Introduction to Real World Evidence Studies

David Morgan
Pharmaceutical Medicine Group, King’s College London
Formerly VP Chief Statistical Scientist, Ipsen

ICR Project Management SIG – 11 October 2019
Forbes Magazine May 2018: Will Real World Performance Replace RCTs as Healthcare's Most Important Standard?

Social Media-sourced Real-world Evidence - A Novel, Cheap, Effective Method

Big Data: Prescription for the Pharmaceutical R&D Plight
THE BENEFITS OF SOCIAL MEDIA SOURCED RWE:

✓ Low-cost solution in comparison to RCTs; often up to a tenth of the cost
✓ Data is collected on a large scale, enabling detailed / granular analysis
✓ Studies can be conducted into areas where conventional healthcare research is not possible, or difficult
✓ The data collection process is not governed by clinical trial regulations, offering greater control
✓ It provides evidence to strengthen existing approaches and develop new ones

SURVEY HIGHLIGHTS

Data science approaches on real world evidence may even replace controlled experimentation

56% RWE supplement controlled experiment
44% RWE replace controlled experimentation
<table>
<thead>
<tr>
<th>Level</th>
<th>Therapy/Prevention, Aetiology/Harm</th>
<th>Prognosis</th>
<th>Diagnosis</th>
<th>Differential diag/symptom prevalence</th>
<th>Economic and decision analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>1a SR (with homogeneity*) of RCTs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SR (with homogeneity*) of inception cohort studies; CDR† validated in different populations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SR (with homogeneity*) of Level 1 diagnostic studies; CDR† with 1b studies from different clinical centres</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SR (with homogeneity*) of prospective cohort studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SR (with homogeneity*) of Level 1 economic studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1b</td>
<td>Individual RCT (with narrow Confidence Interval†)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Individual inception cohort study with &gt; 80% follow-up; CDR† validated in asingle population</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Validating** cohort study with good†† reference standards; or CDR† tested within one clinical centre</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prospective cohort study with good follow-up****</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Analysis based on clinically sensible costs or alternatives; systematic review(s) of the evidence; and including multi-way sensitivity analyses</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1c</td>
<td>All or none§</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>All or none case series</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Absolute SpPins and SnNouts††</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>All or none case-series</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Absolute better-value or worse-value analyses ††††</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>SR (with homogeneity*) of cohort studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SR (with homogeneity*) of either retrospective cohort studies or untreated control groups in RCTs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Levels of Evidence for Therapy

- 1A – Systematic Review (with homogeneity) of Randomised Clinical Trials
- 1B – Individual RCT with narrow Confidence Interval
- 2A - Systematic Review (with homogeneity) of cohort studies
- 2B – individual cohort study
- 3 – Case control studies
- 4 – Case series (and poor quality cohort and case control studies)
- 5 – Expert opinion without explicit critical appraisal
Why are RCTs seen as the “Gold Standard”?

- Randomisation ensures that treatment groups are similar in every respect other than the investigative treatment applied;
  - For factors which are known to be important
    - Could stratify the randomisation to improve precision
  - And for factors whose impact is unknown

- Blinding ensures that assessment of outcome is not affected by knowledge of the patient’s assigned treatment (by patient or investigator or operational staff)
  - Avoids bias
  - High internal validity
But …

• “Your trials include
  • Patients from 18-65
  • Generally healthy apart from disease you are investigating
  • Taking no other drugs
• “The patients I see are
  • Mostly over 65
  • Suffering from several diseases
  • Taking multiple drugs
• “How can I be sure that your trial results are applicable to my patients?”
• Does your drug work in the real world?
RCT vs. RWE

Randomized Control Trial

Patients are randomized to the treatments; physicians’ and patients’ choices are not considered for selection of the treatment.

Non-adherent patients are taken out of the analysis.

Experiment is based on an artificially created homogeneous treatment group. The purpose here is to establish the efficacy of the medication/therapy.

Real-World Evidence

Therapy or medications to patients are determined by doctors’ choices as per the standard practice.

Non-adherent patients can switch the treatment and in such a case are likely to remain included.

Contains heterogeneous patient population reflecting realistic scenario. The study is likely to indicate the effectiveness of the drug/therapy under various conditions.
Sources of RWE

• Patient claims
  • Hospital claims – episode level
  • Provider claims – procedure level
  • Prescription claims

• Patient / Medical registries
  • Focused on target populations

• Electronic Health Record (EHR) / Electronic Medical Record (EMR)
  • Patient-level record from single provider practice (EMR) or multiple practices (EHR)

• Social data
  • Patient interaction on diseases, treatment experiences and side effects
  • Facebook, Twitter, Medhelp, PatientsLikeMe …
Real World studies

• Follow up patients in a normal environment, rather than using artificial clinical trial procedures
• Examine use of the product in patients with
  • Co-morbidities
  • Concomitant medication
• Can be much cheaper with computerised databases
  • Though the data owner may charge you heavily!
• Good external validity
• So why less highly regarded?
Example - Nurses Health Study 1976

- After menopause, rate of heart disease in women rises
- Is it worth giving oestrogen and progesterone after menopause to lower rate of heart disease?
- 127,000 nurses aged between 30 and 55 entered study and completed questionnaires every two years
  - 90% of questionnaires returned
- Grodstein et al (NEJM 1996) showed oestrogen users had 40% fewer heart attacks than those taking no hormones
- Later paper showed similar result in women taking oestrogen and progesterone
- Millions of prescriptions followed
Women’s Health Initiative 1991

• Randomised 16,000 women aged 50-79 to
  • oestrogen and progesterone
  • Or placebo

• Study was stopped 3 years early because of higher rates of
  • Breast cancer, Heart disease, stroke, pulmonary embolism
  • 7 / 8 / 8 / 8 more events per 10,000 women
  • In the active group.

• The well-designed NHS observational study had missed these effects: why?
The women who took hormones in NHS

• Were less likely to
  • Have family history of heart disease
  • Have hypertension or diabetes
  • Smoke

• Were more likely to
  • Take aspirin, birth control pills, vitamins
  • Be Younger
  • Drink more alcohol
  • Eat more saturated fat

• And they were wealthier
• These are called CONFOUNDING VARIABLES
Confounding variables

• Maybe one of the differences in confounding variables masked the effect, eg aspirin or …

• Any one explanation is plausible but their effects are all mixed up together or confounded.
Real-world glycemic, blood pressure, and weight control in patients with type 2 diabetes mellitus treated with canagliflozin—an electronic health-record-based study

Patrick Lefebvre, Dominic Pilon, Marie-Noëlle Robitaille, Marie-Hélène Lafeuille, Wing Chow, Michael Pfeifer & Mei Sheng Duh
Quality questions

• Focussed issue?
  • Defined goals for HbA1c, blood pressure and weight control
  • Simple summary analyses: no claims for anything dramatic

• Recruitment of cohort acceptable = free of bias?
  • IMS Real-World Data
    • Includes primary care / specialists, Medicare / Medicaid / commercial / cash patients
    • March 2012 – September 2014 (product approved by FDA in March 2013)
    • 18 years plus (so includes older patients)
    • Diagnosis for T2DM
    • CANA prescription
    • 12 months plus clinical records before index date
Recognised limitations in paper

• No link to secondary care data
• Prescribing drug does not guarantee it was taken
  • Correct in real world study?
• Ignores use of drug samples to initiate patients
• Missing information may affect the results
• Changes in anti-hypertensive / anti-hyperglycaemic / weight loss medication not tracked
• Duration of CANA treatment under-estimated as it may continue
• Adverse events not reported
Follow-up of subjects – Table 2

- Are patients with two years complete IMS data representative?
  - 20,725 patients with T2DM and CANA
  - 16,163 (78%) had 12 month pre-baseline data and were 18+ years
  - 11,984 had HbA1c assessment pre-index with 10,478 over 7%
  - 6,912 had HbA1c assessment within 45 days of index date
  - 3,612 had HbA1c assessment at 3 month timepoint
    - 17% of original cohort
  - 619 at 12 months - 3% of original cohort
  - Are those with more frequent measurements the more problematic patients?
COMPROMISES BETWEEN RWE AND RCTs: Pragmatic trials

• In cases where genuine uncertainty exists about best treatment
  • Allow medical practitioners to randomise patients to treatments
  • Then follow up as normal
  • Randomisation ensures that treatment groups are comparable

• Or randomise practices to treatment
  • Then treat all patients within that practice the same

• Following up Overall Survival in an oncology trial where patients have crossed over to other treatments uses a technique related to Propensity Score matching
PROBE TRIALS

- Prospective
- Randomised
- Open
  - For patient and investigator
  - Operational study team may be blinded to avoid bias
- Blinded endpoint
  - Assessed by blinded individual not connected to study
- Example of FAST study
BMJ Open  Protocol of the Febuxostat versus Allopurinol Streamlined Trial (FAST): a large prospective, randomised, open, blinded endpoint study comparing the cardiovascular safety of allopurinol and febuxostat in the management of symptomatic hyperuricaemia

Thomas M MacDonald,1 Ian Ford,2 George Nuki,3 Isla S Mackenzie,1 Raffaele De Caterina,4 Evelyn Findlay,1 Jesper Hallas,5 Christopher J Hawkey,6 Stuart Ralston,3 Matthew Walters,7 John Webster,8 John McMurray,7 Fernando Perez Ruiz,9 Claudine G Jennings1
FAST study

- Post Approval Safety Study (PASS) to address CV safety concerns when febuxostat was approved by EMA (2008)

- Recruited patients are
  - Aged over 60
  - Prescribed allopurinol for symptomatic hyperuricemia
  - Have at least one additional CV risk factor

- Patients randomised to optimal dose allopurinol or febuxostat (80-120mg)

- Followed up for at least 2 years through electronic records *
  - average of 3 years
  - Expected 2 year recruitment period
    - * EMA insisted on annual face-to-face assessments
FAST study

• Primary endpoint is occurrence of APTC event
  • Non-fatal MI
  • Non-fatal stroke
  • CV death

• Secondary endpoints
  • All-cause mortality
  • Hospitalisation for specified CV events

• Objective to demonstrate that febuxostat is non-inferior to allopurinol
  • Hazard ratio < 1.30
What to do?

• ‘I keep saying the sexy job in the next ten years will be statisticians … those skills – of being able to access, understand, and communicate the insights you get from data analysis – are going to be extremely important.’ (Google Chief Economist Hal Varian, 2009)

• ‘What sets statisticians apart? … our understanding of and ability to quantify uncertainty … guarding against false discovery, bias, and confounding.’ (Lisa Lavange, FDA Director of Statistics 2014)
Key statistical questions

• Is there scope for bias in
  • Data collection?
  • Data presentation?
  • Selective presentation of results?
• Have all data been included?
  • Does the fact that data is missing tell you something about the treatment?
• Are claims made backed up by
  • Full data disclosure?
  • Estimates of false discovery rates?
RCT or RWE?

- Both bring great value
- RCT assesses efficacy
  - Can the drug work?
- RWE assesses effectiveness
  - Does the drug work in practice?
- If RWE contradicts results of RCT would you change your drug label?
QUESTIONS?