Ahronheim JC, Bernholm AS, Clark WD. Age trends in autopsy rates: striking decline in late life. *JAMA*. 1983;250:1182-1186.
79. Sinarid JH. Factors affecting autopsy rates, autopsy request rates, and autopsy findings at a large academic medical center. *Exp Mol Pathol*. 2001;70:333-343.
80. Rosenbaum GE, Burns J, Johnson J, Mitchell C, Robinson M, Truog RD. Autopsy consent practice at US teaching hospitals: results of a national survey. *Arch Intern Med*. 2000;160:374-380.
81. Sanner MA. Medical students' attitudes toward autopsy: how does experience with autopsies influence opinion? *Arch Pathol Lab Med*. 1995;119:851-858.