STAT 36900: Homework 4

Robert Winter

Table of Contents

1 Introduction

The data for this problem come from a randomized clinical trial to determine the efficacy of the drug auranofin versus placebo in the treatment of rheumatoid arthritis (Bombardier *et al.*, 1986, Auranofin therapy and quality of life in patients with rheumatoid arthritis, *American Journal of Medicine, 81,* 565–578).

The outcome of interest is self-assessment of arthritis, coded as `0` = poor and `1` = good. Individuals were randomized to one of the two treatment groups after baseline self-assessment of arthritis, and then took the treatment for the length of the study. All 294 patients were observed at baseline, and 1 and 3 months after randomization. The self-assessment of arthritis at these three timepoints represent the responses of the dichotomous dependent variable across time. These data are provided at the class Canvas website as `arthritb.dat.txt`. The variables in the dataset are `ID` (subject ID), `ARTHRIT` (dependent variable coded `0` = poor and `1` = good), `TXAURA` (treatment group coded `0` = placebo and `1` = auranofin), and `TIME` (coded `0` = baseline, `1` = Month 1, and `2` = Month 3).

For this problem, use dummy-codes for the effect of time treating baseline as the reference cell. Let's call the resulting two dummy-codes `TIMEA` and `TIMEB`.

Given that randomization was done after the baseline assesment, and therefore subjects are the same at baseline, use the following regressors: `TIMEA`, `TIMEB`, `TX*TIMEA`, and `TX*TIMEB`. Note that, here, not include `TX`, which would represent the group difference at baseline, is a very special case. Normally, one includes all component variables of interactions in the model. However, here, because the groups have not yet received treatment at baseline, and because it is a randomized study, these groups are the same at baseline.

2 Question 1

Based on the observed correlations, select a reasonable working correlation structure and estimate a GEE logistic regression model with the above four regressors. Describe the meaning of the estimated regression coefficients in this model. Is there evidence of a treatment effect? If so, is the effect seen at both follow-ups? How well does the model fit the observed proportions for the two groups at the three timepoints?

In the table below, we present the pairwise Spearman correlations between patients' arthritis ratings at times `0` (baseline), `1` (Month 1), and `2` (Month 3). Notably, the correlations between times `0` and `1` and times `1` and `2` are very similar: the former is roughly 0.435, while the latter is roughly 0.448. Having roughly the same correlation between two different pairs of consecutive timepoints suggests that a stationary m-dependent working correlation structure may be appropriate here. (Without more timepoints to study, we can't tell whether or not there is a geometric relationship between correlations for progressively larger lags, so we cannot say whether an AR(1) structure would be even more appropriate.) It is important to acknowledge that the timepoints in this study were not evenly spaced: there was a one-month gap between times `0` and `1`, but a two-month gap between times `1` and `2`. For the sake of this assignment, we assume that the timepoints were evenly spaced and move forward with a stationary m-dependent structure fo the working correlation matrix.

```
. infile id arthrit txaura time ///
> using arthritb.dat.txt, clear
(882 observations read)
.
. * Reshape to wide format
. quietly reshape wide arthrit, i(id) j(time)
.
. * Generate Spearman correlation matrix
. spearman arthrit0 arthrit1 arthrit2

Number of observations = 294

             | arthrit0 arthrit1 arthrit2
-------------+---------------------------
```

```
    arthrit0 |   1.0000
    arthrit1 |   0.4345   1.0000
    arthrit2 |   0.3172   0.4480   1.0000
```

Thus, we estimate the following GEE model:

$$\begin{aligned}\eta_{ij} = \beta_0 + \beta_1 \cdot TimeA_j + \beta_2 \cdot TimeB_j \\ + \beta_3 \cdot TX_i \times TimeA_j + \beta_4 \cdot TX_i \times TimeB_j\end{aligned}$$

where:

- $i = 1, \ldots, 294$ individuals,
- $j = 1, 2, 3$ observations for each patient i, and
- $\eta_{ij} = \log\left(\frac{\mathbb{P}(Arthrit_{ij}=1)}{1-\mathbb{P}(Arthrit_{ij}=1)}\right)$,

and

- β_0 is the population average log odds of reporting "good" arthritis at baseline (across patients in the placebo and auranofin-treated groups),
- β_1 is the average change in the log odds of reporting "good" arthritis as of Month 1 compared to baseline among patients in the placebo group,
- β_2 is the average change in the log odds of reporting "good" arthritis as of Month 3 compared to baseline among patients in the placebo group,
- β_3 is the difference in the average change in the log odds of reporting "good" arthritis as of Month 1 compared to baseline between patients in the treated and placebo groups, and
- β_4 is the difference in the average change in the log odds of reporting "good" arthritis as of Month 3 compared to baseline between patients in the treated and placebo groups.

```
. infile id arthrit txaura time ///
> using arthritb.dat.txt, clear
(882 observations read)
. generate timea = time==1
. generate timeb = time==2
. generate tx_timea = txaura * timea
. generate tx_timeb = txaura * timeb
.
. xtgee arthrit timea timeb tx_timea tx_timeb, ///
>   fam(bin) link(logit) i(id) t(time) corr(stationary 2) robust

Iteration 1:  Tolerance = .01404028
```

```
Iteration 2:  Tolerance = 6.086e-07

GEE population-averaged model                   Number of obs    =    882
Group and time vars: id time                    Number of groups =    294
Family: Binomial                                Obs per group:
Link:   Logit                                                 min =      3
Correlation: stationary(2)                                    avg =    3.0
                                                              max =      3
                                                Wald chi2(4)     =  38.00
Scale parameter = 1                             Prob > chi2      = 0.0000

                                    (Std. err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
     arthrit | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       timea |   .6840975   .1720079     3.98   0.000     .3469683    1.021227
       timeb |   .3753876   .1962969     1.91   0.056    -.0093473    .7601224
    tx_timea |   .0432509   .2174686     0.20   0.842    -.3829798    .4694815
    tx_timeb |   .6348553   .2321005     2.74   0.006     .1799467    1.089764
       _cons |  -1.220257   .1392718    -8.76   0.000    -1.493225   -.9472897
------------------------------------------------------------------------------
. estat wcorr

Estimated within-id correlation matrix R:

      |        c1         c2         c3
------+----------------------------------
   r1 |         1
   r2 |  .4432533          1
   r3 |  .3202026   .4432533          1
```

As summarized in the Stata output above, we find that:

- At baseline, the population average log odds of reporting "good" arthritis is approximately -1.220.
- One month after baseline, the average log odds of reporting "good" arthritis among patients in the placebo group increased by approximately 0.684 relative to baseline.
- Three months after baseline, the average log odds of reporting "good" arthritis among patients in the placebo group increased by approximately 0.375 relative to baseline; this estimate is (just barely) not statistically significant at the $\alpha = 0.05$ level.
- One month after baseline, the average log odds of reporting "good" arthritis among patients in the treated group was approximately 0.043 higher than the average log

odds of reporting "good" arthritis among patients in the placebo group. This estimate is not statistically significant, as discussed in greater detail below.

- Three months after baseline, the average log odds of reporting "good" arthritis among patients in the treated group was approximately 0.635 higher than the average log odds of reporting "good" arthritis among patients in the placebo group.
- The estimated correlation matrix very closely resembles the empirical correlation matrix shown earlier.

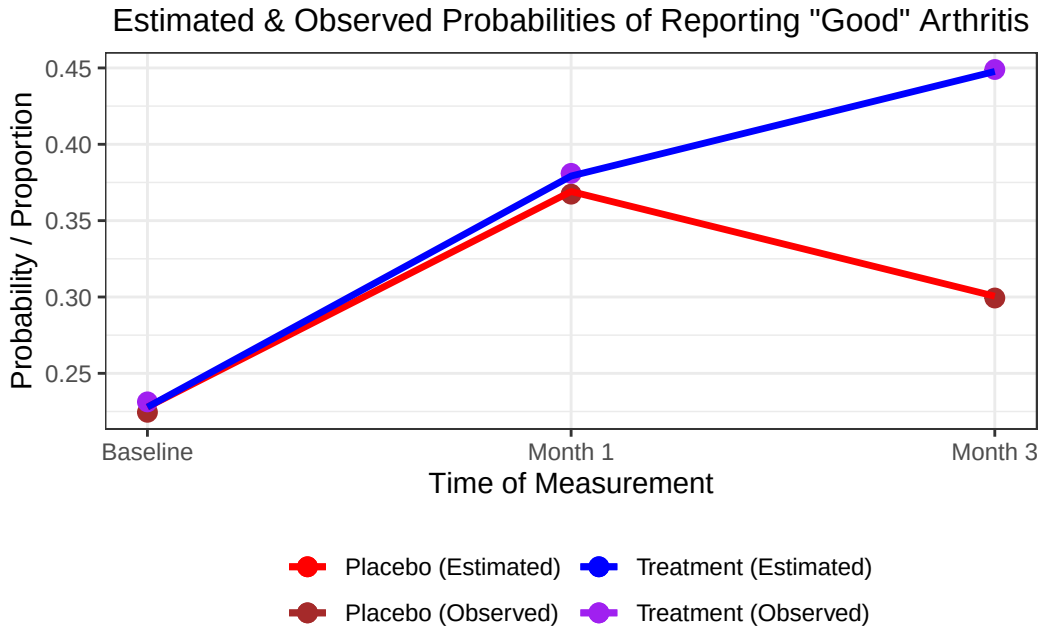
In particular, we find that there is evidence of a treatment effect, but only at the three months after baseline mark. Indeed, our estimate of β_3 is not only small, at around 0.043, but is also not statistically significant, with a p-value of approximately $0.842 > 0.05$. As such, there is not statistical evidence of a significant difference in the log odds of reporting "good" arthritis between the treatment and placebo groups one month after beginning treatment. Meanwhile, our estimate of β_4 is much larger, at around 0.635, and is statistically significant, with a p-value of $0.006 < 0.01$. Thus, there is statistical evidence of a significant difference in the log odds of reporting "good" arthritis between the treatment and placebo groups three months after beginning treatment.

As shown in the table and plot below, our GEE model's fitted/estimated probabilities of reporting "good" arthritis (`pmarg`) very closely match the actual proportions observed in the data (`pobs`).

```
. infile id arthrit txaura time ///
> using arthritb.dat.txt, clear
(882 observations read)
. generate timea = time==1
. generate timeb = time==2
. generate tx_timea = txaura * timea
. generate tx_timeb = txaura * timeb
.
. quietly xtgee arthrit timea timeb tx_timea tx_timeb, ///
>   fam(bin) link(logit) i(id) t(time) corr(stationary 2) robust
.
. predict probs
(option mu assumed; Pr(arthrit != 0))
. collapse (mean) pmarg=probs pobs=arthrit, by(txaura time)
. list time txaura pmarg pobs

     +-----------------------------------+
     | time   txaura      pmarg       pobs |
     |-----------------------------------|
  1. |    0        0   .2278912   .2244898 |
  2. |    1        0   .3690813   .3673469 |
  3. |    2        0   .3005101   .2993197 |
```

```
4. |    0        1    .2278912    .2312925 |
5. |    1        1    .3792085    .3809524 |
   |-----------------------------------------|
6. |    2        1    .4476885    .4489796 |
   +-----------------------------------------+
```



3 Question 2

Estimate a random-intercepts (*i.e.*, mixed) model with the above four parameters. Describe the meaning of the estimated regression coefficients in this model. Is there evidence of a treatment effect? If so, is the effect seen at both follow-ups? How well does the model fit the observed proportions for the two groups at the three timepoints?

We estimate the following mixed model:

Within-Subjects Model:

$$\eta_{ij} = b_{0i} + b_{1i} \cdot TimeA_j + b_{2i} \cdot TimeB_j$$

where:

- $i = 1, \ldots, 294$ individuals,
- $j = 1, 2, 3$ observations for each patient i, and

- $\eta_{ij} = \log\left(\frac{\mathbb{P}(Arthrit_{ij}=1)}{1-\mathbb{P}(Arthrit_{ij}=1)}\right)$,

and:

- b_{0i} is the log odds of reporting "good" arthritis at baseline for patient i,
- b_{1i} is the change in patient i's log odds of reporting "good" arthritis as of Month 1 compared to baseline, and
- b_{2i} is the change in patient i's log odds of reporting "good" arthritis as of Month 3 compared to baseline.

Between-Subjects Model:

$$
\begin{aligned}
b_{0i} &= \beta_0 + \upsilon_{0i} \\
b_{1i} &= \beta_1 + \beta_3 \cdot TX_i \\
b_{2i} &= \beta_2 + \beta_4 \cdot TX_i \\
\upsilon_{0i} &\sim \text{iid } \mathcal{N}(0, \sigma_\upsilon^2)
\end{aligned}
$$

where:

- β_0 is the log odds of reporting "good" arthritis at baseline, given $\upsilon_{0i} = 0$,
- υ_{0i} is the incremental log odds of reporting "good" arthritis at baseline for patient i,
- β_1 is the change in a given patient in the placebo group's log odds of reporting "good" arthritis as of Month 1 compared to baseline, for a given value of υ_{0i} corresponding to that patient,
- β_3 is the difference between a patient in the treatment group and a patient in the placebo groups' log odds of reporting "good" arthritis as of Month 1, assuming both have equal values of υ_{0i},
- β_2 is the change in a given patient in the placebo group's log odds of reporting "good" arthritis as of Month 3 compared to baseline, for a given value of υ_{0i} corresponding to that patient, and
- β_4 is the difference between a patient in the treatment group and a patient in the placebo groups' log odds of reporting "good" arthritis as of Month 3, assuming both have equal values of υ_{0i}.

```
. infile id arthrit txaura time ///
> using arthritb.dat.txt, clear
(882 observations read)
. generate timea = time==1
. generate timeb = time==2
. generate tx_timea = txaura * timea
. generate tx_timeb = txaura * timeb
```

```
.
. melogit arthrit timea timeb tx_timea tx_timeb ///
>   || id:, covariance(unstructured)

Fitting fixed-effects model:

Iteration 0:  Log likelihood = -543.84357
Iteration 1:  Log likelihood = -542.97543
Iteration 2:  Log likelihood = -542.97493
Iteration 3:  Log likelihood = -542.97493

Refining starting values:

Grid node 0:  Log likelihood = -505.70185

Fitting full model:

Iteration 0:  Log likelihood = -505.70185
Iteration 1:  Log likelihood = -486.64269
Iteration 2:  Log likelihood = -483.98059
Iteration 3:  Log likelihood = -483.81198
Iteration 4:  Log likelihood = -483.81075
Iteration 5:  Log likelihood = -483.81073

Mixed-effects logistic regression               Number of obs     =        882
Group variable: id                              Number of groups  =        294

                                                Obs per group:
                                                              min =          3
                                                              avg =        3.0
                                                              max =          3

Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(4)      =      35.55
Log likelihood = -483.81073                     Prob > chi2       =     0.0000
------------------------------------------------------------------------------
     arthrit | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       timea |   1.167807   .3174948     3.68   0.000     .5455288    1.790086
       timeb |   .6333286   .3188323     1.99   0.047     .0084287    1.258228
    tx_timea |   .0836466   .3863711     0.22   0.829    -.6736268      .84092
    tx_timeb |   1.111217   .3955792     2.81   0.005     .3358959    1.886538
       _cons |   -2.13456   .2668461    -8.00   0.000    -2.657569   -1.611551
-------------+----------------------------------------------------------------
```

```
id             |
   var(_cons)|    5.12623    1.16563                      3.282821    8.004772
------------------------------------------------------------------------------
LR test vs. logistic model: chibar2(01) = 118.33      Prob >= chibar2 = 0.0000
```

As summarized in the Stata output above, we find that:

- At baseline, a patient with $\upsilon_{0i} = 0$ has a log odds of reporting "good" arthritis of approximately -2.135.
- One month after baseline, a given patient (i.e., holding υ_{0i} fixed) in the placebo group's log odds of reporting "good" arthritis increase by approximately 1.168 relative to baseline.
- One month after baseline, a patient in the treatment group's log odds of reporting "good" arthritis are approximately 0.084 higher than that of a patient in the placebo group with the same value of υ_{0i}. This estimate is not statistically significant, as discussed in greater detail below.
- Three months after baseline, a given patient (i.e., holding υ_{0i} fixed) in the placebo group's log odds of reporting "good" arthritis increase by approximately 0.633 relative to baseline.
- Three months after baseline, a patient in the treatment group's log odds of reporting "good" arthritis are approximately 1.111 higher than that of a patient in the placebo group with the same value of υ_{0i}.

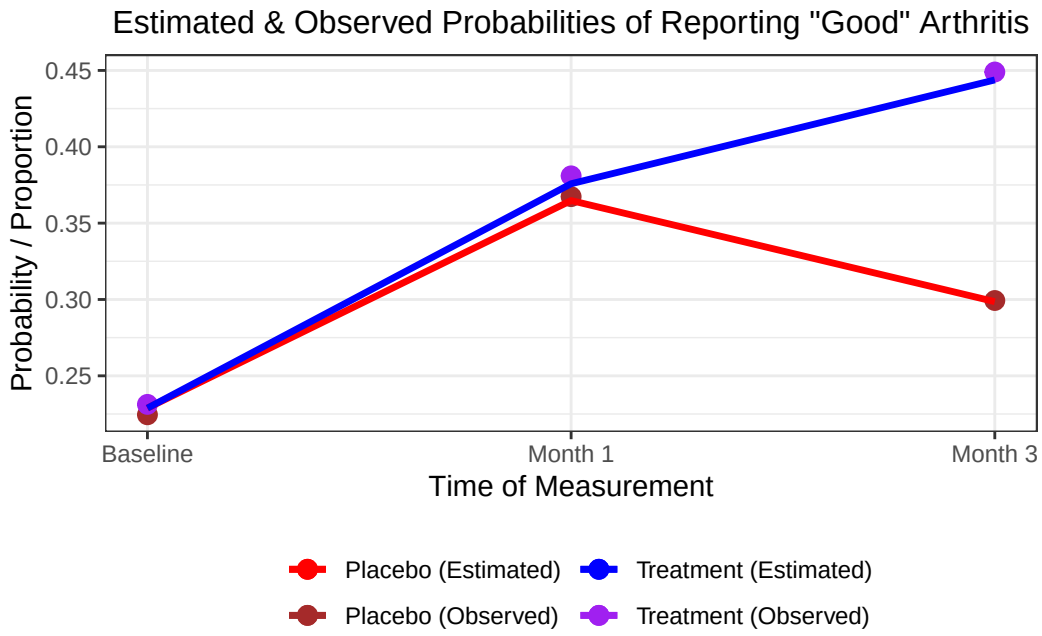
In particular, we find that there is evidence of a treatment effect, but only at the three months after baseline mark. Indeed, our estimate of β_3 is not only small, at around 0.084, but is also not statistically significant, with a p-value of approximately $0.829 > 0.05$. As such, there is not statistical evidence of a significant difference in the log odds of reporting "good" arthritis between a patient with a given value of υ_{0i} prescribed auranofin and a patient with the same value of υ_{0i} prescribed the placebo one month after beginning treatment. Meanwhile, our estimate of β_4 is much larger, at around 1.111, and is statistically significant, with a p-value of $0.005 < 0.01$. Thus, there is statistical evidence of a significance difference in the log odds of reporting "good" arthritis between a patient with a given value of υ_{0i} prescribed auranofin and a patient with the same value of υ_{0i} prescribed the placebo three months after beginning treatment.

As shown in the table and plot below, our mixed model's fitted/estimated probabilities of reporting "good" arthritis (`pmarg`) very closely match the actual proportions observed in the data (`pobs`).

```
. infile id arthrit txaura time ///
> using arthritb.dat.txt, clear
(882 observations read)
. generate timea = time==1
```

```
. generate timeb = time==2
. generate tx_timea = txaura * timea
. generate tx_timeb = txaura * timeb
.
. quietly melogit arthrit timea timeb tx_timea tx_timeb ///
>   || id:, covariance(unstructured)
.
. predict p_marg, marginal
(option mu assumed)
(using 7 quadrature points)
. collapse (mean) pmarg=p_marg pobs=arthrit, by(txaura time)
. list time txaura pmarg pobs

     +-----------------------------------+
     | time   txaura      pmarg       pobs |
     |-----------------------------------|
  1. |    0        0   .2288529   .2244898 |
  2. |    1        0   .3647873   .3673469 |
  3. |    2        0   .2985835   .2993197 |
  4. |    0        1   .2288529   .2312925 |
  5. |    1        1   .3757912   .3809524 |
     |-----------------------------------|
  6. |    2        1   .4437701   .4489796 |
     +-----------------------------------+
```



4 Question 3

How do the estimates from (1) and (2) relate to each other?

The GEE coefficient estimates we found in Question (1) are the marginal log odds effects associated with each regressor. That is, each GEE coefficient estimate is the population average effect of a one-unit increase in the corresponding variable on the log odds of reporting "good" arthritis.

Meanwhile, the MRM coefficient estimates we found in Question (2) are the conditional log odds effects associated with each regressor for a given value of υ_{0i}. That is, each MRM coefficient estimate is the effect of a one-unit increase in the corresponding variable on the log odds of reporting "good" arthritis for a specific patient with a given value of υ_{0i}.

The two sets of coefficient estimates are (approximately) related as follows:

$$\hat{\beta}_{GEE} \approx \hat{\beta}_{MRM} / \sqrt{\left(\frac{16}{15}\right)^2 * \frac{3}{\pi^2} * \sigma^2_{\upsilon_{0i}} + 1}$$

From our Stata output in Question (2), $\sigma^2_{\upsilon_{0i}} \approx 5.12623$. In the table below, we show the conditional coefficient estimates we calculated using the MRM in Question (2), use the formula above to approximate the marginal coefficients, and finally show that these approximations are close to the marginal coefficient estimates we calculated using the GEE in Question (1) (shown by comparing the values in the final two columns).

```
. mata
------------------------------------------------- mata (type end to exit) -----
:
: /* (Constant, TimeA, TimeB, TX*TimeA, TX*TimeB) */
:
: X = (1, 0, 0, 0, 0 \
>      1, 1, 0, 0, 0 \
>      1, 0, 1, 0, 0 \
>      1, 0, 0, 0, 0 \
>      1, 1, 0, 1, 0 \
>      1, 0, 1, 0, 1)

:
: beta_mrm = (-2.13456 \
>              1.167807 \
>              0.6333286 \
>              0.0836466 \
>              1.111217)

:
: var_u = 5.12623
```

```
:
: beta_mrm_std = beta_mrm / sqrt(((16/15)^2) * (3/(pi()^2)) * var_u + 1)

: beta_mrm_std
                 1
    +----------------+
  1 |  -1.281868759  |
  2 |   .7013039267  |
  3 |    .380333252  |
  4 |   .0502323492  |
  5 |   .6673198957  |
    +----------------+

:
: end
-------------------------------------------------------------------------------
```

Coefficient	$\widehat{\beta}_{MRM}$ (Q2)	$\tilde{\beta}_{GEE} = \hat{\beta}_{MRM} / \sqrt{\left(\frac{16}{15}\right)^2 * \frac{3}{\pi^2} * \sigma^2_{\upsilon_{0i}} + 1}$	$\widehat{\beta}_{GEE}$ (Q1)
Constant	-2.135	-1.282	-1.220
$TimeA$	1.168	0.701	0.684
$TimeB$	0.633	0.380	0.375
$TX \times TimeA$	0.084	0.050	0.043
$TX \times TimeB$	1.111	0.667	0.635

This relationship also explains why we were able to estimate nearly identical marginal probabilities for each Time-TX pair in Questions (1) and (2).

5 Question 4

What do you conclude about the efficacy of auranofin in the treatment of arthritis? Descriptive statistics, like marginal proportions or odds ratios, may help elucidate your results. Write up your final conclusions in a way that a medical doctor could understand.

Based on our analyses of the data using both random intercept and GEE models, we can conclude that auranofin is an effective treatment for arthritis (at least as measured by patients' self-reported presence or absence of pain), but requires a period of between one and three months for the effect to set in. At the outset of the clinical trial, the participants

had an average probability of reporting "good" (i.e., minimally painful) arthritis of only around 0.23. One month into the study, participants in the auranofin-treated group had an average probability of reporting "good" arthritis of around 0.37–0.38, but participants in the placebo group had an average probability of reporting "good" arthritis of around 0.36–0.38. So, at the one-month mark, auranofin did not significantly improve patients' self-reported arthritis levels relative to the placebo. However, three months into the study, participants in the auranofin-treated group had an average probability of reporting "good" arthritis of around 0.44–0.45, while participants in the placebo group had an average probability of reporting "good" arthritis of only around 0.30. That is, even though patients in the treated and placebo groups did not have different self-reported arthritis outcomes on average after one month of treatment, by three months of treatment, patients treated with auranofin reported "good" arthritis with a significantly higher probability than their counterparts in the placebo group.

It is also worth noting that patients treated with auranofin didn't just improve relative to the placebo group on average, but also improved relative to their own starting points on average. Indeed, since $0.44 \div 0.23 \approx 1.9 \approx 2$, the probability of a patient treated with auranofin reporting "good" arthritis nearly doubled (on average) from the beginning of the study to the three-month mark!