Published in final edited form as:
Nurs Res. 2011 ; 60(2): 115–123. doi:10.1097/NNR.0b013e3182097813.

Diary Data Subjected to Cluster Analysis of Intake/Output/Void Habits with Resulting Clusters Compared by Continence Status, Age, Race

Janis M Miller, PhD, ANP-BC,
Associate Research Scientist and Assistant Professor School of Nursing and Research Assistant Professor of Obstetrics and Gynecology, Medical School, 400 N Ingalls St Div II, Ann Arbor, MI 48109-5482

Ying Guo, BS, and
University of Michigan School of Public Health, Doctoral Student

Sarah Becker Rodseth, MSW, RN
University of Michigan School of Nursing, Doctoral Student

Abstract

Background—Data that incorporate the full complexity of healthy beverage intake and voiding frequency do not exist; therefore, clinicians reviewing bladder habits or voiding diaries for continence care must rely on expert opinion recommendations.

Objective—To use data-driven cluster analyses to reduce complex voiding diary variables into discrete patterns or data cluster profiles, descriptively name the clusters, and perform validity testing.

Method—Participants were 352 community women who filled out a 3-day voiding diary. Six variables (void frequency during daytime hours, void frequency during nighttime hours, modal output, total output, total intake, and body mass index) were entered into cluster analyses. The clusters were analyzed for differences by continence status, age, race (Black women, n = 196 White women, n = 156), and for those who were incontinent, by leakage episode severity.

Results—Three clusters emerged, labeled descriptively as Conventional, Benchmark, and Superplus. The Conventional cluster (68% of the sample) demonstrated mean daily intake of 45 ±13 ounces; mean daily output of 37 ± 15 ounces, mean daily voids 5 ± 2 times, mean modal daytime output 10±0.5 ounces, and mean nighttime voids 1±1 times. The Superplus cluster (7% of the sample) showed double or triple these values across the 5 variables, and the Benchmark cluster (25%) showed values consistent with current popular recommendations on intake and output (e.g., meeting or exceeding the 8 × 8 fluid intake rule of thumb). The clusters differed significantly (p < .05) by age, race, amount of irritating beverages consumed, and incontinence status.

Discussion—Identification of three discrete clusters provides for a potential parsimonious but data-driven means of classifying individuals for additional epidemiological or clinical study. The clinical utility rests with potential for intervening to move an individual from a high risk to low risk cluster with regards to incontinence.

Keywords

Incontinence; overactive bladder; urination; voiding; voiding diary

Corresponding Author: Janis M. Miller, janismm@umich.edu, (734) 764-4545.

Analysis of intake and output diaries and what they reveal about home behavioral habits has been proposed as informative for evaluating women with urinary incontinence and associated risk factors (Groutz et al., 2000; Stav, Dwyer, & Rosamilia, 2009). Diaries allow clinicians to cluster multiple variables to reveal an individual pattern of behavior. That pattern is then checked against notions about bladder control. However, there are no definitive data concerning healthy population norms and quantitative evidence of association between intake, output, and voiding pattern variance, and bladder control issues are lacking.

A review of previous research indicates that even studies using diaries for capturing bladder habits fail to account for the full breadth of information provided on diaries, the wide variation in output patterns (Larson & Victor, 1988; Kassis & Schick, 1993; Fitzgerald, Stablein, & Brubaker, 2002; Fitzgerald, Ayuste, & Brubaker, 2005; Homma et al., 2000), or intake patterns (Duffey & Popkin, 2007). Instead, reports typically consist of simple one-variable tallies, such as number of voids per 24 hours or liters consumed. No studies were found to have taken into account the overall pattern formed from simultaneous consideration of the numerous variables that account for bladder habits: output frequency, timing, and amount; beverage intake amount; and day vs. night considerations.

If, as is typical in the current literature, only one variable at a time is considered, interpretation is limited. For example, two women might report voiding 10 times per 24 hours. They appear identical in the literature, yet the first woman may drink 2,000 ml per day and the second only 1,000 ml per day. The differences might be better illuminated by providing a second variable to create, for instance, an average volume per void. However, average volume per void does not account for distribution of voids across 24 hours; for example, half of the first woman’s voids might be at night whereas the second woman’s voids are all during daytime hours: important data when considering continence status. Taking into account all of the diary variables leads to both a more realistic interpretation of diary data and provides a more nuanced interpretation of individual data. The question is how to operationalize a composite approach to voiding diary information in large research studies or under high clinical demands. One solution is to have available a few discrete composite groupings from the population-based data to serve as best-match comparatives for individual data. The purpose of this study was to derive those composites using cluster analysis.

Data driven cluster analysis has several advantages over classifying individual data through, for instance, use of *a priori* algorithms or expert opinion (Fan, Brown, Kowaleski-Jones, Smith, & Zick, 2007; Eisen, Spellman, Brown, & Botstein, 1998). In cluster analysis, (a) bias is reduced; (b) missing data issues are avoided since statistical iterations on grouping multiple factors can accommodate for individual overlap on a single variable into a different category; (c) efficiency is gained over hand-classifying within a large database; and (d) clusters can be validated retroactively based on theoretical constructs and by performing external validity checks. Once validated, the clusters can be used for comparative purposes in both clinical and research environments.

Objectives

The objective of this study was to use cluster analysis (data-driven) composites from individual values of intake, output, and voiding frequency; day vs. night voids; and body mass index (BMI) obtained by voiding diary from women in the community. A secondary objective was to test the resulting clusters for their utility by testing for differences by key variables of interest external to the clustering variables: urine leakage status and frequency, influence of beverages deemed irritating to the bladder, and selected demographics.

Method

Data Collection Procedures

This study was part of a parent research project called the Establishing the Prevalence of Incontinence (EPI) involving interview and clinical evaluation in a population of White and Black women in southeast Michigan. The EPI study was approved by the University of Michigan Institutional Review Board, and all participants completed an informed consent document. Data was collected in 2002–2004. Details of the parent study are described elsewhere (Delancey et al., 2010; Fenner et al., 2008; Thomas et al., 2009). In brief, the EPI study was designed in two phases. The first phase of the study involved a telephone interview of women ages 35–64 years to gather self-reported incontinence drawn from a sample of southeastern Michigan community-based women, with over-sampling of Black women (1,922 self-identified Black women and 892 self-identified White women; Fenner et al., 2008). In the second phase, a subset of the women who participated in the telephone interview were invited to undergo testing in the clinic. The subset was recruited to achieve prespecified groups of women with and without urinary incontinence (about half and half) and a minimum of 100 Black and White women of each continence status. Full details of recruitment for the clinical portion of the study has been reported elsewhere (DeLancey et al., 2010).

Before clinic testing, women were asked to fill out a standard voiding diary; data from that diary are the focus of this presentation. The 3-day diary with cover letter instructions was mailed to each EPI study participant invited for entry into the clinical exam portion of the study. The diary included a row for each hour of the day and columns to indicate voided volume, beverage intake volume and type, and a check mark for any episode of urinary incontinence. The cover letter instructions were detailed and included an example page.

Participants

Of the 393 women who returned 3-day voiding diaries, 41 women's diaries were discarded due to consistently missing information. Decision for discard was made by consensus after three research team members reviewed individual diaries. After discarding the 41 incomplete diaries, the final sample for analysis was 352 women: 196 women who self-identified as Black and 156 women who self-identified as White. Mean age did not differ by race and was 49.7 years (*SD* = 8.0). Mean education did not differ by race and was 14.3 years (*SD* = 2.0, range = 7–17).

Data Management and Statistical Analysis

All data from the diaries were entered into an Excel database. Daytime hours were assigned arbitrarily by the researchers as 7:00 am to 11:59 pm and nighttime hours as 12:00 am to 6:59 am. Cluster analysis was selected as an appropriate multivariate statistical technique, as it is used to identify homogeneous groups within sets of data. The aim was to identify relatively distinct groups of subjects based on the natural structure among the variables. That is, there was no preconceived notion of how to cluster variables. Rather, cluster analysis was used to provide high within-cluster homogeneity and high between-cluster heterogeneity (Hair, Anderson, Thatham, & Black, 1992).

Cluster analysis was performed in three stages, according to the guidelines provided by Hair et al. (1992):

Partitioning—In the partitioning stage, the six clustering variables were (a) number of voids during daytime hours, (b) number of voids during nighttime hours, (c) daytime modal output volume in mL, (d) total 24 hours output volume in mL, (e) total 24-hour beverage intake in mL, and (f) BMI (a variable that was speculated to be related to intake).

There are various clustering methods that group subjects with similarity into respective categories. In this study, hierarchical tree clustering was used with bottom-up aggregating of data. First, each of the six variables was considered as a cluster. Second, two or more of the variables were combined into the same cluster based on similarity or distance. The degrees of similarity between objects are represented also in hierarchical tree as tree branch length. Solutions of 3, 4, and 5 clusters were studied and the best solution was selected based on greatest distinction and parsimony.

Euclidean distance was used to determine the distance between objects and between clusters. Considering two objects, *x* and *y* with *N* observed conditions, with each condition representing a feature of grouping objects, *Euclidean distance* is computed as

$$E(x, y)=\{\sum_{i=1}^{N}(x_i - y_i)^2\}^{1/2}.$$

For clustering, a similarity matrix which contained *Euclidean distance* scores was computed for all pairs of *N* objects. Then, the matrix was searched for the highest value reflecting the most similar pair of objects and a cluster was created to join these two objects. Wald's linkage method was used to join the clusters. This method helps determine when two clusters are sufficiently similar to be linked together: Semipartial R-square measure is used to evaluate the loss of homogeneity due to merging two clusters to form a new cluster at a given step. Small values of R-squared imply that the new cluster obtained at a given step is formed by merging two very homogeneous clusters. On the other hand, if the value is large, then it implies that two heterogeneous clusters have been merged to form the new cluster.

Finally, clustering methods were combined with a graphical representation approach, displaying each cluster in a distinct color and to reflect the quantitative features of subjects. The end product provided a view of the numbers of clusters in a natural, intuitive manner.

Interpretation—In the interpretation stage, ANOVA tests were used to identify differences between cluster group means for each of the individual variables. The contribution of each variable to cluster separation was compared using effect size to measure strength of association and by using the *R*-squared value. Clusters were labeled with a descriptive name to assist in distinguishing them in future analysis.

Profiling—In this stage, utility of the clusters was tested by subjecting them to analysis based on the expectation that the intake and output patterns that underlie cluster membership may differ by certain other relevant external variables. Utility of the clusters was tested by profiling on continence status, volume of potential bladder irritants (defined as coffee, tea, carbonated beverages, and citric juices), proportion of total fluids that are potential bladder irritants, race, age, and length of formal education. For those women who were incontinent, severity was evaluated also by mean episodes of leakage per day: none, less than daily, or at least twice. Statistical significance was tested in differences between clusters using both an ANOVA test for continuous data and a Chi-square test for categorical data.

Finally, the entirety of the cluster analysis was repeated, limiting the sample to only incontinent women, and profiling on all of the above listed variables except continence status, since all were incontinent. This additional analysis was conducted to determine robustness of the clustering considering only incontinent women. All analyses were conducted using SAS 9.0.

Results

Partitioning

The sample means of the 6 variables subjected to cluster analysis are shown in Table 1, and Pearson correlation coefficients among variables are displayed in Figure 1. Participants' BMI does not significantly correlate with other variables except very weakly ($r = 0.12$) with number of voids during nighttime hours. Despite lack of correlation, BMI was kept in the cluster analysis as a cluster variable with the consideration that it might provide information beyond intake and output for grouping subjects.

All 352 subjects can be grouped into three clusters (Figure 2). To verify this result as most parsimonious, 4 and 5 clusters were explored, also using a graphical representation. The 3-cluster solution was selected on the basis of good separation of the clusters (Figure 3a, compared to Figures 3b and 3c) and the clinical meaningfulness of the group means for each of the 6 variables (Table 2). The 4 and 5 cluster analyses showed less parsimony and greater overlap between patterns (Figures 3b and 3c) and were abandoned.

Interpretation and Labeling of the Three Clusters

Examination of ANOVA statistics showed that all clustering variables contributed significantly to the 3-cluster separation and do not simply parallel each other (Table 2), with the exception of BMI. Total output volume and beverage intake showed the biggest effect sizes (R-square = 0.688 and 0.616, respectively). *Post hoc* analyses using the Bonferroni correction with a significance level of .05 also indicated that the three clusters are significantly different from each other on all of the formative variables used, except BMI.

Cluster 1: Conventional—Cluster 1 included the largest portion of women ($n = 233$, 66%). This cluster is characterized, relative to the other clusters, by lower intake volume (~1320 mL), fewer voids (on average 5 daytime voids and 1 nighttime void), lower bladder volume emptying (modal output ~290 mL), and lower total output (~1069 mL per 24 hours; Table 2). Cluster 1 was labeled Conventional since all values were consistent with more historical patterns of intake and output recommendations prior to recent trends (Duffey & Popkin, 2007; Nielsen & Popkin, 2004).

Cluster 2: Benchmark—Cluster 2 ($n = 96$, 27%) showed an average of about 2,445 mL of beverage intake per day. Void frequency was 7.0 times during daytime hours and 1.3 times during nighttime hours. Modal daytime urine output was about 444 mL with total daily output averaging 1,907 mL. This profile was higher in all variables compared to Cluster 1 (Table 2), but representative of a pattern that falls within the healthy bladder norms suggested by expert panels such as the International Continence Society and the Wound Ostomy Continence Nurses Association, and by numerous websites or organizations such as the Women's Bladder Health website (www.womensbladderhealth.com) and the National Association for Continence website (www.nafc.org). The term Benchmark was chosen since all values conform to standards that are widely popularized as *healthy* or *normal*.

Cluster 3: Superplus—Cluster 3 ($n = 23$, 7%) showed extreme intake and output patterns (Table 2). Intake averaged 3,774 mL ounces. Voids averaged 10.2 times at daytime and close to 2 times routinely at night, modal output was 647 mL per void, and total output per 24 hours was 3281 mL.

Validation and Profiling

Results comparing whether other variables of interest, external to the formative variables, differ across Cluster 1, 2, or 3 are shown in Table 3. When continence status was analyzed, using the average number of leakage episodes recorded on the 3-day diaries and including women without any recorded leakage (0 leaks on the 3-day diary), frequency of leakage was significantly different among the clusters ($p = .02$), with the Superplus cluster demonstrating the most frequent leakage. When data on frequency of leakage was collapsed into categories of zero, less than daily, or at least twice on the 3-day diary, there was a significant difference by cluster ($p = .05$), again with the Superplus cluster showing the worst leakage.

Also analyzed was volume of beverages consumed that are generally considered to be potential bladder irritants (e.g., carbonated beverages, coffee, tea, alcohol). There was again a significant difference by clusters ($p < .001$), with Superplus consuming the most potential irritants. However, while absolute volume of potential bladder irritants differed by cluster, the clusters showed no difference in irritating beverage intake proportion relative to total beverage intake. That is, in general, those who drank a lower volume of potential bladder irritants also had a lower intake of total beverage volume as well; and those who drank higher volume of potential bladder irritants had a corresponding higher intake of total beverage volume.

In terms of demographics, age was marginally higher in the Superplus cluster ($p = .08$), but educational level did not differ. There was a significant relationship between race and cluster status, with Black women being more represented proportionally in the Conventional group and White women composing nearly two-thirds of the Superplus and Benchmark groups (Table 3).

Analysis Repeated with Only Women with Leakage

There were 188 women with leakage, defined for this analysis as having recorded at least one episode of leakage on the 3-day diary. The cluster analysis was repeated, with findings similar to those revealed with the full sample size. A 3-cluster solution was best and with similar mean profiles of the 6 clustering variables (data not shown). There was a significant difference between the three clusters in average number of incontinence episodes ($p = .02$; Table 4), with the Superplus cluster having the most frequent occurrence of leakage. Significant differences were observed in intake of potential bladder irritants with the Superplus cluster having the most ($p < .001$). Again, proportion of potential bladder irritant beverages to total beverage intake did not differ by cluster ($p = .88$) when the analysis is restricted to only incontinent women, similar to when both continent and incontinent women were included. In this analysis, inclusive only of incontinent women, neither age nor education differed significantly by cluster ($p = .19$ and $p = .20$, respectively), but race was again significant ($p = .002$), with a higher proportion of Black women in the Conventional cluster and a higher proportion of White women in the Superplus cluster.

Discussion

The findings of this study challenge current popular recommendations for healthy bladder care, which often emphasize high intake. Little data exist in the literature to compare with the findings incorporating the full complexity of bladder habit information obtained from 3-day diaries, in a sample representative of both continent and incontinent women, and by racial groupings. An accumulation of data show a wide variation in output patterns within asymptomatic community-dwelling women (Table 5). However, these studies do not track beverage intake, which must be extrapolated indirectly from the output.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

In the current study, which tracked both, intake and output did not match as closely as anticipated. All of the clusters in the study showed an average output that is about 25% less than their average intake, logically due to water loss, such as sweating and exhalation, a known source of daily fluid loss. Most of the studies in the literature do not have a representative sample of Black women, with the exception of that done by Fitzgerald, Stablein, and Brubaker (2002), whose sample included 39% Black women. Their data, tracking beverage intake, concurred with the current study in that White women had higher median fluid intake of 2,160 mL per day compared to Black women's median intake 1,890 mL per day. However, no statistical analysis was provided and the ranges for both groups were broad and overlapping (675–4,980 for White women and 750–4,980 for Black women). To compare the current data directly to Fitzgerald, Stablein, et al. it is reported here as medians of 1,843mL (592–5,678) for White women and 1,404mL (335–4,692) for Black women.

Black women are known also to have lower incontinence prevalence rates. It is possible that their more conventional intake, output, and voiding habits, including modest beverage intake, may provide a partial explanation for the previously documented lower rates of incontinence in these women (Fenner et al., 2008).

A literature search focusing on the years 1990–2009 using various combinations of the keywords *fluid*, *intake*, and *incontinence* produced few systematic studies on beverage intake related to bladder health. However, in comparing the output data from the studies published over 14 years 1988–2002 (Table 5), there was an overall 300 ml increase per day of urine voided, which provides an indirect measure of trend towards increased fluid intake over the past 20 years. Supporting these data, Duffey and Popkin (2007), in studying obesity, showed an increase from 1989 to 2002 from an average of 78.9 oz per day to 100.1 oz per day of beverages consumed.

Popular, research-based, and historical fluid consumption recommendations are at odds as to whether this cultural trend of increased fluid intake is beneficial. An entrenched general health recommendation is intake of eight servings of 8-ounce glasses of water per day, yet no scientific or medical basis for this recommendation can be found in the literature. The Institute of Medicine issued a report in 2004 with guidelines for total water intake for healthy people. That recommendation for women was 2.7 L per day, but was intended to include water consumed both through beverages and food. This food vs. fluid component has been a source of confusion for the public and professionals alike, with a misperception of fluid source as only deriving from beverage intake. The effects of these fluid recommendations on bladder control have not been evaluated systematically. Beverage consumption well beyond 64 ounces (with food fluid source excluded) is a common occurrence, and is associated with incontinence, which raises an interesting conceptual issue of whether overactive bladder is, instead, overactive beverage consumption.

The current data have set the stage for additional intervention research to determine to what extent Superplus or Benchmark women may benefit from instruction to modify their habits towards Conventional patterns as a prevention or treatment for incontinence. It cannot be determined from this study design if women in the Conventional cluster show less frequent leakage because they have a stronger continence mechanism or because they have discovered self-protective habits, such as beverage control.

An interesting conundrum is that although the volume of overall beverage intake (over the 3 days) was a significant contributor to cluster differentiation, and amount of potentially irritating beverages consumed differed by clusters, the proportion of potential irritants to overall beverage intake was relatively the same across the three clusters: roughly 65%. Thus, it remains unclear if a higher absolute volume of potential irritants is simply a byproduct of higher absolute beverage intake in total. It is possible that volume rather than type of beverage is responsible for the finding of greater leakage in the Superplus cluster. It seems reasonable to suggest that the Superplus women who experienced bladder control issues may do well to focus on reducing their overall volume intake, which they may find to be more palatable than reducing certain types of beverages. By reducing overall beverage intake, they may inadvertently achieve the lower potential irritant intake as well. More research is needed to determine the relationship between volume and type of beverage (potentially irritating versus nonirritating) and incontinence.

Strengths and Limitations

The benefit of diaries is that they offer a means for gathering dense, hour-to-hour information within the participant's own environment. Intake, output, and voiding diaries are standard assessment for comprehensive evaluation of women presenting for incontinence treatment. But the data gained is dense, and reduction of that information into meaningful and manageable classifications has been difficult. The strength of the work presented in this presentation is that complex information was simplified into data-driven robust clusters now available for use in future work. That is, for 352 women recording information as often as hourly on 6 variables across 72 consecutive hours, and with each resulting data point entered into the spreadsheet and subjected raw into the overall cluster analysis, robust cluster were formulated, with beginning characterization and profiling by external variables. In the future, clinicians and researchers can exploit this representative cluster analysis by reducing the complex data of an individual's diary information into one of the three cluster labels. For many common uses of diaries in research or clinical practice, this pragmatic 3-category approach to voiding diary information should be workable, but it was first important to derive the cluster groups and labels from data, and not just expert opinion.

One missing component from the diaries was a place to record bedtime and waking hours. Due to this oversight, daytime and nighttime hours were assigned retrospectively to the data, so individualized sleep vs. wake patterns are not reflected in this study. This primarily affects the variable of number of voids at night. This is not thought to be a serious flaw, as the data are close, on average, to other published studies of nighttime voiding frequency; for instance, Lukacz, Whitcomb, Lawrence, Nager, and Luber (2009) reported 33% of women had 2 episodes per night and Parsons et al. (2007) reporting nighttime frequency increases with increasing intake volume.

An additional strength of the study is the unusually high sample of self-identified Black women, who are a traditionally underrepresented group in research. With this adequacy of sample size, it was possible to analyze for racial comparisons across the clusters. The finding that distribution of race is significantly different by cluster group must be interpreted within the constraints of sampling from within a parent study, conducted in southeast Michigan, and with purposeful oversampling for Black women. Also, the sample consisted only of women self-identified as Black or White race, so findings cannot be generalized to other racial designations, nor should the groupings be interpreted as having been formed by more strict criteria. For instance, women were not asked if they were of mixed race.

The research findings are restricted to women, because men were not included in the parent study. The findings were gained from a particular geographic region, southeast Michigan, which has widely varying temperature and humidity across the year. These environmental factors were not controlled for, and may have influenced intake on any particular day. However, the study was conducted across several years and women were entered randomly, so there is no reason to believe that group data was influenced highly by a particular seasonal shift. Also not controlled for were exercise and food intake.

Conclusion

The results offer a simplified data-derived classification system for reducing complex intake, output, and voiding diary information into meaningful categories. This new, data-derived classification system of three clusters expands the potential for the clinician and researcher to categorize an individual's pattern by placing them within the best-fit cluster and fosters testing interventions aimed at moving an individual from a higher risk cluster to a lower risk cluster, and monitor effectiveness accordingly.

Acknowledgments

Caroline Garcia for excellent editorial assistance. NIH R01 HD/AG41123**,** with additional Investigator support from NIH NICHD/ORWH P50 HD044406 and NIH P30NR009000

References

DeLancey JO, Fenner DE, Guire K, Patel DA, Howard D, Miller JM. Differences in continence system between community-dwelling black and white women with and without urinary incontinence in the EPI study. American Journal of Obstetrics and Gynecology. 2010; 202(6):e1–e584.

Duffey KJ, Popkin BM. Shifts in patterns and consumption of beverages between 1965 and 2002. Obesity. 2007; 15(11):2739–2747. [PubMed: 18070765]

Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proceedings of the National Academy of Sciences of the United States of America. 1998; 95(25):14863–14868. [PubMed: 9843981]

Fan JX, Brown BB, Kowaleski-Jones L, Smith KR, Zick CD. Household food expenditure patterns: A cluster analysis. Monthly Labor Review. 2007; 130(4):38–51.

Fenner DE, Trowbridge ER, Patel DL, Fultz NH, Miller JM, Howard D, et al. Establishing the prevalence of incontinence study: Racial differences in women's patterns of urinary incontinence. The Journal of Urology. 2008; 179:1455–1460. [PubMed: 18295278]

Fitzgerald MP, Ayuste D, Brubaker L. How do urinary diaries of women with an overactive bladder differ from those of asymptomatic controls? BJU International. 2005; 96(3):3565–3567.

Fitzgerald MP, Stablein U, Brubaker L. Urinary habits among asymptomatic women. American Journal of Obstetrics and Gynecology. 2002; 187(5):1384–1388. [PubMed: 12439535]

Groutz A, Blaivas JG, Chaikin DC, Resnick NM, Engleman K, Anzalone D, et al. Noninvasive outcome measures of urinary incontinence and lower urinary tract symptoms: A multicenter study of micturition diary and pad tests. The Journal of Urology. 2000; 164(3 Pt 1):698–701. [PubMed: 10953128]

Hair, JF.; Anderson, RE.; Thatham, RI.; Black, WC. Multivariate data analysis. 3. New York, NY: Maxwell MacMillan International Editions; 1992.

Homma Y, Yamaguchi O, Kageyama S, Nishizawa O, Yoshida M, Kawabe K. Nocturia in the adult: Classification on the basis of largest voided volume and nocturnal urine production. The Journal of Urology. 2000; 163(3):777–781. [PubMed: 10687975]

Institute of Medicine and Food and Nutrition Board. Dietary reference intakes for water, potassium, sodium, chloride, and sulfate. Washington, DC: National Academies Press; 2004.

Kassis A, Schick E. Frequency-volume chart pattern in a healthy female population. British Journal of Urology. 1993; 72(5 Pt 2):708–710. [PubMed: 8281399]

Larsson G, Victor A. Micturition patterns in a healthy female population, studied with a frequency/volume chart. Scandinavian Journal of Urology and Nephrology, Supplementum. 1988; 114:53–57. [PubMed: 3201170]

Lukacz ES, Whitcomb EL, Lawrence JM, Nager CW, Luber KM. Urinary frequency in community-dwelling women: What is normal? American Journal of Obstetrics and Gynecology. 2009; 200(5):552, e1–552.e7. [PubMed: 19249726]

Parsons M, Tissot W, Cardozo L, Diokno A, Amundsen CL, Coats AC. Normative bladder diary measurements: Night versus day. Neurourology and Urodynamics. 2007; 26(4):465–473. [PubMed: 17335055]

Stav K, Dwyer PL, Rosamilia A. Women overestimate daytime urinary frequency: The importance of the bladder diary. The Journal of Urology. 2009; 181(5):2176–2180. [PubMed: 19296975]

Thomas A, Low LK, Tumbarello JA, Miller JM, Fenner DE, DeLancey JO. Changes in self-assessment of continence status between telephone survey and subsequent clinical visit. Neurourology and Urodynamics. 2009; 29(5):734–740. [PubMed: 19816917]

NIH-PA Author Manuscript
NIH-PA Author Manuscript

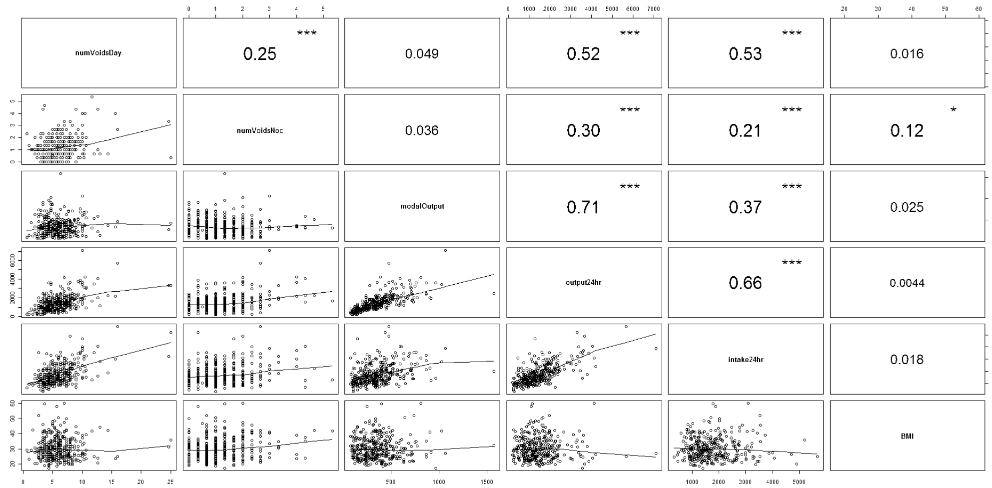


Figure 1.

Scatter plots of six variables and corresponding Pearson correlation coefficients.

Notes. The upper-diagonal figure shows correlation coefficients between variables. * indicates the significance of correlation. *** means that the two variables are highly correlated with P values < .001. The lower-diagonal figure shows scatter plots between variables.

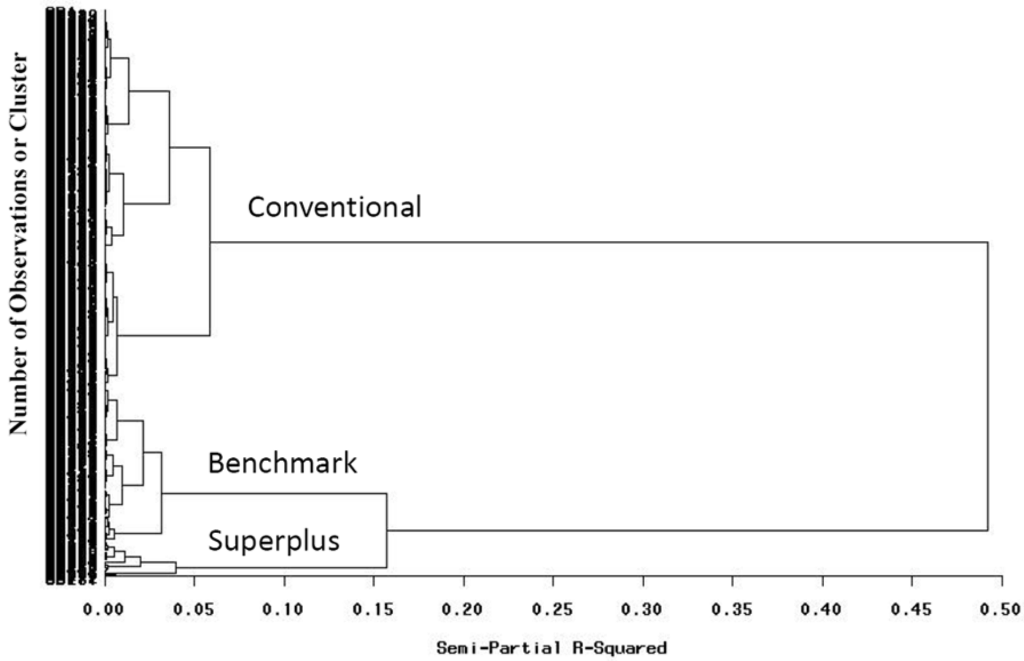
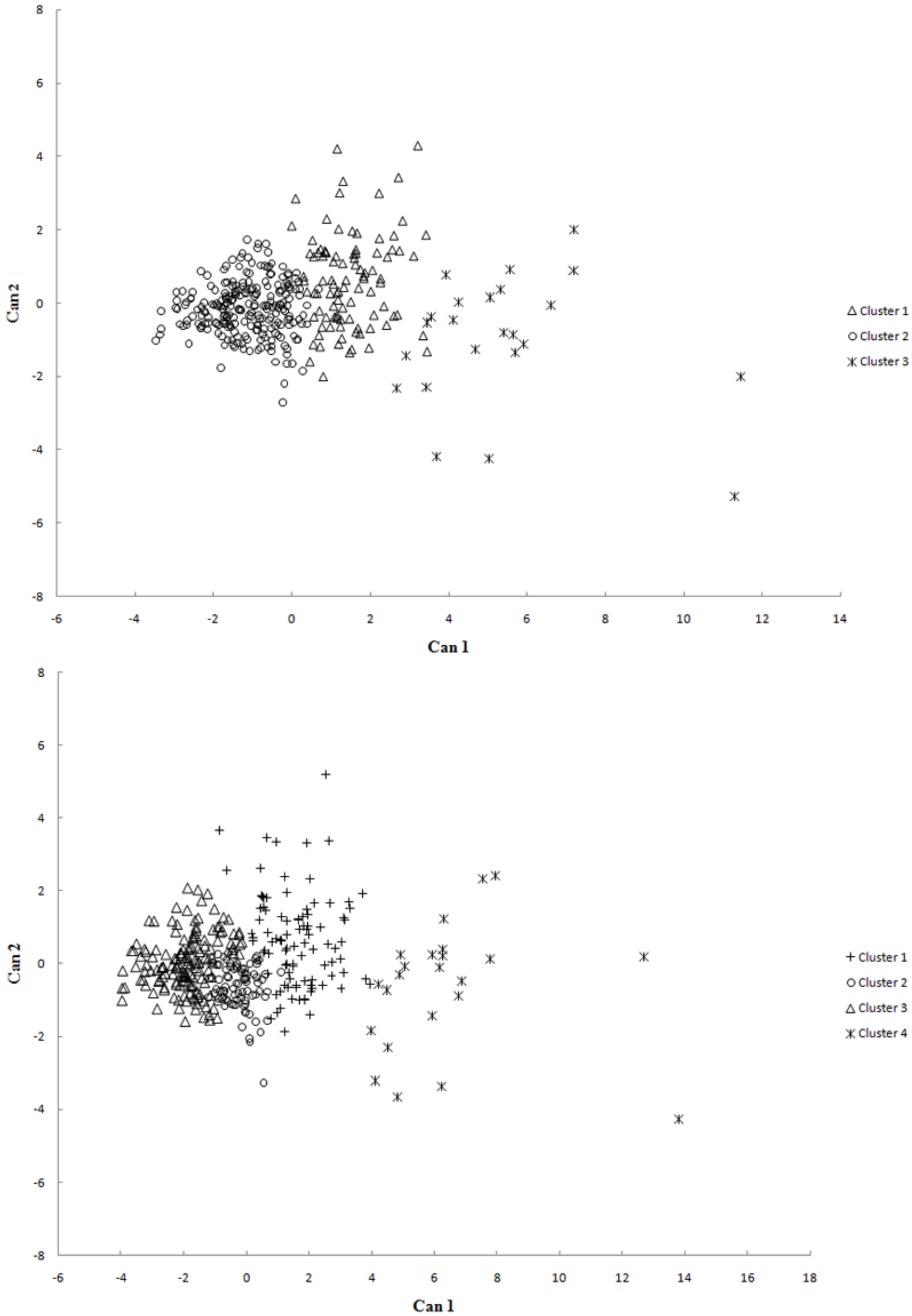


Figure 2.
Hierarchical tree, indicating the relationships among subjects.
Notes. The horizontal axis denotes the linkage distance, with longer lines indicating greater distance between the groups. Three principle clusters are identified with Ward's linkage.

NIH-PA Author Manuscript
NIH-PA Author Manuscript



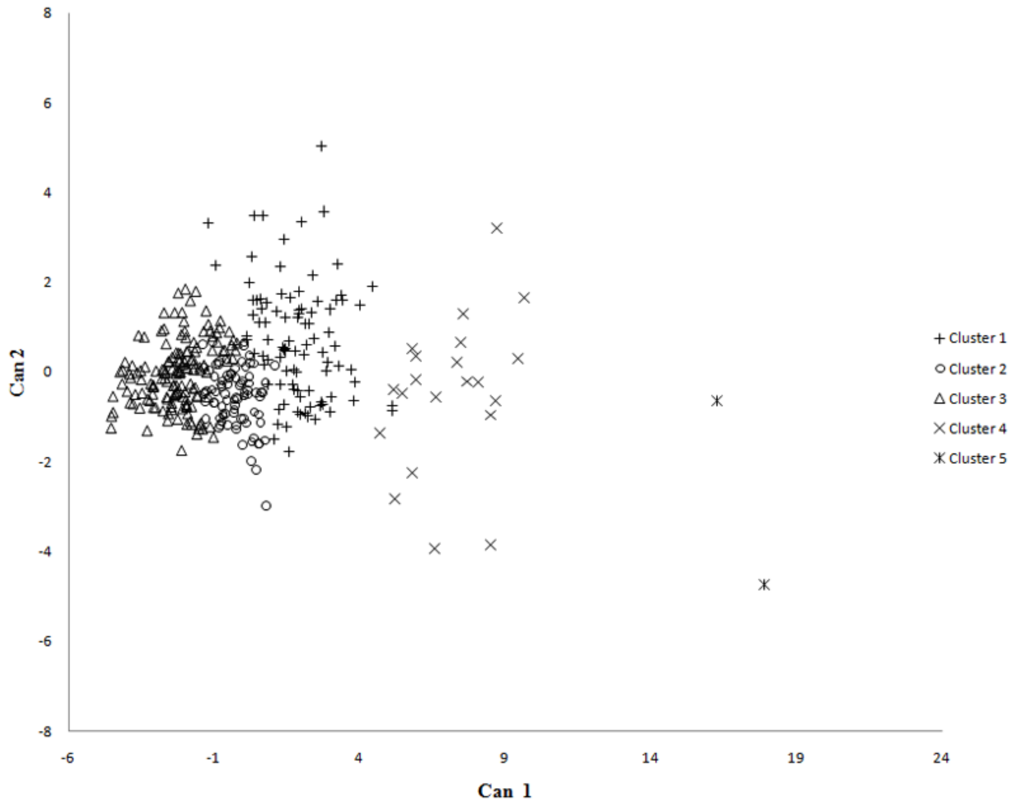


Figure 3.

Figure 3a. The graphical representation of the clusters (number of cluster, *K* =3).

Figure 3b. The graphical representation of the clusters (number of cluster, *K* = 4) shows more undesirable overlap than K=3 (Figure 3a).

Figure 3c. The graphical representation of the clusters (number of cluster, *K* =5) shows more undesirable overlap than when K=4 or K=5 (Figure 3b & 3c).

Notes. Each dot in the figure represents a subject. The subjects within the clusters are represented by the same color. Can1 and Can2 are canonical variables computed by canonical discriminant analysis to help visually identify the groups. Canonical variables are defined by the linear combination of quantitative variables, which aim to maximize between group variance while minimizing within group variance. The quantitative variables used in this study include number of voids at daytime, number of voids at nighttime, daytime modal output volume, total 24 hours output volume, total beverage intake, and BMI. Clusters are linked to Table 2 by the ordinal numbers in Table 2.

Table 1

Means and Standard Deviations of Clustering Variables for All Subjects

Clustering Variables	M	SD
Number of voids: daytime (count)	6.0	2.7
Number of voids: nighttime (count)	1.2	0.9
Modal Output: daytime (ml)	355.2	184.8
Overall output: 24 hr (ml)	1474.0	857.6
Overall beverage intake: 24 hr (ml)	1755.1	807.6
BMI (points)	30.4	7.4

Notes. All variables except BMI were obtained by 3-day diary. Height and weight were measured during a clinic exam.

BMI = body mass index

Table 2

Descriptive Results of all Clustering Variables for All Subjects

	Cluster 1: Conventional (*n* = 233)	Cluster 2: Benchmark (*n* = 96)	Cluster 3: Superplus (*n* = 23)	*p*	R^2	Cluster comparison
Total intake per Daytime (mL)	1320.4±375.3	2444.6±542.9	3773.9±912.5	<.001	0.626	3-2* 2-1* 1–3*
Total output during 24 Hr (mL)	1068.6±434.0	1907.0±437.4	3280.8±1066.3	<.001	0.689	3-2* 2-1* 1–3*
Modal output at daytime (mL)	289.8±143.6	444.0±201.6	646.9±209.9	<.001	0.278	3-2* 2-1* 1–3*
Number of voids, daytime (count)	5.2±1.8	6.9±2.3	10.5±5.3	<.001	0.270	3-2* 2-1* 1–3*
Number of voids, nighttime (count)	1.1±0.8	1.4±0.9	1.9±1.2	<.001	0.056	3-2* 2-1 1–3*
BMI (points)	30.3±7.1	30.7±7.6	30.0±9.3	.81	0.001	3-2 2-1 1–3

Notes. BMI = body mass index. Results are expressed as Mean ± standard error; *p* values are computed by ANOVA test.

* cluster comparison significance at .05 level.

Table 3

Comparison by Clusters of Clinical and Demographic Variables External to Those Used for the Cluster Analysis

	Conventional (*n* = 233)	**Benchmark (*n* = 96)**	**Superplus (*n*= 23)**	***p***
Bladder irritants intake	1310.5±373.2	2417.9±559.2	3318.0±110.4	<.001
Proportion of bladder irritants of total intake	0.64±0.27	0.58±0.27	0.65±0.29	.27
Incontinence episodes/day (average)	0.9±1.3	1.2±2.3	2.0±2.7	.02
Incontinence episodes/day				.05
0 leakages	114(41.7%)	40(48.9%)	10(43.5%)	
0< leakages = 1	66(30.2%)	29(28.3%)	2(8.7%)	
Leakages > 1	53(28.1%)	27(22.8%)	11(47.8%)	
Race				<.001
White	82(35.2%)	58(60.4%)	16(69.6%)	
Black	151(64.8%)	38(39.6%)	7(30.4%)	
Age (years)	49.7±8.1	49.2±7.7	53.3±7.4	0.08
Education (years)	14.2±2.0	14.4±2.1	14.0±2.2	0.60

Notes. For continuous variables, p values were computed by ANOVA test. Results are expressed as mean ± standard error; for categorical variables (i.e., race). p values were computed by chi-square test and results are expressed as numbers (percentages).

Table 4

Descriptive Results of External Variables for Incontinent Women by Cluster

	Conventional (*n* = 89)	Benchmark (*n* = 84)	Superplus (*n* = 15)	*p*
Bladder irritants intake (oz)	859.1±502.9	1286.5±684.4	2264.3±1373.8	<.001
Proportion of irritants of total intake	0.63±0.26	0.59±0.28	0.61±0.31	.88
Incontinence severity (episodes/day)	1.8±1.7	2.2±2.9	3.5±2.4	.02
Age (years)	48.6±8.2	50.1±7.8	52.2±6.5	.19
Education level (years)	14.1±2.0	14.3±2.0	13.3±2.4	.20
Race				.002
Black	58(65.2%)	34(40.5%)	5(33.3%)	
White	31(34.8%)	50(59.5%)	10(66.7%)	

Notes. For continuous variables, p values were computed by ANOVA test. Results are expressed as mean ± standard error; for categorical variables (i.e., race).

p values were computed by chi-square test and results are expressed as numbers (percentages).

Table 5

Comparison of Results of the Study to Other Studies with Asymptomatic Women

	Larsson & Victor (1988)	Kassis & Schick (1993)	Homma et al. (2000)	FitzGerald, Stablein, & Brubaker (2002)	Miller, Guo, & Rodseth-Becker (present study)	
	Studies with Asymptomatic Women	**Continent**	**Incontinent**			
Subjects (*n*)	151	33	32	300	164	188
Mean voided volume per 24 h (mL ± SD)	1430 ± 487	1473 ± 386	1332 ± 59	1759 ± 797	1328±842	1363±873
Mean voids per 24 h (*n* ± SD)	5.8 ± 1.4	5.6 ± 1.3	8 ± 0.4	8.3 ± 2.4	6.8±2.9	7.6±3.2*
Mean Voided volume (mL ± SD)	250 ± 79	Not reported	175 ± 8	216 ± 87	226±111	198±98*
Maximum voided volume (mL ±SD)	460 ± 174	Not reported	277 ± 16	362 ± 161	367±188	345±201
Nocturia (% of subjects who reported ≥ 1 episodes)∞	15	Rare	Not reported	44	38.41	47.87

Notes.

* $p < .01$ comparing continent and incontinent women in present study

∞ Miller, Guo, Rodseth-Becker (present study) used arbitrary “nighttime” designation as between the hours of 12:00 a.m and 6:59 a.m. on 24-hour diary