Efficacy and Effectiveness models

Graziano Onder
Centro Medicina dell’Invecchiamento
Università Cattolica del Sacro Cuore
Rome - Italy

EMA Workshop: Ensuring safe and effective medicines for an ageing population
Definition

**Efficacy** is the capacity to produce an effect. In medicine, it is the ability of an intervention or drug to produce a desired effect in expert hands and *under ideal circumstances*. **Effectiveness** is the capability of producing a desired result. In medicine, effectiveness relates to *how well a treatment works in practice*, as opposed to efficacy, which measures how well it works in RCT or laboratory studies.
Researchers have largely shied away from the complexity of multiple chronic conditions — avoidance that results in expensive, potentially harmful care of unclear benefit.

Tinetti M. NEJM2011
Efficacy and Effectiveness research

**Effectiveness research** addresses practical questions about an intervention as it would occur in routine clinical practice, preserving the ‘ecology’ of care: hypothesis and study design are formulated based on information needed to make a decision.

**Efficacy research** is aimed to better understand how and why an intervention works.
Efficacy and Effectiveness research

3 key features differentiates effectiveness *(pragmatic or practical trials)* and efficacy research *(explanatory trials)*:

1. Population (sample)
Population

Efficacy research
Population with single disease, no complexity

- Generalizability

Effectiveness research
Population that consumes the most health care (comorbidity, behavioral and physical conditions, different settings)

+ Generalizability
- Heterogeneity
Heterogeneity

Heterogeneity resulting from:
- patients’ initial level of risk for a given outcome;
- responsiveness to treatment;
- vulnerability to adverse effect

Treatments compared within homogeneous risk strata

Tinetti M. NEJM2011
## Population

<table>
<thead>
<tr>
<th>Efficacy research</th>
<th>Effectiveness research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population with single disease, no complexity</td>
<td>Population that consumes the most health care (comorbidity, behavioral and physical conditions, different settings)</td>
</tr>
<tr>
<td>+ Retention/adherence</td>
<td>+ Generalizability</td>
</tr>
<tr>
<td>- Generalizability</td>
<td>- Heterogeneity</td>
</tr>
<tr>
<td></td>
<td>Retention/adherence</td>
</tr>
</tbody>
</table>
Leukotriene Antagonists as First-Line or Add-on Asthma-Controller Therapy

**Design:** Two parallel, pragmatic trials to evaluate effectiveness of LTRA

**Study 1:** LTRA vs inhaled glucocorticoid for first-line asthma-controller therapy

**Study 2:** LTRA vs a long-acting beta2-agonist as add-on therapy in patients already receiving inhaled glucocorticoid therapy.
Leukotriene Antagonists as First-Line or Add-on Asthma-Controller Therapy

**Study 1**
- **LTRA**
  - Retention: 92%
  - Adherence: 65%
- Glucocorticoid
  - Adherence: 41%

**Study 2**
- **LTRA**
  - Retention: 97%
  - Adherence: 74%
- **Beta2-agonist**
  - Adherence: 46%

*Price D NEJM 2011*
Poor adherence
Poor retention

→ Dilution of the effect

→ Need of large sample size

Data analysis: ‘... an intention to treat analysis will provide a valid comparison of treatment strategies.’
Poor adherence → Poor retention → Dilution of the effect → Need of large sample size

Data analysis: ‘... in equivalence trials it can create a bias toward a finding of equivalence’
Poor adherence
Poor retention

Dilution of the effect

Need of large sample size

Data analysis: ‘... a pragmatic equivalence trial with a substantial rate on nonadherence may not demonstrate equivalence robustly.’

Ware JH NEJM 2011
Rivaroxaban versus Warfarin in Nonvalvular Atrial Fibrillation

**Design:** Pragmatic clinical trial (ROCKET AF)

**Sample:** 14,264 patients with nonvalvular atrial fibrillation

**Study groups:** rivaroxaban vs. dose-adjusted warfarin
Adherence – Rocket AF

**Inclusion criteria:** history of stroke, transient ischemic attack, or systemic embolism, heart failure or a left ventricular ejection fraction of 35% or less, hypertension, an age of 75 years or more, or the presence of diabetes mellitus

Mean **CHADS score 3.5**

Warfarin dosing evaluated by **time in therapeutic range (TTR) = 55%**

*Patel MR NEJM 2011*
...findings were not adequate to determine whether rivaroxaban was as effective compared with warfarin when the existing treatment is used skillfully... The FDA said the median TTR for warfarin in general use is about 65%, but in ROCKET AF, the TTR was only a “relatively poor” 55%
Efficacy and Effectiveness research

3 key features differentiates effectiveness (pragmatic or practical trials) and efficacy research (explanatory trials):

1. Population (sample)
2. Interventions
Intervention

Efficacy research
- Placebo comparison
- Blinded

Effectiveness research
- Head to head comparisons
- Pharmacological and non-pharmacological interventions
- Unblinded
Interventions in effectiveness research

1. Examination of treatments for common pairs of diseases in which treatment of one may exacerbate or improve the other;
Treatment of pain and behavioural symptoms in NH residents with dementia

Husebo B  BMJ 2011
Interventions in effectiveness research

1. Examination of treatments for common pairs of diseases in which treatment of one may exacerbate or improve the other;
2. Testing interventions that can affect simultaneously multiple conditions;
Exercise and dietary weight loss in obese older adults with knee osteoarthritis: the ADAPT study

<table>
<thead>
<tr>
<th>Study group</th>
<th>Baseline</th>
<th>6 months</th>
<th>18 months</th>
<th>Change from baseline at 18 months (95% confidence interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy lifestyle</td>
<td>434.61 ± 10.96</td>
<td>428.56 ± 12.88</td>
<td>429.89 ± 12.77</td>
<td>-4.72 (-29.75, 20.31)</td>
</tr>
<tr>
<td>Diet only</td>
<td>425.98 ± 10.89</td>
<td>433.68 ± 11.94</td>
<td>435.63 ± 12.88</td>
<td>9.65 (-15.79, 35.09)</td>
</tr>
<tr>
<td>Exercise only</td>
<td>424.15 ± 11.42</td>
<td>465.04 ± 12.13</td>
<td>472.73 ± 13.12†</td>
<td>48.58 (22.87, 74.29)</td>
</tr>
<tr>
<td>Diet plus exercise</td>
<td>416.15 ± 11.34</td>
<td>482.37 ± 12.65</td>
<td>477.76 ± 13.12†</td>
<td>61.61 (35.90, 87.32)</td>
</tr>
</tbody>
</table>

Messier SP Arthritis Rheum 2004
Interventions in effectiveness research

1. Examination of treatments for common pairs of diseases in which treatment of one may exacerbate or improve the other;
2. Testing interventions that can affect simultaneously multiple conditions;
3. Combination of pharmacological and non-pharmacological treatments;
ROT combined with cholinesterase inhibitors in Alzheimer's disease

<table>
<thead>
<tr>
<th>Patients</th>
<th>Treatment group</th>
<th>Control group</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE</td>
<td>0.2 (0.4)</td>
<td>-1.1 (0.4)</td>
<td>0.02</td>
</tr>
<tr>
<td>ADAS–Cog</td>
<td>0.4 (0.8)</td>
<td>-2.5 (0.8)</td>
<td>0.01</td>
</tr>
<tr>
<td>Neuropsychiatric Inventory</td>
<td>0.9 (1.9)</td>
<td>-2.5 (2.1)</td>
<td>0.23</td>
</tr>
<tr>
<td>Barthel Index</td>
<td>-0.9 (1.0)</td>
<td>-2.9 (1.0)</td>
<td>0.18</td>
</tr>
<tr>
<td>Number of impaired IADL</td>
<td>0.0 (0.2)</td>
<td>-0.2 (0.2)</td>
<td>0.34</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Caregivers</th>
<th>Treatment group</th>
<th>Control group</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hamilton Rating Scale for Depression</td>
<td>-0.9 (0.4)</td>
<td>-1.0 (0.4)</td>
<td>0.83</td>
</tr>
<tr>
<td>Hamilton Anxiety Scale</td>
<td>-0.3 (0.4)</td>
<td>-0.5 (0.4)</td>
<td>0.80</td>
</tr>
<tr>
<td>Caregiver Burden Inventory</td>
<td>-2.0 (1.4)</td>
<td>-1.3 (1.5)</td>
<td>0.72</td>
</tr>
<tr>
<td>SF–36</td>
<td>-1.3 (1.4)</td>
<td>-1.1 (1.4)</td>
<td>0.90</td>
</tr>
</tbody>
</table>
Interventions in effectiveness research

1. Examination of treatments for common pairs of diseases in which treatment of one may exacerbate or improve the other;

2. Testing interventions that can affect simultaneously multiple conditions;

3. Combination of pharmacological and non-pharmacological treatments;

4. Comparison of models of care
A RCT of Inpatient and Outpatient Geriatric Evaluation and Management
**Intervention**

<table>
<thead>
<tr>
<th>Efficacy research</th>
<th>Effectiveness research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placebo comparison</td>
<td>Head to head comparisons</td>
</tr>
<tr>
<td>Blinded</td>
<td>Pharmacological and non-pharmacological interventions</td>
</tr>
</tbody>
</table>

- Not informative  
+ Informative for users  
- Blindness
Blindness and outcomes

... the combination of unblinded treatment and patient self-assessment undermines an important element of efficacy trials, creating a potential for bias: patients' expectations may influence their outcomes report ... Effectiveness trials are stronger when they include both objective (e.g., survival, test results) and subjective outcome measures (e.g., quality-of-life surveys).
Efficacy and Effectiveness research

3 key features differentiates effectiveness (pragmatic or practical trials) and efficacy research (explanatory trials):

1. Population (sample)
2. Interventions
3. Outcomes
Outcomes

Efficacy research

*Disease oriented*
(occurrence of a single disease or exacerbation of a single chronic condition)

*Rating scales/test measures*

Effectiveness research

*Universal health outcomes* (symptoms burden, function, health related quality of life, active life expectancy)

*Real-world measure of clinical practice*
Antipsychotics - Outcomes

Efficacy and adverse effects of atypical antipsychotics for dementia: meta-analysis of randomized, placebo-controlled trials.

15 trials met selection criteria ... a total of 3,353 patients were randomized to drug and 1,757 to placebo.

Results: **Efficacy on rating scales** was observed by meta-analysis for aripiprazole and risperidone, but not for olanzapine.

Schenider LS Am J Geriatr Psychiatry 2006
Antipsychotics – CATIE-AD

The **primary end point is an accurate reflection of a clinical event**: the decision to change treatment because the patient's condition is worsening or not improving sufficiently ... The CATIE-AD study is an exemplar of the clinical trial's revolutionary role in shaping therapeutics.
# Outcomes

<table>
<thead>
<tr>
<th>Efficacy research</th>
<th>Effectiveness research</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Disease oriented</strong>&lt;br&gt;(occurrence of a single disease or exacerbation of a single chronic condition)&lt;br&gt;<em>Rating scales/test measures</em>&lt;br&gt;- People at risk for multiple adverse outcomes</td>
<td><strong>Universal health outcomes</strong>&lt;br&gt;(symptoms burden, function, health related quality of life, active life expectancy)&lt;br&gt;<em>Real-world measure of clinical practice</em>&lt;br&gt;+ Informative&lt;br&gt;- Harder to collect</td>
</tr>
</tbody>
</table>
**SHEP - Chlortalidone versus placebo**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>RR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroke</td>
<td>0.67</td>
<td>0.56-0.80</td>
</tr>
<tr>
<td>CHF</td>
<td>0.46</td>
<td>0.33-0.65</td>
</tr>
<tr>
<td>CHD</td>
<td>0.75</td>
<td>0.60-0.94</td>
</tr>
<tr>
<td>Any CVD</td>
<td>0.68</td>
<td>0.58-0.79</td>
</tr>
</tbody>
</table>

SHEP JAMA 1991
Deterioration of ADLs in SHEP

Bar chart showing the percentage of Basic ADLs and Moderate ADLs for Placebo and Active treatment groups. The p-values are 0.20 and 0.30 respectively.

Placebo: Basic ADLs: p=0.20, Moderate ADLs: p=0.30
Active treatment: Basic ADLs: p=0.20, Moderate ADLs: p=0.30

Archieves of Internal Medicine, Applegate W Arch Intern Med 1994
Missing disability assessments in SHEP

% with missing data

Year

1 2 3 4

Placebo

Active treatment

* p<.001
** p=.04

Di Bari Am J Epidemiol 2000
SHEP sensitivity analyses - RR of ADL disability for active treatment vs placebo

Reported

Sensitivity analysis: % disability among missing data

Di Bari Am J Epidemiol 2000
Pragmatic trials are designed to study real-world practice and therefore represent less-perfect experiments than efficacy trials; *they sacrifice internal validity to achieve generalizability*. The challenge is to *keep the balance right* so that the findings are likely to be both correct and applicable to clinical practice or health care delivery.