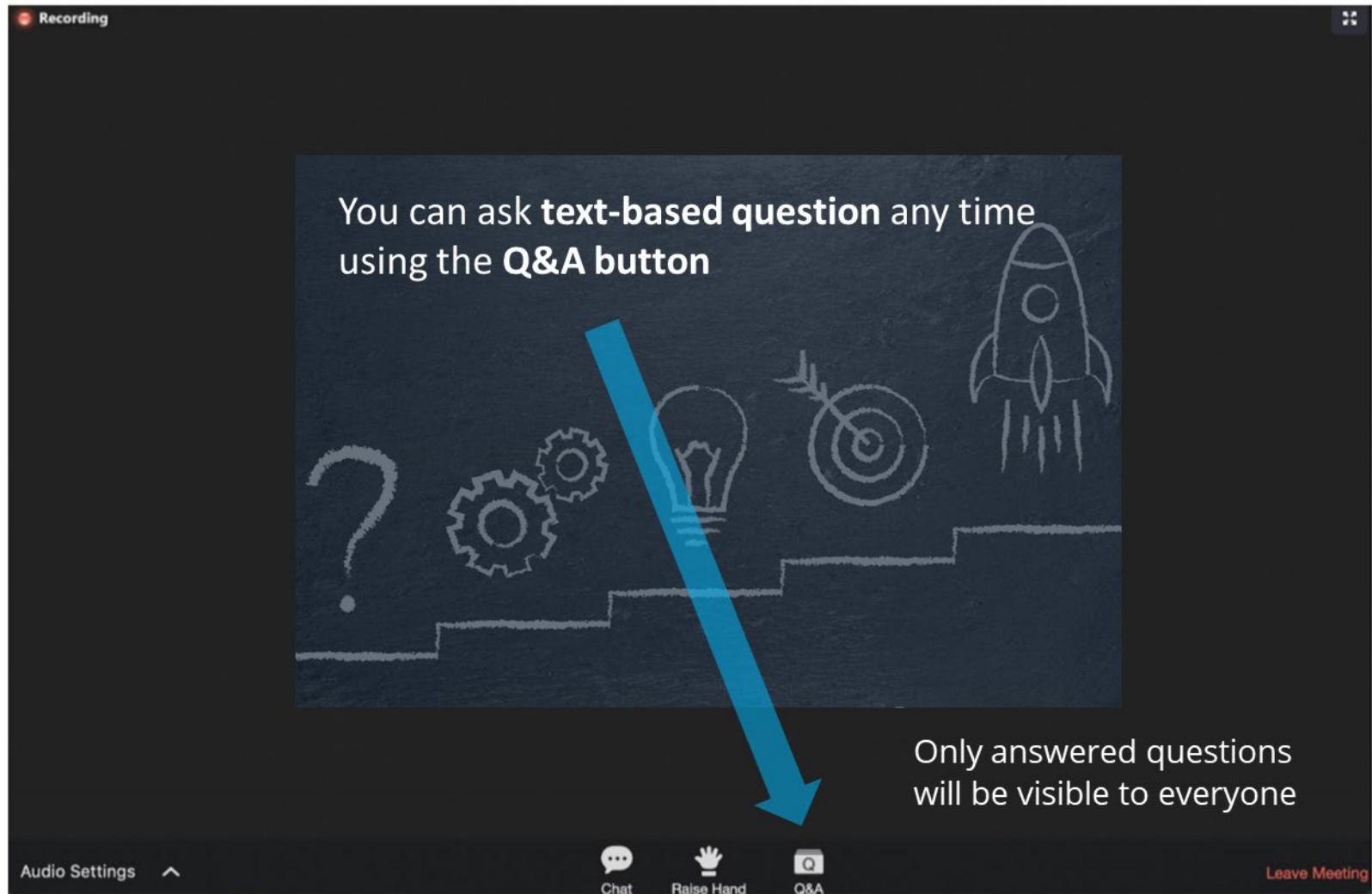


- The webinar starts at **2:00 pm Basel Time**
8:00 am East Coast Time
- Everyone is placed on mute during the webinar
- Webinar material (including presentation) can be downloaded here:
<https://training.intiquan.com/MIDDmodules/M2.1.zip>
- The webinar will be recorded
 - Recording will be made available on the following link:
<https://training.intiquan.com/MIDDmodules/M2.1.mp4>
 - Recording available ~1 day after the webinar



Q&A during Webinar



Recording

You can ask **text-based question** any time using the **Q&A button**

Only answered questions will be visible to everyone

Audio Settings ^

Chat Raise Hand Q&A

Leave Meeting

The screenshot shows a webinar interface with a dark background. At the top left, there is a red dot and the word "Recording". In the center, there is a chalkboard-style illustration with a question mark, gears, a lightbulb, a target, and a rocket, all connected by a series of steps. A large blue arrow points from the text "Q&A button" to the "Q&A" button in the bottom toolbar. The bottom toolbar also includes "Audio Settings", "Chat", "Raise Hand", and "Leave Meeting".

Module 2.1

Efficient Data format for NLME analyses

IntiQuan Webinar Series on efficient support of
Model Informed Drug Development (MIDD)

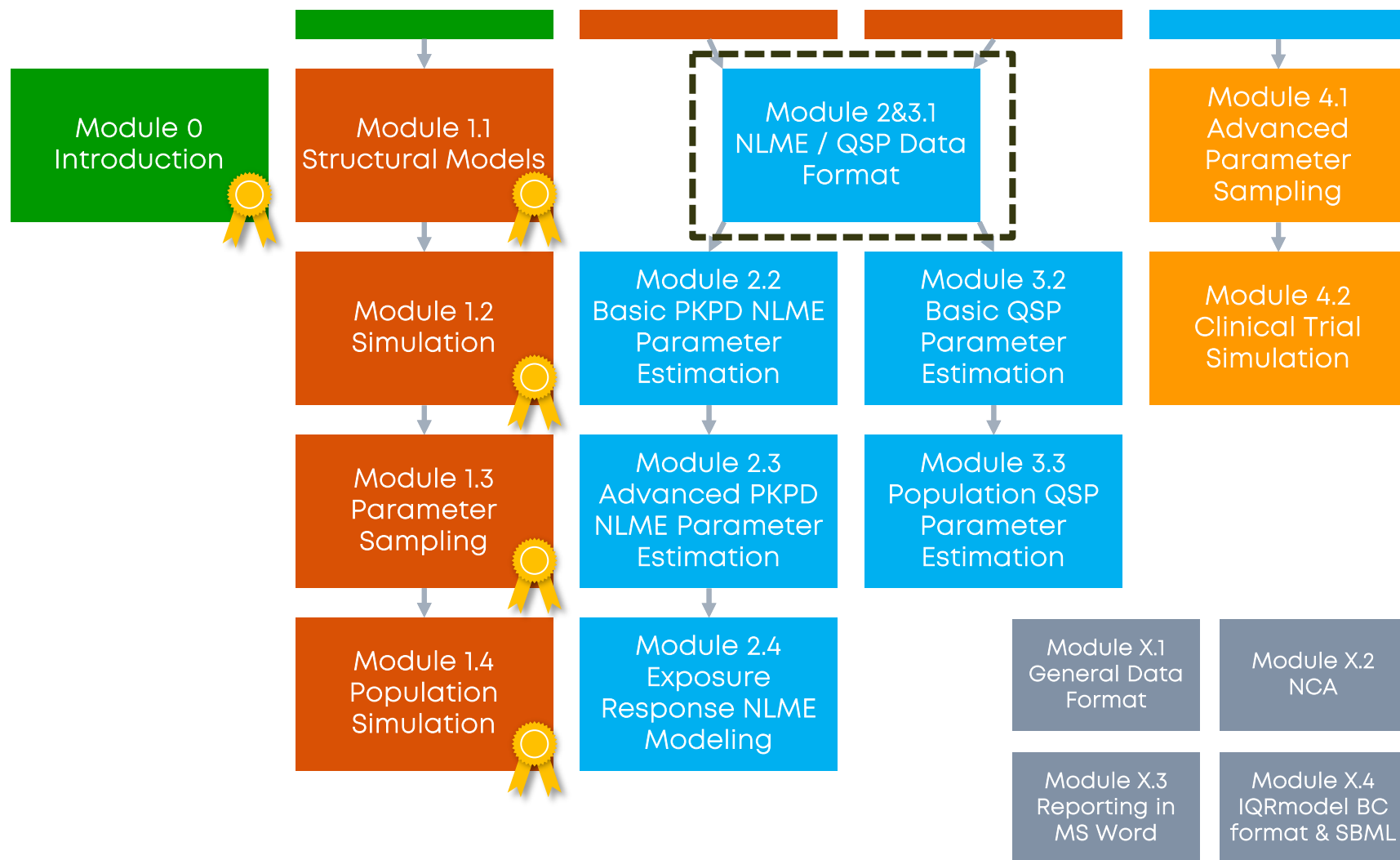
Outline

1. Background
2. NLME data format explained
3. Introduction General Dataset Format (GDF)
4. Conclusions
5. Outlook webinar modules
6. Outlook on tool requirements modules 2.2-2.4
7. Q&A

Background

Overview of Webinar Modules

IntiQuan Webinar Series on efficient support of Model Informed Drug Development (MIDD)



Increasing need for cross-discipline approaches

Successful implementation of MIDD requires several disciplines (e.g., QSP & pharmacometrics) working together and being able to exchange information (e.g., data and models)

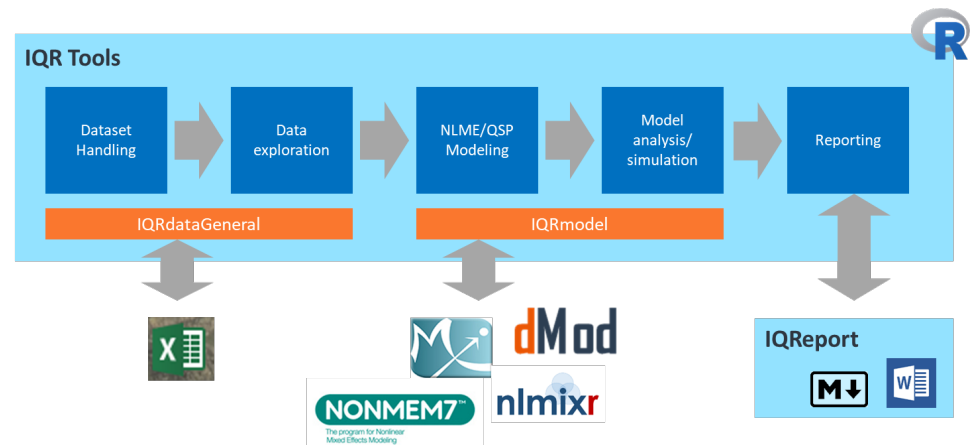
- Different tools, different model and data formats, different approaches, different terminology, different perception, different thinking, different individuals, different groups, ...

IQR Tools Approach: Using same modeling platform, interfacing different tools depending on needs and questions to be addressed

Benefits:

- QSP and pharmacometric analyses on same platform
- Exchangeability of data and models
- Same model and data formats
- Same workflow syntax
- Efficient conduct of analyses

• 7 ...



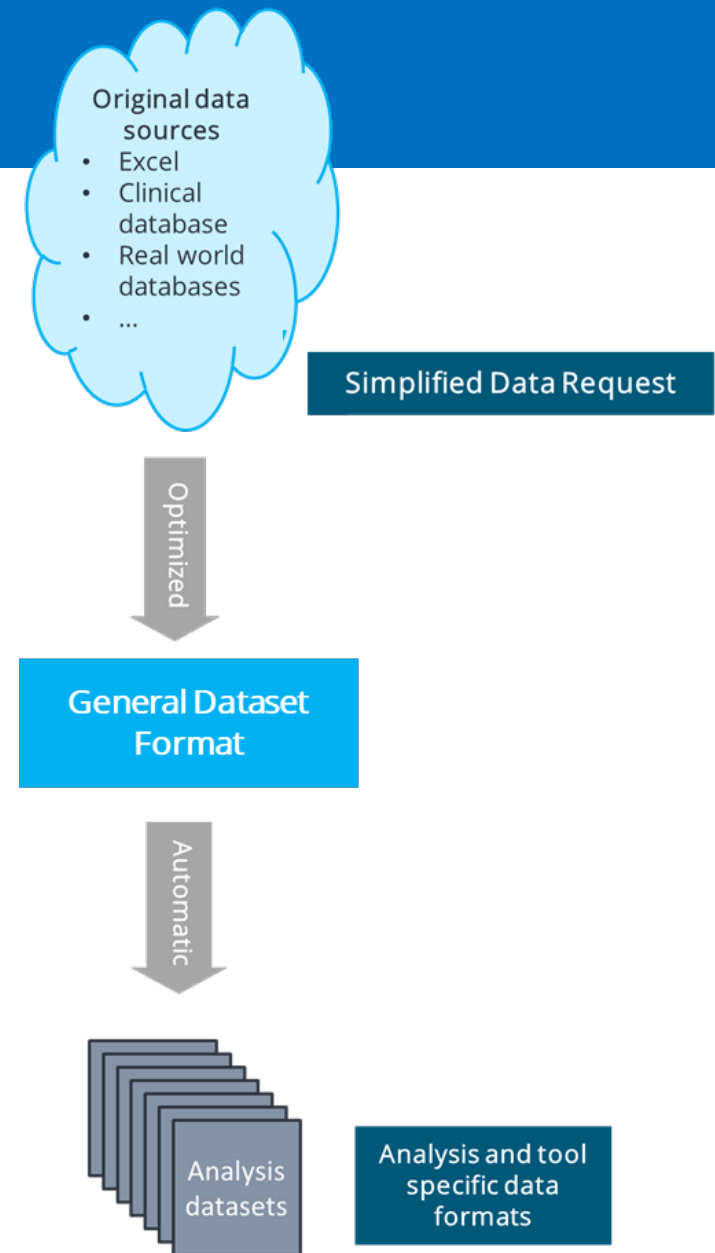
One bottleneck in MIDD is often the data situation

- Data preparation often takes considerable time
- Quality of results often call for many iterations until data can be used
- Often no standard data spec exists
- Analysis datasets are often programmed for one specific analysis and are usually not reusable for other analyses

IQR Tools Approach: Systematic approaches based on a standardized general dataset format allow to considerably streamline data programming

Benefits:

- Simplified specification, same standard format
- Reduced time to analysis data
- Improved quality
- Fewer iterations
- Data reusable across several analyses



Goals of this module

- ✓ You will have been introduced to an NLME dataset format that can be used for NONMEM, MONOLIX, and NLMIXR
- ✓ In addition, this dataset format allows to store critical metadata directly in the dataset – rendering the dataset self-explaining and it's handling less prone to errors than typical “numeric only” NONMEM datasets
- ✓ This format can be generated automatically starting from an even more convenient “general dataset format”

This module will give you the ability to define own datasets in the format used by IQR Tools

Requirements for execution of examples

- R \geq 3.5.0
- On Windows: Rtools (allowing to build R source packages that require compilation of C code)
<https://cran.r-project.org/bin/windows/Rtools/history.html>
- IQR Tools R package

```
install.packages(  
  "IQRtools",  
  repos = c("https://iqrtools.intiquan.com/rrepo",  
            "https://cran.r-project.org/"),  
  dependencies = TRUE  
)  
  
library(IQRtools)  
test_IQRtools()  
  
# If the last command results in a data.frame => you are ready to go!
```

Download of webinar material

- The Webinar material is available as a convenient download
- After installation of IQR Tools, simply type the following

```
library(IQRtools)  
install_MIDDmodule("2.1")
```

- Or download directly from:

```
https://training.intiquan.com/MIDDmodules/M2.1.zip
```

NLME data format explained

Typical Analysis Dataset

	B	D	F	K	M	P	Q	R	S	U	W	X	Y
1	ID	TRT	TIME	YTYPE	DV	MDV	EVID	AMT	ADM	RATE	DOSE	WT0	SEX
2	1	1	0	0	0	1	1	400	1	0	400	79.6	1
3	1	1	0.25	1	2.84	0	0	0	0	0	400	79.6	1
4	1	1	0.57	1	6.57	0	0	0	0	0	400	79.6	1
5	1	1	1.12	1	10.5	0	0	0	0	0	400	79.6	1
6	1	1	2.02	1	9.66	0	0	0	0	0	400	79.6	1
7	1	1	3.82	1	8.58	0	0	0	0	0	400	79.6	1
8	1	1	5.1	1	8.36	0	0	0	0	0	400	79.6	1
9	1	1	7.03	1	7.47	0	0	0	0	0	400	79.6	1
10	1	1	9.05	1	6.89	0	0	0	0	0	400	79.6	1
11	1	1	12.12	1	5.94	0	0	0	0	0	400	79.6	1

- Usually numeric only
- Programmed for one particular analysis – often not easily reusable
- Non-understandable without metadata or a dedicated data specification documentation

=> Room for improvement ...

IQRnlmeData format specification – minimal set of required columns. Suggested columns used as minimal meta data

COLUMN	INFO	FORMAT	MEANING
IXGDF	REQUIRED	numeric	A column with index numbers of the records in the dataset either defined for this dataset alone as 1,2,3, ... or defined in a dataset from which this one here has been derived and the IXGDF entries make the link between the remaining records and the original records.
USUBJID	REQUIRED	string	Unique subject identifier – to ensure mapping of subjects across different datasets and analyses
ID	REQUIRED	numeric	Modeling tool specific subject identifier
TIME	REQUIRED	numeric	Time relative to first dose (per subject)
TIMEPOS	REQUIRED	numeric	Time from first event in the dataset (per subject)
TAD	REQUIRED	numeric	Time after previous dose
TIMEUNIT	SUGGESTED	string	Time unit used in the dataset. No spaces allowed (exchange with ':')
NAME	SUGGESTED	string	Name of the event used for annotation purposes. No spaces allowed (exchange with ':')
YTYPE	REQUIRED	numeric	Number of output if observation, 0/NA/. if dosing event
DV	REQUIRED	numeric	Observed value for observations, 0 for doses, LLOQ value if CENS=1
UNIT	SUGGESTED	string	Unit of event record for annotation purposes. No spaces allowed (exchange with ':')
MDV	REQUIRED	numeric	Missing data flag
EVID	REQUIRED	numeric	Event ID (0 for observations, 1 for dosing records) Values of 3 and 4 can be used ... If desired and needed.
CENS	REQUIRED	numeric	Censoring flag (0 uncensored, 1 left censored/BLOQ)
AMT	REQUIRED	numeric	Dosing amount if dose event, 0 if observation event
ADM	REQUIRED	numeric	Number of the input if a dose, 0/NA/. if an observation
RATE	REQUIRED	numeric	Rate of administration if infusion, 0 if bolus, -2 if estimated
TINF	REQUIRED	numeric	Infusion time of administration if infusion, 0 if bolus (please define even if RATE is defined)
OCC	OPTIONAL	numeric	An OCC column might be present and can be used to define distinct occasions for consideration of IOV in the modeling. An OCC column needs to contain numeric entries (integers only allowed).
"OTHER columns"	OPTIONAL	numeric or string	Covariates, regression parameters, additional documentation, etc.

Datasets in this format are used by IQR Tools for parameter estimation with NONMEM, MONOLIX, NLMIXR, and the QSP algorithm

Example 1

Simple dataset in IQRnlmeData format

data_example_1.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	IXGDF	USUBJID	ID	TIME	TIMEPOS	TAD	TIMEUNIT	NAME	YTYPE	DV	UNIT	MDV	EVID	CENS	AMT	ADM	RATE	TINF	DOSE	WTO	SEX
2	1	Theo_1	1	0	0	0	Hours	Theophylline:::Dose	0	0 mg		1	1	0	400	1	0	0	400	79.6	1
3	2	Theo_1	1	0.25	0.25	0.25	Hours	Theophylline:::Concentration	1	2.84 ug/ml		0	0	0	0	0	0	0	400	79.6	1
4	3	Theo_1	1	0.57	0.57	0.57	Hours	Theophylline:::Concentration	1	6.57 ug/ml		0	0	0	0	0	0	0	400	79.6	1
5	4	Theo_1	1	1.12	1.12	1.12	Hours	Theophylline:::Concentration	1	10.5 ug/ml		0	0	0	0	0	0	0	400	79.6	1
6	5	Theo_1	1	2.02	2.02	2.02	Hours	Theophylline:::Concentration	1	9.66 ug/ml		0	0	0	0	0	0	0	400	79.6	1
7	6	Theo_1	1	3.82	3.82	3.82	Hours	Theophylline:::Concentration	1	8.58 ug/ml		0	0	0	0	0	0	0	400	79.6	1
8	7	Theo_1	1	5.1	5.1	5.1	Hours	Theophylline:::Concentration	1	8.36 ug/ml		0	0	0	0	0	0	0	400	79.6	1
9	8	Theo_1	1	7.03	7.03	7.03	Hours	Theophylline:::Concentration	1	7.47 ug/ml		0	0	0	0	0	0	0	400	79.6	1
10	9	Theo_1	1	9.05	9.05	9.05	Hours	Theophylline:::Concentration	1	6.89 ug/ml		0	0	0	0	0	0	0	400	79.6	1
11	10	Theo_1	1	12.12	12.12	12.12	Hours	Theophylline:::Concentration	1	5.94 ug/ml		0	0	0	0	0	0	0	400	79.6	1
12	11	Theo_1	1	24.37	24.37	24.37	Hours	Theophylline:::Concentration	1	3.28 ug/ml		0	0	0	0	0	0	0	400	79.6	1
13	12	Theo_10	2	0	0	0	Hours	Theophylline:::Dose	0	0 mg		1	1	0	400	1	0	0	400	58.2	1
14	13	Theo_10	2	0.37	0.37	0.37	Hours	Theophylline:::Concentration	1	2.89 ug/ml		0	0	0	0	0	0	0	400	58.2	1
15	14	Theo_10	2	0.77	0.77	0.77	Hours	Theophylline:::Concentration	1	5.22 ug/ml		0	0	0	0	0	0	0	400	58.2	1
16	15	Theo_10	2	1.02	1.02	1.02	Hours	Theophylline:::Concentration	1	6.41 ug/ml		0	0	0	0	0	0	0	400	58.2	1
17	16	Theo_10	2	2.05	2.05	2.05	Hours	Theophylline:::Concentration	1	7.83 ug/ml		0	0	0	0	0	0	0	400	58.2	1
18	17	Theo_10	2	3.55	3.55	3.55	Hours	Theophylline:::Concentration	1	10.21 ug/ml		0	0	0	0	0	0	0	400	58.2	1
19	18	Theo_10	2	5.05	5.05	5.05	Hours	Theophylline:::Concentration	1	9.18 ug/ml		0	0	0	0	0	0	0	400	58.2	1
20	19	Theo_10	2	7.08	7.08	7.08	Hours	Theophylline:::Concentration	1	8.02 ug/ml		0	0	0	0	0	0	0	400	58.2	1
21	20	Theo_10	2	9.38	9.38	9.38	Hours	Theophylline:::Concentration	1	7.14 ug/ml		0	0	0	0	0	0	0	400	58.2	1
22	21	Theo_10	2	12.1	12.1	12.1	Hours	Theophylline:::Concentration	1	5.68 ug/ml		0	0	0	0	0	0	0	400	58.2	1
23	22	Theo_10	2	23.7	23.7	23.7	Hours	Theophylline:::Concentration	1	2.42 ug/ml		0	0	0	0	0	0	0	400	58.2	1
24	23	Theo_11	3	0	0	0	Hours	Theophylline:::Dose	0	0 mg		1	1	0	400	1	0	0	400	65	0
25	24	Theo_11	3	0.25	0.25	0.25	Hours	Theophylline:::Concentration	1	4.86 ug/ml		0	0	0	0	0	0	0	400	65	0
26	25	Theo_11	3	0.5	0.5	0.5	Hours	Theophylline:::Concentration	1	7.24 ug/ml		0	0	0	0	0	0	0	400	65	0
27	26	Theo_11	3	0.98	0.98	0.98	Hours	Theophylline:::Concentration	1	8 ug/ml		0	0	0	0	0	0	0	400	65	0
28	27	Theo_11	3	1.98	1.98	1.98	Hours	Theophylline:::Concentration	1	6.81 ug/ml		0	0	0	0	0	0	0	400	65	0

Datasets in this format are used by IQR Tools for parameter estimation with NONMEM, MONOLIX, NLMIXR, and the QSP algorithm

Example 1

Mapping of doses and observations between dataset and IQRmodels

example_1_data.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	IXGDF	USUBJID	ID	TIME	TIMEPOS	TAD	TIMEUNIT	NAME	YTYPE	DV	UNIT	MDV	EVID	CENS	AMT	ADM	RATE	TINF	DOSE	WT0	SEX
2	1	Theo_1	1	0	0	0	Hours	Theophylline::Dose	0	0 mg		1	1	0	400	1	0	0	400	79.6	1
3	2	Theo_1	1	0.25	0.25	0.25	Hours	Theophylline::Concentration	1	2.84 ug/ml		0	0	0	0	0	0	0	400	79.6	1
4	3	Theo_1	1	0.57	0.57	0.57	Hours	Theophylline::Concentration	1	6.57 ug/ml		0	0	0	0	0	0	0	400	79.6	1
5	4	Theo_1	1	1.12	1.12	1.12	Hours	Theophylline::Concentration	1	10.5 ug/ml		0	0	0	0	0	0	0	400	79.6	1
6	5	Theo_1	1	2.02	2.02	2.02	Hours	Theophylline::Concentration	1	9.66 ug/ml		0	0	0	0	0	0	0	400	79.6	1
7	6	Theo_1	1	3.82	3.82	3.82	Hours	Theophylline::Concentration	1	8.58 ug/ml		0	0	0	0	0	0	0	400	79.6	1
8	7	Theo_1	1	5.1	5.1	5.1	Hours	Theophylline::Concentration	1	8.36 ug/ml		0	0	0	0	0	0	0	400	79.6	1
9	8	Theo_1	1	7.03	7.03	7.03	Hours	Theophylline::Concentration	1	7.47 ug/ml		0	0	0	0	0	0	0	400	79.6	1
10	9	Theo_1	1	9.05	9.05	9.05	Hours	Theophylline::Concentration	1	6.89 ug/ml		0	0	0	0	0	0	0	400	79.6	1
11	10	Theo_1	1	12.12	12.12	12.12	Hours	Theophylline::Concentration	1	5.94 ug/ml		0	0	0	0	0	0	0	400	79.6	1
12	11	Theo_1	1	24.37	24.37	24.37	Hours	Theophylline::Concentration	1	3.28 ug/ml		0	0	0	0	0	0	0	400	79.6	1
13	12	Theo_1	1	0	0	0	Hours	Theophylline::Dose	0	0 mg		1	1	0	400	1	0	0	400	79.6	1

Link between model and data is made by **INPUT1** in model obtaining dose events with **ADM=1** and **OUTPUT1** in model compared to observation record with **YTYPE=1**

No need for a compartment (CMT) column

See also example_1.R to generate and run NONMEM, MONOLIX, and NL MIXR models

```

1 ***** MODEL NAME
2
3 modelPK
4
5 ***** MODEL NOTES
6
7 Simple linear one compartment PK model with 1st order absorption.
8
9 ***** MODEL STATES
10
11 d/dt(Ad) = -ka*Ad + INPUT1
12 d/dt(Ac) = ka*Ad - CL/Vc*Ac
13
14 ***** MODEL PARAMETERS
15
16 ka = 1 # Absorption rate parameter (1/hours)
17 CL = 3 # Apparent clearance (L/hours)
18 Vc = 60 # Apparent central volume (L)
19
20 ***** MODEL VARIABLES
21
22 # Calculate plasma concentration in ug/mL
23 Cc = Ac/Vc
24
25 # Assign output variable for matching with dataset
26 OUTPUT1 = Cc # Plasma concentration (ug/mL)
27

```

Example 2

Handling BLOQ information

example_2_data.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	IXGDF	USUBJID	ID	TIME	TIMEPOS	TAD	TIMEUNIT	NAME	YTYPE	DV	UNIT	MDV	EVID	CENS	AMT	ADM	RATE	TINF	DOSE	WT0	SEX
2	1	Theo_1	1	0	0	0	Hours	Theophylline::Dose	0	0	mg	1	1	0	400	1	0	0	400	79.6	1
3	2	Theo_1	1	0.25	0.25	0.25	Hours	Theophylline::Concentration	1	1.5	ug/ml	0	0	1	0	0	0	0	400	79.6	1
4	3	Theo_1	1	0.57	0.57	0.57	Hours	Theophylline::Concentration	1	6.57	ug/ml	0	0	0	0	0	0	0	400	79.6	1
5	4	Theo_1	1	1.12	1.12	1.12	Hours	Theophylline::Concentration	1	10.5	ug/ml	0	0	0	0	0	0	0	400	79.6	1
6	5	Theo_1	1	2.02	2.02	2.02	Hours	Theophylline::Concentration	1	9.66	ug/ml	0	0	0	0	0	0	0	400	79.6	1
7	6	Theo_1	1	3.82	3.82	3.82	Hours	Theophylline::Concentration	1	8.58	ug/ml	0	0	0	0	0	0	0	400	79.6	1
8	7	Theo_1	1	5.1	5.1	5.1	Hours	Theophylline::Concentration	1	8.36	ug/ml	0	0	0	0	0	0	0	400	79.6	1
9	8	Theo_1	1	7.03	7.03	7.03	Hours	Theophylline::Concentration	1	7.47	ug/ml	0	0	0	0	0	0	0	400	79.6	1
10	9	Theo_1	1	9.05	9.05	9.05	Hours	Theophylline::Concentration	1	6.89	ug/ml	0	0	0	0	0	0	0	400	79.6	1
11	10	Theo_1	1	12.12	12.12	12.12	Hours	Theophylline::Concentration	1	5.94	ug/ml	0	0	0	0	0	0	0	400	79.6	1
12	11	Theo_1	1	24.37	24.37	24.37	Hours	Theophylline::Concentration	1	1.5	ug/ml	0	0	1	0	0	0	0	400	79.6	1
13	12	Theo_10	2	0	0	0	Hours	Theophylline::Dose	0	0	mg	1	1	0	400	1	0	0	400	58.2	1

The entries of the CENS columns are defined by:

0: observation \geq LLOQ

1: observation $<$ LLOQ

See also example_2.R to generate NONMEM models using M3 and M4 method

If CENS=1 then entry in DV column is set to LLOQ

If CENS=1 entries are present, then M3 or M4 method is used automatically (selection during NLME project generation)

Other methods of BLOQ handling implemented during NLME dataset generation from general dataset format (M1, M5, M6, M7).

Summary

- **The IQRnlmeData format**

- Follows the typical NONMEM dataset formats
 - Minimal differences
- Requires the presence of default columns
 - Making it slightly more comprehensive than the NONMEM minimum
- Suggests the presence of additional columns that allow better readability
 - Uses alphanumeric content (spaces should be exchanged with ":::")
- Introduces NONMEM non-typic columns (ADM, YTYPE, CENS)
 - Inspired by MONOLIX
 - Allowing to simplify connection to structural model and automatic support of more advanced BLOQ handling methods (M3 & M4)
- Requires presence of USUBJID column
 - Always painful when subjects in PK analysis datasets cannot be mapped to subjects in efficacy or safety datasets (happens far too often)

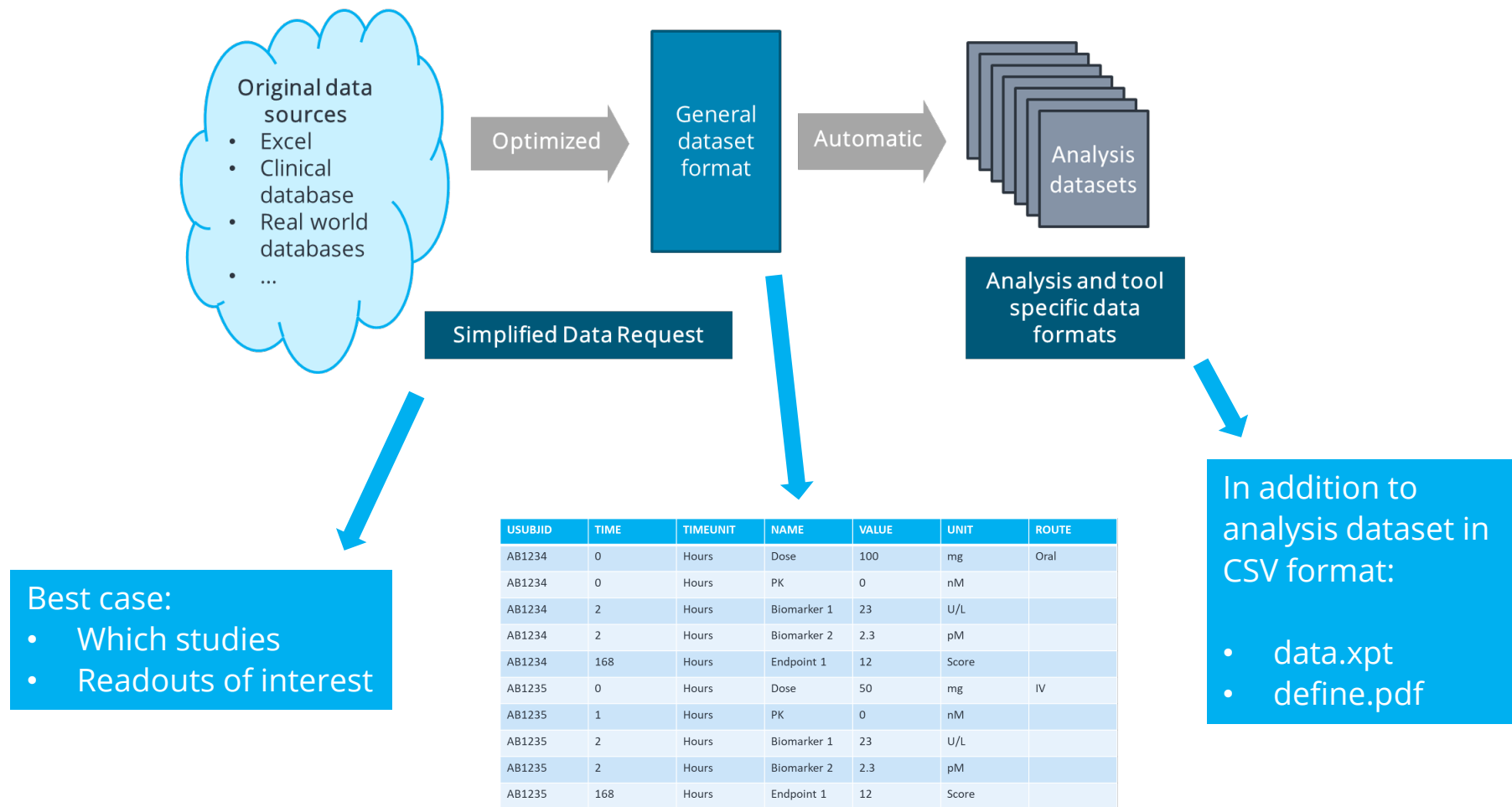
Any legacy NONMEM (or MONOLIX) dataset can be easily brought into this format

Introduction General Dataset Format

Brief introduction

More details in Modul X.1

A “General Dataset Approach” has the potential to considerably simplify data programming



Minimal example dataset in GDF

In contrast to many typical modeling datasets – the “GDF” is self explaining ...

- It is NOT a modeling dataset
- It aims at capturing the clinically relevant information for a considered project
- It is MUCH easier to program from any source than a NONMEM dataset

USUBJID	TIME	TIMEUNIT	NAME	VALUE	UNIT	ROUTE
AB1234	0	Hours	Dose	100	mg	Oral
AB1234	0	Hours	PK	0	nM	
AB1234	2	Hours	Biomarker 1	23	U/L	
AB1234	2	Hours	Biomarker 2	2.3	pM	
AB1234	168	Hours	Endpoint 1	12	Score	
AB1235	0	Hours	Dose	50	mg	IV
AB1235	1	Hours	PK	0	nM	
AB1235	2	Hours	Biomarker 1	23	U/L	
AB1235	2	Hours	Biomarker 2	2.3	pM	
AB1235	168	Hours	Endpoint 1	12	Score	

Minimal set of columns required by IQR Tools in a general dataset

Example 3

Conversion of GDF to analysis dataset

	A	B	C	D	E	F	G	H	I	J	K
1	USUBJID	TIME	TIMEUNIT	DURATION	TYPENAME	NAME	VALUE	VALUETXT	UNIT	ROUTE	LLOQ
2	HS0815-01-001-001	-17.88263889	Days		0 Vital Signs	Height	168.9	.	cm	.	.
3	HS0815-01-001-001	-0.794444444	Days		0 Vital Signs	Weight	51.7	.	kg	.	.
4	HS0815-01-001-001	-0.790972222	Days		0 Lab Values	Albumin	4.5	.	g/dL	.	.
5	HS0815-01-001-001	-0.790972222	Days		0 Lab Values	Creatine Kinase	155	.	U/L	.	.
6	HS0815-01-001-001	-0.006944444	Days		0 PK	HS0815 Concentration	0	.	ug/mL	.	0.2
7	HS0815-01-001-001		0 Days		0 Demographics	Age	34	.	years	.	.
8	HS0815-01-001-001		0 Days		0 Demographics	Gender	0	M	.	.	.
9	HS0815-01-001-001		0 Days		0 Demographics	Race	6	OTHER	.	.	.
10	HS0815-01-001-001		0 Days		0 Demographics	Ethnicity	2	NOT HISPANIC OR LATINO	.	.	.
11	HS0815-01-001-001		0 Days	0.083333333	Dose	HS0815 Dose	0	.	mg	IV	.
12	HS0815-01-001-001	0.084722222	Days		0 PK	HS0815 Concentration	0	.	ug/mL	.	0.2
13	HS0815-01-001-001	0.166666667	Days		0 PK	HS0815 Concentration	0	.	ug/mL	.	0.2
14	HS0815-01-001-001	0.333333333	Days		0 PK	HS0815 Concentration	0	.	ug/mL	.	0.2
15	HS0815-01-001-001	1.006944444	Days		0 PK	HS0815 Concentration	0	.	ug/mL	.	0.2
16	HS0815-01-001-001		2 Days		0 PK	HS0815 Concentration	0	.	ug/mL	.	0.2
17	HS0815-01-001-001	7.147916667	Days		0 PK	HS0815 Concentration	0	.	ug/mL	.	0.2
18	HS0815-01-001-001	14.15069444	Days		0 PK	HS0815 Concentration	0	.	ug/mL	.	0.2
19	HS0815-01-001-001	21.15416667	Days		0 PK	HS0815 Concentration	0	.	ug/mL	.	0.2
20	HS0815-01-001-001	28.17222222	Days		0 PK	HS0815 Concentration	0	.	ug/mL	.	0.2
21	HS0815-01-001-001	42.14375	Days		0 PK	HS0815 Concentration	0	.	ug/mL	.	0.2

example_3_data.csv (dummy data)

- Additional columns can be present to better inform the data
- Some examples shown in red ...

Example 3

Conversion of GDF to NLME analysis dataset (NONMEM, MONOLIX, NLMIXR)

```
# Import as GDF
dataGen <- IQRdataGENERAL(
  input      = "example_3_data.csv",

# Define "NAMEs" of dosing records
doseNAMES = "HS0815 Dose",

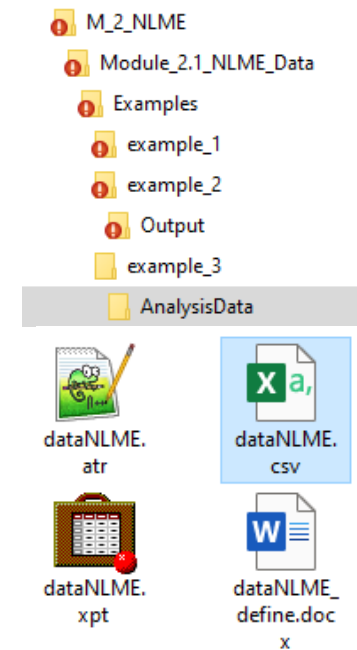
# Define "NAMEs" of observations considered in modeling
obsNAMES  = "HS0815 Concentration",

# Define "observations" to consider as continuous baseline covariates
cov0 = list(
  AGE0 = "Age",
  HT0  = "Height",
  WT0  = "Weight",
  ALB0 = "Albumin",
  CK0  = "Creatine Kinase"
),

# Define "observations" to consider as categorical baseline covariates
cat0 = list(
  SEXF = "Gender",
  RACE = "Race"
),

# Define BLOQ handling method (M1,3,4,5,6,7)
methodBLLOQ = "M1",
)

# Export NLME data (CSV, XPT, define)
exportNLME_IQRdataGENERAL(data = dataClean,
  filename = "AnalysisData/dataNLME.csv",
  FLAGxpt = TRUE,
  FLAGdefine = TRUE)
```



Example 3

Some simple data based analysis

```
# Simple summary  
summary(dataGen)
```

example_3.R

```
> summary(dataGen)
```

INFO	NAME	VALUE
Dose events	HS0815 Dose	Ntotal: 40, Nindiv (min/median/max): 1/1/3)
Observation events (all)	HS0815 Concentration	Ntotal: 519, Nindiv (min/median/max): 7/14/26)
Observation events (MDV=0)	HS0815 Concentration	Ntotal: 306, Nindiv (min/median/max): 1/13/13)
Doses AMT=0 present	ALL dose events	TRUE (N=16)
Placebo subjects present (AMT=0 or no doses)	ALL dose events	TRUE (N=12)
IGNORED (MDV=1) observation records present	ALL observation events	TRUE (N=213)
Subjects without observations (MDV=0) present	ALL observation events	TRUE (N=11)
Total BLOQ information	HS0815 Concentration	N=212 / 40.8%
BLOQ handling method	All observation events	M1
NLME columns containing NA	All events	DV
Issues present in the data	Minor	NONE
Issues present in the data	Warnings	YES (see text below the table for more information)
Issues present in the data	Errors	NONE

WARNINGS in the dataset that should be addressed
=====

GLOBAL LEVEL: Records are present that have neither VALUE nor VALUETXT defined

- Simple summary
- Data checks and informative messages possible – due to standards
- Here: many BLOQ observations present since placebo subjects included in data
- General Dataset functionality in IQR Tools supports easy handling of such cases and automatic generation of logfiles for reporting of “data changes” => more in Module X1

Example 3

Some simple data based analysis

Simple summary

```
summaryCov_IQRdataGENERAL(dataGen,FLAGtotal = TRUE)
```

```
summaryCat_IQRdataGENERAL(dataGen,FLAGtotal = TRUE)
```

example_3.R

```
> summaryCov_IQRdataGENERAL(dataGen,FLAGtotal = TRUE)
Summary of demographic and baseline characteristics for continuous information
=====
```

Characteristic	XYZ [N=36]	TOTAL [N=36]
Age (years)	33 (10.9) [18-53]	33 (10.9) [18-53]
Height (cm)	171 (6.31) [157-182]	171 (6.31) [157-182]
Weight (kg)	65.9 (9.3) [46.8-84.4]	65.9 (9.3) [46.8-84.4]
Albumin (g/dL)	4.47 (0.256) [3.8-5]	4.47 (0.256) [3.8-5]
Creatine Kinase (U/L)	176 (96.7) [48-444]	176 (96.7) [48-444]

N: Number of subjects

Entries represent: Mean (Standard deviation) [Minimum-Maximum]

```
> summaryCat_IQRdataGENERAL(dataGen,FLAGtotal = TRUE)
Summary of demographic and baseline characteristics for categorical information
=====
```

Characteristic	Category	XYZ [N=36]	TOTAL [N=36]
Gender	M	23 (63.9%)	23 (63.9%)
	F	13 (36.1%)	13 (36.1%)
Race	WHITE	9 (25%)	9 (25%)
	OTHER	27 (75%)	27 (75%)

N: Number of subjects

Number of subjects in each category and percentage within this category

Table 1 Summary of demographic and baseline characteristics for continuous information

Characteristic	XYZ [N=36]	TOTAL [N=36]
Age (years)	33 (10.9) [18-53]	33 (10.9) [18-53]
Height (cm)	171 (6.31) [157-182]	171 (6.31) [157-182]
Weight (kg)	65.9 (9.3) [46.8-84.4]	65.9 (9.3) [46.8-84.4]
Albumin (g/dL)	4.47 (0.256) [3.8-5]	4.47 (0.256) [3.8-5]
Creatine Kinase (U/L)	176 (96.7) [48-444]	176 (96.7) [48-444]

N: Number of subjects

Entries represent: Mean (Standard deviation) [Minimum-Maximum]

Table 2 Summary of demographic and baseline characteristics for categorical information

Characteristic	Category	XYZ [N=36]	TOTAL [N=36]
Gender	M	23 (63.9%)	23 (63.9%)
	F	13 (36.1%)	13 (36.1%)
Race	WHITE	9 (25%)	9 (25%)
	OTHER	27 (75%)	27 (75%)

N: Number of subjects

Number of subjects in each category and percentage within this category

- Easy generation of demographic tables for continuous and categorical information
- Provide filename => report in Word

Example 3

Some simple data based analysis

```
# Some generic plots that every analyst would want to see ...  
plot(dataGen,pathname = "AnalysisData/plots")
```

example_3.R

Module_2.1_NLME_Data

Examples

example_1

example_2

Output

example_3

AnalysisData

plots

tables

Name

01_OBSvsTIME_Spaghetti.pdf

02_OBSvsTIME_Individual.pdf

04_ContCovDistributions.pdf

05_ContCovCorrelation.pdf

06_CatCovCorrelation.pdf

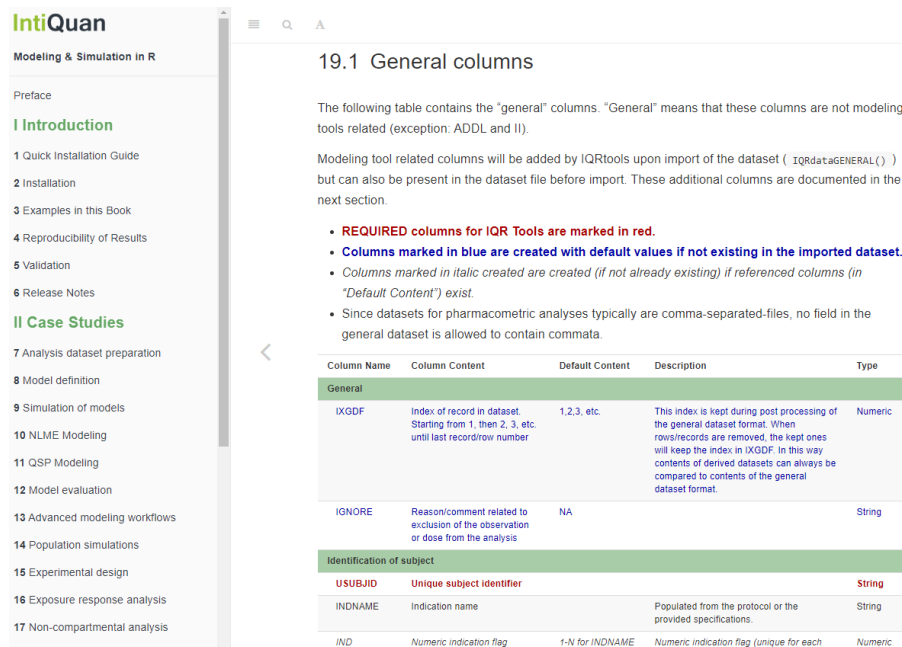
07_ContCatCovCorrelation.pdf

Summary

- General Dataset Format
 - Representation of (pre-)clinically relevant information
 - Humanly readable
 - Easy to program from source data
 - Able to capture all types of data typically informing QSP, Sysbio, and Pharmacometric modeling efforts
 - Automatic conversion to NLME / QSP modeling datasets
 - Automatic generation of XPT and define file for submission
 - Handling of desired BLOQ handling methods
 - Generation of
 - Summary statistics
 - Generic data exploratory graphics
- Considerably speeding up data programming tasks

Legacy NONMEM (MONOLIX) datasets can easily be converted into the GDF

Generalized dataset format - Documentation



The screenshot shows the IntiQuan website with a sidebar on the left containing a table of contents. The main content area is titled '19.1 General columns'. It explains that the following table contains 'general' columns, which are not modeling tools related (exception: ADDL and II). It also notes that modeling tool related columns will be added by IQRtools upon import of the dataset (IQRdataGENERAL()) but can also be present in the dataset file before import. These additional columns are documented in the next section.

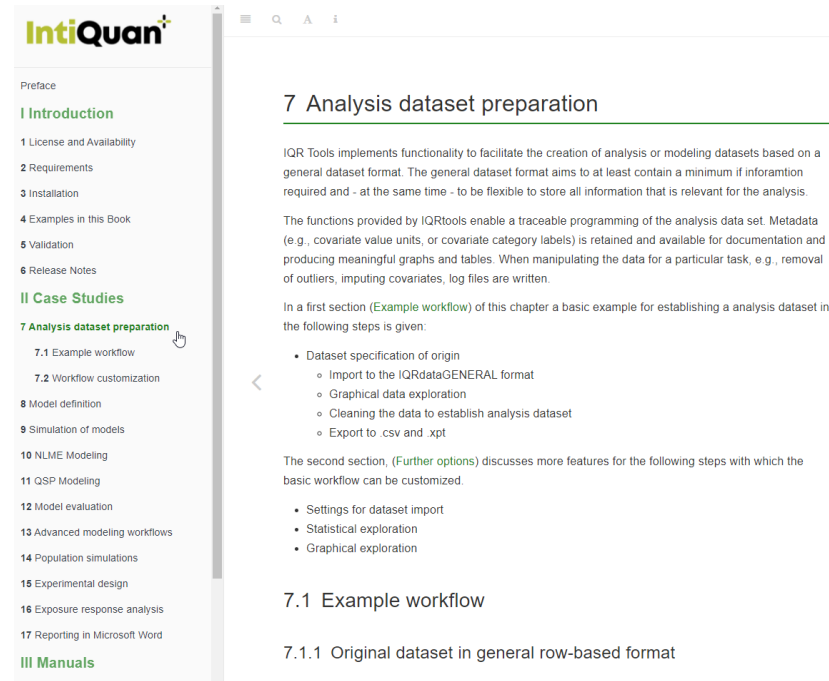
- **REQUIRED columns for IQR Tools are marked in red.**
- **Columns marked in blue are created with default values if not existing in the imported dataset.**
- *Columns marked in italic created are created (if not already existing) if referenced columns (in "Default Content") exist.*
- Since datasets for pharmacometric analyses typically are comma-separated-files, no field in the general dataset is allowed to contain commas.

Column Name	Column Content	Default Content	Description	Type
General				
IXGDF	Index of record in dataset. Starting from 1, then 2, 3, etc. until last record/row number	1,2,3, etc.	This index is kept during post processing of the general dataset format. When rows/records are removed, the kept ones will keep the index in IXGDF. In this way contents of derived datasets can always be compared to contents of the general dataset format.	Numeric
IGNORE	Reason/comment related to exclusion of the observation or dose from the analysis	NA		String
Identification of subject				
USUBJID	Unique subject identifier			String
INDNAME	Indication name		Populated from the protocol or the provided specifications.	String
IND	Numeric indication flag	1-N for INDNAME	Numeric indication flag (unique for each INDNAME)	Numeric

<https://iqrtools.intiquan.com/doc/book/GenDatFormat.html>

Documentation of GDF handling and postprocessing

Full documentation of GDF



The screenshot shows the IntiQuan website with a sidebar on the left containing a table of contents. The main content area is titled '7 Analysis dataset preparation'. It explains that IQR Tools implements functionality to facilitate the creation of analysis or modeling datasets based on a general dataset format. The general dataset format aims to at least contain a minimum of information required and - at the same time - to be flexible to store all information that is relevant for the analysis.

The functions provided by IQRtools enable a traceable programming of the analysis data set. Metadata (e.g., covariate value units, or covariate category labels) is retained and available for documentation and producing meaningful graphs and tables. When manipulating the data for a particular task, e.g., removal of outliers, imputing covariates, log files are written.

In a first section (Example workflow) of this chapter a basic example for establishing a analysis dataset in the following steps is given:

- Dataset specification of origin
 - Import to the IQRdataGENERAL format
 - Graphical data exploration
 - Cleaning the data to establish analysis dataset
 - Export to .csv and .xpt

The second section, (Further options) discusses more features for the following steps with which the basic workflow can be customized.

- Settings for dataset import
- Statistical exploration
- Graphical exploration

7.1 Example workflow

7.1.1 Original dataset in general row-based format

<https://iqrtools.intiquan.com/doc/book/analysis-dataset-preparation.html>

Conclusions

Conclusions

The analysis data format used in IQR Tools

- Follows the typical NONMEM dataset format
- Ensures a minimum amount of documentation within the data itself
- Is the same across all supported parameter estimation tools
 - NONMEM, MONOLIX, NLMIXR

The general dataset format used on IQR Tools

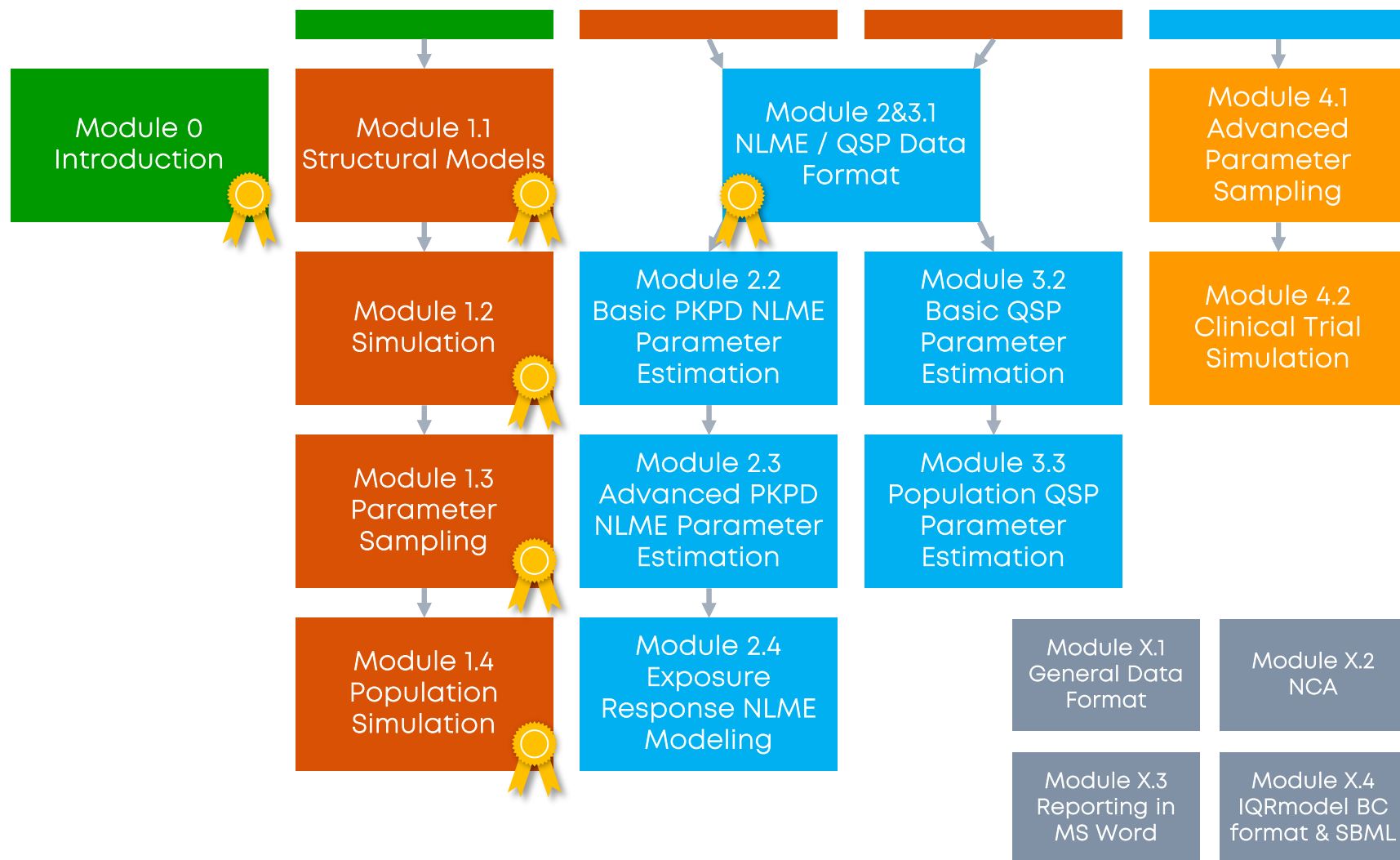
- Supports simplified programming of datasets from any source
- Reduces programming complexity and number of iterations
- Allows export of analysis ready NLME datasets
- Supports generation of standard statistics and data exploratory plots that are needed in any project

**Legacy NONMEM (MONOLIX) datasets
can easily be converted into the formats
used within IQR Tools**

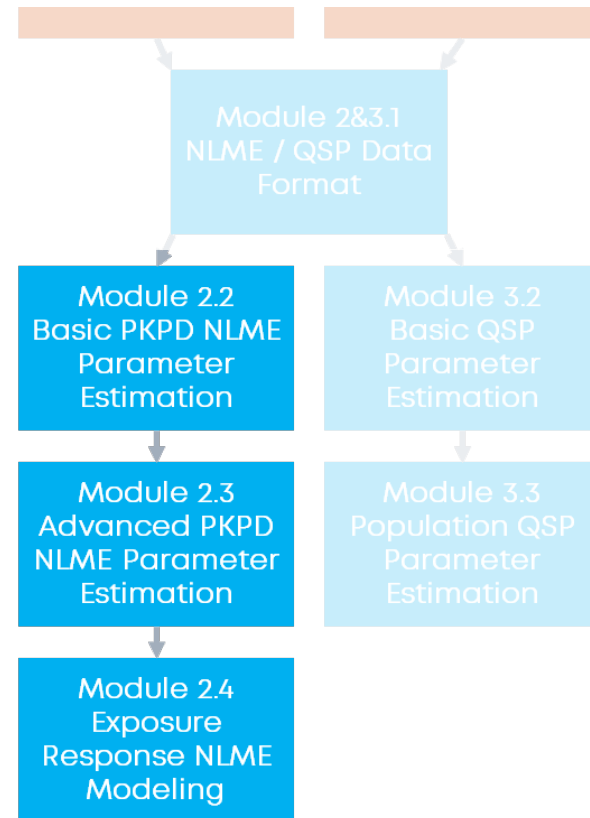
Outlook

Overview of Webinar Modules

IntiQuan Webinar Series on efficient support of Model Informed Drug Development (MIDD)



Outlook on tool requirements



Subsequent modules (2.2, 2.3, 2.4) will require presence of NONMEM and/or MONOLIX on the system to run examples

Convenient setup of IQR Tools options, including interfacing with NONMEM and MONOLIX on the system through the IQR Tools options setup:

```
library(IQRtools)
setup_IQRtools()
```

```
49 # UNIX setup
50 # -----
51 # Name of the NONMEM executable or shell script. Required calling syntax:
52 # "command controlfile outputfile"
53 # NONMEM Version 7.2/7.3/7.4 have been tested with IQR Tools.
54 .PATH_SYSTEM_NONMEM <- list(
55   # First entry is used as default version
56   NM743 = 'nmfe74'
57 )
58
59 # MacOS setup
60 # -----
61 # Name of the NONMEM executable or shell script. Required calling syntax:
62 # "command controlfile outputfile"
63 # NONMEM Version 7.2/7.3/7.4 have been tested with IQR Tools.
64 .PATH_SYSTEM_NONMEM <- list(
65   NM74 = "nmfe74"
66 )
67
68 # WINDOWS setup
69 # -----
70 # Path to NONMEM (Version 7.2/7.3/7.4 have been tested) batch files. The info
71 # is provided as a list. By default the first entry is used but the user can
72 # switch when calling the IQRnlmeProject function.
73 .PATH_SYSTEM_NONMEM <- list(
74   NM74 = "C:/nm74g64/run/nmfe74.bat"
75 )
```



The program for Nonlinear
Mixed Effects Modeling



Parallel NONMEM runs handled by IQR Tools as well – can be set up in options as well.
Requires a shell/batch script available on systems command line with the following calling syntax:

```
command NRCORES controlfile outputfile"
```

IQdesktop

- Helps getting the complete environment set up in ~0 min
- Is freely available
- Requires admin rights on your computer (to install docker)
- *Bring your own NONMEM and/or MONOLIX licenses ...*

- More information:
<https://iqdesktop.intiquan.com>

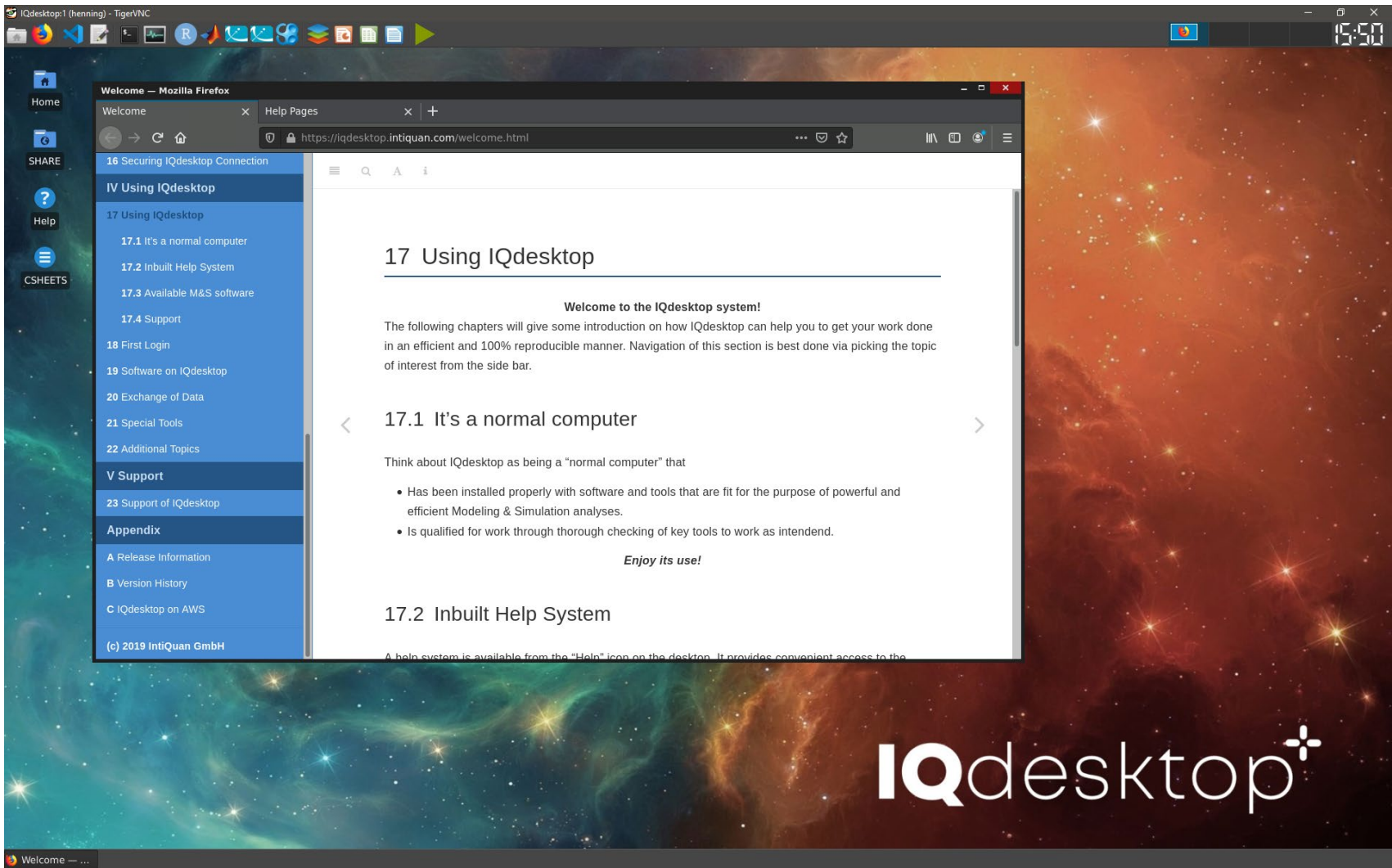
■ Installation guide

- [Windows](#)
- [macOS](#)
- [Linux](#)
- [Cloud](#)

https://iqdesktop.intiquan.com/doc/IQDesktop_Installation_Windows.mp4

A.1 Release V2.1.1 - 30 April 2021

- **Image name:** `intiquan/iqdesktop:2.1.1`
- **Base system:** Ubuntu 18.04
- **Connections:** VNC, SSH (incl. screen)
- **Modeling & Simulation & specific Development Tools**
 - IQR Tools (Version 1.7.0)
 - IQnca (Version 1.0.0)
 - IQReport (Version 1.51)
 - IQSlides (Version 0.3.2)
 - R default (Version 3.6.3)
 - CRAN snapshot dated 2020-03-15
 - URL: <https://cran.intiquan.com/snapshot/2020-03-15>
 - R exploratory (Version 4.0.3)
 - CRAN snapshot dated 2020-12-21
 - URL: <https://cran.intiquan.com/snapshot/2020-12-21>
 - IQR Tools, IQnca, and IQSlides not installed on R 4.0.3
 - MATLAB (Version R2021A) + Simbiology and relevant toolboxes
 - NONMEM (Version 7.4.3)
 - NONMEM (Version 7.5.0)
 - MONOLIX (Version 2019R1)
 - MONOLIX (Version 2020R1)
 - CellDesigner (Version 4.4.2)
 - PsN (Version 5.0.0) (no gls and qa)
 - RHEM (Version 0.1.2)
 - Rstudio (Version 1.2.5042)
 - Jenkins 2.249.2 CI/CD server (including suggested plugins)
 - Population Isoboles R package (1.0.0)
- **Changes**
 - Update to MATLAB R2021A
 - Update to IQSlides 0.3.2



<https://iqdesktop.intiquan.com>

https://iqdesktop.intiquan.com/doc/IQDesktop_Installation_Windows.mp4

- Standard issues with reproducibility:
 - NONMEM: dependent on hardware platform & compiler, etc.
 - R: Highly dependent on package version (constant change on CRAN)
 - Manual conduct vs. fully scripted analyses
- Computer systems evolve, new versions are installed, old are dropped
- To ensure 100% reproducibility you could store computers in a warehouse – one per analysis conducted, including scripts & data
- IQdesktop ensures 100% reproducibility through a “virtual warehouse” approach



Q&A session

Thank
You

Contact information



IntiQuan GmbH
Elisabethenstrasse 23
4051 Basel
Switzerland

Phone: +41 76 603 28 06
E-mail: info@intiquan.com
Web: www.intiquan.com