

Registry-based Diabetes Risk Detection Schema for the Systematic Identification of Patients at Risk for Diabetes in West Virginia Primary Care Centers

by Adam Baus, MA, MPH; Gina Wood, RD, LD; Cecil Pollard, MA; Belinda Summerfield, RN; and Emma White, RN

Abstract

Approximately 466,000 West Virginians, or about 25 percent of the state population, have prediabetes and are at high risk for developing type 2 diabetes. Appropriate lifestyle intervention can prevent or delay the onset of type 2 diabetes if individuals at risk are identified and treated early. The West Virginia Diabetes Prevention and Control Program and the West Virginia University Office of Health Services Research are developing a systematic approach to diabetes prevention within primary care. This study aims to demonstrate the viability of patient registry software for the analysis of disparate electronic health record (EHR) data sets and standardized identification of at-risk patients for early detection and intervention. Preliminary analysis revealed that of 94,283 patients without a documented diagnosis of diabetes or prediabetes, 10,673 (11.3 percent) meet one or more of the risk criteria. This study indicates that EHR data can be repurposed into an actionable registry for prevention. This model supports meaningful use of EHRs, the Patient-Centered Medical Home program, and improved care through enhanced data management.

Keywords: electronic health records; diabetes; West Virginia; registry; patient registry

Introduction

Traditionally, identifying persons at risk for diabetes occurs in venues such as community-level screenings and routine primary care appointments. However, the availability of electronic health record (EHR) data in primary care affords an opportunity for the use of innovative methods of chronic disease screening and prevention. We sought to examine the ability to pair the Chronic Disease Electronic Management System (CDEMS) patient registry software with disparate EHR data sets to standardize identification of at-risk patients. This study aims to demonstrate that registry-based diabetes risk detection queries can identify at-risk patients across care sites and EHR systems for the purposes of early intervention and disease prevention.

Background

Diabetes mellitus is the seventh leading cause of death in the United States and is a major contributor to decreased life expectancy, increased rates of heart disease, stroke, kidney failure, lower limb amputations, retinopathy, and adult-onset blindness.^{1,2} The prevalence of diabetes is particularly high in

West Virginia, which ranks third highest nationwide with 12 percent prevalence.^{3,4} Furthermore, approximately 466,000 West Virginians, or about 25 percent of the state population, have prediabetes and are at high risk for developing type 2 diabetes.⁵ Evidence-based lifestyle intervention such as the National Diabetes Prevention Program (National DPP) can prevent or delay the onset of diabetes if those at risk for the disease are identified and treated early.⁶⁻⁸ Community-level diabetes screenings, such as those that take place at health fairs, are often time consuming and cost prohibitive and may fail to link individuals to care.⁹ The Centers for Disease Control and Prevention supports opportunistic screening in primary care during routine appointments.¹⁰ However, screenings must be sensitive to time constraints in primary care and ideally make use of available patient data.

Given the prevalence of diabetes and the challenges associated with screening, this study examines the utility of analyzing disparate EHR data sets using the CDEMS registry software. The intent is to systematically identify at-risk patients in a way that is sensitive to the time constraints in primary care and makes use of available patient data. Pairing registry functionality with routinely collected EHR data supports quality improvement by creating actionable information to systematically identify at-risk patients across care sites and EHR systems for early intervention, bolstering use of clinical decision support tools, increasing opportunities for patient-provider engagement, and strengthening the foundation for referrals from providers to National DPP sites.^{11, 12}

The analysis is based on de-identified EHR data from 14 West Virginia primary care organizations partnering with the West Virginia Diabetes Prevention and Control Program (DPCP) and the West Virginia University Office of Health Services Research (OHSR). This partnership between primary care, public health, and academia has been shown to be effective in helping to improve diabetes outcomes.^{13, 14} This study aims to demonstrate that registry-based diabetes risk detection queries can be used to identify at-risk patients across sites and EHRs through analysis of patient diagnoses, age, body mass index, and laboratory results.¹⁵⁻²⁰

Methods

This study is a nonexperimental retrospective analysis of established patients without a documented diagnosis of diabetes or prediabetes at 14 West Virginia primary care organizations comprising 23 care sites. Established patients are identified as those receiving care for 12 months or more. These primary care organizations use two different EHR systems; 11 use one EHR system, and 3 use another system. De-identified EHR data sets are shared by these primary care organizations with the OHSR and the DPCP on a quarterly basis. Memorandums of understanding between each organization, the OHSR, and the DPCP outline the de-identified data sharing process for the purpose of improving patient care and evaluation. The de-identified EHR data sets are imported to the CDEMS software to create a common, multisite registry. CDEMS is Microsoft Access-based public-domain patient registry software modified by OHSR for use in chronic disease identification and tracking. Importing the de-identified data to CDEMS allows the data to be checked for consistency and accuracy, while increasing the capacity to use the data for patient tracking, reporting, data sharing, and improving the quality of care. CDEMS has been effective in previous research analyzing EHR data across multiple primary care centers for consistency and accuracy in diagnostic coding.²¹ Table 1 lists the EHR data elements imported into the registry software. These data provide information essential for unique patient identification, demographic information, diagnoses, services received, laboratory results, and vital signs.

Using prediabetes screening guidelines from the American Diabetes Association²² and the Centers for Disease Control and Prevention,²³ queries were built in CDEMS to search the de-identified EHR data to do the following:

1. Identify established patients by primary care organization (i.e., those receiving care for 12 months or more).
2. Exclude patients with a documented diagnosis of type 1 diabetes, type 2 diabetes, or prediabetes.

3. Identify patients at risk for diabetes and in need of follow-up screening. Inclusion criteria used to identify these patients are as follows:
 - a) Patients age 45 years or older with last recorded body mass index (BMI) greater than 25.
 - b) Patients age younger than 45 years with last recorded BMI greater than 25, with one or more of the following documented health conditions: hypertension, hyperlipidemia, gestational diabetes, family history of diabetes, or cardiovascular disease.
 - c) Patients with last fasting blood glucose results in the range of 100–125 mg/dL.

Table 2 lists the queries used to identify patients at risk for diabetes and describes the functions of each query.

Results

A total of 130,021 active patients in the 14 primary care organizations were included in this study. Among those patients, 106,367 (81.8 percent) have received care for 12 months or more. Among those established patients, 94,283 patients (88.6 percent) do not have a documented diagnosis of diabetes or prediabetes and are the focus of this analysis.

Overall, this study indicates that the registry software can identify at-risk patients through analysis of diagnoses, patient age, body mass index, and fasting blood glucose results, in accordance with national guidelines. Among the 94,283 established patients without a documented diagnosis of diabetes or prediabetes, 10,673 (11.3 percent) meet one or more of the inclusion criteria for being at risk for diabetes and therefore in need of targeted follow-up and screening. The majority of these patients (70.9 percent) were identified by the search criteria of being age 45 years or older with last recorded BMI greater than 25. The second greatest number of patients (21.1 percent) was identified by the search criteria of having their most recent fasting blood glucose results in the range of 100–125. Lastly, 8.0 percent of at-risk patients were identified by the search criteria of being younger than 45 years with BMI greater than 25 with a diagnosis of hypertension, hyperlipidemia, gestational diabetes, family history of diabetes, or cardiovascular disease. Standard deviations reveal substantial center-level variability in each of the three identifying criteria. Table 3 provides results for these patients overall and by primary care organization.

Discussion

This study examines the use of the CDEMS patient registry software to analyze disparate EHR data for the purpose of systematically identifying patients at-risk for diabetes and in need of targeted screening. The principal finding of this study is that at-risk patients can be identified retrospectively by importing EHR data into CDEMS for analysis, rather than by conducting a patient-by-patient review of medical information to determine risk status. This study finds that 10,673 (11.3 percent) of the 94,283 established patients in 14 primary care organizations without a documented diagnosis of diabetes or prediabetes meet one or more inclusion criteria for diabetes risk (see Table 3). The strongest identifying criteria is being age 45 years or older with last recorded BMI greater than 25. Given these results, it is reasonable to expect that employing this model in primary care centers is a viable option for early detection and intervention without the expenditure of significant resources or need for additional data collection. This method supports diabetes prevention efforts in a way that is sensitive to the time constraints of providers and the care team, and integrates with data-driven quality improvement efforts such as the Patient-Centered Medical Home.^{24, 25}

These results are supported by a recent study by Nichols et al.,²⁶ which found that patients with a documented diagnosis of diabetes could be identified by searching EHR data from multiple primary care organizations. Although that research contrasts with this study in that existing diabetes cases were the focus of analysis, the rationale for using EHR data to identify patients by health condition is consistent.

The current study builds on the findings of Nichols et al. in two ways. First, data are analyzed to identify patients who do not have a diagnosis of prediabetes or diabetes but are at risk for developing those conditions in the absence of intervention. This analysis offers the potential for primary rather than tertiary prevention, affording the opportunity to improve patients' quality of life and reduce the long-term costs of healthcare. Second, the scope of this study extends into both practice-based quality improvement and academic/public health research realms. The registry software is used in the clinical setting to improve population-level care and facilitate meaningful use of EHR data,²⁷ while de-identified versions of these data sets are used for analysis, surveillance, and program planning by academic and public health institutions. While the DPCP and OHSR receive only de-identified versions of the data, each care site has patient-identified registry data and reporting tools to use for quality improvement efforts and reporting. This dual use of data overcomes limitations in previous research, in which reports used by physicians and medical staff would have been more useful had they included patient identifiers.²⁸ Furthermore, the transfer of key clinical data from EHR to registry software advances the use of registries beyond their traditional role of simply tracking the care and outcomes of patients already diagnosed with chronic health conditions. The registry software becomes a tool for prevention by automating and standardizing the search for patients meeting specific risk-factor criteria.

This study supports the combined use of patient registry software with EHR data to identify patients at risk for diabetes, and helps alleviate concerns cited in previous research regarding the lack of a systematic approach in identifying at-risk patients.²⁹ The methods and data tools used in this study allow for identification of at-risk patients across 14 primary care organizations using two different EHRs. These tools can be modified to meet the unique data structures of other EHRs. Furthermore, the standardized methods and registry software for identifying at-risk patients developed in this study support primary care organizations' adopting and successfully implementing the National DPP, which aims to prevent onset of diabetes among at-risk individuals through lifestyle change.³⁰

Results from this study raise additional questions. First, although differences in health center size can be expected to account for some of the variability observed (see Table 3), what is not explained is the variability in numbers of patients identified according to fasting blood glucose results in the range of 100–125. Procedural differences in documentation of fasting blood glucose results may help to explain the variability but will need to be verified through follow-up contact with these organizations. Understanding more about this variability will be important in presenting findings to and implementing practice and policy changes in these organizations. Second, in follow-up analysis it will be advantageous to include patients with a documented diagnosis of prediabetes in the identification schema to provide a more inclusive account of at-risk patients. This is especially important at the health center level, as these methods can aid providers in easily identifying at-risk patients for referral to a National DPP site. In effect, the proportion of at-risk patients identified in this study is a conservative estimate and underscores the need for intervention in this patient population.

Diabetes risk detection and decision support tools discussed here are currently being used as a catalyst to help spur diabetes prevention efforts in West Virginia in a way that links primary care with community organizations. The DPCP and OHSR are partnering with four communities in West Virginia in pilot efforts that employ these tools to (1) identify at-risk patients in primary care centers, (2) have the patient lists reviewed by the care team, (3) screen patients as needed, (4) develop referral mechanisms to help place at-risk individuals in National DPP intervention sites, and (5) ensure that information from the intervention sites is channeled back to the primary care centers. These efforts also involve development of clinical practice and policy for the care of patients who cannot participate in intensive lifestyle change programs such as the National DPP. Furthermore, the DPCP and OHSR are supporting adoption of these tools among all partnering care centers to help transform clinical data into actionable information that can be used for prevention. Integration of practice and policy on diabetes prevention will help sustain these efforts and help increase primary care centers' capacity to work with community resources to address this pressing public health problem.

Follow-up research is warranted on the sensitivity and specificity of the methods developed in this study. Future research comparing primary data collection on at-risk patients by manual review of EHR

records on a patient-by-patient basis to the results of these more automated methods will help to validate these findings.

Conclusion

This study advances the notion of a patient registry from its traditional use in secondary and tertiary prevention to a role in primary prevention. This model supports meaningful use of EHR data, which aims to improve decision support, increase the application of patient data to care, and ultimately improve health outcomes. Likewise, this model supports Patient-Centered Medical Home efforts, which are data driven and focus on measurement of outcomes among at-risk patient populations. The availability of EHR data in primary care affords an opportunity for innovative efforts in chronic disease prevention. However, from a public health perspective, the diversity of EHR systems in use, the potential difficulty of extracting the necessary data, and the overall expense of these systems present challenges in using EHR data for prevention. In this study, we examined the ability to use the CDEMS patient registry software as a means of analyzing disparate EHR data sets to identify of patients at risk for diabetes for the purpose of early detection and intervention. The crux of the current effort is of practical importance to public health: identifying a sector of the patient population potentially unaware of their risk while providing primary care centers with a tool for more efficient screening.

Adam Baus, MA, MPH, is the assistant director of the Office of Health Services Research at the West Virginia University School of Public Health in Morgantown, WV.

Gina Wood, RD, LD, is the manager of the West Virginia Diabetes Prevention and Control Program at the West Virginia Bureau for Public Health in Charleston, WV.

Cecil Pollard, MA, is the director of the Office of Health Services Research at the West Virginia University School of Public Health in Morgantown, WV.

Belinda Summerfield, RN, is the program coordinator for the West Virginia Diabetes Prevention and Control Program at the West Virginia Bureau for Public Health in Charleston, WV.

Emma White, RN, is the healthcare quality improvement coordinator at the Office of Health Services Research at the West Virginia University School of Public Health in Morgantown, WV.

Acknowledgments

This work was made possible through funding from the West Virginia Bureau for Public Health, Office of Community Health Systems and Health Promotion, Division of Health Promotion and Chronic Disease.

Notes

1. US Department of Health and Human Services. "Healthy People 2020." 2013. Available at <http://healthypeople.gov/2020/topicsobjectives2020/overview.aspx?topicid=8>.
2. US Preventive Services Task Force. "Screening for Type 2 Diabetes Mellitus in Adults: Preventive Services Task Force Recommendation Statement." *Annals of Internal Medicine* 148 (2008): 846–54.
3. West Virginia Department of Health and Human Resources Health Statistics Center. "The Burden of Diabetes in West Virginia." 2009. Available at http://www.wvdhhr.org/bph/hsc/pubs/other/burdenofdiabetes2009/diabetes_burden_mort_2010.pdf.
4. New America Foundation. "The State of State Health." 2012. Available at <http://statehealth.newamerica.net>.
5. Institute for Alternative Futures. "Diabetes 2025 Forecasts." 2011. Available at http://www.altfutures.org/pubs/diabetes2025/WEST_VIRGINIA_Diabetes2025_Overall_BriefingPaper_2011.pdf.
6. US Department of Health and Human Services. "Healthy People 2020."
7. Narayan, Venkat K. M., and David F. Williamson. "Prevention of Type 2 Diabetes: Risk Status, Clinic, and Community." *Journal of General Internal Medicine* 25 (2009): 154–57.
8. Shrot, Richard J., Frances M. Sahebzamani, and James H. Brownlee Jr. "Case Study: Screening and Treatment of Pre-Diabetes in Primary Care." *Clinical Diabetes* 22 (2004): 98–100.
9. American Diabetes Association. "Screening for Diabetes." *Diabetes Care* 25 (2002): S21–S24.
10. Centers for Disease Control and Prevention. "Screening for Type 2 Diabetes." *Diabetes Care* 27, no. 1 (2004): 511–14.
11. Hanna, Kathie E., Susan Anderson, and Sara D. Maddox. *Think Research: Using Electronic Medical Records to Bridge Patient Care and Research*. Washington, DC: Center for Accelerating Medical Solutions, 2005.
12. Tuan, Wen-Jan, Ann M. Sheehy, and Maureen A. Smith. "Building a Diabetes Screening Population Data Repository Using Electronic Medical Records." *Journal of Diabetes Science and Technology* 5 (2011): 514–22.
13. Baus, Adam, Emma White, Gina Wood, Belinda Summerfield, and Cecil Pollard. "Primary Care/Public Health Partnership for Improved Type 2 Diabetes Outcomes at Roane County Family Health Care." *Commune Bonum* (Spring 2012): 34–41.
14. Pollard, Cecil, Kelly Bailey, Trisha Petite, Adam Baus, Mary Swim, and Michael Hendryx. "Electronic Patient Registries Improve Diabetes Care and Clinical Outcomes in Rural Community Health Centers." *Journal of Rural Health* 25, no. 1 (2009): 77–84.
15. US Preventive Services Task Force. "Screening for Type 2 Diabetes Mellitus in Adults: Preventive Services Task Force Recommendation Statement."
16. Narayan, Venkat K. M., and David F. Williamson. "Prevention of Type 2 Diabetes: Risk Status, Clinic, and Community."
17. Shrot, Richard J., Frances M. Sahebzamani, and James H. Brownlee Jr. "Case Study: Screening and Treatment of Pre-Diabetes in Primary Care."
18. Spigt, Mark, Annelies Rikkers, Mirte Doornbos, Evelyn Wouters, Isi Spitz, Ludovic Van Amelsvoort, and Paul Zwietering. "The Effect of Screening on the Prevalence of Diagnosed

Type 2 Diabetes in Primary Care.” *Scandinavian Journal of Primary Health Care* 27 (2009): 232–37.

19. Dube S., V. K. Sharma, T. N. Dubey, A. Mehta, and D. Gour. “Screening for Impaired Fasting Glucose and Diabetes Mellitus among People with Major Risk Factors for Type 2 Diabetes Mellitus.” *International Journal of Diabetes in Developing Countries* 27 (2007): 35–40.

20. Smith-Ray, R. L., F. A. Almeida, J. Bajaj, S. Foland, M. Gilson, S. Heikkinen, and P. A. Estabrooks. “Translating Efficacious Behavioral Principles for Diabetes Prevention into Practice.” *Health Promotion Practice* 10 (2009): 58–66.

21. Baus, Adam, Michael Hendryx, and Cecil Pollard. “Identifying Patients with Hypertension: A Case for Auditing Electronic Health Record Data.” *Perspectives in Health Information Management* (Spring 2012): 1–15.

22. American Diabetes Association. “Prediabetes FAQs.” 2013. Available at

<http://www.diabetes.org/diabetes-basics/prevention/pre-diabetes/pre-diabetes-faqs.html>.

23. Centers for Disease Control and Prevention. “Prediabetes: Am I at Risk?” 2012. Available at <http://www.cdc.gov/diabetes/prevention/prediabetes.htm>.

24. Kabcenell, Andrea I., Jerry Langley, and Cindy Hupke. *Innovations in Planned Care*. Cambridge, MA; Institute for Healthcare Improvement, 2006.

25. Hawkins, Dan, and DaShawn Groves. “The Future Role of Community Health Centers in a Changing Health Care Landscape.” *Journal of Ambulatory Care Management* 34, no. 1 (2011): 90–99.

26. Nichols, Gregory A., Jay Desai, Jennifer E. Lafata, Jean M. Lawrence, Patrick J. O’Connor, Ram D. Pathak, Marsha A. Raebel, Robert J. Reid, Joseph V. Selby, Barbara G. Silverman, John F. Steiner, W. F. Stewart, Suma Vupputuri, and Beth Waitzfelder. “Construction of a Multisite DataLink Using Electronic Health Records for the Identification, Surveillance, Prevention, and Management of Diabetes Mellitus: The SUPREME-DM Project.” *Preventing Chronic Disease* 9 (2012): 110311.

27. Centers for Medicare and Medicaid. “CMS EHR Meaningful Use Overview.” 2012.

Available at http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Meaningful_Use.html.

28. Ornstein, S., P. Nietert, R. Jenkins, A. Wessell, L. Nemeth, and H. Rose. “Improving the Translation of Research into Primary Care Practices: Results of a National Quality Improvement Demonstration Project.” *Joint Commission Journal on Quality and Patient Safety* 34, no. 7 (2008): 379–90.

29. Centers for Disease Control and Prevention. “Screening for Type 2 Diabetes.”

30. National Diabetes Information Clearinghouse. “Diabetes Prevention Program.” 2012.

Available at <http://diabetes.niddk.nih.gov/dm/pubs/preventionprogram>.

Table 1

Chronic Disease Electronic Management System Data Fields Populated from Electronic Health Records

Field Name	Field Description
patient_ID	Auto-generated unique identifier for each patient
clinic_code	Identifies site at which care is being provided
sex	Sex
ethnicity	Ethnicity
age	Age
start_date	Date on which patient was added to the electronic record
raw_code (health)	Health condition
start_date (health)	Date on which diagnosis was added to the patient health profile
end_date (health)	Date on which diagnosis was archived in the patient health profile
raw_code (lab)	Laboratory test
result (lab)	Laboratory result
service_date (lab)	Date on which lab result was received
raw_code (service)	Service
result (service)	Service result
service_date (service)	Date on which service was provided
raw_code (medication)	Medication
start_date (medication)	Date on which medication was prescribed
end_date (medication)	Date on which medication was discontinued
visit_date	Date on which office visit occurred
bp_systolic	Systolic blood pressure (linked to visit_date)
bp_diastolic	Diastolic blood pressure (linked to visit_date)
weight	Weight (linked to visit_date)
height	Height (linked to visit_date)
waist_circ	Waist circumference (linked to visit_date)

Table 2

Query Sequence and Descriptions

Sequence	Description of the Query Function
1	Returns a list of active patients
2	Builds on query 1 to return the number of active patients who are established (defined as having received care for 12 months or more)
3	Returns a list of patients with a documented diagnosis of diabetes or prediabetes
4	Compares the list of established patients (query 2) to the list of patients with a documented diagnosis of diabetes or prediabetes (query 3) and returns only those established patients without a diagnosis of diabetes or prediabetes
5	Builds on query 4 to identify patients who are age 45 years or older
6	Builds on query 4 to identify patients with last BMI greater than 25
7	Compares the list of patients age 45 years or older (query 5) to the list of patients with last BMI greater than 25 (query 6) and returns only those patients meeting both criteria
8	Builds on query 4 to identify patients who are younger than 45 years
9	Compares the list of patients younger than 45 years (query 8) to the list of patients with last BMI greater than 25 (query 6) and returns only those patients meeting both criteria
10	Returns a list of patients younger than 45 years with BMI greater than 25 (query 9) with one or more of the following documented health conditions: hypertension, hyperlipidemia, gestational diabetes, family history of diabetes, or cardiovascular disease
11	Builds on query 4 to identify patients with fasting blood glucose in the range of 100–125
12	Limits results from query 11 to the most recent fasting blood glucose results, by patient

Table 3
Patients Identified as at Risk for Prediabetes, Overall and by Primary Care Organization

Primary Care Center	Patients without a Diagnosis of Diabetes or Prediabetes	Patients Age \geq45 Years with Last Calculated BMI $>$25	Patients Age $<$45 Years with BMI $>$25 with Hypertension, Hyperlipidemia, Gestational Diabetes, Family History of Diabetes, or Cardiovascular Disease	Patients with Last Fasting Blood Glucose between 100 and 125	Total Number and Percentage of Patients Identified as at Risk for Prediabetes, by Location
A	1,546	112	18	1	131 (8.5%)
B	1,682	334	40	4	378 (22.5%)
C	2,068	308	49	1	358 (17.3%)
D	1,050	54	7	70	131 (12.5%)
E	1,110	15	3	0	18 (1.6%)
F	1,849	62	15	2	79 (4.3%)
G	2,068	284	35	11	330 (16.0%)
H	5,517	235	26	21	282 (5.1%)
I	8,407	669	70	0	739 (8.8%)
J	17,792	2,467	288	1,627	4,382 (24.6%)
K	10,026	557	52	504	1,113 (11.1%)
L	9,185	627	91	3	721 (7.8%)
M	19,038	1054	90	2	1,146 (6.0%)
N	12,945	794	69	2	865 (6.7%)
Total	94,283	7,572	853	2248	10,673 (11.3%)
Mean	6,734.5	540.8	60.9	160.6	
Standard deviation	6,307.2	635.2	71.4	442.5	