"Science Stays True Here"
Journal of Mathematics and Statistical Science (ISSN 2411-2518, USA), Vol.12, Issue 11, 15-25 | Science Signpost Publishing

Non-parametric Multivariate Modified Compartmentalization Method for the Meaningful Organization, Decomposition and Analysis of Patient Reported Outcomes

Manasi Sheth[1] and N.Rao Chaganty[2]

1. *University of Wisconsin, Whitewater, WI, USA*
2. *Old Dominion University, Norfolk, VA, USA*

Received: August 16, 2026 / Accepted: September 11, 2026 / Published: November 25 2026.

Abstract: With the increasing integration of health care delivery systems, there is more provision for managing entire episodes of care in a patient-focused manner and to assess the impact of care on patient outcomes, including patient-reported outcomes (PROs). PROs are often relevant in assessing patient's activities of daily living and functionality when comparing the effects of the treatment on the patients. PRO's are often of ordinal in nature and hence are difficult to be detected by other statistics. In this paper, we present a novel approach of analyzing multivariate PROs such as mood quality and sleep quality using the Multivariate Modified Compartmentalization Method (MMCM). This method for multivariate outcomes will help preserve the complexity of the data yet provide a clinically and statistically meaningful organization and analyses of the data. It will also provide the tools necessary for the investigation necessary for large complex data. By breaking down these multivariate complex variables, the analytical methodology generates accurate solutions while upholding the fundamental integrity of our system.

Key words: Compartmentalization Method, Patient Reported Outcomes, Geometry, Transition Proportion Matrix, Longitudinal Data, Tensors

1. Introduction Generally, an ordinal scale such as a Likert scale is used for most health status indicators of interest for which the quantitative differences between levels are unclear or unknown. Characteristics such as perceived health status, functional independence, mobility, and pain are most appropriately obtained using an ordinal scale for patient reported outcomes. However, the data are usually analyzed using a continuous scale such as averages or means. Most of the methods for analyzing ordinal data are done by altering the nature of the data, that is, collapsing the ordinal scale to a dichotomous one or considering it to be continuous. When the ordinality of ordered data is not fully utilized, there is a significant loss of information and a significant bias that is introduced in the methods and the analyses (See Sheth and Chaganty (2025)).

A patient-reported outcome (PRO) is considered to be any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else. PROs are self-reported by the patients on the physical, mental and functioning health status and are useful to evaluate health conditions. In clinical trials, they are also useful to measure benefit or risk of a medical product such as a novel treatment or medical device (See Sheth and Chaganty (2025)). More details about patient reported outcomes are presented in Sheth and Chaganty (2025).

There are an abundant of papers analyzing PROs, but very few papers discusses the methodologies for analyzing patient reported outcomes (See Sheth and Chaganty (2025); Kucukemiroglu and Sheth (2021); Kucukemiroglu and Sheth (2020); Detry (2016); Prakash (2008), and, Black (2013)). In the first three papers, Sheth et al. discusses the application of Compartmentalization Method on ordinal data for univariate PROs. In this paper, we are proposing to quantify the increment or detriment in the ordinal scale of the multivariate patient reported outcomes that is easily interpretable by the clinicians, epidemiologists and statisticians so that it is clinically as well as mathematically/statistically meaningful. In this paper, we introduce a novel approach using the Multivariate Modified Compartmentalization Method (MMCM) is a non-parametric dimension reduction approach to analyzing multivariate categorical data for two or more time points or raters. The geometrical implications of this approach will also be discussed.

2. Methodology Section

2.1 Data Format of Patient Reported Outcomes. Every institution has a different way of collecting patient reported outcomes. Pertinent factors such as patient's baseline characteristics such as patient demographics, socio-economic factors etc. are also collected during an observational or experimental study (See Sheth and Chaganty (2025)). Patient-reported outcomes can be discrete or continuous. In this paper, we will only concentrate on the multivariate patient-reported outcomes that are of ordinal nature, such as those with Likert scale. Usually, the data is of the format that is seen in the Table 1 with $\boldsymbol{y_i}$ as the multivariate ordinal response for subject i. $i = 1, ..., n$. Let $\boldsymbol{x_i}$ be the corresponding set of co-variate(s) associated with the subject i. For this paper, we will especially concentrate on the geometrical representations and non-parametric approach to analyzing of the PROs such as severity level of mood, sleep, pain, nausea, sleep. when there are repeated measurements over time using the Multivariate Modified Compartmentalization Method (MMCM). Note that the bold symbols are used for vectors.

Table 1: General Layout of Multivariate Response Raw Data at one time point

Subject	Response	Covariates		
1	$(y_{11}, ..., y_{mn})$	x_{11}	$\dots$	x_{1l}
2	$(y_{21}, ..., y_{2m})$	x_{21}	$\dots$	x_{2l}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
n	$(y_{n1}, ..., y_{nm})$	x_{n1}	$\dots$	x_{nl}

2.2 Methodology and Illustration The Modified Compartmentalization Method (MCM) was presented in Sheth and Chaganty (2025) for univariate ordinal outcomes such as patient reported outcomes. Here, we present the Multivariate Modified Compartmentalization Method (MMCM) for multivariate ordinal outcomes.

Suppose there are two treatments or groups. In tables 2 and 3, we see that each cell represents a vector of frequencies of the m-dimensional multivariate ordinal responses as we transition from time t to $t + 1$. In Sheth and Chaganty (2025), we had explored tables with similar structure but for a univariate case of

ordinal outcomes. This table is obtained for a set of k x k multivariate ordinal responses, and for $t-1$ contingency tables. Thus, each cell represents $\boldsymbol{n_{ij}}$ where $i = a, b$ stands for the two the treatments given to each patient; j stands for the cell location such as row 2 column 2. Here, we assume that each of the survey question, say about pain or nausea or sleep etc., is based on the same ordinal scale such as a Likert scale.

Table 2: Cross-tabulation of multivariate ordinal responses for Treatment A

	Visit 2				
Visit 1	1	2	...	k	Total
1	$\boldsymbol{n_{a11}}$	$\boldsymbol{n_{a12}}$	...	$\boldsymbol{n_{a1k}}$	$\boldsymbol{n_{a1.}}$
2	$\boldsymbol{n_{a21}}$	$\boldsymbol{n_{a22}}$	...	$\boldsymbol{n_{a2k}}$	$\boldsymbol{n_{a2.}}$
.	...	...	...	...	...
k	$\boldsymbol{n_{ak1}}$	$\boldsymbol{n_{ak2}}$	...	$\boldsymbol{n_{akk}}$	$\boldsymbol{n_{ak.}}$
Total	$\boldsymbol{n_{a.1}}$	$\boldsymbol{n_{a.2}}$	...	$\boldsymbol{n_{a.k}}$	$\boldsymbol{n_{a..}}$

Table 3: Cross-tabulation of multivariate ordinal responses for Treatment B

	Visit 2				
Visit 1	1	2	...	k	Total
1	$\boldsymbol{n_{b11}}$	$\boldsymbol{n_{b12}}$	...	$\boldsymbol{n_{b1k}}$	$\boldsymbol{n_{b1.}}$
2	$\boldsymbol{n_{b21}}$	$\boldsymbol{n_{b22}}$	...	$\boldsymbol{n_{b2k}}$	$\boldsymbol{n_{b2.}}$
.	...	...	...	...	...
k	$\boldsymbol{n_{bk1}}$	$\boldsymbol{n_{bk2}}$	...	$\boldsymbol{n_{bkk}}$	$\boldsymbol{n_{bk.}}$
Total	$\boldsymbol{n_{b.1}}$	$\boldsymbol{n_{b.2}}$	...	$\boldsymbol{n_{b.k}}$	$\boldsymbol{n_{b..}}$

2.3 Dimension Reduction If we take each cell of m-dimensional vector of frequencies across the diagonal of this table, then we can obtain the addition of the vectors and cleverly obtain an m-dimensional multivariate proportion of each subject who have reach consistency $\boldsymbol{(p_{Cit})}$. Similarly, we can obtain an m - dimensional improvement $\boldsymbol{(p_{Iit})}$, and seen deterioration or are in worse condition $\boldsymbol{(p_{Wit})}$ by using the upper as well as lower diagonals of the contingency table. This can be lined up to write each row for the multivariate matrix, which is an alternative to writing a 3-way tensor.

For each multivariate matrix comparing each consecutive time point, the multivariate proportion of each subject or patient who have reached consistency $\boldsymbol{(p_{Cit})}$, seen improvement $\boldsymbol{(p_{Iit})}$, and seen deterioration or are in worse condition $\boldsymbol{(p_{Wit})}$ is calculated. This method is called Multivariate Modified Compartmentalization Method (MMCM). Note here that each of the joint probabilities is of 1 x m vector of multivariate proportions. Also, $\boldsymbol{(p_{Cit})}$ is a m by 1 vector of consistency probabilities and obtained for $t-1$ contingency tables; $\boldsymbol{(p_{Iit})}$ is a m by 1 vector of improvement probabilities and obtained for $t-1$ contingency tables; and, $\boldsymbol{(p_{Wit})}$ is a m by 1 vector of worsening probabilities and obtained for $t-1$ contingency tables. That means, each probability is a matrix and there are three of them, making this a probability cuboid of dimension 3 x m x $t-1$. This is done under the assumption that there is no correlation

between the multivariate responses at this point. Also, note that the conditional probabilities described in the first paper, $(\boldsymbol{p_{Ct}})$, $(\boldsymbol{p_{It}})$, and $(\boldsymbol{p_{Wt}})$, add up to 1 for each contingency table. That is, $(p_{11}, p_{12}, ..., p_{1m})$ for each transitional contingency table. There are $t-1$ tables like this. So the notation is $\boldsymbol{p_{i'j'k'}}$ where i' = for number of treatments; j' = for number of severity levels, $k' = 1, 2, 3, ..., t-1$ for tables. The statistic $\boldsymbol{p_{i'j'k'}}$ is also a multivariate vector with as many variables as in the study such as pain, sleep, nausea, etc.

Thus say for 4 visits, we can write the multivariate transition proportional matrices or transitional proportions tensor product for the two treatments as follows:

$P_A =$

$$\begin{pmatrix} \boldsymbol{p_{a11}} & \boldsymbol{p_{a21}} & \boldsymbol{p_{a31}} \\ \boldsymbol{p_{a12}} & \boldsymbol{p_{a22}} & \boldsymbol{p_{a32}} \\ \boldsymbol{p_{a13}} & \boldsymbol{p_{a23}} & \boldsymbol{p_{a33}} \end{pmatrix}$$

and $P_B =$

$$\begin{pmatrix} \boldsymbol{p_{b11}} & \boldsymbol{p_{b21}} & \boldsymbol{p_{b31}} \\ \boldsymbol{p_{b12}} & \boldsymbol{p_{b22}} & \boldsymbol{p_{b32}} \\ \boldsymbol{p_{b13}} & \boldsymbol{p_{b23}} & \boldsymbol{p_{b33}} \end{pmatrix}$$

Similarly for 5 visits, we can write the multivariate transition proportional matrices or the multivariate transitional proportional tensor products for the two treatments as follows:

$P_A =$

$$\begin{pmatrix} \boldsymbol{p_{a11}} & \boldsymbol{p_{a21}} & \boldsymbol{p_{a31}} \\ \boldsymbol{p_{a12}} & \boldsymbol{p_{a22}} & \boldsymbol{p_{a32}} \\ \boldsymbol{p_{a13}} & \boldsymbol{p_{a23}} & \boldsymbol{p_{a33}} \\ \boldsymbol{p_{a14}} & \boldsymbol{p_{a24}} & \boldsymbol{p_{a34}} \end{pmatrix}$$

and $P_B =$

$$\begin{pmatrix} \boldsymbol{p_{b11}} & \boldsymbol{p_{b21}} & \boldsymbol{p_{b31}} \\ \boldsymbol{p_{b12}} & \boldsymbol{p_{b22}} & \boldsymbol{p_{b32}} \\ \boldsymbol{p_{b13}} & \boldsymbol{p_{b23}} & \boldsymbol{p_{b33}} \\ \boldsymbol{p_{b14}} & \boldsymbol{p_{b24}} & \boldsymbol{p_{b34}} \end{pmatrix}$$

Note here that the tensor is any object that is more than 3 dimensions. Therefore, we could even combine the multivariate transition probabilities matrices for P_A and P_B into a 4-way tensor. But for simplicity and readibility purposes, we will use two of the 3-way tensors instead.

Thus, if we obtain these geometrical representations for $m+1$ visits, we can get a m x 3 x 3 transitional proportion tensor product for each of the treatments; for 5 visits, we get a m x 4 x 3 transitional proportional tensor product for each of the treatments; for 6 visits, we can get a m x 5 x 3 transitional proportional tensor product for each of the treatments; and so on. Note: here we assume that each of the dimensions of the transitional probabilities are independent from each other. That is, we assume that pain variable is independent of the nausea variable and is further independent of the sleep variable in the survey questionnaire.

From combining the tensors for the two treatments, we can also get a m x 3 x 3 x 2 which is a 4-way tensor depicting for all the data. The tensor given by each of the matrix above can be depicted here using the Tesseract as in the Figure 1. Here the tensor is depicted to have a multivariate vector for each cell

unlike the usual notation given in mathematical textbooks. It is beyond the grasp of human brain perhaps to even imagine what the structure looks like for an un-compartmentalized contingency table. Sheth and Chaganty (2025) has already shown in the univariate case that the dimension reduction using Modified Compartmentalization Method (MCM) have proven that this method is quite meaningful using algebraic geometry.

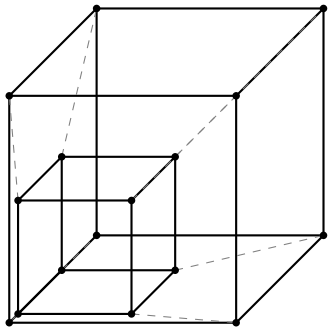


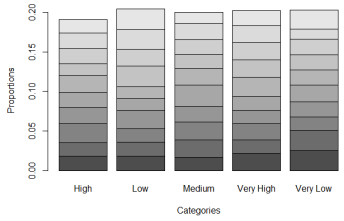
Figure 1: Tesseract or Octachoron

The Canonical Polyadic (CP) Decomposition of a tensor refers to its expression as a sum of r rank-1 components (Ballard, G., & Kolda, T. G. (2025)). Each component is a vector outer product. The book refers to r colloquially as the rank of the decomposition, though this is technically only precise if r is minimal. Each vector is called a factor, and the set of factors of a mode is called a factor matrix. In the 3-way tensor, each component is the outer product of three factors. Specifically, component j is the outer product of column j of the first factor matrix, column j of the second factor matrix, and column j of the third factor matrix. The factors are matched and not interchangeable.

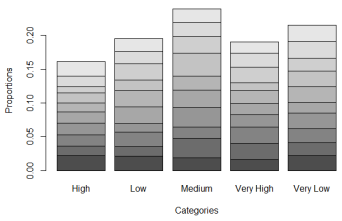
2.4 Method of Analyses

Note that we have p_{ijkl} where we have four dimensions of values such as 2 treatment levels, 3 severity levels, $t-1$ number of contingency tables and k number of multivariate responses. We can easily fit a logit model to this to obtain the treatment and time effect or any set of covariates. In this paper, we will just simply present how to fit the logit models in such situations.

The statistic $\boldsymbol{p_{i'j'k'}}$ is also a multivariate vector with as many variables as in the study such as pain, sleep, nausea, etc. For each probability in the tensor, for example $\boldsymbol{p_{a11}}$, can be written as $(p_{a111}, p_{a112}, ..., p_{a11k})$ with $l' =$ usually for a number of variables such as mood quality, sleep quality, pain, etc. using vector notation. Thus, for p_{a111} we can fit a logit model or a logit regression model given a set of covariates pertaining to that particular combination of treatment, severity levels, and number of contingency tables. Similarly, for p_{a112} we can fit another logit regression model given a set of covariates pertaining to that particular combination of treatment, severity levels, and number of contingency tables. Here, in the context of conditional probabilities, we have compartmentalization levels that refers to the proportions of consistency, improvement and deterioration of the patient's particular health condition. Since in this paper, we are assuming that the outcome variables are not related to each other, one option is to run as many logit models as the number of outcomes variables. Another option is perform nested logit models using levels of severity and number

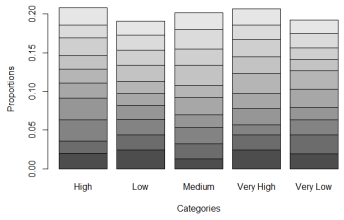


(a) 2a

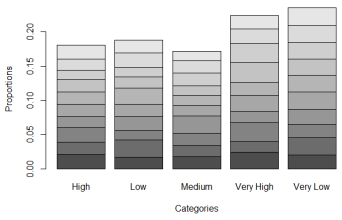


(b) 2b

Figure 2: Plots of Proportions of Mood Quality of Treatments A and B



(a) 3a



(b) 3b

Figure 3: Plots of Proportions of Sleep Quality of Treatments A and B

of reported outcome variables. In case of dependency, this number gets reduced and that will be covered in another paper.

3. Results and Discussion Suppose that we have the following the Mental Health Diagnosis and Treatment Monitoring dataset that contains about 1000 rows of data representing simulated real-world mental health diagnoses, treatment plans, and outcomes. It include patient demographics, symptom severity, medication, therapy types and progress tracking. This dataset is synthetic and created for research and analysis purposes only, and is not to be used for generalizability of the results.

The synthetic dataset includes key features such as patient ID, age, gender, diagnosis, symptom severity $(1-10)$, mood score $(1-5)$, sleep quality $(1-5)$, physical activity measured in hours per week of activity, medication, therapy type, treatment start date, treatment duration, stress level $(1-10)$, treatment progress $(1-10)$, and adherence to treatment.

The following figures 2 and 3 shows the distribution of proportion by therapy types for 10 different visits for the two PRO outcomes of mood source and sleep quality, respectively. The plots provide descriptive results of the outcomes by the two different types for the two different treatments.

The multivariate matrix notation has the proportions for "(Mood Score, Sleep Quality)" in each cell element of the matrix whereas the 3-way tensor in this example has the "mood score" and "sleep quality" separated. There are two treatments, A and B, so that there are two matrices or four frontal slice of tensors. Each column represents three different vectors or multivariate proportions of consistency, improvement and deterioration. Each row represents of

increment from time t to time $t+1$. Since in our synthetic dataset, we have 10 consecutive time periods, we have 9 different rows for 9 different transitional matrices and three columns depicting multivariate proportions of consistency, improvement and deterioration of a patient's condition based on their consecutive responses to "Mood Score" and "Sleep Quality" for each cell of the matrix. Thus, for each treatment or therapy, we have 2 x 9 x 3 tensor. Below are both the notations of the transition probabilities in a multivariate matrix as well as a tensor format.

The frontal slices of the tensor for the transition probabilities for treatment A are as follows: $P_A(:,:,1) =$

$$\begin{pmatrix} 0.22 & 0.39 & 0.39 \\ 0.15 & 0.36 & 0.49 \\ 0.20 & 0.46 & 0.34 \\ 0.18 & 0.43 & 0.39 \\ 0.28 & 0.35 & 0.37 \\ 0.18 & 0.43 & 0.39 \\ 0.13 & 0.47 & 0.40 \\ 0.19 & 0.42 & 0.39 \\ 0.23 & 0.50 & 0.27 \end{pmatrix}$$

and $P_A(:,:,2) =$

$$\begin{pmatrix} 0.26 & 0.45 & 0.29 \\ 0.21 & 0.42 & 0.37 \\ 0.16 & 0.45 & 0.29 \\ 0.19 & 0.43 & 0.38 \\ 0.14 & 0.52 & 0.34 \\ 0.21 & 0.36 & 0.43 \\ 0.16 & 0.36 & 0.48 \\ 0.17 & 0.38 & 0.45 \\ 0.16 & 0.30 & 0.54 \end{pmatrix}$$

The frontal slices of the tensor for the transition probabilities for treatment B are as follows: $P_B(:,:,1) =$

$$\begin{pmatrix} 0.19 & 0.43 & 0.39 \\ 0.19 & 0.37 & 0.49 \\ 0.19 & 0.43 & 0.34 \\ 0.17 & 0.37 & 0.39 \\ 0.22 & 0.44 & 0.37 \\ 0.19 & 0.38 & 0.39 \\ 0.20 & 0.38 & 0.40 \\ 0.23 & 0.33 & 0.39 \\ 0.17 & 0.38 & 0.27 \end{pmatrix}$$

and $P_B(:,:,2) =$

$$\begin{pmatrix} 0.29 & 0.36 & 0.35 \\ 0.18 & 0.34 & 0.44 \\ 0.29 & 0.36 & 0.35 \\ 0.27 & 0.35 & 0.38 \\ 0.19 & 0.36 & 0.45 \\ 0.21 & 0.37 & 0.42 \\ 0.20 & 0.35 & 0.45 \\ 0.13 & 0.34 & 0.53 \\ 0.18 & 0.36 & 0.46 \end{pmatrix}$$

An alternative method for writing 3-way 9 x 3 X 2 tensor is as below for both treatments using the multivariate matrix notation mentioned above: $P_A =$

$$\begin{pmatrix} (0.22,\ 0.26) & (0.39,\ 0.45) & (0.39,\ 0.29) \\ (0.15,\ 0.21) & (0.36,\ 0.42) & (0.49,\ 0.37) \\ (0.20,\ 0.16) & (0.46,\ 0.45) & (0.34,\ 0.29) \\ (0.18,\ 0.19) & (0.43,\ 0.43) & (0.39,\ 0.38) \\ (0.28,\ 0.14) & (0.35,\ 0.52) & (0.37,\ 0.34) \\ (0.18,\ 0.21) & (0.43,\ 0.36) & (0.39,\ 0.43) \\ (0.13,\ 0.16) & (0.47,\ 0.36) & (0.40,\ 0.48) \\ (0.19,\ 0.17) & (0.42,\ 0.38) & (0.39,\ 0.45) \\ (0.23,\ 0.16) & (0.50,\ 0.30) & (0.27,\ 0.54) \end{pmatrix}$$

and $P_B =$

$$\begin{pmatrix} (0.19,\ 0.29) & (0.43,\ 0.36) & (0.38,\ 0.35) \\ (0.19,\ 0.18) & (0.37,\ 0.34) & (0.44,\ 0.44) \\ (0.19,\ 0.29) & (0.43,\ 0.36) & (0.38,\ 0.35) \\ (0.17,\ 0.27) & (0.37,\ 0.35) & (0.46,\ 0.38) \\ (0.22,\ 0.19) & (0.44,\ 0.36) & (0.34,\ 0.45) \\ (0.19,\ 0.21) & (0.38,\ 0.37) & (0.43,\ 0.42) \\ (0.20,\ 0.20) & (0.38,\ 0.35) & (0.42,\ 0.45) \\ (0.23,\ 0.13) & (0.33,\ 0.34) & (0.44,\ 0.53) \\ (0.17,\ 0.18) & (0.38,\ 0.36) & (0.45,\ 0.46) \end{pmatrix}$$

After performing CP decomposition on the tensor of treatment A for mood score and sleep quality, we find that the λ's are 13.221 and 4.948. The norm percent is 90.75% and so this factorization is useful as seen below. The first component is
$[U_A]_1 =$

$$\begin{pmatrix} \text{-}0.088 & 0.164 \\ \text{-}0.110 & 0.122 \\ \text{-}0.087 & 0.146 \\ \text{-}0.103 & 0.131 \\ \text{-}0.106 & 0.141 \\ \text{-}0.115 & 0.110 \\ \text{-}0.124 & 0.077 \\ \text{-}0.131 & 0.069 \\ \text{-}0.134 & 0.051 \end{pmatrix}$$

The second component is $[U_A]_2 =$

$$\begin{pmatrix} 0.148 & -0.306 \\ 0.278 & -0.653 \\ 0.575 & 0.040 \end{pmatrix}$$

The third component is $[U_A]_3 =$

$$\begin{pmatrix} -0.484 & 0.574 \\ -0.515 & -0.426 \end{pmatrix}$$

After performing CP decomposition on the tensor of treatment B for mood score and sleep quality, we find that the λ's are 17.601 and 0.925. The norm percent is 91.20% and so this factorization is useful as seen below. The norms and residuals were found to be similar for the treatments. The first component is
$[U_B]_1 =$

$$\begin{pmatrix} 0.116 & -0.142 \\ 0.115 & 0.069 \\ 0.112 & -0.131 \\ 0.110 & -0.072 \\ 0.114 & 0.080 \\ 0.111 & 0.030 \\ 0.111 & 0.078 \\ 0.107 & 0.256 \\ 0.102 & 0.141 \end{pmatrix}$$

The second component is $[U_B]_2 =$

$$\begin{pmatrix} -0.216 & -0.393 \\ -0.381 & -0.018 \\ -0.403 & 0.971 \end{pmatrix}$$

The third component is $[U_B]_3 =$

$$\begin{pmatrix} -0.497 & -0.029 \\ -0.503 & 0.971 \end{pmatrix}$$

In the above analyses, we had performed CP decomposition on the 3-way tensors for the treatments A and B. However, one can also use this decomposition on a 4-way tensor (not depicted in this paper due to difficulty in readability purposes). See the results below:

The λ's are 20.405 and 44.865. The norm percent is 89.716% and so this factorization is useful as seen below. The norms and residuals were found to be similar when compared to the 3-way decomposition.

The first component is

$[U_{AB}]_1 =$

$$\begin{pmatrix} 0.214 & 0.187 \\ -0.098 & 0.045 \\ 0.251 & 0.202 \\ 0.097 & 0.129 \\ 0.078 & 0.127 \\ 0.078 & 0.127 \\ -0.007 & 0.085 \\ -0.162 & 0.014 \\ 0.048 & 0.104 \end{pmatrix}$$

The second component is $[U_{AB}]_2 =$

$$\begin{pmatrix} 0.168 & -0.200 \\ 0.347 & -0.372 \\ 0.485 & -0.427 \end{pmatrix}$$

The third component is $[U_{AB}]_3 =$

$$\begin{pmatrix} -0.481 & 0.490 \\ -0.519 & 0.509 \end{pmatrix}$$

The fourth component is $[U_{AB}]_4 =$

$$\begin{pmatrix} 0.5 & -0.5 \\ 0.5 & -0.5 \end{pmatrix}$$

The main goal of creating these 4-way tensors is to see if there is an impact of the therapies or treatments on the bivariate ordinal responses of "mood quality" and "sleep quality".

Note that $[U_{AB}]_1$ is a 9 x 2 matrix representing the CP decomposition for the effect of two therapies on the 9 transition times, i.e. 9 contingency tables. Similarly, $[U_{AB}]_2$ is a 3 x 2 matrix representing the CP decomposition for the effect of two therapies on the 3 levels of patient's well being, i.e. consistency, improvement and deterioration. $[U_{AB}]_3$ is a 2 x 2 matrix representing the CP decomposition for the effect of two therapies on the bivariate outcomes, i.e. mood quality and sleep quality.

Performing logit regression models on this synthetic dataset shows no major significance in the bivariate outcomes by the therapy types indicating that the two arbitrary therapies had a similar effect on the multivariate outcomes. Other factors were not considered in the logit regression models fitted to the above data.

4. Conclusions Usually the ordinal data such as PROs come in very large and complex format and are very challenging to understand, organize and analyze. Understanding the data can not only help in providing the appropriate method to analyze, but also reduce time and effort put into the data analysis architecture (See Sheth and Chaganty (2025)).

The multivariate matrix or tensors obtained from the organization of the data provides the transition probability distribution for multiple outcomes, where as the figure provides the dimension reduction logic on the transition probability distribution in our proposed method. These choices represent the

choice made by the patient in determining the best or worst or consistent transition probabilities when dimension reduction is used. Given the changing nature of responses, this proposal includes a time-dependent approach to a response or choice pairs. There is also a need to understand the dependency structure between the multivariate outcomes especially those of patient reported outcomes, but that will be covered in a subsequent paper with further analyses. Another disadvantage noticed here is there is a reduction in the norm percentage when using a more than 3-way tensor for CP decomposition and that will be investigated in a separate paper.

Acknowledgment Dr. Sheth acknowledges the support of Departments of Mathematics and Computer Science, University of Wisconsin - Whitewater, WI, USA. She would also like to thank Dr. Sumen Sen for providing his expertise in user - friendly programming codes which are available upon request.

Dr. Chaganty acknowledges the support of the Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA, USA.

References

1. Sheth, M., & Chaganty, N. R. (2025). Model-free organization of patient reported outcomes data: Geometrical rep-resentation of the Modified Compartmentalization Method. Journal of Mathematics and Statistical Science, 11(11), 87-96.

2. S. Kucukemiroglu and M. Sheth, Nonparametric Analysis of Patient Reported Outcomes using Compartmentalization Method, Joint Statistical Meeting Proceedings, 2021, p. 574-581.

3. S. Kucukemiroglu and M. Sheth, Compartmentalization of Discrete Repeated Measures in Patient-Reported Outcomes Questionnaires, In JSM Proceedings, Biometrics Section, pp.2593-2599.

4. M.A. Detry and Y. Ma, Analyzing Repeated Measures using Mixed Models, JAMA (2016), 315(4), pp. 407-408.

5. A. Prakash, R.C. Risser, and C.H. Mallinckrodt, The impact of the analytical method on the interpretation of results in longitudinal clinical trials, International Journal of Clinical Practice (2008), 62(8), pp. 1147-1158.

6. R. Diestel, Graph theory, Springer Texts (2020), 6th Edition

7. Ballard, G., & Kolda, T. G. (2025). Tensor Decompositions for Data Science. Cambridge University Press.