CODEN [USA]: IAJPBB ISSN: 2349-7750

INDO AMERICAN JOURNAL OF
PHARMACEUTICAL SCIENCES

http://doi.org/10.5281/zenodo.1480872

Available online at: http://www.iajps.com *Review Article*

SYSTEMATIC REVIEW OF EVIDENCE GRADING SYSTEMS USED IN INPATIENT PEDIATRIC GUIDELINES

[1]**Dr. Usman Rasheed**, [2]**Dr. Iqra Khalid and** [3]**Dr. Tasneem Fatima**

[1]Sharif Medical & Dental College, Lahore, Pakistan

[2]Punjab Medical College, Faisalabad, Pakistan

[3]Allama Iqbal Medical College, Lahore, Pakistan

Abstract:

Pediatric guidelines are being developed worldwide. There are a number of different systems in use for grading the evidence used to create guidelines. The objective of this study was to systematically review the overall quality of the evidence grading systems used in pediatric guidelines for the five most prevalent inpatient conditions developed in countries from the Organization for Economic Co-operation and Development (OECD). The paper includes guidelines, guideline summaries, and guideline methodology papers were included. Two reviewers independently assessed the studies for eligibility. Data was extracted to a data abstraction form. Details of the evidence grading systems were extracted directly into tables. There were 14 guidelines and 10 different evidence-grading systems that met inclusion criteria. Three methodologic papers and 5 summaries were also identified. Three different evidence-grading systems were noted to score well in all three domains of quality, quantity, and consistency as previously defined by the Agency for Healthcare Research and Quality (AHRQ). There was greater variability in the domains of quantity and consistency. The main limitation is this systematic review was limited to pediatric guidelines from the OECD countries and English language limiting its generalizability to other guidelines. There is still great variability in evidence grading systems used in pediatric guidelines. Standardization would improve the transparency and clarity of pediatric guidelines and the evidence grading behind them.

Key Words: *guidelines, evidence, pediatric, child, children*

Corresponding author:

Usman Rasheed,

Sharif Medical & Dental College,
Lahore,
Pakistan

QR code

Please cite this article in press Usman Rasheed et al., ***Systematic Review of Evidence Grading Systems Used In Inpatient Pediatric Guidelines.****, Indo Am. J. P. Sci, 2018; 05(11).*

INTRODUCTION:

In response to the variability in physicians' clinical observations, perceptions, reasoning, conclusions, and practices [1], clinical practice guidelines were developed to condense the medical literature into a usable format to help physicians and patients make decisions in specific clinical circumstances [2]. The Institute of Medicine in 2011 defines clinical practice guidelines as, "*statements that include recommendations intended to optimize patient cares that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options [3]"* Guidelines attempt to refine clinical questions and balance trade-offs of benefit vs. risk of an intervention and the alternatives. Guidelines give recommendations with the designed purpose to influence a physician's care of a patient [4].

There has been rapid proliferation of guidelines in the past twenty years, and in the Agency for Healthcare Research and Quality (AHRQ) website, National Guideline Clearinghouse, half of the listed guidelines are indexed as applicable to children [5], though the quality doesn't seem to be improving. In a recent study by Isaac *et al.,* the American Academy of Pediatrics (AAP) guidelines and AAP endorsed guidelines were found to not have significantly improved over the 10 years that they were evaluated by the Appraisal of Guidelines for Research and Evaluation II (AGREE-II) instrument though two different guideline policy statements were published by the AAP in 2004 and 2008.

If summarizing the available evidence is done well using an evidence summary grading system, the summary of the quality of evidence should be similar for different guideline panels reviewing the same disease process or patient population [6]. However, there are a number of different methods available for summarizing the quality of evidence for use in guidelines. This can lead to confusion as there are a large number of alphanumeric systems used to summarize and evaluate the overall quality of evidence [7].

In a review performed by the Agency for Healthcare Research and Quality (AHRQ), there were forty different systems used for grading the strength of a body of evidence available in the literature. Only three specific systems reported are responsible for the development of both adult and pediatric guidelines [8]. A subsequent review by the Canadian Optimal Medication Prescribing and Utilization Service (COMPUS) found over 60 different systems.[9] Due to the high number and variation in the evidence evaluation systems, the message regarding the quality and strength of the body of evidence can often be misunderstood by providers trying to use the guidelines in practice [10].

Grading Recommendations, Assessment, Development and Evaluation (GRADE), a group of over 200 health professionals, researchers, and guideline developers came together in 2000 to develop a better more transparent system for overall evidence quality evaluation and determining the strength of recommendation for clinical practice guidelines [11]. To date, over 70 organizations have endorsed GRADE as a means to assess the overall quality of the body of evidence and provide the strength of recommendation for a clinical care guideline [12]. Though GRADE is gaining in popularity, it is inconsistently used and when utilized it is often modified against the recommendations of the GRADE collaborative hindering the development of a single standardized approach to overall evidence quality assessment. No pediatric specific organizations have endorsed the GRADE guidelines [13]. Pediatric guidelines developed worldwide for common pediatric conditions vary in the evidence assessment systems used to evaluate the quality of evidence available to make a recommendation for a clinical care guideline. Differences in these systems have not been clearly demonstrated or described previously for pediatric guidelines. The number of different systems in use, their differences, and modifications made prior to use, and the validity of each of these tools will be important to illustrate for pediatric guideline developers going forward. It may impact whether pediatric guideline developers continue to use systems developed specifically by the associations or organizations that produce these guidelines or if a more standardized approach without modifications is developed and adopted going forward.

The objective of this study was to conduct a systematic review to assess the evidence grading systems used by various international organizations in development of pediatric inpatient clinical care guidelines on their ability to satisfy the three domains previously described by the AHRQ of quality, quantity and consistency of evidence synthesis.

METHODS:

Protocol and Registration

The protocol for this systematic review was not eligible for registration though it is available for review via contact with the primary author. We used the Preferred Reporting Items for Systematic

Reviews and Meta-Analysis statement for reporting this review [14].

Information Sources and Search

We searched Medline and Embase up to June 4, 2014 using MeSH headings and textwords. The exact search terms for each database are shown in Appendix 1. We limited our searches to children and 2003 and later as we were interested in pediatric guidelines and evidence synthesis systems published in the last 10 years. We also extensively searched the gray literature for additional guidelines including the following sites: Guidelines and Audit Implementation Network, National Guidelines Clearing House (U.S.), National Health and Medical Research Council (Australia), New Zealand Guidelines Group, National Institute for Health and Clinical Excellence (NICE-United Kingdom), Scottish Intercollegiate Guidelines Network (SIGN), Trip Database, Guidelines International Network (G-I-N), American Academy of Pediatrics (AAP), and the Canadian Pediatric Society (CPS). This search was conducted under the guidance of an information scientist.

Eligibility Criteria

Guidelines on the top 5 most prevalent inpatient pediatric conditions of asthma, bronchiolitis, pneumonia, cellulitis, and gastroesophageal reflux were used as a means to identify the evidence synthesis systems used in general pediatric guidelines published within the last 10 years from the 34 countries currently participating in the Organisation for Economic Co-operation and Development (OECD).[15] We also included any guideline summaries and methodology papers that contained information on the evidence grading systems. Guidelines were excluded if no evidence grading system was used, they were regional, the focus was surgical treatment, they were specific to critical care or neonatal care, they were specific to outpatient management, they focused solely on patient education, they were outdated, were symptom focused (i.e. cough rather than asthma), or if they were written in a language other than English.

Study Selection

The results of the literature search were imported into EndNote and duplicates were removed. Two authors independently screened the titles and abstracts for relevant studies. Studies deemed relevant were further screened by the two authors independently and assessed for inclusion based on the inclusion criteria. Any discrepancies were resolved though discussion. If agreement was not reached, a third author was used for arbitration of the decision.

Data Collection Process

Two independent review authors who were not blinded abstracted the data onto structured data abstraction forms including date, reference number, citation, inclusion criteria, exclusion criteria, and the outcomes of quality, quantity, consistency, modifications made, and validation of the system if found. The details of the evidence grading systems were abstracted directly into the summary table by each of the review authors and discrepancies were reviewed. Any discrepancies between the two authors were first resolved through discussion. If resolution could not be found, a third review author was used for arbitration of the decision. The results of the data abstraction from the guidelines and supporting articles were summarized in a table of included studies (Table 1).

Data Items

Recent research prioritization work by the Pediatric Research in the Inpatient Settings (PRIS) network ranked the top 50 most prevalent conditions in hospitalized children, for priority for further comparative effectiveness research in pediatrics within the United States.[16] To start with identification of guidelines, the five most prevalent conditions were used for saturation of guideline evidence synthesis systems used by various guideline organizations worldwide.

The evidence synthesis mechanisms used for any guideline from the OECD countries produced in the last ten years for the top five pediatric inpatient conditions of asthma, bronchiolitis, pneumonia, cellulitis, and gastroesophageal reflux were described and evaluated. The focus was on guidelines produced within the past ten years, as those previous to ten years ago are likely out of date, no longer followed, and less likely to have a robust evidence grading system used in their development. The primary outcome of interest was the quality of the overall guideline evidence grading system as judged using the AHRQ domains: quality of the aggregated studies used based on minimizing bias, quantity (magnitude of the effect, number of studies, sample size or power), and consistency (how well different studies and study types findings agree)[8]. These were used to determine the overall quality of the guideline evidence grading system. The secondary outcomes of interest were 1) the determination of whether the evidence evaluation system was modified from an original version and type of modification and reason(s) for modification and 2) comparing internal and external validity of the different evidence evaluation tools if validation has been explored and reported. Thus, the evidence assessment grading

systems were described, the overall quality of the guidelines evidence grading systems were judged using the domains first developed by the AHRQ, any modifications to the grading system if adapted were explored, and validity if available was reported.

Summary Measures

The methodological quality of the evidence systems was judged on the three domains first described by the AHRQ for assessing systems for grading the strength of a body of evidence: quality, quantity, and consistency. These were graded as yes, no, and partially met during the analysis by the AHRQ. Grading systems that considered at least two of the following criteria for quality: study design, conduct, analysis, or methodologic rigor was given a score of Yes on quality. A Yes for quantity meant the system incorporated at least two of the three elements required: the magnitude of the effect (estimate effects such as mean differences, odds ratios, relative risks, etc.), the number of studies performed on the topic of interest, and the number of individuals studied, combined when feasible to provide confidence interval widths and effect estimates. Consistency was determined as the degree of agreement of studies within the body of scientific evidence. This was treated as a dichotomous variable. A Yes rating was given if the concept of consistency was considered.[8] We used this same technique for our review. The results of the findings of the evidence synthesis tools were reported and explained in table and narrative fashion.

Two review authors independently completed these method assessments. Any discrepancy between the two authors was resolved through discussion. If an agreement was not reached, a third author arbitrated the decision.

Synthesis of Results

The domains of quality, quantity and consistency are shown in table format below as excerpted from the AHRQ report (Table 2) [8]. We followed the scoring guidelines outlined by the AHRQ for use of these domains and the definitions for yes, no and partial.[8] We also commented on if modification of the system occurred prior to use for guideline development as well as on the validity of the evidence synthesis systems if available.

Also, over time, different guideline organizations may have updated their evidence synthesis methods. Any changes or modifications to the evidence synthesis system used over time were highlighted.

Risk of Bias Across Studies

Risk of bias across studies was not separately assessed, as minimizing bias was part of the quality domain of the AHRQ criteria.

Additional Analyses

A subgroup analysis was specified apriori on evidence synthesis systems used in global guidelines compared to national guidelines identified for an inherent difference in the quality of evidence synthesis systems used internationally compared to national organizations. Post hoc qualitative analyses were performed to evaluate changes in quality of evidence evaluation systems over time by year, comparing multi-organization to single organization endorsement guideline grading systems, and comparing pediatric organizations to those organizations responsible for both pediatric and adult guidelines grading systems. No sensitivity analysis was performed.

RESULTS:

Using our search strategy, we identified and screened 7,459 citations, of them 1,558 were duplicates. Twenty-six guidelines, summaries, and methodology documents were identified, and four are still pending assessment once the original guideline is obtained (Figure 1). Characteristics of the included studies were included in Table 1. After initial screening of titles and abstracts, 56 articles were identified and full text was obtained.

Thirty of the articles were excluded after obtaining full text for the following reasons: 17 were guidelines without evidence evaluation systems represented, 10 were outdated with a more recent version available, 1 was not applicable to inpatient treatment, and 2 were guidelines pertaining to more than 1 condition (Appendix 2).

Study Characteristics

A total of 14 guidelines were identified and evaluated. Eight guidelines on asthma were identified [17-24], three on bronchiolitis [25-27], two on pneumonia [28,29], none on cellulitis specifically (one on methicillin resistant staphylococcus aureas (MRSA) was excluded as it included multiple conditions caused by MRSA), and one on gastroesophageal reflux [30] (Table 3). There were a total of 10 different evidence-grading systems used in these guidelines. Five different systems were found for the asthma guidelines as two of the guidelines were developed by the same organization using the same evidence grading system for both guidelines [18-20], and another two of the guidelines used a system adopted from Jadad et al. 2000 [31]. For the bronchiolitis evidence grading systems, two different

evidence-grading systems were used. One guideline organization adopted a system that had already been used by a different organization [22]. There were two different evidence evaluation systems used for the two different pneumonia guidelines, and a different system used by the gastroesophageal reflux guideline. There were three methodology papers identified in the search as well as five summary statements supporting different guidelines (Table 4). Often, supporting evidence tables were found only on the website and not included in the guideline document or summary statements themselves.

RESULTS OF INDIVIDUAL STUDIES

The results of the quality of the evidence grading systems used within the 14 guidelines are summarized in Table 5.

Quality

The quality judgment was based on minimizing bias requiring that two of the following criteria were met: study design, conduct, analysis, or methodologic rigor discussed.

Study design was the most frequently mentioned quality factor within all the grading systems with RCTs and meta-analysis getting the highest grade. Most grading systems also commented on the conduct of the study [17,22,24-27,29,32,33]. All but one system included at least two of these factors in the evidence grading system used [34].

Quantity

The quantity construct was based on incorporating at least two of the three factors into the evidence grading system constituting the quantity domain: the magnitude of effect, number of studies performed, and the number of individuals studied. Seven of the studied guideline grading systems incorporated at least two of these domains into their evidence grading system [17-21,24,32,33]. In some instances, this information was obtained from evidence tables available on the website [17-35]. Two guideline-grading systems received a partial rating as they commented on the number of studies performed, but did not address the magnitude of the effect or number of individuals studied [26,27]. Four evidence grading systems did not address quantity [22,25,29,34].

Consistency

Consistency was evaluated as a dichotomous outcome without a partial option. The consistency domain was evaluated on the extent of similar findings were reported from the body of evidence regardless of the type of study design used. Eight guideline evidence evaluation systems addressed consistency [17,1820,22,24,26,27,32,33]. Five of the systems evaluated did not address consistency [19,21,25,29,34]. This domain was the most frequently omitted of the three domains used by the AHRQ.

All domains

The guidelines with the best overall evidence grading systems were the European Respiratory Society (ERS) and American Thoracic Society (ATS) International ERS/ATS Guideline on the Definition, Evaluation, and Treatment of Severe Asthma, the two guidelines by GINA (Global Initiative for Asthma (GINA) 2012/Global Strategy for the Diagnosis and Management of Asthma in Children 5 Years or Younger), the British Guideline on the Management of Asthma: A National Guideline by the Scottish Intercollegiate Guidelines Network (SIGN) and the British Thoracic Society (BTS), The National Heart, Lung, and Blood Institute (NHLBI), National Asthma Education and Prevention Program (NAEPP), Expert Review Panel-3 (EPR-3): Guidelines for the Diagnosis and Management of Asthma, and The Management of Community –Acquired Pneumonia in Infants and Children Older than 3 Months of Age: Clinical Practice Guidelines by the Pediatric Infectious Disease Society (PIDS) and Infectious Disease

Society of America (IDSA). All of these guidelines scored yes in all three domains of quality, quantity, and consistency as defined by the AHRQ. The evidence grading systems used were Grading Recommendations Assessment, Development, and Evaluation (GRADE) employed by the ERS/ATS guideline and the PIDS/IDSA guideline [10], the SIGN methodology for the SIGN/BTS asthma guideline [36], and adaption of the Jadad et al. 2000 method for the GINA and NHLBI/NAEPP guideline [31]. No studies regarding the validation of these evidence-grading systems were identified in the literature.

The guidelines sponsored solely by the Scottish Intercollegiate Guideline Network (SIGN) and those that used the SIGN methodology (The Spanish National Healthcare System Ministry for Health and Social Policy) were given a Partial rating for quantity as no evidence tables were found for these guidelines so it was more difficult to give a complete score for quantity though it was partially discussed in the body of the guideline text. The Veteran's Affairs (VA)/Department of Defense (DoD) guideline and the Canadian Thoracic Society (CTS) both used different evidence grading systems that did not address consistency. The VA/DoD guideline used the

US Preventative Services Task Force Method (USPSTF) method from 2001 whereas the CTS guideline used a system modified from The American College of Chest Physicians Task Force in 2006.[37] The Spanish Guidelines for Asthma Management (GEMA) guideline evaluation system did not comment on the quantity of studies though this system was reported as loosely based on GRADE methodology.

The guideline evaluation systems that scored the worst in the AHRQ domains were the ones housed within the American Academy of Pediatrics (AAP) guideline on bronchiolitis, the British Thoracic Society (BTS) guideline on pneumonia, and the North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition (NASPGHAN)/European Society for Gastroenterology, Hepatology, and Nutrition (ESPGHAN) guideline on gastroesophageal reflux. The AAP developed it's own evidence grading system as outlined in a report from 2004.[7] The BTS guidelines were based off a system used for the evidence grading in the community acquired pneumonia guidelines for adults published in 2009, explaining the discrepancy in score between this guideline and the one co-sponsored with SIGN. The NASPGHAN/ESPGHAN guideline for gastroesophageal reflux had the most rudimentary system evaluated leading to the poorest overall score. The evidence grading system for this guideline was adopted from the Oxford Center for Evidence-Based Medicine Levels of Evidence though little information is given regarding the scoring done.

Additional Analyses

When comparing global guidelines to national guidelines, the two global guidelines for asthma (GINA, ERS/ATS) scored complete marks for quality, quantity and consistency. The NASPGHAN/ESPGHAN guideline on gastroesophageal reflux scored poorly with only a partial score for quality. However, this was the oldest international guideline created in 2009. There was substantial variation in national guideline evidence scoring system quality.

Additionally, evidence-grading systems seemed to improve over time though this wasn't entirely consistent (Table 6). Those guidelines with multiple societies sponsoring scored better for evidence system grading quality overall than single society sponsorship (Table 7). Two obvious examples are the SIGN and BTS society sponsored guidelines. Those developed by each society individually did not score as well as the guideline developed in tandem using the SIGN methodology. It scored better than the SIGN guideline as it had evidence tables available on the website clearly defining all measures of quality, quantity and consistency whereas these were not available for the SIGN guideline. Overall, the SIGN methodology for evidence grading was superior to the system used by the BTS for development of the pneumonia guideline.

There were two guidelines developed by pediatric focused organizations vs. those organizations encompassing both children and adults. In general, the guidelines from the strictly pediatric societies had less robust evidence grading systems than those developed by societies encompassing both pediatrics and adults (Table 8).

DISCUSSION:

Summary of Evidence

There was wide variation in the quality of the different evidence grading systems used in the pediatric guidelines for the five most prevalent inpatient pediatric conditions with a general trend in improvement with more recent guidelines though this was not completely consistent. Interestingly, there is still substantial variation in evidence grading system quality in pediatric guideline development twelve years after the initial publication of the AHRQ's report in 2002 and fourteen years after the introduction of GRADE collaborative first established in 2000. The AHRQ noted in their report that those evidence grading systems used specifically for guidelines were lagging behind the systems that were not used for guideline development and that this lag seemed to be increasing over time [8]. Though we did not evaluate evidence grading systems used for other purposes than guideline development, the results of our systematic review supports the fact that the systems currently used for guideline development in pediatrics are variable in their inclusion of the important features for an evidence grading system more than 10 years after these findings were originally published by the AHRQ. This was more recently substantiated in the 2011 Institute of Medicine's report, "Guidelines we can Trust", that stated that the major problems with guidelines still includes the lack of rigorous methodology to develop guidelines [3].

The findings that multi-organization sponsored guidelines had more robust evidence grading systems as well as those that were sponsored by organizations encompassing both pediatric and adult care versus solely pediatric care are intriguing. No evidence was found discussing the impact of multi-organization sponsorship on evidence. This may be an area for

further study in the future. An earlier study from 2000 showed that specialty society guidelines were unsatisfactory in their reporting of evidence grading. Eighty-two percent did not give any explicit rating of the evidence for guidelines evaluated from 1988-1998 [38]. Our results suggest that specialty societies such as pediatric specific societies may still be lagging behind other organizations in their use of rigorous evidence grading methods.

Limitations

There are several limitations to our study. The top five most prevalent pediatric inpatient conditions were used as a means to find guidelines from OECD countries for evaluation of evidence grading systems used within the guidelines with the goal of saturation. There is the possibility that saturation was not reached, and further systems would have been identified with inclusion of a greater number of conditions. However, repetition of organizations was noted in our retrieval suggestive of a good sampling of guidelines and thus guideline evaluation systems used in pediatric guideline development. Since our search was limited to pediatric guidelines from OECD countries that were English speaking, these findings would not be generalizable to developing country guidelines or potentially those written in another language. Guideline grading systems are continually evolving as the body of evidence around guideline development continues to advance.

Thus, it is possible that the evidence grading systems have been updated since these guidelines were published. However, this systematic review gives a good overview of the state of evidence grading systems used in pediatric guidelines from 2003 forward. Much of the supporting material for guidelines is housed online. It is possible that we did not find all the supporting documents for the guidelines despite thorough searching.

CONCLUSIONS:

In general, there continues to be variability in the evidence grading systems used in pediatric guidelines. The variability in evidence grading systems continues to make it difficult to interpret the quality of evidence behind recommendations in pediatric guidelines. Most systems have been modified over time and some have been adapted or adopted from different previous publications. A single evidence grading system applied to all guidelines developed in pediatrics would be very beneficial for clarity and transparency. GRADE has aimed to be this system, and 14 years into development, it is slowly spreading. It too is at risk of modification by various organizations rather than full implementation that could potentially lead to further confusion. One unified system is likely the best approach, and we are not there yet with pediatric inpatient guidelines for the five most common conditions.

REFERENCES:

1. Wachter RM, Goldman L. The emerging role of "hospitalists" in the American health care system. *The New England Journal of Medicine.* Aug 15 1996;335(7):514-517.
2. Society of Hospital Medicine. Definition of a Hospitalist and Hospital Medicine. Accessed November 27, 2014; http://www.hospitalmedicine.org/Web/About_SHM/Hospitalist_Definition/About_SHM/Industry/Hospital_Medicine_Hospital_Definition.aspx?hkey=fb083d78-95b8-4539-9c5b-58d4424877aa.
3. Starfield B, Simpson L. Primary care as part of US health services reform. *The Journal of the American Medical Association.* Jun 23-30 1993;269(24):3136-3139.
4. Wachter RM, Goldman L. The hospitalist movement 5 years later. *The Journal of the American Medical Association.* Jan 23-30 2002;287(4):487-494.
5. Wachter RM, Katz P, Showstack J, Bindman AB, Goldman L. Reorganizing an academic medical service: impact on cost, quality, patient satisfaction, and education. *The Journal of the American Medical Association.* May 20 1998;279(19):1560-1565.
6. Diamond HS, Goldberg E, Janosky JE. The effect of full-time faculty hospitalists on the efficiency of care at a community teaching hospital. *Annals of Internal Medicine.* Aug 1 1998;129(3):197-203.
7. Kroger E, Grant M. Are hospitalists truly more efficient? a natural experiment. *Journal of General Internal Medicine.* 1998;13:S66.
8. Stein MD, Hanson S, Tammaro D, Hanna L, Most AS. Economic effects of community versus hospital-based faculty pneumonia care. *Journal of General Internal Medicine.* 1998;13(11):774-777.
9. Craig DE, Hartka L, Likosky WH, Caplan WM, Litsky P, Smithey J. Implementation of a hospitalist system in a large health maintenance organization: the Kaiser Permanente experience. *Annals of Internal Medicine.* Feb 16 1999;130(4 Pt 2):355-359.
10. Freese R. Clinical, logistical, and political issues in creating a hospitalist system. *Annals of*

Internal Medicine. 1999;130:350-354.
11. Rifkin W, Conner D, Silver A, Eichorn A. Comparing hospitalists' and community based primary care physicians' care of pneumonia patients. *Journal of General Internal Medicine.* 2001;16:S215.
12. Rifkin W, Connor D, Eichorn A, Rose D. Comparison of hospitalists and primary care internists in the care of inpatients with pneumonia. *Journal of General Internal Medicine.* 1999;14:S118.
13. Davis KM, Koch KE, Harvey JK, Wilson R, Englert J, Gerard PD. Effects of hospitalists on cost, outcomes, and patient satisfaction in a rural health system. *The American Journal of Medicine.* Jun 1 2000;108(8):621-626.
14. Halpert AP, Pearson SD, LeWine HE, McKean SC. The impact of an inpatient physician program on quality, utilization, and satisfaction. *The American Journal of Managed Care.* May 2000;6(5):549-555.
15. Bellet PS, Whitaker RC. Evaluation of a pediatric hospitalist service: impact on length of stay and hospital charges. *Pediatrics.* Mar 2000;105(3 Pt 1):478-484.
16. Landrigan C, Srivastava R, Muret-Wagstaff S, Graef J, Homer C, Goldmann D. Outcomes of hospitalization in pediatric patients insured by HMOs: comparison of care by hospitalists and traditional academic providers. Paper presented at: Pediatric Research 2000.
17. Srivastava R, Landrigan C, Muret-Wagstaff S, Graef J, Goldmann D, Homer C. Impact of a managed care hospitalist system in academic pediatrics. Paper presented at: Pediatric Research 2000.
18. Srivastava R, Landrigan C, Muret-Wagstaff S, Ross-Degnan D. Cost savings for patients with acute conditions cared for by pediatric hospitalists in a tertiary care center. *Pediatric Research.* 2001;49:125A.
19. Auerbach A, Wachter R, Katz P. Implementation of a hospitalist service at a community teaching hospital: evolution of service utilization, costs, and patient outcomes. Paper presented at: Abstract Volume of the Third Annual Meeting of the National Association of Inpatient Physicians 2000.
20. Hackner D, Tu G, Braunstein GD, Ault M, Weingarten S, Mohsenifar Z. The value of a hospitalist service: efficient care for the aging population? *Chest.* Feb 2001;119(2):580- 589.
21. Kearns PJ, Wang CC, Morris WJ, et al. Hospital care by hospital-based and clinic-based faculty: a prospective, controlled trial. *Archives of Internal Medicine.* Jan 22 2001;161(2):235-241.
22. Maggioni A, Reyes M, Lifshitz F. Evaluation of a pediatric hospitalist service by APR- DRG's: impact on length of stay and hospital charges. *Pediatric Research.* 2001;49:691.
23. Palmer H, Armistead N, Marivannan S. The effect of a hospitalist service on patient care in an academic teaching hospital. *Journal of General Internal Medicine.* 1999;14(suppl):S116.
24. Palmer H, Delamata M, McBride C, Dunsworth T, Evans K, Hobbs G. After discharge effects of a hospitalist service. *Journal of General Internal Medicine.* 2001;16(suppl):S211.
25. Tingle LE, Lambert CT. Comparison of a family practice teaching service and a hospitalist model: costs, charges, length of stay, and mortality. *Family Medicine.* Jul-Aug 2001;33(7):511-515.
26. Molinari C, Short R. Effects of an HMO hospitalist program on inpatient utilization. *American Journal of Managed Care.* 2001;7(11):1051-1060.
27. Meltzer D, Shah M, Morrison J, Jin L, Levinson W. Decreased length of stay, costs and mortality in a randomized trial of academic hospitalists. *Journal of General Internal Medicine.* 2001;16(Suppl 1):S208.
28. *Clinical Practice Guidelines We Can Trust.* The National Academies Press; 2011.
29. Menna VJ. Inpatient care: the general pediatrician's future. *Pediatrics in Review / American Academy of Pediatrics.* Dec 1990;12(6):165-166.
30. Bellet PS, Wachter RM. The hospitalist movement and its implications for the care of hospitalized children. *Pediatrics.* Feb 1999;103(2):473-477.
31. Redelmeier DA. A Canadian perspective on the American hospitalist movement.*Archives of Internal Medicine.* Aug 9-23 1999;159(15):1665-1668.
32. Friedman J, Laxer RM. Providing after-hours on-call clinical coverage in academic health sciences centres: the Hospital for Sick Children experience. *CMAJ : Canadian Medical Association Journal = Journal de l'Association Medicale Canadienne.* Aug 8 2000;163(3):298-299.
33. Urowitz M, Crescenzi AM, Muharuma L. Residents' duty hours in the province of Ontario, Canada. *Academic medicine : Journal of the Association of American Medical Colleges.* Jan 2003;78(1):9-10.
34. Dwight P, MacArthur C, Friedman JN, Parkin PC. Evaluation of a staff-only hospitalist system in a tertiary care, academic children's hospital.

Pediatrics. Dec 2004;114(6):1545- 1549.
35. Beck CE, Parkin PC, Friedman JN. Pediatric hospitalist medicine: an overview and a perspective from Toronto, Canada. *Clinical Pediatrics.* Jul 2008;47(6):546-548.
36. Mahant S, Mekky M, Parkin P. Integrating pediatric hospitalists in the academic health science center: practice and perceptions in a Canadian center. *Journal of Hospital Medicine : an official publication of the Society of Hospital Medicine.* Apr 2010;5(4):228-233.
37. Fisher ES. Pediatric hospital medicine: historical perspectives, inspired future. *Current Problems in Pediatric and Adolescent Health Care.* May 2012;42(5):107-112.
38. Freed GL, Brzoznowski K, Neighbors K, Lakhani I. Characteristics of the pediatric hospitalist workforce: its roles and work environment. *Pediatrics.* Jul 2007;120(1):33-39.
39. Rauch DA, Lye PS, Carlson D, et al. Pediatric hospital medicine: a strategic planning roundtable to chart the future. *Journal of Hospital Medicine: an official publication of the Society of Hospital Medicine.* Apr 2012;7(4):329-334.
40. Keren R, Luan X, Localio R, et al. Prioritization of comparative effectiveness research topics in hospital pediatrics. *Archives of Pediatrics & Adolescent medicine.* Dec 2012;166(12):1155-1164.
41. Yu HW, Lauren M.; Elixhauser, A. *Hospital Stays for Children, 2009.* August 2011 2011.
42. Weisz G, Cambrosio A, Keating P, Knaapen L, Schlich T, Tournay VJ. The emergence of clinical practice guidelines. *The Milbank Quarterly.* Dec 2007;85(4):691-727.
43. Eddy DM. The challenge. *The Journal of the American Medical Association.* Jan 12 1990;263(2):287-290.
44. Institute of Medicine. *Clinical Practice Guidelines: Directions for a New Program.*Washington, DC: National Academy Press; 1990.
45. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *BMJ (Clinical Research Ed.).* Feb 20 1999;318(7182):527-530.
46. Institute of Medicine. *Clinical Practice Guidelines We Can Trust.* The National Academies Press; 2011.
47. Hayward RS, Wilson MC, Tunis SR, Bass EB, Guyatt G. Users' guides to the medical literature. VIII. How to use clinical practice guidelines. A. Are the recommendations valid? The Evidence-Based Medicine Working Group. *The Journal of the American Medical Association.* Aug 16 1995;274(7):570-574.
48. Parikh K, Hall M, Teach SJ. Bronchiolitis management before and after the AAP guidelines. *Pediatrics.* Jan 2014;133(1):e1-7.
49. Lenzer J, Hoffman JR, Furberg CD, Ioannidis JP. Ensuring the integrity of clinical practice guidelines: a tool for protecting patients. *British Medical Journal (Clinical Research Ed.).* 2013;347:f5535.
50. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet.* Nov 27 1993;342(8883):1317-1322.
51. Feuerstein JD, Akbari M, Gifford AE, et al. Systematic analysis underlying the quality of the scientific evidence and conflicts of interest in interventional medicine subspecialty guidelines. *Mayo Clinic Proceedings.* Jan 2014;89(1):16-24.
52. Koh C, Zhao X, Samala N, Sakiani S, Liang TJ, Talwalkar JA. AASLD clinical practice guidelines: a critical review of scientific evidence and evolving recommendations. *Hepatology (Baltimore, Md.).* Dec 2013;58(6):2142-2152.
53. Wilson MC, Hayward RS, Tunis SR, Bass EB, Guyatt G. Users' guides to the Medical Literature. VIII. How to use clinical practice guidelines. B. what are the recommendations and will they help you in caring for your patients? The Evidence-Based Medicine Working Group. *The Journal of the American Medical Association.* Nov 22-29 1995;274(20):1630-1632.
54. Cohen E, Uleryk E, Jasuja M, Parkin PC. An absence of pediatric randomized controlled trials in general medical journals, 1985-2004. *Journal of Clinical Epidemiology.* Feb 2007;60(2):118-123.
55. Isaac A, Saginur M, Hartling L, Robinson JL. Quality of reporting and evidence in American Academy of Pediatrics guidelines. *Pediatrics.* Apr 2013;131(4):732-738.
56. Alexander PE, Bero L, Montori VM, et al. World Health Organization recommendations are often strong based on low confidence in effect estimates. *Journal of Clinical Epidemiology.* Jun 2014;67(6):629-634.
57. Eisenberg JM. Globalize the evidence, localize the decision: evidence-based medicine and international diversity. *Health Affairs (Project Hope).* May-Jun 2002;21(3):166-168.
58. Eddy DM. Clinical decision making: from theory to practice. Resolving conflicts in practice policies. *The Journal of the American Medical Association.* Jul 18 1990;264(3):389-391.

59. Adhyaru BB, Jacobson TA. New Cholesterol Guidelines for the Management of Atherosclerotic Cardiovascular Disease Risk: A Comparison of the 2013 American College of Cardiology/American Heart Association Cholesterol Guidelines with the 2014 National Lipid Association Recommendations for Patient-Centered Management of Dyslipidemia. *Cardiology Clinics.* May 2015;33(2):181-196.
60. Parisi P, Vanacore N, Belcastro V, et al. Clinical guidelines in pediatric headache: evaluation of quality using the AGREE II instrument. *The Journal of Headache and Pain.* 2014;15:57.
61. Dhaliwal R, Madden SM, Cahill N, et al. Guidelines, guidelines, guidelines: what are we to do with all of these North American guidelines? *JPEN. Journal of Parenteral and Enteral Nutrition.* Nov-Dec 2010;34(6):625-643.
62. Tavender EJ, Bosch M, Green S, et al. Quality and consistency of guidelines for the management of mild traumatic brain injury in the emergency department. *Academic*
63. Hussain T, Michel G, Shiffman RN. The Yale Guideline Recommendation Corpus: a representative sample of the knowledge content of guidelines. *International Journal of Medical Informatics.* May 2009;78(5):354-363.
64. McAlister FA, Van Diepen S, Padwal RS, Johnson JA, Majumdar SR. How evidence-based are the recommendations in evidence-based guidelines? *PLoS Medicine.* 2007;4(8):e250.
65. McMurray J, Swedberg K. Treatment of chronic heart failure: a comparison between the major guidelines. *European Heart Journal.* Aug 2006;27(15):1773-1777.
66. Campbell F, Dickinson HO, Cook JV, Beyer FR, Eccles M, Mason JM. Methods underpinning national clinical guidelines for hypertension: describing the evidence shortfall. *BMC Health Services Research.* 2006;6:47.