ontada

Feasibility of Using Oncology Specific Electronic Health Records (EHR) Data to Emulate Clinical Trial Inclusion and Exclusion Criteria

Thomas W. Wilson, PhD, DrPH (presenter)

In collaboration with:

Joseph T. Dye, PhD; Sarah Spark MHA, MBA, BSN, RN; John Bian, PhD; E. Susan Amirian, PhD; Janet L. Espirito, PharmD; and Nicholas J. Robert, MD

Abstract ID: 115525

May 17, 2022

Background: RWE Potential

"Real world evidence (RWE) holds the promise of timely data at a reasonable cost, can offer large sample sizes that enable analysis of subpopulations and less-common effects, and can provide a representation of real-world practice and behaviors."

[1] ISPOR - Top 10 HEOR Trends. https://www.ispor.org/heor-resources/about-heor/top-10-heor-trends

"Real world data (RWD) hold 'great promise' for randomized clinical trials (RCT), especially for development of synthetic [external] control groups."

[2] Thorlund K, Dron L, Park JJH, Mills EJ. Synthetic and External Controls in Clinical Trials - A Primer for Researchers. Clin Epidemiol. 2020;12:457-467. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7218288/

Background: Challenge #1

RCTs often have very restrictive eligibility criteria

"Overly restrictive, and sometimes poorly justified, eligibility criteria are a *key barrier* that leads to *low enrolment in clinical trials*."

[3] Liu R, et al. Evaluating eligibility criteria of oncology trials using real-world data and AI. Nature. 2021 Apr;592(7855):629-633. https://pubmed.ncbi.nlm.nih.gov/33828294/

"Overly restrictive eligibility criteria can *limit access to clinical trials* as part of cancer care, impede enrollment thereby slowing drug development, and cause trials to be *less reflective of the patient population that will eventually use new medicines once approved*"

4] Allen J. The Future OF Medicine: Legislation to Encourage Innovation and Improve Oversight, Testimony of Jeff Allen, President & CEO Friends of Cancer Research. March 17, 2022. https://energycommerce.house.gov/sites/democrats.energycommerce.house.gov/files/documents/Witness%20Testimony_Allen_HE_2022.03.17.pdf

Background: Challenge #2

RWD in structured EHR fields alone may not meet the needs of most RCT eligibility criteria

"Of the 46 trials for which clinical inclusion and exclusion criteria were available on ClinicalTrials.gov, 2 (4%) had at least 80% of their criteria that could be routinely ascertained in EHR^ data."

^This study ascertained availability of data related to eligibility criteria in insurance claims and/or structured EHR data.

5] Wallach JD, et al. Feasibility of Using Real-world Data to Emulate Postapproval Confirmatory Clinical Trials of Therapeutic Agents Granted US Food and Drug Administration Accelerated Approval JAMA Netw Open. 2021 Nov 1;4(11):. https://pubmed.ncbi.nlm.nih.gov/34751763/

Study Objective

This study was designed to investigate whether data needed for oncology clinical trial eligibility criteria listed in clinicaltrials.gov are likely available retrospectively in EHRs – *both* structured and unstructured RWD fields.

Methods

- FDA approvals of oncology drugs in 2020 were identified [6]
- These were matched to phase III RCT trial data from the Aggregate Analysis of Clinical Trials.gov (AACT) database [7]
- AACT is a relational database that contains all protocol and results data elements about all studies registered in ClinicalTrials.gov
- The eligibility fields were parsed into different criteria based on comma placement
- All analyses were done using SAS v9.4

- Informal "Delphi panel" of experts in medicine, pharmacy, epidemiology, and chart abstraction (the authors) reviewed the criteria in two ways:
  - Created five data categories related to:
    - Demographics, Cancer-related Factors, Comorbidities, Functional Status, Trial Operations
      - 29 sub-categories
  - Judged likelihood of information needed to assess criteria being found in the EHR in:

    a) structured data or unstructured data

    b) unstructured data only

[6] ASCO: "2020 FDA Approvals of Drugs for Cancer Treatment" https://ascopost.com/issues/december-25-2020/2020-fda-approvals-of-drugs-for-cancer-treatment/
[7] https://aact.ctti-clinicaltrials.org/schema; https://aact.ctti-clinicaltrials.org/download (downloaded 12/17/2021)

Methods

Definitions:

- *Structured*: Data readily available in standardized EHR fields (e.g., dropdown menus, check boxes, or date fields for age, sex, diagnosis, stage, some biomarkers, etc.)
- *Unstructured*: Free-text data in medical record (e.g., progress notes, scanned documents, external diagnostic reports, etc.)

- Reporting:
  - Primary diagnoses
  - Descriptive statistics of criteria per trial
  - Number of trials with eligibility data available, both overall and with trial specific or operational criteria excluded :
    - *Percent of trials with 100% of the information needed to assess trial criteria available in RWD*
    - *Percent of trials with 80%^ of the information needed to assess trial criteria available in RWD*

---

^ Standard used in [5] Wallach JD, et al. Cited earlier

Methodologic Considerations

Acknowledgement of variability of structured data fields across EHR systems

Specialized EHRs (e.g., oncology-specific EHR) are likely to have different and more disease-specific structured fields than general EHRs

Retrospective RWD collected in routine medical practice

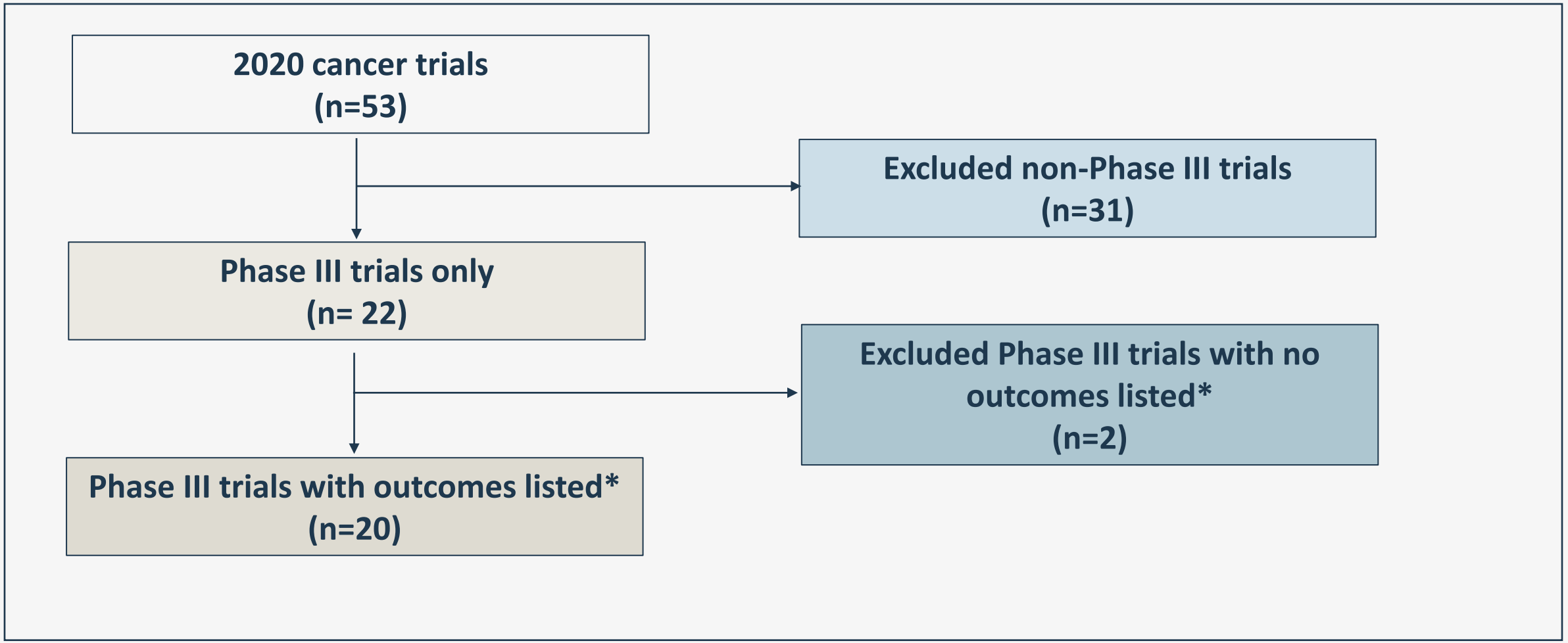
as opposed to data collected solely for research purposes

Criteria important for prospective studies may not be as relevant for retrospective data collection

E.g., Intentions, attestations, investigator discretions, prognoses, etc.

Note: Other methods for external control generation besides matching on clinical trial eligibility criteria are available (e.g., techniques using propensity scores).

Clinical Trial Consort Diagram



*Original intent was to examine RWD for specific outcomes, but this was later considered out of scope for this study.

Summary Statistics

Total of 463 eligibility criteria across 20 trials (mean ~23 criteria per study; median=21)

Type	N trials	Average	SD	Median	Min	Max	Sum
Inclusion criteria	20	10.7	7.1	8.5	5	31	213
Exclusion criteria	20	12.5	7.8	12.0	2	25	250
Total eligibility criteria	20	23.2	12.0	21.0	7	51	463

*Note on range (min to max) of 7-51 criteria per trial:
Items listed in the eligibility field of AACT database are not necessarily an exhaustive list for each trial.

Primary Cancer Diagnoses in 20 Trials

Condition	Number of Trials	Percent of Trials
NSCLC	4	20%
Leukemia	3	15%
Breast Cancer	2	10%
Myeloma	2	10%
Renal Carcinoma, Transitional Cell	1	5%
Colorectal Cancer	1	5%
Hepatocellular Cancer	1	5%
Melanoma	1	5%
Mesothelioma	1	5%
Myelodysplastic Syndromes	1	5%
Ovarian Cancer	1	5%
Prostate Cancer	1	5%
SCLC	1	5%
Total	20	100%

Data Categories

Data Categories (n=5) and Sub-categories (n=29) per expert panel

DEMOGRAPHIC	CANCER RELATED	COMORBIDITY RELATED	FUNCTIONAL STATUS	TRIAL SPECIFIC*
Age & gender	Biomarker or Genetic Marker	Cardiovascular Status	Hematologic Status	Life Expectancy
	Cancer Stage	Comorbidity	Hepatic Status	Donor Status
	Cancer Type	Concomitant Medications	HIV Status	Trial Operations
	Measurable Disease	Contraindication	Immune Status	
	Metastasis	Drug-Drug / Food Interaction	Ocular Status	
	Other Cancer	Hypersensitivity	Performance Status	
	Prior Cancer Treatments/ Procedures	Infection	Pulmonary Status	
	Progressive Disease	Organ Function	Renal Status	
			Reproductive Status	

*Trial specific are those items considered to be low relevance for retrospectively identified external controls.

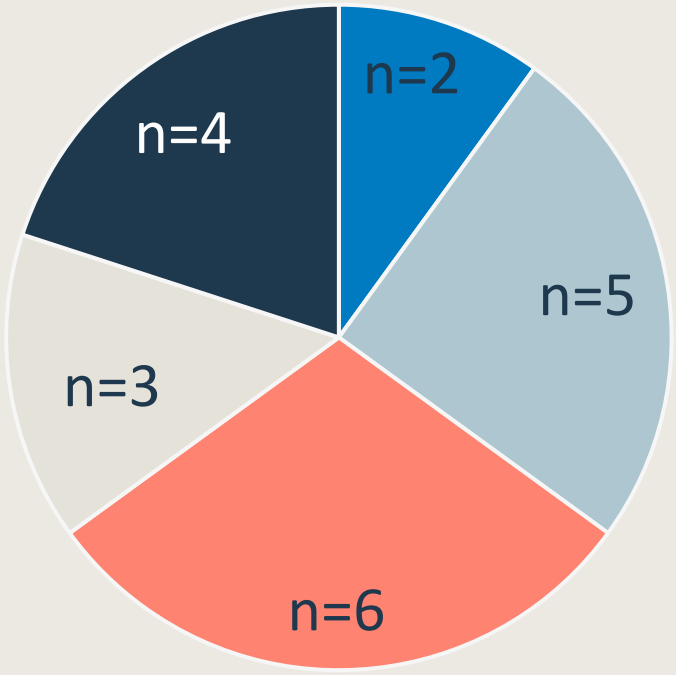
Data Criteria Examples
(of 463)

Demographic	Male or female >55 years of age
Cancer Related	Newly diagnosed AML
Functional Status Related	ECOG PS: 1, 2, or 3
Comorbidity Related	Significant history of CVD
Trial Specific	Donor matching requirements Projected life expectancy Negative pregnancy test Willing and able to comply with scheduled visit and study procedures

Results: Availability of Oncology Trial Eligibility Criteria in RWD

Percent of Trial Criteria Potentially Assessable through RWD:

- < 80%
- 80% to < 85%
- 85% to <90%
- 90% to <95%
- 95% to <100%
- 100%



Total Criteria: n=463

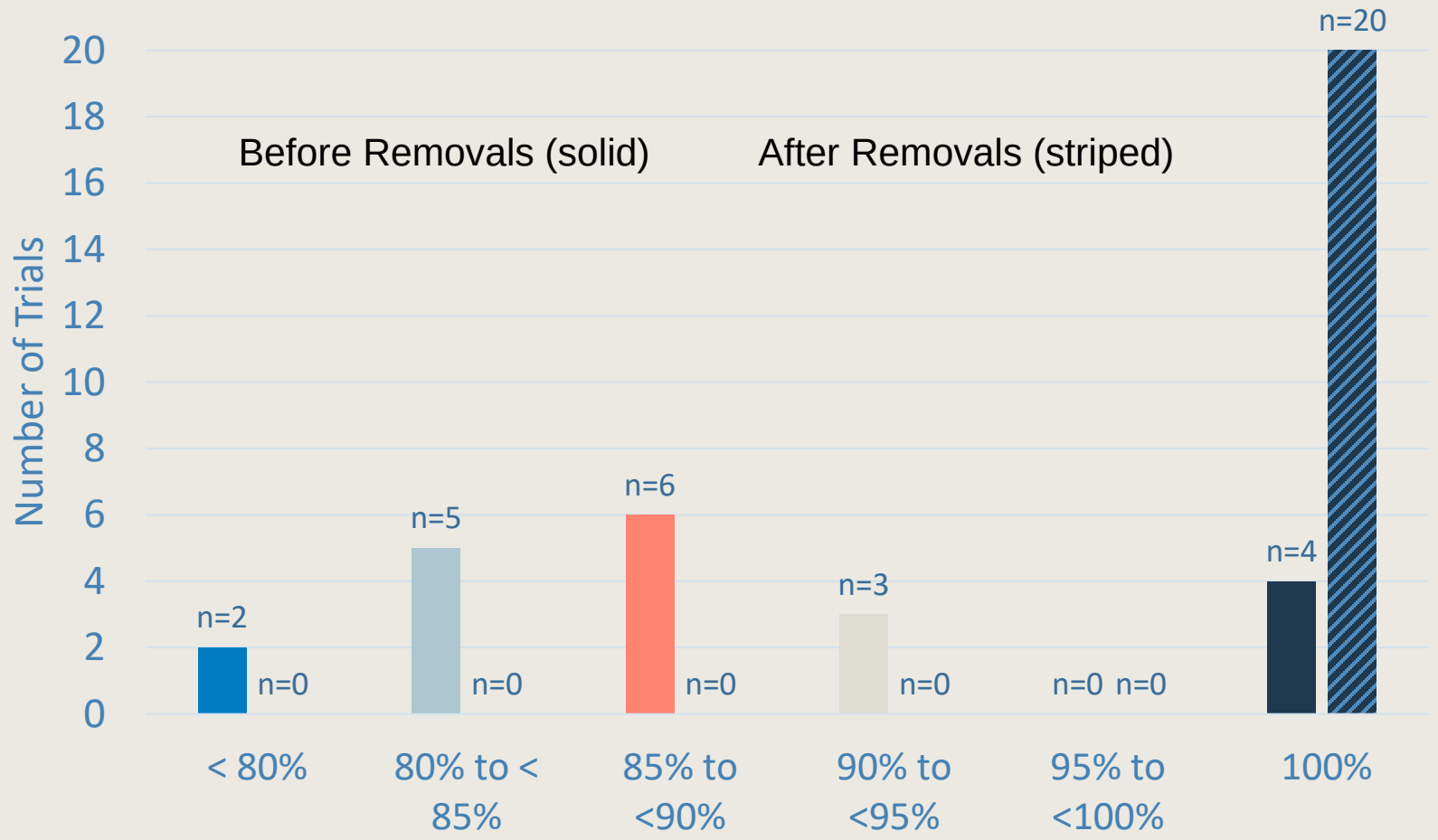
4 of 20 (20%) [dark blue only] of the trials would likely have met the 100% threshold of finding all eligibility criteria in RWD

18 of 20 (90%) met the 80% threshold [all except light blue]

Examples of Trial-Specific Criteria classified as "*Likely Not Available*" in EHR
(structured or unstructured data)

Classification (Total=67: 14% of total)	Examples
Trial operations (n=55)	• Agreement to remain abstinent • Fertile patients must use contraception • Must refrain from donating blood • Ability to comply with study protocol, in investigator's judgement • Have available tumor for central lab analysis
Donor matching (n=8)	• All available first-degree relatives must be HLA typed • No syngenic donors
Life expectancy (n=4)	• Patient must have a projected life expectancy of at least 12 weeks

Results: Comparison of Before and After Removal of Trial Operational Criteria



20 out of 20 would likely meet 100% threshold of finding all eligibility criteria

67 criteria removed;

296 criteria remained

Summary of Results

- Among the 53 trials identified, 20 were Phase III trials and had information required for inclusion in our study (i.e., outcomes available in AACT).
- A total of 463 eligibility criteria was found in the 20 trials.
- The median number of criteria per trial was 21 (range: 7-51). Median inclusion was 8.5 (range: 5-31) and median exclusion was 12.0 (range: 2-25).
- Overall, information needed to assess all trial eligibility criteria were judged likely to be found in RWD for 4 out of 20 trials (20%), while ≥ 80% of the eligibility criteria were likely to be found for 90% of the trials (18 of 20).
- *Removing trial-specific criteria, information needed to assess all remaining eligibility criteria were likely to be found in RWD for 20 of the 20 studies (100%).*

Conclusion

Using both *structured and unstructured* RWD in generating external controls for trials has much higher *potential* yield than reported for EHR *structured-only* data in prior reports.

A significant number of eligibility criteria in the AACT database were deemed by the expert panel to be relevant only to prospective operations of a clinical trial and were not considered relevant to generate retrospective external controls from real-world data.

Removing these operational and logistical trial criteria from consideration improved the potential yield rate to 100% in our study.

Discussion

The combination of structured and unstructured real-world data collected in the day-to-day operations of a community oncology center can serve as a valuable asset for the creation of external controls.

We note that a potential to be found in an EHR, does not necessarily mean it will be found (as noted in the literature).

The number of eligibility criteria is inversely proportional to the actual yield rate in real-world data.

A reduction in the number of eligibility criteria –particularly those that do not substantially impact internal validity of the trial– may improve the feasibility of using RWD to identify external controls and may enhance external validity.

Question: Is it “fit for purpose” for RCT only or for both RCT & RWE?

Differentiating trial specific operation criteria only “fit for purpose” for RCT will facilitate construction of external controls in retrospective data.