NLP-enabled decision support for cervical cancer screening and surveillance

Principal Investigator and Team Members:

1. Rajeev Chaudhry MBBS, MPH[1]. PI
2. Kathy L. MacLaughlin MD[1]. Co-Investigator
3. Ravikumar Komandur Elayavilli[1]. PhD. Co-Investigator
4. Maya E. Kessler MD[1]. Co-Investigator
5. Hongfang Liu, PhD[1]. Co-Investigator
6. Kavishwar B. Wagholikar. MD, PhD[2]. Co-Investigator
7. Marianne R. Scheitel. BA[1]. Co-Investigator
8. Branden C. Hickey. BA[1]. Co-Investigator
9. Walter K. Kremers, Ph.D[1]. Co-Investigator
10. Darrell R. Schroeder. PhD[1]. Co-Investigator

[1]Mayo Clinic, Rochester, MN, USA; [2]Massachusetts General Hospital, Boston, USA

Organization: Mayo Clinic, Rochester, MN

Dates of Project: September 30, 2014 to November 29, 2016

Federal Project Officer: Bryan Kim PhD.

Agency Support: Division of Health Information Technology (IT), Center for Evidence and Practice Improvement (CEPI), Agency for Healthcare Research and Quality (AHRQ).

Grant Award Number: R21 HS22911

Abstract :

1. **Purpose:** To develop, validate and optimize the NLP enabled CDSS for cervical cancer screening and surveillance. To determine the impact of reminders to non-adherent high-risk patients and the impact of real time CDSS to providers.
2. **Scope:** Cervical cancer screening and follow up of abnormal findings Guidelines are complex leading to poor compliance. Currently available CDS in EHR's don't address the follow up for abnormal findings as cytology reports are in text format in the.
3. **Methods:** We utilized national guidelines to develop CDSS which has 54 clinical pathways with 13 pathways for screening and 41 pathways for surveillance. We then identified data required for the CDSS, developed it and tested that for accuracy. Once errors were rectified, we utilized Big data infrastructure to process patient records and deliver list of patients who were overdue for surveillance and later real time screening and surveillance recommendations to providers seeing patients. Patients overdue for surveillance were contacted one time to call for appointment.
4. **Results:** With iterative development and stratified testing we improved accuracy of the CDSS from 93.6% to100 %. For the patients contacted for surveillance rates improved from 5.7 % to 23.7 % (p value < 0.001). From delivery of recommendations to providers the overall rates improved from 32.6% to 60.8% (p value < 0.001). Rates for patients needing only screening improved from 33.5% to 64.8% (p value < 0.001).
5. **Key Words:** Cervical cancer screening and surveillance, Natural language processing, Clinical Decision Support System.

Purpose (Objectives of Study):

1. Develop, validate and optimize the CDS system in the clinical setting.

Develop a natural language processing (NLP) enabled clinical decision support (CDS) system to pull patient information (text and discrete) from electronic health record (EHR) to determine optimal recommendations for screening and surveillance for cervical cancer based on national guideline. Validate the system in non-clinical setting using clinical experts. Then optimize and implement the system in the organization's Electronic Health Record for real time delivery of knowledge to the providers in primary care practices.

2. Determine the impact of reminders to non-adherent high-risk patients.

From the NLP enabled CDS identify population of patients that are due or overdue for surveillance due to prior abnormalities noted on pap smear cytology or Human Papilloma Virus (HPV) testing. Work with practices to send reminders via a patient portal for patients who have portal account or mail for patients to call to schedule visit for pap smears. Study the effects of delivering reminders on completion of services in comparison to historical controls.

3. Determine the impact of CDS alerts to healthcare providers.

Deliver real time CDS alerts to providers when seeing patients who are due for screening or surveillance. Study the effects of delivering alerts on completion of services in comparison to historical controls.

Scope:

Background, incidence and prevalence: The introduction and implementation of cervical cancer screening has reduced cervical cancer incidence and death in the United States by over 60% but there are still 12,000 new cases and 4,000 deaths annually. The progression from precancerous cervical cytologic abnormalities to cervical cancer occurs over many years, which allows screening to be successful as long as appropriate intervention and follow-up occurs in response to abnormal screening results. Highest risk groups for the development of cervical cancer include women never screened, under-screened and those with delayed or no follow-up of abnormal test results.

One factor that may contribute to inadequate care of women with a history of abnormal Pap or HPV results is the complexity of the management guidelines. Cervical cancer screening and management guidelines were updated by the American Society for Colposcopy and Cervical Pathology (ASCCP) in 2012 and then revised in 2013. Studies of clinician application of the ASCCCP screening guidelines reflect low levels of understanding and compliance. Surveys have reported 12.1% of gynecology and primary care clinicians were not aware of the updated guidelines one year post-release and just 5.7% answered all of the presented knowledge questions correctly. Over half of Pacific NW region gynecologists surveyed reported performing screening tests more often than recommended by the updated guidelines. Among the 65% of 1,268 gynecology and primary care clinicians surveyed who endorsed support for the ASCCP

screening guidelines, only 15% recommended correct test type and screening intervals across all age groups.

Fewer studies have focused on adherence to the more complicated guidelines for management of abnormal cervical cytology (Pap test) or Human Papilloma Virus (HPV) tests. A Canadian study of compliance with guidelines developed in Ontario for managing low grade abnormalities found over-utilization of colposcopy referral but more concerning, a lack of recommended follow-up in 13.4-14.0% of women. In a recent publication specifically assessing compliance with ASCCP 2009 guidelines for management of abnormal Pap tests, over half of patients in one of three university-based practices did not receive guideline-adherent intervention or were lost to follow-up. There is acknowledgment by the lead author of the ASCCP guidelines that the management algorithms are complex and will likely only increase in complexity with new test modalities and that information technology must be applied to assist clinicians.

Clinical decision support systems (CDSS) offer the potential to improve appropriate follow-up of high-risk patients by analyzing the electronic health records (EHR) to accurately identify patient populations that are not compliant with the guideline-based recommendations. However, cervical cytology (Pap) reports are in text format and current CDSSs available in EHRs only utilize discrete data for decision making and therefore do not provide decision support for patients with abnormal cytology results. We previously developed a prototype version of CDSS on our institutions research information technology (IT) infrastructure to automate recommendations to clinicians on cervical cancer screening intervals and subsequently demonstrated its potential of an enhanced CDSS that included management of abnormal results by utilizing natural language processing.

Setting:

The study was conducted at three primary care sites (four practices) affiliated with the Mayo Clinic in Rochester, Minnesota, an academic medical center. These practices provide community longitudinal care to patients of Olmsted County. The sites are the Northeast Clinic (NE), Northwest Clinic (NW), Primary Care Internal Medicine and Family practices at Baldwin location. The two Baldwin practices function organizationally as one site as they are under same clinical operations.

Participants:

A total of 25,500 women ages 18-65 receive care at these practice sites, with 63% at Baldwin, 22% at NE and 15% at NW. Each patient is assigned and empaneled to primary care clinicians (physician, nurse practitioner, or physician assistant) at the three participating clinical sites.

Methods:

1. Develop, validate and optimize the CDS system in the clinical setting.

1.1 Development of CDSS

The American Society for Colposcopy and Cervical Pathology (ASCCP) published updated guidelines for screening and managing abnormal cervical cancer screening results in 2013. To accommodate these changes we revised our clinical pathways in our previous prototype (Figure 1). The updated CDSS has 54 clinical pathways with 13 pathways for routine screening and 41 pathways for high risk patients.

We then worked to develop the CDSS workflow to deliver real time CDSS recommendation to the provider. Figure 2 captures the details of the workflow.

We then worked on defining the primary and secondary data elements required for the CDSS. The CDSS is developed via a three step process: (1) extraction of primary and derived data elements from the clinical records as outlined in Table 1; (2) reassembly of the data elements extracted from multiple reports of the patient in a timeline fashion; and (3) delivery of care recommendations on the reassembled data based on the ASCCP guidelines. Figure 3 and Figure 4 illustrate the details of the three step process.

Table1. Data elements required for cervical cancer screening and surveillance.

Report type	Primary data elements	Secondary data elements
Cytology report	LSIL, HSIL, ASCUS	Recent PAP, Previous PAP, Previous to Previous PAP, Any Previous three cytologies either HSIL, ASCH or AGC
HPV test	Positive, Negative	Recent HPV, Prior HPV
Pathology/Histology	CIN2, CIN3	History of CIN2/CIN3, History of Colposcopy
Surgery	Hysterectomy	History of Hysterectomy
Demographics	Age, Sex	Age at recent PAP, Age at recent HPV
Problem List	Immuno-deficiency, HIV, Transplant, in utero DES (di-ethylstilbesterol) exposure, Cervical cancer, Adenocarcinoma in-situ (AIS)	

1.2 Validation of CDSS

We then electronically ran the revised CDSS on all primary care practice patients whose primary care clinician is at the Mayo Clinic Rochester and had visits to their primary care clinician between May 1st and May 15th, 2015. We performed stratified sampling to select patients from a pool of ~4000 patients for the accuracy testing of our CDDS. Stratified sampling was done to ensure that patient from every possible clinical decision pathway were included in the evaluation. About 10% of the population (393 patients) was selected for testing as shown in table 2. We calculated the accuracy of the recommendations for both normal and abnormal test results. The evaluation was carried out by a domain expert supervised by a physician and both were both familiar with the ASCCP guidelines. We then further enhanced the CDSS to address the errors before implementation in the practice.

Table 2. Patient sampling from different recommendation end points

Decision End Point	# of Patients	# of Patients Sampled	Percentage sample
R12	40	4	10.0
R14	31	3	9.7
R15	47	5	10.6
R16	1	1	100.0
R18	5	2	40.0
R19	58	6	10.3
R20	15	2	13.3
R23	96	10	10.4
R25	124	12	9.7
R26	69	7	10.1
R27	11	1	9.1
R28	174	17	9.8
R29	52	5	9.6

Decision End Point	# of Patients	# of Patients Sampled	Percentage sample
R33	58	6	10.3
R35	19	2	10.5
R39	263	26	9.9
R4	717	70	9.8
R41	112	11	9.8
R42	110	11	10.0
R45	167	17	10.2
R46	843	83	9.8
R47	5	5	100.0
R48	151	16	10.6
R49	326	34	10.4
R5	2	2	100.0
R52	5	5	100.0
R6	76	8	10.5
R7	3	3	100.0
R8	8	7	87.5
R9	116	12	10.3
Total	3704	393	10.6

1.3 Optimization of CDSS for Clinical Setting

Our initial system took on an average takes 54 seconds to process the data and compute recommendations for a single patient. To improve performance we worked with our institutional IT group in Big Data infrastructure to scale up the CDSS solution for real time clinical use. The final architecture is shown in figure 5.

2. Determine the impact of reminders to non-adherent high-risk patients

2.1 Study Design

The impact of the intervention was tested by comparing response rates (Pap or Pap/HPV co-test completion) in women sent electronic or letter reminders to historical controls.

2.2 Eligibility Criteria for reminders

Women identified by the CDSS as being at increased risk for cervical cancer and overdue for guideline-based follow-up were included. These were women with a history of abnormal Pap, HPV or colposcopic biopsy who were overdue for follow-up and women with a history of in utero DES exposure, CIN 2-3 or adenocarcinoma in situ in the prior 20 years or cervical cancer ever, solid organ transplant, HIV infection or on chronic immunosuppressant medication.

Patients were excluded if they already had a primary care or gynecology clinic appointment scheduled for a Pap test.

2.3 Intervention

Women identified as appropriate candidates for the intervention were sent electronic reminders for follow-up if they were registered on the clinic patient portal or letter reminders if they were not. The reminders explained that the patient had been identified as being higher risk and overdue for screening or overdue for follow-up of a past abnormal result. The patient was given a phone number to call to schedule an appointment. Letters at the first clinic (NE) were sent staggered over 6 weeks because of concern about appointment access. Later, as appointment availability was not observed to be a problem at the first site, reminders were sent on a single date for the two additional sites (NW and Baldwin).

2.4 Data Sources

From the CDSS running on the population of eligible patients we were able to identify patients who had the recommended service complete at four months after reminder being sent. We also collected demographic information like age, race, ethnicity, insurance, English proficiency from the demographics part of the EHR.

2.5 Limitations

The CDSS could not electronically identify if patients already had a primary care or gynecology clinic appointment scheduled for a Pap test. CDSS also could not identify women who had their Pap test follow-up completed at an outside facility or had exited screening because of medical comorbidities and limited life expectancy. This limitation was overcome by electronic medical record review by the study team (KLM, MEK,

MRS) or by the patient's primary care clinician prior to reminder being sent to the patient.

3. Determine the impact of CDS alerts to healthcare providers

3.1 Study Design

The impact of the intervention was tested by comparing response rates (Pap or Pap/HPV co-test completion) in women seen by primary care providers for whom there was a recommendation to historical controls.

3.2 Intervention

After ensuring accuracy of CDSS with the review by providers during the reminder phase of the study we implemented the output of CDSS as a web service to be consumed by the existing EHR solution for the preventive services and chronic conditions for the real time delivery of recommendations for cervical cancer screening and surveillance.

3.3 Data Sources

From the CDSS running on the population of patients that had an appointment with their primary care provider we were able to identify patients who had the recommended service complete at four months after being seen. We also collected demographic information like age, race, ethnicity, insurance, English proficiency from the demographics part of the EHR.

3.4 Limitation

As we implemented real time CDSS to providers right after the reminder letter intervention part of the study for patients who were due for surveillance was complete, it could negatively affect the impact of this intervention for the surveillance.

Results:

1. Validation of the CDSS in the clinical setting

1.1 CDSS Accuracy

Out of the 393 patients evaluated, the revised system made correct recommendations to 369 patients achieving an accuracy of 93.4%. Table 3 stratifies the results for patients with previous normal and abnormal test results. Out of the 393 patients 307 patients had normal past test results while 86 had a history of abnormal results (ratio of abnormal to normal is approximately 1:3). The accuracy of the CDSS among the patients who had normal Pap history was higher (96.7%) than among patients who had abnormal PAP results (83.7%).

1.2 Error analysis

We performed a detailed analysis of the errors we encountered while testing our new prototype which is summarized in Table 3. The errors can be broadly categorized into four broad categories: (1) data source errors; (2) modeling errors (include two sub-categories); (3) programming errors; and (4) evaluation errors.

As can be seen from the table we were able to address the issues related to modeling, programming and evaluation prior to the implementation of the CDSS.

Table 3. Error Analysis and Categorization

Error type	Specific Error	No. of errors	Ability to address	Solution
Data Source errors	Errors in coded problem list	5	No	Feed back to the data sources
Modeling errors	Clinical decision not clearly captured in the decision logic	6	Yes	Altering the implementation based on expert feedback
	Lack of adherence to ASCCP guidelines in past clinical practice	5	No	Such errors will gradually be eliminated once clinical practice strictly adheres to ASCCP guideline
Programming errors	Determination of correct end points	7	Yes (Partially solvable)	While programming errors of simple kind can be permanently eliminated, certain error of correct next follow-up time may not be possible.
Evaluation errors	Clinician arriving at a wrong decision	4	Yes	Adoption of such CDSS described in this paper has the potential to eliminate such manual errors

2. Impact of reminders to non-adherent high-risk patients

There was significant improvement noted for the intervention group that received one reminder compared to the historical control. This improvement was consistent across all practice sites and overall rates improved from 5.7% to 23.7% (odds ratio 4.71, p value <0.001) as shown in table 4.

Table 4. Impact of reminders for the Surveillance, Overall and by Clinic Site

	Surveillance Test Completion Rate				
Site	Intervention	Control	OR	95% CI	p value
Overall , % (n)	23.7% (61/257)	5.7% (30/529)	4.71	2.90-7.62	<0.001
NE Clinic, % (n)	31.8% (21/66)	10.5% (9/86)	3.89	1.61-9.29	0.002
NW Clinic, % (n)	23.2% (22/95)	2.7 % (2/74)	11.02	2.44-49.74	0.001
Baldwin, % (n)	18.8% (18/96)	5.15% (19/369)	4.12	2.05-8.26	<0.001

Table 5 displays the demographic characteristics of the two populations. The intervention group was not significantly different than the control group.

Table 5. Patient Characteristics for the Patient reminder intervention for surveillance

	Included Patients		
Characteristic	Intervention Group n=257	Control Group n= 529	p value
Age in years, mean (SD)	43.9 (13.1)	42.8 (13.6)	0.92
Race or ethnicity			
Asian, % (n)	2.5% (6/242)	2.0% (10/513)	
Black or African American, not Hispanic or Latino, % (n)	7.0% (17/252)	4.9% (25/513)	
Hispanic or Latino, % (n)	1.2% (3/242)	1.8% (9/513)	
White, not Hispanic or	89.3% (216/242)	91.4% (469/513)	0.58

Characteristic	Intervention Group n=257	Control Group n= 529	p value
Latino (n, %)			
Insurer			
Government, % (n)	21.8% (56/257)	23.3% (123/529)	
Private, % (n)	5.1% (13/257)	4.5% (24/529)	
None, % (n)	73.2% (188/257)	72.2% (382/529)	0.87
Education Level			
High school or less, % (n)	19.5% (48/246)	20.1% (106/506)	
Some college, % (n)	40.1% (100/246)	41.3% (209/506)	
Four year college, % (n)	21.5% (53/246)	20.8% (105/506)	
Post-graduate, % (n)	18.3% (45/246)	17.0% (86/506)	0.94
Limited English Proficiency, % (n)	3.2% (8/248)	1.6% (9/516)	0.13

3. Impact of CDSS alerts to healthcare providers

Table 6 shows the results of the intervention of real time CDSS to providers compared to control period. The overall screening rate went up significantly from 32.6% to 60.8% (odds ratio 2.42, p value <0.001). However the gains were isolated in the patients who were due for screening. There was no significant difference in surveillance rates

Table 6. Impact of CDSS alerts to healthcare providers on Screening and Surveillance Rate

	Test Completion Rate				
Site	Intervention	Control	OR	95% CI	p value
Overall, % (n)	60.8 (794/1307)	32.6% (375/1150)	2.42	2.02-2.91	<0.001
Screening, % (n)	64.8% (771/1189)	33.5% (322/961)	2.75	2.25-3.36	<0.001
Surveillance, % (n)	19.5% (23/118)	28.0% (53/189)	0.62	0.36-1.08	0.09

Table 7 displays the demographic characteristics of the two populations. The intervention group patients were younger, less white, and with limited English proficiency. There was no difference regarding insurance and educational level. The differences in the populations were accounted for in our statistical analysis.

Table 7. Patient Characteristics for CDSS alerts to healthcare providers on Screening and Surveillance Rate

Characteristic	Intervention Group n=	Control Group n=	p value
Age in years, mean (SD)	38.8 (13.7)	43.0 (12.5)	<0.001
Race			
Asian, % (n)	5.5% (70/1278)	3.5% (40/1141)	
Black or African American, not Hispanic or Latino, % (n)	5.2% (67/1278)	3.8% (43/1141)	
Hispanic or Latino, % (n)	4.3% (55/1278)	2.0% (23/1141)	
Other, not Hispanic or Latino, % (n)	4.0% (51/1278)	2.6% (30/1141)	
White, not Hispanic or Latino, % (n)	81.0% (1035/1278)	88.1 (1005/1141)	<0.001
Insurer			
Government, % (n)	20.2%(263/1305)	20.4% (233/1142)	
Private, % (n)	4.3% (56/1305)	2.7% (31/1142)	
None, % (n)	75.6% (986/1305)	76.9% (878/1142)	0.10
Education Level			
High school or less, % (n)	20.5% (228/1112)	17.4% (185/1061)	
Some college, % (n)	35.9% (399/1112)	38.0% (403/1061)	
Four year college, % (n)	27.5% (306/1112)	27.5% (292/1061)	
Post-graduate, % (n)	16.1% (179/1112)	17.1% (181/1061)	0.30
Limited English Proficiency, % (n)	7.0% (90/1282)	3.8% (43/1138)	<0.001

The CDSS was found to be accurate by the providers who have utilized it for real time decision support during patient encounters from March 2016 onwards.

Discussion:

Our study demonstrates the value of NLP to enhance the CDS for cervical screening and surveillance. With incremental development and error resolution we achieved 100% accuracy. We also demonstrate how a Big Data infrastructure is essential for performance for real time decision support to providers seeing patients.

We observed that sending reminders to non-adherent high risk patients significantly improved surveillance rates. Due to need for moving to the fulfillment of the third aim of our study (implementing CDS alerts) we were able to send only one reminder letter. We postulate that the response could have been even higher if a second reminder was sent or a phone call was initiated to invite the patient for the screening that was due.

Lastly we successfully demonstrated improvement in overall screening rates with real time CDS to the providers. Sending reminders to non-adherent high risk patients before we implemented the CDS alerts limited our ability to study the effectiveness of this intervention in this high risk population.

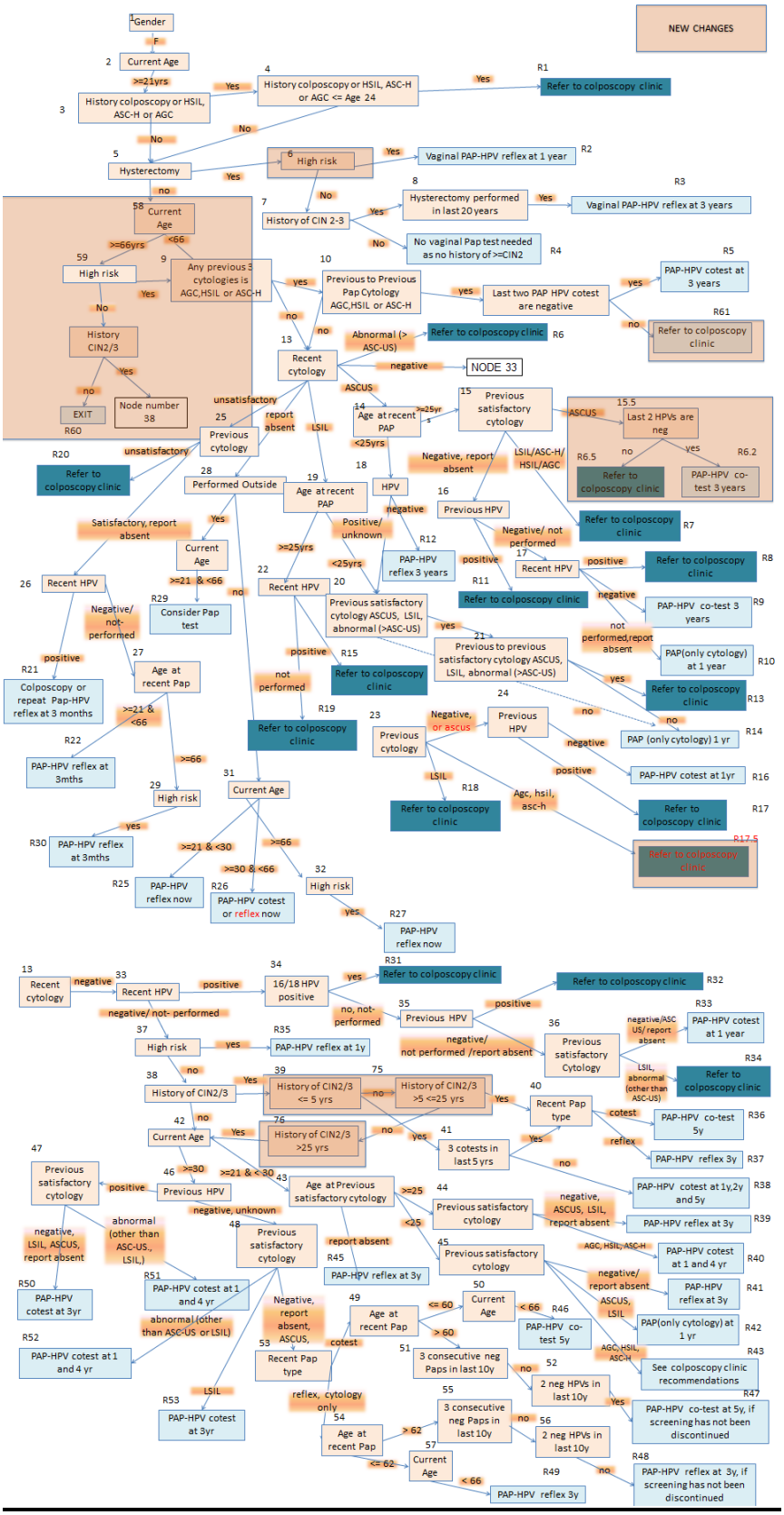
Conclusion:

NLP can significantly improve current CDS by incorporating the findings that currently are recorded in text formats in EMR's. Similar opportunities are there for many other conditions like the surveillance of colon polyps that requires reading of text pathology report and colonoscopy report; lung nodule surveillance where the size of the nodule and previous screenings are needed and many other surveillance opportunities based on the radiology and pathology reports. Enhancing CDS in this manner will be welcome by primary care providers as currently they are spending valuable time looking for information that is stored in the text format. In addition the reminders to high-risk patients who are overdue for their services can significantly improve surveillance rates.

List of publications and products:

1. Computational model for management of abnormal pap smears. (Figure 1).
2. Software code for CDSS for cervical cancer screening and surveillance will be made available at www.ohnlp.org.
3. Working to share the knowledge of developing this CDSS with EMR company, Epic so that it can be replicated across other sites.
4. Two publications are in progress.
5. Abstract accepted at AMIA iHealth 2017 meeting, Philadelphia, May 2-4, 2017. Clinical decision support system improves adherence to appropriate follow-up for patients at higher risk for Cervical Cancer. Kathy L. MacLaughlin, M.D., Maya Kessler MD, Branden C. Hickey, B.S., K.E. Ravikumar, PhD, Marianne R. Scheitel, Hongfang Liu, PhD and Rajeev Chaudhry, MBBS, M.P.H.
6. Presentation by Dr. MacLaughlin on Impact of guideline changes on CDSS for cervical cancer screening and surveillance at the national meeting of American Society for Colposcopy and Cervical Pathology, April 13 - 16, 2016 New Orleans, LA

Figure 1 - Computational workflow for
Cervical cancer surveillance and screening after ASCCP guidelines update



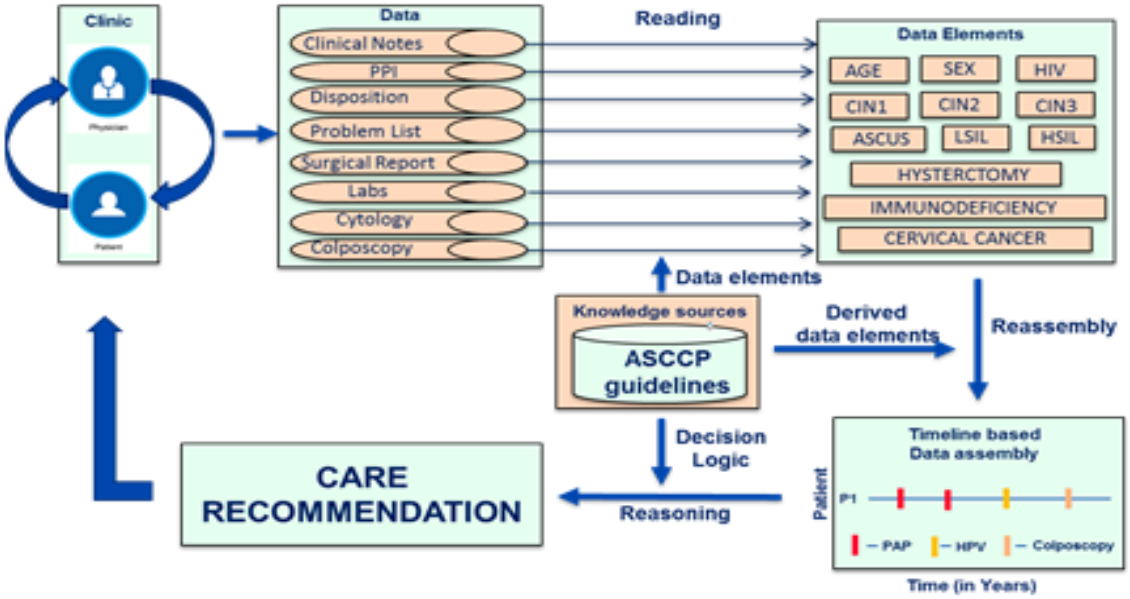


Figure 2. Cervical Cancer Screening and Surveillance CDSS workflow

Data sources and data elements required for the CDSS. Application of the ASCCP guidelines to develop derived data elements (or recommendation) which is then applied to timeline of patients previous services to come up with real time care recommendation.

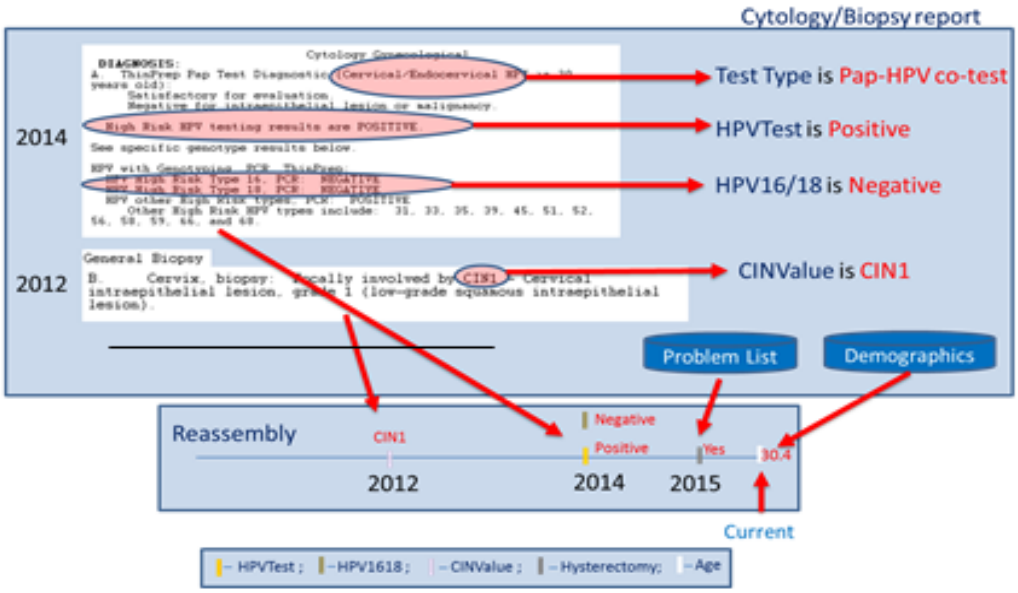


Figure 3. Temporal assembly of primary and secondary data elements related to cervical cancer screening for a given patient

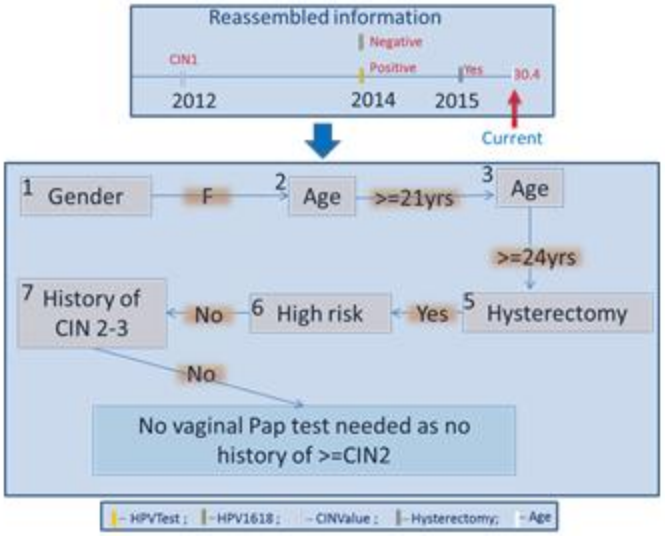


Figure 4. Example of a decision logic rule that computes care recommendation based on data elements and their respective values over a time period by CDSS.

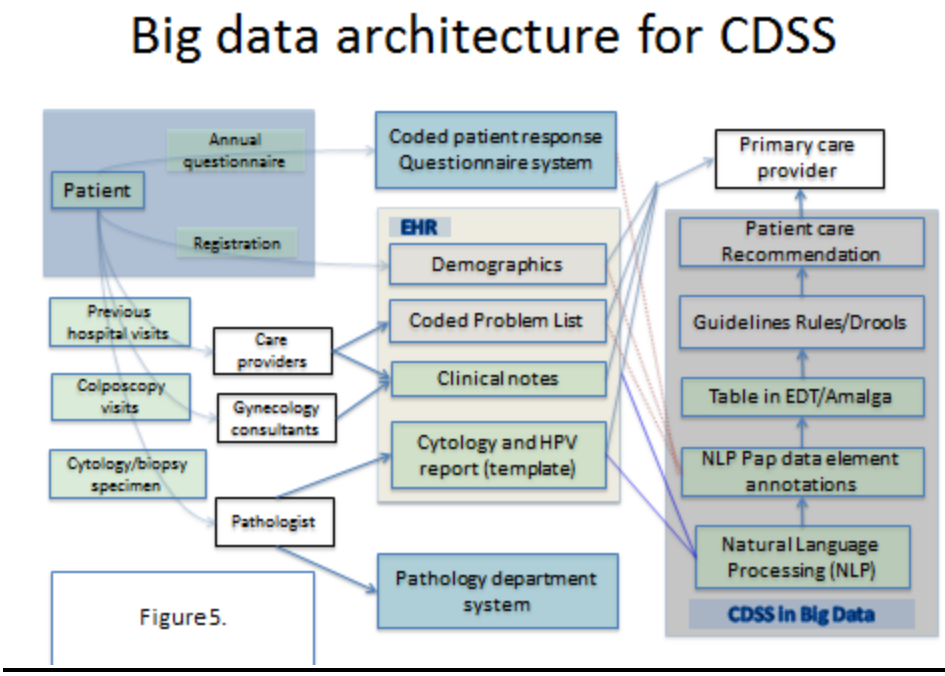

Big data architecture for CDSS
Annual questionnaire
Patient
Registration
Coded patient response Questionnaire system
EHR
Demographics
Coded Problem List
Clinical notes
Cytology and HPV report (template)
Previous hospital visits
Care providers
Colposcopy visits
Gynecology consultants
Cytology/biopsy specimen
Pathologist
Pathology department system
Primary care provider
Patient care Recommendation
Guidelines Rules/Drools
Table in EDT/Amalga
NLP Pap data element annotations
Natural Language Processing (NLP)
CDSS in Big Data

Figure5.