Lasagna plots: A saucy alternative to spaghetti plots

Bruce J. Swihart, Brian Caffo, Bryan D. James,
Matthew Strand, Brian S. Schwartz, Naresh M. Punjabi

Abstract

Longitudinal repeated measures data has often been visualized with spaghetti plots for continuous outcomes. For large datasets, this often leads to over-plotting and consequential obscuring of trends in the data. This is primarily due to overlapping of trajectories. Here, we suggest a framework called lasagna plotting that constrains the subject-specific trajectories to prevent overlapping and utilizes gradients of color to depict the outcome. Dynamic sorting and visualization is demonstrated as an exploratory data analysis tool. Supplemental material in the form of sample R code additional illustrated examples are available online.

1 Introduction

Longitudinal data are a cornerstone to causal inference in epidemiologic studies. As the number of subjects, frequency of measurements, and period of active data collection grows, so does the size of the data set. As the size of the data set grows, so does the burden and complexity of being able to graphically explore and summarize the data for comprehensible viewing without obscuring salient features.

The current gold standard for graphically displaying and exploring longitudinal data is the spaghetti plot, which involves plotting a subject's values for the repeated outcome measure (vertical axis) versus time (horizontal axis) and connecting the dots chronologically. However, a number of limitations to the spaghetti plot present obstacles to the display of longitudinal data. Although useful for fewer subjects, trends and patterns are obscured for the larger numbers of subjects typical to modern epidemiologic studies. For example, trajectories commonly overlap in a spaghetti plot, as both subjects and the magnitude of the outcome measure are displayed on the vertical axis. With large datasets, the resulting plot often succumbs to "over-plotting" and is a confusing jumble of intersecting lines with no discernible patterns. Attempts to handle large datasets have traditionally involved plotting a meaningful subset of the data based on medians or deciles, but this approach fails to utilize the whole dataset.[1] Furthermore, repeated measures data containing different enrollment times, missing data, or loss to follow-up (censoring) typically are difficult to display in the spaghetti plot.

Exploratory plots generally are used to reveal various structures in the data: trends (how do most people respond over time), outliers (are some subjects different from all),

clusters (are there groups of patients responding the same, maybe due to a covariate such as treatment assignment), as well as "gut checks" of reasonable outcome values and sensical collection patterns of multi-center repeated measures data. Spaghetti plots suffering from over-plotting lower the probability of data features being revealed.

In this commentary, we advocate for the use of so-called heatmaps[2] as an alternative, or at least complementary, graphical data exploration technique to the spaghetti plot. Heatmaps enjoy frequent use in the genomics literature and other areas with high-throughput data.[3;4;5] However, in more standard longitudinal studies, they are less popular, as evidenced by the recommendation of spaghetti plots of the raw data[1;6;7] or summarized data[8] in popular longitudinal data analysis textbooks.

To remain consistent with the Italian cuisine-themed spaghetti plot, we refer to heatmaps as "lasagna plots." To summarize the conclusion from this article, when graphically exploring longitudinal data, consider lasagna instead of spaghetti!

The proposed lasagna plot uses color or shading to depict the magnitude of the outcome measurement and fixes the vertical dimension per subject. Thus each subject forms a "layer" in the lasagna plot. The lasagna plot takes advantage of color to provide a third dimension and display information clearly, rather than relying upon the vertical dimension to display overlapping magnitudes of change.

Lasagna plots with discrete colors and color gradients can be used for categorical as well as continuous outcomes, respectively. There are several advantages of the lasagna plot over the spaghetti plot. First, group, cohort, and individual level information are preserved regardless of the number of subjects or time points. Second, dynamic sorting of the data can be used to ascertain group level behavior over time. Third, the clear display of missing data and easy handling of intermittent missing data. Finally, an easier display and visualization of the distribution of onset times and ending times.

2 The Basics

In a spaghetti plot, each subject is a "noodle" across time. This allows the overlapping of trajectories and can lead to over-plotting, which often obscures data features. The lasagna plot displays each subject as a "layer" across time. By utilizing color and a fixed vertical position for each subject, overlapping is prevented. The framework of lasagna plots enables dynamic sorting of data. In the following section, we demonstrate how the information in spaghetti plots are represented in lasagna plots and elaborate on the dynamic sorting.

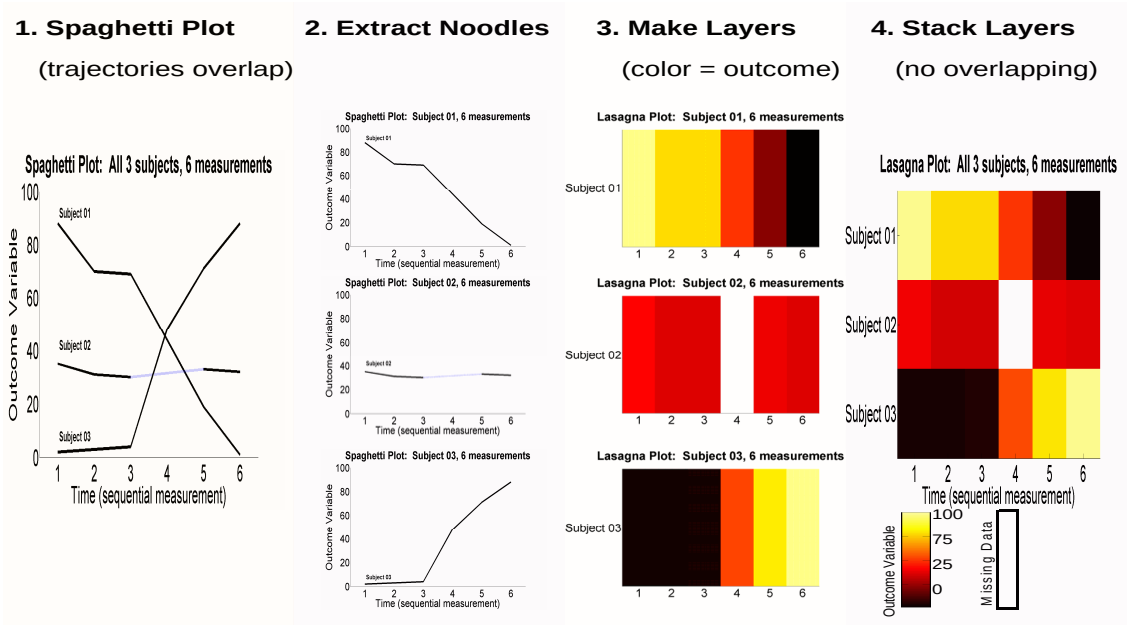
1. Spaghetti Plot
(trajectories overlap)
2. Extract Noodles
3. Make Layers
(color = outcome)
4. Stack Layers
(no overlapping)
Spaghetti Plot: All 3 subjects, 6 measurements
Subject 01
Subject 02
Subject 03
Outcome Variable
Time (sequential measurement)
Spaghetti Plot: Subject 01, 6 measurements
Spaghetti Plot: Subject 02, 6 measurements
Spaghetti Plot: Subject 03, 6 measurements
Lasagna Plot: Subject 01, 6 measurements
Lasagna Plot: Subject 02, 6 measurements
Lasagna Plot: Subject 03, 6 measurements
Lasagna Plot: All 3 subjects, 6 measurements
Missing Data

Figure 1:

2.1 From spaghetti to lasagna

Longitudinal data often are classified by two attributes: the state space X and the time space T, with each space independently being discrete or continuous. Truly continuous time space is usually discretized to the level of sampling repeated measures in epidemiologic studies, so we focus on discrete time data examples. In the case of truly continuous time space or an instance where the discordance of common time points of measurements among subjects in a discretized time space, spaghetti plots may be more feasible than lasagna. With a reasonably large set of mutually shared time points in a discretized time space, lasagna plots work well for continuous and discrete state spaces, allowing for the usual considerations of heatmap coloring for continuous outcomes or discrete outcomes of high dimensions.

Conceptually, the process of lasagna plotting can be constructed from that of a spaghetti plot (Figure 1). However, a matrix construction may be more practical. Epidemiologic longitudinal data following many subjects over time can be represented by a history matrix $\mathbf{H}$. $\mathbf{H}$ is a $m \times n$ matrix where the m rows are the number of subjects in the group and n, the number of columns, is representative of the maximum number of intervals of recording of all groups comparatively. Therefore, an element h_{ij} of $\mathbf{H}$ (where $i = 1, 2, ..., m$ and $j = 1, 2, ..., n$) is a number that corresponds to the state/outcome measurement of the ith individual during the jth interval. In this construct, the matrix $\mathbf{H}$ is a stack of rows, with each row depicting an individual's path over time. Each row of $\mathbf{H}$ contains subject-specific data of repeated measurements. Each column contains cross-sectional data at the group level across subjects.

Transforming the matrix containing numbers into a graphical visualization via "painting by number" gives a snapshot of the data. A lasagna plot provides a simple image with color representing the outcome measure. Each row contains a subject's clustered measurements with each measurement taking place over the column variable, time (or location). Recognizing that the underpinning of the image is a matrix, we dynamically broaden the scope of information that can be obtained via sorting. Given $\mathbf{H}$, five sortings are possible within this framework: within-row, entire-row, within-column, and entire-column, and cluster.

2.2 Dynamic Sorting

The sortings are a way to rearrange data to better visualize potentially obscured information. They can be done on the original $\mathbf{H}$ or on a resultant sorting.

1. Within-row: sorts values of each subject in ascending order. This preserves subject specific information, but no longer displays the temporal ordering.

2. Entire-row: sorts layers of subject in ascending order of a characteristic. These characteristics can be internal or external. An internal characteristic would be a feature in

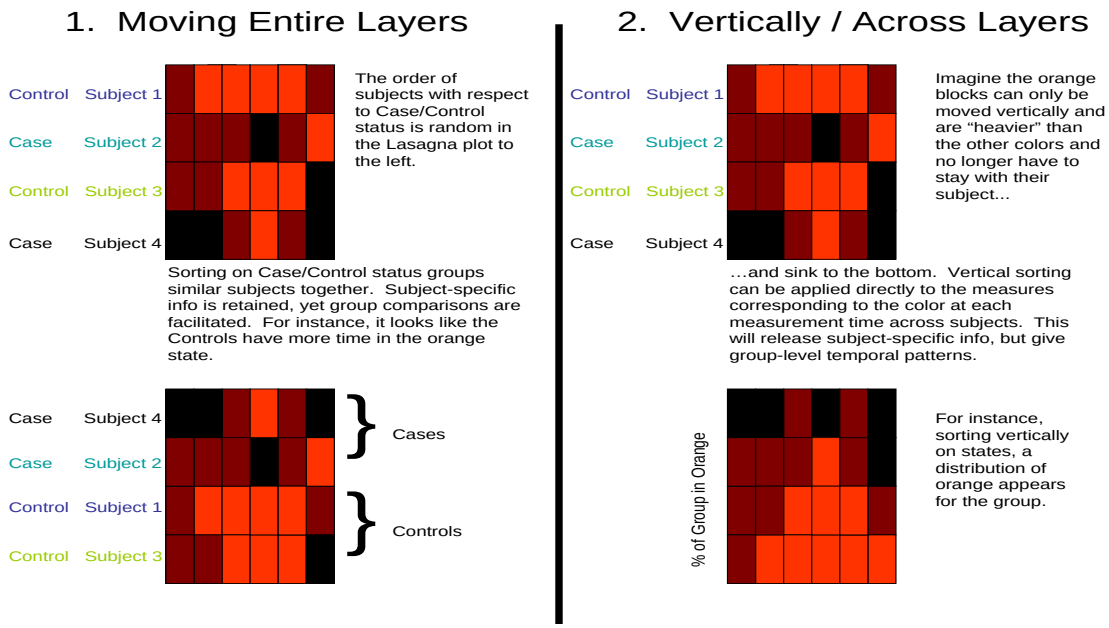

Figure 2:

the lasagna plot, for example, the subject specific means of the outcome. An external characteristic could be baseline age of subjects. Entire-row sorting organizes cohorts together for analysis of cohort effects. This sort preserves subject specific information (Figure 2).

3. Within-column: sorts values across subjects in ascending order within each epoch. This no longer displays subject specific information, but may reveal group-level temporal patterns. This is often called "vertical sorting" (Figure 2).

4. Entire-column: rearranges the columns of a lasagna plot in ascending order based on a characteristic, which can be internal or external. An internal characteristic is one apparent from the information displayed in the graph, such as mean outcome value across subjects at each measurement time. An external characteristic could be measurements of a second outcome variable. This sorting could be useful when the variable is seasonal, for example.

5. Cluster: A specific type of entire-row sorting where the characteristic is internal and a clustering algorithm such as hierarchical clustering or K-means is used so that subjects that have similiar trajectories/layers are grouped together into clusters. Discrete outcomes can be ordered by strata or comparable clustering algorithms, such as correspondence analysis, can be used.

Whereas spaghetti plots are static and in the case of over-plotting stand to obscure trends, outliers, clusters, and "gut checks," a series of lasagna plots with sequential and

cumulative sorting may allow the data to tell its story. Subject-specific trends inherently are easier to see in a lasagna plot due to layers not overlapping as noodles do in a spaghetti plot, and with dynamic sorting trends of different cohorts (entire-row sorting on a classifying variable) and of the entire study population (within-column sorting) are possible. Outliers can be identified with various color spectra corresponding to threshold definitions as well as dynamic (entire-row) sorting. Clusters can be viewed by entire-row sorting on external variables or discovered by using cluster sorting. "Gut checks" are conducted by seeing data collection patterns, for example, changing the time variable from the subject-specfic visit count to calendar month.

3 Motivating Example

Deferring study details (see Document, Supplemental Digital Content 1, which contains study details and corresponding lasagna plots), we display spaghetti plots for one individual, for 120 individuals, the corresponding unordered lasagna plot for 120 individuals, and the corresponding row-sorted lasagna plot.

Addressing Figure 3: a spaghetti plot for one individual is informative, for 120 subjects, less so. The overlapping of multiple trajectories leads to an obscuring of trends for a moderate number of subjects, and consequently the conveyance of intermittent missing data fails. In the lasagna plot for 120 subjects the subjects (rows) appearing in random order, but the intermittent missing data (white) is clearly conveyed. After an entire row sort on disease status and date of EEG recording within disease status, the intermittent missing data is not only conveyed, but the sort allows the exploration of possible trends. After the sort, the darker red region indicates that the diseased have less percent δ sleep, it is seen that only the diseased have missing data, and that the recorder successfully recorded the first 19 SDB subjects, then malfunctioned for the next 11 recording dates in a way where it dropped measurements approximately every hour of sleep from onset. The recorder was righted and operated with full functionality for the next 14 SDB subjects, only to malfunction again by dropping measurements about three hours from sleep onset for the next six SDB subjects. The issue was addressed, and the recorder successfully recorded the rest of the SDB subjects. Comparing the two bottom panels of Figure 3, the same outcome information is in them, but lasagna plots more effectively depict the data because the non-overlapping of trajectories keeps the outcome information uncluttered and its sorting can incorporate more information. As was seen, this extra information can be intermittment missing data and cohort effects, and as outlined in subsequent examples (see Document, Supplemental Digital Content 1, which contains study details and corresponding lasagna plots), it can also be in the form of informative censoring and practice effects.

In summary, we contend that lasagna plots stand to nicely complement spaghetti plots

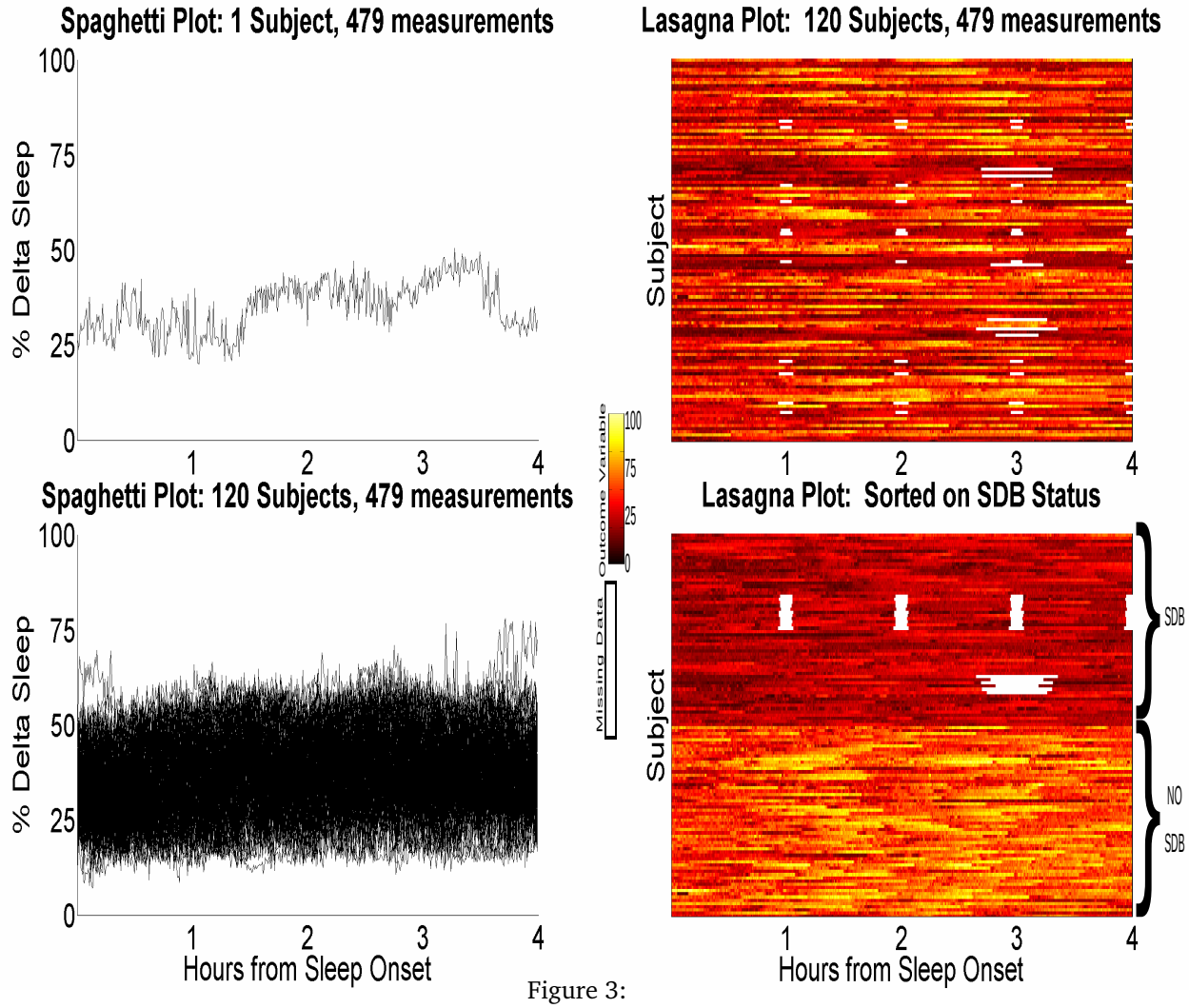

Figure 3:

in many aspects for graphically exploring epidemiologic longitudinal data. Coding for lasagna plots is easily done in R[9] (see Document, Supplemental Digital Content 2, which contains R code snippets).

`Supplemental Digital Content 1. Document with more examples and figures from Sleep Heart Health Study and the Former Lead Worker Study. pdf`

`Supplemental Digital Content 2. Document with R code snippets. pdf`

This document serves as an online supplement to "Lasagna plots: A saucy alternative to spaghetti plots."

4 More Examples

We have used lasagna plots to aide the visualization of a number of unique disparate datasets, each presenting their own challenges to data exploration. Three examples from two epidemiologic studies are featured: the Sleep Heart Health Study (SHHS) and the Former Lead Workers Study (FLWS). The SHHS is a multicenter study on sleep-disordered breathing (SDB) and cardiovascular outcomes.[10] Subjects for the SHHS were recruited from ongoing cohort studies on respiratory and cardiovascular disease. Several biosignals for each of 6,414 subjects were collected in-home during sleep. Two biosignals are displayed here-in: the δ-power electroencephalogram (EEG) and the hypnogram. Both the δ-power EEG and the hypnogram are stochastic processes. The former is a discrete-time contiuous-outcome process representing the homeostatic drive for sleep and the latter a discrete-time discrete-outcome process depicting an individual's trajectory through the rapid eye movement (REM), non-REM, and wake stages of sleep.

The FLWS is a study of age, lead exposure, and other predictors of cognitive decline. The study spans a decade, with up to seven study visits and three separate phases of data collection. During each visit, subjects participated in a battery of cognitive tests, resulting in a longitudinal dataset of repeated measures of test scores for each subject.

In the SHHS, we explore the data by disease status, looking for distinguishing patterns within each disease group. The disease under consideration is sleep apnea (or sleep-disordered breathing (SDB)), a condition characterized with repetitive breathing pauses during sleep. Comparing groups requires carefully selected subsamples, and thus our focus is on 59 SDB subjects and 59 subjects without SDB (no-SDB). In both the continuous and discrete outcome examples, our data assumes a wide format, where the number of measurements far exceed the number of subjects. In the FLWS, all 1,110 subjects are analyzed over a maximum of 7 visits. Cluster sorting will help evaluate the presence of two common problems for longitudinal studies of cognitive function: informative censoring and "practice effects."

4.1 Continuous Longitudinal Data with Intermittent Missingness

For continuous electroencephalogram (EEG) signals derived from the sleep studies, four distinct frequency powers are typically discerned via band-pass filters on the Fourier transform: α,β,δ, and θ. Percent δ sleep is defined as $\frac{\delta}{\alpha+\beta+\delta+\theta} \times 100$. Every 30 seconds during sleep, percent δ sleep was calculated for 59 SDB subjects and 59 no-SDB. In this introductory illustrative example, we look at only the first four hours of data for each individual,

so that everyone has a common onset and stopping point. We also assume that the same device was used to record everyone's sleep and thus no two individuals had sleep recorded on the same date. To showcase the capability of displaying intermittent missing data of the lasagna plot, a pattern of missingness is artificially applied. Via dynamic sorting, the pattern of missingness will be revealed, illustrating how patterns can be uncovered with this exploratory data analysis technique of sorting and visualizing.

We see that a spaghetti plot is a salient display of data for one subject, but not for 118 (Figure 6). The corresponding lasagna plot for 118 subjects shows intermittent missing data, and upon entire-row sorting on the external factors of disease status and date of EEG recording reveals intiguing patterns of the missingness, as well as disease-group differences in percent δ sleep (Figure 7). It appears that only subjects with SDB have missing data and that for a period of recording dates measurements were dropped hourly. Possibly the recording device was malfunctioning, subsequently fixed, and then enjoyed a period of proper functionality only to succumb to dropping measurements 3 hours after sleep onset before being repaired again. To explore the group-level characteristics of percent δ sleep evolution over the course of the night, an additional within-column sorting is conducted within disease status (Figure 8). This highlights a temporal undulation to the signal of the no-SDB group, as well as the no-SDB group having generally higher percent δ sleep than the SDB group. Delta sleep is thought to have an important positive association with cognition and is a maker for homeostatic sleep drive.[11]

4.2 Discrete State-Time Data with a Common Onset

The three classifications of sleep stages (Wake, REM, and Non-REM) are discretizations of several continuous physiologic acquired during sleep. The EEG signals are binned into epochs (often 30 seconds) from sleep onset and collectively used to determine what stage of sleep a subject is in. To accommodate the different lengths of sleep time, an absorbing state is utilized to ensure each subject has an equal number of "measurements," which aids visualization. This example showcases data from the SHHS, where 59 diseased subjects were matched on age, BMI, race and sex to 59 non-diseased subjects. Because the outcome is discrete (the state of sleep), the spaghetti plot is a state-time plot specifically known as the hypnogram to sleep physicians. As in the previous example, for one subject, the spaghetti plot shows the durations in states and transitions among states clearly. The subsequent spaghetti plot for all 118 subjects falls prey to over-plotting, limiting its informativeness (Figure 9). A lasagna plot shows the 1,031 outcomes of each of the 118 subjects' trajectories in random order with respect to SDB status (Figure 10). Applying an entire-row sort on the external characteristic of disease status and the internal characteristic of overall sleep time shows that the groups are well matched on total sleep time (even though the two groups were not explicitly matched on this). Note the degree of

fragmentation and the frequency of short and long-term bouts of WAKE of those with SDB compared to controls. This shows there might be a difference in sleep continuity between the two groups, suggesting SDB fragments sleep. It has been conjectured that sleep continuity may be important in the recuperative effects of sleep, especially in the study of sleep disordered breathing (SDB) and its impact on health outcomes.[12;13;14] Applying an additional within-column sorting within disease status shows the difference in REM temporal evolutions among groups (Figure 11). This dynamic sorting shows the SDB group having an overall weaker REM signal, a bimodal first peak, an absence of a peak at hour 3, the presence of a peak at ~ 7.75 hours. In addition, the peaks widen as time increases, which backs empirical findings of REM duration in state time lengthening as the overall sleep progresses.

4.3 Discretized Longitudinal Data

Lasagna plots are also useful in visualizing and detecting many of the common challenges to population-based longitudinal cohort studies in epidemiological research. The former lead workers study is a study of age, lead exposure, and other predictors of cognitive decline. The study spans a decade, with up to seven study visits and three separate phases (tours) of data collection. This complex dataset is beset with missing data, and both left and right censoring of subjects. Because subjects were enrolled over time in multiple tours, subjects could have as few as 2 or as many as 7 study visits. Study dropout is likely to be dependent upon outcome status (declines in neurobehavioral function) resulting in informative censoring. An additional challenge is the problem of a "practice effect": scores on neurobehavioral tests of cognitive function can become better through practice, masking real declines in cognitive abilities. Lasagna plots provide an unique opportunity to visualize these complex data and detect evidence of both informative censoring and a learning effect. In order to do so, lasagna plots are made with visit as the unit of time as well as tour. Each reveal temporal patterns.

The spaghetti plot (Figure 12) for 1,110 subjects over seven visits is over-plotted, but does show a "thinning" of subjects, suggesting many had three visits, distinctly fewer had three to six, and fewer than that had all seven. In order to facilitate detection of potential informative censoring or a learning effect, neurobehavioral scores were binned based on quintiles of the first visit score distribution. The spaghetti plot of the binned quintile data (Figure 13) is over-plotted and uninformative because the number of subjects for each trajectory is not discernable. A spaghetti plot of discrete outcomes on the Y axis can show possible trajectories, but no indication of how many subjects are in the study due to the exact overlapping of trajectories. The lasagna plot shows the loss to follow up for subjects over time even more clearly than the spaghetti plot (Figure 14) . A cluster sort (sorting within the first column, then the second, etc.) allows us to move entire-rows so that

subjects with similar trajectories are closer to one another (Figure 15). Immediately, we can identify a cluster that did not have a value reported for a first visit, but had values for subsequent visits, indicating missing data. These findings highlight the utility of lasagna plots for exploratory data analysis and data validation. The lasagna plot can also help in examining data for informative dropout and practice effects. Here it appears that if one is in the bottom (worst) quintile on the first visit, the loss to follow up is much worse than if one was in the top (best) quintile at visit 1, indicating informative dropout. A practice effect can be discerned crudely if the subjects have higher test scores on their second study visit than their first, and then scores subsequently decline over time. Overall trends in cognitive function over time are visually apparent as the amount of lighter colors decrease from left to right and the amount of darker colors increase, empirically confirming the overall decline in cognitive function observed with aging.

Finally, if an additional within-column sort was conducted, we would derive the classic stacked bar chart (Figure 16). This removes all subject-specific trajectories and instead summarizes overall distributions of neurobehavioral test scores for each study visit. The practice effect is most visible as scores (based on quintiles of the first visit scores) appear to jump up between the first and second study visits, and then decline over time. However, from Figure 16, we cannot ascertain that the individuals in the top quintile on the first visit are in the top quintile on the second visit. Using stacked bar charts prevents statements on typical pathways, whereas a cluster sorted lasagna plot displays the trajectories for full viewing.

The time structure is complex, for subjects' visit 1 measurement may not have taken place in the same tour. Also, the amount of time lapsed between one subject's adjacent visits may not be the same as another subject's due to visit number being interlinked with what tour they enrolled. Analyzing informative censoring and practice effects is further facilitated by making a lasagna plot with tour as the time variable (Figure 17) and then sorting within each Tour the subject's 1st visit quintile cognitive measure (Figure 18). Comparing those enrolled in Tour 1 of the worst and best quintile, we see that there is more dropout for those starting out in the worst quintile, possibly indicating informative censoring. This pattern of drop out being related to first visit quintile rank holds for the later tours as well. Training effects can be seen when a subject's color lightens when tracking that individual across time. For instance, a fair portion of individuals in Tour 1 were in the 2nd best quintile and advance to the best quintile in their second visit in Tour 2.

4.4 Result Tables and Covariate Selection

Simulations under different conditions often give rise to mulitple tables of output. Identifying trends and comparing tables is often an arduous and obfuscating task. With lasagna

plots, a quick snapshot of the tables are rendered, allowing trends within tables to be identified and compared across tables (Figure 19).

In building regression models, it is important to know what variables have high degrees of missingness. For large epidemiologic datasets modeling an outcome, a lasagna plot can be used to show the proportion of the sample covariate missingness over time (Figure 20). Here, the vertical axis is the variable, the horizontal axis is time, and the darker the plot the greater the proportion of the sample that has a reported value for that layer's variable. This helped guide the inclusion and exclusion of covariates in the model building process.

Lasagna plots work well for data tables that have many numbers and are essentially an image of a matrix. Commonly, the layers are denoting an individual, the columns are denoting times or locations largely in common to all the subjects being visualized, and color to reflect the state occupied or magnitude/intensity of the trait. One exception to this paradigm is diary data for an individual. In diary data, the layers are days, the columns are hours, and the colors reflect activities partaken for a certain time on a particular day. This approach has proven useful in mapping out infant and child ideal sleep patterns.[15] Nutritionist colleagues are implementing lasagna plots to display caloric intake and purge cycles amongst those with eating disorders.

5 Discussion

Lasagna plots have been presented as an effective means to explore data that can be arranged into a matrix. The strengths of lasagna plotting are that it can incorporate a wide platform of data structures, ranging from longitudinal repeated measures to multidimensional temporal-spatial (i.e., FMRI) to gene expression of genes by tissue type (i.e. Barcodes).[4;3] Lasagna plots are visualizations that "above all else show the data" and are more akin to the raw data than a modeling procedure.[16;17] Row and column sorting and clustering are intuitive to a non-technical audience, and visualizations of sequential sortings and/or clustering serve as a way to engage a collaborative analysis of data. Weaknesses include the color-dependency, difficulties handling continuous time, and how growing size of the data make seeing individual layers more difficult. The first is becoming less of an issue as digital publication overtakes traditional paper publishing. The second weakness can be ameliorated by coarsening/binning time, and the third by doing sorts or making plots on subsets of the population.

Often, longitudinal data have been traditionally viewed as either a spaghetti plot or stacked bar chart, which falls prey to over-plotting and aggressive summarization, respectively. Multi-state survival (event history) data can be viewed through a longitudinal repeated measures lens, and historically were viewed with eventcharts, corresponding components of dynamic interaction and linked graphs, as well as event history graphs. These

were important steps in visualizing survival data simultaneously at the group level and individual level.[18;19;20] Limitations of the eventchart included difficulty handling multiple groups, large amounts of individuals, denoting multiple events, and the incorporation of color. These limitations are not present with the lasagna plotting of survival data. Lasagna plots work well in most trivariate and multiway data settings, conveying at least the same information as superposed level regions in color plots and multiway dot plots.[21]

Genomics and pediatric sleep science are currently utilizing plots that are special cases of what we call lasagna plots. Recent graphing techniques in the statistical programming and analysis language R have come to show ingenuity in handling complex data, as evinced by the functions `heatmap.2()`, `hist2d()`, `seas`, and `mvtplot`.[22] Lasagna plotting and cluster sorting is implemented by `heatmap.2()`, grouping similar genes (rows) and tissues (columns) together. The essence of lasagna plotting, within-column and entire-row sorting is captured in `mvtplot` the best, in that it displays not only the data itself but simultaneously group level temporal trends in a smoothed curve below and subject specific summary measures on the right sidebar. Lasagna plotting encompasses `mvtplot` and implements dynamic sorting to further explore the data. Discrete outcomes are not handled well by `mvtplot` because the element being visualized are not necessarily numeric, thus the summarizations on the bottom and right hand panel are not useful when the data measures are nominal. Lasagna plotting and subsequent sorting handles the nominal case.

References

[1] PJ Diggle, PJ Heagerty, KY Liang, and SL Zeger. *The analysis of longitudinal data 2nd ed*. Oxford, England: Oxford University Press, 2002.

[2] L. Wilkinson and M. Friendly. The history of the cluster heat map. *The American Statistician*, 2009.

[3] M.J. Zilliox and R.A. Irizarry. A gene expression bar code for microarray data. *Nature Methods*, 4:911–913, 2007.

[4] R. Baumgartner and R. Somorjai. Graphical display of fMRI data: visualizing multi-dimensional space. *Magnetic resonance imaging*, 19(2):283–286, 2001.

[5] D. Sarkar. *Lattice: multivariate data visualization with R*. Springer Verlag, 2008.

[6] D. Hedeker and R.D. Gibbons. *Longitudinal data analysis*. Wiley-Interscience, 2006.

[7] G.M. Fitzmaurice, M. Davidian, G. Verbeke, and G. Molenberghs. *Longitudinal Data Analysis: A Handbook of Modern Statistical Methods*. Chapman & Hall/CRC, 2008.

[8] G.M. Fitzmaurice, N.M. Laird, and J.H. Ware. *Applied longitudinal analysis*. ASA, 2004.

[9] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2009. ISBN 3-900051-07-0.

[10] SF Quan, BV Howard, C. Iber, JP Kiley, FJ Nieto, GT O'Connor, DM Rapoport, S. Redline, J. Robbins, JM Samet, et al. The Sleep Heart Health Study: design, rationale, and methods. *Sleep*, 20(12):1077–85, 1997.

[11] L. Marshall, H. Helgad, M. Matthias, and J. Born. Boosting slow oscillations during sleep potentiates memory. *Nature*, 444(7119):610–613, 2006.

[12] R.G. Norman, M.A. Scott, I. Ayappa, J.A. Walsleben, and D.M. Rapoport. Sleep continuity measured by survival curve analysis. *Sleep*, 29(12):1625–31, 2006.

[13] N.M. Punjabi, D.J. O'Hearn, D.N. Neubauer, F.J. Nieto, A.R. Schwartz, P.L. Smith, and K. Bandeen-Roche. Modeling hypersomnolence in sleep-disordered breathing a novel approach using survival analysis. *American Journal of Respiratory and Critical Care Medicine*, 159(6):1703–1709, 1999.

[14] M.H. Bonnet and D.L. Arand. Clinical effects of sleep fragmentation versus sleep deprivation. *Sleep Medicine Reviews*, 7(4):297–310, 2003.

[15] R. Ferber. *Solve Your Child's Sleep Problems: New, Revised, and Expanded Edition (Paperback)*. Simon & Schuster, 2006.

[16] E.R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, 2001.

[17] T. Lumley and P. Heagerty. Graphical exploratory analysis of survival data. *Journal of Computational and Graphical Statistics*, pages 738–749, 2000.

[18] A.I. Goldman. Eventcharts: Visualizing survival and other timed-events data. *American Statistician*, pages 13–18, 1992.

[19] E.N. Atkinson. Interactive dynamic graphics for exploratory survival analysis. *American Statistician*, pages 77–84, 1995.

[20] J.A. Dubin, H.G. Muller, and J.L. Wang. Event history graphs for censored survival data. *Statistics in Medicine*, 20(19):2951–2964, 2001.

[21] W.S. Cleveland. *Visualizing data*. Hobart Press, 1993.

[22] RD Peng. A method for visualizing multivariate time series data. *Journal of Statistical Software*, 25(Code Snippet 1):687–713, 2008.

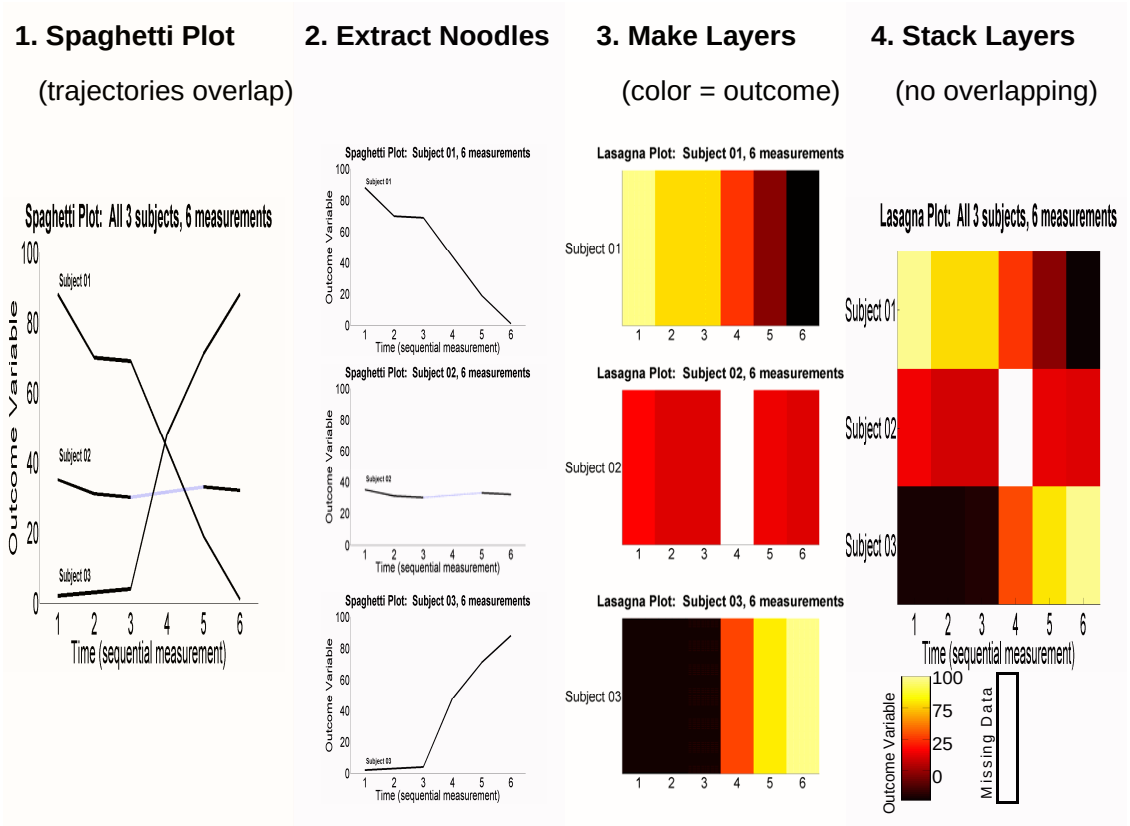

Figure 4: Lasagna plots as derived from spaghetti plots involve making noodles into layers. From left to right, a spaghetti plot with three noodles where trajectories overlap. Extracting each noodle representing repeated measures on a subject, a layer is made by letting color represent the outcome. Individual layers are then stacked to make a lasagna plot, with no overlapping of subject information.

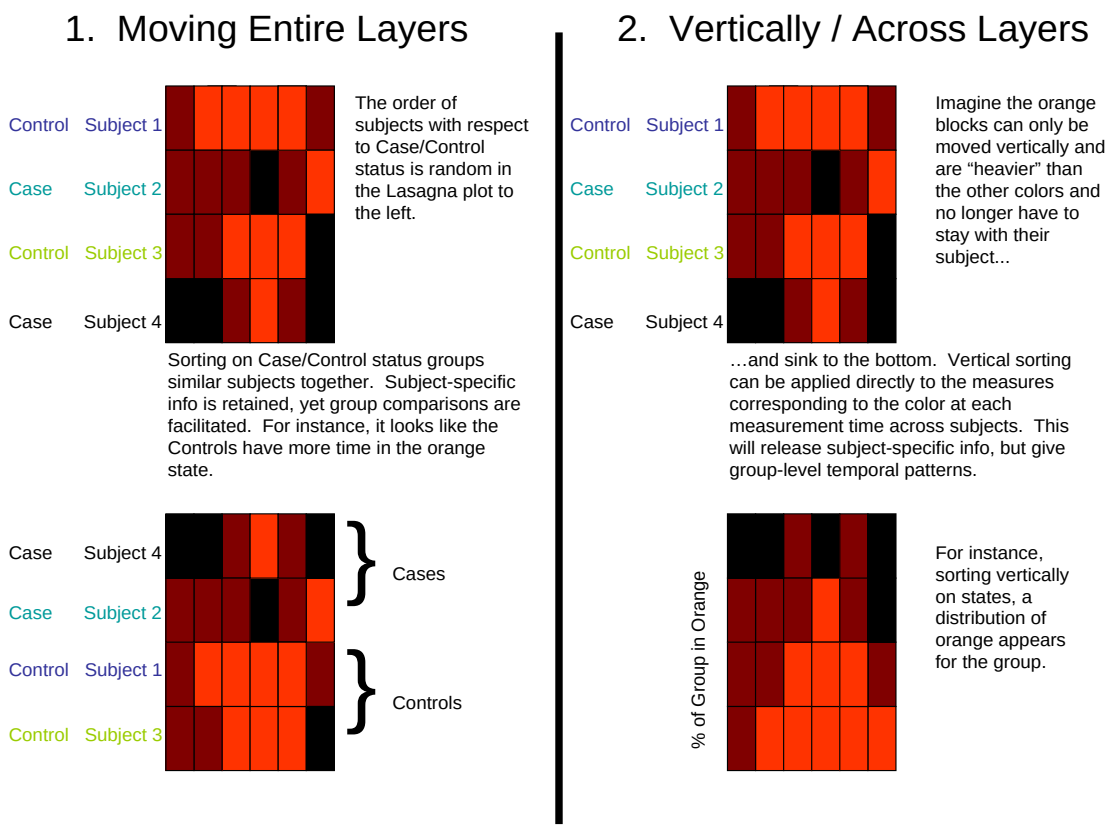

Figure 5: Dynamic sorting of types "entire row" and "within column" are depicted. Often, subjects in a dataset appear in random order. Entire row sorts can organize the data by covariates to explore possible associations between outcomes and covariates. Within column sorting is useful in ascertaining group-level trends of outcomes over the repeated measures.

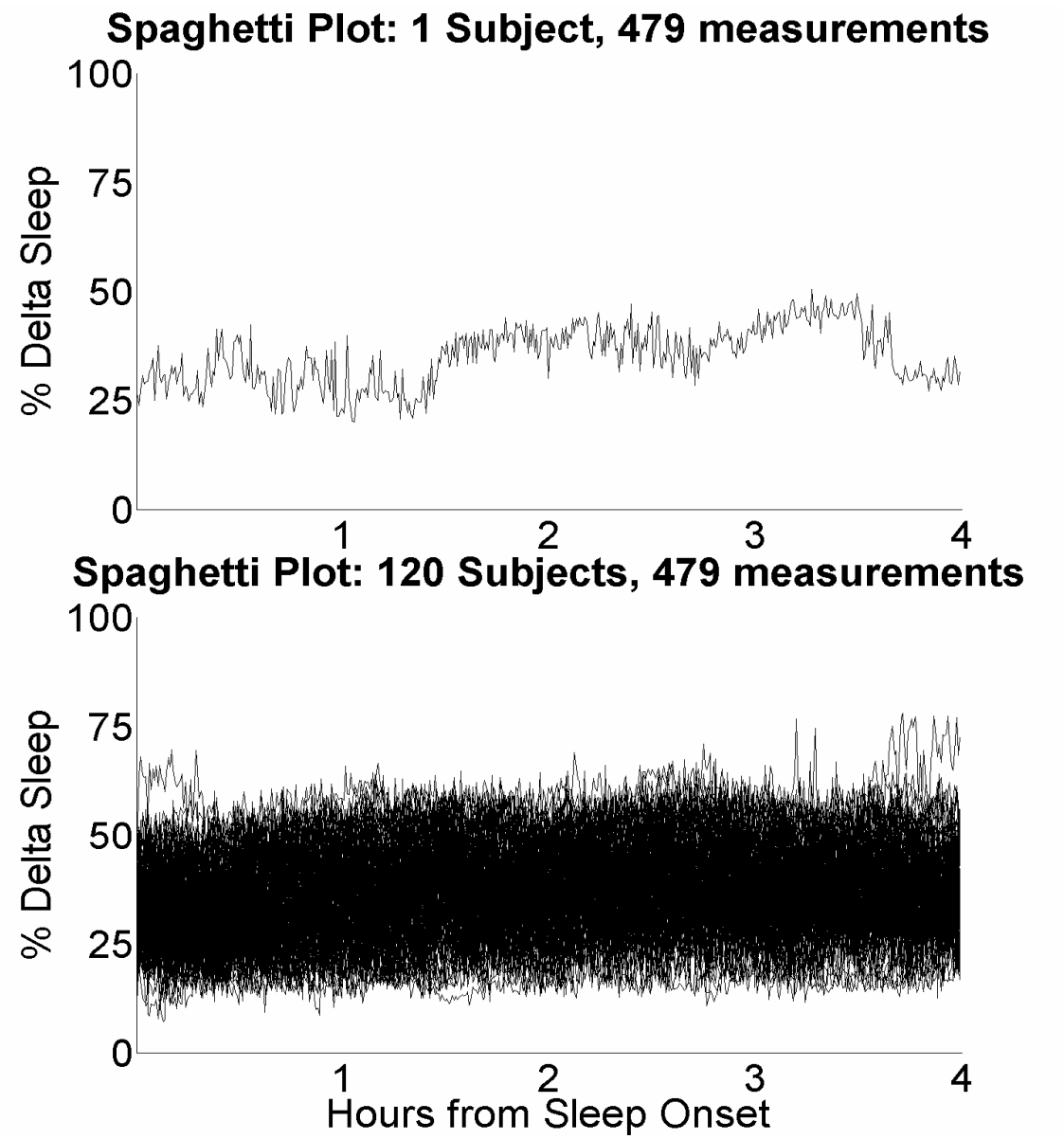

Figure 6: Top panel: Spaghetti plot for one individual. Bottom panel: Spaghetti plot for 120 subjects. The overlapping of multiple trajectories leads to an obscuring of trends for a moderate number of subjects, and consequently the conveyance of intermittent missing data fails.

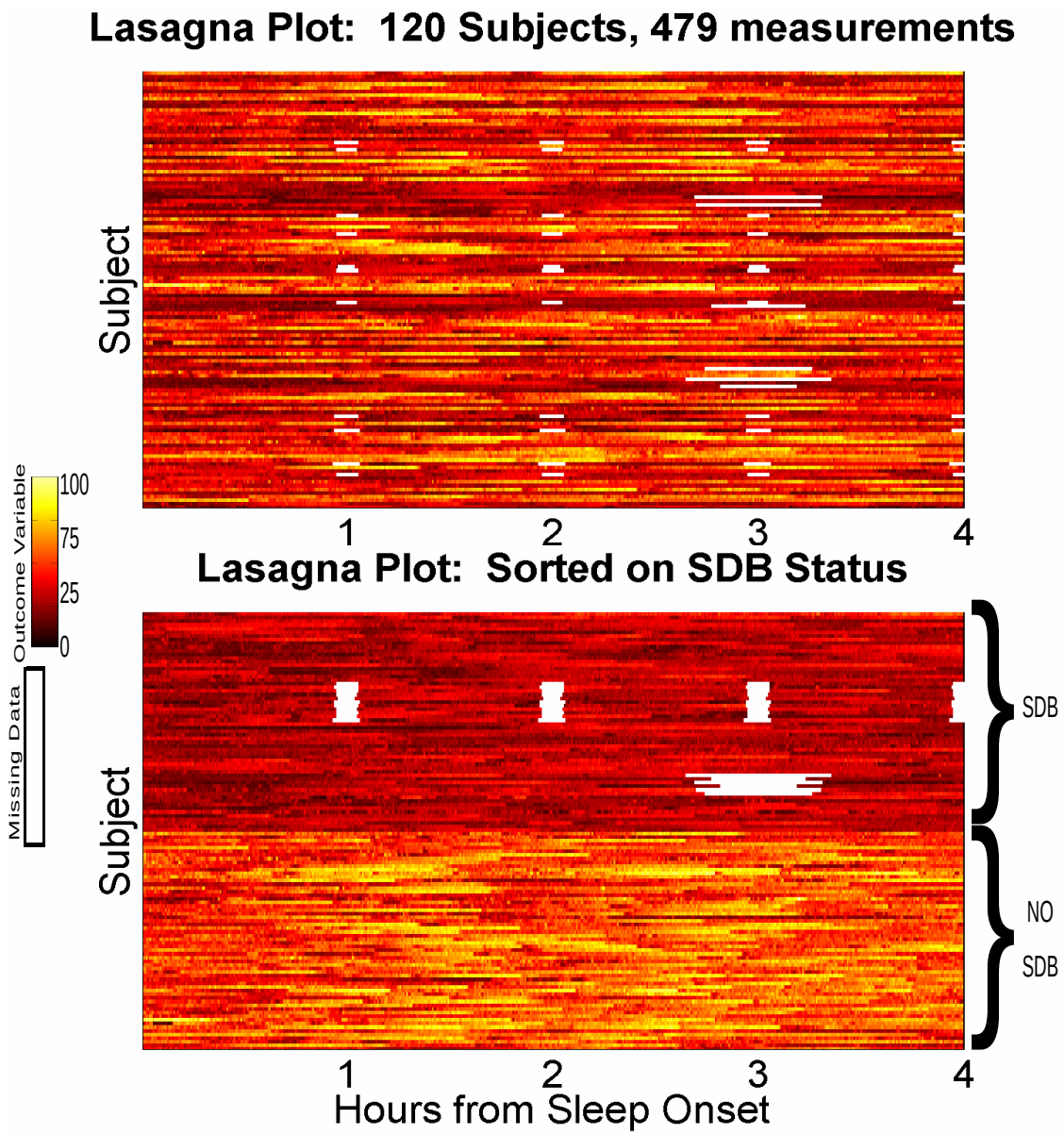

Figure 7: Top panel: Lasagna plot for 120 subjects from the bottom panel of Figure 6. The subjects (rows) appear in random order, but the intermittent missing data (white) is clearly conveyed. Bottom panel: Lasagna plot of the top panel after an entire row sort on disease status and date of EEG recording within disease status. The intermittent missing data is not only conveyed, but the sort allows the exploration of possible trends. After the sort, the darker red region indicates that the disease have less percent δ sleep, it is seen that only the diseased have missing data, and that the recorder successfully recorded the first 19 SDB subjects, then malfunctioned for the next 11 recording dates in a way where it dropped measurements approximately every hour. The recorder was righted and operated with full functionality for the next 14 SDB subjects, only to malfunction again by dropping measurements about three hours from sleep onset for the next six SDB subjects. The issue was addressed, and the recorder successfully recorded the rest of the SDB subjects. Compare the the bottom panel her to that of Figure 6. The same outcome information is in them, but lasagna plots more effectively depict the data because the non-overlapping of trajectories keeps the outcome information uncluttered and its sorting can incorporate more information.

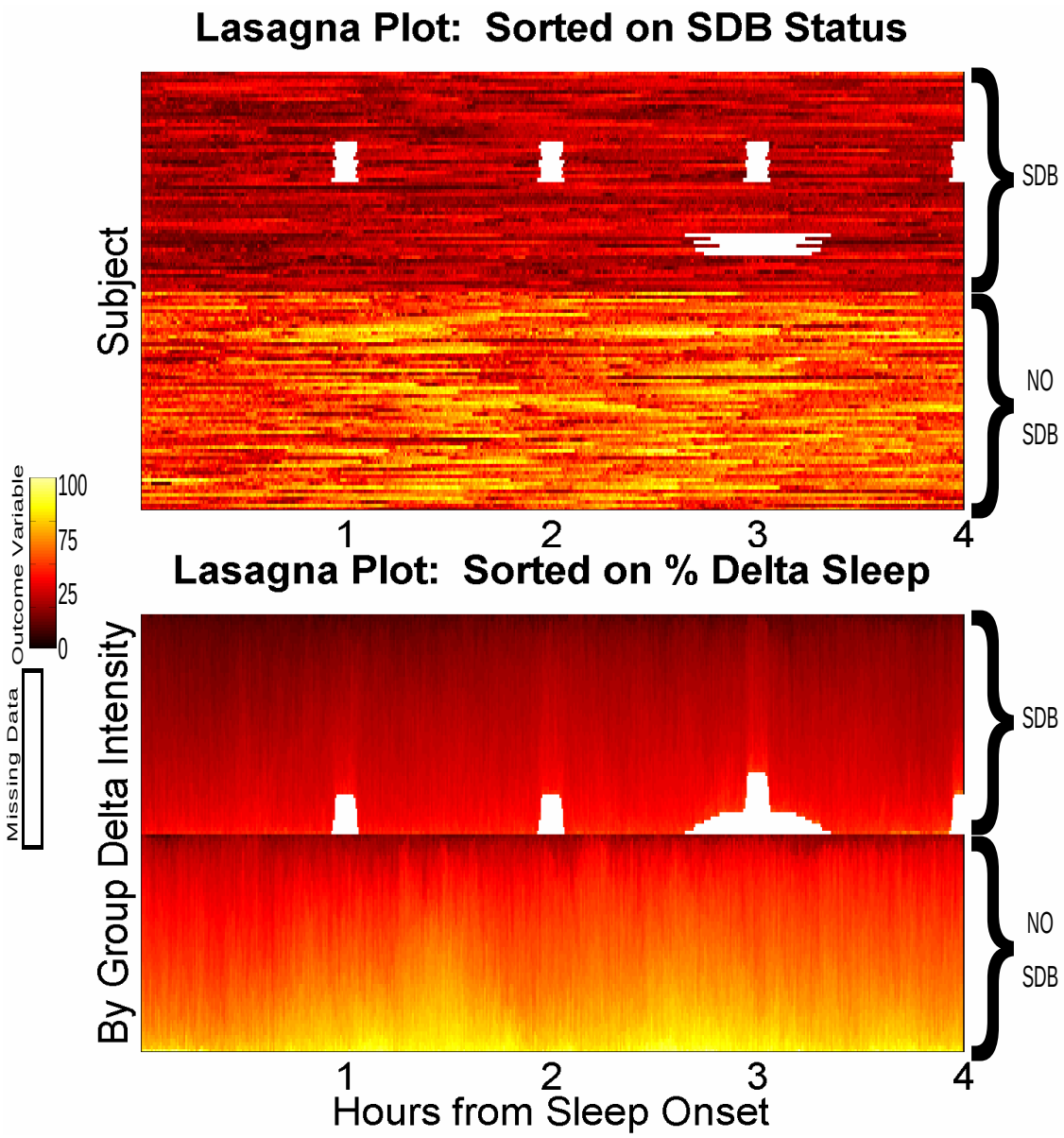

Figure 8: Top panel: The entire row sorted lasagna plot of Figure 7. Bottom panel: A within-column sort applied within disease status to the lasagna plot of the top panel. Note the wax and wane of the yellow in the no-SDB group, depicting the group-level temporal evolution of percent δ sleep in subjects without SDB.

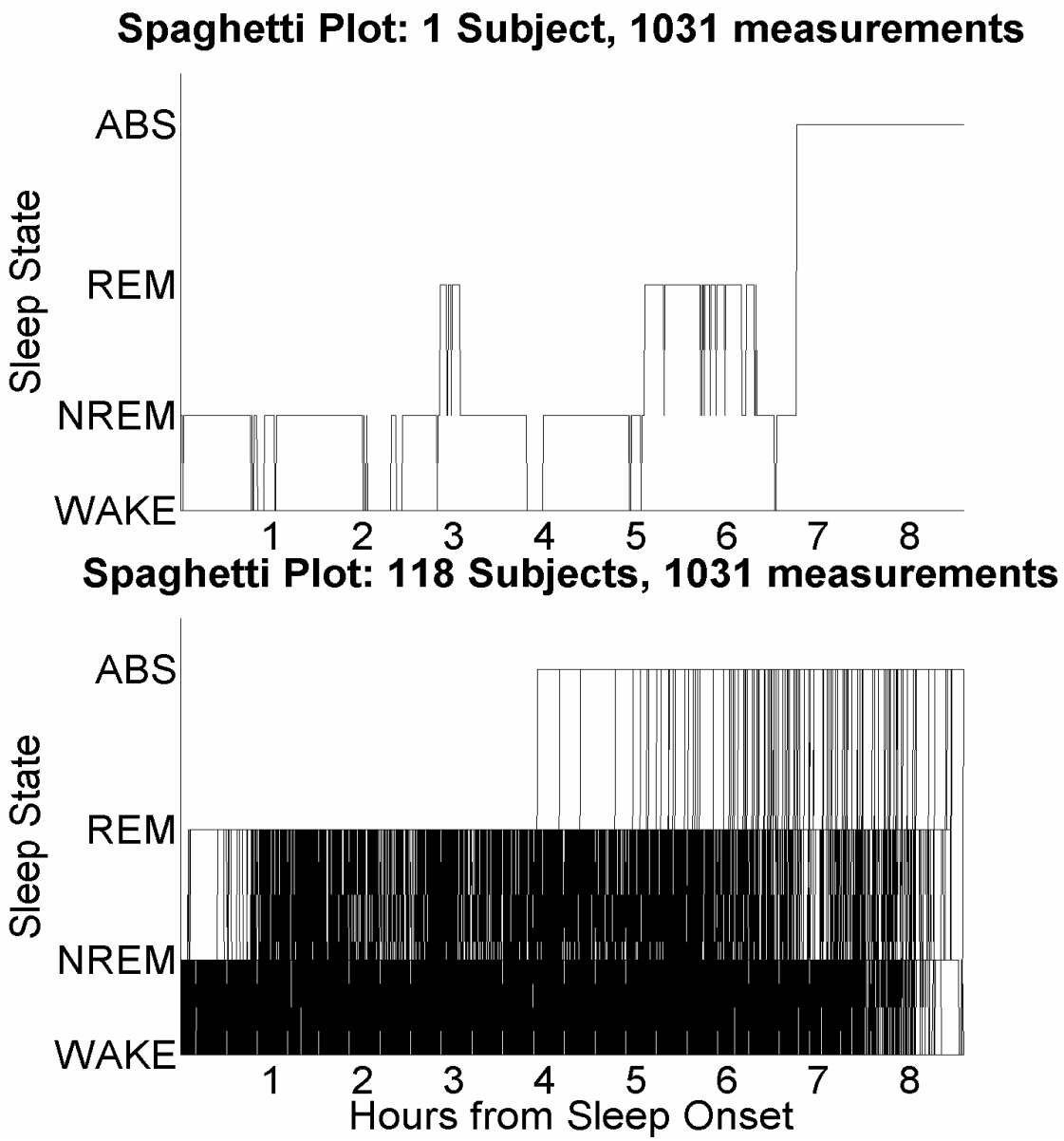

Figure 9: Top panel: a spaghetti plot for a discrete outcome for one subject. Bottom panel: a spaghetti plot for a discrete outcome for 118 subjects. Due to the discrete nature of the outcome, trajectories do not run the risk of merely crossing each other as in continuous outcome cases, but overlapping each other exactly. The informativeness of the spaghetti plot for discrete data on a moderate number of subjects is limited.

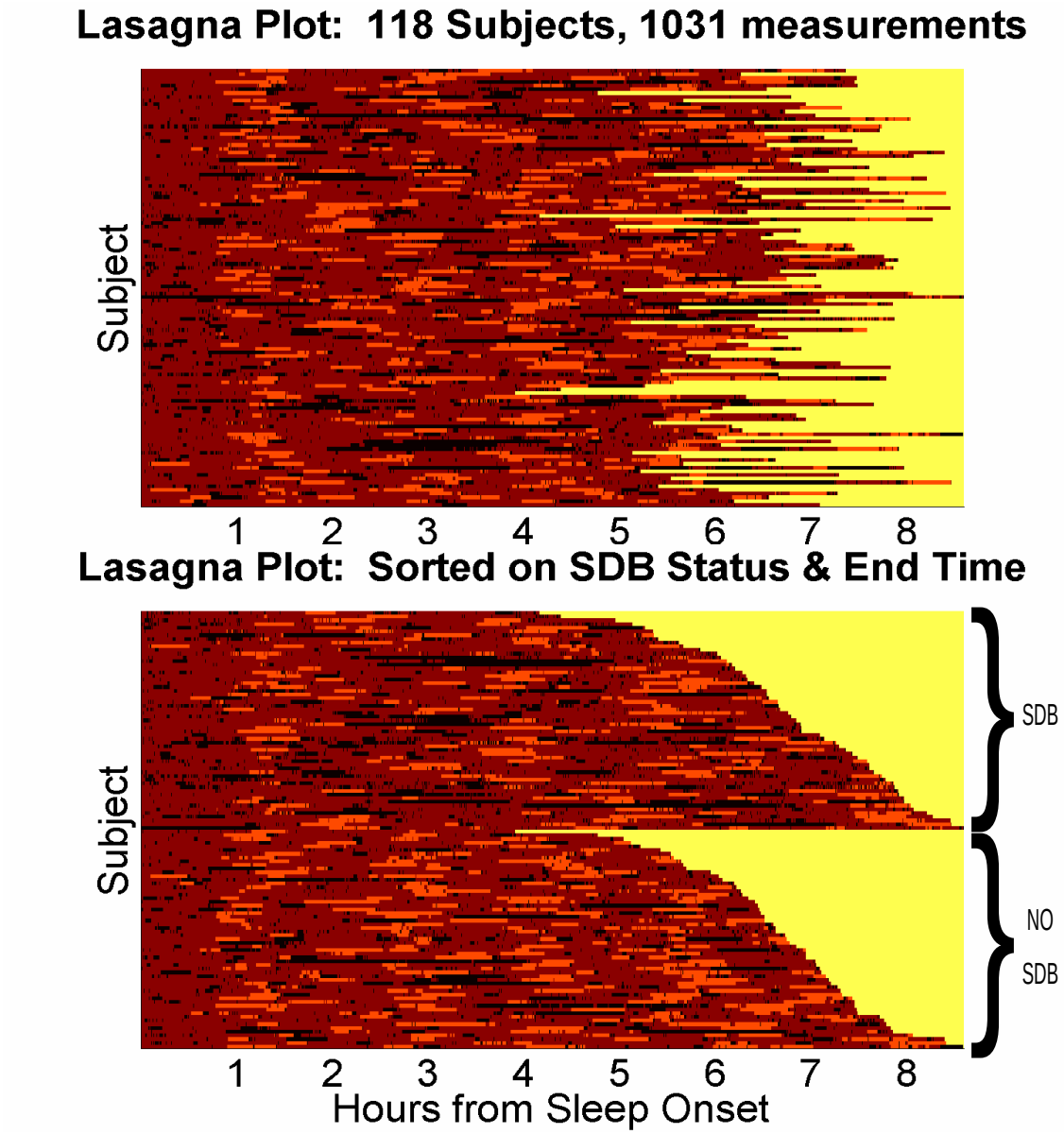

Figure 10: Top panel: corresponding lasagna plot for the spaghetti plot in the bottom panel of Figure 9, with subjects (rows) in random order. Bottom panel: the resulting lasagna plot after an entire row sort of disease status and sleep time recorded. This data organization allows easy comparison of the yellow areas showing that each group has similar distribution of sleep time recorded.

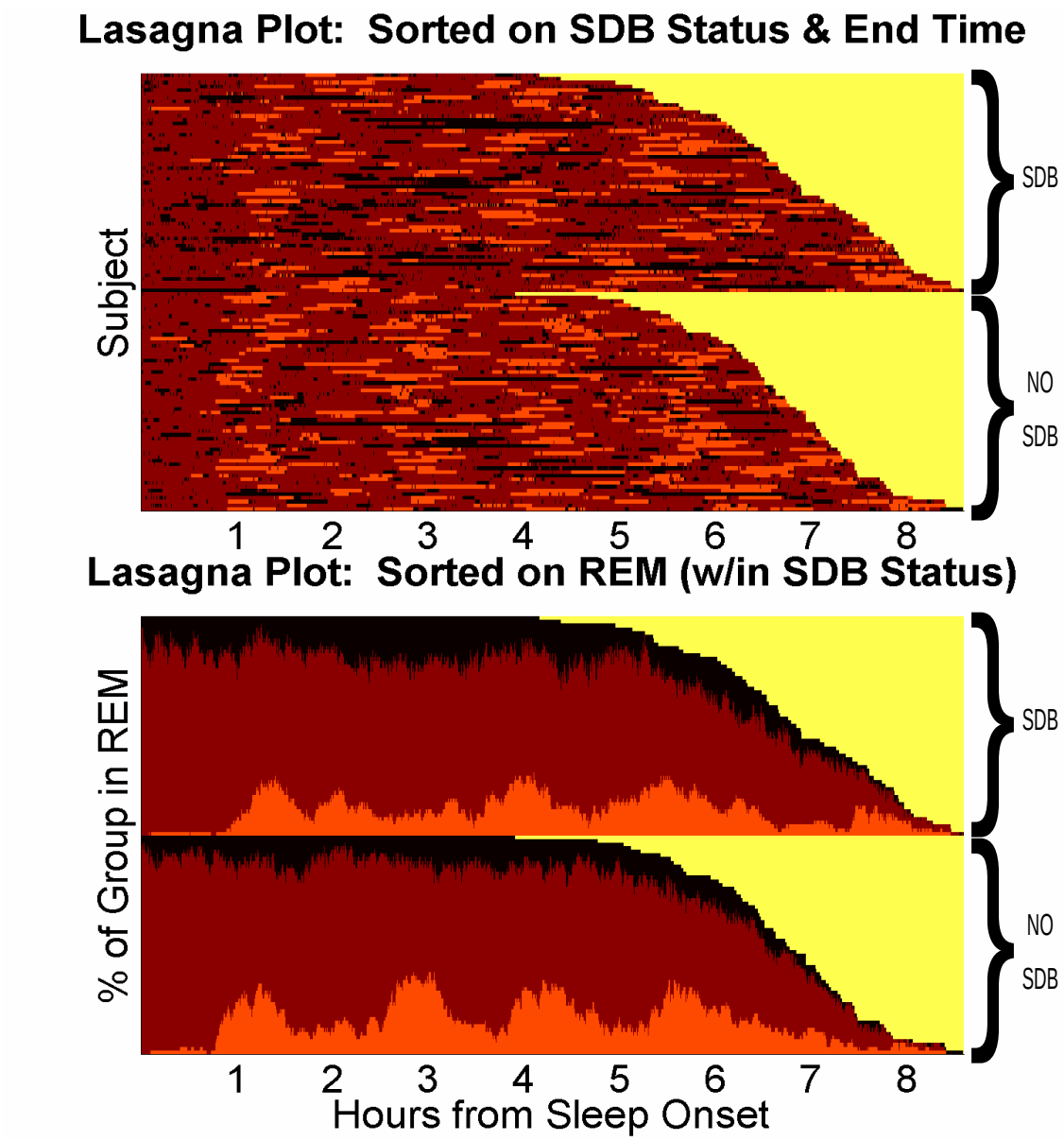

Figure 11: Top panel: same plot as the bottom panel of Figure 10. Bottom panel: the resulting lasagna plot after a within column sort applied within disease status. This data organization shows group-level temporal evolution of REM sleep. The signal seems to be more pronounced in those without SDB.

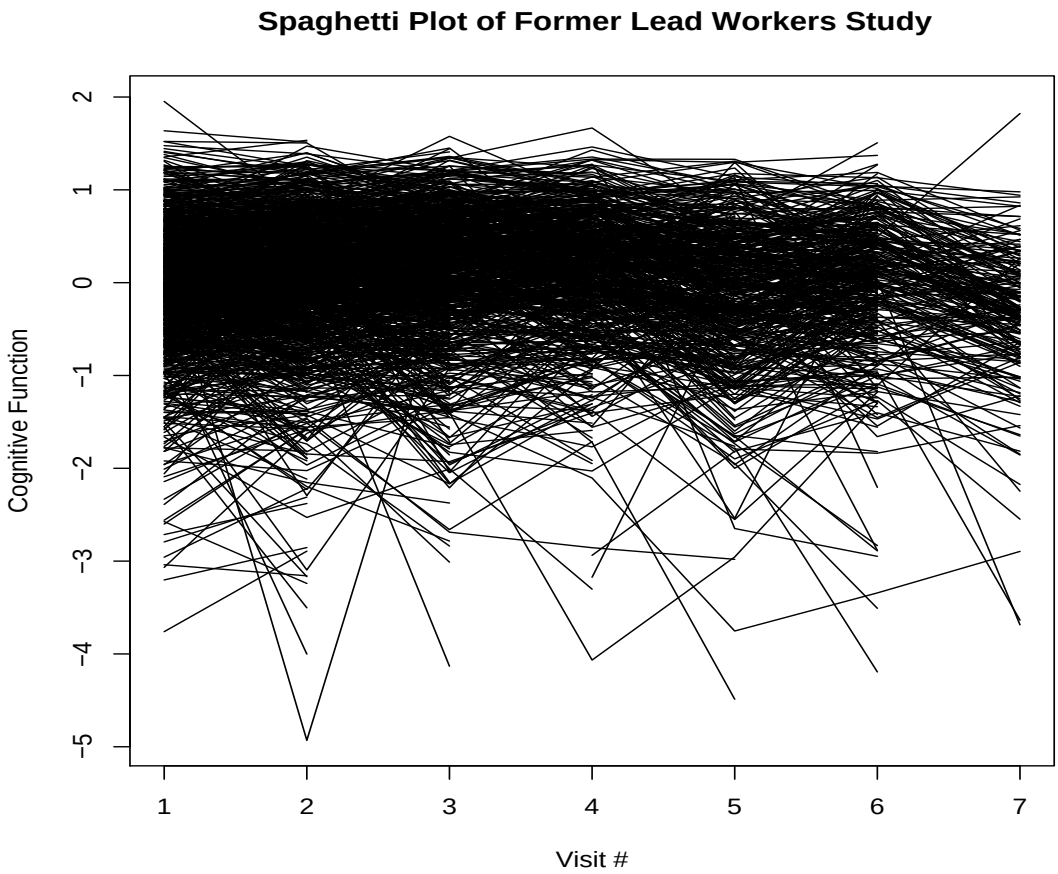

Figure 12: Spaghetti plot for a continuous cognitive measure of 1110 subjects over 7 visits.

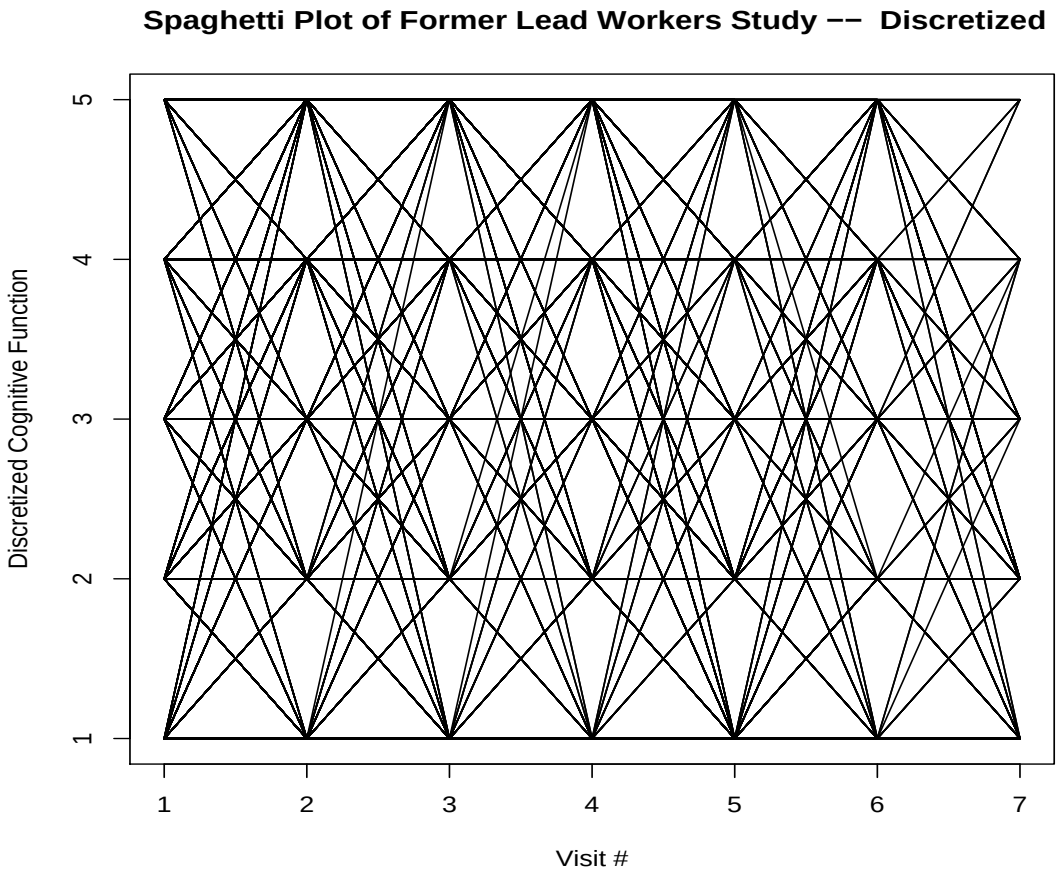

Figure 13: Spaghetti plot for the discretized cognitive measure of Figure 12 for 1110 subjects over 7 visits. The discretization was based on quintiles of the 1st visit outcome measures. The only information this plot can guarantee is that if a line exists between quintile nodes for adjacent visits, at least one subject made that move - it does not show how many subjects took the path, and it cannot show specific paths over multiple nodes for one subject. For instance, notice the absence of a line connecting the 1st visit 5th quintile node to the 1st quintile node of visit 7 (there is no line going from the upper left of the graph to the lower right). This means no subject was recorded on only the first visit and the last visit with no visits between with the measurements recorded having her start out in the top quintile and declining to the bottom quintile. The absence of a line means the path was not taken. However, in a spaghetti plot of discretized data, the presence of a line over multiple nodes does not indicate that the path was taken by a subject. For instance, no individual only had two measurements taken on visit 1 and visit 5 and went from the top quintile to the bottom quintile, yet there is a line between 1st visit 5th quintile node to the 5th visit 5th quintile node. One cannot tell from the spaghetti plot alone if a path is made of one individual between two non-adjacent nodes, or several individuals making the pairwise adjacent transitions. For instance, the line between 1st visit 5th quintile node to the 5th visit 5th quintile node could comprise four individuals: one going from the 5th quintile to the 4th from visit 1 to visit 2, one going from the 4th quintile to the 3rd from visit 2 to visit 3, one going from the 3rd quintile to the 2nd from visit 3 to visit 4, and one going from the 2nd quintile to the 1st from visit 4 to visit 5.

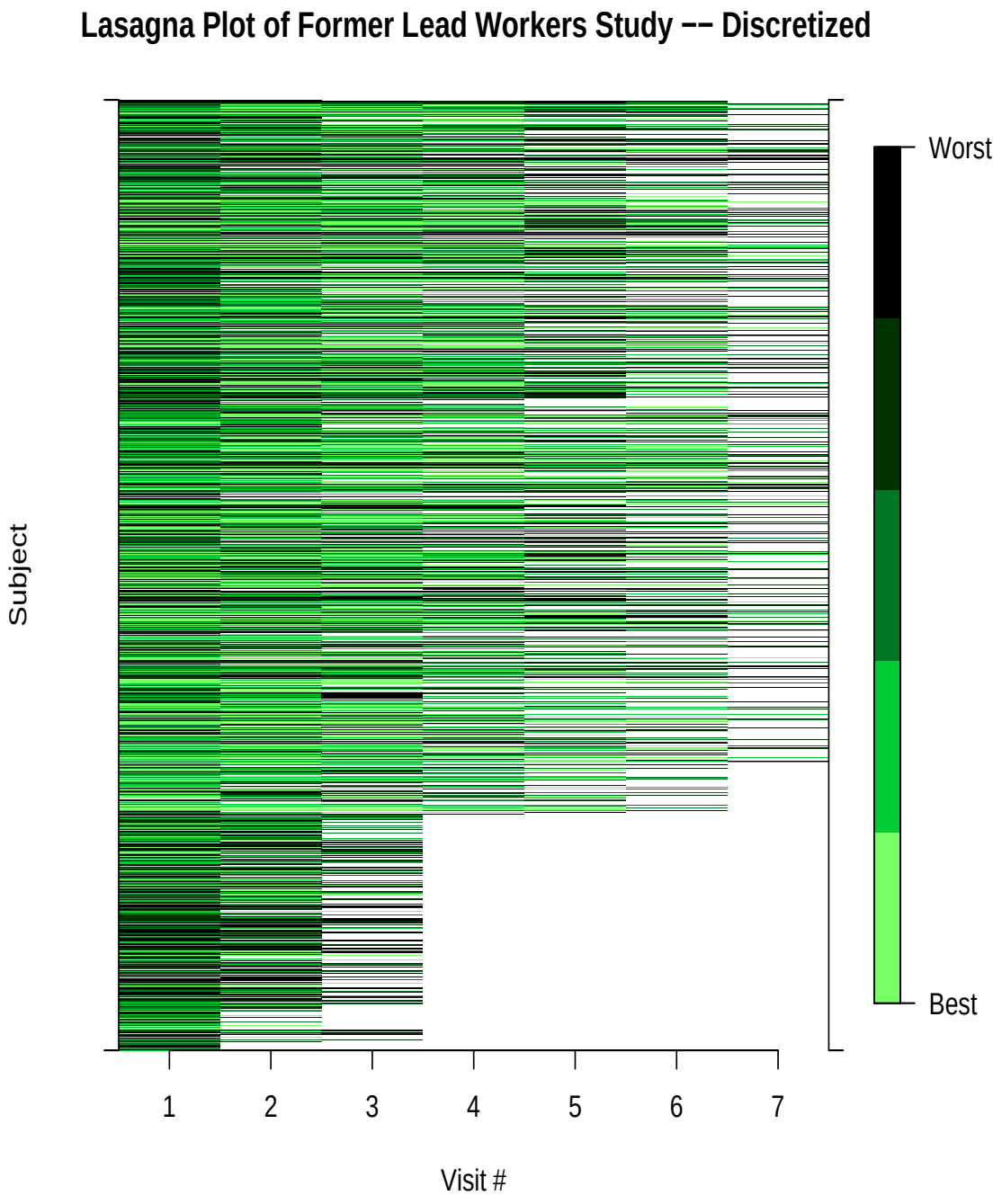

Figure 14: Lasagna plot for 1110 subjects over 7 visits, from Figure 13. This image depicts paths taken by indviduals more clearly than the discretized spaghetti plot.

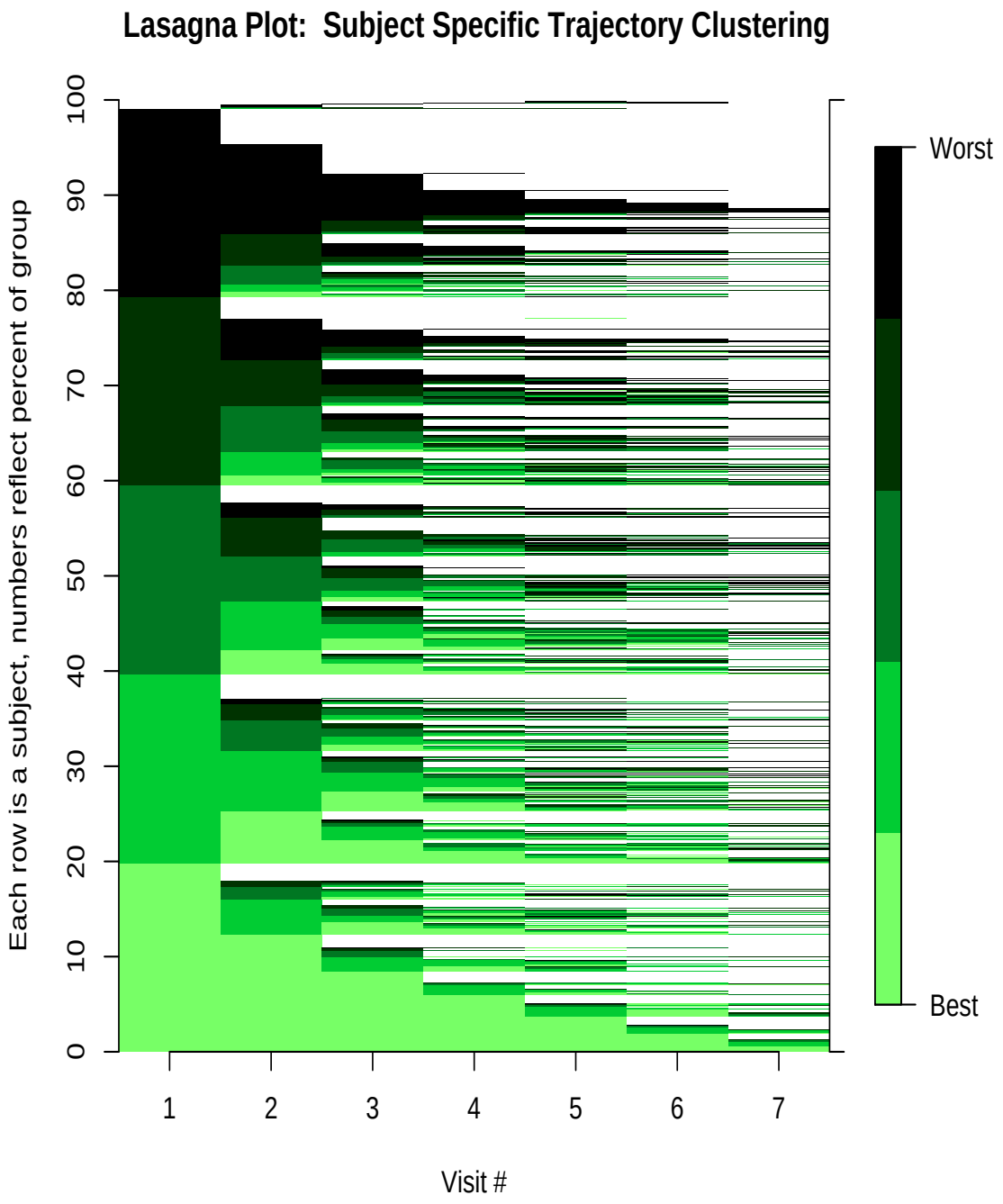

Figure 15: This is a cluster sort of Figure 14. Similar trajectories are grouped together and subject-level information is maintained and the association of cognitive ability by the metric of baseline quintiles across the visit structure can be analyzed.

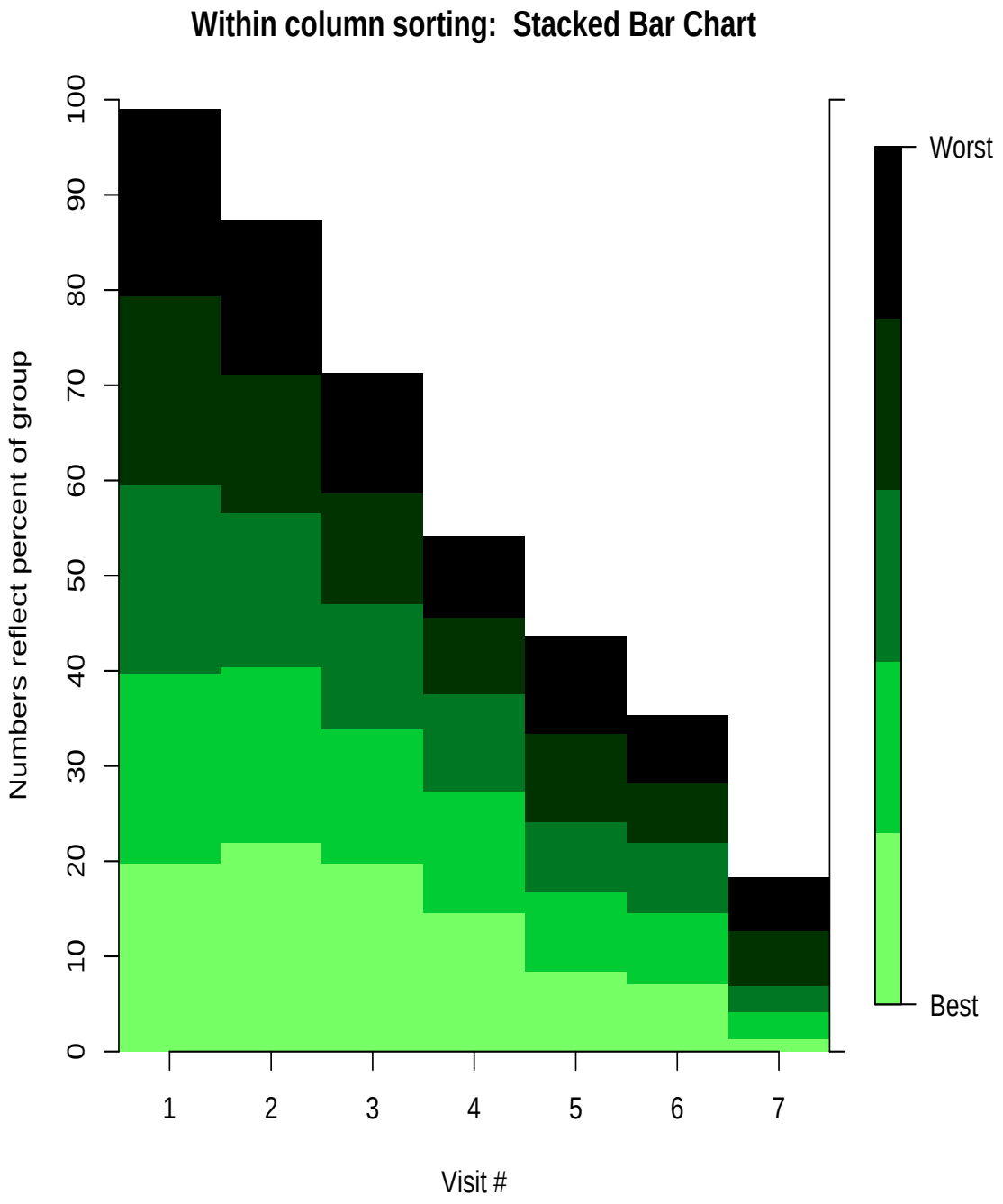

Figure 16: This plot can be thought of as the within column sorting of Figure 15, which shows the derivation of the classic stacked bar chart. This severes the connection of repeated measures within subject completely and is a strong summarization of the data in that it discards a lot of information. We could not see the distributions of 2nd visit best qunitile score conditional on quintile of the score of visit 1. However, we can see a bump of best scores from visit 1 to visit 2, indicating a possible training effect.

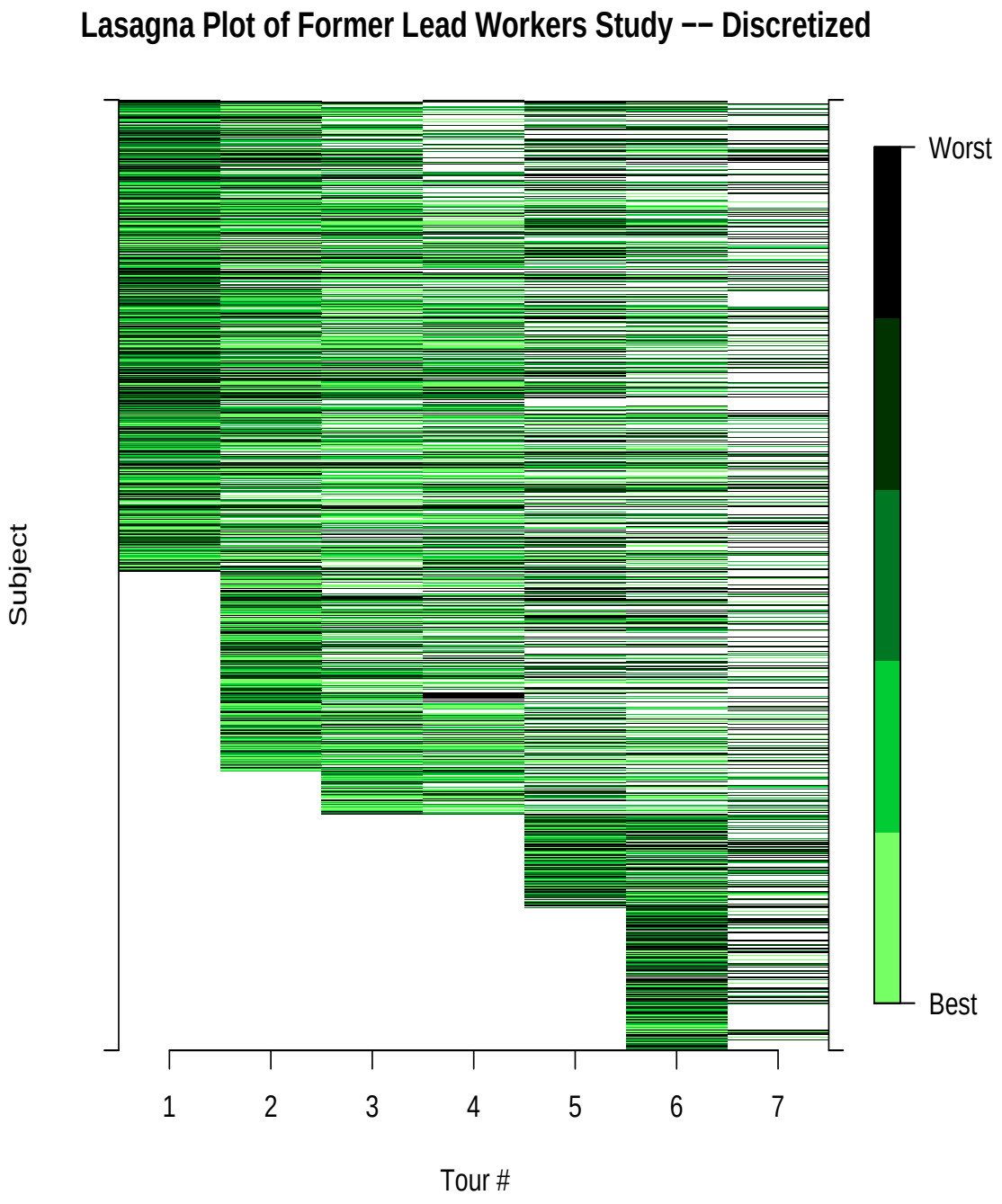

Figure 17: Lasagna plot for 1110 subjects over 7 tours, compare to the same data plotted by visit (Figure 14), with subjects in random order within their tour of enrollment.

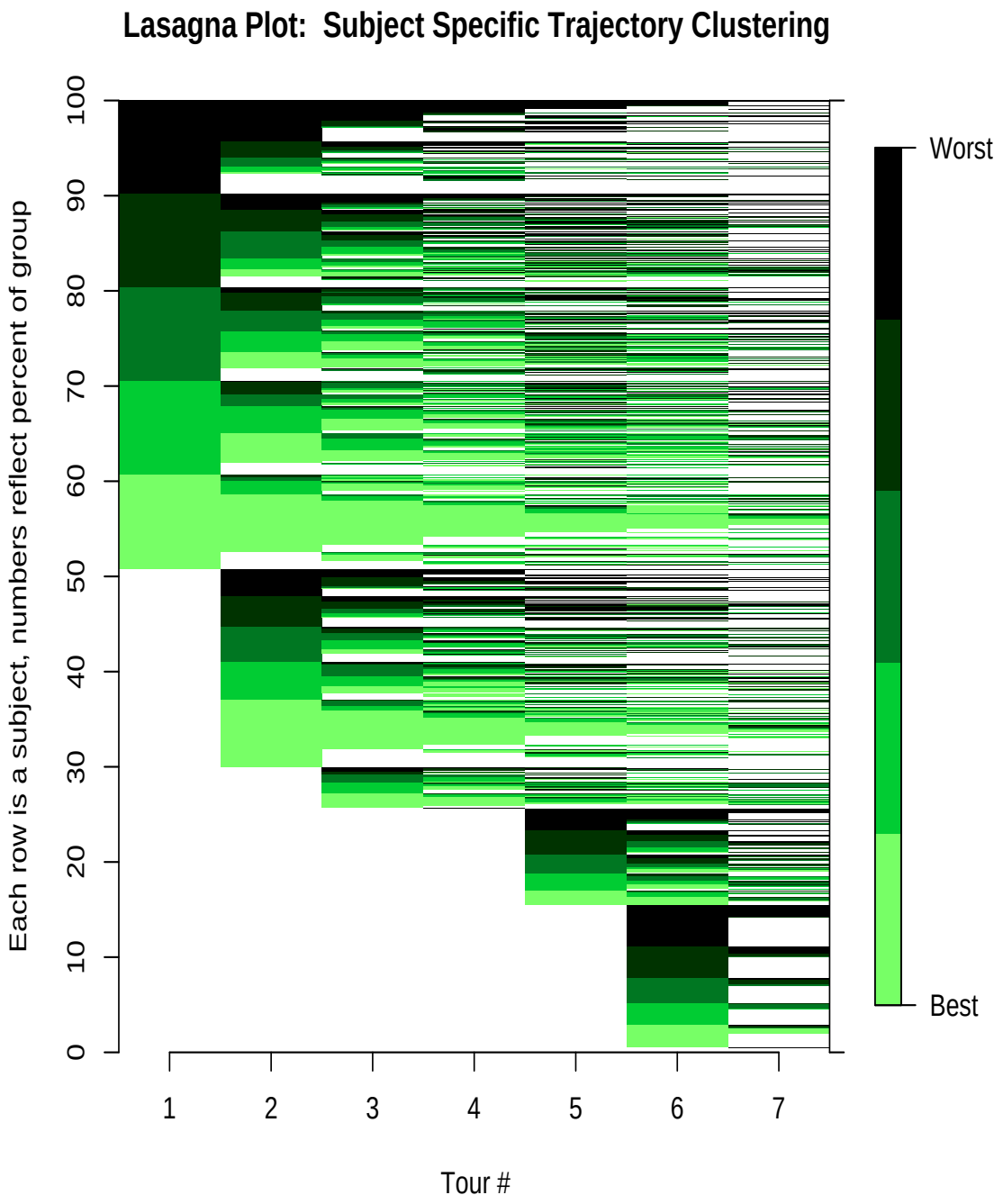

Figure 18: This is a cluster sort of Figure 17. Similar trajectories are grouped together and subject-level information is maintained and the association of cognitive ability by the metric of baseline quintiles over time can be explored. Comparing those enrolled in Tour 1 of the worst and best quintile, we see that there is more dropout for those starting out in the worst quintile, possibly indicating informative censoring. This pattern of drop out being related to first visit quintile rank holds for the later tours as well. Training effects can be seen when a subject's color lightens tracking that individual across time. For instance, a fair portion of individuals in Tour 1 were in the 2nd best quintile and advance to the best quintile in their second visit in Tour 2.

Table 1: Error rate of the confidence interval - simulation

	σ_1^2/σ_2^2						
n_1/n_2	1/8	1/4	1/2	1	2	4	8
1/2	0.012	0.022	0.032	0.056	0.085	0.114	0.148
1	0.059	0.055	0.050	0.050	0.058	0.056	0.058
2	0.149	0.115	0.086	0.053	0.033	0.021	0.014
4	0.252	0.193	0.119	0.058	0.024	0.005	0.003
8	0.364	0.258	0.150	0.056	0.017	0.003	0.000

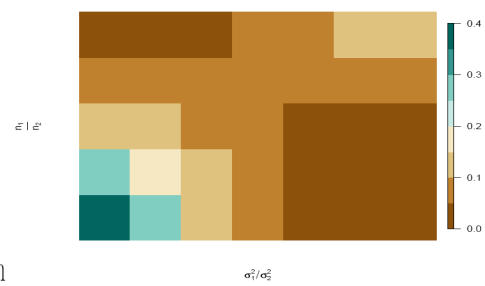

Table 2: Error rate of the confidence interval - normal approximation

	σ_1^2/σ_2^2						
n_1/n_2	1/8	1/4	1/2	1	2	4	8
1/2	0.011	0.016	0.028	0.050	0.080	0.11	0.133
1	0.050	0.050	0.050	0.050	0.050	0.050	0.050
2	0.133	0.110	0.080	0.050	0.028	0.016	0.011
4	0.237	0.179	0.110	0.050	0.016	0.004	0.001
8	0.331	0.237	0.133	0.050	0.011	0.001	0.000

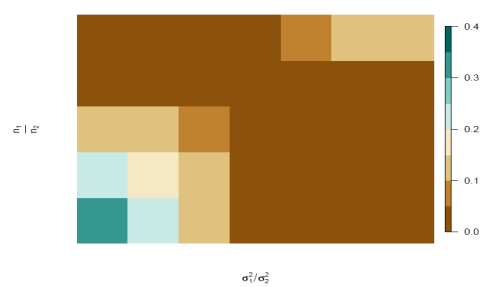

Figure 19: Simulation tables often have outcomes as result of different permutations of parameters. Making a lasagna plot of such a table gives a heatmapthat might convey trends more clearly than the numbers themselves.

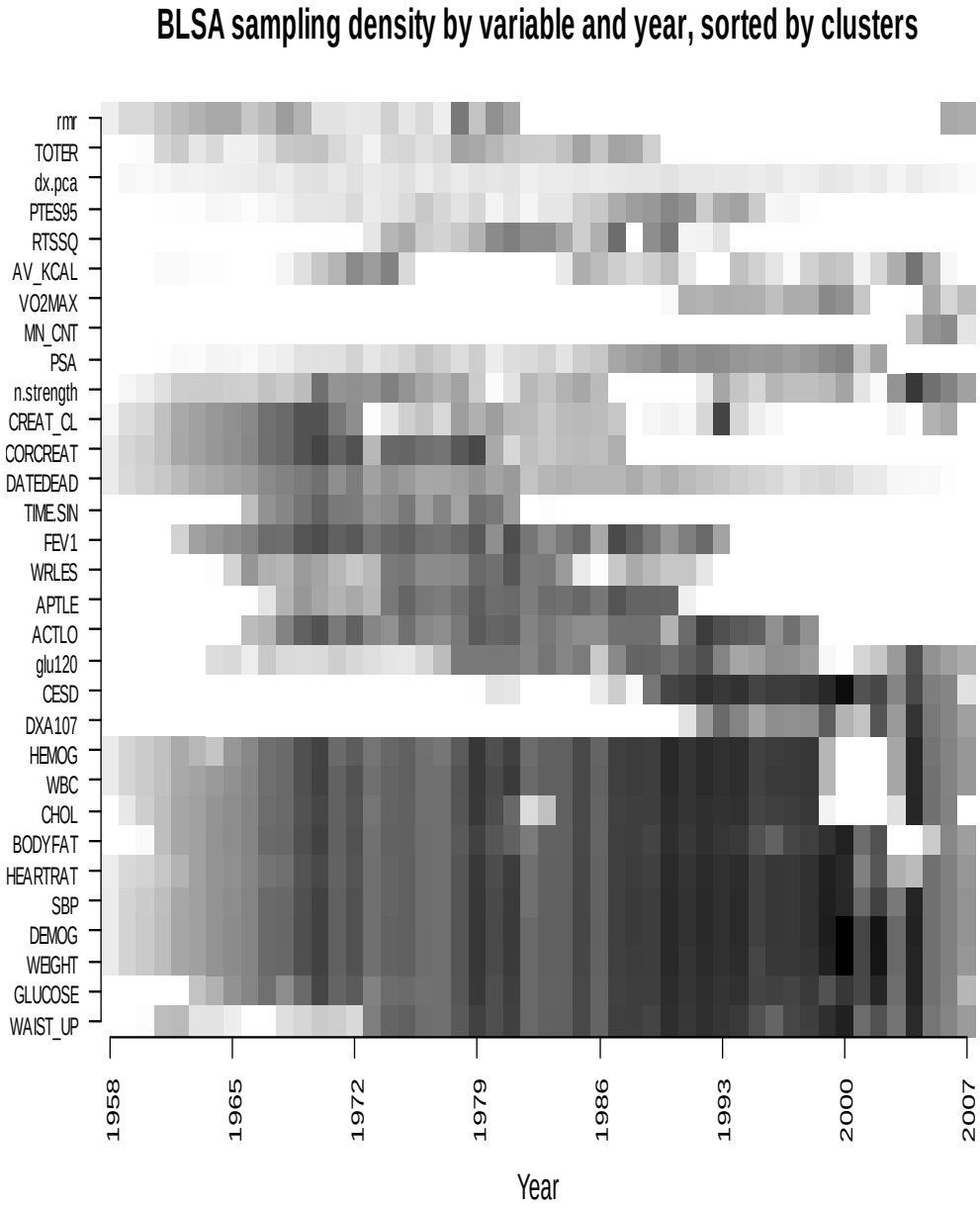

Figure 20: A plot showing the presence of recorded measurements for subjects in an epidemiologic study. The darker the cell, the more subjects that have a non-missing value for that covariate at that time point. This lasagna plot is cluster sorted for similar trajectories and could be useful in model building in trying to limit the inclusion of covariates with a lot of missing data in an effort to maximize number of subjects included in the model. Here, analyzing across all years for the covariates between "HEMOG" and "WAIST_UP" between years 1990 and 1994 would maximize the proportion of subjects used in the model because missing data is minimized.