Patient Similarity-guided Decision Support

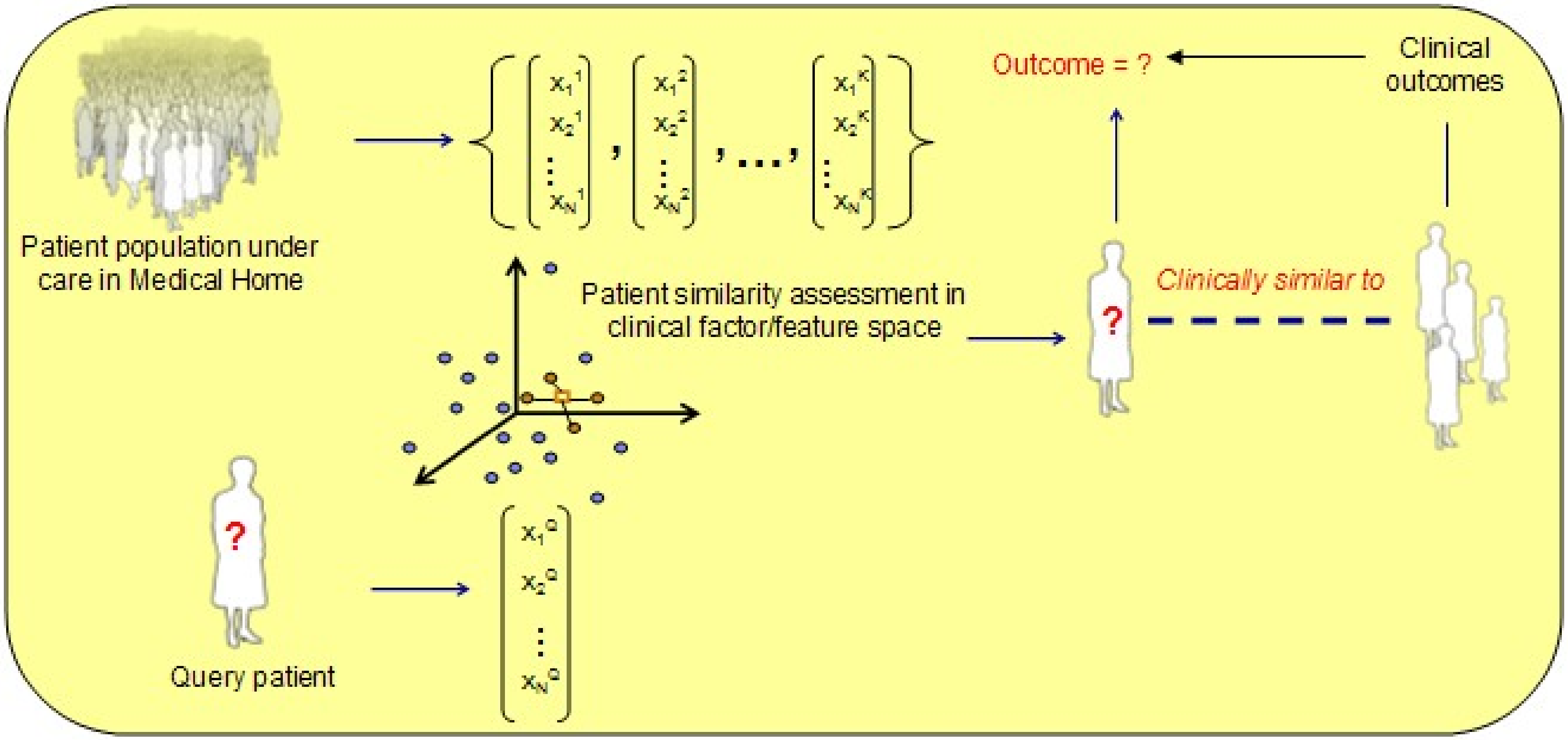
Tanveer Syeda-Mahmood, PhD
IBM Almaden Research Center

What is clinical decision support ?

- Rule-based expert systems – curated by people, inferred by machines for CAD
- Practice guidelines – curated by people and presented by machines
- Rule-based alerts – curated by people, acted upon by machines
- Our approach – statistical decision support from EHR data combined with computable practice guidelines
  - A scalable way to leverage the knowledge in electronic health records
  - Personalized decision support
  - Developed in the domain of cardiology in consultations with cardiologists
  - Test deployed at the catheter lab at Kaiser Permanente's San Francisco Medical Center
    - 3000 patients
  - Running in production at Cedars-Sinai Medical Center for the use case of similar patient cohort retrieval for clinical studies.
    - 1 million patients
  - Patient similarity technology in at least 2 IBM software products/solutions

Patient similarity for clinical decision support

Similar clinical data => similar patients => infer similarity in diagnosis, treatments and outcomes



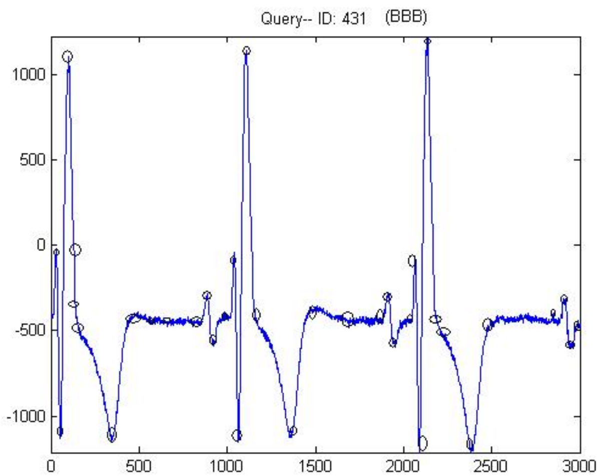
Patient similarity algorithms

- Early fusion approaches using fully-supervised methods
  - Combine normalized features from all modalities into a single vector per sample (patient)
  - Given labeled pairs of similar feature vectors (by clinicians), compute pair-wise distance between samples.
  - Learn a metric in a projected space that ideally separates those vectors from the different similarity groups but groups those that are similar closer in this space
    - Metric learning can be based on many distance measures
      - Mahalanobis distance-based metric learning
      - Information-theoretic metric learning.
- Pros/Cons
  - Pros: Simple to model and use conventional machine learning framework
  - Cons: Cannot handle missing and spurious data well
    - Addition of new data means the metric has to be learned all over again - incremental update of learned distance metric difficult.
  - Computationally complexity and memory requirements can be excessive. Limitation of the size of the distance matrices that can be loaded.
  - Manually intensive as it requires clinicians to compare patients pairwise.
  - Not demonstrated for scalability so far.

Patient similarity algorithms

- Late fusion approaches (semi-supervised)
  - Compute similarity in each modality using separate distance metrics best suited for their respective domain data.
    - Produce separate ranked list of similar patients per modality
  - Fuse ranked lists from modalities to do an overall ranking of the patients
  - Weak supervision by labeling the data based on diagnosis. Use diagnosis as a key element of patient similarity.
    - The diagnosis are either in structured records or can be extracted from reports.
- Pros/Cons
  - Pros: Works even if not all modalities are present for a given patient.
    - Allows for customization of the metric as per modality and clinical content.
      - What makes two EKGs similar is not the same as two echo videos similar.
    - Scalable way to collect ground truth from reports and structured data and doesn't need pairwise similarity comparisons to be made by clinicians.
  - Cons: Doesn't handle time-sensitive aspects.
    - Similarity in modalities have to be over the same period of time in order to be mutually reinforcing
      - Patients who have similar EKGs and similar echocardiograms over a consistent period in time are more similar than those that match at different points in time.

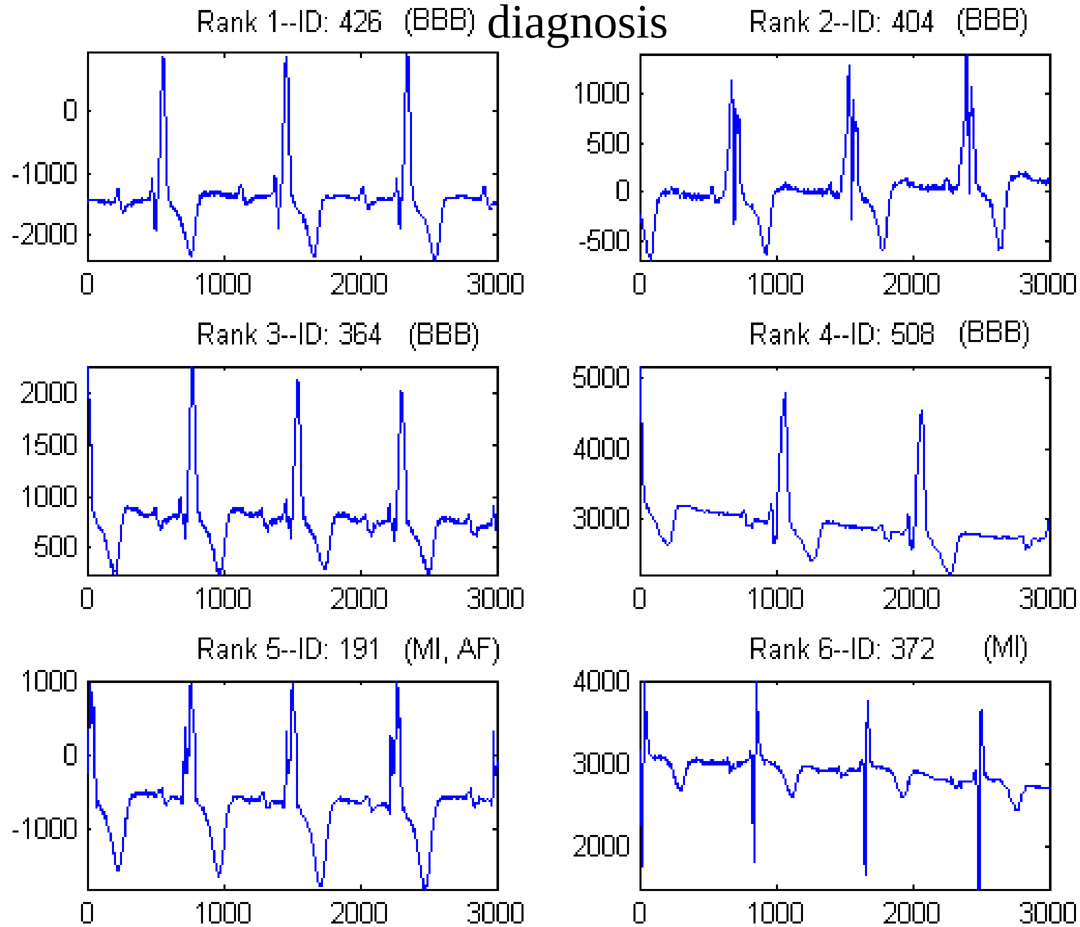
EKG similarity



Query patient EKG

- **EKG similarity search**
  - Invariant to heart rate and signal strength. Computes it automatically
  - Captures perceptual similarity of shapes.
  - No extraction of PR, QT and other intervals needed
  - No built-in rules for recognition of patterns.
  - Uses non-rigid shape matching.

Matching patients retrieved based on EKG and their diagnosis



BBB:67% MI: 33% AF:16%

in EMBC'07

Modality-specific similarity: Text similarity

- Extract concepts from textual reports
    - Large vocabulary-driven extraction of diseases, drugs, symptoms, family history, measurements
    - Vocabularies formed from SNOMED CT, LOINC, RxNorm, ICD9, findings from mining millions of reports
    - Negation and family history references filtered.
- Find similar patients using textual similarity based on cosine distance.

```
Age: 71 Gender:F HT:0.00 WT:0.00 BSA:0.00

Referring MD: XXXXXX
Indication: severe MS

--------------------------------------------------
------------

M-Mode Dimensions Diastole Systole Normal Range
RVID                                < 3.0 cm diast.
LVID                                < 5.0 cm diast.

AO                 2.97             < 4.0 cm diast.
IVS                        4.69     < 1.2 cm diast.
LVPW                                < 1.1 cm diast.
EPSS                                < 0.6 cm
FS                                  25 - 53 %
LV EF                               > 55 %

--------------------------------------------------
------------

2D          Diastole Systole Normal Range
RVAW                         0.2 to 0.7 cm diast.

IVS         0.81     1.71    0.6 to 1.1 cm diast.

Mitral Valve: There is marked mitral annular calcification presen
t which causes restriction of mitral inflow. The mitral valve lea
flets are calcified and restricted. Severe mitral regurgitation i
s present. There is systolic flow reversal in the pulmonary vein.

Tricuspid Valve: The tricuspid valve leaflets are normal (with go
od mobility and coaptation) in appearance. Moderate tricuspid reg
urgitation is present. The spectral doppler derived peak systolic
 velocity across the tricuspid valve is 3.48 m/s, indicating a pe
ak right ventricular pressure of 48.4 mmHg above the right atrial

Pulmonic Valve: The pulmononic valve is not well visualized. No p
ulmonic valvular insufficiency is seen.

Pulmonary Artery: There is moderate secondary pulmonary hypertens
ion, with peak PA pressure, estimated from the systolic tricuspid
 gradient, of 53.4 mmHg.

Pericardium: No pericardial effusion seen.
```

Extracted diseases and severity (positive evidence)

mitral regurgitation, moderate pulmonary hypertension, mitral stenosis

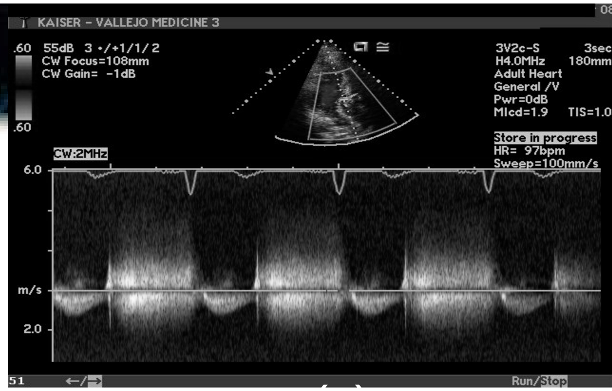
Extracted diseases (negative evidence)

pulmonic valvular insufficiency, pericardial effusion

Extracted measurements

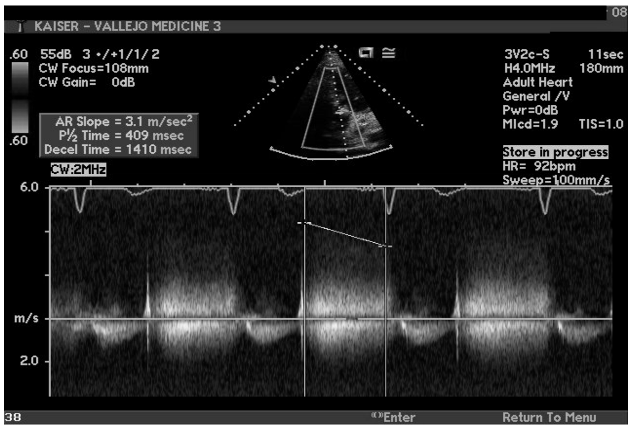
RVID<3.0cm, Age=71, gender=F, etc.

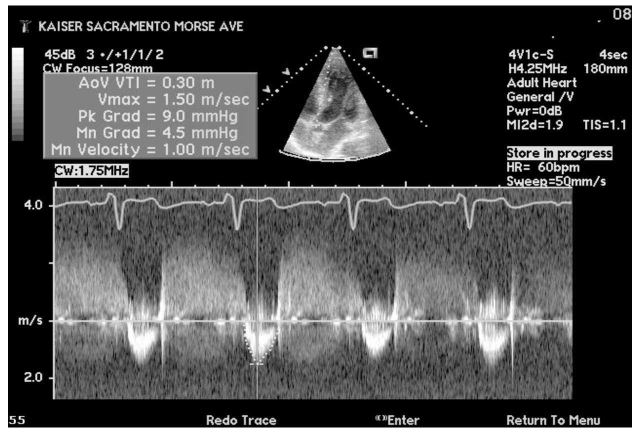
In EMBC '2010

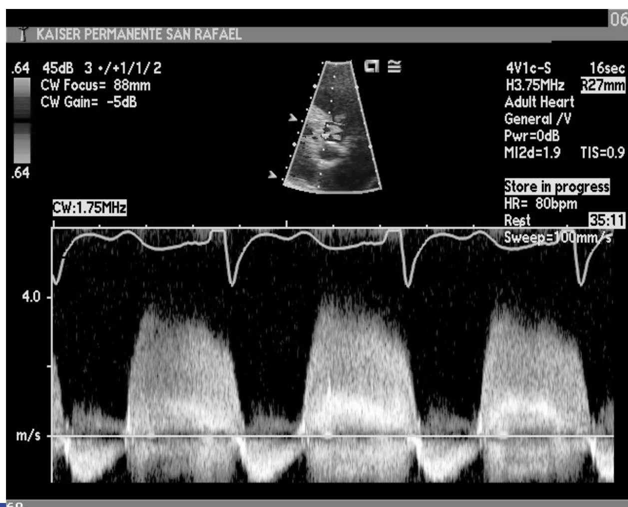


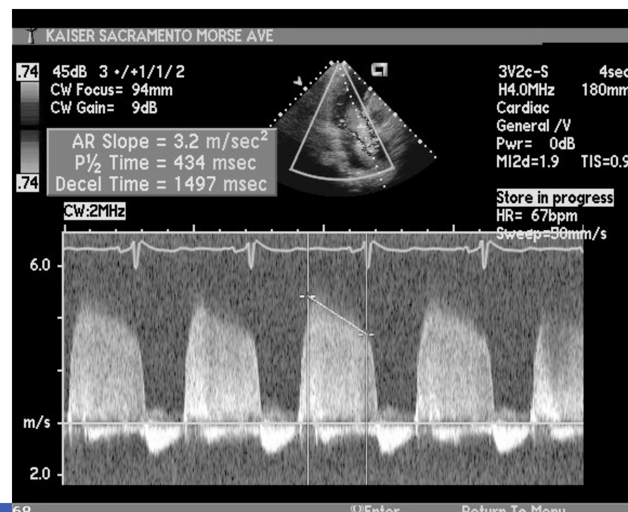
Doppler pattern similarity

CVPR ‘2010







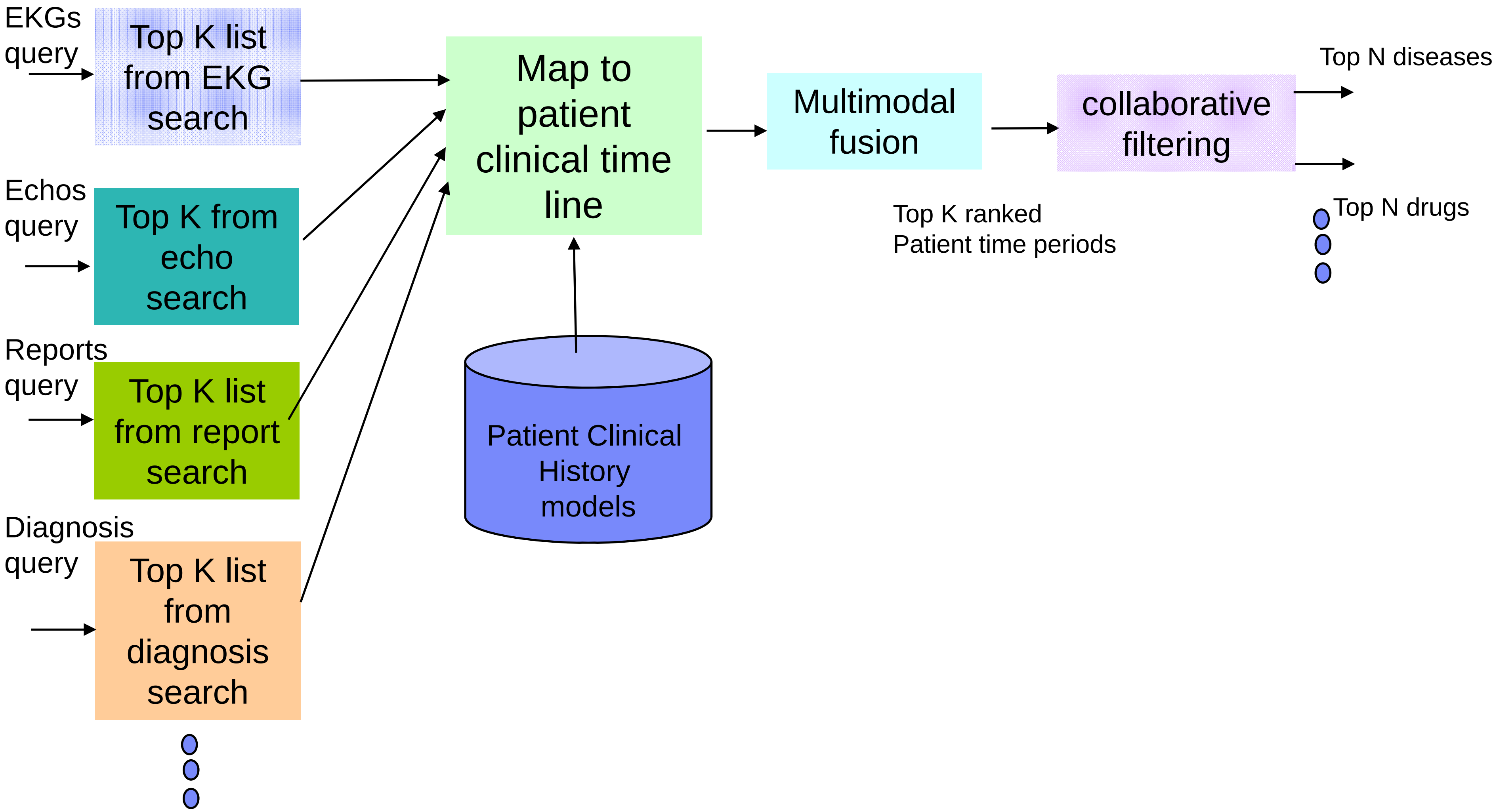


First work to do automatic valvular disease recognition from Doppler

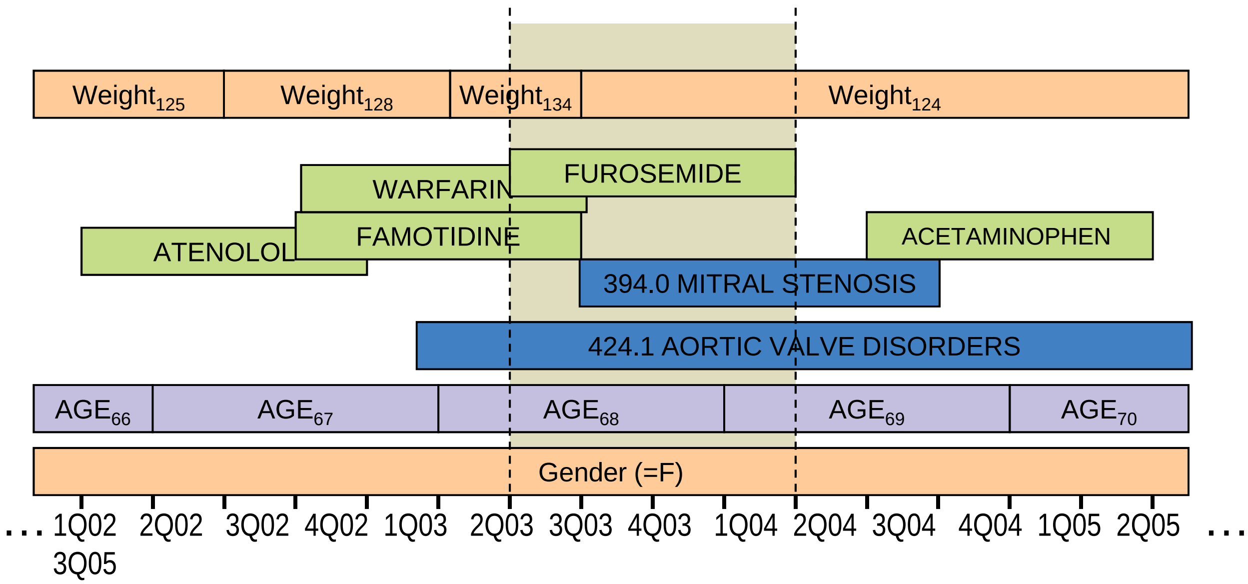
Patient similarity algorithms

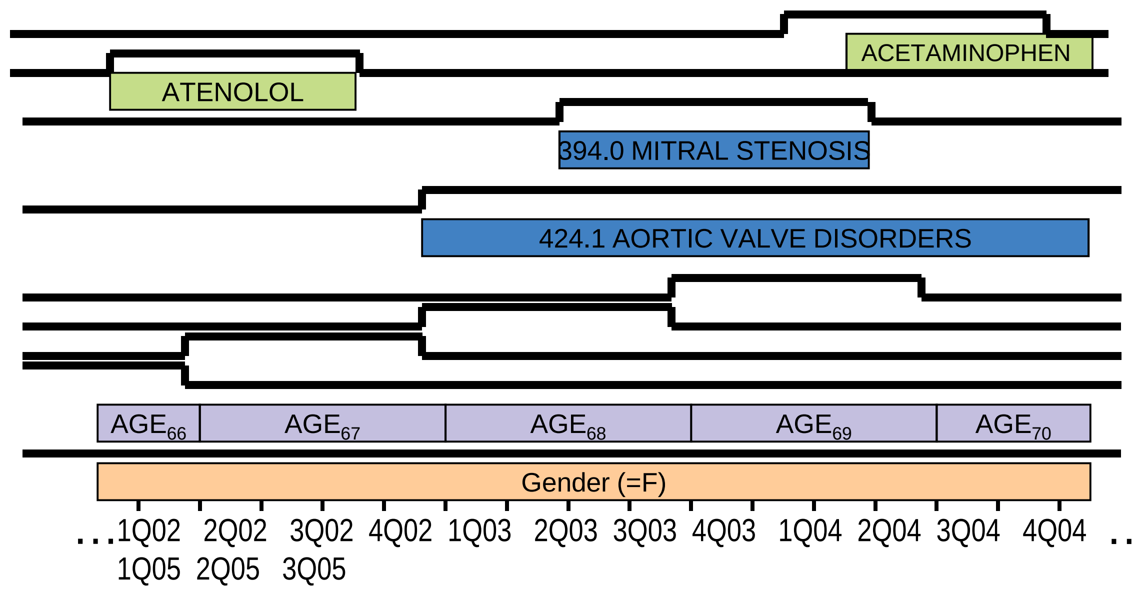
- Multimodal time-sensitive fusion
  - Compute similarity in each modality using separate distance metrics best suited for their respective domain data.
    - Produce separate ranked list of similar patients per modality
  - Model pairwise similarity between patients as a **time-varying similarity function**.
  - Fuse ranked lists from modalities to do an overall ranking of the patients using time-sensitive multimodal fusion
  - Weak supervision by labeling the data based on diagnosis. Use diagnosis as a key element of patient similarity.
    - The diagnosis are either in structured records or can be extracted from reports.
- Pros/Cons
  - Pros: Works even if not all modalities are present for a given patient.
    - Allows for customization of the metric as per modality and clinical content.
      - What makes two EKGs similar is not the same as two echo videos similar.
    - Scalable way to collect ground truth from reports and structured data and doesn't need pairwise similarity comparisons to be made by clinicians.
    - Handles time-sensitive aspects.
  - Cons: Can be difficult to explain.

Multimodal time-sensitive fusion : Overall approach

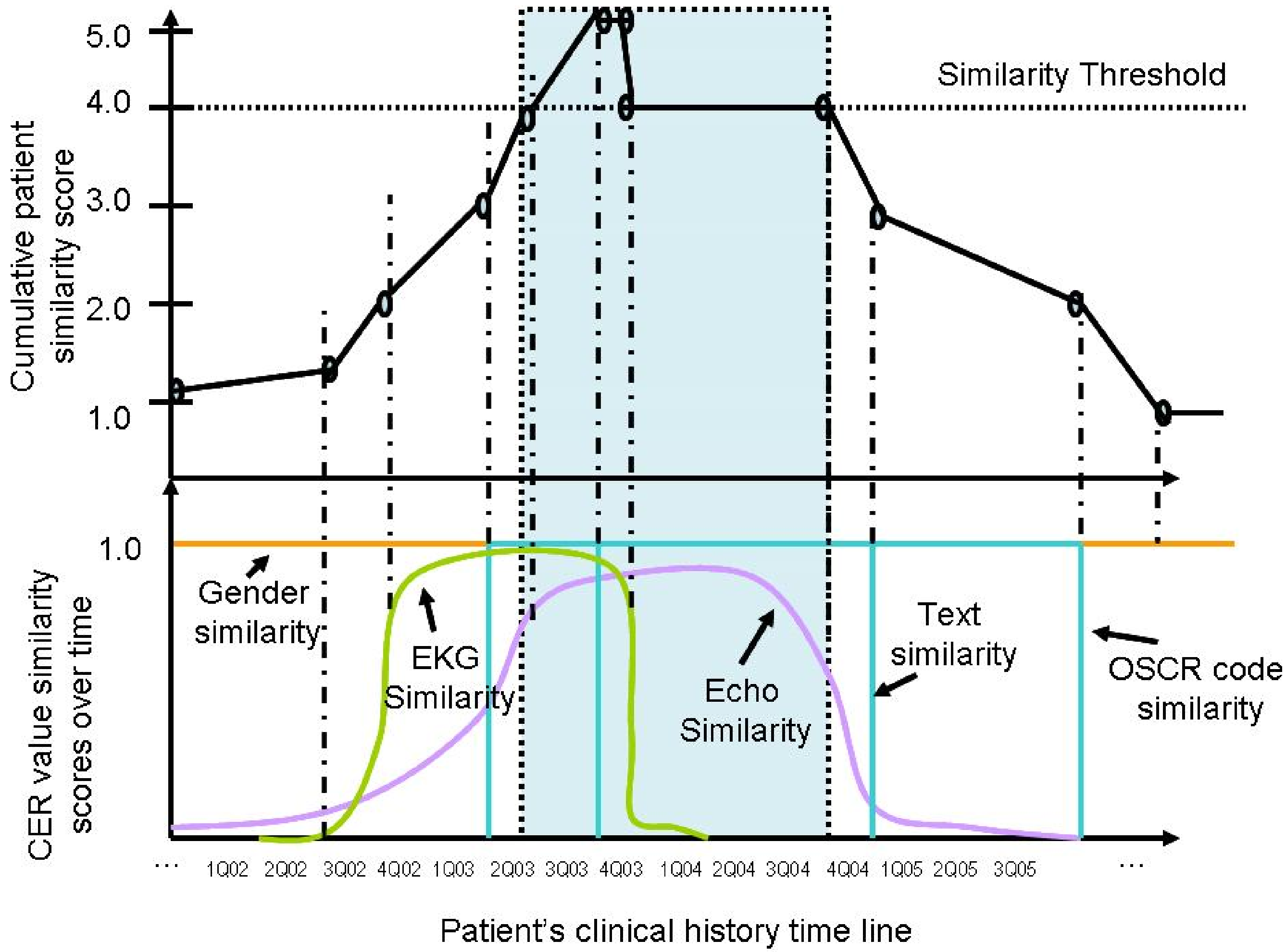


Time-varying modeling in AALIM





Fusing similarity by multi-modal time-sensitive fusion



Each modality is a CER variable

Representing time-varying CER variables for similarity fusion

CER variables : diseases, drugs $V_1, V_2, \ldots V_M$ the set of values taken by a variable $V_i \longrightarrow v_{i1}, v_{i2}, \ldots v_{ik_i}$

A longitudinal patient model (LPM) for a patient P can be denoted by

a set of unit time series

$$LPM(P) = \{S_P(v_{ik}, t) \mid S_P(v_{ik}, t) > 0 \; for \; some \; t, T_{\min}(P) \le t \le T_{\max}(P)\}$$

$$S_P(v_{ik}, t) = \begin{cases} 1 & if \; a \; value \; of \; v_{ik} \; is \; recorded \; for \; CER \; variable \; V_i \; for \; patient \; P \; at \; time \; t \\ 0 & otherwise \end{cases}$$

Eg. a diagnosis of cardiomyopathy over a period of time

$$S_P(v_{disease,\,cardiomyopathy}, t) = 1, \; t_1 \le t \le t_2$$

Multimodal fusion algorithm

Suppose a query patient Q has

$$N_Q = |\vec{V}_{Qi} = \{V_i \mid \exists_{v_{ik}} S_Q(v_{ik}, t) > 0 \;\; for\;\; some\; t\}|$$ active variables.

the number of different values exhibited per CER variable in a patient Q

$$n_Q(V_i) = |\vec{v}_{Qi} = \{v_{ik} \mid S_Q(v_{ik}, t) > 0 \;\; for\;\; some\; t\}|$$

For example, a patient diagnosed with mitral stenosis, hypertension, and aortic stenosis and on medications : warfarin, furosemide, atenolol and amoxicillin during a 10 year period would have $N_Q = 2 \quad n_Q(V_{disease}) = 3 \quad n_Q(V_{drugs}) = 4$

Similarity in the values of a CER variable V_i

exhibited by a candidate patient P and a specific query patient Q can be denoted by $d_{PQ}(v_{ik}, v_{il})$

For each candidate patient P and CER variable V_i

we identify all time points in the patient's time span where a match to one of the values of the query patient Q for variable V_i exists

Multimodal fusion algorithm

For each value v_{ik} of the variable V_i in the set $\vec{V}_{Qi}$ possessed by the query patient Q, we can record all time periods in the timeline of a candidate patient P where there is a match to this value as a function $m_{PQ}(v_{ik}, t)$

$$m_{PQ}(v_{ik}, t) = \begin{cases} d_{PQ}(v_{ik}, v_{il}) & if\ 1 \geq d_{PQ}(v_{ik}, v_{il}) \geq \rho\ and\ S_P(v_{il}, t) = 1 \\ 0 & otherwise \end{cases}$$

Then the extent of match of a patient P to a query patient Q based on the CER variable V_i as:

$$C_{PQ}(V_i, t) = \frac{\sum_{v_{ik}} m_{PQ}(v_{ik}, t)}{n_Q(V_i)}$$

For all CER variables $\vec{V}_{Qi}$ exhibited by Q as

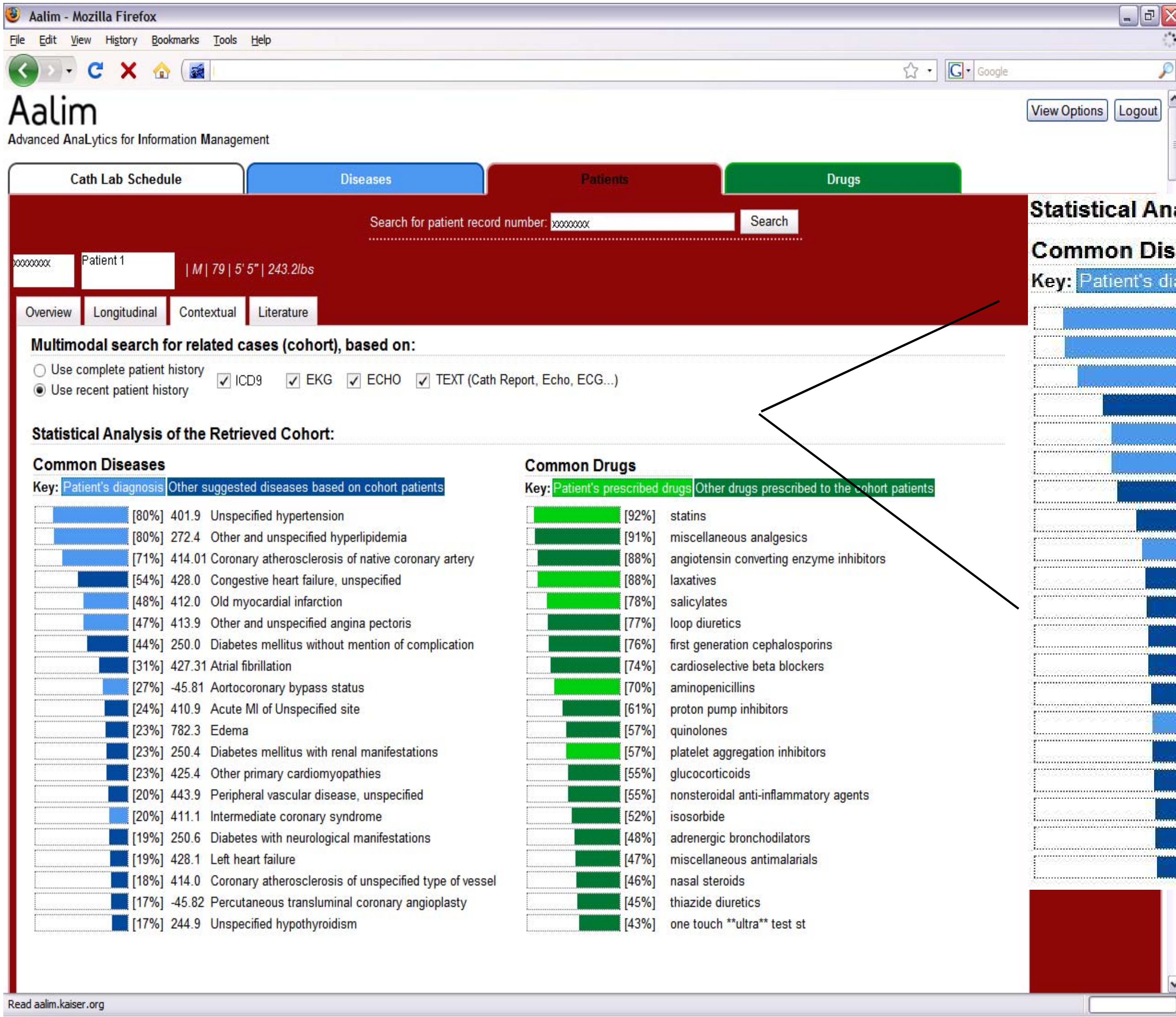
$$U_{PQ}(t) = \frac{\sum_{V_i} C_{PQ}(V_i, t)}{N_Q}$$

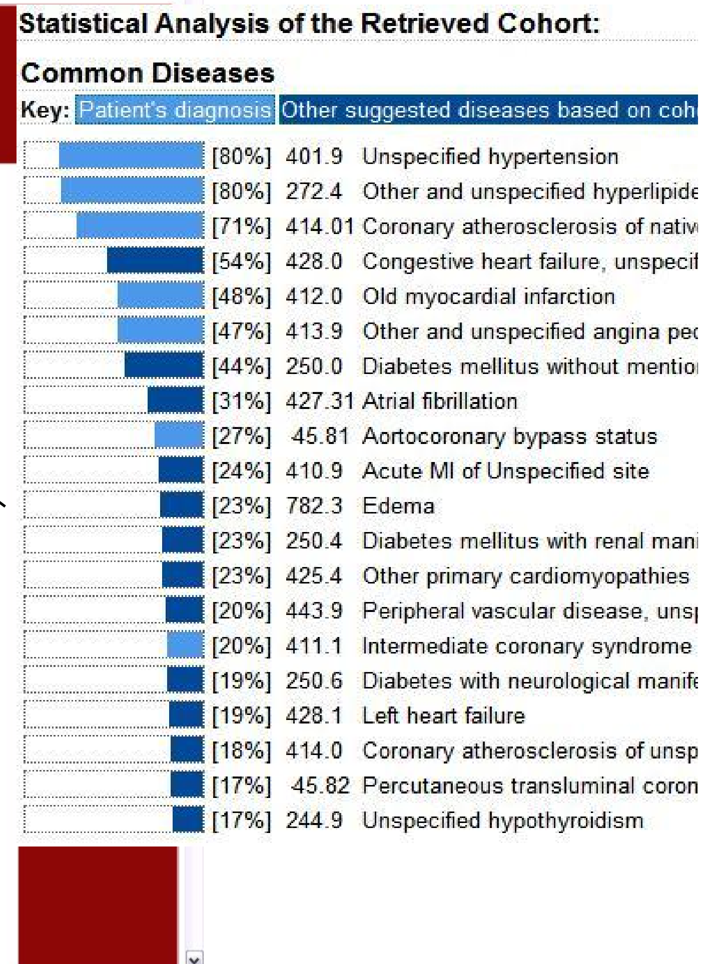
We can now form a patient cohort

$$H(Q) = \{P \mid U_{PQ}(t) > \lambda\ for\ some\ t\}$$

Using similarity by multimodal fusion for clinical decision support

Query patient had:
Hyperlipidemia
Hypertension
Atheroscelorosis





Results

- Patient data set:
  - 1996 patients
  - 12 channel EKG time series: 25,016
  - Echo-cardiographic sequences: 5346
  - CW Doppler images: 34,540
  - Textual reports: 100,042
- Evaluation of Patient similarity by multimodal fusion:
  - Automatic: Comparing against ground truth data about the patients' diseases.
  - Manual: Validation of the top K lists by clinicians.
    - Clinician examines clinical record of the patient
    - Clinician annotates the top K lists returned by the fusion algorithm as

F	factual (present for this person)
NF	non factual
R	Relevant (to consider with similar cardiac patients)
NR	Non-relevant
U	Unlikely, unrelated to patient's cardiac condition

Evaluation of top K disease lists returned by patient similarity

- Given top K matches to a query Q with disease label: $L_Q = (l_{q1}, l_{q2}, .. l_{qN})$
- Let the distribution of disease labels among the top K matches be:

$$L_M = (l_1, l_2, .. l_M)$$

- Metrics used:
  - Recall: Fraction of overlap of the query disease labels with those of the matches

$$\mathrm{Re}\,call = \frac{|L_Q \cap L_M|}{|L_Q|}$$

  - New predictions

$$dis\,\mathrm{cov}\,erability = \frac{|L_M - L_Q|}{|L_M|}$$

  - Valid predictions:

$$validity = \frac{\sum_{l_i \in L_M - L_Q} \max_{l_{qj} \in L_Q}(p(l_i, l_{qj}))}{|L_M - L_Q|}, |L_M - L_Q| > 0$$

Evaluation Studies per modality: EKG

Dataset: 25,990 ECGs

Set of disease labels in the returned matches $L_M = (l_1, l_2, ..l_M)$

Set of disease labels in query: $L_Q = (l_{q1}, l_{q2}, ..l_{qN})$

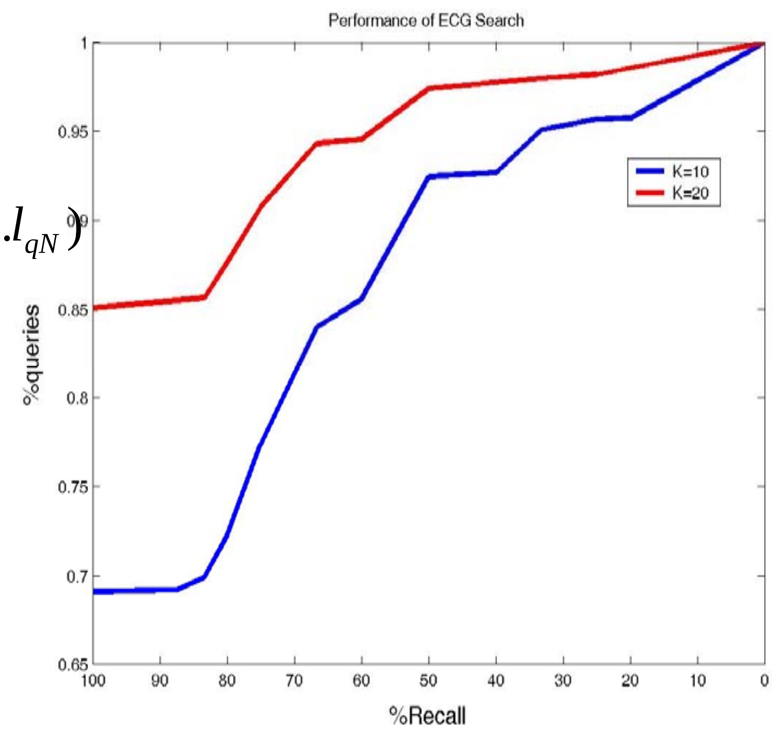
Accuracy: $\mathrm{Re}call = \frac{|L_Q \cap L_M|}{|L_Q|}$

Potential co-morbidity discoveries:

$$discoverability = \frac{|L_M - L_Q|}{|L_M|}$$

Efficacy: :discoverability * likelihood of the co-occurrence of any disease pairs

$$validity = \frac{\sum_{l_i \in L_M - L_Q} \max_{l_{qj} \in L_Q}(p(l_i, l_{qj}))}{|L_M - L_Q|}$$



85% of them were completely consistent with the official human interpretation (100% recall), and nearly all EKGs had at least 50% overlap in their match set with their own disease labels

Evaluation studies on clinical decision support by multimodal fusion

Legend : F: factual (present for this person)	NF: non factual
R: Relevant (to consider with similar cardiac patients)	NR: non-relevant
U: Unlikely, unrelated to patient's cardiac condition	
Patient: 09960-000-07-00007	

272.4Other and unspecified hyperlipidemia (Disorders of lipoid metal	F	NF	R	NR	U
414.01Of native coronary artery (Other forms of chronic ischemic heart disease:Coronary atherosclerosis)	F	NF	R	NR	U
424.1Aortic valve disorders (Other diseases of endocardium)	F	NF	R	NR	U
782.3Edema (Symptoms involving skin and other integumentary tiss	F	NF	R	NR	U
428.0Congestive heart failure, unspecified (Heart failure)	F	NF	R	NR	U
427.31Atrial fibrillation (Cardiac dysrhythmias:Atrial fibrillation and flu	F	NF	R	NR	U
308.9Unspecified acute reaction to stress (Acute reaction to stress)	F	NF	R	NR	U
414.0Of unspecified type of vessel, native or graft (Other forms of chronic ischemic heart disease:Coronary atherosclerosis)	F	NF	R	NR	U
780.4Dizziness and giddiness (General symptoms)	F	NF	R	NR	U
530.81Esophageal reflux (Diseases of esophagus:Other specified di	F	NF	R	NR	U
443.9Peripheral vascular disease, unspecified (Other peripheral vas	F	NF	R	NR	U
424.0Mitral valve disorders (Other diseases of endocardium)	F	NF	R	NR	U
593.9Unspecified disorder of kidney and uretes (Other disorders of k	F	NF	R	NR	U
466.0Acute bronchitis (Acute bronchitis and bronchiolitis)	F	NF	R	NR	U
427.9Cardiac dysrhythmia, unspecified (Cardiac dysrhythmias)	F	NF	R	NR	U
784.0Headache (Symptoms involving head and neck)	F	NF	R	NR	U
412.0Old myocardial infarction ()	F	NF	R	NR	U
780.2Syncope and collapse (General symptoms)	F	NF	R	NR	U
429.9Heart disease, unspecified (Ill-defined descriptions and complic	F	NF	R	NR	U
		NF			

401.9 Unspecified hypertension	F	NF	R	NR	U
427.31 Atrial fibrillation	F	NF	R	NR	U
428.0 Congestive heart failure, unspecified	F	NF	R	NR	U
250.40 250.40 DMII RENL NT ST UNCNTRLD	F	NF	R	NR	U
425.4 Other primary cardiomyopathies	F	NF	R	NR	U
250.60 250.60 Diabetes with neurological manifestations	F	NF	R	NR	U
413.9 Other and unspecified angina pectoris	F	NF	R	NR	U
414.00 Coronary atherosclerosis	F	NF	R	NR	U
272.4 Other and unspecified hyperlipidemia	F	NF	R	NR	U
424.1 Aortic valve disorders	F	NF	R	NR	U
412 Old myocardial infarction	F	NF	R	NR	U
244.9 Unspecified hypothyroidism	F	NF	R	NR	U
V58.61 V58.61 Long-term (current) use of anticoagulants	F	NF	R	NR	U
414.8 Other specified forms of chronic ischemic heart disease	F	NF	R	NR	U
428.1 Left heart failure	F	NF	R	NR	U
424.0 Mitral valve disorders	F	NF	R	NR	U
	F	NF	R	NR	U
	F	NF	R	NR	U
	F	NF	R	NR	U

Clinician annotation of top K lists

Results of manual validation by clinicians of top K lists from multimodal fusion

Patient	Relevant			Non-Relevant	Unlikely/ Unrelated	Total
	Factual	Discovered	Total			
1	3	4	7	0	3	10
2	7	0	7	0	3	10
3	2	3	5	0	5	10
4	8	1	9	0	1	10
5	5	3	8	0	2	10
Average	50%	22%	72%	0%	28%	100%

Summary

- Multi-modal time-sensitive fusion combines similarity of clinical data in multiple modalities while respecting time overlap in their occurrence.
- Patient similarity is a scalable way to achieve clinical decision support without any built-in rules
- It is personalized for the patient-specific conditions
- An example of how knowledge from electronic health records could be leveraged for meaningful use.