Program Validation – Care for a black sheep
Andrea Weber, PRA HealthSciences (weberandrea@prahs.com)

**Situation**

- Oh dear hundreds of differences. That will take me ages

**Problem**

- Plain proc compare
- Ignoring differences in number of observation
- Ignoring sorting order

- Comparing apples with oranges

- Big amount of data / complex data

- Sense of being overwhelmed

- Dependencies

- Issues caused by parent data.

- No information on what differs
- Is the main program incorrect or rather the qc program?
- main programmer must repeat the qc to find out

- Obs 35277 of main dataset is not necessarily the same record as in qc in particular in case of subset or different sorting

- Independent Issues reported dropwise.

**Solution**

- Ensure correct number of observations
- Ensure compare of associated records

- Proc compare with ID statement
- Datasets identically sorted
- Match in ID variables

- Qc subset of data
- Plausibility checks

- Frequency, Min, Max
- Compare values against CT

- Qc parent data first.

- Proc compare for parent variables only
- Child after match of parent

- Explain why considered incorrect

- Hints to specification/ SAP/guidelines
- Assumed route-cause (Optional)

- Expected values

- Provide examples

- Subject ID
- Further key variables

- Report bunch of issues

**Problem Solution**

- No false-positive issues
- Less number of overall issues

- Reduce proc compare outcomes
- Increase visibility of progress
- Parallel work

- Dependent issues disappear
- Less workload

- Quick understanding of issue by main side programmer
- Easy trace of main programming progress

- Less rounds of updates and re-qc.