Get Real

Real-Life Data in Drug Development

TOOLBOX
methods to assess effectiveness before launch

Chris Chinn & Lucien Abenhaim,

on the behalf of WP2

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement no [115546], resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.
www.imi.europa.eu

What problems WP2 tackled?

- The main objectives of WP2 were
  - to better **understand** when/why there is a gap between efficacy and effectiveness
  - and **provide tools** to address it during the drug development process

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement no [115546], resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.
www.imi.europa.eu

What problems WP2 tackled?

- The general questions were
  - Is there an expected "gap" between trial outcomes ("efficacy") and results achieved in actual clinical practice ("effectiveness")?
    - How to **anticipate** it?
    - What are the differences between stu…hat **drive uncertainty** in assessing effectiveness?
  - How/which studies can inform the assessment of effectiveness before launch?
    - Can pre-registration phase 3 trials for …populations, comparators and endpoi…f RE?
    - How to conduct randomized PCTs tackling specific statistic…ssues?

Drivers of effectiveness

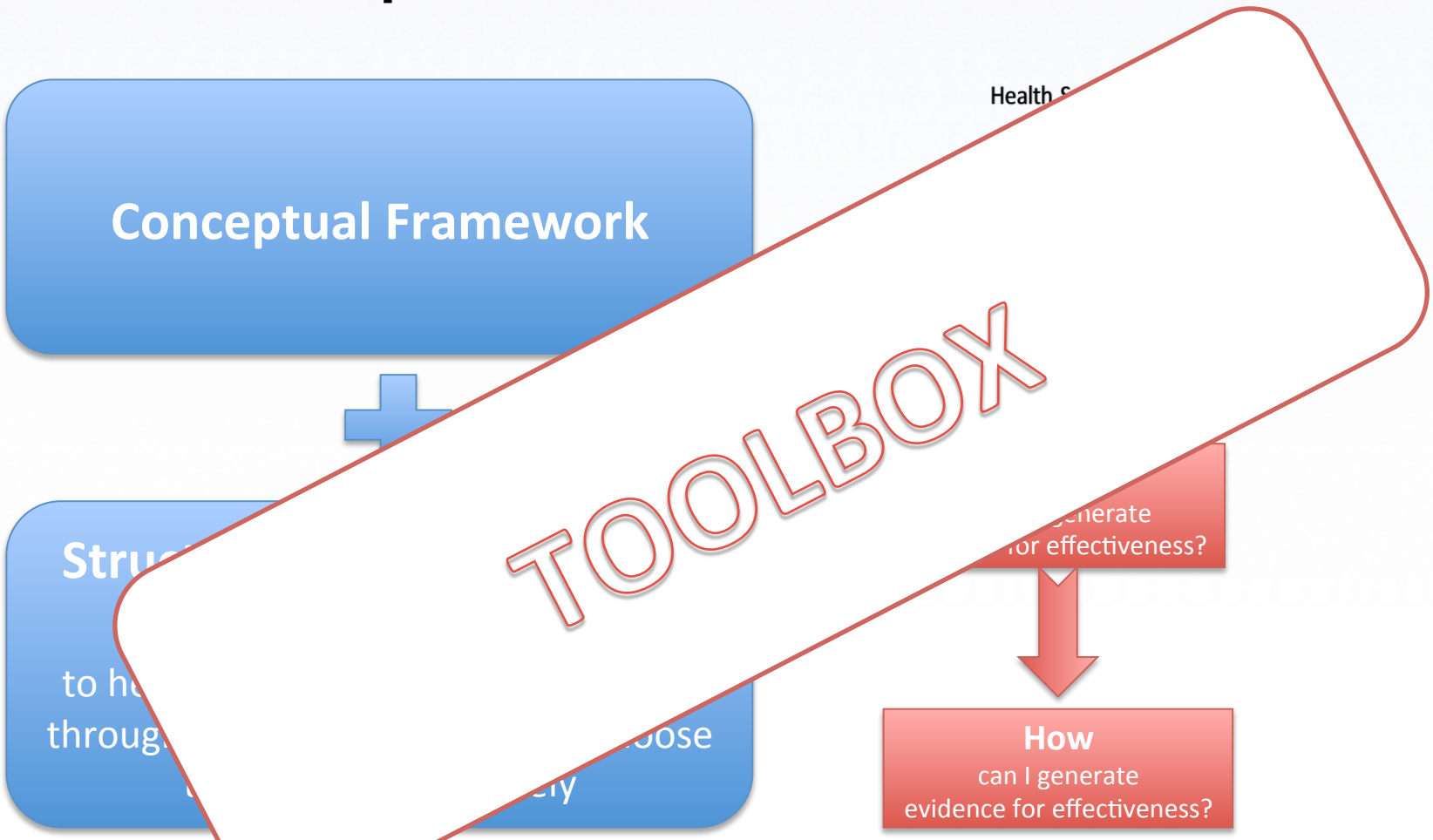
Tools for trials

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement no [115546], resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.
www.imi.europa.eu

Get Real

Real-Life Data in Drug Development

How problems were tackled and solved?



The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement no [115546], resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.
www.imi.europa.eu

Get Real

Real-Life Data in Drug Development

How problems were tackled and solved?
Step 1

Why should I generate evidence for effectiveness?

"What **impact** has the absence of evidence on effectiveness, on **HTA decisions**?"

QUESTION
Is there a compelling need to generate evidence of effectiveness, over and above the evidence plan for registration?

Review of **68 pre and post-authorization dossiers**, submitted to the French HTA body (**HAS**) between 2004 and 2011: the lack of evidence on effectiveness was found to raise specific concerns, when uncertainty rested on:

- The **actual compliance with the "Terms of Use"** and prescription requirements (e.g., identification and description of the population and prescribers, duration and dosage of treatment, adherence)
- The **impact on "morbidity/ mortality"**

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement no [115546], resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.
www.imi.europa.eu

Get Real

Real-Life Data in Drug Development

How problems were tackled and solved?
Step 2

TASKS

- Identify a potential efficacy/ effectiveness gap and the **drivers of effectiveness**
- Assess the **consequences** of not addressing drivers of effectiveness in the development plan

METHODS TO EXPLORE ISSUES

- Review of HTA decisions
- Literature review
- Experts' interviews
- Analysis of available data
- Collect Patient insights
- Seek early scientific advice

"How literature reviews, experts interviews or data analyses can be used to identify **drivers of effectiveness** before launch?"

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement no [115546], resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.
www.imi.europa.eu

Get Real

Real-Life Data in Drug Development

How problems were tackled and solved?
Step 2

TASKS

- Identify a potential efficacy/ effectiveness gap and the **drivers of effectiveness**
- Assess the **consequences** of not addressing drivers of effectiveness in the development plan

METHODS TO EXPLORE ISSUES

- Review of HTA decisions
- Literature review
- Experts' interviews
- Analysis of available data
- Collect Patient insights
- Seek early scientific advice

TOOLS TO IDENTIFY DRIVERS OF EFFECTIVENESS

Structured Literature Reviews can retrieve studies
(1)which explored a "gap" and provide explanations for this gap;
(2)which explored effect-modification on the association between exposure to drug and outcome;
(3)which explored the efficacy of drugs (RCTs) vs. the effectiveness of drugs (observational studies)

Interviews of experts with an extensive clinical experience in the therapeutic field of interest may be useful to
(1)generate hypothesis on potential drivers of effectiveness
(2)or after a literature review, to identify DoE not retrieved by the review and/or weigh the results of literature review with a clinical perspective

Data Analyses may focus on the exploration of
(1)a modification of drug's effect by potential DoE (related to patient's characteristics, actual use of the drug or characteristics of the healthcare system), using simple statistics (sub-group comparison)
(2)statistical interaction between the drug and potential DoE, in regression models
(3)a gap between drugs' effect estimates in RCTs and drugs' effect estimates in observational studies; a comparison of (pooled) results across study types may approximate an efficacy-effectiveness gap and the comparison of patients characteristics may help identifying DoE

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement no [115546], resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.
www.imi.europa.eu

Get Real

Real-Life Data in Drug Development

How problems were tackled and solved? Step 3

Should I generate evidence for effectiveness?

IF YES

How can I generate evidence for effectiveness?

"Are there **pre-authorization RCTs** performed so far, which investigated **effectiveness**?"

What are the **options**?

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement no [115546], resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.
www.imi.europa.eu

Get Real

Real-Life Data in Drug Development

How problems were tackled and solved?
Step 4

TASK
Systematically explore each **PICO** component for (more) pragmatic trial design solutions
- **Patient**
- **Intervention, Comparator,**
- **Outcome**

"Can I **enrich** my trial population to better **predict** effectiveness, without increasing the sample size and compromising my chances of success?"

TASK
Identify technically feasible options for an integrated "effectiveness evidence plan" (trials and analyses; timing, setting)?

"Which are the **statistical issues** raised by making a randomized trial **more pragmatic**?"

"How to measure **effectiveness just after launch**, using observational studies?"

The research leading to th Medicines Initiative Joint Undertaking under grant agreement no [115546], res ution from the European Union's Seventh Framework Programme (FP7/2007-2013) a
www.imi.europa.eu

Get Real

Real-Life Data in Drug Development

How problems were tackled and solved?
Step 4

TASK
Systematically explore each **PICO** component for (more) pragmatic trial design solutions
- **Patient**
- **Intervention, Comparator,**
- **Outcome**

TASK
Identify technically feasible options for an integrated "effectiveness evidence plan" (trials and analyses; timing, setting)?

TOOLS TO ENRICH RCT POPULATION

Step 1: identify the exclusion criteria which might impact the drug's effect estimate in the RCT if applied strictly; consider the possibility to include a subset of these patients ("enrichment subset")
Step 2: determine the good balance in terms of sample size of the "enrichment subset", using modelling techniques: the sample size should be large enough to allow prediction of effectiveness but not too large, to maximize the chances of success of the RCT
Step 3: the RCT is performed exactly as usual; the whole RCT population is used to calculate the primary endpoint
Step 4: The prediction of the drug's effectiveness (when will be prescribed in a real population) will be able to utilize the information based on the "enrichment subset"; usual predictive modelling techniques may be used (Bayesian, regression models, etc.)

efpia imi innovative medicines initiative

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement no [115546], resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.
www.imi.europa.eu

Get Real

Real-Life Data in Drug Development

How problems were tackled and solved?
Step 4

TASK
Systematically explore each **PICO** component for (more) pragmatic trial design solutions
- **Patient**
- **Intervention, Comparator,**
- **Outcome**

TASK
Identify technically feasible options for an integrated "effectiveness evidence plan" (trials and analyses; timing, setting)?

STATISTICAL TOOLS FOR RANDOMIZED PRAGMATIC TRIALS

Consent refusal happens **post-randomization** in TwiCS trials → risk of imbalanced selection bias

•Both an **intention to treat (ITT) and instrumental variable (IV) analysis** should be carried out.

Comparator arm using "standard of care" → managing heterogeneity

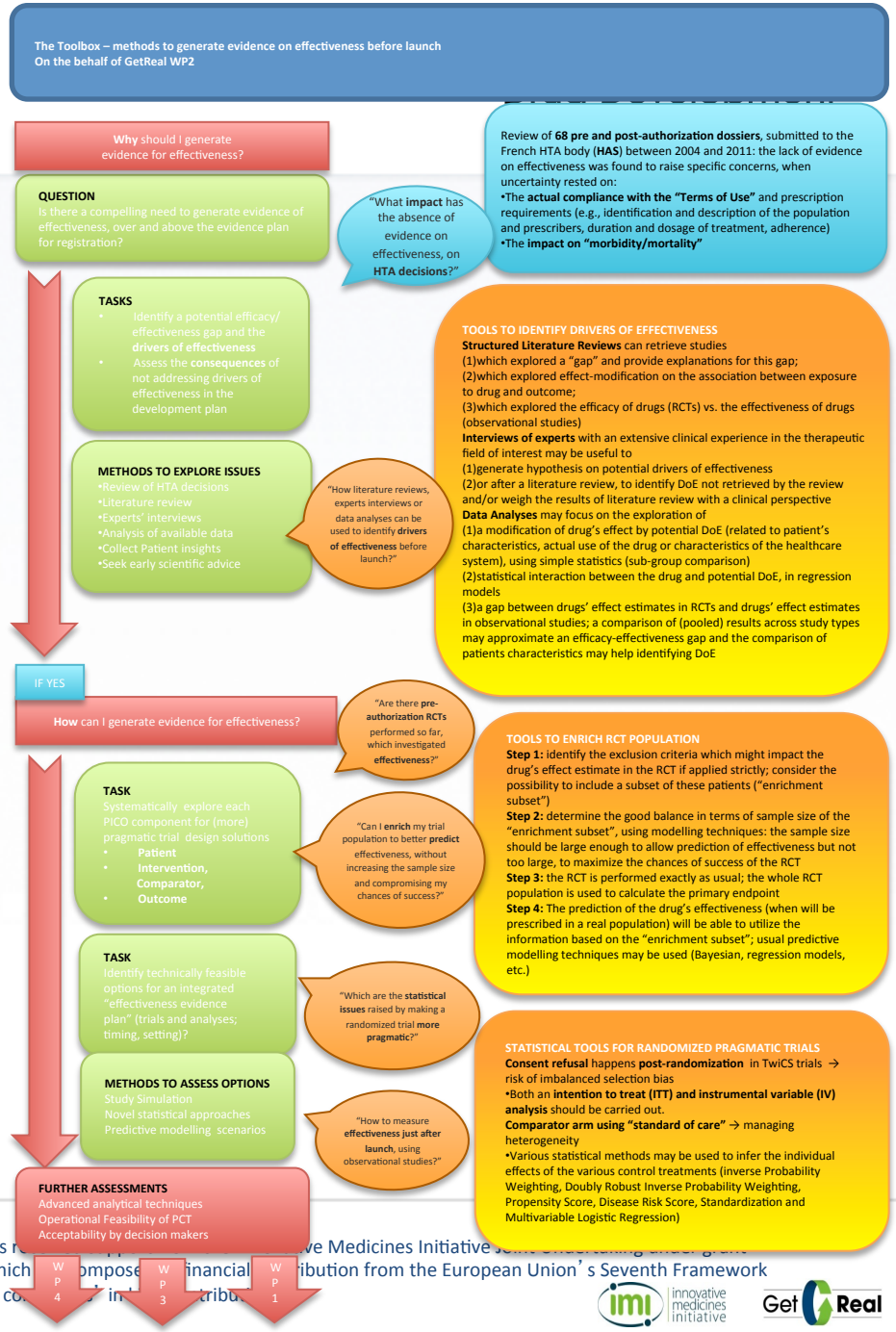
•Various statistical methods may be used to infer the individual effects of the various control treatments (inverse Probability Weighting, Doubly Robust Inverse Probability Weighting, Propensity Score, Disease Risk Score, Standardization and Multivariable Logistic Regression)

efpia imi innovative medicines initiative

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement no [115546], resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.
www.imi.europa.eu

Get Real

Check our poster !



The research leading to these results has r… …ve Medicines Initiative … agreement no [115546], resources of which … …ompose … financial … …ibution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA co… … in… …tribu…
www.imi.europa.eu

innovative medicines initiative

Get Real

Get Real

Real-Life Data in Drug Development

How is this TOOLBOX of use for public?

- For HTA and regulatory bodies
  - Understand what drives the gap between efficacy and effectiveness
  - Be aware of the options available and their specificities (statistics behind!)
  - Be able to assess with background knowledge
- For R&D
  - Utilize the decision-making approach and anticipate!
  - Understand what drives the gap between efficacy and effectiveness
  - Have scientifically sound options while understanding their specificities (randomized (pragmatic) trials)

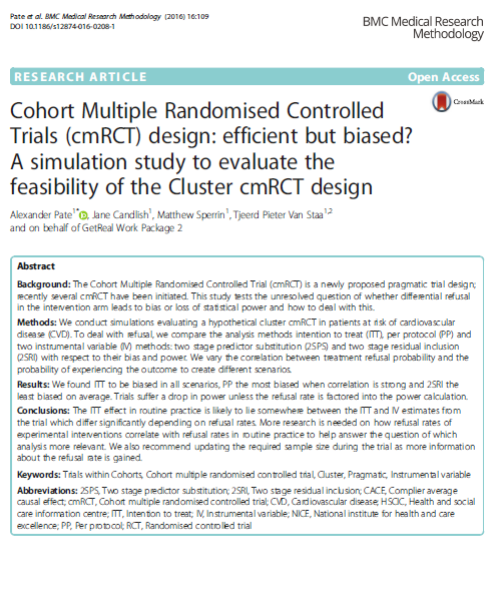
The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement no [115546], resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.
www.imi.europa.eu

Get Real

Real-Life Data in Drug Development

How is this TOOLBOX of use for public?

- The TOOLBOX will be available
  - In the Navigator
  - As a PDF stand-alone document (final development)
  - Through publications (3 published, 11 submitte


Pate et al. BMC Medical Research Methodology (2016) 16:109
DOI 10.1186/s12874-016-0208-1

BMC Medical Research Methodology

RESEARCH ARTICLE Open Access

Cohort Multiple Randomised Controlled Trials (cmRCT) design: efficient but biased? A simulation study to evaluate the feasibility of the Cluster cmRCT design

Alexander Pate[1*], Jane Candlish[1], Matthew Sperrin[1], Tjeerd Pieter Van Staa[1,2] and on behalf of GetReal Work Package 2

Abstract

Background: The Cohort Multiple Randomised Controlled Trial (cmRCT) is a newly proposed pragmatic trial design; recently several cmRCT have been initiated. This study tests the unresolved question of whether differential refusal in the intervention arm leads to bias or loss of statistical power and how to deal with this.

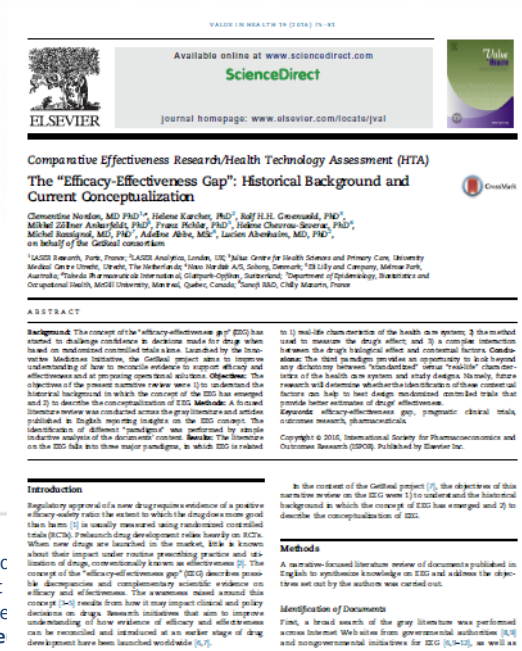
Methods: We conduct simulations evaluating a hypothetical cluster cmRCT in patients at risk of cardiovascular disease (CVD). To deal with refusal, we compare the analysis methods intention to treat (ITT), per protocol (PP) and two instrumental variable (IV) methods: two stage predictor substitution (2SPS) and two stage residual inclusion (2SRI) with respect to their bias and power. We vary the correlation between treatment refusal probability and the probability of experiencing the outcome to create different scenarios.

Results: We found ITT to be biased in all scenarios, PP the most biased when correlation is strong and 2SRI the least biased on average. Trials suffer a drop in power unless the refusal rate is factored into the power calculation.

Conclusions: The ITT effect in routine practice is likely to lie somewhere between the ITT and IV estimates from the trial which differ significantly depending on refusal rates. More research is needed on how refusal rates of experimental interventions correlate with refusal rates in routine practice to help answer the question of which analysis more relevant. We also recommend updating the required sample size during the trial as more information about the refusal rate is gained.

Keywords: Trials within Cohorts, Cohort multiple randomised controlled trial, Cluster, Pragmatic, Instrumental variable

Abbreviations: 2SPS, Two stage predictor substitution; 2SRI, Two stage residual inclusion; CACE, Complier average causal effect; cmRCT, Cohort multiple randomised controlled trial; CVD, Cardiovascular disease; HSCIC, Health and social care information centre; ITT, Intention to treat; IV, Instrumental variable; NICE, National institute for health and care excellence; PP, Per protocol; RCT, Randomised controlled trial


Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

ELSEVIER

Comparative Effectiveness Research/Health Technology Assessment (HTA)

The "Efficacy-Effectiveness Gap": Historical Background and Current Conceptualization

Clementine Nordon, MD PhD[1,*], Helene Karcher, PhD[2], Rolf H.H. Groenwold, PhD[3], Mikkel Zöllner Ankarfeldt, PhD[4], Franz Pichler, PhD[5], Helene Chevrou-Severac, PhD[6], Michel Roseignol, MD, PhD[7], Adeline Abbe, MSc[8], Lucien Abenhaim, MD, PhD[9], on behalf of the GetReal consortium

[1]LASER Research, Paris, France; [2]LASER Analytics, London, UK; [3]Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht, The Netherlands; [4]Novo Nordisk A/S, Soborg, Denmark; [5]Eli Lilly and Company, Melrose Park, Australia; [6]Takeda Pharmaceuticals International, Glattpark-Opfikon, Switzerland; [7]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada; [8]Sanofi R&D, Chilly Mazarin, France

ABSTRACT

Background: The concept of the "efficacy-effectiveness gap" (EEG) has started to challenge confidence in decisions made for drugs when based on randomized controlled trials alone. Launched by the Innovative Medicines Initiative, the GetReal project aims to improve understanding of how to reconcile evidence to support efficacy and effectiveness and at proposing operational solutions. **Objectives:** The objectives of the present narrative review were 1) to understand the historical background in which the concept of the EEG has emerged and 2) to describe the conceptualization of EEG. **Methods:** A focused literature review was conducted across the gray literature and articles published in English reporting insights on the EEG concept. The identification of different "paradigms" was performed by simple inductive analysis of the documents' content. **Results:** The literature on the EEG falls into three major paradigms, in which EEG is related to 1) real-life characterization of the health care system; 2) the method used to measure the drug's effect; and 3) a complex interaction between the drug's biological effect and contextual factors. **Conclusions:** The third paradigm provides an opportunity to look beyond any dichotomy between "standardized" versus "real-life" characteristics of the health care system and study designs. Namely, future research will determine whether the identification of these contextual factors can help to best design randomized controlled trials that provide better estimates of drugs' effectiveness.

Keywords: efficacy-effectiveness gap, pragmatic clinical trials, outcomes research, pharmaceuticals

Copyright © 2016, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

Introduction

Regulatory approval of a new drug requires evidence of a positive efficacy-safety ratio: the extent to which the drug does more good than harm [1] is usually measured using randomized controlled trials (RCTs). Prelaunch drug development relies heavily on RCTs. When new drugs are launched in the market, little is known about their impact under routine prescribing practice and utilization of drugs, conventionally known as effectiveness [2]. The concept of the "efficacy-effectiveness gap" (EEG) describes possible discrepancies and complementary scientific evidence on efficacy and effectiveness. The awareness raised around this concept [3–5] results from how it may impact clinical and policy decisions on drugs. Research initiatives that aim to improve understanding of how evidence of efficacy and effectiveness can be reconciled and introduced at an earlier stage of drug development have been launched worldwide [6,7].

In the context of the GetReal project [7], the objectives of this narrative review on the EEG were 1) to understand the historical background in which the concept of EEG has emerged and 2) to describe the conceptualization of EEG.

Methods

A narrative-focused literature review of documents published in English to synthesize knowledge on EEG and address the objectives set out by the authors was carried out.

Identification of Documents

First, a broad search of the gray literature was performed across Internet Web sites from governmental authorities [8,9] and nongovernmental initiatives for EEG [6,9–12], as well as


Adjusting for Confounding in Early Postlaunch Settings

Going Beyond Logistic Regression Models

Amand F. Schmidt,[a,b,c,d] Olaf H. Klungel,[a,b] and Rolf H. H. Groenwold[a,b]; on behalf of the GetReal Consortium

Background: Postlaunch data on medical treatments can be analyzed to explore adverse events or relative effectiveness in real-life settings. These analyses are often complicated by the number of potential confounders and the possibility of model misspecification.

Methods: We conducted a simulation study to compare the performance of logistic regression, propensity score, disease risk score, and stabilized inverse probability weighting methods to adjust for confounding. Model misspecification was induced in the independent derivation dataset. We evaluated performance using relative bias confidence interval coverage of the true effect, among other metrics.

Results: At low events per coefficient (1.0 and 0.5), the logistic regression estimates had a large relative bias (greater than −100%). Bias of the disease risk score estimates was at most 13.48% and 18.83%. For the propensity score model, this was 8.74% and >100%, respectively. At events per coefficient of 1.0 and 0.5, inverse probability weighting frequently failed or reduced to a crude regression, resulting in biases of −8.49% and 24.55%. Coverage of logistic regression estimates became less than the nominal level at events per coefficient ≤5. For the disease risk score, inverse probability weighting, and propensity score, coverage became less than nominal at events per coefficient ≤2.5, ≤1.0, and ≤1.0, respectively. Bias of misspecified disease risk score models was 16.55%.

Conclusion: In settings with low events/exposed subjects per coefficient, disease risk score methods can be useful alternatives to logistic regression models, especially when propensity score models cannot be used. Despite better performance of disease risk score methods than logistic regression and propensity score models in small events per coefficient settings, bias, and coverage still deviated from nominal.

(*Epidemiology* 2016;27: 133–142)

Nonrandomized studies on (pharmacologic) therapeutics are often conducted to complement results from randomized clinical trials (RCTs). For example, nonrandomized studies might be more appropriate to assess the occurrence of rare, but severe, adverse events.[1–3] Furthermore, nonrandomized studies can be used to estimate the relative effectiveness in real-life clinical practice. Depending on the relationship between the intervention and the outcome, different degrees of confounding can be expected.[1–3] For example, it might be expected that patients who responded poorly to older drugs will cross over to the new drug.[4] Alternatively, as shown by Mack et al.,[5] physicians might be hesitant to prescribe a novel drug to patients with comorbidities. Furthermore, depending (among other factors) on the speed of uptake, differences in patient populations pre- and post-launch or difference between early and late adopters may increase the potential for effect modification, further obstructing comparison of a new drug to older compounds.[6,7]

Frequently, the outcome of interest is dichotomous, such as mortality, in which case multivariable logistic regression[8] is commonly used to adjust for confounding. One (of many) assumption(s) is that associations between confounders and the outcome are sufficiently estimated to adjust for confounding bias. In settings (e.g., nonrandomized early postlaunch studies) where both the number of events and the number of exposed subjects are small, controlling for confounding can be problematic. Further complicating the matter is that it is not uncommon to consider more than 100 potential confounders.[9] Simulation studies showed that for prognostic logistic regression models, 10 or more events per coefficient were needed to

Submitted 28 November 2014; accepted 25 August 2015.

From the [a]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands; [b]Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Utrecht, The Netherlands; [c]Department of Farm Animal Health, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands; and [d]Institute of Cardiovascular Science, Faculty of Population Health, University College London, London, United Kingdom.

Supported by the Innovative Medicines Initiative Joint Undertaking under Grant Agreement No. 115546, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.

A.F.S., R.H.H.G., and O.H.K. contributed to the idea and design of the study. A.F.S. performed the analyses and drafted the manuscript. O.H.K. and R.H.H.G. provided guidance during initial planning of the paper and during critical revision. A.F.S. had full access to all of the data and takes responsibility for the integrity of the data presented.

The research leading to these results was conducted as part of the GetReal consortium. For further information, please refer to www.imi-getreal.eu.

The authors report no conflicts of interest.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

Correspondence: Amand F. Schmidt, Institute of Cardiovascular Science, Faculty of Population Health, University College London, London, United Kingdom. E-mail: amand.schmidt@ucl.ac.uk.

ISSN: 1044-3983/16/2701-0133
DOI: 10.1097/EDE.0000000000000388