# Using Synthetic Control Databases to Accelerate Indication-Specific Safety and Efficacy Evidence

Colin Neate M.Sc.[1], Lisa Ensign PhD[2], Aruna Mani MD[3]
[1]Hoffmann-La Roche Canada; [2]Medidata Solutions (Acorn AI); [3]Genentech, A Member of the Roche Group;

## Introduction

Synthetic control databases (**SCD**[TM]) of recent clinical trial data are one potential tool for **accelerating** development programs.

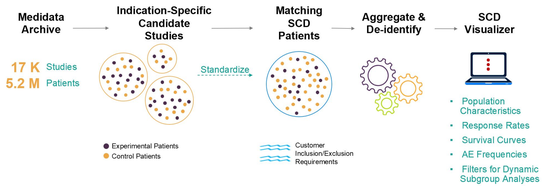
SCDs may provide **rigorous pooled clinical data** beyond that available in published literature and have certain strengths over real-world data (RWD).

SCDs may aid in **planning, executing and interpreting** clinical trials through e.g.,: understanding patient characteristics, defining efficacy targets, identifying background AE rates (including rare events) and expected relationship between variables.

A pilot project has been ongoing to develop and evaluate use of an SCD for **metastatic breast cancer** (mBC). Version 1 was delivered October 2018 and the SCD has subsequently been updated quarterly.

## 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




## Patient Selection Criteria for mBC Pilot

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

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

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

Medidata Rave® studies with data aggregation rights

## Current Uses of Pilot SCD

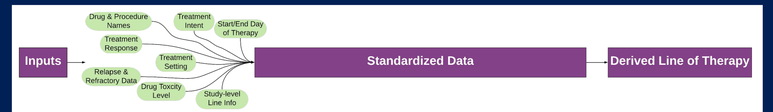
Current focus is on ***exploratory*** questions, including:

- **Dynamic filtering** of SCD to aid understanding contemporary mBC populations. E.g., HR+/HER2- mBC contains heterogeneous subpopulations defined by treatment history (chemo eligible, CDK4/6i treated, etc.) and other factors.
- **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 standard of care increasing this need for external data sources.

DIA 2019
GLOBAL ANNUAL MEETING
SAN DIEGO | JUNE 23-27

## Data Domains, Variables and Example Results from Current mBC SCD (May 2019)

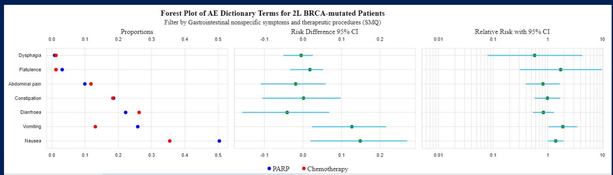
### Complex Derived Variables Included in SCD: Accomplished by Subject Matter Experts, Algorithms & Machine Learning




| | Standardized Data | | | | | |
|---|---|---|---|---|---|---|
| Reported Term | Preferred Term | Treatment Intent | Best Overall Response | Therapy Study Start Day | Therapy Study End Day | Drug Class |
| PACLITAXEL | paclitaxel | Curative | | -2352 | -2268 | Taxanes |
| GOSERELIN | goserelin | Curative | CR | -2170 | -1450 | Endocrine therapy not otherwise speciifed |
| TAMOXIFEN | tamoxifen | Curative | | -2170 | -1294 | SERM |
| CYCLOPHOSPHAMIDE | cyclophosphamide | Adjuvant | | -1268 | -1207 | Chemotherapy not otherwise specified |
| DOCETAXEL | docetaxel | Adjuvant | | -1268 | -1207 | Taxanes |
| LETROZOLE | letrozole | | SD | -1099 | -456 | AI |
| DENOSUMAB | denosumab | | | -988 | | Supportive Care |
| EXEMESTANE | exemestane | | PD | -456 | -343 | AI |
| EVEROLIMUS | everolimus | | PD | -442 | -343 | Protein Kinase Inhibitors (mTOR inhibitor) |
| FULVESTRANT | fulvestrant | | SD | -343 | -175 | SERD |
| NANOPARTICLE ALBUMIN-BOUND PACLITAXEL (ABRAXANE) | paclitaxel | | PD | -133 | -63 | Taxanes |

Curative & Adjuvant → Early Breast Cancer Therapy
1st Line
Filter Out Supportive Care
Start/End within 60 Days & Progressive Disease → 2nd Line
3rd Line
More toxic than previous therapy → 4th Line

### Detailed AE Info Standardized into MedDRA: Supports SMQ Filtering and Investigation of Subpopulations

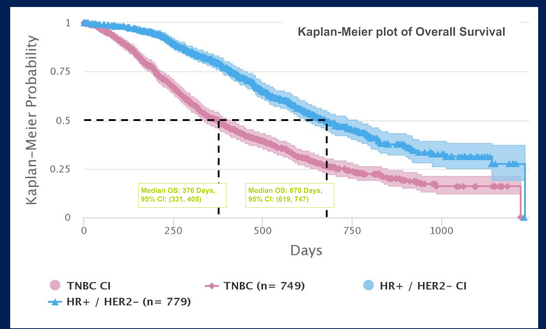



| | PARP Inhibitor TOTAL N = 220 | | Chemotherapy TOTAL N = 76 | | | |
|---|---|---|---|---|---|---|
| Variable | Count | Incidence | Count | Incidence | Difference (Adj. 95% CI) | Adj. Relative Risk (95% CI) |
| Dysphagia | 2 | 0.0091 | 1 | 0.0132 | -0.0041 (-0.05035, 0.0261) | 0.5807 (0.08, 4.32) |
| Flatulence | 7 | 0.0318 | 1 | 0.0132 | 0.0187 (-0.0324, 0.05319) | 1.7421 (0.31, 9.86) |
| Abdominal pain | 22 | 0.1 | 9 | 0.1184 | -0.0184 (-0.10893, 0.05973) | 0.8252 (0.4, 1.68) |
| Vomiting | 57 | 0.2591 | 10 | 0.1316 | 0.1275 (0.02377, 0.2167) | 1.908 (1.04, 3.49) |
| Constipation | 41 | 0.1864 | 14 | 0.1842 | 0.0022 (-0.10463, 0.09839) | 0.9972 (0.58, 1.71) |
| Diarrhoea | 49 | 0.2227 | 20 | 0.2632 | -0.0404 (-0.15674, 0.06873) | 0.8413 (0.54, 1.31) |
| Nausea | 111 | 0.5045 | 27 | 0.3553 | 0.1493 (0.02039, 0.27067) | 1.4127 (1.02, 1.96) |

### Standardized Data Across Multiple Domains: Allows Detailed Exploratory Analyses

| Domain Name | Variables (Subset Shown) |
|---|---|
| **Adverse Events** | Causality, Outcome, Preferred Term (+ all MedDRA levels), SAEs, Severity |
| **Baseline Disease Characteristics** | Age at Diagnosis of Disease, Baseline ECOG, Baseline Lactate Dehydrogenase, Baseline Serum Albumin, BC Subtype (TNBC or HER2- / HR+), Diagnosis Preferred Term, Disease Stage (Including TNM), Line of Therapy, Menopausal Status, Relapsed / Refractory Flags, Tumor Grade |
| **Biomarkers** | BRCA, BRCA1, BRCA2, CA 15-3, CA 17-19, Estrogen Receptor, HER2, Hormone Receptor, Progesterone Receptor |
| **Concomitant Medications** | Ongoing, Preferred Term (+ all ATC levels), Route |
| **Demographics** | Age, Baseline Height/Weight, Ethnicity, Gender, Race, Region |
| **Disease Response** | Best Overall Response, Clinical Benefit, Death, Progression Free Survival, Overall Survival, Time to First Partial or Complete Response |
| **Disposition** | Death Study Day, Duration of All Treatments, End of Efficacy / Treatment / Safety Reason, Lost to Follow Up |
| **Exposure** | Assigned Dose, Dose Unit, Drug Class, Treatment Duration |
| **Medical History** | Dictionary Term (+ all MedDRA levels), Length of the Condition, Ongoing |
| **Prior Anti-Cancer Regimens / Therapies** | Drug Class, Line of Therapy, Preferred Term, Therapy Duration, Treatment Intent, Treatment Setting |
| **Summary of Metastases** | Abdominal, Bone, Brain, Liver, Lung, Visceral |
| **Vital Signs** | Baseline Flag, Numeric Result, Test Parameter, Visit Study Day |

### SCD Insights: Offers Greater Detail and Compares Well with Published Literature Overall Survival for TNBC < HR+/HER2-




### Highly Nuanced and Complete Data in SCD: An Advantage Over many RWD Sources

| | | TNBC (N = 749) | HER2- / HR+ (N = 779) | Total (N = 1528) |
|---|---|---|---|---|
| **BRCA1 or BRCA2** | Positive | 379 (50.6%) | 431 (55.3%) | 810 (53.0%) |
| **Derived Line of Therapy for the Current Treatment** | 1L | 0 ( 0.0%) | 114 (14.6%) | 114 ( 7.5%) |
| | 2L | 328 (43.8%) | 230 (29.5%) | 558 (36.5%) |
| | 2L+ | 265 (35.4%) | 163 (20.9%) | 428 (28.0%) |
| | 3L+ | 156 (20.8%) | 272 (34.9%) | 428 (28.0%) |
| **Assigned Treatment Arm Drug Class** | Antimetabolites | 146 (19.5%) | 63 ( 8.1%) | 209 (13.7%) |
| | Antimetabolites + SERD | 0 ( 0.0%) | 46 ( 5.9%) | 46 ( 3.0%) |
| | Antimetabolites + Targeted-Other | 37 ( 4.9%) | 79 (10.1%) | 116 ( 7.6%) |
| | Chemotherapy not otherwise specified | 41 ( 5.5%) | 52 ( 6.7%) | 93 ( 6.1%) |
| | PARP inhibitors | 270 (36.0%) | 300 (38.5%) | 570 (37.3%) |
| | SERD | 0 ( 0.0%) | 51 ( 6.5%) | 51 ( 3.3%) |
| | SERM + Endocrine therapy not otherwise specified | 0 ( 0.0%) | 157 (20.2%) | 157 (10.3%) |
| | Targeted-Other | 213 (28.4%) | 0 ( 0.0%) | 213 (13.9%) |
| | Missing | 42 ( 5.6%) | 31 ( 4.0%) | 73 ( 4.8%) |
| **Sites of Metastases at Enrollment** | Bone | 138 (31.2%) | 544 (70.0%) | 682 (56.0%) |
| | Brain | 95 (21.5%) | 35 ( 4.5%) | 130 (10.7%) |
| | Liver | 95 (21.5%) | 270 (34.7%) | 365 (29.9%) |
| | Lung | 212 (48.0%) | 258 (33.2%) | 470 (38.6%) |
| | Other | 250 (56.6%) | 471 (60.6%) | 721 (59.1%) |
| **Visceral Metastasis** | Yes | 382 (86.4%) | 638 (82.1%) | 1020 (83.7%) |

## Current Status of Pilot SCD

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

The large majority (>90%) of patients are from trials in 2nd or later line of therapy.

The SCD includes >190 patient characteristic, efficacy and safety outcome variables.

Researchers access the SCD through an interactive Visualizer which allows them to generate tabular and graphical summaries.

To mitigate risk of identification of patient or trial-level results, all subject, site and sponsor data is de-identified, extreme outliers are removed, and filtered subgroups are subject to a 5/2 rule (at least 5 patients and 2 Sponsors).

## Conclusions

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. Current restriction to open-label studies may reduce representativeness of the SCD. 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 (e.g., CRFs, data structure) will facilitate creation and use of SCDs.

If these challenges are addressed, possible future is for evidence from SCDs of external clinical trials 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

## Disclosures

Colin Neate, Employee and shareholder of Hoffmann-La Roche
Lisa Ensign, Employee and shareholder of Medidata
Aruna Mani, Employee of Genentech and shareholder of Hoffmann-La Roche