From data to models: incorporating uncertainty into decision support systems

Wade P Smith, PhD

New York Oncology Hematology
Albany, NY 12206

Outline

- From data to information to models
- Some basic models
- Why do we model?
- Decision support

Probabilistic vs Mechanistic models

A mechanistic model uses fundamental physics to predict an outcome

An example: Monte Carlo dose calculation

A probabilistic model links probabilities of outcomes to initial states

An example: flipping a coin.

When probabilistic?

- No need to model the mechanism
- Mechanistic too difficult
- Not enough info to form a mechanistic model
- Looking for a mechanism

Frequentist and subjective probabilities

- Frequentist:
  - rolling dice
- Subjectivist:
  - will my sports team win?
  - Can be defined with a ‘bet’

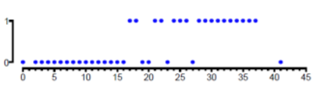
Data

↓

Information

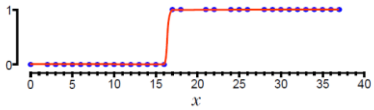
↓

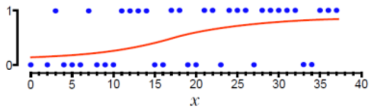
Models



Logistic Regression

$\pi(x) = \exp(\alpha + \beta x) / (1 + \exp(\alpha + \beta x))$





Model Validation

Compare your model vs another data set

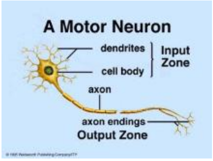
Discover that it doesn’t always work as well as you would like.

Need more variables:
Try a multivariate logistic regression, or a neural net.

Neural network

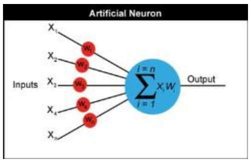
Modeling inspired by structure of neurons

Each node has a series of weighted inputs, the output ‘fires’ as a function of the weighted inputs.



Neural network

Each node has a series of weighted inputs, the output ‘fires’ as a function of the weighted inputs.



$$\text{Output} = f(\sum_{i=1}^{n} weight_i \times input_i)$$

Neural network

$$\text{Output} = f(\sum_{i=1}^{n} weight_i \times input_i)$$

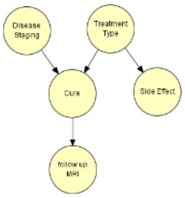
Choose an output function.

Require large amounts of data to learn from

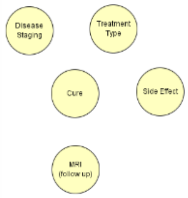
Somewhat of a black box – may make it difficult to learn mechanisms after validating the model.

Bayesian Model

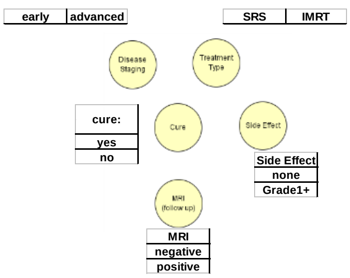
- A directed acyclic graph that represents a set of variables
- Each node contains a set of states (a conditional probability table) for that variable
- Conditional probability tables contain the quantitative dependencies between variables
- Conditional probabilities can come from:
  - Mechanistic models (e.g. TCP)
  - Prospective clinical studies
  - Retrospective analysis
  - Physician beliefs (subjective probabilities)



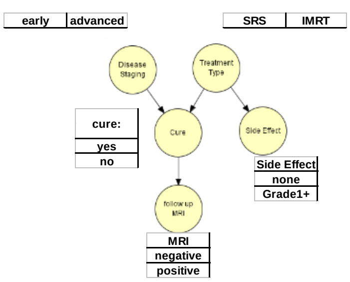
Example Bayesian model



Example Bayesian model



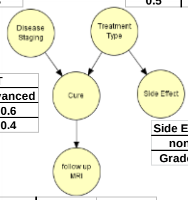
Example Bayesian model



Example Bayesian model

early	advanced
0.7	0.3

SRS	IMRT
0.5	0.5



cure:	SRS		IMRT	
	early	advanced	early	advanced
yes	0.9	0.7	0.8	0.6
no	0.1	0.3	0.2	0.4

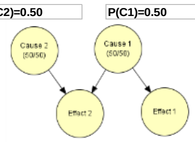
Side Effect	SRS	IMRT
none	0.9	0.7
Grade1+	0.1	0.3

MRI	yes cure	no cure
negative	0.98	0.2
positive	0.02	0.8

P(advanced,SRS,cure,no side effect,MRI+) = 0.3*0.5*0.7*0.9*0.02

Diagnostic Reasoning

P(C2)=0.50 P(C1)=0.50

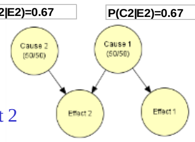


E1	E2	C1	C2
0	0	0	0
1	1	1	0
1	1	1	1
0	1	0	1

Diagnostic Reasoning

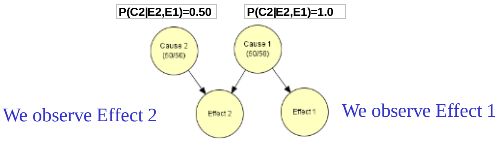
P(C2|E2)=0.67 P(C2|E2)=0.67

We observe Effect 2



E1	E2	C1	C2
~~0~~	~~0~~	~~0~~	~~0~~
1	1	1	0
1	1	1	1
0	1	0	1

Diagnostic Reasoning



E1	E2	C1	C2
~~0~~	~~0~~	~~0~~	~~0~~
1	1	1	0
1	1	1	1
~~0~~	~~1~~	~~0~~	~~1~~

Logical Reasoning
Reasoning under uncertainty

- The Partin tables predict the incidence of extra prostatic extension give PSA, Gleason Score, and Clinical Stage of the patient.
- Built using Logistic Regression
- Widely used

- How well does it perform on another data set?
- Is LR the best performing model?

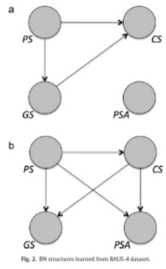
Machine learning for improved pathological staging of prostate cancer: a performance comparison on a range of classifiers.
Regnier-Coudert O, McCall J, Lothian R, Lam T, McClinton S, N'dow J.
Artif Intell Med. 2012 May;55(1):25-35

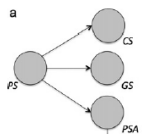
Model comparison using a different data set:
Logistic Regression,
Bayesian Network (3)
Neural Networks (2)
Random Forest
Support Vector Machine
k-nearest neighbors

Machine learning for improved pathological staging of prostate cancer: a performance comparison on a range of classifiers.

Regnier-Coudert O, McCall J, Lothian R, Lam T, McClinton S, N'dow J.
Artif Intell Med. 2012 May;55(1):25-35

Using the 3 clinical variables from the Partin tables the simplest Bayesian Network (naïve Bayes) outperformed all of the other methods.

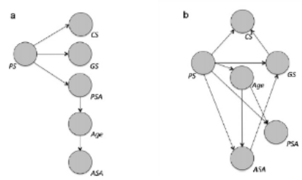




Machine learning for improved pathological staging of prostate cancer: a performance comparison on a range of classifiers.

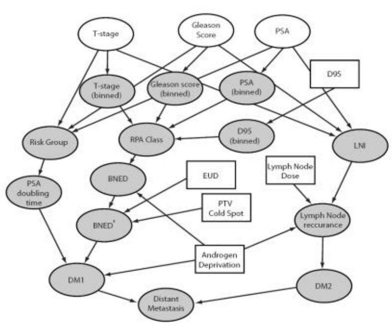
Regnier-Coudert O, McCall J, Lothian R, Lam T, McClinton S, N'dow J.
Artif Intell Med. 2012 May;55(1):25-35

With the addition of 2 clinical variables, the Bayesian Networks which allow interdependencies between variables outperform



A decision aid for intensity-modulated radiation-therapy plan selection in prostate cancer based on a prognostic Bayesian network and a Markov model.

Smith WP et. al., Artif Intell Med. 2009 Jun;46(2):119-30.



P(Gleason=3+4) = 50%
P(Gleason=4+3) = 50%

P(Tstage=T3/T4) = 50%
P(Tstage=T2) = 50%

P(PSA<6) = 5%
P(PSA=6-10) = 15%
P(PSA=10-20) = 30%
P(PSA>20) = 50%

P(DM1) = 14%

P(Gleason=3+4) = 50%
P(Gleason=4+3) = 50%

P(Tstage=T3/T4) = 50%
P(Tstage=T2) = 50%

P(PSA<6) = 5%
P(PSA=6-10) = 15%
P(PSA=10-20) = 30%
P(PSA>20) = 50%

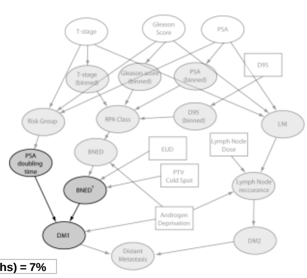
P(DM1) = 14%
P(DM1|BNED=no) = 25%

P(Gleason=3+4) = 50%
P(Gleason=4+3) = 50%

P(Tstage=T3/T4) = 50%
P(Tstage=T2) = 50%

P(PSA<6) = 5%
P(PSA=6-10) = 15%
P(PSA=10-20) = 30%
P(PSA>20) = 50%

P(DM1) = 14%
P(DM1|BNED=no) = 25%
P(DM1|BNED=no,PSA_DT>12months) = 7%

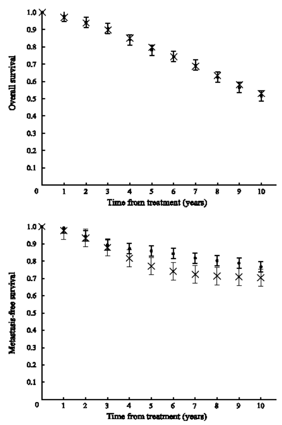


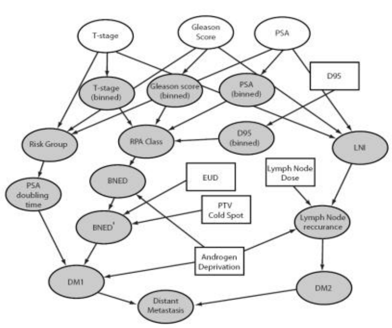
Quality Adjusted Life Expectancy (QALE)

- QALE combines life expectancy with quality of life for different possible outcomes
- Long-term prostate cancer health-states & utilities*:
  - Full health: 1.0
  - Urinary difficulty (grade 1): 0.88
  - Bowel problems (grade 2): 0.71
  - Gr1 Urinary + Gr2 Bowel: 0.70
  - Distant metastasis: 0.25
- In this model there is time-discounting so that each year is valued less than the previous year
- A Markov Model is used to compute QALE

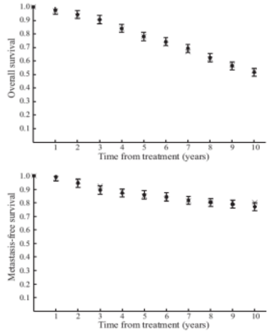
* S. T. Stewart, Med Care, 43 (2005) 347

- Comparison of earlier model prediction (x-marks) with results of a clinical study (solid dots) for Overall Survival and Metastasis-free survival.
- The model over predicts development of Mets.
- RTOG 92-02 (Horwitz JCO '08)



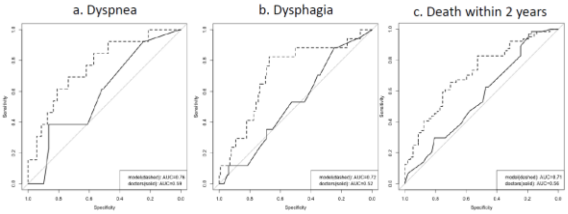


- Comparison of model prediction (x-marks) with results of a clinical study (solid dots) including the effects of androgen deprivation.



A prospective study comparing the predictions of doctors versus models for treatment outcome of lung cancer patients: A step toward individualized care and shared decision making.

Oberije C, Nalbantov G, Dekker A, Boersma L, Borger J, Reymen B, van Baardwijk A, Wanders R, De Ruysscher D, Steyerberg E, Dingemans AM, Lambin P.
Radiother Oncol. 2014 May 17. pii: S0167-8140(14)00175-3.



Decision Support

Data	Data	Data
⬇	⬇	⬇
Information	Information	Information
⬇	⬇	⬇
Models	Models	Models

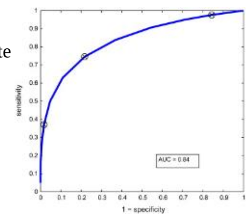
Role of positron emission tomography in the treatment of occult disease in head-and-neck cancer: a modeling approach.
Phillips MH, Smith WP, Parvathaneni U, Laramore GE.
Int J Radiat Oncol Biol Phys. 2011 Mar 15;79(4):1089-95.

Three goals:
Which point to choose on the ROC curve
When to order a PET
Which dose is optimal given the probability of disease (50 Gy or 66Gy)

Input to the model:
Probability of LN involvement for given disease location and staging.
Probability of LN involvement given a PET result (FP vs FN)
Probabilities of control from the literature
Utilities of health states

Role of positron emission tomography in the treatment of occult disease in head-and-neck cancer: a modeling approach.
Phillips MH, Smith WP, Parvathaneni U, Laramore GE.
Int J Radiat Oncol Biol Phys. 2011 Mar 15;79(4):1089-95.

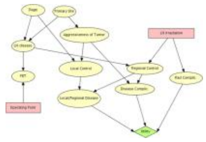
True positive rate



False positive rate

Role of positron emission tomography in the treatment of occult disease in head-and-neck cancer: a modeling approach.
Phillips MH, Smith WP, Parvathaneni U, Laramore GE.
Int J Radiat Oncol Biol Phys. 2011 Mar 15;79(4):1089-95.

Original model differs from physician practice, which leads to a different model, and an understanding of practice.



Projected second tumor risk and dose to neurocognitive structures after proton versus photon radiotherapy for benign meningioma.
Arvold et al, Int J Rad Onc Biol Phys 2012 Jul 15;83(4):e495-500.

Table 4 Projected second tumor risk and late toxicities after proton vs. photon radiotherapy

Late effect	Proton RT	Photon RT	p*
Excess risk of RT-associated tumor[†]	1.34	2.76	0.002
Mean NTCP (%) ± SE[‡]			
Brainstem	0.02 ± 0.02	0.17 ± 0.16	0.36
Temporal lobe			
Left	0.06 ± 0.02	0.14 ± 0.09	0.27
Right	0.16 ± 0.14	0.21 ± 0.18	0.24
Optic chiasm	0.60 ± 0.57	0.48 ± 0.45	0.34
Optic nerve			
Left	0.02 ± 0.01	0.08 ± 0.06	0.34
Right	0.01 ± 0.004	0.01 ± 0.01	0.30

Abbreviations: RT = radiotherapy; NTCP = normal tissue complication probability; SE = standard error.
* Paired *t* test.
† Cases per 10,000 patient-years.

Markov Model

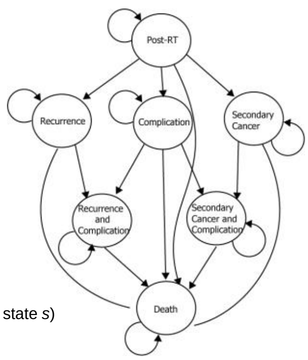
Patients are in one of several health-states

Patients have a probability to make transitions between states during each time-cycle

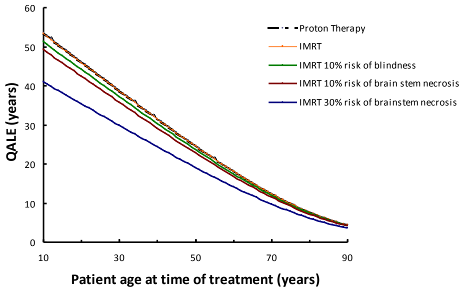
The transition probabilities from:
values from the literature
life expectancy data for a healthy individual comes from the SSA.

A simplified graphical representation is shown at the right

QALE = Σ_s(Time in state *s*) *(Utility of state *s*)



Predicted quality adjusted life expectancy (QALE) for Meningioma patients treated with IMRT or proton therapy



Summary

- The value of model validation
- Many models leads to the need for decision support

- Thank you for your time