CS4980: Computational Epidemiology

Sriram Pemmaraju and Alberto Maria Segre
Department of Computer Science
The University of Iowa

Spring 2020
https://homepage.cs.uiowa.edu/~sriram/4980/spring20/

University of Iowa
comp|epi
computational epidemiology research

Hospital-Acquired Infections

Over the last 10 years, our research group has pursued a program of research that ultimately seeks to better understand *hospital-acquired infections* (HAIs).

Hospital-Acquired Infections

Over the last 10 years, our research group has pursued a program of research that ultimately seeks to better understand *hospital-acquired infections* (HAIs).

We have developed new measurement technology to collect data within the community as well as from patients and healthcare practitioners.

Hospital-Acquired Infections

Over the last 10 years, our research group has pursued a program of research that ultimately seeks to better understand *hospital-acquired infections* (HAIs).

We have developed new measurement technology to collect data within the community as well as from patients and healthcare practitioners.

We have devised algorithmic solutions to a broad range of clinical problems, including, for example, patient outcome prediction and diagnostic delays.

Hospital-Acquired Infections

Over the last 10 years, our research group has pursued a program of research that ultimately seeks to better understand *hospital-acquired infections* (HAIs).

We have developed new measurement technology to collect data within the community as well as from patients and healthcare practitioners.

We have devised algorithmic solutions to a broad range of clinical problems, including, for example, patient outcome prediction and diagnostic delays.

A particular focus of our group is on using mathematical models as a basis for simulations that can help inform useful healthcare interventions or detect as yet unrecognized relationships between practices and outcomes.

Hospital-Acquired Infections

Over the last 10 years, our research group has pursued a program of research that ultimately seeks to better understand *hospital-acquired infections* (HAIs).

We have developed new measurement technology to collect data within the community as well as from patients and healthcare practitioners.

We have devised algorithmic solutions to a broad range of clinical problems, including, for example, patient outcome prediction and diagnostic delays.

A particular focus of our group is on using mathematical models as a basis for simulations that can help inform useful healthcare interventions or detect as yet unrecognized relationships between practices and outcomes.

In short, we think of ourselves as John Snow with loads of data, fancy algorithms, and fast computers.

Why Study Hospital-Acquired Infections?

According to the Centers for Disease Control and Prevention (CDC), HAIs affect about 2 million patients in US hospitals each year and result in an estimated 99,000 deaths.

Why Study Hospital-Acquired Infections?

According to the Centers for Disease Control and Prevention (CDC), HAIs affect about 2 million patients in US hospitals each year and result in an estimated 99,000 deaths.

The estimated direct medical costs of HAIs in US hospitals ranges from $28.4 billion to $45 billion per year.

Why Study Hospital-Acquired Infections?

According to the Centers for Disease Control and Prevention (CDC), HAIs affect about 2 million patients in US hospitals each year and result in an estimated 99,000 deaths.

The estimated direct medical costs of HAIs in US hospitals ranges from $28.4 billion to $45 billion per year.

Infections like influenza and MRSA routinely spread to and among hospitalized patients, often with healthcare workers (HCW) as the vector.

Hospital-Acquired Infections

HAIs, like any infection, are spread through interaction; by direct contact, droplet, or airborne means, depending on the nature of the pathogen.

Hospital-Acquired Infections

HAIs, like any infection, are spread through interaction; by direct contact, droplet, or airborne means, depending on the nature of the pathogen.

Within a hospital, HCW behavior can affect disease transmission, through vaccination, hand hygiene, isolation and use of contact precautions (gowns and gloves), travel restrictions (like the Wuhan coronavirus) and other behavioral changes.

Hospital-Acquired Infections

HAIs, like any infection, are spread through interaction; by direct contact, droplet, or airborne means, depending on the nature of the pathogen.

Within a hospital, HCW behavior can affect disease transmission, through vaccination, hand hygiene, isolation and use of contact precautions (gowns and gloves), travel restrictions (like the Wuhan coronavirus) and other behavioral changes.

For example, hand hygiene is to HAI as vaccination is to communicable diseases, but such measures are only effective if adherence rates are high (remember Bernoulli!), and adherence rates among HCWs typically average less than 50%.

Hospital-Acquired Infections

HAIs, like any infection, are spread through interaction; by direct contact, droplet, or airborne means, depending on the nature of the pathogen.

Within a hospital, HCW behavior can affect disease transmission, through vaccination, hand hygiene, isolation and use of contact precautions (gowns and gloves), travel restrictions (like the Wuhan coronavirus) and other behavioral changes.

For example, hand hygiene is to HAI as vaccination is to communicable diseases, but such measures are only effective if adherence rates are high (remember Bernoulli!), and adherence rates among HCWs typically average less than 50%.

Alternatively, policy interventions, such as risk-based room assignments or deep cleaning of rooms at discharge, could also reduce the (population) burden of infections.

Hospital Acquired *Clostridioides difficile* **Infection (CDI)**

Clostridioides difficile, or *C. diff*, is a leading cause of nosocomial diarrhea in the United States and is associated with significant morbidity and mortality in hospitalized patients.

Hospital Acquired *Clostridioides difficile* **Infection (CDI)**

Clostridioides difficile, or *C. diff*, is a leading cause of nosocomial diarrhea in the United States and is associated with significant morbidity and mortality in hospitalized patients.

Symptoms include diarrhea, fever, and nausea; complications may include pseudomembranous colitis, toxic megacolon, perforation of the colon, and sepsis.

Hospital Acquired *Clostridioides difficile* **Infection (CDI)**

Clostridioides difficile, or *C. diff*, is a leading cause of nosocomial diarrhea in the United States and is associated with significant morbidity and mortality in hospitalized patients.

Symptoms include diarrhea, fever, and nausea; complications may include pseudomembranous colitis, toxic megacolon, perforation of the colon, and sepsis.

CDI is spread via the fecal-oral route.

Hospital Acquired *Clostridioides difficile* Infection (CDI)

Clostridioides difficile, or *C. diff*, is a leading cause of nosocomial diarrhea in the United States and is associated with significant morbidity and mortality in hospitalized patients.

Symptoms include diarrhea, fever, and nausea; complications may include pseudomembranous colitis, toxic megacolon, perforation of the colon, and sepsis.

CDI is spread via the fecal-oral route.

C. diff is a spore-forming bacteria (not all bacteria form spores), which can persist on contaminated surfaces for as much as 30 days, and can be spread via HCW hands.

Hospital Acquired *Clostridioides difficile* **Infection (CDI)**

Clostridioides difficile, or *C. diff*, is a leading cause of nosocomial diarrhea in the United States and is associated with significant morbidity and mortality in hospitalized patients.

Symptoms include diarrhea, fever, and nausea; complications may include pseudomembranous colitis, toxic megacolon, perforation of the colon, and sepsis.

CDI is spread via the fecal-oral route.

C. diff is a spore-forming bacteria (not all bacteria form spores), which can persist on contaminated surfaces for as much as 30 days, and can be spread via HCW hands.

Spores are not harmed by alcohol-based hand rub.

Observed Risk Factors for CDI

Men, infants and older adults (>65), patients with complications or comorbidities, those with multiple hospitalizations or those with extended hospital stays are known to be particularly prone to CDI.

Observed Risk Factors for CDI

Men, infants and older adults (>65), patients with complications or comorbidities, those with multiple hospitalizations or those with extended hospital stays are known to be particularly prone to CDI.

CDI is also associated with use/overuse of antibiotics (especially specific antibiotics), proton pump inhibitors and histamine blockers.

Observed Risk Factors for CDI

Men, infants and older adults (>65), patients with complications or comorbidities, those with multiple hospitalizations or those with extended hospital stays are known to be particularly prone to CDI.

CDI is also associated with use/overuse of antibiotics (especially specific antibiotics), proton pump inhibitors and histamine blockers.

ABX use disrupts the intestinal fauna, giving CDI a chance to take hold.

Observed Risk Factors for CDI

Men, infants and older adults (>65), patients with complications or comorbidities, those with multiple hospitalizations or those with extended hospital stays are known to be particularly prone to CDI.

CDI is also associated with use/overuse of antibiotics (especially specific antibiotics), proton pump inhibitors and histamine blockers.

ABX use disrupts the intestinal fauna, giving CDI a chance to take hold.

Note that CDI can also be asymptomatic; that is, a patient can test positive but have no symptoms: many otherwise healthy people — including HCWs and family members — are colonized with *C. diff*, meaning they can spread the disease.

CDI at UIHC

Variable	cases 1,606 (0.66%)	non-cases 239,642 (99.32%)
Age (median, range)	58 (0-98)	45 (0-105)
Age: < 45	450 (28.02)	111,379 (46.48)
Age: $[45, 64]$	575 (35.80)	76,127 (31.77)
Age: > 64	581 (36.18)	52,136 (21.76)
LOS (median, range)	9 (0-447)	3 (0-562)
LOS: < 4	343 (21.36)	133,568 (55.74)
LOS: $[4, 7]$	371 (23.10)	61,913 (25.84)
LOS: > 7	892 (55.54)	44,161 (18.43)
Male	821 (51.12)	115, 765 (48.31)
White	1411 (87.86)	196,741 (82.10)
At least 1 admit in previous 60 days	587 (36.55)	59,608 (24.87)

CDI at UIHC

Variable	cases 1,606 (0.66%)	non-cases 239,642 (99.32%)
CCI (median, range)	2 (0-51)	0 (0-67)
CCI: $= 0$	503 (31.32)	126,948 (52.97)
CCI: 1-2	502 (31.26)	61,683 (25.74)
CCI: ≥ 3	601 (37.42)	51,011 (21.29)
Histamine 2 Blocker	342 (21.30)	31,509 (13.15)
Proton Pump Inhibitor	917 (57.10)	2630 (36.36)
Low Albumin Level	198 (12.33)	6,952 (2.90)
Amoxicillin/ampicillin	130 (8.09)	18,751 (7.82)
Clindamycin	65 (4.05)	8,676 (3.62)
Third-generation cephalosporin	164 (10.21)	11,231 (4.69)
Fourth-generation cephalosporin	251 (15.63)	7409 (3.09)
Fluoroquinolone	501 (31.20)	27,070 (11.30)

CDI Treatment and Prevention

Antibiotics such as metronidazole, vancomycin or fidaxomicin will cure CDI.

CDI Treatment and Prevention

Antibiotics such as metronidazole, vancomycin or fidaxomicin will cure CDI.

Proactive control of antibiotic use, including changing prescription norms and practices, can help decrease the likelihood of a CDI outbreak.

CDI Treatment and Prevention

Antibiotics such as metronidazole, vancomycin or fidaxomicin will cure CDI.

Proactive control of antibiotic use, including changing prescription norms and practices, can help decrease the likelihood of a CDI outbreak.

Once a CDI has occurred, possible interventions to reduce its spread include improved hand hygiene compliance, deep cleaning, use of UV-emitting robots for room disinfection, improved room assignment policies, etc.

CDI Treatment and Prevention

Antibiotics such as metronidazole, vancomycin or fidaxomicin will cure CDI.

Proactive control of antibiotic use, including changing prescription norms and practices, can help decrease the likelihood of a CDI outbreak.

Once a CDI has occurred, possible interventions to reduce its spread include improved hand hygiene compliance, deep cleaning, use of UV-emitting robots for room disinfection, improved room assignment policies, etc.

Fecal microbiota transplants and probiotics may decrease risk of recurrence.

Reducing CDI

Choosing an effective intervention depends on knowing what the primary pathway of infection is likely to be.

Reducing CDI

Choosing an effective intervention depends on knowing what the primary pathway of infection is likely to be.

Samore (1999) lists 3 mechanisms for CDI transmission: **direct** (*e.g.*, from HCW hands), **environmental** (*e.g.*, from spores left in the environment) and **endongenous** (*i.e.*, self colonized).

Reducing CDI

Choosing an effective intervention depends on knowing what the primary pathway of infection is likely to be.

Samore (1999) lists 3 mechanisms for CDI transmission: **direct** (*e.g.*, from HCW hands), **environmental** (*e.g.*, from spores left in the environment) and **endongenous** (*i.e.*, self colonized).

Each of these pathways can be addressed by a different intervention (*e.g.*, better hand hygiene, deep cleaning at discharge, or improved ABX Rx and patient transfer practices).

CDI Pathways

Unfortunately, evidence is mixed for which is the most common pathway.

CDI Pathways

Unfortunately, evidence is mixed for which is the most common pathway.

There is some statistical evidence for *CDI pressure*, where a concurrent CDI case elsewhere on a patient's unit statistically increases that patient's risk for CDI, suggesting a direct pathway.

CDI Pathways

Unfortunately, evidence is mixed for which is the most common pathway.

There is some statistical evidence for *CDI pressure*, where a concurrent CDI case elsewhere on a patient's unit statistically increases that patient's risk for CDI, suggesting a direct pathway.

There is also some evidence that a prior CDI case in the same room increases risk of subsequent CDI, suggesting an environmental pathway.

CDI Pathways

Unfortunately, evidence is mixed for which is the most common pathway.

There is some statistical evidence for *CDI pressure*, where a concurrent CDI case elsewhere on a patient's unit statistically increases that patient's risk for CDI, suggesting a direct pathway.

There is also some evidence that a prior CDI case in the same room increases risk of subsequent CDI, suggesting an environmental pathway.

However, some genotyping studies have shown that a substantial number of cases in the same unit are not genetically related, suggesting an endongenous pathway, perhaps triggered by overzealous use of antibiotics.

CDI Pathways

Unfortunately, evidence is mixed for which is the most common pathway.

There is some statistical evidence for *CDI pressure*, where a concurrent CDI case elsewhere on a patient's unit statistically increases that patient's risk for CDI, suggesting a direct pathway.

There is also some evidence that a prior CDI case in the same room increases risk of subsequent CDI, suggesting an environmental pathway.

However, some genotyping studies have shown that a substantial number of cases in the same unit are not genetically related, suggesting an endongenous pathway, perhaps triggered by overzealous use of antibiotics.

Can we use our UIHC CDI and knowledge of UIHC and its patients to tease these pathways apart?

The Big Idea

Given a set of UIHC CDI cases located in space and time, can we determine if the observed "clustering" is accidental or the result of some underlying pathway?

The Big Idea

Given a set of UIHC CDI cases located in space and time, can we determine if the observed "clustering" is accidental or the result of some underlying pathway?

Showing the clustering behavior did not occur at random is evidence for other than endogenous CDI (*i.e.*, a direct or environmental) pathway.

Theme: Spatiotemporal Context vs. Random Mixing

Traditionally, epidemiologists use *random mixing* to model disease transmission (imagine a herd of cows out to pasture, randomly interacting: that's random mixing).

Theme: Spatiotemporal Context vs. Random Mixing

Traditionally, epidemiologists use *random mixing* to model disease transmission (imagine a herd of cows out to pasture, randomly interacting: that's random mixing).

Random mixing is easy to model!

Theme: Spatiotemporal Context vs. Random Mixing

Traditionally, epidemiologists use *random mixing* to model disease transmission (imagine a herd of cows out to pasture, randomly interacting: that's random mixing).

Random mixing is easy to model!

In contrast, we believe *spatiotemporal context* matters; that architecture and human behaviors conspire to regularize, and not randomize, agent mixing, and that systematic patterns emerge from individual behaviors.

Theme: Spatiotemporal Context vs. Random Mixing

Traditionally, epidemiologists use *random mixing* to model disease transmission (imagine a herd of cows out to pasture, randomly interacting: that's random mixing).

Random mixing is easy to model!

In contrast, we believe *spatiotemporal context* matters; that architecture and human behaviors conspire to regularize, and not randomize, agent mixing, and that systematic patterns emerge from individual behaviors.

In that way, we're not all that different than John Snow, who paced off his Voronoi boundaries on the map of Soho.

University of Iowa Hospitals and Clinics

The main University of Iowa Hospitals and Clinics (UIHC) complex has 3.2 million sqft on 9+ floors, and is over 0.3 miles long along its major axis. The new 14 floor UIHC Children's Hospital added another 0.5 million sqft.

University of Iowa Hospitals and Clinics

The main University of Iowa Hospitals and Clinics (UIHC) complex has 3.2 million sqft on 9+ floors, and is over 0.3 miles long along its major axis. The new 14 floor UIHC Children's Hospital added another 0.5 million sqft.

Construct a graph model consisting of roughly uniform length edges, with rooms as nodes, and edges representing room adjacency (larger rooms and corridors were segmented into approximately room-size chunks).

University of Iowa Hospitals and Clinics

The main University of Iowa Hospitals and Clinics (UIHC) complex has 3.2 million sqft on 9+ floors, and is over 0.3 miles long along its major axis. The new 14 floor UIHC Children's Hospital added another 0.5 million sqft.

Construct a graph model consisting of roughly uniform length edges, with rooms as nodes, and edges representing room adjacency (larger rooms and corridors were segmented into approximately room-size chunks).

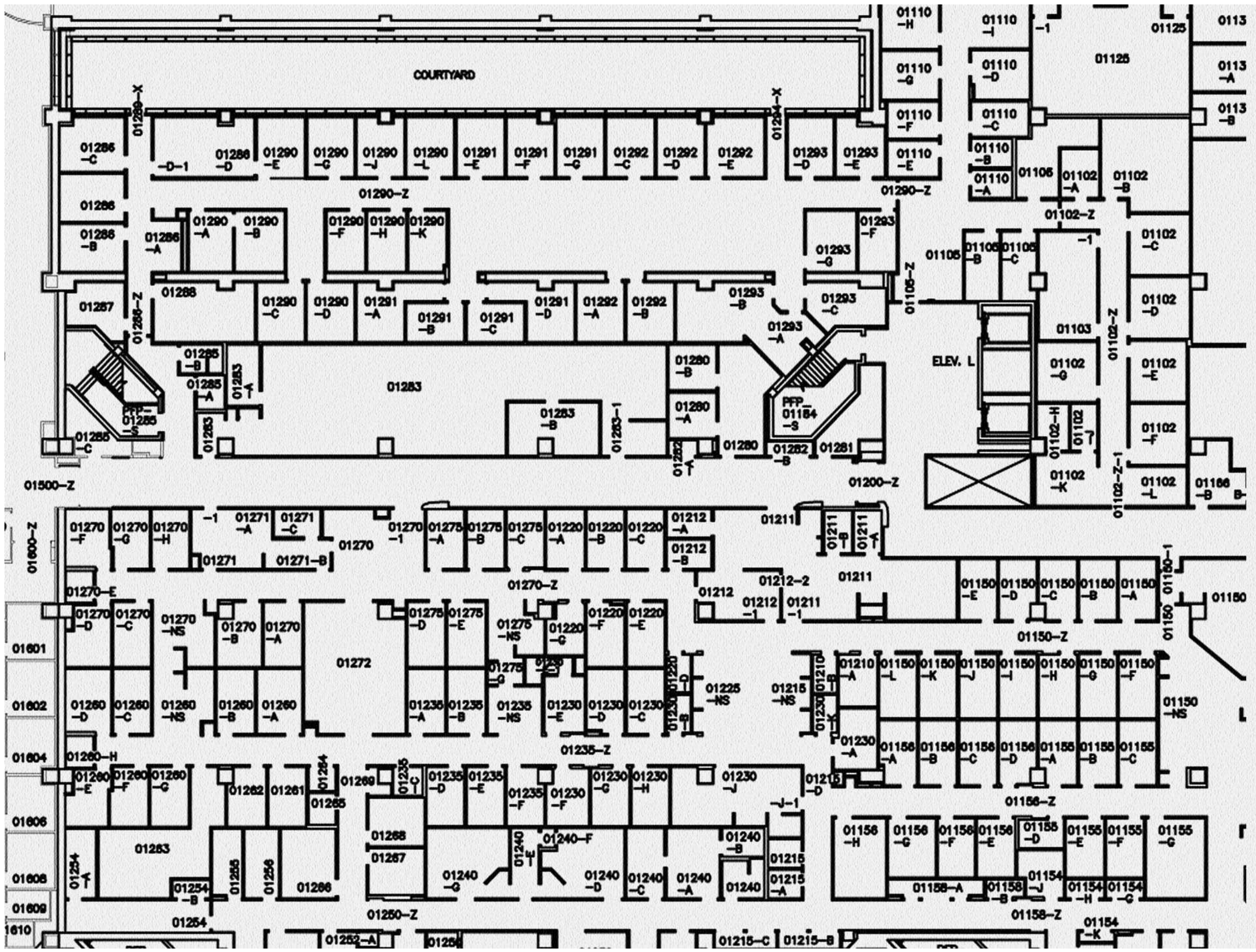
Our UIHC model provides a high-resolution spatial model of proximity and accessibility, consisting of 19,554 nodes and 23,556 edges representing 3.2 million square feet. We also precomputed and cached all 382,339,362 room-to-room shortest paths.

UIHC Model

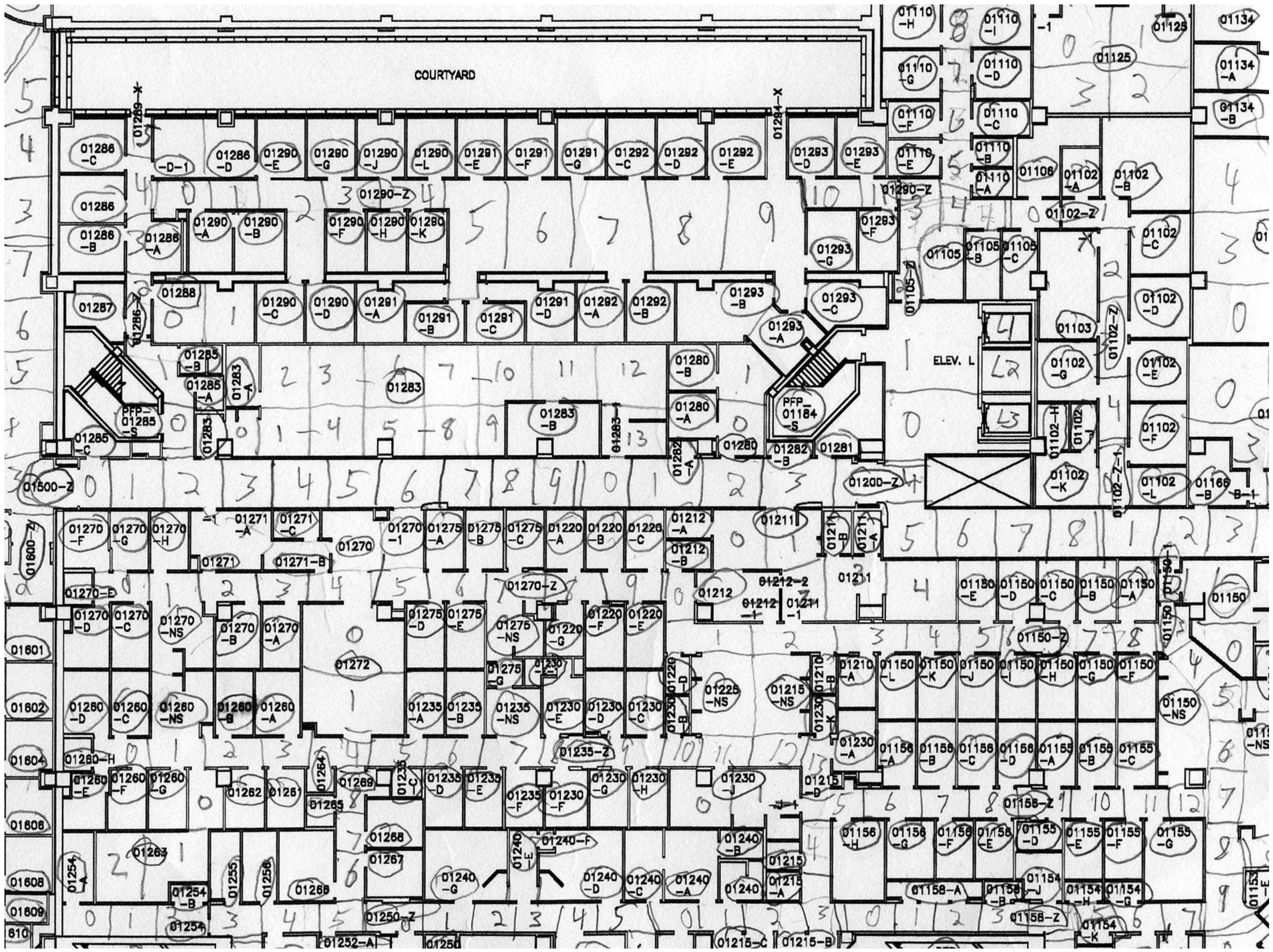
University of Iowa
comp|epi
computational epidemiology research

UIHC Model

UIHC Model

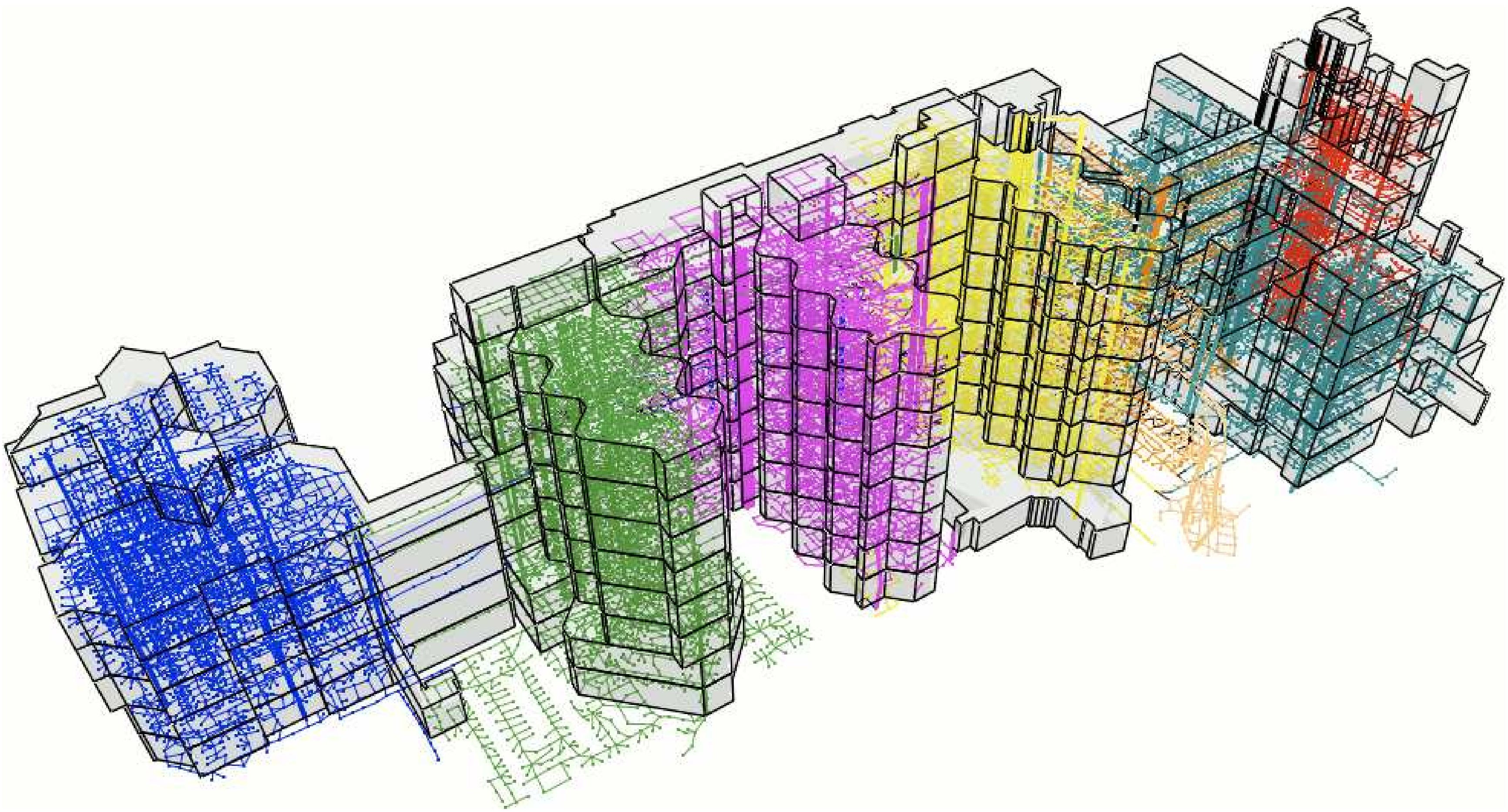


UIHC Model



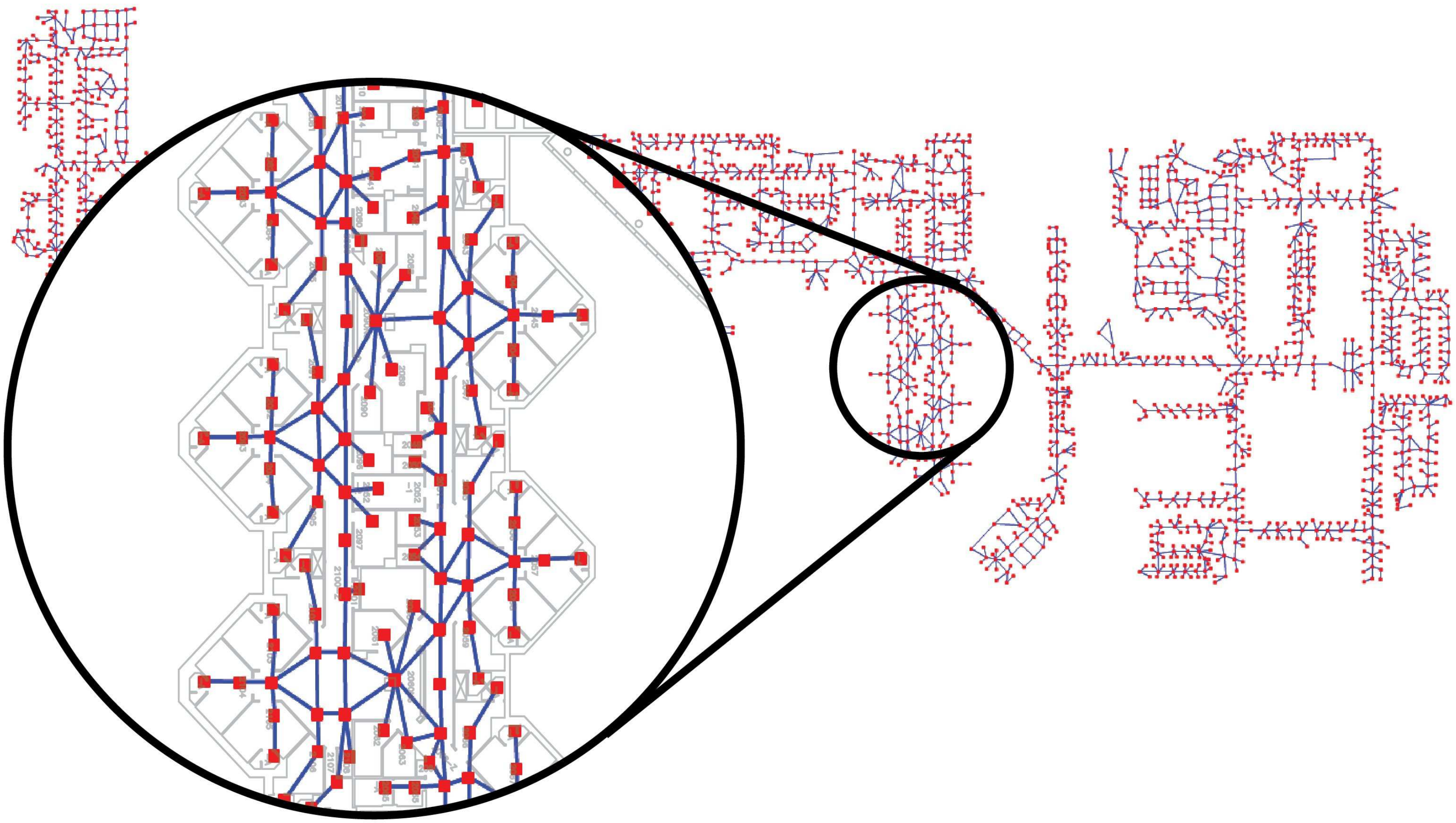
University of Iowa
comp|epi
computational epidemiology research

UIHC Model



University of Iowa
comp|epi
computational epidemiology research

UIHC Model



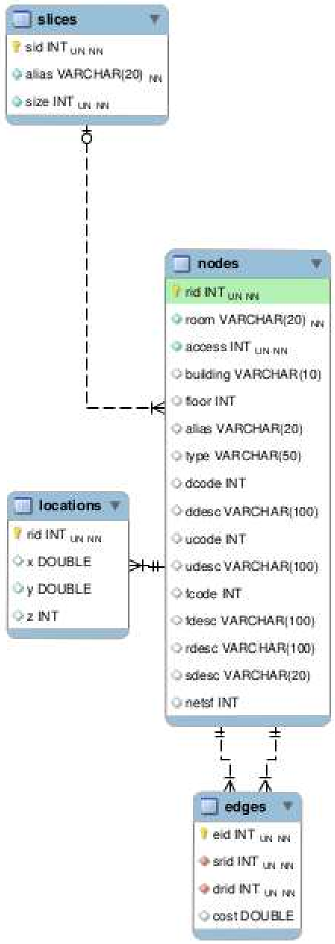
University of Iowa
comp|epi
computational epidemiology research

UIHC Physical Model

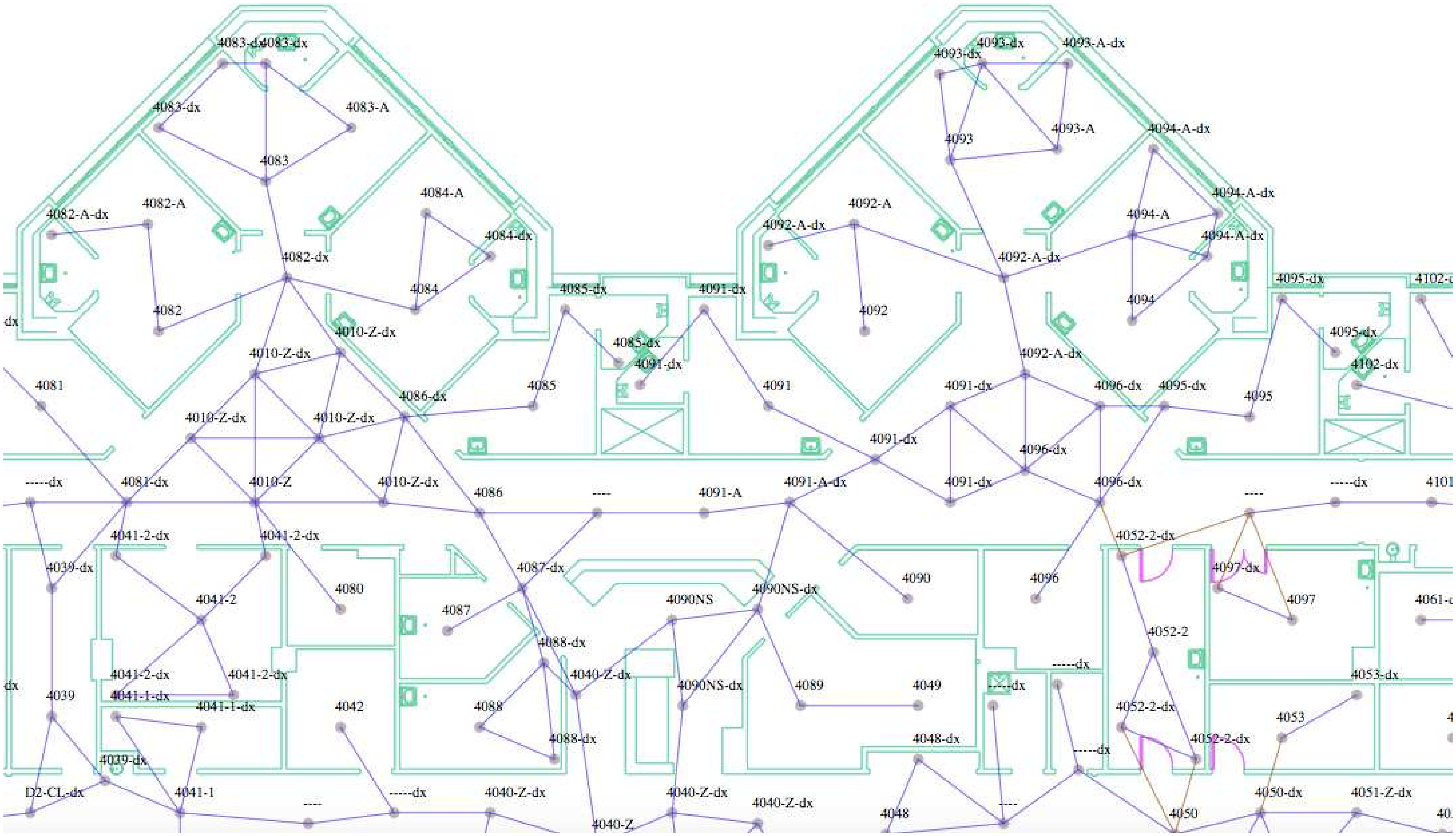
Final graph contains 19,554 nodes and 23,556 edges.

Provides high-resolution spatial model of proximity and accessibility.

UIHC Physical Model



Digression: Automated Graph Extraction from CAD Files



University of Iowa
comp|epi
computational epidemiology research

Add HCW Logins to Electronic Medical Record

Combining the spatial model with de-identified Electronic Medical Record (EMR) login records for the 22 months between 9/1/2006 and 6/21/2008 yields insight into HCW movement.

records	days	users	job types	departments	devices	locations
19.8 million	660	14,595	404	80	17,522	4,379

login date & time	logout date & time	device	location	user ID	job type & department
2006-09-01, 0:00:00.40	2006-09-01, 0:24:17.29			A00012	STAFF NURSE I, NURSING
2006-09-01, 0:00:00.43	2006-09-01, 0:00:21.76	M95089	JPP 6750	A00029	STAFF NURSE II, NURSING
2006-09-01, 0:00:01.23	2006-09-01, 0:03:55.21			J00023	STAFF NURSE II, NURSING
2006-09-01, 0:00:02.29	2006-09-01, 0:00:14.81	MA1458	RCP 1100	C00112	HOUSE STAFF III, NEUROLOGY
2006-09-01, 0:00:02.54	2006-09-01, 0:46:37.82	B71118	RCP 1047	M00018	HOUSE STAFF I, ETC

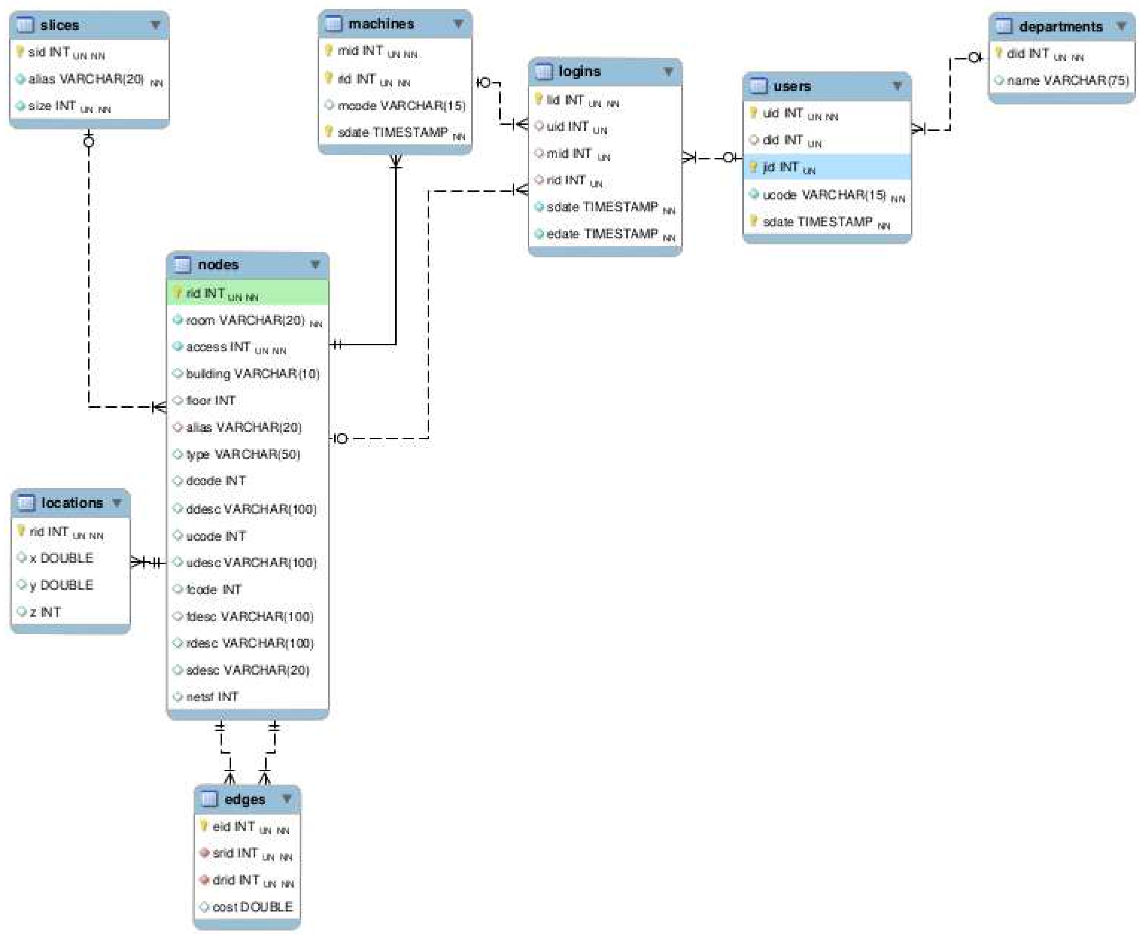
UIHC Login Records

19,800,955 login records between 9/1/2006 and 6/21/2008.

22,996 machines.

29,862 users in 91 departments.

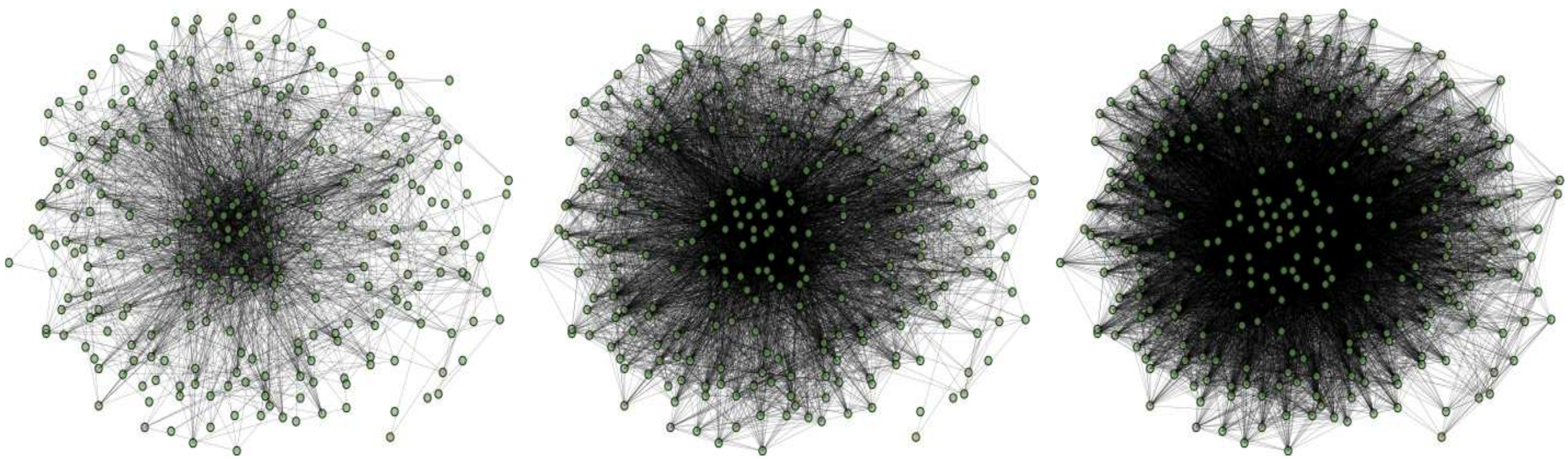
UIHC Login Records



University of Iowa
comp|epi
computational epidemiology research

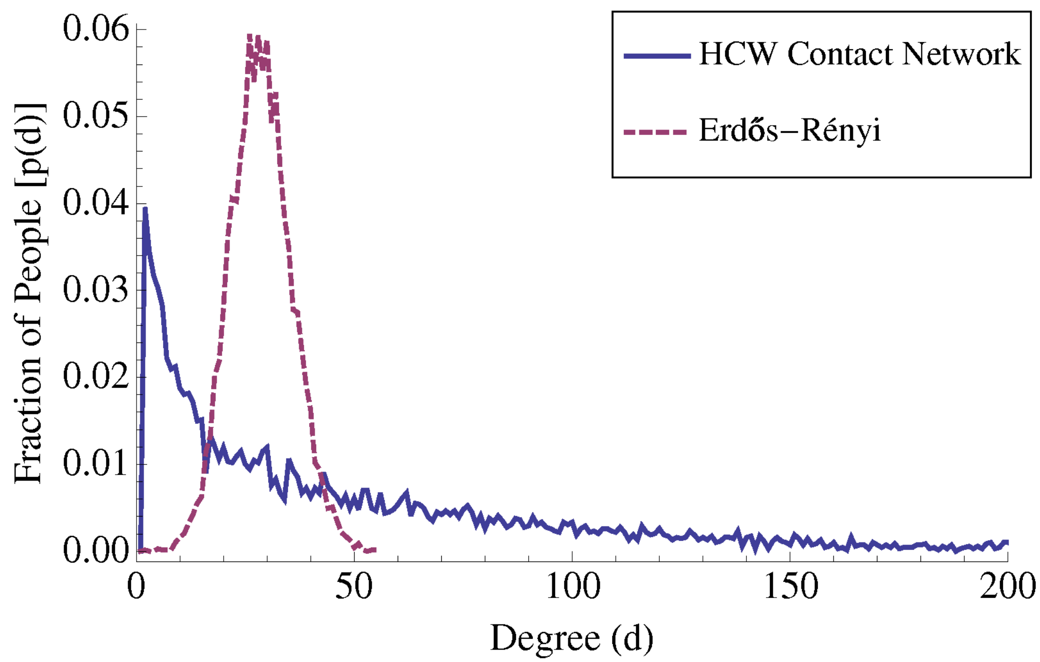
Inferring Contact Networks from EMR Data

We can easily generate contact networks from these EMR data, using *hop distance d* and *time interval t* as parameters.



Example: Three sample 295 node subgraphs (from 6,875 nodes) for a 4-week time window ($d = 1, t = 1$; $d = 3, t = 15$; and $d = 5, t = 30$) starting September 10, 2006.

Observation: HCW Contacts are Heavy Tailed



Compared with an Erdös-Rènyi random graph having same number of vertices ($n = 6,875$), edges/average degree ($m = 174,739$, $\delta = 50.83$) as the inferred 4-week HCW contact graph ($d = 3$, $t = 15$) starting on September 10, 2006.

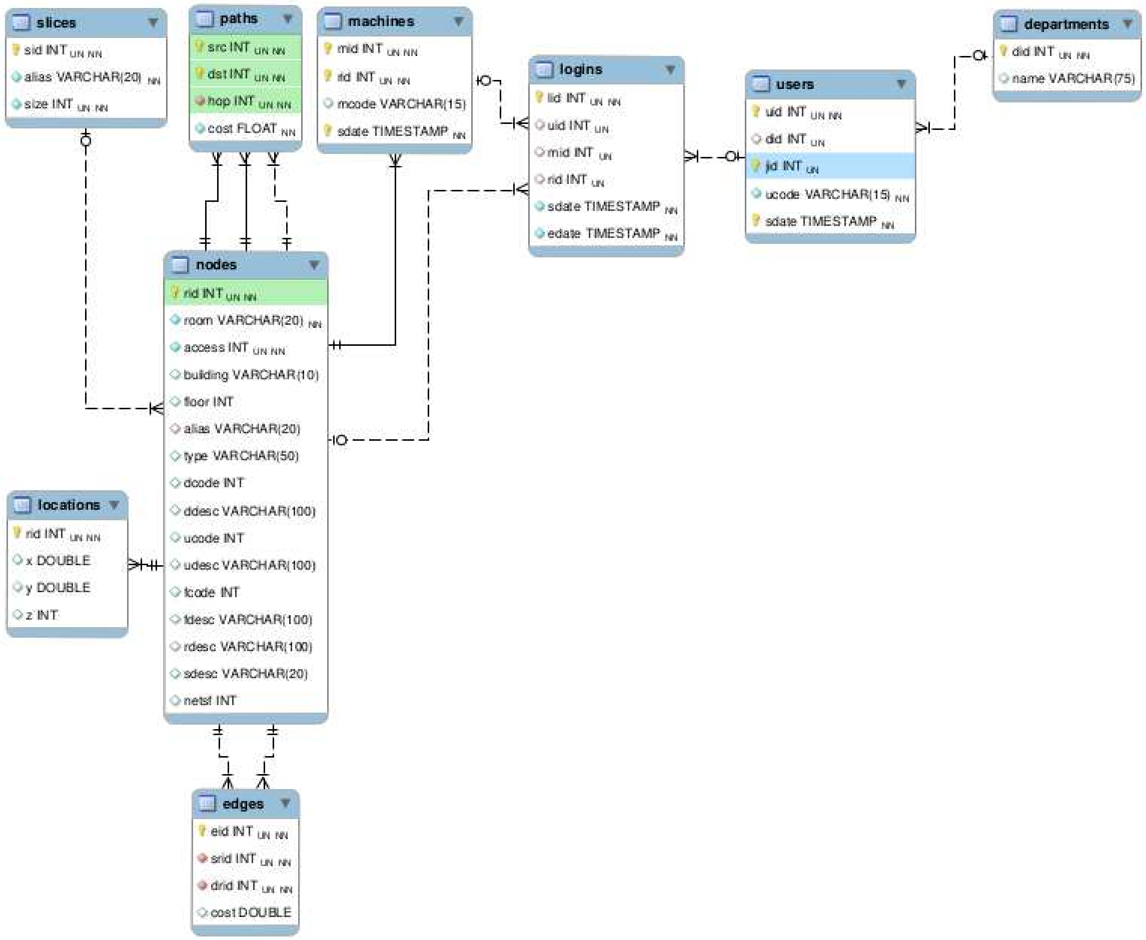
UIHC Connectivity

Precomputed 382,339,362 room-to-room shortest paths.

Cost model:

- Standard edges 1.0
- Corridors 0.8
- Elevators 5.0
- Stair up 4.0
- Stair down 3.0

UIHC Connectivity



University of Iowa
comp|epi
computational epidemiology research

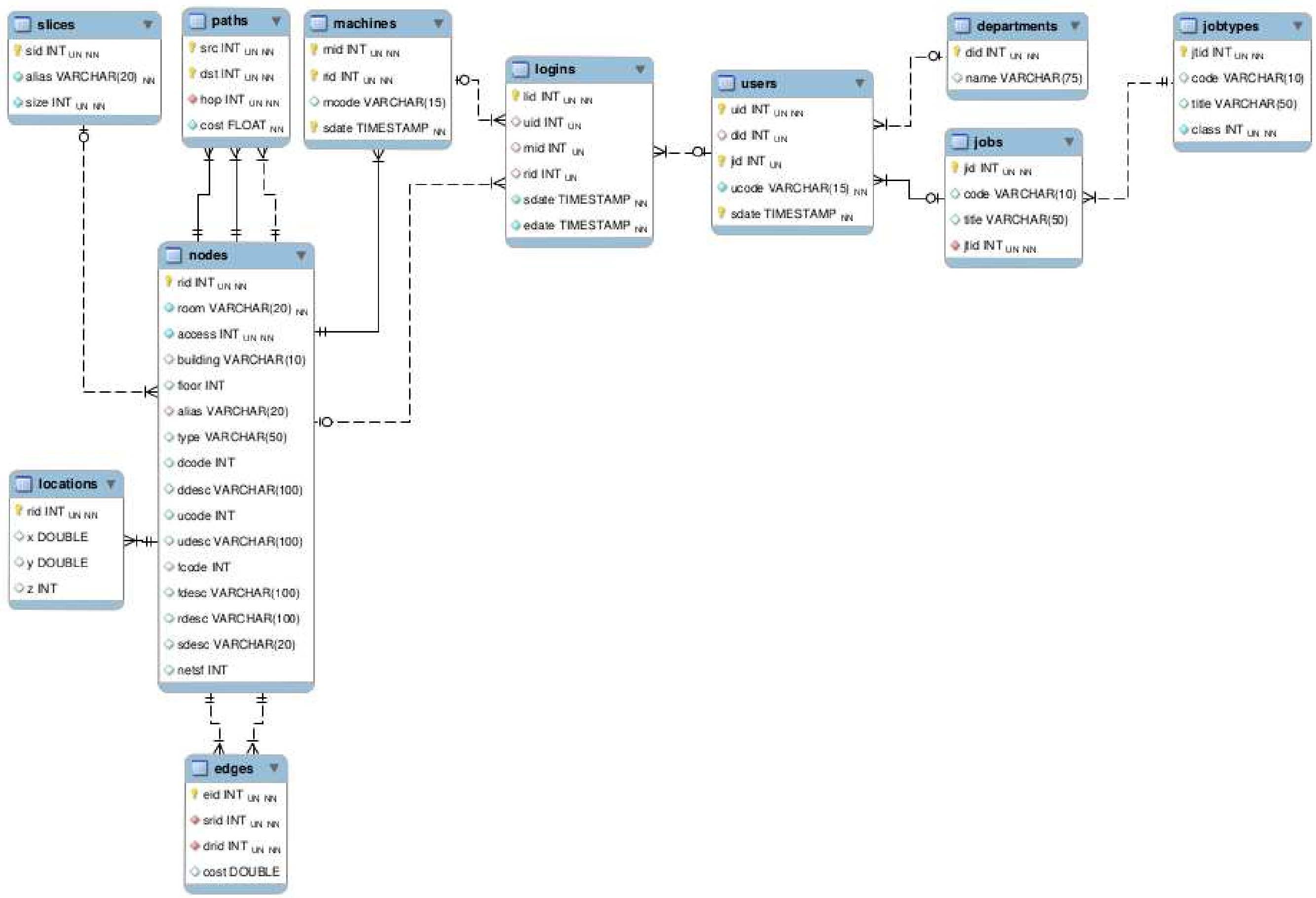
UIHC Human Resource Data

Obtained job catagories and job types from HR.

Used to augment original job data from logins.

477 jobs assorted into 36 HR job types.

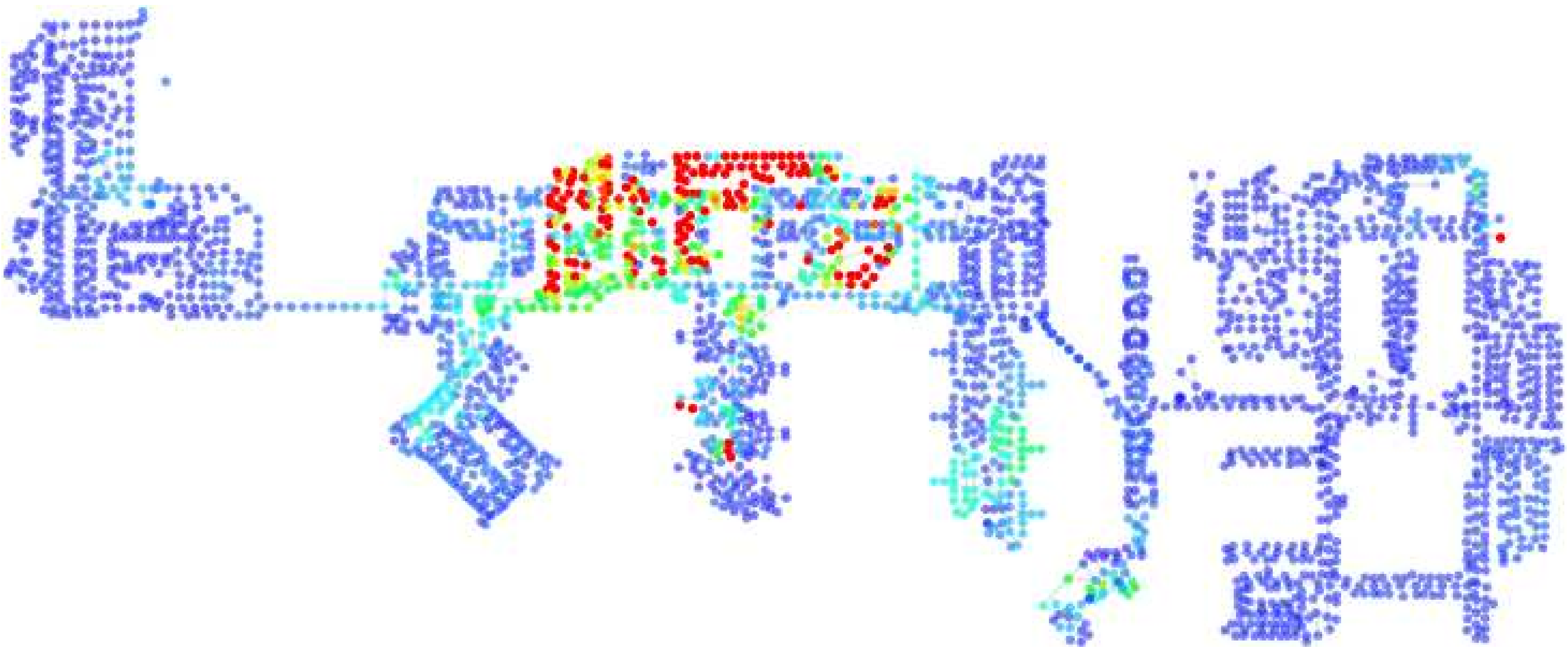
UIHC Human Resource Data



University of Iowa
comp|epi
computational epidemiology research

Observation: These Data Contain Lots of Information

Using EMR login data based on machine location, we investigated how to infer HCW distribution models from EMR login data.



Pediatric staff logins centered in 2nd floor pediatric unit (March 2007).

University of Iowa
comp|epi
computational epidemiology research

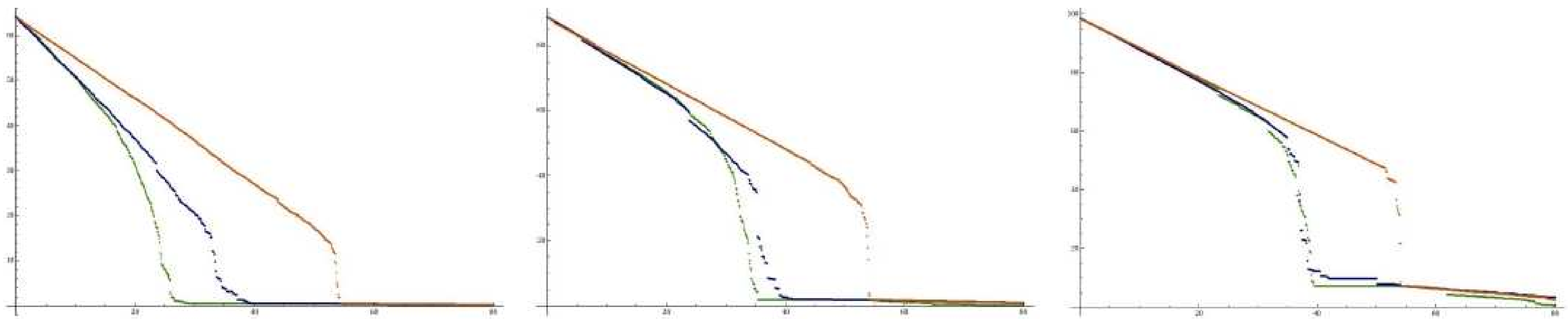
Observation: Not all HCWs are Created Equal

Job category	Average 0.8-radius
CT Service Tech	1.50
Secretary	5.06
Unit Clerk	7.20
Nurse Manager	11.0
Sonographer	13.6
Pharmacy Tech	14.0
Clinical Lab Scientist	16.5
Professor	20.1
Social Worker	21.2
Dietician	21.4
Imaging Tech	25.6
Respiratory Therapist	25.8
House Staff	30.3

Job category	Average 0.8-radius
House Staff I	35.6
House Staff II	31.3
House Staff III	31.8
House Staff IV	25.0
House Staff V	29.6

Mobility varies by job type and seniority, with important implications for disease diffusion.

Application: Who to Vaccinate?



Simulations based on EMR contact networks can be used to inform practical decisions during vaccine shortages.

We proposed a mobility-based vaccination strategy, which approximates omniscient degree-based vaccination strategy and is easy to implement in practice (use a pedometer!).

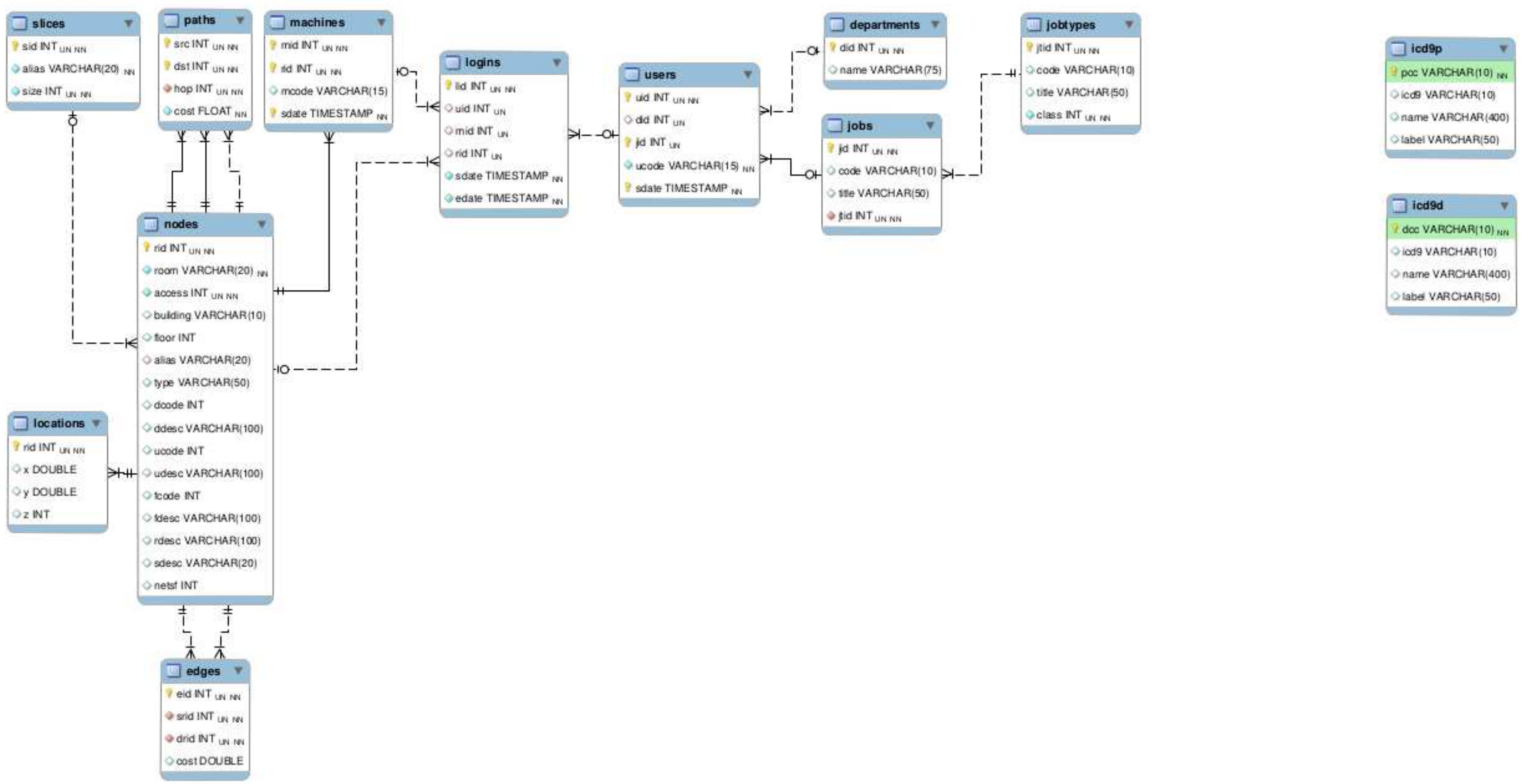
UIHC ICD9 Codes

Obtained ICD9 diagnostic and procedure codes from Medicare.

14,614 ICD9 diagnostic codes.

3,877 ICD9 procedure codes.

UIHC ICD9 Codes



UIHC ADT Data

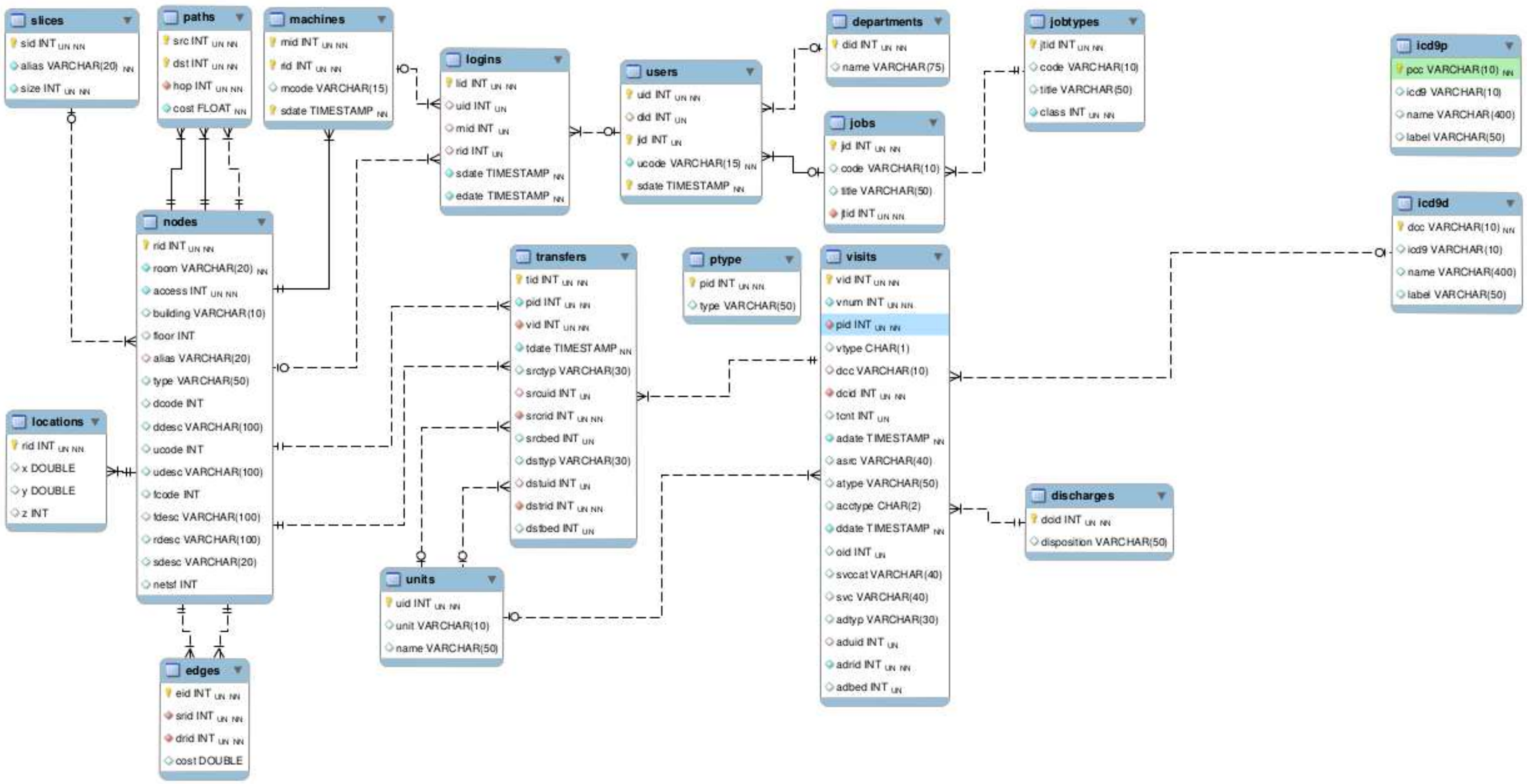
Integrated UIHC admission/discharge/transfer data from 2005-2013.

273,285 inpatient records for 160,322 patients including about 500,000 room transfer entries.

Gives best spatial information about where each patient is when.

Additional information about where they came from (LTC?), where they were discharged from, diagnosis at admission, etc.

UIHC ADT Data

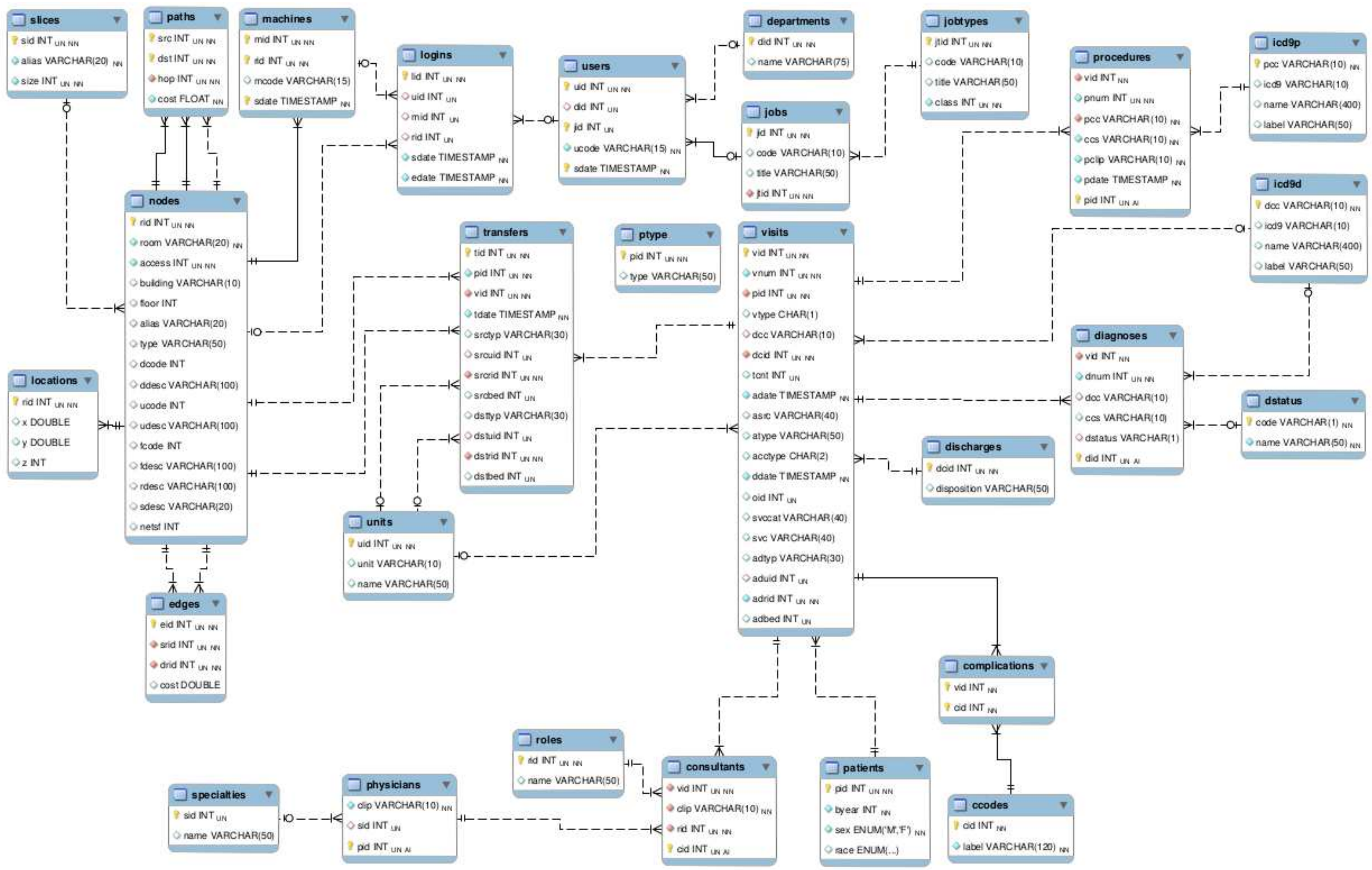


University of Iowa
comp|epi
computational epidemiology research

UIHC Quality Data

The quality data contain much richer information about the individual patient's demographics, diagnoses, procedures, physicians, complications, outcomes and so on.

UIHC Quality Data



University of Iowa
comp|epi
computational epidemiology research

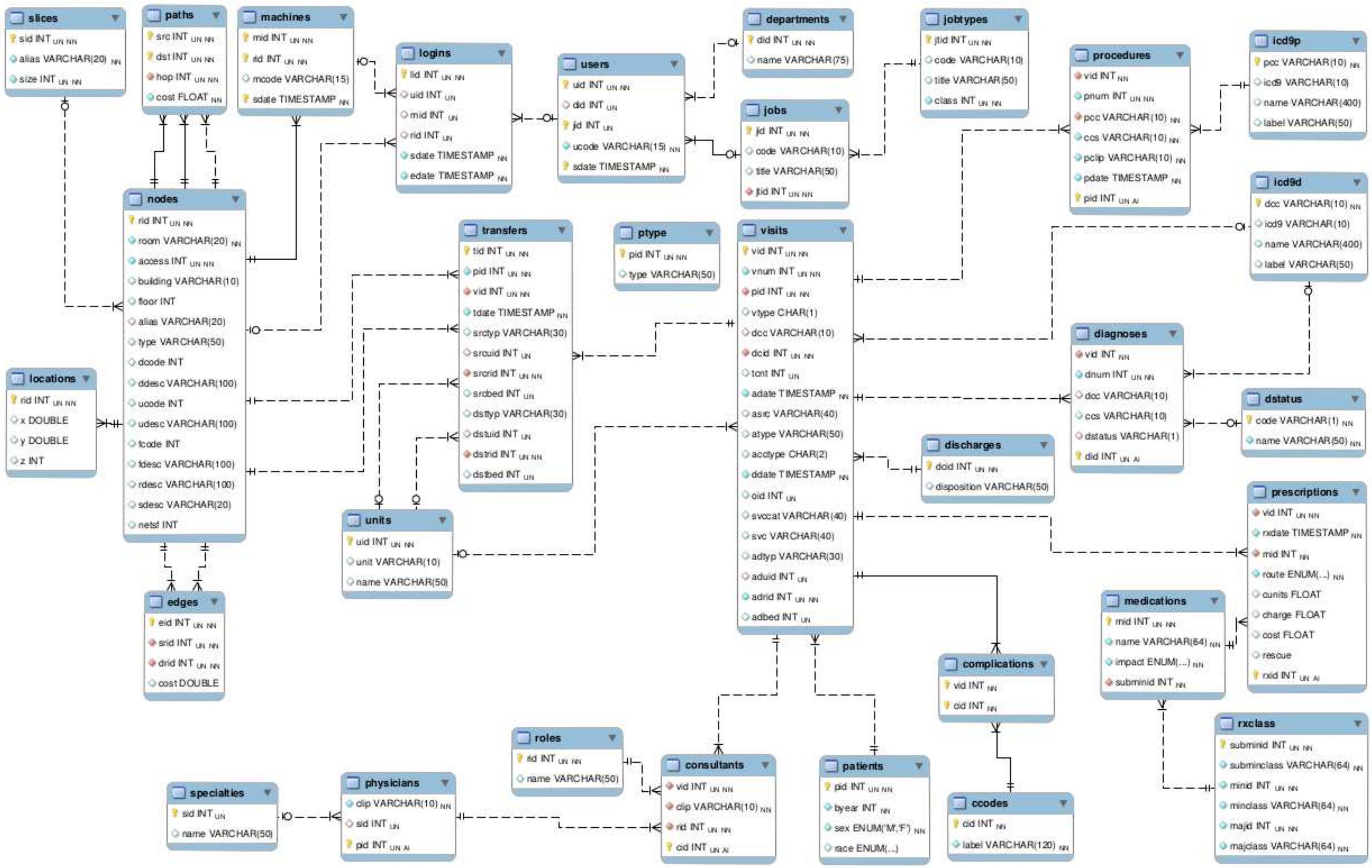
UIHC Pharmacy Data

7,788,703 prescriptions for UIHC patients.

Documents what was prescribed and when/how it was given.

Also contains information about drugs and drug types.

UIHC Pharmacy Data



University of Iowa
comp|epi
computational epidemiology research

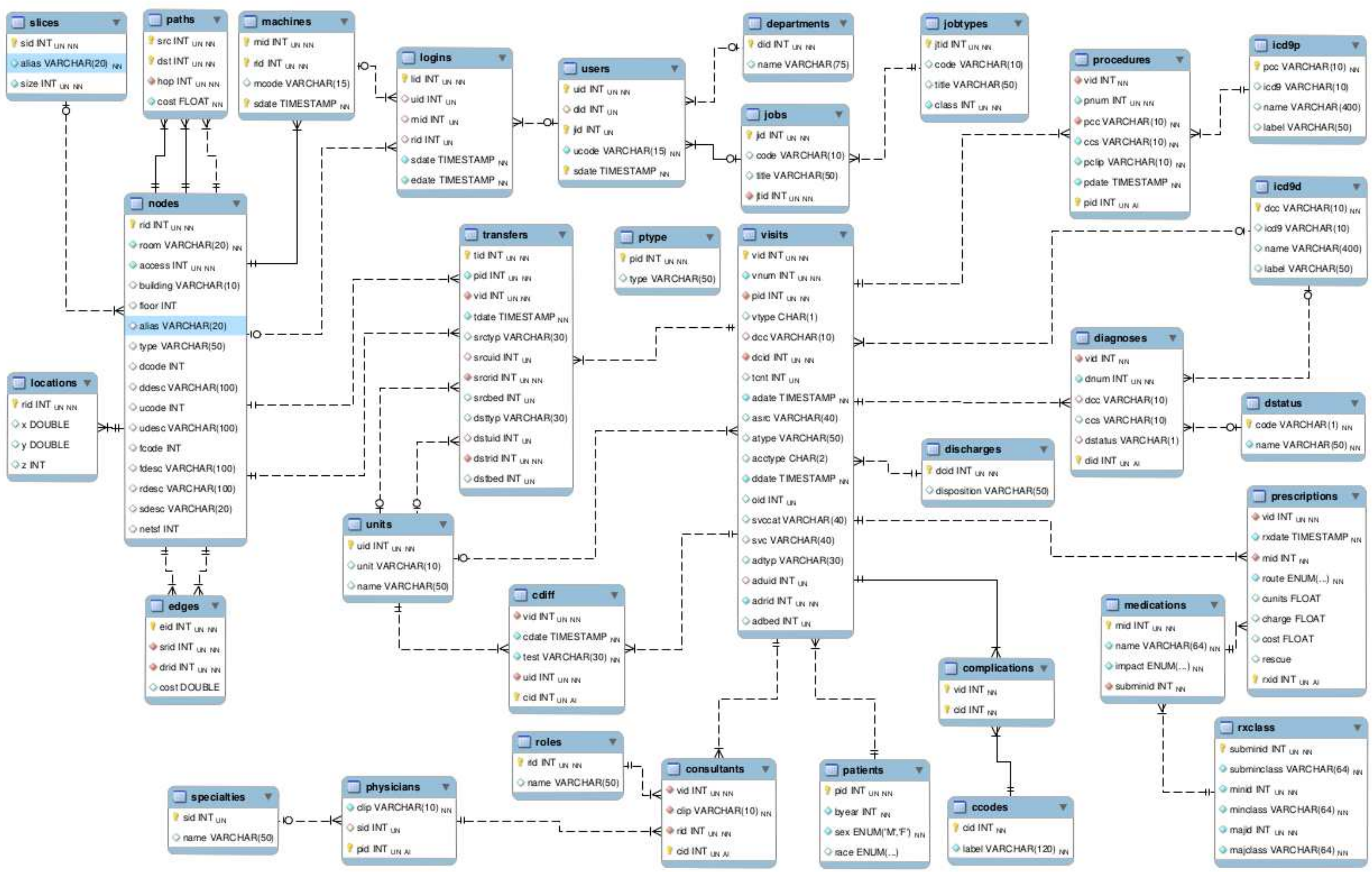
UIHC CDiff Data

Information about CDiff cases at UIHC.

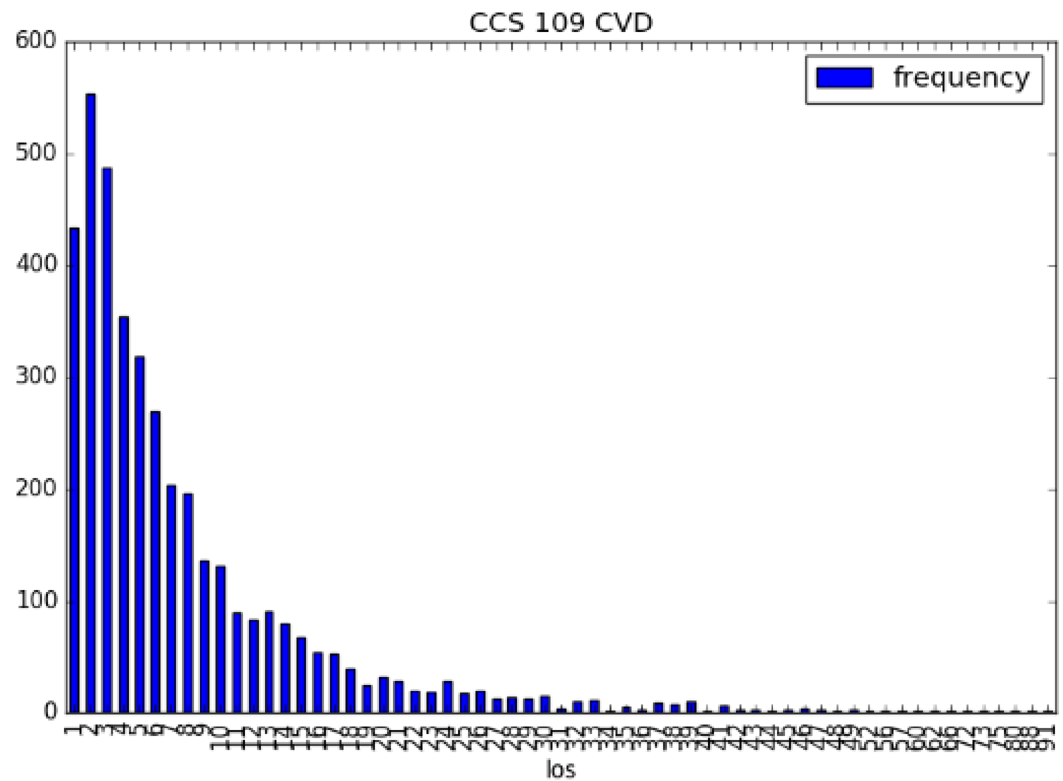
Contains date and time of diagnosis linked to visit.

Also contains which CDiff test was used to confirm diagnosis.

UIHC CDiff Data



Digression: This Rich and Diverse Data Set Supports our Research



With Osteoarthritis, Acute CVD (CCD code 109) is the most common admission diagnostic; 3,992 patients stayed an average of 7.8 days (median 5, max 91).

University of Iowa
comp|epi
computational epidemiology research