Research Paper

PAIN 166 (2025) 1859–1870

PAIN

OPEN

Are people with chronic pain more diverse than we think? An investigation of ergodicity

Felicia T. A. Sundström*, Amani Lavefjord, Monica Buhrman, Lance M. McCracken

Abstract

This study investigates whether data from people with endometriosis (n = 58) and fibromyalgia (n = 58) exhibit what is called "ergodicity," meaning that results from analyses of aggregated group data can be used to support conclusions about the individuals within the groups. The variables studied here are commonly investigated in chronic pain: pain intensity, pain interference, depressive symptoms, psychological flexibility, and pain catastrophizing. Data were collected twice daily for 42 days from each participant and analyzed in 2 ways: as separate cross-sectional group studies using the timepoints as the separate data sets (between-person) and as individual longitudinal studies using each person's time series data (within person). To confirm ergodicity, the results from the 2 analyses should agree. However, this is not what was observed in several respects. The between-person data showed substantially less variability compared with within-person data. This was evident in both the summary statistics involving single variables and in the correlational analyses. Overall, between-person correlations were relatively restricted in range, while within-person correlations varied widely. These findings have potentially profound implications for the field of chronic pain research. Because ergodicity was not found, this raises doubts around the assumption that aggregated data collected from groups can accurately represent the range of individual experiences in chronic pain. These results advocate for a shift toward inclusion of more individual person-focused approaches as an addition to group-based approaches. This shift could lead to more personalized and effective treatments by better capturing and then clarifying the heterogeneous nature of chronic pain, including the processes that underlie it.

Keywords: Chronic pain, Endometriosis, Fibromyalgia, Group-to-individual generalizability, Individual variability, Ergodicity

1. Introduction

Psychological treatments for chronic pain can increase well-being and reduce disability.[3,10,12,19,42] Yet, some individuals benefit greatly from treatments, while others do not. Perhaps this happens because we currently do not deliver treatments that are personalized.[6] Although personalized treatments are recommended,[25] delivering these requires further research.[18] It seems that the knowledge and methods needed to deliver personalized treatment for chronic pain will require an improved understanding of processes or mechanisms of change for individual people.[17,21,24,25,28]

The logic of group research designs is familiar, with the emphasis on large aggregated samples of people and few observations for each, and a focus on variance between people. It is less familiar to consider individually focused designs, using a large number of observations and a small number of people, focusing on variance within people.

A typically unstated assumption underlying group-based research is that findings derived from aggregated group data are generalizable to the individual only if the examined processes are *ergodic*.[8,30,31] For this to hold, 2 conditions are necessary: (1) Observations of events in the overall sample must align consistently with observations obtained for each subset or individual within the group and (2) these observations of events should remain stable over time. If these conditions are met, statistical values such as means, standard deviations, and correlations based on measures of these events, for any and all individuals, should match those of the overall sample.[31]

Numerous studies now demonstrate a lack of ergodicity in psychological data.[1,2,4,7,9,23,30,34] Illustrations of this failure to demonstrate ergodicity, based on actual research data, are shown in **Figure 1**. In previous research designed specifically to examine ergodicity, Fisher et al.[7] examined data from 5 separate studies, including depressed mood, worry, fear, and positive and negative affect. In each data set, variation within participants was compared with the variation across participants, using as many within-person data points as the number of participants included. The results indicated little group-to-individual consistency.[7] We note that inconsistency is not simply an error or unreliability of measures—these differences are substantial. This clearly indicates that the results from data derived from group-based research are not automatically generalizable to the particular people to whom they might be applied[7,11]. If data from people with chronic pain fail to demonstrate ergodicity, group-based methods may need to be supplemented with other approaches to improve our understanding of chronic pain and to produce

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

Department of Psychology, Uppsala University, Uppsala, Sweden

**Corresponding author. Address: Department of Psychology, Uppsala University, Box 1225, 751 42 Uppsala, Sweden. E-mail address: felicia.sundstrom@psyk.uu.se (F. T. A. Sundström).*

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.painjournalonline.com).

Copyright © 2025 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the International Association for the Study of Pain. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

http://dx.doi.org/10.1097/j.pain.0000000000003573

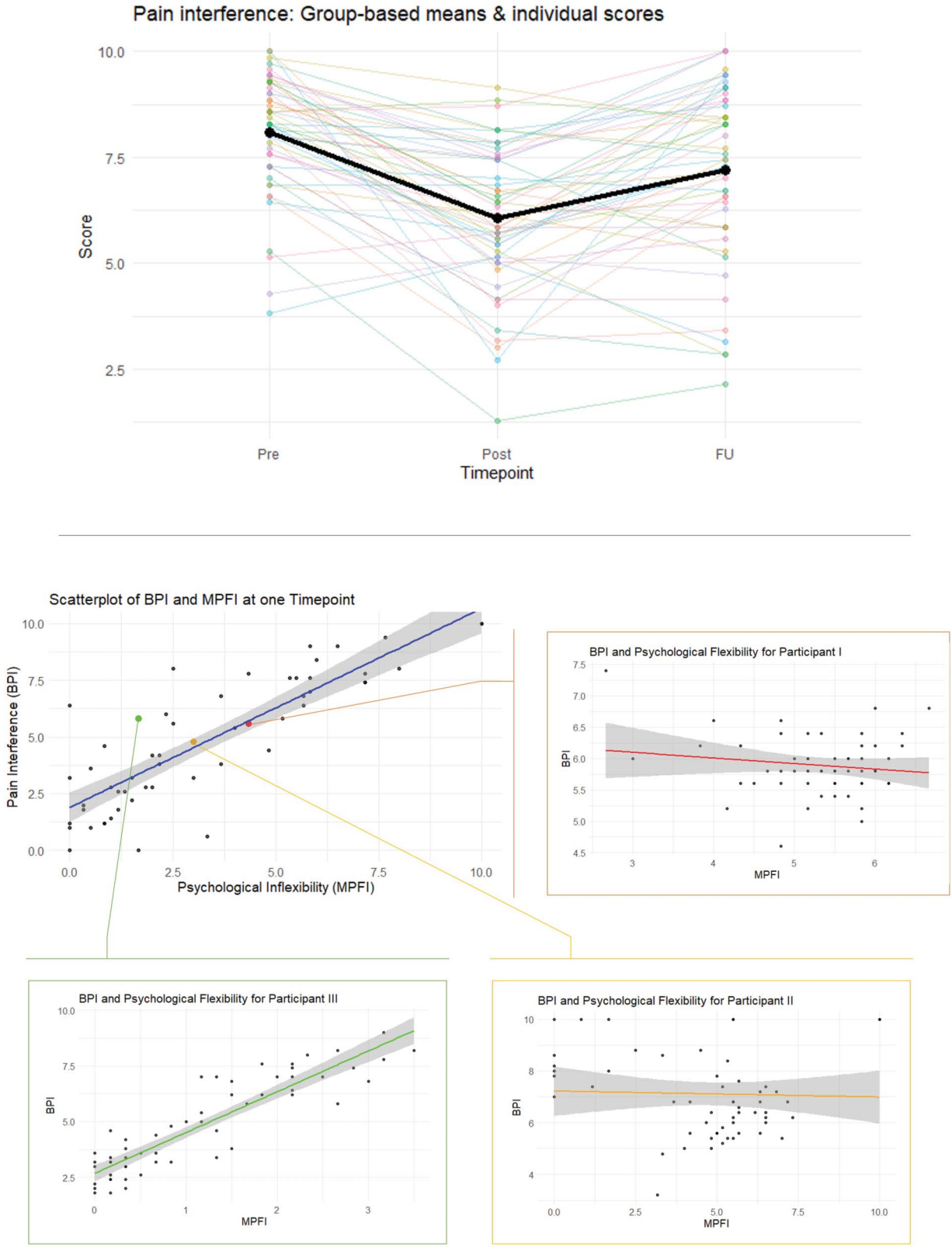

Figure 1. The top panel of Figure 1 shows pre-, post-, and follow-up results for pain interference in a group of participants from a treatment study. The black line represents the group mean at each timepoint, while the colored lines show individual trajectories, highlighting the variability among participants within the same sample. The bottom panel illustrates the relationship between pain interference and psychological inflexibility: the top-left plot shows the cross-sectional relationship between these variables at a single timepoint, while the time series plots for 3 participants demonstrate how an observed correlation at one timepoint may not consistently reflect individual patterns over time.

knowledge that is both more precise and more generally applicable. First, however, we need to determine whether this is the case.

The purpose of this study is to test the assumption of ergodicity in data from people with 2 chronic pain conditions: endometriosis and fibromyalgia. This will be conducted by examining commonly studied pain variables, including pain intensity, pain interference, depression, psychological flexibility, and pain catastrophizing. We compare results obtained from 2 separate perspectives: the conventional group data approach vs an approach that examines data intensively over time for individual people. If the assumption of ergodicity is supported, the results from these 2 perspectives should match.

2. Methods

2.1. Participants and procedure

Participants eligible for the study were adults aged 18 or older suffering from endometriosis or fibromyalgia with persistent or recurrent pain for 3 months or more. Fluency in written Swedish and access to an internet-enabled phone were required. Participants were also asked to agree to be contacted during the study to improve their response rate and to follow-up on their responses if needed.

After providing informed consent, self-report measures were administered through REDCap, a secure online survey platform.[13,14] Participants were then contacted with information about how to download and use m-path,[29] a phone application for daily measures to be administered. After downloading the application, participants received a notification on their phone when it was time to complete the daily measures. Participants were notified at the start of a 180-minute response window for each of the 2 daily assessment occasions. A reminder was sent within 30 minutes if the participant did not fill out the measure. The daily measures were administered twice daily for 42 consecutive days with default times set to 10 AM and 9 PM, although these could be adjusted upon individual request.

In addition to the daily measures, participants received a short weekly measure through m-path. At the end of the 42 days, participants were invited through e-mail to once again respond to the full-length standardized measures through REDCap. The responses from these weekly measures and the full-length standardized measures are not the focus of this study and will be presented elsewhere. Ethical approval from the Swedish Ethical Review Authority was received before data collection (DNR: 2022-02032-01).

Participants provided informed consent through REDCap after being recruited through Facebook posts in groups focused on chronic pain, Instagram accounts focused on endometriosis, patient organizations, and through information given to people at health care clinics working with endometriosis and fibromyalgia. In total, 198 participants with endometriosis and 228 participants with fibromyalgia began providing data. The number of participants who continued the study through m-path, regardless of pain condition or rates of missing data, was 249. Because the proportion of participants dropping out when moving from REDCap to m-path was large, the 2 groups were analyzed for potential differences in key variables (age, years of pain, pain condition, pain intensity last week, pain interference, depressive symptoms, psychological inflexibility, and pain catastrophizing). No differences were found.

According to a predetermined data collection and analysis plan (ClinicalTrials registration: NCT05518630), an equal number of participants and completed daily surveys were needed to investigate whether the data fulfilled the criteria for ergodicity. Thus, for each of the 2 samples, 58 participants who filled out at least 58 daily surveys (70% response rate) were randomly selected for the analyses described below after removing 6 participants with insufficient variability in their data. Then, for each participant, 58 completed surveys were randomly selected after removing timepoints with missing data.

2.2. Measures

2.2.1. Demographics and pain characteristics

Participants were asked to provide data on age, sex, relationship status, family origin, work status, whether they self-identify as a member of a minority group, education level, and financial situation. They were also asked to indicate their primary pain condition, whether a medical doctor had provided this diagnosis, their pain sites, pain duration, number of health care visits during the past 6 months, additional pain or health conditions, psychological and medical treatments, and which recruitment route led them to enter the study.

2.2.2. Daily measures

At each twice-daily assessment, 18 items were administered. These items were selected from or based on standardized and validated questionnaire items, including from the Brief Pain Inventory (BPI),[5] Patient Health Questionnaire-9 (PHQ-9),[22] Multidimensional Psychological Flexibility Inventory (MPFI),[33,37] and Pain Catastrophizing Scale (PCS).[35] All items, except for current pain intensity, were focused on a specific time frame, the time elapsed since the previous measurement prompt (approximately a half day). The pain item focused on the current moment. All items were rated on an 11-point scale, with 0 corresponding to not being in pain, not experiencing pain interference, or not agreeing with the statement in an item, while 10 corresponded to being in the worst possible pain, pain interfering completely, or agreeing very much with the statement in an item.

The first item in the daily measure was taken from the BPI, the item assessing current pain intensity. The next several items, also selected from the BPI, included pain interference with general activities, mood, relationships with other people, and ability to enjoy life, again since the time of the past measurement prompt. The BPI also includes an item on how pain interfered with sleep. In the evening, this item was rephrased as how the pain interfered with feeling well-rested. An additional item, created by the research team, was about how the pain interfered with one's ability to enjoy sex. The rest of the daily assessments consisted of 2 items from the PHQ-9[22] assessing a lack of interest in doing things and feeling down, depressed, and hopeless; 6 items assessing psychological inflexibility from the MPFI[33,37]; and 3 items from the PCS[35] on pain catastrophizing. The full set of items is provided in the supplementary materials, http://links.lww.com/PAIN/C242.

2.3. Statistical analyses

The univariate statistics and plots were calculated using the R packages "dplyr"[41] and "ggplot2,"[40] and frequency distributions were subsequently constructed. Univariate within-person plots were produced by calculating the mean and standard deviation of each participant's time series. Univariate between-person plots were calculated by calculating the mean and standard deviation

using each timepoint (T = 1 to T = 58) as a complete between-person group data set of 58 participants each. This resulted in distributions of means and standard deviations of 58 different group data sets, with 58 participants each for each of the 2 pain samples. The same was performed for the 58 individual within-person data sets. Summary statistics of the within-person and between-person means and standard deviations were then calculated. For the within-person summary statistics, the mean of the individual means across days was calculated as well as the standard deviation. For the between-person statistics, the mean was calculated by deriving the mean of the between-people data sets' means. The standard deviation was derived by calculating the standard deviation of the means.

Pearson's correlations were calculated using both raw data and data where control over temporal dependency was applied. For between-person correlations, the correlations were calculated using each timepoint as a separate data set across 58 participants, resulting in 58 sets of correlation coefficients. For within-individual correlations, the correlations were calculated using the data for each person across 58 occasions in time for each of the variables in each individual. This also resulted in 58 sets of correlation coefficients per pain sample. The correlation coefficients were then plotted using "ggplot2."[40] In addition, the mean, standard deviation, and range were calculated for both the within- and between-person results. The results from correlations using raw data are available in the supplementary materials, http://links.lww.com/PAIN/C242.

Two approaches for correcting for temporal dependency were tested. First, Autoregressive Integrated Moving Average models were applied using the R-package "forecast,"[20] which are provided in the supplementary materials, http://links.lww.com/PAIN/C242. Second, a method similar to Fisher et al.[7] was used in which both autoregressive and cross-lagged effects are accounted for in the within-person data. This means that for each individual, the first variable X (eg, pain) was modeled as a function of both its own previous value (autoregressive effect) and the previous value of the second variable Y (eg, pain interference) (cross-lagged effect). Similarly, Y was modeled as a function of its own prior value and the prior value of X. Next, the residuals were retained and tested using a Durbin-Watson test to check whether a single lag was sufficient to correct for the autocorrelation for all participants. If not, further lag was added for that specific participant. Finally, the relationship between variables X and Y were analyzed using Pearson's correlations for every individual.

For most participants, controlling for 1 lag was sufficient to result in a nonsignificant Durbin-Watson test, indicating temporal dependency was successfully whitened. Controlling for an additional lag was needed for 6 participants in the endometriosis sample and 4 in the fibromyalgia sample across the 4 reported bivariate pairs.

To provide examples of individual trajectories and trends in the data set, 2 sets of pairs with similar mean were selected for individual plotting of the individual time series. The mean pain interference score was calculated using the "dplyr"[41] package. Then, a threshold for similarity between the 2 means was set before an iteration was created to find similar means respecting the threshold. The 4 participants were then selected and plotted using "ggplot2."

3. Results: endometriosis sample

3.1. Participant characteristics

The mean age for the 58 participants with endometriosis was 33.31 years (SD = 8.43). A majority identified as female (98.28%), 33 participants (56.90%) held full-time employment, the mean number of health care visits due to the condition in the past 6 months was 6.74 visits (SD = 14.13), and 21 participants (36.21%) described suffering from generalized pain. See **Table 1** for further details.

3.2. Univariate plots and statistics

On examining the univariate plots that depict both between-person and within-person distributions of mean values (**Fig. 2**), a consistent pattern is observed across all variables. Compared with the aggregated group data where variability is generally quite small, there is substantially greater variability in the individuals' within-person data. This is evidenced in both the point-by-point frequency counts of means and the density lines. Overall, in the individual-based within-person data, all variables have means extending across the scale, ranging from around zero to approximately 7 or 9 in pain intensity, pain interference, and psychological inflexibility, but not in group-based between-person data. This is despite the distributions consisting of the exact same datapoints.

Looking at the between-person data regarding depressive symptoms, as shown in **Figure 2**, the between-person data cluster around a mean of 5 to 7. A remarkably similar pattern is shown in the between-person data for all variables featured in **Figure 2**. Furthermore, for the within-person distributions, there are rarely more than 7 people who achieve the same value regardless of variable analyzed, and the distributions are generally quite flat. However, the data for catastrophizing and depressive symptoms could be seen as deviations from the pattern because of the within-person data featuring relatively large proportions of people who obtaining fairly low values, which is itself interesting.

In analyzing the range of means and SDs summarized in **Table 2**, it is evident that even if the means are identical across all variables because of the same data points being used for all calculations, both the SDs and the range of SDs are substantially smaller in the between-person data.

In the within-person data, most variables have a relatively low average mean, but there is also a notable variability observed in both the within-person means and the SDs as compared with the between-person data. This indicates differing degrees of variability over time among individuals. Specifically, the data show that while some individuals exhibit significant variability in their time series, others demonstrate more stability.

3.3. Distributions of correlations

After controlling for temporal dependency in the within-person time series data, the histograms depicting the magnitude of correlations between variables show only a small range in the between-person data, but a wide range in the within-person correlations (**Fig. 3**), similar to the distributions featuring single variables (**Fig. 2**). Overall, all pairs examined resulted in smaller ranges of correlations for the between-person calculations, while within-person calculations resulted in large ranges, in some instances ranging from small negative correlations for some participants to large positive correlations for others (**Table 3**). For example, the between-person correlations between pain intensity and depressive symptoms only ranged from r = 0.33 to r = 0.77. By contrast, the within-person correlations ranged from a small negative correlation (r = −0.05) in one individual to a large positive correlation (r = 0.87) in another. In addition, there is a pattern of very dissimilar mean correlations when comparing between-person data and within-person data; when data are analyzed as between-person data sets, mean correlations are greater by 25.37% in psychological inflexibility and depressive symptoms,

Table 1

Participant characteristics.

	Mean (SD) or n (%)	
	Endometriosis	**Fibromyalgia**
Age	33.31 (8.43)	50.93 (9.95)
Gender (women)	57 (98.28%)	56 (98.25%)
Identify as part of a minority group	7 (12.07%)	49 (84.48%)
Education (highest completed)		
Elementary	2 (3.45%)	3 (5.17%)
Secondary/higher vocational education	27 (46.55%)	28 (48.28%)
College/university	29 (50.00%)	35 (60.34%)
Work status		
Employed (full-time)	33 (56.90%)	11 (18.97%)
Employed (part-time)	8 (13.79%)	11 (18.97%)
Self-employed	2 (3.45%)	1 (1.72%)
Sick leave	7 (12.07%)	20 (34.48%)
Other (retired, student, etc.)	8 (13.79%)	15 (25.86%)
Financial situation		
Very good	9 (15.52%)	3 (5.17%)
Good	18 (31.03%)	22 (37.93%)
Sufficient	28 (48.28%)	23 (39.66%)
Bad	3 (5.17%)	7 (12.07%)
Very bad	0 (0.00%)	3 (5.17%)
Relationship status (married/in a relationship)	38 (65.52%)	41 (70.69%)
Diagnosed pain condition	56 (96.55%)	57 (1.72%)
Pain duration (>2 y)	54 (93.10%)	58 (100.00%)
Generalized pain	21 (36.21%)	54 (93.10%)
Pain sites (most frequent)		
Abdomen	51 (87.93%)	16 (27.59%)
Pelvic region	41 (70.69%)	27 (46.55%)
Back region	32 (55.17%)	51 (87.93%)
Neck	4 (6.90%)	46 (79.31%)
Arms or shoulders	3 (5.17%)	47 (81.03%)
Prescribed opioids	32 (55.17%)	9 (15.52%)
Health care visits due to pain (past 6 mo)	6.74 (14.13)	2.14 (3.89)
Recruitment site		
Social media	50 (86.21%)	50 (86.21%)
Hospital or health care clinic	4 (6.90%)	0 (0.00%)
Other	3 (5.17%)	6 (10.34%)
Patient organizations	1 (1.72%)	2 (3.45%)

Pain sites were not mutually exclusive.

28.33% in pain intensity and pain interference, and 51.35% in pain and depressive symptoms, compared with when analyzed as within-person data. However, correlations including pain catastrophizing as a variable (**Fig. 3** and supplementary materials, http://links.lww.com/PAIN/C242) show another pattern, with somewhat more similar means for between- and within-person correlations compared with other variable pairs, leading to an increase of 7.81% when analyzed as between-person data. Nevertheless, the range of correlations in both the between- and within-person data is similar to the other bivariate combinations. Furthermore, although the mean correlation seems similar for the between-person and within-person distributions, the SDs are highly dissimilar (**Table 3**).

3.4. Individual time series

Of course, any calculation of an individual's mean is a form of aggregation, even when accompanied by SDs, and this will not fully capture the varying nature of some participants' data. This becomes obvious when finding 2 participants with both similar means and SDs of their pain interference ratings over time, Participant 1 with a mean of 6.83 (SD = 1.55) and Participant 2 with a mean of 6.82 (SD = 1.47).

Figure 4 illustrates the 2 time series plotted in the same graph for both participants and 2 additional plots separating the 2 participants and including the time series of pain intensity and psychological inflexibility.

Here, one sees how much information is lost when relying on means and SDs, despite similarity across participants from a summary point of view. Participant 1 has a seemingly more extreme values at both ends of the spectrum while the pattern of Participant 2 over the time series seem similar for all variables, namely a gradual increase. However, consistently demonstrated for both participants is a lack of stability in all variables and, perhaps even more concerning, inconsistent patterns within the series data, meaning that the relationship between variables

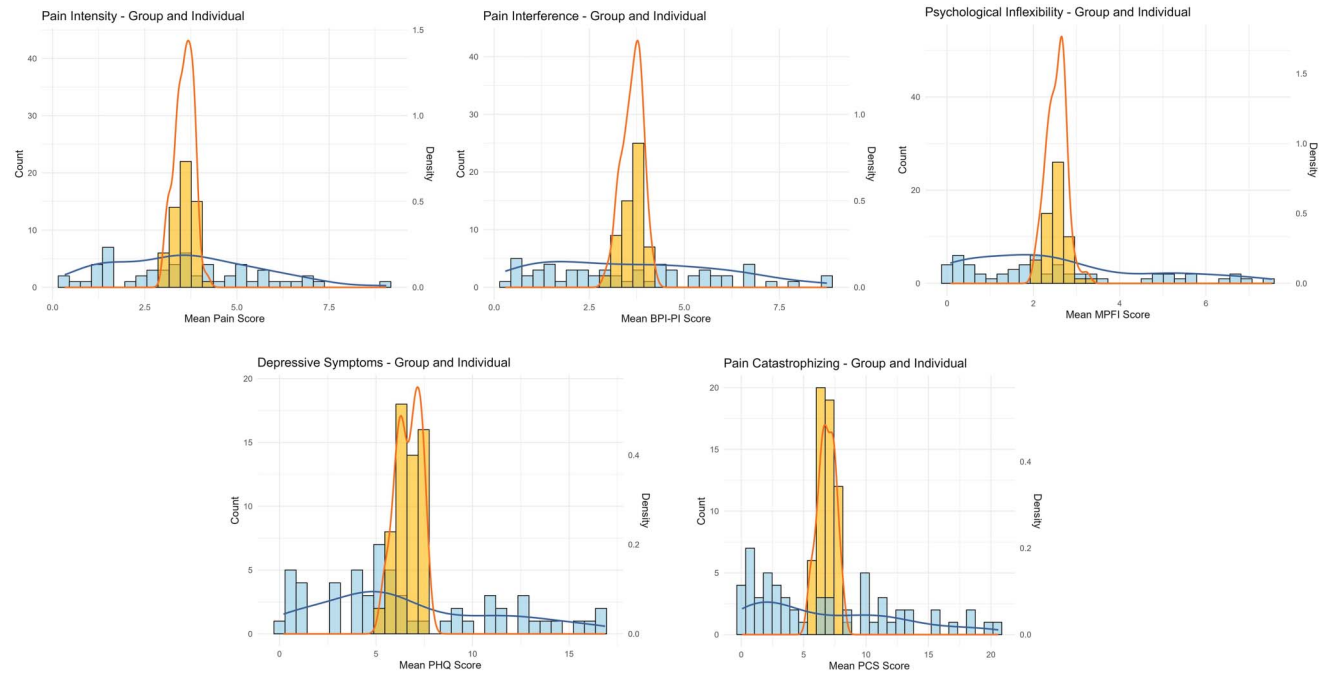

Figure 2. Endometriosis: Depicting 5 histograms and density plots of 5 variables for individuals in blue (means of 58 individual time series) and for groups in orange (means of 58 cross-sectional data sets with 58 participants).

seems to change over time. We note that it occurs in the absence of intervention on our part.

4. Results: fibromyalgia sample

4.1. Participant characteristics

The mean age for the 58 participants with fibromyalgia was 50.93 years (SD = 9.95). A majority identified as female (98.25%), 11 participants (18.97%) held full-time employment, the mean number of health care visits because of the condition in the past 6 months was 2.14 visits (SD = 3.89), and 54 participants (93.10%) described suffering from generalized pain. See **Table 1** for further details.

4.2. Univariate plots and statistics

Similar to the patterns in the endometriosis sample, histograms featuring single variables in **Figure 5** show substantially less variability when analyzed as between-person data compared with within-person data. This is shown in both the point-by-point frequency counts of means and the density lines.

The between-person means cluster around higher scores in all variables (**Table 4**), a different pattern compared with the results from people with endometriosis. For example, for depressive symptoms, the data are clustered around a mean of 7 to 9. This is similar to pain interference, which also differs between the 2 samples. However, the within-person distributions for fibromyalgia are similar to the endometriosis data in most variables. If anything, they are flatter with less clustering (**Fig. 5**).

Comparing the between-vs within-person analysis methods, it is evident that the relatively narrow variability seen in the between-person statistics is not mirrored in the within-person statistics. Quite the contrary, there is substantial inconsistency between the 2 forms of analysis despite the inclusion of the same data. Even if the means of the 2 analysis methods are the same across all variables, which again is to be expected due to the same data being used for both methods, the pooled standard deviations differ in both range and mean for both methods.

Table 4 summarizes the range of means and standard deviations for both analyses. Similar to the endometriosis sample, the within-person means have a large range in average score over time, but also that there is varying variability within each time series.

Table 2

Univariate within-person and between-person statistics—endometriosis.

Variable	Within-person			Range (SD)	Between-person			Range (SD)
	Mean	SD	Range (mean)		Mean	SD*	Range (mean)	
Pain	3.57	1.95	0.33-9.05	0.62-3.30	3.57	0.25	3.05-4.17	2.19-2.89
BPI	3.63	2.28	0.29-8.75	0.52-2.80	3.63	0.27	2.96-4.17	2.41-3.12
MPFI	2.54	2.07	0.07-7.53	0.21-2.75	2.54	0.23	2.06-3.24	2.07-3.07
PHQ	6.67	4.63	0.21-16.86	0.77-6.11	6.67	0.63	5.36-7.74	5.00-6.93
PCS	6.84	5.83	0.07-20.59	0.28-9.26	6.84	0.69	5.41-8.16	5.79-8.80

* The SD of the group-based mean.

BPI, Brief Pain Inventory; MPFI, Multidimensional Psychological Flexibility Inventory; PCS, Pain Catastrophizing Scale; PHQ, Patient Health Questionnaire.

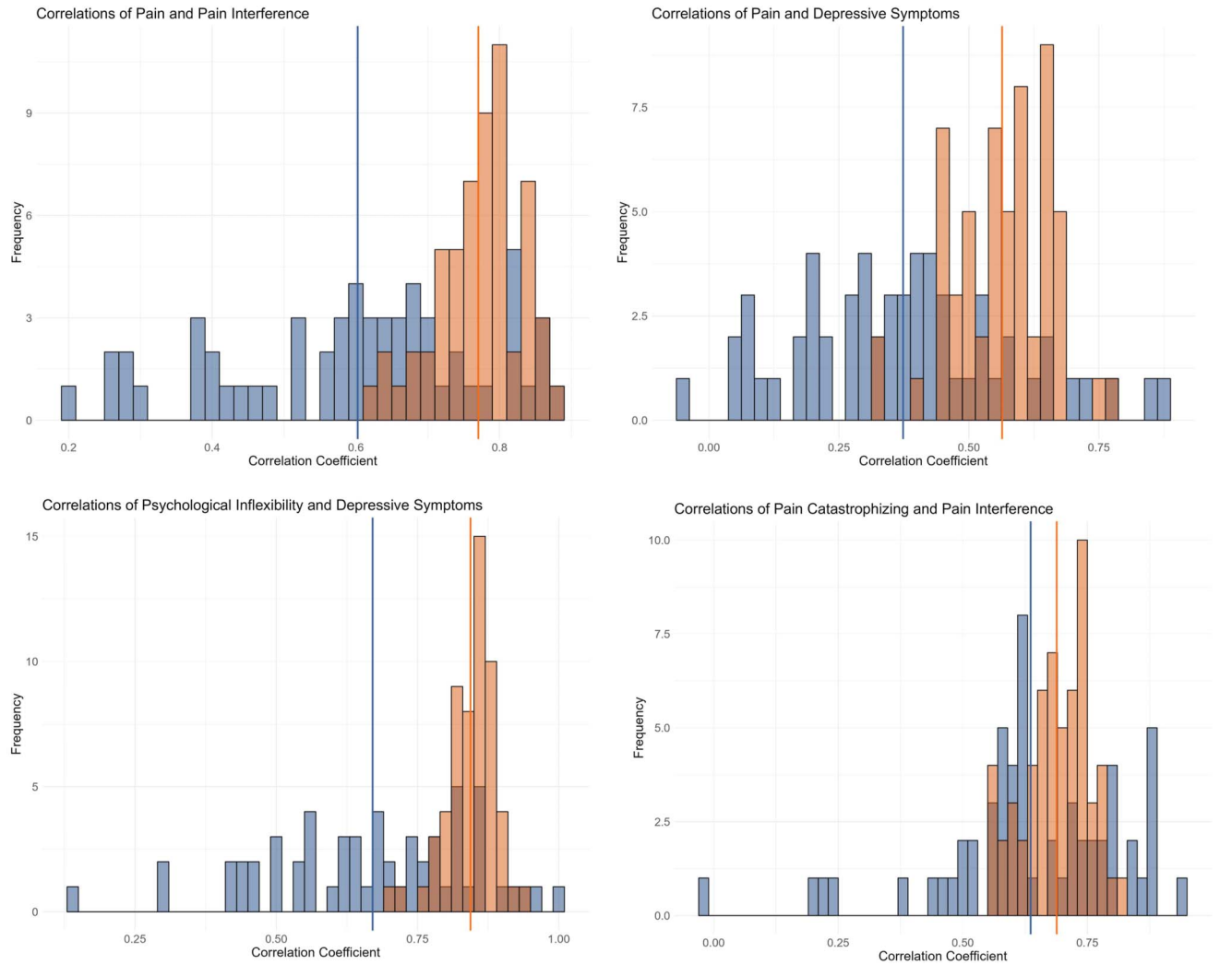

Figure 3. Endometriosis: Distributions of correlations between 6 variable pairs using within-person correlations (blue) corrected for temporal dependency and between-person correlations (orange). The solid lines mark the mean correlational coefficient for between- and within-person correlations, respectively.

As summarized in **Table 4**, some participants have a low standard deviation meaning the scores varied little over time for that specific participant, while others have a relatively large SD showing a less stable day-to-day pattern within the individual over time.

4.3. Distributions of correlations

After controlling for temporal dependency in the within-person data, correlations were computed for the variable pairs, depicted in **Figure 6**, revealing a similar pattern to that of endometriosis. In all variable pairs, the between-person calculations not only resulted in a larger mean correlation but also a smaller range of actual correlations compared with the within-person calculations. For example, correlations between pain catastrophizing and pain interference (**Table 5**) ranged from a small negative correlation ($r = -0.10$) in one participant to a large positive correlation ($r = 0.86$) in another. As another example, the between-person correlations between pain intensity and depression ranged only from $r = 0.49$ to $r = 0.78$. However, for the within-person analyses, these correlations ranged from $r = -0.15$ for one participant over time to $r = 0.84$ for another (**Table 5**).

Comparing the distributions of correlations, we can also see a different pattern as compared with the results from the endometriosis sample. Namely, all distributions show a similar range of difference between the mean magnitude of the correlations in the within-person and between-person analyses, ranging from 35.00% in psychological inflexibility and depressive symptoms, and 60.98% in pain intensity and depressive symptoms.

4.4. Individual time series

By finding 2 participants with similar means and SDs of pain interference of their individual time series, we can gain deeper

Table 3

Bivariate correlations between common process and outcome variables—endometriosis.

Variable	Within-person			Between-person		
	Mean	SD	Range (mean)	Mean	SD	Range (mean)
Pain—BPI	0.60	0.18	0.20 to 0.88	0.77	0.06	0.62-0.87
Pain—PHQ	0.37	0.21	−0.05 to 0.87	0.56	0.09	0.33-0.77
MPFI—PHQ	0.67	0.18	0.14 to 0.99	0.84	0.04	0.70-0.93
PCS—BPI	0.64	0.19	−0.02 to 0.95	0.69	0.06	0.56-0.81

Residuals from correction for temporal dependency are used for the within-person analyses.
BPI, Brief Pain Inventory; MPFI, Multidimensional Psychological Flexibility Inventory; PCS, Pain Catastrophizing Scale; PHQ, Patient Health Questionnaire.

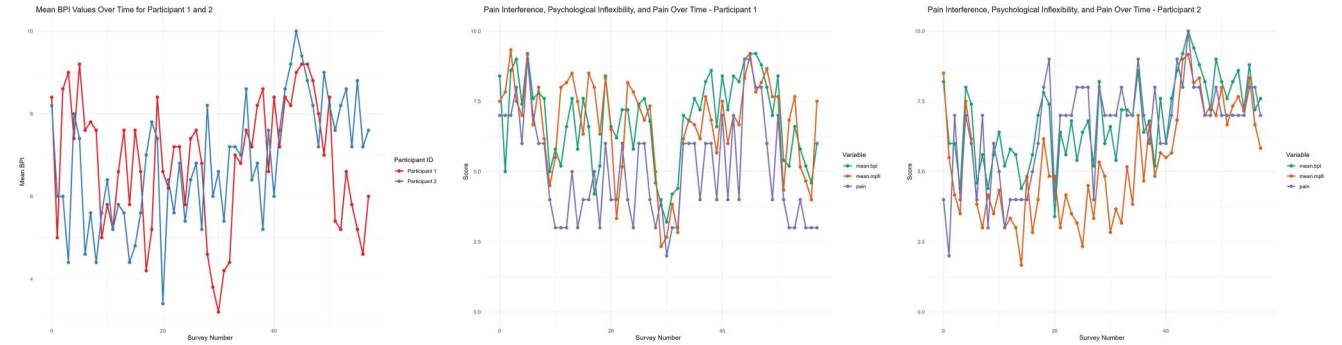

Figure 4. Endometriosis: Time series for 2 participants. The first graph shows the pain interference values of Participants 1 (red) and 2 (blue) plotted in the same graph. The second graph depicts the time series of 3 variables (pain interference, psychological inflexibility, and pain) in Participant 1. The third graph depicts the time series of 3 variables (pain interference, psychological inflexibility, and pain) in Participant 2.

insights into the importance of supplementing our typical focus on means, even if accompanied by SDs, and few points in time. In the first graph of **Figure 7**, 2 participants' time series are plotted: Participant 3 having a mean of 3.67 (SD = 2.03) and Participant 4 having an identical mean of 3.67 (SD = 0.96).

Compared with the endometriosis sample, 2 now familiar problems are again evident here. First, there is not much stability for any of the variables in either of the cases. Second, based on visual inspection, it seems that relationships between the variables such as pain interference, pain intensity, and psychological inflexibility change over time for Participant 3 despite the data being purely observational.

5. Discussion

This study compares the results obtained from a conventional group data approach focusing on variability between people vs an approach that focuses on individual people and how they vary within themselves over time. In both cases, the same commonly studied chronic pain variables were considered. What we observe is that these 2 perspectives do not provide symmetrical results. When examining individual data points over time (within-person), we see remarkable individual, person-to-person, variability to a degree that is not captured in the group data (between-person). These results, including summary statistical indicators of variability and bivariate association, all show essential inconsistency and indicate an absence of ergodicity.[8,30,31] These results were observed in people reporting endometriosis and replicated in people reporting fibromyalgia.

The means of between-person and within-person distributions from the individual variables are mathematically identical. This is because the raw data are the same and the mathematical calculations are simple, although the perspective applied differs.[31] At the same time, the variations in ranges and SDs reveal substantial differences when the perspectives are compared. When analyzing single variable frequency distributions, the within-person perspective reveals results that are far more variable than expected compared with the between-person perspective. Perhaps a more interesting and

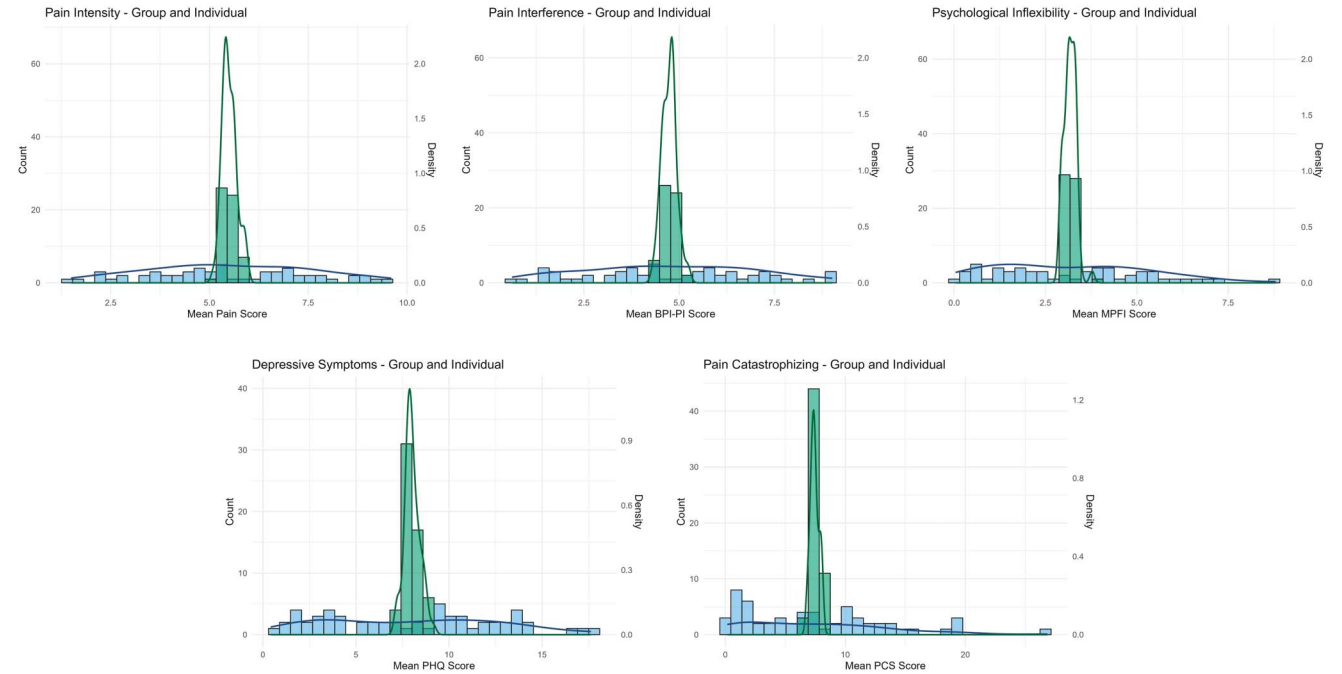

Figure 5. Fibromyalgia: Depicting 5 histograms and density plots of 5 variables for individuals in blue (means of 58 individual time series) and for groups in green (means of 58 cross-sectional data sets with 58 participants).

Table 4

Univariate within-person and between-person statistics—fibromyalgia.

Variable	Within-person			Range (SD)	Between-person			Range (SD)
	Mean	SD	Range (mean)		Mean	SD*	Range (mean)	
Pain	5.51	2.07	1.50-9.60	0.22-2.33	5.51	0.18	5.10-5.95	2.09-2.71
BPI	4.74	2.22	0.62-9.01	0.47-2.81	4.74	0.19	4.31-5.25	2.28-2.84
MPFI	3.17	2.11	0.04-8.80	0.15-2.82	3.17	0.16	2.90-3.78	2.15-2.73
PHQ	8.02	4.67	0.41-17.62	0.88-7.08	8.02	0.41	7.16-9.16	5.13-6.30
PCS	7.44	6.07	0.16-26.84	0.47-7.46	7.44	0.35	6.78-8.14	6.19-7.59

* The SD of the group-based mean.

BPI, Brief Pain Inventory; MPFI, Multidimensional Psychological Flexibility Inventory; PCS, Pain Catastrophizing Scale; PHQ, Patient Health Questionnaire.

unexpected finding is that within-person variability is itself variable over time as well. This is an important concept to understand. When people report pain interference, as an example, their data vary in mean levels between people, they can also vary between people in overall relative variability within any particular time frame, and people can vary within themselves in how variable this variability is, looking across multiple time frames. Variability is multidimensional, between people, within people, and in time. In addition, typical group-based research designs do not capture this variability, particularly the variability within people in time.

Looking further at the results, here we see additional sources of variability where we are perhaps not used to seeing them. In the analyses of relations between pairs of variables, very few distributions of correlations exhibit similar mean correlations when comparing between-person and within-

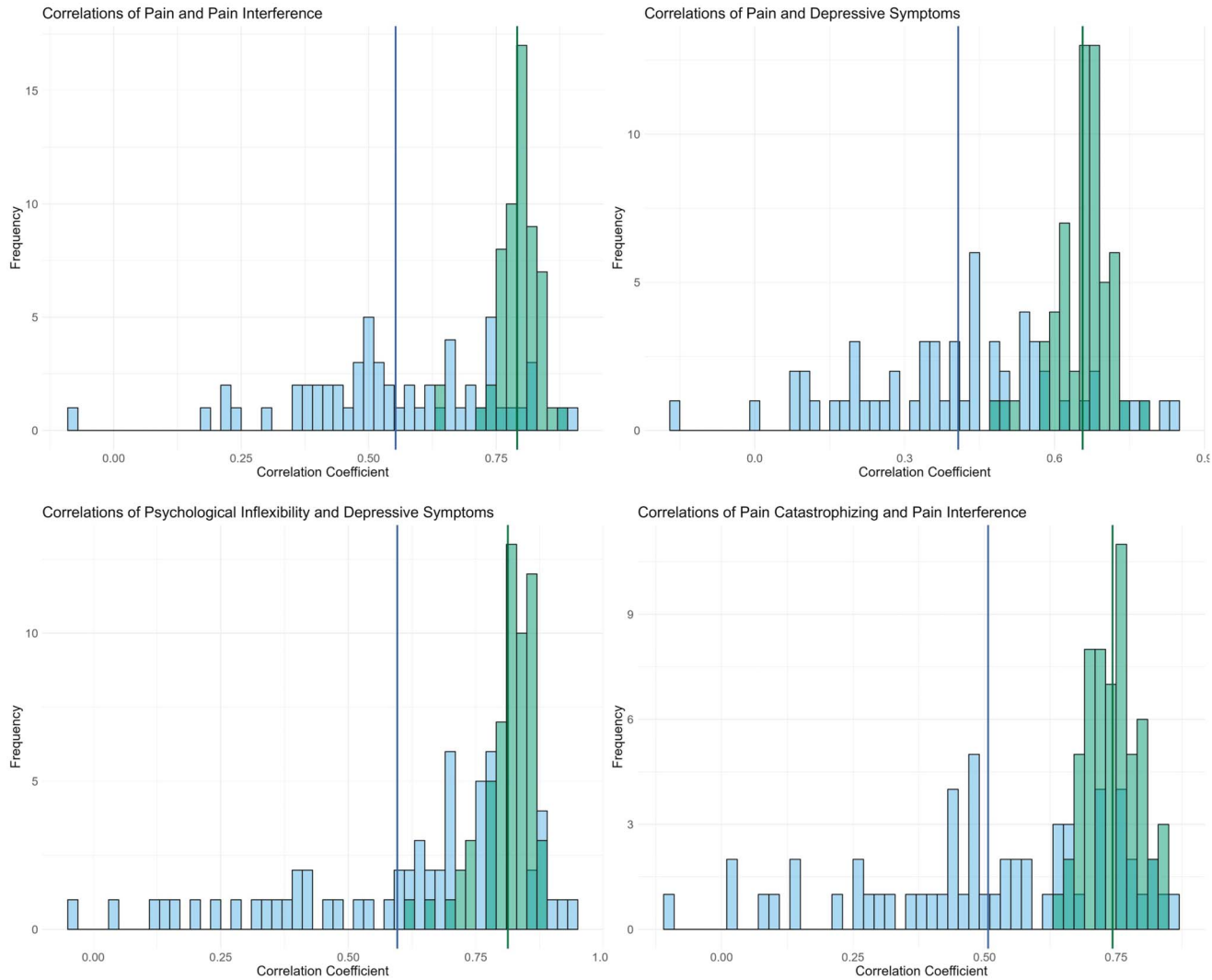

Figure 6. Fibromyalgia: Distributions of correlations between 6 variable pairs using within-person correlations (blue) corrected for temporal dependency and between-person correlations (green). The solid lines mark the mean correlational coefficient for between- and within-person correlations, respectively.

Table 5

Bivariate correlations between common process and outcome variables—fibromyalgia.

Variable	Within-person			Between-person		
	Mean	SD	Range (mean)	Mean	SD	Range (mean)
Pain—BPI	0.55	0.20	−0.09 to 0.91	0.79	0.04	0.64-0.88
Pain—PHQ	0.41	0.21	−0.15 to 0.84	0.66	0.05	0.49-0.78
MPFI—PHQ	0.60	0.24	−0.04 to 0.94	0.81	0.05	0.63-0.88
PCS—BPI	0.51	0.24	−0.10 to 0.86	0.74	0.05	0.64-0.84

Residuals from correction for temporal dependency are used for the within-person analyses.
BPI, Brief Pain Inventory; MPFI, Multidimensional Psychological Flexibility Inventory; PCS, Pain Catastrophizing Scale; PHQ, Patient Health Questionnaire.

person statistics. Similarly, the ranges observed in these correlations across various variable pairs vary, ranging from negligible or nonexistent to strong when analyzed as within-person data, a result not captured by the between-person correlations. This means that by relying only on group data, even when individual correlations are reported, we risk overlooking cases where some individuals show no relationship between variables, contrary to what our theories would lead us to expect. We might also miss relationships that emerge at different time lags, such as lag 2 or 3, if we do not examine these relations over time.

When comparing the 2 samples, participants with endometriosis and fibromyalgia, different patterns emerged here as well. Participants with fibromyalgia generally reported higher scores across measures and showed a narrower range of pain intensity compared with those with endometriosis. However, we might add that variability between people seems much large than between conditions, a result we see as consistent with our overall conclusions from this study.

The implications of the results are manifold. First and foremost, the findings from this study challenge a fundamental assumption of group-based methods used in pain research, namely that we can rely on aggregated data as a way to precisely understand the individuals contained within the aggregation. The results from this study show this may not be fully justified or may need to be done with caution. These same kinds of results have been demonstrated in other populations and other variables.[1,2,4,7,9,23,31,34]

The results from this study show that aggregating data may create problems *even* if one begins by capturing change within an individual over time. It seems to be common sense, once one thinks about it, that relations between variables and influences on those variables will be situation- and time-dependent. Although the data here are observational, with no particular interventions administered to any of the individuals, the absence of an intervention does not mean that circumstances do not intervene or that day-to-day life does not include influences on the data, which it certainly will. It is expected that for many people, pain, pain interference, and mood will change as events in life unfold. In a similar fashion, potential drivers of change in outcome processes, such as pain catastrophizing and psychological inflexibility, will also vary. Once again, this is likely to happen in a context-dependent, and therefore, time-dependent, fashion, for many individuals. These observed features of the series data clearly violate current assumptions required in the statistics we often use to analyze them (eg, linear regression), but also clinical and theoretical assumptions.[26,27,32,36,38,39] If data exhibit no functional relationship between what is seen as typical processes of change with cognitive behavioral therapy or acceptance and commitment therapy and common outcome variables, do we have the sensitivity to detect this and to orient treatment methods accordingly? Based on the results observed here, the finer-grained relations in the individual data deserve our attention and ought not to be buried in a process of summarizing groups. Instead of aggregating data and obscuring the individual patterns, we ought to first find the patterns and then summarize by pulling together the people who have similar patterns.[15]

Highlighting the value of individual data need not imply disinterest in group data. It does not make sense to entirely abandon group data. It is rather the case that an expansion in the methods used seems needed. A more discerning approach in selecting between between-person and within-person methods, and a greater inclusion of intensive, longitudinal, individually focused methods may enhance the discoveries we make and the knowledge we need to produce better applications. Although group-based methods may be suitable when the primary objective is to discuss groups or populations, individual-based methods may be more informative when the aim is to understand the process of change in treatment or inform specific individual treatment decisions. These specific decisions include determining the particular treatment methods or components needed to move a particular ultimate outcome, by moving a particular process of change, in a particular time and context, during an ongoing series of treatment encounters.[16]

The results here indicate that data from pain-related variables cannot be deemed ergodic. These results, however, cannot be regarded as the final answer, and the study is not without its

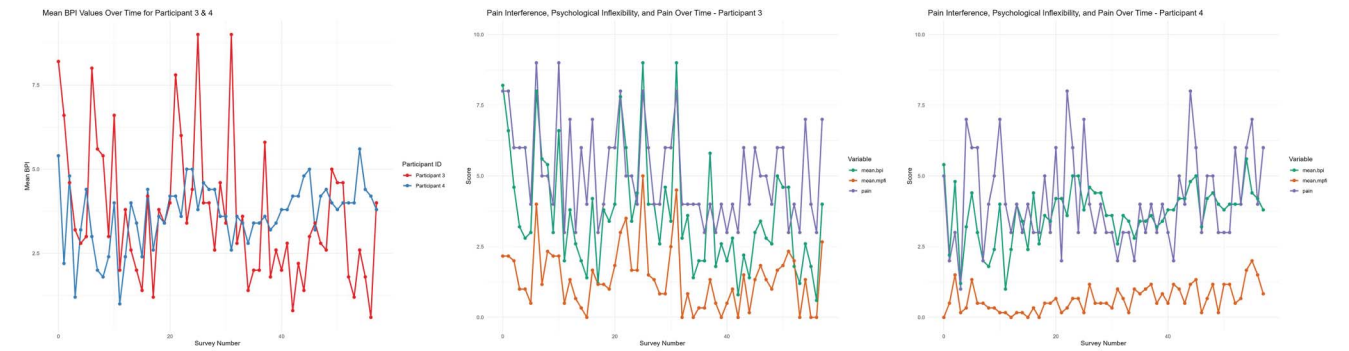

Figure 7. Fibromyalgia: Time series for 2 participants. The first graph shows the pain interference values of Participant 3 (red) and 4 (blue) plotted in the same graph. The second graph depicts the time series of 3 variables (pain interference, psychological inflexibility, and pain) in Participant 3. The third graph depicts the time series of 3 variables (pain interference, psychological inflexibility, and pain) in Participant 4.

limitations. First, the study only includes 2 types of pain conditions: endometriosis and fibromyalgia. Future studies should aim to replicate the results in other pain conditions. A focus on these conditions also led to a skewed sample in gender. In addition, the measures used are selected from many that could have been used, and it is thus fair to ask whether proof for this theorem would be found if they had been different. Finally, this study focused solely on ergodicity in observational data. Future studies could also examine ergodicity in treatment trials including people with chronic pain.

To summarize, it seems that self-report data widely used in chronic pain research are not ergodic. As such, one cannot assume precise applicability of the results from studies of people that use conventional group designs to individual people with chronic pain. The results here challenge our reliance, or sole reliance, on conventional group-based methods as a way to understand individual people. Not only do the results reveal significant variability in individual pain experiences between people and over time but they also highlight that studies using intensive longitudinal data collection seem to be needed.

We may need to critically evaluate the prevailing reliance on group-based methods involving aggregated group data, particularly where the primary aim is to translate research findings into individualized care. It may be that failure to fully appreciate the individuality and attendant heterogeneity of people with chronic pain has held us back. If this is the case, a shift toward methodologies that capture the dynamic and individualized nature of pain may pave the way for more effective treatments. As we say, this may be an opportunity or turning point. We may need to change the way we think about behavior, processes of behavior change, and data to ensure that our research methods are as personalized as the treatments we want to design.

Conflict of interest statement

The authors have no conflicts of interest to declare.

Acknowledgements

This research paper was made possible with the support of Anna-Sofia Melin, PhD, and Associate Professor Nina Bohm-Starke, PhD, who assisted with participant recruitment. The authors thank Jennie Hallberg for her help during data collection and Yukai Yang, PhD, for statistical consultation.

This study was funded by Insamlingsstiftelsen Kvinnor och Hälsa (F.S.). The funder had no role in the design and conduct of the study, data collection, management, analysis, interpretation, manuscript preparation, review or approval, or decision to submit the manuscript for publication.

Data availability: The data used in this study are not publicly available. Requests should be directed to the corresponding author and will be considered on a case-by-case basis in accordance with the relevant Swedish and European Union data protection and privacy legislation. The code used to analyze the data in the current study can be accessed upon request to the corresponding author.

Supplemental digital content

Supplemental digital content associated with this article can be found online at http://links.lww.com/PAIN/C242.

Article history:
Received 10 July 2024
Received in revised form 9 November 2024
Accepted 10 December 2024
Available online 18 March 2025

References

[1] Brose A, Schmiedek F, Lövdén M, Molenaar PCM, Lindenberger U. Adult age differences in covariation of motivation and working memory performance: contrasting between-person and within-person findings. Res Hum Develop 2010;7:61–78.
[2] Brose A, Voelkle MC, Lövdén M, Lindenberger U, Schmiedek F. Differences in the between–person and within–person structures of affect are a matter of degree. Eur J Personal 2015;29:55–71.
[3] Buhrman M, Gordh T, Andersson G. Internet interventions for chronic pain including headache: a systematic review. Int Interv 2016;4:17–34.
[4] Ciarrochi J, Sahdra B, Hayes SC, Hofmann SG, Sanford B, Stanton C, Yap K, Fraser MI, Gates K, Gloster AT. A personalised approach to identifying important determinants of well-being. Cogn Ther Res 2024;48:1–22.
[5] Cleeland C, Ryan K. Pain assessment: global use of the brief pain inventory. Singapore: Annals, Academy of Medicine, 1994.
[6] Elbers S, Wittink H, Konings S, Kaiser U, Kleijnen J, Pool J, Köke A, Smeets R. Longitudinal outcome evaluations of Interdisciplinary Multimodal Pain Treatment programmes for patients with chronic primary musculoskeletal pain: a systematic review and meta-analysis. Eur J Pain 2022;26:310–35.
[7] Fisher AJ, Medaglia JD, Jeronimus BF. Lack of group-to-individual generalizability is a threat to human subjects research. Proc Natl Acad Sci U S A 2018;115:E6106–15.
[8] Gates KM, Chow S-M, Molenaar PC. Intensive longitudinal analysis of human processes. Boca Raton: CRC Press; 2023.
[9] Gloster AT, Nadler M, Block V, Haller E, Rubel J, Benoy C, Villanueva J, Bader K, Walter M, Lang U, Hofmann SG, Ciarrochi J, Hayes SC. When average isn't good enough: identifying meaningful subgroups in clinical data. *Cogn Ther Res.* 2024;48:537–551.
[10] Gloster AT, Walder N, Levin ME, Twohig MP, Karekla M. The empirical status of acceptance and commitment therapy: a review of meta-analyses. J Contextual Behav Sci 2020;18:181–92.
[11] Gomes CMA, De Araujo J, Do Nascimento E, Jelihovschi EG. Routine psychological testing of the individual is not valid. Psychol Rep 2019;122:1576–93.
[12] Gómez-de-Regil L, Estrella-Castillo DF. Psychotherapy for physical pain in patients with fibromyalgia: a systematic review. Pain Res Manag 2020;2020:3408052.
[13] Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, McLeod L, Delacqua G, Delacqua F, Kirby J, Duda SN, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019;95:103208.
[14] Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377–81.
[15] Hayes SC, Ciarrochi J, Hofmann SG, Chin F, Sahdra B. Evolving an idionomic approach to processes of change: towards a unified personalized science of human improvement. Behav Res Ther 2022;156:104155.
[16] Hayes SC, Hofmann SG, Ciarrochi J. A process-based approach to psychological diagnosis and treatment:The conceptual and treatment utility of an extended evolutionary meta model. Clin Psychol Rev 2020;82:101908.
[17] Hayes SC, Hofmann SG, Stanton CE, Carpenter JK, Sanford BT, Curtiss JE, Ciarrochi J. The role of the individual in the coming era of process-based therapy. Behav Res Ther 2019;117:40–53.
[18] Hofmann SG, Hayes SC. The future of intervention science: process-based therapy. Clin Psychol Sci 2019;7:37–50.
[19] Hughes LS, Clark J, Colclough JA, Dale E, McMillan D. Acceptance and commitment therapy (ACT) for chronic pain: a systematic review and meta-analyses. Clin J Pain 2017;33:552–68.
[20] Hyndman RJ, Khandakar Y. Automatic time series forecasting: the forecast package for R. J Stat Softw 2008;27:1–22.
[21] Kravitz RL, Duan N, Vohra S, Li J. The DEcIDE Methods Center N-of-1 Guidance Panel. Introduction to N-of-1 trials: indications and barriers. Design and implementation of N-of-1 trials: a user's guide. Rockville: Agency for Healthcare Research and Quality; 2014:1–11.

[22] Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Int Med 2001;16:606–13.
[23] Lowie WM, Verspoor MH. Individual differences and the ergodicity problem. Lang Learn 2019;69:184–206.
[24] McCracken LM. Beyond therapy types: mindful self-compassion and the future of process-based therapy for chronic pain. Eur J Pain 2021;25:729–30.
[25] McCracken LM. Personalized pain management: is it time for process-based therapy for particular people with chronic pain? Eur J Pain 2023; 27:1044–55.
[26] McCracken LM. Psychological flexibility, chronic pain, and health. Annu Rev Psychol 2024;75:601–24.
[27] McCracken LM, Morley S. The psychological flexibility model: a basis for integration and progress in psychological approaches to chronic pain management. J Pain 2014;15:221–34.
[28] McDonald S, Quinn F, Vieira R, O'Brien N, White M, Johnston DW, Sniehotta FF. The state of the art and future opportunities for using longitudinal n-of-1 methods in health behaviour research: a systematic literature overview. Health Psychol Rev 2017;11:307–23.
[29] Mestdagh M, Verdonck S, Piot M, Niemeijer K, Kilani G, Tuerlinckx F, Kuppens P, Dejonckheere E. m-Path: an easy-to-use and highly tailorable platform for ecological momentary assessment and intervention in behavioral research and clinical practice. Front Digital Health 2023;5:1182175.
[30] Molenaar PCM. A manifesto on psychology as idiographic science: bringing the person back into scientific psychology, this time forever. Meas Interdiscip Res Perspect 2004;2:201–18.
[31] Molenaar PCM, Campbell CG. The new person-specific paradigm in psychology. Curr Dir Psychol Sci 2009;18:112–7.
[32] Quartana PJ, Campbell CM, Edwards RR. Pain catastrophizing: a critical review. Expert Rev Neurotherapeut 2009;9:745–58.
[33] Rolffs JL, Rogge RD, Wilson KG. Disentangling components of flexibility via the hexaflex model: development and validation of the multidimensional psychological flexibility inventory (MPFI). Assessment 2018;25:458–82.
[34] Sahdra BK, Ciarrochi J, Klimczak KS, Krafft J, Hayes SC, Levin M. Testing the applicability of idionomic statistics in longitudinal studies: the example of "doing what matters". J Contextual Behav Sci 2024;32:100728.
[35] Sullivan MJ, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. Psychol Assess 1995;7:524–32.
[36] Sullivan MJ, Lynch ME, Clark A. Dimensions of catastrophic thinking associated with pain experience and disability in patients with neuropathic pain conditions. PAIN 2005;113:310–5.
[37] Sundström FT, Lavefjord A, Buhrman M, McCracken LM. Assessing psychological flexibility and inflexibility in chronic pain using the multidimensional psychological flexibility inventory (MPFI). J Pain 2023; 24:770–81.
[38] Sundström FT, Lavefjord A, Buhrman M, McCracken LM. Associations between psychological flexibility and daily functioning in endometriosis-related pain. Scand J Pain 2024;24:20220157.
[39] Trompetter HR, Bohlmeijer ET, Fox J-P, Schreurs KMG. Psychological flexibility and catastrophizing as associated change mechanisms during online Acceptance & Commitment Therapy for chronic pain. Behav Res Ther 2015;74:50–9.
[40] Wickham H. ggplot2. WIREs Comput Stat 2011;3:180–5.
[41] Wickham H, François R, Henry L, Müller K, Vaughan D. dplyr: a grammar of data manipulation. R package version 1.1.4. Computer software, 2023. Available at: https://github.com/tidyverse/dplyr.
[42] Williams ACC, Fisher E, Hearn L, Eccleston C. Psychological therapies for the management of chronic pain (excluding headache) in adults. Cochrane Database Syst Rev. 2020;8:CD007407.