



Evidence-based anomaly detection in clinical domains

**Milos Hauskrecht, Michal Valko, Branislav
Kveton, Shyam Visweswaram, Gregory Cooper**

Computer Science Department and Department of
Biomedical Informatics, University of Pittsburgh



Detection of unusual patient management decisions

Patient medical records today:

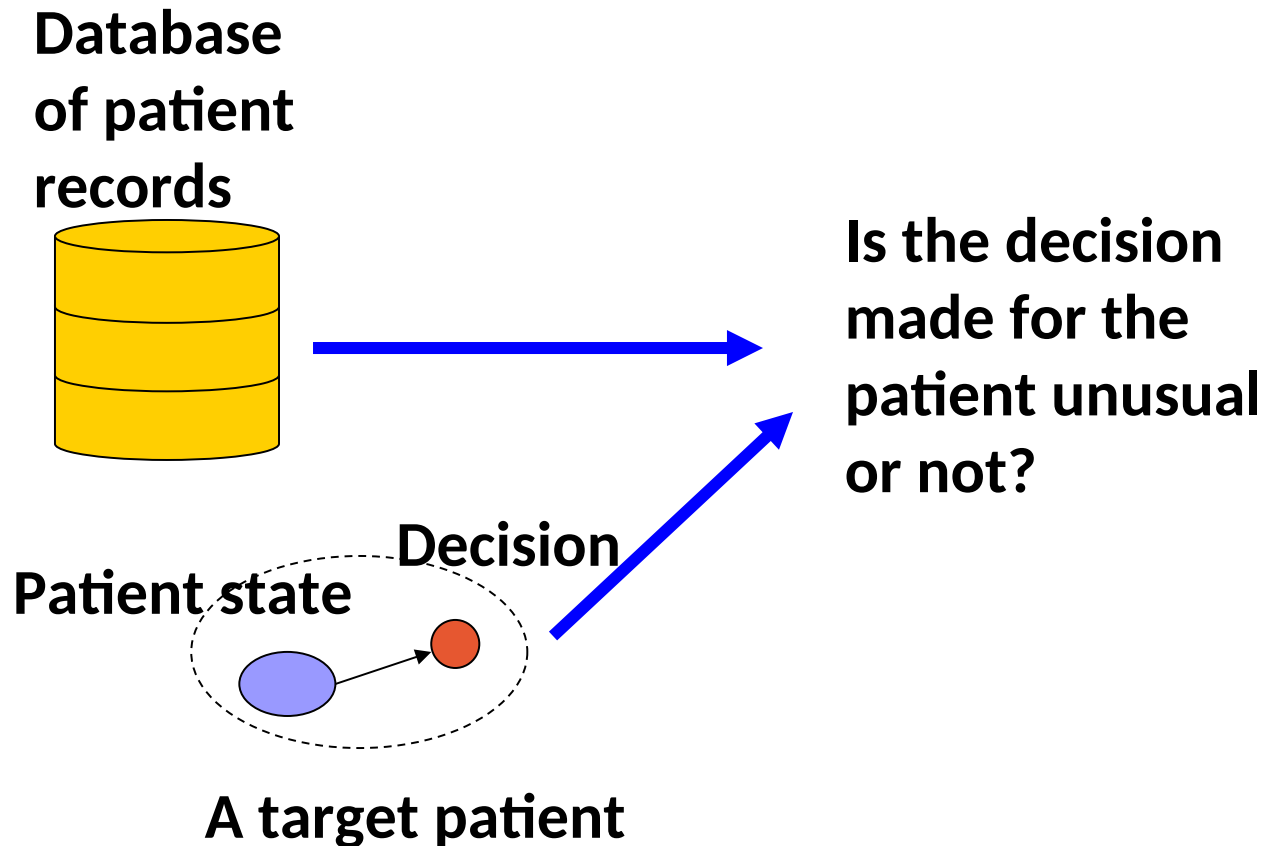
- Include thousands of electronic entries related to patient conditions and treatments
- Can be utilized in predictions, decision support, discovery or acquisition of new clinical knowledge

Objective of this work:

- Develop computational tools that utilize previously collected patient data to detect unusual (anomalous) patient-management decisions

Detection of unusual patient-management decisions

- **Goal:** Detect unusual (anomalous) patient-management decision patterns using past evidence





Application to detection of patient-management errors

Application:

- The design of on-line systems that alert clinicians to potential management/ treatment errors
- **Main assumption to be tested:**
 - Anomalies suggest errors often enough to justify alerting
- **Benefits:**
 - Automated anomaly detection methods require minimal or no expert input and are evidenced based
 - As more data is available electronically, these automated methods become ever more applicable
 - The automated methods apply widely to clinical care
 - Complementary to **knowledge-based alerting systems**
- **Limitations:**
 - Some errors will not be detected
 - Some (many?) anomalies will not represent errors



Detection of unusual patient-management decisions

- A decision that is perfectly normal for one patient may be unusual for the other
 - **Example:** Application of an anticoagulant treatment is normal for the majority of patients but not for a patient with a hemorrhagic stroke
- The fact that the patient management decision is unusual **depends on the patient's *clinical context***
- **Conditional anomaly detection (CAD):**
 - Methods for detecting the presence of unusual data pattern depending on the values of other data features

Conditional anomaly detection

- Data entry:



Condition (context) variables
(patient state)

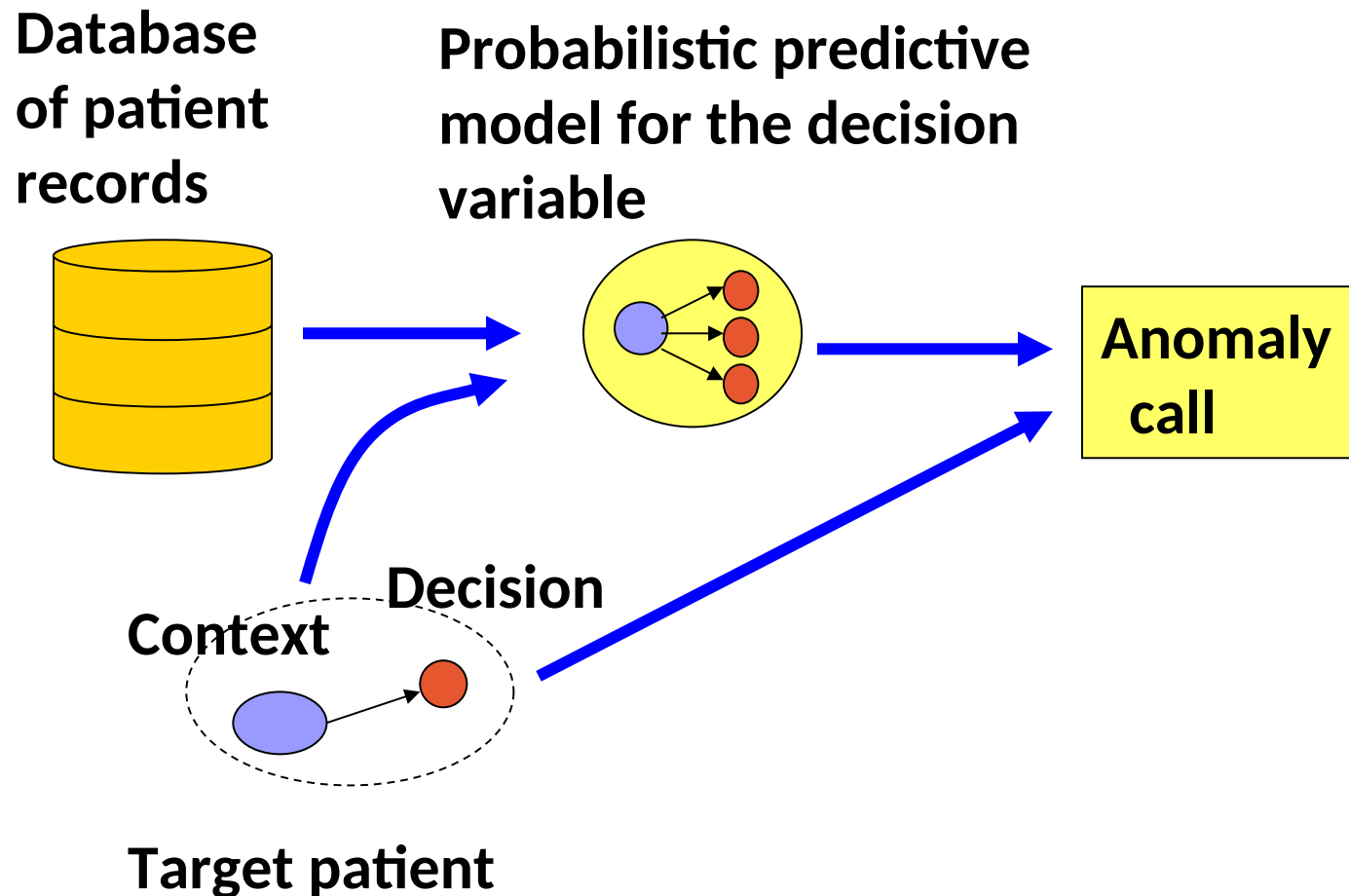
Target variables
(decision)

Question we ask:

- Given the values of context variables for the current patient are the values of the target variables for that patient unusual?

Probabilistic conditional anomaly detection

- Builds an auxiliary probabilistic predictive model

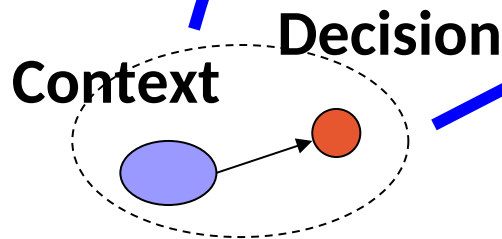
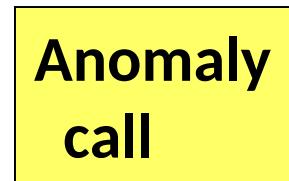
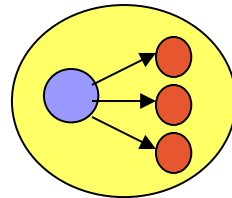
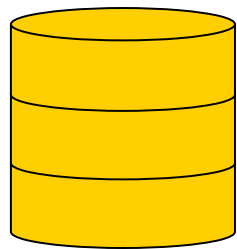


Probabilistic conditional anomaly detection

- Builds an auxiliary probabilistic predictive model

Database of patient records

Probabilistic predictive model for the decision variable



Target patient

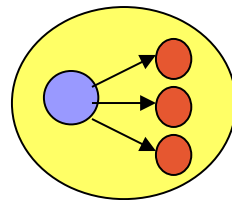
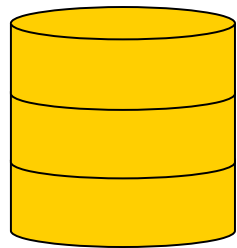
Step1:
Predictive model building

Probabilistic conditional anomaly detection

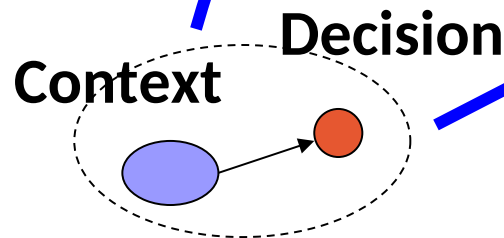
- Builds an auxiliary probabilistic predictive model

Database of patient records

Probabilistic predictive model for the decision variable



Anomaly call

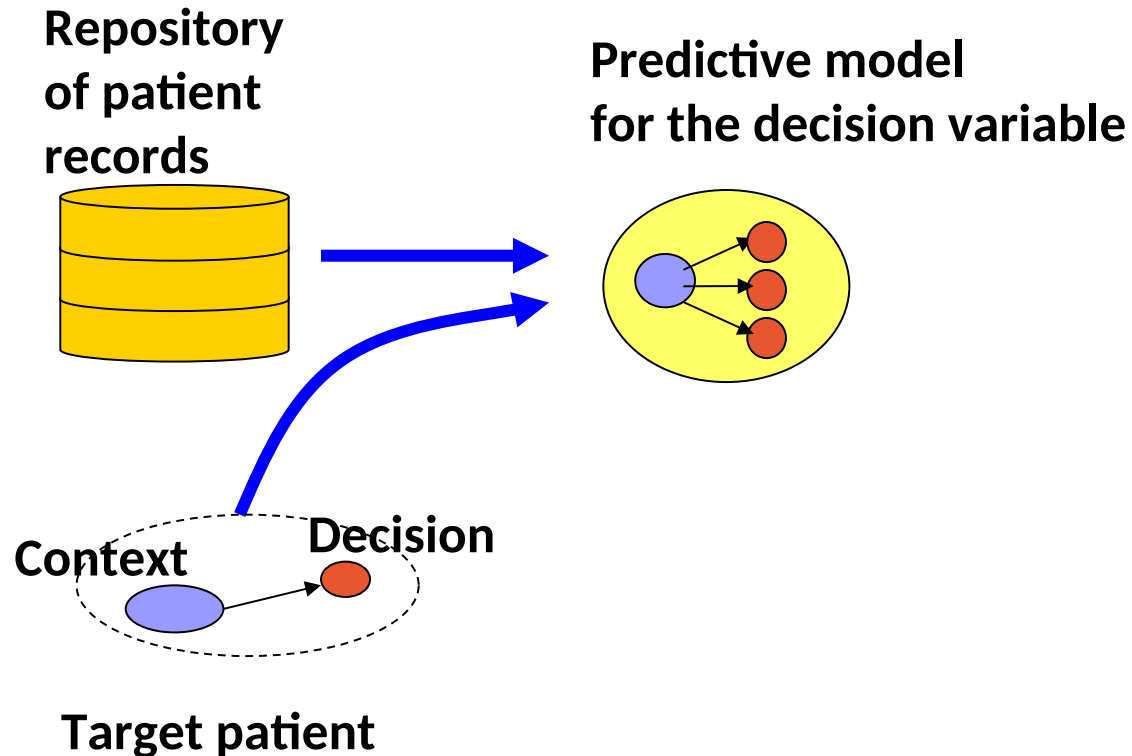


Target patient

Step2:
Anomaly detection (call)

CAD: model building

- **Step 1: Building a predictive model**
- **Goal:** build (learn) a model that defines a distribution of decisions for the current patient based on data for past patients





CAD: model building

- **We rely on models from statistical machine learning:**
 - Naïve Bayes model
 - Linear discriminant analysis
 - Logistic regression
 - Bayesian belief networks
 - Many others
- The learning task is similar to the classification learning problem
- The difference is that we want to learn **an accurate predictive distribution for any patient state**



CAD: model building

We would like to have:

- a model that is rich enough to give us a good estimate of **the probability of a decision for any patient state**
- May require a very large dataset of past patient cases

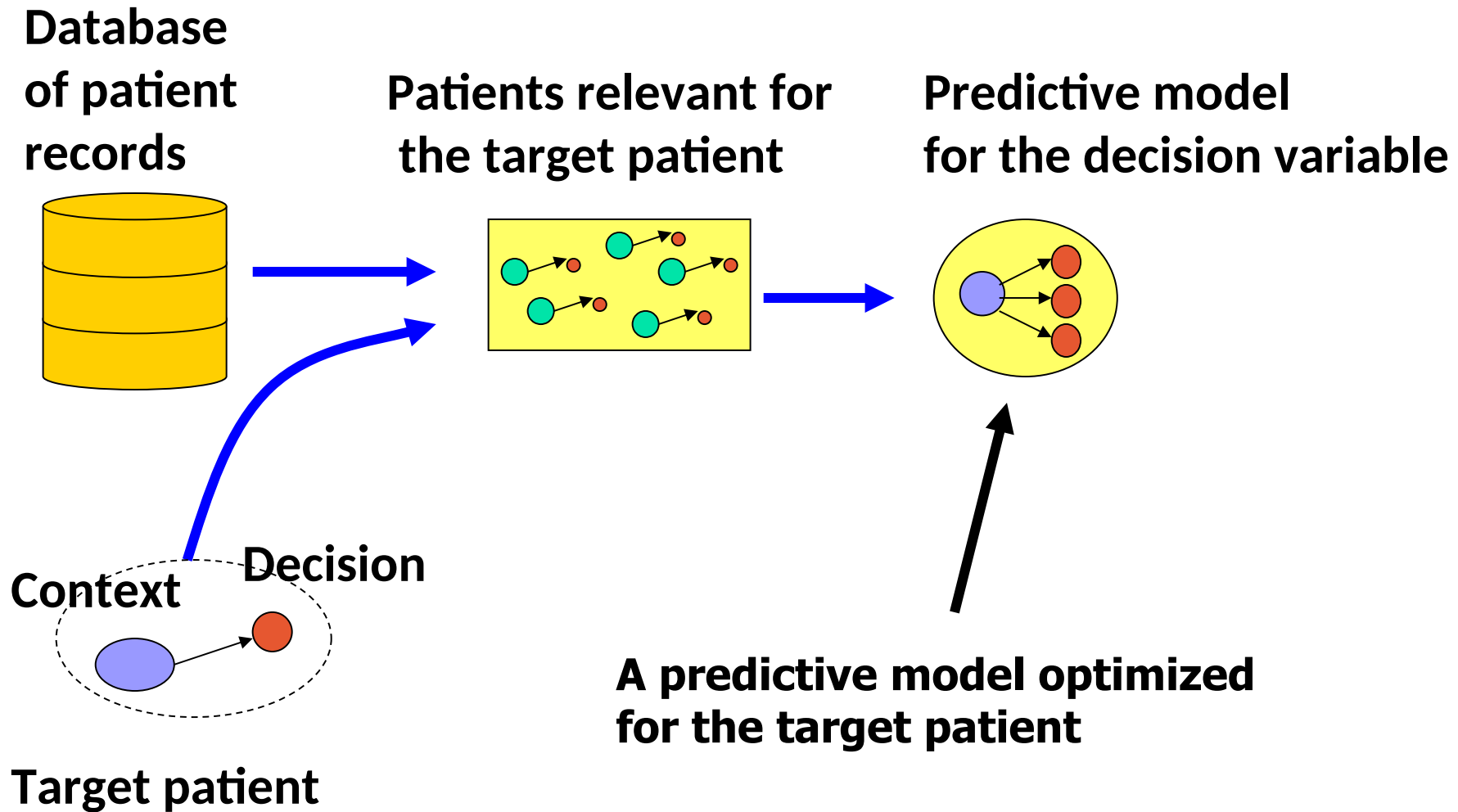
Reality:

- The number of samples in the database is limited and may be small relative to the number of patient-state features

Solution:

- Use **instance-specific models**
 - Models can be simpler (require less examples)
 - Models are tuned to individual patients

Instance-specific models



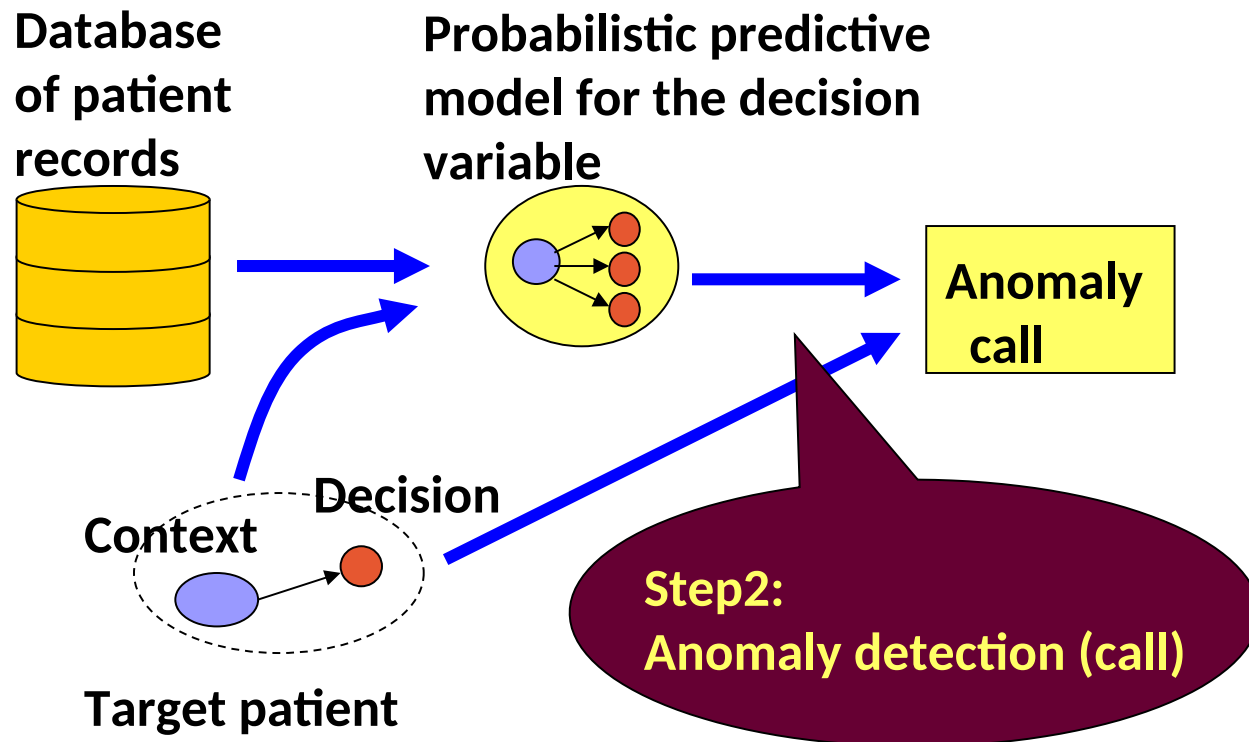


Selection of the relevant subpopulation

- **Exact match** on all patient's attributes
 - May not yield a (statistically) sufficient population
- **All patients**
 - Population biased by the prior
- **Similarity-based methods**
 - Euclidean metric
 - Ignores feature correlates (double counting)
 - Mahalanobis metrics
 - Features unimportant for the decision are counted
 - Weighted Mahalanobis metrics
 - Accounts for the importance of the decision

Conditional anomaly detection

- **Step 2. Anomaly detection**
- Given a probabilistic model and the target patient, decide if the decision made for the target patient is anomalous





CAD: Anomaly detection

Step 2. Anomaly detection

- **Approach: The anomaly call is made with the help of a probability threshold α**
 - Absolute criterion: $P(y \mid \mathbf{x}, \mathbf{M}) < \alpha$
 - Given the predictive model (\mathbf{M}) calculate the probability of observing a given treatment decision (y) in the current patient (who has values of \mathbf{x} for a set of clinical context variables)
 - If the probability falls below the threshold α then report an anomaly



Preliminary experimental evaluation

PORT dataset: Patients diagnosed with the community acquired pneumonia (2287 patient cases, 19 attributes)

Management decision:

Hospitalization

Demographic factors

Age > 50

Gender (male, female)

Co-existing illnesses:

Congestive heart failure

Cerebrovascular disease

Neoplastic disease

Renal disease

Liver disease

Physical-examination:

Pulse \geq 125 per min

Respiratory rate \geq 30 per min

Sys. blood pressure < 90mm Hg

Temperature < 35C or > 40C

Lab & radiological findings:

Blood urea nitrogen (mg/dl)

Glucose (mg/dl)

Hematocrit < 30 %

Sodium < 130 mmol/l

Art. O₂ pressure < 60 mm Hg

Arterial pH < 7.35

Pleural effusion



Experimental evaluation

100 patients selected for the evaluation:

- 21 patient cases out of the 2287 available were found to be anomalous by a simple Naïve Bayes detector ($\alpha = 0.05$)
- Of the remaining $2287 - 21$ patient cases, 79 cases were selected randomly

Selected patients presented to a panel of three physicians:

- Agreement with the hospitalization decision
- Unsure about the hospitalization decision
- Disagreement with the hospitalization decision

The case was labeled as “anomalous” if and only if:

- Two experts on the panel disagreed with the actual hospitalization decision that was made in a given case, or
- All three were unsure about the advisability of the decision



Experimental evaluation: comparison

Analysis: Leave-one-out analysis of hospitalization for each of the 100 cases

Predictive models tried:

- A naïve Bayes classifier
- A Bayesian belief network

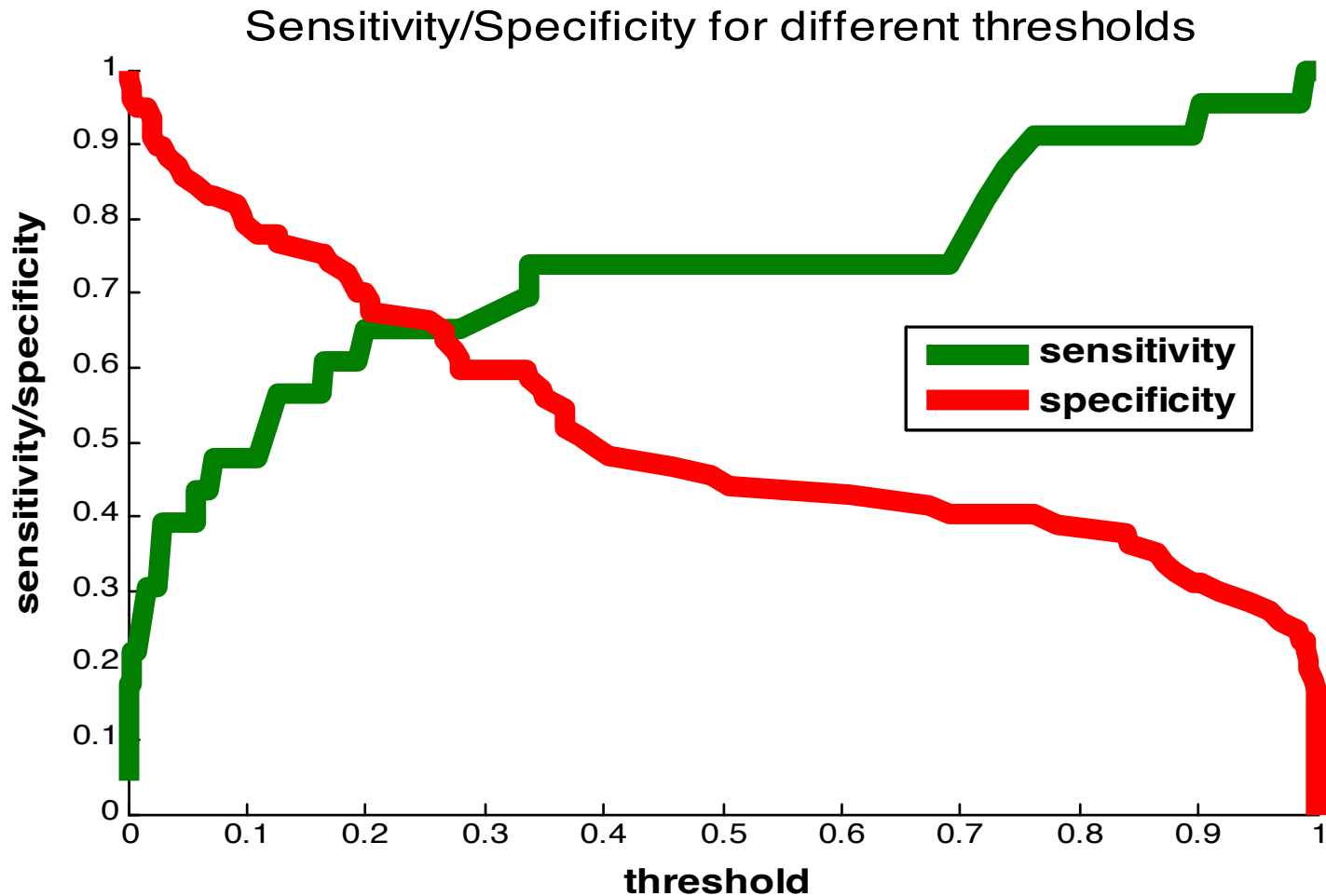
■ **Subpopulations:**

- All cases
- Top 40 Mahalanobis matches
- Top 40 weighted Mahalanobis matches
 - weights based on the univariate AUC

- **Anomaly calls:** Varied the anomaly threshold α to generate tradeoffs of sensitivity versus specificity

Experimental evaluation: comparison

- **Example:** NB detector





Experimental evaluation: Results

Prob. model	Population selection method	AUC ROC
NB	All cases	0.74
NB	Mahalanobis	0.77
NB	Weighted Mahalanobis	0.80
BBN	All cases	0.81
BBN	Mahalanobis	0.79
BBN	Weighted Mahalanobis	0.80



Experimental evaluation: Results

Prob. model	Population selection method	AUC ROC
NB	All cases	0.74
NB	Mahalanobis	0.77
NB	Weighted Mahalanobis	0.80
BBN	All cases	0.81
BBN	Mahalanobis	0.79
BBN	Weighted Mahalanobis	0.80

Naïve Bayes model:

- Subpopulation selection helps
- We achieve a better “local” fit, but we depend on the quality of the subpopulation



Experimental evaluation: Results

Prob. model	Population selection method	AUC ROC
NB	All cases	0.74
NB	Mahalanobis	0.77
NB	Weighted Mahalanobis	0.80
BBN	All cases	0.81
BBN	Mahalanobis	0.79
BBN	Weighted Mahalanobis	0.80

Bayesian belief network (BBN):

- data are sufficient to learn a model that is rich enough to capture key feature (in)dependencies



Experimental evaluation: Results

Prob. model	Population selection method	AUC ROC
NB	All cases	0.74
NB	Mahalanobis	0.77
NB	Weighted Mahalanobis	0.80
BBN	All cases	0.81
BBN	Mahalanobis	0.79
BBN	Weighted Mahalanobis	0.80

BBN benefits less from the subpopulation selection

- Model is already rich enough
- More parameters need to be estimated



Results: Limitation of the experiment

Analysis of false negatives

- The evaluators only saw the case summaries available to the anomaly detection system (rather than the full medical record)
- Many missing attributes
- Disagreements:
 - Physicians did not see the reason to hospitalize the patient
 - Our system saw examples of similar cases and did not call the anomaly



Discussion

- Clinical anomaly detection that is based on machine learning is a relatively unexplored method that potentially has wide clinical applicability
- It is particularly important as more clinical data become available electronically
- The results of a preliminary evaluation of the method to the hospitalization decision for patients with community acquired pneumomonia suggest that the method is promising and should be explored further



Current and Future work

- We are evaluating the ability of the method to detect clinically useful anomalies in post-surgical cardiac patients (~ 5000 medical records)
- We are refining the conditional anomaly framework
 - Metric learning methods
 - Non-linear similarity metrics
- We are planning to explore related applications, including:
 - Detecting anomalous patient outcomes
 - Detecting anomalous care that is associated with especially favorable or unfavorable patient outcomes



Acknowledgments

People who contributed to the project:

- Melissa Saul
- Amy Seibert
- Tomas Singliar
- James Harrison
- Andrew Post
- Michael Fine

- Funded by the National Library of Medicine:
 - Grant R21-LM009102



Evidence-based anomaly detection in clinical domains

**Milos Hauskrecht, Michal Valko, Branislav
Kveton, Shyam Visweswaram, Gregory Cooper**

Computer Science Department and Department of
Biomedical Informatics, University of Pittsburgh



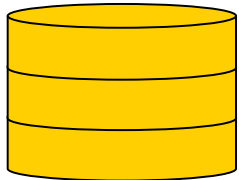
Additional slides

CAD: instance-specific models

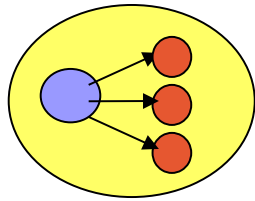
Basic idea:

- one rich predictive model vs many predictive models tuned to individual patients

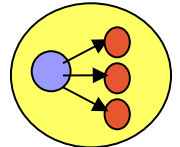
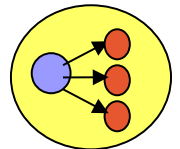
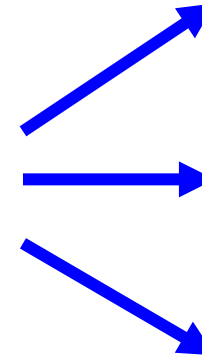
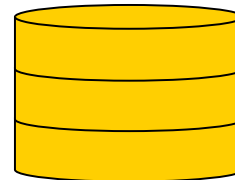
Database
of patients



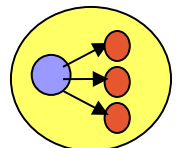
Predictive model



VS

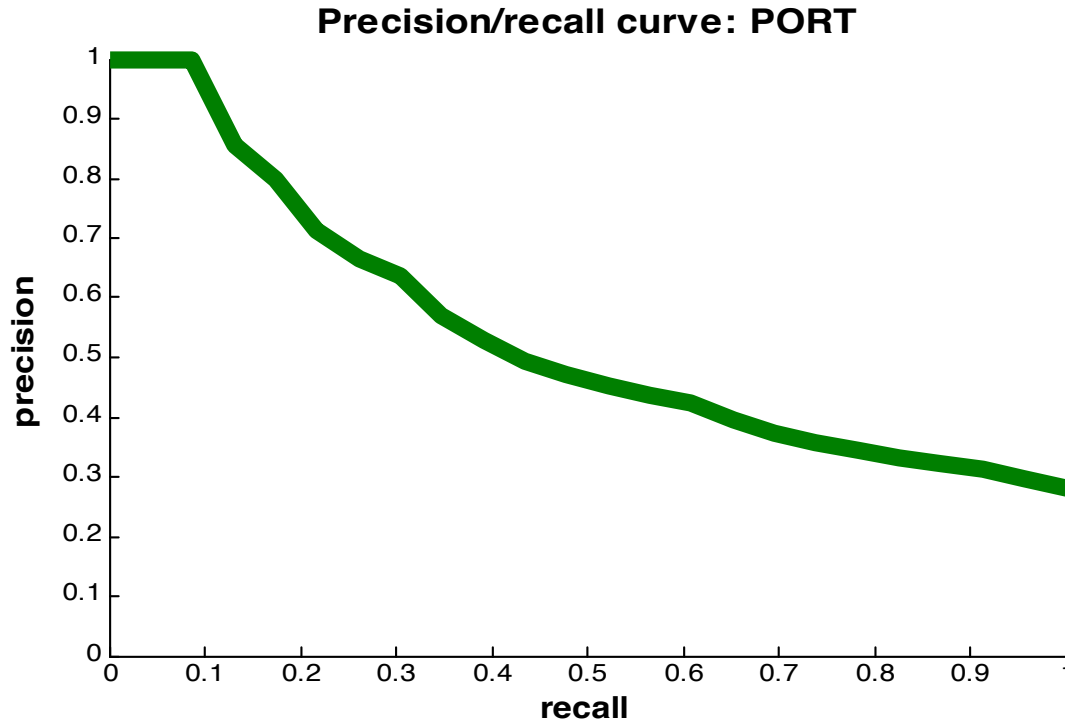


...



Experimental evaluation: comparison

- Varied detection thresholds were used to construct the **Precision-Recall curve** (used in the information retrieval)





CAD: subpopulation selection

Caveats of similarity-based subpopulation selection

- **How to determine the size of the relevant subpopulation?**
 - top k patient matches
 - drop in the distance
 - or weight patients proportionally to their distance
- **Is the population sufficient to make a sound anomaly call?**
 - Exclude outliers based on the average subpopulation distance.



Are “unusual” decisions equal to mistakes

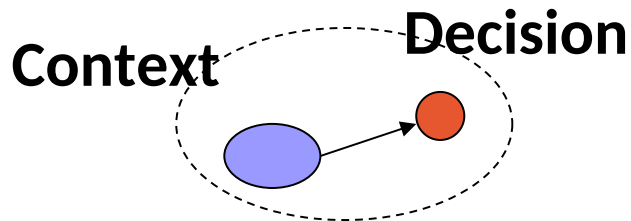
- Some unusual decisions are not equal to errors
 - Differences in management choices across hospitals, or for different physicians may exist.
 - Our system may alert to these decisions.
- **Two solutions:**
 - Flag the alert so that the mistake is not made again.
 - **Abstractions.** Filter out some false positives by incorporating prior knowledge that lets us assess the “equivalence” of some decisions (say two different drugs are equivalent in terms of the goals)



End

Conditional anomaly detection

- **Our goal:** Detect unusual patient management decisions
- **Formulation of the problem in CAD framework:**
 - **Condition (context) variables:**
 - define the patient state
 - **Target variable:**
 - a patient-management decision

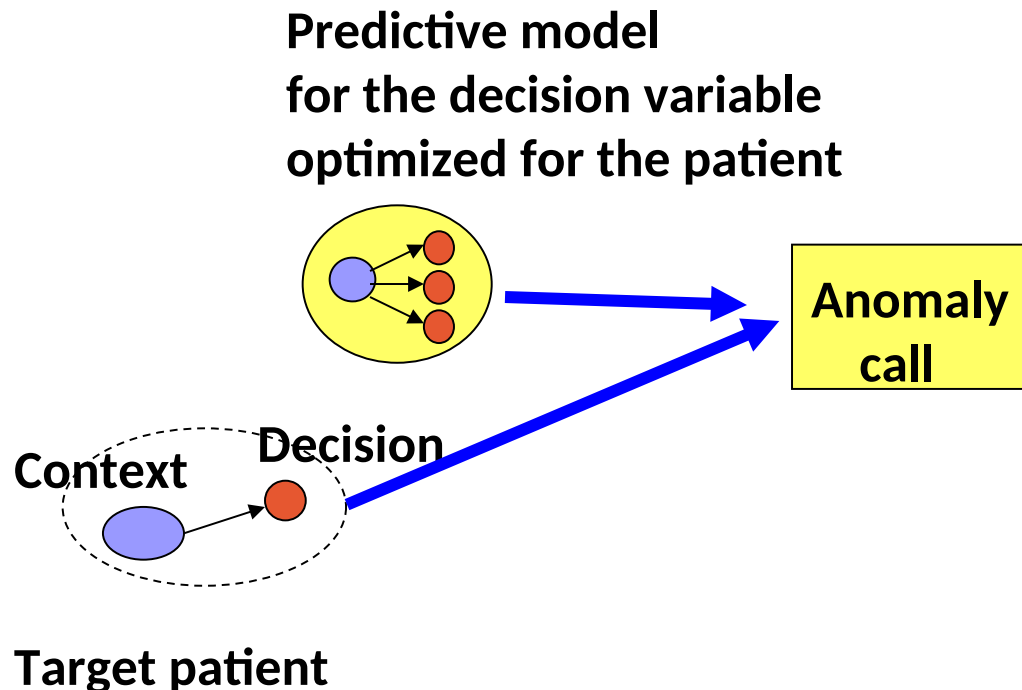


A patient case

CAD: anomaly detection

Step 2. Anomaly detection

- **Goal:** Given a probabilistic model and the target patient, decide if the decision made for the target patient is anomalous



Instance-specific model

Selection of a relevant patient subpopulation

- Identify patients similar to the target patient with respect to the decision variable of interest

