Identifying Longitudinal Patterns for Individuals and Subgroups: An Example with Adherence to Treatment for Obstructive Sleep Apnea

Article *in* Multivariate Behavioral Research · February 2015

DOI: 10.1080/00273171.2014.958211

CITATIONS
36

READS
280

4 authors, including:

Steven F Babbin
Geisel School of Medicine at Dartmouth
22 PUBLICATIONS **277** CITATIONS

SEE PROFILE

Wayne F Velicer
University of Rhode Island
244 PUBLICATIONS **52,672** CITATIONS

SEE PROFILE

Clete A Kushida
Stanford University
232 PUBLICATIONS **21,479** CITATIONS

SEE PROFILE

Multivariate Behavioral Research
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hmbr20

Identifying Longitudinal Patterns for Individuals and Subgroups: An Example with Adherence to Treatment for Obstructive Sleep Apnea

Steven F. Babbin[a], Wayne F. Velicer[b], Mark S. Aloia[c] & Clete A. Kushida[d]
[a] Department of Psychology, Geisel School of Medicine at Dartmouth
[b] Cancer Prevention Research Center, University of Rhode Island
[c] National Jewish Health and Philips/Respironics, Inc.
[d] Stanford Sleep Medicine Center
Published online: 17 Feb 2015.

Multivariate Behavioral Research, 50:91–108, 2015

ISSN: 0027-3171 print / 1532-7906 online
DOI: 10.1080/00273171.2014.958211

Routledge
Taylor & Francis Group

Identifying Longitudinal Patterns for Individuals and Subgroups: An Example with Adherence to Treatment for Obstructive Sleep Apnea

Steven F. Babbin
Department of Psychology, Geisel School of Medicine at Dartmouth

Wayne F. Velicer
Cancer Prevention Research Center, University of Rhode Island

Mark S. Aloia
National Jewish Health and Philips/Respironics, Inc.

Clete A. Kushida
Stanford Sleep Medicine Center

To improve complex behaviors such as adherence to medical recommendations, a better understanding of behavior change over time is needed. The focus of this study was adherence to treatment for obstructive sleep apnea (OSA). Adherence to the most common treatment for OSA is poor. This study involved a sample of 161 participants, each with approximately 180 nights of data. First, a time series analysis was performed for each individual. Time series parameters included the mean (average hours of use per night), level, slope, variance, and autocorrelation. Second, a dynamic cluster analysis was performed to find homogenous subgroups of individuals with similar adherence patterns. A four-cluster solution was found, and the subgroups were labeled: Great Users (17.2%; high mean and level, no slope), Good Users (32.8%; moderate mean and level, no slope), Low Users (22.7%; low mean and level, negative slope), and Slow Decliners (moderate mean and level, negative slope, high variance). Third, participants in the identified subgroups were compared to establish external validity. These steps represent a Typology of Temporal Patterns (TTP) approach. Combining time series analysis and dynamic cluster analysis is a useful way to evaluate longitudinal patterns at both the individual level and subgroup level.

Tracking individuals over time creates unique and valuable opportunities to study complex behaviors. Adherence to medical recommendations is a complicated behavior that involves changes over time at both the group and individual level. To better understand behaviors such as adherence, analytic methods should be employed that best evaluate the intricacies of the behavior. Traditionally, comparisons between groups of individuals are emphasized (Molenaar, 2004). However, individual-level longitudinal patterns reveal trends that cannot be assessed with group-level comparisons. In this article, a hybrid approach is described that assesses both individual-level and group-level patterns using previously established statistical tools. More specifically, time series analysis is used to identify patterns of adherence to treatment with a sample of participants with obstructive sleep apnea (OSA). Then participants with similar adherence trajectories are grouped together with a type of cluster analysis. Finally, demographic, psychosocial, and sleep apnea variables are evaluated to both validate the approach and reveal which factors are associated with subgroup membership.

Adherence to treatment is poor in some areas, particularly for sleep disorders. In an extensive review of adherence to medical treatments that spanned 50 years and 569 studies,

Correspondence concerning this article should be addressed to Steven F. Babbin, Department of Psychiatry, Geisel School of Medicine at Dartmouth, Rivermill Complex, 85 Mechanic Street, Suite B3–1, Lebanon, NH 03766. E-mail: steven.f.babbin@dartmouth.edu

average adherence rates across 17 disease groupings ranged from 65.5% to 88.3% (DiMatteo, 2004). The poorest average adherence rate (65.5%) was for treatments for sleep disorders. Better understanding adherence is vital to reducing preventable mortality and wasted health care money (Oldridge, 2001). OSA is a significant sleep disorder with multiple consequences. At least 2–4% of middle-aged adults suffer from OSA, with higher rates in men than in women (Young et al., 1993). The disorder is characterized by fragmented sleep resulting from repeated cessations of breathing during sleep. Widely known consequences of OSA include daytime sleepiness, decreased cognitive performance, irritability, and depression. Higher rates of mortality are also associated with the disorder due to its impact on the cardiovascular system, including an increased risk of hypertension (Lavie, Herer, & Hoffstein, 2000; Young et al., 1997) and other cardiovascular disease (Shahar et al., 2001). The usual treatment for OSA is nasal positive airway pressure (PAP), and treatment is delivered by a PAP device. Such machines are connected to masks that the patient wears while sleeping, and they effectively reduce symptoms of OSA by administering pressure through the nose to keep the airway open. There is evidence of a dose-response relationship with nightly PAP use; at least 4 hours of PAP use is generally necessary for symptom reduction, and improvements continue up to 7 hours of PAP use (Weaver & Grunstein, 2008; Weaver et al., 2007). PAP devices are configured to deliver the prescribed amount of pressure, and they automatically record nightly usage data. Studying adherence to treatment for OSA is important because adherence rates are poor, with many failing to use their PAP device every night or discontinuing use completely (Engleman & Wild, 2003; McArdle et al., 1999). To improve poor adherence rates, a better understanding of adherence to PAP is needed. Since PAP machines automatically record usage data, they provide excellent monitoring of adherence to treatment. These data include the amount of time per night they are in use, which provides an ideal variable for adherence. The present study utilizes data from a clinical trial of variations of PAP (Kushida et al., 2011) where participants were monitored for 180 days.

One approach to better understanding adherence is to consider individual differences in patterns of adherence rather than just group-level statistics (Aloia et al., 2008; Dunbar-Jacob & Sereika, 2001), since group-level statistics do not fully capture the complexity of within-individual variability. The value of group-level analysis versus individual-level analysis involves the distinction between nomothetic and idiographic methods (Hamaker & Dolan, 2009; Molenaar, 2004; 2008). With a nomothetic approach, the focus is to analyze data from a large number of individuals on one or a small number of occasions (inter-participant). The goal of the nomothetic approach is to describe general laws, or general descriptions of what is always common to all (Lamiell, 1998; Windelband, 1998). With an idiographic approach, the focus is to analyze a single individual on a large number of occasions (intra-participant variability). The goal of the idiographic approach is to intensely describe the individual and the unique properties of the individual (Molenaar, 2004; Windelband, 1998). Both nomothetic and idiographic approaches have strengths and weaknesses. Idiographic methods are used extensively in a number of fields, particularly business and econometrics (Domowitz & El-Gamal, 2001; Judge, 1988; Lütkepohl & Krätzig, 2004), as well as engineering, biology, chemistry, and physics (Bendat & Piersol, 1980; Shumway & Stoffer, 2010; Strogatz, 1994). Nomothetic, group-level methods are used extensively in psychological research; methods such as analysis of variance, multiple regression, and logistic regression are well known and are ubiquitous in the literature. One reason for the popularity of these methods is that they produce results that are widely believed to be generalizable to the population (Hamaker & Dolan, 2009; Molenaar, 2004; 2008). Idiographic methods are less popular (Molenaar, 2004), and since they focus on individuals, they are not as generalizable to the population. However, while nomothetic methods may relate to the population, they may not relate to individuals. Nomothetic approaches focus on group averages, and as a result individual differences represent noise that is removed. A calculated summary statistic (such as a group mean) may not describe any individual in the data since group-level statistics blend individuals together. Inferring from group to individual constitutes an ecological fallacy (Molenaar, 2008). Thus, both nomothetic and idiographic approaches have issues with generalizability.

The importance of considering individuals is highlighted by the ergodic theorems. These theorems specify the general mathematical conditions that must be met by measurable dynamic processes in order for the analysis of interindividual variation (group-level analyses) to yield equivalent results to the analysis of intraindividual variation (individual-level analyses) (Keller, 1998; Molenaar, 2004; 2008; Petersen, 1983). The two conditions of the ergodic theorems are: (1) each individual trajectory has to be stationary, and (2) each individual trajectory must have the same statistical characteristics. Meeting these conditions is extremely challenging in practice. The first condition refers to the generating function of the process across time; to be ergodic, each individual must have a consistent (stationary) generating function over time (constant level with no time-varying trends). An example of a statistic that describes a generating function is an autocorrelation, or the sequential dependence of a process. The second condition refers to whether or not individuals can be considered equivalent in terms of statistical characteristics (e.g., slope, mean, level, and variance). This condition implies that a population must be homogenous, with no individual differences in parameters. If these conditions are violated, then only considering group-level analyses will not provide a comprehensive overview of the data. Molenaar (2004) argues that most psychological variables are non-ergodic. Developmental processes, for

example, are never ergodic (Molenaar, 2008) as such processes always have time-varying generating functions. Thus to fully explore the complexities of adherence to OSA, a complex behavior that often varies greatly among individuals (Aloia et al., 2008), statistical techniques reflecting an idiographic approach are warranted. Idiographic methods offer unique ways to think about research questions because they emphasize intraindividual variation rather than just interindividual variation.

While all individuals may have unique behavior patterns over time, there may be homogenous subgroups of individuals that better address the ergodic theorems. Identifying individuals that share similar statistical trajectories will create subgroups that are more generalizable to a particular population. The nomothetic approach of grouping homogenous individuals together is valuable for applications such as adherence because characteristics of subgroups can be studied to reveal larger trends. In the present study, a hybrid approach will be utilized that involves both idiographic and nomothetic approaches. With this approach, individuals are first evaluated at the individual level, using an idiographic technique such as time series analysis, and then similar temporal patterns are aggregated, using a nomothetic technique, such as cluster analysis. In this article we are referring to this general approach as a Typology of Temporal Patterns (TTP). There are many examples of TTP in the literature including studies of patterns of daily smoking (Chandra, Shiffman, Scharf, Dang, & Shadel, 2007; Hoeppner, Goodwin, Velicer, Mooney, & Hatsukami, 2008), daily alcohol use (Harrington, Velicer, & Ramsey, 2014), and OSA (Aloia et al., 2008).

The first part of the present study involved time series analysis. It was used to characterize individual patterns of adherence over time. Time series methodology has previously been utilized to study participants prescribed PAP machines for OSA. In Aloia et al. (2008), the authors used time series analysis to examine adherence patterns in 71 participants with OSA. A separate time series analysis was performed for each participant, and the authors looked at four parameters, which are also considered in the present study: level of the series, variance of the series, slope of the series, and autocorrelation. Each individual's time series were graphed, and the authors used these graphs to classify the participants into groups of similar adherence patterns. Seven groups were identified: (1) Good Users, (2) Slow Improvers, (3) Slow Decliners, (4) Variable Users, (5) Occasional Attempters, (6) Early Drop-outs, and (7) Non-Users. A goal of the present study was to replicate some findings from this previous use of time series methodology by examining adherence to PAP in a new sample.

The second part of the present study involved a type of cluster analysis known as dynamic cluster analysis. It was used to identify homogenous subgroups of individuals with similar patterns of adherence. Cluster analysis is an exploratory technique used to identify subgroups in data that are not known a priori. In traditional cluster analysis, homogenous subgroups are created from a number of variables measured on one occasion (Everitt, Landau, Leese, & Stahl, 2011). In dynamic cluster analysis, homogenous subgroups are created from one variable measured across a number of occasions (Hoeppner et al., 2008; Norman, Velicer, Fava, & Prochaska, 1998; Prochaska, Velicer, Guadagnoli, Rossi, & DiClemente, 1991). In the present study, homogenous subgroups were identified by clustering hours of PAP use over time. Thus, rather than using qualitative visual identification to classify participants into subgroups with similar trajectories across time, as in Aloia et al. (2008), a quantitative approach was used.

The third and final part of the present study compared demographic, sleep apnea, and psychological variables to both establish external validity of the clusters and to explore potential predictors of adherence. There are many predictors of PAP adherence, and while such variables have been investigated in many past studies (e.g., Aloia, Arnedt, Stepnowsky, Hecht, & Borrelli, 2005; Collard, Pieters, Aubert, Delguste, & Rodenstein, 1997; Pelletier-Fleury, Rakotonanahary, & Fleury, 2001; Stepnowsky, Marler, & Ancoli-Israel, 2002), such variables have never been used to compare homogeneous subgroups of participants. Identification of factors that influence specific subgroups will lead to a better understanding of what affects adherence. Investigating the impact of demographics is particularly important for race and ethnicity, especially as it relates to the areas of OSA and PAP adherence (Weaver & Grunstein, 2008; Young, Peppard, & Gottlieb, 2002). Results from studies comparing adherence to PAP among ethnicities have been mixed. For example, there is evidence that Blacks or African Americans demonstrate less adherence to PAP than Whites (Joo & Herdegen, 2007), but others have found no difference (Scharf, Seiden, DeMore, & Carter-Pokras, 2004). The issue of PAP adherence among different ethnicities is complex due to differences in socioeconomic status and cultural differences in lifestyle (Weaver & Grunstein, 2008; Young et al., 2002). Psychological variables are also of particular importance because they are dynamic and can be changed; significant predictors can become the focus of future interventions to increase adherence rates to PAP treatment. Predicting subgroups with variables not included in the clustering also establishes external validity.

The goals of the present study are to describe individual patterns of PAP use over time, quantitatively create subgroups of individuals with similar patterns, and identify variables that are related to the subgroups. Part of the study will also replicate a previous study of adherence to treatment for sleep apnea (Aloia et al., 2008) in a larger sample. Ultimately, the present study assesses the utility of combining time series analysis and dynamic cluster analysis to identify patterns over time.

METHOD

Participants

The sample ($n = 161$) in the present study was from an international, multicenter, double-blinded, randomized clinical trial examining the influence of variations of PAP on efficacy, adherence, and sleep outcomes (Kushida et al., 2011). Participants had already been diagnosed with OSA after a routine clinical visit and PAP titration polysomnography (PSG). Participants were recruited from five different sleep centers: Stanford University, Stanford, CA; Shands and University of Florida Sleep Disorders Center, Gainesville, FL; Charité-Universitätsmedizin Berlin Interdisciplinary Center of Sleep Medicine, Berlin, Germany; Gaylord Hospital, Wallingford, CT; and the Sleep Disorders Center of Alabama, Birmingham, AL. For a description of the inclusion and exclusion criteria, please refer to Kushida et al. (2011). Eligible participants were required to provide written informed consent to participate in the study. The participants were randomly selected to receive either: (1) A-Flex, which is automatically adjusted positive airway pressure (APAP) with a comfort feature, (2) continuous positive airway pressure (CPAP), and (3) APAP-derived optimal pressure for CPAP ($CPAP_{APAP}$).

The sample was 72.7% male. Participants ranged in age from 21–75 years (mean age = 48.9). The sample included multiple ethnicities: White or Caucasian (79.1%); Black or African American (11.0%); Hispanic or Latino (2.3%); Asian (2.3%); other (1.2%), and no response (4.1%).

Procedure

After recruitment and randomization, participants underwent a full-night PSG to document the baseline efficacy of the treatment. While different variations of PAP were used, all participants had the same PAP device (REMstar Auto M-Series, Philips Respironics, Murrysville, PA). Demographic, sleep apnea, and psychological variables were taken at baseline. Adherence data including demographic, sleep apnea, and psychological variables were also collected during follow-up visits at 30, 90, and 180 days after the start of treatment. The authors of the original study did not find any significant differences between groups on primary study outcomes, including adherence rates and mean hours of PAP use per night (Kushida et al., 2011). The study lasted for 180 days; some participants used their machines longer, and the PAP machines did not stop recording after Day 180. There were 168 participants randomized into the study. Some participants dropped out early (attrition), with 140 participants completing the study (83.3%). For this secondary data analysis, data were de-identified.

Measures

The primary measure of interest in this study was the amount of PAP use, specifically, the amount of time (in hours) that a user was connected to the machine while it was administering the prescribed pressure. In addition to this dependent variable for adherence, there were a number of other measures involved in this study. Important demographics that were collected included: age, height, weight, body mass index (BMI), blood pressure, heart rate, education, and ethnicity. Important sleep apnea variables included: severity, as determined by baseline apnea-hypopnea index (AHI, a widely accepted index of apneas); prescribed PAP pressure (cm H_2O); sleepiness, measured by the Epworth Sleepiness Scale (ESS; Johns, 1992); functional outcomes related to sleepiness, measured by the functional outcomes of sleep questionnaire (FOSQ; Weaver et al., 1997); and vigilance, measured by the psychomotor vigilance task (PVT; Dinges & Powell, 1985). Important psychosocial variables included: attitudes toward use or self-efficacy, measured by the attitudes toward use questionnaire (ATUQ; Aloia et al., 2005; Stepnowsky, Marler, & Ancoli-Israel, 2002), and perceived sleep quality, satisfaction, comfort, and benefit, as measured by a Visual Analog Scale (VAS; Kushida et al., 2011).

Typology of Temporal Patterns

The present study is an example of a TTP with three phases. First, time series analysis was used to assess individual patterns. Second, dynamic cluster analysis was used to aggregate similar patterns. Third, group comparisons were performed to validate the clustering. In this example, sample size (n) represents the number of observations (days); the number of replications (s) represents the number of participants, and the number of subgroups (k) represents the number of clusters.

Time Series Analysis

The first part of the present study employed time series analysis. Time series analysis is an idiographic method appropriate for single participants measured repeatedly at regular intervals (Glass, Willson, & Gottman, 1975; Velicer & Molenaar, 2013). It is a regression-based technique that utilizes autoregressive moving average (ARIMA) models of the order (p, d, q) to model dependence. The p (the autoregressive component) involves dependency, the d (the differencing component) involves the amount a series has to be differenced to make it stationary, and the q (the moving average component) involves the amount of shock to the time series that results from previous observations. All available data were used for the time series analyses.

Analyses employed SAS Proc ARIMA with the General Transformation Approach (Velicer & McDonald, 1984). This approach uses an ARIMA (5, 0, 0) model for all analyses, which specifies the estimation of five autoregressive parameters (series is partially predictable from previous five observations; $p = 5$) and zero differencing and moving average estimates ($d = 0$, $q = 0$). This approach was utilized for multiple reasons. First, this approach avoids the often-problematic identification step; even among the highly

trained, accuracy for ARIMA model specification is often very poor (Velicer & Harrop, 1983). Second, the General Transformation Approach has been shown to be appropriate for most data encountered in the behavioral sciences (Harrop & Velicer, 1985). Third, this approach was used in the previously described study that used time series analysis to examine adherence to PAP (Aloia et al., 2008). Fourth, using the same model for all individuals increases the interpretability across individuals. While the primary study lasted 180 days, the *n* (number of days) was often larger because participants continued to use the devices.

Dynamic Cluster Analysis

The second part of the present study employed dynamic cluster analysis. Cluster analysis groups individuals into categories based on the level (means), scatter (variances), and shape of the patterns (Cronbach & Gleser, 1953). Typically, cluster analysis is used to group multiple variables on a single occasion. In this study, which employed dynamic cluster analysis, homogenous subgroups were created from one variable (hours of PAP use) across time; the 180 days were treated as 180 variables for clustering. Note that the time series parameters were not used for the clustering; this was because the time series results do not include the shapes of the raw patterns, which are essential to clustering with a distance metric. The clusters, based entirely on the hours of PAP use and estimated independently from the time series results, provide both a meaningful way to organize the time series results and a degree of cross-validation; observing similar time series results within each cluster provides a validity check for both the time series analyses and for the clustering, as observing dramatically different individual patterns would indicate either an issue with the individual-level time series or with the clustering method. The method of dynamic cluster analysis has been shown to be appropriate for identifying subpopulations across time, although it is sensitive to changes in level and less sensitive to changes in shape (Dumenci & Windle, 2001).

Analyses employed SAS Proc CLUSTER using Ward's minimum variance clustering (Ward, 1963) and a squared Euclidean distance metric. This clustering method has demonstrated good performance in several simulation studies (Blashfield, 1976; Milligan, 1980; Overall, Gibson, & Novy, 1993). Ward's minimum variance method is a type of agglomerative hierarchical clustering (Everitt et al., 2011; Jain, Murty, & Flynn, 1999) where *agglomerative* refers to the algorithmic process of starting with all distinct patterns and then successively merging clusters together, and where *hierarchical* refers to how the method joins clusters in a nested series. With Ward's method, the distance between two clusters is defined by the sum of squares between these clusters, summed across all clustering variables; at each level of the hierarchy, the goal of the merging algorithm is to minimize the increase in the total within-cluster error sum of squares (Everitt et al., 2011; Ward, 1963). Ward's minimum variance method was not modified in any way to perform the dynamic clustering; the distinction of "dynamic" clustering involved specifying the 180 days as the clustering variables (rather than multiple variables at a single time point). Multiple methods were utilized to determine the number of clusters. The cluster number solution was chosen based on statistical indices, including the inverse scree test (Lathrop & Williams, 1987), the cubic clustering criterion (Sarle, 1983), and the pseudo-F test (Calinski & Harabasz, 1974). Visual inspections of cluster means over time were also used to compare cluster solutions. Agglomerative hierarchical clustering methods, in general, lack any strong, reliable indices of the number of clusters (Dumenci & Windle, 2001; Everitt et al., 2011; Mun, Windle, & Schainker, 2008; von Eye, Mun, & Indurkhya, 2004). In the present study, the validity of the number of clusters was examined by comparing variables not involved in the clustering.

Group Comparisons

The third part of the present study involved comparing variables among the identified clusters. Analyses were performed to explore the predictive usefulness of adherence variables and, as outlined above, to establish external validity of identified clusters. To test for differences between the clusters at each time point, multivariate analysis of variance (MANOVA) was performed with cluster membership as the independent variable and the adherence-related variables as the dependent variables. For statistically significant MANOVAs, differences across clusters on the variables were tested with a series of one-way ANOVAs (with follow-up Tukey HSD tests). Chi-square tests were performed for categorical variables. Variables that were examined include: demographic variables collected at baseline (age, height, weight, BMI, blood pressure, heart rate, gender), physiological sleep apnea variables at baseline (AHI, PAP pressure, snoring, nasal allergies, nasal congestion), other sleep apnea variables at 0, 30, 90, and 180 days (ESS, FOSQ, PVT), and psychosocial variables at 0, 30, 90, and 180 days (ATUQ, VAS; VAS not taken at baseline). Psychosocial variables are of particular interest because they are dynamic and have the possibility of being changed. Effect size magnitudes were interpreted using Cohen's (1992) guidelines for practical significance (small = .010, medium = .059, large = .138).

Missing Data

Missing data for participants was the result of a technical problem in recording use data or a failure to power the PAP machine. Since missing data may reflect use or nonuse, missing data estimation approaches were employed. For the time series analyses, missing data were estimated with the maximum likelihood (ML) procedure (Jones, 1980) in SAS. This procedure has been shown to be very accurate for ARIMA time series analysis, even with up to 40% of the data

missing (Velicer & Colby, 2005). Missing data estimation for dynamic cluster analysis required a different procedure due to the details of the clustering algorithm. Since Ward's method involves a distance metric, actual data points were necessary for clustering; thus, imputation was necessary to estimate missing data (Everitt et al., 2011; Wang, Smith, & Hyndman, 2006; Jain et al., 1999). Multiple imputation (MI; Rubin, 1987; Schafer & Graham, 2002) was used in SAS to estimate missing data for clustering with Ward's method. This was performed as a step prior to clustering. Both ML and MI procedures assume that data are missing at random (MAR). For the dynamic cluster analysis, the validity of MI was assessed by both comparing the results of the complete cases to the results with imputed data and by evaluating the consistency of classification of individuals into the same clusters across the imputed data sets.

RESULTS

Time Series Analysis

A separate time series analysis was performed for 161 participants ($s = 161$) using SAS Proc ARIMA. Summary statistics for all individual time series analyses are presented in Table 1; aggregated time series results are presented in Table 2. Five summary statistics are included in both tables: mean, level, slope, autocorrelation (r_1), and variance. The mean can be interpreted as the average hours of PAP use per night. The level can be interpreted as the intercept of a straight line fit to the data, or the number of hours of PAP use at the beginning of the trial. The slope estimates the rate of change in PAP use over time. The slope of the series reveals whether or not an individual is increasing, decreasing, or consistently adhering to PAP. The autocorrelation (first order autoregressive parameter), is a measure of dependency. Autocorrelations, like product–moment correlation coefficients, range from –1 to 1. Table 2 also includes estimates of cumulative change in PAP use over the study based on the product of the average number of days of PAP and the slope.

Time series results varied greatly among the participants. Due to this heterogeneity, trends in mean, level, slope, and variance will be discussed in terms of identified subgroups. At the level of the whole sample, slope and autocorrelation are of particular importance; the slope describes the trend over time, and the autocorrelation provides information about the generating function, i.e., what is driving the trend. The slope was significant ($p < 0.05$) for nearly half of the sample (75 of 161; 46.6%). Of the statistically significant slope parameters, approximately 25.3% (19 of 75) were positive (increasing use over time) and 74.7% (56 of 75) were negative (decreasing use over time). The differences in slope at the group level suggest there are different underlying longitudinal patterns of PAP adherence. The autocorrelation (r_1) was significant ($p < 0.05$) for 39.1% (63 of 161) of the sample. However, of the participants that had a significant autocorrelation, only 27% (17 of 63) had an autocorrelation greater than $r_1 > 0.30$. Thus, for most participants (89.4%; 144 of 161), the amount of PAP use on one night was not meaningfully influenced by prior nights.

Dynamic Cluster Analysis

Dynamic cluster analysis was used to identify homogenous subgroups using SAS Proc CLUSTER. Cluster analyses were performed with and without first estimating missing data. All analyses had the inclusion criteria of $n = 180$ days. First, only complete cases were used. Out of 161 possible participants, 74 participants (46.0%) that had $n = 180$ days without missing data. Based on the statistical indices (inverse scree test, cubic clustering criterion, pseudo-F test) and visual inspection of the cluster profiles, a four-cluster solution ($k = 4$) was chosen to best describe the data. Alternative numbers of clusters (two, three, five, and six) were also considered, but the four-cluster solution provided the best fit and produced the most interpretable graph of cluster means. Next, SAS Proc MI was employed to estimate missing data in the sample. Participants that caused convergence problems in the MI algorithm, due to excessive missing data, were removed from the dynamic cluster analysis and not used in any future analyses. Out of 161 possible participants, 128 participants (79.5%) were able to be clustered with missing data estimated. Five datasets with imputed missing values were generated. Dynamic cluster analysis was performed on the five samples, and the results consistently replicated the results of the complete case analysis: for all samples, a four-cluster solution was chosen ($k = 4$). Classification of individuals into the same clusters was also extremely consistent; 127 of the 128 participants (99.2%) were classified into the same clusters across all five imputations. The consistency of the number of clusters, as well as the consistency of individual classifications, provides evidence of the validity of the missing data estimation, and participants were assigned to clusters based on the datasets with imputed data. The 1 participant with some variation in classification (ID 222–36) was included in the Low Users cluster (see below) because it was classified into this cluster in 4 out of 5 iterations. Final cluster membership (which was also used for group comparisons) is included in Table 1.

The means of the four clusters are graphed in Figure 1. These clusters represent four different patterns of adherence to PAP. The cluster profiles are similar to ones identified visually by Aloia et al. (2008). The four clusters are: Great Users, Good Users, Low Users, and Slow Decliners. These descriptive labels are based primarily on mean, level, and slope parameters. A plot of the average for the entire sample is also included in Figure 1 to demonstrate that the clusters reveal distinct trends from the overall mean. The overall mean appears to be decreasing slightly, and it completely obscures two important subgroups. The Great

TABLE 1
Time Series Parameter Estimates (N, Mean, Level, Variance, Slope, and Autocorrelation r_1) for All Individuals

ID	N	Mean	Level	Slope	A. (r_1)	Var.
			Great Users ($n_1 = 22$)			
111–26	187	6.99	6.37	0.007*	0.231*	2.20
111–16	179	7.01	6.72	0.003	−0.146	2.61
111–9	180	7.11	6.16	0.011*	0.080	1.86
222–34	175	7.76	7.76	−0.0001	0.115	0.68
222–25	186	7.24	6.87	0.004	0.144	2.07
222–10	192	6.95	7.01	−0.001	0.124	0.89
222–4	188	6.98	7.88	−0.009*	0.146	5.82
222–3	207	7.12	7.34	−0.002*	−0.002	0.65
222–2	156	8.15	8.65	−0.006	−0.019	2.50
333–10	181	7.85	6.87	0.011*	0.027	4.91
444–36	181	7.08	7.13	−0.001	−0.033	1.35
444–32	181	7.26	6.74	0.006*	−0.105	1.14
444–25	180	7.44	7.50	−0.001	0.243*	1.17
444–23	185	6.60	6.89	−0.003*	0.157*	0.78
444–14	180	7.26	7.54	−0.003	−0.047	0.91
444–2	167	6.90	6.67	0.002	−0.094	5.38
555–35	181	6.93	6.27	0.007*	−0.044	2.69
555–32	205	7.77	6.93	0.008*	0.184*	4.34
555–30	196	6.91	6.53	0.004	−0.097	2.43
555–27	179	6.93	6.42	0.006*	0.056	1.42
555–26	182	8.00	7.44	0.006*	0.054	1.53
555–22	213	8.22	8.00	0.002	0.015	1.49
			Good Users ($n_2 = 42$)			
111–12	174	6.00	6.45	−0.005	−0.116	4.3
111–3	181	5.15	4.98	0.002	0.069	6.37
111–4	192	5.17	4.37	0.008*	0.028	7.62
111–7	189	6.70	7.43	−0.008*	−0.003	3.96
222–14	184	6.78	6.62	0.002	−0.075	2.96
222–15	163	5.76	5.96	−0.003	−0.060	0.97
222–18	181	5.28	5.57	−0.003	−0.028	7.06
222–19	187	6.34	6.27	0.001	0.156*	2.95
222–20	181	6.48	5.64	0.008*	0.004	4.13
222–21	186	5.65	4.67	0.011*	0.111	2.53
222–24	186	4.91	4.26	0.008	0.190*	5.2
222–31	198	4.88	5.83	−0.010*	0.146*	4.99
222–35	196	5.66	5.68	−0.0001	0.122	4.78
222–39	182	6.10	6.83	−0.008*	0.047	4.92
222–40	183	5.03	2.74	0.025*	0.361*	5.47
333–13	184	4.86	4.41	0.005	0.065	4.5
333–18	181	5.33	3.89	0.016*	0.127	7.46
333–22	181	6.30	6.75	−0.005	0.510*	3.7
333–25	180	5.87	4.85	0.012	0.284*	7.89
333–27	179	5.37	4.84	0.006	0.130	8.92
333–30	181	5.36	4.48	0.010	0.033	8.6
333–31	223	5.94	5.38	0.005*	0.119	1.69
333–5	181	5.03	5.23	−0.002	−0.023	8.13
333–7	172	5.32	4.96	0.004	0.101	1.62
444–10	181	5.65	6.16	−0.006*	−0.097	4.78
444–16	181	5.60	5.94	−0.004	0.522*	3.7
444–27	185	5.82	7.39	−0.017*	0.289*	2.96
444–33	189	5.51	5.23	0.003	0.132	2.21
444–39	177	5.64	6.49	−0.009*	0.168*	2.26
444–8	181	6.39	4.97	0.016*	0.217*	5.85
555–12	184	5.97	5.97	−0.00004	0.063	1.42
555–14	183	5.34	5.52	−0.002	0.064	1.34
555–15	182	6.42	6.53	−0.001	0.057	1.63
555–2	183	5.69	5.55	0.001	0.079	3.75
555–20	173	5.68	5.76	−0.001	0.134	3.42
555–21	181	6.40	6.71	−0.003*	−0.027	1.16

(Continued on next page)

TABLE 1
Time Series Parameter Estimates (N, Mean, Level, Variance, Slope, and Autocorrelation r_1) for All Individuals *(Continued)*

ID	N	Mean	Level	Slope	A. (r_1)	Var.
555–24	179	5.39	5.51	−0.001	0.171*	5.84
555–29	193	4.68	4.22	0.005*	0.193*	1.07
555–3	179	6.65	5.99	0.007*	0.366*	2.27
555–31	191	6.43	7.32	−0.009*	0.083	3.19
555–4	179	6.02	6.39	−0.004	0.171*	7.11
555–6	179	5.51	4.83	0.008*	0.220*	2.7
			Low Users ($n_3 = 29$)			
111–1	185	1.81	2.19	−0.004	0.061	2.67
111–10	180	3.71	4.94	−0.013*	0.197*	5.22
111–18	181	1.73	3.65	−0.018*	0.391*	3.94
111–2	191	2.49	3.82	−0.014*	0.400	5.83
111–20	205	1.67	3.77	−0.020*	0.134	1.36
111–21	177	3.38	3.25	0.002	0.218*	4.72
111–5	209	0.50	1.93	−0.013*	0.113	1.14
222–05	193	2.78	2.38	0.004	0.067	4.10
222–06	181	2.67	3.35	−0.007	0.096	6.80
222–1	187	3.59	6.62	−0.030*	0.413*	3.38
222–16	169	2.97	3.28	−0.004	0.143	6.82
222–28	195	2.13	3.93	−0.018*	0.170*	2.15
222–30	230	2.15	4.26	−0.018*	0.178*	3.36
222–32	181	3.34	3.39	−0.002	0.084	6.67
222–36	188	3.31	4.57	−0.012	0.249*	7.59
222–38	187	2.01	3.19	−0.012*	0.142	4.53
333–1	178	0.88	3.12	−0.023*	0.417*	0.95
333–15	185	2.24	2.67	−0.005	0.180*	4.25
333–16	154	2.25	2.31	−0.001	0.099	5.18
333–17	181	3.37	6.97	−0.039*	0.019	5.04
333–19	184	3.44	4.21	−0.008	0.005	3.60
333–29	182	2.95	4.25	−0.014*	0.133	3.49
333–32	175	1.83	2.53	−0.009	0.323*	4.75
333–34	173	2.24	4.70	−0.028*	0.139	4.47
444–28	175	1.63	2.07	−0.004	0.281*	3.17
444–29	195	1.25	2.27	−0.010*	0.096	2.14
444–40	184	3.06	5.08	−0.022*	0.056	4.84
555–28	176	2.86	5.61	−0.030*	0.238*	5.46
555–34	188	2.74	2.02	0.008	0.115	4.36
			Slow Decliners ($n_4 = 35$)			
111–14	181	5.19	4.95	0.003	0.184*	11.65
111–28	182	5.10	6.33	−0.014*	0.199*	3.05
111–8	170	3.88	6.00	−0.025*	0.111	5.98
222–12	180	4.45	5.73	−0.014*	0.097	5.38
222–17	200	4.74	5.86	−0.010	0.176*	13.07
222–22	194	4.12	5.78	−0.017*	0.097	3.51
222–23	180	4.03	3.51	0.006	0.049	4.02
222–27	185	4.42	4.88	−0.001	0.341*	5.07
222–29	180	3.12	4.18	−0.011	0.335*	3.11
222–33	203	4.16	5.59	−0.015*	0.184*	2.63
222–37	182	4.12	5.83	−0.019*	0.133	1.61
333–12	181	4.27	3.90	0.004	−0.016	8.34
333–14	182	4.55	5.13	−0.006	0.236*	4.15
333–20	178	4.37	4.76	−0.004	0.069	6.84
333–23	179	4.51	4.87	−0.004	0.167*	3.24
333–24	182	4.12	5.07	−0.010*	−0.018	7.42
333–26	181	4.83	4.68	0.002	0.361*	5.58
333–33	180	3.63	4.51	−0.010	0.233*	4.56
333–4	190	5.22	7.89	−0.028*	0.142	7.81
333–9	181	3.61	4.56	−0.010*	−0.165*	3.48
444–13	186	5.04	5.76	−0.006	0.355*	6.37

(Continued on next page)

TABLE 1
Time Series Parameter Estimates (N, Mean, Level, Variance, Slope, and Autocorrelation r_1) for All Individuals *(Continued)*

ID	N	Mean	Level	Slope	A. (r_1)	Var.
444–17	179	4.16	5.17	−0.011*	0.091	4.95
444–18	179	5.06	6.26	−0.013*	0.213*	4.35
444–24	180	4.55	5.74	−0.013*	0.240*	4.10
444–26	181	4.14	5.59	−0.015	0.292*	7.89
444–4	181	4.13	4.61	−0.005	0.233*	6.66
444–7	181	3.25	2.93	0.003	0.012	2.70
444–9	183	4.29	4.85	−0.006	0.030	4.34
555–13	170	3.95	2.27	−0.019*	0.416*	3.42
555–16	180	4.38	5.61	−0.013*	0.297*	2.73
555–17	184	4.55	4.33	0.002	0.089	4.80
555–18	183	4.72	3.39	0.015	0.226*	4.97
555–19	183	5.41	5.82	−0.005	−0.012	6.95
555–36	178	3.82	4.25	−0.004*	−0.049	2.95
555–7	180	3.91	5.98	−0.023*	0.258*	3.37
			Not clustered due to excessive missing data (n = 33)			
111–11	87	1.67	2.93	−0.023	0.365*	5.13
111–17	163	1.87	2.06	−0.002	−0.167*	2.18
111–19	20	1.25	1.72	0.025	0.268	1.54
111–22	34	2.83	4.62	−0.091*	−0.075	2.58
111–24	129	3.76	3.58	0.003	0.223*	5.22
111–27	31	3.59	4.36	−0.055	−0.051	3.96
111–6	119	0.43	2.17	−0.013	0.075	0.94
222–11	28	3.57	7.58	−0.316	−0.071	8.45
222–13	181	1.24	1.81	−0.014*	−0.003	2.55
222–26	12	6.17	5.53	0.045	−0.569	2.90
222–8	185	2.94	—	—	—	—
333–11	13	6.13	—	—	—	—
333–2	90	5.49	4.17	0.029	0.446*	4.11
333–21	178	1.29	2.01	−0.009*	0.127	2.25
333–28	163	1.67	2.94	−0.015*	0.058	5.05
333–3	90	5.75	2.59	0.060*	0.407*	2.92
333–6	7	4.91	—	—	—	—
333–8	87	4.50	6.10	−0.036*	0.253*	7.51
444–11	89	5.83	7.66	−0.041*	−0.082	4.02
444–12	89	2.97	5.44	−0.056*	0.166	3.73
444–21	15	1.81	2.16	−0.041	−0.24	1.37
444–22	42	6.45	7.21	−0.036	−0.002	3.89
444–30	101	0.47	2.62	−0.033	0.455*	0.53
444–31	8	3.62	—	—	—	—
444–38	33	0.46	—	—	—	—
444–5	30	2.50	5.54	−0.190*	0.135	2.45
555–1	26	2.75	3.24	−0.028	−0.082	2.49
555–10	20	1.48	3.15	−0.170*	−0.041	3.25
555–11	22	2.18	3.51	−0.088*	−0.555	0.30
555–23	41	5.13	4.70	0.022	0.427*	6.22
555–25	116	3.47	5.98	−0.047*	0.220*	3.07
555–33	83	3.29	4.60	−0.010	0.310*	4.24
555–8	149	1.21	2.77	−0.022*	0.258*	4.16

Note. Estimation algorithm did not converge.
*$p < 0.05$.

Users and the Low Users make up nearly 40% of the 128 participants that were clustered, yet these consistently high and consistently low users are obscured when the sample is observed as a whole. Figure 2 features exemplars of these profiles.

Great Users ($n_1 = 22$; 17.2%) were characterized by a high mean (7.29 hours) and a high level (7.08 hours). The average slope was virtually zero (0.002 hours per night), which suggests the participants were not changing their usage. While some participants increased use and some decreased use,

TABLE 2
Summary of the Four Clusters (and Individuals Not Clustered Due to Excessive Missing Data) Aggregated From the Individual Results From Time Series Analyses (Tabled Values Are Means, Standard Deviations in Parentheses)

	Great Users (n = 22)	Good Users (n = 42)	Low Users (n = 29)	Slow Decliners (n = 35)	Not Clustered (n = 33)
Days	184.59 (12.60)	183.45 (8.86)	185.14 (13.60)	182.26 (6.38)	75.18 (57.81)
Mean	7.29 (0.46)	5.72 (0.55)	2.45 (0.83)	4.34 (0.54)	3.11 (1.83)
Level	7.08 (0.63)	5.59 (1.01)	3.67 (1.33)	5.04 (1.08)	4.07 (1.78)
Slope	0.002 (0.005)	0.001 (0.008)	−0.013 (0.011)	−0.007 (0.010)	−0.056 (0.108)
Autocorrelation (r_1)	0.045 (0.113)	0.122 (0.143)	0.178 (0.118)	0.160 (0.134)	0.030 (0.379)
Variance	2.22 (1.55)	4.22 (2.28)	4.21 (1.71)	5.17 (2.49)	3.35 (1.99)
Cumulative Change in Average Use	0.37 hrs or 22 min	0.18 hrs or 11 min	−2.41 hrs or −145 min	−1.28 hrs or −77 min	—

Note. Days = number of days of use; Mean = average hours of PAP use per night; Level = *y*-intercept or number of hours of PAP at the beginning of the trial; Slope = rate of change in PAP use over time; Autocorrelation (r_1) = the influence of the previous night's PAP use on the present night; Variance = variability in hours of PAP use; and Cumulative Change in Average Use = approximate change in PAP use based on the average Days and Slope (calculated from Days*Slope), not calculated for Not Clustered due to the heterogeneity of the group.

the slopes were very small. Cumulative change in average use was an increase of 0.37 hours, or 22 minutes. The variance ($s^2 = 2.22$) was the smallest out of all of the groups. Out of the 22 Great Users, only 4 had significant autocorrelations (18.2%). All of these autocorrelations were very small ($r_1 < 0.30$).

Good Users ($n_2 = 42$; 32.8%) were characterized by a moderate mean (5.72 hours) and a moderate level (5.59 hours). The average slope was virtually zero (0.001 hours per night), similar to the Great Users. Cumulative change in average use was an increase of 0.18 hours, or 11 minutes. The variance was moderate ($s^2 = 4.22$) compared to the other groups. Out of the 42 Good Users, 15 had significant autocorrelations (35.7%). Eleven of these autocorrelations were very small ($r_1 < 0.30$).

Low Users ($n_3 = 29$; 22.7%) were characterized by a low mean (2.45 hours) and a low to moderate level (3.67 hours). The average slope suggested a decrease in use over time (–0.013 hours per night). This was the largest average slope of any of the clusters. Cumulative change in average use was a decrease of 2.41 hours, or 145 minutes. The variance was moderate ($s^2 = 4.21$), and was similar to that of the Good Users. Out of the 29 Low Users, 12 had significant autocorrelations (41.4%). Nine of these autocorrelations were very small ($r_1 < 0.30$).

Slow Decliners ($n_4 = 35$; 27.3%) were characterized by a low to moderate mean (4.34 hours) and a moderate level (5.04). The average slope suggested a decrease in use over time (–0.007). The magnitude of this slope was only exceeded by the Low Users. Cumulative change in average use was a decrease of 1.28 hours, or 77 minutes. The variance was the highest of any of the clusters ($s^2 = 5.17$). Out of the 35 Slow Decliners, 20 had significant autocorrelations (57.1%). The Slow Decliners cluster was the only cluster where the majority of participants demonstrated a significant autocorrelation. Regardless, 15 of these autocorrelations were very small ($r_1 < 0.30$).

Group Comparisons

Participants in the four identified subgroups were compared on a number of available variables (demographics, sleep apnea measures, and psychosocial measures, as described above) that were not involved in the clustering to establish external validity. Categorical variables (e.g., ethnicity, education) were tested with chi-square tests. A chi-square goodness-of-fit indicated that ethnicities were not equally distributed across the groups, χ^2 (12, $n = 128$) = 25.106, $\varphi^2 = 0.196$. Participants that were Black or African American were more likely to be in the Low Users group. Baseline characteristics for the sample, including both continuous and categorical variables, are included in Table 3. Omnibus differences with continuous variables were tested with a series of MANOVAs with cluster as the independent variable. To avoid potential issues with dependency or multicollinearity, a separate MANOVA was performed at each time point (0, 30, 90, and 180 days). Wilks' Lambda criteria for significance are reported. MANOVAs revealed non-significant differences among clusters at baseline ($F(60, 308.13) = 0.979$, NS, $\eta^2 = 0.159$); Day 30 ($F(42, 315.21) = 0.876$, NS, $\eta^2 = 0.103$); and Day 90 ($F(42, 321.15) = 1.14$, NS, $\eta^2 = 0.129$). The MANOVA at Day 180 revealed significant differences among the clusters, ($F(42, 324.11) = 2.14$, $p < 0.001$, $\eta^2 = 0.215$). This is a large effect.

Variables at Day 180 were compared among the four identified subgroups at the univariate level with a series of one-way ANOVAs (with follow-up Tukey HSD tests). Multiple variables were found to have statistically significant differences across the clusters at Day 180, and these are summarized in Table 4. Good Users reported feeling more vigilant (FOSQ Vigilance) than Low Users and Slow Decliners at Day 180 ($F(3, 124) = 5.712$, $p = 0.001$, $\eta^2 = 0.121$). Good Users reported more general productivity (FOSQ Productivity) than Low Users and Great Users at Day 180 ($F(3, 124) = 3.761$, $p = 0.013$, $\eta^2 = 0.083$). Low Users reported more

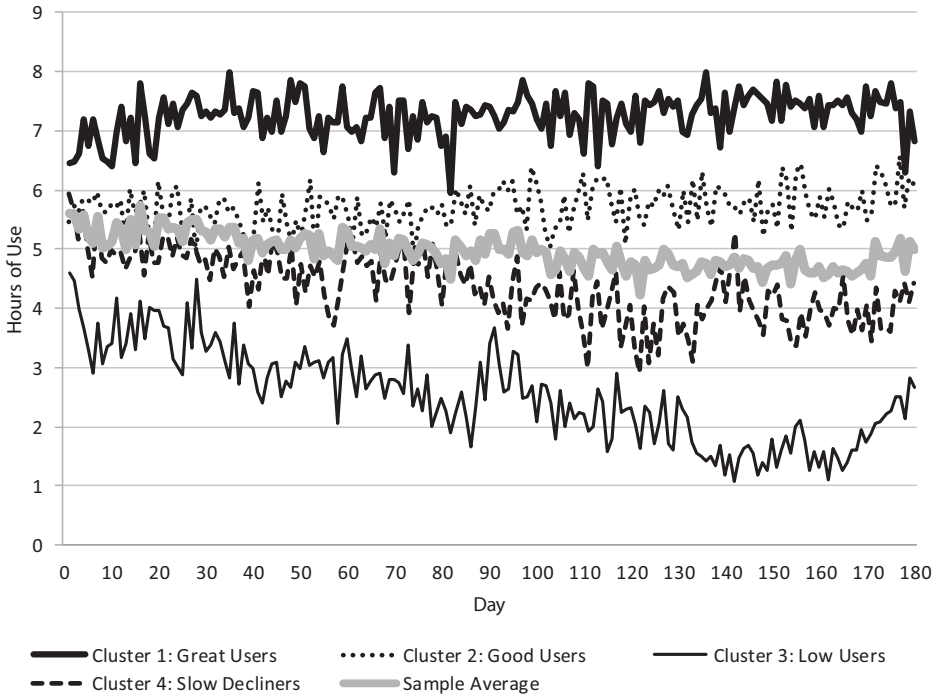

FIGURE 1 Nightly average PAP use of the four identified clusters; the mean of the overall sample is also included to demonstrate that the average use across participants varies greatly from the Great Users and from Low Users.

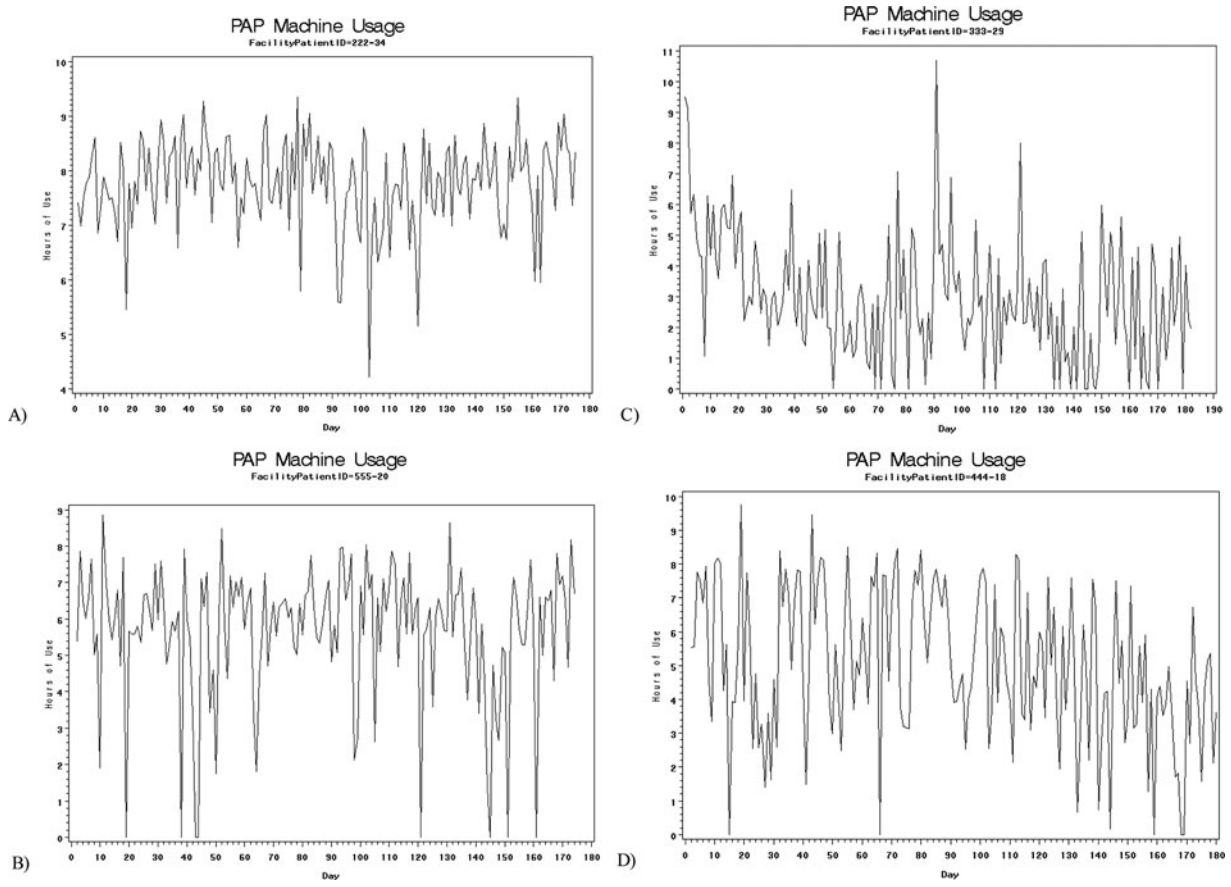

FIGURE 2 Exemplars of clusters. (A) Great Users. (B) Good Users. (C) Low Users. (D) Slow Decliners.

TABLE 3
Baseline Characteristics by Identified Subgroup; For Continuous Variables, Tabled Values Are Means, Standard Deviations in Parentheses; For Categorical Variables, Tabled Values Are Counts, Column Percentages in Parentheses

	Great Users (n = 22)	Good Users (n = 42)	Low Users (n = 29)	Slow Decliners (n = 35)
		Demographics		
Age	47.68 (10.02)	49.55 (11.34)	47.28 (9.41)	50.69 (11.91)
Height (in.)	68.17 (5.55)	69.16 (4.29)	69.11 (3.54)	68.41 (3.85)
Weight (lb.)	231.82 (60.87)	242.85 (50.13)	244.73 (47.07)	215.68 (50.72)
BMI	35.02 (8.17)	35.92 (8.05)	35.95 (6.31)	32.82 (6.63)
BP Systolic	131.73 (10.92)	128.07 (11.82)	130.79 (16.07)	126.00 (16.78)
BP Diastolic	78.00 (11.32)	79.52 (7.81)	82.66 (8.69)	78.11 (10.27)
Heart Rate	75.05 (14.61)	73.64 (10.74)	82.90 (13.69)	73.69 (13.39)
Gender (male)	17 (77.30%)	34 (81.00%)	20 (69.00%)	29 (82.90%)
		Education (3 levels)		
High School	6 (27.30%)	5 (11.90%)	12 (41.40%)	9 (25.70%)
College	11 (50.00%)	29 (69.00%)	14 (48.30%)	20 (57.10%)
Post Graduate	5 (22.70%)	8 (19.00%)	3 (10.30%)	6 (17.10%)
		Ethnicity (5 levels)[1]		
Asian	1 (4.50%)	1 (2.40%)	0 (0%)	0 (0%)
Black or African American	0 (0%)	2 (4.80%)	9 (31.00%)	3 (8.60%)
Hispanic or Latino	0 (0%)	0 (0%)	1 (3.40%)	2 (5.70%)
Other	0 (0%)	0 (0%)	0 (0%)	1 (2.90%)
White or Caucasian	21 (95.50%)	39 (92.20%)	19 (65.50%)	29 (82.90%)
		Sleep Apnea Related		
AHI (events/hour)	53.62 (40.10)	42.01 (33.58)	40.33 (24.88)	32.25 (25.73)
PAP Pressure (cm H_2O)	11.57 (3.53)	10.24 (2.37)	10.59 (2.61)	10.66 (2.65)
		Sleep Center (5 levels)		
Charité-Universitätsmedizin	6 (27.30%)	12 (28.60%)	2 (6.90%)	7 (20.00%)
Gaylord Hospital	6 (27.30%)	11 (26.20%)	9 (31.00%)	8 (22.90%)
Sleep Center of Alabama	1 (4.50%)	9 (21.40%)	8 (27.60%)	9 (25.70%)
Stanford University	6 (27.30%)	6 (14.30%)	3 (10.30%)	8 (22.90%)
University of Florida	3 (13.60%)	4 (9.50%)	7 (24.1%)	3 (8.60%)
		Study Group (3 Levels)		
APAP	9 (40.90%)	12 (28.60%)	11 (37.90%)	11 (31.40%)
Auto-CPAP	6 (27.30%)	19 (45.20%)	10 (34.50%)	9 (25.70%)
CPAP	7 (31.80%)	11 (26.20%)	8 (27.60%)	15 (42.90%)

Note. BMI = body mass index; BP Systolic = systolic blood pressure; BP Diastolic = diastolic blood pressure; AHI = apnea–hypopnea index (a measure of severity) in events per hour; and PAP Pressure = prescribed therapeutic airway pressure. (Detailed information on all variables available upon request.)
[1] $\chi 2(12) = 25.11$, $p < 0.05$, $\varphi 2 = 0.20$.

sleepiness than Great Users and Good Users at Day 180 ($F(3, 124) = 2.735$, $p = 0.046$, $\eta^2 = 0.062$). Great Users reported more self-efficacy (ATU-A Score or confidence) to use PAP than Low Users at Day 180 ($F(3, 124) = 11.875$, $p < 0.001$, $\eta^2 = 0.223$). Great Users and Good Users reported a higher quality of sleep than Low Users Day 180 ($F(3, 124) = 7.26$, $p < 0.001$, $\eta^2 = 0.149$). Great Users reported more of a benefit of treatment (VAS Benefit) than Low Users at Day 180 ($F(3, 124) = 3.262$, $p = 0.024$, $\eta^2 = 0.073$).

DISCUSSION

Time series analysis and dynamic cluster analysis were combined in a Typology of Temporal Patterns (TTP) to assess patterns of adherence to PAP in a sample of adults with OSA. First, a separate time series analysis was conducted for each participant to evaluate individual-level patterns over time. Second, dynamic cluster analysis was employed to identify homogenous subgroups of participants in the sample. These subgroups represent different adherence patterns, and they were used to organize the time series results. Third, a series of one-way MANOVAs and then one-way ANOVAs and chi-square tests was performed with demographic, psychosocial, and sleep apnea variables to establish external validity. Multiple variables were found to be statistically significant with medium and large effect sizes, and these findings demonstrate the usefulness of this procedure.

Each phase of the present study revealed important information regarding adherence patterns. The time series results demonstrated both the complexity of individual-level patterns and the great heterogeneity among participants. At the individual level, participants varied on all time series parameters, including mean, level, slope, autocorrelation, and

TABLE 4
ANOVA Tests Among Clusters for Measures Taken at Day 180; Tabled Values Are Means, Standard Deviations in Parentheses

	Great Users (n = 22)	Good Users (n = 42)	Low Users (n = 29)	Slow Decliners (n = 35)	F(3, 124)	$\eta 2$
FOSQ – Vigilance	3.38 (0.78)	3.66 (0.40)	3.12 (0.52)	3.24 (0.69)	5.71*	0.12
FOSQ – Productivity	3.51 (0.85)	3.84 (0.25)	3.50 (0.44)	3.63 (0.39)	3.76*	0.08
ESS	5.78 (4.67)	6.17 (4.27)	8.90 (4.87)	7.37 (4.59)	2.74*	0.06
ATU-A Score	24.09 (1.93)	23.26 (2.52)	19.34 (4.78)	22.17 (2.90)	11.88*	0.22
VAS – Sleep	75.05 (14.19)	74.86 (12.20)	58.17 (22.86)	67.57 (14.12)	7.26*	0.15
VAS – Benefit	82.09 (13.92)	78.21 (16.98)	69.03 (21.63)	79.06 (16.79)	3.26*	0.07

Note. FOSQ – Vigilance = Functional Outcomes of Sleep (Vigilance Subscale); FOSQ – Productivity = Functional Outcomes of Sleep (Productivity Subscale); ESS = Epworth Sleepiness Scale score; ATU-A Score = Attitudes Toward Use Questionnaire (Self-Efficacy Subscale); VAS – Sleep = Visual Analog Scale for perceived sleep quality; and VAS – Benefit = Visual Analog Scale for perceived benefit of PAP use. Only variables that demonstrated statistical significance are included here (detailed information on all variables available upon request).

*$p < 0.05$.

variance. Notably, the slope, which describes the degree to which usage is changing, varied greatly among participants. The slope as a function of time was significant for nearly half the sample; of those with a significant slope, about 25% demonstrated increases in use over time and about 75% demonstrated decreases in use over time. Such differences are masked if the participants are viewed as a group and the individual differences are ignored. The heterogeneity of the time series results was simplified with dynamic cluster analysis, which identified homogenous subgroups. Four subgroups were found: Great Users, Good Users, Low Users, and Slow Decliners. The individuals that comprise these subgroups had strong similarities in their patterns over time, and these similarities were identified quantitatively.

This TTP relates to the ergodic theorems (Molenaar, 2004). Time series analysis assesses the two conditions of the ergodic theorems relevant to whether the analysis of inter-individual variation will be equivalent to the analysis of intra-individual variation. To satisfy the first condition, the trajectory of each participant has to be stationary across time. Stationarity refers to autocorrelations in time series analysis, which offer information about the generating function. For this sample of adults with OSA, this condition was violated. Approximately 40% of the participants demonstrated a statistically significant lag 1 autocorrelation, and these estimates for the lag 1 autocorrelation varied at the individual level (in terms of both directionality and magnitude). These results suggest that the processes driving adherence are not the same for each individual. To satisfy the second condition, the trajectory of each participant must have the same statistical characteristics. This condition, which refers to the time series summary statistics such as the mean, level, slope, and variance, was also violated in the sample. Individuals demonstrated very different means, levels, and variances, and slopes that were positive and negative, both significant and nonsignificant. The differences in the statistical characteristics were why subgroups were able to be identified. Looking at the sample as a whole violated the conditions of the ergodic theorems. In order to fully account for individual differences, participant data needed to be analyzed individually.

Dynamic cluster analysis identified homogenous subgroups that represent a compromise between idiographic and nomothetic methods and better meet the conditions of the ergodic theorems. Again, as in the entire sample, the first condition was violated. The underlying dynamic processes were not consistent across individuals, as represented by the autocorrelations. However, the subgroups have similar trajectories with similar statistical characteristics. Thus, the subgroups arguably satisfy the second condition. This improvement in ergodicity is important to generalizability. More specifically, the consideration of interindividual and intraindividual variation clarifies generalization between levels of analysis. Results of the present study demonstrated that it is not correct to generalize between an individual and the group as a whole. For example, a Great User, with a high mean and level, is not the same as the overall group average. However, when homogenous subgroups are identified, it is more correct to generalize between an individual and a subgroup. For example, a Great User has a number of statistical similarities to the Great Users subgroup. The subgroups can be compared to each other because the statistical characteristics of the individuals in the subgroups are similar.

Ultimately, neither the entire sample nor the homogenous subgroups fully meet the two conditions of the ergodic theorems. Thus, the group-level analyses do not yield equivalent conclusions to the individual-level analyses. The most accurate descriptions of each individual are the time series analyses (Table 1). In the tradition of idiographic research, the time series analysis for each individual could be considered a separate study (Velicer & Molenaar, 2013). However, combining time series analysis with dynamic cluster analysis in a TTP offers a useful middle ground between idiographic, individual-level analyses and nomothetic, group-level analyses. This typology represents an idiographic-nomothetic hybrid that has some important benefits. First, each individual was evaluated separately, and this revealed important information regarding intraindividual variation. Second, individuals with similarities were grouped together, and this revealed important information regarding interindividual variation. The use of only idiographic methods is problematic

because the findings are only generalizable to each individual over time. Additionally, broad trends across individuals cannot be evaluated. The use of only nomothetic methods is problematic because creating groups without any consideration of individual differences can hide important trends. In the present study, combining all of the individuals into one group mean over time completely obscured the Great Users and Low Users. These users made up nearly 40% of the sample, but their patterns were completely hidden in the overall trend.

The results of the time series analyses and the dynamic cluster analysis offer an important replication of a study by Aloia et al. (2008). In the study, PAP usage was tracked for 365 days. A time series analysis was performed for each participant, and subgroups of similar patterns were formed with visual inspection. The present study replicates both time series and subgroup findings in a larger sample. Participants in both studies demonstrated individual differences in series mean, level, slope, and variance. The majority of autocorrelations were non-significant in both studies. Aloia et al. (2008) identified seven subgroups (Good Users, Slow Improvers, Slow Decliners, Variable Users, Occasional Attempters, Early Drop-outs, and Non-Users) compared to the four subgroups (Great Users, Good Users, Low Users, and Slow Decliners) identified in the present study. These subgroups are different for two important reasons. First, the subgroups were identified in different ways. Aloia et al. identified subgroups qualitatively with visual inspection whereas the present study identified subgroups quantitatively with dynamic cluster analysis. This difference is the reason why two of the subgroups (Early Drop-outs and Non-Users) were not found in the present study; due to the extreme lack of data, these types of users were taken out of the dynamic cluster analysis. Dynamic cluster analysis is also sensitive to changes in level (Dumenci & Windle, 2001), which may explain the differentiation between "Great Users" and "Good Users" in the present study. Second, the length of the study was different. Participants in the Aloia et al. study were tracked for 365 days compared to 180 days in the present study. Aloia et al. (2008) emphasize that some of their subgroups (e.g., Slow Improvers) were characterized by changes late in the study, such as increases or decreases after 6 months. This suggests that the present study identified fewer subgroups because some potential subgroups simply did not form within the 180 days of the study.

The external validity of the groups identified in the present study was tested with a number of demographic, psychosocial, and sleep apnea variables. Most of the demographic variables (at baseline) were found to be non-significant. An important finding at baseline, however, was that ethnicity was significantly different among the groups, with a large effect size. Participants that identified as Black or African American were disproportionately located in the Low Users group. Over one third of the Low Users were Black or African American. This is consistent with past studies (e.g., Joo & Herdegen, 2007) that have found that Blacks or African Americans adhere less to PAP. This is an important issue that warrants future investigation. One limitation of the present study is that the vast majority of the sample was White. Future studies need to involve greater proportions of ethnicities to determine whether other adherence discrepancies exist. If some racial or ethnic subgroups consistently demonstrate lower PAP use, interventions could be targeted specifically to this population to help increase adherence and thus improve OSA outcomes.

There were differences among the subgroups on multiple variables collected at the end of the study (Day 180). These differences both validate the methodological approach of dynamic typology and offer insight into potential interventions that could be employed to increase long-term PAP adherence. Six measures were found to significantly differentiate groups at Day 180 of the study. They were vigilance (FOSQ – Vigilance), productivity (FOSQ – Productivity), sleepiness (ESS), confidence or self-efficacy to use PAP (ATU-A), sleep quality (VAS – Sleep), and perceived benefit of the treatment (VAS – Benefit).

Low Users reported the least amount of self-efficacy to use PAP at Days 90 and 180. This is consistent with prior research demonstrating that self-efficacy is an important predictor of PAP adherence (Aloia et al., 2005; Stepnowsky et al., 2001). This self-efficacy finding is particularly important for creating interventions to increase adherence. There were no significant differences among the groups at baseline or at Day 30, but by Day 180 the self-efficacy measure had the largest effect size of any of the measures. At the beginning of the study, all users reported confidence to use PAP, and the results of the present study suggest that users do not start to significantly lose confidence until months after treatment has started. Developing an intervention administered at 3 and 6 months after PAP treatment has begun may boost the self-efficacy of users and help them maintain confidence to use PAP every night. Intervening to change this behavioral variable has potential to greatly increase adherence.

Low Users reported the most sleepiness at Day 180, and this is consistent with prior research; improvement in sleepiness has been shown to be associated with continued adherence to PAP (Pelletier-Fleury et al., 2001). Low Users also reported the worst quality of sleep, lowest vigilance, lowest productivity, and rated the benefit of the treatment the lowest at Day 180. The value of the vigilance finding is questionable; while self-reported vigilance (FOSQ –Vigilance) was found to be significantly different among the groups, there were no differences among the groups for the objective measures of vigilance (PVT – Mean Response Time and Lapses).

The users that were not clustered due to excessive missing data ($n = 33$) were not quantitatively grouped together and were not included in any group comparisons. The most salient characteristic about these users was that they had a high percentage of missing data for the hours of use per night. During the estimation of missing data with MI, these

participants caused convergence problems in the algorithm. These users included participants that were lost to attrition and participants that had large gaps of data within the 180-day period. Users with a high count of days typically had data at the beginning and the end of the 180-day trial, but were missing large periods of data in the middle of the study. Users with a low count of days appeared to have dropped out of the study early (attrition). The exact causes of the missing data cannot be determined from the data. Participants with a large amount of missing data could have decided to drop out of the study, or there could have been technical problems with the PAP machine's ability to function and/or record data.

The best way to resolve usage questions and better understand reasons for use or non-use would be to include follow-up interviews with participants. This qualitative approach could answer a number of questions that cannot be captured by a set of common variables. Participants could explain specific reasons why they did or did not adhere to PAP. For example, someone could explain that the noise of the machine kept them up at night and prevented them from getting restful sleep. This would be particularly important to better understanding the users that had excessive missing data. With participant-reported explanations for the missing data, more distinct subgroups could be created, such as "non-users due to attrition" and "non-users due to technical problems."

Comparing the homogenous subgroups produced results that both validate the dynamic typology approach and offer insight into variables associated with maintaining PAP adherence. Self-efficacy appears to be the most important variable that separates PAP users from PAP non-users. Intervening on this variable is especially meaningful, as it is theory-based; the measure was developed from the transtheoretical model of behavior change (Stepnowsky et al., 2002). Increased perceived sleep quality and benefit of treatment may be related to an unknown higher order construct of positive attitudes toward PAP. Developing an intervention that emphasizes boosting self-efficacy and benefits of treatment such as improved vigilance, productivity, sleep quality, and reduced sleepiness could significantly increase PAP adherence. Based on the results of the present study, such interventions should continue late into treatment, at least until 6 months after the start of treatment. These interventions could be brief, and they could be delivered efficiently and at a low cost via phone or internet. Alternatively, these interventions could start approximately 2 months after the start of treatment. At this point, both Slow Decliners and Low Users demonstrated decreases in use that could be observed visually in a graph. Identifying participants that are starting to decrease in usage and targeting these individuals would be an efficient way to administer interventions.

There are three limitations to the present study that could be explored in future studies. First, since it was a secondary data analysis, the present study was limited to data collected in the primary study. Psychosocial variables, such as self-efficacy, are extremely important because they can potentially become the basis of interventions to increase adherence. Only a small number of psychosocial variables were available in the present study. Future studies could remedy this by including more measures to assess different theory-based constructs, such as decisional balance (Stepnowsky et al., 2002). The second limitation involved the sensitivity to mean changes with dynamic cluster analysis. When organized by the clusters, the individuals demonstrated remarkably similar means, levels, and variances. While all of the slopes were somewhat similar (all are very small), the directionality of the slopes sometimes varied within subgroup. For example, most individuals in the Great Users and Good Users subgroups that had a significant slope had a positive slope. However, there were some Great Users and Good Users that had a significant negative slope. This decreasing trend seems to contradict the classification, since optimal users should be either increasing or maintaining a stable level of adherence. The individuals that demonstrated a statistically significant decrease in use over time while maintaining a high amount of use per night were clustered based on this high mean. The third limitation involved users that did not have enough data for analysis. As a result, these users were left out of the group comparisons. This is problematic because these users, the non-users, are the most severe failures from an adherence perspective. While these users could not be included in the quantitative process described in the present study, they may represent users that need the most help to increase their adherence and therefore cannot be ignored.

Alternative approaches to data analysis have the potential to replicate or extend some of the findings from the present study. The combination of time series and dynamic cluster analysis represents only one variant of a Typology of Temporal Patterns approach that assesses both individual and subgroup similarities and differences across time. In the present study, the cluster analysis was based on the raw hours per night; alternative approaches to clustering could utilize results from time series analysis for clustering, such as serial correlation, seasonality, and non-linearity (Wang et al., 2006). Clustering based on such time series features may reveal new patterns. Additionally, alternative clustering methods (rather than Ward's method) could have been utilized to test for subgroups. Multiple clustering methods have been proposed that can estimate missing data without the need for prior imputation, such as mixture models that utilize the expectation maximization (EM) algorithm to estimate missing data (Fraley & Raftery, 1998; Jain et al., 1999; Titterington, Smith, & Makov, 1985).

An entirely different approach to the methods in the present study is growth mixture modeling (Muthén & Khoo, 1998; Muthén & Muthén, 2000). Growth mixture modeling represents another approach to exploring subgroups of individuals that have different trajectories over time and describing individual variation in growth. When applied to the data in the present study, this type of modeling could accommodate separate subgroups, or it could analyze

individual trajectories by treating the entire sample as a single, heterogeneous population. A notable and unique advantage of using time series and dynamic cluster analysis is that they can easily accommodate a large number of measurement occasions. Growth curve modeling may be problematic with such a large number of time points. Growth mixture modeling would likely produce some complimentary results to the approach in the present study, and comparing these approaches could be the focus of a future investigation.

The approach that resulted in a typology of PAP users was primarily exploratory. As with the details of the statistical methods (time series analysis, cluster analysis, etc.), the TTP approach can involve exploratory and/or confirmatory analyses. Continuing with the example of adherence to PAP, a researcher could hypothesize that a sample contains a certain number of adherence subtypes. Time series analysis could be used to describe the individuals, and the researcher could specify the hypothesized number of subgroups in the clustering algorithm. Statistical indices associated with the clustering could provide evidence of fit, and these clusters could be cross-validated by using them to organize the time series results (as in the present study).

The TTP approach described in the present study is not limited to studying adherence to PAP or to adherence in general. When data are collected at regular intervals, time series analysis can be used to explore trends over time. Such idiographic, individual-level analyses provide unique insight into the data. When a number of individuals are tracked over time, dynamic cluster analysis can be used to create subgroups of these individuals that share similar longitudinal trends. Variables not involved in the clustering can then be used to validate the process. This procedure respects both individual differences and the need to explore whether similar groups exist within a population. With the advent of modern computing and new technologies to collect longitudinal data (Goodwin, Velicer, & Intille, 2008), researchers have more ways than ever before to explore complex behaviors. The results of the present study demonstrate the usefulness of combining time series and cluster analysis, which is a process that can be applied to other problems in the behavior sciences.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of the authors of the primary study (in addition to Clete Kushida): Richard Berry, M.D. (Shands and University of Florida Sleep Disorders Center, Gainesville, FL, USA), Alexander Blau, M.D. (Charité-Universitätsmedizin Berlin Interdisciplinary Center of Sleep Medicine, Berlin, Germany), Tami Crabtree, M.S. (Santa Rosa, CA, USA), Ingo Fietze, M.D., Ph.D. (Charité), Meir Kryger, M.D. (Gaylord Hospital, Wallingford, CT, USA), Samuel T. Kuna, M.D. (University of Pennsylvania and the Philadelphia VAMC, Philadelphia, PA, USA), G. Vernon Pegram, Jr., Ph.D. (Sleep Disorders Center of Alabama, Birmingham, AL, USA), and Thomas Penzel, Ph.D. (Charité). Steven Babbin would also like to thank Colleen Redding, Ph.D., Bryan Blissmer, Ph.D., and Matthew Goodwin, Ph.D., for help with revising this article, and Andrea Paiva, Ph.D. for help with statistical programming.

FUNDING

The data for this study comes from a clinical trial funded by Philips/Respironics, Inc. Mark S. Aloia is an employee of Philips/Respironics, Inc. and owns stock in Philips. This secondary data analysis was not funded by Philips/Respironics, Inc., and Philips/Respironics, Inc. had no role in study design, analysis, interpretation of the data, writing of the manuscript, or the decision to submit the manuscript for publication.

REFERENCES

Aloia, M. S., Arnedt, J. T., Stepnowsky, C., Hecht, J., & Borrelli, B. (2005). Predicting treatment adherence in obstructive sleep apnea using principles of behavior change. *Journal of Clinical Sleep Medicine*, *1*, 247–254.

Aloia, M. S., Goodwin, M. S., Velicer, W. F., Arnedt, T., Zimmerman, M., Skrekas, J. et al. (2008). Time series analysis of treatment adherence patterns in individuals with obstructive sleep apnea. *Annals of Behavioral Medicine*, *36*, 44–53.

Bendat, J. S., & Piersol, A. G. (1980). *Engineering applications of correlation and spectral analysis*. New York: Wiley.

Blashfield, R. K. (1976). Mixture model tests of cluster analysis: Accuracy of four agglomerative hierarchical methods. *Psychological Bulletin*, *83*(3), 377–388.

Calinski, R. B., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics*, *3*, 1–27.

Chandra, S., Shiffman, S., Scharf, D., Dang, Q., & Shadel, W. G. (2007). Daily smoking patterns, their determinants, and implications for quitting. *Experimental and Clinical Psychopharmacology*, *15*, 67–80.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159.

Collard, P., Pieters, T., Aubert, G., Delguste P., & Rodenstein, D. O.(1997). Compliance with nasal CPAP in obstructive sleep apnea patients. *Sleep Medicine Reviews*, *1*, 33–44.

Cronbach, L. J., & Gleser, G. (1953). Assessing similarities between profiles. *Psychological Bulletin*, *50*, 456–473.

DiMatteo, M. R. (2004). Variation in patients' adherence to medical recommendations: A quantitative review of 50 years of research. *Medical Care*, *42*, 200–209.

Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Sleep*, *20*, 835–843.

Domowitz, I., & El-Gamal, M. A. (2001). A consistent nonparametric test of ergodicity for time series with applications. *Journal of Econometrics*, *102*, 365–398.

Dumenci, L., & Windle, M. (2001). Cluster analysis as a method of recovering types of intraindividual growth trajectories: A Monte Carlo study. *Multivariate Behavioral Research*, *36*, 501–522.

Dunbar-Jacob, J., & Sereika, S. (2001). Measurement of compliance: Conceptual and methodological problems. In: L. E. Burke & I. S. Ockene (Eds.), *Compliance in Healthcare and Research* (pp. 93–104). Armonk, NY: Futura.

Engleman, H., & Wild, M. R. (2003). Improving CPAP use by patients with the sleep apnoea/hypopnoea syndrome (SAHS). *Sleep Medicine Reviews*, *7*, 81–99.

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis* (5th ed.). Padstow, Cornwall, UK: Wiley.

Fraley, C., & Raftery, A. E. (1998). How many clusters? Which clustering method? Answers via model-based cluster analysis. *The Computer Journal*, *41*(8), 578–588.

Glass, G. V., Willson, V. L., & Gottman, J. M. (1975). *Design and analysis of time series experiments*. Boulder, Colorado: Associate University Press.

Goodwin, M. S., Velicer, W. F., & Intille, S. S. (2008). Telemetric monitoring in the behavior sciences. *Behavior Research Methods*, *40*, 328–341.

Hamaker, E. L., & Dolan, C. V. (2009). Idiographic data analysis: Quantitative methods—from simple to advanced. In J. Valsiner, P. C. M. Molenaar, M. C. D. P. Lyra, & N. Chaudhary (Eds.), *Dynamic process methodology in the social and developmental sciences* (pp. 191–216). New York: Springer.

Harrington, M., Velicer, W. F., & Ramsey, S. (2014). Typology of alcohol users based on longitudinal patterns of drinking. *Addictive Behaviors*, *39*, 607–621.

Harrop, J. W., & Velicer, W. F. (1985). A comparison of alternative approaches to the analysis of interrupted time-series. *Multivariate Behavioral Research*, *20*, 27–44.

Hoeppner, B. B., Goodwin, M. S., Velicer, W. F., Mooney, M. E., & Hatsukami, D. K. (2008). Detecting longitudinal patterns of daily smoking following drastic cigarette reduction. *Addictive Behaviors*, *33*, 623–639.

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys*, *31*(3), 264–323.

Johns, M. W. (1992). Reliability and factor analysis of the Epworth Sleepiness Scale. *Sleep*, *15*, 376–381.

Jones, R. H. (1980). Maximum likelihood fitting of ARIMA models to time series missing observations. *Technometrics*, *22*, 389–396.

Joo, M. J., & Herdegen, J. J. (2007). Sleep apnea in an urban public hospital: Assessment of severity and treatment adherence. *Journal of Clinical Sleep Medicine*, *3*, 285–288.

Judge, G. G. (1988). *Introduction to the theory and practice of econometrics* (2nd ed.). New York: Wiley.

Keller, G. (1998). *Equilibrium states in ergodic theory*. New York: Cambridge University Press.

Kushida, C. A., Berry, R. B., Blau, A., Crabtree, T. C., Fietze, I., Kryger, M. H., et al. (2011). Positive airway pressure initiation: A randomized controlled trial to assess the impact of therapy mode and titration process on efficacy, adherence, and outcomes. *Sleep*, *34*, 1083–1092.

Lamiell, J. T. (1998). "Nomothetic" and "idiographic": Contrasting Windelband's understanding with contemporary usage. *Theory and Psychology*, *8*, 23–38.

Lathrop, R. G. & Williams, J. E. (1987). The reliability of inverse scree tests for cluster analysis. *Educational and Psychological Measurement*, *47*, 953–959.

Lavie, P., Herer, P., & Hoffstein, V. (2000). Obstructive sleep apnoea syndrome as a risk factor for hypertension: Population study. *BMJ*, *320*, 479–482.

Lütkepohl, H., & Krätzig, M. (Eds.). (2004). *Applied time series econometrics*. New York: Cambridge University Press.

McArdle, N., Devereaux, G., Heidarnejad, H., Engleman, H. M., Makay, T. W., & Douglas, N. J. (1999). Long-term use of CPAP therapy for sleep apnea/hypopnea syndrome. *American Journal of Respiratory and Critical Care Medicine*, *159*, 1108–1114.

Milligan, G. W. (1980). An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika*, *45*(3), 325–342.

Molenaar, P. C. M. (2004). A manifesto on psychology as an idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, *2*, 201–218.

Molenaar, P. C. M. (2008). On the implications of the classical ergodic theorems: Analysis of developmental processes has to focus on intraindividual variation. *Developmental Psychobiology*, *50*(1), 60–69.

Mun, E. Y., Windle, M., & Schainker, L. M. (2008). A model-based cluster analysis approach to adolescent problem behaviors and young adult outcomes. *Development and Psychopathology*, *20*(01), 291–318.

Muthén, B. O., & Khoo, S. T. (1998). Longitudinal studies of achievement growth using latent variable modeling. *Learning and Individual Differences*, *10*(2), 73–101.

Muthén, B., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, *24*(6), 882–891.

Norman, G. J., Velicer, W. F., Fava, J. L., & Prochaska, J. O. (1998). Dynamic typology clustering within the stages of change for smoking cessation. *Addictive Behaviors*, *23*, 139–153.

Oldridge, N. B. (2001). Future directions: what paths do researchers need to take? What needs to be done to improve multilevel compliance? In L. E. Burke & I. S. Ockene (Eds.), *Compliance in Healthcare and Research* (pp. 331–347). Armonk, NY: Futura.

Overall, J. E., Gibson, J. M., & Novy, D. M. (1993). Population recovery capabilities of 35 cluster analysis methods. *Journal of Clinical Psychology*, *49*(4), 459–470.

Pelletier-Fleury, N., Rakotonanahary, D., & Fleury, B. (2001). The age and other factors in the evaluation of compliance with nasal continuous positive airway pressure for obstructive sleep apnea syndrome. A Cox's proportional hazard analysis. *Sleep Medicine*, *2*, 225–232.

Petersen, K. E. (1983). *Ergodic theory*. New York: Cambridge University Press.

Prochaska, J. O., Velicer, W. F., Guadagnoli, E., Rossi, J. S., & DiClemente, C. C. (1991). Patterns of change: Dynamic typology applied to smoking cessation. *Multivariate Behavioral Research*, *26*, 83–107.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.

Sarle, W. S. (1983). *Cubic clustering criterion* (SAS Tech. Rep. A-108). Cary, NC: SAS Institute.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*, 147–177.

Scharf, S. M., Seiden, L., DeMore, J., & Carter-Pokras, O. (2004). Racial differences in clinical presentation of patients with sleep-disordered breathing. *Sleep and Breathing*, *8*, 173–183.

Shahar, E., Whitney, C. W., Redline, S., Lee, E. T., Newman, A. B., Javier-Nieto, F., et al. (2001). Sleep-disordered breathing and cardiovascular disease: Cross-sectional results of the Sleep Heart Health Study. *American Journal of Respiratory and Critical Care Medicine*, *163*, 19–25.

Shumway, R. H., & Stoffer, D. S. (2010). *Time series analysis and its applications: With R examples*. New York: Springer.

Stepnowsky, C. J., Marler, M. R., & Ancoli-Israel, S. A. (2002). Determinants of nasal CPAP compliance. *Sleep Medicine*, *3*, 239–247.

Strogatz, S. H. (1994). *Nonlinear dynamics and chaos, with applications to physics, biology, chemistry and engineering*. Cambridge, MA: Perseus Books.

Titterington, D. M., Smith, A. F., & Makov, U. E. (1985). *Statistical analysis of finite mixture distributions* (Vol. 7). New York: Wiley.

Velicer, W. F., & Colby, S. M. (2005). A comparison of missing-data procedures for ARIMA time-series analysis. *Educational and Psychological Measurement*, *63*, 1–21.

Velicer, W. F. & Harrop, J. (1983). The reliability and accuracy of time series model identification. *Evaluation Review*, *7*, 551–560.

Velicer, W. F. & McDonald, R. P. (1984). Time series analysis without model identification. *Multivariate Behavioral Research*, *19*, 33–47.

Velicer, W. F., & Molenaar, P. (2013). Time Series Analysis. In J. Schinka & W. F. Velicer (Eds.), *Research Methods in Psychology: Handbook of Psychology* (Vol. 2, 2nd Ed., pp. 628–660). New York: Wiley.

von Eye, A., Mun, E. Y., & Indurkhya, A. (2004). Typifying developmental trajectories: A decision making perspective. *Psychology Science*, *46*, 65–98.

Wang, X., Smith, K., & Hyndman, R. (2006). Characteristic-based clustering for time series data. *Data Mining and Knowledge Discovery*, *13*(3), 335–364.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*, 236–244.

Weaver, T. E., & Grunstein, R. R. (2008). Adherence to continuous positive airway pressure therapy: The challenge to effective treatment. *Proceedings of the American Thoracic Society*, *5*, 173–178.

Weaver, T. E., Laizner, A. M., Evans, L. K., Maislin, G., Chugh, D. K., Lyon, K., et al. (1997). An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep*, *20*, 835–843.

Weaver, T. E., Maislin, G., Dinges, D. F., Bloxham, T., George, C. F. P., Greenburng, H., et al. (2007). Relationship between hours of CPAP use and achieving normal levels of sleepiness and daily functioning. *Sleep*, *30*, 711–719.

Windelband, W. (1998). History and natural science. *Theory and Psychology*, *8*, 5–22.

Young, T., Palta, M., Dempsey, J., Skatrud, J., Weber, S., & Badr, S. (1993). The occurrence of sleep-disordered breathing among middle-aged adults. *The New England Journal of Medicine*, *328*, 1230–1235.

Young, T., Peppard, P. E., & Gottlieb, D. J. (2002). Epidemiology of obstructive sleep apnea: A population health perspective. *American Journal of Respiratory and Critical Care Medicine*, *165*, 1217–1239.

Young, T., Peppard, P., Palta, M., Hla, K. M., Finn, L., Morgan, B., et al. (1997). Population-based study of sleep-disordered breathing as a risk factor for hypertension. *Archives of Internal Medicine*, *157*, 1756–1752.