Supplementary Online Content

Sherman KJ, Cherkin DC, Wellman RD, et al

A randomized trial comparing yoga, stretching, and a self-care book for chronic low back pain. *Arch Intern Med*. 2011. doi:10.1001/archinternmed.2011.524.

eAppendix. Sensitivity analysis for missing data.

eTable 1. Missing data patterns for primary outcomes* and covariates used for adjustment.
eTable 2. Missing data patterns for primary outcomes* and covariates used for adjustment.
eTable 3. Differences in mean Roland Disability Questionnaire (RDQ) and bothersomeness outcomes expressed in terms of the parameters and estimated differences and 95% CIs from model (1) comparing missing data categories and holding all other variables constant.
eTable 4. Comparison of adjusted complete-case analysis and imputed analysis for Roland Disability Questionnaire (RDQ) and bothersomeness outcomes.

eReference

This supplementary material has been provided by the authors to give readers additional information about their work.

eAPPENDIX: Sensitivity Analysis for Missing Data

Description of Missing Data

Overall, 86% (197/228) of participants had complete outcome data at all three follow-up times (eTable 1). In addition, 96% (218/228) of participants had outcome data collected for at least one follow-up time point. The proportion of individuals with complete outcome data across all time points was 86% in the yoga group, 82% in the stretching group and 96% in the usual care group (eTable 1). Of the 31 participants with missing primary outcome values, 10 were missing values only at one time point, 11 were missing values at two time points, and 10 were missing values at all three time points (eTable 1). The sensitivity analyses required participants to have at least one follow-up time point (10 excluded) and all covariate information (3 excluded) for a total of 215 participants included in the analysis.

Sensitivity Analysis Methods

To assess the potential effect of differential missing data across treatment groups, we used a method proposed by Wang and Fitzmaurice (2006) for data imputation in longitudinal studies with non-ignorable non-response[1]. This method requires the specification of a model for the conditional mean of the outcome given the missing data patterns, as well as a model for the marginal mean of the response (in this case, marginal refers to the covariate-adjusted model used in the main analysis). Equations (1) and (2) below give the former and the latter model, respectively, where R_{j1} and R_{j2} are indicators of whether individual i (i=1,…,n) was missing 1 or 2 outcome measurements respectively, Trt_{i1} and Trt_{i2} are indicators for exercise and yoga, respectively, and T_{ij} indicates the follow up time j (j=1,2) for 12 and 26 weeks. Additionally, $\vec{\mathbf{Z}}_i$ for i=1…n is a vector of baseline covariates used for adjustment. In both models (1) and (2), $\vec{\mathbf{Z}}_i$ includes baseline Roland Disability Questionnaire (RDQ) score, baseline symptom bothersomeness score, age, sex, body mass index (BMI), days of lower back pain in the last six months, pain travelling down the leg and employment-related exertion.

$$E(Y_{ij} \mid R_{i1}, R_{i2}, Trt_{i1}, Trt_{i2}, T_{ij}, \vec{\mathbf{Z}}_i) = \alpha_0 + \alpha_1 Trt_{i1} + \alpha_2 Trt_{i2} + \alpha_3 T_{i1} + \alpha_4 T_{i2} + \alpha_5 R_{i1} + \alpha_6 R_{i2} + \alpha_7 Trt_{i1} \times T_{i1} + \alpha_8 Trt_{i1} \times T_{i2} + \alpha_9 Trt_{i2} \times T_{i1} + \alpha_{10} Trt_{i2} \times T_{i2} + \alpha_{11} Trt_{i1} \times R_{i1} + \alpha_{12} Trt_{i2} \times R_{i1} + \alpha_{13} Trt_{i2} \times R_{i2} + \vec{\boldsymbol{\alpha}} \vec{\mathbf{Z}}_i \quad (1)$$

$$E(Y_{ij} \mid Trt_{i1}, Trt_{i2}, T_{ij}, \vec{\mathbf{Z}}_i) = \beta_0 + \beta_1 Trt_{i1} + \beta_2 Trt_{i2} + \beta_3 T_{i1} + \beta_4 T_{i2} + \beta_5 Trt_{i1} \times T_{i1} + \beta_6 Trt_{i1} \times T_{i2} + \beta_7 Trt_{i2} \times T_{i1} + \beta_8 Trt_{i2} \times T_{i2} + \vec{\beta} \vec{\mathbf{Z}}_i \quad (2)$$

Data were imputed with the predicted values from model (1) and then analyzed using model (2). It is worth noting that certain comparisons between groups with different numbers of missing data points are not possible. For example, the usual care group is only observed to be missing data at one of the three follow up times. Therefore, there is no estimable comparison of the difference in mean outcome resulting from missing more than one outcome measure for either the yoga group or the exercise group compared to the usual care group. From model (1), the effects (coefficients) of having one or two missing outcome measurements on mean outcome holding other variables constant are given by equations (3) and (4) respectively.

$$(\alpha_5 + \alpha_{11} Trt_{i1} + \alpha_{12} Trt_{i2}) R_{i1} \quad (3)$$

$$(\alpha_6 + \alpha_{13} Trt_{i2}) R_{i2} \quad (4)$$

eTable 3 provides a summary of the effects of missing data patterns on mean outcomes in terms of model parameters and provides estimates and 95% CIs obtained from fitting model (1) for both the RDQ and the bothersomeness outcomes. In equation (3), α_5 is the difference in mean outcome for those having missing data at one follow-up time point versus those having no missing outcomes within the usual care group. Accordingly, $\alpha_5 + \alpha_{11}$ and $\alpha_5 + \alpha_{12}$ are the differences in mean outcome for those with missing data at one time point versus those with no missing outcome data for the exercise and yoga groups, respectively. From equation (4), we can see that α_6 is the difference in mean outcome

for those missing two follow-up measurements versus those with fully observed outcomes within the stretching group. Correspondingly, $\alpha_6 + \alpha_{13}$ quantifies the difference in mean outcome for the yoga group attributable to missing two outcome measurements compared to having no missing outcomes. Differences in average outcomes for groups missing two outcomes measurements versus those missing one outcome measurement within the yoga and stretching groups are given by $\alpha_6 + \alpha_{13} - (\alpha_5 + \alpha_{12})$ and $\alpha_6 - (\alpha_5 + \alpha_{11})$, respectively. For consistency with the complete case analysis, all omnibus tests were conducted using the score test, while all pairwise comparisons and confidence intervals are Wald-based.

Sensitivity Analysis Results

eTable 4 shows adjusted estimates of mean RDQ and bothersomeness scores estimated with the 2-step GEE imputation method versus the complete case analysis. Overall, the estimates and corresponding 95% CIs obtained from the imputed data are the same, or very slightly lower, than those obtained from the complete-case analysis. The omnibus score statistics vary between the two sets of results but the inference does not change. Pairwise comparisons computed in the presence of a significant omnibus test also yield comparable inference in each case. Estimates of the effects of the missing data patterns on mean outcomes are given in eTable 3 and show that the small amount of missing data observed in this study leads to very little precision for estimating the effect of the missing data. For example, having one missing outcome point versus having fully observed outcomes in the yoga group is consistent with a change anywhere between a decrease in mean RDQ of 6.1 points to an increase of 3 points. Based on this analysis we conclude that imputed analysis does not add substantially to the understanding of the data and we have therefore chosen to report the simpler complete-case results in the main body of this manuscript following our a-priori specified analysis plan.

eTable 1. Missing data patterns for primary outcomes* and covariates used for adjustment.

	Yoga	Stretching	Usual Care	Total
Number of missing outcomes				
0	79 (85.9%)	75 (82.4%)	43 (95.6%)	197 (86.4%)
1	3 (3.3%)	5 (5.5%)	2 (4.4%)	10 (4.4%)
2	5 (5.4%)	6 (6.6%)		11 (4.8%)
3	5 (5.4%)	5 (5.5%)		10 (4.4%)
	92	91	45	228

eTable 2. Missing data patterns for primary outcomes* and covariates used for adjustment.

Missing Data Pattern (1=Missing)					Group, n (%)			
Baseline	6 weeks	12 weeks	26 weeks	Covariates	Yoga	Stretching	Usual Care	Total
0	0	0	0	0	78 (84.8)	74 (81.3)	42 (93.3)	194 (85.0)
0	0	0	0	1	1 (1.1)	1 (1.1)	1 (2.2)	3 (1.3)
0	1	0	0	0	2 (2.2)	4 (4.4)	1 (2.2)	7 (3.1)
0	0	1	0	0	1 (1.1)	0 (0.0)	1 (2.2)	2 (0.9)
0	0	0	1	0	0 (0.0)	1 (1.1)	0 (0.0)	1 (0.4)
0	1	1	0	0	1 (1.1)	1 (1.1)	0 (0.0)	2 (0.9)
0	0	1	1	0	4 (4.4)	4 (4.4)	0 (0.0)	8 (3.5)
0	1	1	1	0	5 (5.4)	4 (4.4)	0 (0.0)	9 (4)
0	1	0	1	0	0 (0.0)	1 (1.1)	0 (0.0)	1 (0.4)
0	1	1	1	1	0 (0.0)	1 (1.1)	0 (0.0)	1 (0.4)
					92	91	45	228

*Notes: 1) RDQ and bothersomeness outcomes have the same pattern of missing data. 2) Covariates refer to the set of adjustment variables used through all presented analyses: baseline value of outcome, age, sex, baseline Roland Disability Questionnaire (RDQ) and bothersomeness scores, body mass index, physical work demands, and whether pain travels down the leg.

eTable 3. Differences in mean Roland Disability Questionnaire (RDQ) and bothersomeness outcomes expressed in terms of the parameters and estimated differences and 95% CIs from model (1) comparing missing data categories and holding all other variables constant.

Number Missing Outcome Measures			Yoga	Stretching	Usual Care
1	vs.	0	$\alpha_5 + \alpha_{12}$	$\alpha_5 + \alpha_{11}$	α_5
RDQ			-1.6 (-6.1, 3.0)	0.2 (-2.0, 2.4)	1.0 (-0.2, 2.1)
Bothersomeness			-0.8 (-2.7, 1.2)	-0.3 (-1.6, 1.0)	-1.3 (-2.2, -0.4)
2	vs.	0	$\alpha_6 + \alpha_{13}$	α_6	NA
RDQ			3.4 (-1.2, 8.1)	1.4 (-0.6, 3.4)	
Bothersomeness			-0.03 (-3.1, 3.0)	1.2 (-0.8, 3.3)	
2	vs.	1	$\alpha_6 + \alpha_{13} - (\alpha_5 + \alpha_{12})$	$\alpha_6 - (\alpha_5 + \alpha_{11})$	NA
RDQ			5.0 (-1.5, 11.5)	1.2 (-1.7, 4.0)	
Bothersomeness			0.7 (-2.8, 4.3)	1.5 (-0.9, 3.9)	

eTable 4. Comparison of adjusted complete-case analysis and imputed analysis for Roland Disability Questionnaire (RDQ) and bothersomeness outcomes.

	Yoga (Y)			Stretching (S)			Self- Care (SC)			P-value (Y-S,Y-SC,S-SC)	Omnibus P-value
TOTAL, N	92			91			45				
6-week Outcomes											
RDQ Score											
Adjusted mean 2-step (95% CI)	5.8	(4.9,	6.7)	5.3	(4.7,	6.0)	7.1	(5.9,	8.2)	0.40, 0.089, 0.009	0.023
Adjusted mean (95% CI)	6.0	(5.1,	6.9)	5.5	(4.9,	6.1)	7.3	(6.1,	8.4)	0.36, 0.098, 0.009	0.040
12-week Outcomes											
RDQ Score											
Adjusted mean 2-step (95% CI)	4.3	(3.4,	5.1)	4.5	(3.7,	5.2)	6.6	(5.6,	7.6)	0.71, 0.002, 0.001	<0.001
Adjusted mean (95% CI)‡	4.3	(3.6,	5.1)	4.6	(3.9,	5.3)	6.8	(5.8,	7.8)	0.57, <.0001, <.001	0.001
26-week Outcomes											
RDQ Score											
Adjusted mean 2-step (95% CI)	4.1	(3.1,	5.0)	4.3	(3.5,	5.2)	5.7	(4.7,	6.8)	0.61, 0.026, 0.036	0.040
Adjusted mean (95% CI)	4.1	(3.3,	5.0)	4.5	(3.6,	5.3)	5.9	(4.9,	7.0)	0.56, 0.007, 0.026	0.026

	Yoga (Y)			Stretching (S)			Self- Care (SC)			P-value (Y-S,Y-SC,S-SC)	Omnibus P-value
TOTAL, N	92			91			45				
6-week Outcomes											
Bothersomeness Score											
Adjusted mean 2-step (95% CI)	3.8	(3.4,	4.3)	3.8	(3.4,	4.2)	4.0	(3.4,	4.6)	NA	0.745
Adjusted mean (95% CI)	3.9	(3.5,	4.4)	3.9	(3.5,	4.2)	4.1	(3.5,	4.7)	NA	0.829
12-week Outcomes											
Bothersomeness Score											
Adjusted mean 2-step (95% CI)	3.1	(2.6,	3.5)	3.6	(3.2,	4.1)	4.1	(3.6,	4.7)	0.045, <.001, 0.107	0.003
Adjusted mean (95% CI)	3.2	(2.8,	3.6)	3.7	(3.2,	4.1)	4.3	(3.7,	4.8)	0.095, 0.002, 0.100	0.010
26-week Outcomes											
Bothersomeness Score											
Adjusted mean 2-step (95% CI)	3.4	(2.9,	3.9)	3.4	(2.9,	3.9)	3.8	(3.1,	4.4)	NA	0.553
Adjusted mean (95% CI)	3.5	(3.0,	3.9)	3.4	(3.0,	3.9)	3.9	(3.2,	4.5)	NA	0.509

References

1. Molin Wang, Garrett M. Fitzmaurice. A Simple Imputation Method for Longitudinal Studies with Non-ignorable Non-response. Biometrical Journal, Volume 48, Issue 2, pages 302-318, April 2006.