

## Synthetic Controls In Clinical Trials

#### When is it appropriate to use a SCA?

Rare disease	As disease understanding improves and diseases become sub- stratified they can effectively become "rare" diseases	
Difficult Recruitment	Fear of being enrolled to a control treatment is a major determinant of non-participation in clinical trials <sup>1</sup>	
Treatment Switching	E.g. NCT00428597: sunitinib vs. placebo in pancreatic cancer; median PFS at interim analysis was 11.4 vs. 5.5 months (HR = 0.42); 69% of patients crossed-over after interim analysis <sup>2</sup>	
Ethical Concerns	Breakthroughs can challenge notions of clinical equipoise. E.g. drug targets an established driver, was studied within a population that has the driver, there are no major safety concerns, extremely high remission rates in early phase trials <sup>3</sup>	

<sup>1</sup> Gaddipati H, Liu K, Pariser A, Pazdur R. Clin Cancer Res 2012; 18: 5172-5178. <sup>2</sup> Raymond E, Niccoli P, Castellano D, et al. J Clin Onc 2016 34:4\_suppl, 309-309. <sup>3</sup> Sharma MR, Schilsky RL. Nat Rev Clin Oncol. 2012;9(4):208.

Cytel

## For further details

#### **Clinical Epidemiology**

**Dovepress** open access to scientific and medical research

8 Open Access Full Text Article

REVIEW

# Synthetic and External Controls in Clinical Trials – A Primer for Researchers

This article was published in the following Dove Press journal: *Clinical Epidemiology* 

Kristian Thorlund (1,2) Louis Dron<sup>2</sup> Jay JH Park<sup>2,3</sup> Edward J Mills<sup>1,2</sup>

<sup>1</sup>Department of Health Research Methods, Evidence & Impact (HEI), McMaster University, Hamilton, ON, Canada; <sup>2</sup>MTEK Sciences, Vancouver, BC, Canada; <sup>3</sup>Department of Medicine, University of British Columbia, **Abstract:** There has been a rapid expansion in the use of non-randomized evidence in the regulatory approval of treatments globally. An emerging set of methodologies have been utilized to provide greater insight into external control data used for these purposes, collectively known as synthetic control methods. Through this paper, we provide the reader with a set of key questions to help assess the quality of literature publications utilizing synthetic control methodologies. Common challenges and real-life examples of synthetic controls are provided throughout, alongside a critical appraisal framework with which to assess future publications. **Keywords:** synthetic control, RCTs, real-world evidence



### **Text Rich Checklist**

Table I Synthetic Control Quality Checklist

ltem Number	Key Question	Criteria for Judgement			
External C	ontrol Data Sources				
I	Was the original data collection process similar to that of the clinical tria?	State whether patients are from large well-conducted RCT(s) or high quality prospective cohort studies, and whether patient characteristic similar to the target population			
2	Was the external control population sufficiently similar to the clinical trial population?	State how the external population is similar with regards to key characteristics, such as (but not limited to): age, geographic distribution, performance status, treatment history, sex etc.			
3 Did the outcome definitions of the external control match those of that clinical trial?		State whether the outcomes are measured similarly or not			
4 Was the synthetic control data set sufficiently reliable and comprehensive?		State whether there is sufficient sample sizes and covariates that can crea comparable control groups			
5	Were there any other major limitations to the dataset?	State any other potential limitations of the dataset that would limit the reliability and validity of comparisons			
Synthetic C	Control Methods				
6	Did the clinical trial include a concurrent control arm, or is the synthetic control data the only control data?	State the size of the concurrent control arm and whether the external data set is the only dataset being used or is being used to complement concurrent control arm(s)			
7	How was the synthetic control data matched to the intervention group?	State the analytical method(s) – eg propensity matching scores – used t create the synthetic control arm			
8 Were the results robust to sensitivity assumptions and potential biases?		State whether the sensitivity analyses were undertaken or reasons for r conducting sensitivity analyses, and compare whether the sensitivity analyses were comparable to the primary analyses.			
9	Were synthetic control comparisons possible for all clinically important outcomes?	State if all clinically important outcomes were considered for analyses. not, state justifications for not including all important outcomes			
10	Are the results applicable to your patients?	State whether the synthetic control group created are similar to the patient group of interest			
П	Were there any other major limitations to the synthetic control methods?	State any other potential limitations of the statistical methods that would limit the reliability and validity of comparisons			



4





### **The Success Story of Blinatumomab**

•

OPEN

Citation: Blood Cancer Journal (2016) 6, e473; doi:10.1038/bcj.2016.84

#### www.nature.com/bc

#### ORIGINAL ARTICLE

#### Blinatumomab vs historical standard therapy of adult relapsed/ refractory acute lymphoblastic leukemia

N Gökbuget<sup>1</sup>, M Kelsh<sup>2</sup>, V Chia<sup>2</sup>, A Advani<sup>3</sup>, R Bassan<sup>4</sup>, H Dombret<sup>2</sup>, M Doubek<sup>9</sup>, AK Fielding<sup>7</sup>, S Giebel<sup>8</sup>, V Haddad<sup>9</sup>, D Hoeizer<sup>1</sup>, C Holland<sup>10</sup>, N Ifrah<sup>11</sup>, A Katz<sup>2</sup>, T Maniar<sup>12</sup>, G Martinelli<sup>13</sup>, M Morgades<sup>14</sup>, S O'Brien<sup>15</sup>, J-M Ribera<sup>14</sup>, JM Rowe<sup>16</sup>, A Stein<sup>17</sup>, M Topp<sup>19</sup>, M Wadleigh<sup>9</sup> and H Kantarjian<sup>15</sup>

We compared outcomes from a single-arm study of blinatumomab in adult patients with B-precursor Ph-negative relapsed/ refractory acute lymphoblastic leukemia (R/R ALL) with a historical data set from Europe and the United States. Estimates of complete remission (CR) and overall survival (OS) were weighted by the frequency distribution of prognostic factors in the blinatumomab trial. Outcomes were also compared between the trial and historical data using propensity score methods. The historical cohort included 694 patients with CR data and 112 patients with OS data compared with 189 patients with CR and survival data in the blinatumomab trial. The weighted analysis revealed a CR rate of 24% (95% CI: 26–36%) and a median OS of 3.3 months (95% CI: 28–3.6) in the blinatumomab trial. Propensity score analysis estimated increased odds of CR/CRh (OR=2.68, 95% CI: 167–4.31) and improved OS (HR=0.536, 95% CI: 0.394–0.730) with blinatumomab. The analysis demonstrates the application of different study designs and statistical methods to compare novel therapies for R/A LL with historical data.

Blood Cancer Journal (2016) 6, e473; doi:10.1038/bcj.2016.84; published online 23 September 2016



U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research Office of Translational Sciences Office of Biostatistics

#### STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

BLA/Serial Number #:	#: BLA 125557 / 00	
Supplement #:	Original Biologics License Application	
Drug Name:	Blincyto (blinatumomab) for continuous intravenous infusion	
Indication(s):	Treatment of adults with Philadelphia chromosome negative relapsed or refractory B-precursor acute lymphoblastic leukemia	
Applicant:	Amgen	
Date(s):	Submission date: 19 September 2014	
	PDUFA date: 19 May, 2015	

- Accelerated FDA approval in 2014
- Ph-negative B-precursor cell relapsed/refractory acute lymphoblastic leukemia
- Single-arm phase II trial (MT103-211; n = 189\*)
- Synthetic control arm of clinical trial data from 11 databases in the US and EU (n = 694 for CR and 1112 for OS)
- Estimated the control rate and treatment effect using synthetic control arm

\*NOTE: 4 patients who were not within the definition of relapsed/refractory ALL were excluded from the FDA submission



### **Estimation of the Control Rate for Complete Response<sup>1</sup>**

**Table 2.** Stratified and weighted analysis results: comparison of historical data and blinatumomab clinical trial data: CR by strata and weighted to blinatumomab clinical data

Stratum Definition		Blinatumomab trial (MT103-211)				Historical data set			
Age	Disease status	Ν	Stratum proportion (%)	Number with CR <sup>a</sup>	CR/CRh % (95% Cl)	Ν	Stratum proportion (%)	Number with CR	CRsg % (95% Cl)
< 35	Prior alloHSCT <sup>b</sup>	40	21.2	15	38 (23, 54)	48	6.9	14	29 (17, 44)
< 35	In 1st salvage <sup>c</sup>	10	5.3	7	70 (35, 93)	119	17.1	52	44 (35, 53)
< 35	In 2nd or greater salvage <sup>c</sup>	40	21.2	17	43 (27, 59)	150	21.6	27	18 (12, 25)
≥35	Prior alloHSCT <sup>b</sup>	24	12.7	14	58 (37, 78)	41	5.9	11	27 (14, 43)
≥35	In 1st salvage <sup>c</sup>	19	10.1	5	26 (9, 51)	187	26.9	57	30 (24, 38)
≥35	In 2nd or greater salvage <sup>c</sup>	56	29.6	23	41 (28, 55)	149	21.5	25	17 (11, 24)
Com	bined weighted summary	189	100	81	43 (35, 50)	694	100	186	24 (20, 27)

Note: 186/694 = 27%; estimated 95% confidence intervals using bootstrapping; also conducted sensitivity analyses by time period

<sup>1</sup> Gökbuget N, Kelsh M, Chia V, Advani A, Bassan R, Dombret H, Doubek M, Fielding AK, Giebel S, Haddad V, Hoelzer D. Blinatumomab vs historical standard therapy of adult relapsed/refractory acute lymphoblastic leukemia. Blood cancer journal. 2016;6(9):e473.



### **Estimation of treatment effect<sup>1</sup>**

#### Results from IPTW weighted logistic regression and Cox PH models



<sup>1</sup> Gökbuget N, Kelsh M, Chia V, Advani A, Bassan R, Dombret H, Doubek M, Fielding AK, Giebel S, Haddad V, Hoelzer D. Blinatumomab vs historical standard therapy of adult relapsed/refractory acute lymphoblastic leukemia. Blood cancer journal. 2016;6(9):e473.



### Summary of Blinatumomab Case Study

- Synthetic control arms can lead to regulatory approval without a head-to-head comparison in a phase III trial
- Success of Blinatumomab can be attributable to the:
  - 1) Appropriateness of the indication
  - 2) Proper identification of external control
  - 3) Robustness of the statistical analyses







### **Dynamic borrowing**



- Information from one or more external data sources are **borrowed** to augment the control arm and allow for unequal randomization (e.g. 3:1)
- Heterogeneity between the control arm and the external data are assessed
- The external data are dynamically down-weighted proportional to the degree of similarity between the external and control arms



### **Estimation via dynamic Bayesian borrowing**



- Bayesian methods can dynamically borrow from historical data sources using informative priors
- Viele (2014) and Wadsworth (2019) provide an overview of statistical methods<sup>1,2</sup>

<sup>1</sup> Viele K, Berry S, Neuenschwander B, Amzal B, Chen F, Enas N, Hobbs B, Ibrahim JG, Kinnersley N, Lindborg S, Micallef S. Use of historical control data for assessing treatment effects in clinical trials. Pharmaceutical statistics. 2014 Jan;13(1):41-54. <sup>2</sup> Wadsworth I, Hampson LV, Jaki T. Extrapolation of efficacy and other data to support the development of new medicines for children: a systematic review of methods. Statistical methods in medical research. 2018 Feb;27(2):398-413.





Characteristic	INTEREST	ZODIAC	PROCLAIM	Study 57
Overall sample size ( <i>n</i> )	1433	1391	598	1240
Control group		10		
median overall	8	10	25	8
survival, months	(4, 14)	(4, 13)	(10, 35)	(4, 12)
(Q1, Q3)				
Stage III, (%)	38	15	100	17
Stage IV, (%)	53	85	0	83
Average age, (years)	61	59	59	61
Adenocarcinoma histology, (%)	54	60	75	60
Two or more				
prior	16	0	0	25
chemotherapy	10	0	0	35
regimens, (%)				
Radiotherapy			60 66 CV	
sequence, dose,	None	None	Concurrent	None
(control arm)			Concurrent	

- We used IPD from four second line non-small cell lung cancer trials for patients with stage III/IV disease.
- Used the INTEREST trial as the 'anchor' trial and simulated its performance with fewer patients randomized to control, but *dynamically borrowing* from the three other trials' control groups



Dron L, Golchi S, Hsu G, Thorlund K. Minimizing control group allocation in randomized trials using dynamic borrowing of external control data–An application to second line therapy for non-small cell lung cancer. Contemporary clinical trials communications. 2019 Dec 1;16:100446.

### Case Study<sup>1</sup>





- In the original trial (INTEREST) 734 patients were recruited to control
- With a 1:1, a 1:2 or a 1:4 ratio of control patients to borrowed controls, we demonstrate a relatively minor change in precision
- Testing various scenarios showed that the control arm can be reduced by up 60% without jeopardizing the validity of statistical inferences
- The findings also applied where external control data was heterogeneous, but not inconsistent with the concurrent control



Dron L, Golchi S, Hsu G, Thorlund K. Minimizing control group allocation in randomized trials using dynamic borrowing of external control data–An application to second line therapy for non-small cell lung cancer. Contemporary clinical trials communications. 2019 Dec 1;16:100446.

### The FDA recognizes the role of dynamic borrowing<sup>1,2</sup>

"This document provides guidance to sponsors and applicants on interacting with the FDA on complex innovative trial design (CID) proposals for drugs or biological products. FDA is issuing this guidance to satisfy, in part, a mandate under section 3021 of the 21<sup>st</sup> Century Cures Act [...] Some examples of trial designs that might be considered novel or CID are those that formally borrow external or historical information or borrow control arm data from previous studies to expand upon concurrent controls."2



<sup>1</sup> 2019 FDA guidance for Rare Diseases: Natural History Studies for Drug Development.
<sup>2</sup> 2019 FDA guidance for Interacting with the FDA on Complex Innovative Clinical Trial Designs for Drugs and Biological Products



### When ethics impede data utility in placebo cross-over designs

- Placebo cross-over can result in substantial reductions in 'true' placebo sample size or duration
- Common in:
  - Auto-immune diseases
  - Cancers with poor prognosis
  - Chronic diseases with 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, ... lines of Tx



#### When ethics impede data utility in placebo cross-over designs





### **Regulatory landscape – an overview**



Trial designs using real-world data: The changing landscape of the regulatory approval process

Elodie Baumfeld Andre<sup>1</sup> | Robert Reynolds<sup>1,2</sup> | Patrick Caubel<sup>1</sup> | Laurent Azoulay<sup>3,4</sup> | Nancy A. Dreyer<sup>5,6</sup> |

TABLE 2 Examples of recent regulator-supported initiatives that drive forward the application of RWE

	Details of the Initiative			
Guidance to submit documents using real-world data and real world evidence to FDA for drugs and biologics (2019) <sup>24</sup>	This guidance is intended to encourage sponsors and applicants who are using RWD to generate RWE as part of a regulatory submission to FDA to provid information on their use of RWE in a simple, uniform format.			
HMA/EMA Joint Task Force on Big Data (2019) <sup>25</sup>	This document provides recommendations for a path towards understanding the acceptability of RWE from "Big data" sources to support regulatory evaluation and monitoring.			
Framework for FDA: Real-world evidence program (2018) $^{\rm B}$	The FDA published a framework to help evaluate the potential use of RWE to support new indications for drugs already approved or to satisfy post-approva study requirements. This framework demonstrates a commitment to transforming the drug development cycle.			
Guidance on the use of Electronic Health Records (FDA) $(2018)^{26}$	The focus of this guidance is on data integrity. It emphasizes the need to cite the "data originator" and preserve the audit trail. It also reinforces that RWE may be used to inform approval of new indications for approved drugs and to satisfy post-approval study requirements.			

٠

- Baumfeld Andre et al. 2019 is a high quality of review of trial designs utilizing real-world data as part of their submissions.
- They provide examples of 22 submissions to the EMA and FDA which utilized real-world evidence as part of their submission process for approvals (either accelerated or full).
  - They identified most applications were for either rare metabolic disease or rare oncology applications



### **Regulatory landscape – complexity ≠ success**



An illustration of the Brineura intracerebroventricular infusion process.

- Cerolipinase alfa (Brineura) is a treatment for Batten disease a rare inherited nervous system disorder.
- Owing to the disease rarity and the associated treatment administration, a small, single-arm trial of 22 patients was conducted.
- Efficacy data was compared (using a 1:1 matching algorithm) to a well-defined historical natural history cohort of 42 patients.
- FDA/EMA approval was granted on the basis of this comparison.







#### When things get a bit too messy

#### So what have we learned?





## Thank You!