USING ELECTRONIC HEALTH RECORDS AS A REAL WORLD COMPARATOR

Using Flatiron's electronic health records as a real-world comparator in the setting of single-arm trials

Jessica Davies
In collaboration with Michael Martinec, Gracy Crane, Mathieu Coudert, Rachel Rosenthal, Dietrich Knoerzer, Thomas Kuenzel, Reynaldo Martina
PSI Conference 2018

Effectiveness evidence comparing new therapies to SOC SOC is needed at the time of reimbursement

- Accelerated access to novel targeted therapies based on Phase Ib or II data has preceded availability of phase III data for patient populations in the late-stage oncology setting.

- Indirect treatment comparisons (ITC) are commonly used in economic decision making to compare randomized trial data when there is one common treatment arm. However, in a single arm trial setting, an ITC may not be feasible.
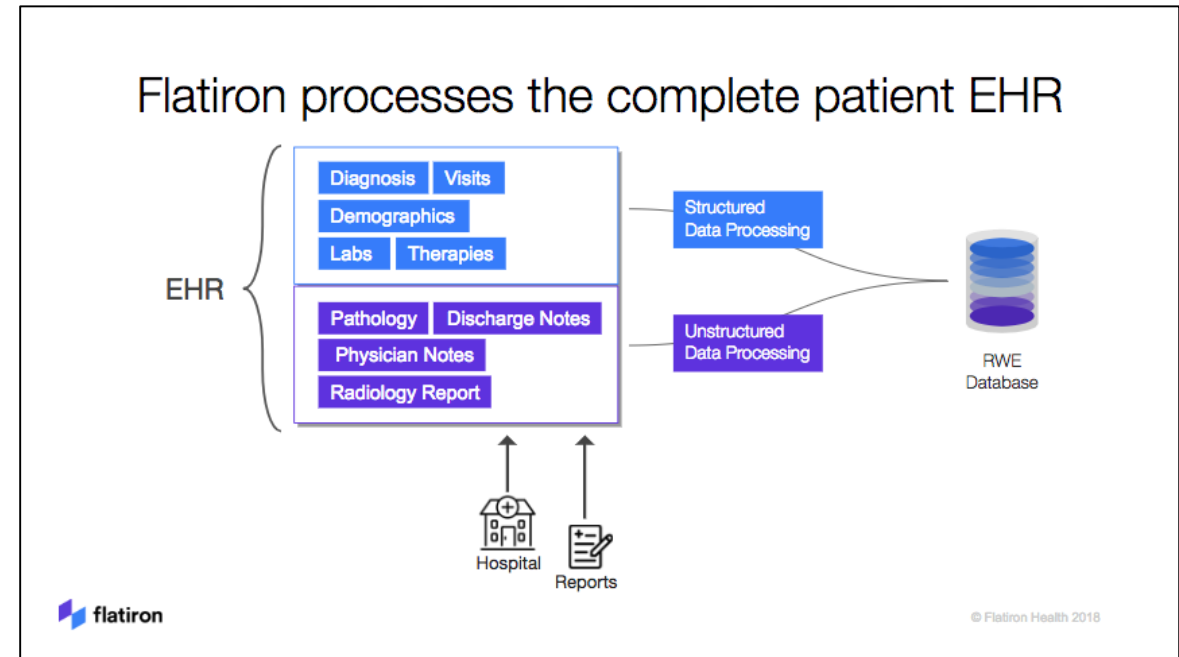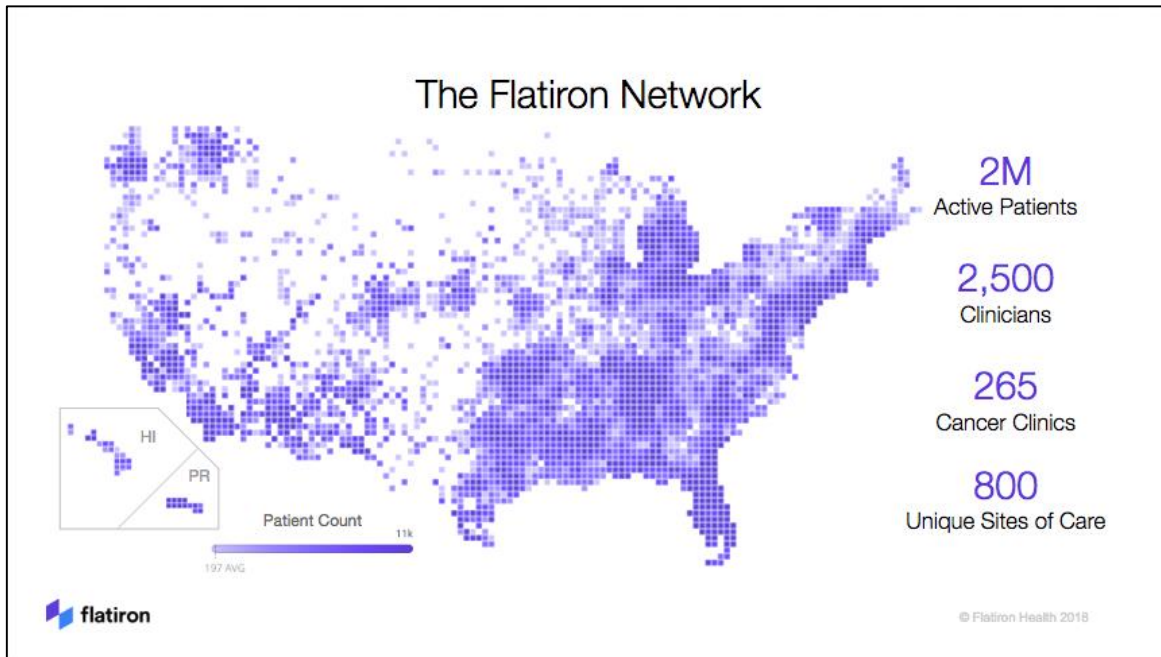
Real world data (RWD) could provide a source to complement trial data in this setting by building a comparison with an external control.
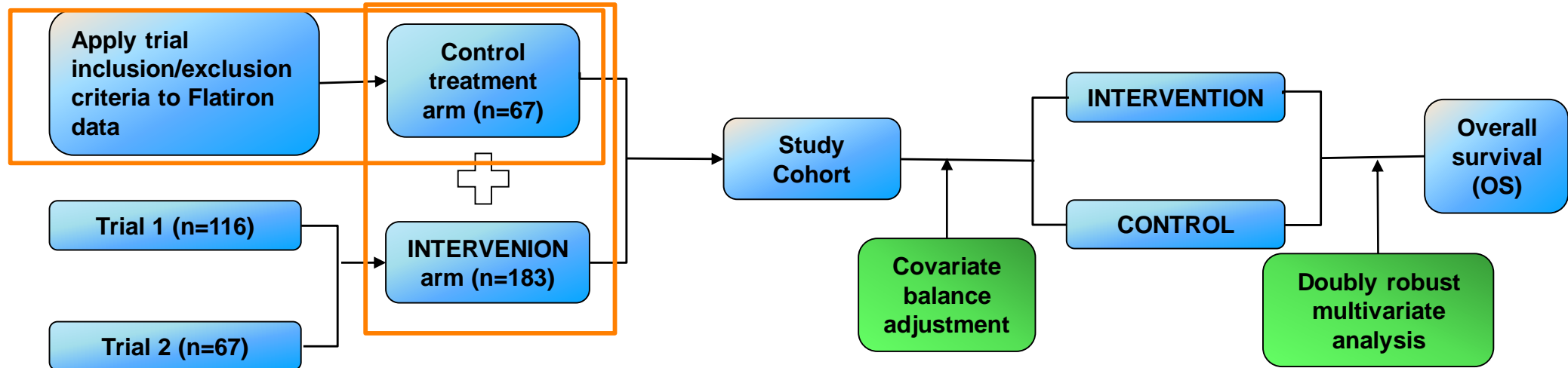
In this illustrative example, we used individual patient level data (IPD) extracted from electronic health records (EHR) to estimate comparative effectiveness of a single-arm trial.

Flatiron provides patient level data to enable contemporary external controls

The Flatiron Network

2M Active Patients

2,500 Clinicians

265 Cancer Clinics

800 Unique Sites of Care

Patient Count
197 AVG — 11k

HI
PR

© Flatiron Health 2018

Flatiron processes the complete patient EHR

EHR

Diagnosis | Visits
Demographics
Labs | Therapies

Pathology | Discharge Notes
Physician Notes
Radiology Report

Structured Data Processing

Unstructured Data Processing

RWE Database

Hospital

Reports

© Flatiron Health 2018

Comparative effectiveness with an external control requires additional analytical steps to address indication bias
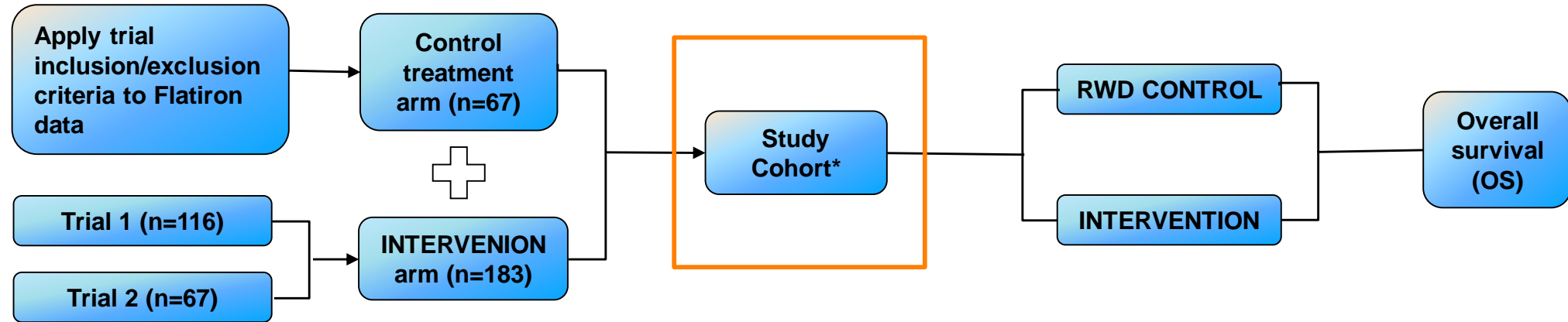

Study Population:

- Apply trial inclusion and exclusion criteria
- IPD from three different data sources were pooled for a study population

Statistical Analysis:

- 3 adjusted comparisons were made to enable covariate balance:
  - Inverse propensity treatment weighting (IPTW)
  - Propensity score matching (PSM)
  - Genetic Matching (GenMatch)
- Assess standardized mean differences
- Overall survival was calculated from initiation of treatment to the end of follow-up and/or death
- Multivariate Cox proportional hazards model to compare OS and estimate the average treatment effect
- Comparison with TRIAL CONTROL data & Sensitivity analyses

Comparative effectiveness with an external control requires additional analytical steps to address indication bias

Study Population:

- Apply trial inclusion and exclusion criteria
- IPD from three different data sources were pooled for a study population

Statistical Analysis:

- 3 adjusted comparisons were made to enable covariate balance:
    - Inverse propensity treatment weighting (IPTW)
    - Propensity score matching (PSM)
    - Genetic Matching (GenMatch)
- Assess standardized mean differences
- Overall survival was calculated from initiation of treatment to the end of follow-up and/or death
- Multivariate Cox proportional hazards model to compare OS and estimate the average treatment effect
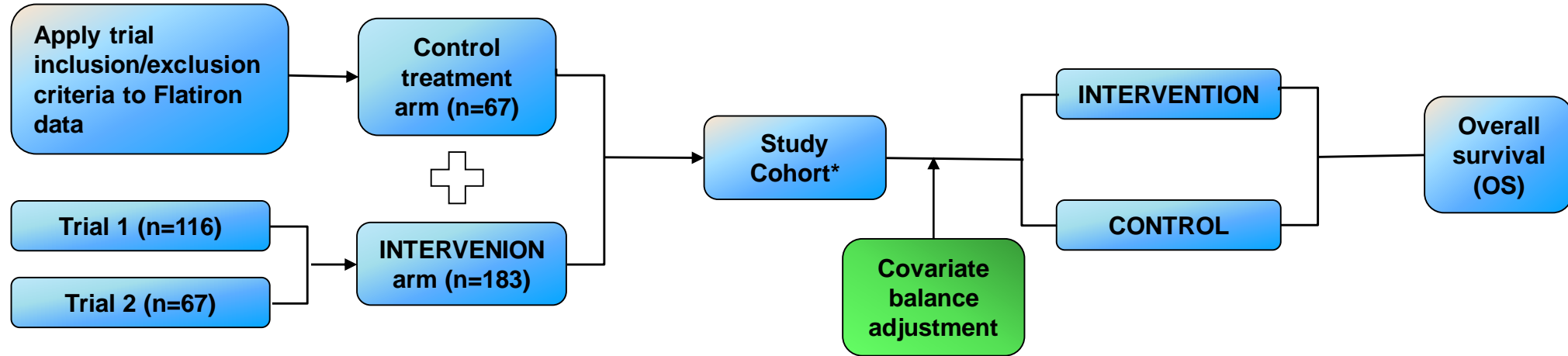- Comparison with TRIAL CONTROL data & Sensitivity analyses

Characteristics were imbalanced prior to covariate balance

Table 1. Patient demographics and clinical characteristics prior to covariate balance

Characteristics	INTERVENTION (n=183)	RWD CONTROL (n=67)
Median age, years (IQR)	55 (44–61)	61 (53–67)
<65 years, n (%)	160 (87)	41 (61)
Female, n (%)	98 (54)	37 (55)
White, n (%)	133 (73)	49 (73)
ACA histology, n (%)	175 (96)	61 (91)
Stage IV, n (%)	170 (93)	59 (88)
No smoking history, n (%)	121 (66)	37 (55)
CNS metastasis, n (%)	111 (61)	23 (34)
Prior line of treatment, n (%)		
1	52 (28)	38 (57)
2	66 (36)	20 (30)
≥3	65 (36)	9 (13)

ACA, adenocarcinoma; CNS, central nervous system; IQR, interquartile range

Comparative effectiveness with an external control requires additional analytical steps to address indication bias

Study Population:

- Apply trial inclusion and exclusion criteria
- IPD from three different data sources were pooled for a study population

Statistical Analysis:

- 3 adjusted comparisons were made to enable covariate balance:
    - Inverse propensity treatment weighting (IPTW)
    - Propensity score matching (PSM)
    - Genetic Matching (GenMatch)
- Assess standardized mean differences
- Overall survival was calculated from initiation of treatment to the end of follow-up and/or death
- Multivariate Cox proportional hazards model to compare OS and estimate the average treatment effect
- Comparison with CONROL TRIAL data & Sensitivity analyses

INTERVENTION and CONTROL more closely overlapped after covariate balance adjustments with IPTW and GenMatch

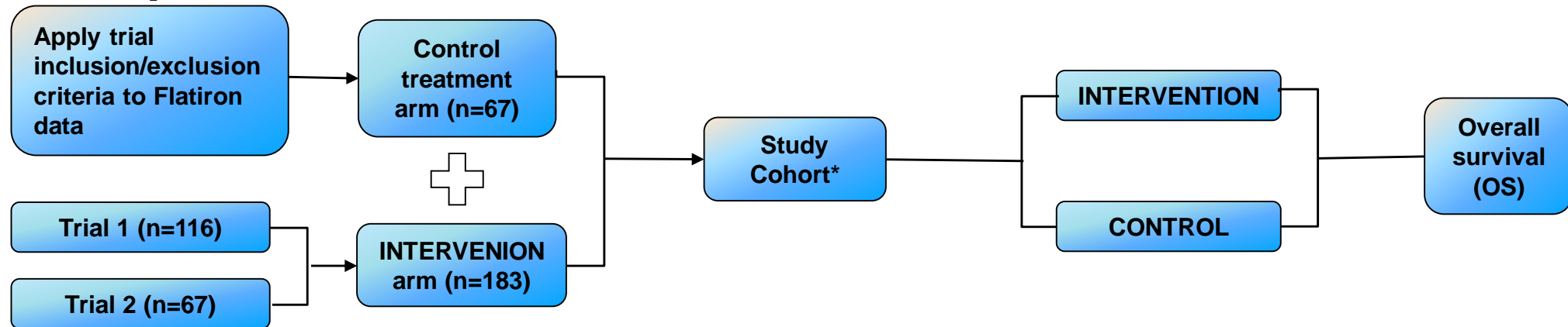| Table 2. Distribution and absolute difference after balance adjustments											
Characteristic	**Unadjusted**			**IPTW**			**PSM**			**GenMatch**		
	I	**C**	**Diff**	**I**	**C**	**Diff**	**I**	**C**	**Diff**	**I**	**C**	**Diff**
Age <65	87	61	-26	79.9	79.5	0.4	87.4	68.1	-19.3	89.8	89.8	0.0
Female	54	55	-1	54.4	53.1	1.3	53.6	48.9	4.7	45.9	45.9	-2.7
White	73	73	0	72.5	72.4	0.1	72.7	80.9	-8.2	84.7	84.7	0.7
Stage IV	93	88	5.0	91.8	92.4	-0.6	92.9	91.5	1.4	96.3	96.3	1.4
Prior line of treatment												
1	28	57	-29	36.2	37.5	-1.3	28.4	51.1	-22.7	52.6	52.6	0.0
2	36	30	6	34.4	36.0	-1.6	36.1	31.9	4.2	34.6	34.6	0.0
>3	36	13	23	29.4	26.6	2.8	35.5	17.0	18.5	12.7	12.7	0.0

After balance adjustment IPTW & GenMatch achieved covariate balance
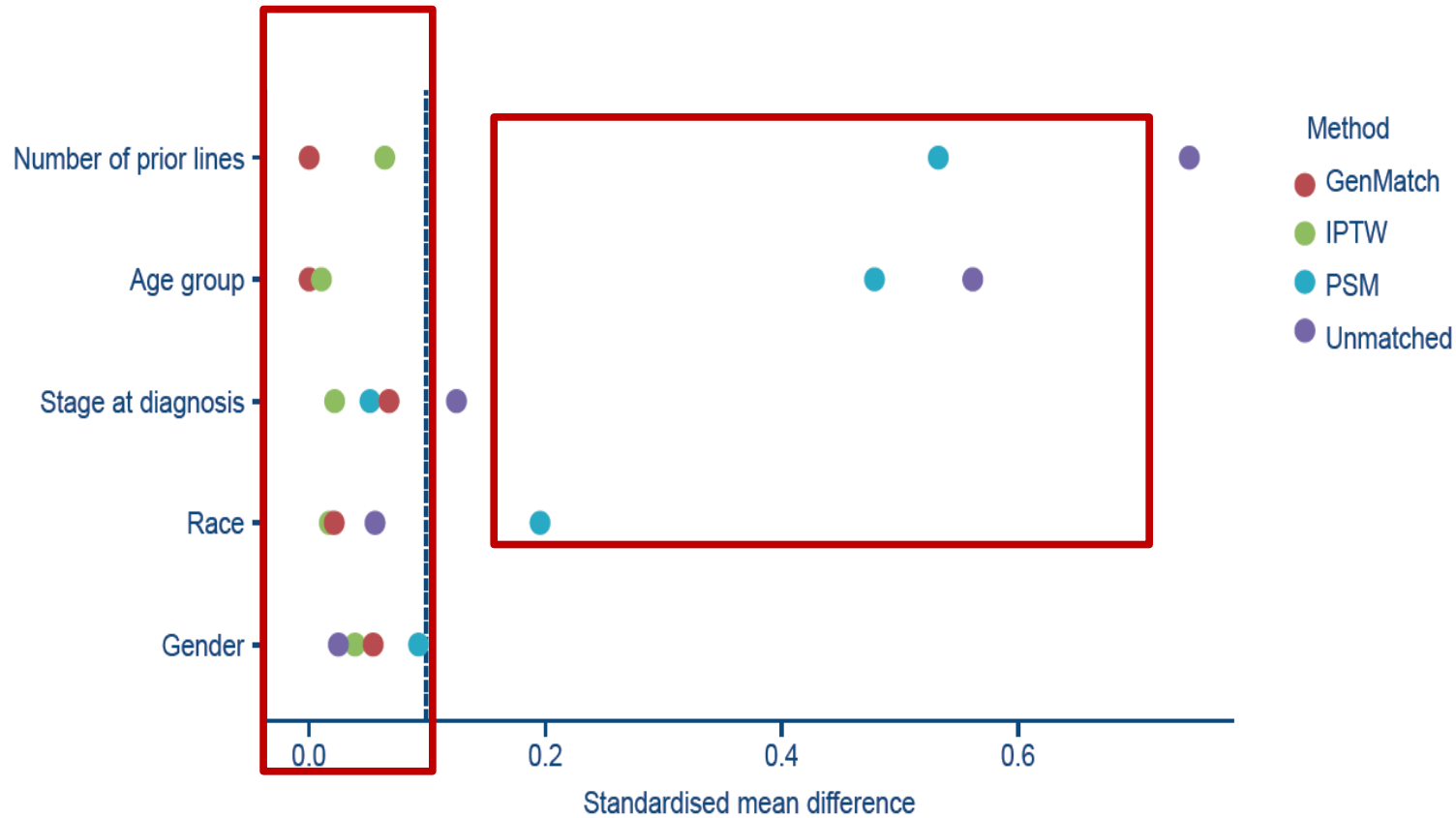
- After balancing adjustments:
  - IPTW and GenMatch improved balance (SMD <10%)
  - PSM did not balance measured confounders (SMD ≥25%)

9

Comparative effectiveness with an external control requires additional analytical steps to address indication bias
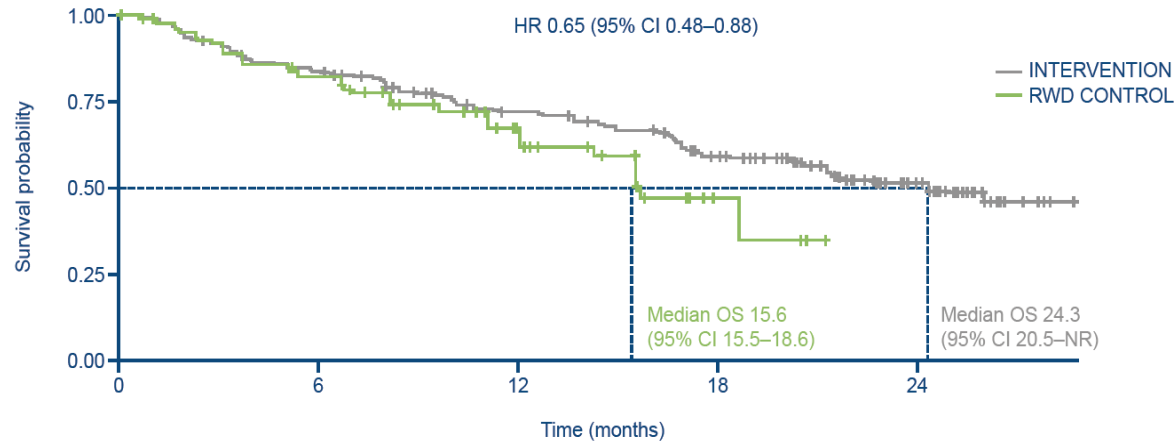
Study Population:

- Apply trial inclusion and exclusion criteria
- IPD from three different data sources were pooled for a study population

Statistical Analysis:

- 3 adjusted comparisons were made to enable covariate balance:
  - Inverse propensity treatment weighting (IPTW)
  - Propensity score matching (PSM)
  - Genetic Matching (GenMatch)
- Assess standardized mean differences
- Overall survival was calculated from initiation of treatment to the end of follow-up and/or death
- Multivariate Cox proportional hazards model to compare OS and estimate the average treatment effect
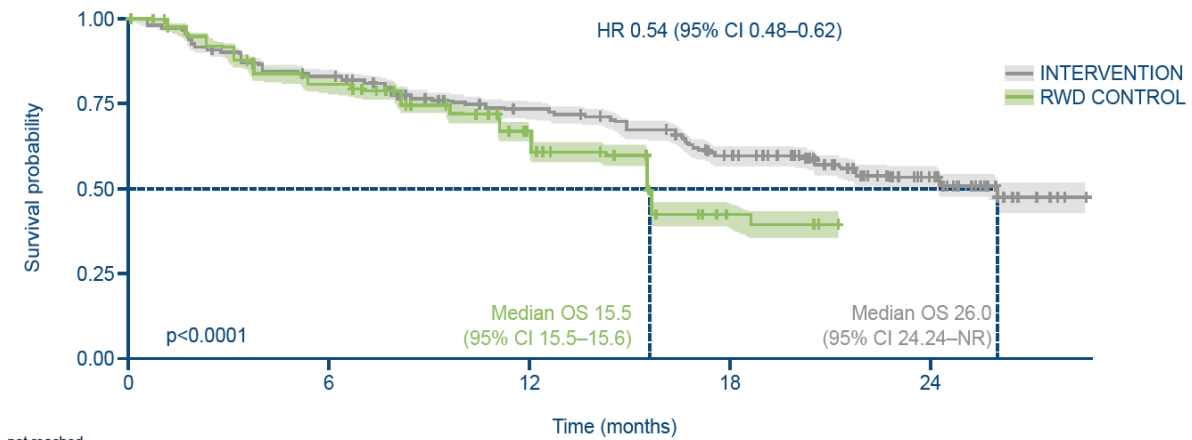- Comparison with TRIAL CONTROL data & Sensitivity analyses

Survival by IPTW and GenMatch analysis

Figure 2. Survival by IPTW analysis

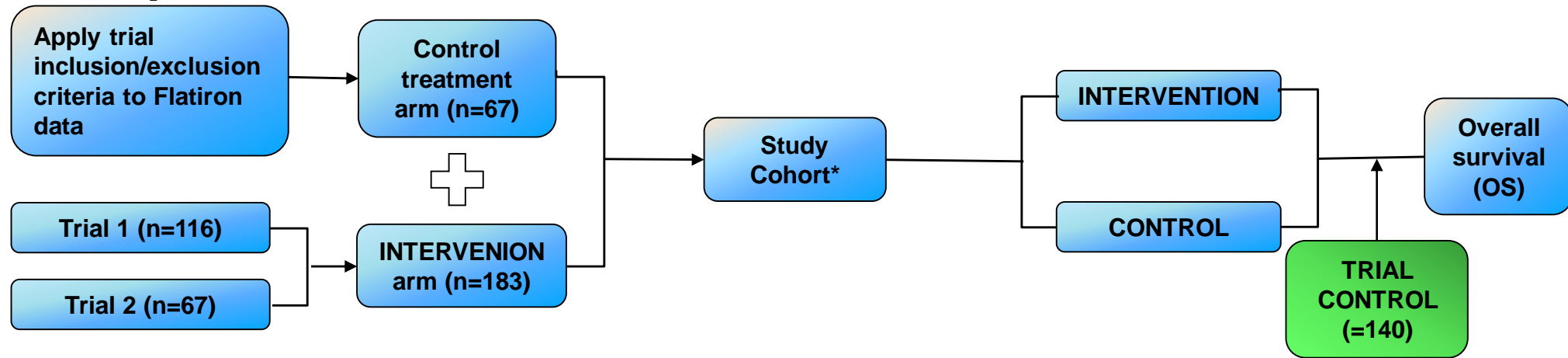

Figure 3. Survival by GenMatch analysis

NR, not reached

The observed treatment effect on the risk of death:
- IPTW cohort (I=251, C=236; HR 0.64; 95% CI 0.48–0.88)
- GenMatch (I=1,241, C=1,241; HR 0.54; 95% CI 0.48–0.62)

Note: Population sizes are larger due to weighting. Stabilized weights were used in order to have an average weight close to 1 (and pseudo sample size close to original sample size)

11

Comparative effectiveness with an external control requires additional analytical steps to address indication bias


Study Population:

- Apply trial inclusion and exclusion criteria
- IPD from three different data sources were pooled for a study population

Statistical Analysis:

- 3 adjusted comparisons were made to enable covariate balance:
  - Inverse propensity treatment weighting (IPTW)
  - Propensity score matching (PSM)
  - Genetic Matching (GenMatch)
- Assess standardized mean differences
- Overall survival was calculated from initiation of treatment to the end of follow-up and/or death
- Multivariate Cox proportional hazards model to compare OS and estimate the average treatment effect
- Comparison with TRIAL CONTROL data & Sensitivity analyses

Median OS was similar between the unadjusted RWD CONTROL cohort (15.6 months) and TRIAL CONTROL cohort (14.9 months)

Figure 4. Survival of RWD CONTROL and re-digitised TRIAL CONTROL
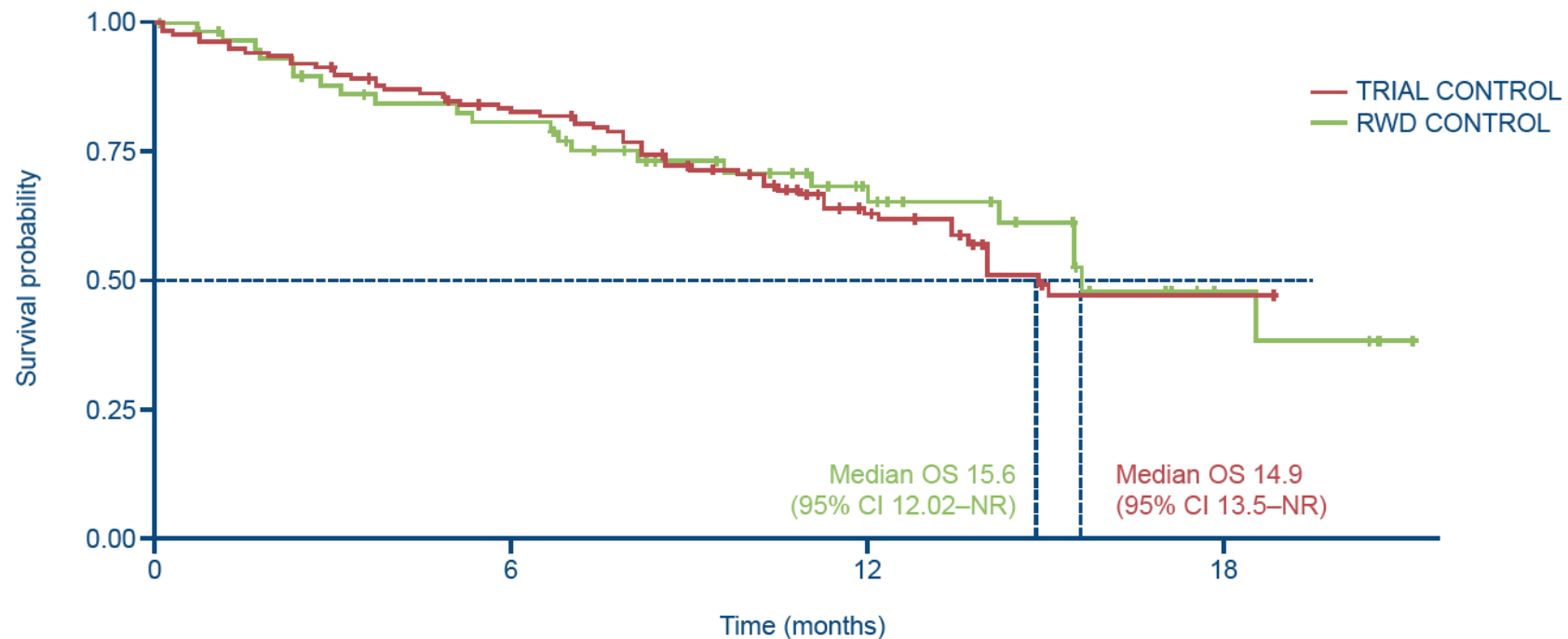
NR, not reached

Case Study Learnings

- Our results demonstrate the utility of EHR data to estimate comparative effectiveness for a single-arm trial, and as a tool to provide comparative data when randomised controlled trials are not available due to swiftly changing standards of care.

- Additional measures to reduce bias from known confounding factors should be taken when comparing data from different sources.

- The PSM analysis result did not adequately balance the two treatment cohorts. However, the other methodologies allowed us to achieve balance and therefore produce an unbiased comparison between the INTERVENTION cohort and RWD CONTROL cohort. The similarity between the TRIAL CONTROL cohort and RWD CONTROL cohort adds credibility to the results observed with the RWD TRIAL cohort comparison.

Back-up