IRT, CAT, Item Banks and "Classical" Approaches to QOL Assessment: Can't we all just get along?

Jeff. A. Sloan, Ph.D.
Mayo Clinic Comprehensive Cancer Center
Rochester, MN

MAYO CLINIC

Philosophical Question

Are the controversies and challenges to IRT and item banks unique?

Philosophical Answer

No

Philosophical Question #2

How do we overcome the challenges facing the development of a national item bank and usage of IRT and CAT?

Philosophical Answer #2

Look back

Voices of wisdom…..

- "Automobiles will never replace the horse"

- "The telephone is a devil's device that will suck your soul into it"

- "There is no need for most people to have a computer in their home"

- "The real money is in the hardware"

- "I have to get to the bank before it closes"

- "You have a cell number? When were you in prison?"

What is being said about item banks and IRT?

- "IRT will lead to a common definition of human symptoms in 10 years"

- "IRT/CAT are exciting tools…at times has led to excessive enthusiasm"

How do we pick the "ultimate" measure from the crowd?

How do we measure pain?

McGill Pain Questionnaire (20)

Brief pain inventory (BPI) (13/4/1)

Subscales of EORTC/FACT/SF (Varies)

Neuropathic pain scale (4)

Memorial pain thermometer (2)

FACES scale (1-3)

Single-item (LASA) (1)

How about a pain item bank?

It can be something simple

Please rate your pain

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

**No
Pain**

**Pain
as bad as
it can be**

JCAHO (2000) mandated intake pain assessment

How do we measure depression?

Centre for Epi Studies – Depression (CES-D) (20)

Hamilton anxiety depression scale (HADS) (30)

State-trait anxiety inventory (STAI) (60)

Profile of mood states (POMS) (30/60)

Beck depression inventory (BDI) (22)

Zung depression scale (ZDS) (12)

Yale geriatric Scale (Yale) (1)

Single-item (LASA) (1)

All are useful, None is the accepted leader

Item bank or something simple?

Please rate how depressed you were over the past week

 0   1   2   3   4   5   6   7   8   9   10

Not
Depressed

Depression
as bad as
it can be

Demonstrate the value added of an item bank

Lessons from History:
can we make lightning strike twice?

The Application of Survival Curves in Health Care Research

- **1927 William Gossett develops the t-test to make better Guinness (Student)**

- **1950's (Non)parametric alternatives appear to solve some problems**

- **1958 Kaplan-Meier present estimates to correct for missing data**

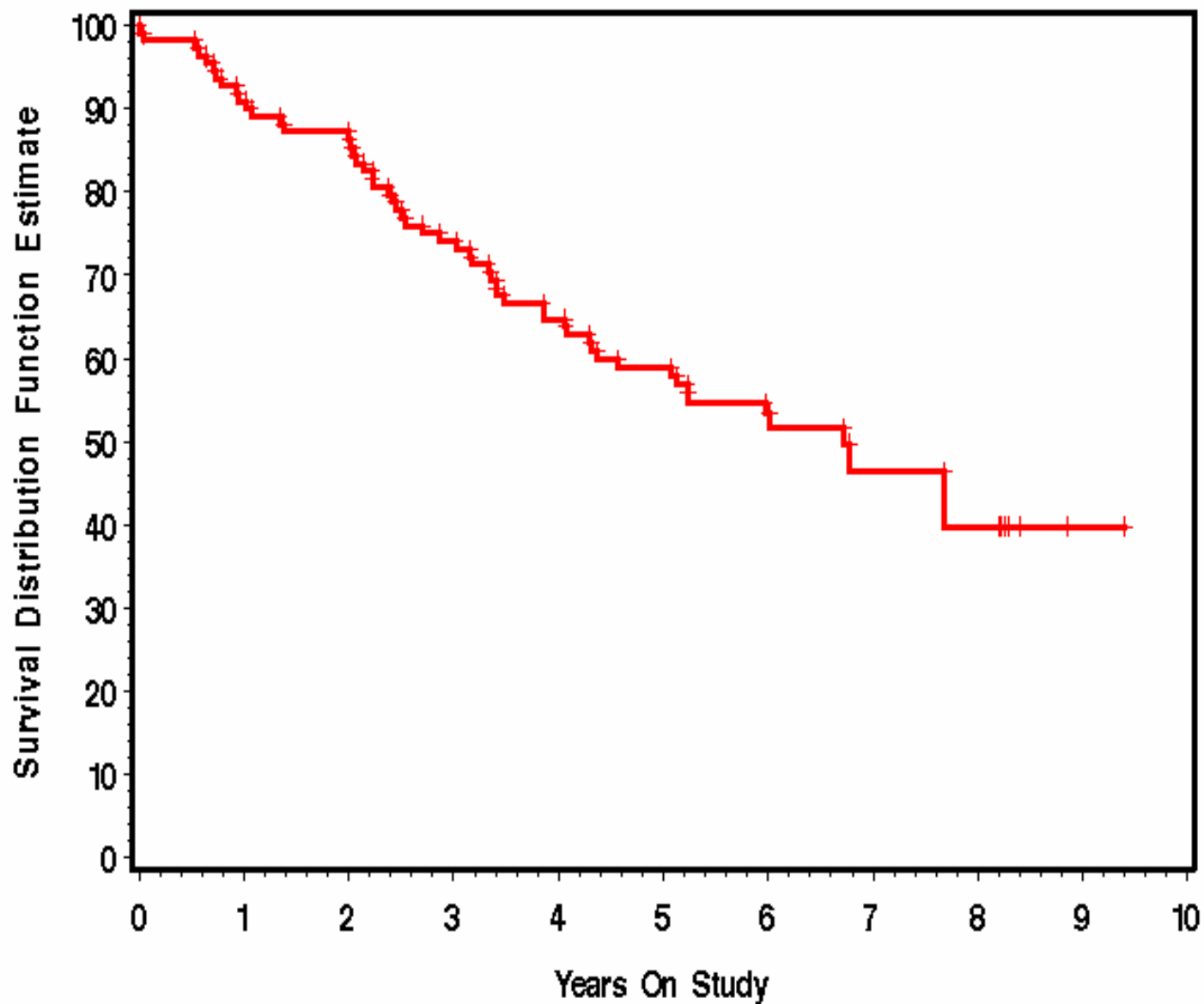- **1972 Cox proportional hazards model**
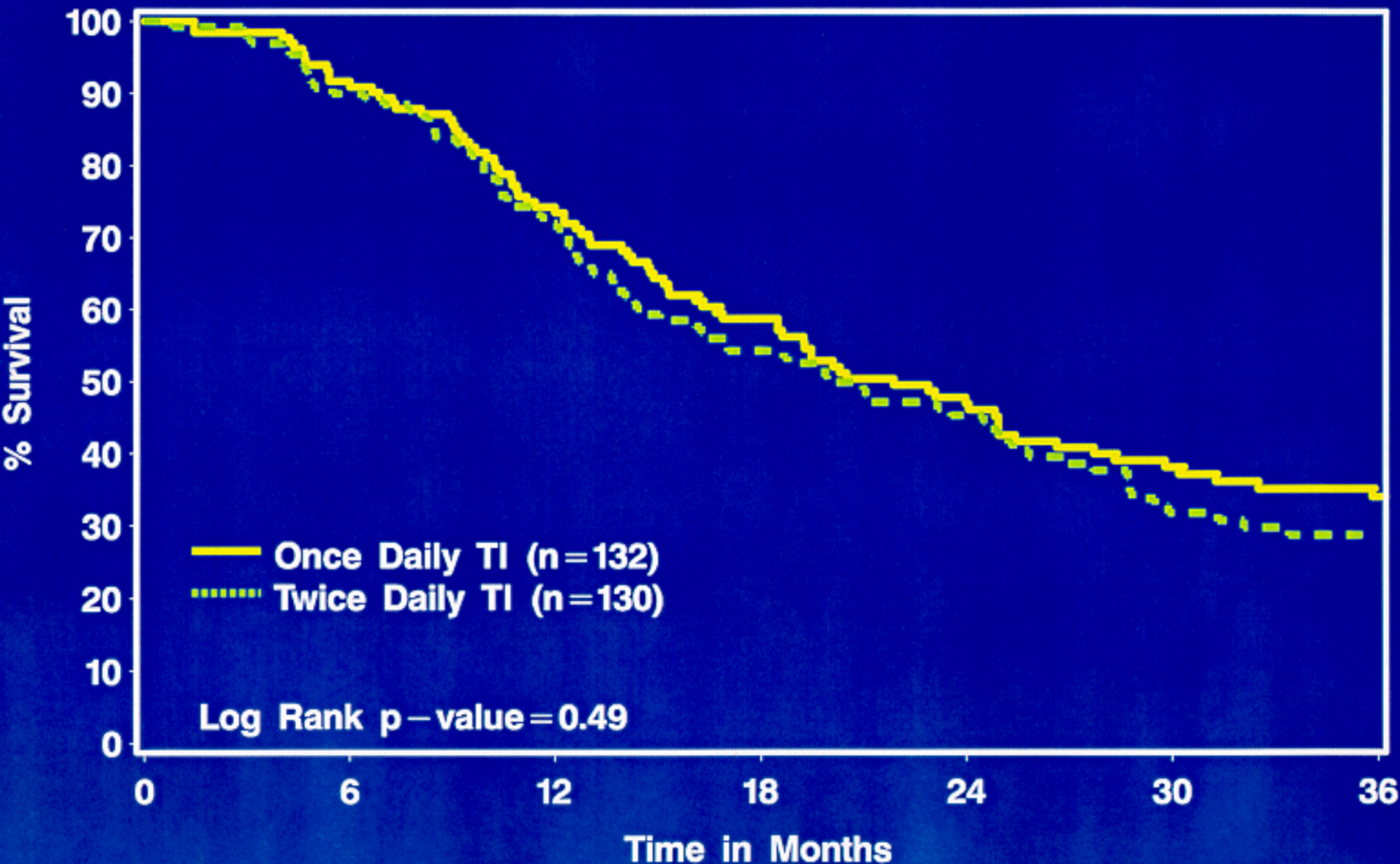
Survival Curve

No Missing Data
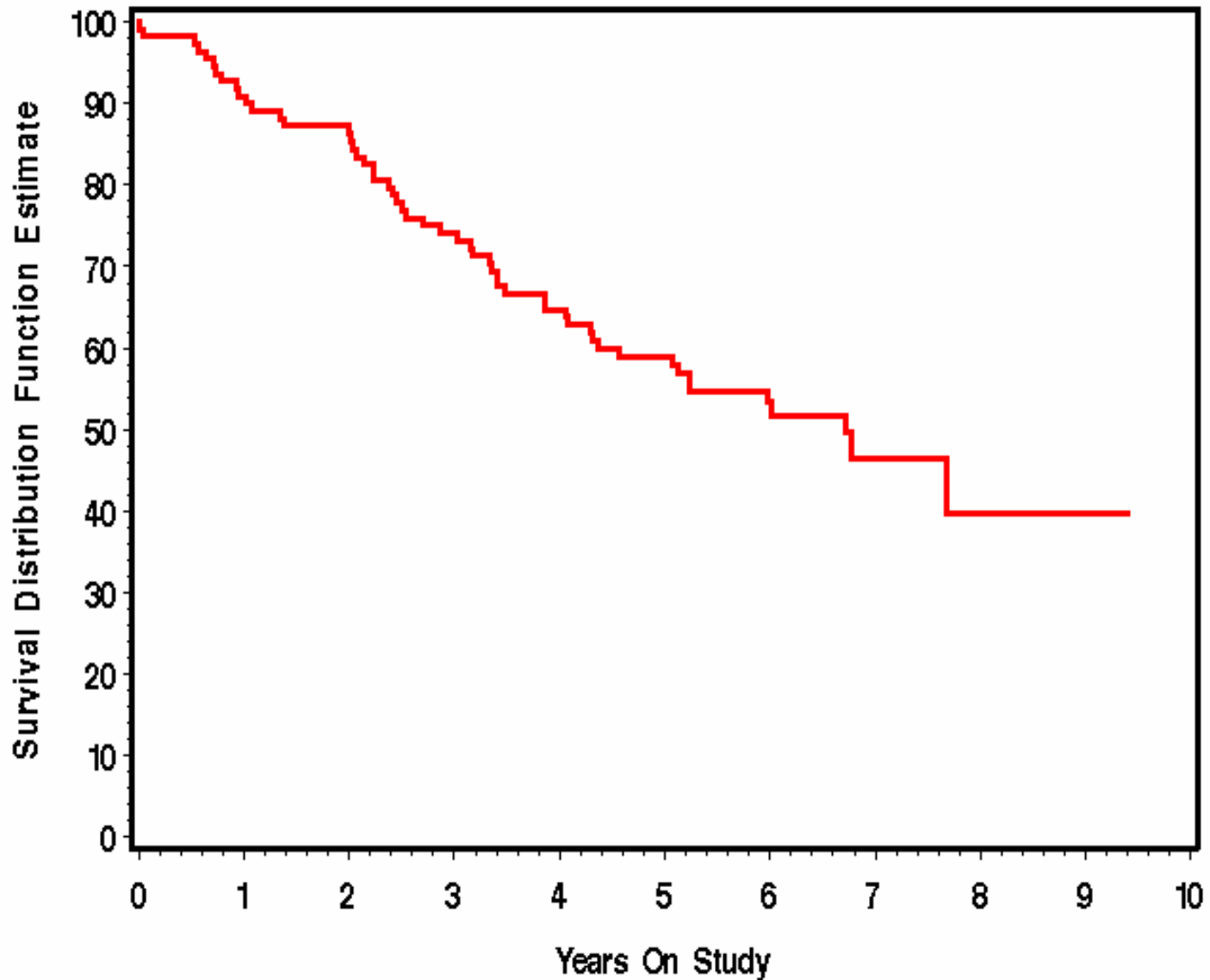
Survival Curve

Missing Data shown as Censored Values

Cancer Statistics 2003

Primary Site	New Cases (no.)	Deaths (no.)	5-year survival	
			1974-76	1992-98
Lung	171,900	157,200	12%	15%
Colorectal	147,500	57,100	50%	62%
Breast	212,600	40,200	75%	86%
Pancreas	30,700	30,000	3%	4%
Prostate	220,900	28,900	67%	97%

Survival Comparison Between Once Daily
and Twice Daily Thoracic Radiation

Once Daily TI (n=132)
Twice Daily TI (n=130)

Log Rank p−value=0.49

% Survival

Time in Months
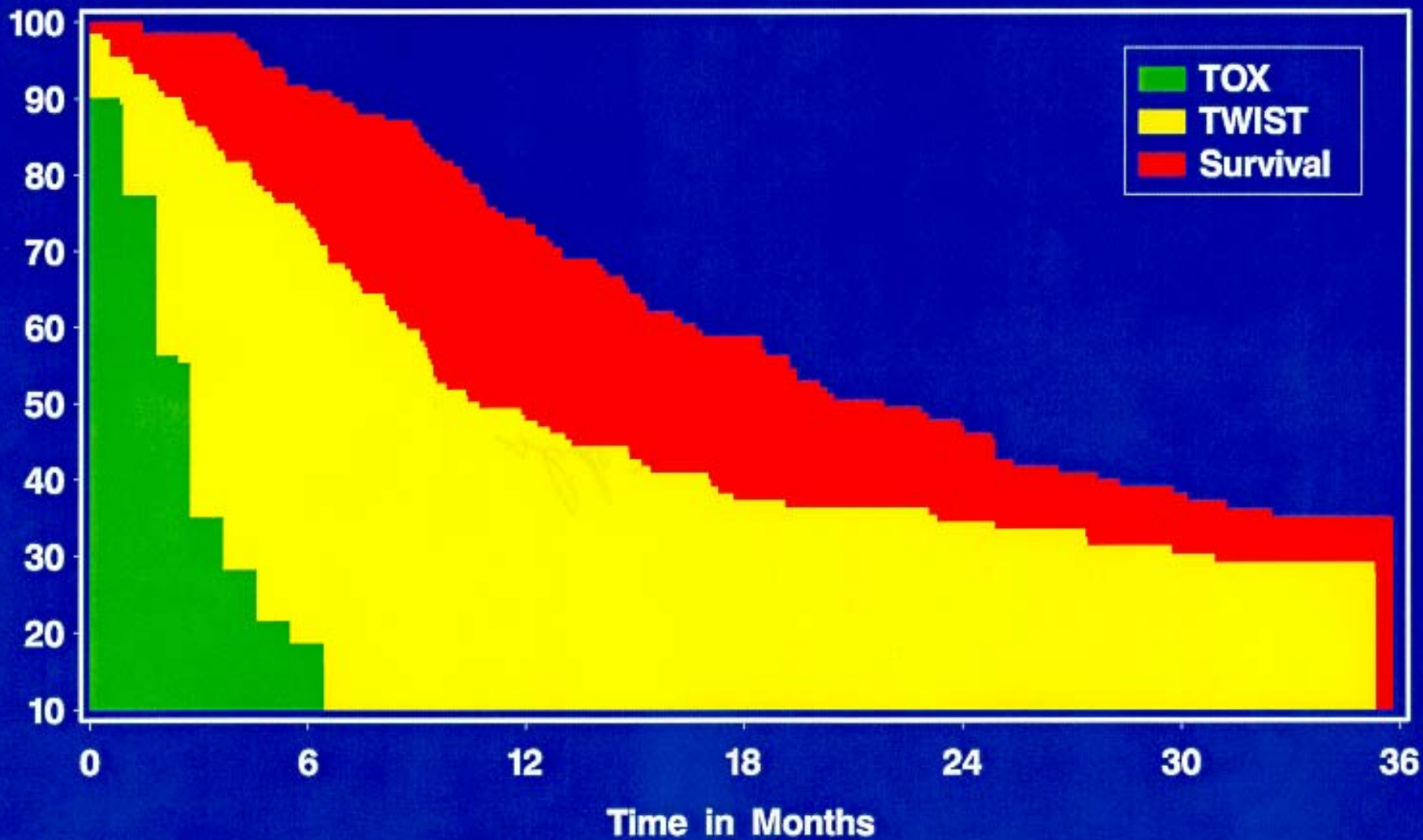
Survival Analysis History: Lessons Learned

- **The math was complex**

- **The application was accessible**

- **The software was made widely available and usable (PHREG)**

- **"I don't need a statistician to interpret these survival results"**

Historical tale #2:
Arguing with yourself can be self-defeating….

The Application of Bayesian Methods in Health Care Research

- 1763 Rev. Thomas Bayes proposed using conditional probability to express the likelihood of events

- Simple concept, complex math

- 1990 Lindley's approximation

- 1991 Gelfand & Smith Gibbs Sampling

- 1990's MCMC

Bayes Timeline Continued

- **Lindley 1980? "Classical statistics will go the way of the dinosaur by the year 2000"**

- **1990's CRM proposed for phase I trials**

- **2003 Berry, Few clinical trials use Bayes methods**

MAYO CLINIC

Bayesian Analysis History: Lessons Learned

- **The math was complex**

- **Protagonists spent as much time arguing in the literature as building**

- **The application was NOT accessible**

- **The software was NOT made widely available and usable**

History tale #3:
Sometimes the solution is close at hand

The Application of Factor Analysis in Health Care Research

- 1904 Spearman invented / applied to IQ testing

- 1955 Rao provided estimation and significance tests

- 1968 "Factor analysis will revolutionize the way in which psychological assessments are developed…" (Anon.)

- 1970, 1974 Kaiser's Little Jiffy (I-IV)

Factor Analysis Timeline Continued

- **1990's Many QOL assessments still developed without factor analysis**

- **Little Jiffy is the method of choice**

- **2004 Factor analysis proposed as a screening method for IRT**

- **June 24-25, 2004 Factor analysis and IRT compliment each other**

Factor Analysis History: Lessons Learned

- **The math was complex, options myriad**

- **The application was accessible, simple?**

- **The software was made widely available and usable**

- **Factor analysis found a niche as <u>A</u> useful tool development method**

How do we move ahead hand in hand....

Goals for item banking

- Acknowledge that an item bank will not solve all problems, and will create some

- Acknowledge vested interests

- Make the application of the methods accessible to non-experts, especially clinicians

- Make the tools accessible to researchers (PROC IRT)

- Give people a reason to participate (like this meeting) rather than mandate

How do we make it accessible?

- "A score of 75 on the near vision scale resulted in a 99% probability of having no difficulty shaving but only a 30% probability of having no difficulty in reading small print"

- "This methodology yielded an instrument that was 25% shorter yet retained the same information"

How do we make it INaccessible?

- "Samples of 500 - 1,000 are probably sufficient"

- Simple power analysis guidelines are needed

- Calibration of clinical meaning

How do we make it INaccessible? The Iowa Test of Basic Skills

- **Two weeks of class time to assess if Johnny can read, write, and do math**

- **Students not allowed to see their own tests, affidavit required for parents**

- **Teachers do not understand the results, use their own systems**

- **Stanines**

MAYO CLINIC

Expect people to have differing opinions

How do we make it INaccessible?

- "IRT models are sample independent"

- "It is not strictly correct to say that IRT models provide parameter estimates that are sample independent"

- "and your point is…."

Don't Call My Baby Ugly

- "The 2PLM model is vastly superior to the 1PLM model"

- Box "All models are incorrect, but some are useful"

Voices of Consensus and Reason

- "IRT modeling should not replace CTT, both methods can be used to inform each other"

- "Involving content experts and eliciting their feedback is extremely important"

- "Consumers will come along if the producers agree on the basic quality of the products"

Promising Evidence from this Meeting

- **25 posters:**
  - **6 discussed IRT methodology**
  - **6 showed how IRT and CCT could be integrated**
  - **9 indicated IRT promising**
  - **6 demonstrated CAT reduces items**

  - **3 had confrontational titles**
  - **1 IRT better than CCT**

Summary

- **An item bank CAN be built**

- **Gaps in knowledge need to be acknowledged and addressed**

- **Positive motivation works better in the long run**

- **Incorporate alternative approaches, celebrate the differences**

The quality of the output is a function of the input

MAYO CLINIC

Thank you

References: jsloan@mayo.edu