Primer on Decision Trees

James G. Kahn, MD, MPH
Professor of Health Policy, Epidemiology, & Global Health
University of California, San Francisco

CFAR
CENTER FOR AIDS RESEARCH
UNIVERSITY OF WASHINGTON
FRED HUTCHINSON CANCER RESEARCH CENTER
SEATTLE BIOMED
SEATTLE CHILDREN'S

Overview

- What is a decision tree?
- When should I use a decision tree?
- How to construct a decision tree
- How to analyze a decision tree
- Software options

What is a decision tree?

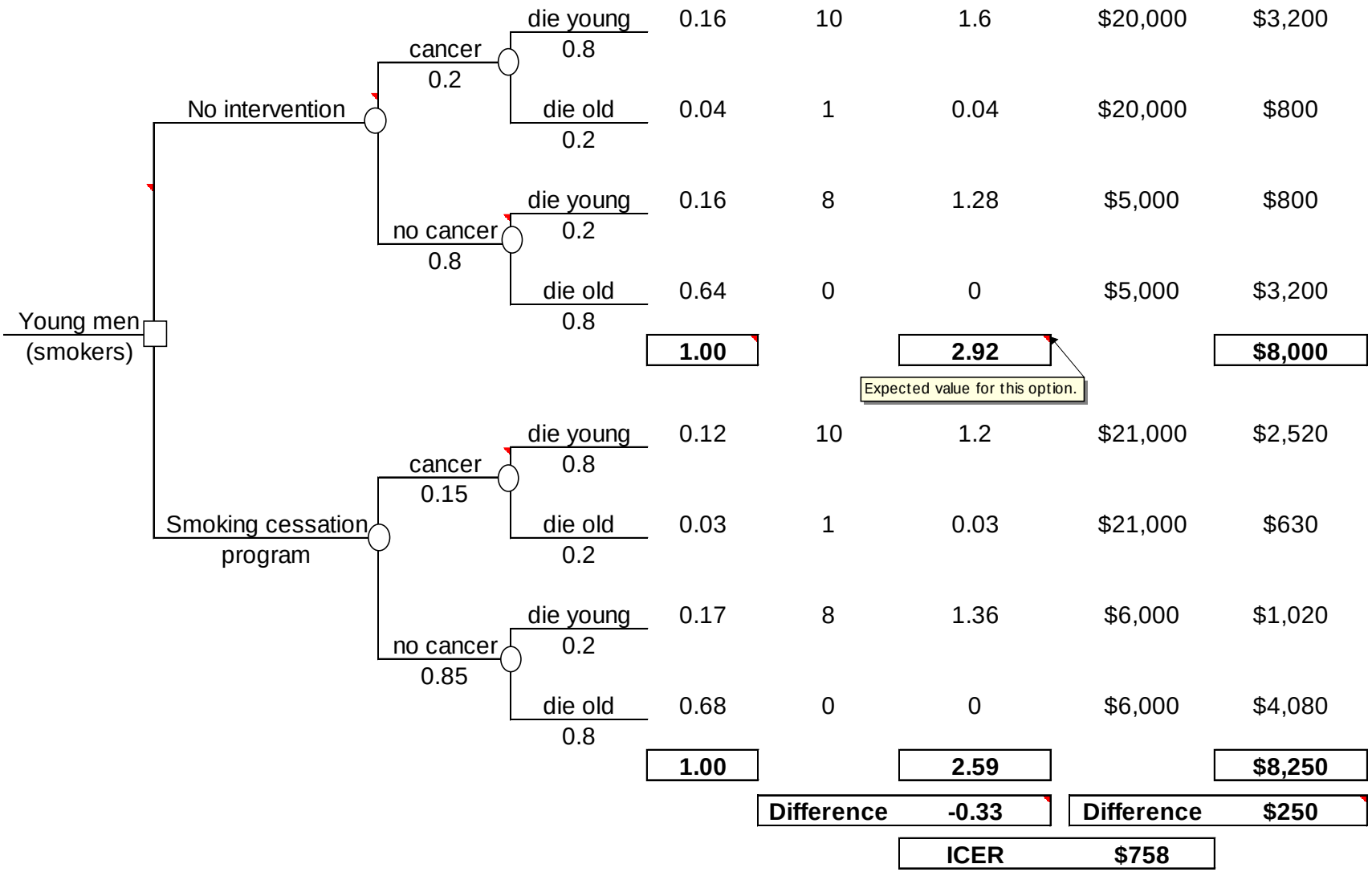
- A branching structure that leads from a **choice** (among competing courses of action) through a probability net of possible **consequences** (temporary and final) …
- … in which each path of consequences has an associated **probability** and set of **outcomes** of interest (e.g., cost and health status) such that …
- … each course of action can be assigned **expected values** for the outcomes (as the weighted mean of relevant paths) that can be compared and used to **guide decisions** among the actions.

Decision tree structure to evaluate interventions

Population	Decision node / Course of Action / Options	Chance nodes / Consequences	Path Probability	DALYs		Costs	
				For terminal nodes	Intermediate product	For terminal nodes	Intermediate product
Young men (smokers)	No intervention	cancer 0.2	die young 0.8	0.16	10	1.6	$20,000	$3,200
			die old 0.2	0.04	1	0.04	$20,000	$800
		no cancer 0.8	die young 0.2	0.16	8	1.28	$5,000	$800
			die old 0.8	0.64	0	0	$5,000	$3,200
				1.00		**2.92**		**$8,000**

Expected value for this option.

Decision tree structure to evaluate interventions



Population	Decision node / Course of Action / Options	Chance nodes / Consequences	Path Probability	DALYs		Costs	
				For terminal nodes	Intermediate product	For terminal nodes	Intermediate product
		die young	0.16	10	1.6	$20,000	$3,200
		cancer 0.2 / die old	0.04	1	0.04	$20,000	$800
		die young	0.16	8	1.28	$5,000	$800
		no cancer 0.8 / die old	0.64	0	0	$5,000	$3,200
			1.00		**2.92**		**$8,000**
					Expected value for this option.		
		die young	0.12	10	1.2	$21,000	$2,520
		cancer 0.15 / die old	0.03	1	0.03	$21,000	$630
		die young	0.17	8	1.36	$6,000	$1,020
		no cancer 0.85 / die old	0.68	0	0	$6,000	$4,080
			1.00		**2.59**		**$8,250**
				Difference	**-0.33**	**Difference**	**$250**
					ICER	**$758**	

When should I use a decision tree?

- **Conceptualizing: Almost always.** Extremely useful to develop and portray the structure of a cost-effectiveness analysis … clarify thinking, tighten logic, avoid omissions of possible paths. Can be used in conjunction with other visual portrayals of model dynamics.
- **Operationalizing: Often.** Assures that conceptual approach is reflected in implementation. Often used in conjunction with other calculation tools. Balance of tree & other calculation structures is personal preference.
- **Presenting: Sometimes.** Some analyses done with trees are presented with trees, some not.

How to construct a decision tree

- **Population & context**
- **Decision node (square)** – the question under study, 2 or more action options – all plausible (judgment call). Later decisions brought to front.
- **Chance nodes (circles)** – in each node probabilities sum to 100%. Mutually exclusive & exhaustive. Dichotomous easiest to manipulate. Markov can be incorporated.
- **Terminal node utilities = outcomes** – health, costs (direct, time)
- **Expected values** for health and costs, for each action option as weighted mean of paths.
- **Iterative revision** – unlike RCTs, the approach can (and nearly always does) change with early results and better understanding. The trick is knowing when to stop refining, and balancing completeness with transparency.

How to analyze a decision tree

- **Comparisons across options** – compare expected values for costs and health outcomes ... ordered (least to most expensive) & step-wise incremental ... then incremental cost-effectiveness ratios (ICERs)

	Cost	Δ Cost	DALYs	Δ DALYs (averted)	ICER	
Option A	$1,000		10			
Option C	$1,500	$500	8	2.0	$250	
Option B	$1,700	$200	8.5	-0.5	Dominated	
Option D	$2,500	$800	7.5	1.0	$2,000	[vs C]

- **Sensitivity analyses** – 1-way, 2-way, scenarios, thresholds, multivariate (eg Monte Carl0).

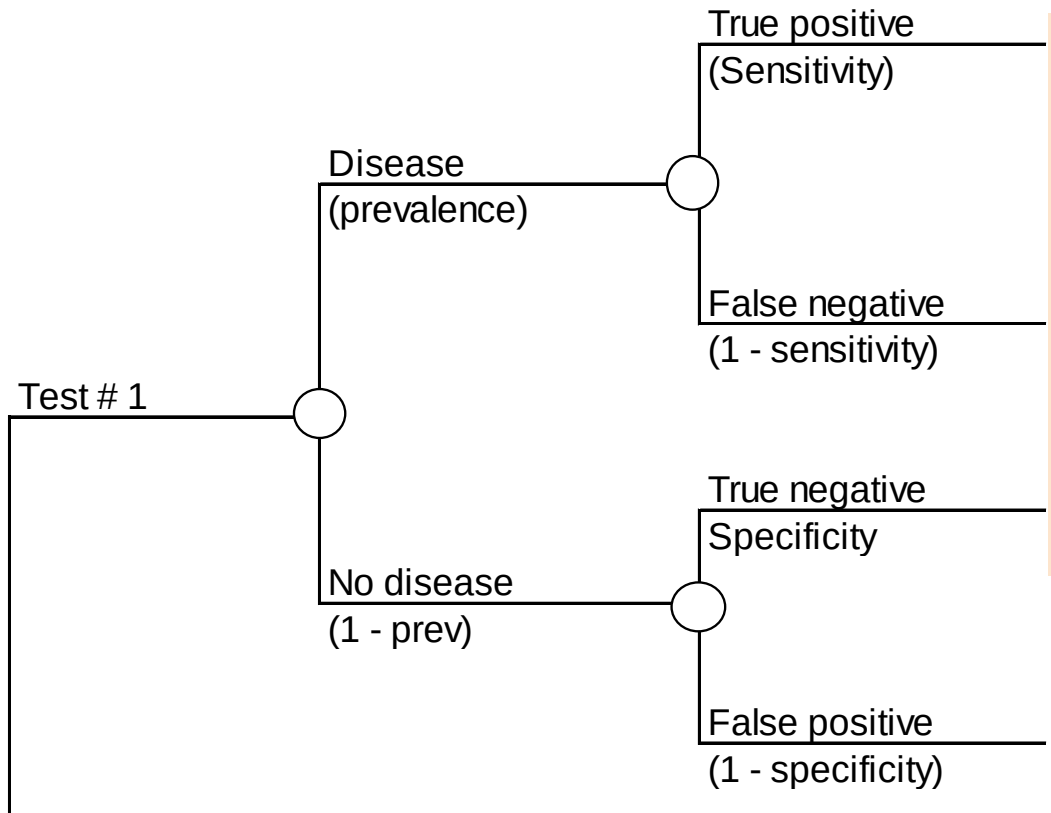
Software

- **Excel** – familiar, generic, flexible (eg incorporate epidemic and cost models), has sensitivity analysis add-ons (Crystal Ball, @Risk). My favorite. Consider starting with template.
- **TreeAge** – new, specialized, efficient for set CEA tasks, less flexible, quirky manual and implementation.
- **@Risk** – newer, specialized, efficient for set CEA tasks, powerful, complex, narrow market.

Extra credit: testing analysis *prevalence before test performance*

Tree structure to evaluate diagnostic tests

Population	Decision: Test or not	True disease prevalence (risk of disease)	Test performance*



Why?

Forces equal disease prevalence in all arms, o/w source of error

Avoid ppv, npv, etc which are more complex (prevalence-dependent)

Contact

jgkahn@ucsf.edu

CFAR
CENTER FOR AIDS RESEARCH
UNIVERSITY OF WASHINGTON
FRED HUTCHINSON CANCER RESEARCH CENTER
SEATTLE BIOMED
SEATTLE CHILDREN'S