

1 Retrieval of Individual Participant Data for a Meta-Analysis

2

3 Running Head: Retrieval of Individual Participant Data

4

5 Type of Article: Short Communication

6

7 Abstract Word Count: 192

8 Manuscript Word Count (Introduction, Methods, Results, Discussion, Conclusions: 1519

9

10 Corresponding Author: Dr. George A. Kelley<sup>a</sup>, FACSM, Professor and Director,  
11 Meta-Analytic Research Group, <sup>a</sup>School of Public Health, Department of Biostatistics,  
12 Robert C. Byrd Health Sciences Center, West Virginia University, Morgantown, WV  
13 26506-9190, Office Phone: 304-293-6279, Fax: 304-293-5891, E-mail:  
14 [gkelley@hsc.wvu.edu](mailto:gkelley@hsc.wvu.edu)

15 Co-Author: Kristi S. Kelley<sup>a</sup>, M.Ed., Research Instructor, <sup>a</sup>School of Public Health,  
16 Department of Biostatistics, Robert C. Byrd Health Sciences Center,  
17 West Virginia University, PO Box 9190, Morgantown, WV 26506-9190, Office Phone:  
18 304-293-6280, Fax: 304-293-5891, E-mail: [kskelley@hsc.wvu.edu](mailto:kskelley@hsc.wvu.edu)

## 19 **Abstract**

20 **Purpose:** Despite the potential for individual participant data (IPD) meta-analysis to  
21 yield more valid results than aggregate data (AD) meta-analysis, the feasibility of  
22 obtaining IPD is not well established. The purpose of this study was to add to that  
23 literature. **Methods:** Using data from a previous meta-analysis of 29 studies on exercise  
24 in adults with arthritis and other rheumatic disease, the percentage of studies in which  
25 useable IPD was provided was calculated. Exact logistic regression was used to  
26 examine the association between the retrieval of IPD with year of publication and  
27 country in which the study was conducted (USA versus Other) included as potential  
28 predictors. A two-tailed alpha value  $\leq 0.05$  was considered statistically significant.

29 **Results:** Eight of 29 authors (27.6%, 95% CI = 11.3% to 43.8%) provided useable IPD.  
30 Neither year of publication (odds ratio = 1.05, 95% CI = 0.90 to 1.27,  $p = 0.58$ ) nor  
31 country (odds ratio = 1.36, 95% CI = 0.20 to 10.9,  $p = 1.00$ ) was significantly associated  
32 with the obtainment of IPD. **Conclusions:** Aggregate data meta-analysis may be  
33 preferable to IPD meta-analysis. However, further research is warranted before any  
34 definitive recommendations can be made.

35

36 **Key words:** meta-analysis; methods; aggregate data; individual participant data;  
37 individual patient data; systematic review

38

## 39 **Abbreviations**

40 IPD – Individual participant data

41 AD – aggregate data

## 42 **1. Introduction**

43 The prevalence of meta-analyses has increased substantially over approximately the  
44 past 25 years. For example, a recent PubMed search by the first author on February 21,  
45 2015 using the keyword “meta-analysis” found that the number of citations increased  
46 from 331 in 1990 to 14,329 in 2014 (unpublished results). While aggregate data (AD)  
47 meta-analysis, an approach in which summary statistics from eligible studies are  
48 pooled, is still the most common type of meta-analysis, individual participant data (IPD)  
49 meta-analysis pools the results from different studies based on data from each  
50 participant [1]. Two of the potential advantages of IPD meta-analyses are the ability to  
51 examine covariates at the individual level as well as a potential reduction in reporting  
52 and publication biases [1-4]. However, a major disadvantage may be the ability to  
53 retrieve IPD from eligible studies, thus leading to a smaller number of included trials and  
54 potentially biased results on the topic of interest. The investigative team has previously  
55 reported the retrieval of IPD from only 29 of 76 (38.2%) eligible studies dealing with the  
56 effects of exercise on bone mineral density in adults [5]. However, this study was  
57 conducted approximately 13 years ago. Since that time, technological advances have  
58 improved one’s ability to store and share data, thereby making it easier to share de-  
59 identified IPD with others. In addition, knowledge regarding the feasibility of obtaining  
60 IPD, an important factor when making decisions about whether an AD or IPD meta-  
61 analysis should be conducted, is not well established [6-8]. The purpose of this short  
62 communication was to add to this literature.

## 63 **2. Materials and Methods**

64 *2.1. Data source.* Data were derived from a recently published AD meta-analysis in  
65 which an IPD meta-analysis was originally planned, details of which have been  
66 described elsewhere [9]. Briefly, studies were included if they were randomized  
67 controlled trials examining the effects of exercise (aerobic, strength training, or both) on  
68 depressive symptoms in adults with arthritis and other rheumatic diseases [9]. Twenty-  
69 nine studies representing 2,449 participants (1,470 exercise, 979 control) were included  
70 [10-38].

71 *2.2. Retrieval of IPD.* Using a pre-defined form letter (see Supplementary File 1), de-  
72 identified IPD was requested by having the second author contact the corresponding  
73 author of each eligible study via electronic mail asking if they would be interested in  
74 providing IPD. A response was requested within two weeks, regardless of interest, with  
75 the authors being informed that they would be listed in the acknowledgements section  
76 of each published study derived from the project if they provided their IPD. If no  
77 response was received within two weeks, up to two additional requests were sent via  
78 electronic mail. For those who responded but chose not to contribute, reasons given for  
79 not participating, if any, were recorded. For those authors who expressed interest in  
80 providing IPD, a second electronic mail was sent that included an attachment consisting  
81 of a pre-defined list of IPD requested (see Supplementary File 2). Investigators were  
82 asked to provide IPD in a format that was convenient for them within four weeks. If IPD  
83 was not received within four weeks, as many as two additional reminders were sent via  
84 electronic mail. The dates of all communications were recorded.

85 *2.3. Statistical Analysis.* Descriptive statistics were used to describe the number of  
86 responses to initial electronic mail requests for IPD, number of days to respond to initial

87 electronic mail requests, number of authors who provided useable IPD, and number of  
88 days from initial requests to receipt of IPD. Reasons for not providing IPD from authors  
89 who were willing to supply such information were also recorded. Furthermore,  
90 descriptive statistics were calculated for the two potential predictors included in the  
91 regression model.

92 Because of the small sample size, exact logistic regression was used to examine for  
93 potential predictors with respect to whether or not IPD was received [39]. Based on  
94 previous research [5], the two potential predictors included in the model were country in  
95 which the study was conducted and year that the study was published. The chi-square  
96 distribution ( $\chi^2$ ) was used to examine the overall model. The alpha level for statistical  
97 significance was set at  $p \leq 0.05$ .

### 98 **3. Results**

99 *3.1. Descriptive Statistics for Retrieval of IPD.* The authors from 20 of 29 studies  
100 (69.0%, 95% CI = 52.2% to 85.8%) responded to initial electronic mail requests for IPD  
101 while 9 (31.0%, 95% CI = 14.2% to 47.8%) never responded despite multiple requests.  
102 The response time to initial requests varied widely from 1 to 181 days ( $\bar{X} \pm SD = 54.2 \pm$   
103  $74.8$ , 95% CI = 21.4 to 87.0, Mdn = 17). Eight of 29 authors (27.6%, 95% CI = 11.3% to  
104 43.8%) provided useable IPD. The number of days from initial requests for data to  
105 receipt of IPD ranged from 36 to 179 ( $\bar{X} \pm SD = 74.4 \pm 46.4$ , 95% CI = 42.0 to 106.8, Mdn  
106 = 64). Reasons for not providing IPD included not having the data any longer (n=4) and  
107 time (n=1). Another author said they would supply IPD if a consortium was formed and  
108 in which they were included as a co-author. Year of publication ranged from 1989 to  
109 2011. Thirteen studies (44.8%, 95% CI = 26.7% to 62.9%) were conducted in the United

110 States while the remaining 16 (55.2%, 95% CI = 37.1% to 73.3%) were conducted in  
111 countries other than the United States.

112 *3.2. Potential Predictors in the Obtainment of IPD.* Results for exact logistic regression  
113 are shown in Table 1. The overall model was not statistically significant ( $\chi^2 = 0.62$ ,  $p =$   
114  $0.71$ ) and neither year of publication nor country were significant predictors for the  
115 receipt of IPD ( $p > 0.05$  for both).

#### 116 **4. Discussion**

117 *4.1. Overall Findings.* The current study suggests that the obtainment of IPD was low,  
118 with less than one third of eligible studies providing such. The inability to obtain IPD  
119 from the majority of eligible studies may bias results and limit one from conducting  
120 analyses based on individual versus grouped data, one of the very reasons for  
121 conducting an IPD meta-analysis [1,4]. While methods exist for combining IPD and AD  
122 [8,40], the investigative team was not comfortable using those methods because of the  
123 inability to obtain the data necessary to replicate the results reported in this previous  
124 work (Dr. Richard Riley, personal electronic mail communication, July 11, 2012, Dr. Jan  
125 Staessen, personal electronic mail communication, July 30, 2012) [40].

126 The response rates observed are either similar to [5,6], lower [7,8] or higher [41] than  
127 previous research. Most notably, this latter study was only able to retrieve IPD for 15%  
128 of eligible participants [41]. Possible reasons for the lower response rates observed in  
129 the current study include the lack of a consortium, larger number of eligible studies and  
130 older publication years for eligible studies. The lack of association between year and  
131 country with the obtainment of IPD are in contrast with previous research where a trend  
132 was found for both to be associated with the retrieval of IPD [5]. One possible reason

133 for this discrepancy may have been the smaller number of studies included (29 versus  
134 76) [5].

135 Finally, as previously reported [9], no statistically significant or clinically important  
136 differences were found in depressive symptoms between those studies that supplied  
137 IPD versus those that did not, a finding consistent with previous research [5].

138 *4.2. Implications for Research and Practice.* Given that this was a case study, a need  
139 exists for additional work that includes multiple IPD meta-analyses. This should include  
140 data regarding the time and costs involved in conducting an IPD meta-analysis,  
141 something that was beyond the scope of the current investigation.

142 Given the inability to obtain the majority of IPD, an AD meta-analysis may be  
143 preferable. However, if one chooses to conduct an IPD meta-analysis, the length of  
144 time to retrieve IPD needs to be considered. In addition, while the development of a  
145 consortium between authors of the original studies may increase the amount of IPD  
146 obtained, this should be balanced with the additional time involved. Most notably, any  
147 potential benefit of an IPD meta-analysis should be considered with respect to the  
148 increased costs. For example, Steinberg et al., estimated that the costs associated with  
149 conducting a meta-analysis of 12 studies was more than 5 times greater using the IPD  
150 versus AD approach [42] while others estimated the costs of this same study to be at  
151 least 8 times greater given that the investigative team continued to work on the study  
152 after funding for the project ended [43].

153 *4.3. Strengths and Potential Limitations Strengths.* To the best of the authors'  
154 knowledge, this is one of the most recent studies regarding the feasibility of retrieving  
155 IPD. However, since the study focused on one attempted IPD meta-analysis, the

156 findings may not be generalizable to other IPD meta-analyses. In addition, no cost data  
157 were collected or analyzed, thereby limiting the applicability of results. Finally, the  
158 collapsing of countries other than the US into one category because of the small  
159 number of studies available for each country could have biased the results.

## 160 **5. Conclusions**

161 The results of the current study suggest that an AD meta-analysis may be preferable to  
162 an IPD meta-analysis. However, further research is warranted before any definitive  
163 recommendations can be made.

## 164 **Acknowledgments and Funding**

165 The authors would like to thank the following people for supplying us with IPD: Susan  
166 Buckelew, Jennifer L. Etnier, Marlene Fransen, A. Hakkinen, Geri B. Neuberger, Sheila  
167 C. O'Reilly, Michael Doherty, Joanna Romawski, Mike E. Miller, Brenda Penninx,  
168 William B. Applegate, Candace L. Schachter, Ayse Karan, Sureyya Sencan, Jamil  
169 Natour, Valerie Valim, and H. Valkeinen.

170  
171 This work was supported by the National Institute of Arthritis and Musculoskeletal and  
172 Skin Diseases of the National Institutes of Health (grant number RO1 AR061346) and  
173 the National Institute of General Medical Sciences of the National Institutes of Health  
174 (grant number U54 GM104942). The content is solely the responsibility of the authors  
175 and does not necessarily represent the official views of the National Institutes of Health.

176

## 177 **Conflict of Interest Statement**

178 The Authors declare that there is no conflict of interest.

179 **References**

- 180 [1] Stewart LA, Parmar MKB. Meta-analysis of the literature or of individual patient  
181 data: is there a difference? *The Lancet* 1993;341:418-22.
- 182 [2] Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data:  
183 rationale, conduct, and reporting. *BMJ* 2010;340:c221.
- 184 [3] Riley RD. Commentary: Like it and lump it? Meta-analysis using individual  
185 participant data. *Int J Epidemiol* 2010 Jul 26;39(5):1359-61.
- 186 [4] Riley RD, Kauser I, Bland M, et al. Meta-analysis of randomised trials with a  
187 continuous outcome according to baseline imbalance and availability of individual  
188 participant data. *Stat Med* 2013;32(16):2747-66.
- 189 [5] Kelley GA, Kelley KS, Tran ZV. Retrieval of individual patient data for an exercise  
190 meta-analysis. *Am J Med Sports* 2002;4:350-4.
- 191 [6] Duval S, Vazquez G, Baker WL, et al. The collaborative study of obesity and  
192 diabetes in adults (CODA) project: meta-analysis design and description of  
193 participating studies. *Obes Rev* 2007;8(3):263-76.
- 194 [7] Piepoli MF, Davos C, Francis DP, et al. Exercise training meta-analysis of trials  
195 in patients with chronic heart failure (ExTraMATCH). *BMJ* 2004 Jan  
196 24;328(7433):189.
- 197 [8] Riley RD, Simmonds MC, Look MP. Evidence synthesis combining individual  
198 patient data and aggregate data: a systematic review identified current practice  
199 and possible methods. *J Clin Epidemiol* 2007;60(5):431-9.
- 200 [9] Kelley G.A., Kelley KS, Hootman JM. Effects of exercise on depression in adults  
201 with arthritis: A systematic review with meta-analysis of randomized controlled  
202 trials. *Arthritis Res Ther* 2015;in press.

- 203 [10] Alentorn-Geli E, Padilla J, Moras G, et al. Six weeks of whole-body vibration  
204 exercise improves pain and fatigue in women with fibromyalgia. *J Altern*  
205 *Complement Med* 2008 Oct;14(8):975-81.
- 206 [11] Buckelew SP, Conway R, Parker J, et al. Biofeedback/relaxation training and  
207 exercise interventions for fibromyalgia: a prospective trial. *Arthritis Rheum* 1998  
208 Jun;11(3):196-209.
- 209 [12] Daltroy LH, Robb-Nicholson C, Iversen MD, et al. Effectiveness of minimally  
210 supervised home aerobic training in patients with systemic rheumatic disease. *Br*  
211 *J Rheumatol* 1995 Nov;34(11):1064-9.
- 212 [13] Etnier JL, Karper WB, Gapin JI, et al. Exercise, fibromyalgia, and fibrofog: a pilot  
213 study. *J Phys Act Health* 2009 Mar;6(2):239-46.
- 214 [14] Fransen M, Nairn L, Winstanley J, et al. Physical activity for osteoarthritis  
215 management: a randomized controlled clinical trial evaluating hydrotherapy or  
216 Tai Chi classes. *Arthritis Rheum* 2007 Apr 15;57(3):407-14.
- 217 [15] Fontaine KR, Conn L, Clauw DJ. Effects of lifestyle physical activity on perceived  
218 symptoms and physical function in adults with fibromyalgia: results of a  
219 randomized trial. *Arthritis Res Ther* 2010;12(2):R55.
- 220 [16] Gowans SE, deHueck A, Voss S, et al. Effect of a randomized, controlled trial of  
221 exercise on mood and physical function in individuals with fibromyalgia. *Arthritis*  
222 *Rheum* 2001 Dec;45(6):519-29.
- 223 [17] Haak T, Scott B. The effect of Qigong on fibromyalgia (FMS): a controlled  
224 randomized study. *Disabil Rehabil* 2008;30(8):625-33.

- 225 [18] Hakkinen A, Hakkinen K, Hannonen P, et al. Strength training induced  
226 adaptations in neuromuscular function of premenopausal women with  
227 fibromyalgia: comparison with healthy women. *Ann Rheum Dis* 2001  
228 Jan;60(1):21-6.
- 229 [19] Ide MR, Laurindo IMM, Rodrigues AL, et al. Effect of aquatic respiratory  
230 exercise-based program in patients with fibromyalgia. *Int J Rheum Dis*  
231 2008;11(2):131-40.
- 232 [20] Jones KD, Burckhardt CS, Deodhar AA, et al. A six-month randomized controlled  
233 trial of exercise and pyridostigmine in the treatment of fibromyalgia. *Arthritis*  
234 *Rheum* 2008 Feb;58(2):612-22.
- 235 [21] Komatireddy GR, Leitch RW, Cella K, et al. Efficacy of low load resistive muscle  
236 training in patients with rheumatoid arthritis functional class II and III. *J*  
237 *Rheumatol* 1997 Aug;24(8):1531-9.
- 238 [22] Mannerkorpi K, Nyberg B, Ahlmen M, et al. Pool exercise combined with an  
239 education program for patients with fibromyalgia syndrome. A prospective,  
240 randomized study. *J Rheumatol* 2000 Oct;27(10):2473-81.
- 241 [23] Minor MA, Hewett JE, Webel RR, et al. Efficacy of physical conditioning exercise  
242 in patients with rheumatoid arthritis and osteoarthritis. *Arthritis Rheum* 1989  
243 Nov;32(11):1396-405.
- 244 [24] Neuberger GB, Aaronson LS, Gajewski B, et al. Predictors of exercise and  
245 effects of exercise on symptoms, function, aerobic fitness, and disease outcomes  
246 of rheumatoid arthritis. *Arthritis Rheum* 2007 Aug 15;57(6):943-52.

- 247 [25] O'Reilly SC, Muir KR, Doherty M. Effectiveness of home exercise on pain and  
248 disability from osteoarthritis of the knee: a randomised controlled trial. *Ann*  
249 *Rheum Dis* 1999 Jan;58(1):15-9.
- 250 [26] Patrick DL, Ramsey SD, Spencer AC, et al. Economic evaluation of aquatic  
251 exercise for persons with osteoarthritis. *Med Care* 2001 May;39(5):413-24.
- 252 [27] Penninx BW, Rejeski WJ, Pandya J, et al. Exercise and depressive symptoms: a  
253 comparison of aerobic and resistance exercise effects on emotional and physical  
254 function in older persons with high and low depressive symptomatology. *J*  
255 *Gerontol B Psychol Sci Soc Sci* 2002 Mar;57(2):124-32.
- 256 [28] Rooks DS, Gautam S, Romeling M, et al. Group exercise, education, and  
257 combination self-management in women with fibromyalgia: a randomized trial.  
258 *Arch Int Med* 2007 Nov 12;167(20):2192-200.
- 259 [29] Sanudo B, Galiano D, Carrasco L, et al. Effects of a prolonged exercise program  
260 on key health outcomes in women with fibromyalgia: a randomized controlled  
261 trial. *J Rehabil Med* 2011 May;43(6):521-6.
- 262 [30] Schachter CL, Busch AJ, Peloso PM, et al. Effects of short versus long bouts of  
263 aerobic exercise in sedentary women with fibromyalgia: a randomized controlled  
264 trial. *Phys Ther* 2003 Apr;83(4):340-58.
- 265 [31] Sencan S, Ak S, Karan A, et al. A study to compare the therapeutic efficacy of  
266 aerobic exercise and paroxetine in fibromyalgia syndrome. *J Back Musculoskelet*  
267 *Rehabil* 2004 Apr;17(2):57-61.

- 268 [32] Tomas-Carus P, Gusi N, Leal A, et al. [The fibromyalgia treatment with physical  
269 exercise in warm water reduces the impact of the disease on female patients'  
270 physical and mental health]. *Rheumatol Clin* 2007 Jan;3(1):33-7.
- 271 [33] Tomas-Carus P, Gusi N, Hakkinen A, et al. Eight months of physical training in  
272 warm water improves physical and mental health in women with fibromyalgia: a  
273 randomized controlled trial. *J Rehabil Med* 2008 Apr;40(4):248-52.
- 274 [34] Valim V, Oliveira L, Suda A, et al. Aerobic fitness effects in fibromyalgia. *J*  
275 *Rheumatol* 2003 May;30(5):1060-9.
- 276 [35] Valkeinen H, Alen M, Hannonen P, et al. Changes in knee extension and flexion  
277 force, EMG and functional capacity during strength training in older females with  
278 fibromyalgia and healthy controls. *Rheumatol* 2004 Feb;43(2):225-8.
- 279 [36] Wang C, Schmid CH, Hibberd PL, et al. Tai Chi is effective in treating knee  
280 osteoarthritis: a randomized controlled trial. *Arthritis Rheum* 2009 Nov  
281 15;61(11):1545-53.
- 282 [37] Wang C, Schmid CH, Rones R, et al. A randomized trial of tai chi for  
283 fibromyalgia. *N Engl J Med* 2010 Aug 19;363(8):743-54.
- 284 [38] Wigers SH, Stiles TC, Vogel PA. Effects of aerobic exercise versus stress  
285 management treatment in fibromyalgia. A 4.5 year prospective study. *Scan J*  
286 *Rheumatol* 1996;25(2):77-86.
- 287 [39] Mehta CR, Patel NR. Exact logistic-regression - Theory and examples. *Stat Med*  
288 1995;14(19):2143-60.

- 289 [40] Riley RD, Lambert PC, Staessen JA, et al. Meta-analysis of continuous  
290 outcomes combining individual patient data and aggregate data. *Stat Med* 2008  
291 May 20;27(11):1870-93.
- 292 [41] Villain B, Dechartres A, Boyer P, et al. Feasibility of individual patient data meta-  
293 analyses in orthopaedic surgery. *BMC Med* 2015;13:131.
- 294 [42] Steinberg KK, Smith SJ, Stroup DF, et al. Comparison of effect size estimates  
295 from a meta-analysis of summary data from published studies and from a meta-  
296 analysis using individual patient data for ovarian cancer studies. *Am J Epidemiol*  
297 1997;145:917-25.
- 298 [43] Cooper H, Patall EA. The relative benefits of meta-analysis conducted with  
299 individual participant data versus aggregated data. *Psychol Methods*  
300 2009;14(2):165-76.

301

302

303 Table 1. Results for exact logistic regression for receipt of IPD (n = 29).

Variable	OR	SE	p	95% CI
Year	1.05	0.09	0.58	0.90, 1.27
Country	1.36	1.12	1.00	0.20, 10.90

304 Notes: IPD, Individual participant data; OR, odds ratio; SE, standard error;  
305 p, alpha value, calculated as 2 \* the probability of the sufficiency statistic,  
306 a statistic derived from single-parameter tests of the null hypothesis that  
307 the coefficient equals zero versus a 2-sided alternative; 95% CI, 95%  
308 confidence interval; Alpha (p) and 95% CI calculated from exact  
309 conditional distributions; both independent variables (year and country)  
310 calculated separately with the other variable conditioned out of the  
311 calculation.

312