As I remember, the design has

5 interventions, which I’ll call Baseline, A, B, C and D

And 5 weeks, which I’ll call 0,1, 2,3 and 4

Roughly, the number of observations in each cell should be proportional to the numbers below, assuming N total subjects.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | **Week** | | | | |
| **0** | **1** | **2** | **3** | **4** |
| **Intervention** | **Baseline** | N | 0 | 0 | 0 | 0 |
| **A** | 0 | N/4 | N/4 | N/4 | N/4 |
| **B** | 0 | N/4 | N/4 | N/4 | N/4 |
| **C** | 0 | N/4 | N/4 | N/4 | N/4 |
| **D** | 0 | N/4 | N/4 | N/4 | N/4 |

I’m not sure of the actual coding for Week, which can make a difference in interpreting some of the coefficients. The problem is that differences between the baseline group and the active interventions may be due to an effect of an intervention or to an effect between Week 0 and Weeks 1-4. There is no data to differentiate the time effect (between week 0 and the other weeks) from the intervention effect (between baseline and the other interventions).

This is why you might have seen intervention effects look negative sometimes and positive at others, depending on how time was handled. The problem is that the significance or non-significance of any intervention effect (relative to baseline) depends entirely on the size of the time effect.

**What can be done?**

1. Assume there is not time effect. This is a little worrisome given there may be some evidence of one.
2. Test whether there is a time effect. This can be done within the baseline week using a day variable. IT can also be done within weeks 1-4. Again, I would code time in terms of days for this. Given the simplifications of the model, the results may be different than before. If there is no statistical evidence for a time effect, you could omit it in the primary analysis.
3. Assume that the time effect is linear or at least smooth for all five weeks. This is not implausible, but leads to the following.
4. a. Analyze only weeks 1-4. For example, you could use the delta (intervention – baseline) as the dependent variable and look at whether it differs from 0 on average. However, this is sidestepping the problem as the delta conflates both the time and intervention effects. Reviewers will want to know if there is the change relative to baseline is an intervention effect or a time effect (i.e. a sort of regression to the mean).

b. An alternative I would suggest here is to use the % increase in number of steps as the dependent variable. That might take care of the some of the heteroscedasticity issue (i.e. the variance of the residuals varying according to the fitted value). Personally, I would not know how to interpret a 500 increase in the number of steps, say. It doesn’t sound big, but is it? If you told me an intervention caused a 5% increase, it might be more meaningful.

c. IF you use % increase for weeks 1-4, then you could include a day effect and test whether the intercept equals one, which it would if there were no effect at baseline. You have to be careful about how the variables are coded here. (As above, use a day variable for time in study.)

d. Using the number of steps as in 4a with a day effect, you could also test whether the intercept goes through 0. This is parallel to 4c, but using change in number of steps instead of % change.

1. You could include some sort of sensitivity analysis. That is, you could formally consider whether there would still be an intervention effect for some hypothetical possible time effects. It’s a little difficult to know what those effects should be however.

My recommendation is to use option 4c above.