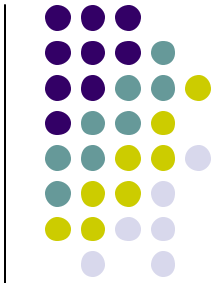
A Smart Health Care Data Analysis System: What Might It Look Like?

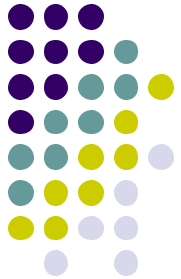


K. Ponnambalam

Professor, Department of Systems Design Engineering

University of Waterloo

October 11, 2006
Davis Centre 1304, University of Waterloo



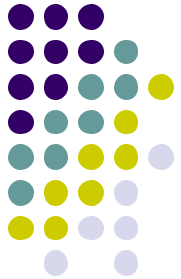
Abstract

Smart cards are now used in many countries to store the health data of an individual. This is normally a personal history of medical facts and data. In addition, regulations require many jurisdictions to capture specific health data of all patients, for example, with cancer or other specific conditions. In this talk, we give a general system architecture that can take advantage of such data to support improved medical care to the individual, and also help improve the system efficiency. We will present also a case study using data mining.

Outline

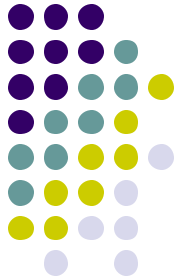
- Health Record
  - Smart card
  - Electronic Health Record
- What can we do with ***all*** the data?
- Case Study – Mental Health

Smart Health Card



- Unlike most cards (e.g. credit card) “Smart” cards are highly secure and has an embedded microprocessor.
- Allows only authenticated readers to read the information on the card.
- User normally carries the card
- Can be connected to the Web servers through the card reader

What data it may contain?



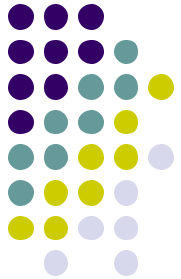
Usually has four kinds of data:

- Device data (identifying the device(s) and functions);
- User data (unique ID of the device holder and related parties);
- System data (funding relationships, insurance carriers, public or private healthcare providers); and
- Clinical data (information about the cardholder's health, events, appraisal and labeling by a healthcare professional, and related actions planned or performed).

Electronic Health Record

"An Electronic Health Record (EHR) is a health record of an individual that is accessible online from many separate, interoperable automated systems within an electronic network.", Health Canada – Health Care System, 2006

Canadian Health Infoway



“Vision
A high-quality, sustainable and effective Canadian healthcare system supported by an infostructure that provides residents of Canada and their healthcare providers with timely, appropriate and secure access to the right information when and where they enter into the healthcare system. Respect for privacy is fundamental to this vision.”

eHealth

Institution: administration systems; laboratory and radiology information systems; electronic messaging systems; and, telemedicine -- teleconsults, telepathology, and teledermatology, and others
Home care : teleconsults and remote vital signs monitoring systems used for diabetes medicine, asthma monitoring and home dialysis systems
Primary care : use of computer systems by general practitioners and pharmacists for patient management, medical records and electronic prescribing.
The above are from Heatlh Canada, 2006

I add further three items:
Patient: continuous monitoring and guidance for better health and preventive care.
Professional practice: Improve best practices, College of Physicians and Surgeons.
Governance: Improve regional, provincial, national, and global health governance.

WIHIR (Univ. of Waterloo eHealth Seminars for the last 12 months)

September 28, 2005 - Standing on a Burning Platform: Implementing a Province Wide Wait-time Information System

October 26, 2005 - Getting to Filmless: A Shared Digital Imaging Initiative

November 23, 2005 - Accelerating the Development and Implementation of Electronic Health Records (EHR)

January 25, 2006 - Drowning in Data: What's a Healthcare Provider to Do?

February 22, 2006 - In Pursuit of the EHR Holy Grail: A Critical Situational Review?

March 22, 2006 - The Ontario Laboratories Information System (OLIS) System Project A strategic change initiative for eHealth in Ontario

April 26, 2006 - New Strategies for Health Care in the Home

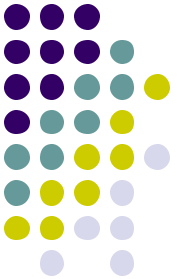
May 24, 2006 - Ontario's Public Health E-Health Strategy - Supporting Public Health Renewal

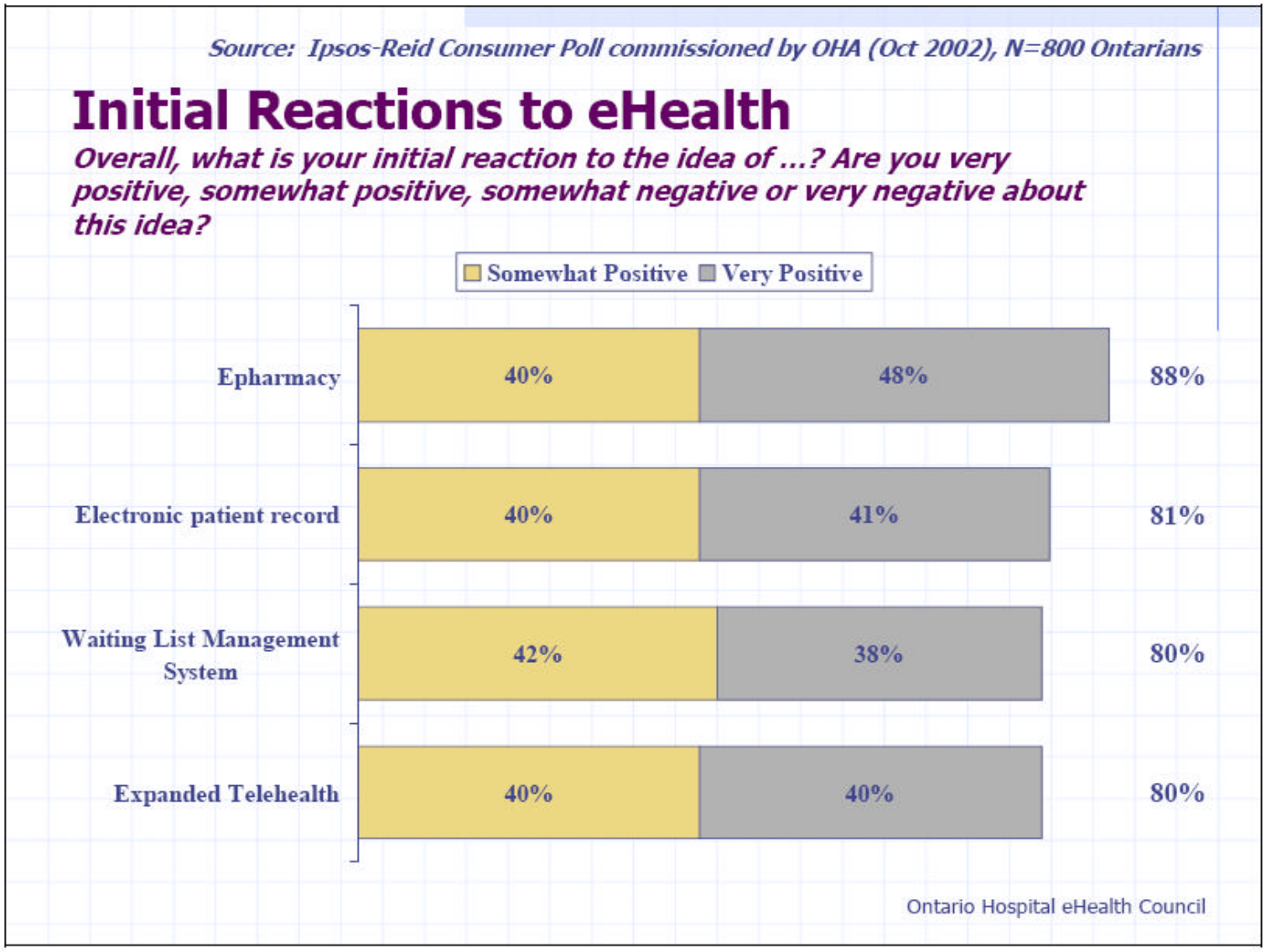
June 21, 2006 - The Physician in eHealth: The Missing Link

September 27, 2006 - Engaging Physicians in Ontario's eHealth Strategy

October 25, 2006 - eHealth Strategies in Support of Psychiatry

Why we do need yet another eHealth seminar now?





From The Road to Electronic Health Records for Ontario: *Enabling Regional Service Delivery,* The Ontario Hospital eHealth Council Electronic Health Record Working Group, February 2003

July 20, 2006 10:52 AM

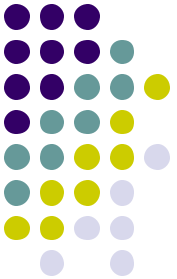
US government: by 2014 all medical records will be electronic

Rand Corporation estimates that in efficiency costs alone a paperless medical records infrastructure could save US$77 billions each year. Aiming a 2014 target, the goverment is creating a series of projects to fund the electronic medical records initiative, including a certification process, that will qualify vendors and products.

Full electronic records are not only cheaper, but far more reliable and fast. Think no more overnight delivery of piles of dust-covered folders or those always-misplaced dental x-rays.

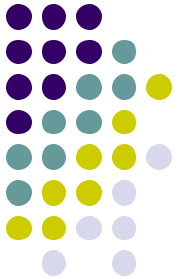
Source: Ars Technica

Current Objectives of EHR



- Better prescribing practices
- Reduced waiting times
- Improved access in remote communities
- Complete and accurate clinical data for diagnosis and treatment
- Less administrative burden and more time with patients

Diagnostic Tools or Toys?



- Raja's Palm Library.htm
  (http://abusharr.net/palm/)

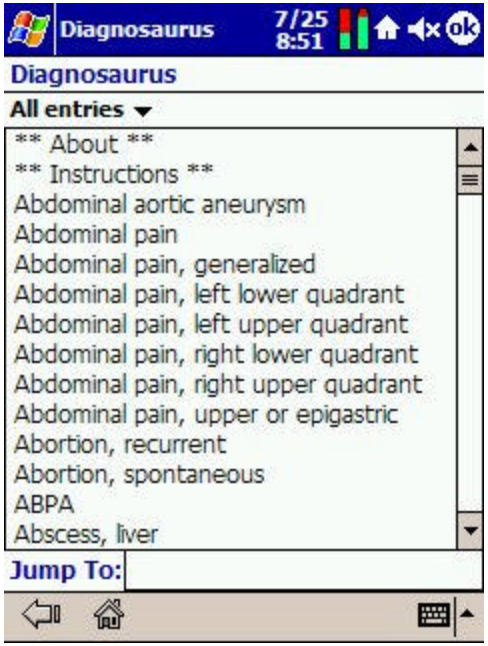
- Tech4doctors.htm
  (http://www.canhealth.com/doctors.html)

- Ectopic Brain - What's New.htm
  (http://pbrain.hypermart.net/)

Diagnosaurus: a free differential diagnosis tool for PalmOS and Pocket PC

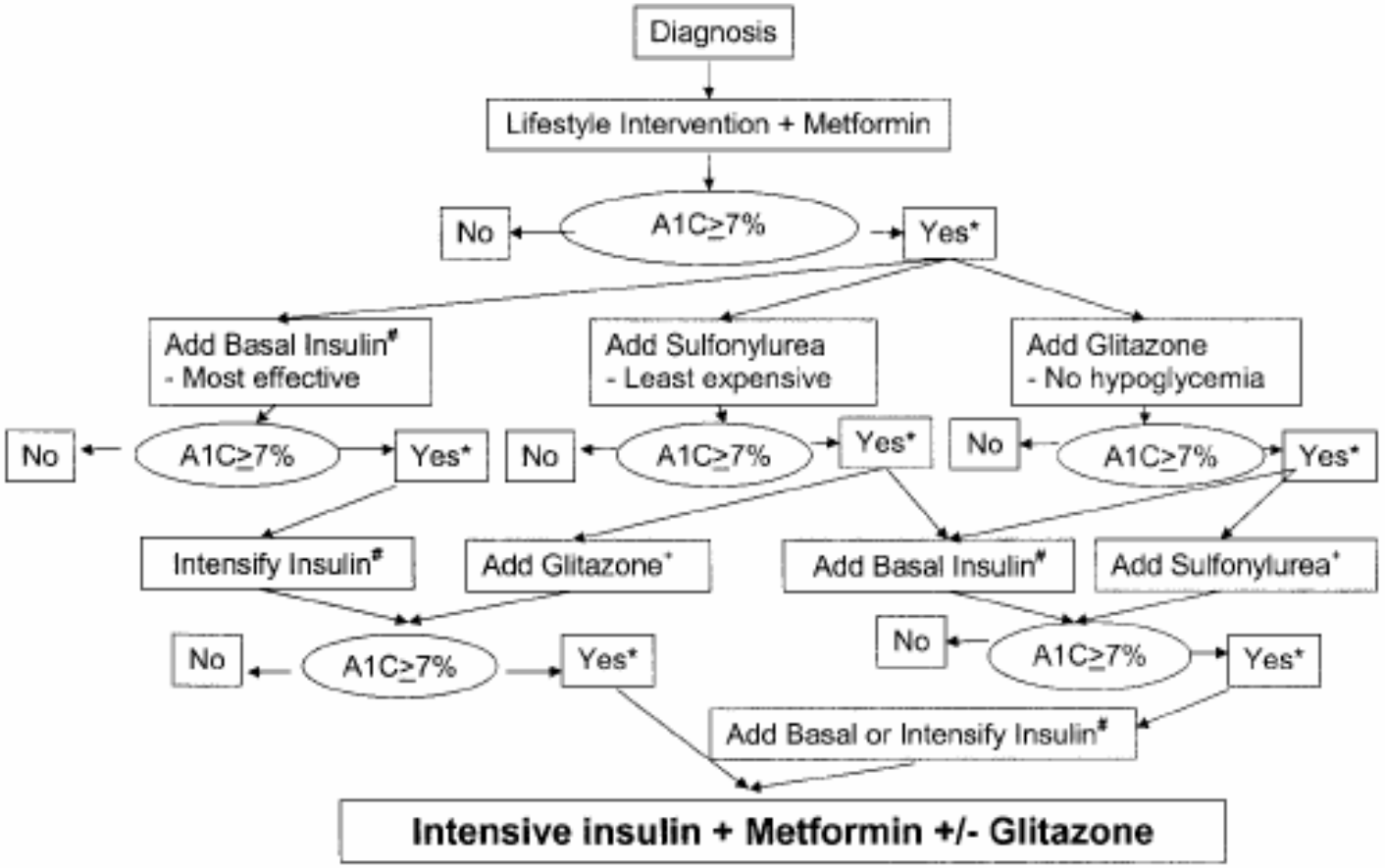
(McGraw Hill Inc.)

This software is a freeware designed to provide differential diagnosis for



more than 1000 common presentations. ….

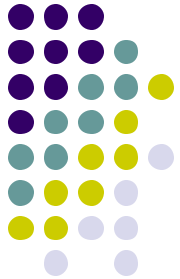
Management of hyperglycemia in type 2 diabetes



From Management of Hyperglycemia in Type 2 Diabetes: A Consensus Algorithm for the Initiation and Adjustment of Therapy: A consensus statement from the American Diabetes Association and the European Association for the Study of Diabetes

David M. Nathan, John B. Buse, Mayer B. Davidson, Robert J. Heine, Rury R. Holman, Robert Sherwin, and Bernard Zinman
Diabetes Care 2006 29: 1963-1972.

Modern tools that a health care system depends on

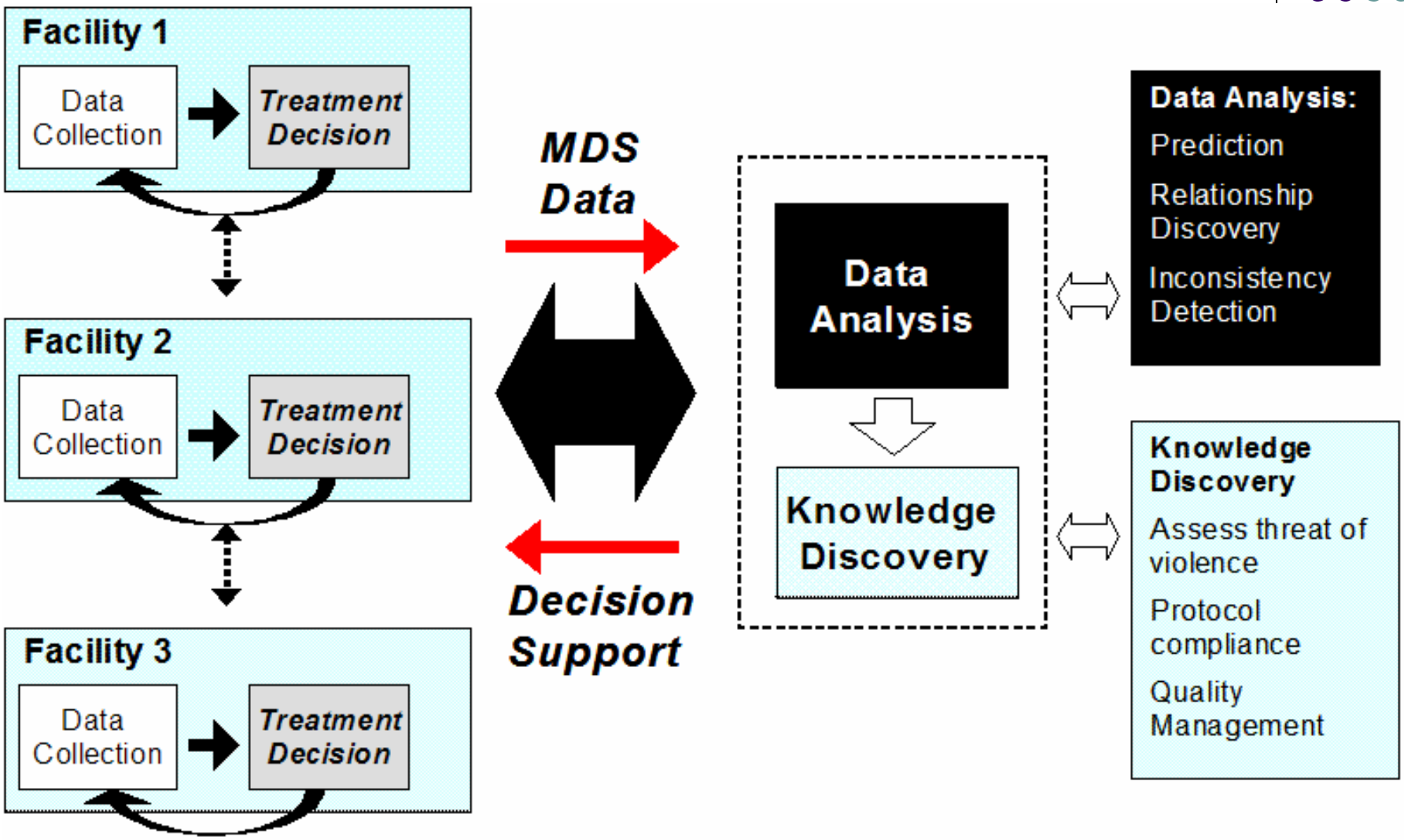


- Various Electronic Monitors
- Sophisticated Laboratories (e.g.DNA analysis)
- MRI
- CATSCAN
- Radiation Therapy
- Various Robotic Devices

In the future, it will be many different software based on new algorithms and knowledge

None of which can replace a well trained practitioner!

Most practitioners would agree that a well designed tool could save time, save lives!

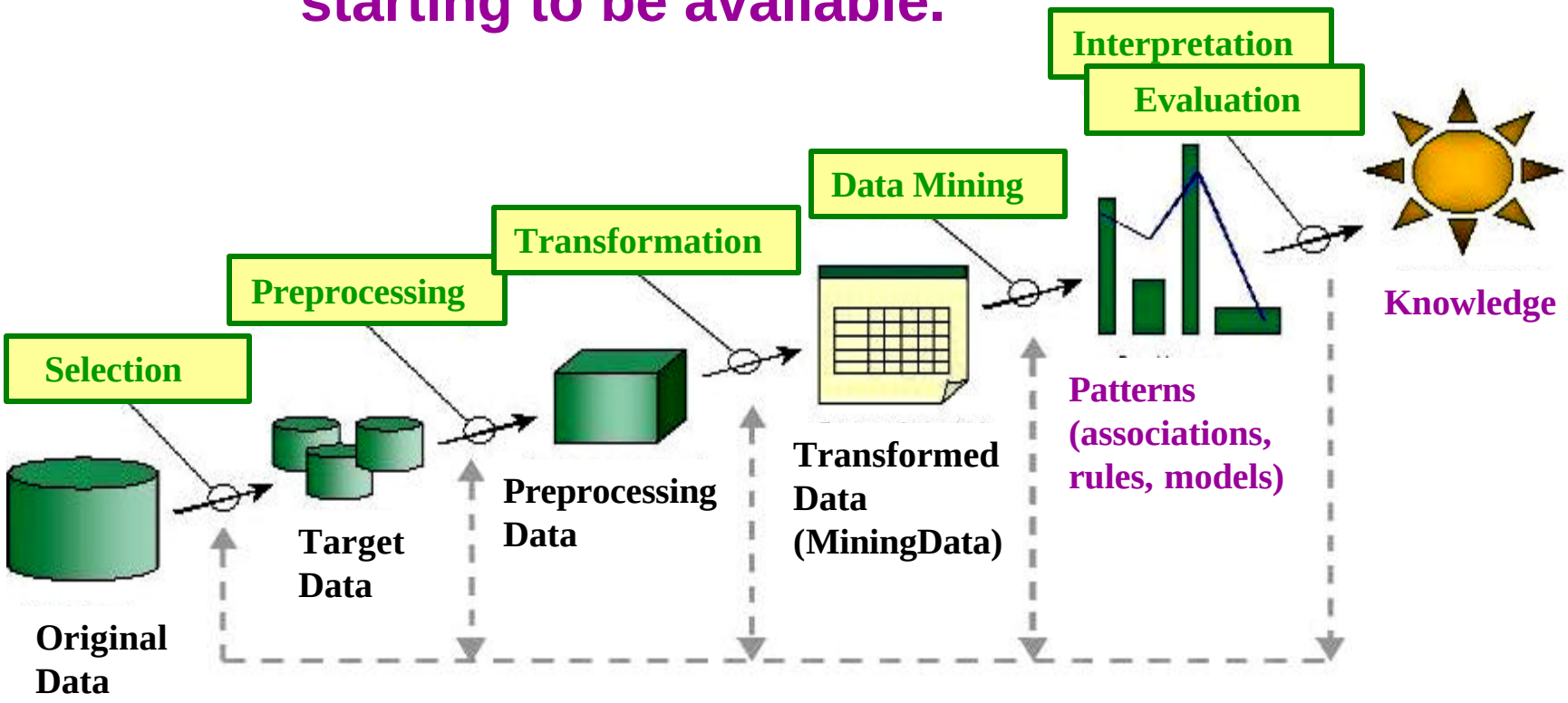


System Architecture for Mental Health Care

*Source: Discover*e tool PDS, Waterloo*

The Data Mining Process

A tool to use thousands of records starting to be available.



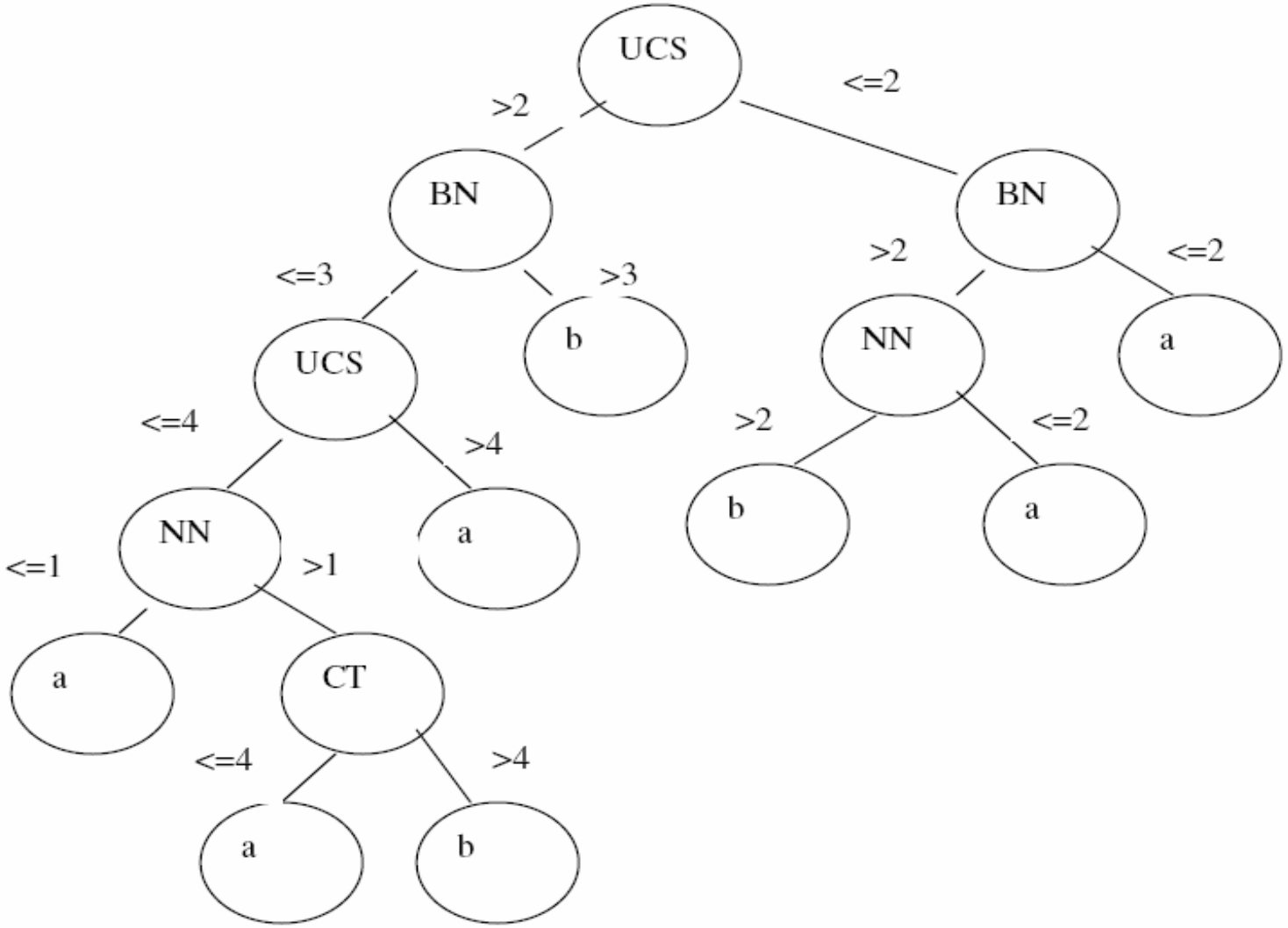


Figure: Decision tree that is created using the WEKA software for Breast Cancer Diagnosis.

ACCURACY OF DATA MINING TOOLS: WISCONSIN BREAST CANCER DATA

WEKA		CRUISE		DISCOVER*E		LVQ	
METHOD	ACCURACY	METHOD	ACCURACY	METHOD	ACCURACY	METHOD	ACCURACY
ZeroR	77.88%	Univariate Split	98.49%	Decision Tree	99.49%	OLVQ	98.99%
Decision Tree	98.995%			Dependence Tree	97.98%		
Decision Table	98.55%	Linear Split	98.49%	Rule Classification	99.49%		
Naïve Bayes	98.49%						

Source: Eapen, A.G.(2004) "Application of Data mining in Medical Applications" M.A.Sc. Thesis, Department of Systems Design Engineering University of Waterloo, Canada.

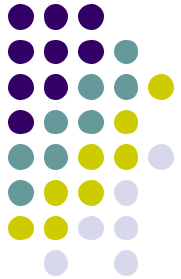
Case Study: Data mining in mental health

- A.G. Eapen[1], ***K. Ponnambalam[1]***, J.F. Arocha[2], R. Shioda[3], T.F. Smith[2], J. Poss[2], & J. Hirdes[2]

- [1]Systems Design Engineering, [2]Health Studies,[2]Combinatorics & Optimization, University of Waterloo

Acknowledgments

- Homewood Healthcare Research Institute, Guelph, ON
- Pattern Discovery Systems Inc., Waterloo, ON
- interRAI

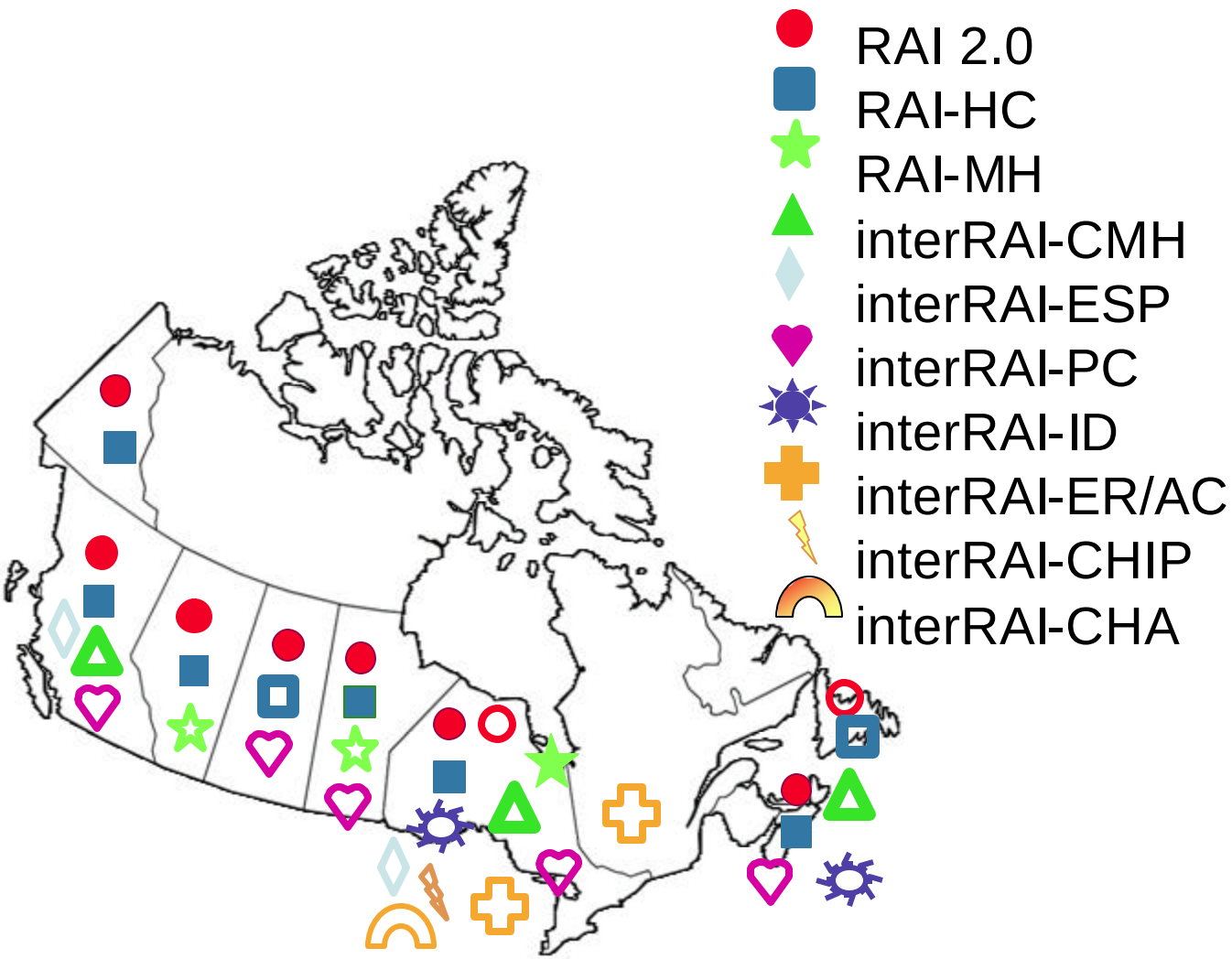


- The InterRAI MDS organization
- The data set
  - Functions
  - Scope
  - The instrument
- Data mining tools
- Experimental results

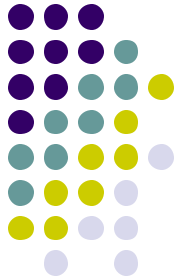
interRAI

- interRAI Minimum Data Set
  - Resident Assessment Instrument
  - Smallest number of pieces of information required from each patient
- Collection of tools for gathering health data
- Integration of data from different sources
- Thousands of patient records have already been collected

Largest health care database in Canada



Data categories



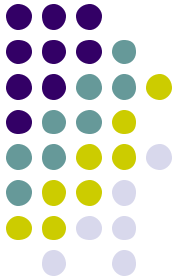
- Over 250 data items per instrument
  - Personal items
  - Health service use, history
  - Substance use
  - Cognitive state, history
  - Nutrition
  - Medications
  - Diagnoses
  - etc....,

RAI Integrated System

- Common language
- Common theoretical/conceptual basis
- Common clinical emphasis
- Common data collection methods
- Common core elements

Hirdes et al., *Healthcare Management Forum*, 1999)

RAI data quality



- Collection and evaluation of patient-specific assessment data with
  - A common language
  - Multiple uses
  - Permits direct comparisons
  - Allows for timely access to data

Resident _____ Numeric Identifier _____

MINIMUM DATA SET (MDS) — *VERSION 2.0*
FOR NURSING HOME RESIDENT ASSESSMENT AND CARE SCREENING
FULL ASSESSMENT FORM
(Status in last 7 days, unless other time frame indicated)

SECTION A. IDENTIFICATION AND BACKGROUND INFORMATION

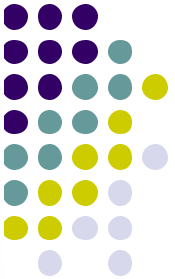
1.	RESIDENT NAME	a. (First) b. (Middle Initial) c. (Last) d. (Jr/Sr)
2.	ROOM NUMBER	
3.	ASSESSMENT REFERENCE DATE	a. *Last day of MDS observation period* Month Day Year b. Original (0) or corrected copy of form (enter number of correction)
4a.	DATE OF REENTRY	**Date of reentry from most recent temporary discharge to a hospital in last 90 days (or since last assessment or admission if less than 90 days)** Month Day Year
5.	MARITAL STATUS	1. Never married 3. Widowed 5. Divorced 2. Married 4. Separated
6.	MEDICAL RECORD NO.	
7.	CURRENT PAYMENT SOURCES FOR N.H. STAY	(*Billing Office to indicate;* ***check all that apply in last 30 days***) a. Medicaid per diem b. Medicare per diem c. Medicare ancillary part A d. Medicare ancillary part B e. CHAMPUS per diem f. VA per diem g. Self or family pays for full per diem h. Medicaid resident liability or Medicare co-payment i. Private insurance per diem (including co-payment) j. Other per diem
8.	REASONS FOR ASSESSMENT	***[Note—if this is a discharge or reentry assessment, only a limited subset of MDS items need be completed]*** **a.** Primary reason for assessment 1. Admission assessment (required by day 14) 2. Annual assessment 3. Significant change in status assessment 4. Significant correction of prior full assessment 5. Quarterly review assessment 6. Discharged—return not anticipated 7. Discharged—return anticipated 8. Discharged prior to completing initial assessment 9. Reentry 10. Significant correction of prior quarterly assessment 0. NONE OF ABOVE ***b. Codes for assessments required for Medicare PPS or the State*** *1. Medicare 5 day assessment* *2. Medicare 30 day assessment* *3. Medicare 60 day assessment* *4. Medicare 90 day assessment* *5. Medicare readmission return assessment* *6. Other state required assessment* *7. Medicare 14 day assessment* *8. Other Medicare required assessment*
9.	RESPONSIBILITY/ LEGAL GUARDIAN	(***Check all that apply***) a. Legal guardian b. Other legal oversight c. Durable power of attorney/health care d. Durable power attorney/financial e. Family member responsible f. Patient responsible for self g. *NONE OF ABOVE*
10.	ADVANCED DIRECTIVES	(*For those items with supporting* ***documentation*** *in the medical record,* ***check all that apply***) a. Living will b. Do not resuscitate c. Do not hospitalize d. Organ donation e. Autopsy request f. Feeding restrictions g. Medication restrictions h. Other treatment restrictions i. *NONE OF ABOVE*

SECTION B. COGNITIVE PATTERNS

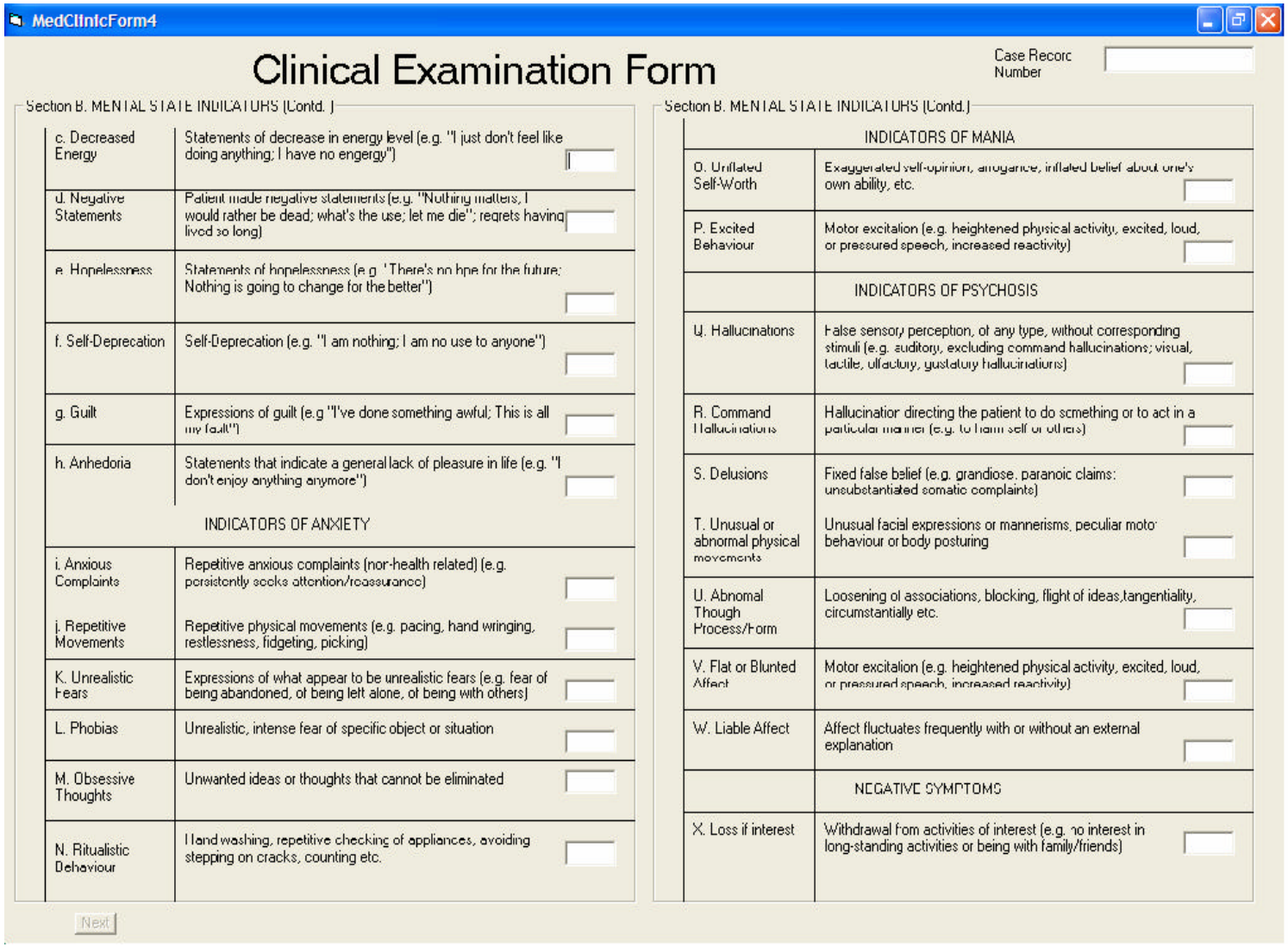
1.	COMATOSE	(*Persistent vegetative state/no discernible consciousness*) 0. No 1. Yes (**If yes, skip to Section G**)
2.	MEMORY	(*Recall of what was learned or known*) a. Short-term memory OK—seems/appears to recall after 5 minutes 0. Memory OK 1. Memory problem b. Long-term memory OK—seems/appears to recall long past 0. Memory OK 1. Memory problem
3.	MEMORY/ RECALL ABILITY	(***Check all*** *that resident was* ***normally able to recall during last 7 days***) a. Current season b. Location of own room c. Staff names/faces d. That he/she is in a nursing home e. *NONE OF ABOVE* are recalled
4.	COGNITIVE SKILLS FOR DAILY DECISION-MAKING	(*Made decisions regarding tasks of daily life*) 0. *INDEPENDENT*—decisions consistent/reasonable 1. *MODIFIED INDEPENDENCE*—some difficulty in new situations only 2. *MODERATELY IMPAIRED*—decisions poor; cues/supervision required 3. *SEVERELY IMPAIRED*—never/rarely made decisions
5.	INDICATORS OF DELIRIUM—PERIODIC DISORDERED THINKING/ AWARENESS	(*Code for behavior in the* ***last 7 days.***) ***[Note: Accurate assessment requires conversations with staff and family who have direct knowledge of resident's behavior over this time].*** 0. Behavior not present 1. Behavior present, not of recent onset 2. Behavior present, over last 7 days appears different from resident's usual functioning (e.g., new onset or worsening) a. EASILY DISTRACTED—(e.g., difficulty paying attention; gets sidetracked) b. PERIODS OF ALTERED PERCEPTION OR AWARENESS OF SURROUNDINGS—(e.g., moves lips or talks to someone not present; believes he/she is somewhere else; confuses night and day) c. EPISODES OF DISORGANIZED SPEECH—(e.g., speech is incoherent, nonsensical, irrelevant, or rambling from subject to subject; loses train of thought) d. PERIODS OF RESTLESSNESS—(e.g., fidgeting or picking at skin, clothing, napkins, etc; frequent position changes; repetitive physical movements or calling out) e. PERIODS OF LETHARGY—(e.g., sluggishness; staring into space; difficult to arouse; little body movement) f. MENTAL FUNCTION VARIES OVER THE COURSE OF THE DAY—(e.g., sometimes better, sometimes worse; behaviors sometimes present, sometimes not)
6.	CHANGE IN COGNITIVE STATUS	Resident's cognitive status, skills, or abilities have changed as compared to status of **90 days ago** (or since last assessment if less than 90 days) 0. No change 1. Improved 2. Deteriorated

SECTION C. COMMUNICATION/HEARING PATTERNS

1.	HEARING	(*With hearing appliance, if used*) 0. *HEARS ADEQUATELY*—normal talk, TV, phone 1. *MINIMAL DIFFICULTY* when not in quiet setting 2. *HEARS IN SPECIAL SITUATIONS ONLY*—speaker has to adjust tonal quality and speak distinctly 3. *HIGHLY IMPAIRED*/absence of useful hearing
2.	COMMUNICATION DEVICES/ TECHNIQUES	(***Check all that apply*** *during last 7 days*) a. Hearing aid, present and used b. Hearing aid, present and not used regularly c. Other receptive comm techniques used (e.g., lip reading) d. *NONE OF ABOVE*
3.	MODES OF EXPRESSION	(***Check all used*** *by resident to make needs known*) a. Speech b. Writing messages to express or clarify needs c. American sign language or Braille d. Signs/gestures/sounds e. Communication board f. Other g. *NONE OF ABOVE*
4.	MAKING SELF UNDERSTOOD	(*Expressing information content—however able*) 0. *UNDERSTOOD* 1. *USUALLY UNDERSTOOD*—difficulty finding words or finishing thoughts 2. *SOMETIMES UNDERSTOOD*—ability is limited to making concrete requests 3. *RARELY/NEVER UNDERSTOOD*
5.	SPEECH CLARITY	(*Code for speech in the* ***last 7 days***) 0. *CLEAR SPEECH*—distinct, intelligible words 1. *UNCLEAR SPEECH*—slurred, mumbled words 2. *NO SPEECH*—absence of spoken words
6.	ABILITY TO UNDERSTAND OTHERS	(*Understanding verbal information content—however able*) 0. *UNDERSTANDS* 1. *USUALLY UNDERSTANDS*—may miss some part/detail of message 2. *SOMETIMES UNDERSTANDS*—responds adequately to simple, direct communication 3. *RARELY/NEVER UNDERSTANDS*
7.	CHANGE IN COMMUNICATION/ HEARING	Resident's ability to express, understand, or hear information has changed as compared to status of **90 days ago** (or since last assessment if less than 90 days) 0. No change 1. Improved 2. Deteriorated



Electronic Health Record – Mental Health



MedClinicForm4

Clinical Examination Form

Case Record Number

Section B. MENTAL STATE INDICATORS (Contd.)

c. Decreased Energy	Statements of decrease in energy level (e.g. "I just don't feel like doing anything; I have no engergy")
d. Negative Statements	Patient made negative statements (e.g. "Nothing matters; I would rather be dead; what's the use; let me die"; regrets having lived so long)
e. Hopelessness	Statements of hopelessness (e.g. 'There's no hpe for the future; Nothing is going to change for the better")
f. Self-Deprecation	Self-Deprecation (e.g. "I am nothing; I am no use to anyone")
g. Guilt	Expressions of guilt (e.g "I've done something awful; This is all my fault")
h. Anhedoria	Statements that indicate a general lack of pleasure in life (e.g. "I don't enjoy anything anymore")
	INDICATORS OF ANXIETY
i. Anxious Complaints	Repetitive anxious complaints (nor-health related) (e.g. persistently seeks attention/reassurance)
j. Repetitive Movements	Repetitive physical movements (e.g. pacing, hand wringing, restlessness, fidgeting, picking)
K. Unrealistic Fears	Expressions of what appear to be unrealistic fears (e.g. fear of being abandoned, of being left alone, of being with others)
L. Phobias	Unrealistic, intense fear of specific object or situation
M. Obsessive Thoughts	Unwanted ideas or thoughts that cannot be eliminated
N. Ritualistic Behaviour	Hand washing, repetitive checking of appliances, avoiding stepping on cracks, counting etc.

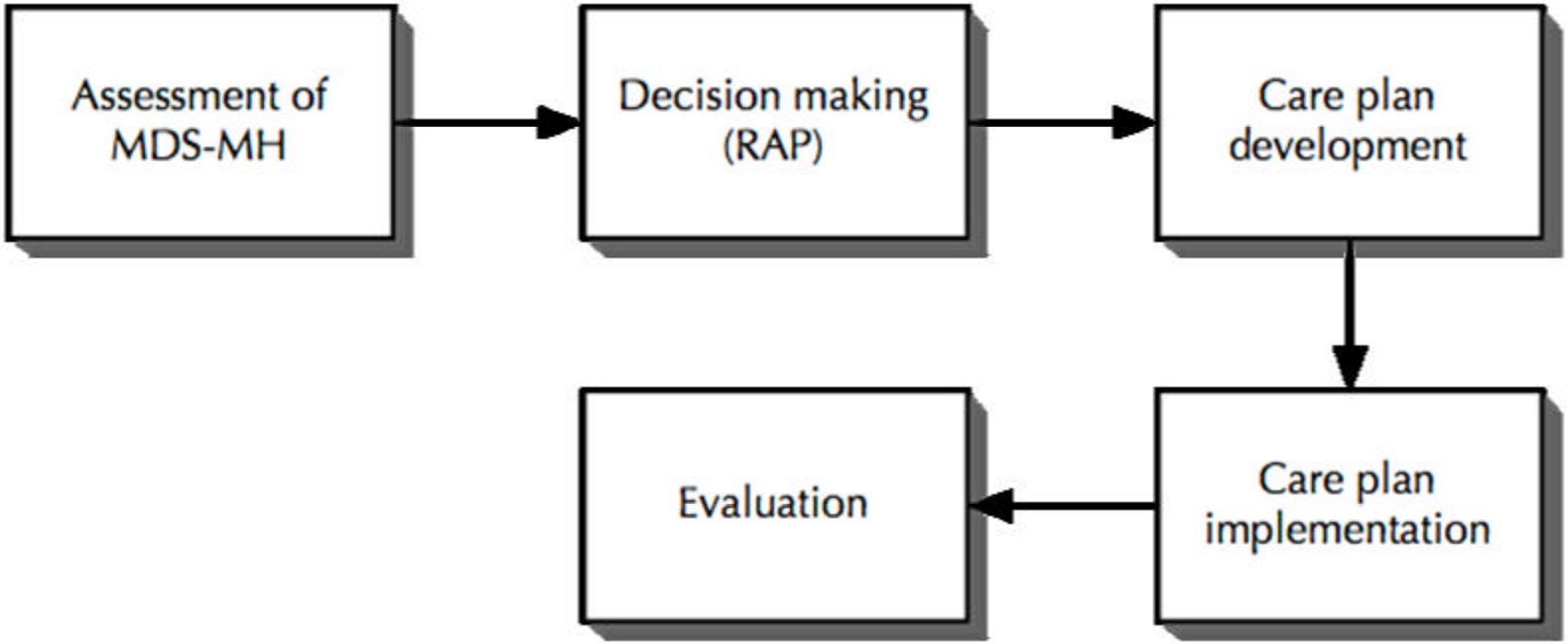
Section B. MENTAL STATE INDICATORS (Contd.)

	INDICATORS OF MANIA
O. Inflated Self-Worth	Exaggerated self-opinion, arrogance, inflated belief about one's own ability, etc.
P. Excited Behaviour	Motor excitation (e.g. heightened physical activity, excited, loud, or pressured speech, increased reactivity)
	INDICATORS OF PSYCHOSIS
Q. Hallucinations	False sensory perception, of any type, without corresponding stimuli (e.g. auditory, excluding command hallucinations; visual, tactile, olfactory, gustatory hallucinations)
R. Command Hallucinations	Hallucination directing the patient to do something or to act in a particular manner (e.g. to harm self or others)
S. Delusions	Fixed false belief (e.g. grandiose, paranoic claims; unsubstantiated somatic complaints)
T. Unusual or abnormal physical movements	Unusual facial expressions or mannerisms, peculiar motor behaviour or body posturing
U. Abnormal Though Process/Form	Loosening of associations, blocking, flight of ideas,tangentiality, circumstantially etc.
V. Flat or Blunted Affect	Motor excitation (e.g. heightened physical activity, excited, loud, or pressured speech, increased reactivity)
W. Liable Affect	Affect fluctuates frequently with or without an external explanation
	NEGATIVE SYMPTOMS
X. Loss if interest	Withdrawal from activities of interest (e.g. no interest in long-standing activities or being with family/friends)

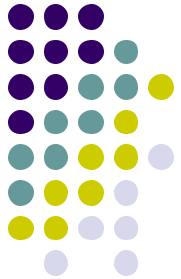
Next

Pedarla, P.(2004) "E-Intelligence form design and Data Preprocessing in Health Care", M.A.Sc. Thesis, Department of Systems Design Engineering University of Waterloo, Canada.

InterRAI assessment process

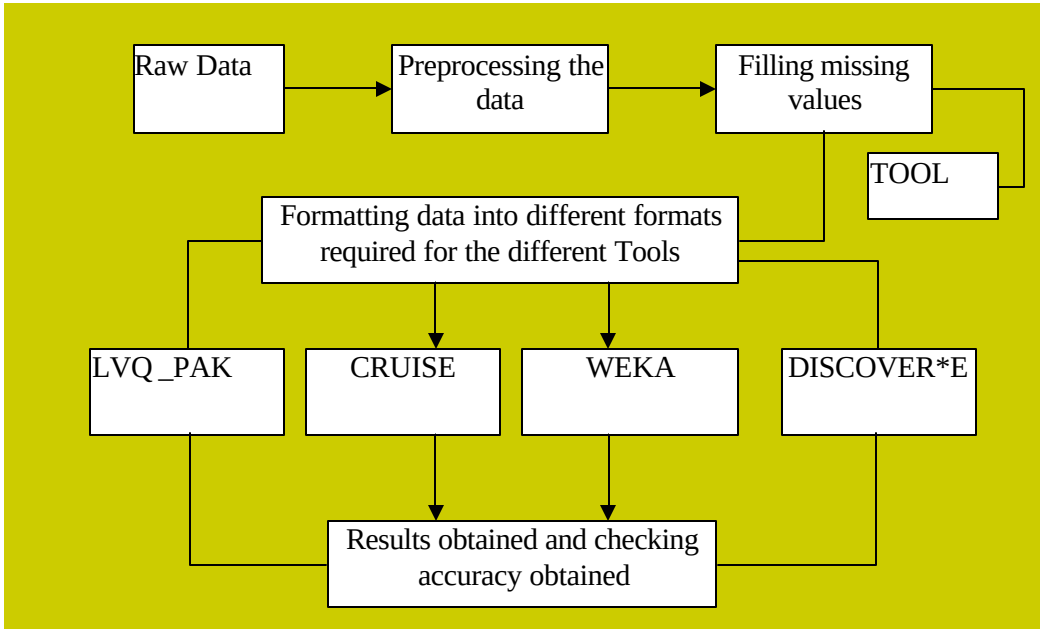


Data Mining Software & Methods to Help Decision Making

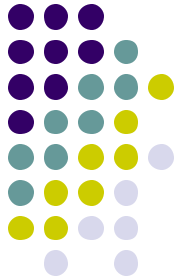


- WEKA: ZeroR, decision tree (J 4.8), decision table, naïve Bayes
- CRUISE: Univariate split decision tree; bivariate split decision tree
- LVQ-PAK: Optimized learning vector quantization
- Discover*E: Decision tree, dependence tree, rule based classification

Research Process

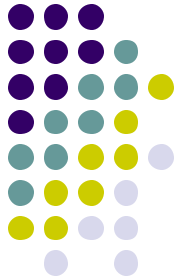


Data set



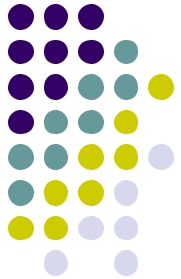
- interRAI Mental Health Data set (Ontario)
- 4,000 cases were used
- 455 variables per record
- 3,500 training cases, 500 test cases

Experiment 1



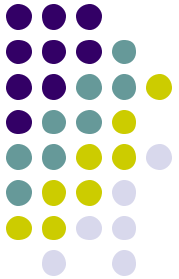
- Classifying the patients into 4 categories
  - Acute care
  - Longer-term patient
  - Forensic patient
  - Psychogeriatic patient

Experiments 2 & 3



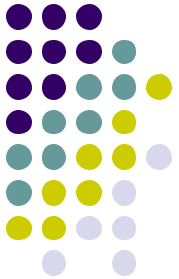
- Classifies if patient is likely to be a threat to himself or not
- Classifies if patient is likely to be threat to others

Experiments 4 & 5



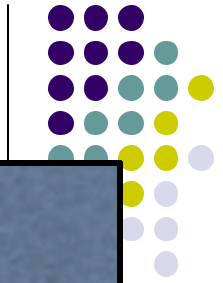
- Classifies if patient has injured himself or herself (actual event, postdiction)
- Classifies if patient has been violent to others (actual data, postdiction)

Experiments 8 & 9



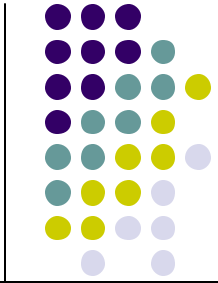
- Similar to experiments 4 &5
- Violence to self and others; removing information about past violence
  - Increased difficulty

Base case: Weka's ZeroR



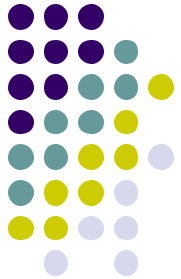
Experiment	Accuracy
1	75.75 %
2	70.74 %
3	75.69 %
4	62.72 %
5	66.53 %
6	62.72 %
7	66.53 %
8	70.74 %
9	75.75 %

Decision tree results



Experiment	Cruise univariate split	Cruise linear combination split	Discover*E decision tree	Weka J4.8 decision tree
1	80.56 %	87.71 %	87.71 %	78.16 %
2	76.55 %	77.15 %	73.74 %	73.35 %
3	78.48 %	88.84 %	81.48 %	82.27 %
4	77.35 %	80.96 %	67.07 %	71.94 %
5	83.66 %	82.16 %	80.33 %	84.17 %
6	88.17 %	86.77 %	85.11 %	86.77 %
7	86.17 %	83.97 %	83.90 %	81.16 %
8	79.16 %	77.95 %	69.92 %	71.74 %
9	79.63 %	88.77 %	79.45 %	80.76 %

Why we do need yet another eHealth seminar now?

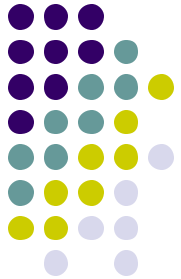


- Suggest ways to use data of ***all*** patients to do:
  - Knowledge discovery
  - Discover systemic deficiencies and efficiencies
  - Optimally allocate resources
  - Monitor and continuously predict health of patients, practices, institutions, and regions

Ideally, for the patient

- Continuous monitoring (not every minute but may be event based) and guidance for preventive care (for example, obesity, blood pressure, etc indicating problems and asking for medical visit)

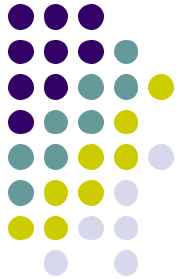
Ideally, for example, for the primary health (practitioner)



- Continuous updating of protocols, diagnostic rules, available specialist resources, billing, etc

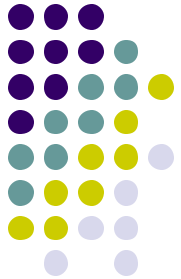
Both of the above require Knowledge discovery and other systems analysis tools

Ideally, Institutions could use



- Optimal allocation of resources
- Performance indices
- New directions (problem areas identification and requiring new resource allocations and research)

Challenges



- Secure but easy access (for practitioners and researchers)
- Large-scale data
- Need for new algorithmic developments
- Need for supercomputing (large memory and large CPU time)
- Validation (need the help of practitioners)
- $$$$