PRA Health Sciences

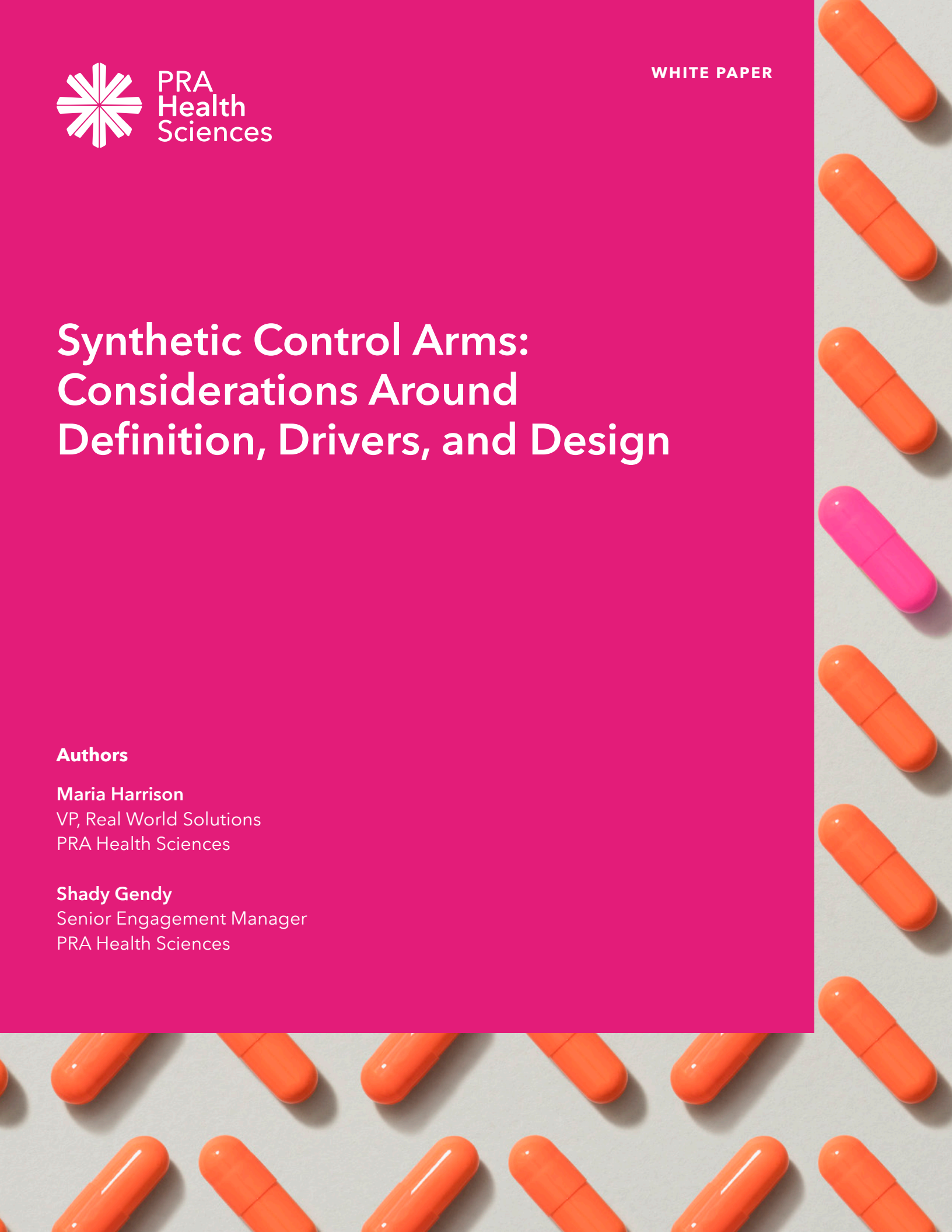
WHITE PAPER

Synthetic Control Arms: Considerations Around Definition, Drivers, and Design

Authors

Maria Harrison
VP, Real World Solutions
PRA Health Sciences

Shady Gendy
Senior Engagement Manager
PRA Health Sciences



Executive Summary

Those that work in the biopharmaceutical or device industry are likely familiar with the term *synthetic controls*. But what exactly are synthetic controls, and why (where, when, how) would one use them? Although there's still some debate around their value in clinical research, high-quality synthetic controls can increase study efficiency, lower costs, and get life-changing treatments to patients faster.

Definitions

Before we define a synthetic control, let's talk about controls in general. A control arm is a group of individuals in a clinical study who do not receive the treatment of interest in order to determine the efficacy (or effectiveness) of the treatment in question. Controls are important because they allow the researcher to minimize differences in all other variables except the one being tested.

There are many types of traditional controls, including placebo, no treatment, active treatment, and dose-comparison controls. An active treatment control is usually the gold standard treatment. In all of the aforementioned, the control arm is both internal and concurrent. This means the control arm is part of the trial and is assessed or followed at the same time as the treatment arm. Traditional controls include historical controls, which are different from placebos and the like. Historical controls are outside of the trial (we say 'external') and were evaluated at an earlier time (that is, they are 'non-concurrent').

With the rise of real-world data (RWD) and real-world study design, we have added a few new control types to our lexicon. This is where synthetic controls come in. Synthetic controls are sourced from electronic health records, administrative claims data, patient-generated data (such as from a fitness tracker or Apple Watch), disease registries, and even historical clinical trial or study data. Another term that we're hearing lately is *in silico*. *In silico* refers to computer-modeled or simulated controls. These are control patients who are created from 'big data' sources. Controls can be real patients – based on real, patient-level data – or they can be simulated via a computer model.

Although the terms synthetic, historical, and *in silico* controls are sometimes used interchangeably, there are important differences between the three. The broadest term you can use is Synthetic Control: it encompasses both historical and *in silico* *in silico* controls. What do they all have in common? They're all external to any treatment trial or study. How are they distinct? It depends on whether they're concurrent or non-concurrent with the treatment group, and if they're based on real or simulated patients.

Term	Origin	Temporality	Real or Simulated Patients
Synthetic	Multiple Sources	Concurrent or Non-concurrent	Real or Simulated
Historical	Previous Trial or Study	Non-concurrent	Real
In silico	Computer-modeled	Concurrent or Non-concurrent	Simulated

Table 1: Synthetic Control

In this discussion, we'll focus on synthetic controls that are sourced from real-world settings and real patients.

Design

Data Source

How does one stand up a synthetic control arm (SCA)? First, one selects the appropriate data source for an SCA. This includes assessing the quality and the scope of the data collected, along with the availability and the reliability of study endpoints. As mentioned above, RWD can come from many sources such as electronic health records, administrative claims, historical clinical trials, etc. Every source has its pros and cons that must be considered on a study-by-study basis. Next, one must ensure that the control arm is similar to the treatment arm. In a randomized controlled trial, this similarity is accomplished by randomization. In the absence of randomization, one must match controls based on characteristics such as age, gender, concomitant illness, severity of disease, etc. – all the things that form the inclusion and exclusion criteria in a clinical study.

One reason historical controls are so inviting is that if they are sourced from previous clinical trials, they are sure to come with a great deal of coded and validated clinical data that don't come with a patient sourced from a clinic or hospital EMR. Historical controls have their own limitations, which include changes over time in medical care and the subsequent introduction of what we call *temporal bias*.

Besides matching, another important consideration in the design of a synthetic control arm is the availability of the endpoint of interest. Let's say you're conducting an asthma study and your primary endpoint involves exacerbations. The synthetic control arm must be comprised of asthmatics from another study or data source in which exacerbation information was (or can be) captured.

Minimizing Bias

Epidemiologists have found an efficient way to minimize bias by matching patients based on a summary score of all of the patient characteristics that might influence treatment outcomes. This summary score, called a *propensity score*, ranges from a value of 0 to 1, and is calculated via a type of logistic regression. Treated patients are matched to control patients based on their propensity score using an approach termed *nearest neighbor matching*.

For example, if I'm a treated patient and my propensity score is .22, an ideal match for me would be a control with a score of .22 – but she may not exist in my control pool. In that case, I might be matched to a control whose score is .25. One limitation of propensity score matching is that you have to throw out (set aside) those controls whose propensity scores are highly divergent from those of the treated patients. Because of this, we generally have to assemble a potential pool of controls that is twice the size of the pool needed for analysis. This entire process enables virtual recruitment of eligible patients for a control arm. The control arm feeds into the study database where it can be analyzed just as it would in a conventional trial.

How many variables feed into the propensity score? It depends, but we generally include as much information as we can to produce the best match and to reduce bias as much as possible. In the absence of randomization, bias is never completely eliminated; epidemiologists refer to this leftover bias as *residual bias*.

Drivers

SCAs are helpful when a placebo isn't ethical or practical. Perhaps there is no existing effective treatment for a condition, or you have a breakthrough therapy for a rare disease that's fast-tracked for approval. SCAs work best in therapeutic areas for which information is easy to mine from real-world data sources, particularly when the standard of care is stable, and a large treatment effect is expected. Why does a large treatment effect help? If you have concerns about the possible dampening effects of bias due to residual or historical bias, a large treatment effect can partially overcome that dampening effect. Oncology studies are well-suited to SCAs because the data tend to be highly standardized and plentiful.

As stated in the introduction, synthetic controls benefit the sponsor because they can increase study efficiency, reduce study delays, and lower study costs. SCAs benefit the patient because they eliminate the fear of assignment to placebo. SCAs benefit society because they accelerate lifesaving therapies to market!

Are there real-life, concrete examples of SCA use? Yes! Here are a few:

Roche's Alecensa[1] (alectinib) received accelerated FDA approval in Dec 2015 as a treatment for a specific form of lung cancer. In Feb 2017, it was conditionally approved in the EU. The EU required additional evidence of effectiveness relative to the standard of care (ceritinib). Rather than waiting for Phase 3 results, Roche used an SCA of 67 patients (Flatiron data) to provide the necessary evidence. This advanced coverage of *Alecensa* by 18 months in 20 European countries.

The PD-L1 inhibitor Bavencio[2] (avelumab) from Pfizer and Merck KGaA for Merkel cell carcinoma employed an SCA because short patient survival times precluded the recruitment of a prospective control group. The control arm used data from EMRs in community and academic centers. This led to approval of Bavencio in 2018.

Blincyto, from Amgen, received an additional approval for Philadelphia chromosome-negative relapsed and refractory B-cell precursor acute lymphoblastic leukemia based on a single-intervention group trial. The results were compared with historical data from 694 comparable patients extracted from 2,000 patient records in the US and EU.

Let's look at a more theoretical example, set in oncology, where single arm trials abound because precision oncology drugs increasingly receive accelerated regulatory approval based on uncontrolled trials.

Carrigan et al.[3] looked rather recently at the performance of electronic health records (EHR) to derive control arms for early phase singe-arm lung cancer trials, specifically in non-small cell lung cancer. Their results were impressive. Patients from a US oncology EHR database were aligned with patients from randomized control trials (RCTs), and trial-specific eligibility criteria were applied to the EHR dataset. Overall survival (OS) in the EHR-derived control arm (the Y axis) was compared with OS in the RCT experimental arm (the X axis).

The primary outcome was OS, defined as time from randomization or treatment initiation (in the EHR) to death. EHR-derived hazard ratio (HR) estimates aligned closely with those from the corresponding RCT with one exception (NCT01519804). Note how the values cluster around the diagonal line of perfect agreement. Comparing log HRs among all RCT and EHR results gave a Pearson correlation coefficient of 0.86 (perfect agreement being 1.0). The authors concluded that "properly selected control arms from contemporaneous EHR data could be used to put single-arm trials of OS in advanced non-small cell lung cancer into context."

An inherent feature of single arm designs is that a standard-of-care control arm is not included, which can lead to challenges in interpretation of efficacy. It should come as no surprise that the FDA has stated that external concurrent controls – as in synthetic controls – are better than no control at all. The FDA has cautioned, however, "that we should not…say that [external controls are] going to be the new standard and [forget] about randomized, controlled studies."[4]

As more use cases demonstrate the value of synthetic controls in bringing life-saving therapies to market faster, we will no doubt gain a larger degree of comfort with a real-world control arm. The FDA, along with some other government agencies, is funding a project called RCT Duplicate[5]. The RCT Duplicate research team is within the Division of Pharmacoepidemiology and Pharmacoeconomics at the Brigham and Women's Hospital in Boston. Their advisory board includes experts in the design and analysis methods for nonrandomized studies as well as randomized trials. They are funded to demonstrate, via 30 different projects, that real world data—when properly curated and analyzed—can replicate the results of an RCT.

[1] https://www.statnews.com/2019/02/05/synthetic-control-arms-clinical-trials/

[2] https://www.statnews.com/2019/03/06/whats-scary-and-appealing-about-real-world-evidence/

[3] Carrigan G, Whipple S, Capra WB, et al. Using Electronic Health Records to Derive Control Arms for Early Phase Singe-Arm Lung Cancer Trials: POC in Randomized Controlled Trials. Clinical Pharmacology & Therapeutics Vol 107(2) Feb 2020.

[4] External Control Arms: Better Than Single-Arm Studies But No Replacement For Randomization. Pink Sheet, 2 January 2019.

[5] https://www.rctduplicate.org

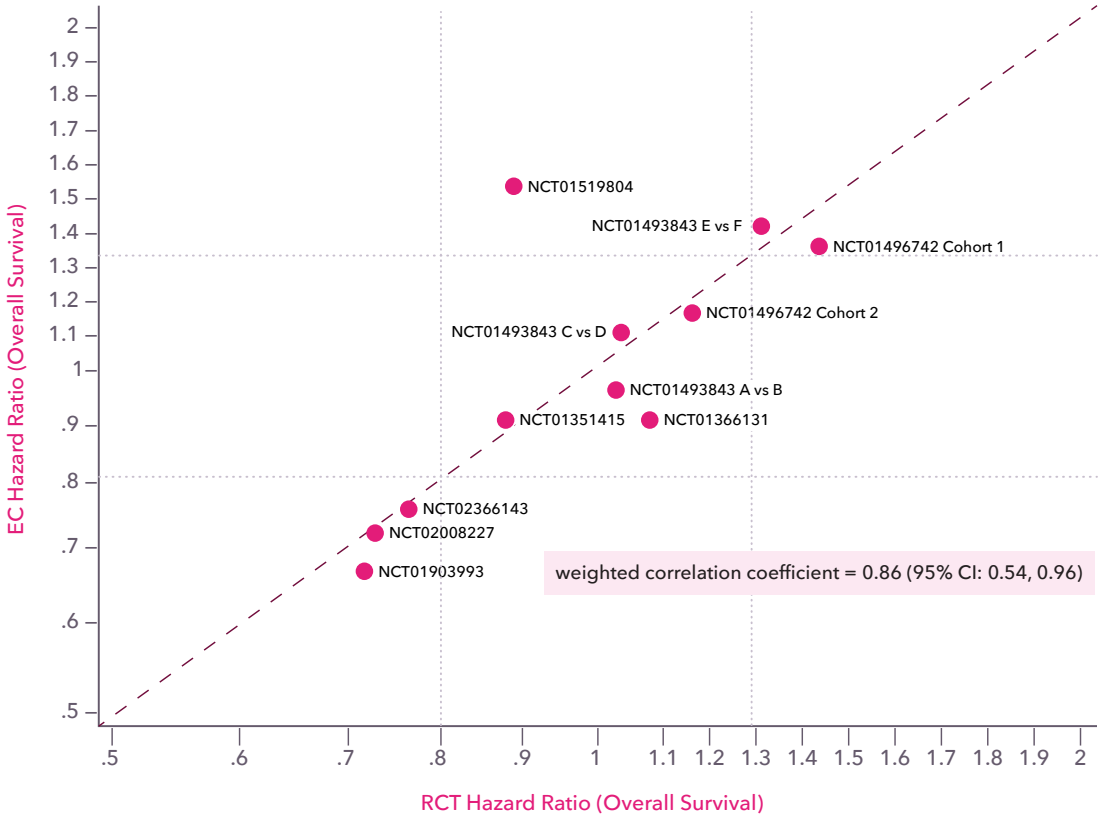


Chart 1: Summary Results

Here are seven examples of real-world data studies that matched the results of a randomized, controlled trial. For each example, the study that was released first is shaded[6].

The RCT Duplicate Team's goal is "to understand for what types of clinical questions real world data analyses can be conducted with confidence, and with which designs and analysis methods." If you unpack that statement, you grasp that they are not suggesting the use of real-world data every time. They wish to identify where, when, and how real-world data perform best. Their website notes, "If principled nonrandomized study approaches based on healthcare databases can consistently match the results of published trials and predict the results of ongoing trials, then we gain confidence in the validity of future real-world data analyses that may be performed in the absence of randomized trial evidence."

These authors agree wholeheartedly with that statement; you might just say that we are a 'perfect match.'

[6] Franklin JM , Glynn RJ, Martin D, Schneeweiss S. Evaluating the Use of Nonrandomized Real-World Data Analyses for Regulatory Decision Making. Clinical Pharmacology & Therapeutics. Vol 105(4) April 2019.

		Real World Data Analysis	Randomized Controlled Trial
Studies of harmful effects (Safety)	1	Schneeweiss S, et al. **Aprotinin during coronary-artery bypass grafting and risk of death,** NEJM 2008; 358; 771-83*	Fergusson DA, et al. **A comparison of aprotinin and lysine analogues in high-risk cardiac surgery**. NEJM 2008, 358-2319-31
	2	Kim, SC, et al, **Cardiovaslular safety of tocillzumab versus tumor necrosis factor inhibitatiors in patients with rheumatoid arthritis** - a multi-database cohort study. Arthritis Rheumatol 2017; 69(6): 1154-64*	Giles JT, et al. **Comparative cardiovascular safety of tocillizumab vs etanercept in rheumatoid arthritis:** Results of a randomized parallel-group, multicenter, non-inferiority, phase 4 clinical trial [abstract]. Arthritis Rheumatol 2016; 68 (suppl 10)
	3	Zhang, MA, et al. **Risk of cardiovascular events inolder patients with gout initiating febuxostat vs allopurinol:** A population-based cohort study. Circulation 2018; 138-1116-26*	White WB, et al. **Cardiovascular safety of febuzostat or allopurinal in patients with gout.** NEJM 2018; 378:1200-10
Studies of beneficial effects (Effectiveness)	4	Patorno E, et al. **Cardiovascular safety of canagilflozin versus other non-gillflozin antidiabetic agents:** A population-based cohort study. BMJ 2018; 360-k119*	Neal B, et al. **Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes.** NEJM 2017; 377:644-57
	5	Gokbudget N, et al. **Bilnatumonmab vs historical standard therapy of adult relapsed/refractory acute lymphoblastic leukemia**. Blood Cancer J 2016; 6:e473	Katarjian H, et al. **Bilnatumonmab versus chemotherapy for advanced acute lymphoblastic leukemia.** NEJM 2017; 376:836-47
	6	Seeger JD, et al. **Safety and effectiveness of dabigatran and warfarin in routine care of patients with artrial fibrillation**. Thromb Haemostasis 2015; 114:1277-89	Connolly SJ, et al. **Dabigatran versus warfarin in patients with artrial fibrillation**. NEJM 2009;361(12):1139-57
	7	Fralick M, et al. Using healthcare claims databases to identify supplemental indications of approved medications. JAMA Internal Medicine 2018; 178:55-63	ONTARGET Investigators, Yusuf S, et al. T**elmisartan, rampril, or both i patients at high risk for vascular events.** NEJM 2008; 358:1547-59

* We provide citation to published manuscript here. Initial results were presented at FDA or scientific conferences before the corresponding RCT findings were known.

Table 2: Real World Data Studies

Contact Information

For further information, or to discuss any aspect of PRA's services offered in this field, please contact your Business Development Manager or the employees listed below:

Maria Harrison
Vice President, Real World Solutions

PRA Health Sciences
731 Arbor Way, Suite 100,
Blue Bell, PA 19422

Phone: +1 (215) 591 1179
Mobile: +1 (484) 919 2855
HarrisonMaria@prahs.com

Shady Gendy
Senior Engagement Manager

PRA Health Sciences
Phone: +1 (202) 227 9842
GendyShady@prahs.com

World Headquarters
4130 ParkLake Avenue, Suite 400
Raleigh, North Carolina 27612 USA

Phone: +1 (919) 786 8200
Fax: +1 (919) 786 8201
www.prahs.com