Making the Most of your (and Others’) Data

Sherman C. Stein, MD
Clinical Professor of Neurosurgery
Every Student’s Goal

- Be involved in high quality clinical research (RCT)
- Published in high impact journal
- Get authorship credit
- Use as basis for brilliant career
Part-time Student Research

- Time limited by other demands
- Clinical years too busy; summers too short
- Rotation demands interfere with subject availability
- Share work with many others
  - diminished satisfaction, expertise
  - graduate too soon
  - authorship credit
Additional Constraints in Neuroscience

- Low volume diseases – not common cold, breast cancer, etc.
  - accrual, trial length
  - compliance
  - can’t fund

- Finish study, but too small to report
Few RCT’s in Neurosurgery

Only 14 high-quality RCT’s in all of Neurosurgery


Recent Medline search – 28 additional studies, only $\frac{1}{2}$ high quality (and 8/13 negative!)
Multicenter collaborative trials

- Increase sample size
- Reduce time to complete
- But introduces heterogeneity
- Reduce your contribution

Multicenter Trial of Early Hypothermia in Severe Brain Injury

Guy L. Clifton, Pamala Drever, Alex Valadka, David Zygun, and David Okonkwo
"Just how many ways are there to skin a cat?"
Alternative approaches to address clinical questions when your own data limited or lacking altogether

- Meta-analysis
- Decision analysis
- Comparative effectiveness
- Cost-Effectiveness
- Threshold analysis

Incorporating your own Data set, or purely from Literature review
Meta-Analysis

• Statistical means to combine multiple RCTs

• Pool results to estimate effect size (mean odds or risk ratio, 95% CI)

• Useful to:
  – Clarify a group of non-dispositive studies
  – Increase Power to Achieve Statistical Significance
  – Help Resolve Conflicting Results
Example: Estrogen for Menopause

Data plentiful but contradictory

- Traditional – estrogen decreases risks of heart attack, stroke, ovarian & breast cancer (& hot flashes)
- 2002 Women’s Health Initiative – increased risks of breast cancer, heart attacks and strokes, PE’s
- 2004 Million Women Study – stroke risk reduced
- 2004 WHI follow-up - breast cancer risk reduced
- 2006 WHI follow-up - breast cancer risk not significantly reduced
I USED TO BE INDECISIVE... NOW I'M NOT SURE.
Meta-analysis Example – 6 studies comparing Treatments A & B

• Calculate % success with each treatment
• Calculate Odds ratio of A vs. B
• Odds ratio (measure of effect size) =
  \[
  \left( \frac{\text{# successes A}}{\text{# failures A}} \right) / \left( \frac{\text{# successes B}}{\text{# failures B}} \right)
  \]
• If ratio > 1, A is better treatment
• If 95% confidence interval all above 1, the difference is significant
Weighting
are some data more equal than others?

• Is study with 20 observations = to one with 2000?

• **Unweighted** – each study counted same

• **N–weighted** – each observation counted same

• **Variance–weighted** – study’s Weight $1/\alpha$ variance
### Odds Ratios – Treatment A/B

<table>
<thead>
<tr>
<th>Series</th>
<th>N</th>
<th>Odds Ratio</th>
<th>Var (*1000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>50</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>B</td>
<td>20</td>
<td>0.15</td>
<td>2</td>
</tr>
<tr>
<td>C</td>
<td>25</td>
<td>0.1</td>
<td>1.6</td>
</tr>
<tr>
<td>D</td>
<td>50</td>
<td>0.2</td>
<td>0.8</td>
</tr>
<tr>
<td>E</td>
<td>500</td>
<td>0.6</td>
<td>0.00005</td>
</tr>
<tr>
<td>F</td>
<td>1000</td>
<td>0.9</td>
<td>0.0025</td>
</tr>
</tbody>
</table>

Here N = # of cases in each study, Var. = Variance
<table>
<thead>
<tr>
<th>Analysis</th>
<th>Odds Ratio - Pooled Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unweighted</td>
<td>0.49 (signif)</td>
</tr>
<tr>
<td>N-weighted</td>
<td>0.77 (not signif)</td>
</tr>
<tr>
<td>Inverse variance-weighted</td>
<td>0.57 (signif)</td>
</tr>
<tr>
<td>Study</td>
<td>ES (95% CI)</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
</tr>
<tr>
<td>A</td>
<td>1.00 (0.61, 1.39)</td>
</tr>
<tr>
<td>B</td>
<td>0.15 (-0.24, 0.54)</td>
</tr>
<tr>
<td>C</td>
<td>0.10 (-0.29, 0.49)</td>
</tr>
<tr>
<td>D</td>
<td>0.20 (-0.19, 0.59)</td>
</tr>
<tr>
<td>E</td>
<td>0.60 (0.59, 0.61)</td>
</tr>
<tr>
<td>F</td>
<td>0.90 (0.80, 1.00)</td>
</tr>
<tr>
<td>Pooled</td>
<td>1.00 (0.61, 1.39)</td>
</tr>
</tbody>
</table>
Meta-analytic Pooling – Observational Data

- Meta-analytic Techniques to Weight each study (inverse-variance) & pool means
- Example: Interested in what Literature says about quality of life (QOL) after minor stroke
- Several studies, varying scores
- Works same as standard meta-analysis, except gives pooled mean QOL instead of odds ratio
- Can do same with probabilities
<table>
<thead>
<tr>
<th>Study ID</th>
<th>QOL in Mild Stroke</th>
<th>ES (95% CI)</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>0.85 (0.77, 0.93)</td>
<td>5.61</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.85 (0.77, 0.93)</td>
<td>5.61</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.80 (0.80, 0.80)</td>
<td>8.13</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.90 (0.88, 0.92)</td>
<td>8.00</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.90 (0.88, 0.92)</td>
<td>8.00</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>0.90 (0.88, 0.92)</td>
<td>8.00</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>0.90 (0.88, 0.92)</td>
<td>8.00</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>0.81 (0.67, 0.95)</td>
<td>3.65</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>0.89 (0.86, 0.92)</td>
<td>7.75</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.83 (0.81, 0.85)</td>
<td>7.97</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>0.70 (0.65, 0.75)</td>
<td>6.83</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>0.91 (0.88, 0.94)</td>
<td>7.70</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>0.88 (0.84, 0.92)</td>
<td>7.40</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>0.78 (0.74, 0.82)</td>
<td>7.40</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td>0.85 (0.82, 0.89)</td>
<td>100.00</td>
</tr>
</tbody>
</table>

NOTE: Weights are from random effects analysis
Weaknesses Of Meta-analysis

• Only resolves 1 Endpoint at a Time
• Can only compare 2 Treatments at a time
• If Endpoints compete, cannot definitively choose better Treatment
# Competing Endpoints

<table>
<thead>
<tr>
<th>Complication Type</th>
<th>Higher Complication Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment A</td>
</tr>
<tr>
<td>complication 1</td>
<td>X</td>
</tr>
<tr>
<td>complication 2</td>
<td>X</td>
</tr>
<tr>
<td>complication 3</td>
<td>X</td>
</tr>
<tr>
<td>complication 4</td>
<td></td>
</tr>
<tr>
<td>complication 5</td>
<td>X</td>
</tr>
<tr>
<td>complication 6</td>
<td>X</td>
</tr>
</tbody>
</table>

Which Treatment is Better, A or B?
### Impact on Quality of Life Important

<table>
<thead>
<tr>
<th>Complication Type</th>
<th>Higher Complication Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment A</td>
</tr>
<tr>
<td>halitosis</td>
<td>X</td>
</tr>
<tr>
<td>bad hair</td>
<td>X</td>
</tr>
<tr>
<td>sniffles</td>
<td>X</td>
</tr>
<tr>
<td>Death</td>
<td></td>
</tr>
<tr>
<td>rare headache</td>
<td>X</td>
</tr>
<tr>
<td>occasional flatulence</td>
<td></td>
</tr>
</tbody>
</table>

**Now which Treatment is Better?**
## Relative Frequency Important

<table>
<thead>
<tr>
<th>Complication Type</th>
<th>Higher Complication Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment A</td>
</tr>
<tr>
<td>halitosis</td>
<td>.5</td>
</tr>
<tr>
<td>bad hair</td>
<td>.5</td>
</tr>
<tr>
<td>sniffles</td>
<td>.5</td>
</tr>
<tr>
<td>Death</td>
<td>.00010</td>
</tr>
<tr>
<td>rare headache</td>
<td>.5</td>
</tr>
<tr>
<td>flatulence</td>
<td>.5</td>
</tr>
</tbody>
</table>

Now which Treatment is Better?
Conclusion: Both frequencies of All outcomes and their impacts on a patient’s quality of life are important.

Meta-analysis can only deal with one impact at a time.
Decision Analysis

“Virtual” Clinical Trial

1. Alternative when cannot do (or haven’t done) real trial
2. Predicts outcome of trial
3. Incorporates all outcomes at once — reports single global outcome
4. Allows comparisons of >2 treatments
Decision Analysis

Mathematical analysis of existing data to anticipate results of clinical trial and to aid in decision-making
Elements of Decision Analysis

- 2 or more alternate choices
- list of all likely outcomes
- probability of each outcome
- impact of each outcome on quality of life (QOL)
- individual elements usually already in literature
Pascal’s Wager

How to justify belief in God in the Age of Enlightenment?

• Costs –
  – Church attendance
  – Good deeds
  – Charitable contributions
  – Live scandal-free
  – Etc.

• Reward or Payoff – Heaven

Blaise Pascal
Wager as Decision Tree

Believers
Potential Gains: Infinite
Potential Costs: Finite
Pascal: belief = safe bet
Game Theory (It’s in the game!)
Win or Lose $1 per Flip
Don’t need formal Analysis
With 2 of 3, win 5 of every 8 times you play
Who picks that gamble?
Model Decision Tree

- **Decision Node**: 2 of 3?
- **Chance Node**: 1 flip
  - Correct: 0.5
  - Wrong: 0.5
- **Terminal Node**: Payoff (in this case $)
  - Correct: 3
  - Wrong: -3

- **Probability**: 0.5

---

Penn Neurosurgery
Rollback Analysis

• Multiply probability of each branch by its payoff
• For each decision, add the products
• Sum = expected payoff associated with that decision (how it will turn out on average)
• The best decision is the one with the greatest expected payoff
With 2 of 3, win 5 of every 8 times you play
But: 4/5 wins only $1, every loss $3
Roll-back analysis shows average game loses $0.25
Decision Analysis

• Popular in business in 50’s

• Developed for gov. strategy by Rand Corp. in 60’s
Medical Decision Analysis

• This approach readily adapted for medical decision-making (Lusted 1971)

• Odds become Transition Probabilities and Payoffs become Health Utilities (measures of quality of life)

• These numbers obtainable from Med. literature
Rules for Medical Decision Models

- **Decision Node**: ≥ 2 branches (treatments, strategies)
- **Transition Probabilities**:  
  - Exhaustive – probabilities add up to 1  
  - Mutually exclusive – no overlap
- **Terminal Node**: process final (no later developments)
- **Payoff**: utility-based QOL  
  - Preference-based – relative importance to patient  
  - Parametric & between 0 (dead) & 1 (perfect health)
Chance in Medical Decisions

- Dr’s make decisions (or advise patients to decide) every day
- But based on implicit and unstructured reasoning
- Evidence-based choices uncommon
- Actual calculations, comparing likelihood or importance of possible outcomes, even rarer
Medical Decisions Involve Risk. Why rely on gut feelings?
Advantages of Decision Analysis

- Forces systematic examination of problem
- Forces assignment of explicit values
- Controls complexity and thus minimizes processing errors
- Allows combining data from multiple sources
- Allows approach to problems for which no one study addresses all aspects
Disadvantages of Decision Analysis

• Time consuming (a lot for an individual patient)
• Model simplified, calculations complex
• Uses point estimates for values, but may be great uncertainty—conclusions suspect
• Results often difficult to explain
• Methods not well understood or trusted by uninitiated
If Successful, Hard Work Pays Off
• Model simplified
• Probabilities Arbitrary
• Utilities: cure = 1; death = 0 (by convention)
• Don’t need formal analysis – superiority of Surgery obvious
Successively better Models

+ surgical complications
Utility of morbidity = 0.5

+ delayed surgery if antibiotic fails
but lower odds of success
Expected Utility

• Rollback analysis – multiply probability of each outcome of a particular treatment by its utility
• Add the utilities for each possible outcome of each treatment choice
• This is expected utility of that treatment – the average outcome a patient with that treatment can expect
• The treatment resulting in the highest expected utility is preferred
Abscess – Expected Utility

- **Brain abscess**: 0.80
  - Antibiotic: 0.80
    - Surgery: 0.75
      - Cure: 0.700
        - Nonfatal: 0.50
        - Fatality: 0.00
      - Non-cure: 0.200
        - 1.00; $P = 0.500$
    - No cure, late surgery: 0.200
      - Nonfatal: 0.60
        - Non-cure: 0.40
          - Fatality: 0.300
          - 0.50; $P = 0.100$
          - 0.50; $P = 0.150$

Penn Neurosurgery
Where do We Get Data for Transition Probabilities?

• Our Own Prospective Trial
• Retrospective Review of Our Own Cases
• Literature: 1 or More Case series
• (Expert Opinion, Educated Guess last Resort)
Utility vs. Quality of Life (QOL)

• Many QOL instruments (many specific for a disease; scores based on disease severity)
• However, not all based on patient’s preference
• Not all parametric (halve the utility, double the severity)
• Not anchored at 0 & 1, or can’t be converted
• Utility (definition) fulfills all requirements
Obtaining Utilities

- Own Cases
- Literature Review
Obtaining Own Utilities

e.g. Stroke

Plusses & minuses with each

• Healthcare Community
• Patients with Illness in Question
• Patients at Risk of Illness (e.g. heart disease for stroke)
• General Community
• Proxy (e.g. Caregiver)
Utility Instruments

- Preference-Based
  - Standard Gamble
  - Time Tradeoff
  - Visual Analog Scale (VAS)
- Proprietary Questionnaire
  - EuroQual-5D
  - Health Utilities Index
  - SF-6D
Example of Decision Analysis

Decision analysis of treatment options for vestibular schwannoma

Clinical article

ROBERT G. WHITMORE, M.D.,† CHRISTOPHER URBAN, B.S.,† EPHRAIM CHURCH, B.A.,† MICHAEL RUCKENSTEIN, M.D.,† SHERMAN C. STEIN, M.D.,† and JOHN Y. K. LEE, M.D.†

Departments of †Neurosurgery and †Otornolaryngology, Hospital of the University of Pennsylvania Philadelphia, Pennsylvania

Base case: typical patient covered by decision analysis –

In this study, tumor is small, useful hearing is still preserved
Rollback Analysis: Limitation

- Conclusion point estimate – single #
- Considering many component calculations, this must be suspect
- Role of uncertainty – if your estimates are off, can your conclusions be trusted?
- No single measure of overall uncertainty, as in individual trial
Dealing with Uncertainty

Sensitivity Analysis

Monte Carlo Simulation
- Calculate 95% CI for each probability & utility value
- Random walks: a bunch of simulated trials in which each value is varied
- Gives mean outcome & SD

Vary parameters which have greatest impact on outcome
Comparative Effectiveness

Form of Decision Analysis

- **Efficacy vs. Effectiveness**
  - *efficacy* requires careful controls, randomization
  - strict inclusion/exclusion criteria
  - “intent to treat” analysis
  - *effectiveness* = how treatment works in “real world”
  - often uses uncontrolled, observational data
  - “as treated” analysis
  - data need not come from publications
Choosing the best operation for chronic subdural hematoma: a decision analysis

Clinical article

*Bradley C. Lega, M.D.,¹ Shabbar F. Danish, M.D.,¹ Neil R. Malhotra, M.D.,¹ Seema S. Sonnad, Ph.D.,² and Sherman C. Stein, M.D.¹

Departments of ¹Neurosurgery and ²Surgery, University of Pennsylvania, Philadelphia, Pennsylvania

Included unpublished databases
Quantity of Life

- As important as quality of life (QOL)
- e.g. advantage of radiosurgery probably lost at 10 years & reversed later in vestibular schwannomas
- Usually measured in years of life
- Combine quality/quantity using quality-adjusted life years (QALYs)
- e.g. 2 years in perfect health = 2 QALYs = 4 years at utility of 0.5 or 10 years at 0.2
Role of Costs

“Try to focus less on a cure and more on a treatment you can afford.”
Cost-Effectiveness

• Compares both effectiveness & costs of 2 or more treatments
• Calculate expected costs of a treatment same way as expected utilities
• Compare cost/effectiveness of 2 or more treatments by calculating incremental cost-effectiveness ratio (difference in costs, divided by difference in effectiveness)
Is aggressive treatment of traumatic brain injury cost-effective?

Clinical article

Robert G. Whitmore, M.D.,¹ Jayesh P. Thawani, M.D.,¹ M. Sean Grady, M.D.,¹ Joshua M. Levine, M.D.,¹⁻³ Matthew R. Sanborn, M.D.,¹ and Sherman C. Stein, M.D.¹

Departments of ¹Neurosurgery, ²Neurology, and ³Anesthesiology and Critical Care, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania

How many QALYs does the severe TBI patient gain from aggressive treatment? Is it worth the added cost?
Threshold Study

• What if you are considering a new treatment without a track record?

• Can compare it to an established study and ask how good it needs to be to yield better results than standard care (effectiveness threshold)
Deep brain stimulation compared with methadone maintenance for the treatment of heroin dependence: a threshold and cost-effectiveness analysis

James H. Stephen¹, Casey H. Halpern¹, Cristian J. Barrios¹, Usha Balmuri¹, Jared M. Pisapia¹, John A. Wolf¹, Kyle M. Kampman², Gordon H. Baltuch¹, Arthur L. Caplan³ & Sherman C. Stein¹

¹Department of Neurosurgery, Centers for Functional and Restorative Neurosurgery, Hospital of the University of Pennsylvania, Philadelphia, PA, USA; ²Department of Psychiatry, Center for Study of Addiction, Hospital of the University of Pennsylvania, Philadelphia, PA, USA; ³Department of Medical Ethics, Center Study of Bioethics, Hospital of the University of Pennsylvania, Philadelphia, PA, USA.
Questions?

Thank You