

This is your research on
informatics.

Emily Pfaff

Administrative Director, Informatics & Data Science

NC TraCS Institute



TraCS Informatics: Hiding in plain sight?

- TraCS informatics provides services to researchers at UNC and across the CTSA consortium.
- We've got our hands in a lot of initiatives.



one patient ID, one problem list, one medication list, one bill



CLARK!



EMERSE



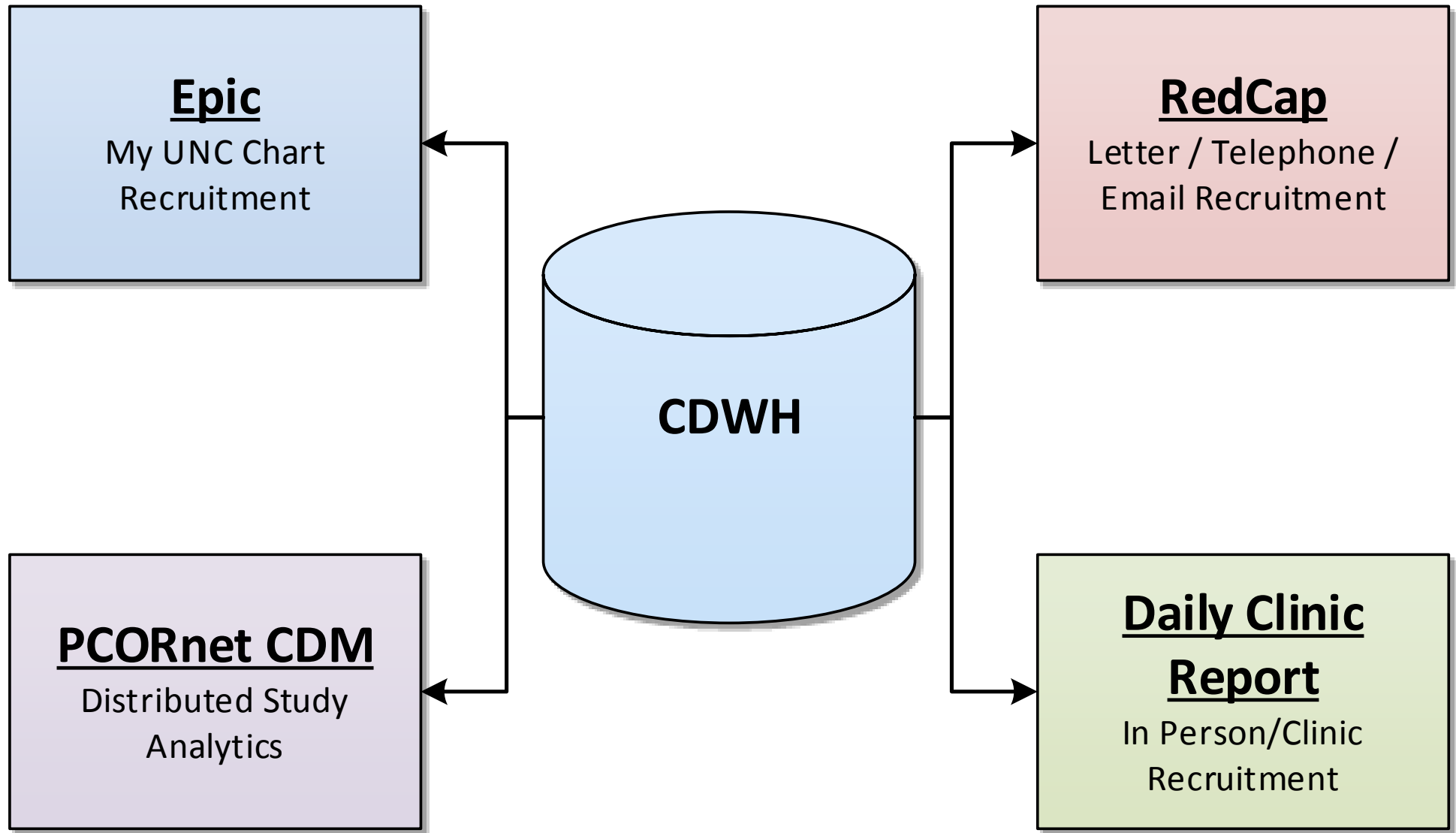
Carolina
Data Warehouse
Research Data
Service



Success Story: ADAPTABLE Trial Recruitment

The ADAPTABLE Trial

- Objective: To perform a large, pragmatic clinical trial to compare the effectiveness and safety of two doses of aspirin (81 mg and 325 mg) in high-risk patients with coronary artery disease
 - And recruit LOTS of patients (20,000 patients across all sites), FAST.
- UNC site PI: Darren DeWalt, MD
- Which informatics tools and tricks could TraCS provide to support recruitment?



✕
Bulk Communication

1. Mail (0 patients)
2. Phone (0 patients)
3. MyChart (582 patients)

Subject:	Help us find the right dose of aspirin for heart disease	Template:	UNCH RSCH ADAPTABLE IRB16-2762	Reply Options	<input checked="" type="checkbox"/> Do not allow patient reply <input type="checkbox"/> Allow reply directly to me
----------	--	-----------	--------------------------------	---------------	---

abc						<input type="text" value="Insert SmartText"/>					<input checked="" type="checkbox"/> Allow all current proxies to view this message
-----	--	--	--	--	--	---	--	--	--	--	--

Dear @NAME@,

You could be part of an exciting national research study about aspirin (the ADAPTABLE study). We know aspirin helps people with heart disease. What we do not know is what dose works best. This study hopes to figure out the best dose of aspirin to reduce the risk of heart attacks and strokes in patients like you. If we find the best dose, we believe we could prevent thousands of deaths from heart disease each year in the United States.

What does the ADAPTABLE study do?
 The study compares regular strength aspirin (325 mg) to "baby" aspirin (81 mg). Both of these doses are effective and routinely prescribed for heart disease.

To participate in the study

- You would continue taking a daily aspirin, perhaps at an adjusted dose, and
- You would fill out a few online surveys.
- You will not have to visit a doctor's office and can participate using the study's secure website.

After you enroll

Patient Outreach Tracking	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 20%; border-right: 1px solid gray; padding: 5px;">Contacted about:</td> <td style="padding: 5px;"> <input type="text" value="Non-Clinical Communication"/> </td> </tr> <tr> <td style="border-right: 1px solid gray; padding: 5px;">Next contact:</td> <td style="padding: 5px;"> <div style="display: flex; gap: 5px;"> <div style="border: 1px solid gray; padding: 2px; display: flex; align-items: center; gap: 5px;"> </div> <div style="border: 1px solid gray; padding: 2px; width: 40px; text-align: center;">1 Week</div> <div style="border: 1px solid gray; padding: 2px; width: 40px; text-align: center;">1 Month</div> <div style="border: 1px solid gray; padding: 2px; width: 40px; text-align: center;">3 Months</div> <div style="border: 1px solid gray; padding: 2px; width: 40px; text-align: center;">6 Months</div> </div> </td> </tr> </table>	Contacted about:	<input type="text" value="Non-Clinical Communication"/>	Next contact:	<div style="display: flex; gap: 5px;"> <div style="border: 1px solid gray; padding: 2px; display: flex; align-items: center; gap: 5px;"> </div> <div style="border: 1px solid gray; padding: 2px; width: 40px; text-align: center;">1 Week</div> <div style="border: 1px solid gray; padding: 2px; width: 40px; text-align: center;">1 Month</div> <div style="border: 1px solid gray; padding: 2px; width: 40px; text-align: center;">3 Months</div> <div style="border: 1px solid gray; padding: 2px; width: 40px; text-align: center;">6 Months</div> </div>
Contacted about:	<input type="text" value="Non-Clinical Communication"/>				
Next contact:	<div style="display: flex; gap: 5px;"> <div style="border: 1px solid gray; padding: 2px; display: flex; align-items: center; gap: 5px;"> </div> <div style="border: 1px solid gray; padding: 2px; width: 40px; text-align: center;">1 Week</div> <div style="border: 1px solid gray; padding: 2px; width: 40px; text-align: center;">1 Month</div> <div style="border: 1px solid gray; padding: 2px; width: 40px; text-align: center;">3 Months</div> <div style="border: 1px solid gray; padding: 2px; width: 40px; text-align: center;">6 Months</div> </div>				

Information in this report will not be refreshed automatically, because you are sending bulk communications for more than 100 patients. Manually refresh rows as necessary.

✓ Accept

✕ Cancel



UNC Health Care
04/04/2017 01:34 PM

Print Delete

Help us find the right dose of aspirin for heart disease

Dear Puccini Research,

You could be part of an exciting national research study about aspirin (the ADAPTABLE study). We know aspirin helps people with heart disease. What we do not know is what dose works best. This study hopes to figure out the best dose of aspirin to reduce the risk of heart attacks and strokes in patients like you. If we find the best dose, we believe we could prevent thousands of deaths from heart disease each year in the United States.

What does the ADAPTABLE study do?

The study compares regular strength aspirin (325 mg) to "baby" aspirin (81 mg). Both of these doses are effective and routinely prescribed for heart disease.

To participate in the study

- You would continue taking a daily aspirin, perhaps at an adjusted dose, and
- You would fill out a few online surveys.
- You will not have to visit a doctor's office and can participate using the study's secure website.

After you enroll

- The ADAPTABLE research team will e-mail a survey to you.
- After you complete that survey, you will receive \$25 to thank you for your time.
- Then, the ADAPTABLE research team will e-mail short health surveys to you every 3 to 6 months, for up to 2 years.

Your input may help researchers and doctors learn how to best treat heart disease. Your participation is voluntary. If you decide not to participate, your care at UNC Health Care will not be affected.

To see if you are eligible and to learn more about the ADAPTABLE, please answer our single-questions survey here: [SURVEY](#)

You can also contact us at adaptable@unc.edu or call our team's research coordinator, Kristie Thompson, at (919) 537-9881.

Thank you for your time.

Darren DeWalt, MD
UNC Health Care

https://redcap.unc.edu/plugins/adaptable_survey.php?sid=f/HZYXuzJNnzYuyMXLcxkA==

ADAPTABLE Aspirin study to help reduce heart attacks and strokes



Dear _____,



Would you like to see if you are eligible to participate in the ADAPTABLE Aspirin study?

- Yes, I am interested.
- I'd like more information.
- I'm not ready at this time.

Please choose one.

Submit



Adaptable

The Aspirin Study

Let's get started!

Thank you for taking the time to find out more details about the ADAPTABLE aspirin study. With your help, we hope to find out what is the right dose of aspirin for people with heart disease.

Got a code?

Please enter in the special code that was included in your invitation:

VLRPW

ENTER

No code? No problem!

You can still learn more about this study even if you have not been asked to participate.



CONTINUE

Already have a profile? [Login](#)

So, did it work?

- When accompanied by a follow-up phone call, My UNC Chart messages resulted in **11% of patients visiting the study website** to learn more, and **5.5% of patients actually enrolling in the study**.
- Letters, the traditional low-touch study recruitment method, only resulted in 3.4% of patients visiting the website, and 1.5% enrolling—even when accompanied by a follow-up phone call.

Contact Method	# Contacted	# Visited Study Website	% Visited Study Website	#Enrolled	% Enrolled
Letter and Phone Follow-Up	409	14	3.4	6	1.5
My UNC Chart Total	1640	212	13	90	5.5
Message Only		30	1.8	12	0.7
Message and Phone Follow-Up		82	11	78	4.8

Sending portal messages without a follow-up phone call (literally, a “no touch” method of recruitment) was less effective, but still shows potential for trials.



New Directions: Text Mining

Consider all the data we're NOT using.

- The CDWH contains a little over 6 terabytes of data
 - For scale: about the same size as the entirety of Wikipedia in 2010.
- The vast majority of that space? Full-text clinical notes.
- A rich source of data—almost useless for analytics in current form.
 - Symptoms, social history, rare diseases, diagnostics, incidental findings...
- Beginning in 2016, TraCS and ISD started working together to allow users to take advantage of notes for research, QI, and hospital operations.

Enter EMERSE



- EMERSE (Electronic Medical Record Search Engine) allows users to search free text (unstructured) clinical notes from the electronic health record (EHR).
- Like "Google" for clinical notes, developed at the University of Michigan in 2005
- EMERSE can support a variety of tasks, including clinical and translational research, internal quality improvement and quality assurance initiatives, as well as hospital operational support tasks.
- We have loaded EMERSE with clinical notes from UNCHCS.

“Natural language” searching

Search Terms to include
Please enter search terms to include one at a time
Term: Add

Click on the term to edit

- cigarette
- migraine
- headache
- painful
- "chest pain"
- cardiac

Drag and drop search terms to obtain desired sort order.

Search Terms to exclude
Please enter search terms to exclude one at a time
Term: Add

Click individual terms to highlight or de-highlight

packs per day tob pipe pack-year smoker cigarettes [spelling alternative] pk-years pk-year cigars pack year smoker cigarette [spelling alternative] PPD pack-years pack year pk year pk years pipes packs cigarettes per day cigaretties cigar py pack years pack-year tobacco smoker pack history

Terms cigarette migraine headache painful "chest pain" cardiac

Search

Color Palette Mosaic

Relevance Low High

MRN	Patient Name	MiChart	CareWeb	Radiology	Pathology	Other
		5 of 11	222 of 451			
			1 of 9			
		9 of 11	89 of 275			
			7 of 36			
			37 of 267			
		5 of 13	34 of 152			
			6 of 22			

That's nice, but you really want me to read through 10,000 notes?

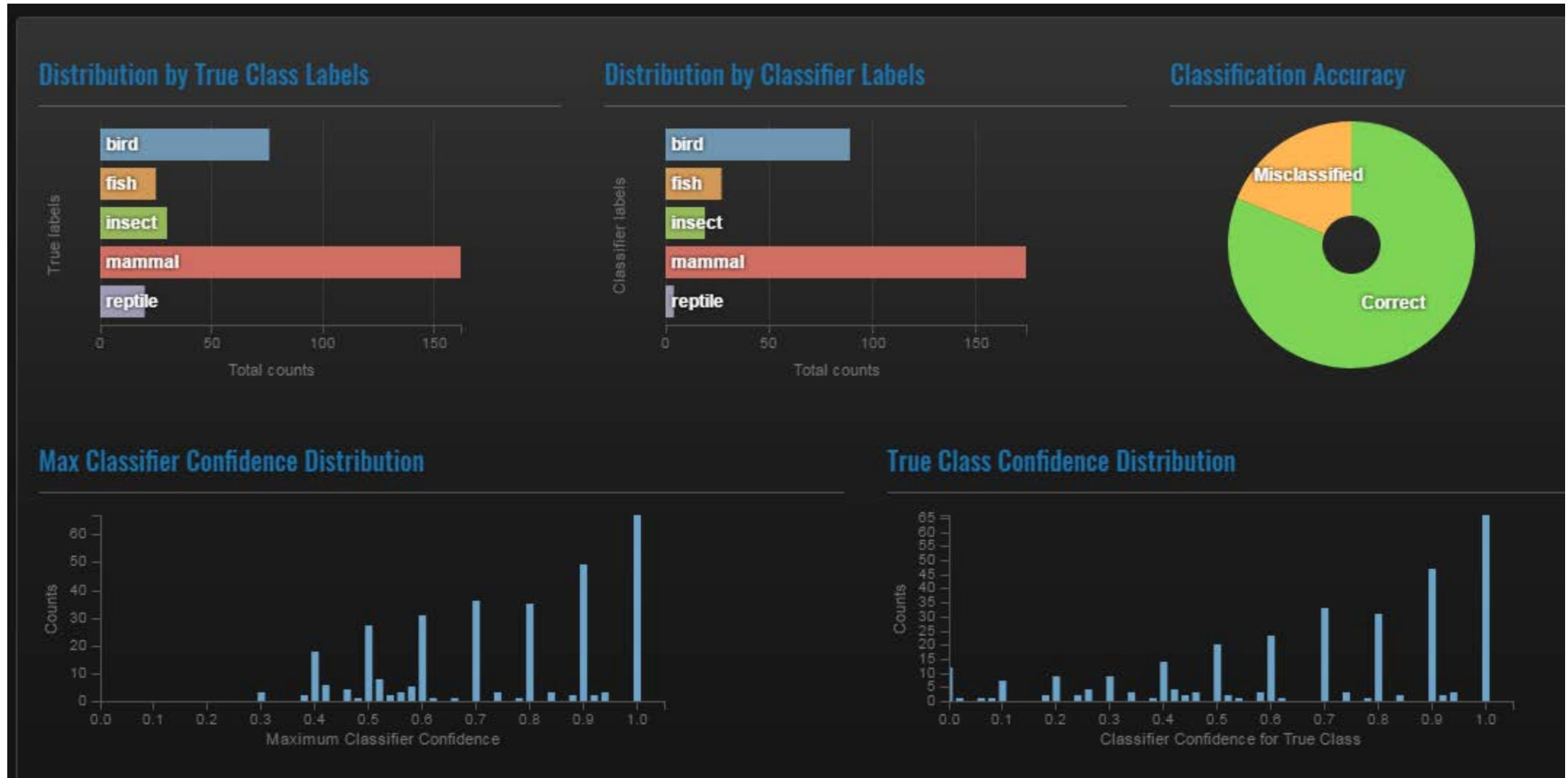
- Well, not necessarily.
- If you're more interested in using clinical notes to find patterns and clusters, or to identify/subset large cohorts, you need a different strategy.
- Basic premise: Clinical subject matter expert (that's you) identifies **features** (that's words and phrases) in clinical notes that serve as clues for what you're looking for.
- An analyst will use that information to help train a *machine learning classifier* to try and recognize those patterns on a large scale.

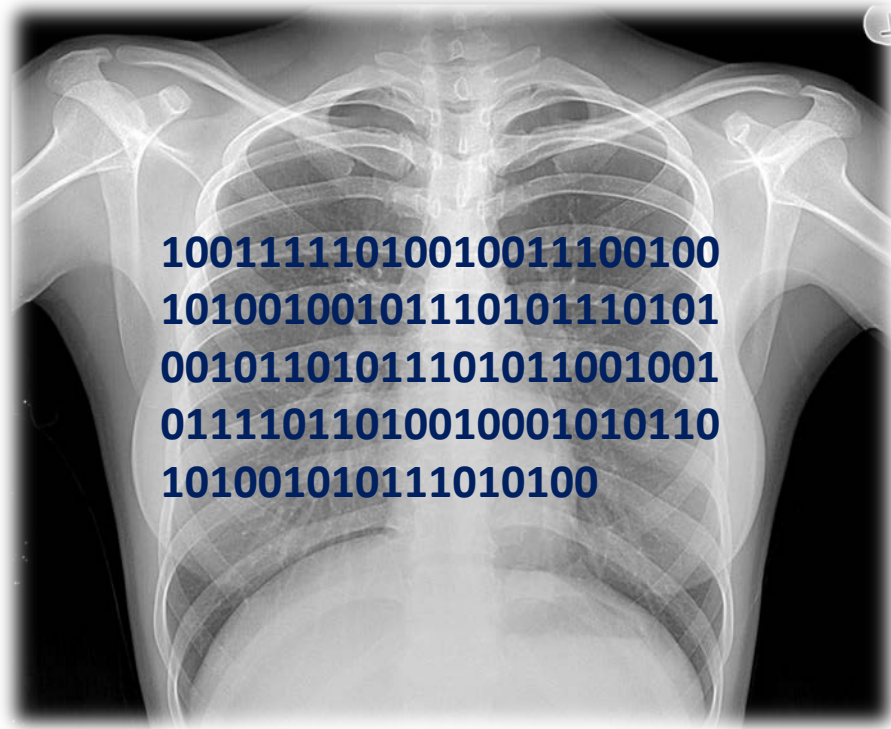
Enter CLARK!



- CLARK! (**CL**inical **A**nnotation **R**esearch **K**it) was developed in collaboration with a local company, CoVar, specializing in machine learning.
- The steps required to use CLARK:
 - Identify a cohort of interest, preferably one with an attainable “gold standard”
 - Identify relevant features in the text that can be used to teach CLARK! which patterns to recognize.
 - Train CLARK! to classify the patients based on the gold standard.
 - Give CLARK! a totally new set of patients/notes to classify.
 - Have a human reviewer check how well CLARK! performed on the new set of notes.
 - Rinse and repeat!

Interface is non-programmer friendly





Pulling it all together: Computable
Phenotype Development

Sounds like tech jargon to me.

- Let's break it down.
- [A phenotype](#) is understood as measurable biological (physiological, biochemical, and anatomical features), behavioral (psychometric pattern), or cognitive markers that are found more often in individuals with a disease or condition than in the general population.
- A [computable phenotype](#) is a clinical condition, characteristic, or set of clinical features that can be determined solely from the data in EHRs and ancillary data sources and does not require chart review or interpretation by a clinician. These can also be referred to as *EHR condition definitions*, *EHR-based phenotype definitions*, or simply *phenotypes*.

What's in it for me?

- With a computable phenotype in hand, you can:
 - Identify a group of patients that potentially qualify for your trial or study, and possibly target them for recruitment.
 - Identify a cohort and variables of interest for secondary data analysis.
 - Where appropriate, fire off alerts in the EHR based on your phenotype (e.g., page a coordinator when a patient arrives meeting criteria x , y , but not z).
 - Within a CDRN, share code for your phenotype with collaborators at other institutions to create a consistent case definition across sites.

And TraCS can help me make one of these things?

- Yes! It's kind of our thing.
- Our Comparative Effective Research (CER) Service can help with initial definitions from an epi perspective.
- Once you work that out, a friendly TraCS analyst from the Informatics Service will help you translate that definition into code.
- Help from our Biostatistics and/or Regulatory service is also available where needed.

How can TraCS support research informatics needs, from



to



?

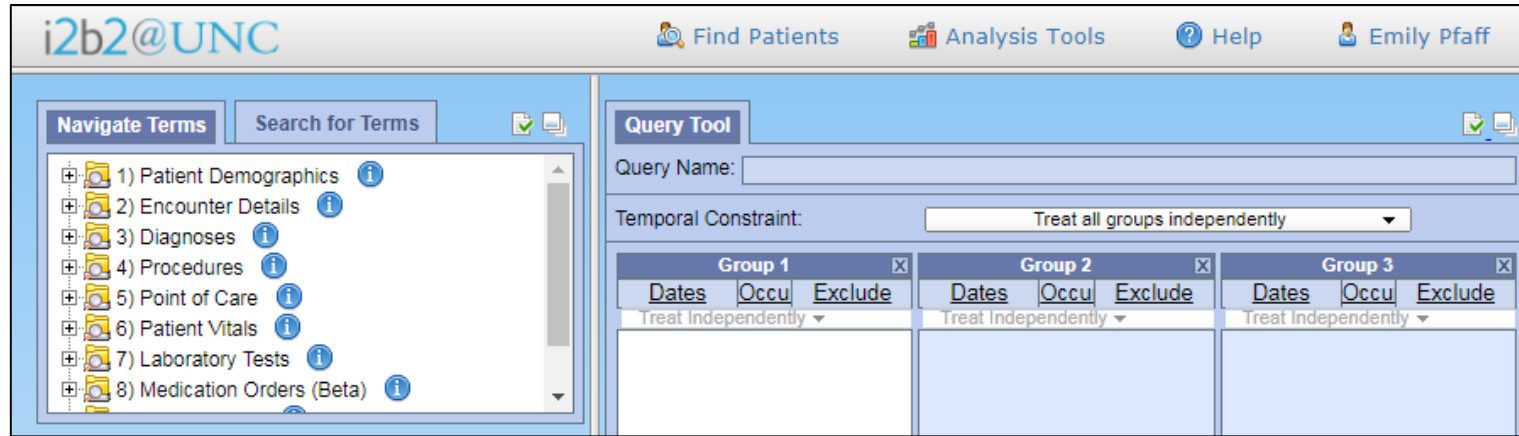
How can TraCS support research analytics needs, from

Structure Data Table							
Structure Number	Size (inches)	Description	Length (feet)	Up Stream Elevation	Down Stream Elevation	Slope %	AD/MH=Rm or CI=FG Elevation
CI#101		Double Curb Inlet, STA 0+99.25			423.44		426.44
PI#102	15	RCP Class III (MIN)	28.0	423.44	422.94	1.79	
CI#103		Double Curb Inlet, STA 0+99.25		422.94	422.94		426.44
PI#104	18	RCP Class III (MIN)	131.0	422.94	413.90	6.90	
AD#105		30"x30" "E" Box		413.90	413.30		417.10
PI#106	24	RCP Class III (MIN)	56.0	413.30	412.90	0.71	
AD#107		Shallow Manhole w/ E.J. 6488 Grate or 48"x48" Box w/ E.J. 6610 Grate		412.90	412.50		416.80
PI#108	30	RCP Class III (MIN)	131.0	412.50	412.20	0.23	
CI#109		Double Curb Inlet, STA 5+91.08		412.20	411.80		415.73
PI#110	36	36" Equivalent Elliptical RCP Class IV (MIN)	28.0	411.80	411.76	0.12	
CI#111		Double Curb Inlet, STA 5+91.08		411.76	411.76		415.73
PI#112	36	36" Equivalent Elliptical RCP Class IV (MIN)	121.0	411.76	411.53	0.19	
CI#116		24x36 Curb Inlet, STA 5+31.08			413.16		416.16
PI#117	15	RCP Class III (MIN)	80.0	413.16	412.69	0.76	
CI#111		Double Curb Inlet, STA 5+91.08		412.69	411.53		415.73
FES#113	15	Concrete Flared End Section with Trash Grate	6.1	410.75	410.70	0.83	
PI#114	15	RCP Class III (MIN)	18.0	410.70	410.65	0.83	
FES#115	15	Concrete Flared End Section	6.1	410.55	410.50	0.83	

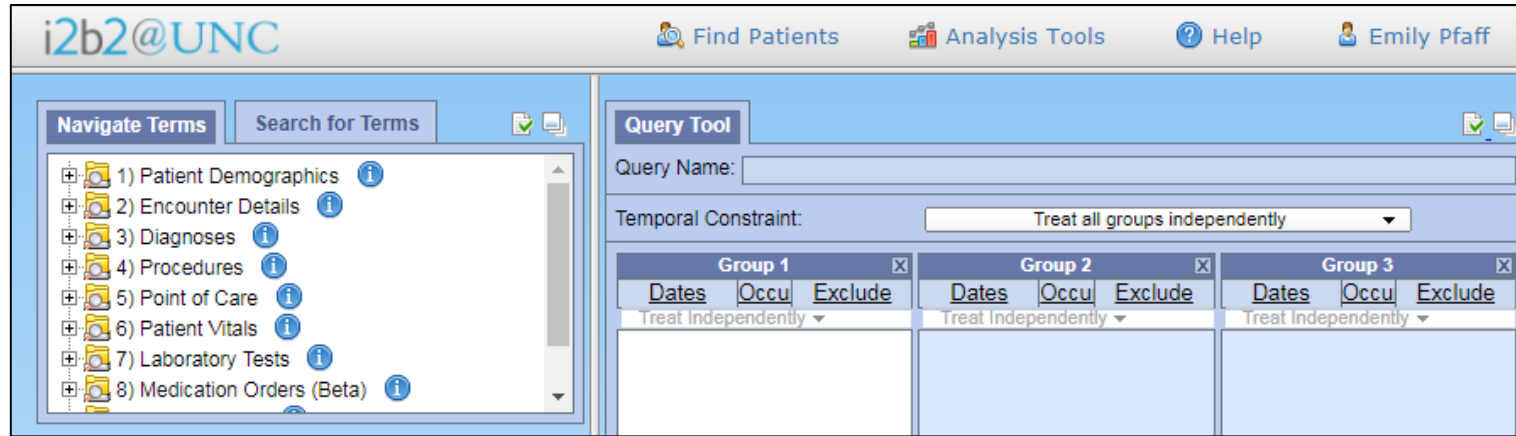
to



?



Use i2b2 to identify a cohort, using as much discrete data as you can.



Use i2b2 to identify a cohort, using as much discrete data as you can.

Patients Patients that Snore (94 Patients)

Dates All Dates: 02/01/2008 through 10/21/2016

Terms overweight obese CPAP apnea

Overview

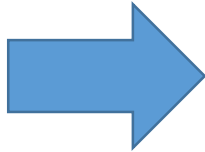
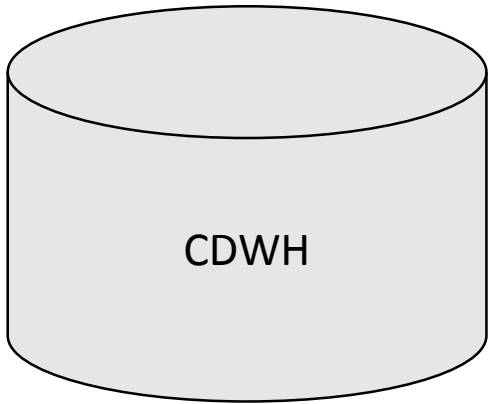
Overview

Sorted By:

Numbers Grayscale Mosaic

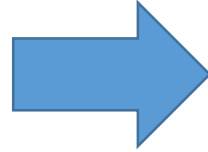
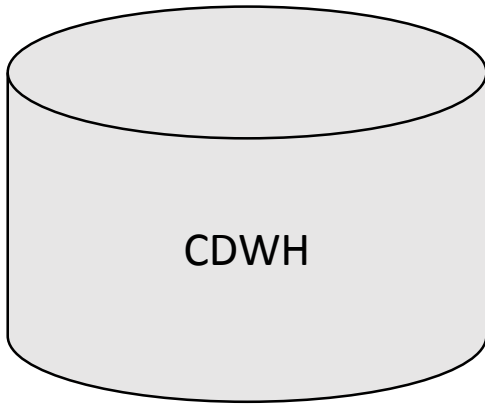
MRN	Patient Name	EPIC
1050444768	PATIENT, PATIENT	

After IRB approval, browse their notes in EMERSE.



Structure Data Table							
Structure Number	Size (Inches)	Description	Length (feet)	Up Stream Elevation	Down Stream Elevation	Slope %	AD/MI=Rim or CIP-CG Elevation
CM#101		Double Curb Inlet, STA 0+99.25			423.44		426.44
PF#102	15	RCP Class III (MIN)	28.0	423.44	422.94	1.79	
CM#103		Double Curb Inlet, STA 0+99.25		422.94	422.94		426.44
PF#104	18	RCP Class III (MIN)	131.0	422.94	413.90	6.90	
AD#105		30"x30" E' Box		413.90	413.30		417.10
PF#106	24	RCP Class III (MIN)	58.0	413.30	412.90	0.71	
AD#107		Shallow Manhole w/ E.J. 6488 Grate or 48"x48" Box w/ E.J. 6610 Grate		412.90	412.50		416.80
PF#108	30	RCP Class III (MIN)	131.0	412.50	412.20	0.23	
CM#109		Double Curb Inlet, STA 5+91.08		412.20	411.80		415.73
PF#110	36	36" Equivalent Elliptical RCP Class IV (MIN)	28.0	411.80	411.76	0.12	
CM#111		Double Curb Inlet, STA 5+91.08		411.76	411.76		415.73
PF#112	36	36" Equivalent Elliptical RCP Class IV (MIN)	121.0	411.76	411.53	0.19	
CM#116		24x36 Curb Inlet, STA 6+31.08			413.16		416.16
PF#117	15	RCP Class III (MIN)	60.0	413.16	412.69	0.78	
CM#111		Double Curb Inlet, STA 6+91.08		412.69	411.53		415.73
FES#113	15	Concrete Flared End Section with Trash Gate	6.1	410.75	410.70	0.83	
PF#114	15	RCP Class III (MIN)	18.0	410.70	410.55	0.83	
FES#115	15	Concrete Flared End Section	6.1	410.55	410.50	0.83	

Extract data (discrete and full-text) from the CDWH.



Structure Data Table						
Structure Number	Size (feet)	Description	Length (feet)	Up Stream Elevation	Down Stream Elevation	Slope %
Ci#101		Double Curb Inlet, STA 0+99.25			423.44	426.44
P#102	15	RCP Class III (MIN)	28.0	423.44	422.94	1.79
Ci#103		Double Curb Inlet, STA 0+99.25		422.94	422.94	426.44
P#104	18	RCP Class III (MIN)	131.0	422.94	413.90	6.90
AD#105		30"x30" E' Box		413.90	413.30	417.10
P#106	24	RCP Class III (MIN)	58.0	413.30	412.90	0.71
AD#107		Shallow Manhole w/ EJ 6488 Grate or 48"x48" Box w/ EJ 6510 Grate		412.90	412.50	416.80
P#108	30	RCP Class III (MIN)	131.0	412.50	412.20	0.23
Ci#109		Double Curb Inlet, STA 5+91.08		412.20	411.80	415.73
P#110	36	36" Equivalent Elliptical RCP Class IV (MIN)	28.0	411.80	411.76	0.12
Ci#111		Double Curb Inlet, STA 5+91.08		411.76	411.76	415.73
P#112	36	36" Equivalent Elliptical RCP Class IV (MIN)	121.0	411.76	411.53	0.19
Ci#116		24x36 Curb Inlet, STA 5+31.08			413.16	416.16
P#117	15	RCP Class III (MIN)	80.0	413.16	412.69	0.78
Ci#111		Double Curb Inlet, STA 5+91.08		412.69	411.53	415.73
FES#113	15	Concrete Flared End Section with Trash Grate	6.1	410.75	410.70	0.83
P#114	15	RCP Class III (MIN)	18.0	410.70	410.55	0.83
FES#115	15	Concrete Flared End Section	6.1	410.55	410.50	0.83

Extract data (discrete and full-text) from the CDWH.

CLARK! Clinical Annotation Research Kit

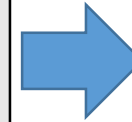
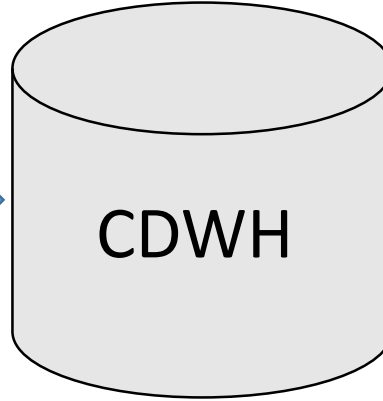
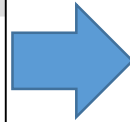
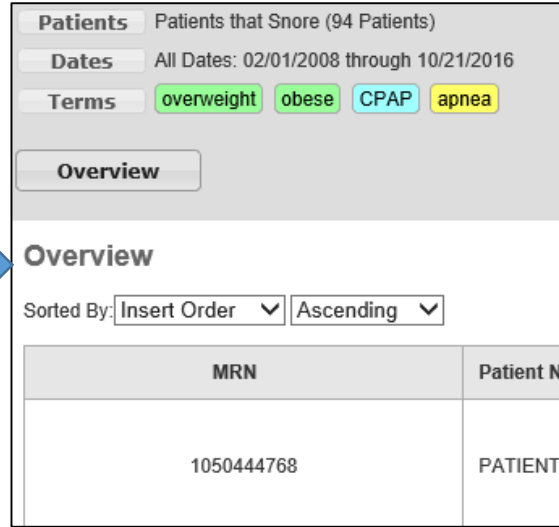
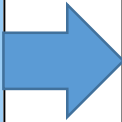
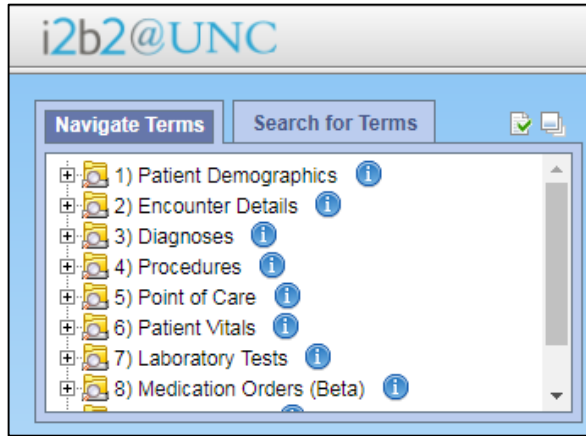
Algorithm Setup

- TRAINING CORPUS 313
- REGULAR EXPRESSIONS LIBRARY 0
- ACTIVE REGULAR EXPRESSIONS 0
- SECTION DEFINITIONS

Training Corpus 313

PATIENTS	NOTES
Aardvark	2016-11-13 19:45:00
Aardwolf	
African buffalo	
African elephant	
African leopard	
Albatross	
Alligator	
Alpaca	
American bison	
American robin	

Perform in-depth cohort classification in CLARK.



The TraCS Informatics pipeline