PAIN®

Investigating intraindividual pain variability: methods, applications, issues, and directions

Chung Jung Mun[a,*], Hye Won Suk[b], Mary C. Davis[c], Paul Karoly[c], Patrick Finan[a], Howard Tennen[d], Mark P. Jensen[e]

Abstract

Pain is a dynamic experience subject to substantial individual differences. Intensive longitudinal designs best capture the dynamical ebb and flow of the pain experience across time and settings. Thanks to the development of innovative and efficient data collection technologies, conducting an intensive longitudinal pain study has become increasingly feasible. However, the majority of longitudinal studies have tended to examine average level of pain as a predictor or as an outcome, while conceptualizing intraindividual pain variation as noise, error, or a nuisance factor. Such an approach may miss the opportunity to understand how fluctuations in pain over time are associated with pain processing, coping, other indices of adjustment, and treatment response. The present review introduces the 4 most frequently used intraindividual variability indices: the intraindividual SD/variance, autocorrelation, the mean square of successive difference, and probability of acute change. In addition, we discuss recent development in dynamic structural equation modeling in a nontechnical manner. We also consider some notable methodological issues, present a real-world example of intraindividual variability analysis, and offer suggestions for future research. Finally, we provide statistical software syntax for calculating the aforementioned intraindividual pain variability indices so that researchers can easily apply them in their research. We believe that investigating intraindividual variability of pain will provide a new perspective for understanding the complex mechanisms underlying pain coping and adjustment, as well as for enhancing efforts in precision pain medicine. Audio accompanying this abstract is available online as supplemental digital content at http://links.lww.com/PAIN/A817.

Keywords: Pain variability, Pain dynamics, Intraindividual variability, Instability, Chronic pain, Intensive longitudinal design, Daily diary, Autocorrelation, Dynamic structural equation modeling

1. Introduction

Chronic pain is the experience of noxious sensory and affective arousal that extends over a long period of time (eg, more than 3 months[58,90]). Although the term *chronic* conveys a sense of stable resistance to change, pain is fundamentally *dynamic* rather than static. In fact, numerous studies have demonstrated that, among individuals with chronic pain, pain intensity levels vary substantially over time periods ranging from moments to hours to days.[13,28,47,60,66,72,84]

One of the best ways to understand individuals' dynamical pain experience is to use an intensive and repeated assessment of pain-relevant variables in real-time and in the real-world settings.[77] A majority of the extant research on chronic pain, however, has relied on single retrospective assessments of average, worst, or least pain intensity levels. Although such a strategy is widely used to evaluate the efficacy of pain treatment on usual pain experience, it does not allow for an examination of dynamical pain experiences and processes, and introduces recall error and bias.[8,82–84] Hence, to capture more accurately the dynamic ebb and flow of individuals' pain experiences, investigators must turn to multiple and intensive assessment. Unfortunately, a barrier to such an approach is a lack of resources to guide researchers on how they can best take advantage of these rich data to examine the variable nature of pain experiences. To address this issue, the present review provides an introduction to a number of statistical approaches that allow for a nuanced understanding of changing pain processes. In addition, we address important methodological issues, illustrate a real-world example of the application of variability indices, and provide a discussion of future research directions.

1.1. A brief overview of intensive longitudinal design

An intensive longitudinal design (ILD) refers to a procedure that includes numerous repeated measurements of individuals' experiences, cognition, mood, behavior, and/or physiology over time.[6] Daily diaries, experience sampling methods, ecological momentary assessment, and ambulatory assessment are all considered special cases of ILD. Intensive longitudinal data can be collected using a variety of methods including paper-and-pencil diaries, personal digital assistants, interactive voice

recordings, smartphone applications, text messages or emails, and wearable devices (eg, actigraphy and wrist heart rate monitor).[77] In short, ILD allows for measuring individuals' various experiences, as they occur in ecologically valid contexts.

Intensive longitudinal design can be largely divided into 3 different sampling types[77,81,95]: (1) time-contingent sampling; (2) event-contingent sampling; and (3) hybrid sampling (combination of the 2). *Time-contingent sampling* requires participants to complete assessments at fixed, random, or a combination of random and fixed times each day. In the case of fixed-prompt assessments, participants are typically provided with a specific time window during the day (eg, 20 minutes upon awakening or 30 minutes before going to bed) to complete their assessments (**Fig. 1A**). Random-prompt assessments, on the other hand, require participants to complete assessments at randomly determined time points throughout the day in order for researchers to capture a more representative and less biased momentary state or experience of a participant[77] (**Fig. 1B**). A combination of both fixed and random time assessments can also be used. For instance, participants might be asked to rate pain in each of the predetermined time windows in the morning, afternoon, or evening, with random prompts occurring within each fixed-time window (**Fig. 1C**).

The other popular ILD sampling scheme is called *event-contingent sampling*. Rather than providing alerts (eg, beeps, phone calls, or text prompts) signaling participants to assess their current state or experience, participants are instead asked to voluntarily enter their assessments whenever an important event (eg, experiencing more than the usual level of pain, use of opioids, or an argument with a partner) occurs. This method requires researchers to provide participants a clear definition for determining a targeted event and sufficient training for them to participate in this type of assessment. A major limitation of this sampling strategy is that it is impossible to monitor whether participants entered an assessment when an event occurred. This design is depicted in **Figure 1D**.

Finally, *hybrid sampling* includes both time- and event-contingent sampling. These methods tend to complement each other. This sampling approach can provide a more comprehensive picture of an individual's experience and of how one's state is associated with or precedes a target event.[77] This design is depicted in **Figure 1E**.

A daily diary study with fixed-time assessment (eg, at the end of the day) requires the simplest statistical approach because the time interval between measurements is approximately equal. However, daily diary studies using fixed assessments are less likely to capture participants' experiences in "real time" and, thus, are more susceptible to recall error and bias.[76] Other study designs such as ecological momentary assessment or experience sampling method that incorporate random prompts and/or event-contingent reports address this limitation. However, the statistical modeling required for these study designs is much more complex, especially when investigators are interested in examining lagged effects (ie, the effect of one variable on another that is measured at a later time point), because the time interval can vary widely across measurements. We discuss this issue as well as existing methodological solutions further in the section on "Notable methodological issues."

1.2. Why we need to investigate pain variability

Variation in pain intensity is an important clinical target, as it is closely associated with an individual's moment-to-moment cognition, affect, behavior, and motivation. For example,

Litcher-Kelly et al.[54] found that, among individuals with chronic pain, momentary changes in pain intensity were associated with both affective distress and activity limitations. Similarly, in a community sample of individuals with chronic pain, greater than usual pain intensity in the morning was inversely associated with anticipatory work and lifestyle goal cognitions in the morning, and both of these factors predicted afternoon goal pursuit.[43] Recently, a daily diary study of individuals with fibromyalgia showed that experiencing more than typical experience of pain intensity in the morning was associated with higher afternoon pain catastrophizing, which, in turn, predicted elevated pain intensity later in the day.[87]

However, studies such as these provide an incomplete understanding of pain variability because their analytic approaches focus primarily on how daily or momentary pain deviating from individuals' typical level of pain is associated with outcomes. The majority of previous studies have not investigated an important source of information that can be gained from ILD, namely individual differences in *intraindividual pain variability*. Intraindividual pain variability denotes how one's experience of pain fluctuates over time. Some previous studies have examined this variability,[25,28,47,72,73,84,97] illustrating the importance of investigating intraindividual differences. In fact, it has been recently suggested that assessing intraindividual pain variability and the temporal features of pain could facilitate more accurate classification of chronic pain conditions and enhance precision pain medicine.[21,26] Below, we summarize 3 streams of the literature that reveal why it is important to evaluate pain variability. Throughout our discussion, "intraindividual pain variability" and "pain variability" are used interchangeably.

First, individual differences in pain variability have been examined in the context of pain recall error and bias. Individuals with greater intraindividual pain variability tend to rate recalled pain higher relative to their average momentary pain ratings over time.[28,47,84] Stone et al.[84] explain that this effect could be due to the fact that individuals with higher pain variability are more likely to be exposed to experiences of higher pain. Based on the "peak-end effect" (ie, individuals are more likely to remember their peak pain experience or most recently experienced pain), those with greater variability may therefore be more likely to attend to (and recall) high pain experiences and provide overall higher pain recall ratings.[84] More recently, Lefebvre and Keefe[52] investigated whether some individual difference factors, including neuroticism and depressive symptoms, are associated with recall of pain variability. They found that higher neuroticism levels were related to more accurate recall of the variability of pain unpleasantness over time, whereas depressive symptoms were not associated with greater accuracy in pain variability recall.

Second, researchers have also been investigating the associations among intraindividual pain variability, personality characteristics, and physical and psychosocial functioning. Schneider et al.[72] reported on 2 daily diary studies of individuals with rheumatological illnesses and osteoarthritis, and found that individuals with greater intraindividual pain variability reported higher depression levels and lower arthritis self-efficacy. Similarly, in an older sample of adults who completed 30-day daily diary, Zakoscielna and Patricia[97] found that baseline depressive symptoms and pain were significant predictors of higher pain variability. More recently, some investigators used a "regime-switching" model that focuses on recurring shifts between different states (or "regimes") and found that persistence of average pain duration and dominance of higher pain states across consecutive time points significantly predicted emotional and physical functioning, respectively.[73]
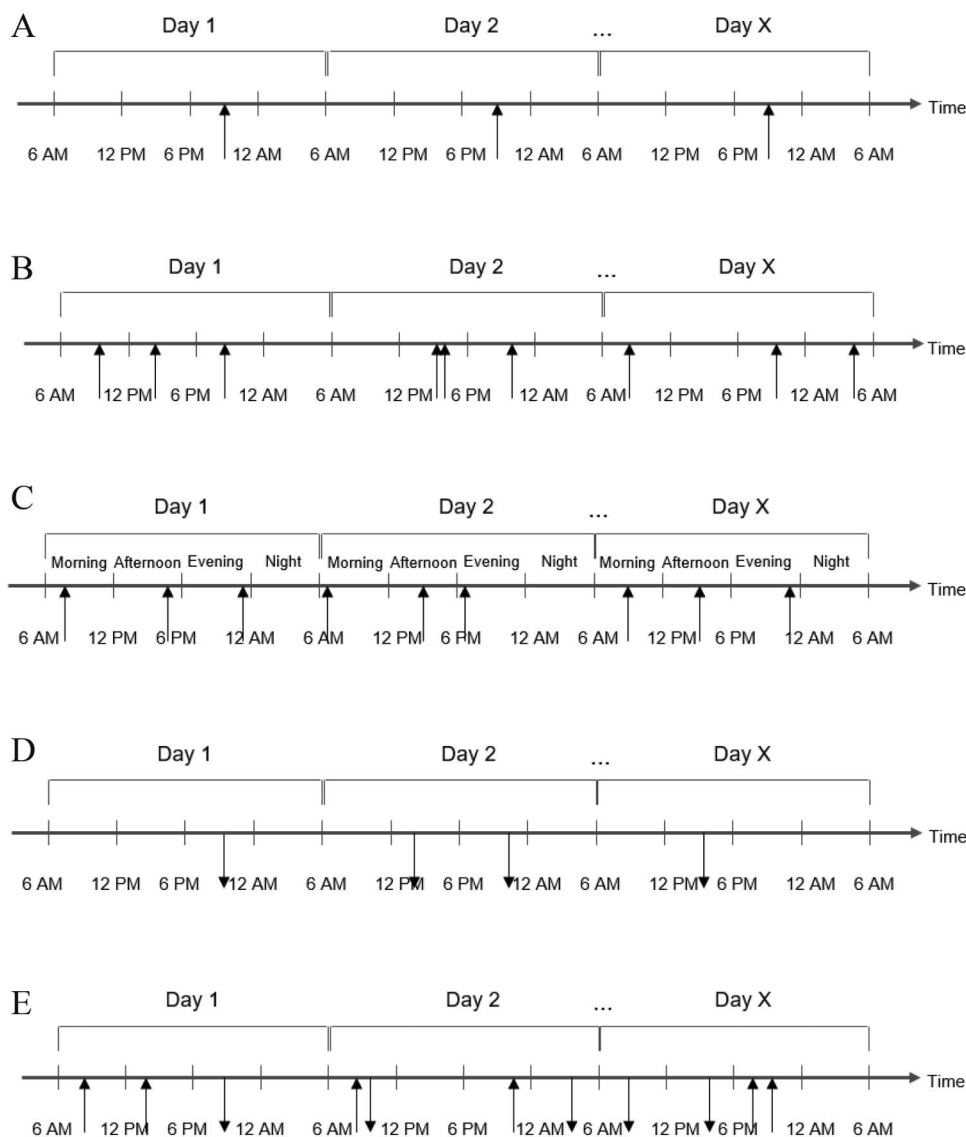
Figure 1. (A) Illustration of fixed-prompt assessment. Arrows going up indicate prompts. (B) Illustration of random-prompt assessment. (C) Illustration of combination (fixed-time and random-prompt) assessment. (D) Illustration of event-contingent assessment. Arrows going down indicate event-contingent report. (E) Illustration of hybrid sampling assessment.

Finally, individual differences in baseline pain variability are also recognized as important predictors of clinical trial outcomes. In a randomized, placebo-controlled clinical trial of milnacipran (ie, a dual serotonin and norepinephrine reuptake inhibitor) for individuals with fibromyalgia, those with higher baseline pain variability were more likely to respond to placebo but not the active drug.[35] This pattern was further replicated by pooled analyses of 12 different placebo-controlled gabapentin and pregabalin clinical trials in postherpetic neuralgia and painful diabetic peripheral neuropathy. Farrar et al.[25] demonstrated that individuals with higher pain variability, which was measured using data from a 7-day daily diary at baseline, were more likely to respond to a placebo but not to active medications. On the other hand, using a pooled analysis of 4 double-blind, randomized controlled trials that examined the efficacy of 8% topical capsaicin in patients with postherpetic neuropathic pain, a study by Martini et al.[56] reported a different pattern. They found that individuals with higher intraindividual pain variability during the 14 days before treatment were more likely to fall into the treatment-responder trajectory

group that showed the largest pain intensity reductions during the treatment.

Findings from these studies collectively point to individual differences in pain variability as both an interesting and important area of pain research that deserves greater attention. Identifying factors that are associated with individual differences in pain variability and examining how pain variability predicts individuals' psychosocial and physical functioning, as well as treatment response, holds promise for informing clinical efforts to develop more personalized and effective chronic pain remediation.

1.3. The present review

Methods to assess momentary pain experience over time have become widely accessible to researchers. However, relative to the dramatic growth of ILD methods in pain research, few studies have successfully applied quantitative methods to calculate intraindividual pain variability. Even when pain variability has been explored, most previous studies[25,28,71,84,97] have relied on

a single variability index (ie, the intraindividual SD). One of the major barriers to examining individual differences in pain variability is the complexity of data analytic skills required to analyze rich intensive longitudinal data. The present review provides conceptual background and hands-on guidance in the application of statistical methods to quantify intraindividual pain variability to encourage researchers to more widely use measures of intraindividual pain variability when addressing questions of interest. We review different methods of calculating intraindividual variability, note some important methodological issues that need to be addressed in future studies, and provide an introduction to a new statistical modeling framework (ie, dynamic structural equation modeling [DSEM]) that allows for more effective investigation of pain variability. We also offer an actual demonstration using empirical data. Finally, we provide statistical software syntax that we developed for computing a range of intraindividual variability indices, to allow researchers to readily calculate intraindividual pain variability using their own data.

We limit our scope to individual differences in intraindividual pain variability. Investigating some other research questions such as "How much time does it take for an individual's sudden spike of pain experience to return to its usual equilibrium?" and "How does the dynamic association between pain and depressive symptoms change over time during a pain intervention?" requires other more complex modeling approaches (eg, differential equation modeling and time-varying effect modeling) that are beyond the scope and aim of the present review.

Before moving on to more detailed descriptions of each pain variability index, we note that the term "pain variability" used in the present review is based on measures of *pain intensity* measured by either a numerical rating scale (NRS) or visual analogue scale. Pain experience is fundamentally subjective and cannot be observed by those who are not experiencing it.[96] Hence, to better understand pain, it is important to use multimodal assessments and investigate the context and process of pain reports (see Refs. 39 and 96 for a review). In this review, we focus on pain intensity ratings when discussing pain variability simply for illustrative purposes. Pain intensity is the most common pain domain assessed in pain research and clinical settings[26]; therefore, it provides the most straightforward and relevant indicator through which intraindividual pain variability is understood. It is certainly possible to calculate the variability of other pain-related measures. We hope that the present article provides momentum for researchers to go beyond the assessment of pain intensity when researching pain variability in the future.

1.4. Methods of calculating intraindividual pain variability

The basic concept of variability refers to the degree of spread manifested in a group of data. Three indices of variability are *range* (ie, the lowest score subtracted from the highest score), *variance* (ie, how close the scores in the distribution are to the middle of the distribution), and SD (ie, the square root of the variance). When using single retrospective assessments of pain, these variability measures capture "interindividual" difference (eg, how pain levels spread out across the sample). However, these indices do not capture the "intraindividual" variability of pain, which involves the fluctuation in pain that an individual experiences over a certain time interval.

Capturing the "intraindividual" variability of pain is more complex than capturing the "interindividual" variability. The good news, however, is that the methods we describe here are not overwhelmingly complex. As long as one has a firm understanding of the basic concepts of variability and correlation, one should be able to understand and apply the methods that follow.

Before providing detailed descriptions of each intraindividual pain variability index, we note that we are not providing an exhaustive review of all possible methods to calculate intraindividual variability. Instead, we focus on methods that have been used most often in relevant domains of research, particularly in affective dynamics (cf., Refs. 33, 38, and 91). Below, we discuss 4 relatively well-known intraindividual variability measures and note their strengths and limitations, as well as some issues involved in their use.

1.4.1. Intraindividual SD (magnitude of fluctuations)

The intraindividual SD or intraindividual variance[62,68] (iSD or iSD^2) is the simplest and most popular index of intraindividual variability. This is probably because iSD has the most intuitive appeal in that it captures each participant's SD across all his or her observations. A high iSD indicates a high amplitude of fluctuations. This method is represented in Equation 1. iSD_i indicates the iSD for the ith individual. $X_{i,t}$ denotes the observed score of the ith individual at the tth occasion. The mean of the individual's scores is denoted as $\bar{X}_i$, and n_i is the number of measurement occasions for the individual. In sum, what iSD measures is simply the overall magnitude of the observed fluctuations.

$$iSD_i = \sqrt{\frac{1}{n_i - 1}\sum_{t=1}^{n_i}\left(X_{i,t} - \bar{X}_i\right)^2}. \qquad (1)$$

The major drawback of this method, however, is that iSD does not possess temporal sensitivity and does not capture the frequency of fluctuations. For example, let us say that participants were asked to rate their pain level every day for 10 days on a 0 to 10 NRS. Participant A reported having pain level of 3 for the first 5 days and a pain level of 7 for next 5 days (ie, 3, 3, 3, 3, 3, 7, 7, 7, 7, and 7). Participant B reported that his/her pain level fluctuated every day from 3 to 7 across 10 days (ie, 3, 7, 3, 7, 3, 7, 3, 7, 3, and 7). Although participants A and B show very different intraindividual pain variability patterns across days, their iSD values are identical (iSD = 2.11 for both cases). Because important information regarding differences between individuals is lost, iSD values should be interpreted with caution.

1.4.2. Autocorrelation (temporal dependency)

Another method of calculating the intraindividual variability of pain is autocorrelation, an index that measures the extent to which current observations can be predicted from previous observations. This index captures temporal dependency. The mathematical formula for autocorrelation is shown in Equation 2, where $\rho_i(\tau)$ indicates the autocorrelation with lag τ for the ith individual. $\gamma_i(\tau)$ is the autocovariance at lag τ for the ith individual as defined in Equation 3. The autocovariance at lag τ indicates the covariance between the observation at the tth measurement occasion and the observation at $(t + \tau)^{th}$ measurement occasion. $\gamma_i(0)$ indicates the variance of the observations for the individual, and the square root of this is identical to the iSD (Equation 1). The notation ˆ indicates an estimate; for instance, $\hat{\rho}_i(\tau)$ is an estimate of $\rho_i(\tau)$.

$$\hat{\rho}_i(\tau) = \frac{\hat{\gamma}_i(\tau)}{\hat{\gamma}_i(0)}. \qquad (2)$$

$$\hat{\gamma}_i(\tau) = \frac{1}{n_i - \tau - 1}\sum_{t=1}^{n_i - \tau}\left(X_{i,t} - \bar{X}_i\right)\left(X_{i,t+\tau} - \bar{X}_i\right). \qquad (3)$$

A positive value of an autocorrelation indicates that if a person shows a higher (or lower) pain level than his or her mean at

a particular occasion, it is likely that this person also shows higher (or lower)-than-average pain level at a later occasion with lag τ. For example, a positive autocorrelation at lag 1 [AR(1)] indicates that the pain experience is more likely to persist over time when it is above or below the mean pain level. By contrast, a negative value of an AR(1) implies a back-and-forth pattern of pain ratings. If a person reported higher than his or her average level of pain yesterday, he or she will likely report a lower-than-average level of pain today, and a higher-than-average level of pain tomorrow. Values of an autocorrelation that are close to 0 indicate that the pain level at a particular occasion does not predict the pain level at a later time point (at lag τ). AR(1) close to 0 would mean that yesterday's pain experience is not predictive of today's pain experience.

The major advantage of using the autocorrelation over iSD is that the former captures the order effects of observations. For instance, returning to the previous example, participant A, with a 10-day pain pattern of 3, 3, 3, 3, 3, 7, 7, 7, 7, and 7, has an estimated autocorrelation of 0.7 at lag 1 [AR(1)]. By contrast, participant B, with a 10-day pain pattern of 3, 7, 3, 7, 3, 7, 3, 7, 3, and 7, has an estimated AR(1) of −0.9. These findings suggest that participant A shows a greater resistance to change when his or her pain is elevated or attenuated than does participant B.

The downside of autocorrelation index, however, is that it only captures temporal dependency; it does not provide information regarding the overall magnitude of the observed fluctuations. Furthermore, the result can be significantly different depending on which lag (ie, order of autocorrelation) is selected. For instance, if a participant had a 10-day pain pattern of 3, 3, 7, 7, 3, 3, 7, 7, and 3, the estimated lag 1 [AR(1)] is 0.01, indicating almost no temporal dependence between today's and tomorrow's pain levels. However, if we choose lag 2, the estimated second-order autocorrelation [AR(2)] is −0.79, indicating a strong and negative temporal dependence in the pain level (ie, a back-and-forth pattern of pain ratings) between today and the day after tomorrow. Researchers have primarily used the lag 1 autocorrelation because of its simplicity. In addition, providing a good rationale for choosing a different of lag of autocorrelation is conceptually challenging.

Autocorrelation has been a very popular intraindividual variability index in the study of affective dynamics.[51,70,85] In affective dynamics, autocorrelation—more specifically AR(1)—is interpreted as reflecting an individual's *affective inertia*.[51,85] Continuously adapting emotions in response to situational demands is important for psychological well-being, and absence of flexibility and contextual sensitivity often presages psychopathology.[44] In the same vein, higher affective inertia has been found to be associated with maladaptive emotion-regulation functioning.[49,51] On the other hand, lower affective inertia has been thought to represent flexible and successful regulation of emotion in changing environments.[48]

Whether this concept of "inertia" can be directly applied to the pain experience is largely unknown. The mechanisms that cause temporal dependence of pain might be different from those that relate to emotion-regulation; and high autocorrelations in pain experience can have different implications from high affective inertia. Higher autocorrelations reflect not only greater resistance to change and inflexibility but also higher predictability. If a pain experience becomes highly fluctuating with increased uncertainty, individuals may feel more anxious and helpless about controlling their pain. Therefore, lower autocorrelations in pain may be related to decreased psychological well-being and physical functioning. These thoughts are merely our speculation, however. We encourage empirical studies that carefully and thoroughly examine how autocorrelation of pain may map onto individuals' personal characteristics and daily functioning.

1.4.3. Mean square of successive difference (temporal instability)

The mean square of successive difference (MSSD[53,63]), which is also known as *temporal instability*,[38] can address some of the limitations of iSD and autocorrelation. This method combines the qualities of both the iSD (*magnitude of fluctuations*) and the autocorrelation (*temporal dependency*). The equation is as shown below (Equation 4). $MSSD_i$ indicates the MSSD score for the ith individual and is obtained by averaging the squared difference between observations at $(t + 1)^{th}$ and t^{th} occasions. n_i represents the total number of measurement occasions for the individual.

$$MSSD_i = \frac{1}{n_i - 1}\sum_{t=1}^{n_i - 1}\left(X_{i,t+1} - X_{i,t}\right)^2. \qquad (4)$$

The MSSD reflects the average of the squared successive changes between 2 adjacent observations; and the major advantage for its use is that it integrates both magnitude of change and temporal dependency. The MSSD is less influenced by systematic intraindividual trends across time, which can lead to spuriously high variance and high autocorrelation.[38]

However, the MSSD is not free from limitations. First, compared with iSD or autocorrelation, the interpretation of the MSSD is less clear, as it incorporates both amplitude of fluctuations and temporal dependency. The relations among the MSSD, iSD, and AR(1) are shown below (Equation 5).[75] Basically, the expected value of the MSSD is a product of the population variance (iSD^2) and 1 − AR(1). Hence, a large expected value of the MSSD can be attributed to high iSD and/or low autocorrelation. However, if both iSD and autocorrelation are high (eg, as is the case for participant A), the MSSD will have a small expected value, as the effects of iSD and autocorrelation will cancel each other out. In fact, in the affective dynamics literature, the use of only the MSSD has been quite problematic; previous studies have shown that high iSD and high autocorrelation of negative affect are both associated with poorer psychosocial outcomes.[34,51,91] Consequently, some researchers suggest that rather than merely reporting the MSSD, both iSD and autocorrelation should be reported.[93]

$$E(MSSD_i) = 2\left(iSD_i^2\right)\left(1 - \rho(1)\right). \qquad (5)$$

1.4.4. Probability of acute change

The probability of acute change (PAC) is the frequency of acute changes (eg, sudden change in pain magnitude) divided by the total number of successive changes. Individual investigators determine the cutpoint that defines the acute change magnitude. The PAC score is particularly useful when the focus of research is on examining the likelihood of significant changes across observations. For instance, pain researchers might be interested in the frequency of either sudden elevations or decreases in pain intensity. A specific threshold or a cutpoint for an acute change can be set by the investigator (eg, a change in magnitude greater than 2 on a 0-10 NRS pain ratings from one observation to the next), allowing for the calculation of how frequently an individual experiences moment-to-moment (or day-to-day) changes in pain above (or below) a given cutpoint.

$$PAC_i = \frac{1}{n_i - 1}\sum_{t=1}^{n_i - 1}AC_{i,t+1}, \qquad (6)$$

where $AC_{i,t+1} = 1$, if $|X_{i,t+1} - X_{i,t}| > k$ (k indicates a predetermined cutpoint), and $AC_{i,t+1} = 0$, otherwise.

approach ignores the distinction between different time points falling into the same segment and, therefore, can lead to some loss of information about the exact timing of observations. Using a finer grid of time segments to avoid loss of such information would require adding a large amount of missing data. However, Asparouhov et al.[2] performed simulation studies demonstrating that using shorter segments produced reliable results despite a large amount (eg, more than 80%) of missing data inserted. de Haan-Rietdijk et al.[31] also found that this approach yielded results that were not substantially different from those obtained using a continuous time modeling approach. The discretization method has been recently adopted in Mplus, which is one of the most widely used structural equation modeling software programs. Users in Mplus can easily draw on this procedure by adding one line of extra syntax when using dynamic structural equation models (see more detailed instruction at http://www.statmodel.com/TimeSeries.shtml).

1.5.4. Handling missing data

Typically, about 20% to 25% of data are missing in intensive longitudinal studies. Yet, dealing with multilevel missing data is a relatively new topic. Recently, a multiple imputation method has been developed to handle multilevel missing data (see Ref. 23). Specifically, Keller and Enders[46] developed a free software called "Blimp" that can handle multilevel missing data (up to 3 levels) based on multiple imputation. However, it is not yet possible to use this software to handle missing data when computing intraindividual variability indices.

In the case of the DSEM framework, Bayesian estimation with a Markov Chain Monte Carlo (MCMC) algorithm is used to handle missing data, in which missing data are treated as unknown parameters.[2] In the Bayesian approach, unknown parameters are considered as random variables that have distributions, and the posterior distributions of the random variables are obtained for statistical inferences. This estimation method consists of an iterative process that begins with providing starting values for all unknown parameters. Subsequently, it samples each unknown from its conditional distribution given the starting values of the other unknowns and the data. Finally, it repeats this process with the generated values replacing the starting values. After a sufficient number of iterations, the chain of the values produced by this iterative procedure converges to a stationary distribution, which approximates the joint posterior distribution of the unknown parameters. After convergence, further iterations continue to form the joint posterior distribution by sampling and are used to make inferences about the unknown parameters. As Hamaker et al.[32] pointed out, this method yields consistent estimation when missing data are missing at random (MAR). If participants fail to respond because of their high level of pain, missing data would not meet the MAR assumption and the results would not be dependable. However, it is difficult to determine whether the MAR assumption has been met. Thus, how to properly handle missing data for calculating intraindividual variability should be rigorously examined in future studies. Specifically, it would be important to investigate (1) how different missing data patterns (ie, missing completely at random vs MAR vs missing not at random) influence the estimation of intraindividual variability; and (2) how much missing data are acceptable (eg, less than 40% of data missing) to reliably estimate intraindividual variability.

1.5.5. Detrending data

In the time series literature, analyses frequently rely on the assumption of stationarity, which indicates the mean, variance, and autocorrelation of the time series are constant over time.[12] However, time series or longitudinal data in everyday life often show trends that cast doubt on the stationarity assumption. Ignoring systematic trends or cycles in time series data may lead to erroneous inferences when examining the relationships between 2 time-varying variables. Thus, it is important to control for the time effects by removing trends or cycles unrelated to the experimental manipulations or the study designs.[55,94]

Failing to address systematic trends in one's data has serious consequences for calculating some variability measures. The MSSD is known to be only slightly affected by systematic trends[38,93] and so is the PAC.[38] However, if existing trends are not removed, these indices may yield dramatically increased variances and spurious autocorrelations.[38,93] Therefore, researchers have advocated detrending data before calculating iSD and AR(1).[38,93]

Detrending refers to a statistical operation of eliminating the trend from a time series data.[94] It may appear to be conceptually straightforward. However, in practice, identifying and handling trends are highly subjective processes.[93] Hamaker et al.[32] also pointed out that "It [how to handle potential trends and cycles in the data] should therefore reflect our conviction of the underlying mechanism, and whether we believe trends and cycles are something that happen separately from the dynamics, or as an intrinsic part of it" (p. 17). Unfortunately, the underlying mechanism of a trend is often unknown, and it is not easy to determine whether a trend in one's data reflects the variability and dynamics of the phenomenon under study, or irrelevant processes that should be removed from the analysis. The augmented Dickey–Fuller (ADF[15,16]) test can guide an investigator in determining if the data in question need to be detrended. A nonsignificant result of this test indicates that the data are not stationary and may need to be detrended. The ADF test can also be used to examine whether the detrended data are stationary. However, it is not clear what steps should be taken if the detrended data were still not stationary. For example, if removing a linear trend fails to yield stationary data, should the investigator try removing a quadratic or cubic trend, or use more complex detrending methods? Using a complex detrending model may help the data become stationary in some cases. However, such a model should not be used without a specific rationale because it may introduce artificial autocorrelations into the data.[93]

Selecting an optimal method of detrending is also challenging. There are a variety of methods for estimating a trend, such as polynomial curve fitting (eg, linear, quadratic, and cubic) and nonparametric curve fitting (eg, kernel smoothing, LOWESS, and smoothing splines) (cf., Shumway and Stoffer,[78] Chapter 2 for more details about detrending methods). Other systematic effects, such as weekend effect or day of the week effect, can also be removed before calculating variability measures.[38,72] However, to the best of our knowledge, there is no clear guidance for determining when detrending is needed, and if so, which detrending method should be used. In sum, how to properly handle existing trends in data before calculating intraindividual variability measures and how different detrending strategies affect intraindividual variability measures should be more rigorously examined in future methodological studies.

1.5.6. Intraindividual variability can be confounded by the mean

It has also been reported that intraindividual variability measures can be confounded with the individual's mean value for the measure in question, especially with bounded measurement

tools such as NRS.[4,22,42] In other words, the possible values of variability may depend on the mean in situations where the measurements are bounded. For instance, when using a 0 to 10 NRS for pain intensity, if an individual's 30-day mean pain level was 10, this indicates that all pain ratings were 10, and hence, no variability exists. However, if the mean pain level was 5, there are a number of attainable iSD, MSSD, AR(1), and PAC values. Clearly, scores of these variability indices depend in part on the mean value. Mestdagh et al.[59] recently developed the Relative Variability Index (RVI), which allows researchers to examine the role of iSD and MSSD independent of the mean level. The RVI is defined as "the proportion of variability that is observed, relative to the maximum possible variability which can be observed given a certain mean" (p. 694).[59] More technical details about this new approach and software tools (R and MATLAB packages) for estimating the RVI are provided by Mestdagh et al.[59]

1.6. Introduction to dynamic structural equation modeling

With the exponential growth of intensive longitudinal studies, there has also been a rapid development in statistical techniques and software that can better handle the data that are derived from such studies. Dynamic structural equation modeling is a promising recent development in modeling that can help researchers effectively analyze intensive longitudinal data of pain experiences. Below, we provide a brief overview of how the DSEM framework applies to research in intraindividual pain variability. More technical details on DSEM are discussed by Asparouhov et al.[2]

Dynamic structural equation modeling is an extension of single-level time series analysis to 2-level, multivariate time series analysis in the structural equation modeling framework.[2] Although traditional time series analysis was constrained to a single case (N = 1), DSEM allows for multiple individuals (N > 1) and can handle multiple dependent variables in a model. As shown in **Figure 2**, DSEM decomposes the hierarchical (nested) data of an observed variable to within- and between-person portions using latent variable centering. At the within-person level, researchers can estimate intraindividual variability indices, specifically iSD^2 (intraindividual variance) and autocorrelation. Methods for estimating the MSSD and PAC are not yet available in this framework. If interested, researchers can also examine the cross-lagged relations between 2 variables. The intraindividual variability estimates (ie, iSD^2 and autocorrelation) are allowed to vary across individuals (ie, random effect) and can be used as a predictor or an outcome variable at the between-person level.[32] Dynamic structural equation modeling is based on Bayesian estimation with the MCMC procedure, which allows for specifying a model with a joint distribution of multiple variables.[45]

Dynamic structural equation modeling has a number of advantages over standard multilevel modeling (MLM) when calculating intraindividual variability indices. First, in DSEM, both iSD^2 and autocorrelation are estimated for the underlying latent variables (free from measurement errors) rather than the observed scores. This is an important strength because, as mentioned previously, it is difficult to achieve good reliability with autocorrelation, even with a large number of assessments. Second, DSEM in Mplus adopts the discretization approach (ie, adding missing values between observations to artificially render the time interval between observations approximately equal) and adequately handles the unequal spacing problem in ILD. Third, DSEM can also effectively handle missing data when the MAR assumption is met using Bayesian estimation.[2] Fourth, it uses latent variable centering and can eliminate negative bias (ie, underestimating model parameters) that can occur when

centering a lagged predictor with person-mean centering (a widely used centering method in the regular MLM framework). Fifth, DSEM also allows level-1 residual variance, also known as "innovation variance," to vary across individuals. This innovation variance represents a variance that cannot be explained by previous scores or states.[41] For example, innovation variance can be viewed as the variance due to some day-specific factors other than yesterday's pain experience, such as mood, sleep disturbance, pain medication use, hormone levels, and social interaction. Using DSEM, researchers can easily explore individual difference in innovation variance—which is increasingly recognized as an additional source of information that could be clinically relevant.[32,41] Finally, all of the variability indices that are derived from DSEM can be included as a predictor, a mediator, or an outcome in a model simultaneously with other variables (ie, one-step approach). In a traditional MLM framework, this one-step approach is only available when intraindividual variability measures are used as outcomes. If researchers were interested in using variability indices as either predictors or mediators, a 2-step approach (ie, computing variability indices for each participant and then using these values for subsequent analyses) was the only option that was previously available. In summary, recent developments in DSEM provide a great resource for researchers, enabling them to more readily and flexibly examine intraindividual pain variability.

1.7. Which intraindividual variability measure is best to use and report?

As we have noted, various indices of intraindividual variability are available. Unfortunately, there is no consensus regarding which variability measure is the most meaningful one to use and report. Jahng et al.[38] proposed that researchers focus on using the MSSD as a global index of temporal instability of fluctuations and use the rest of the indices (ie, iSD, autocorrelation, and the PAC) as supplements. However, this approach has some important limitations. First, the interpretation of the MSSD is quite challenging because, as noted above, it is a function of the relation between iSD and autocorrelation. Second, as demonstrated in their study (and also in our example below), the MSSD is strongly correlated with iSD and PAC.[38] This raises a concern about whether the MSSD is distinguishable from iSD and PAC. Finally, related to the second point, it is also difficult to include iSD and PAC in the same model as supplements because of potential multicollinearity.

Having reviewed the pros and cons of each intraindividual variability index, we cautiously note the relative advantages of using both iSD^2 and autocorrelation derived from DSEM. First, both are the most commonly used intraindividual variability indices and have intuitive appeal for both researchers and clinicians. Second, using the new DSEM approach, these 2 indices can be calculated easily and used as both predictor and outcome variables. Moreover, calculating the MSSD and PAC is not yet available through DSEM. Third, by using the DSEM framework, both indices are estimated for latent variables free from measurement errors and, thus, should be more reliable. Fourth, the correlation between iSD and autocorrelation tends to be weak (see Ref. 38 and also from our example below). This reduces concerns over multicollinearity when using these values as predictors in the same model. Moreover, the weak association between these measures suggests that they reflect different aspects of variability. As more is learned about the variability factors they represent, the use of both measures may provide greater insights when interpreting findings.
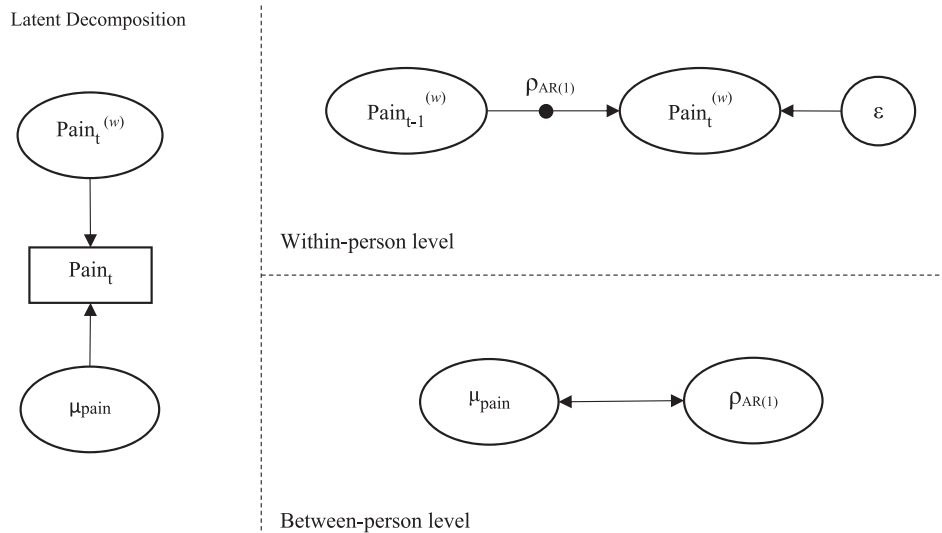
Latent Decomposition

Figure 2. Illustration of autoregressive lag 1 [AR(1)] dynamic structural equation model. t = measurement occasion; (w) = within-person estimate (ie, pain level of an individual at occasion t) decomposed by latent centering; μ = between-person estimates (ie, within-person mean of pain) decomposed by latent centering; filled black dot = random effect of AR(1); ε = innovation/dynamic error. This figure was adapted from Hamaker et al.[32]

Nonetheless, to achieve concrete recommendations regarding which variability index is the best one to report under various situations, both simulation studies and empirical studies that longitudinally examine the predictive validity of each intraindividual pain variability are required. This is one of the primary reasons that we provide the statistical syntax for calculating a range of intraindividual variability indices—so that future researchers can test the validity of these measures.

1.8. A real-world example of an investigation of intraindividual pain variability

To provide an example of how to apply the above-mentioned indices of intraindividual variability using real-world data, we conducted secondary analysis using previously collected daily diary data (see Ref. 27 for more detail) from individuals with rheumatoid arthritis (RA).

2. Methods

2.1. Participants

Individuals with RA (N = 231) participated in a daily diary study. The study inclusion criteria were (1) age 18 years or older; (2) written confirmation of RA diagnosis from a rheumatologist; (3) not taking any cyclical estrogen replacement therapy medication; (4) not diagnosed with systematic lupus erythematosus; and (5) not pursuing litigation related to their pain condition. A large proportion of participants were women (70%), Caucasian (89%), and middle-aged, with approximately 64% of the sample being either married or living with a partner. The average duration of RA was 11.5 years for females and 13.6 years for males.

2.2. Procedures

Participants completed a 30-day paper-and-pencil daily diary, a method that is comparable with that of electronic diaries.[7,89] They completed a diary each evening 30 minutes before going to bed and deposited the completed diary in the next day's mail using the postage-paid envelopes. Most of the diary reports were returned in a timely manner (ie, 82.3% of diaries were mailed on the next morning). The diary completion rate was 97% (ie, 6708 of 6930 possible observations were completed across the 30 days). Participants completed an average of 29.31 of 30 diaries (SD = 1.64; range 18-30).

2.3. Measures

2.3.1. Pain intensity

Daily average pain intensity was measured using a 0 (*No pain*) to 100 (*Pain as bad as it can be*) standard NRS.[40]

2.3.2. Depressive symptoms

The level of daily depressive symptoms was measured using 5 items assessing common symptoms of depression drawn from the Patient Health Questionnaire.[50] Items were rated on a 3-point scale (1 = no, 2 = yes, slightly, and 3 = yes, very much). This daily measure of depressive symptoms was used and validated in previous daily diary studies.[10] In this study, we obtained the mean depressive level for each participant to examine its relationship with level and variability of pain.

2.4. Data analytic plan

First, the mean, MSSD, and PAC for each participant were calculated using raw pain ratings. For calculating the mean, all nonmissing pain ratings were used. When calculating the MSSD, we used all pairs of nonmissing adjacent observations. For the PAC, a change in magnitude greater than 20 (on a 0-100 NRS of pain intensity) between the 2 adjacent pain ratings was defined as an acute change following the IMMPACT guidance for defining meaningful clinical change in pain experience,[20] and all pairs of nonmissing adjacent observations were used for calculation. iSD^2 (intraindividual variance) and AR(1) for each participant were derived by using DSEM in Mplus Version 8.2.[61] To obtain participant-specific iSD^2 values, we fit the multilevel model for heterogeneous variances.[37,72] Note that Mplus provides log-transformed variances [log (iSD^2)], which we used for subsequent analyses. We then fit the

multilevel AR(1) model with random innovation variances (see Hamaker et al's.[32] model 2) to obtain participant-specific AR(1) values. Calculating the PAC and MSSD is not yet feasible in the DSEM framework. Hence, to examine relations among variability indices, mean pain level, and depressive symptom severity, we exported factor scores of log (iSD^2) and AR(1) from Mplus and merged them with other measures into a single data file. The statistical syntax that was used to derive the 4 different intraindividual variability indices is available as an supplemental materials (available at http://links.lww.com/PAIN/A816).

Second, to detrend the data, we controlled for the linear effect of time when we fit the models in Mplus by including time as a within-person level predictor. We considered the linear trend simply for illustrative purposes. As discussed above, handling trends is complex and an appropriate detrending method should be carefully chosen. Future methodological studies should investigate and provide guidance for pain researchers regarding the best method(s) for detrending their intensive longitudinal data.

Third, to obtain person-specific autocorrelations and variances, it was necessary to use Bayesian estimation in Mplus, assuming data were MAR. Based on Hamaker et al.,[32] we used 50,000 MCMC iterations with a thinning of 10 iterations (ie, only 1 of 10 iterations was used), and thus, our results are based on a total of 5000 iterations. We ensured model convergence by checking the proportional scale reduction and trace plots, which showed absence of trends, spikes, or other irregular patterns. The Mplus codes that we used are provided in supplemental materials (available at http://links.lww.com/PAIN/A816).

Fourth, we provide an illustration of the pain ratings of some participants with different patterns of change in pain over time. Specifically, we contrasted the pain ratings of participants with different levels of each intraindividual variability measure.

Last, using zero-order correlations, we examined associations among mean pain level, intraindividual variability indices, and depressive symptom severity. Before conducting correlational analyses, we log-transformed the MSSD because it was substantially skewed and leptokurtic, as shown in **Table 1**. Note that to deal with 0 MSSD values, we log-transformed MSSD + 1 instead of MSSD. Then, using partial correlations, we further examined the associations between each of the pain variability indices [log(iSD^2), AR(1), log(MSSD), and PAC] and depressive symptom severity while controlling for the mean pain level.

3. Results

Table 1 shows the basic descriptive statistics for the mean pain level and intraindividual pain variability indices. **Figure 3** illustrates the pain ratings of 10 participants with different patterns of intraindividual pain variability. Panels (A) and (B) contrast the pain ratings of 2 participants who have different mean levels but similar

profiles of the 4 variability measures. On average, the pain ratings in panel (B) are higher than those shown in panel (A). Panels (C) and (D) show the pain ratings of 2 participants who have similar mean levels but different iSD levels. The pain ratings in panel (D) fluctuate with a higher amplitude around the linear trend compared with those in panel (C). Panels (E) and (F) contrast the pain ratings of 2 participants who had similar mean levels but different levels of autocorrelation (time dependency). The participant in panel (F) has a high and positive AR(1). For this participant, being above (or below) the linear trend of pain at a time point is likely to be followed by being above (or below) the linear trend of pain at an immediately subsequent time point. By contrast, the participant in panel (E) manifests a low positive AR(1). For this participant, being above (or below) the linear trend of pain at a time point is not systematically related to being above (or below) the linear trend of pain at the next time point. Panels (G) and (H) contrast the pain ratings of 2 participants who have similar mean levels but different MSSD levels. The participant in panel (G) shows small successive changes, whereas the participant in panel (H) has greater and more frequent successive changes. Panels (I) and (J) show the pain ratings of 2 participants who have similar mean levels but different PAC levels. The pain ratings of the participant in panel (I) do not show a sudden change in any of the consecutive observations, yielding PAC = 0. However, the pain ratings of the participant in panel (J) show a sudden jump or drop greater than 20 more frequently, yielding PAC = 0.52.

Table 2 presents the zero-order correlations among the 5 pain measures. Correlation values lower than 0.4 indicate "weak" correlation, whereas values between 0.4 and 0.59 represent "moderate" correlation, and values greater than 0.6 indicate "strong" correlation.[24] The mean pain level was weakly-to-moderately and positively correlated with log(iSD^2), log(MSSD), and PAC. In other words, participants with a higher (or lower) mean pain level tended to experience a higher (or lower) amplitude of pain fluctuations and more (or less) frequent successive and acute pain changes. However, the mean pain level and autocorrelation (temporal dependency of pain) were very weakly associated. This result suggests that mean pain level and variability indices may capture different aspects of pain dynamics in this example. The 3 variability measures, log (iSD^2), log (MSSD), and the PAC, were strongly and positively correlated. However, AR(1) was moderately correlated with log(iSD^2), weakly correlated with log(MSSD), and very weakly correlated with the PAC. This result suggests that AR(1) may reflect different aspects of pain dynamics than those reflected by the other 3 variability measures in this example.

As shown in **Table 3**, zero-order correlations, mean pain level, and all intraindividual pain variability measures are positively related to mean depressive symptom severity. To examine whether variability measures are uniquely associated with depressive symptoms beyond the mean pain level, we also

Descriptive statistics for the 5 pain measures (n = 231).

	Mean pain level	Log(iSD^2)	AR(1)	MSSD	Log(MSSD)	PAC
Mean	35.09	4.60	0.38	234.93	4.83	0.19
SD	18.41	1.26	0.14	243.72	1.36	0.17
Minimum	0.21	−0.27	0.02	0.00	0.00	0.00
Maximum	85.27	6.76	0.79	1536.11	7.34	0.79
Skewness	0.21	−1.00	0.30	2.04	−1.07	0.67
Kurtosis	−0.70	1.29	0.08	5.90	1.47	−0.24

MSSD, mean square of successive difference; PAC, probability of acute change.
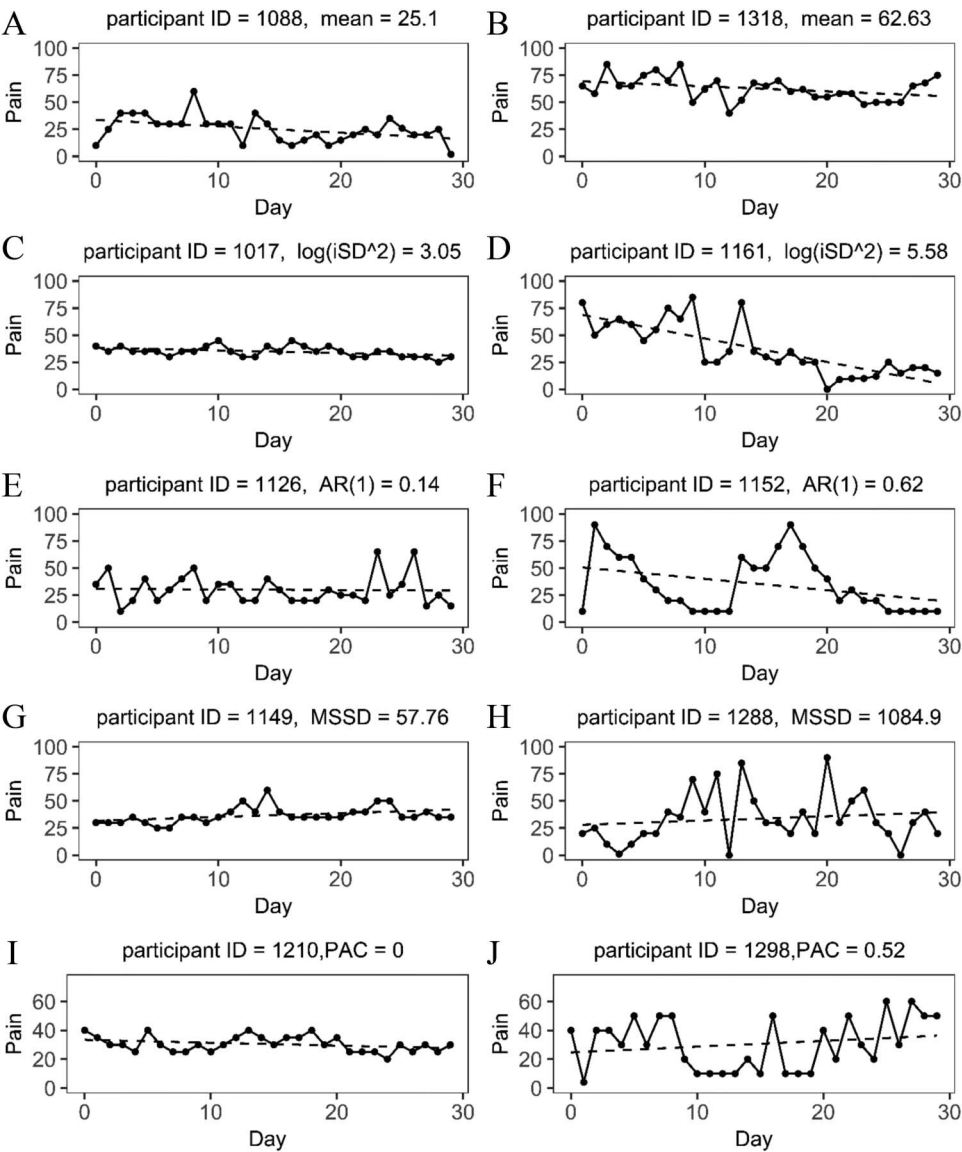
Figure 3. Illustration of the pain ratings of 10 participants with different patterns of intraindividual pain variability over time. The dots indicate the observed pain ratings, and the dotted lines indicate linear trends. MSSD, mean square of successive difference; PAC, probability of acute change. (A) a participant with similar pain variability as participant B but with low mean pain level; (B) a participant with similar pain variability as participant A but with high mean pain level; (C) a participant with similar mean pain level as participant D but with low iSD; (D) a participant with similar mean pain level as participant C but with high iSD; (E) a participant with similar mean pain level as participant F but with low first order autocorrelation; (F) a participant with similar mean pain level as participant E but with high first order autocorrelation; (G) a participant with similar mean pain level as participant H but with low MSSD; (H) a participant with similar mean pain level as participant G but with high MSSD; (I) a participant with similar mean pain level as participant J but with low PAC; (J) a participant with similar mean pain level as participant I but with high PAC.

calculated partial correlations between each of the 4 variability indices and depressive symptom severity while controlling for the mean pain level. As shown in **Table 3**, all of the 4 variability indices

were somewhat weakly correlated with depressive symptom severity even after accounting for the correlation between mean pain level and depressive symptom severity. This result implies that each of the 4 variability indices may further explain individual differences in depressive symptom severity above and beyond the variability in depressive symptom severity that is explained by the mean pain level.

3.1. Future research directions

Because advances in technology now permit the collection of intensive longitudinal data, researchers are far more capable of examining the dynamic ebb and flow of an individual's pain experience, as well as the functional domains (eg, pain interference, disability, and emotional distress) associated with changes in that experience. Investigating intraindividual pain

Table 2

Zero-order correlations among the 5 pain measures.

	Mean pain level	Log(iSD^2)	AR(1)	Log(MSSD)	PAC
Mean pain level	1.00				
Log(iSD^2)	0.42***	1.00			
AR(1)	0.09	0.42***	1.00		
Log(MSSD)	0.42***	0.92***	0.16	1.00	
PAC	0.34***	0.76***	−0.01	0.82***	1.00

***P < 0.001, **P < 0.01, and *P < 0.05.
MSSD, mean square of successive difference; PAC, probability of acute change.

Zero-order correlations between depressive symptoms, mean pain level, and pain variability measures and partial correlations between depressive symptoms and the 4 pain variability measures controlling for the mean pain level.

	Mean pain level	Log(iSD^2)	AR(1)	Log(MSSD)	PAC
Depressive symptoms					
Zero-order correlations	0.40***	0.32***	0.20*	0.29***	0.25***
Partial correlations	N/A	0.18**	0.17**	0.15*	0.13*

***$P < 0.001$, **$P < 0.01$, and *$P < 0.05$.
MSSD, mean square of successive difference; PAC, probability of acute change.

variability, a process that can be quite easily computed from intensive longitudinal data, has the potential to further expand our nascent understanding of pain processing, coping, and adjustment, as well as treatment response. Below, we discuss a number of important and viable future research avenues built around the prior literature and existing theoretical models.

3.1.1. Predictors and consequences of pain variability differences

An important research direction would be to examine the predictors and consequences of individual differences in intraindividual pain variability.[72,97] First, diverse biopsychosocial factors, such as chronic pain conditions, sex, race, age, comorbid psychopathology (eg, major depressive disorder or generalized anxiety disorder), genetic polymorphisms, sleep quality, personality, social support system, and financial stress, may serve important roles in influencing individuals' pain variation. Very few studies have investigated how some of these factors (eg, depression, genetic polymorphisms, and acute low back pain) are associated with pain variability.[57,86,97] Exploring more factors such as these may provide new insights into how clinicians can optimize their treatments and thereby help individuals achieve a stronger sense of pain control.

Second, individual differences in pain variability may serve as important predictors of subsequent physical and psychosocial functioning in addition to pain intensity and pain interference levels. Previous studies have suggested that highly variable pain may increase the unpredictability of the pain experience and, thus, could potentially lower individuals' sense of control over pain and increase distress and helplessness.[1,65,86] Most previous studies that examined the association between pain variability and physical/psychosocial functioning relied on a cross-sectional design, a practice that makes it difficult to interpret the direction of the obtained associations. Researchers should consider using a more appropriate research design, such as *measurement burst design*, to address this gap in the literature. Measurement burst designs incorporate "bursts" of intensive longitudinal assessments within a short period of time (eg, multiple times per day for 2 weeks) that is repeated longitudinally with wider time intervals (eg, every 6 months for 2 years).[67,79] Such a research design would permit a more nuanced understanding of the predictors and consequences of pain variability over time.

3.1.2. Intraindividual pain variability as an end point in clinical trials

Most meta-analyses and systematic reviews reveal that evidence-based nonpharmacological interventions for chronic pain such as Cognitive-Behavioral Therapy, Acceptance-Commitment Therapy, and Mindfulness-Based Stress Reduction show small-to-moderate effect sizes in reducing pain severity.[3,17,29,36,69,92] As pointed out by Stone et al.,[84] evaluating pre–post intervention improvements in pain intensity through one-time retrospective assessment may not capture the true efficacy of those

interventions because of recall error and bias. It is possible that study participants who did not benefit from a treatment could be those who had higher pain variability during the treatment owing to their tendency to focus more on high pain experiences during pain recall.[84] In addition to using the average of momentary assessments of pain, investigating pre–post intervention changes in intraindividual pain variabilities may provide even more useful information on the effects of an intervention. For example, one of the key elements of Cognitive-Behavioral Therapy for chronic pain involves assisting individuals to successfully identify and alter cognitive distortions (eg, "My pain will never end") that emerge when they experience sudden elevations in pain. Cognitive-Behavioral Therapy for chronic pain also targets helping individuals to develop adaptive activity pacing strategies so that they can prevent themselves from experiencing unnecessary injury or exacerbation of pain when their pain is less severe. Mindfulness-based interventions aid individuals to detach from automatic reactions (eg, thoughts, interpretations, and behaviors) to momentary changes in pain. Thus, it would appear that a potentially important therapeutic mechanism of chronic pain intervention involves training people to effectively manage pain fluctuations, over and above helping people to manage their pain severity. However, to the best of our knowledge, (1) treatment-related changes in pain variability and (2) treatment-related changes in the associations between pain variability and coping have not yet been examined in pain intervention clinical trials. This knowledge gap could be addressed by the use of ILD to assess pain intensity and pain-coping responses before, during, and after the treatment interventions in the context of randomized control trials.[88] A handful of previous studies have implemented this type of clinical design[14,98] but have yet to examine pain variability as an outcome.

3.1.3. Intraindividual pain variability as a predictor of treatment response

Previous studies have demonstrated that individual differences in baseline pain variability serve as an important predictor of treatment response in medication trials.[25,35,56] In fact, the IMMPACT recently suggested that pain variability is one of the main phenotypic characteristics that can advance personalized precision pain medicine.[21] To the best of our knowledge, however, all previous studies that investigated the predictive effect of pain variability on treatment response have been medication trials. It is therefore important to replicate and extend these findings to clinical trials of psychosocial pain interventions. Based on previous findings from medication trials, we speculate that individuals with higher pain variability at baseline are less likely to benefit from psychosocial pain interventions.

We may also consider identifying potentially meaningful subgroups of individuals with distinct pain variability patterns through advanced statistical approaches, such as finite mixture modeling or machine learning. Along these lines, Fillingim et al.[26] have suggested that examining pain variability could facilitate

more accurate classification of chronic pain conditions. Different patterns of pain variability may serve as clinically important transdiagnostic features of chronic pain. Examining how these potential subgroups respond differently to pain treatment over and above other baseline phenotypes would be an important future research direction in precision pain medicine.

4. Conclusion

The present review provided a nontechnical overview of a range of intraindividual variability indices (ie, iSD, autocorrelation, MSSD, and PAC), important methodological issues, and recent development in DSEM, as well as a data-based example from a secondary analysis of previously published data. Although we cautiously note some strengths of using both iSD^2 and autocorrelation obtained by using the DSEM framework over other variability indices, the determination of which variability index is best still requires further support from future simulations and empirical studies. A number of methodological issues also need to be addressed in future studies, such as adequately handling missing data, determining adequate sample size and the number of assessments required for measuring pain variability, and how best to detrend data.

Although there remain a number of additional methodological issues to consider, the evaluation of intraindividual pain variability in future studies of chronic pain can be a productive focus of investigation. Specifically, identifying factors that shape pain variability and determining how pain variability influences individuals' physical and emotional functioning are viable future research objectives that could provide insights in our effort to help individuals successfully adjust to chronic pain. Furthermore, testing whether changes in intraindividual pain variability can serve as a main end point in clinical trials and whether pain variability modulates efficacy of treatment outcomes over and above other phenotypes will be particularly important in improving existing pain interventions and in enhancing precision pain medicine.

Conflict of interest statement

The authors have no conflicts of interest to declare.

Acknowledgements

Appendix A. Supplemental digital content

Supplemental digital content associated with this article can be found online at http://links.lww.com/PAIN/A817 and http://links.lww.com/PAIN/A816.

References

[1] Allen KD. The value of measuring variability in osteoarthritis pain. J Rheumatol 2007;34:2132–3.

[2] Asparouhov T, Hamaker EL, Muthén B. Dynamic structural equation models. Struct Equ Model 2018;25:359–88.

[3] Astin JA, Beckner W, Soeken K, Hochberg MC, Berman B. Psychological interventions for rheumatoid arthritis: a meta-analysis of randomized controlled trials. Arthritis Care Res 2002;47:291–302.

[4] Baird BM, Le K, Lucas RE. On the nature of intraindividual personality variability: reliability, validity, and associations with well-being. J Pers Soc Psychol 2006;90:512–27.

[5] Boker SM, Laurenceau JP. Dynamical systems modeling: an application to the regulation of intimacy and disclosure in marriage. In: Walls TA, Schafer JL, editors. Models for intensive longitudinal data. New York: Oxford University Press, 2006. p. 195–218.

[6] Bolger N, Laurenceau JP. Intensive longitudinal methods: an introduction to diary and experience sampling research. New York: Guilford Press, 2013.

[7] Bolger N, Shrout PE, Green AS, Rafaeli E, Reis HT. Paper or plastic revisited: let's keep them both–Reply to Broderick and Stone (2006); Tennen, Affleck, Coyne, Larsen, and DeLongis (2006); and Takarangi, Garry, and Loftus (2006). Psychol Methods 2006;11:123–5.

[8] Broderick JE, Schwartz JE, Vikingstad G, Pribbernow M, Grossman S, Stone AA. The accuracy of pain and fatigue items across different reporting periods. PAIN 2008;139:146–57.

[9] Cleveland WS, Devlin SJ. Locally weighted regression: an approach to regression analysis by local fitting. J Am Stat Assoc 1988;83:596–610.

[10] Conner TS, Tennen H, Zautra AJ, Affleck G, Armeli S, Fifield J. Coping with rheumatoid arthritis pain in daily life: within-person analyses reveal hidden vulnerability for the formerly depressed. PAIN 2006;126:198–209.

[11] Crocker L, Algina J. Introduction to classical and modern test theory. New York: Holt, Rinehart & Winston, 1986.

[12] Cryer JD, Chan KS. Time series analysis with applications to R. 2nd ed. New York: Springer, 2008.

[13] Davin S, Wilt J, Covington E, Scheman J. Variability in the relationship between sleep and pain in patients undergoing interdisciplinary rehabilitation for chronic pain. Pain Med 2014;15:1043–51.

[14] Davis MC, Zautra AJ. An online mindfulness intervention targeting socioemotional regulation in fibromyalgia: results of a randomized controlled trial. Ann Behav Med 2013;46:273–84.

[15] Dickey DA, Fuller WA. Distribution of the estimators for autoregressive time series with a unit root. J Am Stat Assoc 1979;74:427–31.

[16] Dickey DA, Fuller WA. Likelihood ratio statistics for autoregressive time series with a unit root. Econom J Econom Soc 1981;1057–72.

[17] Dixon KE, Keefe FJ, Scipio CD, Perri LM, Abernethy AP. Psychological interventions for arthritis pain management in adults: a meta-analysis. Health Psychol 2007;26:241–50.

[18] Driver CC, Oud JHL, Voelkle MC. Continuous time structural equation modeling with R package ctsem. J Stat Softw 2017;77:1–35.

[19] Du H, Wang L. Reliabilities of intraindividual variability indicators with autocorrelated longitudinal data: implications for longitudinal study designs. Multivariate Behav Res 2018;53:502–20.

[20] Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, Haythornthwaite JA, Jensen MP, Kerns RD, Ader DN. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain 2008;9:105–21.

[21] Edwards RR, Dworkin RH, Turk DC, Angst MS, Dionne R, Freeman R, Hansson P, Haroutounian S, Arendt-Nielsen L, Attal N. Patient phenotyping in clinical trials of chronic pain treatments: IMMPACT recommendations. PAIN 2016;157:1851–71.

[22] Eid M, Diener E. Intraindividual variability in affect: reliability, validity, and personality correlates. J Pers Soc Psychol 1999;76:662–76.

[23] Enders CK, Keller BT, Levy R. A fully conditional specification approach to multilevel imputation of categorical and continuous variables. Psychol Methods 2017;23:298–317.

[24] Evans JD. Straightforward statistics for the behavioral sciences. Belmont, CA: Thomson Brooks/Cole Publishing Co, 1996.

[25] Farrar JT, Troxel AB, Haynes K, Gilron I, Kerns RD, Katz NP, Rappaport BA, Rowbotham MC, Tierney AM, Turk DC, Dworkin RH. Effect of variability in the 7-day baseline pain diary on the assay sensitivity of neuropathic pain randomized clinical trials: an ACTTION study. PAIN 2014;155:1622–31.

[26] Fillingim RB, Loeser JD, Baron R, Edwards RR. Assessment of chronic pain: domains, methods, and mechanisms. J Pain 2016;17:T10–20.

[27] Finan P, Zautra A, Tennen H. Daily diaries reveal influence of pessimism and anxiety on pain prediction patterns. Psychol Health 2008;23:551–68.

[28] Gavaruzzi T, Carnaghi A, Lotto L, Rumiati R, Meggiato T, Polato F, De Lazzari F. Recalling pain experienced during a colonoscopy: pain expectation and variability. Br J Health Psychol 2010;15:253–64.

[29] Glombiewski JA, Sawyer AT, Gutermann J, Koenig K, Rief W, Hofmann SG. Psychological treatments for fibromyalgia: a meta-analysis. PAIN 2010;151:280–95.

[30] Gollob HF, Reichardt CS. Taking account of time lags in causal models. Child Dev 1987;58:80–92.

[31] de Haan-Rietdijk S, Voelkle MC, Keijsers L, Hamaker EL. Discrete- vs. continuous-time modeling of unequally spaced experience sampling method data. Front Psychol 2017;8:1–19.

[32] Hamaker EL, Asparouhov T, Brose A, Schmiedek F, Muthén B. At the frontiers of modeling intensive longitudinal data: dynamic structural equation models for the affective measurements from the COGITO study. Multivariate Behav Res 2018;3171:1–22.

[33] Hamaker EL, Ceulemans E, Grasman R, Tuerlinckx F. Modeling affect dynamics: state of the art and future challenges. Emot Rev 2015;7:316–22.

[34] Hardy J, Segerstrom SC. Intra-individual variability and psychological flexibility: affect and health in a National US sample. J Res Pers 2017;69:13–21.

[35] Harris RE, Williams DA, McLean SA, Sen A, Hufford M, Gendreau RM, Gracely RH, Clauw DJ. Characterization and consequences of pain variability in individuals with fibromyalgia. Arthritis Rheum 2005;52:3670–4.

[36] Hoffman B, Papas R, Chatkoff D, Kerns R. Meta-analysis of psychological interventions for chronic low back pain. Health Psychol 2007;26:1–9.

[37] Hoffman L. Multilevel models for examining individual differences in within-person variation and covariation over time. Multivariate Behav Res 2007;42:609–29.

[38] Jahng S, Wood PK, Trull TJ. Analysis of affective instability in ecological momentary assessment: indices using successive difference and group comparison via multilevel modeling. Psychol Methods 2008;13:354–75.

[39] Jensen MP, Karoly P. Self-report scales and procedures for assessing pain in adults. Handbook of pain assessment. New York: The Guilford Press, 1992.

[40] Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. PAIN 1986;27:117–26.

[41] Jongerling J, Laurenceau JP, Hamaker EL. A multilevel AR(1) model: allowing for inter-individual differences in Trait-scores, inertia, and innovation variance. Multivariate Behav Res 2015;50:334–49.

[42] Kalmijn W, Veenhoven R. Measuring inequality of happiness in nations: in search for proper statistics. J Happiness Stud 2005;6:357–96.

[43] Karoly P, Okun MA, Enders C, Tennen H. Effects of pain intensity on goal schemas and goal pursuit: a daily diary study. Health Psychol 2014;33:968–76.

[44] Kashdan TB, Rottenberg J. Psychological flexibility as a fundamental aspect of health. Clin Psychol Rev 2010;30:865–78.

[45] Kelava A, Brandt H. A nonlinear dynamic latent class structural equation model. Struct Equ Model 2019;00:1–20.

[46] Keller BT, Enders CK. Blimp user's manual (version 1.1). Los Angeles, CA, 2018.

[47] Kikuchi H, Yoshiuchi K, Miyasaka N, Ohashi K, Yamamoto Y, Kumano H, Kuboki T, Akabayashi A. Reliability of recalled self-report on headache intensity: investigation using ecological momentary assessment technique. Cephalalgia 2006;26:1335–43.

[48] Koval P, Butler EA, Hollenstein T, Lanteigne D, Kuppens P. Emotion regulation and the temporal dynamics of emotions: effects of cognitive reappraisal and expressive suppression on emotional inertia. Cogn Emot 2015;29:831–51.

[49] Koval P, Kuppens P, Allen NB, Sheeber L. Getting stuck in depression: the roles of rumination and emotional inertia. Cogn Emot 2012;26:1412–27.

[50] Kroenke K, Spitzer RL, Williams JBW. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med 2002;64:258–66.

[51] Kuppens P, Allen NB, Sheeber LB. Emotional inertia and psychological maladjustment. Psychol Sci 2010;21:984–91.

[52] Lefebvre JC, Keefe FJ. The effect of neuroticism on the recall of persistent low-back pain and perceived activity interference. J Pain 2013;14:948–56.

[53] Leiderman PH, Shapiro D. Application of a time series statistic to physiology and psychology. Science 1962;138:141–2.

[54] Litcher-Kelly L, Stone AA, Broderick JE, Schwartz JE. Associations among pain intensity, sensory characteristics, affective qualities, and activity limitations in patients with chronic pain: a momentary, within-person perspective. J Pain 2004;5:433–9.

[55] Liu Y, West SG. Weekly cycles in daily report data: an overlooked issue. J Pers 2016;84:560–79.

[56] Martini CH, Yassen A, Krebs-Brown A, Passier P, Stoker M, Olofsen E, Dahan A. A novel approach to identify responder subgroups and predictors of response to low-and high-dose capsaicin patches in postherpetic neuralgia. Eur J Pain 2013;17:1491–501.

[57] Martire LM, Wilson SJ, Small BJ, Conley YP, Janicki PK, Sliwinski MJ. COMT and OPRM1 genotype associations with daily knee pain variability and activity induced pain. Scand J Pain 2016;10:6–12.

[58] Merskey H, Bogduk N. Classification of chronic pain, IASP Task Force on Taxonomy. Seattle: WA Int. Assoc. Study Pain Press, 1994.

[59] Mestdagh M, Pe M, Pestman W, Verdonck S, Kuppens P, Tuerlinckx F. Sidelining the mean: the relative variability index as a generic mean-corrected variability measure for bounded variables. Psychol Methods 2018;23:690–707.

[60] Mun CJ, Karoly P, Okun MA. Effects of daily pain intensity, positive affect, and individual differences in pain acceptance on work goal interference and progress. PAIN 2015;156:2276–85.

[61] Muthén LK, Muthén BO. Mplus user's guide. 8th ed. Los Angeles, CA: Muthén & Muthén, 2017.

[62] Nesselroade JR, Salthouse TA. Methodological and theoretical implications of intraindividual variability in perceptual-motor performance. J Gerontol B Psychol Sci Soc Sci 2004;59:P49–55.

[63] Von Neumann J. Distribution of the ratio of the mean square successive difference to the variance. Ann Math Stat 1941;12:367–95.

[64] Oud JHL. Continuous time modeling of reciprocal relationships in the cross-lagged panel design. Data Anal Tech Dyn Syst 2007:87–129.

[65] Parry E, Ogollah R, Peat G. Significant pain variability in persons with, or at high risk of, knee osteoarthritis: preliminary investigation based on secondary analysis of cohort data. BMC Musculoskelet Disord 2017;18:1–11.

[66] Peters ML, Sorbi MJ, Kruise DA, Kerssens JJ, Verhaak PFM, Bensing JM. Electronic diary assessment of pain, disability and psychological adaptation in patients differing in duration of pain. PAIN 2000;84:181–92.

[67] Ram N, Gerstorf D. Time-structured and net intraindividual variability: tools for examining the development of dynamic characteristics and processes. Psychol Aging 2009;24:778–91.

[68] Ram N, Rabbitt P, Stollery B, Nesselroade JR. Cognitive performance inconsistency: intraindividual change and variability. Psychol Aging 2005;20:623–33.

[69] Reiner K, Tibi L, Lipsitz JD. Do mindfulness-based interventions reduce pain intensity? A critical review of the literature. Pain Med 2013;14:230–42.

[70] Rovine MJ, Walls TA. Multilevel autoregressive modeling of interindividual differences in the stability of a process. In: Walls TA, Schafer JL, editors. Models for intensive longitudinal data. New York: Oxford University Press, 2006. p. 124–47.

[71] Sandhu SS, Leckie G. Orthodontic pain trajectories in adolescents: between-subject and within-subject variability in pain perception. Am J Orthod Dentofacial Orthop 2016;149:491–500.

[72] Schneider S, Junghaenel DU, Keefe FJ, Schwartz JE, Stone AA, Broderick JE. Individual differences in the day-to-day variability of pain, fatigue, and well-being in patients with rheumatic disease: associations with psychological variables. PAIN 2012;153:813–22.

[73] Schneider S, Junghaenel DU, Ono M, Stone AA. Temporal dynamics of pain: an application of regime-switching models to ecological momentary assessments in patients with rheumatic diseases. PAIN 2018;159:1346–58.

[74] Schultzberg M, Muthén B. Number of subjects and time points needed for multilevel time-series analysis: a simulation study of dynamic structural equation modeling. Struct Equ Model 2018;25:495–515.

[75] Segerstrom SC, Sephton SE, Westgate PM. Intraindividual variability in cortisol: approaches, illustrations, and recommendations. Psychoneuroendocrinology 2017;78:114–24.

[76] Shiffman S. Ecological momentary assessment (EMA) in studies of substance use. Psychol Assess 2009;21:486–97.

[77] Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1–32.

[78] Shumway RH, Stoffer DS. Time series analysis and its applications: with R examples. New York, NY: Springer, 2017.

[79] Sliwinski MJ. Measurement-burst designs for social health research. Soc Personal Psychol Compass 2008;2:245–61.

[80] Spearman C. The proof and measurement of association between two things. Am J Psychol 1904;15:72–101.

[81] Stone AA, Shiffman S, Atienza A, Nebeling L. The science of real-time data capture: self-reports in health research. New York, NY: Oxford University Press, 2007.

[82] Stone AA, Broderick JE. Real-time data collection for pain: appraisal and current status. Pain Med 2007;8(suppl 3):S85–93.

[83] Stone AA, Broderick JE, Shiffman SS, Schwartz JE. Understanding recall of weekly pain from a momentary assessment perspective: absolute agreement, between- and within-person consistency, and judged change in weekly pain. PAIN 2004;107:61–9.

[84] Stone AA, Schwartz JE, Broderick JE, Shiffman SS. Variability of momentary pain predicts recall of weekly pain: a consequence of the peak (or salience) memory heuristic. Personal Soc Psychol Bull 2005;31:1340–6.

[85] Suls J, Green P, Hillis S. Emotional reactivity to everyday problems, affective inertia, and neuroticism. Personal Soc Psychol Bull 1998;24:127–36.

[86] Suri P, Rainville J, Fitzmaurice GM, Katz JN, Jamison RN, Martha J, Hartigan C, Limke J, Jouve C, Hunter DJ. Acute low back pain is marked by variability: an internet-based pilot study. BMC Musculoskelet Disord 2011;12:220.

[87] Taylor SS, Davis MC, Yeung EW, Zautra AJ, Tennen HA. Relations between adaptive and maladaptive pain cognitions and within-day pain exacerbations in individuals with fibromyalgia. J Behav Med 2017;40:458–67.

[88] Tennen H, Affleck G, Armeli S, Carney MA. A daily process approach to coping: linking theory, research, and practice. Am Psychol 2000;55:626–36.

[89] Tennen H, Affleck G, Coyne JC, Larsen RJ, DeLongis A. Paper and plastic in daily diary research: comment on Green, Rafaeli, Bolger, Shrout, and Reis (2006). Psychol Methods 2006;11:112–18.

[90] Treede RD, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, Cohen M, Evers S, Finnerup NB, First MB. A classification of chronic pain for ICD-11. PAIN 2015;156:1003–7.

[91] Trull TJ, Lane SP, Koval P, Ebner-Priemer UW. Affective dynamics in psychopathology. Emot Rev 2015;7:355–61.

[92] Veehof MM, Oskam MJ, Schreurs KMG, Bohlmeijer ET. Acceptance-based interventions for the treatment of chronic pain: a systematic review and meta-analysis. PAIN 2011;152:533–42.

[93] Wang LP, Hamaker E, Bergeman CS. Investigating inter-individual differences in short-term intra-individual variability. Psychol Methods 2012;17:567–81.

[94] Wang LP, Maxwell SE. On disaggregating between-person and within-person effects with longitudinal data using multilevel models. Psychol Methods 2015;20:63–83.

[95] Wheeler L, Reis HT. Self-recording of everyday life events: origins, types, and uses. J Pers 1991;59:339–54.

[96] Wideman TH, Edwards RR, Walton DM, Martel MO, Hudon A, Seminowicz DA. The multimodal assessment model of pain: a novel framework for further integrating the subjective pain experience within research and practice. Clin J Pain 2019;35:212–21.

[97] Zakoscielna KM, Parmelee PA. Pain variability and its predictors in older adults: depression, cognition, functional status, health, and pain. J Aging Health 2013;25:1329–39.

[98] Zautra AJ, Davis MC, Reich JW, Nicassario P, Tennen H, Finan P, Kratz A, Parrish B, Irwin MR. Comparison of cognitive behavioral and mindfulness meditation interventions on adaptation to rheumatoid arthritis for patients with and without history of recurrent depression. J Consult Clin Psychol 2008;76:408–21.