

The effect of nasal mucus viscosity on annual nosebleed rates

David Ruau*

December 2017

Bayer is inlicensing a new treatment (superdupripine) for the unmet clinical need of recurrent serious nosebleeds¹. The drug has just undergone a small proof-of-concept trial in which groups of patients have been assigned to either superdupripine or placebo groups, have been followed for some time, and we have counted, for each patient, the number of nosebleeds that have required hospitalization. Bayer is considering whether or not to fund an expensive large-scale Phase III clinical trial for this compound.

It is suggested that the viscosity of nasal mucus is an important predictor of the effect of superdupripine, with the **drug much more effective in patients with high nasal mucus viscosity**. Accordingly, the project team wish to focus their Phase III trial on patients for whom superdupripine is most efficacious. The project team want to only select patients with **nasal mucus viscosity greater than some value**. Also, the most medically serious cases are those where the patient buys a **large amount of paper tissues**. The project team may also wish to restrict the Phase III study to just these patients. They have already produced a number of tables of the results of classical statistical hypothesis tests, but they have not yet really looked at the data or gained any understanding of it.

Your job is to help the project team understand the data that they have so that they can **explain it to the people in Bayer who decide whether we'll fund an expensive Phase III program**. You have been provided with some datasets from their study (described in table 1 below) and an R installation. Your task is to look at the data and produce what **summaries and visualizations** you think reasonable before you go into a meeting with the **project team** to discuss what can be done.

Selecting from the following questions may help you focus your work.

- How might you show how the treatment effect depends on nasal mucus viscosity? What about the effect of paper tissues?
- **Hospitalization for nosebleed may depend on local medical practice**. Does this have any impact? How can you understand this?
- How might you predict the rate of nosebleed from the data that you have? What might a statistical model for this look like?

*david.ruau@bayer.com

¹This is a purely fictitious example designed only for candidate assessment in recruitment.

file name	field	description
randomization.csv	subject	the subject identifier
	arm	the assigned treatment arm
subject.csv	subject	the subject identifier
	country	the country of the subject
	eye.colour	eye colour of subject
	mucus.viscosity	the nasal mucus viscosity at baseline
	tissue.use	whether the subject uses lots of paper tissues (assessed at baseline)
efficacy.csv	previous.year	number of nosebleeds requiring hospitalization in previous year
	subject	the subject identifier
	nosebleeds	the number of nosebleeds observed on study
	duration	the time that the subject was on the study

Table 1: Data that you have available.

- How can you use such a statistical model to simulate a Phase III trial? What inputs would it need, how would you generate them, and what outputs would it have?