Identifying patients for clinical trials using fuzzy ternary logic expressions on HL7 messages

Raphael W. Majeed, Rainer Röhrig
Section Medical Informatics in Anesthesiology and intensive care medicine
Justus-Liebig University Gießen (Germany)

August 30th 2011, Oslo
Background

• Failure of clinical trials often due to not enough participants
• Government funded joint project to assist patient recruitment
• Communication infrastructure based on HL7 at the university Giessen
Aim: Assist patient recruitment

- *passively* listen to passing HL7 messages
  - 20000 per day, peak rate of 30 messages/s
- focus on speed and scalability
  - 40000 patients x 50 trials = 2m processes
- preferably close to real-time recruitment
- formal description of eligibility criteria needed
Methods (1)

4 distinct groups for eligibility criteria

a) Completely decidable facts (e.g. MPR, lab results)
   “age between 3 months and 30 years”

b) Partially decidable facts
   “medulloblastoma or ependymoma”

c) Undecidable facts
   “no renal disease”

d) Unprocessable facts
   “subject unlikely to comply with protocol”
Methods (2)

• Analysis of eligibility criteria
  – nation-wide recruiting trials from ClinicalTrials.gov

• Formal description of trial criteria

• Prototype implementation

• Feasibility study
  – Data structures and algorithms appropriate for real-time processing?
Results (1): Formal description

• Conjunctive normal form (CNF) expressions:
  – natural representation of eligibility criteria
  – e.g. age > 16 and hemoglobin. > 10g/dl
    and diagnosis A or B

• Three valued (ternary) logic to allow missing information
  – e.g. true AND unkn. = unkn., false AND unkn. = false

• Fuzzy logic for flexible recommendations
  – NOT a, a AND b, a OR b = (1-a), min(a,b), max(a,b)
Results (2)

• 45% of eligibility criteria are decidable
  – sufficient according to practitioners

• prototype implementation:
  – yet without evaluating expressions
  – processing MPR, diagnoses, therapies, lab results
  – 250 messages/s on desktop pc
Conclusion

• Eligibility criteria can be formally described as fuzzy ternary logic expressions

• Low descriptional complexity → good
  – all relevant problems decidable

• Limitations
  – Not all information immediately available
    (e.g. delayed entry by practitioner)
Future directions

• Complete implementation and evaluation
• Data warehouse integration
• From logic formulas to finite state machines

(a OR b) AND (c OR d OR e)

– requesting missing examinations, lab tests
– suitability for general decision support?
Thank you for your attention

Raphael W. Majeed, Rainer Röhrig:

**Identifying patients for clinical trials using fuzzy ternary logic expressions on HL7 messages**

Contact information:

raphael.majeed@chiru.med.uni-giessen.de

rainer.roehrig@chiru.med.uni-giessen.de

phone: +49 (0)641-985-44494