

Expanding Use of Real-World Evidence A National Academies Workshop Series

Gregory Simon MD MPH – Kaiser Permanente Washington Health
Research Institute

Acknowledgements to:

Co-Chair: Mark McClellan

National Academies

Amanda Wagner Gee

Carolyn Shore

Planning Committee, Speakers, Panelists, Discussants:

Adam Haim

Douglas Dumont

Rachael Fleurence

Andy Bindman

Jeff Hurd

Rory Collins

David Madigan

Issac Reeser

Morgan Romine

Elliot Levy

Jaqueline Corrigan-Curry

Sebastien Schneeweiss

Joanne Waldstreicher

Melissa Robb

Sharon Terry

John Graham

Rob Califf

Jeff Allen

Marcus Wilson

Rachel Sherman

Richard Kuntz

Rich Platt

Deven McGraw

Martin Landray

Michael Horberg

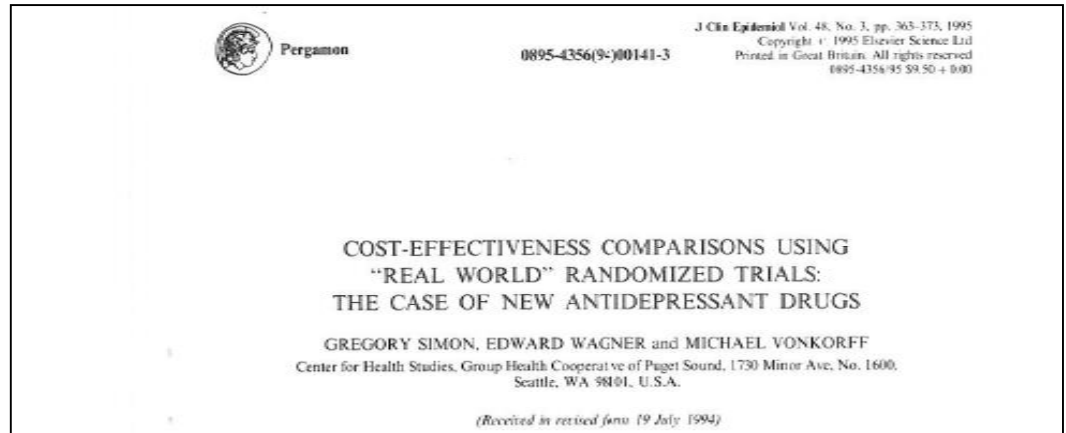
Gracie Lieberman

David Martin

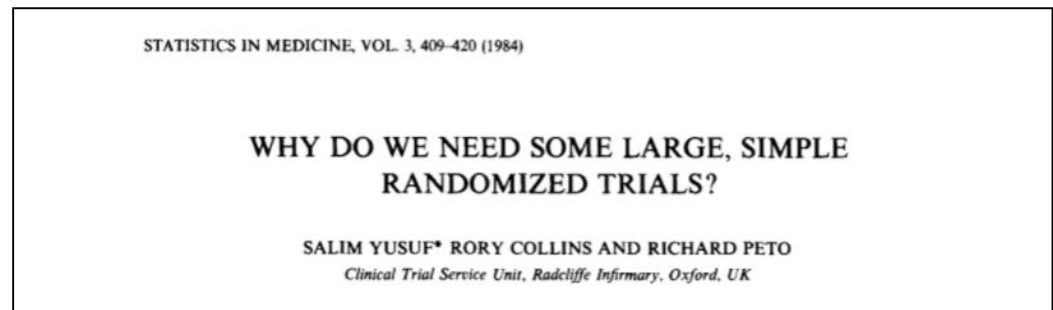
Bill Potter

Ancient history:

1994:



1984:



Recent history:

- NASEM Drug Forum workshop on Real World Evidence (Oct 2016)
 - Stakeholder priorities, Variety and value of real-world data, promising examples
- FDA Workshop on Real-World Evidence Generation (Dec 2016)
 - Enabling developments, use cases, infrastructure, models for implementation
- Duke Margolis Center Collaborative to Advance Real-World Evidence
 - Stake holder engagement to promote use of RWE in regulatory decisions
 - Focus on concept of “fit for purpose”
- NASEM Workshop Series on Real-World Evidence and Medical Product Development
 - Sept. 19-20, 2017 – Identifying barriers, aligning incentives, re-examining traditions
 - March 6-7, 2018 – Turning real-world data into evidence: 3 key questions
 - July 17-18, 2018 – Test-Driving Useful Tools

NASEM Workshop 1: Incentives & Barriers

<http://bit.ly/RWEworkshop1>

<http://bit.ly/RWEproceedings1>

Our research traditions can be:

- Vital anchors to our central purpose
- Or just anchors that keep us stuck

How might we know the difference?

Five dialectics:

- Definitions: RWD vs. RWE
- Regulators: Arbiters vs. Curators
- Traditions: Icons vs. Idols
- Departures from Tradition: Virtues vs. Necessities
- Value of Information: Validity vs. Credibility

RWD vs. RWE

- RWD = Data derived FROM the real world:
 - Routine health care clinical or business operations
 - Observation of free-living humans
- RWE = Evidence relevant TO the real world
- RWD does not always make RWE
- RWE usually starts with RWD

Arbiter vs Curator

- Scott Gottlieb: As data become more diverse (to match diverse purposes), FDA may become a curator rather than an arbiter.
- But... what model of curation should we follow?
 - Sundance (Restricted entry, refereed by elites)
 - YouTube (Free entry, refereed by the crowd)

Icons vs. Idols

- Icon: An exemplar that illuminates or animates
- Idol: A surface appearance that distracts

“Good Clinical Practice” called out as our Golden Calf

Virtues vs. Necessities

- The RWE mantra is “faster, better, cheaper”
- Generating evidence faster and cheaper is necessary
 - We ask: What might we lose? Is it good enough?
- Departures from tradition are sometime virtuous
 - We ask: What might we gain? Is it actually better?

Validity vs Credibility

- Credible = Simple, but could be misleading
- Valid = Accurately predicts, but may be obscure
- Two examples in our discussion:
 - Clinical data vs. traditional evidence
 - Traditional RCT vs. more complex methods

What is RWE? – Core Qualities

- Generalizable
- Relevant
- Adaptable
- Efficient

RWE is Generalizable

- Generalizability is more about prediction than resemblance
- Prediction is context-specific, but that's testable
- Predictions are accountable (A scary thought!)

RWE is Relevant

- Grounded in stakeholder priorities
- Directly addresses decisional needs
- “Fit for purpose” presumes diverse purposes

RWE is Adaptable

- Must embrace (and attempt to understand) heterogeneity of patients, providers, and systems.
- Answers not expected to apply everywhere and for all time (But how do you regulate with that?)

RWE is Efficient

- Because answers may be disposable, they should be fast and cheap to create.
- Economy can promote clarity (if we do it right)

NASEM Workshop 2: Specific Questions

<http://bit.ly/RWEworkshop2>

What is Real World Evidence?

Two Challenges

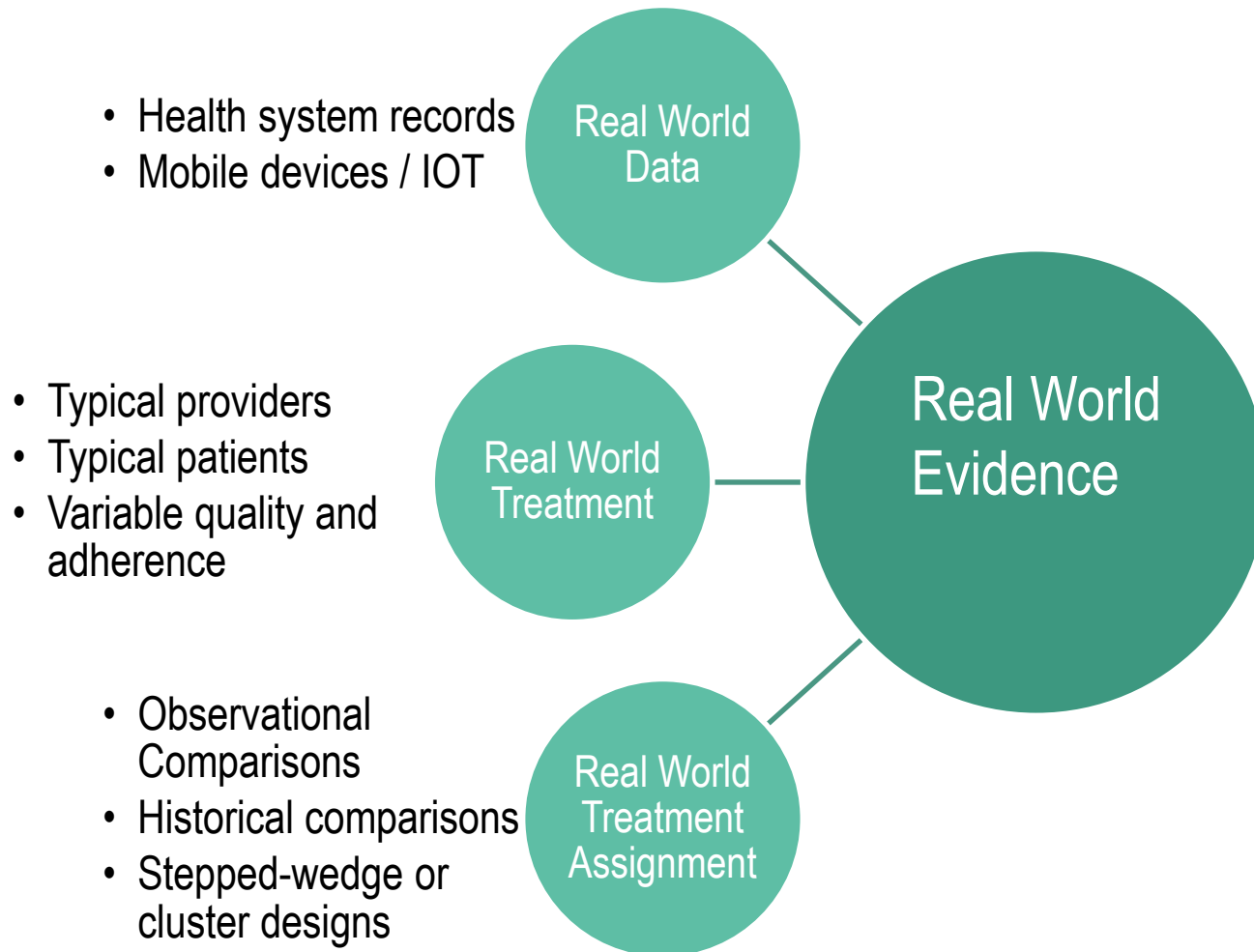
- Mark McClellan: If we're still defining the term, have we made any progress?
- Rory Collins: The term "real world evidence" has so many meanings that it's not useful any more. We should retire it.

What is Real World Evidence? All sorts of things.



What is Real World Evidence?

Three concepts



Can we rely on Real World Evidence?

Three questions:

- Can we rely on real world data?
- Can we rely on real world treatment?
- Can we learn from real world treatment assignment?

WHEN can we rely on Real World Evidence?

Three better questions:

- WHEN can we rely on real world data?
- WHEN can we rely on real world treatment?
- WHEN can we learn from real world treatment assignment?

WHEN can we rely on Real World Evidence?

Three answers:

- WHEN can we rely on real world data?
 - It depends.
- WHEN can we rely on real world treatment?
 - It depends.
- WHEN can we learn from real world treatment assignment?
 - It depends.

WHEN can we rely on Real World Evidence?

Three answers:

- WHEN can we rely on real world data?
 - It depends.
- WHEN can we rely on real world treatment?
 - It depends.
- WHEN can we learn from real world treatment assignment?
 - It depends.

Depends on what?

When can we rely on real-world data?

- When can we rely on EHR data from real-world practice to accurately assess study eligibility, key prognostic factors, and study outcomes?
- When can we rely on data generated outside of clinical settings (e.g. mobile phones, connected glucometers or blood pressure monitors)?
- Does adjudication or other post-processing of real-world data add value or just add cost?

When can we rely on real-world data?

- The pathway from a clinical phenomenon to a study dataset includes several distinct steps, each of which can introduce error.
- Distinct steps in the data “chain of custody” require distinct methods for assessing data quality/integrity.
- Timing of assessments in practice-generated data can be a significant (and unrecognized) source of bias.
- Random error is not always “conservative” (e.g. In a non-inferiority design, random error biases toward finding equivalence).
- Transparency regarding methods and (when possible) intermediate data steps is necessary for credibility.
- Data collection processes of traditional clinical trials are far from a “gold standard”.

When can we trust real-world treatment?

- Safety

- Can community clinicians safely deliver study treatments and monitor/respond to adverse events?
- What reporting and monitoring is useful (rather than wasteful)?

- Effectiveness

- What level of treatment quality/fidelity/adherence is necessary for valid inference?
- When is variation in fidelity or adherence signal instead of noise?

When can we trust real-world treatment?

- Selection of patients, clinicians, and/or practice settings may influence differences between treatments – especially when treatment quality/fidelity or adherence is variable.
- Placing a “floor” under treatment quality can introduce a tension between generalizability and participant safety.
- Controlling treatment quality or adherence is a choice – assessing and reporting it is not.
- Blinding providers and/or patients may reduce some biases, but it can distort true differences between treatments – and add to cost.
- The purpose of monitoring for adverse events is quite different for new treatments vs. established treatments.
- “Enrichment” designs (selectively enrolling participants with specific clinical characteristics) can inform personalized treatment selection.

When can we learn from real-world treatment assignment?

- When can we rely on inference from cluster-randomized or stepped-wedge study designs?
- Under what conditions can we rely on inference from observational or naturalistic comparisons?
- How could we judge the validity of observational comparisons in advance - rather than waiting until we've observed the result?

When can we learn from real-world treatment assignment?

- The fundamental question concerns confounding by indication – and this is distinct from questions regarding data quality.
- Some sources of bias must be addressed in study design (rather than accounted for in the analytic phase).
- Observational comparisons should assess and report sensitivity analyses estimating the magnitude of unmeasured confounding necessary to change the qualitative result.
- Transparency regarding analytic methods is always expected, and use of standard tools strongly preferred.

NASEM Workshop 3: Useful Tools

<http://bit.ly/RWEworkshop3>

Decision Aids: Useful tools for producers and consumers of Real-World Evidence

- Analogy to clinical tools for shared decision-making:
 - There is no right answer for all situations (It depends).
 - But there are useful questions to ask (It depends on what)
 - Additional obligation for transparent reporting
- Four candidate decision aids:
 - Are data from practice fit for a specific research purpose?
 - Blinding in effectiveness or pragmatic trials: why, who and when?
 - Controlling treatment quality and adherence
 - Assessing and addressing potential for unmeasured confounding

Are data from practice fit for purpose?

Potential points of error in “chain of custody”

- Ascertainment: presentation to this clinical setting
- Assessment: accuracy of diagnosis
- Recording: influences on data entry at point of care
- Extraction: completeness and de-duplication
- Harmonization: translation to common data model
- Reduction: specifications or computable phenotypes

Decision Aid Questions:

Are data from practice fit for purpose?

- Is ascertainment reasonably complete (or at least unbiased)?
- Can real-world clinicians accurately assess the clinical phenomenon of interest?
- Are those real-world assessments consistently recorded across time, setting, etc.?
- Can those data be accurately and efficiently extracted?
- Can data from different sources/systems be combined and harmonized?
- Does data reduction introduce error or bias?

Next Steps

- Brief summary of NASEM Workshop 2 (*anticipated release in early July*)
- NASEM Workshop 3: Application (*July 17-18, 2018*)
- Capstone summary of all three workshops