Assessment of Placebo Response in Objective and Subjective Outcome Measures in Rheumatoid Arthritis Clinical Trials

Jan Vollert, PhD; Nancy R. Cook, ScD; Ted J. Kaptchuk; Shiv T. Sehra, MD; Deirdre K. Tobias, PhD; Kathryn T. Hall, PhD

Abstract

IMPORTANCE Large placebo responses in randomized clinical trials may keep effective medication from reaching the market. Primary outcome measures of clinical trials have shifted from patient-reported to objective outcomes, partly because response to randomized placebo treatment is thought to be greater in subjective compared with objective outcomes. However, a direct comparison of placebo response in subjective and objective outcomes in the same patient population is missing.

OBJECTIVE To assess whether subjective patient-reported (pain severity) and objective inflammation (C-reactive protein [CRP] level and erythrocyte sedimentation rate [ESR]) outcomes differ in placebo response.

DESIGN, SETTING, AND PARTICIPANTS The placebo arms of 5 double-blind, randomized, placebo-controlled clinical trials were included in this cross-sectional study. These trials were conducted internationally for 24 weeks or longer between 2005 and 2009. All patients with rheumatoid arthritis randomized to placebo (N = 788) were included. Analysis of data from these trials was conducted from March 27 to December 31, 2019.

INTERVENTION Placebo injection.

MAIN OUTCOMES AND MEASURES The difference (with 95% CIs) from baseline at week 12 and week 24 on a 0- to 100-mm visual analog scale to evaluate the severity of pain, CRP level, and ESR.

RESULTS Of the 788 patients included in the analysis, 644 were women (82%); mean (SD) age was 51 (13) years. There was a statistically significant decrease in patient-reported pain intensity (week 12: −14 mm; 95% CI, −12 to −16 mm and week 24: −20 mm; 95% CI, −16 to −22 mm). Similarly, significant decreases were noted in the CRP level (week 12: −0.51 mg/dL; 95% CI, −0.47 to −0.56 mg/dL and week 24: −1.16 mg/dL; 95% CI, −1.10 to −1.25 mg/dL) and ESR (week 12: −11 mm/h; 95% CI, −12 to −10 mm/h and week 24: −25 mm/h; 95% CI, −24 to −26 mm/h) (all P < .001).

CONCLUSIONS AND RELEVANCE The findings of this study suggest that improvements in clinical outcomes among participants randomized to placebo were not limited to subjective outcomes. Even if these findings could largely demonstrate a regression to the mean, they should be considered for future trial design, as unexpected favorable placebo responses may result in a well-designed trial becoming underpowered to detect the treatment difference needed in clinical drug development.
Introduction

Variability in the magnitude of clinical responses in the placebo arms of randomized clinical trials (RCTs) can result in effective drugs failing to reach the clinic. Placebo arms in RCTs are designed to control for the nonspecific effects of enrollment, including standard of care, regression to the mean, natural history of the disease, and placebo effects. Placebo arms should be kept as small as possible to avoid withholding potentially effective treatment from patients. It is important to distinguish between the placebo effect (a distinct physiological response to the therapeutic context and treatment delivery) and the placebo response in a clinical trial (improvement of patients in a placebo arm for any reason). Placebo effects are particularly strong in subjective outcomes (eg, pain), functional conditions (eg, irritable bowel syndrome), and fatigue, leading to the idea that placebo responses in clinical trials could be reduced by using objective biomarkers as primary outcomes rather than patient-reported outcomes. This approach, however, is indirect, since little is known about how objective and subjective outcomes vary in the placebo arms of RCTs and the approach ignores the role of other factors in the generation of the placebo response observed in an RCT. To directly compare placebo responses in objective biomarkers and subjective patient-reported outcomes, rheumatoid arthritis RCTs are ideal, as they usually collect data for assessment of patient-reported pain levels as well as changes in objective biomarkers (C-reactive protein [CRP] levels or erythrocyte sedimentation rate [ESR]), enabling insights into the magnitude of placebo effects within the placebo response in a homogeneous patient population. In the present study, we analyzed outcomes in the placebo arms of 5 RCTs of rheumatoid arthritis, comparing CRP level, ESR, and subjective pain levels. We hypothesized that, if the placebo effect is a leading contributor to the placebo response in these trials, the placebo response will be higher in pain levels, which are considered to be subjective compared with objective biomarkers. If, however, the placebo effect is superimposed by natural history disease fluctuation, regression to the mean phenomena, and other contributors to the placebo response in clinical trials, one would expect to note only minor differences between objective and subjective outcomes.

Methods

All included studies were separately approved by the respective ethical boards. Per the Common Rule, this study was exempt from institutional review board review owing to use of database information. We extracted individual patient- and study-level data from the placebo and standard of care arm of available RCTs of rheumatoid arthritis providing pain, CRP, and ESR measures in the TransCelerate Biopharma placebo and standard of care database as of October 30, 2018. At the time it was shared, this database included contributions of anonymized or pseudonymized data from Allergan, Amgen, AstraZeneca, Eli Lilly, GlaxoSmithKline, Johnson & Johnson, Pfizer, Roche, and UCB Pharma. Contributions to the database were performed on a voluntary basis and did not follow comprehensive inclusion criteria. We included all studies in the database performed involving patients with rheumatoid arthritis of at least 24 weeks’ duration, providing individual patient-level data on at least (1) pain levels at baseline, week 12, and week 24; (2) CRP levels at baseline, week 12, and week 24; and (3) ESR at baseline, week 12, and week 24. Five studies fulfilled these inclusion criteria, all conducted between 2005 and 2009, with results being reported between 2007 and 2015. Analysis of data from these trials was conducted from March 27 to December 31, 2019. Study characteristics and size of the patient cohorts can be found in Table 1. Using the study identifiers extracted from the TransCelerate database to identify ClinicalTrials.gov records, information on general study design, study location, and treatment vehicle were extracted from ClinicalTrials.gov and study publications (Table 1, publicly available data). Because these data are proprietary, the data of this analysis are not shared publicly, and no public involvement in this study was possible. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline for cross-sectional studies as applicable.
Trial level data included study design and setting, intervention duration, outcomes at week 12 and week 24, general location of the trial (US, European Union, or both), number of treatment arms in the study, original overall trial size, concomitant use of methotrexate or other disease-modifying anti-rheumatic drugs (DMARDs), approximate years the trial was conducted, and trial design elements (i.e., placebo run-in or crossover to active treatment). Patient-level demographic data were basic characteristics of participants (age and sex). The subjective outcome measure was a standard pain severity assessment (0- to 100-mm visual analog scale). Objective measures were levels of the inflammation biomarkers CRP and ESR. For all 3 outcome measures, a negative change from baseline indicated clinical improvement.

In some trials patients initially randomized to placebo treatment were allowed to cross over to the active treatment after week 12. Thus, we a priori selected 12 weeks as a point that was available for most participants preceding the decision point to drop out or change treatment arms. In addition, we compared the outcome at week 12 with the outcome at week 24. Patients who did not continue treatment were excluded from the main analyses because differential effects between weeks 12 and 24 were part of the research question; however, to address the robustness of the findings, the results were compared in a last-observation-carried-forward analysis.

### Statistical Analysis

Log-normally distributed data were log transformed, and statistical tests in these cases were performed in log space. Comparisons between baseline, week 12, and week 24 were performed using paired t tests, either in normal space or in log-normal space, if warranted by the data. Baseline data are presented as means (SDs) or, in case of log-normal distribution, as median and interquartile range. Changes from baseline are presented as arithmetic or geometric mean, along with 95% CIs. To investigate whether effects on outcomes were associated, Pearson correlation analyses were performed to test for associations between placebo response (1) at week 12 and week 24, (2) across

### Table 1. Placebo Arms From the 5 Rheumatoid Arthritis Trials Included in This Study

| Variable                        | Value               | Trial 1 | Trial 2 | Trial 3 | Trial 4 | Trial 5 |
|---------------------------------|---------------------|---------|---------|---------|---------|---------|
| Data set of this analysis       |                     |         |         |         |         |         |
| Total participants, including  | 2996                | 950     | 590     | 194     | 512     | 750     |
| treatment arms, No.             |                     |         |         |         |         |         |
| Participants randomized to     | 788                 | 199     | 101     | 88      | 156     | 244     |
| placebo in this analysis, No.  |                     |         |         |         |         |         |
| Women, No. (%)                  | 644 (82)            | 167 (84)| 84 (83) | 69 (78) | 134 (86)| 190 (78)|
| Age at baseline, mean (SD), y   | 51 (13)             | 52 (11) | 52 (12) | 54 (12) | 52 (12) | 48 (13) |
| Dropout week 12, No. (%)        | 57 (7)              | 13 (7)  | 5 (5)   | 11 (13) | 12 (8)  | 16 (7)  |
| Dropout week 24, No. (%)        | 261 (33)            | 155 (78)| 5 (5)   | 11 (13) | 62 (40) | 28 (11) |
| Pain at baseline, mean (SD)     | 59 (22)             | 64 (20) | 61 (22) | 37 (19) | 58 (19) | 64 (23) |
| CRP level at baseline, median   | 1.90 (0.82-3.63)    | 2.19 (0.80-4.20)| 1.89 (0.80-3.22) | 1.20 (0.80-1.90) | 1.25 (0.55-3.22) | 2.47 (1.45-4.02)| mg/dL |
| ESR at baseline, median (IQR),  | 43 (32-65)          | 45 (34-64)| 38 (32-53) | 30 (28-36) | 40 (26-55) | 60 (40-85)| mm/h |
| Publicly available data         |                     |         |         |         |         |         |
| Region                          | NA                  | International | International | Central Europe | International | International |
| Duration, wk                    | NA                  | 24       | 24       | 24       | 52       |
| Treatment vehicle               | NA                  | Injection | Injection | Injection   | Injection   |
| Randomization ratio             | NA                  | 4:1      | 4:1      | 1:1       | 2:1       | 2:1       |
| Start date, y                   | NA                  | 2005     | 2005     | 2008      | 2005      | 2006      |
| Washout period, d               | NA                  | 28       | 28       | 28        | 14        | 14        |
| Inclusion/exclusion criteria    | NA                  | Methotrexate, stable for at least 6 mo | Methotrexate, stable for at least 6 mo | Methotrexate, stable for at least 6 mo | Methotrexate, stable for at least 6 mo | No previous methotrexate treatment allowed |

Abbreviations: CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; IQR, interquartile range; NA, not applicable.

SI conversion factor: To convert CRP to milligrams per liter, multiply by 10.

---

\[P_{\text{placebo}} \times (1 - R_{\text{placebo}}) \times (1 - R_{\text{active}}) \]
outcomes at week 12, and (3) between baseline and mean of baseline and week 12 and placebo response at week 12 to analyze influence of baseline on outcomes. No P values were calculated for correlations.

The association between the placebo response and demographic and study variables was investigated using mixed-effects models, with change $Y_i$ in pain (outcome − baseline), CRP level, or ESR (ratio of outcome to baseline) at week 12 or 24 as the dependent variables, trial as random effect $\alpha$, and baseline, age, sex, washout length, randomization ratio, and use of DMARDs as fixed effects $\beta_1-6$. For the $j$th patient in the $i$th trial, the model was

$$Y_{ij} = \alpha_0 = \alpha_i + \beta_1 \text{baseline}_{ij} + \beta_2 \text{age}_{ij} + \beta_3 \text{sex}_{ij} + \beta_4 \text{washout}_{ij} + \beta_5 \text{Rand}_{ij} + \beta_6 \text{DMARD}_i + \varepsilon_{ij}.$$  

The estimated adjusted effects of the variables were calculated with their 95% CIs, and findings were considered significant if the 95% CI did not cross zero.

All analyses were performed in R, version x64 3.4.1 (R Project for Statistical Computing), using basic functions and the package lme4 for mixed-effects models. Given the low number of hypothesis tests performed in this analysis, we did not correct for multiple testing. $P < .05$ was considered significant.

**Results**

The data from 5 randomized clinical trials representing 788 participants randomized to placebo control arms were included in this study (644 women [82%]; 144 men [18%]; mean [SD] age, 51 [13] years) (Table 1). All investigations were parallel, double-blind trials of at least 24 weeks' duration with double-blind placebo administration by vehicle-matched injection. There were statistically significant decreases from high baseline levels in pain intensity at week 12 (−14 mm; 95% CI, −12 to −16 mm) and week 24 (−20 mm; 95% CI, −16 to −22 mm), CRP level at week 12 (−0.51 mg/dL; 95% CI, −0.47 to −0.56 mg/dL) and week 24 (−1.16 mg/dL; 95% CI, −1.03 to −1.30 mg/dL) (to convert to milligrams per liter, multiply by 10), and ESR at week 12 (−11 mm/h; 95% CI, −10 to −12 mm/h) and week 24 (−25 mm/h; 95% CI, −22 to −26 mm/h) (all $P < .001$) (Figure 1). When trial 5, which had the greatest placebo response and included patients naive to DMARDs, was excluded from the analysis (Figure 2), effects were smaller yet remained statistically significant. Pain levels decreased by 7 mm (95% CI, 5-9 mm) at week 12 and 12 mm (95% CI, 10-14 mm) at week 24, CRP levels decreased by 0.18 mg/dL (95% CI, 0.17-0.19 mg/dL) at week 12 and 0.91 mg/dL (95% CI, 0.84-1.00 mg/dL) at week 24, and ESR decreased by 8 mm/h (95% CI, 7-8 mm/h) at week 12 and 22 mm/h (95% CI, 21-23 mm/h) at

---

**Figure 1. Improvement in Pain Severity, C-Reactive Protein Level, and Erythrocyte Sedimentation Rate in Placebo Arm of 5 Randomized Clinical Trials**

![Graph A: Pain scale](image1.png)  
![Graph B: C-reactive protein](image2.png)  
![Graph C: Erythrocyte sedimentation rate](image3.png)

Values are means for pain (measured by a visual analog scale with values 0-100 mm) (A) and geometric means for C-reactive protein level (B) and erythrocyte sedimentation rate (C) to account for log-normal distribution. All mean changes were significant ($P < .001$).
week 24 (all \( P < .001 \)). Results did not change substantially when last observation carried forward was used instead of exclusion of dropouts, and all results remained significant at \( P < .001 \).

Correlations between week 12 and week 24 outcomes were moderate to high (pain: \( r = 0.73 \), CRP: \( r = 0.39 \), ESR: \( r = 0.59 \)), and those between subjective and objective improvement were lower than those between objective measures (pain/CRP: \( r = 0.27 \), pain/ESR: \( r = 0.25 \), CRP/ESR: \( r = 0.48 \)). While correlations between change and baseline were moderate (pain: \( r = 0.47 \), CRP: \( r = 0.39 \), ESR: \( r = 0.25 \)), correlations between change and average of baseline and week 12 assessment were poor to nonexistent (pain: \( r = 0.16 \), CRP: \( r = 0.08 \), ESR: \( r = 0.02 \)).

In the mixed-effects models, previous treatment with DMARDs (naive vs stable) was identical for all but 1 trial (trial 5) and could therefore not be disentangled from random study effects (Table 2). For all outcomes, baseline values showed a significant influence on the outcome at both weeks 12 and 24. No other variables tested had statistically significant effects, except for washout length, which had an unexpectedly high effect on both CRP level and ESR at week 24 only. This result could be biased: washout length was associated in this data set with earlier use of DMARDs, the only study with no previous use of DMARDs making up 61% of the group of patients with a shorter washout length.

Figure 2. Improvement in Pain Severity, C-Reactive Protein Level, and Erythrocyte Sedimentation Rate for Placebo Arms of Trials 1 Through 4

Values are means for pain (measured by a visual analog scale with values 0-100 mm) (A) and geometric means for C-reactive protein level (B) and erythrocyte sedimentation rate (C) to account for log-normal distribution. All mean changes were significant (\( P < .001 \)).

Table 2. Effects on Subjective and Objective Outcomes at Week 12 and 24 of Various Factorsa

| Variable                  | Pain levelb | C-Reactive protein levelc | Erythrocyte sedimentation ratec |
|---------------------------|-------------|---------------------------|---------------------------------|
|                           | Week 12     | Week 24                   | Week 12                         | Week 24                         | Week 12                         | Week 24                         |
| Baseline                  | 0.6 (0.5-0.7)d | 0.6 (0.5-0.7)d            | \(-0.001 (-0.002 to -0.001)d\) | \(-0.001 (-0.002 to -0.001)d\) | \(-0.006 (-0.007 to -0.004)d\) | \(-0.005 (-0.007 to -0.002)d\) |
| Sex                       | 0.7 (-30.5 to 50.0) | 0.5 (-40.6 to 50.5) | 0.101 (-0.137 to 0.340) | \(-0.066 (-0.432 to 0.301)\) | \(-0.081 (-0.189 to 0.026)\) | \(-0.050 (-0.183 to 0.084)\) |
| Age                       | 0.1 (-0.1 to 0.3) | 0.1 (-0.1 to 0.3) | \(-0.004 (-0.014 to 0.005)\) | \(-0.007 (-0.020 to 0.006)\) | \(-0.003 (-0.008 to 0.001)\) | \(-0.002 (-0.007 to 0.003)\) |
| Randomization ratio       | \(-4.3 (-10.0 to 15.5)\) | 0.5 (-120.1 to 130.1) | 0.004 (-0.164 to 0.172) | 0.108 (-0.076 to 0.292) | 0.042 (-0.007 to 0.090) | 0.025 (-0.038 to 0.089) |
| Washout length            | \(-0.7 (-10.7 to 0.3)\) | \(-0.7 (-20.9 to 10.5)\) | 0.014 (-0.014 to 0.041) | 0.036 (0.009 to 0.062)d | \(-0.002 (-0.010 to 0.005)\) | 0.011 (0.001 to 0.023)d |

a Values are presented as effect on the 0- to 100-mm pain scale or on the ratio outcome to baseline for C-reactive protein level and erythrocyte sedimentation rate to account for log-normal distribution.
b Positive factors indicate improvement.
c Owing to ratio rather than difference calculation, negative factors indicate improvement.
d Statistically significant finding.
Discussion

Contrary to our hypothesis, we found that objective and subjective outcome measures in rheumatoid arthritis trials improved to a clinically meaningful extent within the 5 trials in this analysis, showing that the placebo responses observed in this study are more than a psychological placebo effect. Correlation analyses of baseline level to placebo response showed moderate correlations; however, correlation between mean of baseline and week 12 and placebo response was poor or nonexistent. This low level could be caused by inclusion biased toward high baseline values, for example, minimum disease severity. A recent systematic review reinforces this view, finding increased placebo responses in clinical trials of rheumatoid arthritis over the past 30 years and partly attributing these to inflated baseline measures. On the other hand, a meta-analysis suggested that the method of placebo delivery (eg, oral vs intra-articular) has a significant effect on placebo response in osteoarthritis trials, showing that the psychological dimension of the placebo response cannot be dismissed. Minimum levels at enrollment for these trials were between 0.63 and 1.47 mg/dL for CRP level or between 28 and 30 mm/h for ESR, which might be the main factors affecting these findings. A minimum pain level was not set as an inclusion criterion for any of the trials; however, the mean baseline level of 59 mm can be considered high and is linked to inflammation severity. Patients were stable with DMARD therapy across trials 1 to 4, while potentially naive patients were introduced to DMARDs in trial 5. Thus, we expected a higher response in the placebo arms of these trials due to the drug effect of DMARDs. We performed a separate analysis excluding this trial, yet the results stayed significant. While the magnitude of placebo response was clinically meaningful across pain, CRP level, and ESR, these findings are not necessarily in the same patients: correlations between improvement of pain and CRP level and ESR were only approximately 0.25. Even for the 2 objective inflammation markers (ESR and CRP level), the correlation was below 0.5. The correlation between week 12 and week 24 was moderate to high, showing that natural history, standard of care, and within-patient variability might explain another important part of the placebo response.

The clinical phase of rheumatoid arthritis in its natural history progresses slowly, making it difficult to compare with progression within a clinical study. However, CRP levels seem to increase even before clinical manifestation and improvement of the degree found in this study seems unlikely to happen as spontaneous, untreated remission.

Limitations

Given that this was a retrospective study based on available deidentified data, there are several limitations, such as demographic characteristic variables. Only a comparison of placebo arms with natural history arms could have clearly demonstrated the influence of regression to the mean, and such arms were missing in the trials included in our analysis. A no-treatment group would have yielded important information on the natural history of outcome fluctuation in association with placebo response. Most patients were randomized to placebo plus standard of care, that is, DMARDs such as methotrexate, which represents a typical situation in rheumatoid arthritis trials. Thus, we cannot rule out DMARDs increasing the placebo response; however, the placebo response noted with pain severity was similar to other causes of chronic pain, which mirrors the typical situation in rheumatoid arthritis trials.

Conclusions

The results of this study suggest that, to reduce the placebo response in clinical trials, replacing subjective with objective outcomes will not necessarily lead to clearer results. Within our data set, high baseline values at enrollment due to minimum levels as inclusion criteria and natural history seem to overlay the psychological placebo effects. Appropriate means to account for nonpsychological elements of the placebo response could be using alternative study designs, for
example, with multiple baseline measures,\textsuperscript{22} or using methods to statistically account for the regression to the mean in the analysis.\textsuperscript{23} Including no treatment arms in the RCTs to understand the role of natural history and regression to the mean more clearly would be beneficial for future trials, and similar comparisons between objective and subjective outcomes in other diseases may corroborate or challenge our findings. For new drug development, efforts to understand how baseline covariates and confounding factors in the placebo responses via data sharing initiatives\textsuperscript{24,25} can be helpful to study design considerations, expedite clinical drug development, and ultimately bring effective treatments to patients in need.

### ARTICLE INFORMATION

**Accepted for Publication:** June 1, 2020.

**Published:** September 16, 2020. doi:10.1001/jamanetworkopen.2020.13196

**Open Access:** This is an open access article distributed under the terms of the CC-BY License. © 2020 Vollert J et al. JAMA Network Open.

**Corresponding Author:** Jan Vollert, PhD, Brigham and Women’s Hospital, 900 Commonwealth Ave, Boston, MA 02215 (jvollert@bwh.harvard.edu).

**Author Affiliations:** Brigham and Women’s Hospital, Harvard Medical School, Boston, Massachusetts (Vollert, Cook, Tobias, Hall); Pain Research, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, United Kingdom (Vollert); Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts (Kaptchuk); Mount Auburn Hospital, Harvard Medical School, Cambridge, Massachusetts (Sehra).

**Author Contributions:** Dr Vollert had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

- **Concept and design:** Vollert, Kaptchuk, Tobias, Hall.
- **Acquisition, analysis, or interpretation of data:** Vollert, Cook, Sehra, Tobias, Hall.
- **Drafting of the manuscript:** Vollert, Kaptchuk.
- **Critical revision of the manuscript for important intellectual content:** Cook, Sehra, Tobias, Hall.
- **Statistical analysis:** Vollert, Cook.
- **Obtained funding:** Hall.
- **Administrative, technical, or material support:** Kaptchuk, Tobias, Hall.

**Supervision:** Hall.

**Conflict of Interest Disclosures:** Dr Vollert reported receiving personal fees from Casquar outside the submitted work. No other disclosures were reported.

**Funding/Support:** Drs Vollert, Cook, Kaptchuk, Tobias, and Hall received funding from TransCelerate Biopharma for this work.

**Role of the Funder/Sponsor:** Neither TransCelerate Biopharma nor the data providers contributed to design and conduct of the study; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. Neither TransCelerate Biopharma nor the data providers contributed to the analysis and interpretation of the data, but they collected most of the data and contributed to managing the data, both before this study.

**Additional Contributions:** Florence Yong, PhD (committee lead), Edward Bowen, PhD, and Lanju Zhang, PhD (TransCelerate-Harvard Collaboration Committee), contributed to interpretation of the data and review of the manuscript. There was no financial compensation outside of salary.

**Additional Information:** This study was based in part on data from the TransCelerate BioPharma Inc Placebo and Standard of Care Data Sharing Initiative, which, at the time the data were shared, included contributions of anonymized or pseudonymized data from Allergan, Amgen, AstraZeneca, Eli Lilly, GlaxoSmithKline, Johnson & Johnson, Pfizer, Roche, and UCB Pharma (“data providers”).

### REFERENCES

1. Finniss DG, Kaptchuk TJ, Miller F, Benedetti F. Biological, clinical, and ethical advances of placebo effects. *Lancet*. 2010;375(9715):686-695. doi:10.1016/S0140-6736(09)61706-2
2. US Food and Drug Administration, US Dept of Health and Human Services. Rare Diseases: Natural History Studies for Drug Development: Guidance for Industry. US Food and Drug Administration; 2019.

3. Kaptchuk TJ, Miller FG. Placebo effects in medicine. N Engl J Med. 2015;373(1):8-9. doi:10.1056/NEJMps14023

4. Shaibani A, Frisaldi E, Benedetti F. Placebo response in pain, fatigue, and performance: possible implications for neuromuscular disorders. Muscle Nerve. 2017;56(3):358-367. doi:10.1002/mus.25635

5. Patel SM, Stason WB, Legedza A, et al. The placebo effect in irritable bowel syndrome trials: a meta-analysis. Neurogastroenterol Motil. 2017;19:332-340. doi: 10.1111/j.1365-2982.2015.06650.x

6. Hoenemeyer TW, Kaptchuk TJ, Mehta TS, Fontaine KR. Open-label placebo treatment for cancer-related fatigue: a randomized-controlled clinical trial. Sci Rep. 2018;8(1):2784. doi:10.1038/s41598-018-20993-y

7. McCoy CE. Understanding the use of composite endpoints in clinical trials. West J Emerg Med. 2018;19(4):631-634. doi:10.5811/westjem.2018.4.38383

8. American College of Rheumatology Committee to Reevaluate Improvement Criteria. A proposed revision to the ACR20: the hybrid measure of American College of Rheumatology response. Arthritis Rheum. 2007;57(2):193-202. doi:10.1002/art.22552

9. Emery P, Deodhar A, Rigby W, et al. Efficacy and safety of different doses and treatment of rituximab: a randomised, placebo-controlled trial in patients who are biological naive with active rheumatoid arthritis and an inadequate response to methotrexate (Study Evaluating Rituximab’s Efficacy in MTX Inadequate Responders (SERENE)). Ann Rheum Dis. 2010;69(9):1629-1635. doi:10.1136/ard.2009.119933

10. Keystone E, Heijde D, Mason D Jr, et al. Certolizumab pegol plus methotrexate is significantly more effective than placebo plus methotrexate in active rheumatoid arthritis: findings of a fifty-two-week, phase III, multicenter, randomized, double-blind, placebo-controlled, parallel-group study. Arthritis Rheum. 2008;58(11):3319-3329. doi:10.1002/art.23964

11. Smolen JS, Landewé RB, Mease P, et al. Efficacy and safety of certolizumab pegol plus methotrexate in active rheumatoid arthritis: the RAPID 2 study, a randomised controlled trial. Ann Rheum Dis. 2009;68(6):797-804. doi:10.1136/ard.2008.101659

12. Smolen JS, Emery P, Ferraccioli GF, et al. Certolizumab pegol in rheumatoid arthritis patients with low to moderate activity: the CERTAIN double-blind, randomised, placebo-controlled trial. Ann Rheum Dis. 2015;74(5):843-850. doi:10.1136/annrheumdis-2013-204632

13. Tak PP, Rigby W, Rubbert-Roth A, et al. Sustained inhibition of progressive joint damage with rituximab plus methotrexate in early active rheumatoid arthritis: 2-year results from the randomised controlled trial IMAGE. Ann Rheum Dis. 2012;71(3):351-357. doi:10.1136/annrheumdis-2011-200170

14. Tu YK, Gilthorpe MS. Revisiting the relation between change and initial value: a review and evaluation. Stat Med. 2007;26(2):443-457. doi:10.1002/sim.2538

15. Alam J, Jantan I, Bukhari SNA. Rheumatoid arthritis: recent advances on its etiology, role of cytokines and pharmacotherapy. Biomed Pharmacother. 2017;92:615-633. doi:10.1016/j.biopha.2017.05.055

16. Deane KD. Learning about the natural history of rheumatoid arthritis development through prospective study of subjects at high risk of rheumatoid arthritis-related autoimmunity. Arthritis Rheum. 2012;64(6):1708-1712. doi:10.1002/art.34445

17. Deane KD, Holers VM. The natural history of rheumatoid arthritis. Clin Ther. 2019;41(7):1256-1269. doi:10.1016/j.clinthera.2019.04.028

18. Pincus T, Callahan LF. What is the natural history of rheumatoid arthritis? Rheum Dis Clin North Am. 1993;19(1):123-151.

19. Vase L, Vollert J, Finnerup NB, et al. Predictors of the placebo analgesia response in randomized controlled trials of chronic pain: a meta-analysis of the individual data from nine industrially sponsored trials. Pain. 2015;156(9):1795-1802. doi:10.1097/j.pain.00000000000000217

20. Bechman K, Yates M, Norton S, Cope AP, Galloway JB. Placebo response in rheumatoid arthritis clinical trials. J Rheumatol. 2020;47(1):28-34. doi:10.3899/jrheum.190008

21. Bannuru RR, McAlindon TE, Sullivan MC, Wong JB, Kent DM, Schmid CH. Effectiveness and implications of alternative placebo treatments: a systematic review and network meta-analysis of osteoarthritis trials. Ann Intern Med. 2015;163(5):365-372. doi:10.7326/M15-0623

22. Barnett AG, van der Pols JC, Dobson AJ. Regression to the mean: what it is and how to deal with it. Int J Epidemiol. 2005;34(1):215-220. doi:10.1093/ije/dyh299
23. Thomas DM, Clark N, Turner D, et al. Best (but oft-forgotten) practices: identifying and accounting for regression to the mean in nutrition and obesity research. *Am J Clin Nutr.* 2020;111(2):256-265. doi:10.1093/ajcn/nqz196

24. Selker HP, Eichler HG, Stockbridge NL, et al. Efficacy and effectiveness too trials: clinical trial designs to generate evidence on efficacy and on effectiveness in wide practice. *Clin Pharmacol Ther.* 2019;105(4):857-866. doi:10.1002/cpt.1347

25. Kuntz RE, Antman EM, Califf RM, et al. *Individual Patient-Level Data Sharing for Continuous Learning: A Strategy for Trial Data Sharing.* NAM Perspectives; 2019.