Data Mining using the Omaha System

Oral Health in Dakota Co.

Gyorgy J. Simon
Jeanne Pfeiffer
Laura L. Fitzsimmons
Karen A. Monsen

Oral Health

- Oral health is a window to overall health
  - Gum disease may allow bacteria to enter the bloodstream
  - Chronic periodontitis can contribute to cardiovascular disease
  - Gum disease have been linked to premature birth
- Few databases contain information about oral health
  - One of the 42 central concepts in the Omaha System

Oral Health in Dakota Co.

- Familial data for 6,425 clients
  - Problems, Signs and Symptoms, Interventions,
  - 1,781 (27.7%) have oral problems
  - Study period: 2009 – 2011

- Find patterns in clients that are predictive of oral health problems
  - Initially, patterns in oral health data
  - Patterns in mothers predictive of oral health problem in children
  - Successful interventions

Data Mining

Data Mining	***Statistics***
Exploratory • Find novel, interesting patterns Hypothesis generation • Large number of hypotheses • Filter down to a smaller set No guarantees about results Large number of predictors	Confirmatory • Confirm hypothesis Hypothesis testing • Few hypotheses Rigorous Few, very relevant predictors

Association Rule Mining

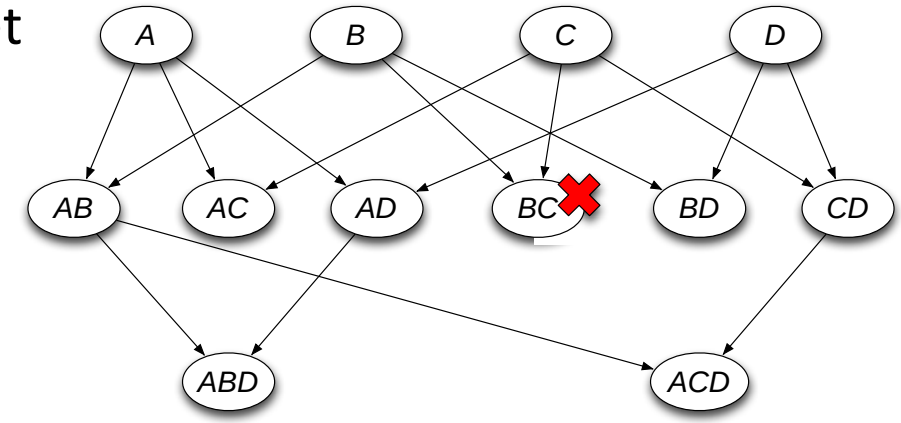
- Origins from sales data
- *Items*: articles carried by a store
- *Transactions*: items in the same shopping cart
- *Itemsets*: sets of items
- *Goal*: find *all* itemsets that are *frequently* purchased together

Cart #	*A*	*B*	*C*	*…*	*D*
001	Y	Y	Y		Y
002	Y	Y	Y		Y
003		Y	Y		
004		Y			
005	Y	Y	Y		

- *Support*: # of transactions the itemset *I* appeared in
  - Support(ABC)=3
- *Frequent*: an itemset *I* is frequent, if support(*I*)>*minsup*

Application to Oral Health

- Items are problems
- Rows are patients



X: infrequent

Method

- Create a binary matrix
  - columns are Problems
  - rows are clients
  - Entry indicates whether the particular client has had the particular Problem during the study period and received Intervention for it
- Extract all combinations of Problems that exist in at least 5 clients with oral health problem
  - 2,900 combinations were found

Method 2

- 2,900 patterns are difficult to interpret

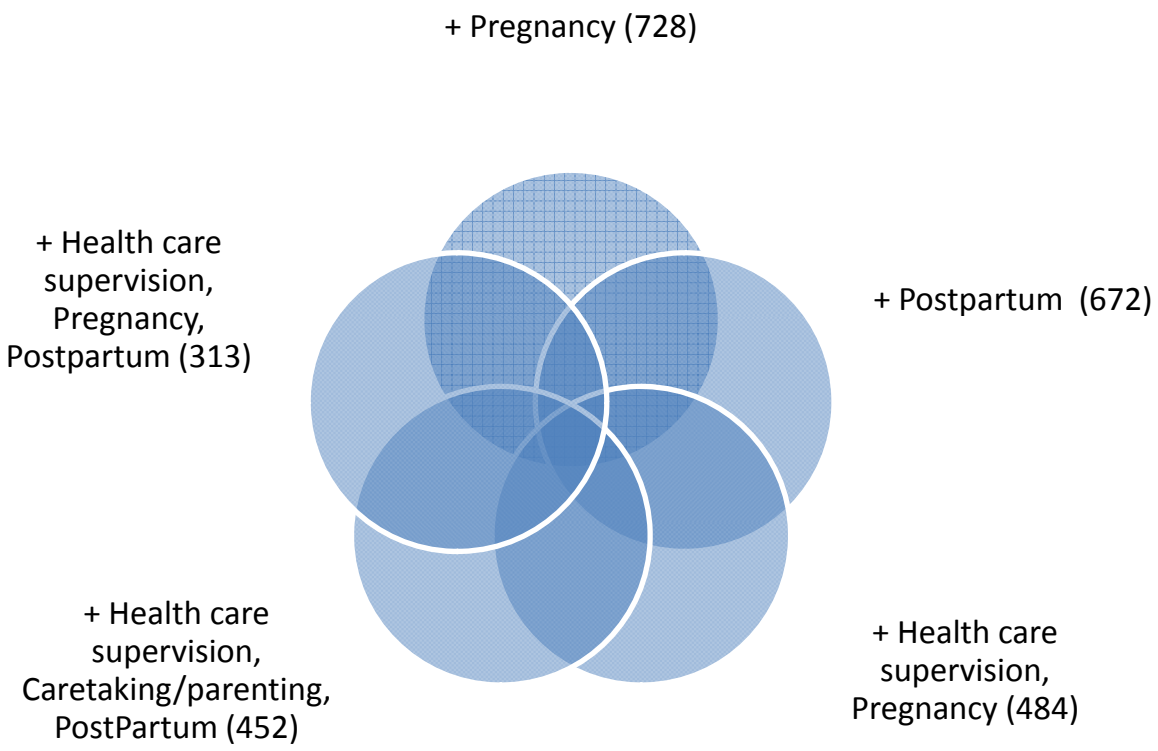
- Filtering based on predictive capability
  - Patterns that are not significantly predictive of oral health are irrelevant
- Filtering based on independence of items
  - Items co-occur more frequently than expected under the assumption that they co-occur at random
- Filtering based on independence of sub-patterns
- Summarize the patterns

Filtered and Summarized Patterns

supOH Pattern	
728	Mentalhealth Pregnancy Income Residence Abuse FamPlanning Substance
672	Mentalhealth Income Postpartum Residence Abuse FamPlanning Substance
484	Mentalhealth Pregnancy Income Residence HealthCareSuprv Abuse FamPlanning Substance
452	Mentalhealth Income Postpartum Residence Caretaking HealthCareSuprv Abuse FamPlanning Substance
316	Mentalhealth Pregnancy Income Postpartum Residence HealthCareSuprv Abus FamPlanning Substance
14	Nutrition Residence
13	Mentalhealth Antepartum Caretaking
13	Antepartum FamPlanning
13	Mentalhealth Antepartum FamPlanning

Focus of PHN Practice

- Mental health, Income, Residence, Abuse, Family planning, Substance Use



Summary

- Considering just Problems for individuals (as opposed to families) findings are consistent with use of evidence-based care plans that are in place and are being used in Dakota County
  - Demonstrates quality of PHN care/documentation
  - Suggests possible client types
  - May indicate evidence-based changes in PHN practice over time

Next Steps

- Select patients who have interventions for the Oral health
  - Consider family a unit (not a client)
  - Take signs/symptoms into account
  - Assess the effects of interventions
- Deidentification blinds us to dates
  - Temporal sequence of events is unclear

Acknowledgements

Jeanne Pfeiffer
University of Minnesota

Laura L. Fitzsimmons
Dakota County Public Health

Karen A. Monsen
University of Minnesota

Thank you !