In a diabetes study, several traits and information were collected, including (see data in the attached file “diabetes\_study.txt”):

Sex: male/female

Age: age of the study subject

bmi: body mass index

fbg: fasting blood glucose

fins: fasting insulin

hba1c: hemoglobin A1c

tg: total glyceride

tcho: total cholesterol

hdl: high density lipoprotein

ldl: low density lipoprotein

The investigators are interested in the effects of different factors on hba1c. You, as a biostatistician, is helping in the data analysis. Based on prior knowledge, the focus will be put on fbg, tg, fins, and bmi. For the following issues, please briefly summarize and report what you have done and what you have found out. Make sure that your report is not just a simple aggregate of software outputs!

1. Using the data in their original scales, please find an appropriate regression model. The regression assumptions should be evaluated and considered by the fitted model if necessary.
2. Conduct appropriate data transformation and find a corresponding fitted model.
3. Assess which of the two models established in a and b is more appropriate for the data. Provide a brief justification for your conclusion.