Navigating Data Package Submission in the Chinese Regulatory Landscape

Rebecca Moretti, Lead Statistical Programmer, Chiesi Farmaceutici S.p.A
Emanuele Rocco Calabrò, Lead Data Manager, Chiesi Farmaceutici S.p.A

Meet the Speakers

Rebecca Moretti

Title: Lead Statistical Programmer

Organization: Chiesi Farmaceutici S.p.A

Rebecca graduated in Biostatistics Master of Science at University of Milano-Bicocca and is currently a Lead Statistical Programmer with more than 8 years of experience in the pharmaceutical sector. Before joining Chiesi, she worked in different CROs both as a Statistical Programmer and as a Real-World Evidence Analyst. She is currently Lead Statistical Programmer and SDTM Subject Matter Expert in Chiesi.

Emanuele Calabrò

Title: Lead Data Manager

Organization: Chiesi Farmaceutici S.p.A

Emanuele has a Master of Science in Biostatistics, and he is a data management expert with over 13 years of experience in the pharmaceutical sector. He is a member of the CDISC Controlled Terminology Protocol Deviation sub-team, a member of Chiesi Protocol Deviation Organization Working Group, and a referent point in Chiesi on submission in Chinese Medicines Approval Authority (NMPA).

Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the authors and do not necessarily reflect the official policy or position of CDISC.*

- *The authors have no real or apparent conflicts of interest to report.*

Agenda

1. Setting the scene
2. Regulatory package
3. Clinical data package
4. Walls / Wins / Wisdoms

Setting the scene

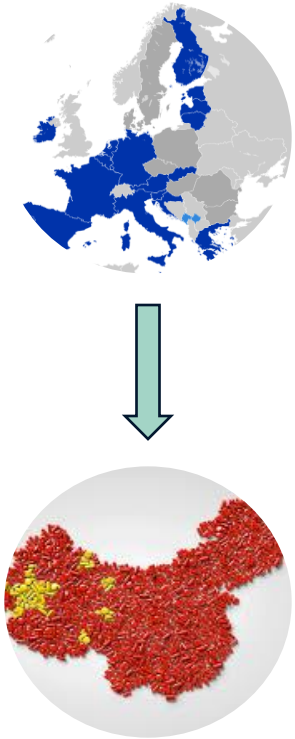
Overview on our Drug Development process

The drug was developed and approved in Europe and other countries (like Australia)

Chiesi expansion into the Chinese market

Two studies were designed on a Chinese population: PK study (on HV) and pivotal study

Application of the Chinese Clinical Technical Requirements

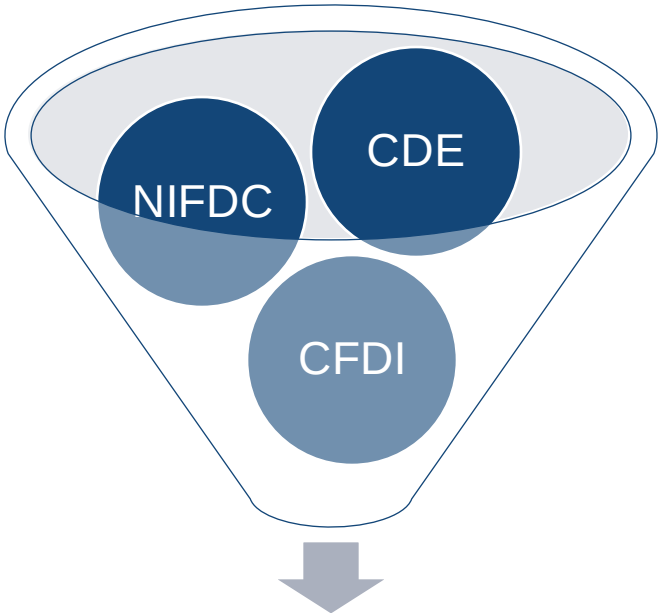


PK: pharmacokinetics
HV: healthy volunteers

Source:
Drugs marketed overseas but not marketed in China – Clinical technical requirements

- *https://www.cde.org.cn/zdyz/opinioninfopage?zdyzIdCODE=4832fe1bef75686610c58cc092e0f911&rddt=1*
- *http://english.nmpa.gov.cn/2020-11/18/c_568155.htm*

Chinese national authorities involved in the NDA review

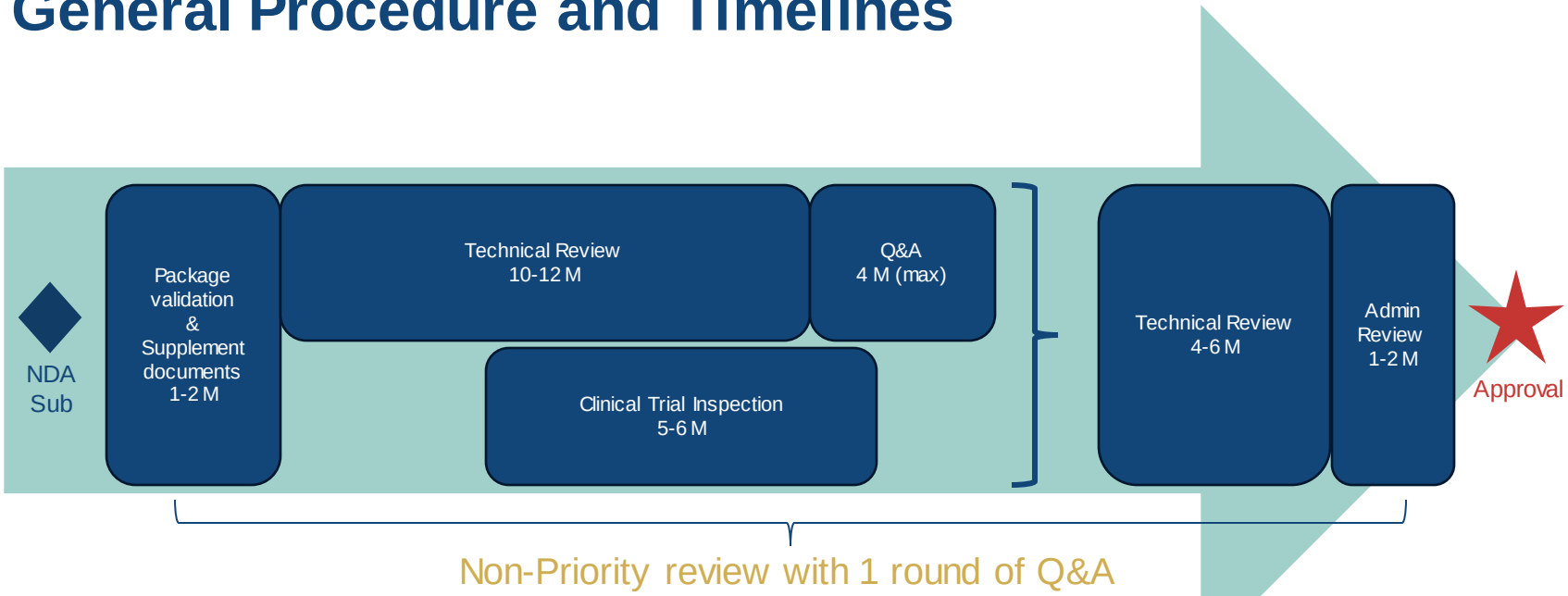


Role	Official name	Responsibility
Main authority	National Medical Products Administration (NMPA)	Administrative approval for NDA application
Affiliated institutions	Center for Drug Evaluation (CDE)	Technical review
	National Institutes for Food and Drug Control (NIFDC)	Quality test conduction and review of quality tests
	Center for Food and Drug Inspection (CFDI)	Inspection of drug clinical trial, non-clinical research institutions and sponsors

NATIONAL MEDICAL PRODUCTS ADMINISTRATION
国家药品监督管理局

https://english.nmpa.gov.cn/

NDA submission
General Procedure and Timelines

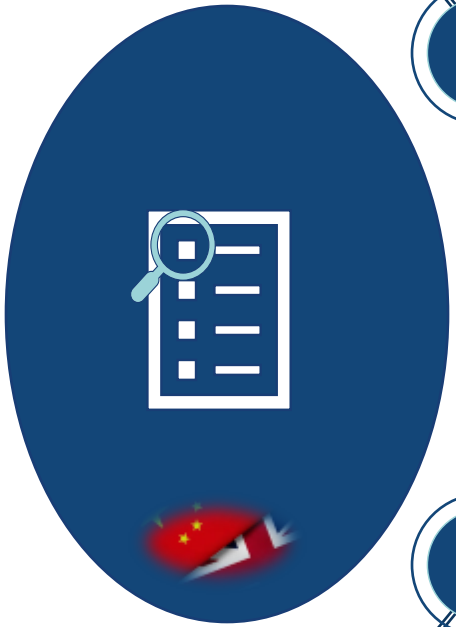


Notes:
Q&A Timeframe highly depends on the duration of the preparation of response document;
8M extension if a 2nd round of Q&A is required

This time may vary due to the time frame of Q&A and inspection.

Regulatory package

Regulatory package



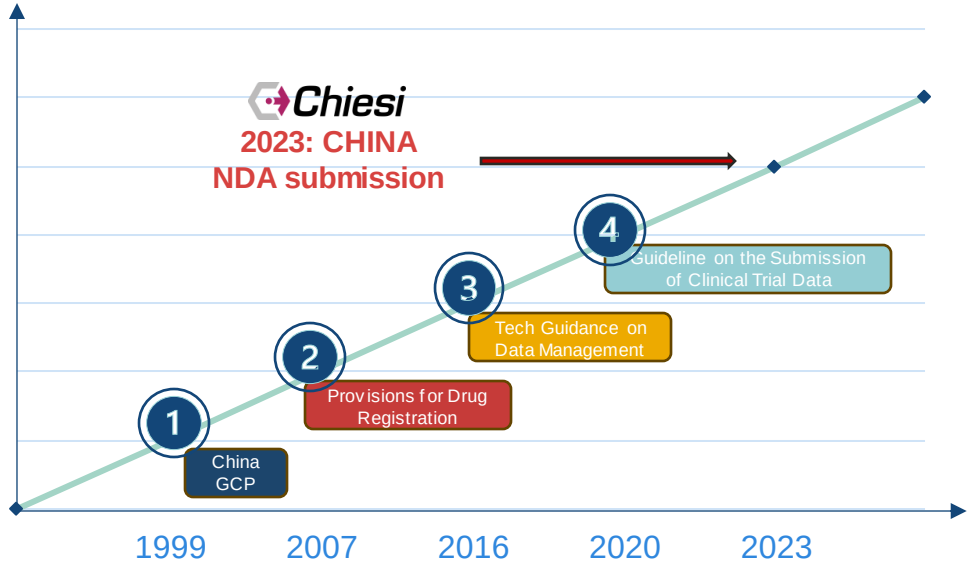
1. eCTD Clinical Modules on Europeans population

eCTD: electronic Common Technical Document

2. Clinical Study Reports of Chinese studies
3. Additional documents on clinical development in China
4. Clinical Data Package

Regulatory package: Clinical Data

Clinical Data Submission in China



Source:
https://data42.cn/c3c/webinar/20200527/C3C_NMPA.mp4

Guideline on the Submission of Clinical Trial Data (Jul 2020)

Version:	• July 2020
Field of application	• This guideline is **primarily applicable to pivotal clinical trials** for the purpose of drug registration and marketing, as well as to clinical trials for non-registration purposes
Submission components of Clinical Trial Data:	• Study database • Analysis database • Data definition file • Data reviewer's guide • Annotated CRF • Programming code

Table of Contents

Guideline on the Submission of Clinical Trial Data (Jul 2020)

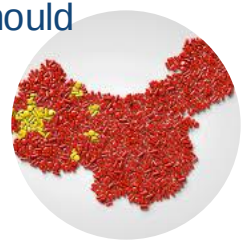
Aligns technical requirements for clinical data package supporting submission with international CDISC standards

Details on:

- Components of the data package
- Format and conventions
- Additional details

Which studies should be submitted as part of the data-package?

- Studies conducted in China should be submitted



- At the moment PK studies in Chinese HV are <u>not</u> required
- Data package related to studies conducted abroad (i.e., EU only) are <u>not</u> required

PK: pharmacokinetics
HV: healthy volunteers

Clinical data package

Clinical data package: required elements

DATABASES: SDTM and ADaM

XPT V5 or above* DM and ADSL are mandatory

ANNOTATED CRF

aCRF.pdf*

PROGRAMMING CODE

Readable and understandable
Do not include external program calls; Avoid to use nested macros.
.txt as the file extension

DATA DEFINITION FILES

.xml* or .pdf*
.pdf is not required when .xml is used for submission

DATA REVIEWER'S GUIDE

.pdf*
Recommended but not mandatory

**Chinese translation requirement for foreign language database*

XML:Extensible Mark-up Language
PDF: portable document format

Clinical data package: annotated CRF

NMPA requirements

- Pdf format

Chinese translation:

- questions designed to collect data;
- values or codes list of efficacy indicators.

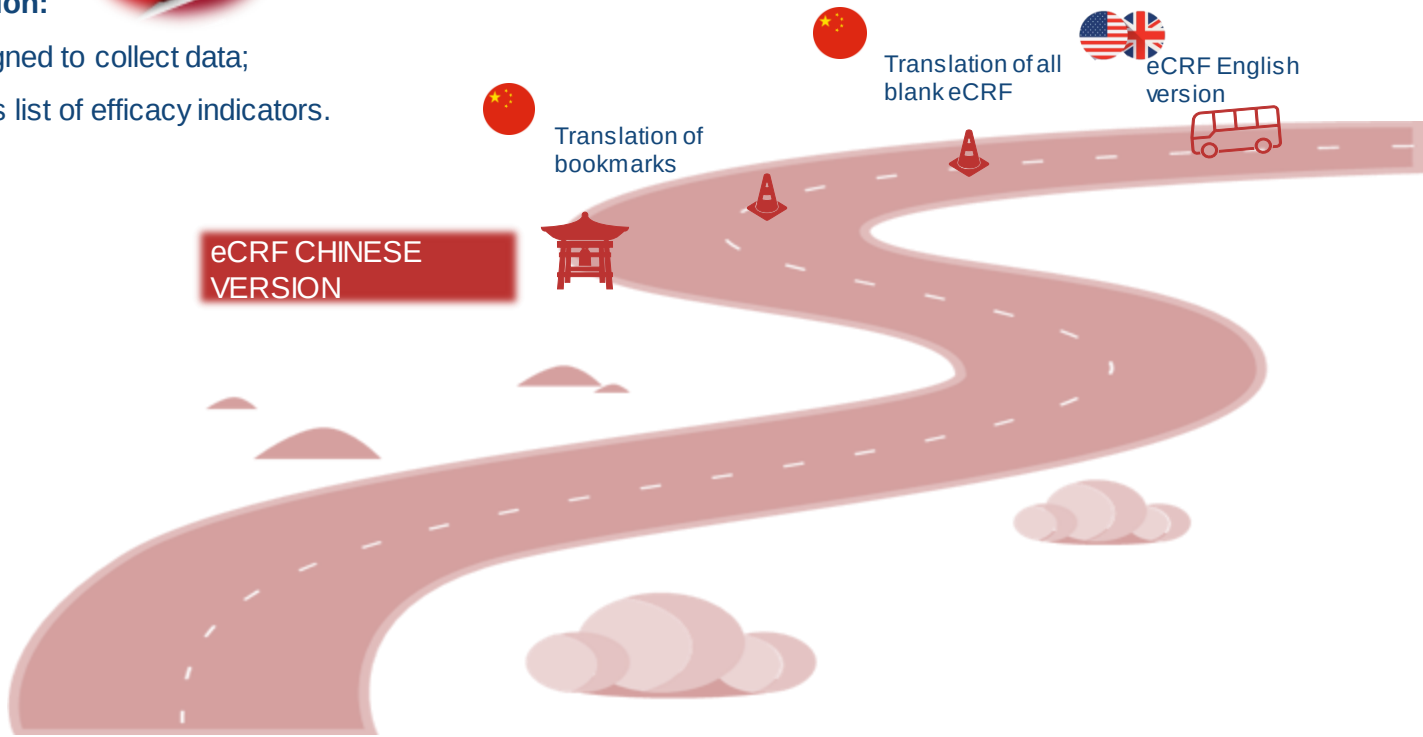
Clinical data package: annotated CRF

NMPA requirements

- Pdf format

Chinese translation:

- questions designed to collect data;
- values or codes list of efficacy indicators.



Clinical data package: annotated CRF

NMPA requirements

- Pdf format

Chinese translation:

- questions designed to collect data;
- values or codes list of efficacy indicators.

Translation of External data Data Transfer Specifications (DTS) because of external data not integrated in EDC

Translation of bookmarks

eCRF CHINESE VERSION

Translate dataset description from annotation (translation conformed to Chinese version of SDTM IG)

Translation of all blank eCRF

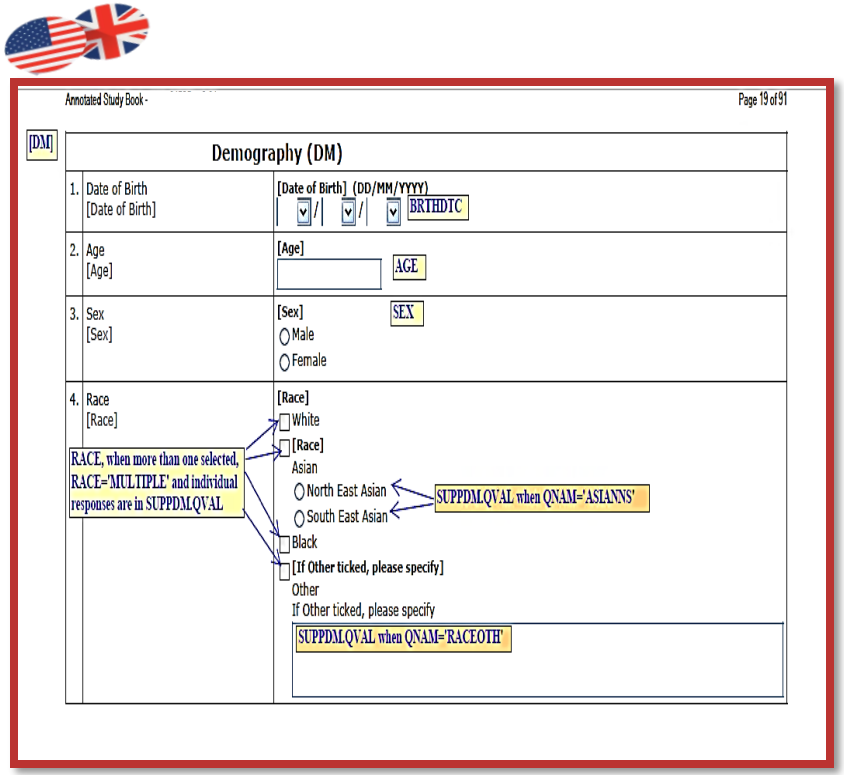
eCRF English version

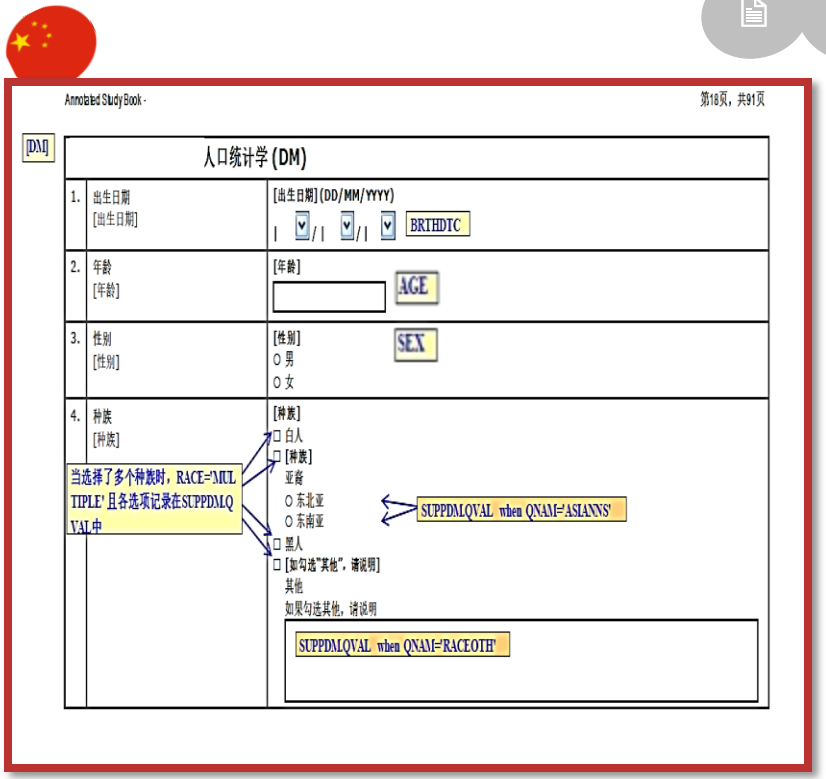
aCRF English version

aCRF CHINESE VERSION

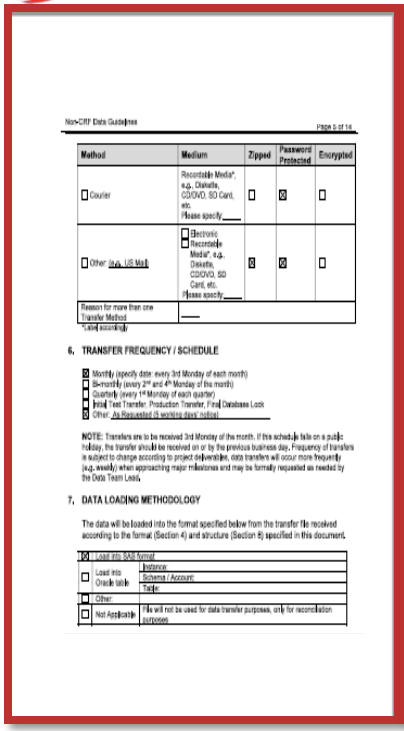
The CRF text should be identical with the Chinese eCRF translated

Clinical data package: annotated CRF





Clinical data package: annotated CRF



Non-CRF Data Guidelines — Page 5 of 14

Method	Medium	Zipped	Password Protected	Encrypted
☐ Courier	Recordable Media*, e.g., Diskette, CD/DVD, SD Card, etc. Please specify:___	☐	☒	☐
☐ Other: (e.g., US Mail)	☐ Electronic ☐ Recordable Media*, e.g., Diskette, CD/DVD, SD Card, etc. Please specify:___	☒	☒	☐
Reason for more than one Transfer Method:	___			

*Label accordingly

6. TRANSFER FREQUENCY / SCHEDULE

☒ Monthly (specify date: every 3rd Monday of each month)
☐ Bi-monthly (every 2nd and 4th Monday of the month)
☐ Quarterly (every 1st Monday of each quarter)
☐ Initial Test Transfer, Production Transfer, Final Database Lock
☒ Other: As Requested (5 working days' notice)

NOTE: Transfers are to be received 3rd Monday of the month. If this schedule falls on a public holiday, the transfer should be received on or by the previous business day. Frequency of transfers is subject to change according to project deliverables; data transfers will occur more frequently (e.g., weekly) when approaching major milestones and may be formally requested as needed by the Data Team Lead.

7. DATA LOADING METHODOLOGY

The data will be loaded into the format specified below from the transfer file received according to the format (Section 4) and structure (Section 8) specified in this document.

☒	Load into SAS format	
☐	Load into Oracle table	Instance: Schema / Account: Table:
☐	Other:	
☐	Not Applicable	File will not be used for data transfer purposes, only for reconciliation purposes

[QS] ECCAT = "STUDY MEDICATION"
[EC] ECSCAT = "Recorded in the patient diary"
ECROUTE="RESPIRATORY (INHALATION)

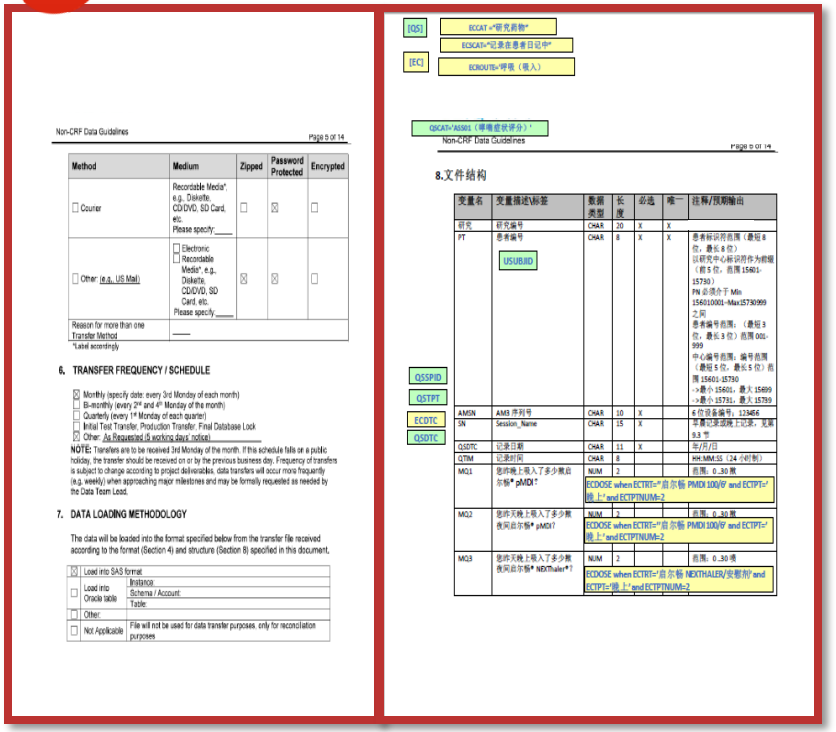
QSCAT='ASS01(Asthma Symptom Score)'

Non-CRF Data Guidelines — Page 6 of 14

8. FILE STRUCTURE

Variable Name	Variable Description / Label	Data Type	Size	Mandatory	Unique	Comment / expected output
STUDY	Study identifier	CHAR	20	X	X	
PT	Patient identifier USUBJID	CHAR	8	X	X	Patient identifier range (Min 8, Max 8) Site identifier as the prefix (first 5 digits, range 15001-15730) PN must between Min 15001001 - Max15730999 Patient number range: (Min 3, Max 3) range 001 –999 Center number range: number range (Min 5, Max 5) range 15601- 15730 •-> Min 15601, Max 15699 •-> Min 15731, Max 15736
AMSN	AM3 Serial Number	CHAR	10	X		6 digit device number:123456
SN	Session_Name	CHAR	15	X		Morning Session or Evening Session, see section 8.3
QSDTC	Session Date	CHAR	11	X		YYYY/MM/DD
QTM	Session Time	CHAR	8			HH:MM:SS (24h format)
MQ1	How many puffs from the Foster® pMDI did you take yesterday evening?	NUM	2			Range: 0 - 30 puff(s) ECDOSE when ECTRT='FOSTER PMDI 100/6' and ECTPT='EVENING' and ECTPTNUM=2
MQ2	How many puffs from the Moon Foster® pMDI did you take yesterday evening?	NUM	2			Range: 0 - 30 puff(s) ECDOSE when ECTRT='FOSTER PMDI 100/6' and ECTPT='EVENING' and ECTPTNUM=2
MQ3	How many inhalations from the Moon Foster® NEXThaler® did you take yesterday evening?	NUM	2			Range: 0 - 30 inhalation(s) ECDOSE when ECTRT='FOSTER NEXTHALER/ PLACEBO' and ECTPT='EVENING' and ECTPTNUM=2

QSSPID, QSTPT, ECDTC, QSDTC



Non-CRF Data Guidelines — Page 5 of 14

Method	Medium	Zipped	Password Protected	Encrypted
☐ Courier	Recordable Media*, e.g., Diskette, CD/DVD, SD Card, etc. Please specify:___	☐	☒	☐
☐ Other: (e.g., US Mail)	☐ Electronic ☐ Recordable Media*, e.g., Diskette, CD/DVD, SD Card, etc. Please specify:___	☒	☒	☐
Reason for more than one Transfer Method:	___			

*Label accordingly

6. TRANSFER FREQUENCY / SCHEDULE

☒ Monthly (specify date: every 3rd Monday of each month)
☐ Bi-monthly (every 2nd and 4th Monday of the month)
☐ Quarterly (every 1st Monday of each quarter)
☐ Initial Test Transfer, Production Transfer, Final Database Lock
☒ Other: As Requested (5 working days' notice)

NOTE: Transfers are to be received 3rd Monday of the month. If this schedule falls on a public holiday, the transfer should be received on or by the previous business day. Frequency of transfers is subject to change according to project deliverables; data transfers will occur more frequently (e.g., weekly) when approaching major milestones and may be formally requested as needed by the Data Team Lead.

7. DATA LOADING METHODOLOGY

The data will be loaded into the format specified below from the transfer file received according to the format (Section 4) and structure (Section 8) specified in this document.

☒	Load into SAS format	
☐	Load into Oracle table	Instance: Schema / Account: Table:
☐	Other:	
☐	Not Applicable	File will not be used for data transfer purposes, only for reconciliation purposes

[QS] ECCAT ="研究药物"
ECSCAT="记录在患者日记中"
[EC] ECROUTE="呼吸（吸入）

QSCAT='ASS01（哮喘症状评分）'

Non-CRF Data Guidelines — Page 5 of 14

8.文件结构

变量名	变量描述/标签	数据类型	长度	必选	唯一	注释/预期输出
研究	研究编号	CHAR	20	X	X	
PT	患者编号 USUBJID	CHAR	8	X	X	患者标识符范围（最短8位，最长8位） 站研究中心标识符作为前缀（前5位，范围15601-15730） PN必须介于Min 15601001-Max15730999之间 患者编号范围：（最短3位，最长3位）范围001-999 中心编号范围，编号范围（最短5位，最长5位）范围15601-15730 ->最小15601，最大15699 ->最小15731，最大15739
AMSN	AM3 序列号	CHAR	10	X		6位设备编号：123456
SN	Session_Name	CHAR	15	X		早晨记录或晚上记录，见第8.3节
QSDTC	记录日期	CHAR	11	X		年/月/日
QTM	记录时间	CHAR	8			HH:MM:SS（24小时制）
MQ1	您昨晚上吸入了多少揿启尔畅® pMDI？	NUM	2			范围：0..30揿 ECDOSE when ECTRT='启尔畅 PMDI 100/6' and ECTPT='晚上' and ECTPTNUM=2
MQ2	您昨天晚上吸入了多少揿优风启尔畅® pMDI？	NUM	2			范围：0..30揿 ECDOSE when ECTRT='启尔畅 PMDI 100/6' and ECTPT='晚上' and ECTPTNUM=2
MQ3	您昨天晚上吸入了多少揿优风启尔畅® NEXThaler®？	NUM	2			范围：0..30揿 ECDOSE when ECTRT='启尔畅 NEXTHALER/安慰剂' and ECTPT='晚上' and ECTPTNUM=2

QSSPID, QSTPT, ECDTC, QSDTC

Clinical data package: DATABASES 1/6

NMPA requirements

Data from CRF and external sources

Sponsor is encouraged to submit SDTM according to CDISC

Chinese translation

Traceability

Analysis ready

Analysis metadata

Sponsor is encouraged to submit ADaM

Chinese translation

- Dataset and variable labels
- Adverse events terms
- Generic name of concomitant medications
- Medical history in CSR and other documents

Clinical data package: DATABASES 2/6

Chiesi strategy

DATASET AND VARIABLE LABELS

- Dataset and variables labels be translated in an excel file
- Dataset and variables labels conformed to Chinese version of SDTM and ADAM IGs
- SAS programs to update the datasets with the Chinese text. Double-programming to ensure correctness

VERBATIM

- All unique verbatim terms in an excel file (--TERM, --TRT) translated in Chinese in a different column than the verbatim terms in English
- Excel file imported using SAS to update --TERM and --TRT variables with the translated term. Independent double-programming applied to ensure correctness.

CODING TERMS

- For datasets coded using MedDRA dictionary, --LLT, --DECOD, --HLT, --HLGT, --BODSYS and –SOC variables have been translated.
- For CM, coded using WHO Drug dictionary, CMDECOD, CMCLASS, ATCNAME (1-4) have been translated.

Done programmatically using the Chinese version of the corresponding dictionary (the term codes are unique and identical between English and Chinese versions).

Translated terms extracted to check consistency with CSR tables

Clinical data package: DATABASES 3/6

AESEQ	AESPID	AETERM	AELLT	AELLTCD
1	1	Upper respiratory tract infection	Upper respirator...	10046306
2	2	Upper respiratory tract infection	Upper respirator...	10046306
1	7	Alanine amino acid transferase increased	Alanine aminotr...	10001551
2	11	Upper respiratory tract infection	Upper respirator...	10046306
3	9	glucose increased	Glucose increased	10018421

AESEQ	AESPID	AETERM	AELLT	AELLTCD
1	1	上呼吸道感染	上呼吸道感染	10046306
2	2	上呼吸道感染	上呼吸道感染	10046306
1	7	丙氨酸氨基转...	丙氨酸氨基转...	10001551
2	11	上呼吸道感染	上呼吸道感染	10046306
3	9	葡萄糖升高	葡萄糖升高	10018421

Clinical data package: DATABASES 4/6

	CMSEQ	CMSPID	CMTRT	ACAT1	ACAT2	CMINDC	CMDECOD
1	19		Salmeterol Xin...	Asthma		Asthma	FLUTICASONE PROPIONATE;SALMETEROL X...
2	3		Acarbose Tabl...	Non-Asthma		MH: Type 2 dia...	ACARBOSE
3	1		Gliclazide Mod...	Non-Asthma		MH: Type 2 dia...	GLICLAZIDE
4	2		Rosiglitazone...	Non-Asthma		MH: Type 2 dia...	ROSIGLITAZONE
5	9		Thioctic Acid f...	Non-Asthma		MH: Type 2 dia...	THIOCTIC ACID

	CMSEQ	CMSPID	CMTRT	CMDECOD
1	19		沙美特罗昔萘酸和丙酸氟替卡松粉吸入剂250/50 ug	丙酸氟替卡松;昔萘酸沙美特罗
2	3		阿卡波糖片	阿卡波糖
3	1		格列齐特缓释片	格列齐特
4	2		罗格列酮钠片	罗格列酮

Clinical data package: DATABASES 5/6

NMPA databases general requirement

If dataset needs to be split → Sponsor can just submit the split dataset* → Details to be included in cSDRG

NO INDICATION OF SUBMISSION REQUIREMENTS →

Chiesi strategy

FDA REQUIREMENTS (5GB) HAVE BEEN CONSIDERED FOR THE SUBMISSION

**Different approach from FDA that requires also the non-split dataset to be submitted*

Clinical data package: DATABASES 6/6

⚠ POINTS OF ATTENTION

Additional challenges in Chinese translation:

- Chinese character = 3 bytes
- English character = 1 byte (generally)

→
- Length of variable might exceed the limit imposed by SAS xport format and CDISC requirements
- Chinese characters need UTF-8 encoding

Medications Dictionary used for English package creation IQ2018Q3

→
- Chinese WHO Drug Dictionary version Q12020 used for translation (first Chinese version available). In case a Chinese term was not available, the current English term has been displayed

Source:
http://xml4pharma.com/publications/Poster_Jozef_Aerts_Chinese_characters_XPT.pdf
https://www.lexjansen.com/phuse-us/2021/dh/PAP_DH12.pdf

Clinical data package: Data Definition Files

NMPA requirements

Data Definition files (DDF)

- Both SDTM and ADaM definition files should be submitted
- Good traceability between data to facilitate regulatory review
- Sponsor needs to specify the dictionaries and their versions in DDF
- DDF is generally in XML or PDF format

Chinese translation

- Description/label and specification of each dataset
- Description/label and derivation progress of variables
- Values or codes list of efficacy indicators

Chiesi strategy

English SDTM and ADaM data definition files have been completely translated in Chinese

XML:Extensible Mark-up Language
PDF: portable document format

Clinical data package: Data Definition Files

DM (人口学) - SPECIAL PURPOSE

相关补充修饰语数据集: SUPPDM (DM补充限定符)

变量	标签/描述	类型	角色	长度或显示格式	受控术语或ISO格式	来源/源/方法/注释
STUDYID	研究标识符	text	Identifier	15		Protocol
DOMAIN	域名缩写	text	Identifier	2	SDTM Domain Abbreviation [31 术语]	Assigned 两个字符的域名缩写。
USUBJID	受试者唯一标识符	text	Identifier	24		Derived USUBJID由STUDYID和SUBJID连接组成，用'_'分隔，例如，'
SUBJID	受试者标识符	text	Topic	8		CRF Annotated Case Report Form [80]

POINTS OF ATTENTION

The pagination in aCRF has changed due to translation, page numbering in SDTM define has been updated accordingly

Define.xml and data have been validated through P21 community with CDISC engine

Clinical data package: Analysis Results Metadata (ARM)

Analysis Results Metadata include **statistical displays or inferential statements such as p-values or estimates of treatment effect**.

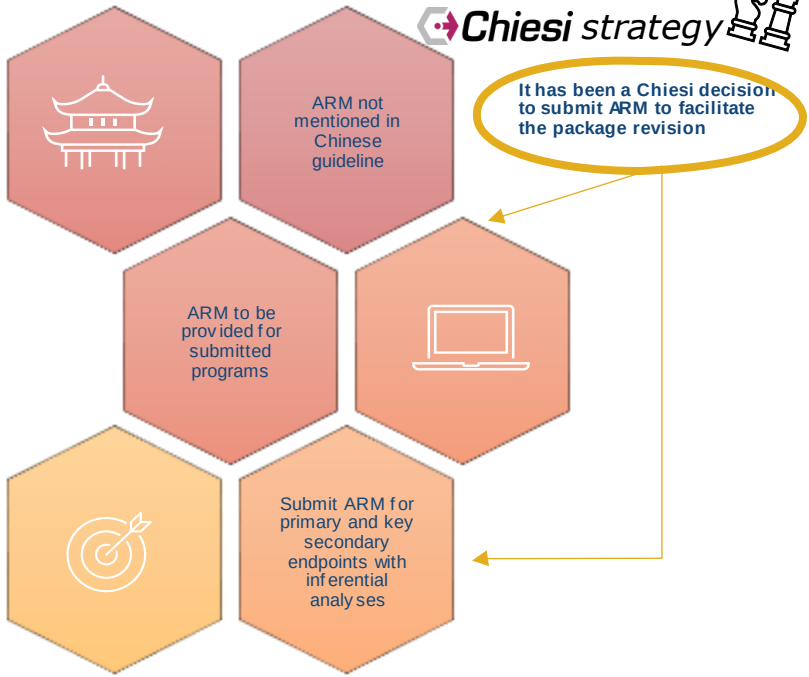
Analysis results metadata are not needed or even advisable for every analysis included in a clinical study report or submission. **The sponsor determines which analyses should have analysis results metadata**.

Table 14.2.1.1.2

显示	Statistical Analysis:Change from Baseline over the Entire Treatment Period in Average Pre-Dose Morning PEF (L/Min).ANCOVA Model (Intention-To-Treat Set)
分析结果	按治疗组在模型中考虑前定试验的数值 (n)
分析参数	PARAMCD = "PEFAMETP" (整个治疗期早晨给药前PEF均值 (L/min))
分析变量	ADPEF.TRT01P (阶段01计划治疗)
分析原因	SPECIFIED IN SAP
分析目的	PRIMARY OUTCOME MEASURE
数据参考（包括选择标准）	ADPEF [PARAMCD = "PEFAMETP" and BASE ≠ missing and CHG ≠ missing and ITTFL = "Y" and REGION1 ≠ "missing" and SEX ≠ "missing"]
文献资料	意向治疗(ITT)分析集中包含的受试者的唯一计数，其中每个治疗组内协变量（基线值）未缺失，固定效应（区域和性别）未缺失，并且从随机治疗期开始日期后第二天的早晨时刻到随机治疗期结束日期的早晨时刻之间的时间间隔内至少有一个可评估的访视间期。 SAP第9.1.1节 [35] SAP第14节 [58]
编程语句	[SAS version 9.4] proc sql; select distinct TRT01P, count(distinct USUBJID) as n from ADPEF where ITTFL="Y" and PARAMCD="PEFAMETP" and missing(CHG)=0 and missing(BASE)=0 and REGION1 and SEX group by TRT01P; quit; t14_2_1_1_2

NMPA Data definition files requirements

Chiesi strategy



Submission package: Reviewer's guide

NMPA requirements

Reviewer's guide

- Should be submitted in Chinese
- Supplement to data definition files for the reviewers
- Provides information in addition to what we have in data definition file
- Submitted in .pdf

Chiesi strategy

Both English SDTM and ADaM reviewer's guides have been translated

Clinical data package: Programming code

NMPA programming requirements

Submitted as txt files

- Readable with comments
- Understandable
- No external macro calls

No clear definition within the Chinese guideline of which programs should be submitted

Chiesi strategy

ADaM
- Dataset that contains primary endpoint
- Dataset that contains selected secondary endpoints

TABLES
- Tables used for primary endpoint
- Tables used for selected secondary endpoints

LISTINGS and FIGURES
- Not submitted

MACROS
- Only study specific macros

```
ods exclude all;
ods output ModelANOVA=_ModelANOVA ContrastEstimate=_ContrastEstimate;

proc phreg data = adam.adtte;
        CLASS TRTPN SITEGR1 EXP1YG2 FV1PPG1 SMSTT;
        MODEL AVAL*cnsr(1) = TRTPN SITEGR1 EXP1YG2 FV1PPG1 SMSTT / ties=exact type3(wald) RL=wald;
        CONTRAST "Study treatment vs Placebo" TRTPN 1 / estimate=exp;
        where ittfl = 'Y' and paramcd="TFMSEX";
RUN;

ods output close;
ods exclude none;

proc freq data=adam.adsl noprint;
        where ITTFL="Y";

        table TRT01PN*TRT01P / out=trtcnt_01 (drop=percent);
run;

data trtcnt_02;
        set trtcnt_01;
        format text $80.;
        printord="01";
        sort=TRT01PN;
        text=strip(put(TRT01P,$30.))||" (N="||strip(put(count,4.))||")";
run;
```

No translation is needed!!

Walls / Wins / Wisdoms

Walls

1. **Chinese requirements**: evolving regulations with continuous updates and China-specific activities

2. **It is not all about translation**: Chinese SDTM IG and ADaM IG should be applied and consistency between documents should be ensured

3. Complex review of the **Chinese translation** of e-submission package

The Great China Wall, II century before Christ

Wins

1. **Strong internal know-how** on FDA package preparation

2. **Cross-functional team-work is the key:**
   - Close cooperation between **clinical, regulatory and company team based in China** to ensure full alignment on the strategy

3. Constant **support to translation**, regulatory and medical writers to ensure consistency

Gengis Khān, 1162-1227 after Christ

Wisdoms

1. **Plan ahead:**
   - **Define internal standards** for Chinese data submission

2. Chinese Guideline to be more and more **aligned with requirements from other countries**

3. **…and a dream:** that in the mid term Chinese authority will not longer ask the translation of the database!

Confucius, 551-479 before Christ

A big big thank to..

Chiesi data package team:

- Paola Vaghi (Head of Statistical Programming)
- Glauco Cappellini (Lead Statistician)

Thank You!

Rebecca Moretti: r.moretti@chiesi.com

Emanuele Calabrò: e.calabro@chiesi.com