Individual Program Level

- Always include a program header
- Write comments that clearly explain the rationale of the code
- Follow consistent coding conventions, e.g.
  - Do not overwrite existing datasets
  - Use ‘dataset’ option in procedure statements
  - When converting character variables to numeric or vice versa, use the put and input functions to explicitly convert the variables
- Write programs to be reusable
  - Do not include study-specific values in programs
  - Use macros parameters to allow flexibility

Study Level

Study Programmer: Responsible for overseeing all programs for a study and ensuring they work together

**Principle 1: Understand the study and data**

- Study Protocol
- Annotated CRF
- SAP & Shells
- Data Specifications

PhUSE GPP Steering Board recommends “Defensive programming”

- Write programs that anticipate future changes of data
- Understanding data helps with defensive programming

**Principle 2: Organize the Programs**

- One main program per analysis dataset
- Analysis-ready datasets so TLFs can be produced with minimal programming
- One main program per TLF

**Principle 3: Follow industry and company standards**

**CDISC**

- Run OpenCDISC checks early, prior to final CSR output

**Programming Standards**

- Know and use standard tools and macros

**Principle 4: Consistency across programs and data cuts**

- Naming conventions
- Consistent algorithms across programs and data cuts
- Common variables have same values in all datasets

**Principle 5: Document data and programs**

**Documentation is just as important as programs!**

- Program header and comments
- Documentation of all data cuts for study (purpose, date, location)
- Documentation of all data sources for study, especially non-CRF sources
- Analysis dataset specifications
  - Can use spec as the basis for Define file
  - Good practice to complete specs before programs are QCed so QCer can QC both programs and specs at same time

Compound or Therapeutic Area Level

Project/Compound Programmer: Responsible for compound-wide consistency in data structure and programming algorithms

**Additional Responsibilities at Compound or TA Level**

- Compound or Therapeutic Standards
  - Coalition for Accelerating Standards and Therapies (CFAST) developing TA standards
- Compound-level macros
- Cross-Compound interrogations of data

The Wiki is a place to share your ideas and post your contributions. Please provide your comments!