The Case for APBD Clinical Trials with Two Arms

by Larry Schwartz, 15 January 2018

Introduction

Earlier, I quantified that the 22 APBD patients who participated in the Triheptanoin Phase 2 clinical trial had very dissimilar conditions.[1] Dr. Raphael Schiffmann, the PI of this trial, said, "*This study…emphasizes the difficulty of conducting trials…with a wide clinical heterogeneity.*" [2]

In principle, cluster analysis can break down APBD heterogeneity into relatively homogeneous sub-classes, thereby making clinical trials less difficult to conduct. In this paper, I show that cluster analysis does in deed bring APBD heterogeneity under control.

This paper is organized as follows:

Section 1. Executive Summary

I demonstrate statistical clustering with the recently completed Triheptanoin Phase 2 clinical trial for APBD. That cluster process created two relatively homogeneous sub-classes based upon differences in the baseline six-minute walking distances (the primary endpoint) and ages of the participants. As a result, this trial's participants would have been assigned to either a "mild" or "severe" APBD sub-class----reducing heterogeneity to under nine percent, meeting the strict standards of assay lab work.

For future Phase 2 clinical trials regarding APBD, I document this clustering process to obtain similar results. It too would be based upon the universal baseline characteristics of a primary endpoint (for clustering purposes, does not matter which is selected) and the age of its trial participants. To take advantage of this approach, the protocol for a future trial would have two arms, one for each of the relatively homogeneous sub-classes.

However, about 50 APBD patients would have to participate in such a future trial; feasible with about 100 in the APBD Research Foundation's registries (CAP and FAN). This many participants are necessary to meet sample size requirements for each trial arm.

Section 2. APBD Heterogeneity for the Triheptanoin Trial

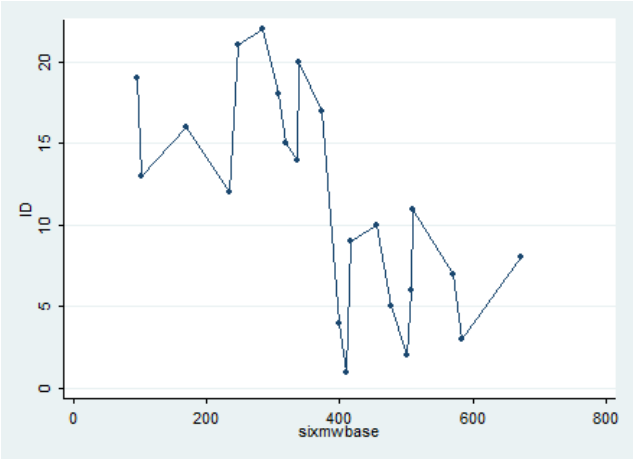
---

[1] **Schwartz, L. "*APBD Heterogeneity: The Case of the Triheptanoin Trial,*" January 2018, APBD Research Foundation Web site.**

[2] **Schiffmann, R. and others, "*A double-blind, placebo-controlled trial of triheptanoin in adult polyglucosan body disease and open-label, long-term outcome*," Journal of Inherited Metabolic Disease, 6 November 2017. We are indebted to Dr. Rafael Schiffmann, PI of the Triheptanoin trial, for providing the data used in this study. Errors of omission or commission solely remain with the author.**

As shown earlier, Graph 1 illustrates the heterogeneity of the baseline distances walked in six minutes for the Triheptanoin trial. On the vertical axis is the ID numbers assigned to each of its 22 APBD participants. On the horizontal axis is the baseline six-minute walking distances, in meters. Visually, the trial participants seem to have dissimilar six-minute walking distances.

Graph 1. Baseline Six Minute Walk for Trial Participants



In statistical terms, the mean for this baseline six-minute walking is 377 meters with a standard deviation of 153 meters. As a summary measure, the coefficient of variation (CV) takes the standard deviation and normalizes it with the mean, useful for comparing results to other medical areas. On this measurement basis, the CV for this APBD trial is 41 percent (153/377), much higher than what is acceptable for *inte*r-assay lab work (15 percent) or *intra*-assay lab work (10 percent).[3]

Section 3. Standard of Tolerable Heterogeneity

By all accounts, this CV of 41 percent led to difficulties in the conduct of the Triheptanoin Phase 2 trial for APBD. The process of clustering can reduce such heterogeneity, but it requires a standard of tolerable heterogeneity to do so. For this purpose, I adopt the strict assay lab standard of 10 percent.

Section 4. Clustering with the Triheptanoin Trial Example

In this section, I define clustering and its process as well as document the results with the Triheptanoin trial example; leaving the more technical material to Appendix B.

[3] **Steven, "*How to Calculate the Coefficients of Variation*, "Top Tip Bio, April 5, 2017.**
Salimetrics, "*Inter- and Intra Assay Coefficients of Variation,*" about 2010.

Definition: What is a cluster? In general, a cluster is a grouping of things that are similar with respect to one or more characteristics. Each grouping is dissimilar to any other grouping based on those characteristics. In medicine, cluster analysis is commonly used to identify sub-categories of illnesses or conditions.[4]

Process: Consider the scenario for clustering with the Triheptanoin trial based solely on its baseline six-minute walking.[5] The goal is to create two relatively homogeneous sub-classes. The first clustering step randomly selects two of the participants' baseline walking distances among those of the 22 trial participants. One of the selections becomes the median seed for cluster 1, the other the median seed for cluster 2.

The process continues with the selection of any third baseline six-minute walking distance, drawn from the remaining 20 unassigned participants. The third participant is assigned to either cluster 1 or cluster 2 depending on its closeness to the two established cluster medians; "closeness" is measured as the sum of squared differences between the additional six-minute walking distance and that of the two standing clusters. With each addition to a cluster, its median is updated.

The process ends when the assignment of a participant to a sub-class, or reassignment from one sub-class to another, no longer reduces walking-distance differences.

I validate the cluster approach for APBD with both an internal and external test.

Results: In fact, I go through two cluster scenarios. First continuing with the above six-minute-walking scenario, its clustering process creates a milder APBD sub-class that represents 10 six-minute walking distances that average 510 meters, while the more severe APBD sub-class represents 12 six-minute walking distances that average only 266 meters. Visionally, Graph 1 suggests these two distinct clusters.

But the cluster for the mild APBD sub-class has a CV of 17 percent, while the cluster for the more severe APBD sub-class has a CV of 37 percent. Although this is a reduction from the overall CV of 41 percent (before clustering), it is not nearly good enough to meet assay lab work standards.

---

[4] **Peters, J.B. and others, "*Integral Health Status-Based Cluster Analysis in Moderate–Severe COPD Patients Identifies Three Clinical Phenotypes: Relevant for Treatment as Usual and Pulmonary Rehabilitation*," International Journal of Behavioral Medicine, 2017**

Axén, I. and others, "*Clustering patients on the basis of their individual course of low back pain over a six-month period*," BMC Musculoskeletal Disorders, *17 May 2011.*

Mulroy, S. and others, *"Use of cluster analysis for gait pattern classification of patients in the early and late recovery phases following stroke*," Gait & Posture, August 2003.

[5] **There are several clustering methods, but I chose the so-called Kmedians approach for two reasons. First with such wide variability in the APBD condition, this approach appropriately uses the median to distinguish clusters. Second with the limited number of APBD patients available to participate in a trial, the Kmedians approach allows me to pre-determine two clusters.**

In clustering analysis, you can do better by introducing more baseline characteristics. For this second scenario, therefore, I add the age of the APBD participants (at time of consent) to the six-minute-walking distances---combining the two into an index.[6] Before clustering, the coefficient of variation for the entire walking-age index is 13.6 percent (standard deviation of 11 index values relative to a mean of 84 index values). After clustering with the walking-age index, the CV for the mild APBD cluster is reduced to 8.2 percent, the CV for the severe APBD cluster to 5.6 percent----both meeting the 10 percent assay lab standard. Table 1 summarizes this second scenario.

Table 1. APBD-Triheptanoin *Clusters Based on the Walking-Age Index*

Statistic	***Cluster 2 (Mild APBD)***	***Cluster 1 (Severe APBD)***
Index:		
CV	8.2%	5.6%
Mean	71.7	91.2
Standard Deviation	5.9	6.1
Participant IDs	2,3, 5, 6 7,8,9,10	1, 4, 11, 12,13,14 15, 16, 17,18,19 20, 21, 22
Associated Medians:		
Age of participant	54	60
SM Walking distance, meters	505	313

As mentioned earlier, Cluster 2 may be labelled as “mild APBD”, Cluster 1 as “severe APBD”. The mild APBD cluster has considerably longer median six-minute walking distances (505 meters vs. 313 meters) and somewhat younger median-aged participants (54 years old vs. 60 years old).

But the two walking-age clusters have very uneven number of IDs assigned to them: The more severe APBD cluster contains 63 percent of the IDs (14/22), while the mild APBD cluster contains 37 percent of them (8/22). In contrast the clustering with the sole baseline characteristic of the six-minute walk contains a more even split of IDs, but does not reduce APBD heterogeneity nearly enough.

[6] **Statistically, principal component analysis is well-suited for this purpose. In this paper, I use the first principal component to develop the walking-age index.**

I undertake two validations of the clustering---one internal, the other external. For the internal validation, I fit a regression equation between the walking-age index and the cluster variable (with its values of 1 or 2). It shows that these two clusters explain 74 percent of the variation of the walking-age index values. This is a high number for such an analysis—giving us confidence in this cluster solution.[7]

For the external validation, I test clustering with a different primary endpoint, Rand's Physical Health Summary (PHS). It also successfully produced two relatively homogeneous sub-classes that meet assay lab working standards of tolerable heterogeneity; see Appendix A.

Section 5. Guaiacol Trial Considerations

There are (at least) five considerations regarding clustering with a Guaiacol Phase 2 trial for APBD in the United States:

1. Primary Endpoint: The Guaiacol Phase 2 trial for APBD may or may not designate the six-minute walk as the primary endpoint. Other primary endpoints are under consideration----including the biomarker of measuring glycogen levels in the liver and the composite ALS Functional Rating Scale. The clustering process demonstrated in this paper would work well regardless of the selected primary endpoint.[8]

2. Maximum APBD Participants: Currently, no more than 50 APBD patients can be expected to participate in a Phase 2 Guaiacol trial. There is about 100 APBD patients in the APBD registries (CAP and FAN), but meeting inclusion criteria, recruiting and other factors can reduce participation by as much as 50 percent. (Dr. Schiffmann started out with 45 potential participants with the Triheptanoin Phase 2 trial for APBD but ended up with half that number for these reasons.)

3. Sample Sizes: Accepting Dr. Schiffmann's minimum sample requirement of 18 APBD participants (actually worked with 22) for the single-armed Triheptanoin Phase 2 trial, the sample size requirement for each arm of the Guaiaciol Phase 2 trial would be the same. However, Guaiacol clustering with the (baseline) endpoint-age characteristics naturally yields uneven IDs in the two sub-classes: 37 percent fall into the mild APBD sub-class, 63 percent into the severe APBD sub-class. Based on these proportions, the Gauaiacol trial would require 50 participants to ensure at least 18 were in each cluster: The mild APBD sub-class obtaining 19 (0.37x50) and the severe APBD sub-class 31(.63 x 50).

---

[7] **The regression equation is as follows: SMW-Age Index=117.6-19.5 x Cluster (1 or 2). The t-ratio for the Cluster coefficient is -7.7, significant at the 95 percent confidence level; and the R-square for the equation is 0.74.**

Note that the mean index values can be derived from this equation: For the cluster with value =1, the mean SMW-Age=117.6-19.5 x 1=98.1. And for the cluster with the value =2, the mean SMW-Age=117.6-19.5 x 2=78.6. The same mean index values are shown in Table 1, but they were calculated directly from the underlying data.

[8] **This assumes that the choice of a primary endpoint has reasonable sensitivity to differences in APBD conditions, thereby showing the inherent heterogeneity of APBD.**

4. Age of Participants: This clustering process requires documentation for the ages of the participants (as well as a primary endpoint). In the Triheptanoin Phase 2 trial for APBD, the ages of the participants were obtained at the time of consent. The same could be done for the future Guaiacol trial.[9]

5. Clustering Effort: With the detailed instructions in Appendix B, the statistician should be able to do this statistical work and document its results within days.

Appendix A. External Validation Test

INTRODUCTION

Appendix A works with Rand's composite physical component summary (PCS) to further test the clustering template.

The results of this test were positive—a further validation of its general use.

ASSUMPTIONS

For purposes of testing the clustering template, I make the following four assumptions:

(1) The composite PCS would serve as the baseline primary endpoint for purposes of this test;

(2) The goal of the cluster template is to create two relatively homogeneous sub-classes;

(3) The standard of tolerable heterogeneity is the assay lab work of 10 percent;

(4) The clustering begins with the lone characteristic of the composite PCS; if its results are not satisfactory, then the ages of the participants are added as a second characteristic.

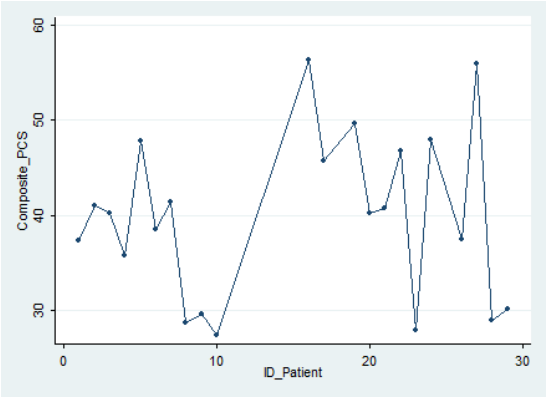
A positive test result would minimize heterogeneity and meet the assay lab standard.

DEGREE of HETEROGENEITY

Appendix A Graph 1 illustrates the heterogeneity of the baseline composite PCS in the Triheptanoin Phase 2 trial for APBD. On the vertical axis is the ID numbers assigned to each of its 22 APBD participants. On the horizontal axis is the baseline composite PCS measure, in percent; the higher the PCS percent, the more serious the physical condition. Visually, the trial participants have dissimilar PCS scores.

[9] **In my review of the CAP registry, the birth dates of the participants are not included.**

Appendix A Graph 1. PCS of the Rand Short Survey



In statistical terms, the mean for the baseline PCS measure was 39.8 percent, with a standard deviation of 8.7 percent. As a summary measure, the coefficient of variation (CV) takes the standard deviation and normalizes it with the mean, useful for comparing results to other medical areas. On this measurement basis, the CV for the PCS is 22 percent (.087/.398), considerably higher than what is acceptable for the strict assay lab standard of 10 percent.

CLUSTER ANALYSIS

I first perform cluster analysis with the lone characteristic of PCS, and it yields a mild APBD sub-class and a more severe one.[10] The mild APBD sub-class has a coefficient of variation (CV) of 3.5 percent, but the more severe APBD sub-class has a CV of 14.4 percent. This result is not good enough to meet the assay lab work standard.

In clustering analysis, you can do better by introducing more (baseline) characteristics to distinguish sub-classes. Accordingly, I add the age of the APBD participants (at time of consent) to the composite PCS---combining the two into an index.[11] Before clustering, the coefficient of variation for the entire PCS-age index is 12.2 percent. After clustering of the PCS-age index, the CV is 5.8 percent for the mild APBD cluster, 8.9 percent for the severe one----both meeting the 10 percent assay standard.

[10] **There are several clustering methods, but I chose the so-called Kmedians approach for two reasons. First with such wide variability in the APBD condition, this approach appropriately uses the median to distinguish clusters. Second with the limited number of APBD patients available to participate in a trial, the Kmedians approach allows me to pre-determine two clusters that can be populated for a trial.**

[11] **Statistically, principal component analysis is well-suited for this purpose. In this paper, I use the first principal component to develop the walking-age index.**

SUMMARY of FINDINGS

This clustering approach created two relatively homogeneous APBD sub-classes with the baseline characteristics of the PCS (assumed primary endpoint) and the ages of the patients.

The resultant two APBD sub-classes meet the strict standards of assay lab work, 10 percent.

CONCLUSION

The application of the clustering template to the external PCS case further validates its general use.

Appendix B. Statistician Instructions

The statistician should be well versed in principal component analysis, cluster analysis, and regression analysis to undertake the clustering for a new APBD trial. Further, have access to a statistical software package that offers these statistical procedures; the author used Stata, but SAS and SPSS also offer this capability.

The end product of the statistician's work is to fill in Table 1 (page 4 of the paper) for the new clinical trial, and briefly write it up.

In this Appendix, I will go through the seven necessary steps to create two clusters with the future trial's baseline primary endpoint and the ages of its participants. Along the way, Stata printouts are shown for the Triheptanoin case.

***Step 1*:** Compile the APBD data for patient ID numbers, the baseline primary endpoint and the ages of the participants (at consent signing or after selecting the trial participants). In clustering, there should be about 50 observations for the Phase 2 trial for APBD.

Triheptanoin case: Here is the Stata printout of these data for the 22 patients that participated in the Triheptanoin Phase 2 trial for APBD:

```
     +--------------------------+
     | ID   sixmwb~e   agebegin |
     |--------------------------|
  1. |  1        409         73 |
  2. |  2        502         56 |
  3. |  3        583         54 |
  4. |  4        398         62 |
  5. |  5        476         39 |
```

```
     |--------------------------|
  6. |  6        508         44 |
  7. |  7        571         55 |
  8. |  8        672         60 |
  9. |  9        416         55 |
 10. | 10        455         35 |
 11. | 11        509         70 |
 12. | 12        235         60 |
 13. | 13        102         59 |
 14. | 14        335         63 |
 15. | 15        318         57 |
     |--------------------------|
 16. | 16        170         65 |
 17. | 17        373         59 |
 18. | 18        307         66 |
 19. | 19         95         60 |
 20. | 20        338         57 |
 21. | 21        247         55 |
 22. | 22        285         50 |
     +--------------------------+
```

Step 2: Prepare to cluster by first applying principal component analysis to the data. Principal component analysis transforms the different dimensions of the (baseline) primary endpoint and participant age variables into standard normal variables or pure numbers----with means of zero and variances equal to unity. Then it assigns weights to the standard normal variables so that their inherent variances are reflected as much as possible in the resultant composite endpoint-age index. For this purpose, I use the first principal component (as long as it captures at least 60 percent of the variation); shown as comp1 in the Triheptanoin printout.

Triheptanoin case

```
Principal components/correlation                  Number of obs    =        22
                                                  Number of comp.  =         2
                                                  Trace            =         2
    Rotation: (unrotated = principal)             Rho              =    1.0000
--------------------------------------------------------------------------
       Component |   Eigenvalue   Difference         Proportion   Cumulative
 ----------------+---------------------------------------------------------
```

```
           Comp1 |      1.22935        .4587             0.6147       0.6147
           Comp2 |       .77065            .             0.3853       1.0000
    ------------------------------------------------------------------------------

Principal components (eigenvectors)

---------------------------------------------
        Variable |    Comp1     Comp2 | Unexplained
    -------------+--------------------+-------------
       sixmwbase |  -0.7071    0.7071 |           0
        agebegin |   0.7071    0.7071 |           0
```

Step 3. Transform the first principal component result into an index form with a value of 100 for one of the ID numbers. The index form is necessary because the CV for the raw component index values (Comp1) is undefined with its mean value of zero and standard deviation of unity.

Triheptanoin case

Below I show the raw Comp1 values; and in two steps, convert it to an index with ID number 16=100 (the highest pure number in the series). First, I multiply all the Comp1 values by 10 (the result shown in the column labelled Com1prelim). Then I add 84.14872 to all of the Comp1prelim numbers---bringing the index value of ID # 16 to a value of 100, as shown in the column labelled Comp1index.

```
     +---------------------------------------+
     | ID     Comp1   Comp1prelim   comp1iindex
     |---------------------------------------|
  1. |  1    1.106481    11.06481   95.21352 |
  2. |  2   -.6539913   -6.539913   77.60881 |
  3. |  3   -1.184959   -11.84959   72.29913 |
  4. |  4    .2963713    2.963713   87.11243 |
  5. |  5   -1.864375   -18.64375   65.50497 |
     |---------------------------------------|
  6. |  6   -1.620954   -16.20954   67.93919 |
  7. |  7   -1.051219   -10.51219   73.63653 |
  8. |  8   -1.126756   -11.26756   72.88116 |
  9. |  9   -.3347176   -3.347176   80.80154 |
 10. | 10   -2.080377   -20.80377   63.34496 |
     |---------------------------------------|
 11. | 11    .4094152    4.094152   88.24287 |
 12. | 12    .8933155    8.933155   93.08188 |
```

```
13. | 13    1.429851    14.29851   98.44723 |
14. | 14    .6658633    6.658633   90.80735 |
15. | 15    .2748339    2.748339   86.89706 |
    |---------------------------------------|
16. | 16    1.585128    15.85128        100 |
17. | 17    .1771293    1.771293   85.92001 |
18. | 18    1.030102    10.30103   94.44975 |
19. | 19    1.540478    15.40478    99.5535 |
20. | 20    .1823821    1.823821   85.97254 |
    |---------------------------------------|
21. | 21    .4464999    4.464999   88.61372 |
22. | 22    -.120503    -1.20503   82.94369 |
```

Step 4. Calculate the mean, standard deviation, and coefficient of variation (CV) for the entire baseline endpoint-age index (Comp1index). Most importantly, this CV shows the APBD heterogeneity before clustering.

Triheptanoin case: The CV for the entire index, before clustering, is 13.6 percent (11.1/84.1):

```
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
  comp1index |        22    84.14872     11.0876   63.34496        100
```

Step 5. Apply Kmedians clustering to the endpoint-age index. This step assigns the ID numbers of the participants to either Cluster 1 or Cluster 2, the ultimate goal.

Triheptanoin case: Assigns the participants to the two clusters.

```
    +---------------+
    | clusters   ID |
    |---------------|
 1. |        1   16 |
 2. |        1   14 |
 3. |        1    1 |
 4. |        1   17 |
 5. |        1   11 |
    |---------------|
 6. |        1   19 |
 7. |        1   15 |
```

```
 8. |        1   12 |
 9. |        1   20 |
10. |        1   22 |
    |---------------|
11. |        1   18 |
12. |        1   13 |
13. |        1    4 |
14. |        1   21 |
15. |        2    5 |
    |---------------|
16. |        2    8 |
17. |        2    2 |
18. |        2   10 |
19. |        2    7 |
20. |        2    3 |
    |---------------|
21. |        2    9 |
22. |        2    6 |
    +---------------+
```

Step 6. For the two clusters, calculate their coefficients of variation. Based upon the associated baseline median primary endpoint and participant ages, refer to the two clusters as either “severe APBD” or “mild APBD.”

Triheptanoin Case. In Table 1, page 4, the cluster of mild APBD is associated with considerably longer walking distances and somewhat younger aged participants than that of the severe APBD cluster. Most importantly, the coefficient of variations for both clusters are within the strict assay lab work standard.

Step 7. For internal validation, regress COMP1 as the dependent variable and the cluster variable as the independent variable.

Triheptanoin case: With this regression, the cluster variable is significant at the 95 percent confidence level—providing confidence in the clustering solution.

```
      Source |       SS       df       MS              Number of obs =      22
-------------+------------------------------           F(  1,    20) =   59.47
       Model |  1931.94907     1  1931.94907           Prob > F      =  0.0000
    Residual |   649.68517    20  32.4842585           R-squared     =  0.7483
-------------+------------------------------           Adj R-squared =  0.7358
       Total |  2581.63424    21  122.934964           Root MSE      =  5.6995
```

comp1index	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
clusters	-19.4805	2.526032	-7.71	0.000	-24.74971	-14.21129
_cons	110.713	3.652636	30.31	0.000	103.0938	118.3323