Evaluating Patient Count Vs Hospitalization Risk for Common Clinical Trial Eligibility Criteria: A Case Study for Relapsed/Refractory Lymphoma/Leukemia

James R. Rogers[1]; Casey N. Ta[1]; Cong Liu[1]; Ali Soroush[1,2]; Ying Kuen Cheung[3]; George Hripcsak[1,4]; Chunhua Weng[1]

[1]Department of Biomedical Informatics, Columbia University, New York, NY, USA; [2]Division of Gastroenterology, Department of Medicine, Columbia University Irving Medical Center, New York, NY, USA; [3]Department of Biostatistics, Columbia University, New York, NY, USA; [4]Medical Informatics Services, New York-Presbyterian Hospital, New York, NY, USA

Background

Clinical trials are important for generating medical evidence. An essential section of a clinical trial protocol is eligibility criteria, which define the study population. Although these criteria are assumed to be chosen for safety precautions and to minimize study population heterogeneity, explicit criteria justification is typically minimal.[1,2] A result of poorly selected criteria is unnecessarily limiting the number of patients that could benefit from trial participation. To explore the extent of this concern, one approach is to apply common eligibility criteria (CEC) extracted from a series of clinical trials for a particular disease to existing electronic health record (EHR) data and then assess the effects of the criteria on available patient counts and safety events. Using a convenience sample of trials for adult relapsed/refractory (r/r) lymphoma/leukemia trials conducted at one large academic medical center, the objective of this study is to assess the tradeoff in patient count vs hospitalization risk when using different CEC sets.

Methods

The following procedure was applied: (1) extract and identify CEC from an available sample of trials conducted at Columbia University Irving Medical Center (CUIMC); (2) build cohort pairs; (3) analyze effects of CEC. For extraction and identification, we utilized a database storing all eligibility criteria from all ClinicalTrial.gov entries; database contents were populated by automating a natural language processing tool that encodes criteria into OMOP standardized concepts[3]. From our trial sample, we focused on concepts that appeared in at least 25% of our trials. Criteria were reviewed, consolidated and selected if they could be reasonably represented in EHR data. Each individual criterion selected in this way is a CEC, while a CEC set is a unique combination of these CEC.

After CEC selection, cohort pairs were built using CUIMC EHR data stored in the OMOP CDM v05 format (**Figure 1**). All cohorts started from an initial group of r/r lymphoma/leukemia adult patients (i.e., baseline), with entry predicated on a therapeutic treatment for lymphoma/leukemia (e.g., one portion of a chemotherapy regimen) occurring within a condition era for lymphoma/leukemia; treatments were selected based on clinical input and UpToDate resources[4]. At least two diagnoses for lymphoma/leukemia had to exist prior to entry, with the latest prior era's end date occurring at least 180 days before the condition era containing entry. Consecutive condition eras occurring less than 180 days from one another were combined. Calendar time for cohort entry was bounded by the earliest and latest recruitment dates from our trial sample. From baseline setup, cohort pairs for analysis were created based on CEC sets, with each pair having two mutually exclusive cohorts: a qualifying cohort, defined by qualification of the chosen CEC set, and a non-qualifying cohort. When possible, validated phenotypes were used for CEC representation. The outcome of interest was all-cause hospitalization (with length of stay > 1 day), with censoring events being end of follow-up (180 days) or end of patient data. Hospitalization was chosen

because it is considered a serious adverse event and is well captured in EHR data[5]. Because of this, patient entries that were inpatient were excluded. Likewise, if the hospitalization event was related to a hematopoietic cell transplantation procedure (within 180 days), the record was also excluded.
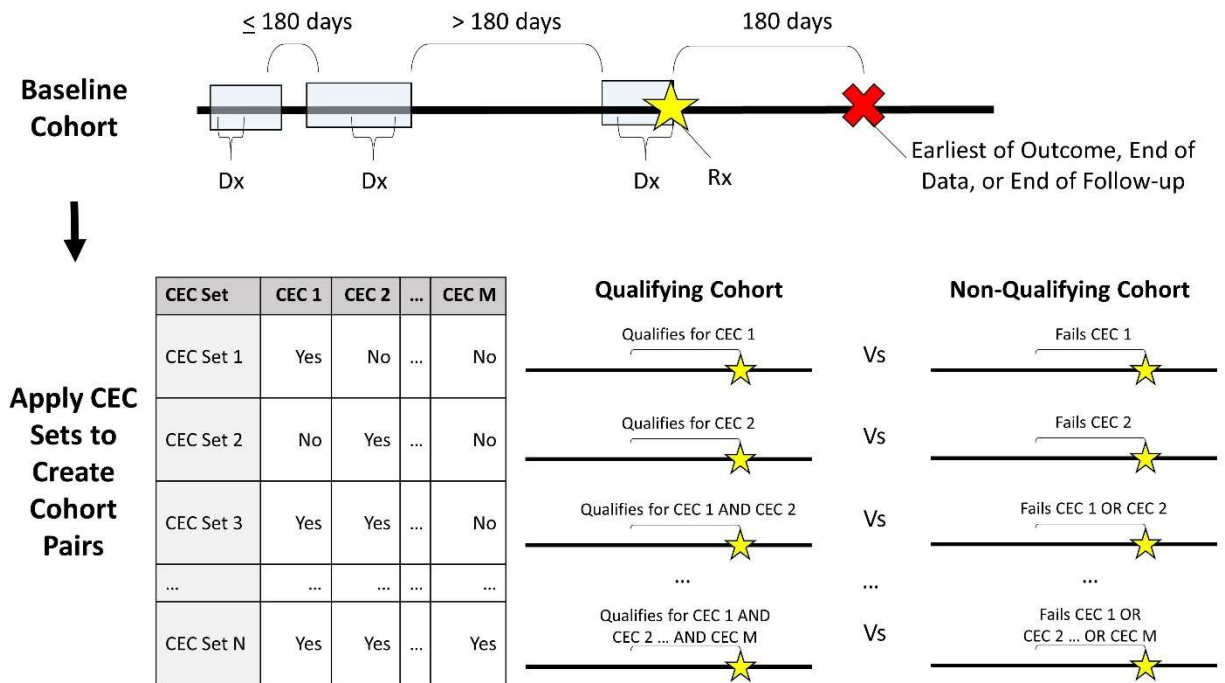
Figure 1. Basic cohort setup and construction of cohort pairs per CEC sets, with star representing the index, blue boxes representing condition eras, and the red "X" representing an endpoint: Dx = diagnosis for lymphoma/leukemia; Rx = medication treatment for lymphoma/leukemia; CEC = common eligibility criteria; HCT = hematopoietic cell transplantation

Analysis was driven by patient count changes and estimations of hospitalization risk for each cohort pair. Patient counts were based on number of patients in the qualifying cohort. Risk of hospitalization was based on a hazard ratio (HR) between cohort pairs. Because some CEC sets might be underpowered, only cohort pairs with both cohorts meeting the minimum sample size to detect a change of 10% in the proportion of hospitalizations at 80% power and alpha of 0.05 were included. To visualize tradeoffs between qualifying cohort size and HRs, a scatterplot with k-means clustering analysis was applied to examine CEC patterns.

Results

Our sample contained 23 r/r lymphoma/leukemia trials, with 623 available patients at baseline. From our trials, we identified 9 CEC (**Table 1**). From an individual effect standpoint, no prior malignancy led to the greatest change in available patients (only 50% of the baseline qualified). The 9 CEC led to 511 possible CEC sets (the empty set of no CEC is excluded); only 256 (50%) were powered.

CEC Label	CEC Description	Number of Trials N (%)	Patient Count N (%)
Start	N/A	23 (100)	623 (100)
No HIV	No HIV within the past 365 days	20 (86.96)	614 (98.56)
No HBV/HCV	No HBV/HCV within the past 365 days	19 (82.61)	613 (98.39)
Not pregnant	No evidence of current pregnancy within the past 60 days	19 (82.61)	622 (99.84)
No prior chemo/rad	No prior chemotherapy or radiotherapy within the past 14 days (excludes index)	18 (78.26)	590 (94.70)

No prior malignancy	No prior malignancy (beside lymphoma, leukemia, non-melanoma skin cancer, melanoma in situ, carcinoma in situ of the cervix, benign tumor, or lipomatous tumor) within the past 1095 days	17 (73.91)	313 (50.24)
Adequate eGFR	Most recent eGFR measure within the past 180 days > 30 mL/min/1.73m^2 (per MDRD equation)	11 (47.83)	525 (84.27)
No infection	No active infection within the past 30 days	10 (43.48)	604 (96.95)
Adequate ANC	Most recent ANC measure within the past 180 days > 1000/mm3	9 (39.13)	587 (94.22)
No corticosteroid	No prior corticosteroid use within the past 7 days (excludes index)	9 (39.13)	612 (98.23)

Table 1. Individual CEC found with brief description, number of trials, and number of qualifying patients from baseline; *if no lab value was present, patients were assumed to have an adequate value and thus were not excluded. CEC = common eligibility criteria

Figure 2 displays a scatterplot of all 256 powered CEC sets, with the x-axis representing the number of patients in the qualifying cohort and the y-axis representing the risk of hospitalization for those qualifying vs those non-qualifying. Using k-means clustering, characteristics of 6 clusters (represented by the colors) were found. All clusters applied no prior malignancy. CEC sets within Cluster 5 generally resulted in the lowest risk of hospitalization, but lowest available amount of patients. When comparing Cluster 5 to the other clusters: all of them apply no infection (in contrast to Clusters 1, 4, and 6), while the majority of CEC sets in Cluster 5 apply no prior chemotherapy/radiotherapy (in contrast to Clusters 2 and 3).
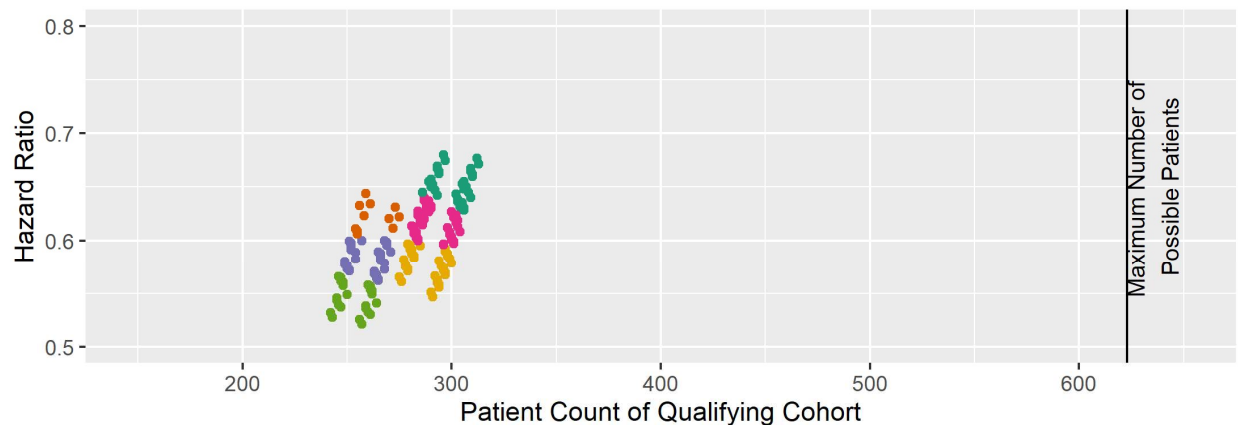
	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
N of CEC Sets	33	22	58	52	48	43
No HIV	8 (24%)	2 (9%)	22 (38%)	22 (42%)	40 (83%)	34 (79%)
No HBV/HCV	12 (36%)	12 (55%)	28 (48%)	25 (48%)	24 (50%)	27 (63%)
Not pregnant	17 (52%)	11 (50%)	29 (50%)	27 (52%)	24 (50%)	20 (47%)
No prior chemo/rad	0 (0%)	0 (0%)	24 (41%)	24 (46%)	40 (83%)	40 (93%)
No prior malignancy	33 (100%)	22 (100%)	58 (100%)	52 (100%)	48 (100%)	43 (100%)
Adequate eGFR	13 (39%)	8 (36%)	24 (41%)	25 (48%)	32 (67%)	26 (60%)
No infection	0 (0%)	22 (100%)	58 (100%)	0 (0%)	48 (100%)	0 (0%)
Adequate ANC	9 (27%)	4 (18%)	28 (48%)	29 (56%)	32 (67%)	26 (60%)
No corticosteroid	13 (39%)	14 (64%)	26 (45%)	32 (62%)	24 (50%)	19 (44%)

Figure 2. Scatter plot of patient counts of the qualifying cohort vs hospitalization risk for CEC sets that are appropriately powered. Each dot represents a CEC set applied while colors represent a cluster found from k-means clustering. The table provides a breakdown of how many CEC sets within each cluster has a specific CEC applied. CEC = common eligibility criteria

Conclusion

For our sample, forgoing no prior malignancy CEC led to the largest number of patients available, but led to underpowered CEC sets – of note, the phenotype for this criterion might be too inclusive as some cancer diagnoses are likely either: (1) differential diagnoses; or (2) specified by body location not readily confirmed as r/r lymphoma or leukemia until later in the record. Minimization of hospitalization risk was primarily tied to the no infection criterion and no prior chemotherapy/radiotherapy (i.e. Cluster 5). Requiring no infection likely excludes sicker patients while requiring no prior chemotherapy/radiotherapy might be removing those that had an adverse event leading to hospitalization. The remaining criteria should not necessarily be discounted despite their muted effect in our sample – specifically, a different sampling environment may have found more individuals with these characteristics, and thus may have led to different CEC patterns. This procedure can be promising for other diseases, although different criteria are likely to emerge and relevant patients will likely have different comorbidity patterns. Regardless, this study demonstrates a possible approach to assessing the tradeoff between patient counts and hospitalization risk to strengthen criteria justification.

References/Citations

1. Jin S, Pazdur R, Sridhara R. Re-Evaluating Eligibility Criteria for Oncology Clinical Trials: Analysis of Investigational New Drug Applications in 2015. J Clin Oncol. 2017 Nov 20;35(33):3745–52.

2. Van Spall HGC, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. JAMA. 2007 Mar 21;297(11):1233–40.

3. Liu H, Chi Y, Butler A, Sun Y, Weng C. A knowledge base of clinical trial eligibility criteria. Journal of Biomedical Informatics. 2021 May 1;117:103771.

4. About Us [Internet]. UpToDate. [cited 2021 Jun 1]. Available from: https://www.uptodate.com/home/about-us

5. What is a Serious Adverse Event? | FDA [Internet]. [cited 2021 May 2]. Available from: https://www.fda.gov/safety/reporting-serious-problems-fda/what-serious-adverse-event