Published in final edited form as:

J Emerg Med. 2012 December ; 43(6): 1160–1166. doi:10.1016/j.jemermed.2011.11.002.

ER vs. ED: A Comparison of Televised and Real-Life Emergency Medicine

Brian A. Primack, MD, EdM, MS*,

Departments of Medicine and Pediatrics, University of Pittsburgh School of Medicine, Pittsburgh, PA

Traci Roberts, BS,

University of Pittsburgh School of Medicine, Pittsburgh, PA

Michael J. Fine, MD, MSc,

Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA

Francesca R. Dillman Carpentier, PhD,

School of Journalism and Mass Communication, University of North Carolina, Chapel Hill, NC

Kristen R. Rice, MPH, and

Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA

Amber E. Barnato, MD, MPH

Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA

Abstract

Background—Although accurate health-related representations of medical situations on television can be valuable, inaccurate portrayals can engender misinformation.

Objective—The purpose of this study was to compare socio-demographic and medical characteristics of patients depicted on television vs. actual U.S. Emergency Department (ED) patients.

Methods—Two independently working coders analyzed all 22 programs in one complete year of the popular emergency room drama ER. Inter-rater reliability was excellent, and all initial coding differences were easily adjudicated. Actual health data were obtained from the National Health and Ambulatory Medical Care Survey data from the same year. We used Pearson's chi-squared test to compare televised vs. real distribution across key socio-demographic and medical variables.

Results—Ages at the extremes (e.g., 4 and 45) were less commonly represented on television compared with reality. Compared with reality, characters on television were less commonly women (31.2% vs. 52.9%), African American (12.7% vs. 20.3%) or Hispanic (7.1% vs. 12.5%). The two most common acuity categories for television were the extreme categories “non-urgent” and “emergent,” whereas the two most common categories for reality were the middle categories “semi-urgent” and “urgent.” Compared with reality, televised visits were most commonly due to injury (63.5% vs. 37.0%), and televised injuries were less commonly work-related (4.2% vs. 14.8%).

Conclusions—Comparison of represented and actual characteristics of ED patients may be valuable in helping us determine what types of patient misperceptions may exist as well as what types of interventions may be beneficial in correcting that potential misinformation.

* Address for Reprints: Brian A. Primack, MD, EdM, MS, University of Pittsburgh School of Medicine, 230 McKee Place Suite 600, Pittsburgh, PA 15213, bprimack@pitt.edu, 412-586-9789 (phone); 412-692-4838 (fax).

Keywords

Television; emergency department; content analysis; NHAMCS; health disparities

INTRODUCTION

Americans obtain health-related information from mass media more than any source other than health care professionals [1]. Cultivation theory, widely accepted by scholars of communication, suggests that exposure to television in particular over time substantially “cultivates” viewers’ perceptions of reality [2]. Empiric data support the application of this theory to health care. For example, of those who watch the popular medical drama *ER*, 1 in 3 report that they utilize information from the program to make their own health care choices [3].

Accurate information presented in narrative mass media can be highly valuable and educational. After viewing a particular episode of *ER* describing emergency contraceptives, the proportion of viewers who knew that birth control pills could be used as emergency contraceptives increased from 10% to 33% [3]. Televised representations of medical situations, however, can also be highly inaccurate. Roughly 75% of televised cardiac arrest victims on television survive, much more than the 5% who survive in reality [4]. Similarly, traumatic coma victims on television are far more likely to return to normal function than real traumatic coma victims (89% vs. 7%) [5]. Although television is not intended to mirror reality with precise accuracy—and indeed few viewers would enjoy watching a program that plods along with slavish devotion to reality [6]—exposure to such representations may nevertheless engender misinformation. For example, it has been estimated that 96% of Americans exaggerate the potential benefit of cardiopulmonary resuscitation, and such misinformation can lead to unnecessary litigation, false expectations of medical treatments, and other communication difficulties between patients and providers [4, 7].

In addition to shaping expectations and behaviors among patients, televised medical dramas can influence the expectations and behaviors of future and current health professionals [6]. Medical students who watched *ER* incorporated attitudes and beliefs from the show into their professional careers and strongly influenced students’ career choices [8]. After the first season of *ER*, students pursuing the specialty doubled, and this was followed by a national increase over 25% during initial years of the program [8].

Because television increasingly provides a considerable portion of the public’s medical knowledge, it is valuable to study the differences between televised health and reality. The purpose of the current study was to move beyond the comparison of rare medical situations such as CPR or coma to compare the full spectrum of hospital-based emergency medicine. Thus, we performed a comprehensive content analysis of the socio-demographic and medical characteristics of patients depicted on a complete season of a popular medical television drama (*ER*) to compare these data with those of actual U.S. emergency department (ED) patients.

MATERIALS AND METHODS

Design

We conducted a descriptive comparison between a qualitatively coded content analysis from one complete year of a popular emergency room drama and actual National Health and Ambulatory Medical Care Survey (NHAMCS) data from the same year.

Procedures

Approvals—The University of Pittsburgh Institutional Review Board approved the protocol, which met the criteria for exemption from written informed consent.

Selection of Text to Code—We purchased the most recently available DVD for a complete season of the television program *ER*. This program was selected for this analysis for several reasons. First, it focuses on the ED—our setting of interest—whereas other popular programs such as *Grey's Anatomy* involve multiple parts of the hospital. Second, *ER* is highly popular: it was the longest-running television program of this type and it ranked top 5 for its first nine seasons. Finally, the program *ER* aims to achieve a relatively high measure of “reality,”[6] whereas other popular programs such as *Scrubs* and *House* specifically distance themselves from reality through highly bizarre situations and quirky production techniques. Because cultivation theory suggests that viewers of these other programs would be less likely to integrate those televised situations into their vision of real life health and medicine, we were less interested in assessing their content. Although we considered assessing other programs in addition to *ER*, no other programs satisfied our criteria of being set primarily in the ED, highly popular, and aiming to accurately mirror reality.

Coding procedures—All 22 programs contained in the season were analyzed in their entirety by two coders working independently from one another. Coders were advanced medical students trained by a physician in utilizing published algorithms to code important aspects of emergency room visits assessed by the NHAMCS, such as visit diagnosis and visit acuity. Coders were of opposite gender and different racial backgrounds. Coders used structured forms to organize all information. Coders attempted to assign codes to every patient based on whatever information was available. For example, pain might be assessed based on facial expression or dialogue, triage level might be influenced by observed vital signs or number of tests ordered, and final diagnostic category might be determined based on the most commonly stated medical problem for that patient. After both coders analyzed the complete season, we computed Cohen's κ to determine inter-rater reliability for each coded variable. We defined *a priori* sufficient inter-rater reliability to be at least “moderate” according to the Landis-Koch framework (i.e., $\kappa > 0.4$). After achieving this sufficient level of agreement after their first attempt, the coders adjudicated all initial differences. The rare differences of opinion (i.e., <1%) at this point were adjudicated with the assistance of the Principal Investigator and repeat viewing when necessary.

National comparison data—The National Hospital Ambulatory Medical Care Survey (NHAMCS) collects data on the utilization and provision of ambulatory care services in hospital emergency departments. Findings are based on a national sample of visits to the emergency departments of noninstitutional general and short-stay hospitals—exclusive of Federal, military, and Veterans Administration hospitals—located in the 50 States and the District of Columbia. The survey uses a four-stage probability design with samples of geographically defined areas, hospitals within these areas, clinics within the outpatient departments and emergency service areas within the emergency departments of these hospitals, and patient visits to these clinics and emergency services areas. We accessed NHAMCS data for the same year as the televised program (2000) via the Web link provided by the Centers for Disease Control and Prevention [9].

Measures

Definition of a patient—Because NHAMCS data are based on the individual patient encounter, our unit of analysis for coding was the patient encounter. We defined *a priori* a patient in a television episode as any individual who either (1) spoke at least two lines; or (2)

had at least two lines spoken about him/her. We did this because we wanted to avoid giving extras the same weight as characters integral to the story lines. We included the provision for characters who had lines spoken about them because occasionally integral characters do not speak, such as those who are unconscious. Specific patient data coded included their sociodemographic and medical characteristics.

Sociodemographic characteristics—Sociodemographic information coded included the primary sociodemographic information recoded by NHAMCS, including age, gender, race, and ethnicity. We used age, race, and ethnicity categorizations as specified by NHAMCS.

Medical characteristics—Medical information coded included urgency, level of consciousness, level of pain, if the visit was work-related, and major disease category by ICD-9 code. Urgency was defined as emergent, urgent, semi-urgent, or non-urgent based on established criteria of the Emergency Severity Index (ESI) [10]; each coder was provided with a copy of the ESI algorithm and trained in its application. Level of pain was categorized as none, mild, moderate, or severe using definitions provided by NHAMCS [9]. We also coded whether an injury-related visit was work-related; if there was no mention of work-relatedness we assumed that the visit was not work related. Finally, we used the current ICD-9 coding guidelines to assign a major disease category to each patient encounter (e.g., infectious, neoplastic, endocrine). For the purposes of this analysis, we determined *a priori* not to utilize supplemental or administrative “V” codes as the primary major disease category. Therefore, in order to be consistent, we did not include NHAMCS visits utilizing a “V” code as the major disease category in our analyses. Our systematic protocol involved assigning each NHAMCS patient a single category code (e.g., cardiac, respiratory) associated with the first ICD-9 code listed.

Analysis

We used κ statistics to compare coding of televised representations between the two independent coders. To calculate population-based ED estimates, we used the sampling weights provided with the NHAMCS data [9]. We used Pearson’s chi-squared test to compare televised vs. real distribution across each key sociodemographic and medical characteristic of interest.

RESULTS

Sample and Missing Data

The televised *ER* sample included 192 patients depicted during the 2000 season. For rare observations, coders agreed that specified variables were indeterminate. For example, three characters were not physically shown on screen and race and ethnicity could not be determined from the script. Similarly, pain could not be assessed for two characters. In these cases, these variables were coded as missing. However, data were missing in this way <1% of the time and thus did not affect analyses. The real ED sample included 25,622 patients, representing 108,016,777 ED visits in 2000.

Content Analysis Agreement

According to the Landis-Koch framework [11], initial inter-rater agreement was “substantial” ($\kappa = 0.61$ – 0.80) or “almost perfect” ($\kappa = 0.81$ – 1.00) for nearly all variables (Table 1). Although initial agreement was somewhat lower for more subjective codes with four or more categories (e.g., acuity and level of pain), all levels of agreement were at least “moderate” ($\kappa = 0.41$ – 0.60) and 100% of segments were coded with agreement after adjudication.

Sociodemographic characteristics

Patient age differed between television and reality ($P<.001$); in particular, ages at the extremes (e.g., 4 and 45) were less commonly represented on television compared with reality (Table 1). Men were represented more frequently on television compared with reality (68.8% vs. 47.1%, $P<.001$). Compared with reality, individuals on television were less commonly African American (12.7% vs. 20.3%, $P=.043$) or Hispanic (7.1% vs. 12.5%, $P=.021$).

Medical Characteristics

Visit urgency differed between television and reality ($P<.001$). In particular, the two most common categories in reality were the middle categories “semi-urgent” (41.9%) and “urgent” (22.7%), whereas the two most common categories for television were the extreme categories “non-urgent” (29.7%) and “emergent” (28.1%). Pain was most commonly absent on television (42.1%) but most commonly “mild” or “moderate” in real life (35.4% and 29.2%, $P<.001$). Injuries were far less commonly work-related on television (4.2% vs. 14.8%, $P<.001$). Injury was overrepresented among televised patients (63.5% vs. 37.0%), and respiratory conditions were underrepresented (3.2% vs. 15.1%).

DISCUSSION

This manuscript presents a comprehensive comparison of the case mix of televised and real emergency medicine. Prior work has more commonly focused on rare, dramatic events, such as CPR and coma [4, 5]). Not surprisingly, we found that the demographics and medical characteristics of patients in the television program *ER* did not mirror the reality of the American ED. In particular, television patients were disproportionately young and middle-aged white adults with apparently painless non-work-related injuries. Cultivation theory suggests that the televised representations we found may influence perceptions of both patients and providers: many medical students may have been surprised to find sick elders, babies, chronic illness, and pain when they experienced their first shift of real emergency medicine.

Our findings with regard to the underrepresentation of very young, very old, female, black, and Hispanic patients corroborate previous research that examines both hospital dramas and prime time television as a whole. Some have suggested that these disparities occur because there is a small percentage of women, minorities, or old/young writers in Hollywood. Regardless of why this occurs, however, the trend is problematic because viewers may underestimate the ED-related needs of infants, young children, and the elderly. Indeed, it is ironic that the particular groups of individuals who were underrepresented on the television program are more likely in real life to experience health disparities: the elderly, females, blacks, and Hispanics.

Whereas televised visits were usually due to injury (63.5%), in reality only 37% of visits were for injuries, with other disease categories (e.g., respiratory, gastrointestinal, and dermatologic) more heavily represented. These differences between televised and actual utilization of the emergency department may influence the types of conditions patients feel are appropriate for emergency department care. For example, because of the underrepresentation of respiratory diagnoses on television, patients with exacerbations of chronic respiratory conditions may be less likely to seek care even if it is warranted. Conversely, viewers might learn from television programming that the ED is a suitable treatment facility for certain conditions which are actually more appropriate for primary care (such as minimally sprained ankles). This is particularly noteworthy because of our finding that the most level of urgency in our televised sample was “non-urgent,” representing nearly

one-third of visits. It may be interesting for future investigations to assess whether these educational gaps are present, and whether they may be associated with these differences in representation as would be predicted by cultivation theory.

Our analysis was limited in that it focused on only one year of one particular medical program. Although we attempted to select a program that would be most relevant to influencing perceptions—because of its popularity and its purported attempt to mirror reality—it is possible that analysis of other seasons or programs may have produced different results. It should also be noted that there is inherent subjectivity in assessing factors such as “pain,” “triage level,” and “final discharge diagnosis” based only on dramatic information. However, this is why we chose the rigorous protocol of double-coding every patient encounter rather than the more customary procedure of merely verifying adequate inter-rater reliability among a subsample.

CONCLUSION

In conclusion, this study illustrates ways in which televised emergency medicine differs from real emergency medicine. Our findings indicate that the sociodemographic and medical characteristics of patients portrayed in the television program *ER* did not mirror those of patients accessing emergency department services in reality. These differences in representation may affect audiences’ health beliefs and behaviors, healthcare choices and expectations of medical treatment outcomes in the ED.

Acknowledgments

Dr. Primack is supported in part by a Physician Faculty Scholar Award from the Robert Wood Johnson Foundation and a career development award from the National Cancer Institute (K07-CA114315). Dr. Primack had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Hesse BW, et al. Trust and sources of health information: the impact of the Internet and its implications for health care providers: findings from the first Health Information National Trends Survey. *Arch Intern Med.* 2005; 165(22):2618–24. [PubMed: 16344419]
2. Gerbner, G., et al. Growing up with television: The cultivation perspective. In: Bryant, J.; Zillman, D., editors. *Media effects: Advances in theory and research.* Lawrence Erlbaum Associates, Inc; Hillsdale, NJ: 1994. p. 61-90.
3. Brodie M, et al. Communicating health information through the entertainment media. *Health Aff (Millwood).* 2001; 20(1):192–9. [PubMed: 11194841]
4. Diem SJ, Lantos JD, Tulsy JA. Cardiopulmonary resuscitation on television. Miracles and misinformation. *N Engl J Med.* 1996; 334(24):1578–82. [PubMed: 8628340]
5. Casarett D, et al. Epidemiology and prognosis of coma in daytime television dramas. *BMJ.* 2005; 331(7531):1537–9. [PubMed: 16373744]
6. Baer NA. Cardiopulmonary resuscitation on television. Exaggerations and accusations. *N Engl J Med.* 1996; 334(24):1604–5. [PubMed: 8628346]
7. Jones GK, Brewer KL, Garrison HG. Public expectations of survival following cardiopulmonary resuscitation. *Acad Emerg Med.* 2000; 7(1):48–53. [PubMed: 10894242]
8. O'Connor MM. The role of the television drama *ER* in medical student life: entertainment or socialization? *JAMA.* 1998; 280(9):854–5. [PubMed: 9730001]
9. Centers for Disease Control and Prevention. *National Ambulatory Health Medical Care Survey.* Centers for Disease Control and Prevention; Atlanta, GA: 2007.
10. Gilboy, N., et al. *Emergency Severity Index, Version 4: Implementation Handbook.* Agency for Healthcare Research and Quality; Rockville, MD: 2005.

11. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33(1):159–174. [PubMed: 843571]

Table 1

Comparison of Demographics and Diagnoses

Characteristic	Televised Data		National Data (%)	P value [†]
	N = 192 (%)	κ^*	N = 108,016,777 Weighted	
Age	N = 189	0.74	N = 108,016,777	< .001
<1	1.6		3.8	
1–4	1.6		7.8	
5–14	14.8		10.1	
15–24	20.1		16.4	
25–44	37		30.0	
45–64	13.2		17.0	
65–74	9		6.1	
75+	2.7		9.0	
Gender	N = 192	0.93	N = 108,016,777	< .001
Male	68.8		47.1	
Female	31.3		52.9	
Race	N = 189	0.90	N = 107,934,141	.043
White	85.2		77.0	
Black	12.7		20.3	
Asian	2.1		2.1	
American Indian	0.0		0.6	
Hispanic	N = 189	0.72	N = 88,862,103	.021
Yes	7.1		12.5	
No	92.7		87.5	
Urgency	N = 192	0.46	N = 80,405,526	< .001
Non-Urgent	29.7		21.1	
Semi-Urgent	24.5		41.9	
Urgent	17.7		22.7	
Emergent	28.1		14.3	
Level of Pain	N = 190	0.48	N = 53,076,746	< .001
None	42.1		25.2	
Mild	23.7		35.4	
Moderate	22.6		29.2	
Severe	11.6		10.3	
Work-Related	N = 120	0.56	N = 28,363,560	.001
Yes	4.2		14.8	
No	95.8		85.2	
Diagnostic Group	N = 189	0.65	N = 84,058,196	< .001
Infectious	5.8		3.8	
neoplasm	2.1		0.3	
endocrine	1.6		1.9	
blood	1.1		0.7	

Characteristic	Televised Data		National Data (%)	P value [†]
	N = 192 (%)	κ [*]	N = 108,016,777 Weighted	
mental	5.8		3.6	
nervous	1.6		8.9	
circulatory	5.8		5.2	
respiratory	3.2		15.1	
digestive	1.6		7.5	
genitourinary	1.1		5.9	
pregnancy	2.7		1.4	
skin	0.5		3.6	
musculoskeletal	3.2		6.9	
congenital	0.5		0.03	
perinatal	0.0		0.10	
injury	63.5		37.0	

^{*}Decimals in this column refer to Cohen's κ , a measure of inter-rater reliability between the two coders, after initial coding. According to an established framework [11], inter-rater agreement is "almost perfect" for $\kappa = 0.81-1.00$, "substantial" for $\kappa = 0.61-0.80$ and "moderate" for $\kappa = 0.41-0.60$. After adjudication, all codes were in agreement.

[†]Pearson's chi-squared to test the difference between televised and weighted national data.