

Assessing appropriateness of pediatric emergency department visits: is it even possible?

Jessica E. Paul, MD*[‡]; Katie Y. Zhu, MD[†]; Garth D. Meckler, MD, MSHS*[‡]; David K. Park , MN[‡]; Quynh Doan , MDCM, PhD*

CLINICIAN'S CAPSULE

What is known about the topic?

Emergency department (ED) visits deemed inappropriate are often targeted to reduce ED use, but most evaluations use clinician judgment to establish appropriateness.

What did this study ask?

How reliable are clinicians in characterizing appropriateness of an ED return visit, if guided by standardized criteria to reflect on?

What did this study find?

Despite guidance, clinicians agree poorly on appropriateness of a scheduled return pediatric ED visit.

Why does this study matter to clinicians?

Future efforts targeting “inappropriate” ED visits should not rely on clinical judgment to determine visit appropriateness.

random subset of 90 return visits reviewed by all three clinicians, agreement was 67% with a Fleiss' Kappa of 0.30 (0.17–0.44). Using a three-question algorithm based on objective criteria, agreement with the index treating provider increased to 115/207 (55.6%).

Conclusions: Although an important contributor to pediatric ED overcrowding, unnecessary or inappropriate visits are difficult to identify. We demonstrated poor reliability of clinician judgment to determine appropriateness of ED return visits, likely due to variability in clinical decision-making and risk-tolerance, social and systems factors impacting access and use of health care. We recommend that future studies evaluating the appropriateness of ED use standardized, objective criteria rather than clinician judgment alone.

RÉSUMÉ

Contexte: Bon nombre d'études font état de la fréquence des consultations inutiles ou non pertinentes au service des urgences (SU), ou encore de facteurs qui y ont associés, et les résultats reposent sur le jugement des cliniciens, mesure de référence de la pertinence. Aussi l'étude visait-elle à évaluer la constance du jugement des cliniciens dans l'appréciation de la pertinence des consultations au SU pédiatriques.

Méthode: Il s'agit d'une étude de cohortes rétrospective, dans laquelle a été comparée la pertinence des consultations au SU, déterminée par 3 cliniciens, à l'aide ou non d'un algorithme structuré, composé de 3 questions. L'équipe a constitué une cohorte de patients chez qui une visite de suivi prévue au SU avait été jugée pertinente par le clinicien traitant au moment de la consultation de référence donnée entre le 1^{er} mai 2012 et le 30 avril 2013. A suivi une mesure du degré de concordance des 3 cliniciens-chercheurs quant à la pertinence des consultations, déterminée à l'aide ou non de l'algorithme structuré.

Résultats: Sur un total de 207 visites de suivi prévues au SU et soumises au jugement du clinicien-examineur principal, 79 (38,2%) ont fait l'objet de concordance avec le clinicien

ABSTRACT

Objectives: Numerous studies reported on the frequency of, and factors associated with inappropriate or unnecessary emergency department (ED) visits using clinician judgment as the gold standard of appropriateness. This study evaluated the reliability of clinician judgment for assessing appropriateness of pediatric ED visit.

Methods: We conducted a retrospective cohort study comparing 3 clinicians' determination of ED visit appropriateness with and without guidance from a three-question structured algorithm. We used a cohort of scheduled ED return visits deemed appropriate by the index treating clinician between May 1, 2012, and April 30, 2013. We measured the level of agreement among three clinician investigators with and without use of the structured algorithm.

Results: A total of 207 scheduled ED return visits were reviewed by the primary clinician reviewer who agreed with the index treating clinician for 79/207 visits (38.2%). Among a

From the *University of British Columbia, Division of Emergency Medicine, Department of Pediatrics, BC Children's Hospital Vancouver, BC; †University of British Columbia, Faculty of Medicine, Vancouver, BC; and ‡University of British Columbia, Department of Pediatrics, BC Children's Hospital, Vancouver, BC.

Correspondence to: Dr. Quynh Doan, BC Children's Hospital, 4480 Oak Street, Shaughnessy Building B428-B, Vancouver, BC, V6H 3N1; Email: qdoan@bcchr.ca

traitant de référence. Par ailleurs, dans un sous-groupe aléatoire comptant 90 visites de suivi soumises à l'examen des 3 cliniciens, le taux de concordance était de 67% et la valeur de kappa, de 0,30 (0,17–0,44). Enfin, lorsque l'algorithme de 3 questions, fondé sur des critères objectifs a été appliqué, le taux de concordance avec le clinicien traitant de référence est passé à 55,6% (115/207).

Conclusion: Si les consultations inutiles ou non pertinentes jouent un rôle important dans l'encombrement des SU pédiatriques, il est toutefois difficile de les cerner. En effet, les résultats de l'étude démontrent la faible constance du jugement des

cliniciens dans la détermination de la pertinence des visites de suivi au SU, probablement en raison de la variabilité des prises de décision clinique et de la tolérance au risque, ainsi que de l'existence de facteurs sociaux et de facteurs systémiques qui ont une incidence sur l'accès aux soins de santé et sur leur utilisation. Aussi est-il recommandé qu'à l'avenir les études visant à évaluer la pertinence des consultations au SU s'appuient sur des critères objectifs, uniformes, plutôt que sur le seul jugement des cliniciens.

Keywords: Pediatrics, health services, return visits

INTRODUCTION

Pediatric emergency departments (EDs) face increasing visit volumes, often leading to overcrowding.¹ ED overcrowding is associated with increased wait times, reduced quality of care, decreased satisfaction, and adverse health outcomes.² While delay in transfer of boarded patients is a common cause of crowding in general EDs, increased input is a primary cause in pediatric EDs.³ As a result, efforts to mitigate ED crowding have sought to identify factors associated with “inappropriate” use.⁴

Unfortunately, there is no consensus definition of ED visit appropriateness. A wide range of methods have been used to evaluate appropriateness of ED visits, but few have been validated. A 2011 systematic review of ED literature, counted 51 different approaches, including implicit criteria (patient's self-assessment or physician assessment); explicit criteria; triage score; symptoms or tests and procedures; and hospital admission. This variability highlights the complexity of defining appropriate emergency use.⁵

Despite widespread use of implicit clinician judgment to classify ED visit appropriateness, little is known about the reliability of this methodology. This brief report sought to measure agreement among clinician reviewers on the appropriateness of a scheduled ED return visit with and without the aid of a standardized three-question guide.

METHODS

Study design

We conducted a retrospective cohort study of all ED visits from May 1, 2012, to April 30, 2013, with a return

visit within 7 days of the index visit. The study protocol was approved with waived consent by the institution's Research Ethics Board as a sub-analysis of a larger study.⁶

To evaluate the reliability of clinician judgment to determine visit appropriateness, we compared the determination of a primary clinician reviewer guided by survey questions against that of the index visit treating clinician. We then assessed interrater agreement among three clinician reviewers on a random subset of 90 scheduled ED return visits.

Study setting and population

We conducted this study at a tertiary pediatric ED that cares for >40,000 children age <17 annually. We limited the study cohort to a subset of scheduled ED return visits determined appropriate by the index visit clinician. We excluded scheduled returns for parenteral antibiotic therapy as data from this cohort were previously published.⁷

Study protocol

The primary reviewer reviewed the entire cohort of scheduled ED return visits guided by a two-part survey and was asked to determine, using clinical judgment, whether the visit was necessary (Supplemental Online Appendix 1). Two additional blinded reviewers subsequently reviewed a random subset of 90 visits using the same survey guide and were asked to evaluate the necessity of the visit. Finally, a standardized algorithm of three dichotomous questions (did the visit result in an admission; were pediatric ED-specific diagnostic/therapeutic interventions administered; and did the visit need to be

Table 1. Visit characteristics for scheduled first ED return visits judged by clinician as appropriate, stratified by method of determination (index visit clinician, primary clinician reviewer, and algorithm)

Scheduled 1st return visit characteristics	Index visit clinician	Primary clinician reviewer	Algorithm	p-Value
N deemed appropriate	207	79	115	
Return scheduled on weekday N (%)	122 (58.9)	34 (43.0)	54 (47.0)	0.21
Median length of stay in minutes (IQR)	133 (72.5, 229.5)	157 (94, 251.5)	184 (101, 311)	0.029
RTED scheduled by off-service N (%)	26 (12.6)	11 (1.4)	19 (16.5)	0.62
Consultant involvement in care N (%)	39 (18.8)	14 (17.7)	23 (20.0)	0.92

IQR = interquartile range; RTED = return to ED.

scheduled during nonoffice hours) was compared with clinical judgment regarding visit appropriateness.

Patient visit information sets and evaluations were created online and by means of REDCap and exported for analyses.

Measures

The primary objective was to measure agreement with the index treating clinician on the appropriateness of the return visit by the primary reviewer and among all three clinical reviewers. The secondary objective was to evaluate agreement between the primary reviewer's clinical judgment and the three-question structured algorithm.

Data analysis

We used descriptive statistics to report on all outcomes as reviewed by the primary clinician reviewer. Agreement between the 3 reviewers was reported using Fleiss' kappa statistic. The primary objective sample size was determined by the total number of scheduled ED return visits meeting inclusion criteria. To determine with 95% confidence, an estimated 85% agreement with $\pm 7.5\%$ precision, 90 randomly selected visits were assessed by three reviewers. We used Microsoft Excel for Mac 2017 Version 15.4 and Stata Version 11.0 for analyses.

RESULTS

Study population

Between May 1, 2012, and April 30, 2013, there were a total of 42,413 pediatric ED visits: 2,962 (6.98%) index visits had one or more associated returns to the ED

and 669 (22.6%) were scheduled. After excluding scheduled outpatient parenteral antibiotic therapy visits, the final cohort included 207 (30.9%) index visits with 232 scheduled returns.

The cohort's mean age was 4.6 years (95% confidence interval [CI], 3.3–5.9) and 54.1% were male. More than 75% were triaged Canadian Triage and Acuity Scale (CTAS) 3 or 4.

Appropriateness using clinical judgment

Among the 207 index visits associated with at least one scheduled ED return visit, the primary reviewer identified 79 (38.2%, 95% CI, 32.6–45.2%) as appropriate using clinical judgment (Table 1). For the subset of 90 cases evaluated by the two additional reviewers, there was 67.4% (95% CI, 60.4–74.4%) agreement among reviewers with a Fleiss' Kappa of 0.30 (95% CI, 0.17–0.44).

Appropriateness using standardized algorithm

The three-question standardized algorithm based on the primary reviewer's survey responses identified 115/207 (55.6%) appropriate return visits (Supplementary Online Appendix 2). Agreement among the 3 investigators increased to 83.7% (95% CI, 77.7–89.7%), Fleiss' kappa 0.67 (95% CI, 0.55–0.79) using the algorithm.

DISCUSSION

In our pediatric ED cohort of scheduled ED return visits, we found poor reliability of clinician judgment to determine visit appropriateness, despite guidance from survey questions highlighting specific patient and visit characteristics. Although pediatric emergency clinicians

surveyed in 2005 agreed that implicit or physician chart review was the most effective way to evaluate visit appropriateness, our results suggest that clinician judgment is unreliable.⁸

Our findings were consistent with earlier studies of clinician judgment in adult emergency visits. O'Brien and colleagues compared the use of triage complaint, 10 explicit criteria, and consensus of two ED clinicians.⁹ Kappas ranged from fair to moderate: triage and explicit, 0.39; triage and physicians, 0.42; and explicit and physicians, 0.42; agreement between the two emergency physicians, 0.4.⁹

No previous studies have reported interrater reliability for clinician assessment of ED visit appropriateness in the pediatric population. De Angelis described the development of a set of criteria based primarily on diagnoses and complaints applied by pediatricians to determine appropriateness of ED visits made by patients in their practice.¹⁰ Neither the original study, nor subsequent publications using these criteria, however, have reported criteria reliability or validity.¹⁰

While our study did not use validated explicit criteria, we found an increase in agreement when limiting reviewers' input to three objective criteria, which improved agreement from 32% to 52% over clinical judgment alone. These results further support the use of explicit rather than implicit approaches in future quality improvement or research.

Limitations

This retrospective review used abstracted data from the medical chart and limited detail may have contributed to the poor agreement between reviewers and the index treating clinician. Our algorithm is neither validated nor intended as a gold standard for visit appropriateness, but rather used to compare a structured approach using dichotomous questions with clinical judgment. Finally, our cohort of scheduled ED return visits at a single center was not meant to define appropriate visits and may not be generalizable. This should not affect the primary objective, which was to measure agreement among clinician reviewers using clinical judgment of appropriateness.

CONCLUSIONS

Although increasing visit volumes contribute to pediatric ED crowding, and unnecessary visits have been the

subject of prior efforts to mitigate overcrowding, our results demonstrated poor reliability of clinician judgment to determine appropriateness of ED visits. Given the complexities in defining visit appropriateness and poor reliability of clinician judgment, future projects targeting appropriateness of ED utilisation should favor objective or explicit criteria rather than clinician judgment alone.

Supplemental material: The supplemental material for this article can be found at <https://doi.org/10.1017/cem.2019.473>.

Acknowledgments: The authors acknowledge Karly Stillwell, the divisional research coordinator for her generous assistance in this study.

Competing interests: The authors have no conflicts of interest to disclose.

REFERENCES

1. Doan Q, Genuis ED, Yu A. Trends in use in a Canadian pediatric emergency department. *CJEM* 2014;16(05):405–10.
2. Bond K, Ospina MB, Blitz S, et al. Frequency, determinants and impact of overcrowding in emergency departments in Canada: a national survey. *Healthc Q* 2007;10(4):32–40.
3. Stang AS, McGillivray D, Bhatt M, et al. Markers of overcrowding in a pediatric emergency department. *Acad Emerg Med* 2010;17(2):151–6.
4. Hilditch JR, Scheftsk-Pedery A, Swain JR, Dyson A, Wright M. Judging the appropriateness of patients' visits to the emergency department. *Can Fam Physician* 1982;28:686–8.
5. Durand AC, Gentile S, Devictor B, et al. ED patients: how nonurgent are they? Systematic review of the emergency medicine literature. *Am J Emerg Med* 2011;29:333–45.
6. Doan Q, Goldman RD, Meckler GD. Management practice-related and modifiable factors associated with paediatric emergency return visits. *Pediatr Child Health* 2018;24(1):e1–e7.
7. Xu M, Doan Q. Outpatient parenteral antimicrobial therapy and judicious use of pediatric emergency resources. *Pediatr Emerg Care* 2017. doi: 10.1097/PEC.00000000000001215.
8. Brousseau DC, Mistry RD, Alessandrini EA. Methods of categorizing emergency department visit urgency: a survey of pediatric emergency medicine physicians. *Pediatr Emerg Care* 2006;22(9):635–9.
9. O'Brien GM, Shapiro MJ, Woolard RW, O'Sullivan PS, Stein MD. "Inappropriate" emergency department use: a comparison of tree methodologies for identification. *Acad Emerg Med* 1996 Mar 1 [cited October 30, 2018];3(3):252–7. Available at: <http://doi.wiley.com/10.1111/j.1553-2712.1996.tb03429.x> (accessed November 19, 2019).
10. Deangelis C, Fosarelli P, Duggan AK. Use of the emergency department by children enrolled in a primary care clinic. *Pediatr Emerg Care* 1985;1(2):61–5.