Open Journal of Statistics, 2024, 14, 492-517
https://www.scirp.org/journal/ojs
ISSN Online: 2161-7198
ISSN Print: 2161-718X

Adaptive Modeling of Monthly Depression Levels in Terms of Daily Assessments of Opioid Medications Taken and Pain Levels for Cancer Patients

George J. Knafl[1], Ryan Quinn[2], Andrew Robinson[2], Salimah H. Meghani[2,3]

[1]School of Nursing, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[2]Department of Biobehavioral Health Sciences, New Courtland Center for Transitions and Health, School of Nursing, University of Pennsylvania, Philadelphia, PA, USA
[3]Leonard Davis Institute for Health Economics, University of Pennsylvania, Philadelphia, PA, USA
Email: gknafl@unc.edu

How to cite this paper: Knafl, G.J., Quinn, R., Robinson, A. and Meghani, S.H. (2024) Adaptive Modeling of Monthly Depression Levels in Terms of Daily Assessments of Opioid Medications Taken and Pain Levels for Cancer Patients. *Open Journal of Statistics*, **14**, 492-517.
https://doi.org/10.4236/ojs.2024.145022

Received: September 27, 2024
Accepted: October 25, 2024
Published: October 28, 2024

Abstract

A research study collected intensive longitudinal data from cancer patients on a daily basis as well as non-intensive longitudinal survey data on a monthly basis. Although the daily data need separate analysis, those data can also be utilized to generate predictors of monthly outcomes. Alternatives for generating daily data predictors of monthly outcomes are addressed in this work. Analyses are reported of depression measured by the Patient Health Questionnaire 8 as the monthly survey outcome. Daily measures include numbers of opioid medications taken, numbers of pain flares, least pain levels, and worst pain levels. Predictors are averages of recent non-missing values for each daily measure recorded on or prior to survey dates for depression values. Weights for recent non-missing values are based on days between measurement of a recent value and a survey date. Five alternative averages are considered: averages with unit weights, averages with reciprocal weights, weighted averages with reciprocal weights, averages with exponential weights, and weighted averages with exponential weights. Adaptive regression methods based on likelihood cross-validation (LCV) scores are used to generate fractional polynomial models for possible nonlinear dependence of depression on each average. For all four daily measures, the best LCV score over averages of all types is generated using the average of recent non-missing values with reciprocal weights. Generated models are nonlinear and monotonic. Results indicate that an appropriate choice would be to assume three recent non-missing values and use the average with reciprocal weights of the first three recent non-

missing values.

Keywords

Adaptive Regression, Cancer, Depression, Intensive Longitudinal Data, Factional Polynomials, Opioid Medications, Pain Levels

1. Introduction

Large amounts of repeated measures data collected for individuals over time are called intensive longitudinal data and are analyzed using time series methods [1]. However, those data can also be used to predict non-intensively collected longitudinal outcomes for the same individuals. The Role of Opioid Adherence Profiles in Cancer Pain Self-Management and Outcomes Study (1R01NR017853) examined the long-term effects of opioid use in cancer patients. This study collected both monthly survey data and intensive longitudinal daily data on numbers of opioid medications taken and pain levels for patients. An important data analysis issue for research studies like this one is how best to integrate the intensive longitudinal daily data into analyses as predictors of the monthly outcome data.

The purpose of this article is to assess alternatives for combining daily data into effective predictors of monthly outcomes through an analysis of a specific set of data. The example outcome for this assessment is depression as measured by the Patient Health Questionnaire (PHQ) 8 [2] collected on surveys at baseline and at approximately 1- , 2- , 3- , and 5-months post-baseline. Four measures are considered based on daily data reported by participants using an app: the number of opioid medications taken, the number of pain flares, the least pain level, and the worst pain level. These daily measures are considered as predictors of depression in cancer patients.

Reported analyses consider nonlinear relationships, which can provide substantial improvements over standard linear relationships [3]. For this reason, fractional polynomial modeling [4] [5] is considered in reported analyses to address possible nonlinearity in the relationship between monthly measured depression and daily measured predictors.

2. Methods

2.1. Measures

2.1.1. The PHQ 8

The PHQ 8 is a depression scale [6] based on eight items scored from 0 = Not at all to 3 = Nearly every day. It was tested on 1165 subjects reporting an average value of 8.63 with standard deviation 5.52 and Cronbach's alpha 0.86. In the current study, the PHQ 8 was collected at baseline and at approximate 1- , 2- , 3- , and 5-months post-baseline.

2.1.2. Daily Data

Daily data during study participation are collected from participants using the mEMA app [7]. Four daily measures are collected for participants including the number of opioid medications taken, the number of pain flares, the least pain level, and the worst pain level.

Participants report the number of opioid medications taken for each type they are currently taking, which are summed to generate the number of opioids taken of all types on a given day. The other three measures have values 0-10. For the number of pain flares, a value of 10 represents 10 or more pain flares.

Daily data are not always reported on the same dates on which monthly data are collected. Also, there may be a limited number of daily measurements prior to the date on which monthly data are collected. For these reasons, incorporation of daily data into analyses of monthly data requires consideration of the possibility of missing prior daily measurements as well as how much earlier than a monthly collection date a daily measurement occurs.

2.2. Data Analyses

Depression level measured by the PHQ 8 is the longitudinal outcome for all reported analyses. Depression levels are predicted using recent non-missing daily measurements, that is, those measured on the associated survey date or on days just prior to that date, for each of the four daily measures of Section 2.1.2. Only post-baseline depression levels are used in analyses due to insufficient daily data collected prior to baseline. Depression is analyzed using linear mixed models accounting for exchangeable correlation and constant variance as used in standard repeated measures modeling. Predictors for mean depression are based on averages of recent non-missing daily measurements as described in Section 2.2.1, possibly power transformed with real-valued powers, that is, using fractional polynomials. The four daily measures can have zero values which is problematic for computing arbitrary power transforms, and so these measures are adjusted in reported analyses to have positive integer values by adding a value of 1 to the observed value. Models are evaluated using 10-fold likelihood cross-validation (LCV) scores as described in Section 2.2.2. Power transforms determining these models are generated adaptively using heuristic search through alternative powers as described in Section 2.2.3.

2.2.1. Averages of Recent Daily Measurements

Let s denote a participant and i a post-baseline survey with values 2, 3, 4, and 5 at approximate 1- , 2- , 3- , and 5-months post-baseline. Let $y_{s,i}$ denote the depression value for participant s at survey i. For j a positive integer, let $x_{s,i,j}$ denote the j^{th} most recent daily measurement for one of the four daily measures on or prior to the date of survey i for participant s. Note that $x_{s,i,j}$ are adjusted as described above to be positive integer-valued but can have missing values. Let $J_{s,i}$ denote the number of non-missing $x_{s,i,j}$ at survey i for participant s. Let $d_{s,i,j}$ denote the nonnegative number of days between the date of survey i for participant s and the date of measurement of $x_{s,i,j}$.

Restrict to surveys i for participants s satisfying $J_{s,i} \ge J_* > 1$ for some minimal integer J_*, that is, with at least J_* recent values that are non-missing. Consider the following possible means computed using the most recent non-missing daily measurements. For $1 \le k \le J_*$, the average of the k most recent non-missing daily measurements with unit weights satisfies

$$m_{1,s,i,k} = \frac{\sum_{j=1}^{k} x_{s,i,j}}{k}.$$

The average of the k most recent non-missing weighted daily measurements with reciprocal weights

$$w_{s,i,j} = \left(1 + d_{s,i,j}\right)^{-1}$$

satisfies

$$m_{2,s,i,k} = \frac{\sum_{j=1}^{k} w_{s,i,j} \cdot x_{s,i,j}}{k}.$$

The weighted average of the k most recent non-missing daily measurements with reciprocal weights satisfies

$$m_{3,s,i,k} = \frac{\sum_{j=1}^{k} w_{s,i,j} \cdot x_{s,i,j}}{\sum_{j=1}^{k} w_{s,i,j}}.$$

The average of the k most recent non-missing weighted daily measurements with exponential weights

$$w'_{s,i,j} = \exp\left(-d_{s,i,j}\right)$$

satisfies

$$m_{4,s,i,k} = \frac{\sum_{j=1}^{k} w'_{s,i,j} \cdot x_{s,i,j}}{k}.$$

The weighted average of the k most recent non-missing daily measurements with exponential weights satisfies

$$m_{5,s,i,k} = \frac{\sum_{j=1}^{k} w'_{s,i,j} \cdot x_{s,i,j}}{\sum_{j=1}^{k} w'_{s,i,j}}.$$

2.2.2. Likelihood Cross-Validation

Denote by SI the set of index pairs s,i for which depression levels $y_{s,i}$ have non-missing values and for which $J_{s,i} \ge J_*$. Let m denote the number of measurements indexed by SI. Participants s can have missing depression values for some of the four post-baseline surveys i. For subsets SI' of SI, let $\boldsymbol{\theta}(SI')$ denote the estimate of the parameter vector $\boldsymbol{\theta}$ for a specific linear mixed model for depression, including parameters for the mean, variance, and correlation components. This estimate is obtained by maximizing the likelihood $f\left(SI';\boldsymbol{\theta}\right)$ in $\boldsymbol{\theta}$. The likelihood is the standard multivariate normal density.

Randomly partition the set SI into folds $SI(h)$ of varying sizes for $1 \le h \le 10$. Arbitrary numbers of folds are possible, but reported analyses use 10 folds. For $1 \le h \le 10$, let

$$U(h) = \bigcup_{h'=1}^{h} SI(h')$$

denote the union of all folds $SI(h')$ for $1 \le h' \le h$ and $U(0)$ the empty fold with likelihood set to 1. Define the 10-fold likelihood cross-validation score LCV as

$$\mathrm{LCV} = \prod_{h=1}^{k} f^{1/n}\left(SI(h) \mid U(h-1); \boldsymbol{\theta}\left(SI \setminus SI(h)\right)\right)$$

where

$$f\left(SI(h) \mid U(h-1); \boldsymbol{\theta}\left(SI \setminus SI(h)\right)\right) = \frac{f\left(U(h); \boldsymbol{\theta}\left(SI \setminus SI(h)\right)\right)}{f\left(U(h-1); \boldsymbol{\theta}\left(SI \setminus SI(h)\right)\right)}$$

is the conditional likelihood for the fold $SI(h)$ conditioned on the union $U(h-1)$ of the prior folds computed with the parameter vector estimated using the data in the complement of the fold. This LCV definition accounts for the possibility of missing survey measurements for some participants. It is a geometric average of deleted fold likelihoods normalized by the number m of measurements. This definition extends readily to handle other likelihoods. The same initial seed is used in generating folds for models of the same data set so that associated LCV scores are comparable.

A larger LCV score indicates a better model but not necessarily a substantially better model. LCV ratio tests (generalizing likelihood ratio tests) are used to assess whether models with larger LCV scores provide substantial improvements, or equivalently whether models with smaller LCV scores are substantially inferior. These tests are based on cutoffs for a substantial percent decrease (PD) in the LCV score for models with smaller scores compared to models with larger scores. This cutoff is computed using the χ^2 distribution and decreases with the number m of measurements. In reported analyses, cutoffs are computed using 1 degree of freedom and a significant level of 5%, but these choices can be adjusted. A detailed formulation is provided in Knafl (2023). LCV ratio tests are more conservative than standard tests for zero coefficients, that is, the removal from a model of a significant parameter using a standard test does not always result in a substantial PD in the LCV score.

2.2.3. Adaptive Regression Modeling

Adaptive regression modeling (for details see [3] [8]) is based on a heuristic search through power transforms of an arbitrary set of primary (untransformed) predictors for modeling a variety of types of possibly correlated outcomes. A base model is systematically expanded adding in transforms. The expanded model is then systematically contracted by removing extraneous transforms and adjusting the powers of remaining transforms. LCV scores and tolerances for allowable decreases in the LCV score control the process. An LCV ratio test is used to set the tolerance for stopping the contraction so that the removal of any one of the transforms from the final contracted model generates a substantial PD in the LCV score. The expansion can generate geometric combinations, that is, products of power transforms of different primary predictors, generalizing standard interactions. The contraction considers removal of the intercept.

The process can be applied to modeling the mean component of the model as well as the variance/dispersion component. It has recently been extended to handle adaptive random effects/coefficient modeling of the covariance structure [9]. However, the full adaptive modeling process is not required to conduct reported analyses. For those analyses, only the mean component is modeled while assuming exchangeable correlations and constant variances. Also, the mean component is modeled in terms of just one average of the five mean types defined in Section 2.2.1. However, the adaptive modeling process considers models based on multiple transforms of that average.

Two subsets of the data are analyzed for each of the four daily measures of Section 2.1.2: surveys with at least J_* recent non-missing daily measurements with $J_* = 3$ and $J_* = 5$. For each of these subsets, possible predictors are averages of k most recent non-missing daily measurements for $k = 1, 2, \cdots, J_*$ using each of the five mean types, and so 15 alternate averages for $J_* = 3$ and 25 for $J_* = 5$.

All computations are conducted in SAS® Version 9.4 (SAS Institute Inc., Cary, NC). A SAS macro for conducting general adaptive modeling is available upon request from the first author (gknafl@unc.edu).

3. Results

This section provides results for analyses of depression in terms of the four daily measures of Section 2.1.2. Section 3.1 provides analyses using the number of opioid medications taken, Section 3.2 analyses using the number of pain flares, Section 3.3 analyses of the least pain level, and Section 3.4 analyses of the worst pain level. These analyses use either simple averages of weighted values, that is, normalized by the number of measurements, or weighted averages, that is, normalized by the sum of the weights.

3.1. Analyses of Numbers of Opioid Medications Taken

3.1.1. Means of Three Recent Non-Missing Numbers of Opioid Medications Taken

For $J_* = 3$, a total of 865 post-baseline depression values $y_{s,i}$ for 276 participants are non-missing with three recent non-missing numbers of opioid medications taken on or prior to the associated survey date. The cutoff for a substantial percent decrease (PD) in the LCV score is 0.22%. Fold sizes range from 60 to 115 measurements for 83 to 96 participants. Since fold complements contain at least 750 measurements, they should generate reliable parameter estimates for computing LCV scores.

Table 1 contains generated adaptive models for depression and LCV scores for the five mean types averaging the three recent non-missing numbers of opioid

Table 1. Adaptive models for depression in terms of alternate types of means of three recent non-missing numbers of opioid medications taken[a].

Mean type	Weights	Number most recent values averaged	Power transforms of the mean[b]	10-Fold LCV score	PD for constant model[c]	PD versus best[d]
average	unit	1	−0.01	0.057535	0.00%	0.95%
		2	0.03	0.057543	0.02%	0.94%
		3	0.06	0.057570	0.06%	0.89%
average	reciprocal	1	0.06	0.057988	0.78%	0.17%
		2	0.07	0.058058	0.90%	0.05%
		3	0.08	0.058088	0.96%	0.00%
weighted average	reciprocal	1	−0.01	0.057535	0.00%	0.95%
		2	0.02	0.057535	0.00%	0.95%
		3	0.02	0.057539	0.01%	0.95%
average	exponential	1	0, 0.09	0.057978	0.77%	0.19%
		2	0, 0.09	0.058004	0.81%	0.14%
		3	0, 0.1	0.057993	0.79%	0.16%
weighted average	exponential	1	−0.01	0.057535	0.00%	0.95%
		2	0	0.057533	0.00%	0.96%
		3	0.01	0.057534	0.00%	0.95%

LCV: likelihood cross-validation; PD: percent decrease. [a]Using data for 865 surveys of 276 participants with 0.22% the cutoff for a substantial PD in the LCV score. Assuming exchangeable correlations and constant variances. [b]A power of 0 corresponds to an intercept. [c]Using the constant model with LCV score 0.057533. [d]Compared to model with the largest LCV score.

medications taken. For each of the five mean types, associated powers for generated transforms are similar across the three averages of recent non-missing numbers of opioid medications taken. For the three cases of the average with unit weights ($m_{1,s,i,k}$), the weighted average with reciprocal weights ($m_{3,s,i,k}$), and the weighted average with exponential weights ($m_{5,s,i,k}$), depression is essentially constant in that mean type for all three averages of recent non-missing numbers of opioid medications taken (*i.e.*, the PD for the constant model is less than the cutoff 0.22% for a substantial PD). For the other two cases of the average with reciprocal weights ($m_{2,s,i,k}$) and the average with exponential weights ($m_{4,s,i,k}$), depression is substantially non-constant in that mean type for all three averages of recent non-missing numbers of opioid medications taken. The average with reciprocal weights of three recent non-missing numbers of opioid medications taken generates the best overall LCV score of 0.058088. Competitive LCV scores are generated for averages with reciprocal or exponential weights for all three averages of recent non-missing numbers of opioid medications taken (*i.e.*, the PD compared to the best overall LCV score is less than or equal to the cutoff 0.22%). All models for these two cases are monotonic since they are either constant or based on a single power transform either with or without an intercept.

Figure 1 contains the plot of estimated mean depression as a function of the average of the three recent non-missing numbers of opioid medications taken with reciprocal weights generating the best overall LCV score. Estimated mean depression increases nonlinearly with larger values of this average. The estimated exchangeable correlation equals 0.70 while the estimated constant standard deviation equals 5.5.

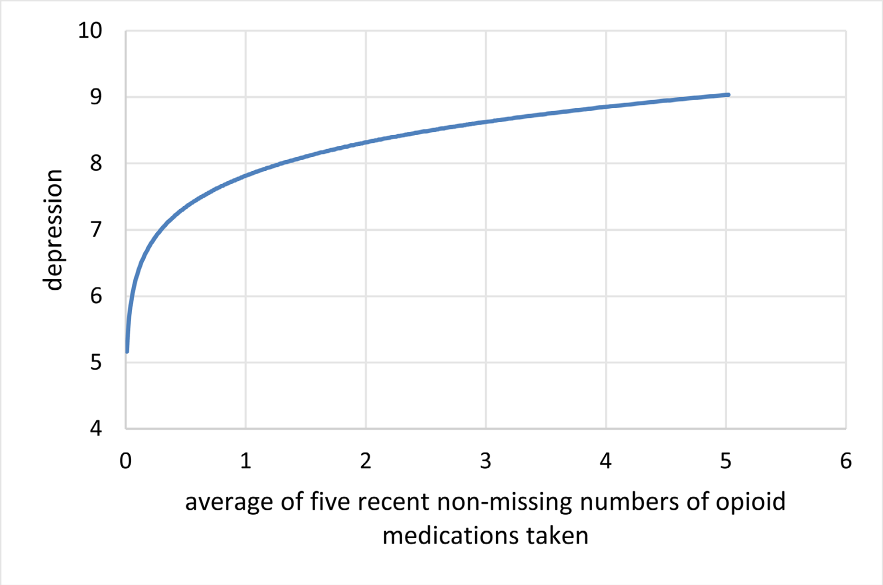

Figure 1. Estimated mean depression as a function of the average of the three recent non-missing numbers of medications taken weighted with reciprocal weights.

3.1.2. Means of Five Recent Non-Missing Numbers of Opioid Medications Taken

For $J_* = 5$, a total of 844 post-baseline depression values $y_{s,i}$ for 269 participants are non-missing with five recent non-missing numbers of opioid medications taken on or prior to the associated survey date. The cutoff for a substantial percent decrease (PD) in the LCV score is 0.23%. Fold sizes range from 59 to 112 measurements for 51 to 93 participants. Since fold complements contain at least 732 measurements, they should generate reliable parameter estimates for computing LCV scores.

Table 2 contains generated adaptive models for depression and LCV scores for the five mean types averaging the five recent non-missing numbers of opioid medications taken. For each of the five mean types, the power transforms of the associated mean are similar across the five averages of recent non-missing numbers of opioid medications taken. For the three cases of the average with unit weights ($m_{1,s,i,k}$), the weighted average with reciprocal weights ($m_{3,s,i,k}$), and the weighted average with exponential weights ($m_{5,s,i,k}$), depression is essentially constant in that mean type for all five averages of recent non-missing numbers of

Table 2. Adaptive models for depression in terms of alternate types of means of five recent non-missing numbers of opioid medications taken[a].

mean type	weights	number most recent values averaged	power transforms of the mean[b]	10-Fold LCV score	PD for the constant model[c]	PD versus best[d]
average	unit	1	−0.02	0.057323	0.00%	0.77%
		2	0.03	0.057328	0.01%	0.77%
		3	0	0.057321	0.00%	0.78%
		4	0	0.057321	0.00%	0.78%
		5	0.05	0.057354	0.06%	0.72%
average	reciprocal	1	0.05	0.057674	0.61%	0.17%
		2	0.07	0.057732	0.71%	0.07%
		3	0.07	0.057757	0.75%	0.02%
		4	0.07	0.057760	0.76%	0.02%
		5	0.08	0.057770	0.78%	0.00%
weighted average	reciprocal	1	−0.02	0.057323	0.00%	0.77%
		2	0.01	0.057323	0.00%	0.77%
		3	0.02	0.057327	0.01%	0.77%
		4	0.03	0.057329	0.01%	0.76%
		5	0.03	0.057330	0.02%	0.76%
average	exponential	1	0, 0.09	0.057714	0.68%	0.10%
		2	0, 0.09	0.057739	0.72%	0.05%
		3	0, 0.1	0.057723	0.70%	0.08%
		4	0, 0.1	0.057719	0.69%	0.09%
		5	0, 0.1	0.057706	0.67%	0.11%
weighted average	exponential	1	−0.02	0.057323	0.00%	0.77%
		2	0	0.057321	0.00%	0.78%
		3	0.01	0.057321	0.00%	0.78%
		4	0.01	0.057321	0.00%	0.78%
		5	0.01	0.057321	0.00%	0.78%

LCV: likelihood cross-validation; PD: percent decrease. [a]Using data for 844 surveys of 269 participants with 0.23% the cutoff for a substantial PD in the LCV score. Assuming exchangeable correlations and constant variances. [b]A power of 0 corresponds to an intercept. [c]Using the constant model with LCV score 0.057321. [d]Compared to model with the largest LCV score.

opioid medications taken. For the other two cases of the average with reciprocal weights ($m_{2,s,i,k}$) and the average with exponential weights ($m_{4,s,i,k}$), depression is substantially non-constant in that mean type for all five averages of recent non-missing numbers of opioid medications taken. The average with reciprocal weights of five recent non-missing numbers of opioid medications taken generates the best overall LCV score of 0.057770. Competitive LCV scores are generated for averages with reciprocal or exponential weights for all five averages of recent non-missing numbers of opioid medications taken. All models for these two cases are monotonic since they are either constant or based on a single power transform either with or without an intercept.

Figure 2 contains the plot of estimated mean depression as a function of the average of the five recent non-missing numbers of opioid medications taken with reciprocal weights generating the best overall LCV score. Estimated mean depression increases nonlinearly with larger values of this average. The estimated exchangeable correlation equals 0.70 while the estimated constant standard deviation equals 5.5.

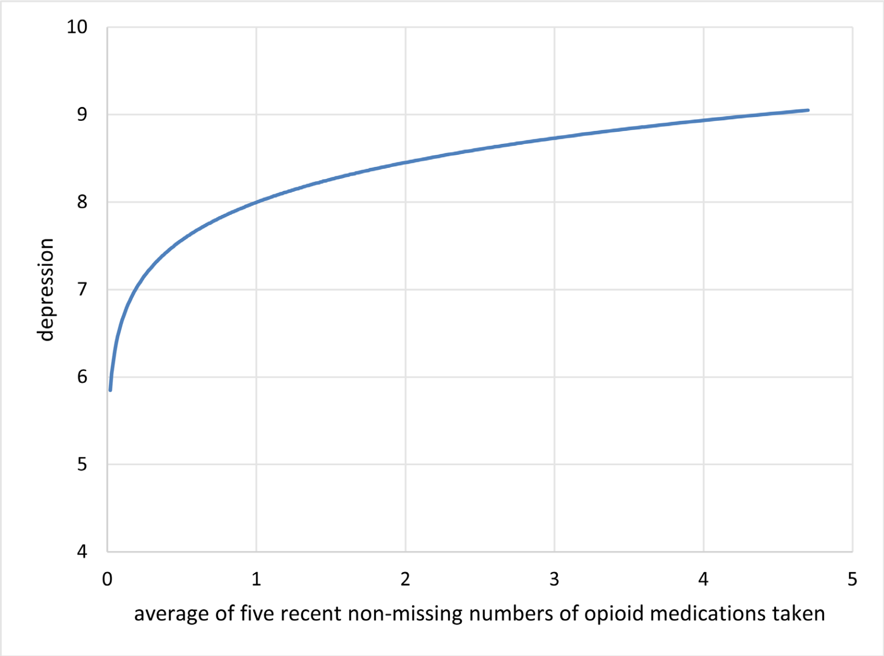

Figure 2. Estimated mean depression as a function of the average of the five recent non-missing numbers of medications taken weighted with reciprocal weights.

3.1.3. Comparison of Results for Means of Three and Five Recent Non-Missing Numbers of Opioid Medications Taken

Averages of non-unit weighted recent non-missing numbers of opioid medications taken generate better models for depression than averages of unit weighted recent non-missing numbers of opioid medications taken and weighted averages of recent non-missing numbers of opioid medications taken. Reciprocal weights generate the best overall LCV scores but exponential weights generate competitive models. Results restricting to five recent non-missing numbers of opioid medications taken are similar to those restricting to three recent non-missing numbers of opioid medications taken.

3.2. Analyses of Numbers of Pain Flares

3.2.1. Means of Three Recent Non-Missing Numbers of Pain Flares

For $J_* = 3$, a total of 970 post-baseline depression values $y_{s,i}$ for 298 participants are non-missing with three recent non-missing numbers of pain flares on or prior to the associated survey date. The cutoff for a substantial percent decrease (PD) in the LCV score is 0.20%. Fold sizes range from 63 to 121 measurements for 63 to 102 participants. Since fold complements contain at least 849 measurements, they should generate reliable parameter estimates for computing LCV scores.

Table 3 contains generated adaptive models for depression and LCV scores for the five mean types averaging the three recent non-missing numbers of pain flares. For each of the five mean types, associated powers for generated transforms are similar across the three averages of recent non-missing numbers of pain flares. For the four cases of the average with unit weights ($m_{1,s,i,k}$), the average with reciprocal weights ($m_{2,s,i,k}$), the average with exponential weights ($m_{4,s,i,k}$), and

Table 3. Adaptive models for depression in terms of alternate types of means of three recent non-missing numbers of pain flares[a].

mean type	weights	number most recent values averaged	power transforms of the mean[b]	10-Fold LCV score	PD for the constant model[c]	PD versus best[d]
average	unit	1	0.07	0.059112	0.21%	1.08%
		2	0.09	0.059155	0.28%	1.01%
		3	0.11	0.059223	0.39%	0.90%
average	reciprocal	1	0.07	0.059679	1.15%	0.13%
		2	0.08	0.059731	1.24%	0.05%
		3	0.09	0.059758	1.29%	0.00%
weighted average	reciprocal	1	0.07	0.059112	0.21%	1.08%
		2	0.04	0.059035	0.08%	1.21%
		3	0.02	0.059004	0.02%	1.26%
average	exponential	1	0, 0.21	0.059645	1.10%	0.19%
		2	0, 0.21	0.059637	1.08%	0.20%
		3	0, 0.2	0.059643	1.09%	0.19%
weighted average	exponential	1	0.07	0.059112	0.21%	1.08%
		2	0.08	0.059137	0.25%	1.04%
		3	0.08	0.059124	0.23%	1.06%

LCV: likelihood cross-validation; PD: percent decrease. [a]Using data for 970 surveys of 298 participants with 0.20% the cutoff for a substantial PD in the LCV score. Assuming exchangeable correlations and constant variances. [b]A power of 0 corresponds to an intercept. [c]Using the constant model with LCV score 0.058990. [d]Compared to model with the largest LCV score.

the weighted average with exponential weights ($m_{5,s,i,k}$), depression is substantially nonconstant in that mean type for all three averages of recent non-missing numbers of pain flares (*i.e.*, the PD for the constant model is larger than the cutoff 0.20% for a substantial PD). For the case of the weighted average with reciprocal weights ($m_{3,s,i,k}$), this only holds for the first most recent non-missing value but not for averages of the first two and three most recent non-missing values. The average with reciprocal weights of three recent non-missing numbers of pain flares generates the best overall LCV score of 0.059758. Competitive LCV scores are generated for averages with reciprocal or exponential weights for all three averages of recent non-missing numbers of pain flares (*i.e.*, the PD compared to the best overall LCV score is less than or equal the cutoff 0.20%). The other three mean types generate distinctly inferior models. All models for these two cases are monotonic since they are based on a single power transform either with or without an intercept.

Figure 3 contains the plot of estimated mean depression as a function of the average of the three recent non-missing numbers of pain flares with reciprocal weights generating the best overall LCV score. Estimated mean depression increases nonlinearly with larger values of this average. The estimated exchangeable correlation equals 0.70 while the estimated constant standard deviation equals 5.3.

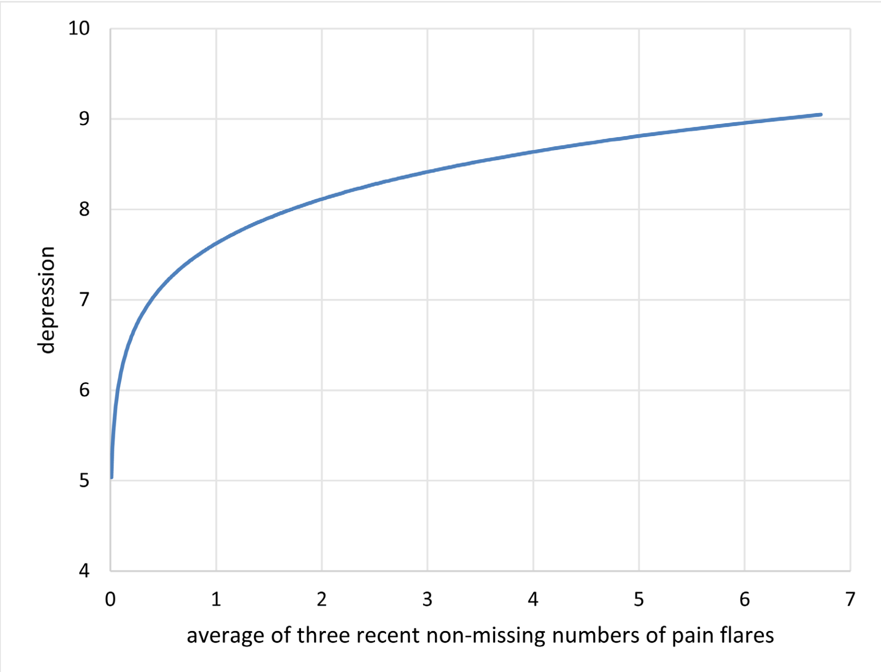

Figure 3. Estimated mean depression as a function of the average of the three recent non-missing numbers of pain flares weighted with reciprocal weights.

3.2.2. Means of Five Recent Non-Missing Numbers of Pain Flares

For $J_* = 5$, a total of 938 post-baseline depression values $y_{s,i}$ for 294 participants are non-missing with five recent non-missing numbers of pain flares on or prior to the associated survey date. The cutoff for a substantial percent decrease (PD) in the

LCV score is 0.20%. Fold sizes range from 60 to 119 measurements for 60 to 100 participants. Since fold complements contain at least 819 measurements, they should generate reliable parameter estimates for computing LCV scores.

Table 4 contains generated adaptive models for depression and LCV scores for the five mean types averaging the five recent non-missing numbers of pain flares.

Table 4. Adaptive models for depression in terms of alternate types of means of five recent non-missing numbers of pain flares[a].

mean type	weights	number most recent values averaged	power transforms of the mean[b]	10-Fold LCV score	PD for the constant model[c]	PD versus best[d]
average	unit	1	0.07	0.059101	0.25%	0.95%
		2	0.1	0.059162	0.35%	0.85%
		3	0.12	0.059239	0.48%	0.72%
		4	0.13	0.059278	0.54%	0.65%
		5	0.13	0.059260	0.51%	0.68%
average	reciprocal	1	0.06	0.059579	1.05%	0.15%
		2	0.07	0.059619	1.11%	0.08%
		3	0.08	0.059641	1.15%	0.04%
		4	0.09	0.059654	1.17%	0.02%
		5	0.09	0.059667	1.19%	0.00%
weighted average	reciprocal	1	0.07	0.059101	0.25%	0.95%
		2	0.05	0.059018	0.11%	1.09%
		3	0.03	0.058978	0.04%	1.15%
		4	0.01	0.058960	0.01%	1.18%
		5	0	0.058955	0.00%	1.19%
average	exponential	1	0, 0.21	0.059435	0.81%	0.39%
		2	0, 0.21	0.059428	0.80%	0.40%
		3	0, 0.22	0.059418	0.78%	0.42%
		4	0, 0.2	0.059440	0.82%	0.38%
		5	0, 0.2	0.059437	0.81%	0.39%
weighted average	exponential	1	0.07	0.059101	0.25%	0.95%
		2	0.08	0.059116	0.27%	0.92%
		3	0.08	0.059099	0.24%	0.95%
		4	0.08	0.059093	0.23%	0.96%
		5	0.08	0.059095	0.24%	0.96%

LCV: likelihood cross-validation; PD: percent decrease. [a]Using data for 938 surveys of 294 participants with 0.20% the cutoff for a substantial PD in the LCV score. Assuming exchangeable correlations and constant variances. [b]A power of 0 corresponds to an intercept. [c]Using the constant model with LCV score 0.058955. [d]Compared to model with the largest LCV score.

For each of the five mean types, associated powers for generated transforms are similar across the five averages of recent non-missing numbers of pain flares. For the four cases of the average with unit weights ($m_{1,s,i,k}$), the average with reciprocal weights ($m_{2,s,i,k}$), the average with exponential weights ($m_{4,s,i,k}$), and the weighted average with exponential weights ($m_{5,s,i,k}$), depression is substantially nonconstant in that mean type for all five averages of recent non-missing numbers of pain flares (*i.e.*, the PD for the constant model is larger than the cutoff 0.20% for a substantial PD). For the case of the weighted average with reciprocal weights ($m_{3,s,i,k}$), this only holds for the first most recent non-missing value but not for averages of the first two to five most recent non-missing values. The average with reciprocal weights of five recent non-missing numbers of pain flares generates the best overall LCV score of 0.059667. Competitive LCV scores are generated for the average with reciprocal weights for all five averages of recent non-missing numbers of pain flares (*i.e.*, the PD compared to the best overall LCV score is less than or equal the cutoff 0.20%). The other four mean types generate distinctly inferior models. All models for these two cases are monotonic since they are either constant or based on a single power transform either with or without an intercept.

Figure 4 contains the plot of estimated mean depression as a function of the average of the five recent non-missing numbers of pain flares with reciprocal weights generating the best overall LCV score. Estimated mean depression increases nonlinearly with larger values of this average. The estimated exchangeable correlation equals 0.70 while the estimated constant standard deviation equals 5.3.

3.2.3. Comparison of Results for Means of Three and Five Recent Non-Missing Numbers of Pain Flares

Averages of non-unit weighted recent non-missing numbers of pain flares generate better models for depression than averages of unit weighted recent non-missing numbers of pain flares and weighted averages of recent non-missing numbers of pain flares. Reciprocal weights generate the best overall LCV scores. Results restricting to five recent non-missing numbers of pain flares are similar to those restricting to three recent non-missing numbers of pain flares with one exception. Averages with exponential weights generate competitive models for means of the three recent non-missing numbers of pain flares but not for means of the five recent non-missing numbers of pain flares.

3.3. Analyses of Numbers of Least Pain Levels

3.3.1. Means of Three Recent Non-Missing Least Pain Levels

For $J_* = 3$, a total of 970 post-baseline depression values $y_{s,i}$ for 298 participants are non-missing with three recent non-missing least pain levels on or prior to the associated survey date. The cutoff for a substantial percent decrease (PD) in the LCV score is 0.20%. Fold sizes range from 63 to 121 measurements for 63 to 102 participants. Since fold complements contain at least 849 measurements, they should generate reliable parameter estimates for computing LCV scores.

Table 5 contains generated adaptive models for depression and LCV scores for

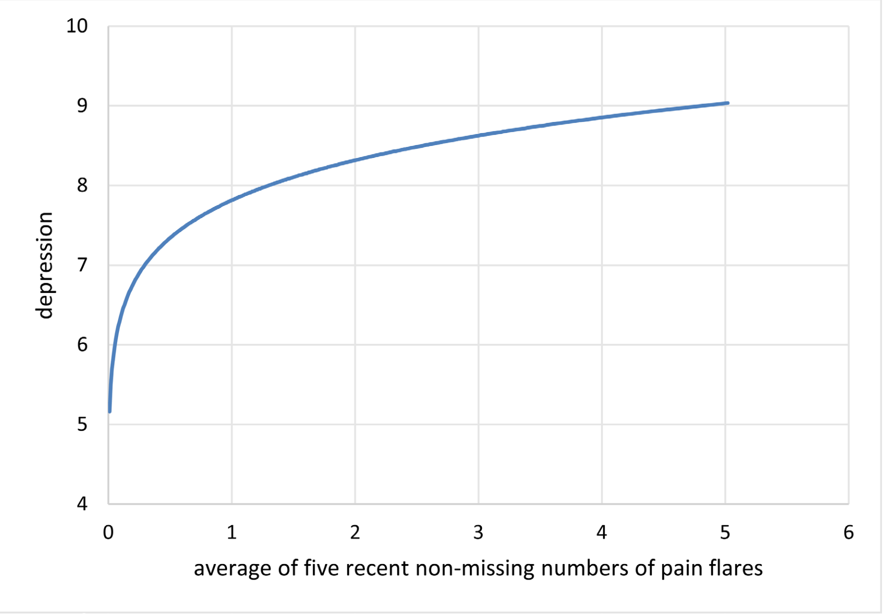

Figure 4. Estimated mean depression as a function of the average of the five recent non-missing numbers of pain flares weighted with reciprocal weights.

Table 5. Adaptive models for depression in terms of alternate types of means of three recent non-missing least pain levels[a].

mean type	weights	number most recent values averaged	power transforms of the mean[b]	10-Fold LCV score	PD for the constant model[c]	PD versus best[d]
average	unit	1	0.05	0.059527	0.10%	0.80%
		2	0.07	0.059560	0.15%	0.74%
		3	0.09	0.059603	0.22%	0.67%
average	reciprocal	1	0.06	0.059967	0.83%	0.06%
		2	0.06	0.059989	0.87%	0.03%
		3	0.07	0.060006	0.89%	0.00%
weighted average	reciprocal	1	0.05	0.059527	0.10%	0.80%
		2	0.02	0.059479	0.02%	0.88%
		3	−0.01	0.059466	--	0.90%
average	exponential	1	0, 0.4	0.059921	0.75%	0.14%
		2	0, 0.5	0.059917	0.75%	0.15%
		3	0, 0.4	0.059927	0.76%	0.13%
weighted average	exponential	1	0.05	0.059527	0.10%	0.80%
		2	0.06	0.059539	0.12%	0.78%
		3	0.06	0.059544	0.12%	0.77%

LCV: likelihood cross-validation; PD: percent decrease. [a]Using data for 982 surveys of 299 participants with 0.20% the cutoff for a substantial PD in the LCV score. Assuming exchangeable correlations and constant variances. [b]A power of 0 corresponds to an intercept. [c]Using the constant model with LCV score 0.059470. [d]Compared to model with the largest LCV score.

the five mean types averaging the three recent non-missing least pain levels. For each of the five mean types, associated powers for generated transforms are similar across the three averages of recent non-missing least pain levels. For the two cases of the average with reciprocal weights ($m_{2,s,i,k}$) and the average with exponential weights ($m_{4,s,i,k}$), depression is substantially nonconstant in that mean type for all three averages of recent non-missing least pain levels (*i.e.*, the PD for the constant model is larger than the cutoff 0.20% for a substantial PD). For the case of the average with unit weights ($m_{1,s,i,k}$), this only holds for the first most recent non-missing value but not for averages of the first two and three most recent non-missing values. For the two cases of the weighted average with reciprocal weights ($m_{3,s,i,k}$) and the weighted average with exponential weights ($m_{5,s,i,k}$), depression is essentially constant in that mean type for all three averages. The average with reciprocal weights of three recent non-missing least pain levels generates the best overall LCV score of 0.060006. Competitive LCV scores are generated for averages with reciprocal or exponential weights for all three averages of recent non-missing least pain levels (*i.e.*, the PD compared to the best overall LCV score is less than or equal the cutoff 0.20%). The other three mean types generate distinctly inferior models. All models for these two cases are monotonic since they are based on a single power transform either with or without an intercept.

Figure 5 contains the plot of estimated mean depression as a function of the average of the three recent non-missing least pain levels with reciprocal weights generating the best overall LCV score. Estimated mean depression increases nonlinearly with larger values of this average. The estimated exchangeable correlation equals 0.71 while the estimated constant standard deviation equals 5.4.

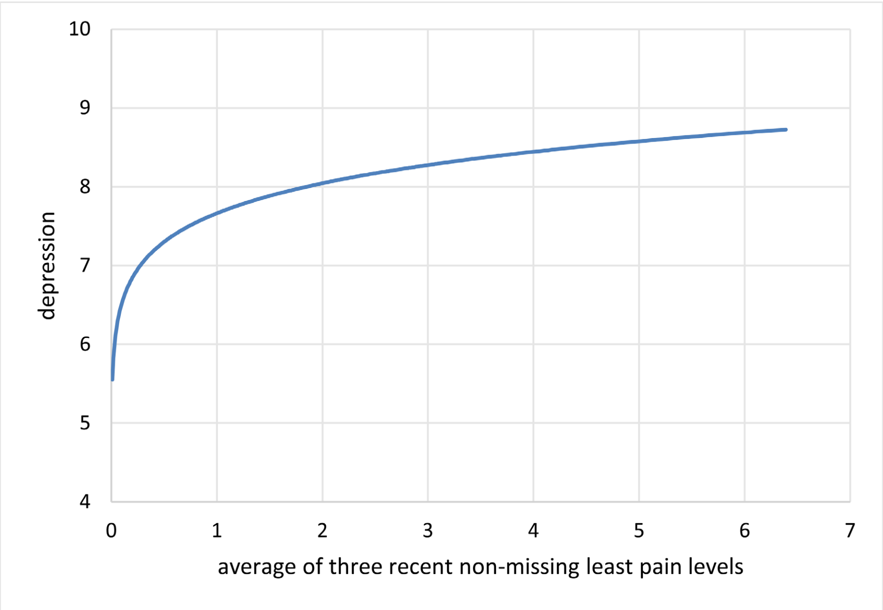

Figure 5. Estimated mean depression as a function of the average of the three recent non-missing least pain levels weighted with reciprocal weights.

3.3.2. Means of Five Recent Non-Missing Least Pain Levels

For $J_* = 5$, a total of 954 post-baseline depression values $y_{s,i}$ for 296 participants are non-missing with five recent non-missing least pain levels on or prior to the associated survey date. The cutoff for a substantial percent decrease (PD) in the LCV score is 0.20%. Fold sizes range from 73 to 121 measurements for 63 to 105 participants. Since fold complements contain at least 833 measurements, they should generate reliable parameter estimates for computing LCV scores.

Table 6 contains generated adaptive models for depression and LCV scores for the five mean types averaging the five recent non-missing least pain levels. For each of the five mean types, associated powers for generated transforms are similar across the five averages of recent non-missing least pain levels. For the two cases of the average with reciprocal weights ($m_{2,s,i,k}$) and the average with exponential weights ($m_{4,s,i,k}$), depression is substantially nonconstant in that mean type for all five averages of recent non-missing least pain levels (*i.e.*, the PD for the constant model is larger than the cutoff 0.20% for a substantial PD). For the case of the average with unit weights ($m_{1,s,i,k}$), this only holds for the averages of the first one and two most recent non-missing values but not for averages of the first three to five most recent non-missing values. For the two cases of the weighted average with reciprocal weights ($m_{3,s,i,k}$) and the weighted average with exponential weights ($m_{5,s,i,k}$), depression is essentially constant in that mean type for all five averages. The average with reciprocal weights of three recent non-missing least pain levels generates the best overall LCV score of 0.059917. Competitive LCV scores are generated for averages with reciprocal or exponential weights for all five averages of recent non-missing least pain levels (*i.e.*, the PD compared to the best overall LCV score is less than or equal the cutoff 0.20%). The other three mean types generate distinctly inferior models. All models for these two cases are monotonic since they are either constant or based on a single power transform either with or without an intercept.

Figure 6 contains the plot of estimated mean depression as a function of the average of the five recent non-missing least pain levels with reciprocal weights generating the best overall LCV score. Estimated mean depression increases nonlinearly with larger values of this average. The estimated exchangeable correlation equals 0.71 while the estimated constant standard deviation equals 5.4.

3.3.3. Comparison of Results for Means of Three and Five Recent Non-Missing Least Pain Levels

Averages of non-unit weighted recent non-missing least pain levels generate better models for depression than averages of unit weighted recent non-missing least pain levels and weighted averages of recent non-missing least pain levels. Reciprocal weights generate the best overall LCV scores but exponential weights generate competitive models. Results restricting to five recent non-missing least pain levels are similar to those restricting to three recent non-missing least pain levels.

Table 6. Adaptive models for depression in terms of alternate types of means of five recent non-missing least pain levels[a].

mean type	weights	number most recent values averaged	power transforms of the mean[b]	10-Fold LCV score	PD for the constant model[c]	PD versus best[d]
average	unit	1	0.05	0.059518	0.08%	0.67%
		2	0.06	0.059544	0.13%	0.62%
		3	0.09	0.059602	0.22%	0.53%
		4	0.1	0.059606	0.23%	0.52%
		5	0.11	0.059652	0.31%	0.44%
average	reciprocal	1	0.05	0.059886	0.70%	0.05%
		2	0.06	0.059893	0.71%	0.04%
		3	0.06	0.059910	0.74%	0.01%
		4	0.07	0.059907	0.73%	0.02%
		5	0.07	0.059917	0.75%	0.00%
weighted average	reciprocal	1	0.05	0.059518	0.08%	0.67%
		2	0.02	0.059473	0.01%	0.74%
		3	0	0.059468	0.00%	0.75%
		4	0	0.059468	0.00%	0.75%
		5	0	0.059468	0.00%	0.75%
average	exponential	1	0, 0.4	0.059799	0.55%	0.20%
		2	0, 0.5	0.059812	0.58%	0.18%
		3	0, 0.5	0.059821	0.59%	0.16%
		4	0, 0.5	0.059823	0.59%	0.16%
		5	0, 0.5	0.059821	0.59%	0.16%
weighted average	exponential	1	0.05	0.059518	0.08%	0.67%
		2	0.05	0.059526	0.10%	0.65%
		3	0.06	0.059529	0.10%	0.65%
		4	0.06	0.059529	0.10%	0.65%
		5	0.06	0.059531	0.11%	0.64%

LCV: likelihood cross-validation; PD: percent decrease. [a]Using data for 954 surveys of 296 participants with 0.20% the cutoff for a substantial PD in the LCV score. Assuming exchangeable correlations and constant variances. [b]A power of 0 corresponds to an intercept. [c]Using the constant model with LCV score 0.059468. [d]Compared to model with the largest LCV score.

3.4. Analyses of Numbers of Worst Pain Levels

3.4.1. Means of Three Recent Non-Missing Worst Pain Levels

For $J_* = 3$, a total of 990 post-baseline depression values $y_{s,i}$ for 301 participants

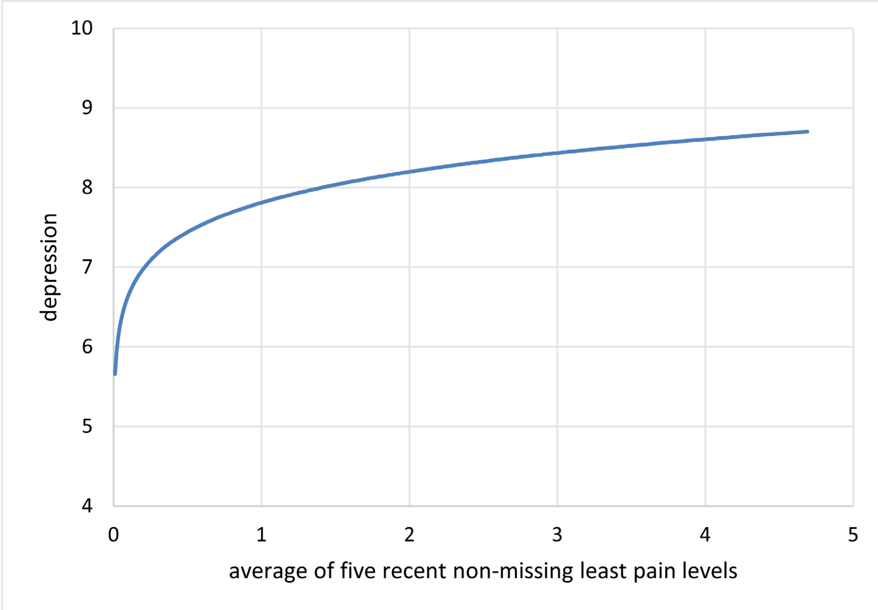

Figure 6. Estimated mean depression as a function of the average of the five recent non-missing least pain levels weighted with reciprocal weights.

are non-missing with three recent non-missing worst pain levels on or prior to the associated survey date. The cutoff for a substantial percent decrease (PD) in the LCV score is 0.19%. Fold sizes range from 63 to 121 measurements for 60 to 110 participants. Since fold complements contain at least 869 measurements, they should generate reliable parameter estimates for computing LCV scores.

Table 7 contains generated adaptive models for depression and LCV scores for the five mean types averaging the three recent non-missing worst pain levels. For each of the five mean types, associated models are based on a single transform across the three averages of recent non-missing worst pain levels with two exceptions. For the three cases of the average with reciprocal weights ($m_{2,s,i,k}$), the average with exponential weights ($m_{4,s,i,k}$), and the weighted average with exponential weights ($m_{5,s,i,k}$), models for all three averages of recent non-missing worst pain levels are based on a single transform. For the two cases of the average with unit weights ($m_{1,s,i,k}$) and the weighted average with reciprocal weights ($m_{3,s,i,k}$), the model for the average of the first two most recent non-missing values is based on two transforms while models for the other two averages are based on a single transform. For each of the five mean types, depression is substantially nonconstant in that mean type for all three averages of recent non-missing worst pain levels (*i.e.*, the PD for the constant model is larger than the cutoff 0.19% for a substantial PD). The average with reciprocal weights of three recent non-missing worst pain levels generates the best overall LCV score of 0.059649. Competitive LCV scores are generated for averages with reciprocal weights for all three averages of recent non-missing worst pain levels (*i.e.*, the PD compared to the best overall LCV score is less than or equal the cutoff 0.19%). The other four mean types generate distinctly inferior models. All models for averages with reciprocal

Table 7. Adaptive models for depression in terms of alternate types of means of three recent non-missing worst pain levels[a].

mean type	weights	number most recent values averaged	power transforms of the mean[b]	10-Fold LCV score	PD for the constant model[c]	PD versus best[d]
average	unit	1	0.12	0.059223	0.46%	0.71%
		2	9, 0.08	0.059419	0.79%	0.39%
		3	0.16	0.059265	0.53%	0.64%
average	reciprocal	1	0.06	0.059608	1.10%	0.07%
		2	0.07	0.059643	1.16%	0.01%
		3	0.08	0.059649	1.17%	0.00%
weighted average	reciprocal	1	0.12	0.059223	0.46%	0.71%
		2	0.08	0.059079	0.21%	0.96%
		3	2.6, 0.13	0.059159	0.35%	0.82%
average	exponential	1	0, 2.5	0.059487	0.90%	0.27%
		2	0, 0.5	0.059470	0.87%	0.30%
		3	0, 0.4	0.059476	0.88%	0.29%
weighted average	exponential	1	0.12	0.059223	0.46%	0.71%
		2	0.13	0.059254	0.51%	0.66%
		3	0.12	0.059236	0.48%	0.69%

LCV: likelihood cross-validation; PD: percent decrease. [a]Using data for 990 surveys of 301 participants with 0.19% the cutoff for a substantial PD in the LCV score. Assuming exchangeable correlations and constant variances. [b]A power of 0 corresponds to an intercept. [c]Using the constant model with LCV score 0.058952. [d]Compared to model with the largest LCV score.

weights are monotonic since they are based on a single power transform either with or without an intercept. The two models based on two power transforms are distinctly inferior but suggest the possibility of non-monotonicity. Figure 7 contains the plot of estimated mean depression for the model based on the average of the first two most recent non-missing values with unit weights and the larger LCV score. Even though it is based on two transforms, the generated curve is still monotonic.

Figure 8 contains the plot of estimated mean depression as a function of the average of the three recent non-missing worst pain levels with reciprocal weights generating the best overall LCV score. Estimated mean depression increases nonlinearly with larger values of this average. The estimated exchangeable correlation equals 0.70 while the estimated constant standard deviation equals 5.4.

3.4.2. Means of Five Recent Non-Missing Worst Pain Levels

For $J_* = 5$, a total of 954 post-baseline depression values $y_{s,i}$ for 296 participants are non-missing with five recent non-missing worst pain levels on or prior

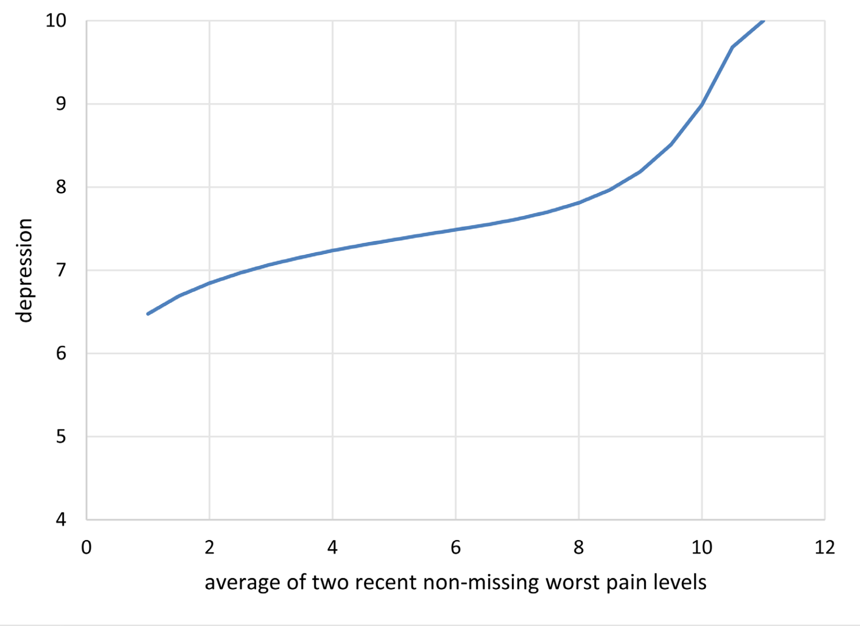

Figure 7. Estimated mean depression as a function of the unit-weighted average of the two most recent non-missing worst pain levels.

Figure 8. Estimated mean depression as a function of the average of the three recent non-missing worst pain levels weighted with reciprocal weights.

to the associated survey date. The cutoff for a substantial percent decrease (PD) in the LCV score is 0.20%. Fold sizes range from 73 to 121 measurements for 63 to 105 participants. Since fold complements contain at least 833 measurements, they should generate reliable parameter estimates for computing LCV scores.

Table 8 contains generated adaptive models for depression and LCV scores for the five mean types averaging the five recent non-missing worst pain levels. For each of the five mean types, associated models are based on a single transform

Table 8. Adaptive models for depression in terms of alternate types of means of five recent non-missing worst pain levels[a].

mean type	weights	number most recent values averaged	power transforms of the mean[b]	10-Fold LCV score	PD for the constant model[c]	PD versus best[d]
average	unit	1	8, 0.06	0.059107	0.88%	0.27%
		2	9, 0.09	0.059178	1.00%	0.15%
		3	0.17	0.058915	0.56%	0.59%
		4	0.18	0.058943	0.61%	0.55%
		5	0.2	0.059013	0.73%	0.43%
average	reciprocal	1	0.06	0.059218	1.07%	0.08%
		2	0.07	0.059238	1.10%	0.05%
		3	0.08	0.059238	1.10%	0.05%
		4	0.08	0.059247	1.12%	0.03%
		5	0.09	0.059266	1.15%	0.00%
weighted average	reciprocal	1	8, 0.06	0.059107	0.88%	0.27%
		2	0.09	0.058720	0.23%	0.92%
		3	0.1, 2.9	0.058779	0.33%	0.82%
		4	2.2, 0.14	0.058865	0.48%	0.68%
		5	1.9, 0.14	0.058909	0.55%	0.60%
average	exponential	1	0, 3	0.059236	1.10%	0.05%
		2	0, 1.5	0.059199	1.04%	0.11%
		3	0, 1.5	0.059201	1.04%	0.11%
		4	0, 1.5	0.059195	1.03%	0.12%
		5	0, 1.5	0.059196	1.03%	0.12%
weighted average	exponential	1	8, 0.06	0.059107	0.88%	0.27%
		2	0.12	0.058898	0.53%	0.62%
		3	0.13	0.058881	0.50%	0.65%
		4	0.12	0.058867	0.48%	0.67%
		5	0.12	0.058871	0.49%	0.67%

LCV: likelihood cross-validation; PD: percent decrease. [a] Using data for 968 surveys of 300 participants with 0.20% the cutoff for a substantial PD in the LCV score. Assuming exchangeable correlations and constant variances. [b]A power of 0 corresponds to an intercept. [c]Using the constant model with LCV score 0.058585. [d]Compared to model with the largest LCV score.

across the five averages of recent non-missing worst pain levels with seven exceptions. For the two cases of the average with reciprocal weights ($m_{2,s,i,k}$) and the average with exponential weights ($m_{4,s,i,k}$), models for all five averages of recent non-missing worst pain levels are based on a single transform. For the two cases of the average with unit weights ($m_{1,s,i,k}$) and the weighted average with reciprocal weights ($m_{3,s,i,k}$), the model for the average of the first two most recent non-missing values is based on two transforms. For the weighted average with reciprocal weights ($m_{3,s,i,k}$), all but the average of two most recent non-missing worst pain levels are based on two transforms. For the weighted average with exponential weights ($m_{5,s,i,k}$), the first most recent non-missing worst pain level is based on two transforms. For each of the five mean types, depression is substantially nonconstant in that mean type for all five averages of recent non-missing worst pain levels (*i.e.*, the PD for the constant model is larger than the cutoff 0.20% for a substantial PD). The average with reciprocal weights of five recent non-missing worst pain levels generates the best overall LCV score of 0.059266. Competitive LCV scores are generated for averages with reciprocal weights and for averages with exponential weights for all five averages of recent non-missing worst pain levels (*i.e.*, the PD compared to the best overall LCV score is less than or equal the cutoff 0.20%). The other three mean types generate distinctly inferior models with one exception: the average of two most recent non-missing worst pain levels with unit weights. All models for averages of reciprocal and exponential weights are monotonic since they are based on a single power transform either with or without an intercept. Six of the seven models based on two power transforms are distinctly inferior. The exception is the average of two most recent non-missing worst pain levels with unit weights, which also generates the best LCV score for the seven models based on two transforms. The plot generated by this model is similar to that of **Figure 7** and so has not been displayed. However, even though it is based on two transforms, the generated curve is still monotonic.

Figure 9 contains the plot of estimated mean depression as a function of the average of the five recent non-missing worst pain levels with reciprocal weights generating the best overall LCV score. Estimated mean depression increases nonlinearly with larger values of this average. The estimated exchangeable correlation equals 0.70 while the estimated constant standard deviation equals 5.4.

3.4.3. Comparison of Results for Means of Three and Five Non-Missing Worst Pain Levels

Averages of non-unit weighted recent non-missing worst pain levels generate better models for depression than averages of unit weighted recent non-missing worst pain levels and weighted averages of recent non-missing worst pain levels. Reciprocal weights generate the best overall LCV scores. Results restricting to five recent non-missing worst pain levels are similar to those restricting to three non-missing recent non-missing worst pain levels with two exceptions. Exponential weights generate competitive models for five recent non-missing worst pain levels but not for three recent non-missing worst pain levels. Also, more models are based on two transforms for five recent non-missing worst pain levels than for three recent non-missing worst pain levels.

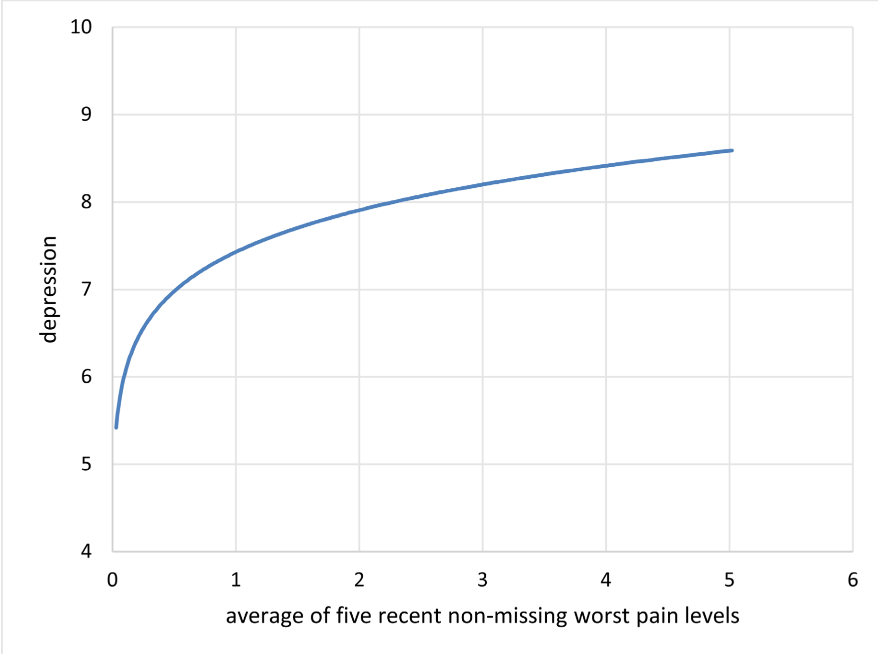

Figure 9. Estimated mean depression as a function of the average of the five recent non-missing worst pain levels weighted with reciprocal weights.

4. Discussion

Analyses were conducted of post-baseline survey data on depression measured by the PHQ 8 using linear mixed models assuming constant variances and exchangeable correlations. Mean depression was modeled in terms of four measures collected on a daily basis including the number of opioids taken, the number of pain flares, the least pain level, and the worst pain level. Each of these four daily measures was incorporated into analyses of mean depression using averages of recent non-missing daily values, that is, measurements on or prior to the associated survey date. Five types of averages were considered including averages with unit weights, averages with reciprocal weights, weighted averages with reciprocal weights, averages with exponential weights, and weighted averages with exponential weights. Nonlinearity in each of these averages was addressed using fractional polynomials, that is, power transforms with real-valued powers. Fractional polynomials were identified using adaptive regression modeling based on heuristic search controlled by likelihood cross-validation (LCV) scores. LCV ratio tests were used to identify substantial changes in LCV scores. A comparison was conducted of results for averages assuming five non-missing recent values with those assuming three recent non-missing values.

Results of analyses indicate that, for all four daily measures, averages of non-unit weighted recent non-missing values generated better models for mean depression than averages of unit weighted recent non-missing values and weighted averages of recent non-missing values. Most models were based on a single power transform of an average of recent non-missing values sometimes with an intercept and sometimes without, indicating that mean depression depended monotonically on those averages. Otherwise, models were based on two power transforms of averages of recent non-missing values without an intercept, but mean depression also depended monotonically on those averages. For all four daily measures, averages of reciprocal weighted recent non-missing values generated the best overall LCV scores. Mean depression under these models increased nonlinearly with increases in averages of reciprocal weighted recent non-missing values. Averages of exponential weighted recent non-missing values generated competitive LCV scores with two exceptions. This did not hold for averages based on three recent non-missing numbers of pain flares and for averages based on three recent non-missing pain worst levels. For the number of opioids taken and the least pain level, results restricting to five recent non-missing values were similar to those restricting to three recent non-missing values. This also held in most cases for the other two daily measures with exceptions related to the effects of exponential weights and related to numbers of models with multiple power transforms. These exceptions did not provide substantial support for considering averages of five recent non-missing values over averages of three recent non-missing values.

Reported analyses addressed only the special case of predicting depression for cancer patients. However, the analysis approach can be effectively applied to a wide variety of monthly outcomes for any type of patients. The analyses considered only the effects of averages of recent daily data on a monthly outcome, but adaptive regression methods can be applied to identify more general sets of possible predictors including averages of recent daily data among others.

Averages are a natural choice for combining recent daily measurements into predictors of monthly outcomes and the five averages considered in reported analyses provided for an effective set of alternatives. However, the results were limited by the fact that more general combinations were possible but not considered, for example, more general linear and geometric averages of recent measurements. Future research is needed to address such more general averages.

5. Conclusion

In summary, results reported in Section 3 supported the conclusion that an effective way to incorporate an intensive longitudinal measure as a predictor of a non-intensive longitudinal outcome is to use the average of the three most recent non-missing values for that intensive longitudinal measure weighting those recent non-missing values using reciprocal weights computed using the number of days between measurement of the daily measure and of the outcome. For the daily measures used in reported analyses, insufficient values were collected to compute recent non-missing averages for baseline outcome values, and so analyses addressed only post-baseline outcome values. A preferable approach when designing studies collecting both intensive and non-intensive data would be to postpone the collection of non-intensive data until after sufficient intensive longitudinal data were collected so that baseline non-intensive outcome values can be addressed in analyses.

Acknowledgements

This work was supported in part by the National Institutes of Health/National Institute of Nursing Research Awards 1R01NR017853 and RC1NR011591.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Ariens, S., Ceulemans, E. and Adolf, F.K. (2020) Time Series Analysis of Intensive Longitudinal Data in Psychosomatic Research: A Methodological Overview. *Journal of Psychosomatic Research*, **137**, 110191. https://doi.org/10.1016/j.jpsychores.2020.110191

[2] Kroenke, K., Strine, T.W., Spitzer, R.L., Williams, J.B., Berry, J.T. and Mokdad, A.H. (2009) The PHQ-8 as a Measure of Current Depression in the General Population. *Journal of Affective Disorders*, **114**, 163-173. https://doi.org/10.1016/j.jad.2008.06.026

[3] Knafl, G.J. and Ding, K. (2016) Adaptive Regression for Modeling Nonlinear Relationships. Springer International Publishing, Switzerland. https://doi.org/10.1007/978-3-319-33946-7_20

[4] Knafl, G.J. (2018) Adaptive Fractional Polynomial Modeling. *Open Journal of Statistics*, **8**, 159-186. https://doi.org/10.4236/ojs.2018.81011

[5] Royston, P. and Altman, D.G. (1994) Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modeling. *Applied Statistics*, **43**, 429-467. https://doi.org/10.2307/2986270

[6] Self-Management Resource Center, SMRC - English_-_PHQ-8.pdf (selfmanagementresource.com) https://selfmanagementresource.com/wp-content/uploads/2022/06/English_-_PHQ-8-1.pdf

[7] Ilumivu Software for Humanity. https://ilumivu.com/solutions/ecological-momentary-assessment-app/

[8] Knafl, G.J. (2023) Modeling Correlated Outcomes Using Extensions of Generalized Estimating Equations and Linear Mixed Modeling. Springer Nature Switzerland AG, Switzerland. https://doi.org/10.1007/978-3-031-41988-1

[9] Knafl, G.J. (2024) Adaptive Random Effects/Coefficients Modeling. *Open Journal of Statistics*, **14**, 179-206. https://doi.org/10.4236/ojs.2024.142009