

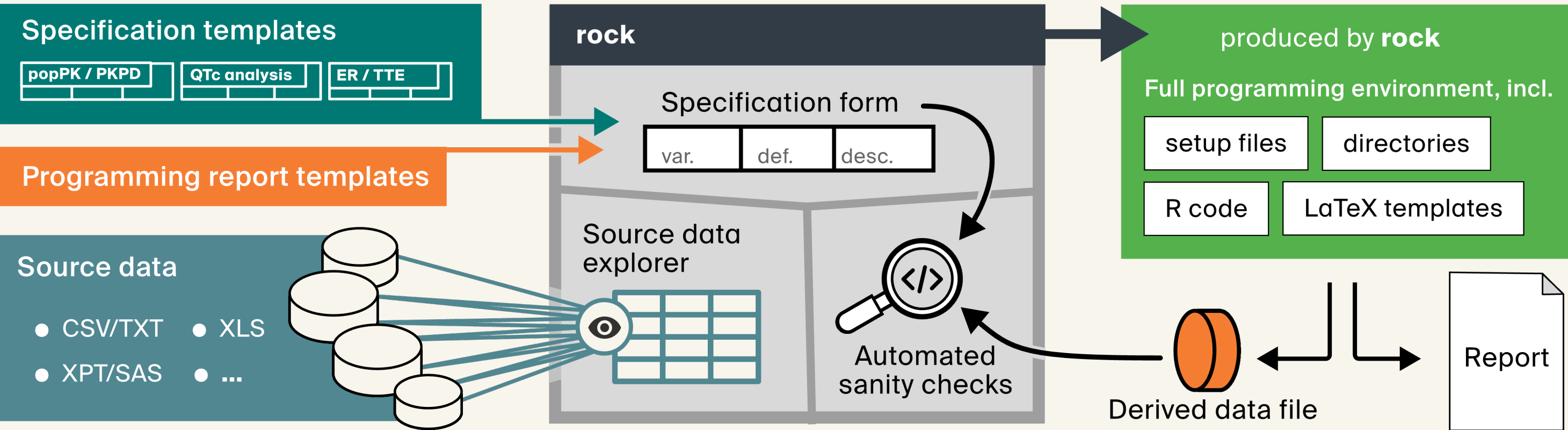
rock and roll ...
Deliverables driven by standardized data specifications

Eric Sandström, Joakim Hellner, Georges Ette, Arnaud Gaby, Svetlana Freiberga, Pēteris Prūsis, Mounika Suddapalli
Pharmetheus AB, Sweden. Contact: dpteam@pharmetheus.com

Objective

rock is a web application built using R shiny and is the starting point of any data request to be initiated. Its primary purpose is to provide a flexible framework to design data request specifications. It also provides several additional functionalities meant to accompany the data programmer through the journey of creating data request deliverables, enforcing adherence to standards and high quality.

The rock framework



Data specification generation

Figure 1. Variable attribute selection and editing – name, definition, unit, possible values, description, rounding and programming dependencies.

Figure 2. Variable source editing – study, file and variables, record selection, derivation details.

Data management report					
Creation of DAT-1a-PMX-CONF-PMX-1.csv					
Appendix 2. Data definition table					
Table A2-1: Data definition table					
Variable	Definition	Possible values	Description	Unit	Source
REF	Unique record number	Integer ≥ 1	Consecutive numbers starting from 1.	N/A	Produced
STUDYIDN	Study identifier	1, 2	1 = STUDY1 2 = STUDY2	N/A	STUDY1 Derived from dm.sas7bdat: STUDYID STUDY2 Derived from dm.xpt: STUDYID
ID	Unique subject identifier within current file	Integer ≥ 1	Consecutive numbers starting from 1 to max number of subjects within current file.	N/A	Produced
TYPE	Type of record	0, 1	0 = Dose event 1 = PK analyze plasma concentration	N/A	Produced
EVID	NONMEM event identification	0, 1	0 = Observation event 1 = Dose event	N/A	Produced
DV	NONMEM dependent variable	Numeric	Observation records (EVID=0): Observed value. BLQ=1: Set to the lower limit of quantification (LLOQ) divided by 2. Dose records (EVID=1): 0	ug/ml	STUDY1 Derived from pc.sas7bdat: PCSTRESN where PCSPEC=PLASMA and PCTESTCD=XAN STUDY2 Derived from pc.xpt: PCSTRESN where PCSPEC=PLASMA and PCTESTCD=X105
BLQ	Flag for concentration reported as below LLOQ in source data	0, 1	0 = No 1 = Yes	N/A	STUDY1 Derived from pc.sas7bdat: PCSTRESN, PCSTRESC STUDY2 Derived from pc.xpt: PCSTRESN, PCSTRESC
LNDV	Log-transformed dependent variable	Numeric, DV	Natural logarithm of positive pharmacokinetic (PK) plasma observations. Rounded to 4 decimals. DV = Set to DV for other record	N/A	Produced

Figure 3. Data specification output.

Conclusions

Using **rock** as a foundation for creating deliverables offers the following advantages:

- Standardized code and structure, facilitating easier handover
- Reproducibility, allowing for source data updates and re-execution
- Enforced traceability through versioning
- Extensibility, with support for additional specification templates
- Diversity, with support for various programming report templates

Data programming report generation

rock creates the R/LaTeX programming environment required to produce the deliverables, including

- Standard directory structure and setup files to load R libraries and source data according to the specifications.
- R script placeholders for dedicated sections (dosing, pk, pd, baseline covariates etc),
- Ready-to-use standard LaTeX templates for generating the programming report.

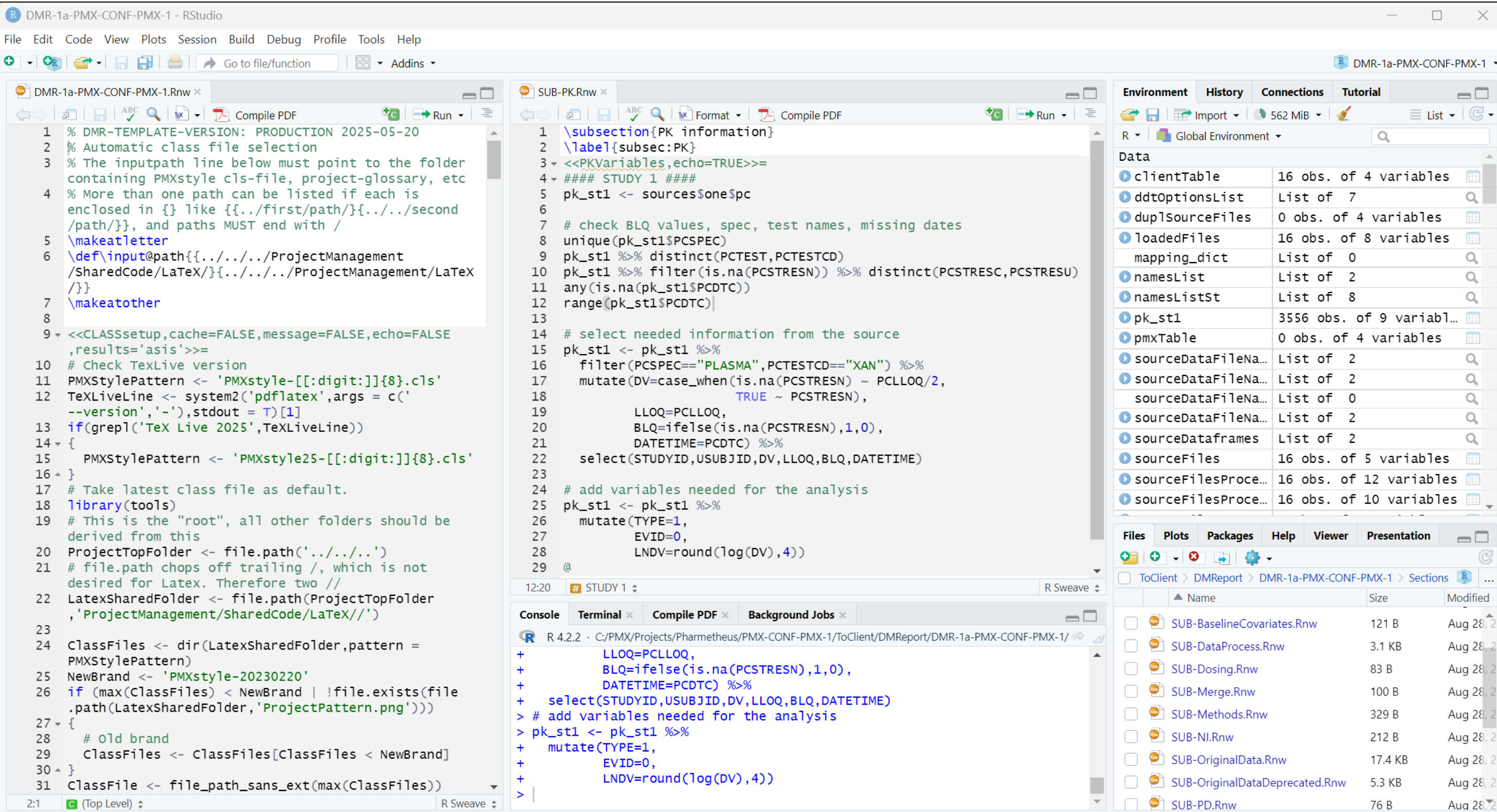


Figure 4. Programming environment framework.

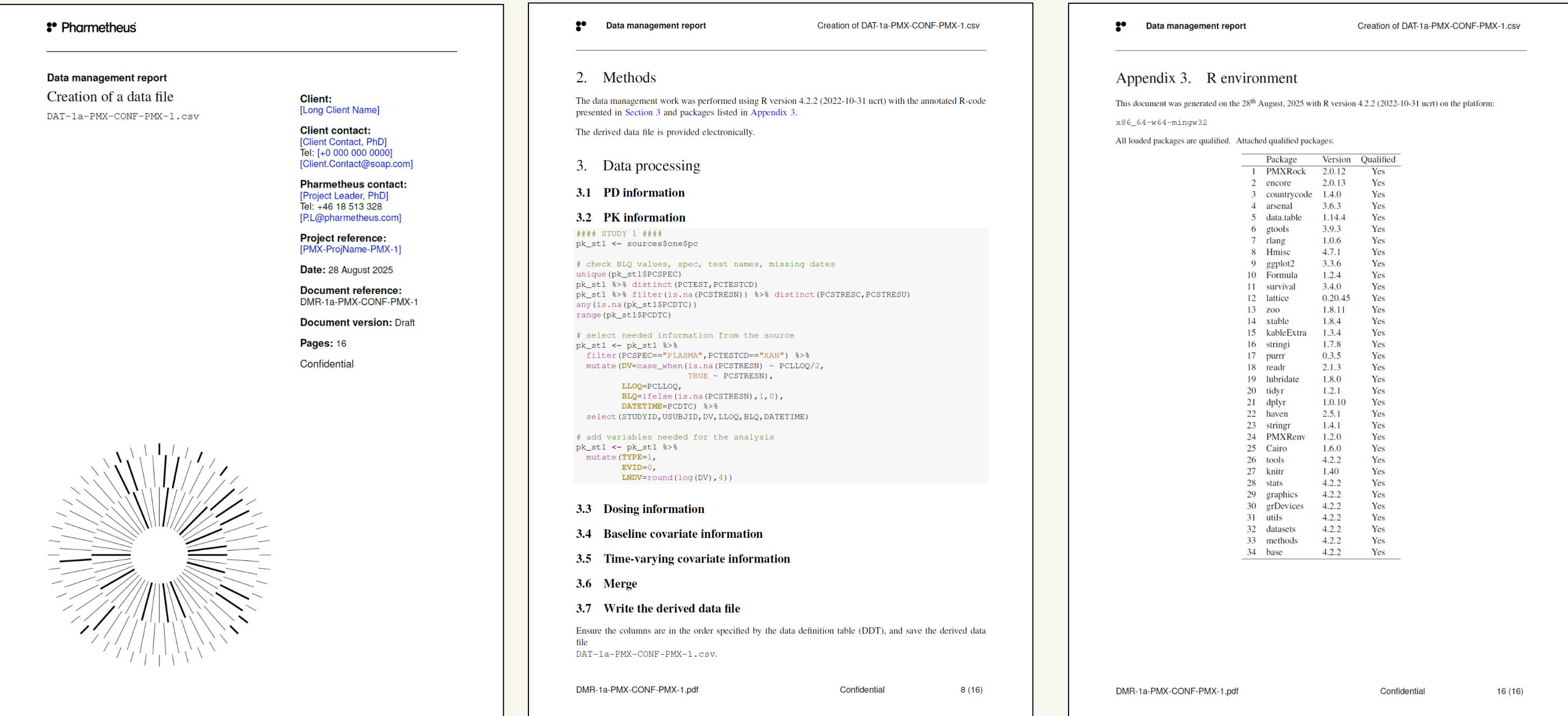


Figure 5. Data programming report with the programming code as a pdf file.

Automated sanity checks

- Summary statistics
- Compare output dataset against specifications
- Dependencies checks
- Generic and compliance checks
- Analysis specific tests

Figure 6. Automated sanity checks.

Variable	Source	Source viewer	DAT vs DDT	LaTeX macros	Global info	Export	Help	Preferences
DAT-1a-PMX-CONF-PMX-1.csv								
Compare file to DDT								
Constants to use in auto-QC								
LLOQ=1, LLOQ=1								
Variable name of unique patient identifier (e.g. ID or URUB.ID)								
ID								
Add continuity check to compare file to DDT?								
There are no columns with empty values.								
Variable	Possible Values	Values in DAT	Classification					
REF	Integer >= 1	min 1 median 3556.5 max 7112	numeric					
STUDYIDN	1, 2	1, 2	numeric					
ID	Integer >= 1	min 1 median 254.5 max 508	numeric					
TYPE	0, 1	1	all_values_identical					
EVID	0, 1	0	all_values_identical					
DV	Numeric	min 0 median 0.0133377195403065 max 1.9380948594273	numeric					
BLQ	0, 1	0, 1	numeric					
LNDV	Numeric, DV	min -inf median -4.31715 max 0.6617	numeric					
LLOQ	Numeric >= 0	0.01	all_values_identical					
SEAN	1, 2	1, 2	numeric					
RACEN	1, 2, 3, 4, 5, 6, 7, 8, 9	(4 unique values): 5, ..., 1	character					
AGEYBL	Numeric >= 0, .99	(37 unique values): 63, 64, 71, 74, 77, 85, 68, 81, 84, 52, 75, 57, 79, 96, ... (23 more unique values)	character					
USUBJID	Character	(508 unique values): 01-701-1015, 01-701-1023, 01-701-1028, 01-701-1033, 01-701-1034, 01-701-1047, 01-701-1097, 01-701-1111, 01-701-1115, 01-701-1116, 01-701-1130, 01-701-1133, 01-701-1146, 01-701-1148, ... (484 more unique values)	character					
DATETIME	Character	lubridate_yrmd_tm	lubridate_yrmd_tm					