

CLINICAL/NARRATIVE REVIEW

A Primer on Effectiveness and Efficacy Trials

Amit G. Singal, MD, MS^{1,2}, Peter D.R. Higgins, MD, PhD³ and Akbar K. Waljee, MD, MS^{3,4}

Although efficacy and effectiveness studies are both important when evaluating interventions, they serve distinct purposes and have different study designs. Unfortunately, the distinction between these two types of trials is often poorly understood. In this primer, we highlight several differences between these two types of trials including study design, patient populations, intervention design, data analysis, and result reporting.

Clinical and Translational Gastroenterology (2014) 5, e45; doi:10.1038/ctg.2013.13; published online 2 January 2014

Subject Category: Clinical Review

INTRODUCTION

Intervention studies can be placed on a continuum, with a progression from efficacy trials to effectiveness trials. Efficacy can be defined as the performance of an intervention under ideal and controlled circumstances, whereas effectiveness refers to its performance under 'real-world' conditions.¹ However, the distinction between the two types of trial is a continuum rather than a dichotomy, as it is likely impossible to perform a pure efficacy study or pure effectiveness study.²

There are several steps that must occur for an efficacious intervention to be effective in clinical practice; therefore, an efficacy trial can often overestimate an intervention's effect when implemented in clinical practice. An efficacious intervention must be readily available, providers must identify the target population and recommend the intervention, and patients must accept and adhere to the intervention.³ For example, several studies highlight how underutilization of colorectal cancer and hepatocellular carcinoma screening contribute to poor effectiveness in clinical practice.^{4–8} In fact, poor access, recommendation, acceptance, and adherence rates can lead to highly efficacious interventions being less effective in practice than less-efficacious interventions. For example, ultrasound has a sensitivity of 63% for detecting hepatocellular carcinoma at an early stage in prospective efficacy studies and is regarded as being more efficacious than alpha fetoprotein. However, in a recent effectiveness study, ultrasound only had a sensitivity of 32%, comparable to that of alpha fetoprotein (sensitivity 46%).^{9,10} This gap was related to the low utilization rates of ultrasound and its operator-dependent nature. Similarly, hepatitis C and hepatocellular carcinoma therapy can also be highly efficacious in reducing morbidity and mortality but are limited by low rates of access, recommendation, and acceptance.^{11–14}

Although efficacy research maximizes the likelihood of observing an intervention effect if one exists, effectiveness research accounts for external patient-, provider-, and

system-level factors that may moderate an intervention's effect. Therefore, effectiveness research can be more relevant for health-care decisions by both providers in practice and policy-makers.¹⁵ The distinction between these two types of trials is important but often poorly understood. In fact, an analysis of product evaluations for Health Technology Assessments found that efficacy data is often assumed to be effectiveness data.¹⁶ The aim of this primer is to highlight differences between these two types of trials (Table 1) and how these differences affect study design.

STUDY DESIGN

Efficacy studies investigate the benefits and harms of an intervention under highly controlled conditions. Although this has multiple methodologic advantages and creates high internal validity, it requires substantial deviations from clinical practice, including restrictions on the patient sample, control of the provider skill set and limitations on provider actions, and elimination of multimodal treatments.² A placebo-controlled randomized controlled trial (RCT) design is ideal for efficacy evaluation because it minimizes bias through multiple mechanisms, such as standardization of the intervention and double blinding. RCTs generally eliminate issues of access (intervention is provided free), provider recommendation, and patient acceptance and adherence.

Effectiveness studies (also known as pragmatic studies) examine interventions under circumstances that more closely approach real-world practice, with more heterogeneous patient populations, less-standardized treatment protocols, and delivery in routine clinical settings. Effectiveness studies may also use a RCT design; however, the intervention is more often compared with usual care, rather than placebo. Minimal restrictions are placed on the provider actions in modifying dose, the dosing regimen, or co-therapy, allowing tailored therapy for each subject. Although effectiveness studies

¹Department of Internal Medicine, UT Southwestern Medical Center, Dallas, Texas, USA; ²Department of Clinical Sciences, University of Texas Southwestern, Dallas, Texas, USA; ³Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan, USA and ⁴Veterans Affairs Center for Clinical Management Research, Ann Arbor, Michigan, USA

Correspondence: Amit G. Singal, MD, MS, Dedman Scholar of Clinical Care, Division of Digestive and Liver Diseases, University of Texas Southwestern Medical Center, 5959 Harry Hines Blvd, POB 1, Suite 420, Dallas, Texas 75390-8887, USA. E-mail: amit.singal@utsouthwestern.edu

Received 19 June 2013; accepted 19 August 2013

sacrifice some internal validity, they have higher external validity than efficacy studies.² Effectiveness trials without a witnessed effect may be related to one of several factors including an ineffective intervention, poor implementation, lack of provider acceptance, or lack of patient acceptance and adherence.

PATIENT POPULATION

Efficacy trials use strict inclusion and exclusion criteria to enroll a defined, homogenous patient population. Inclusion criteria confirm that patients truly have the disease of interest, whereas exclusion criteria exclude those who are unlikely to respond to the intervention. For example, efficacy studies may exclude patients who are at low risk for the primary outcome, those who are deemed likely to be non-compliant, or those with significant comorbid medical conditions. However, these strict inclusion and exclusion criteria can limit the generalizability of the results to patients seen in clinical practice. Effectiveness trials typically have limited exclusion criteria and involve a more heterogeneous population, including higher rates of non-compliant patients and more subjects with significant comorbid conditions.¹⁷ However, effectiveness trials can still exclude patients for safety concerns, as these patients would not be expected to get the intervention in usual practice.¹⁸ For example, a recent RCT demonstrated that rectal indomethacin significantly reduced the risk of post-endoscopic retrograde cholangiopancreatography (ERCP) pancreatitis; however, only high-risk patients, such as those with sphincter of Oddi dysfunction, were included.¹⁹ Effectiveness studies would help clarify if these results can be generalized to low-risk and medium-risk patients undergoing ERCP in everyday practice.

THE INTERVENTION

In efficacy trials, interventions are delivered in a highly standardized way, including timing and dosage of medications and perhaps even the associated patient education. The use of concurrent medications or interventions is often restricted, so any witnessed effect can be attributed to the intervention of interest. Furthermore, efficacy trials are conducted with top-quality equipment and highly experienced providers, who are often provided training in the intervention and measurement of outcomes prior to the study. Finally, intensive resources are often dedicated to maximize provider uptake and patient

compliance with the intervention.¹⁷ Research assistants can provide intense counseling, education, and even reminders for scheduled medications or clinic appointments. This intensive attention can in part explain the high placebo effect seen in some trials, such as those in irritable bowel syndrome and inflammatory bowel disease.^{20,21}

Effectiveness trials standardize the availability of the intervention in the study sample but do not go to extremes to reinforce implementation by providers or participation by patients. There are no requirements regarding provider expertise, and equipment quality may be variable. Similarly, providers are not restricted in terms of offering concurrent therapies or crossing over patients on-and-off therapy, which can lead to higher rates of drug–drug interactions and make it less clear if any effect was truly related to the intervention of interest. Finally, additional study resources, such as reminder phone calls or study coordinators, are not available to augment provider and/or patient compliance.

ANALYSIS

Both efficacy and effectiveness trials typically use an intention-to-treat approach for statistical analysis. However, given that efficacy trials aim to address if interventions work under ideal circumstances, secondary analyses using a per-protocol approach may be informative. Alternative techniques that have been proposed to account for differences between efficacy and effectiveness include contaminated adjustment intention to treat and voting with their feet analyses.^{22–24} Effectiveness trials often have higher rates of missing data than efficacy trials.^{17,18} There are several methods for handling missing data, with details beyond the scope of this primer.^{25–27}

REPORTING DATA

The applicability of results from both efficacy and effectiveness studies depend on the context of the trial and the situation to which the data are being applied. It is crucial for any study to provide sufficient data regarding the trial's setting, participants, and intervention. A trial with an insufficient description regarding the intervention is effectively rendered useless, as external implementation and validation is impossible. Guidelines for reporting results of efficacy and effectiveness studies should be followed to standardize reporting of results.^{28,29}

Table 1 Differences between efficacy and effectiveness studies

	Efficacy study	Effectiveness study
Question	Does the intervention work under ideal circumstance?	Does the intervention work in real-world practice?
Setting	Resource-intensive 'ideal setting'	Real-world everyday clinical setting
Study population	Highly selected, homogenous population Several exclusion criteria	Heterogeneous population Few to no exclusion criteria
Providers	Highly experienced and trained	Representative usual providers
Intervention	Strictly enforced and standardized No concurrent interventions	Applied with flexibility Concurrent interventions and cross-over permitted

COMPARATIVE EFFECTIVENESS RESEARCH

Clinicians have historically been frustrated by the lack of consideration of external validity in RCTs, other efficacy studies, and guidelines.³⁰ Accordingly, there has been a call for studies whose results can be more readily applied to everyday clinical practice.¹⁵ This culminated in the American Recovery and Reinvestment Act, which allotted more than \$1 billion to support comparative effectiveness research (CER). The Institute of Medicine has defined CER as “the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition, or to improve the delivery of care.”³¹ The purpose of CER is to assist patients, providers, and policy-makers in making informed decisions that can improve health care both at the individual and population levels. As suggested by the name, CER places an emphasis on effectiveness studies, conducted in settings similar to real-world clinical practice, to maximize external validity of any results. With increased funding support for effectiveness research, the number of effectiveness studies will likely increase over the next several years.

CONCLUSION

An understanding of the distinction between efficacy and effectiveness research is not only crucial when conducting research but also interpreting results from studies and deciding how applicable it may be to clinical practice and patients who may have less access and less adherence to medications. Given a growing focus on evidence-based medicine and pay-for-performance measures, providers must base clinical decisions on the best available evidence. However, defining the best available evidence may not always be clear. Although some prioritize efficacy data from RCTs, others view effectiveness data as more pertinent to real-world clinical practice decisions.² There are at least two tools, which can help clinicians judge where a trial may lie on the efficacy–effectiveness continuum.^{18,32} Gartlehner and colleagues identified criteria to distinguish efficacy and effectiveness studies, with a sensitivity and specificity of 72% and 83%, respectively. Similarly, PRECIS is a tool with 10 domains (for example, sample exclusion criteria, intervention flexibility, and follow-up intensity) that can help categorize studies as efficacy or effectiveness trials. Although both types of studies are important when evaluating interventions, they serve different purposes and provide different data.

CONFLICT OF INTEREST

Guarantors of the article: Amit G. Singal, MD, MS and Akbar K. Waljee, MD, MS.

Specific author contributions: Study concept and design, drafting of the manuscript, critical revision of the manuscript for important intellectual content, and study supervision: Amit G. Singal and Akbar K. Waljee; critical revision of the manuscript for important intellectual content: Peter D.R. Higgins.

Financial support: None.

Potential competing interests: None.

Study Highlights

WHAT IS CURRENT KNOWLEDGE

- Although efficacy and effectiveness studies are both important when evaluating interventions, they serve distinct purposes and provide different data.

WHAT IS NEW HERE

- Efficacy research maximizes the likelihood of observing an intervention effect, whereas effectiveness research better accounts for external patient-, provider-, and system-level factors that may moderate an intervention's effect in clinical practice.
- Tools exist to help clinicians judge where a trial may lie on the efficacy–effectiveness continuum.

1. Revicki DA, Frank L. Pharmacoeconomic evaluation in the real world. Effectiveness versus efficacy studies. *Pharmacoeconomics* 1999; **15**: 423–434.
2. Fritz JM, Cleland J. Effectiveness versus efficacy: more than a debate over language. *J Orthop Sports Phys Ther* 2003; **33**: 163–165.
3. El-Serag HB, Talwalkar J, Kim WR. Efficacy, effectiveness, and comparative effectiveness in liver disease. *Hepatology* 2010; **52**: 403–407.
4. Klabunde CN, Brown M, Ballard-Barbash R *et al.* Cancer Screening—United States, 2010; 27 January 2012 2012.
5. Singal AG, Yopp A, Skinner C *et al.* Utilization of hepatocellular carcinoma surveillance among American patients: a systematic review. *J Gen Intern Med* 2012; **27**: 861–867.
6. Singal AG, Yopp AC, Gupta S *et al.* Failure rates in the hepatocellular carcinoma surveillance process. *Cancer Prev Res (Phila)* 2012; **5**: 1124–1130.
7. Steinwachs D, Allen JD, Barlow WE *et al.* National Institutes of Health state-of-the-science conference statement: enhancing use and quality of colorectal cancer screening. *Ann Intern Med* 2010; **152**: 663–667.
8. Singal AG, Nehra M, Adams-Huet B *et al.* Detection of hepatocellular carcinoma at advanced stages among patients in the HALT-C trial: where did surveillance fail? *Am J Gastroenterol* 2013; **108**: 425–432.
9. Singal A, Volk ML, Waljee A *et al.* Meta-analysis: surveillance with ultrasound for early-stage hepatocellular carcinoma in patients with cirrhosis. *Aliment Pharmacol Ther* 2009; **30**: 37–47.
10. Singal AG, Conjeevaram HS, Volk ML *et al.* Effectiveness of hepatocellular carcinoma surveillance in patients with cirrhosis. *Cancer Epidemiol Biomarkers Prev* 2012; **21**: 793–799.
11. Singal AG, Volk ML, Jensen D, Di Bisceglie A, Schoenfeld PS. A sustained viral response is associated with reduced liver-related morbidity and mortality in patients with hepatitis C virus. *Clinical Gastroenterology and Hepatology* 2010; **8**: 280–288.
12. Volk ML, Tocco R, Saini S, Lok AS. Public health impact of antiviral therapy for hepatitis C in the United States. *Hepatology* 2010; **50**: 1750–1755.
13. Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. *Hepatology* 2011; **53**: 1020–1022.
14. Tan D, Yopp AC, Beg MS, Gopal P, Singal AG. Meta-analysis: underutilization and disparities of treatment among patients with hepatocellular carcinoma in the United States. *Alimentary Pharmacology & Therapeutics* 2013; **38**: 703–712.
15. Treweek S, Zwarenstein M. Making trials matter: pragmatic and explanatory trials and the problem of applicability. *Trials* 2009; **10**: 37.
16. Abraham J. Use of ‘efficacy’ and ‘effectiveness’ often misleading and may skew reimbursement decisions. *Context Matters* advance online publication, 6 November 2012; <http://globenewswire.com/news-release/2012/11/06/502781/10011322/en/Use-of-Efficacy-and-Effectiveness-Often-Misleading-and-May-Skew-Reimbursement-Decisions.html>.
17. Streiner D, Norman G. Efficacy and effectiveness trials. *Community Oncology* 2009; **6**: 472–474.
18. Thorpe KE, Zwarenstein M, Oxman AD *et al.* A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol* 2009; **62**: 464–475.
19. Elmunzer BJ, Scheiman JM, Lehman GA *et al.* A randomized trial of rectal indomethacin to prevent post-ERCP pancreatitis. *N Engl J Med* 2012; **366**: 1414–1422.
20. Akehurst R, Kaltenthaler E. Treatment of irritable bowel syndrome: a review of randomised controlled trials. *Gut* 2001; **48**: 272–282.
21. Su C, Lewis JD, Goldberg B *et al.* A meta-analysis of the placebo rates of remission and response in clinical trials of active ulcerative colitis. *Gastroenterology* 2007; **132**: 516–526.
22. Elmunzer BJ, Hayward RA, Schoenfeld PS *et al.* Effect of flexible sigmoidoscopy-based screening on incidence and mortality of colorectal cancer: a systematic review and meta-analysis of randomized controlled trials. *PLoS Med* 2012; **9**: e1001352.

23. Rangwala SC, Waljee AK, Higgins PD. Voting With their Feet (VWF) endpoint: a meta-analysis of an alternative endpoint in clinical trials, using 5-ASA induction studies in ulcerative colitis. *Inflamm Bowel Dis* 2009; **15**: 422–428.
24. Sussman JB, Hayward RA. An IV for the RCT: using instrumental variables to adjust for treatment contamination in randomised controlled trials. *BMJ* 2010; **340**: c2073.
25. Ibrahim JG, Chu H, Chen MH. Missing data in clinical studies: issues and methods. *J Clin Oncol* 2012; **30**: 3297–3303.
26. Kaambwa B, Bryan S, Billingham L. Do the methods used to analyse missing data really matter? An examination of data from an observational study of Intermediate Care patients. *BMC Res Notes* 2012; **5**: 330.
27. Waljee AK, Mukherjee A, Singal AG *et al.* Comparison of modern imputation methods for missing laboratory data in medicine. *BMJ Open* 2013; **3**: e002847.
28. Flay BR, Biglan A, Boruch RF *et al.* Standards of evidence: criteria for efficacy, effectiveness and dissemination. *Prev Sci* 2005; **6**: 151–175.
29. Zwarenstein M, Treweek S, Gagnier JJ *et al.* Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008; **337**: a2390.
30. Rothwell PM. External validity of randomised controlled trials: 'to whom do the results of this trial apply?' *Lancet* 2005; **365**: 82–93.
31. Sox HC, Greenfield S. Comparative effectiveness research: a report from the Institute of Medicine. *Ann Intern Med* 2009; **151**: 203–205.
32. Gartlehner G, Hansen RA, Nissman D *et al.* A simple and valid tool distinguished efficacy from effectiveness studies. *J Clin Epidemiol* 2006; **59**: 1040–1048.



Clinical and Translational Gastroenterology is an open-access journal published by **Nature Publishing Group**. This work is licensed under a **Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License**. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>