Paul C. McCormick, MD, MPH
Columbia University College of Physicians and Surgeons

EVIDENCE-BASED SPINE CARE IN THE ERA OF COMPARATIVE EFFECTIVENESS RESEARCH

Disclosures
- None

Blood-letting
- Take 500 poor people with fevers
- Divide them in halves
- I will cure them without blood-letting
- You do, as ye know
- We shall see how many funerals both of us shall have

Evidence-Based Medicine
A New Approach to Teaching the Practice of Medicine

Evidence-Based Medicine Working Group
JAMA, November 4, 1992—Vol 268
12 sailors with scurvy
- 2- quart of cyder
- 2- elixir vitriol
- 2- sea water
- 2- vinegar
- 2- acidulated gargle
- 2- oranges, lemon

“lemons and oranges were the most effectual remedies for this distemper at sea”

---

**Pioneer of Evidenced-based Medicine**

- ‘The end result idea’
- 5x8 card every surgery
- Review card, results at one year
- Errors classification
- Bone sarcoma registry

---

**Modern Evidence-based Medicine**

---

**British National Health Service (NHS)- 1948**

- “My job is to give you all the facilities, resources and help I can, and then to leave you alone… to use your skill and judgment without hindrance.”

---
Modern Evidence-based Medicine

- I read a pamphlet on the advantages of the freedom of British doctors to do what they want.
- I found it ridiculous
- I would sacrifice all my medical freedom for some hard evidence when to do a pneumothorax

Archie Cochrane

Medical statisticians (aka epidemiologists) influence decisions of individual physicians

- Epidemiology—distribution and determinants of disease in populations, not individual patients

Bradford Hill

Modern Evidenced-based Medicine

- Effective—only if demonstrated (RCT) to do more good than harm
- Efficient—uses available resources to maximize the delivery of effective interventions
- Equality—equal access to care, remove socioeconomic barriers

1972
The Cochrane Collaboration

- 28,000 from 100 countries
- Develop, maintain and disseminate up to date systematic reviews of RCTs of health care
- Publish in Cochrane Library (>4,000)

Evidenced-based Medicine

- Canadian Health Sciences
- Evaluate effectiveness of interventions in national Medicare program
- Structure, methods, analysis
- Implementation into training and practice
- “Scientific-based medicine.”

Systematic reviews

- Search literature
- Objective, rigorous, validated quantitative methods
- Assess/grade evidence
- Combine/synthesize total evidence
- Consistent output: Clinical practice guideline, technology assessment, coverage policies

EBM: Hierarchy of Evidence

- Randomized Controlled Double Blind Studies
- Cohort Studies
- Case Control Studies
- Case Series
- Case Reports
- Ideas, Opinions

6/3/2011
Evidence-based Medicine (EBM) in the US

- Rise in health care costs 1970s
- Health care costly and ineffective
- Health care utilization studies

Public Policy

- Congress mandate “patient outcome assessment research program.”
- Agency for Health Care Policy and Research (AHCPR)-1990 (AHRQ-1999)
- “encourage physicians to change practice patterns” by using outcomes research findings
- Developed clinical practice guidelines (1992-1996 (e.g. Lumbar stenosis))

- Unexplained regional variations in treatments (prostate, spine surgery)
- Uncertainty of value of treatments
- Inadequate knowledge of what works

- 5 year contracts with EPCs
- Produce evidence reports and technology assessments
- Coverage decisions (CMS), guidelines, quality measures
Evidenced-based Medicine

- Not enough evidence
  - Doesn’t generate new evidence
  - Synthesis of existing evidence
- Rapid pace of innovation, knowledge growth
- Inconclusiveness of evidence

Evidence-based Medicine

- Reliance on RCTs
  - Expensive
  - Time consuming
  - Lack of generalizability to real world
  - Average truth
    - Subgroups, individuals
  - Ecological fallacy
    - Individual response same as aggregate

Limitations of EBM: The assumptions

- RCT as gold standard of study design (Class I evidence)
- Randomized treatment assignment
- Eliminates bias, confounding
- Any difference in outcome is attributable to treatment

Limitations of EBM: RCT assumptions

- Frequentist assumptions and inference
  - The patient as a random variable
  - Outcomes expressed as relative frequency
  - The truth as a mean
- Patient heterogeneity (i.e. prognostic variables, factors associated with outcome) as bias
  - Factor out (randomization) rather than assess effect on outcome
The frequentist method of inference

- Assume underlying truth (i.e., null hypothesis)
- No difference in averaged ODI score
- Deductively calculate the frequency of all possible outcomes under null

Frequentist inference: Assumptions

- Individual data points (e.g., patient outcome) are random variables (RV)
- RVs take on different values by chance
- Individual RV value unknown but distribution and probability of different possible values objectively calculated by mathematical formula
- Probabilities as relative frequencies

Frequentist inference: Assumptions

- Single outcome (RV) not real or true, just random scatter
- Probability curve, mean, SD are real-calculated, objective, measurable
- Truth resides in the mean, not individual outcomes

Is patient outcome from surgery a random event?

1. Every patient has known probability of achieving every possible outcome
2. Each patient just as likely to gain 29.4 or lose 6.6 points on ODI score
3. If gain 29.4 ODI in trial #1, just as likely to lose 6.6 ODI in trial #2
Is patient outcome from surgery a random event?

4. If trial repeated infinite amount of times on same patient, this patient would eventually experience all possible outcomes.

RCT assumptions

- Record all measured outcomes
- Calculate mean, standard deviation of all outcomes
- Single mean (average) value is “true treatment effect (TTE).”
- RCT: seeks unbiased estimate of “true treatment effect.”

RCT assumptions: Truth as an average

- TTE: Arithmetic average (mean) of each outcome from RCT (e.g. ODI, VAS)
- Outcome each patient can expect from procedure (same for every patient)
- Each treatment has single TTE value (i.e. constant) for every treated patient
- Any variation in actual outcome occurs by chance alone

Evidence based medicine: Average based Medicine

- Evidence for an average truth
- Truth as an average or an average truth
- Calculated mean as truth is a fallacy
The true average temperature in NYC is 56°F.

56°F is not the truth in January, February, July, August.

ECRI- “no evidence fusion provides MCID (>15 ODI).”

Not the truth for 42% of patients.
The truth is not a mean—calculated from an average of all possible outcomes that include both successes and failures. The truth resides in each individual patient. Individual patient success or failure is not a random event that we can never know. Uncertainty of individual outcome due to our insufficient knowledge of patient, not random variation.

Is patient outcome (e.g. success, fail) uncertain because it is a random event, occurs by chance alone, and we can’t know it? (stochastic)

Is patient outcome uncertain because we have insufficient knowledge of specific patient factors predictive of outcome? (deterministic)

Effectiveness is an intrinsic property of treatment (surgical procedure, drug)

Average Based Medicine: The tyranny of the mean

- The truth is not a mean—calculated from an average of all possible outcomes that include both successes and failures.
- The truth resides in each individual patient.
- Individual patient success or failure is not a random event that we can never know.
- Uncertainty of individual outcome due to our insufficient knowledge of patient, not random variation.
**Evidence-based Medicine**

- “Integrating clinical expertise with best available clinical evidence from systematic research.”
  - David Sackett
- “Art kills”
  - David Sackett

**The limitations of RCT design for spine**

- Heterogeneity of spinal conditions, patient factors, treatment response
- Random allocation nullifies identification/assessment of any prognostic factors
- Strong patient preferences
- Gain little knowledge from RCTs, EBM about individual patients to improve future care
  - SPORT
  - Tech assessment lumbar fusion

**Physicians’ and patients’ choices in evidence based practice**

- Evidence does not make decisions, people do

**Population-based epidemiology**

- Overlook subgroup, individual treatment response
- No way to incorporate patient preferences, unique circumstances, physician expertise in individual decisions
- Policy/coverage decisions

**Net True treatment effect of surgery**

- Net True treatment effect of surgery = mean surgery outcome - mean PT outcome
- Average BP β-blocker - Average BP Diuretic = net TTE of β-blocker

**ECRI Institute**

Spinal Fusion and Discography for Chronic Low Back Pain and Uncomplicated Lumbar Degenerative Disc Disease

Pre-post change can be used as a surrogate for the between-group difference. We view the control group as a surrogate for the active-treatment group: what change would the surgical patients have experienced if they had received non-surgical treatment? Taking this view, the between-group difference at follow-up is therefore an unbiased estimate of the surgical group’s change score after factoring out the non-surgical treatment effect. Thus, our use of 10 minus our

**ECRI Institute**

Spinal Fusion and Discography for Chronic Low Back Pain and Uncomplicated Lumbar Degenerative Disc Disease

Pre-post change can be used as a surrogate for the between-group difference. We view the control group as a surrogate for the active-treatment group: what change would the surgical patients have experienced if they had received non-surgical treatment? Taking this view, the between-group difference at follow-up is therefore an unbiased estimate of the surgical group’s change score after factoring out the non-surgical treatment effect. Thus, our use of 10 minus our
Patients with more severe symptoms refused participation in SPORT RCT
- Patients with more severe disease assigned to non-operative Rx. crossed over to have surgery
- Patient with less severe disease assigned to surgery crossed over to have no-operative Rx.

“Conclusions about the superiority or equivalence of the treatments under study are not warranted based on the intent-to-treat analysis.”

Individual response to treatment varies according to:
- Unique patient characteristics
- Genetic, molecular profile
- Environmental exposure

Effectiveness not an intrinsic property of treatment (drug, surgery)

“It’s far more important to know what person the disease has than what disease the person has.”
Hippocrates

Personalized Medicine

- Braco-2- Breast Ca.
- Colon Ca - KRAS gene mutation - Erbutix
- Warfarin - gene variation for CYP2C9 metabolizing enzyme

Ineffectiveness of particular drug by category
Personalized Medicine

- Tailor treatment to individual characteristics of each patient
- Classify individuals into subpopulations that differ in response to a specific treatment
- Therapeutic interventions can then be concentrated on those who will benefit, sparing expense, side effects, and complications of those who will not

Comparative Effectiveness Research (CER)

- Identify
  - What works
  - For which patients
  - Under what circumstances

Comparative Effectiveness Research

- Systematic reviews
- Funding for new research
- Expanded study design, data sources
  - RCTs, practical RCTs, observational studies
  - Registries, administrative databases, EMR
  - High throughput technologies (e.g., genetics)
- Cost, value effectiveness
- Direct comparison to alternative treatment
  - Benefit, harm, costs

Comparative Effectiveness Research

- 2009 ARRA-$1.1 Billion dollars
- ACA-$500 million/yr. trust fund
- Tax on every covered life
- Exempt from Congress
- NIH- primary research (RCTs)
- AHRQ- systematic reviews, registries
- DHHS- data infrastructure, dissemination, translation
Comparative Effectiveness Research in Spine

- Direct comparison of effective treatments for specific condition
- A reductive approach (either/or but not both)
- Non-operative and operative treatment for spine conditions not directly comparable

Comparative Effectiveness Research in Spine

- Heterogeneous conditions
  - Some improve with non-op Rx., some don’t
  - Some improve with surgery, some don’t
- Nearly all patients initially treated with non-operative measures
- If patient improves- no further Rx.
- If no improvement- consider surgery
- Complementary- Applied to different subgroups, under different circumstances, at different times

We do want to discriminate (i.e. select for surgery) on the basis of patient characteristics that predict outcome
Comparative Effectiveness Research in Spine

- CER comparisons of net benefit of non-op Rx. in one group to benefit of surgery in another group (who initially did not respond to non-op Rx.) will always favor non-op Rx.

- Increase risk and cost associated with surgery

- Surgery – treatment of last resort
  - Most disabled, psychosocial, chronic pain

Personalized Medicine for Spine

- Condition heterogeneity
  - Structural, clinical, psychosocial, natural history

- Treatment options
  - Surgical and non-surgical

- Patient preferences

- Response to treatment
  - Patient to patient (e.g. herniated disc)
  - Surgeon to surgeon

- Define individual patient profile

Evidence we need

- High fidelity prospective patient outcome data in real world practice

- Administered, collected and analyzed by neurosurgeons

- Sustainable, updated, learning model

- Patient profiles
  - Identify, define variables in individuals, subgroups that predict treatment response

- Disease modeling, real world effectiveness, risk adjusted outcomes

NeuroPoint Alliance, LLC.

- Need for broad based prospective practice data collection at national and individual practice level

- NPA- develop internet based data submission and management infrastructure and platform

- Supported by AANS, CNS, Joint Spine Section

- Registries, PQRI, MOC, CER

- Projects: ABNS MOC, PQRI, NeuroPoint SD, N²QOD
Record all measured outcomes
- Calculate mean, standard deviation
- Single mean (average) value is “true treatment effect (TTE).”

RCT assumptions: Truth as an average

- Assume null
  - Can never prove a theory (Falsification)
  - Hard to prove a negative (ineffectiveness)
- Deductive (statistical) inference: objective, logical (Frequentist)
- Truth as a population average
- Probability of truth - a relative frequency
- How frequently inference of the average truth will be wrong

From evidence to truth: Logic and decision rules (Science)