Toxicity Cluster--Complexity Reduction and Approximation Through Non-linear Principal Component Analysis With Multivariate Analysis with Optimal scaling

Satheesh kumar Poolakkad Sankaran, DDS, MDS, Pg Diploma, ICRT, MSc, MMSc; Minu Ponnamma Mohan, DDS, MDS, MPH; Stephen Sonis, DMD, DMSc; Sreelakshmi Panginikkode, MD
Venu Gopalakrishnan, MD; Roberto Pili, MD;.
Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, NY, USA; spoolakk@buffalo.edu

Abstract

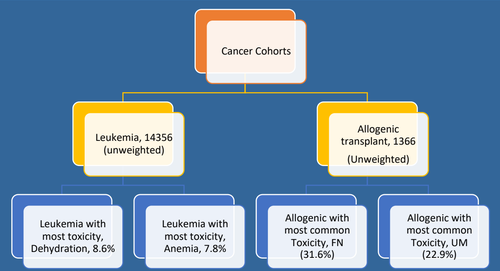
Cancer treatment-induced toxicities (CIT) are a comprehensive indicator of health outcome status among cancer patients and are valuable in informing treatment options for complex clusters. This study sought to determine the appropriateness of identifying CIT clusters utilizing a leukemia cohort from a large inpatient population treated with chemotherapy and hematopoietic stem cell transplants (HSCT).
We used US national inpatient data and applied appropriate weight to explore the original patient numbers. We applied the principal component analysis (PCA) with Multivariate Analysis with Optimal Scaling (MVAOS) method. MVAOS allows the categorical transformations computed concurrently with the PCA to maximize the variance explained by each component. PCA was employed to identify significant principal components from each treatment cohort, and we weighted each contributing variable. To produce CIT scores, variable-specific factor scores were applied to standardized Allogenic HSCT and Chemotherapy values . Weight - explanation
NIS are provided with numbers (unweighted) vs. Original numbers (weighted) we use weights to get original numbers in all our analysis.
We are using further analysis classifying allogenic HSCT and chemotherapy cases/numbers and assigning numbers to it. (i.e., converting numerical values to logical and Nominal categories of the MVOAS)
1,658,134 leukemia patients, with a mean [SD] age 57.8 [20.4]. 69.9% were Whites, 12.9% Blacks, 11.2% Hispanics, and 5.9% Others. 71,780 toxicities with oral, gastric (GI), dermatologic, and constitutional symptoms were reported. MVOS-PCA (Fig 1) demonstrated positive correlation. Principle component analysis 1 (PC1; Fig 2) of GI toxicities shows positive correlation, whereas PC2 (Fig 2) contrasts anorexia, nausea, constipation, and diarrhea (positive PC2) with distension-bloating, and dyspepsia as not being correlated in leukemia cohorts. The proportion of transformed variance explained (Cumulative Variance Accounted For) by each PC, in this case (GI), was 18% for PC1 and 36% for PC2.
Conclusion: The application of MVOS-PCA to a large claims database successfully identified CIT clusters and confirmed the clinical observation that patients suffer multiple toxicities, often path biologically related, simultaneously. And this innovation enhance the adaptability of research and planning for supportive care, which is crucial for determining the optimal treatment approach for CIT.
Need to define PC1.

Objectives

Cancer therapy-induced toxicities (CIT) serve as a complete measure of the health condition of cancer patients and provide significant information for determining treatment options for complicated groups of symptoms. This study aimed to assess the suitability of finding CIT clusters using a leukemia cohort from a large population of hospitalized patients who had chemotherapy and hematopoietic stem cell transplantation (HSCT).

Methods

We used US national inpatient data and applied appropriate weight to explore the original patient numbers. We applied the principal component analysis (PCA) with Multivariate Analysis with Optimal Scaling (MVAOS) method. MVAOS allows the categorical transformations computed concurrently with the PCA to maximize the variance explained by each component. PCA was employed to identify significant principal components from each treatment cohort, and we weighted each contributing variable. To produce CIT scores, variable-specific factor scores were applied to standardized Allogenic HSCT and Chemotherapy values.
NIS are provided with numbers (unweighted) vs. Original numbers (weighted) we use weights to get original numbers in all our analysis.
We are using further analysis classifying allogenic HSCT and chemotherapy cases/numbers and assigning numbers to it. (i.e., converting numerical values to logical and Nominal categories of the MVOAS). We've taken the concept of PCA and applied it to categorical variables. That is, converting a categorical variable into a smaller number of numerical variables, ranked by the amount of variability explained. For this to operate, the original variables had to be converted to a numerical scale with Multivariate Analysis with Optimal Scaling (MVAOS) method (De Leeuw et al., 2016).
MVAOS approaches quantify categorical values with real numbers, allowing PCA to be applied subsequently. However, as indicated in the sources, the transformations are computed concurrently with the PCA to maximize the variance explained by each component in turn.



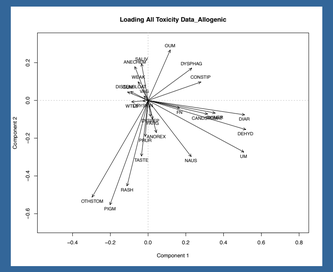


Figure 1. Toxicities from Allogenic cohort

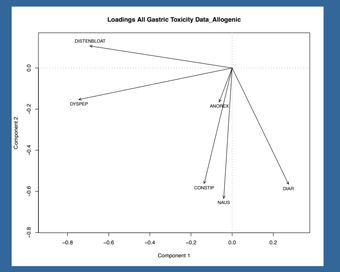


Figure 2. Gastro-toxicities from Allogenic cohort

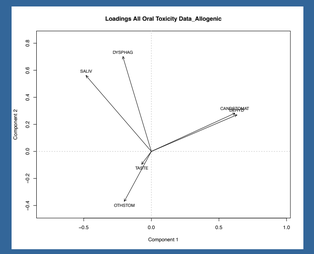


Figure 3. Oral-toxicities from Allogenic cohort

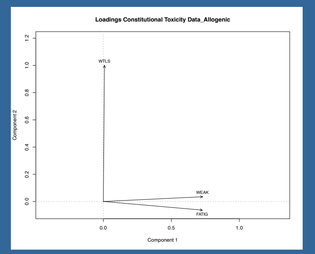


Figure 4. Constit-toxicities from Allogenic cohort

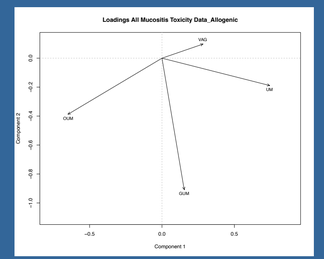


Figure 5. Mucositis toxicities from Allogenic cohort

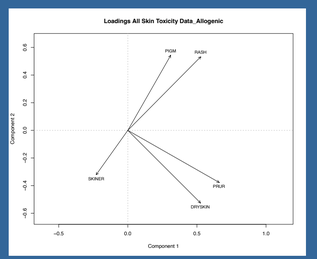


Figure 6. Skin-toxicities from Allogenic cohort

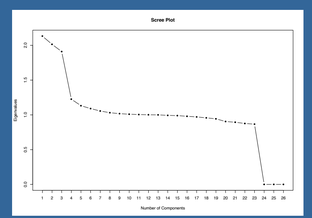


Figure 7 . Scree plot.

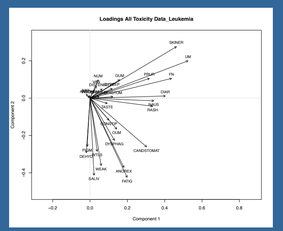


Figure 8. Toxicities from Leukemia cohort

Results

6830 Allogenic HSCT patients, with a mean [SD] age 52 [15.14]. 69.1% were Whites, 8.3% Blacks, 11.3% Hispanics, and 11.2% Others. 8,660 toxicities with oral, gastric (GI), dermatologic, and constitutional symptoms were reported. In MVOS-PCA (Fig 1) ulcerative mucositis, diarrhea, and dehydration showed a correlation in PC1. The proportion of transformed variance explained (Cumulative Variance Accounted For) by each PC, in this case (GI), was 5.42 for PC1 and 10.55 for PC2. The first component has an eigen value of 1.41 and 90% of the variance is explained. Stratifying further to (Fig 2) GI toxicities shows- distension-bloating, and dyspepsia being correlated. The proportion of transformed variance explained (Cumulative Variance Accounted For) by each PC, in this case (GI), was 18% for PC1 and 37% for PC2. The first component has an eigen value of 1.129, and 90% of the variance is explained. The second component has an eigen value of 1.09, and 85% of the variance is explained. In the skin toxicities (Fig 6), there is a perfect contrast between dry skin, pruritis, pigmentation and rash in PCA 1. The proportion of transformed variance explained (Cumulative Variance Accounted For) by each PC, in this case (GI), was 22.7% for PC1 and 44.6% for PC2. The first component has an eigen value of 1.14, and 90% of the variance is explained. In the constitutional toxicities (Fig 4), weakness and fatigue are corelated. The proportion of transformed variance explained (Cumulative Variance Accounted For) by each PC, in this case (GI), was 35% for PC1 and 68.7% for PC2. The first component has an eigen value of 1.06, and 90% of the variance is explained.

Conclusions

Applying MVOS-PCA to an extensive claims database effectively detected CIT clusters and validated the clinical finding that patients experience various toxicities, frequently with a shared biological connection, at the same time. This innovation improves the flexibility of research and planning for supportive care, which is essential for establishing the best treatment approach for CIT.

References

Gifi, A. (1990). Nonlinear Multivariate Analysis. Wiley.
De Leeuw, J., Mair, P., & Groenen, P. (2016, September 26). Multivariate Analysis with Optimal Scaling. https://bookdown.org/jandeleeuw6/gifi/
Kenett, R. S., & Salini, S. (2012). Modern Analysis of Customer Surveys with Applications in R. New York: Wiley
Linting, M., Meulman, J. J., Groenen, P. J. F., & van der Kooij, A. J. (2007). Nonlinear principal components analysis: Introduction and application. Psychological Methods, 12(3), 336–358. https://doi.org/10.1037/1082-989X.12.3.336