Trauma Mortality Prediction Model: TMPM

Turner Osler MD M.Sc. FACS, Research Professor of Surgery, U of VT
Laurent G Glance MD, Associate Professor of Anesthesia, U of Rochester
Jeff Buzas PhD, Professor, Department of Math and Statistics, U of VT
Andrew Dick, PhD, RAND-University of Pittsburg Health Institute, Pittsburg, Pennsylvania
Dana Mukamel, PhD, School of Medicine, University of California, Irvine California Jacob Wagner, MD, PhD, Department of Surgery, University of Vermont, Burlington, Vermont

With thanks to David Hosmer and Bill Jefferys

October 7, 2007

- Quantification of trauma is required for:
  - The scientific study of injury
  - The objective evaluation of trauma care
- We compromise our research and our patient care if we fail to develop and *use* accurate measures of trauma
- The AIS/ISS methodology has served as the standard measure of trauma for 30 years, but…

Serious Concerns with AIS/ISS

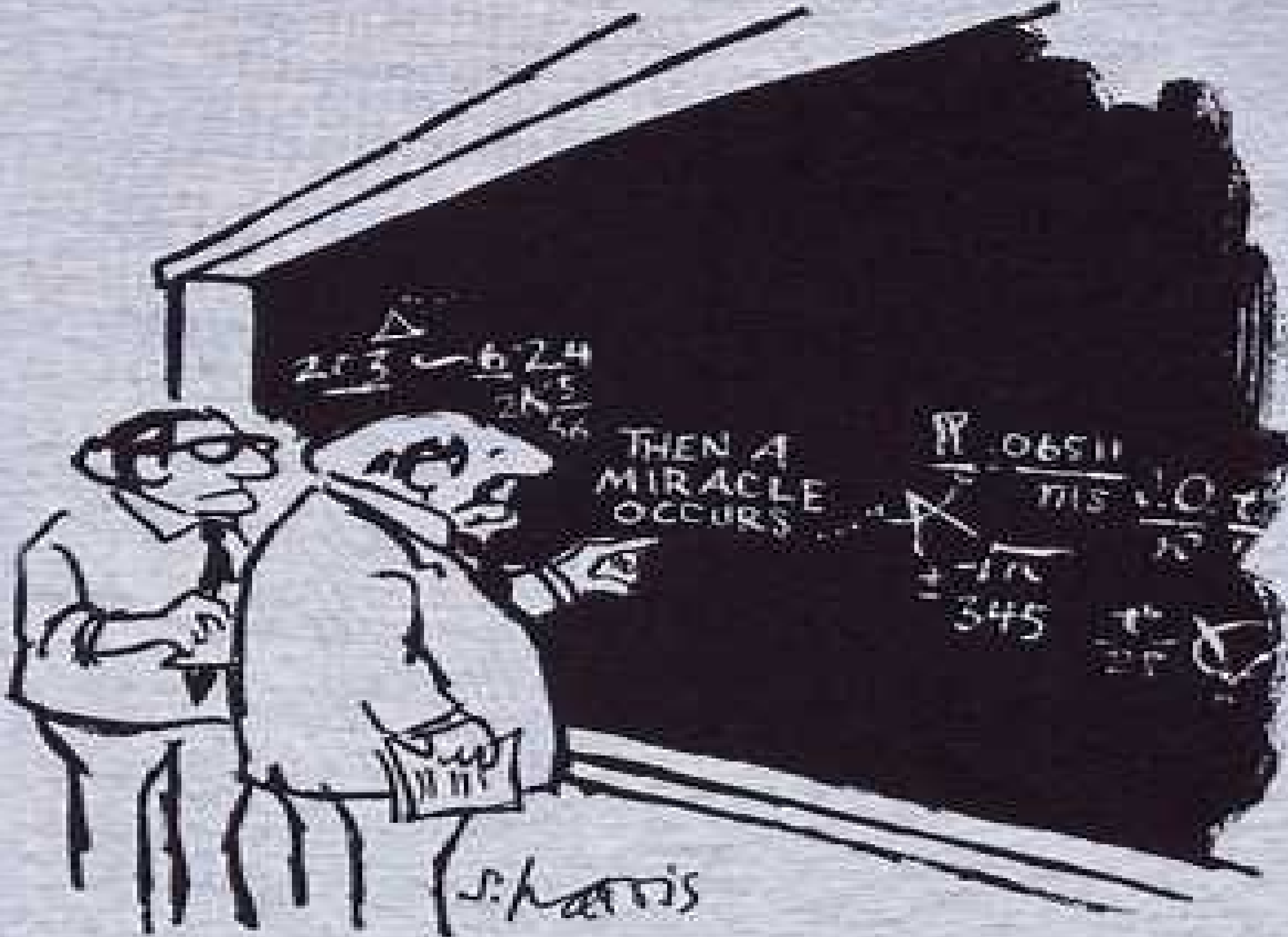
- Structural concerns:
  - **AIS:**
    - Only 6 possible severities
    - Estimated not measured
  - **ISS:**
    - “Square and sum” model (why?)
    - Uses *at most* three injuries
- Performance concerns in large data sets
  - Higher scores sometimes have lower mortality
  - Single worst injury predicts outcome more accurately

Can we do better than ISS? Starting with a clean sheet of paper…

- The AIS lexicon remains an appealing lexicon
  - Divides the continuous landscape of trauma into 1300 plausible injuries
  - Assigns approximate severities
  - Large data sets are available
- Suppose our new model
  - Used ***measured*** severities of individual AIS injuries
  - Used ***a more stringent model*** than “square and sum”?

Measuring severity of individual injuries

- Not a new idea
- Surprisingly difficult, because the data doesn't play fair in two ways:
  - Many injuries are rare, making accurate mortality estimates difficult to obtain
    - The single patient shot through the #@^&* died.
    - What is the mortality for GSW to #@^&*?
      - Probably not 100%, but what is a better estimate?
  - The injury of interest is usually contaminated by concurrent injuries, making it difficult to determine which injury is actually responsible for death
    - Now say our patient shot through #@^&* is also shot through the hypothalamus.
      - Should this change our estimate of the mortality for a GSW to the #@^&*?
  - These issues, although long recognized, have been marginalized, probably because they are hard, intrinsically mathematical, problems
  - I can personally claim to be in the forefront of ignoring this problem



"I THINK YOU SHOULD BE MORE EXPLICIT HERE IN STEP TWO."

Measuring individual injury severity

- Historically
  - Ignore contaminating injuries & low volume (Osler 1996)
    - Result: uniform overestimation of individual injury severities
  - Use only patients with a single injury (Meredith 2003)
    - Clever, fortuitous exploitation of "Experiments of nature"
    - But:
      - Ignores 80% of available data (exacerbates rare injuries problem)
      - Data is no longer representative of the entire data set
- How might we "unscramble" multiple injuries in individual patients?

A model to “unscramble” injuries

- Single, large scale, regression model:

$$\text{Survival } (0/1) = f\,(\text{Inj}_1 * \text{Inj}_2 * \ldots * \text{Inj}_{1386})$$

Outcome	110009.1	113000.6	116000.3	....	....	(1322 of these)	....		919404.5
0	0	0	1	0	0	...	1	...	0
0	1	0	0	0	0	...	...	...	1
0	0	1	0	0	1	...	...	...	0
1	1	1	0	0	0	...	...	...	1
...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...
0	0	0	0	1	1	...	...	...	0

A model to “unscramble” injuries

- A big problem requires substantial resources
  - Very large data sets: NTDB 6.1
    - 2,177,127 injuries
    - 702,299 patients
    - 1,322 AIS codes
  - Large computer
    - Multiprocessor workstation with 4G RAM
  - Patience
    - 100 hours
  - Money
    - 100 hrs X 300 watts X $0.14/KWH = $4.20

A model to "unscramble" injuries

- This model should yield 1,322 coefficients, one for each injury in the AIS lexicon
- But coefficients can be estimated for only 70% of AIS codes:
  - Perfect separation / Collinearity problems
- Average this model (model 1) with a second logistic model at the region/severity level
  - "Rely on data when available, but expert opinion when necessary"

How to Average Models
(MARC: Model Averaged Regression Coefficients)

$$\frac{\frac{1}{\mathrm{Var(Model\ 1)}}}{\frac{1}{\mathrm{Var(Model\ 1)}}+\frac{1}{\mathrm{Var(Model\ 2)}}}\mathrm{Coef}_{\mathrm{Model\ 1}}+\frac{\frac{1}{\mathrm{Var(Model\ 2)}}}{\frac{1}{\mathrm{Var(Model\ 1)}}+\frac{1}{\mathrm{Var(Model\ 2)}}}\mathrm{Coef}_{\mathrm{Model\ 2}}$$

$$Var(RScoef_m) = \frac{1}{N_i - 1}\sum_{j=1}^{Ni} N_i * W_j * [(AIScoef_j - E(AIScoef))]^2$$

Where:

N_i is the number of AIS codes in RS region m

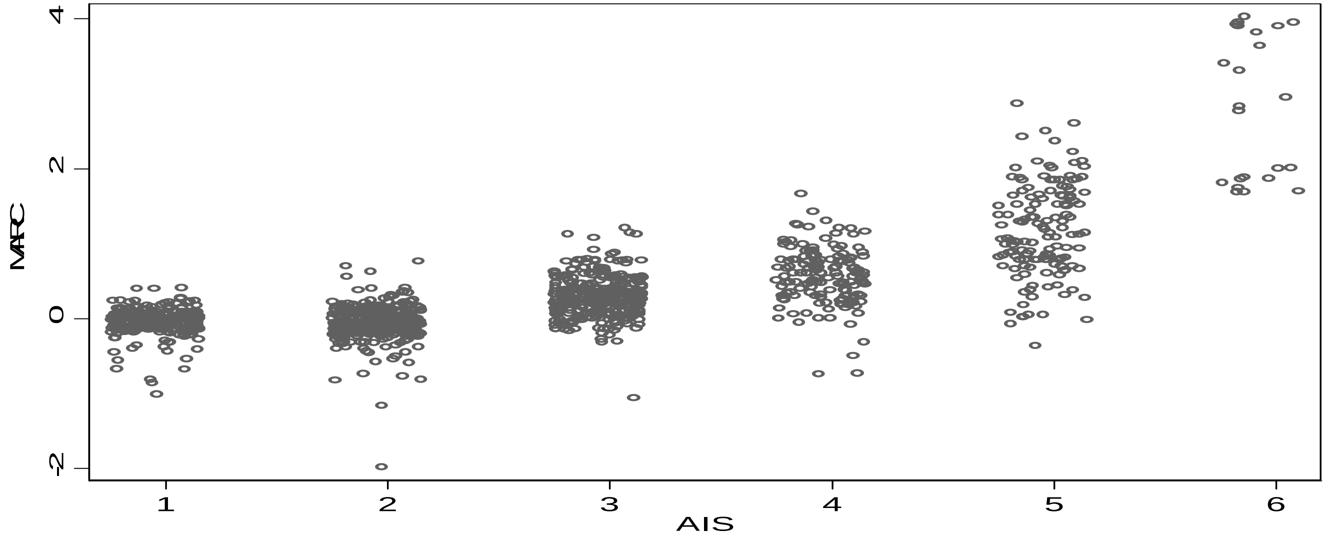
W_j is the weight of the contribution of $AIScoef_j$ to the variance of $RScoef_m$

$$W_j = \frac{\frac{1}{Var\_AIScoef_j}}{\sum_{r=1}^{Ni}\frac{1}{Var\_AIScoef_r}}$$

$$\mathrm{E(AIScoef)} = \sum_{j=1}^{Ni} W_j * AIScoef_j$$

MARCs (all 1,322 of them)

- Face validity
  - "High end" (2 <MARC< 4)
    - Brainstem penetration, hepatic avulsion, aortic transection
  - "Low end" (-1 <MARC< 0)
    - Lumbar sprain, finger nfs, awake on admission, ear injury
- Different from AIS severities:



“Are we there yet?”

- **Not yet**: We still need a prediction model:
  - A list of empirical injury severities is a good start
- What to include?
  - Just Injuries (allows ‘head to head’ comparison with ISS)
    - (All? Even trivial ones? What’s trivial?)
  - Interactions?
  - Other variables?
    - Injury related: Mechanism, safety equipment, …
    - Patient related: Age, gender, BMI, …, eye color, political party…
    - Physiology related: Blood pressure, GCS, GCS motor component…
  - What kind of model? (logit, probit, cloglog, …)

The black art of model building…

Plausible predictors, intelligently incorporated

Survival Models

- **ISS**: Probit model with 1 predictor
  - Sum of squared worst AIS severities for 3 worst injured body regions
- **SWI**: Probit model with 1 predictor
  - MARC value of worst injury
- **TMPM**: Probit model with 7 predictors
  - Worst injury MARC
  - Second worst injury MARC
  - Third worst injury MARC
  - Fourth worst injury MARC
  - Fifth worst injury MARC
  - Indicator: worst and second worst in same body region
  - Interaction of worst and second worst MARC values

Trauma Mortality Prediction Model: TMPM:

Predictor	Coefficient	Robust Std. Error	Z	P>|z|	95% CI
MARC1	1.3138	0.0210	62.43	0.000	(1.2725, 1.3550)
MARC2	1.5136	0.0498	30.42	0.000	(1.4161, 1.6112)
MARC3	0.4435	0.0431	10.29	0.000	(0.3591, 0.5280)
MARC4	0.4240	0.0456	9.29	0.000	(0.3346, 0.5134)
MARC5	0.6284	0.0707	8.89	0.000	(0.4898, 0.7669)
Same Region	-0.1377	0.0114	-12.07	0.000	(-0.1600, -0.1153)
MARC1*MARC2	-0.6506	0.0270	-24.08	0.000	(-0.7036, -0.5976)
Constant	-2.3281	0.0222	-104.70	0.000	(-2.3717, -2.2845)

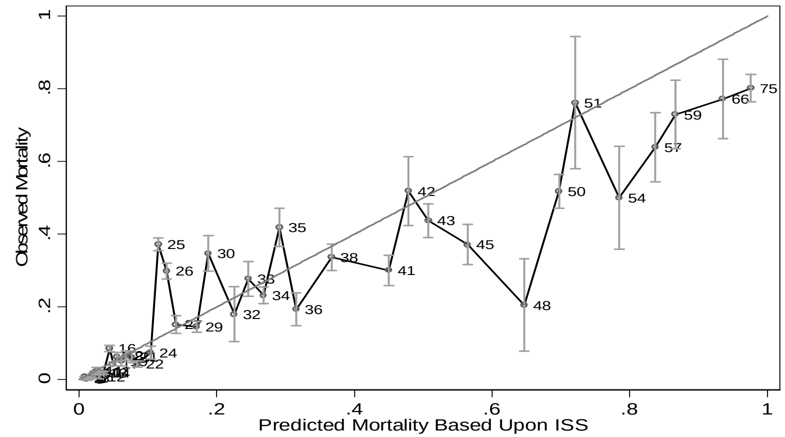
TMPM vs. Other AIS Models

Model	ROC**	HL Stat*	AIC*
ISS	0.872	296	37225
SWI	0.891	314	34059
TMPM	0.902	58	32003

Model	ROC**	HL Stat*	AIC*
ISS age+gender+mech	0.915	54	33773
SWI age+gender+mech	0.921	128	31770
TMPM age+gender+ mech	0.928	19	29645

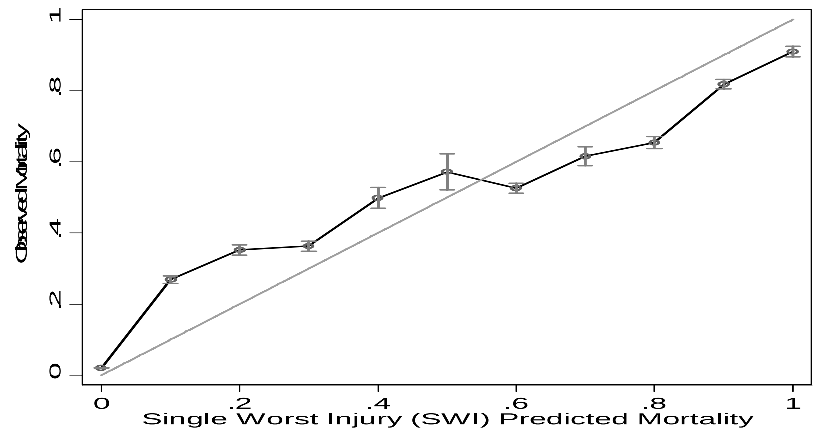
** Higher value = better * Lower value = better CI based upon resampling

ISS



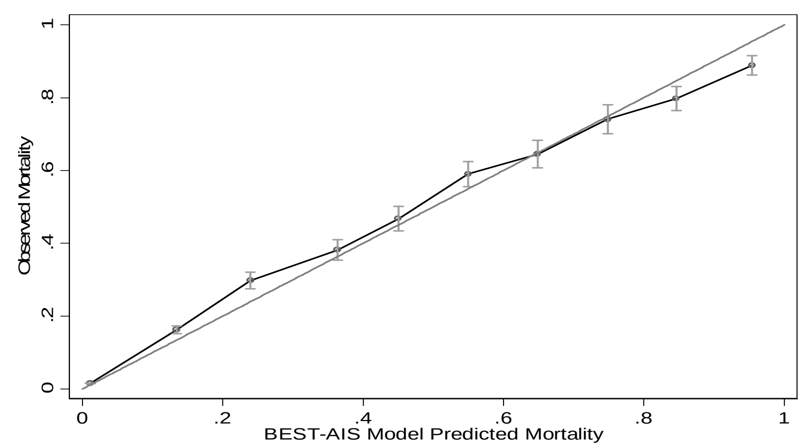
Observed Mortality
0
.2
.4
.6
.8
1
Predicted Mortality Based Upon ISS

SWI



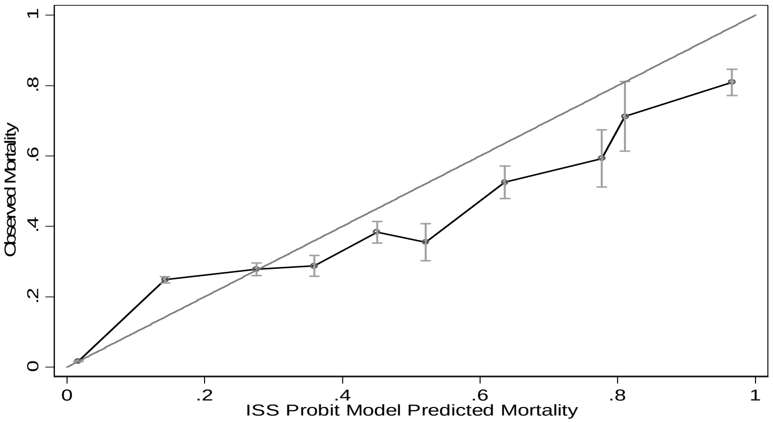
Observed Mortality
0
.2
.4
.6
.8
1
Single Worst Injury (SWI) Predicted Mortality

TMPM

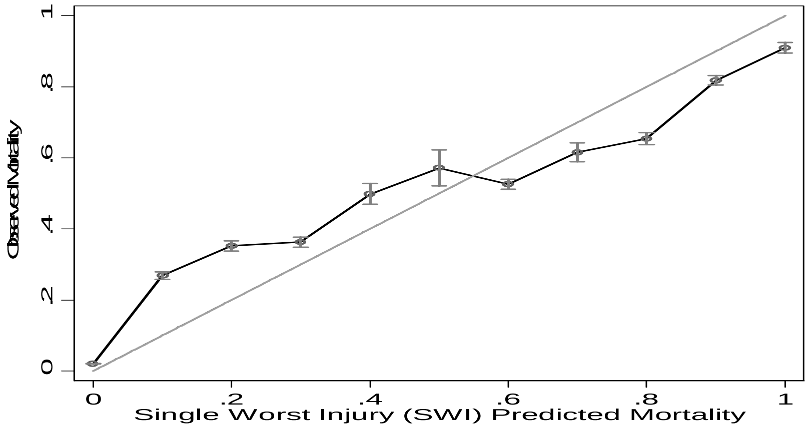


Observed Mortality
0
.2
.4
.6
.8
1
BEST-AIS Model Predicted Mortality

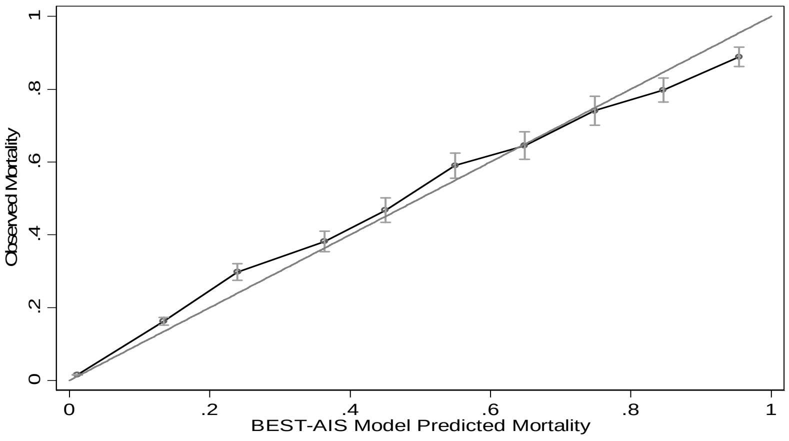
ISS


Observed Mortality
ISS Probit Model Predicted Mortality
0
.2
.4
.6
.8
1

SWI


Observed Mortality
Single Worst Injury (SWI) Predicted Mortality
0
.2
.4
.6
.8
1

TMPM


Observed Mortality
BEST-AIS Model Predicted Mortality
0
.2
.4
.6
.8
1

Practically speaking:
How does one actually Calculate TMPM?

- It's a regression model:
  - Pencil
  - Paper
  - 1,322 MARC values (1 for each AIS code)
  - (normal distribution table)
- Or, do it the easy way:
  - Hand a table of Patient ID and AIS values to an Excel spreadsheet, and push the "calculate" button:
    - (65,536 record limit: Thanks, Bill.)

Parting Thought

- What I have described is a way of getting from the AIS lexicon to survival prediction
- One could do this with any lexicon, say ICD9
  - There are technical problems (priors?)
- But:
  - It works pretty well (better than ICISS)
  - (But not as well as the best AIS based models)

ICD9 BASED MODELS

Model	ROC	HL Stat	AIC
ICISS	0.846	432	44,071
SWI: SRR	0.861	833	42,813
SWI: MARC	0.872	112	42,716
All Injury: MARC	0.878	357	42,688
TMPM-ICD9	0.880	19	41,251

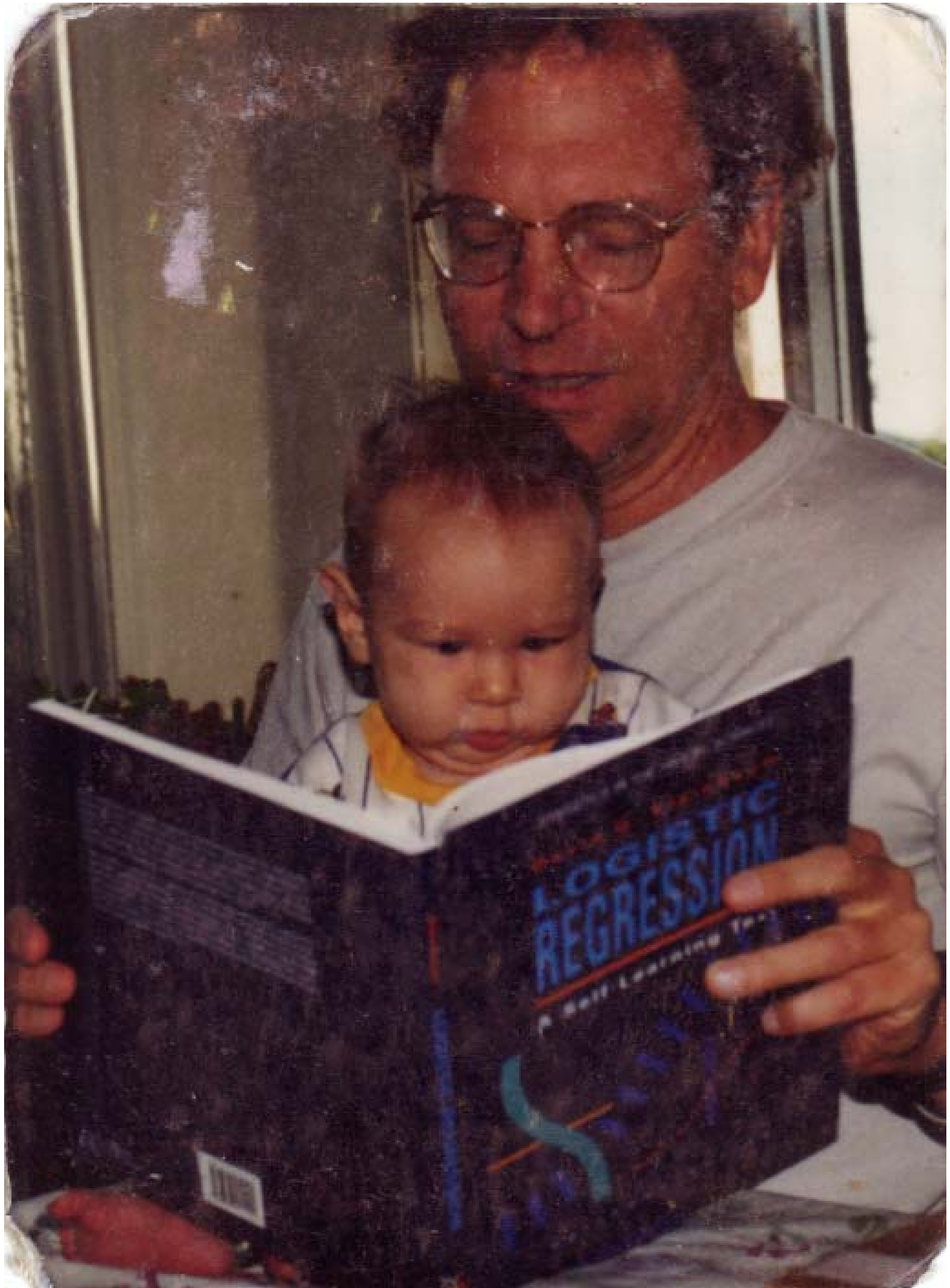
Summary

- Currently available AIS severities are approximate
- Large scale logistic regression -> Better estimates
  - MARC:
    - Face validity
    - Differ substantially from previous AIS severity
- Prediction models based upon these more precise severities are far superior to the venerable ISS.
- TMPM substantially outperforms ISS and SWI
  - ROC, HL, AIC, Calibration curves
- One can do this for ICD9 data with similar results

Conclusion

- We can use empirically derived severities to get better mortality predictions
  - AIS or ICD9 lexicon (but AIS>ICD9)
  - For either lexicon, MARC based models provide better discrimination
- These models can include interaction terms
  - Improves model calibration
- It is time to retire the venerable, but outdated ISS

Logistic Regression



It's

Child's Play

ICD9 BASED MODELS

Model	ROC	HL
ICISS: SRR age gender mechanism	0.891	72
SWI: SRR age gender mechanism	0.893	93
SWI: MARC age gender mechanism	0.903	67
TMPM-ICD9: MARC age gender mechanism	0.911	11

Model Description	C statistic	Hosmer-Lemeshow Statistic
AIS BASED MODELS		
Injury Severity Score + age +gender +mechanism	0.904	54
TMPM-AIS + age + gender + mechanism	0.925	19
ICD9 BASED MODELS		
ICISS + age +gender +mechanism	0.891	72
SWI SRR + age +gender +mechanism	0.893	93
TMPM-ICD9 + age +gender +mechanism	0.911	10