**Goal:** investigate variables which are correlated with treatment respons (=outcome)

**Outcomes (not decided which one yet, but preference)**

1. Binair\_outcome: <50% en ≥50%
2. Grainne: compleet (0 RBC), partial (iets van een reductie), on (0 of toename)
3. **Klimova: compleet (o RBC), partial (≥50%) en non (<50%)**
4. Nieuw toegevoegd – 3cat\_75%: (o RBC), partial (≥75%) en non (<75%)

**Analysis without imputation**

1. Uni-variate correlation with outcome ‘binair’ and the predictors: variables ‘study’, age, gender, stomach etc. tested (see Summary)
2. Multivariate model: full model with backward selection . Study included as variable to correct for differences between studies.

* Problem here is the not normal output of OR’s in the multinominal model (for the outcomes ‘Grainne and Klimova’).

**Imputation:**

-Analyse missings: only imputed data with <50% missings and used all predictors for the multiple imputation

- Result: 20 dataset (of which 1 original included)

- Imputation variables: age, gender, Hb prior and end study, prior APC, medicatie anticoagulantia etc

Explanation of the imputation method:

<http://www.ibm.com/support/knowledgecenter/SSLVMB_21.0.0/com.ibm.spss.statistics.help/mi_data.htm>

**Analysis with imputation:**

1. Univariate correlation with uitcome ‘binair’: variables ‘study’, age, gender, stomach etc. tested (see samenvatting)
2. Multivariate model: full model en backward selection . Study included as variable to correct for differences between studies.

* Problem here is the not normal output of OR’s in the multinominal model (for the outcomes ‘Grainne and Klimova’).

**Solutions?? - things I tried or suggested as possible options.**

* Generalized estimating equations: extends the generalized linear model to allow for analysis of repeated measurements or other correlated observations, such as clustered data 🡪 bij mij niet het geval? Of is studie de geclusterde data?

Als subject variable (study) invoert en rest hetzelfde doet, dan geen warnings bij output. Geanalyseert als geimputeerde data. **Geen random effects!!**

* Handleiding – multilevel modeling of categorical outcomes using SPSS 🡪 mijn data is geen multilevel data, tevens alleen toegang hfd1 en niet hfd 4 (binary outcomes)
* **Generalized linear mixed models (GLMM) 🡪 covers logistic models for binary outcomes (generlazied gedeelte), kan ook met mixed models werken (mixed gedeelte)**

Echter niet imputatie teken ervoor – kan het met geimputeerde data?

Generalized linear model 🡪 kan geen random effects toevoegen (study), tevens warnings als zonder study bekijkt.

* General linear model 🡪 alleen continue uitkomstmaat

**Main problem:**

* Clustered (the studies) and imputed data (20 datatsets)
* Preferred outcome: 3 categories
* Preferred analyses: multivariate regression model
* There are errors, strange OR’s in the multivariate models in SPSS for the multinominal outcomes
* **I think we have to do a GLMM, however this is in SPSS not possible with imputated data.**