

Reassessing the Methods of Medical Record Review Studies in Emergency Medicine Research

Andrew Worster, MD, MSc

R. Daniel Bledsoe, MD

Paul Cleve, MD

Christopher M. Fernandes, MD

Suneel Upadhye, MD

Kevin Eva, PhD

From the Division of Emergency Medicine (Worster, Cleve, Fernandes, Upadhye) and Department of Clinical Epidemiology and Biostatistics (Worster, Eva), McMaster University, Hamilton, Ontario, Canada; and the Emergency Department, York Hospital, York, PA (Bledsoe).

Study objective: An assessment of the methods of medical record review studies published in emergency medicine journals during a 5-year period ending in 1993 provided strategies for improvements. We assess and quantify the current methodologic quality of medical record review studies in emergency medicine journals using published guidelines and compare these results among journals and with those of 10 years previously.

Methods: Independent, systematic searches of emergency medicine journals identified all medical record review studies published in 2003. Methodology assessments of each selected study were conducted independently by 2 other researchers, and disagreements were resolved by arbitration.

Results: We identified 79 (14%) medical record review studies in 563 original research articles in 6 emergency medicine journals. The highest adherence to methodologic standards was found for sampling method (99%; 95% confidence interval [CI] 93% to 100%), and the lowest was for abstractor blinding to hypothesis (4%; 95% CI 1% to 11%). Interobserver agreement for the 12 criteria ranged from 57% to 95%. A comparison of these results with those of 10 years ago revealed significant improvements in 3 of the 8 original criteria assessed: data abstraction forms, mentioning interobserver performance, and testing interobserver performance.

Conclusion: Medical record review studies continue to comprise a substantial proportion of original research in the emergency medicine literature. Important improvements are noted in some criteria, but adherence remains below 50% for 7 of the 12 criteria assessed. [Ann Emerg Med. 2005;45:448-451.]

0196-0644/\$-see front matter

Copyright © 2005 by the American College of Emergency Physicians.

doi:10.1016/j.annemergmed.2004.11.021

SEE RELATED ARTICLE, P. 444, AND EDITORIAL, P. 452.

INTRODUCTION

In medical record review studies, researchers extract data from documents that were not created for the purpose of the study.^{1,2} Although medical record review studies make up 25% of all scientific studies published in peer-reviewed, emergency medicine journals and 53% of emergency medical services studies, there are no universally accepted standards for conducting or reporting them.^{1,3-5}

In this article, we use a definition of medical record as a document containing patient-focused medical information. We examine to what extent studies published in the 6 major, MEDLINE-indexed, peer-reviewed, emergency medicine journals using medical records as the primary data source are

compliant with published methodologic criteria. Eight of these criteria were previously assessed in a review of the quality of medical record reviews 10 years ago, and we compare those results with an assessment of current medical record reviews to determine whether quality has improved in this period.³

MATERIALS AND METHODS

Objectives

The objectives were (1) to assess and quantify the current method quality of medical record review studies published in 6 MEDLINE-indexed, peer-reviewed, emergency medicine journals, and (2) to compare these results among journals and with those of medical record review studies published 10 years previously.

Editor's Capsule Summary*What is already known on this topic*

When 3 emergency medicine journals were studied 10 years ago, articles that used medical record review had poor compliance with 8 proposed quality indicators for such methods.

What question this study addressed

Has the adherence to quality indicators increased in the past 10 years? Is there heterogeneity of adherence among emergency medicine journals?

What this study adds to our knowledge

Despite modest improvements in 7 of the 8 original criteria, adherence remains under 50% for most. There is heterogeneity of adherence among criteria and among journals.

How this might change clinical practice

This will not change clinical practice. Authors and journal editors may wish to examine the proposed criteria to determine whether their research and research publications should be structured to adhere to them.

Ethics

This library-based study was found to be exempt from local, formal ethics review and approval at the Canadian site and exempt from both the Health Insurance Portability and Accountability Act regulations and institutional review board approval at the US site.

Journal Selection

We selected the 6 MEDLINE-listed, peer-reviewed, emergency medicine journals with the highest impact factor published during 2003 except 2: *Journal of Burn Care and Rehabilitation* (content too specific) and *Emergency Medicine Clinics of North America* (no original research articles).⁶ These were replaced with the *Journal of Emergency Medicine* and the *Emergency Medicine Journal*. This list of 6 includes *Resuscitation* and the 3 emergency medicine journals on which the original medical record review methods assessment study by Gilbert et al³ was conducted: *American Journal of Emergency Medicine*, *Annals of Emergency Medicine*, and *Journal of Emergency Medicine*.

Data Selection and Abstraction

Two authors (AW, CMF) independently performed computerized and hand searches of all articles in the 2003 issues of these 6 journals (Table 1).

We selected only those studies that met our definition of medical record and, in keeping with the methods described by Gilbert et al,³ excluded all retrospective studies based on aggregate patient data and computerized databases, case reports

Table 1. Emergency medicine journal impact factor and proportion of medical record review studies.

Emergency Medicine Journal 2003	Impact Factor*	Research Articles, No.	Medical Record Review, %
<i>Annals of Emergency Medicine</i> [†]	2.640	103	11.7
<i>Academic Emergency Medicine</i>	1.844	144	7.6
<i>American Journal of Emergency Medicine</i> [†]	1.489	74	17.6
<i>Resuscitation</i>	1.375	127	15.0
<i>Journal of Emergency Medicine</i> [†]	0.652	42	28.6
<i>Emergency Medicine Journal</i>	0.633	73	16.4

*Impact factors from Journal Citation Reports 2003, Institute for Scientific Information.⁶
[†]Journals assessed by Gilbert et al.³

and case series, letters, editorials, subject reviews, systematic reviews, prospective clinical trials, and studies published in abstract form only. Disagreements about selection were resolved by consensus.

Article Assessment

Abstractors were trained using a list of methodologic evaluation criteria created from the 8 previously published criteria (1 to 8) and 4 additional criteria (9 to 12)^{1,3,7} (Table 2). Two authors, each using a computerized data abstraction form, independently evaluated each article. Discussions were held to clarify issues as they arose. To keep the abstractors blinded to each other's findings, disagreements were arbitrated by a third researcher.

Analysis

We calculated the percentage of research articles in each journal that used medical record review. For each of the 12 criteria, we calculated interobserver reliability (as agreement percentage) and the percentage of articles that were in compliance. We calculated the difference in compliance with 8 original criteria by Gilbert et al³ for the 1989 to 1993 and 2003 medical record review studies in the 3 journals used in the study by Gilbert et al.³

RESULTS

We identified 79 (14%) medical record review studies in 563 original research articles in 6 emergency medicine journals (see Appendix E1, available at <http://www.mosby.com/AnnEmergMed>). The proportion of medical record review studies in 2003 in the original 3 journals reviewed by Gilbert et al³ was 27% versus 25% from 1989 to 1993. The proportion of medical record review studies found in each of the 6 reported emergency medicine journals ranged from 7.6% to 28.6% (Table 1).

Interobserver agreement for the 12 criteria ranged from 57% to 95%. The highest adherence to methodologic standards was

Table 2. Description of methods criteria.*

Method Criterion	Method Criterion Description
1. Abstractors training	Were the abstractors trained before the data collection?
2. Case selection criteria	Were the inclusion and exclusion criteria for case selection defined?
3. Variable definition	Were the variables defined?
4. Abstraction forms	Did the abstractors use data abstraction forms?
5. Performance monitored	Was the abstractors' performance monitored?
6. Blind to hypothesis	Were the abstractors aware of the hypothesis/study objectives?
7. IRR mentioned	Was the interobserver reliability discussed?
8. IRR tested	Was the interobserver reliability tested or measured?
9. Medical record identified	Was the medical record database identified or described?
10. Sampling method	Was the method of sampling described?
11. Missing-data management plan	Was the statistical management of missing data described?
12. Institutional review board approval	Was the study approved by the institutional or ethics review board?

IRR, Interrater reliability.
*Method criteria 1 to 8 from Gilbert et al.³

found for sampling method (99%; 95% confidence interval [CI] 93% to 100%) and case selection criteria (96%; 95% CI 89% to 99%) (Table 3). The lowest adherence to methodologic standards was found for abstractor blinding to hypothesis (4%; 95% CI 1% to 11%) and abstractor performance monitoring (9%; 95% CI 4% to 17%). There was heterogeneity of adherence to the 12 criteria among journals (Table E1, available online at <http://www.mosby.com/AnnEmergMed>).

A comparison of the adherence to methodologic standards in the 3 journals in the original study revealed improvements in 3 of the 8 criteria assessed: data abstraction forms, mentioning interobserver performance, and testing interobserver performance (Table 4).³

Table 4. Adherence to methodologic standards in medical record review studies from 3 emergency medicine journals published in 1989 to 1993 and 2003.*

Method Criterion for Assessment	Adherence %, 1989–1993 [†]	Adherence %, 2003	Difference % (95% CI)
1. Abstractors trained	17.6 (43/244)	21.6 (8/37)	4.0 (–10.1 to 18.1)
2. Case selection criteria	98.4 (240/244)	97.3 (36/37)	–1.1 (–6.5 to 4.4)
3. Variable definition	73.4 (179/244)	86.5 (32/37)	13.1 (0.8–25.5)
4. Abstraction forms	10.7 (26/244)	37.8 (14/37)	27.1 (11.0–43.2)
5. Performance monitored	4.1 (10/244)	8.1 (3/37)	4.0 (–5.1 to 13.1)
6. Blind to hypothesis	3.3 (8/244)	5.4 (2/37)	2.1 (–5.5 to 9.7)
7. IRR mentioned	5.0 (12/244)	29.3 (11/37)	24.8 (9.8–39.8)
8. IRR tested	0.4 (1/244)	18.9 (7/37)	18.5 (5.9–31.2)

Difference %, Percentage adherence in studies published in 2003 less the percentage adherence in studies published from 1989 to 1993.
*Parenthetical values represent the proportion of studies.
[†]Results from Gilbert et al.³

Table 3. Adherence to methodologic standards in medical record review studies from 6 major emergency medicine journals in 2003.*

Method Criterion for Assessment	Agreement % (95% CI)	Adherence % (95% CI)
1. Abstractors training	95 (88–98)	18 (11–28)
2. Case selection criteria	94 (86–97)	96 (89–99)
3. Variable definition	72 (61–81)	77 (67–85)
4. Abstraction forms	81 (71–88)	27 (18–37)
5. Performance monitored	86 (68–79)	9 (4–17)
6. Blind to hypothesis	95 (88–98)	4 (1–11)
7. IRR mentioned	95 (88–98)	22 (14–32)
8. IRR tested	92 (84–96)	13 (7–22)
9. Medical record identified	86 (68–79)	91 (83–96)
10. Sampling method	87 (78–93)	99 (93–100)
11. Missing-data management	57 (46–67)	32 (22–43)
12. Institutional review board approval	95 (88–98)	58 (47–68)

Agreement %, Percentage of all reviewed studies in which 2 abstractors agreed on the criterion's presence or absence; adherence %, percentage of all reviewed studies in which the criterion was deemed present.
*Adherence data for all criteria for each journal available in Table E1 at <http://www.mosby.com/AnnEmergMed>.

LIMITATIONS

The primary limitation to this study is that the 8 methodologic criteria identified by Gilbert et al³ and the 4 identified here have not been validated as indicators of medical record review study quality. Another limitation is the abstractors' awareness of the objectives of the study. However, given the assumption of the authors that improvements would be found, bias in the abstraction process would likely yield higher adherence rates.

DISCUSSION

Gilbert et al³ proposed 8 strategies for improving the quality of medical record review studies, which have since been adopted by *Annals of Emergency Medicine* as guidelines for medical record review reporting. We have added 4 criteria to the original 8, of which all but 1 (institutional review board approval)

potentially affect the reliability or validity of the results of the study. Institutional review board approval (#12), although it has no impact on results, is a requirement of all studies and has great importance in light of recent privacy protection laws. The identification or description of the medical record (#9) is necessary for reproducibility, specifically, to allow others to repeat the study in other settings to determine the validity of the results.⁸ A description of the sampling method (#10) is important because some sampling methods are less subject to bias than others.^{1,2} Missing and conflicting data (#11) is considered one of the greatest weaknesses of this design type, and so appropriate management is crucial.^{1,3,7,9} It is this criterion with which the abstractors had the lowest level of agreement (57%), which we subsequently attributed to the poor quality of reporting in the Methods section of many of the reviewed articles.

Adherence to methodologic standards in medical record review studies is greatest for sampling method and lowest for abstractor blinding to hypothesis. We believe that the latter is due to the study investigators often being the data abstractors, as they were in this study.

In conclusion, medical record review studies continue to comprise a substantial proportion of original research in the emergency medicine literature. Although improvements in adherence to methodologic standards in medical record review studies are noted in 3 of the 8 original criteria assessed—data abstraction forms, mentioning interobserver performance, and testing interobserver performance—adherence remains below 50% for 7 of the 12 criteria assessed. Assuming that the 12 criteria are valid indicators of study quality, our findings suggest that greater effort should be made to improve the quality and reporting of medical record review studies.

Author contributions: AW conceived, designed, and supervised the study. RDB, CMF, and KE contributed to study design; PC contributed to database design. AW and CMF conducted the article searches. RDB, PC, and SU performed data collection, and AW and KE performed data analysis. AW, RDB, CMF, and KE drafted the manuscript, and AW, RDB,

CMF, KE, and SU contributed substantially to its revision. AW takes responsibility for the paper as a whole.

Funding and support: The authors report this study did not receive any outside funding or support.

Publication dates: Received for publication January 29, 2004. Revisions received August 31, 2004; November 1, 2004; and November 15, 2004. Accepted for publication November 16, 2004. Available online February 12, 2005.

Presented as an abstract at the Society for Academic Emergency Medicine annual meeting, Orlando, FL, May 2004.

Reprints not available from the authors.

Address for correspondence: Andrew Worster, MD, MSc, Hamilton Health Sciences, 237 Barton Street East, Hamilton, Ontario L8L 2X2, Canada; 905-527-4322 ext. 75848, fax 905-527-7051; E-mail worster@hhsc.ca.

REFERENCES

1. Worster A, Haines T. Advanced statistics: understanding medical record review studies. *Acad Emerg Med.* 2004;11:187-192.
2. Hulley SB, Cummings SR, Browner WS, et al. *Designing Clinical Research.* 2nd ed. Philadelphia, PA: Lippincott, Williams & Wilkins; 2001:336.
3. Gilbert EH, Lowenstein SR, Kozoi-McLain J, et al. Chart reviews in emergency medicine research: where are the methods? *Ann Emerg Med.* 1996;27:305-308.
4. Lerner EB, Zachariah BS, White LJ. Conducting retrospective emergency medical services research. *Prehosp Emerg Care.* 2002;6(2 Suppl):S48-51.
5. Rangel SJ, Kelsey J, Colby CE, et al. Development of a quality assessment scale for retrospective clinical studies in pediatric surgery. *J Pediatr Surg.* 2003;38:390-396.
6. Journal Citation Reports: JCR 2003 [CD-ROM]. Philadelphia, PA: Institute for Scientific Information; 2003.
7. Worster A, Haines T. Medical record review studies: an overview. *Israeli J Trauma Intensive Care Emerg Med.* 2002;2:21-26.
8. Polgar S, Thomas SA. *Introduction into Research in the Health Sciences.* 3rd ed. New York, NY: Churchill Livingstone; 1995:393.
9. Burnum JF. The misinformation era: the fall of the medical record. *Ann Intern Med.* 1989;110:482-484.