

Longitudinal Modeling Approaches to Assess the Association Between Changes in 2 Clinical Outcome Assessments

Therapeutic Innovation
& Regulatory Science

1-7

© The Author(s) 2017

Reprints and permission:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/2168479017731584

tirs.sagepub.com

Dawn Odom, MS¹, Lori McLeod, PhD¹, Bintu Sherif, MS¹,
Lauren Nelson, PhD¹, and David McSorley, MPH¹

Abstract

Background: Understanding how one clinical outcome assessment (COA) (eg, a patient-reported outcome [PRO]) relates to a second COA (eg, a clinician-reported outcome [ClinRO]) may provide insights into disease burden or treatment efficacy. We aimed to briefly review commonly used cross-sectional methods to evaluate the association between a PRO and a ClinRO and to demonstrate the advantages of longitudinal modeling approaches, particularly a joint mixed model for repeated measures (MMRM), to evaluate this association. **Methods:** We generated an example longitudinal data set that included a PRO measured on an 11-point numeric rating scale and a binary ClinRO. The association between change in PRO score and ClinRO response at each time point was examined using 2 cross-sectional analyses: point biserial correlation and logistic regression. We conducted longitudinal analyses of the association between the 2 COAs across time points using MMRM and joint MMRM approaches. **Results:** Point-biserial correlation and logistic regression analyses correctly captured the “built in” associations between the 2 COAs that strengthened over time, but each association was applicable only for a single time point. The MMRM approach provided correlations over time but only for a single outcome variable. The joint MMRM approach modeled the relationship between both outcome variables simultaneously, allowing for evaluation of the correlations both within and between the variables over time. **Conclusion:** Each analysis demonstrated the relationship between PRO score changes and ClinRO response. Longitudinal analysis methods, particularly the joint MMRM, allow for a more thorough examination of the correlations among the 2 outcomes than cross-sectional analysis methods.

Keywords

clinical outcome assessment, clinician-reported outcome, patient-reported outcome, longitudinal analysis

Background

Clinical outcome assessments (COAs), such as patient-reported outcome (PRO) measures and clinician-reported outcome (ClinRO) measures, are routinely used alongside other types of biological markers to assess a patient’s condition in a number of therapeutic areas, including cardiovascular, neurologic, respiratory, dermatologic, and gastrointestinal diseases.^{1,2} Understanding the longitudinal relationships among the variety of COAs included in a clinical trial across time can offer valuable insight into disease burden and the benefits of treatment. While regression-based longitudinal models are commonly used in the evaluation of the treatment effects on a single COA in clinical trials, we posited that regression-based longitudinal models are less commonly used to examine the relationship (typically correlation) between PRO scores and scores on a ClinRO.

We conducted a targeted search of the literature using the keywords “clinical,” “patient-reported outcomes,” and “associations” in PubMed on October 14, 2016. Few articles that were relevant to our topic of interest were identified.

Among studies examining the association between scores on the PRO measure and a ClinRO measure, cross-sectional regression-based models and cross-sectional correlational analyses were the most commonly used approaches. Research examining correlations over time was seldom used; across all the articles reviewed, we found only 2 articles that examined correlation longitudinally. The first used a temporal correlation function, a method that is limited to time-to-event endpoints,³ and the second used latent growth modeling, a method that is traditionally used in the social sciences.⁴ It is well recognized that cross-sectional evaluations of the relationship between

¹ RTI Health Solutions, Research Triangle Park, NC, USA

Submitted 8-Mar-2017; accepted 9-Aug-2017

Corresponding Author:

Dawn Odom, MS, RTI Health Solutions, 200 Park Offices Drive, Research Triangle Park, NC 27709, USA.

Email: dodom@rti.org

2 outcomes have certain limitations: the results typically have lower power and less precision than longitudinal analyses, and they do not allow for an evaluation of the association across time points. Further, these methods assume any missing data is missing completely at random; this assumption rarely holds, and therefore to avoid potential bias, imputation of the missing values prior to analysis may be required.

The mixed model for repeated measures (MMRM) is increasingly being used by the pharmaceutical industry and regulatory agencies⁵ to estimate treatment effects that account for the covariance among within-subject repeated measures. Specification of the covariance structure allows the mixed model to take missing data into account when estimating treatment effects.⁶ Beyond the examination of treatment effects, MMRM can be used to examine the association between 2 outcomes, most simply constructed with one outcome as the dependent variable and the other as a fixed or independent covariate. A key advantage is that all available longitudinal data are used to evaluate the association. More complex longitudinal MMRM models, such as a joint MMRM, can treat both outcomes as random variables (ie, as dependent variables), allowing for a direct estimation of the correlation between the outcomes over time in a single model.

For this study, we aimed to demonstrate the advantages of using longitudinal methods (MMRM and joint MMRM) versus 2 commonly used cross-sectional approaches (regression-based model and simple correlation) when examining the relationship between changes in PRO scores (continuous) and changes in ClinRO scores (dichotomized) using an example data set generated with SAS software. Although we present results based on an example data set, the results are only used as a tool to illustrate the advantages and disadvantages of these methods.

Methods

Example Data Set

We generated an example data set in a manner that allowed us to easily compare the results from cross-sectional analyses with longitudinal analyses (Figure 1). The data set was based on our experience with dermatology clinical studies, which commonly include comparing the change in a PRO measure (continuous) and changes in ClinRO measure (dichotomized). This framework mirrors typical psoriasis studies, in which the primary ClinRO clinical outcome response is defined as a 75% or greater improvement from baseline in the Psoriasis Area Severity Index (PASI) score, and is defined dichotomously as the PASI 75 (with 0 indicating that a 75% or greater improvement from baseline in the PASI score has not been achieved [non-responder] and 1 indicating that a 75% or greater improvement from baseline in the PASI score has been achieved [responder]), and the PRO is the change from baseline (cfb) in pain severity, a continuous variable that is measured using an 11-point numeric rating scale (with 0 indicating no pain and 10 indicating worst imaginable pain). The data set generated in our example included 200 patients, with complete data available (ie, no

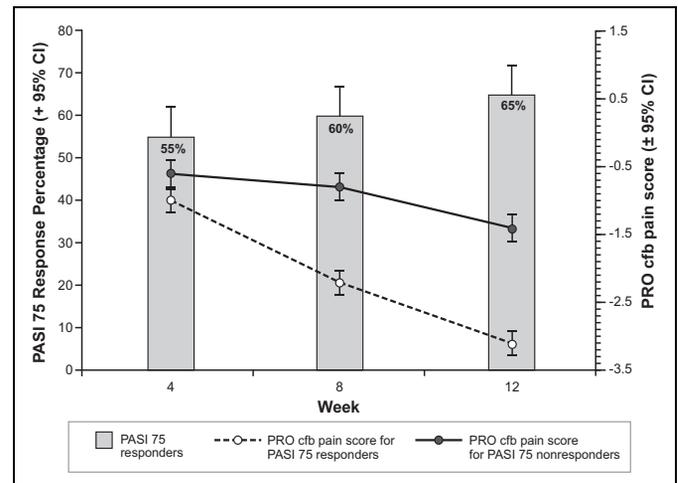


Figure 1. Example data set (N = 200). cfb, change from baseline; CI, confidence interval; PASI, Psoriasis Area Severity Index; PRO, patient-reported outcome.

missing data) for both outcomes measured at baseline and weeks 4, 8, and 12. No treatment was included, and the data set was constructed to easily demonstrate the implications of the different analysis approaches. For the ClinRO, we set the response rate for the binary PASI 75 to be 55% of patients at week 4, 60% at week 8, and 65% at week 12. For the PRO, we randomly generated baseline and cfb values from normal distributions assuming a linear relationship, whereby negative cfb values would represent an improvement in pain from baseline and positive cfb values would represent a worsening of pain from baseline. Patients with a PASI 75 response (PASI 75 responders) were set to have larger decreases in PRO cfb pain scores than patients without a PASI 75 response (PASI 75 nonresponders) at each week. PRO cfb pain scores and PASI 75 response could change from one time point to the next; for instance, a patient could be a PASI 75 responder at one time point but not the following one. There was built-in correlation, both between the ClinRO and the PRO and across time points, as illustrated in Figure 1.

For PASI 75 responders, the mean week 12 PRO cfb pain score was -3.3 (variance = 1.1), and week 12 scores were correlated with week 8 PRO cfb pain scores (correlation $\cong 0.7$) and week 4 PRO cfb pain scores (correlation $\cong 0.5$). For PASI 75 nonresponders, the mean week 12 PRO cfb pain score was -1.4 (variance = 1.15), and week 12 scores were correlated with week 8 PRO cfb pain scores (correlation $\cong 0.7$) and week 4 PRO cfb pain scores (correlation $\cong 0.5$). Fixing these parameters resulted in the following correlations in the example data set between the two outcomes (PASI 75 and PRO cfb pain score at week 4 = -0.36 , week 8 = -0.60 , week 12 = -0.65) and within each outcome (PASI 75 between week 4 and 8 = 0.78, week 4 and 12 = 0.54, week 8 and 12 = 0.75; and PRO cfb pain score between week 4 and 8 = 0.83, week 4 and 12 = 0.74, week 8 and 12 = 0.83).

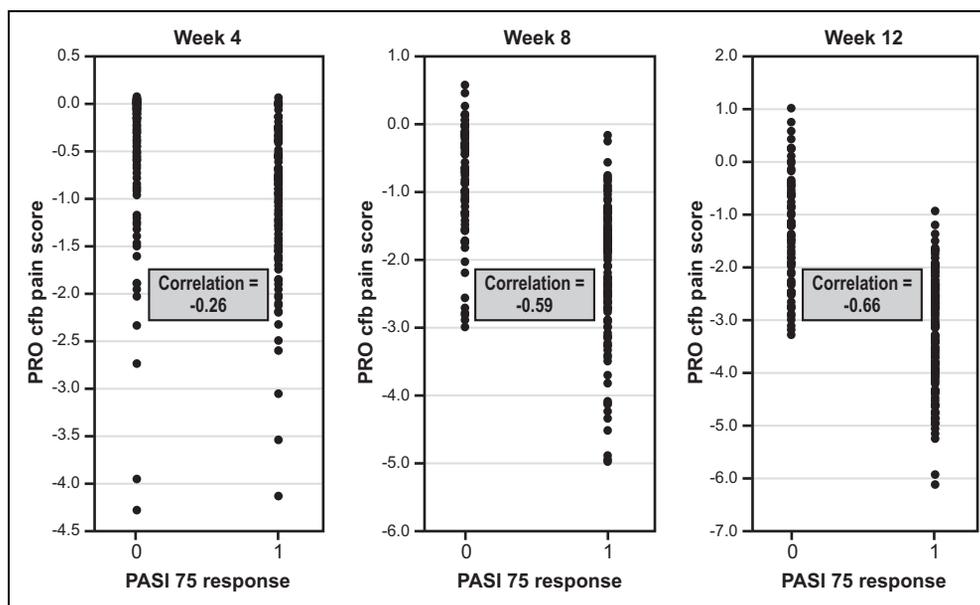


Figure 2. Point-biserial correlation results. cfb, change from baseline; PASI, Psoriasis Area Severity Index; PRO, patient-reported outcome.

Statistical Methods

All analyses were conducted using SAS, version 9.3.

Cross-sectional analyses

We conducted 2 cross-sectional analyses of the example data set at specific time points. These analyses included a point-biserial correlation analysis and a logistic regression analysis. Although other analytic methods were considered (eg, contingency tables; other types of correlation, such as Pearson, Spearman, or Hotelling; other types of regression), point-biserial correlation and logistic regression analyses were chosen because of their frequent application in the literature.

The point-biserial correlation analysis was conducted for each time point (ie, weeks 4, 8, and 12). For the logistic regression analysis, we focused on week 12, although other time points could have been examined (eg, week 4 or week 8). The logistic regression analysis modeled PASI 75 response as the dependent variable and PRO cfb pain score as the independent variable at week 12 to yield predicted probabilities of PASI 75 responses.

Longitudinal analyses

We conducted a longitudinal analysis using an MMRM based on all data in the example data set. This approach accounts for the fact that observations over time from the same patient are correlated via a covariance structure using either a residual error structure (“R side” for repeated measurements) and/or random effects to allow patient-specific intercepts (baseline values) and/or slopes (changes over time). In our example, we used a generalized linear model (GLM), defined such that the dependent variable was PASI 75 response and the independent variables were time point (weeks 4, 8, and 12), cfb in PRO pain scores, and their interaction. A binary response distribution and logit link function was specified for the PASI 75 response. The covariance among within-subject

repeated measurements was estimated using an unstructured covariance pattern, which produced a correlation matrix of the PASI 75 response variable across all the time points adjusted for the other covariates, including the PRO cfb pain variable, in the model. Moreover, the predicted probability of PASI 75 response adjusted for these covariates could be obtained at each time point.

Finally, we conducted a joint MMRM analysis, modeling both outcome variables as the dependent variables together in a longitudinal setting. In a joint MMRM analysis, the key is that the covariance among joint outcomes is accounted for, which then allows for estimation and statistical comparisons of their correlations over time. There are a number of approaches to joint modeling, such as GLMs, conditional models, shared parameter models, and random effects models.⁷ A marginal GLM is one of the simpler approaches because it accounts for the repeated measures through the random error. That is, the marginal GLM incorporates the association between joint outcomes at each time point and across time points in the residual error term using an unstructured covariance matrix. In our example, a marginal GLM was defined such that the joint dependent variables were PASI 75 response and PRO cfb pain score, and the independent variable was time point (weeks 4, 8, and 12). The response distribution and link function were specified for each outcome (ie, binary and logit for PASI 75 response and normal and identity for PRO cfb pain score). The covariance among within-subject repeated measures was estimated using an unstructured covariance pattern. From our joint model, correlation matrices were available that examined (1) the correlation between joint response variables across time point and (2) the correlation within each individual response variable across time point. Moreover, the predicted probability of PASI 75 response and the predicted mean PRO cfb pain value can be obtained at each time point.

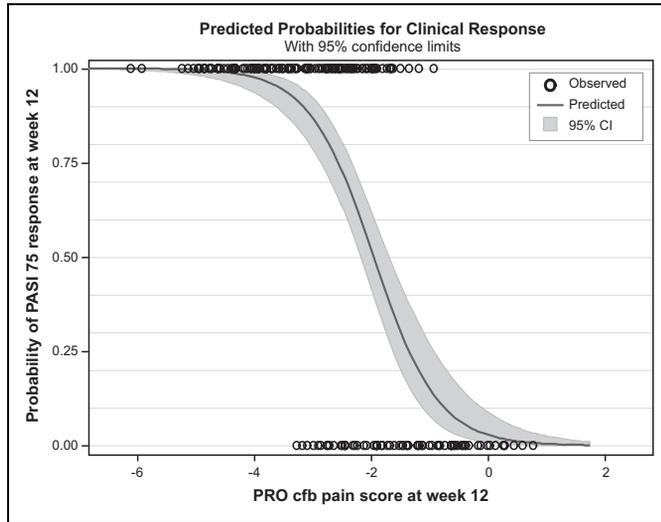


Figure 3. Logistic regression at week 12 results. cfb, change from baseline; CI, confidence interval; PASI, Psoriasis Area Severity Index; PRO, patient-reported outcome.

Results

Cross-Sectional Analyses

Point-biserial correlation analysis

Figure 2 presents the results of the point-biserial correlation analysis. The point-biserial correlation was -0.26 at week 4 and strengthened to -0.66 by week 12, which is very similar to the correlation that was “built in” to the example data set (ie, week 4 = -0.36 and week 12 = -0.65). These results show the relationship between the 2 outcomes at each time point (week 4, week 8, and week 12), but not the relationship of the 2 outcomes across time points.

Logistic regression analysis

Figure 3 presents the predicted probabilities of the cross-sectional logistic regression analysis at week 12. The figure shows the predicted probability of PASI 75 response for different PRO cfb pain scores. With a reduction in the PRO cfb pain score (ie, lower values), there was a higher probability of PASI 75 response. For example, with a PRO cfb in pain score to week 12 of -4.0 , the predicted probability of PASI 75 response at week 12 was 0.98 . Results for weeks 4 and 8 could be obtained similarly, but like the point-biserial correlation analysis above, the methodology is applicable for one time point at a time, not across time points.

Longitudinal Analyses

Mixed model for repeated measures

Figure 4 presents the probability of PASI 75 response predicted by the PRO cfb pain score at each time point from the MMRM model. As a reminder, the data set was created so that the correlations between outcome assessments were greater as time progressed. The MMRM results demonstrate this relationship. For example, a PRO cfb pain score of 1 (worsening) indicated a

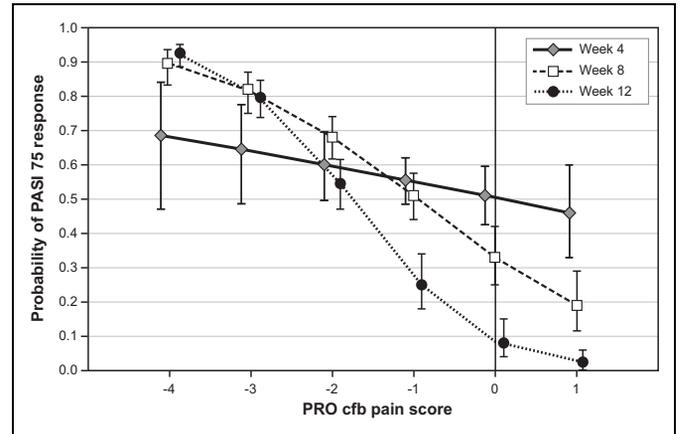


Figure 4. Probability of PASI 75 response by change from baseline (cfb) in PRO pain score at each time point (\pm 95% CI). The results are from the MMRM model. cfb, change from baseline; CI, confidence interval; MMRM, mixed model for repeated measures; PASI, Psoriasis Area Severity Index; PRO, patient-reported outcome.

low-to-moderate probability of PASI 75 response (probability 0.46) at week 4 but very low probability (0.02) at week 12. A PRO change of -4 (improvement) indicated a reasonably high probability of PASI 75 response at week 4 (0.69), but the estimate lacks precision, with a wide confidence interval (CI); in contrast, the probability of PASI 75 response at week 12 was very high (0.94), with good precision and a very tight CI.

Table 1A is a model-based MMRM correlation matrix presenting the correlation of PASI 75 response at each time point for the same set of subjects, adjusted for PRO cfb pain score. For example, the model-based correlation of PASI 75 response between weeks 8 and 12 was 0.53 , adjusted for the PRO cfb pain score. While this analysis provides model-adjusted correlations for PASI 75 response over time, it does not provide the correlation of PRO cfb in pain score at each time point (see Table 1B) or the correlation between PASI 75 response and PRO cfb in pain score over time (see Table 1C).

Joint mixed model for repeated measures

The joint MMRM models the relationship between the 2 clinical outcome variables at each time point, such that both the ClinRO and the PRO are treated as dependent variables. This is similar in concept to the cross-sectional point-biserial correlation analyses (shown in Figure 2). However, unlike the cross-sectional point-biserial analysis, modeling the 2 variables under a joint MMRM model permits a statistical comparison of the correlation between outcome variables at different time points. From the joint MMRM model, 3 model-based correlation matrices across time points are available and are presented in Table 1D-F. The diagonal elements in the joint correlation matrix between the dependent variables across time points shown in Table 1F (ie, week 4 = -0.26 , week 8 = -0.58 , week 12 = -0.66) are similar in magnitude to the point-biserial correlation coefficients in Figure 2; these values are also very

Table 1. MMRM Correlation Matrices.

MMRM Correlation Matrices				Joint MMRM Correlation Matrices			
A. MMRM correlation matrix within PASI 75 response				D. Joint MMRM correlation matrix within PASI 75 response			
	PASI 75 response				PASI 75 response		
PASI 75 response	4	8	12	PASI 75 response	4	8	12
4	1.0			4	1		
8	0.76	1.0		8	0.78	1	
12	0.40	0.53	1.0	12	0.54	0.75	1
B. MMRM correlation matrix within PRO cfb pain score				E. Joint MMRM correlation matrix within PRO cfb pain score			
	PRO cfb pain score				PRO cfb pain score		
PRO cfb pain score	4	8	12	PRO cfb pain score	4	8	12
4				4	1		
8		NA		8	0.80	1	
12				12	0.67	0.84	1
C. MMRM correlation matrix between dependent variables				F. Joint MMRM correlation matrix between dependent variables			
	PRO cfb pain score				PRO cfb pain score		
PASI 75 response	4	8	12	PASI 75 response	4	8	12
4				4	-0.26	-0.38	-0.38
8		NA		8	-0.47	-0.58	-0.56
12				12	-0.59	-0.68	-0.66

Abbreviations: cfb, change from baseline; MMRM, mixed model for repeated measures; NA, not available from the model; PASI, Psoriasis Area Severity Index; PRO, patient-reported outcome.

similar to the correlation between the 2 outcome variables that was “built in” to the example data set at each time point (ie, “built-in”: week 4 = -0.36, week 8 = -0.60, week 12 = -0.65). Based on the joint MMRM, a statistical test of the null hypothesis of equal correlation parameters provides an indication that the coefficients vary across week 4, week 8, and week 12 ($P < .001$).

In addition to the correlation between the 2 dependent variables, the joint MMRM estimates the correlation across time points within each dependent variable. This is in contrast to the previous MMRM model, with PASI 75 response as the dependent variable and PRO cfb in pain score and time points as independent variables, for which only the estimated model-adjusted correlation matrix of PASI 75 response across time points was available (see Table 1A). Table 1 also shows the estimated correlation matrix for PASI 75 response between time points (Table 1D) and the estimated correlation matrix for PRO cfb pain scores between time points (Table 1E). The correlation coefficients for PASI 75 response between time points are identical to the within correlation for the PASI 75 response that was “built in” to the example data set at each time point (ie, week 4 and 8 = 0.78, week 4 and 12 = 0.54, week 8 and 12 = 0.75), and the correlation coefficients for changes in PRO value between time points are very similar to

the within correlation for the PRO cfb pain score that was “built in” to the example data set (ie, week 4 and 8 = 0.83, week 4 and 12 = 0.74, week 8 and 12 = 0.83).

Although one should not statistically compare the results from different, non-nested models, as anticipated, the estimates of the within PASI 75 response correlations reported in Table 1A and the within PASI 75 response correlation matrix from the joint MMRM in Table 1D appear similar based on visual review, with the exception of the correlation between week 8 and week 12 (0.75 and 0.53, respectively).

Lastly, Figure 5 presents the joint MMRM model-based (A) estimated probability of PASI 75 response at each time point and (B) estimated mean PRO cfb pain score at each time point. These results mirror the values defined in the example data set. Because the PASI 75 response and PRO cfb pain score variables are both treated as dependent variables, the joint model does not allow one variable to predict the other variable as available in the traditional MMRM.

Discussion

With clinical trial research placing increasing emphasis on evaluating the patient’s perspective, stakeholders are often interested in understanding the relationship between PRO

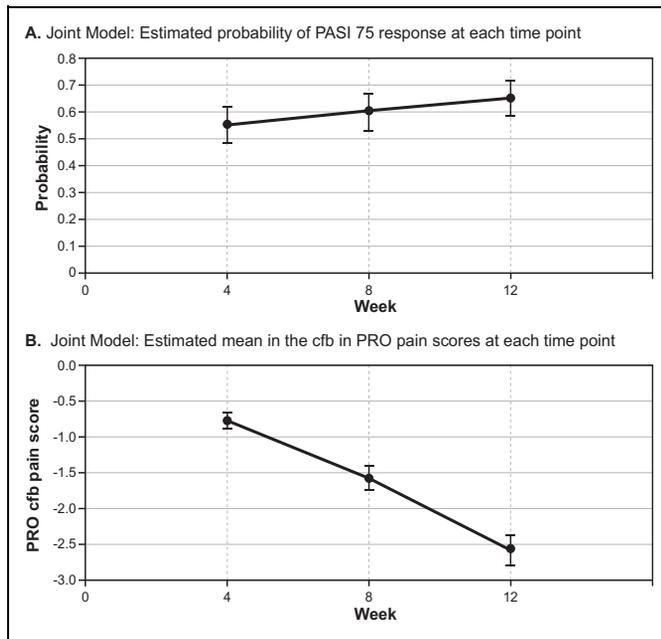


Figure 5. Estimated probability of PASI 75 response at each time point and estimated mean change from baseline (cfb) in PRO pain scores at each time point. The results from the joint MMRM model. cfb, change from baseline; CI, confidence interval; MMRM, mixed model for repeated measures; PASI, Psoriasis Area Severity Index; PRO, patient-reported outcome.

scores and other outcomes such as a response on a ClinRO. Cross-sectional analyses, which are commonly used to evaluate the relationship between PRO measures and ClinRO measures, have certain advantages and limitations. In general, these analyses are simple to execute, are simple to interpret and communicate to various stakeholders, and can be conducted with widely available statistical software. However, the analyses offer only an evaluation of the data at individual time points and do not account for the longitudinal design of most trials. As a result, cross-sectional analyses cannot statistically assess change in correlations across time points in a trial. In our example, the correlation between PASI 75 response and PRO cfb pain scores strengthens over time—a relationship that only the joint MMRM can formally test. With this knowledge, stakeholders can properly communicate that patients with later PASI response are more likely to see improvement in pain. These methods also ignore the type of missing data or require imputation of the missing values prior to analysis. Thus, cross-sectional analyses may not be the most informative method for examining the relationships among various types of measures within a typical clinical trial framework, where data are collected over multiple visits and may be missing. Moreover, there is a risk that results from cross-sectional methods may be misinterpreted by examining the relationship over time based on visual review of the individual time points.

Longitudinal analyses of PRO measures and ClinRO measures also have certain advantages and limitations. In general,

longitudinal analyses use all of the available data and can account for missing data that are missing at random. They can be conducted with established and readily available statistical software such as SAS, which was used to conduct our analyses. However, longitudinal models can be complex to develop and to interpret. MMRM analyses estimate correlations over time for a single dependent variable. These analyses offer predictive modeling over time in which one outcome is the dependent variable and the other is the independent variable. However, a shortcoming of this approach is that correlations between the ClinRO and the PRO cannot be assessed over time. Joint MMRM analyses estimate correlations over time for more than one dependent variable and within one model. For instance, in our example, both the ClinRO and the PRO were treated as dependent variables, each with its own distribution. Importantly, joint MMRM analyses enable statistical testing of correlations over time.

The aim of this study was not to focus on the actual results from the generated example data set, but instead to contrast the cross-sectional statistical methods used to evaluate associations between 2 COAs commonly applied in the literature with longitudinal MMRM methods. In general, the literature evaluating these associations is limited, and we recommend that future analyses incorporate more robust approaches, such as longitudinal MMRM analysis methods. In tandem, we recommend that researchers work to provide clear guidance to help broader audiences better understand the results based on these more sophisticated methods so that the benefits are accessible to stakeholders. For the current study, we generated an example data set and designed analyses in a straightforward manner to clearly portray the relationship of the 2 outcome assessments over time as illustrated through the analyses and the benefit of these proposed alternative methods. Ideally, a more realistic data set could be created that includes both intermittent and monotonically missing data to mirror the typical clinical trial setting, but this was not necessary for our purposes of demonstrating the advantages of MMRM analyses versus cross-sectional methods. Another limitation is that our analysis imposed a linear relationship between the PRO and the ClinRO; however, the use of a linear relationship in our example data set does not limit the general applicability of the methodology. The methodology can be applied to outcome variables with different measurement scales and different relationships between them by choosing the appropriate response distributions (eg, binomial, multinomial, Poisson, negative binomial, lognormal, exponential) and link functions (eg, identity, logit, log) based on the measurement scales and functional forms for each of the respective outcome variables. Therefore, when constructing a model using real-world data, important steps will be to determine the appropriate measurement scale of the outcome variables and to investigate the functional form of the relationship between them and then choose the model, response distributions, and link functions that would best correspond to this relationship.

Conclusions

In addition to clinician-based outcome assessments, there is an increasing regulatory focus on incorporating patient perspectives in drug development.^{8,9} To help stakeholders understand the relevance of PROs, assessing the association of a PRO with a ClinRO is often an important step. We recommend conducting MMRM analyses, particularly the joint MMRM, in addition to the standard cross-sectional analyses in order to provide a better understanding of the relationship of these outcomes over the time course of the study. By doing so, the relationship between both outcomes can be modeled simultaneously, allowing for an examination of the correlations between the 2 dependent variables over time not possible with the cross-sectional methods. In addition, the joint MMRM estimates the correlation across time points within each dependent variable and also allows for statistical comparisons among these correlations. Although the MMRM methodology we recommend is not novel, its application has not been widely used in the literature to examine the association between PROs and ClinROs.

Author Note

A preliminary version of this article was presented in a podium presentation at the International Society for Pharmacoeconomics and Outcomes Research 20th Annual International Meeting; May 16-20, 2015; Philadelphia, PA.

Acknowledgments

Kate Lothman of RTI Health Solutions provided medical writing assistance.

Declaration of Conflicting Interests

No potential conflicts were declared.

Funding

RTI Health Solutions supported this research and the preparation of this publication.

References

1. Flynn KE, Lin L, Ellis SJ, et al; HF-ACTION Investigators. Outcomes, health policy, and managed care: relationships between patient-reported outcome measures and clinical measures in outpatients with heart failure. *Am Heart J*. 2009;158(4 suppl):S64-S71.
2. Mangel AW, Hahn BA, Heath AT, et al. Adequate relief as an endpoint in clinical trials in irritable bowel syndrome. *J Intl Med Res*. 1998;26:76-81.
3. Gong Q, Fang L. A temporal correlation function for the association between patient-reported outcome and clinical endpoints. *Pharm Stat*. 2015;14:377-381.
4. Stull DE, McBride D, Houghton K, Finlay AY, Gnanasakthy A, Balp MM. Assessing changes in chronic spontaneous/idiopathic urticaria: comparisons of patient-reported outcomes using latent growth modeling. *Adv Ther*. 2016;33:214-224.
5. National Research Council. The prevention and treatment of missing data in clinical trials. Panel on Handling Missing Data in Clinical Trials. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press; 2010.
6. Fairclough DL. *Design and Analysis of Quality of Life Studies in Clinical Trials*. Boca Raton, FL: Chapman and Hall/CRC; 2010.
7. Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G. *Longitudinal Data Analysis*. Boca Raton, FL: Chapman and Hall/CRC; 2008.
8. Food and Drug Administration (FDA). Enhancing benefit-risk assessment in regulatory decision-making. <http://www.fda.gov/ForIndustry/UserFees/PrescriptionDrugUserFee/ucm326192.htm>. Accessed February 12, 2016.
9. Peretto EM, Burke L, Oehrlein EM, Epstein RS. Patient-focused drug development: a new direction for collaboration. *Med Care*. 2015;53:9-17.