Basole RC, et al. J Am Med Inform Assoc 2015;0:1–6. doi:10.1093/jamia/ocu016, Brief Communication

Understanding variations in pediatric asthma care processes in the emergency department using visual analytics

RECEIVED 1 May 2014
REVISED 28 September 2014
ACCEPTED 30 October 2014

Rahul C. Basole[1,2], Mark L. Braunstein[1,2], Vikas Kumar[2,3], Hyunwoo Park[2,4], Minsuk Kahng[3], Duen Horng (Polo) Chau[3], Acar Tamersoy[3], Daniel A. Hirsh[5,6], Nicoleta Serban[4], James Bost[5], Burton Lesnick[5], Beth L. Schissel[5,6], Michael Thompson[5]

ABSTRACT

Health care delivery processes consist of complex activity sequences spanning organizational, spatial, and temporal boundaries. Care is human-directed so these processes can have wide variations in cost, quality, and outcome making systemic care process analysis, conformance testing, and improvement challenging. We designed and developed an interactive visual analytic process exploration and discovery tool and used it to explore clinical data from 5784 pediatric asthma emergency department patients.

Key words: visual analytics, process mining, asthma, emergency care, pediatric hospital

BRIEF COMMUNICATION

INTRODUCTION

A recent Institute of Medicine report concluded that "systematic, evidence-based process improvement methods applied in various sectors to achieve often striking results in safety, quality, reliability, and value can be similarly transformative for health care".[1] Process improvement begins with understanding what is currently being done.[2] Correlating variations in current practices to cost and quality can lead to ideas for process improvement or redesign that can be tested in actual practice.

Domains such as manufacturing and retail have long analyzed digital data to understand and improve processes using various process mining approaches.[3] However, these business sectors have attributes often lacking in healthcare delivery. The existent processes may already be standardized. Automated systems are frequently deeply imbedded in them and report highly accurate subprocess data, including time and date information.

Healthcare delivery, however, is a very human-directed activity that is subject to wide variation, even in patients who appear similar.[4–6] Moreover, until recently, detailed digital care delivery data has generally not been available. This is changing because of the success of the federal government's HITECH program to foster electronic health record (EHR) adoption. Many observers believe that EHR data can and will increasingly be used to better understand and refine clinical processes to improve outcomes and reduce costs.[7,8]

Visual analytics can play a fundamental role in all IT-enabled healthcare transformation but particularly in healthcare delivery process improvement.[9,10] Interactive visual approaches are valuable as they move beyond traditional static reports and indicators to mapping, exploration, discovery, and sensemaking of complex processes. This enables decision makers to digest care process data, see patterns, spot trends, and identify outliers thereby improving comprehension, memory, and decision making. Notable work on healthcare data visualization includes patient education,[11] symptom evolution,[12] patient cohort analysis,[13] EHR data and design,[14,15] and patient care plans.[16] Care process visualization is still largely unexplored.

We report our experiences and findings in designing and developing an interactive visual analytic process exploration and discovery tool for understanding the state of pediatric asthma emergency department (ED) care. Detailed explanations of the methods, including data preprocessing and analysis, design, implementation, and evaluation can be accessed in Supplementary Appendix. Asthma is the most common pediatric chronic disease, affecting 9.3% (6.8 million) of children in the United States.[17] It is the third leading cause of hospitalization among children under the age of 15 years, a group with approximately 774 000 emergency room visits for asthma in 2009.[18]

We collaborated with the Scottish Rite Emergency Department of Children's Healthcare of Atlanta (Children's), the largest pediatric care system in the United States. Georgia's pediatric asthma rate is one of the highest in the nation, affecting more than 200 000 children age 18 or under. As a result, Scottish Rite's 2013 asthma claims were $22.3 million, with nearly 5500 ED visits resulting in over 1600 hospitalizations, 135 to intensive care. We conclude with implications and future research opportunities.

Correspondence to Associate Professor Rahul C. Basole, School of Interactive Computing & Tennenbaum Institute, Georgia Institute of Technology, 85 Fifth Street NW, Atlanta, Georgia 30332, USA; basole@gatech.edu; tel: +1.404.385.6269

For numbered affiliations see end of article.

METHOD

Data

We obtained de-identified ED clinical records of 5784 children seen for a primary diagnosis of asthma from January 2013 to January 2014 through Population Discovery, Children's state-of-the-art data warehouse. This included patient and provider information, administrative events, clinical observations, medications, laboratory tests, and charges in a relational database format (see Supplementary Figure S1). A nurse assigns entering ED patients a clinical severity score of 1 (low) to 5 (high) using a modified version of the Emergency Severity Index (ESI) system.[19] A nurse or respiratory therapist often also assigns a locally modified clinical respiratory score (CRS) consisting of six elements with two possible points per element and a maximum score of 12.[20] Patients are grouped into CRS ranges of 0–2, 3–5, 6–8, $\geq$9 for selecting an appropriate clinical pathway. As CRS are not yet routinely obtained, our sample reduced to 1496 patients with at least two CRS, one assessed early, the other late in their ED visit. This allowed us to compute a change in CRS (ΔCRS), providing a numeric, but somewhat subjective, outcome surrogate. Table 1 provides several patient summaries.

We used custom MATLAB code to convert date and time-stamped event data into a graph-based "activity log" (see Supplementary Algorithm 1). Nodes represent activities; a directed edge between nodes denotes a patient-traversed activity dyad (a "careflow"). Edges are weighted by their frequency. Edge direction is visually encoded using small arrows as well as clockwise edge direction.

Design Requirements

Based on discussions with practitioners, and Shneiderman's visual information seeking mantra, we selected an overview-first followed by zoom and filter, and details on demand interaction paradigm,[21] an approach recognized as an effective means to deal with scale and complexity.[22]

To facilitate differentiation between event characteristics and rapid analysis we used both spatial layout and size and color encodings to represent different care process elements. We considered several different layouts (including circle, force-directed), but, based on user feedback and importance in usability, utility, and aesthetics, decided to use a semantic substrate design approach.[23] We grouped nodes into nonoverlapping vertically arranged regions and then sequentially ordered nodes within each region horizontally.

We grouped nodes into nonoverlapping vertically arranged regions and then sequentially ordered nodes within each region horizontally. Within each substrate, horizontal node position represents the sequence or magnitude of activities. In the "CRS" substrate, for instance, we position nodes according to the numerical score, in increasing order from left to right. In the other substrates, we position nodes based on the sequence of events. In the "Administrative" substrate, for instance, nodes are equally spaced representing the semantic sequence of events. In the "Lab Results" and "Medications" substrates, the horizontal node position is determined by in-degrees from and out-degrees toward other substrates. If a node has many incoming flows from other types of activities, it is positioned further to the left. If one has many outgoing flows to other types of activities, it is positioned further to the right. We did these adjustments for visual clarity and reduction of edge overlaps. Layout algorithm details are provided in the Supplementary Appendix.

Since users prefer to rapidly explore and compare care processes, we created a simple, point-and-click interface. Our tool is organized into the three regions. The slide-in left panel filter

Table 1: Descriptive Summary of Patient Population Data ($n = 5784$)

Gender	n (%)
Male	3575 (61.8)
Female	2209 (38.2)
Age	
0–18 months	562 (9.7)
18–36 months	1048 (18.1)
3–6 years	1682 (29.1)
>6 years	2492 (43.1)
Race	
Am. Indian/Alaska Native	23 (0.4)
Asian	176 (3.0)
Black	2344 (40.5)
Nat. Hawaiian/Pac. Isl.	5 (0.1)
Other	496 (8.6)
White	2740 (47.4)
Triage acuity	
ESI 1	3 (0.1)
ESI 2	1516 (26.2)
ESI 3	2913 (50.4)
ESI 4	1283 (22.2)
ESI 5	62 (1.1)
Unknown	7 (0.1)
ED disposition	
Discharge	3995 (69.1)
Admit to ward	1598 (27.6)
Admit to ICU	140 (2.4)
Admit to OR	47 (0.8)
Transfer	4 (0.1)

ICU = intensive care unit; OR = operating room; ESI = emergency severity index; ED = emergency department.

section contains logically organized dynamic filters for rapid exploration of the data.[24] The visualization panel provides care process visualizations and layout manipulation, edge and label visibility, and zooming controls. The slide-in bottom panel provides summarized information on visualization-associated outcome measures with details on demand.

SYSTEM IMPLEMENTATION

We prototyped the tool with Gephi for rapid iterations, feedback, and refinement.[25] We implemented it using a data-driven document (D3) approach.[26] On the backend, a python-based web server contains a portable database containing patient visits and care histories.

RESULTS

Our browser-based visual analytic tool enables decision makers to interactively explore care processes for a patient population or a dynamically filtered subset (see Figure 1), based on patient, provider, and care process characteristics. User interface (UI) details are provided in the Supplementary Appendix.

Figure 1 shows source-colorized processes for patients with an ESI of 2. Other patients' processes are gray and faded for context. This figure shows (from bottom to top using a semantic substrate layout) the care processes patients traversed in four activity groups (administrative processes, CRS scoring, medications, and lab tests) in chronologic order from left to right. Each node is a clinical activity. Its size corresponds to the percentage of patients that had it done. The colored central part represents how often that activity was done for patients with the selected properties (i.e., ESI of 2), the thickness of the gray edge shows how often it was done for other patients. Process metrics, such as total charges, outcome improvements, and disposition, are provided in expandable summary charts below the care process visualization.

As users explore and filter data, they often identify subgroups of particular interest, which they want to examine later or use in comparison to other subgroup(s). We therefore enabled users to save and retrieve filtered datasets (i.e., care process graphs) for subsequent analyses. In Figure 2 patients are constrained to an

Figure 1: Our tool consists of a filter pane, consisting of (**A**) patient (e.g., age, sex, race), (**B**) provider, and (**C**) care process filter characteristics (e.g., visit duration, charge, disposition), and (**D**) a main visualization pane containing a time-ordered semantic substrate representation with zoom + pan, search, label on/off, and edge on/off controls. A collection of analytical tools including care process outcome summary charts (disposition distribution shown here) are also provided in (**E**). Care of patients with an ESI of 2 is colored, while other patients' care activities and processes are grayed out.

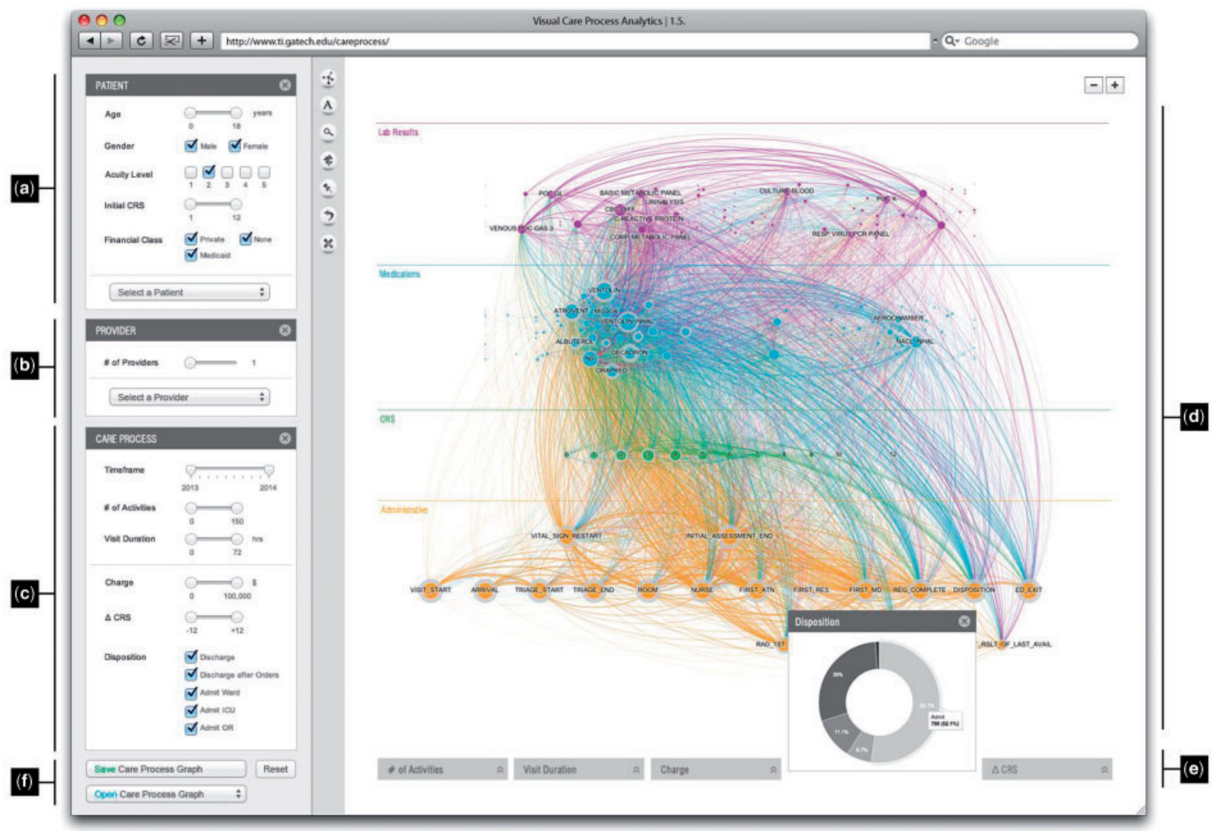
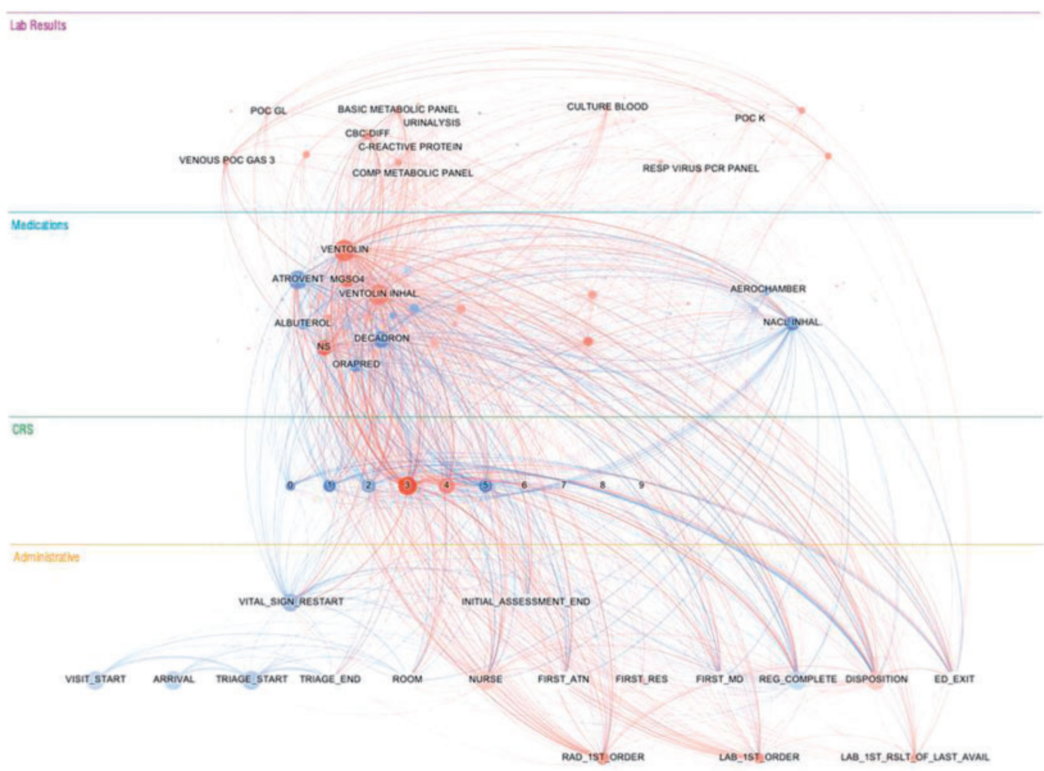


Figure 2: Patients with an initial CRS of 3–5. Reddish/bluish nodes and edges represent the (labeled) clinical activities patients with small/large CRS improvement went through more frequently than the other group, respectively.



BRIEF COMMUNICATION

initial CRS of 3–5 (i.e., the largest CRS subgroup). Users might be interested in comparing subgroups with better "outcomes" (i.e., ΔCRS -2 or less) to those with poorer "outcomes" (i.e., ΔCRS -2 or more) so blue indicates clinical activities for the first subgroup while red indicates those for the second.

While visual inspection reveals process variations between subgroups and summary charts provide insight into the differences, clinical users often prefer to have access to tabular results. Clicking on the analysis icon provides descriptive and statistical details of the filtered patient population. A summary of Figure 2 is provided in Table 2.

Patients with similar CRS (or who are similar in other respects) may have different clinical problems. As a result, in most clinical research, precise cohort identification is needed in order to infer meaningful results. To support this Figure 3 zooms into the medication and lab test regions and illustrates how two "matched" patients with the same gender, age, triage score, entering CRS, and similar total charges can have different care processes because of different clinical problems. The blue color-coded patient presumably has an infection based on their lab work while the red color-code patient may have an isolated asthma exacerbation diagnosed on clinical observations. Hovering over nodes provides detailed activity information including preceding and succeeding activities.

CONCLUDING REMARKS

Visual analytics can play an important role in healthcare process analysis. Our interactive visual approach enables users to gain insight into the complexity of pediatric asthma care processes. We believe it could help with care quality improvement programs, provider comparison and benchmarking, and analysis of conformance to existing care protocols. An extension that "matches" patients based on a selected care process could potentially make cohort identification far more efficient and possibly even more accurate. Similarly, if certain activities are deemed to be "markers" for a specific clinical condition analysis of the care patterns of all patients with those markers might be used to identify clinical care process variations and their relative impacts on outcomes and costs.

FUNDING

This work was supported by Children's Healthcare of Atlanta grant number GT-Q1604.

COMPETING INTERESTS

None.

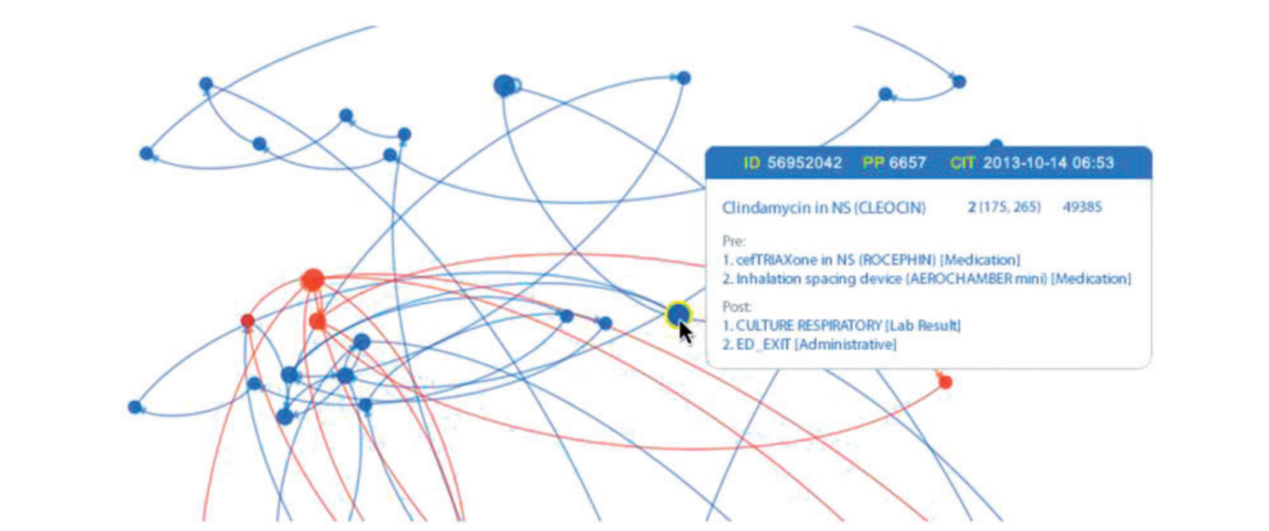
CONTRIBUTORS

Rahul C. Basole (the guarantor) accepts full responsibility for the work and/or the conduct of the study, had access to the

Table 2: Group Comparison of Care Processes

	Group blue (ΔCRS −2 or less)	Group red (ΔCRS −2 or more)
Sample size, *n*	464	441
Age, mean (SD) (months)	76.98 (50.02)	78.84 (47.48)
Gender (%)		
Male	292 (62.9)	279 (63.3)
Female	172 (37.1)	162 (36.7)
Triage acuity		
ESI1	0 (0.0)	2 (0.5)
ESI2	141 (30.4)	182 (41.3)
ESI3	296 (63.8)	241 (54.6)
ESI4	23 (5.0)	16 (3.6)
ESI5	2 (0.4)	0 (0.0)
Initial CRS, mean (SD)	4.00 (0.8)	3.59 (0.71)
ΔCRS, mean (SD)	−2.69 (0.82)	−0.18 (0.98)
Disposition		
Discharge	333 (71.8)	181 (41.0)
Admit to ward	126 (27.2)	236 (53.5)
Admit to ICU	5 (1.1)	24 (5.4)
Charge, mean (SD) ($)	3855 (4423)	7351 (8060)
No. of activities, mean (SD)	23.78 (5.01)	25.34 (6.93)
Visit duration, mean (SD) (h)	238.62 (91.00)	266.86 (101.45)

Figure 3: Two "matched" patients have different diagnoses and very different care processes (red vs. blue) suggesting the potential for finding patient cohorts. Labels are turned off to reduce visual clutter. Hovering over a node provides activity details.



BRIEF COMMUNICATION

data, and controlled the decision to publish. All authors included on the paper fulfill the criteria of authorship.

SUPPLEMENTARY MATERIAL

Supplementary material is available online at http://jamia.oxfordjournals.org/.

REFERENCES

1. Smith M, Saunders R, Stuckhardt L, eds. *Best Care at Lower Cost: the Path to Continuously Learning Health Care in America*. Washington, DC: National Academies Press; 2013.
2. Basole RC, DeMillo RA. Enterprise IT and transformation. In: Rouse WB, ed. *Enterprise Transformation: Understanding and Enabling Fundamental Change*. New York, NY: Wiley; 2006:223–251.
3. Van der Aalst WM. *Discovery, Conformance and Enhancement of Business Processes*. Berlin, Heidelberg: Springer; 2011.
4. Wennberg DE. Variation in the delivery of health care: the stakes are high. *Ann Intern Med.* 1998; 128 (10): 866–868.
5. Noon CE, Hankins CT, Cote MJ. Understanding the impact of variation in the delivery of healthcare services. *J Healthc Manag.* 2002;48(2):82–97.
6. Swensen SJ, Meyer GS, Nelson EC, *et al*. Cottage industry to postindustrial care—the revolution in health care delivery. *New Engl J Med*. 2010;362(5):e12.
7. Porter ME, Teisberg EO. *Redefining Health Care*. Boston, MA: Harvard Business Press; 2006.
8. Agarwal R, Gao G, DesRoches C, *et al*. Research commentary-The digital transformation of healthcare: current status and the road ahead. *Inform Sys Res.* 2010;21(4):796–809.
9. Bigus JP, Campbell M, Carmeli B, *et al*. Information technology for healthcare transformation. *IBM J Res Dev*. 2011; 55(5):1–6:14.
10. Mans RS, van der Aalst WM, Vanwersch RJ, *et al*. Process mining in healthcare: data challenges when answering frequently posed questions. In: Lenz R, Miksch S, Peleg M, eds. *Process Support and Knowledge Representation in Health Care*. Berlin, Heidelberg: Springer; 2013: 140–153.
11. Cao N, Sun J, Lin YR, *et al*. Facetatlas: multifaceted visualization for rich text corpora. *IEEE T Vis Comput Gr.* 2010; 16(6):1172–1181.
12. Perer A, Sun J. Matrixflow: temporal network visual analytics to track symptom evolution during disease progression [conference]. In: *AMIA, November, Chicago, IL*; 2012, p. 716.
13. Zhang Z, Gotz D, Perer A. Iterative cohort analysis and exploration. Inform Vis Published Online First: March 19, 2014. doi:10.1177/1473871614526077.
14. Zhang Z, Wang B, Ahmed F, *et al*. The five W's for information visualization with application to healthcare informatics. *IEEE T Vis Comput Gr* 2013; 19(11):1895–1910.
15. Rind A, Wang T, Aigner W, *et al*. *Interactive information visualization for exploring and querying electronic health records: a systematic review*. Foundations Trends Human-Comput Interact 2013; 5(3):207–298.
16. Perer A, Gotz D. Data-driven exploration of care plans for patients, April, Paris, France. In: *ACM CHI*; 2013, pp. 439–444.
17. CDC FactStats Asthma. http://www.cdc.gov/nchs/fastats/asthma.html. Accessed April 29, 2014.
18. Asthma & Children Fact Sheet. http://www.lung.org/lung-disease/asthma/resources/facts-and-figures/asthma-children-fact-sheet.html. Accessed April 29, 2014.
19. Emergency Severity Index (ESI). http://www.ahrq.gov/professionals/systems/hospital/esi/esihandbk.pdf. Accessed April 29 2014.
20. *Clinical Respiratory Score in Asthma/Bronchiolitis Guidelines: Emergency Department*. Atlanta, GA: Children's Healthcare of Atlanta; 2013:1.
21. Shneiderman B. The eyes have it: a task by data type taxonomy for information visualizations. In: *Proceedings of the IEEE Symposium on Visual Languages, September, Boulder, CO*; 1996, pp. 336–343.
22. Heer J, Shneiderman B. Interactive dynamics for visual analysis. *Queue.* 2012;10(2):30.
23. Aris A, Shneiderman B. Designing semantic substrates for visual network exploration. *Inform Visual.* 2007;6(4):281–300.
24. Keim DA. Information visualization and visual data mining. *IEEE T Vis Comput Gr* 2002;8(1):1–8.
25. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks [conference]. In: *ICWSM, May, San Jose, CA*; 2009, pp.361–362.
26. Bostock M, Ogievetsky V, Heer J. D^3 data-driven documents. *IEEE T Vis Comput Gr* 2011;17(12):2301–2309.

AUTHOR AFFILIATIONS

[1]School of Interactive Computing, Georgia Tech, Atlanta, GA 30332, USA

[2]Tennenbaum Institute, Georgia Tech, Atlanta, GA 30332, USA

[3]School of Computational Science & Engineering, Georgia Tech, Atlanta, GA 30332, USA

[4]School of Industrial & Systems Engineering, Georgia Tech, Atlanta, GA 30332, USA

[5]Children's Healthcare of Atlanta, Atlanta, GA 30322, USA

[6]Pediatric Emergency Medicine Associates, LLC, Atlanta, GA 30342, USA