Using Synthetic Control Databases to Accelerate Indication-Specific Safety & Efficacy Evidence
Lisa Ensign PhD*, Colin Neate M.Sc.**

1: INTRODUCTION
Synthetic control databases (SCD) of recent clinical trial data are one potential tool for accelerating development programs.

SCDs may provide rigorous pooled clinical data beyond that available in published literature and have certain strengths over real-world data (RWD).

SCDs may aid in understanding patient populations, target response rates, background AE rates including rare events, and relationship between variables.

A pilot has been ongoing to develop an SCD for metastatic breast cancer (mBC).

2: SCD CONCEPT
An aggregated dataset of hundreds (or thousands) of patients from multiple recent trials chosen to match a researcher’s inclusion/exclusion criteria for an indication

3: SCD PILOT in mBC

Inclusion criteria:
- mBC subtypes:
  - Hormone Receptor positive, HER2 negative (HR+/HER2-)
  - Triple Negative (TNBC)

Lines of therapy:
- 2nd or later line
- high-risk 1st line

Completed Phase II & Phase III open-label trials with RECIST endpoints

Medidata Rave® studies with data aggregation rights

Version 1 delivered Oct 2018 and updated quarterly

4: CURRENT STATUS
SCD as of May 2019 includes 1528 patients from 5-10 Phase II & III clinical trials
- 779 (51%) HR+/HER2-
- 749 (49%) TNBC

Large majority (>90%) of patients from trials in 2nd or later line of therapy

Includes >190 patient characteristic, efficacy and safety outcome variables

User accesses SCD by generating tabular and graphical summaries in interactive Visualizer. To protect re-identification, filtered subgroups are subject to 5/2 rule (at least 5 patients and 2 Sponsors)

5: EXAMPLE EXPLORATORY SUMMARIES USING THE SCD VISUALIZER

6: CHALLENGES
Terminology Standardization

Simpson’s Paradox - Unadjusted comparisons can lead to misleading results

Ph 2 trial, 1:1 randomization Overall AE rate 1%

Ph 3 trial, 2:1 randomization Overall AE rate 25%

Simple Pooling of both trials

Other Challenges
- Aggregating data to protect patient identification
- Different response assessment schedules
- Limited to clinical data captured in database - excludes data maintained externally (e.g., molecular)
- Possible publication bias regarding companies allowing data to be shared
- Blinded studies - cannot link patient characteristics/outcomes to treatment
- Small sample size in subpopulations

7: CURRENT USES OF PILOT SCD
Currently focus is on exploratory questions via:
- Dynamic filtering of SCD to aid understanding contemporary breast cancer populations. E.g., HR+/HER2-mBC contains heterogeneous subpopulations defined by treatment history (chemo eligible, CDK4/6i treated, etc.) and other characteristics.
- Benchmarking efficacy and safety outcomes for interpreting Ph 1b (often single-arm) studies and designing pivotal studies. HR+/HER2- mBC in particular has seen many recent shifts in SoC increasing this need.

8: CONCLUSIONS / FUTURE VISION
SCDs have potential to aid research programs and future study designs by providing data-driven benchmarks of clinical response, outcomes and safety, insights into disease subsets and correlations between short and long-term endpoints.

Arguably SCDs provide higher level data quality and patients who are more similar to clinical trial participants than real world data (RWD). However, questions addressed by SCDs are limited to the data collected in the underlying studies. Due to heterogeneity across studies in the type/format of data collected, substantial manual mapping and standardisation are required when creating the SCD. Industry standardisation (eg CRFs, data structure) will facilitate creation and use of SCDs.

If above challenges are addressed, possible future is for synthetic data to be incorporated in regulatory decisions:
- External control arms in uncontrolled trials or supplementary controls in partially randomised trials
- Contemporary benchmarks (in combination with other sources of evidence) for interpreting safety signals and quantifying benefit/risk in support of regulatory submissions

9: DISCUSSION QUESTIONS
- Bias in excluding blinded trials from SCD?
- Synthetic control database vs Synthetic control arms?
- Vision for use in submissions or other use cases?
- Standardisation benefits?
- Combining with RWD?