

Using electronic medical record data to report laboratory adverse events

Tamara P. Miller,^{1,2,4} Yimei Li,^{1,3} Kelly D. Getz,^{1,2} Jesse Dudley,⁴ Evanette Burrows,⁴ Jeffrey Pennington,⁴ Azada Ibrahimova,⁵ Brian T. Fisher,^{6,2,7,3} Rochelle Bagatell,^{1,7} Alix E. Seif,^{1,7} Robert Grundmeier^{4,7} and Richard Aplenc^{1,2,7,3}

¹Division of Oncology, The Children's Hospital of Philadelphia, ²Center for Pediatric Clinical Effectiveness, The Children's Hospital of Philadelphia, ³Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania School of Medicine, ⁴Department of Biomedical and Health Informatics, The Children's Hospital of Philadelphia, ⁵Albert Einstein Medical Center, ⁶Division of Infectious Diseases, The Children's Hospital of Philadelphia and ⁷Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, PA, USA

Received 6 October 2016; accepted for publication 23 November 2016

Correspondence: Tamara P. Miller, Children's Hospital of Philadelphia, 3501 Civic Center Boulevard, CTRB 3022, Philadelphia, PA 19104, USA.
Email: millertp@email.chop.edu

Side effects on oncology clinical trials are captured in adverse event (AE) reports using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE). (https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm) Clinical research associates (CRAs) manually review the medical record to report AEs. Although comprehensive and accurate AE ascertainment is crucial for oncology trials, our prior work has demonstrated dramatic under-reporting of AEs (Miller *et al*, 2016).

We and others have sought to improve AE reporting with automated data extraction from the electronic medical record (EMR) (Mandelblatt *et al*, 2014; Lencioni *et al*, 2015). Mandelblatt *et al* (2014) ascertained AEs in breast cancer patients with an overall sensitivity of 89% (range 0–100%), but did not report positive predictive value (PPV). Lencioni *et al* (2015) incorporated EMR AE capture in one hospital system to streamline AE collection, but did not

Summary

Despite the importance of adverse event (AE) reporting, AEs are under-reported on clinical trials. We hypothesized that electronic medical record (EMR) data can ascertain laboratory-based AEs more accurately than those ascertained manually. EMR data on 12 AEs for patients enrolled on two Children's Oncology Group (COG) trials at one institution were extracted, processed and graded. When compared to gold standard chart data, COG AE report sensitivity and positive predictive values (PPV) were 0–21.1% and 20–100%, respectively. EMR sensitivity and PPV were >98.2% for all AEs. These results demonstrate that EMR-based AE ascertainment and grading substantially improves laboratory AE reporting accuracy.

Keywords: paediatrics, adverse events, electronic medical record, epidemiology, accuracy.

provide a methodology for improving AE reporting on multi-centre trials.

This study sought to automate EMR data extraction and grading for laboratory-based AEs at a single institution and to compare this approach to manual AE assessment. We hypothesized that EMR-based AE ascertainment and grading would have a higher sensitivity and PPV than manual review.

Methods

Cohort and data source

Retrospective EMR data from all patients enrolled on two phase III Children's Oncology Group (COG) trials for acute myeloid leukaemia (AML: AAML0531, AAML1031) between December 2006 and March 2015 at the Children's Hospital of

Philadelphia (CHOP) were included (Gamis *et al*, 2014). COG permitted use of trial data. Prospectively and manually collected COG trial AE data were received from COG in a de-identified spread-sheet listing each reported AE by course. Data were extracted from the vendor-supplied EMR (Epic Systems, Inc., Verona, WI). Laboratory data are stored in the Epic Clarity clinical data warehouse. For each patient, EMR data during the on-protocol period through August 2015 were extracted; this end date was determined from available COG AE reports. Nine chemistry and haematology laboratory results were evaluated: potassium, glucose, sodium, alanine aminotransferase (ALT), aspartate aminotransferase (AST), bilirubin, haemoglobin, platelet count and absolute neutrophil count (ANC). Manual chart abstraction was performed by a paediatrician to determine the AE gold standard.

EMR data extraction

Structured Query Language queries extracted data on nine laboratory tests that correspond to twelve CTCAEs from Clarity into comma-separated values (CSV) files for each patient. Data included: medical record number; the ordered, collected and resulted times and dates; the test component (e.g. "potassium") and name (e.g. "comprehensive metabolic panel"); the results, including units, narrative and comment fields; and the reference range. An iterative process of data extraction and review to verify completeness, and re-extraction ensured all results were extracted.

Processing and grading

Data extracted from different Clarity data tables were concatenated. Reference units were standardized across all types of test for the same laboratory component. Appropriate numerical data in comment fields were extracted programmatically. Each laboratory result was graded using automated code based on cut-off values in CTCAE v4 definitions (https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm) CTCAE definitions were interpreted as follows: non-fasting glucose values were used because fasting glucose is not routinely obtained in paediatric inpatients and, given that Grade 1 "neutrophil count decreased" is defined as a lower limit of normal (LLN) – $1.5 \times 10^9/L$, (https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm), when ANC LLN was absent, an ANC of 2000/mm³ was used. Grade 4 anaemia is not numerically-based and could not be graded (https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm). Finally, total bilirubin was graded.

Electronic data quality filtering was performed after grading to ensure that abnormal results were not due to laboratory error. Filtering removed results with the following error types: (i) haemolysed specimens based on text fields; (ii) tests with multiple laboratory abnormalities, indicating improper specimen collection (e.g. inadequate discard volume prior to

Table 1. Accuracy of clinical trial AE report and EMR-based AE ascertainment compared to chart abstraction data.

	Chart abstraction	COG		EMR		PPV	Specificity	NPV	Sensitivity	PPV	Specificity	NPV
		N	N	N	N							
Sodium	Hypo	36	3	8.3 (2.2, 23.6)	100 (31.0, 100)	100 (96.7, 100)	100 (96.7, 100)	81.1 (74.4, 86.5)	100 (88.0, 100)	100 (88.0, 100)	100 (96.6, 100)	100 (96.6, 100)
	Hyper	2	1	50.0 (2.7, 97.3)	100 (5.5, 100)	99.4 (96.4, 100)	99.4 (96.4, 100)	99.4 (96.4, 100)	100 (19.8, 100)	100 (19.8, 100)	100 (97.2, 100)	100 (97.2, 100)
Potassium	Hypo	57	9	15.8 (7.9, 28.4)	45.0 (23.8, 68.0)	100 (96.2, 100)	100 (96.2, 100)	75.6 (68.1, 81.9)	98.2 (89.4, 99.9)	98.2 (89.4, 99.9)	100 (96.0, 100)	100 (96.0, 100)
	Hyper	57	1	1.8 (0.09, 10.6)	20.0 (1.1, 70.1)	97.2 (93.1, 98.9)	97.2 (93.1, 98.9)	98.8 (95.4, 99.8)	100 (19.8, 100)	100 (19.8, 100)	100 (97.2, 100)	100 (97.2, 100)
Glucose	Hypo	2	0	0 (0, 80.2)	–	100 (97.4, 100)	100 (97.4, 100)	99.4 (96.4, 100)	100 (5.5, 100)	100 (5.5, 100)	100 (97.3, 100)	100 (97.3, 100)
	Hyper	62	13	21.0 (12.1, 33.5)	81.3 (53.7, 95.0)	99.1 (94.6, 100)	99.1 (94.6, 100)	71.2 (63.5, 77.8)	100 (92.7, 100)	98.4 (90.3, 99.9)	99.1 (94.3, 100)	100 (95.8, 100)
Total	Increased	6	0	0 (0, 48.3)	–	100 (97.2, 100)	100 (97.2, 100)	96.6 (92.5, 98.6)	100 (51.7, 100)	100 (51.7, 100)	100 (96.9, 100)	100 (96.9, 100)
Bilirubin	Increased	19	4	21.1 (7.0, 46.1)	66.7 (24.1, 94.0)	100 (97.1, 100)	100 (97.1, 100)	91.9 (86.5, 95.3)	100 (79.1, 100)	100 (79.1, 100)	100 (96.6, 100)	100 (96.6, 100)
ALT	Increased	13	2	15.4 (2.7, 46.3)	50.0 (9.2, 90.8)	99.4 (96.2, 100)	99.4 (96.2, 100)	94.3 (89.4, 97.1)	100 (71.7, 100)	100 (71.7, 100)	100 (96.7, 100)	100 (96.7, 100)
AST	Decreased	170	NR	–	–	–	–	–	99.4 (96.3, 100)	98.3 (94.6, 99.5)	–	–
ANC	Decreased	160	NR	–	–	–	–	–	100 (97.1, 100)	100 (97.1, 100)	100 (69.9, 100)	100 (69.9, 100)
Haemoglobin	Decreased	170	NR	–	–	–	–	–	99.4 (96.3, 100)	98.8 (95.4, 99.8)	87.5 (46.7, 99.3)	100 (56.1, 100)
Platelets	Decreased	170	NR	–	–	–	–	–	–	–	–	–

No hypoglycaemia or blood bilirubin increased were reported in COG data.

NR, COG does not require reporting of Grade 3 and 4 haematology adverse events; ALT, alanine aminotransferase; ANC, absolute neutrophil count; AST, aspartate aminotransferase; COG, Children's Oncology Group; EMR, electronic medical record; NPV, negative predictive value; PPV, positive predictive value.

collection); (iii) results that normalized on repeat testing within 1 h.

Statistical analyses

The highest AE grade in each course was determined. The distribution (frequency, percentage) of results and of courses with each grade was tabulated for each AE at each processing step.

Sensitivity, specificity, PPV and negative predictive value (NPV) were determined for COG AE reports and for EMR-based AE ascertainment using chart data as the gold standard. As the trials did not require haematological AE reporting (Gamis *et al*, 2014), analyses for COG AE reports were restricted to non-haematological AEs. Chart review evaluated COG AE reporting error types. Institutional Review Board approval was obtained.

Results

EMR data on 49 patients (173 chemotherapy courses) were extracted. Three courses were excluded because the patient received chemotherapy at another hospital prior to transfer to CHOP. The number of total tests for each AE ranged from 946 to 6326 (944–6322 after false positive results were removed). As expected, >94% of courses had grade 3 or 4 haematological AEs. The percentage of courses with each chemistry AE ranged from 0% to 20.6% for grade 3 and 0 to 16.5% for grade 4.

The sensitivity of COG AE reports ranged from 0% to 21.1% and PPV ranged from 20% to 100% while the sensitivity and PPV of EMR-based AE ascertainment were >98.2% for all AEs (Table I). The median number of AEs missed per patient per course was one to two. COG AE reporting errors were due to one of three categories: incorrect reporting, reporting of false positive laboratory results, or incorrect grading (Table II).

Discussion

This study demonstrates that EMR-based AE ascertainment at a single institution using an automated data extraction,

cleaning and grading process dramatically improves the accuracy of AE laboratory ascertainment. Notably, the automated processing steps to remove potential laboratory errors, which were not previously reported by Lencioni *et al* (2015) are crucial to ensuring accurate laboratory AE identification.

These results have several direct implications. First, automated laboratory AE ascertainment and grading may be able to replace the current manual ascertainment process. Given the extensive time required for manual ascertainment, the limited CRA time available for AE reporting, and the limited resources available to reimburse manual reporting, such an automated system may be of substantial benefit (Roche *et al*, 2002; Nass *et al*, 2010).

Secondly, such course-level data can provide a more accurate estimate of true laboratory AE rates. Given that COG AE rates were substantially lower than in EMR data, published laboratory AE rates probably substantially underestimate the true laboratory AE rates with AML therapy. Accurate rates could serve as baseline estimates in the evaluation of new agents combined with standard AML chemotherapy backbones.

Finally, automated EMR-based laboratory AE reporting may substantially decrease the variation in clinical trial AE reports (Thomas *et al*, 2002; Huynh-Le *et al*, 2014; Lencioni *et al*, 2015). Many CTCAE definitions use “life-threatening consequences” and a numerical cut-off (https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm). In our experience, some CRAs report laboratory AEs only when accompanied by life-threatening consequences, while others report solely based on numerical values. EMR-based AE ascertainment can standardize grading by enforcing a uniform interpretation of CTCAE numerical criteria. Furthermore, the processing code could transparently be adjusted to a desired sensitivity and specificity for a particular clinical trial. Moreover, multiple thresholds could be applied to evaluate a potential laboratory data toxicity signal. Additional processing incorporating other EMR components, such as medication interventions (i.e. kayexalate), may also improve AE ascertainment.

The primary limitation of this study is that it was performed at one hospital. Future work needs to demonstrate

Table II. Types of adverse event (AE) reporting errors.

Area of error	Types of error
Incorrect reporting	Reported the wrong AE (e.g. hypokalaemia instead of hyperkalaemia) Laboratory data was not identified to match the AE report
Reporting of false positive laboratory results	Reported AE was due to haemolysed potassium specimen Reported hyperkalaemia or hyperglycaemia AE had concurrent hyperglycaemia or hyperkalaemia, indicating improper specimen collection Reported AE normalized when re-checked within 1 h
Incorrect grading	Reported the first date of any grade 3 or 4 AE in the course rather than highest grade for the course Incorrect grading according to numerical cut-offs Grading definition includes calculations compared to reference range and was not correctly graded

that EMR-based laboratory AE ascertainment is generalizable across institutions using Epic and other EMRs. This should be feasible based on the work of other investigators (Deakyne *et al*, 2015; Fiks *et al*, 2015). Furthermore, this study did not incorporate non-numerical CTCAE definitions. However, automated extraction of medical intervention data should be possible. Finally, this study did not include laboratory results reported as free text that would require natural language processing (NLP) for extraction. Querying free text using NLP has substantial challenges, however we believe that processing AE results with NLP should be feasible (Friedlin *et al*, 2008; DeLisle *et al*, 2010).

In summary, this single institution study shows that EMR-based AE ascertainment and grading substantially improves laboratory AE reporting. If incorporated into prospective clinical trials, EMR-based laboratory AE ascertainment may substantially improve laboratory AE reporting efficiency. This may free CRAs to ascertain complex AEs, which may improve the accuracy of complex AE reporting as well. Work is ongoing to test this EMR-based AE ascertainment at other institutions.

References

- Deakyne, S.J., Bajaj, L., Hoffman, J., Alessandrini, E., Ballard, D.W., Norris, R., Tzimenatos, L., Swietlik, M., Tham, E., Grundmeier, R.W., Kuppermann, N., & Dayan, P.S.; Pediatric Emergency Care Applied Research, N. (2015) Development, evaluation and implementation of chief complaint groupings to activate data collection: a multi-center study of clinical decision support for children with head trauma. *Applied Clinical Informatics*, **6**, 521–535.
- DeLisle, S., South, B., Anthony, J.A., Kalp, E., Gundlapalli, A., Curriero, F.C., Glass, G.E., Samore, M. & Perl, T.M. (2010) Combining free text and structured electronic medical record entries to detect acute respiratory infections. *PLoS One*, **5**, e13377.
- Fiks, A.G., Grundmeier, R.W., Steffes, J., Adams, W.G., Kaelber, D.C., Pace, W.D. & Wasserman, R.C.; Comparative Effectiveness Research Through Collaborative Electronic Reporting, C. (2015) Comparative effectiveness research through a collaborative electronic reporting consortium. *Pediatrics*, **136**, e215–e224.
- Friedlin, J., Grannis, S. & Overhage, J.M. (2008) Using natural language processing to improve accuracy of automated notifiable disease reporting. *AMIA Annual Symposium Proceedings*, **2008**, 207–211.
- Gamis, A.S., Alonzo, T.A., Meshinchi, S., Sung, L., Gerbing, R.B., Raimondi, S.C., Hirsch, B.A., Kahwash, S.B., Heerema-McKenney, A., Winter, L., Glick, K., Davies, S.M., Byron, P., Smith, F.O. & Aplenc, R. (2014) Gemtuzumab ozogamicin in children and adolescents with de novo acute myeloid leukemia improves event-free survival by reducing relapse risk: results from the randomized phase III Children's Oncology Group trial AAML0531. *Journal of Clinical Oncology*, **32**, 3021–3032.
- Huynh-Le, M.P., Zhang, Z., Tran, P.T., DeWeese, T.L. & Song, D.Y. (2014) Low interrater reliability in grading of rectal bleeding using National Cancer Institute Common Toxicity Criteria and Radiation Therapy Oncology Group Toxicity scales: a survey of radiation oncologists. *International Journal of Radiation Oncology Biology Physics*, **90**, 1076–1082.
- Lencioni, A., Hutchins, L., Annis, S., Chen, W., Ermisoglu, E., Feng, Z., Mack, K., Simpson, K., Lane, C. & Topaloglu, U. (2015) An adverse event capture and management system for cancer studies. *BMC Bioinformatics*, **16**, S6.
- Mandelblatt, J.S., Huang, K., Makgoeng, S.B., Luta, G., Song, J.X., Tallarico, M., Roh, J.M., Munneke, J.R., Houlston, C.A., McGuckin, M.E., Cai, L., Clarke Hillyer, G., Hershman, D.L., Neugut, A.I., Isaacs, C. & Kushi, L. (2014) Preliminary development and evaluation of an algorithm to identify breast cancer chemotherapy toxicities using electronic medical records and administrative data. *Journal of Oncology Practice*, **11**, e1–e8.
- Miller, T.P., Li, Y., Kavcic, M., Troxel, A.B., Huang, Y.S., Sung, L., Alonzo, T.A., Gerbing, R., Hall, M., Daves, M.H., Horton, T.M., Pulsipher, M.A., Pollard, J.A., Bagatell, R., Seif, A.E., Fisher, B.T., Luger, S., Gamis, A.S., Adamson, P.C. & Aplenc, R. (2016) Accuracy of adverse event ascertainment in clinical trials for pediatric acute myeloid leukemia. *Journal of Clinical Oncology*, **34**, 1537–1543.
- Nass, S.J., Moses, H.L. & Mendelsohn, J. (2010) A National Cancer Clinical Trials System for the 21st Century: Reinvigorating the NCI Cooperative Group Program. The National Academies Press, Washington, DC.
- Roche, K., Paul, N., Smuck, B., Whitehead, M., Zee, B., Pater, J., Hiatt, M.-A. & Walker, H. (2002) Factors affecting workload of cancer clinical trials: results of a multicenter study of the National Cancer Institute of Canada clinical trials group. *Journal of Clinical Oncology*, **20**, 545–556.
- Thomas, E.J., Lipsitz, S.R., Studdert, D.M. & Brennan, T.A. (2002) The reliability of medical record review for estimating adverse event rates. *Annals of Internal Medicine*, **136**, 812–816.

Acknowledgements

This work was supported by the National Institutes of Health R01 CA165277, National Institutes of Health K12 CA076931, Statistics and Data Center Grant U10CA098413, Chair's Grant U10CA098543, National Clinical Trials Network (NCTN) Operations Center Grant U10CA180886, NCTN Statistics & Data Center U10CA180899, and the St. Baldrick's Foundation.

Author contributions

TPM, KDG, BTF, RB, AES, RG and RA designed the research study. TPM, JD, EB, AZ, YL and KDG performed the research. TPM, YL and KDG analyzed the data. JP and RG contributed essential research methods. TPM and RA wrote the paper.

Conflicts of Interest

The authors have no conflicts of interest to disclose.