
Research and Applications

Implementation of a scalable, web-based, automated clinical decision support risk-prediction tool for chronic kidney disease using C-CDA and application programming interfaces

Lipika Samal,^{1,2} John D D'Amore,^{3,4} David W Bates,^{1,2,5} and Adam Wright^{1,2}

¹Division of General Internal Medicine and Primary Care, Brigham and Women's Hospital, Boston, MA, USA, ²Harvard Medical School, Boston, MA, USA, ³Diameter Health Inc, Newton, MA, USA, ⁴Boston University, Boston, MA, USA and ⁵Partners Health-Care System, Boston, MA, USA

Corresponding Author: Lipika Samal, 1620 Tremont St, Boston, MA 02120, USA. E-mail: lsamal@bwh.harvard.edu. Phone: 617-732-7812

Received 3 April 2017; Revised 25 May 2017; Accepted 19 June 2017

ABSTRACT

Background and Objective: Clinical decision support tools for risk prediction are readily available, but typically require workflow interruptions and manual data entry so are rarely used. Due to new data interoperability standards for electronic health records (EHRs), other options are available. As a clinical case study, we sought to build a scalable, web-based system that would automate calculation of kidney failure risk and display clinical decision support to users in primary care practices.

Materials and Methods: We developed a single-page application, web server, database, and application programming interface to calculate and display kidney failure risk. Data were extracted from the EHR using the Consolidated Clinical Document Architecture interoperability standard for Continuity of Care Documents (CCDs). EHR users were presented with a noninterruptive alert on the patient's summary screen and a hyperlink to details and recommendations provided through a web application. Clinic schedules and CCDs were retrieved using existing application programming interfaces to the EHR, and we provided a clinical decision support hyperlink to the EHR as a service.

Results: We debugged a series of terminology and technical issues. The application was validated with data from 255 patients and subsequently deployed to 10 primary care clinics where, over the course of 1 year, 569 533 CCD documents were processed.

Conclusions: We validated the use of interoperable documents and open-source components to develop a low-cost tool for automated clinical decision support. Since Consolidated Clinical Document Architecture-based data extraction extends to any certified EHR, this demonstrates a successful modular approach to clinical decision support.

Key words: interoperability, chronic disease, HL7 consolidated clinical document architecture, application programming interface, clinical decision support

BACKGROUND AND OBJECTIVES

Clinical decision support (CDS) “provides clinicians with knowledge and person-specific information, intelligently filtered and presented at appropriate times, to enhance the quality and efficiency of health care.”¹ Patient-specific risk predictions are a form of CDS that improve care by helping physicians to organize and synthesize various clinical data into a single metric.² Many popular risk-prediction models, such as the Framingham Cardiovascular Risk Score, the Atherosclerotic Cardiovascular Disease Risk Score, and the Gail Breast Cancer Risk Score, have been made publicly available through websites, spreadsheets, and smartphone apps.^{3–5} These methods, however, require manual data entry and workflow interruptions. Ambulatory practices also present many challenges with respect to information technology adoption, including minimal budget for infrastructure and service.⁶ Sixty percent of ambulatory practices have <5 physicians and limited personnel available to develop CDS not provided by the electronic health record (EHR).⁷

Given the widespread adoption of EHRs, CDS could alternatively leverage structured clinical data, such as laboratory results, to enable the automation of risk-prediction models. Automation could increase model use, save clinician time, and reduce data entry errors while protecting patient privacy. We sought to develop a modular application to extract data, perform calculations, and display CDS recommendations in a manner that would address real-world constraints of ambulatory practices.⁸ In this study, we developed a stand-alone CDS application for kidney failure risk prediction as a case study of CDS using Health Level 7 (HL7) interoperability standards.^{9,10}

Clinical case study: kidney failure risk prediction

The majority of the 26 million Americans with chronic kidney disease (CKD) are managed in primary care practices where early recognition of progression to kidney failure could significantly impact cost, morbidity, and mortality.^{11,12} It is difficult for primary care physicians (PCPs) to recognize kidney disease progression, since deterioration can be nonlinear and rapid.^{13,14} Therefore, we identified a risk-prediction model with the hypothesis that risk estimates could greatly enhance PCPs’ ability to recognize kidney failure progression. After the technology was developed, we sought to validate the performance of our risk-prediction tool against risk estimates produced by a tool accompanying the original article.¹⁵

METHODS

Requirements and technology selection

The development of the application was divided into 3 stages. The first stage was the development of risk-prediction calculation and a user interface, which was developed in a single-page application (SPA). The second stage required a server to host the SPA and provide long-term data storage. The last stage included EHR integration using application programming interfaces (APIs). (Further details about Stages 1–3 and how we addressed missing and out-of-range data are provided in Supplementary Appendix S1.)

Stage 1: Single-page application development

We identified an 8-variable risk-prediction model incorporating age, sex, estimated glomerular filtration rate (eGFR), urine albumin-to-creatinine ratio, serum albumin, serum bicarbonate, serum calcium, and serum phosphorus to estimate a 5-year risk of progression to kidney failure.¹⁵ A similar version of this model has been validated

in 31 cohorts internationally.¹⁶ We also calculated estimated eGFR with the more recently validated Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation rather than relying on older Modification of Diet in Renal Disease (MDRD) equations reported by our laboratory.^{17,18} The application was designed to diagnose CKD stage according to guidelines incorporating eGFR and albuminuria. If the risk was >10%, the application recommended a nephrology referral. If the risk was >1.5% and any laboratory results were missing, the application recommended orders for any missing laboratory results.

We wanted to ensure that the application could be used by PCPs across multiple practices, operating systems, and devices, so we followed the example set by other projects that have demonstrated the popularity, scalability, flexibility, and security of web-based applications.¹⁹ In addition, we achieved low-cost implementation through the use of open-source libraries.²⁰

We built upon our prior experience with the HL7 Version 2 Messaging and Clinical Document Architecture standards.^{25–28} Since generation of HL7 Consolidated Clinical Document Architecture (C-CDA) documents is a requirement for certified EHRs, providers incur no incremental cost to export data.²⁹ The Continuity of Care Documents (CCDs) that are created according to this standard export structured data, such as laboratory results, encoded with Logical Observation Identifiers Names and Codes (LOINC).

Knowing that web browsers would host our user interface, we identified 2 options widely used in online risk prediction tools: a static HyperText Markup Language (HTML) page, where the risk calculation is performed by a server-based application (Figure 1A), and a JavaScript web application, where the risk is calculated directly in the web browser (Figures 1B and C). This second method uses JavaScript to update without refreshing the HTML page content so that users can change values in the web browser and immediately recalculate risk. This approach allowed rapid testing and prototyping without a hosted environment. We developed a SPA that was able to parse a CCD, extract the data necessary for the 8-variable model, and return CDS recommendations (Figure 2).

First, we validated the application using CCDs from vendors that were available in a public repository (https://github.com/chb/sample_ccdas). We established cross-EHR compatibility by testing fictional samples from multiple EHRs: Allscripts, Cerner, Greenway, Partners HealthCare, and Vitera. Second, we requested the creation of 5 test patients with laboratory results indicating various levels of risk from Partners HealthCare. The demographics and laboratory values were added to these additional test patients in the same manner as if they were generated for real patients. Then, CCDs were generated for these test patients in a non-production environment. We then loaded these CCDs into the SPA and confirmed that the calculated risk was accurate.

Third, after receiving approval from the Partners HealthCare Institutional Review Board, we performed additional testing using documents from a cohort of real patients with stage 3, 4, or 5 CKD assembled for previous studies of CKD care.^{21,22} We manually verified demographic data and serum and urine laboratory results in the patients’ CCDs by comparing them with data in the EHR. We then validated our risk estimates using a spreadsheet calculator supplied with the original article.¹⁵

Stage 2: Server-based application

The next stage was to develop the server programming to automatically ingest CCDs and return a prepopulated version of the SPA. To

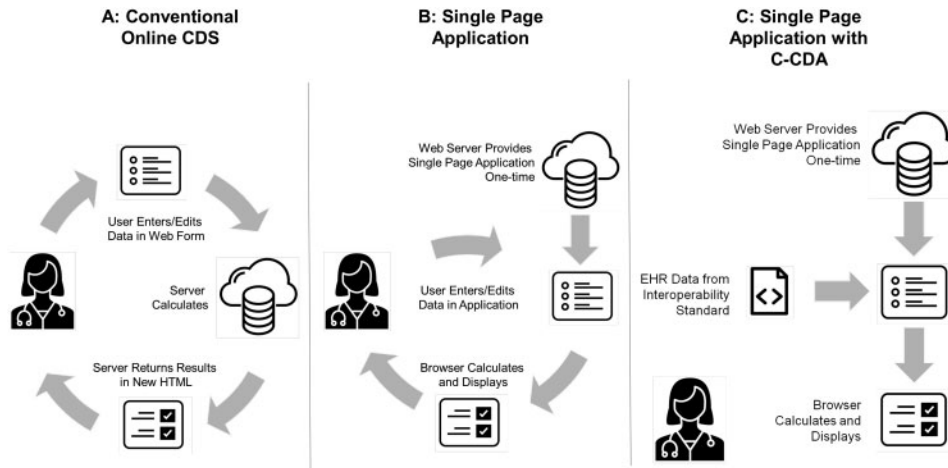


Figure 1. Methods for delivering online clinical decision support.

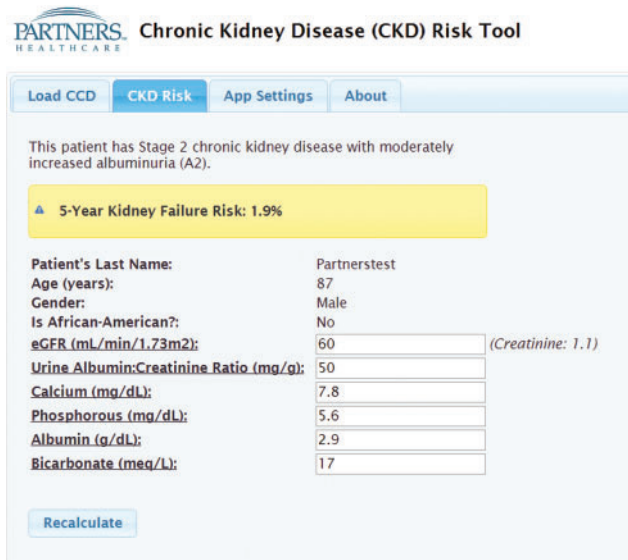


Figure 2. Chronic kidney disease risk tool application.

maintain the security and privacy of health records, the application required the capability of limiting user access and maintaining an activity log. We selected Node.js as the server platform so that we could reuse JavaScript code on both the browser and the server with minimal editing.²³ This period spanned from March to July 2015, during which time the application was again tested for browser compatibility.

Stage 3: EHR integration via API

The final stage was to set up the services to request CCDs automatically and return a prepopulated version of the SPA to the EHR via API. Integration with the Partners HealthCare homegrown EHR (Longitudinal Medical Record) required the use of 3 APIs.

Documents were requested using a CCD Data Portability specification, which had been implemented by Partners to provide 5 years' worth of prior laboratory data. The resulting data were parsed, risk was calculated, and the estimate was stored in a MongoDB database.

The final step was to return a CDS message regarding a patient's risk, which was shown on the clinical summary tab of Partners'

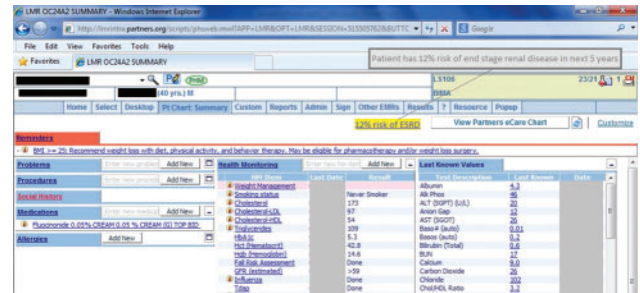


Figure 3. Clinical decision support link and rollover text on LMR patient summary page.

EHR (Figure 3). If users rolled over the blue underlined text, they would see the longer version of the CDS message (eg, “12% risk of end stage renal disease in the next 5 years”).

We requested schedules and CCDs for patients before production deployment to ensure that the application accurately diagnosed patients with stage 3, 4, or 5 CKD. Finally, we coordinated how long hyperlinks would remain present in the EHR, which was determined to be 8 days after appointment date. This period spanned August to November 2015.

RESULTS

Single-page application validation results

As described in the Methods section, we performed several validation steps to ensure that our SPA worked as intended. We processed CCDs for 115 known CKD patients. Several terminology issues were reconciled. Specifically, certain laboratory tests had multiple LOINC codes that could be used. For example, we needed to disambiguate the LOINC code for serum bicarbonate. Though we initially used the LOINC code 1963-8 (Bicarbonate in Serum) based on the name “serum bicarbonate” used in the original article and in common parlance among physicians, we determined that Partners HealthCare actually reports 2028-9 (Carbon Dioxide in Serum) as part of the basic metabolic panel. In addition, we found that not all laboratory results returned numeric results in all circumstances. For urine albumin-to-creatinine ratio, a normal result would be reported by the text string “below assay,” which required additional logic for

us to incorporate into risk calculation. We validated the performance of the application including imputed values (Supplementary Appendix A1; Table 1).

Server-based application validation results

During end-to-end testing, we tested the application's performance for an additional 58 CKD patients in a non-production environment. These patients were selected from the schedules of the 10 primary care clinics by the application. We validated that the application performed eGFR and risk estimate calculations correctly. The links to the EHR were functioning. The clinical recommendations were populating as expected. However, during the initial service implementation, we had to debug several issues with the generation and receipt of CCDs. Specifically, using APIs requires precise management of namespace issues to function reliably. We also discovered that some patients could generate >1 clinical document if they had multiple appointments, and that new patients might not have an enterprise master patient index number, so the application would not be able to request a CCD for them. Steps within the file processing were added to accommodate these scenarios.

The final validation included 82 patients scheduled at 1 primary care clinic, some of whom did not have CKD. Out of those 82 scheduled patients, our application correctly diagnosed the 14 patients who had stage 3, 4, or 5 CKD. The application performed eGFR and risk estimate calculations correctly.

This final phase of testing presented several challenges that we addressed before go-live. This included recognizing that many users would be using Internet Explorer 8. HTML elements were added to create backwards compatibility before the application would function with this browser version. Additionally, database indexing was performed based on query parameters to speed website generation.

The system was placed into production utilization in support of a clinical trial on December 3, 2015. System logs were monitored and no unexpected downtime was experienced. The system requested and processed 569 533 documents during the period between December 4, 2015, and December 3, 2016. During this time frame, 41 842 (7.3%) of processed CCDs diagnosed patients with stage 3, 4, or 5 CKD.

DISCUSSION

We developed and implemented an application to automatically calculate 5-year risk of kidney failure for primary care patients. We determined that it is feasible to use C-CDA documents in a stand-alone CDS application. Moreover, this application was developed without a data warehouse and can be adapted to any certified EHR with minor incremental cost. These advantages are meaningful for cost-conscious ambulatory care providers, such as PCPs, who desire CDS and analytics functionality beyond what their EHR can provide.

We could have created the same tool within an EHR, because many commercial EHRs represent the data necessary for this risk-prediction model in a structured, coded manner. There are several advantages of an externally hosted CDS application like this one, such as the opportunity to deploy a single application across multiple EHRs. Also, knowledge managers have more flexibility to update recommendations and references within a web-based application versus a commercial EHR.

Previous research has demonstrated that substantial heterogeneity and complexity are present within C-CDA documents, and observations from this research study related to LOINC codes substantiate those findings.²⁴ These issues, however, were surmountable. While we do not expect that small practices would pursue application development independently, this research demonstrates that modern development approaches leveraging open-source technology and interoperability standards can provide a disruptive force to this sector.⁸

LIMITATIONS

While our application delivered a reliable service to diagnose a condition and calculate risk across 10 primary care practices and thousands of patients, our findings are limited in several ways. First, our development, design, and findings were specific to CKD. Second, our asynchronous approach to requesting clinical information the night before a patient visit may not support scenarios for clinical decision support that require data collection or immediate clinician response during the office visit. Third, validation of the server-based application was limited to the Partners HealthCare environment. Lastly, the CKD risk-prediction model relied primarily on structured data (demographics and lab results). This risk-prediction model may be less accurate than one that incorporates data that are only recorded in free-text notes, such as elements of the social history (eg, "Patient has started smoking again."). Natural language processing would be one way to enrich the free-text data so that it can be used for risk prediction.

CONCLUSION AND IMPLICATIONS

This research study is the first to establish the feasibility of low-cost stand-alone CDS using modern web technology and C-CDA documents. Future advances in standards and technology are likely to further reduce cost and challenges to implementation.

ACKNOWLEDGMENTS

We would like to acknowledge Joy Gulla for her role as research assistant, Katherine Zigmont, RN, for data validation, Julie Fiskio for cache database queries, and Beatriz Rocha, MD, PhD, for manual creation of CCDs plus technical support for assistance with the Partners HealthCare CCD Factory Service.

FUNDING

Research reported in this publication was supported by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health under award no. K23DK097187. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

ETHICS

This study was reviewed and approved by the Partners HealthCare Human Research Committee.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

REFERENCES

1. Osheroff JA, Teich JM, Middleton B, Steen EB, Wright A, Detmer DE. A roadmap for national action on clinical decision support. *J Am Med Inform Assoc.* 2007;14(2):141–45.
2. Feblowitz JC, Wright A, Singh H, Samal L, Sittig DF. Summarization of clinical information: a conceptual model. *J Biomed Inform.* 2011;44(4):688–99.
3. Cardiovascular Disease (10-year risk). <https://www.framinghamheartstudy.org/risk-functions/cardiovascular-disease/10-year-risk.php>. Accessed July 6, 2017.
4. Breast Cancer Risk Assessment Tool. <https://www.cancer.gov/bcrisktool/>. Accessed July 6, 2017.
5. Goff DC Jr, Lloyd-Jones DM, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation.* 2014;129 (25 Suppl 2):S49–73.
6. MGM Association. *MGMA Cost Survey: 2014 Report Based on 2013 Data – Key Findings Summary Report*. Washington, DC: Medical Group Management Association; 2014.
7. Hsiao CJ, Hing E, Ashman J. Trends in electronic health record system use among office-based physicians: United States, 2007–2012. *Natl Health Stat Report.* 2014;75:1–18.
8. Mandl KD, Kohane IS. No small change for the health information economy. *N Engl J Med.* 2009;360(13):1278–81.
9. Goldberg HS, Paterno MD, Rocha BH, et al. A highly scalable, interoperable clinical decision support service. *J Am Med Inform Assoc.* 2014;21(e1):e55–62.
10. Chipman J, Drohan B, Blackford A, Parmigiani G, Hughes K, Bosinoff P. Providing access to risk prediction tools via the HL7 XML-formatted risk web service. *Breast Cancer Res Treat.* 2013;140(1):187–93.
11. Coresh J, Selvin E, Stevens LA, et al. Prevalence of chronic kidney disease in the United States. *JAMA.* 2007;298(17):2038–47.
12. Go AS, Chertow GM, Fan D, McCulloch CE, Hsu CY. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. *N Engl J Med.* 2004;351(13):1296–305.
13. Li L, Astor BC, Lewis J, et al. Longitudinal progression trajectory of GFR among patients with CKD. *Am J Kidney Dis.* 2012;59(4):504–12.
14. O'Hare AM, Batten A, Burrows NR, et al. Trajectories of kidney function decline in the 2 years before initiation of long-term dialysis. *Am J Kidney Dis.* 2012;59(4):513–22.
15. Tangri N, Stevens LA, Griffith J, et al. A predictive model for progression of chronic kidney disease to kidney failure. *JAMA.* 2011;305(15):1553–59.
16. Tangri N, Grams ME, Levey AS, et al. Multinational Assessment of Accuracy of Equations for Predicting Risk of Kidney Failure: A Meta-analysis. *JAMA.* 2016;315(2):164–74.
17. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med.* 2009;150(9):604–12.
18. Levey AS, Stevens LA. Estimating GFR using the CKD Epidemiology Collaboration (CKD-EPI) creatinine equation: more accurate GFR estimates, lower CKD prevalence estimates, and better risk predictions. *Am J Kidney Dis.* 2010;55(4):622–27.
19. Mandl KD, Mandel JC, Murphy SN, et al. The SMART Platform: early experience enabling substitutable applications for electronic health records. *J Am Med Inform Assoc.* 2012;19(4):597–603.
20. Suner A, Karakulah G, Dicle O, Sokmen S, Celikoglu CC. CorRECTreatment: a web-based decision support tool for rectal cancer treatment that uses the analytic hierarchy process and decision tree. *Appl Clin Inform.* 2015;6(1):56–74.
21. Samal L, Linder JA, Bates DW, Wright A. Electronic problem list documentation of chronic kidney disease and quality of care. *BMC Nephrol.* 2014;15:70.
22. Samal L, Wright A, Waikar SS, Linder JA. Nephrology co-management versus primary care solo management for early chronic kidney disease: a retrospective cross-sectional analysis. *BMC Nephrol.* 2015;16:162.
23. Morrison JJ, Hostetter J, Wang K, Siegel EL. Data-driven decision support for radiologists: re-using the National Lung Screening Trial dataset for pulmonary nodule management. *J Digital Imaging.* 2015;28(1):18–23.
24. D'Amore JD, Mandel JC, Kreda DA, et al. Are Meaningful Use Stage 2 certified EHRs ready for interoperability? Findings from the SMART C-CDA Collaborative. *J Am Med Inform Assoc.* 2014;21(6):1060–8.
25. Tahir D. Providers are sharing more data than ever. So why is everyone so unhappy? *Mod Healthc.* 2015. <http://www.modernhealthcare.com/article/20150422/NEWS/150429965>. Accessed July 6, 2017.
26. D'Amore JD, Sittig DF, Ness RB. How the continuity of care document can advance medical research and public health. *Am J Public Health.* 2012;102(5):e1–4.
27. Klann JG, Mendis M, Phillips LC, et al. Taking advantage of continuity of care documents to populate a research repository. *J Am Med Inform Assoc.* 2015;22(2):370–79.
28. CDA® Release 2. http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7. Accessed July 6, 2017.
29. HL7 Implementation Guide for CDA® Release 2: IHE Health Story Consolidation, Release 1.1 - US Realm. http://www.hl7.org/implement/standards/product_brief.cfm?product_id=258. Accessed July 6, 2017.