Quantifications in Internal Decision Making Process

Yin Yin

BASS XV, November 3 - 7, 2008

Savannah, GA

The Call

- Background of statisticians working in the pharmaceutical industry
- Failed Phase III trials
- FDA's new initiatives
- New role for statisticians

Part I: Sample Size Calculation for PoC (Proof of Concept)

Part II: Decision Issues

Part III: Statisticians' Role in Internal Decision Making Process

Part I

Sample Size Calculation for PoC

A Bayesian Approach

Overview

- Objectives of PoC
- Current practice
- Proposed method
- Example
- Summary

The Objectives of PoC

- **Assessing the probability of success** to make a Go/No-Go decision
  (quantified definition of success)

- **A learning experience –** usually not clearly discussed
  (More than a hypothesis testing and more than one hypothesis)

Assessing the Probability of Success

Making a GO decision

$$\text{if } P(Y \geq t \mid X_{PoC} = x_{PoC}) \geq P_1$$

Y - endpoint in confirmatory study, e.g. $HbA1c_{6mon}$ ⬇

t - target for Y in confirmatory study, e.g. 0.7% - success

X - marker of Y, e.g. Fasting Glucose (much faster than Y)

X_{PoC} – X from a PoC

x_{PoC} – A realization of X_{PoC}

P_1 - the smallest probability of success entailing a GO Decision (can be smaller than 1-a used in a confirmatory study)

The Current Practice

- H_0: $T_{active} = T_{placebo}$
- Identify a 'clinically significant' change for the endpoint of PoC, X,
  - e.g. 30mg/dL drop in 1 month fasting glucose
- Use it (30mg/mL) as the alternative hypothesis
- Calculate a sample size such that with a 90% chance we will see a significant p-value (<0.05)
  - There is only < 5% chance to see
  - There is >=90% chance to see th

Recall the objectives of PoC are:
- Assessing PoS (S: HbA1c reduction >0.7%)
- Learning for later phases

Does this address the objective of a PoC?

α and β

- α = Probability of False Positive (PoF+)
- β = Probability of False Negative (PoF-)

In a PoC, if h is the hurdle for Go/No-Go,

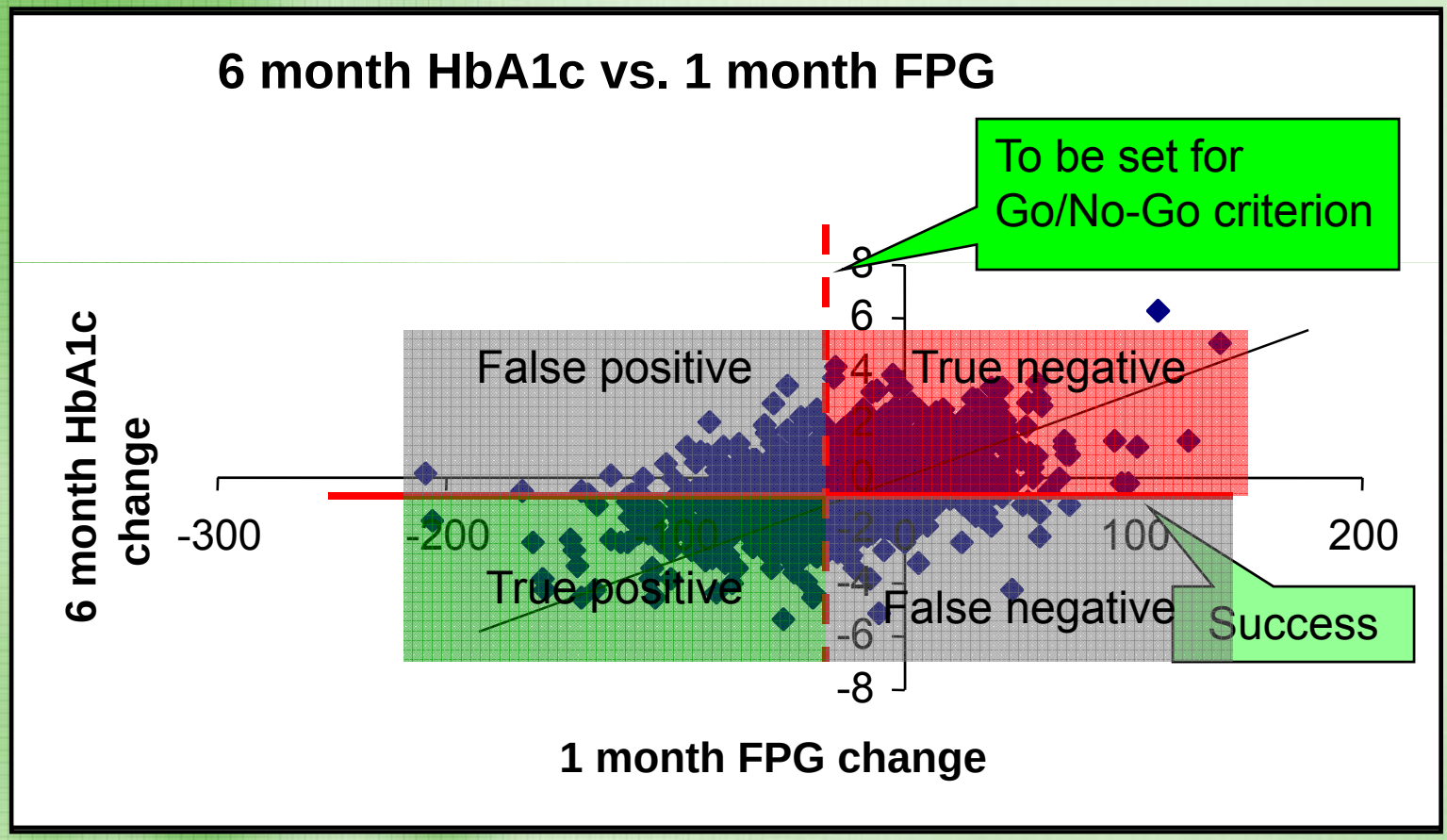
i.e. Go if $X_{PoC} \geq h$ then

- $\alpha = P(X_{PoC} \geq h | Y<t)$ (= probability of Making a wrong Go decision)
- $\beta = P(X_{PoC} < h | Y \geq t)$ (= probability of Making a wrong No-Go decision)

- α and β can be derived from the conditional distribution of Y given X_{PoC} and the distribution of X_{PoC} (both can be learned from datamining).
- α and β should be chosen to minimize the loss:
  - e.g. Loss = α * Cost of "Go" (Dev. cost)
    + β*Cost of "No-Go" (Value of the drug)

under certain constraints, e.g. sample size, not necessarily 0.05 and 0.1

Example



Relationship between Y and X

To calculate the posterior distribution of Y given X

Example (cont.)

Based on safety and/or other factors

Success: ≥0.7% (t) reduction in HbA1C at 6 months

Y: HbA1C at 6 months in the confirmatory study (N = 100)

X: Fasting Glucose at 1 month in the confirmatory study (N=100)

$$Y_i = a + bX_i + \varepsilon_i$$

Assume (not real) a = -0.1, b = 0.015.

X_{PoC} *:* Fasting Glucose at 1 month in the PoC study (N_{PoC} = ?)

If the endpoint used in Phase III is the same as in PoC, then a=0, b=1 and σ_Y =0.

Example (cont.) (Conditional mean and variance)

$$\mu_{\bar{Y}|\bar{X}_{PoC}} = E(\bar{Y} \mid \bar{X}_{PoC} = \bar{x}_{PoC}) = E(E(\bar{Y} \mid \bar{X}) \mid \bar{X}_{PoC} = \bar{x}_{PoC})$$

$$= E(a + b\bar{X} \mid \bar{X}_{PoC} = \bar{x}_{PoC})$$

$$= a + b\left(\bar{x}_{PoC} - (\bar{x}_{PoC} - \mu_{0,X}) \frac{\sigma_X^2}{N_{PoC}\sigma_{0,X}^2 + \sigma_X^2}\right)$$

$$\sigma^2_{\bar{Y}|\bar{X}_{PoC}} = Var(\bar{Y} \mid \bar{X}_{PoC} = \bar{x}_{PoC})$$

$$= \frac{\sigma_Y^2}{N} + b^2\left(\frac{\sigma_X^2}{N} + \frac{\sigma_{0,X}^2\sigma_X^2}{N_{PoC}\sigma_{0,X}^2 + \sigma_X^2}\right)$$

Note: Sample size and prior variance are always together, which means when the prior is non-informative (large prior variance), the effect on the posterior is the same as having a large N_{PoC}, that makes the data more "believable".

If the endpoint used in Phase III is the same as in PoC, then a=0, b=1 and $\sigma_Y = 0$.

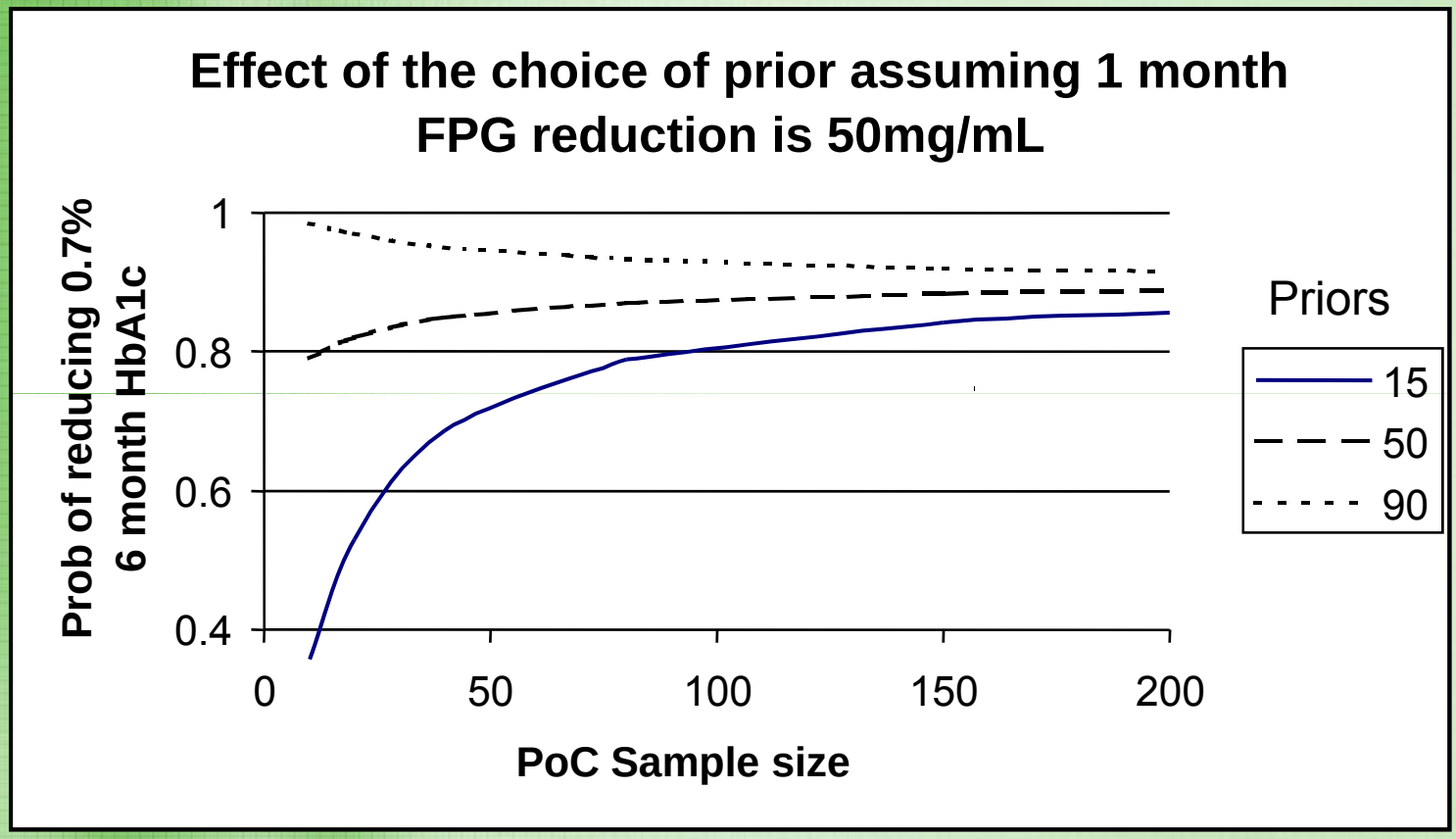
Example (cont.)

GO decision will be made if

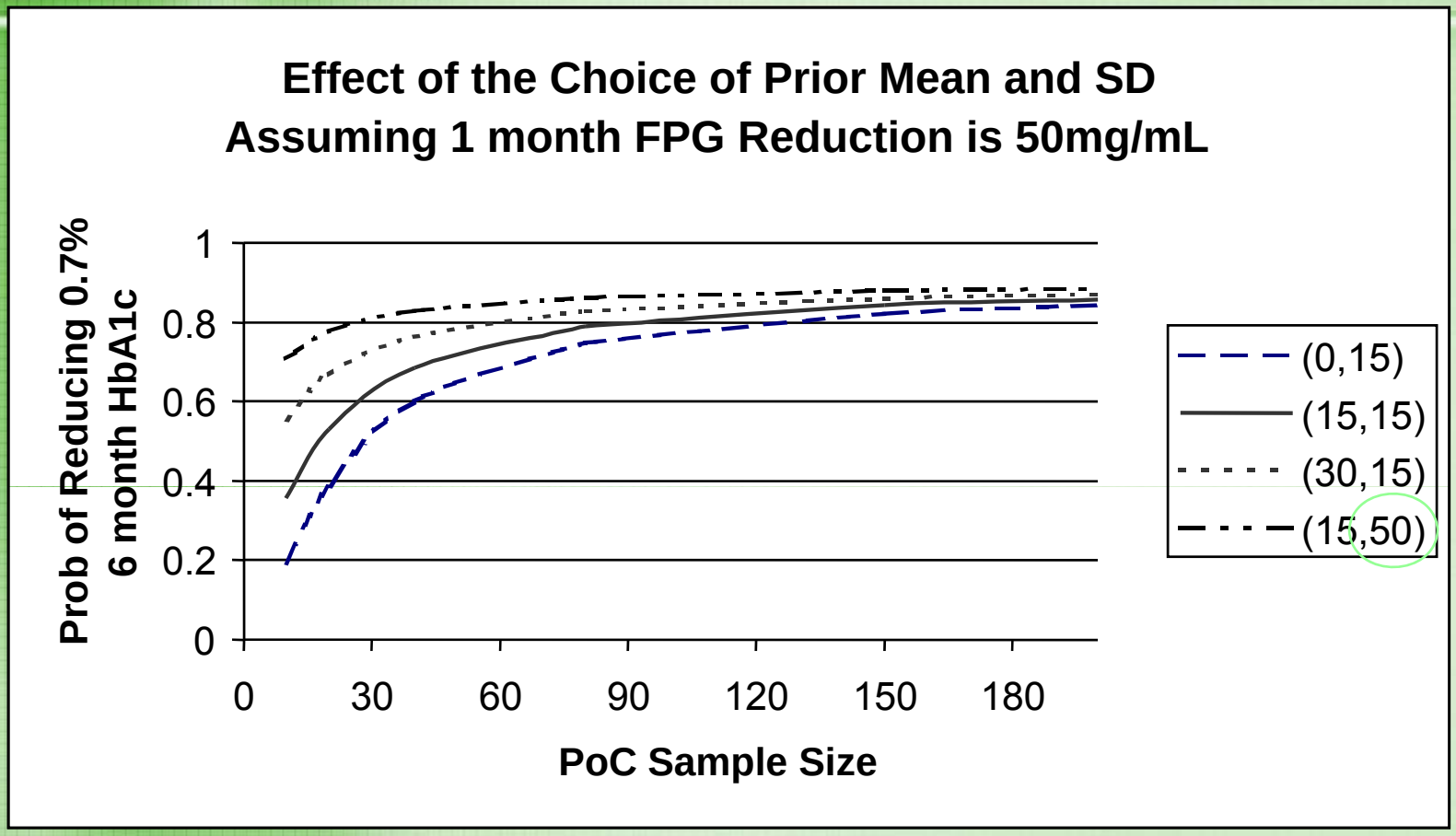
$$P(\overline{Y} \geq t \mid \overline{X}_{PoC}) = P\left(\frac{\overline{Y} - \mu_{\overline{Y}|\overline{X}_{PoC}}}{\sigma_{\overline{Y}|\overline{X}_{PoC}}} \geq \frac{t - \mu_{\overline{Y}|\overline{X}_{PoC}}}{\sigma_{\overline{Y}|\overline{X}_{PoC}}} \middle| \overline{X}_{PoC} \right) \geq P_1$$

Or to solve for N_{PoC} or x_{PoC}

$$\frac{t - \mu_{\overline{Y}|\overline{X}_{PoC}}}{\sigma_{\overline{Y}|\overline{X}_{PoC}}} \leq z_{1-P_1}$$

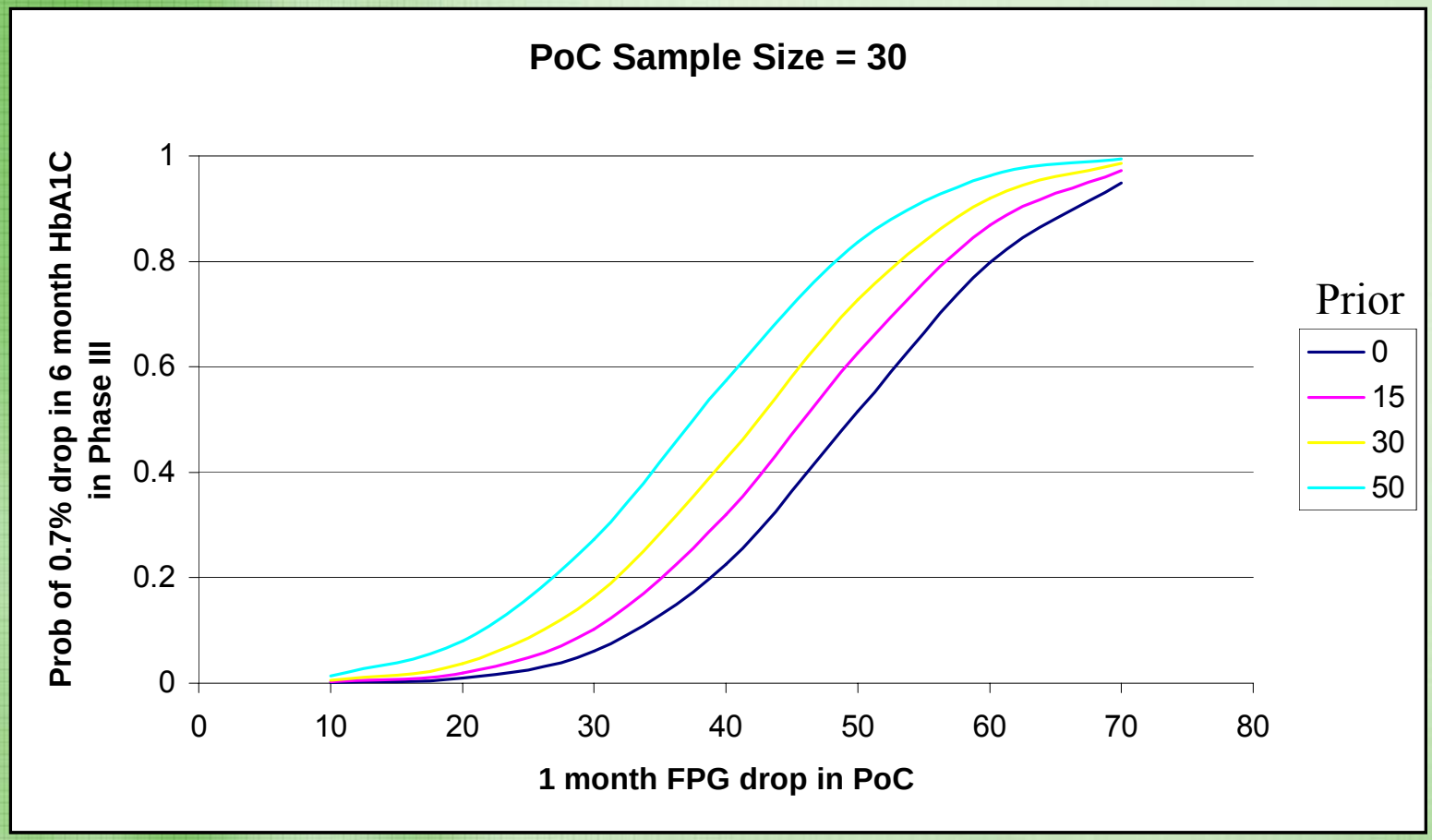


Similarly, when the prior (90) is much better than the outcome, the more patients in the PoC, the more believable the outcome is than the prior, so the lower the conditional probability is.



When $\sigma_{0,x}$ is large, the effect on the conditional mean is equivalent to having a large N_{PoC}, see Slide 12

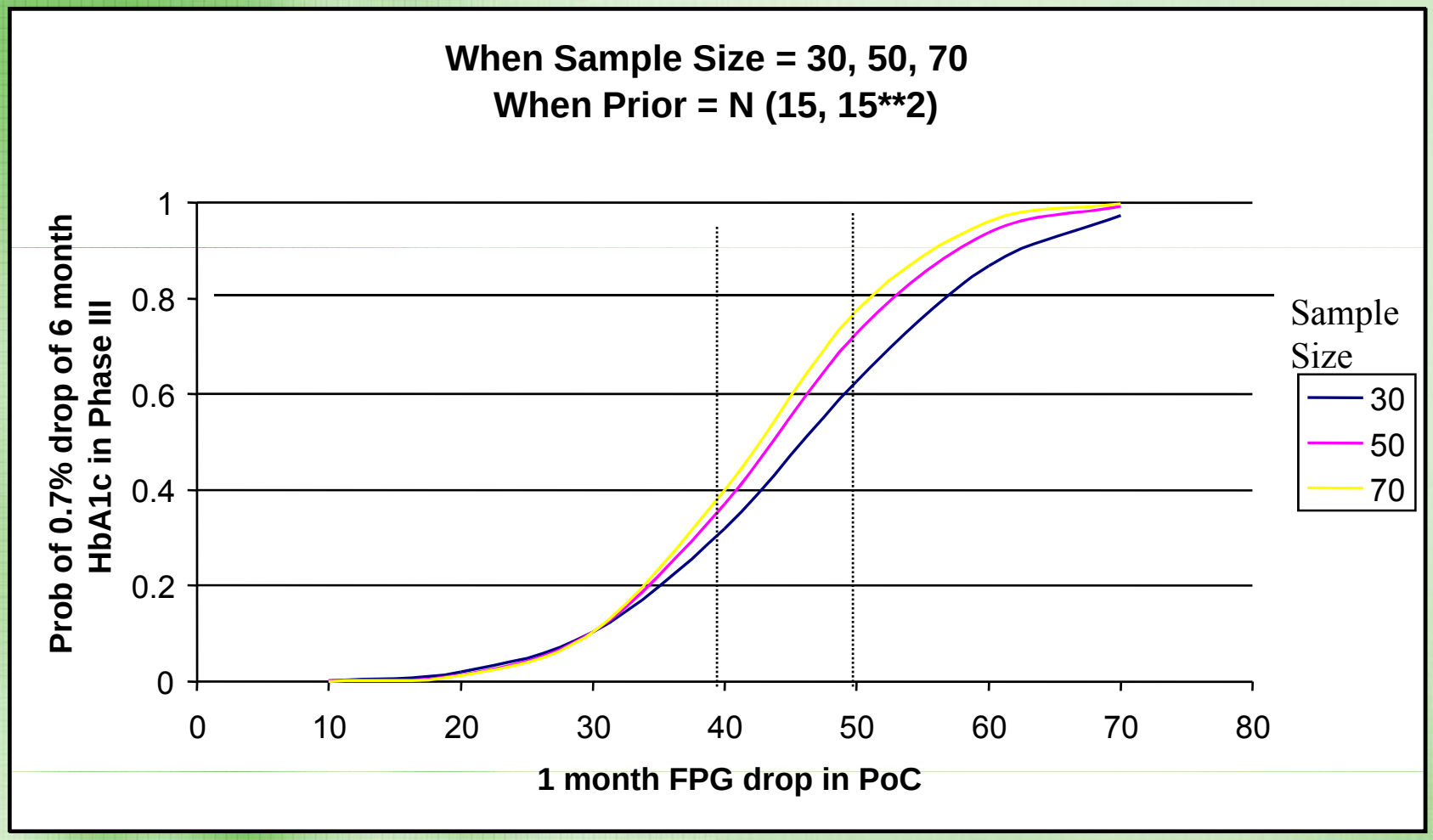
How does the outcome of PoC affect our confidence (probability of success) given the sample size of PoC?



The better the PoC outcome, the higher the probability

Since sample size is small, priors make huge differences

How does the outcome of PoC affect our confidence (probability of success) given the sample size of PoC? (cont.)



Sample size can make a big difference for some PoC outcomes

Is it worth it to get 0.1 higher prob by doubling the sample size?

Probability of success with the sample size calculated for hypothesis testing

Prior (Mean, SD)	FPG reduction in PoC[1]	N_{PoC}[2]	Probability[3]
(0, 15)	20	36	0.009
(15, 15)	20	36	0.017
(30, 15)	20	36	0.032
(0, 15)	40	12	0.108
(15, 15)	40	12	0.219
(30, 15)	40	12	0.378

[1] Used as alternative hypothesis and assumed being the outcome of PoC

[2] Sample size needed for the PoC at 5% significant level and with 90% power assuming FPG reduction in the PoC in the second column to be the alternative hypothesis

[3] The probability of reducing the mean HbA1c of 100 patients by 0.7% in the 6 month study based on the assumptions listed in the example

By-Product

The information needed for

- reasonable priors
- quantified relationship between the endpoints
- Etc.

drives discovery and motivation for information gathering from the literature and other data sources.

Every time when this type of questions were asked, more articles were circulated within the project team

PoC as a learning experience

Especially when the result is in the gray area: $0 < P(Y \geq t | X_{PoC} = x_{PoC}) < P_1$

Need hypotheses before the PoC (scenario analyses)

Hopefully PoC answers some questions

Summary

The objective of PoC is different from that of confirmatory studies, therefore sample size calculation method might be different. Bayesian provides the probability of success.

Precision required for assumptions (e.g. priors) in using this method drives more aggressive information searching from every functional area – *the byproduct*

Part II

Decision Issues

Motivation: How to choose P_1 (the Minimum PoS for Go)

Decision trees?

Decision Issues

Questions we ask everyday

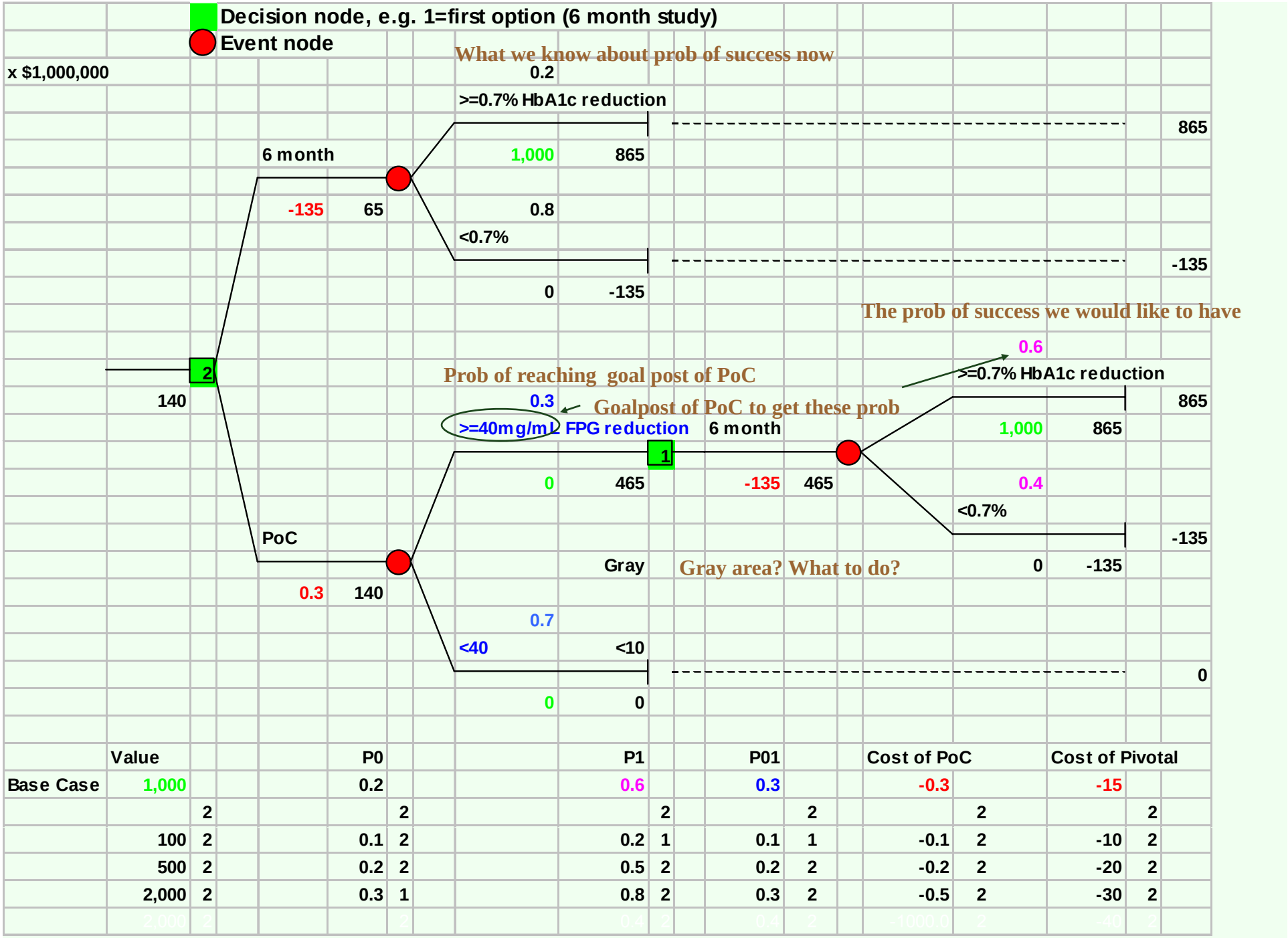
- One month study or two month study?
  - 1 month, it is cheaper and faster
  - 2 months, it is more informative, adds more ‘value’
- Which patient population should be used?
- How to balance cost, time and quality?
- Go/ No-Go
- Which biomarker?
- ……

Decision Issues

Don't

Questions we ask everyday

- One month study or two month study?
  - 1 month, it is cheaper and faster
  - 2 months, it is more informative, adds more 'value'
    - How much more cost or more time (quantified) ?
- Which patient population should be used?
  - What would be the consequences (quantified) ?
- How to balance cost, time and quality?
  - Objectives (quantified goalpost, risk, return)?
  - How to measure quality (quantified) ?
- Go/ No-Go
  - Whose decision (Upper management or project team)?
  - (Quantified) Criteria?
- Which biomarker?
  - What will be the endpoints for the later study and what is the quantitative relationship between them?



	Value			P0			P1			P01			Cost of PoC			Cost of Pivotal		
Base Case	1,000			0.2			0.6			0.3			-0.3			-15		
		2			2			2			2			2			2	
	100	2		0.1	2		0.2	1		0.1	1		-0.1	2		-10	2	
	500	2		0.2	2		0.5	2		0.2	2		-0.2	2		-20	2	
	2,000	2		0.3	1		0.8	2		0.3	2		-0.5	2		-30	2	
	2,000	2			2		0.4	2		0.4	2		-1000.0	2		-40	2	

What should be P_1?

It depends on many factors:

- **Value of the drug**
- **Cost of the pivotal study**
- **Cost of PoC**
- **Time**
- **Psychology**
- **etc.**

Many parameters on the tree are related.

What happened in the real world? The by-product

- Physiological models are needed for the specific mechanisms
- However, from extensive datamining, we learned that
  - the washout period should be eliminated
  - Different patient populations should be enrolled to support different labeling.

Summary

- Need more quantified questions
- Need a big picture
- Need criteria

Decision tree can set up a structure for doing all these

Summary (cont.)

Decision trees

- promote inquiry
- set criteria
- create a big picture for the whole team and for the upper management
- help identify all options
- impel accurate quantitative information collection from
  - Pre-clinical/Clinical database/Literature
  - Commercial
  - etc.

*By the end of the day, it will not be about the tree, it will be all about **HOW** you get the tree and the learning along the way…*

Part III

Statisticians' Roles in Internal Decision Making Process

Statisticians' Role

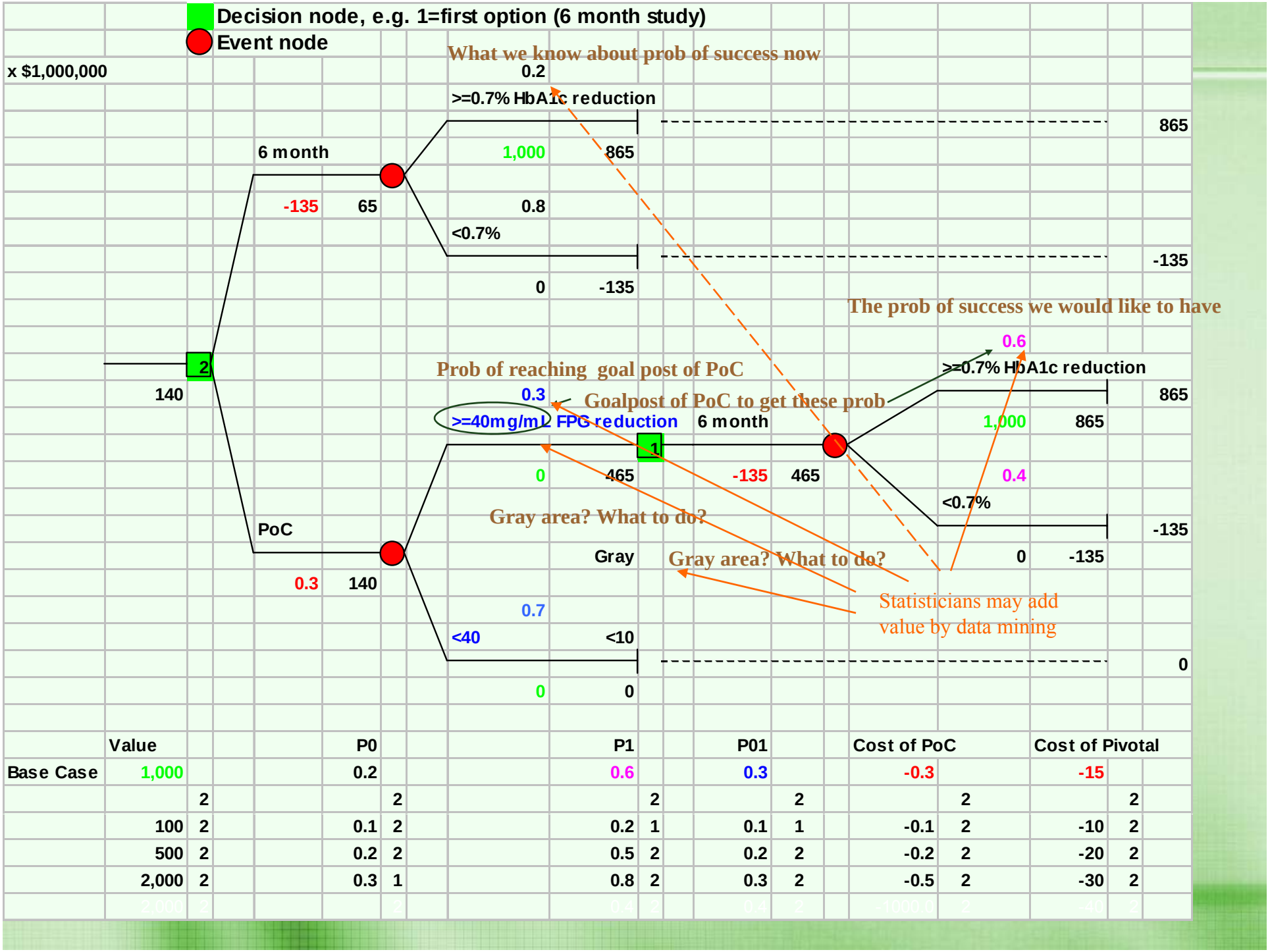
Advantages

- Logical thinking
  - Asking quantitative questions
  - Making hypotheses – the drive for scientific discoveries
- Background
  - Probability (conditional/Bayesian)
  - Fast learner of new tools
- Data ('The Final Product')
  - Data oriented
  - Knowledge and experience in dealing with data
  - Access to database
- Interactions/Connections

Statisticians' Role (cont)

Improvement Needed

- From rejecting (or not rejecting) a hypothesis to programming a decision process by using decision theory
- From doing individual data analysis to data mining, planned database and data warehouse building
- From meeting report time lines to contemplating/proposing strategy
- From being innocently blind to scientifically informed (e.g. pharmacology and physiological modeling)
- From providing services to taking leadership (which is doing the homework and providing information to influence decision makings)

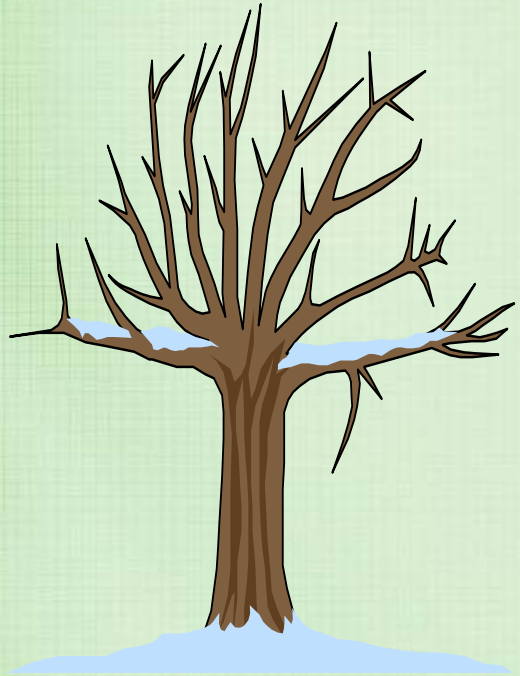


	Value		P0		P1		P01		Cost of PoC		Cost of Pivotal	
Base Case	1,000		0.2		0.6		0.3		-0.3		-15	
		2		2		2		2		2		2
	100	2	0.1	2	0.2	1	0.1	1	-0.1	2	-10	2
	500	2	0.2	2	0.5	2	0.2	2	-0.2	2	-20	2
	2,000	2	0.3	1	0.8	2	0.3	2	-0.5	2	-30	2
	2,000	2		2	0.4	2	0.4	2	-1000.0	2	-40	2

Summary

- Today's change of the scope and nature of problems and the understanding of them
- Efficiency measured quantitatively
- Information
- Quantified decision making process
- Opportunities for statisticians

A TREE



A DECISION ONE?

A TREE + A STATISTICIAN

AFTER ADDING INFORMATION - A LOT OF WORK)

A TREE + A STATISTICIAN + OTHERS

INFORMATION SHARING

References

- Y. Yin: Sample size calculation for a proof of concept study, Jour. of Biopharm.Stat. Vol12 No. 2, 2002
- Gelman, Carlin, Stern and Rubin: Bayesian Data Analysis, Chapman & Hall/CRC reprint 2000
- R.Clemen and T. Reilly: Making Hard Decisions with DecisionTools, Duxbury Thomson Learning
- F. Rockhold: Strategic Use of Statistical Thinking in Drug Development, Statistics in Medicine, 2000;19