PharmaSUG China 2021 – DS-047

Challenges and solutions for e-data submission to NMPA under new regulation

Grace Sun, Astrazeneca

ABSTRACT

As a global biopharmaceutical company, previously the FDA submission could be directly used for NMPA submission in Astrazeneca. With the implementation of new guideline on the submission of clinical trial data from October 1, 2020, there are lots of China specific requirements for submission package, including readable programs and language requirements for dataset in XPT format, aCRF, define and data reviewer's guide. To provide the best practice for NMPA submission, our company streamlined the whole process and developed the working instruction for guiding e-submission under new regulations. The instruction emphasized the working sequence to ensure the consistencies of Chinese translation between datasets and e-submission documents, especially the cross functional collaboration on proofreading of translated aCRF, define and data reviewer's guide, and the translation of concomitant medication. Besides, to improve efficacy of translation and proofreading, datasets are translated with programs, , AI tools are under development to optimize solution for document translation. Overall, the current practice achieves effective cross-functional cooperation and meets regulatory compliance.

INTRODUCTION

The "Guideline on the Submission of Clinical Trial Data" was published on Jul 20th, 2020. NMPA(National Medical Products Administration) requires chemical drugs and biological products implement this new guideline since Oct 1st, 2020. This new guideline highlights the language requirement for the submission components, including study database, analysis database, data definition file, data reviewer's guide and annotated CRF. We faced lots of challenges in the process of translation and tried multiple ways to address these challenges. This paper describes the current best practice of Astrazeneca to make e-submission under new guideline, especially the working sequence to make sure the consistencies of Chinese translation between database and other documents in submission package.

LANGUAGE REQUEST ON DATA SUBMISSION PACKAGE

DATASET

In the data submission package, at least the following items should be in Chinese: dataset label, variable label, adverse events terms, generic name of concomitant medications and medical history that occurred in CSR and other documents.

CRFS/ACRFS, DEFINE.XML, AND DATA REVIEWER'S GUIDES

In aCRF, at least the following terms should be in Chinese: questions designed for data collection, efficacy related value list or code list. In current AZ China practice, entire CRF is translated.

In Define.xml, at least the following terms should be in Chinese: datasets description/label and definition; variables description/label and derivation; efficacy related parameter value list or code list.

Data reviewer's guide including SDRG and ADRG should be in Chinese.

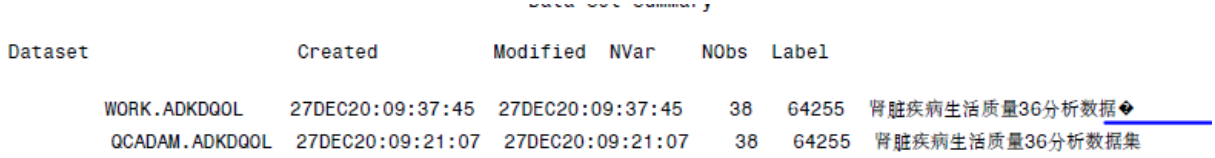
WORKING INSTRUCTION

GENERAL PRINCIPLE

Dataset and variable label translation should be completed before the translation of aCRF, reviewer's guide and define excel spec. In this way, the translation of dataset and variable label could be consistent between database and submission documents. In current practice, AZ internal programmers conduct database translation with two-side programming.

DATASET AND VARIABLE LABEL TRANSLATION

For CDISC dataset and variable labels' translation, guidance documents from CDISC Coordinating Committee(C3C) are used as reference, since common datasets and variables' label have been translated into Chinese in these documents. The dataset and variable labels will be organized into a library. It occurs some variables cannot find corresponding Chinese labels from the latest version. In this case, the other versions should be used, especially for ADAM metadata. Note the length of these label couldn't exceed 40 characters, which will cause truncation during converting XPT V5. When the library set up, the program codes could be used to map Chinese labels from library to SDTM and ADAM datasets.

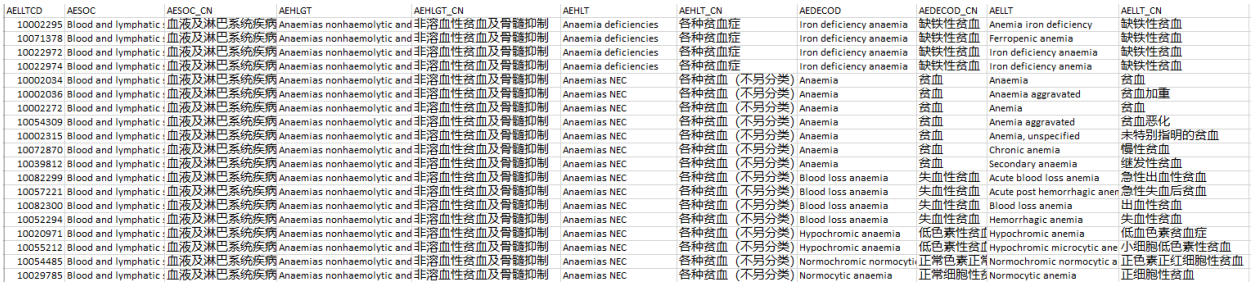


Dataset	Created	Modified	NVar	NObs	Label
WORK.ADKDQOL	27DEC20:09:37:45	27DEC20:09:37:45	38	64255	肾脏疾病生活质量36分析数据
QCADAM.ADKDQOL	27DEC20:09:21:07	27DEC20:09:21:07	38	64255	肾脏疾病生活质量36分析数据集

CHINESE VERSION OF MEDDRA AND WHODRUG MAPPING

MedDRA mapping logic

Select corresponding Chinese validated version of MedDRA dictionary to ensure the consistency between English version and Chinese version. Currently, the regulation requires that adverse events terms and medical history should be in Chinese, which means variables (--BODSYS, --DECOD, --LLT, etc) in AE, MH and CE (optional) domain will be translated by merging Chinese MedDRA with common variable –LLCD. In further ADaM translation, it's recommended to get translated variables from SDTM directly. After completing MedDRA translation, Chinese adverse event and medical history terms should be extracted in excel file as below, which could be used to check the consistencies with CSR in-text tables.



AELLTCD	AESOC	AESOC_CN	AEHLGT	AEHLGT_CN	AEHLT	AEHLT_CN	AEDECOD	AEDECOD_CN	AELLT	AELLT_CN
10002295	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemia deficiencies	各种贫血症	Iron deficiency anaemia	缺铁性贫血	Anemia iron deficiency	缺铁性贫血
10071378	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemia deficiencies	各种贫血症	Iron deficiency anaemia	缺铁性贫血	Ferropenic anemia	缺铁性贫血
10022972	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemia deficiencies	各种贫血症	Iron deficiency anaemia	缺铁性贫血	Iron deficiency anaemia	缺铁性贫血
10022974	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemia deficiencies	各种贫血症	Iron deficiency anaemia	缺铁性贫血	Iron deficiency anemia	缺铁性贫血
10002034	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemias NEC	各种贫血（不另分类）	Anaemia	贫血	Anaemia	贫血
10002036	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemias NEC	各种贫血（不另分类）	Anaemia	贫血	Anaemia aggravated	贫血加重
10002272	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemias NEC	各种贫血（不另分类）	Anaemia	贫血	Anemia	贫血
10054309	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemias NEC	各种贫血（不另分类）	Anaemia	贫血	Anemia aggravated	贫血恶化
10002315	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemias NEC	各种贫血（不另分类）	Anaemia	贫血	Anemia, unspecified	未特别指明的贫血
10072870	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemias NEC	各种贫血（不另分类）	Anaemia	贫血	Chronic anemia	慢性贫血
10039812	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemias NEC	各种贫血（不另分类）	Anaemia	贫血	Secondary anaemia	继发性贫血
10082299	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemias NEC	各种贫血（不另分类）	Blood loss anaemia	失血性贫血	Acute blood loss anemia	急性出血性贫血
10057221	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemias NEC	各种贫血（不另分类）	Blood loss anaemia	失血性贫血	Acute post hemorrhagic anem	急性失血后贫血
10082300	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemias NEC	各种贫血（不另分类）	Blood loss anaemia	失血性贫血	Blood loss anemia	出血性贫血
10052294	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemias NEC	各种贫血（不另分类）	Blood loss anaemia	失血性贫血	Hemorrhagic anemia	失血性贫血
10020971	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemias NEC	各种贫血（不另分类）	Hypochromic anaemia	低色素性贫血	Hypochromic anemia	低血色素贫血症
10055212	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemias NEC	各种贫血（不另分类）	Hypochromic anaemia	低色素性贫血	Hypochromic microcytic ane	小细胞低色素性贫血
10054485	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemias NEC	各种贫血（不另分类）	Normochromic normocyti	正常色素正常	Normochromic normocytic a	正色素正红细胞性贫血
10029785	Blood and lymphatic :	血液及淋巴系统疾病	Anaemias nonhaemolytic and	非溶血性贫血及骨髓抑制	Anaemias NEC	各种贫血（不另分类）	Normocytic anaemia	正常细胞性贫	Normocytic anemia	正细胞性贫血

WHODRUG(B3 format) mapping logic

Generally, the medication code （DRUGCODE）linked to dictionary drug text and ATC code (ATCCODE) linked to the dictionary ATC text description which the reported medication/therapy is coded. The format of the DRUGCODE is: Drug Record Number || Sequence Number 1 (Seq 1) || Sequence

Number 2 (Seq 2). Note the preferred name corresponding to the DRUGCODE when Seq2 = 1. When merging with Chinese WHODrug dictionary, DRUGCODE and ATCCODE are key variables.

Choose appropriate Chinese validated version of WHODrug to ensure the consistency between English version and Chinese version. Not all CMDECOD's value can be programmable translated by mapping with Chinese WHODrug dictionary with GRUGCODE, since not all preferred terms are fully translated in Chinese dictionary. Variable ING_LIST will be considered as preferred term to map into CMDECOD firstly as long as its value has been translated in Chinese WHODrug, then variable DRUGNAME will be used to map the rest of CMDECOD. Finally, for the rest of CMDECOD highlighted in yellow which couldn't be translated to Chinese, study team should decide whether they need to be translated manually.

CMCAT	CMDECOD	CMDECOD_CN	ATCCD	ATCDTXT	ATCDTXT_CN
MEDICATIONS CKD AND DIAB	METHYLDOPA	甲基多巴	C02AB	METHYLDOPA	甲基多巴
MEDICATIONS OTHER	METHYLDOPA	甲基多巴	C02AB	METHYLDOPA	甲基多巴
MEDICATIONS OTHER	DIPHENHYDRAMINE	苯海拉明	R06AA	AMINOALKYL ETHERS	氨基烷基醚类
MEDICATIONS OTHER	DIPHENHYDRAMINE	苯海拉明	D04AA	ANTIHISTAMINES FOR TOPICAL USE	外用抗组胺药
MEDICATIONS OTHER	DIPHENHYDRAMINE	苯海拉明	A04AD	OTHER ANTIEMETICS	其他止吐剂
MEDICATIONS OTHER	DIPHENHYDRAMINE	苯海拉明	N05CM	OTHER HYPNOTICS AND SEDATIVES	其他催眠药和镇静药
MEDICATIONS OTHER	DIPHENHYDRAMINE HYDROCHLORIDE	盐酸苯海拉明	R06AA	AMINOALKYL ETHERS	氨基烷基醚类
MEDICATIONS OTHER	DIPHENHYDRAMINE HYDROCHLORIDE	盐酸苯海拉明	D04AA	ANTIHISTAMINES FOR TOPICAL USE	外用抗组胺药
MEDICATIONS OTHER	DIPHENHYDRAMINE HYDROCHLORIDE	盐酸苯海拉明	A04AD	OTHER ANTIEMETICS	其他止吐剂
MEDICATIONS OTHER	DIPHENHYDRAMINE HYDROCHLORIDE	盐酸苯海拉明	N05CM	OTHER HYPNOTICS AND SEDATIVES	其他催眠药和镇静药
MEDICATIONS OTHER	DIPHENHYDRAMINE LAURILSULFATE	DIPHENHYDRAMINE LAURILSU	D04AA	ANTIHISTAMINES FOR TOPICAL USE	外用抗组胺药

SUBMISSION DOCUMENTS TRANSLATION

aCRF Translation

Per AZ current practice, entire CRF will be translated, and SDTM annotations are not needed to translate. General format and page layout need to be consistent with English version. It is recommended to use doc file for proofreading, while pdf file will be delivered with SDTM annotation finally.

When finalized aCRF is available, Programmer will inform China DM to coordinate translation and conduct proofreading of contents(may need physician's support), format and page layout to ensure the consistency with English version. While programmer will finally check the original SDTM annotation are in the correct place.

Define.xml, Reviewer's Guide translation

Convert the data definition file from define.xml to define spec. Use the define spec to translate variables derivation, efficacy related parameter value list or code list. When excel file translation completed, covert the Chinese excel file back to xml with Pinnacle 21.

These two reviewer's guide need to be fully translated into Chinese. And the common contents should be consistent with protocol and SAP,

CHALLENGES

TRUNCATION OF VARIABLE VALUES WHEN CONVERTING ENCODING

When converting encoding from WLATIN to UTF8, it may cause value truncation for long variables, such as DVTERM. To avoid truncation, the length of variables should be extended before converting encoding.

CLAB-Blood C...	Retest because sample was out of stability.
ADJUDICATIO...	Data are not complete but definately a Cx ocklusion with lear TnT elevation. Proa♦ bably chest pain but dates are confusing. No ECG.
ADJUDICATIO...	Re-adjudication due to unwanted character to be removed.

WHODRUG NOT FULLY TRANSLATED INTO CHINESE

When following this instruction to translate e-submission package, validation should be conducted for each step to check if all required terms have been fully translated into Chinese. After merging with Chinese WHODrug, some concomitant medication terms may still keep English. It's recommended to extract these English terms in excel file and send to study team for decision whether to translate them manually.

CONSISTENCIES OF CHINESE TRANSLATION ACROSS E-SUBMISSION DOCS

It's recommended to extract translated adverse event, concomitant medication and medical history in excel file. We could refer to these excel file to translate e-submission documents and ensure the consistencies between database and docs.

When proofreading aCRF, reviewer's guide and define, the Chinese protocol and statistical analysis plan should be used to keep same wording between study docs.

CONCLUSION

Effective cross functional collaboration could ensure the high quality submission package. With making working plan in advance and opening the kick off meeting, the team can achieve timelines and regulatory requirements. In addition, this working instruction could help make successful e-submission package under new regulation.

REFERENCES

National Medical Products Administration. 2020. "Guideline on the Submission of Clinical Trial Data." July 20, 2020. http://www.cde.org.cn/news.do?method=largeInfo&id=7a43c3abfde95950.

Uppsala Monitoring Centre. "WHODrug User Guide".

CDISC Submission Data Standards Team. 2018. "Study Data Tabulation Model Implementation Guide: Human Clinical Trials". November 20, 2018. https://www.cdisc.org/system/files/members/standard/foundational/SDTMIG_v3.3_FINAL.pdf.

CDISC Submission Data Standards Team. 2019. "Analysis Data Model Implementation Guide". October 03, 2019. https://www.cdisc.org/standards/foundational/adam/adamig-v1-2-release-package.

ACKNOWLEDGMENTS

The authors would like to thank their respective management team for their time in reviewing the paper and providing with valuable comments.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Grace Sun
Astrazeneca
13386087290
Grace.sun1@astrazeneca.com