# Will the real subject-specific odds ratio please stand up?

## or: Why the logistic-Normal is not my favorite model.

Suppose we are evaluating an anti-smoking intervention and we have the outcome variable $Y$ indicating whether the person smoked during the past week and $X$ indicating whether they received the intervention.

The logistic regression model is

$$\text{logit} E\left[Y_i\right] = \alpha + \beta X_i.$$

The effect of treatment can be measured by the odds ratio $\exp(\beta)$. Everything is fine.

But I forgot to tell you that each person is evaluated three times. We now have two regression models

$$\begin{aligned} \text{logit}E\left[Y_{it}\right] &= \alpha + \beta X_{it} \\ \text{logit}E\left[Y_{it}|\epsilon_i\right] &= \alpha^* + \beta^* X_{it} + \epsilon_i \end{aligned}$$

The first is a marginal model, the second is a conditional model. Here $\exp(\beta^*)$ is the *subject-specific* odds ratio. In general $|\beta^*| > |\beta|$. Now we might say that $\beta^*$ measures the actual treatment effect, and $\beta$ has been attenuated.

But I forgot to tell you that this is a group discussion intervention and the groups may be different. We now have

$$\begin{aligned}\text{logit}E\left[Y_{git}\right] &= \alpha + \beta X_{git} \\ \text{logit}E\left[Y_{git}|\epsilon_i, \eta_g\right] &= \alpha^{**} + \beta^{**} X_{git} + \epsilon_i + \eta_g\end{aligned}$$

Now exp($\beta^{**}$) is the real *subject-specific* odds ratio, and we realise that exp($\beta^{*}$) was an attenuated version of it — it was only the group-specific odds ratio.

But I forgot to tell you that the group discussion was facilitated by the primary care physician, so the study was actually randomised by medical practice. We need a random effect for doctor, so we have

$$\begin{aligned} \text{logit} E\left[Y_{dgit}\right] &= \alpha + \beta X_{dgit} \\ \text{logit} E\left[Y_{dgit} | \epsilon_i, \eta_g, \zeta_d\right] &= \alpha^{***} + \beta^{***} X_{dgit} + \epsilon_i \\ &\quad + \eta_g + \zeta_d \end{aligned}$$

Now the subject-specific odds ratio is really $\exp(\beta^{***})$ and it's even bigger than we thought. The marginal odds is still boringly stuck at $\exp(\beta)$.

Note that we haven't even started to consider

- how to model the random effects

- what estimators to use

- how to fit the model

- what happens if the random effects are misspecified