

# Medical Research from Medical Records

Mikel Aickin, PhD  
Family & Community Medicine  
University of Arizona, USA

# The Problem

- Medical researchers believe that therapeutic knowledge comes from randomized clinical trials (RCT)
- Physicians and other practitioners believe that therapeutic knowledge comes from their experiences with their patients, and the experiences of other practitioners.



RCT is everything

Patient is everything

Me

# Why Not RCTs?

- Expensive – not sustainable
- Slow – delay in improving practice
- Artificial – not faithful to practice
- Restricted – many important clinical questions will never be answered
- Irrelevant – no answers for specific patients

*The main technical limitation of clinical trial methodology is that it assumes the homogeneity of the research populations which it studies. It is looking for similarities of response to a treatment but it does this in a way which obscures differences of response. The assumption of homogeneity is enshrined in the use of averages and the statistics based on their comparison....For the individuals in the trial a comparison of means may hide much more than it reveals.*

*John Heron, 1986*

# Terminology

- RCT = randomized clinical trial
- EMR = electronic medical record system
- registry = data collected on patients with a particular condition (like cancer) who may be followed forward in time (in US English)
- register (in European sense) = EMR system in US sense (?)
- CAM = complementary & alternative medicine

# Why EMRs for Research?

- Faithful to medicine, as it is practiced
- Representative of patients who are actually seen in clinics
- Genuine – in real rather than research settings
- Economical – the data are already being collected, and the datasets are huge

# Why Not EMRs?

- Most biomedical researchers do not believe that you can do valid research without a research intervention.
- But the funding for non-interventional research has been rather small.
- Therefore, it is not surprising that methods have not been developed and we do not see published success stories.

# Why Not EMRs?

- If you do not know why a treatment was chosen, how can you know that there is no subtle bias in the assignment of treatment?
- Matching patients on many factors may reduce this objection very substantially.
- But perhaps the real message is that EMRs should try to include information on why a treatment was prescribed.

# Why Not EMRs?

- EMRs do not collect outcome data
- This is the single most important objection to EMRs for research.
- The ways of dealing with it in terms of existing EMRs will be questioned.
- But the real point is, if EMRs could include outcome data, would that not make them very, very attractive for research?

# Why Not EMRs?

- If a patient leaves a practice, what does that mean?
- Not relevant in “closed” patient systems.
- Implies that follow-up needs to go beyond clinical practice in “open” systems.

# EMR Approach – fake RCTs

- Synthetic RCTs
  - Try to extract data from EMRs that are “like” that which would have been obtained from an RCT
- Always an issue about what “like” means
- Does not answer the questions about patient-specific therapeutic recommendations.

# EMR Approach – MCGs

- MCG = matched comparison group
  - patients with a particular condition are matched on a relatively long list of criteria
    - ▶ if they are all given any particular therapy, they should all respond the same (because they are matched so closely) ◀
  - randomization becomes irrelevant; we will learn the same regardless of how therapies are chosen

# EMR Approach – MCGs

- Depends on having a large EMR data system, so that good MCGs can actually be formed.
- Also requires that there be different therapeutic choices within MCGs.

The EMR approach is visionary.

And therefore under-funded.

However...



- Studies in Patient-centered Informatics for Comparative Effectiveness Research
- Funded by the NIH, NCCAM under the “stimulus” umbrella.
- Aim: To compare integrative medicine approaches to conventional medicine approaches in areas of clinical importance.



- Collaboration between the Marino Clinics in Boston and the University of Arizona
- Marino clinics provide both conventional and CAM services.
- Integrative medicine = introducing CAM into conventional practice
- Marino has an excellent EMR for about 8 years.

# SPICER targets

- Back pain ◀ *first topic*
- Irritable bowel syndrome
- Fibromyalgia
- Peri-menopausal symptoms

# Acquiring Data

- Always a problem with EMR systems
- Solution: simplification of the interface between IT professionals and the analysts:
  - based on events, that occur at times to patients
  - events can have multiple characteristics
- Research data is independent of the complexity of the EMR system
- Minimizes the work on the IT side

# What Data?

- Demographics (age, gender)
- Diagnoses (ICD9)
- Visits (type of service)
- Procedures (CPT)
- Claims (utilization) (per CPT & ICD9)
- Providers (who saw the patient?)
- Narratives (chart notes and so on...)
- Medications

# Medications

- Especially important for pain conditions
- Prescribed vs. filled vs. taken
- Different EMR systems represent drugs in very different and inconsistent ways
  - ATC (anatomic, therapeutic, chemical)
  - NDC (US) (identifies a drug product)
  - generic name (several systems)
  - trade name (very complicated)

# What is Back Pain?

- 7,000 – 8,000 BP patients in a clinic with an effective pool of 35,000-55,000 patients.
- 30 ICD9 back pain diagnoses
- Reduced to 6 categories
  - 1. Lumbar symptom
  - 2. Lumbar anatomical
  - 3. Thoracic symptom
  - 4. Cervical symptom
  - 5. General
  - 6. Brachial neuritis

# Variety of BP presentations

Patients	BP diagnosis	
2410	1.....	lumbar symptom
1356	....5.	general
1296	...4..	cervical symptom
552	.2....	lumbar anatomic
442	..3...	thoracic symptom
394	..34..	thoracic symptom + cervical symptom
328	12....	lumbar symptom + lumbar anatomic
179	.....6	brachial neuritis
178	1..4..	lumbar symptom + cervical symptom
150	12.4..	...and so on...
129	1.3...	
94	...45.	
91	1...5.	
83	1.34..	
66	12..5.	
54	123...	
46	..34.6	
45	1234..	
41	...4.6	

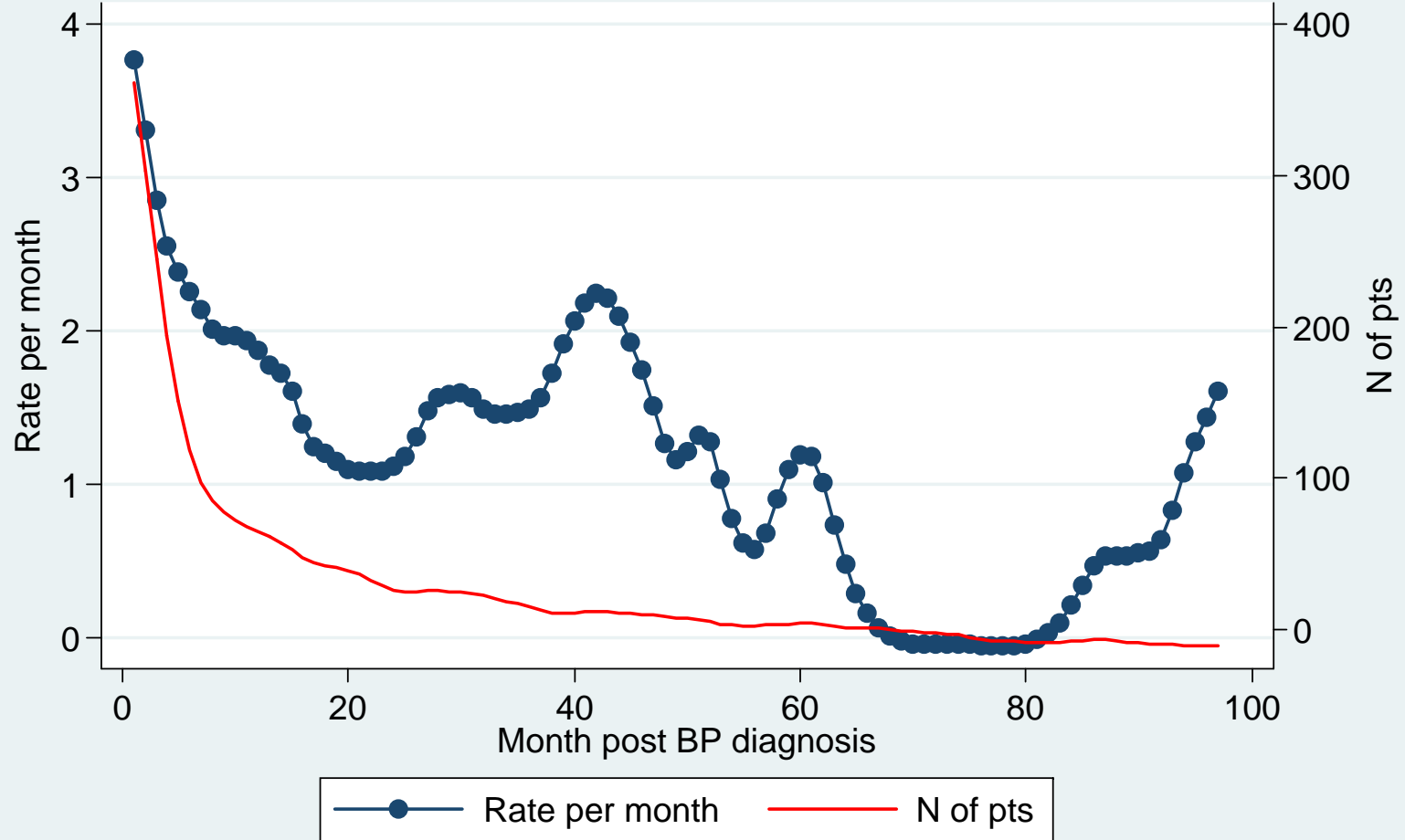
...and so on ...

# Variety of BP treatments

IM	Ever Conventional		Total
	0	1	
....	0	4,130	4,130
...m	17	3	20
..c.	1,034	1,111	2,145
..cm	1	0	1
.b..	125	93	218
.b.m	1	0	1
.bc.	36	59	95
a...	290	62	352
a..m	2	0	2
a.c.	32	42	74
a.cm	0	1	1
ab..	9	4	13
abc.	1	10	11
Total	1,548	5,515	7,063

a = acupuncture   b = body work   c = chiropractic   m = massage

## Acupuncture among Lumbar Symptom Pts



# What We Don't Know Yet

- Can we assess outcomes in the data we have?
- Can we really distinguish different therapeutic pathways?
- Can we produce patient-specific recommendations?

# Critical Issues for Research EMRs

- Outcomes must be recorded
- Reasons for choosing therapies must be recorded
- Follow-up information must be complete (which may mean closed patient populations)

# Hope

- EMRs can be used for research now.
- But they need to be improved for research in the future.
- And if they are, then we can raise the quality of medical care, informed by medical research.

# Medical Research from Medical Records

Mikel Aickin, PhD  
Family & Community Medicine  
University of Arizona, USA

[maickin@earthlink.net](mailto:maickin@earthlink.net)