PCN246
A SYSTEMATIC APPROACH FOR QUANTIFYING HETEROGENEITY WITHIN CLINICAL TRIAL POPULATIONS

Ransom, Joshua F PhD[1];Berger, Marc MD[1]; Galaznik, Aaron MD MBA[1]; Lempernesse, Bruno MS[1]; Shilnikova, Alexandra[1]; Rusli, Emelly, MPH[1]

SHYFT Analytics, a Medidata Company, Boston, MA, USA

Introduction

Heterogeneity in clinical trial populations can contribute to variability in observed treatment effect. According to the Cochrane Handbook, sources can be either clinical or methodological diversity.[1] Assessing the impact of heterogeneity in observed intervention effects is an important part of interpreting trial results, whether determining the presence of unanticipated clinical subgroups in a given trial or designing pooled analyses of trial results, even for those with ostensibly similar methodologies. Two proven quantitative approaches to addressing heterogeneity include use of clustering algorithms and of propensity score modeling techniques.[2,3]

We propose here a quantitative, systematic, scalable approach to assessing clinical trial population heterogeneity, leveraging both approaches, to estimate the likely presence and impact of heterogeneity on observed trial results, and guide future analysis.

Methods

Analysis was conducted in a pooled dataset of 7 clinical trials (n=719) for relapsed/refractory AML, conducted within the past 5 years, from the Medidata MEDS archive of >3000 trials.[4]

Clinical Outcome Heterogeneity Assessment

To quantify the baseline level of heterogeneity in skewness in clinical outcomes, two ratios were adapted from the work of Kent et al, the extreme quartile risk ratio (EQRR, ratio of outcomes rates in lowest quartile to the highest), and median-to-mean risk ration (MMRR, ratio of median outcome to mean outcome).[5]

- Prior analysis within this cohort demonstrated a correlation between treatment duration and Complete Remission (CR).[4]
- To facilitate comparison with real-world data sources, time to progression was approximated as time on treatment for patients receiving first line therapy and Time-to-next-Treatment for those progressing to second line.

Treatment Propensity Heterogeneity Assessment

- Based on NCCN guidelines for AML, top 2 treatment choices (**cytarabine, daunorubicin**) were identified by patient count and propensity scores for treatment receipt calculated for individual respondents.[6]
- Frequency distributions of propensity scores were generated, and percent area of overlap between distributions was calculated.

Population Cluster Heterogeneity Assessment

- K-nearest neighbor classification was conducted. Covariates included age, transfusion frequency, G-CSF administration, infection frequency (ICD9 codes 001-139), and time on first-line treatment.
  - As a starting point, k was defined as √n.
  - Number and size of potential patient clusters was calculated.
- As the purpose was to identify potential classifications, no testing/training split was employed, and advanced feature reduction was not required.

Real-World Data Replication

To demonstrate applicability of this approach to real world data sources, the above methodology was also applied to a US medical claims source(s) from HealthVerity® Marketplace platform of data suppliers from 2/1/2014 – 12/31/2018.

- Data Transformation and Analysis
  - Data was transformed into the OMOP Common Data Model, version 5
  - Analyses were conducted using the SHYFT Quantum V6.7.0 solution
- Inclusion criteria
  - Patients with ≥1 AML diagnosis (ICD-10: C92.0, C92.4, C92.5, C92.6, C92.A, or ICD-9: 250.0) with evidence of treatment with any medical claims-reimbursed oncologic agent (i.e., injectable, infusible agents)
  - **Index date: first treatment administration code within observation period**
  - Age ≥18 at index
  - ≥12-month pre- and post-index continuous enrollment
  - Exclusion Criteria:
    - Evidence of other malignancy
- Clinical Outcome Heterogeneity Assessment
  - Given the lack of response information available, time-to-treatment-discontinuation (TTD) and time-to-next-treatment (TTNT) from index date were used to approximate clinical response
    - New line of therapy was defined as a gap >60 days between treatment administrations
    - Patients with last date of treatment <60 days before end of observation period are censored
  - EQRR and MMRR was calculated using quartile averages of time on therapy (TTNT for patients progressing to 2nd line and TTD for patients receiving first line only) in lieu of risk ratios
- Propensity score distribution assessment
  - Propensity score distribution overlap was calculated using top 2 treatments in each data source (**Azacitidine, Decitabine**)
- Clustering Assessment
  - K-nearest neighbor classification was conducted with covariates included, matching those from clinical trial data analysis

Figure 1: Attrition Table

Claims dataset

Total patients in the Claims data (N=5,010)	Exposure to Azacitidine or Decitabine during study period (N=2,031)	Age >= 18 (N=2,030)	6 months continuous enrollment pre- and post-index (N=644)	Patients included in the analysis (N=644)

Clinical trial dataset

Total patients in the Clinical Trial data (N=719)	Exposure to Cytarabine or Daunorubicin during study period (N=95)	Age >= 18 (N=83)	Patients included in the analysis (N=83)

Table 1: Baseline demographics

		Clinical Trial Data	HealthVerity® Claims
Age At Index	Mean (SD)	66 (28.6)	72.9 (7)
	Median	69.0	73
Gender (N, %)	FEMALE	31 (37%)	271 (42%)
	MALE	52 (63%)	359 (56%)
	Unknown	0 (0%)	14 (2%)
Patients in 1st line of therapy only	N (%)	76 (92%)	391 (61%)
Patients progress to the 2nd line of therapy	N (%)	65 (78%)	48 (7%)
Time to discontinuation — 1st line of therapy	Mean (SD)	48.43 (37.91)	85.2 (94.6)
	Median	36.5	59
Time to discontinuation — 2nd line of therapy	Mean (SD)	694.2 (548.5)	315.8 (187.1)
	Median	605	273
Count of Infections at anytime during study period	Mean (SD)	1 (1.4)	8.32 (183)
	Median	1	2.00
Count of Blood Transfusion at anytime during study period	Mean (SD)	7.1 (16.7)	23.3 (38.7)
	Median	0	10
Length of 1st line therapy	Mean (SD)	48.4 (37.9)	85.2 (94.6)
	Median	36.5	59
Length of 2nd line therapy	Mean (SD)	64.5 (71.7)	72.7 (59.4)
	Median	38	48

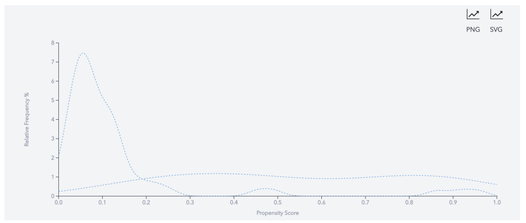
Table 2: Clinical Outcome Heterogeneity Assessment

	Clinical trial MEDS	HealthVerity® Claims	Kent et al. (Median)
EQRR	3.07	5.3	4.3
MMRR	0.87	0.69	0.86

Methods (cont')

Figure 2: Propensity Score Distribution (by Age at Index and Gender)

Clinical trial MEDS (Cytarabine and Daunorubicin)



HealthVerity® Claims (Azacitidine and Decitabine)

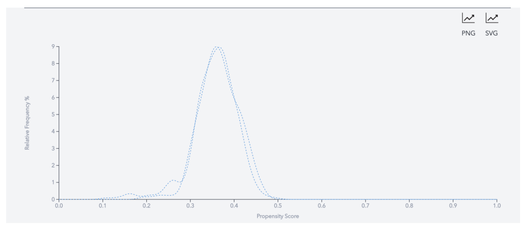


Figure 3: Propensity Matching Balance - Azacitidine and Decitabine (by Age at Index and Gender)

	Mean: treated	Standard deviation: treated	Mean: untreated	Standard deviation: untreated	Difference of means	Standardized difference of means
Age at Index	73.9	5.8	72.4	7.6	1.5	0.2

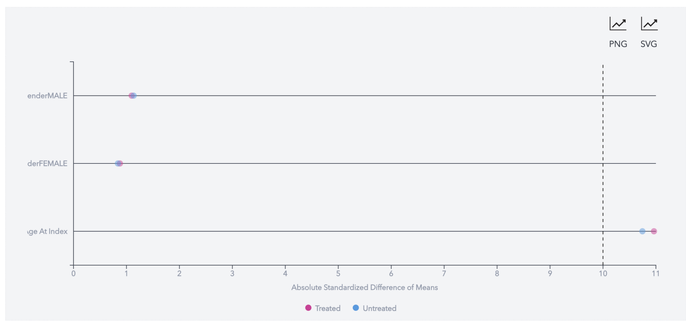


Table 3: Clustering Assessment

Clinical trial MEDS (K=3)

		Age At Index	Days In Line 1	Count of Blood transfusion at anytime during study period	Count of G-CSF at anytime during study period	Count of infections at anytime during study period	Size	Within SS
Cluster	1	37.4	50.4	40.1	5.7	1.7	11	15,262
	2	70.4	33.3	2.7	0.3	0.8	54	48,184
	3	84.3	120.7	0	0	1.1	11	30,637

HealthVerity® Claims (K=3)

		Age At Index	Days In Line 1	Count of Blood transfusion at anytime during study period	Count of G-CSF at anytime during study period	Count of infections at anytime during study period	Size	Within SS
Cluster	1	72.9	35.3	23	5	8.9	247	590,046
	2	74	373.2	13	3.6	4.7	26	275,747
	3	73.8	126.1	31.4	2.9	10.7	118	581,798

Results

With respect to Clinical Outcome Heterogeneity, EQRR and MMRR were relatively low, at 3.07 and 0.87 respectively (**Table 2**). For Treatment Propensity, the distribution overlap between the two most frequent treatments (cytarabine- based regimens and daunorubicin-treated patients) was relatively low (**Figure 2**). For Population Heterogeneity, using a cutoff of cluster sizes of at least 5%, K-means indicated presence of 3 clusters, representing 14%, 22%, and 63% of assessable patients. Although the number of clustering attributes was limited, drivers appear related to differences in infection rate and hematologic recovery.

Overall, when compared to the real-world assessment in insurance claims data, results were similar. For Clinical Outcome Heterogeneity, EQRR was slightly higher at 5.3, with slightly lower MMRR at 0.69. For both datasets, although calculation methodology differed slightly, results were comparable to the average trial ratios from Kent, et. Al. Treatment Propensity assessment found overlap between the two most frequent treatments to be >90%. For Population heterogeneity, K-means, using same criteria, found 3 clusters, representing 63%, 7%, and 30% of assessable patients, respectively.

Conclusion

We outline here an approach to rapidly assess, both visually and quantitatively, heterogeneity due to differences in clinical outcome, treatment assignation propensity, and underlying population characteristics. This can greatly facilitate planning the use of advanced analytics to correct for underlying heterogeneity particularly in populations where heterogeneity may not be anticipated. Standardized and routine estimation of study population heterogeneity has potential application for sub-population identification, clinical trial simulation, and cross-study comparison where differences in intervention effects may be present.

This study also demonstrated applicability of this approach to real-world data cohorts (with some modification due to less availability of clinical detail in the claims data). While further research is needed, the initial similarity in results indicate relatively low levels of heterogeneity in the pooled clinical trial and real-world data cohorts, although one may need to consider treatment propensity, depending on the regimens being assessed.

Next steps include expansion of this assessment to explore heterogeneity of other clinical endpoints (Response, PFS, OS), propensity score distributions for additional treatments, incorporate further patient demographic and comorbidity characteristics into clustering assessments. Application of this approach to additional diseases, in both clinical trial and real-world datasets, will further understanding of drivers of patient and disease heterogeneity.

References

1. Higgins JPT, Green S. 9.5.1 What is Heterogeneity? Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0. March 2011. https://handbook-5-1.cochrane.org/chapter_9/9_5_1_what_is_heterogeneity.htm, downloaded 1/7/2019
2. Shilling D et al. Serum proteomics reveals distinct subtypes associated with treatment response in idiopathic multicentric Castleman disease. Poster presented at American Society of Hematology, San Diego, CA., December 2018.
3. D`Agostino RB. Tutorial in Biostatistics: Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Statist. Med. 1998, 17, 2265-2281.
4. Berry DA et al. Creating a Synthetic Control Arm™ from previous clinical trials: Application to establishing early end points as indicators of overall survival in acute myeloid leukemia (AML). Journal of Clinical Oncology 2017, 35:15_supplSynthetic Control Arm™ , 7021-7021
5. Kent DM et al. Risk and treatment effect heterogeneity: re-analysis of individual participant data from 32 large clinical trials. Int J of Epidemiology, 2016, 45(6):2075-2088
6. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines): Acute Myeloid Leukemia. Version 3.2018 – November 30, 2018.https://www.nccn.org/professionals/physician_gls/pdf/aml.pdf, downloaded 1/7/2019