

Analysis Table Translation Tool

Authors: Jessie Zhao, Mina Chen, Shengying Cheng, Yanyun Shen

Abstract:

Background: Due to the requirement of SFDA, all clinical study report (CSR) will need to be translated into Chinese. This requires that all the in text tables need to be translated into Chinese too. The company used to hire external vendors to perform this task. The translated product needs another round of internal review which is very time consuming and labor intensive.

Our Solution: We established a project to instruct two interns to generate a tool using SAS to automatically convert analysis tables from English into Chinese. With the use of this tool, all in text analysis tables will not need to be manually translated but automatically translated into Chinese in text tables. This greatly increased the efficiency of translation part of the analysis tables used for CSR and reduced the risk of incidental mistakes which can potentially occur during the manual translation process. The other significance of this tool is that it can be modified easily to translate from English to other languages too as long as SAS software and the operation systems support these other languages. To implement this tool in practice, it requires a close collaboration between SAS programmers and biostatisticians. This paper also proposed a working process to ensure effective application of this tool in project work. In summary, we feel this has been a great project as an internship program as it not only provided a challenging opportunity for our interns to develop sophisticated SAS programming skills but also independent research and problem solving skills.

Below is an example of the original analysis in text table and a Chinese translated version of it, which will be include the Chinese CSR to be submitted to sFDA.

aell_c DRAFT - Summary of Adverse Events by Body System and Trial Treatment Summary of All Adverse Events by System/Preferred Term Protocol(s): DUMM49 Analysis: ALL RANDOMIZED POPULATION Center: ALL CENTERS Adverse Event Onset between Time of Very First Drug Intake and Study Day 9999, Time 23:59				aell_c 草稿-不良事件按身体器官系统和治疗分类总结 不良事件按身体器官或首选术语分类总结 方案: DUMM49 分析: 所有随机人群 中心: 所有中心 试验 Day 1, Time 00:00 至试验 Day 9999, Time 23:59 发生的不良事件			
Body System/ Adverse Event	T_CEFTRIAXONE 1 000 mg QD N = 95 No. (%)	T_R04908463 150 0 mg BID N = 104 No. (%)	T_R04908463 750 mg BID N = 103 No. (%)	身体系统/ 不良反应事件	T_CEFTRIAXONE 1 000 mg QD N = 95 No. (%)	T_R04908463 150 0 mg BID N = 104 No. (%)	T_R04908463 750 mg BID N = 103 No. (%)
ALL BODY SYSTEMS				所有身体系统			
Total Pts with at Least one AE	39 (41)	49 (47)	51 (50)	至少发生一次不良事件的患者合计	39 (41)	49 (47)	51 (50)
Total Number of AEs	85	97	108	不良事件合计	85	97	108
GASTROINTESTINAL DISORDERS				消化系统疾病			
Total Pts With at Least one AE	9 (9)	12 (12)	13 (13)	至少发生一次不良事件的患者合计	9 (9)	12 (12)	13 (13)
DIARRHOEA	5 (5)	3 (3)	6 (6)	腹泻	5 (5)	3 (3)	6 (6)
GASTRITIS	1 (1)	1 (<1)	3 (3)	胃炎	1 (1)	1 (<1)	3 (3)
ABDOMINAL PAIN UPPER	1 (1)	3 (3)	-	上腹痛	1 (1)	3 (3)	-
VOMITING	-	3 (3)	1 (<1)	呕吐	-	3 (3)	1 (<1)
CONSTIPATION	1 (1)	-	2 (2)	便秘	1 (1)	-	2 (2)
NAUSEA	1 (1)	2 (2)	1 (1)	恶心	1 (1)	2 (2)	-
TOOTHACHE	-	1 (<1)	-	牙疼	-	1 (<1)	1 (<1)
ABDOMINAL PAIN	-	1 (<1)	-	腹痛	-	1 (<1)	-
DYSPEPSIA	1 (1)	-	-	消化不良	1 (1)	-	-
RECTAL HAEMORRHAGE	-	1 (<1)	-	直肠出血	-	1 (<1)	-
SMALL INTESTINAL HAEMORRHAGE	1 (1)	-	-	小肠出血	1 (1)	-	-
Total Number of AEs	11	15	13	不良事件合计	11	15	13
Percentages are based on N. Multiple occurrences of the same adverse event in one individual counted only once. Output : \$MARSOUT/cdplay2/dummy49/aell_c.dat AE11 28OCT2010:13:40:45 (1 of 8)				以 N 为分母计算百分比 对于同一患者发生多条相同不良事件, 仅计数一次。 输出 : \$MARSOUT/cdplay2/dummy49/aell_c.dat AE11 28OCT2010:13:40:45 (1 of 8)			

Why is it necessary to develop this tool?

According to sFDA requirement, all CSR including in-text tables should be in Chinese. Before this tool is generated, the whole CSR including the in-text tables are manually translated by vendor companies. After the CSR is translated, it requires another round of internal review. Different vendors may translate the same adverse event terms differently due to lack of standardized practice of medical translation in China. If the operating system of SAS supports Chinese, we could have generated analysis tables in Chinese directly. Unfortunately the company's operating system doesn't support Chinese, therefore we need to translate them into Chinese afterwards. In addition to the potential differences in the translated terms of the same contents, the format of the original tables may need to be justified as the translation will result in a different length of the same analysis term for each individual row on the table. The modification of the space for each row is tedious and can easily

cause potential inadvertent changes for analysis numbers, which will jeopardize the overall quality of the analysis tables.

What's the benefit of this tool?

- Produce high quality tables in Chinese
 - ✓ We can guarantee the accuracy of the translation for all Adverse Event (AE) terms because we utilize the MedDRA (the Medical Dictionary for Regulatory Activities) and its Chinese translation published by MSSO (Maintenance and Support Services Organization). Our scientists don't need to check the Chinese translation for AE terms if we use this tool.
 - ✓ We will create a dictionary for some other frequent used terms. Once we performed quality check for those terms, we can guarantee the quality and there is no need to manually check each translated term from the original.
- Accelerate the submission
 - ✓ It takes about 6 weeks for vendors to translate CSR. If we don't ask them to translate the in-text tables, it will reduce the time spent on the whole CSR translation as well as the overall translation cost. And we can perform table translation in parallel.
 - ✓ The time spent on quality check of the translated in-text tables will be minimized as the all the analysis terms are translated in standardized and automated fashion.
 - ✓ There is no need to spend any time on reformatting for the translated tables manually because we can generate the Chinese tables with the same format as those tables in English
- Efficiency Gained
 - ✓ This tool can be used on any analysis tables we generated.
 - ✓ This tool is user friendly. It's very easy to learn how to use it.
- Future application of this tool
 - ✓ It's applicable to translate tables from English to other languages as long as the operation system support and SAS support other languages.

Challenges

- We can't translate medication terms as the same way as we translate adverse event terms.
 - ✓ We are using in-house dictionary which is only in English to do medical coding for medication.
 - ✓ The company didn't purchase WHODD (WHO Drug Dictionary) which has Chinese translation for medication terms
 - ✓ There are two ways to solve this issue, either to create our own Chinese translation for all medications used in our studies or purchase any commercial dictionary. It needs to be further investigated.

The design of the tool and its application in daily work

The tool is developed via a series of macros. This paper will not share the coding as they are the company proprietary information. This paper will share the design of the tool at a high level and

also our recommended working process in its use to daily work to ensure effective collaboration between programmers and biostatisticians.

Design of the tool

The only information this tool requires is the original analysis table. It reads in the original analysis table as an external file and tease out the columns of analysis terms (usually the first one to several columns from left), which need to be translated. The tool leaves out the columns of summary information of all treatment arms as they do not need to be translated. The other information which needs to be translated is the titles and footnotes.

Translation of the analysis terms

The analysis terms of all the tables will be stored in an excel spreadsheet. For example, for an adverse event table, the analysis term will be: 'All body systems', 'Total patients with at least one AE', 'Gastrointestinal Disorders', 'Diarrhea', etc; for a baseline characteristic table, 'Age group', 'gender', etc, will be the analysis terms. These analysis terms will be cumulated as more and more tables need to be translated. Then for each analysis term, the Chinese translation will be added in as another column in the excel spreadsheet. These excel sheet will be converted into a data base which will include a set of unique analysis terms and their Chinese translated terms. The final translation of any of the analysis table can then be easily completed via a merge procedure with the Chinese value replacing the English value for each analysis term.

Translation of titles and footnotes

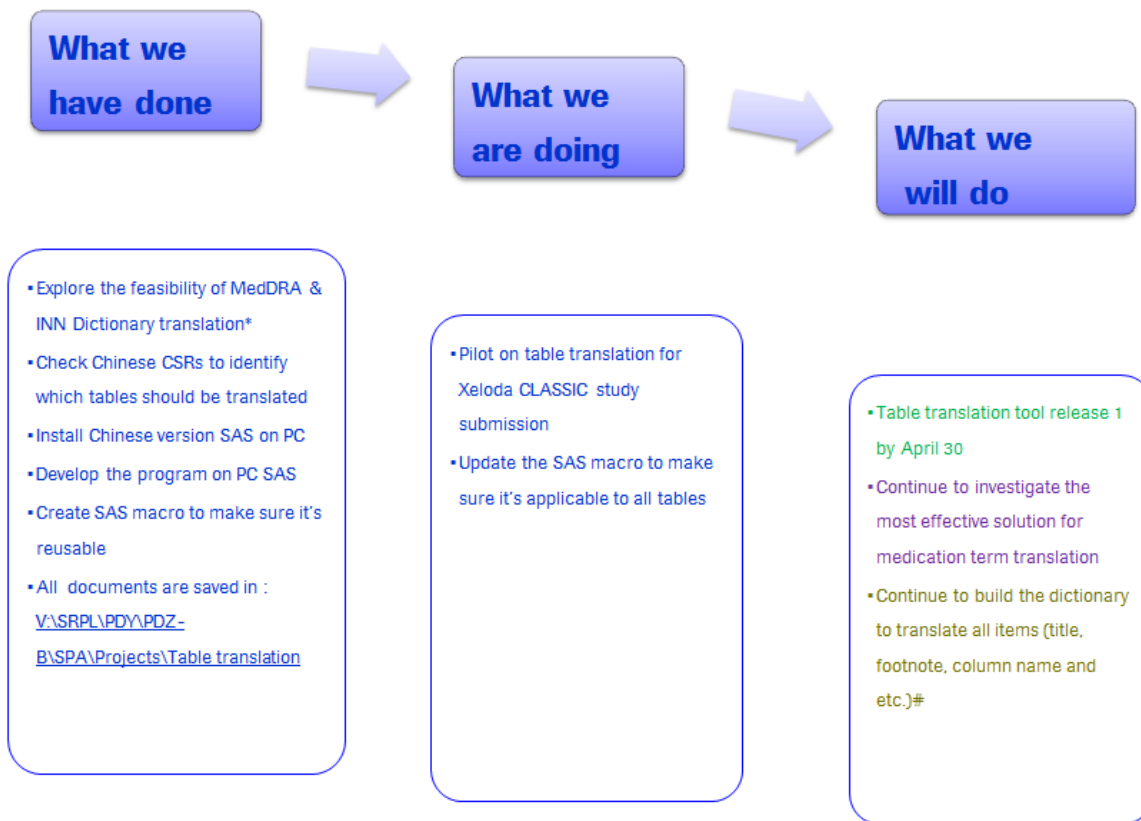
Translation of the titles and footnotes are not via building an excel file as they are not standardized yet from one study to another one for the same type of analysis table. The way it is translated is via reading them in as text strings and applying them via macro variable values. In the future, when footnotes and titles are standardized, building them into another excel table is a more efficient way to perform the translation.

Working process when applying the tool

When we start to apply the tool for a study, it will require a close collaboration between statisticians and programmers to ensure effective use of programmers' time so that only the final in-text tables are translated. In reality, statisticians and programmers usually would like to start the translation as early as possible, but sometimes the decision on selection of in-text tables may not be finalized. In that case, some of the tables may be translated at risk as they are not selected to be included in the final set of in-text tables. These efforts should really be minimized so that time is not wasted. So we ask that statisticians work with clinical scientists at the stage of designing data analysis plan to identify those tables as in-text tables and document them as part of the analysis plan. The programmers then can start the translation once analysis tables are generated. Later on, if more tables are needed to be included as in-text tables, they can be added. We hope the selection of in-text tables are carefully evaluated at each stage mentioned above to avoid unnecessary work.

Summary

In summary, we feel this has been a great project as an internship program as it not only solved a business problem but also provided a challenging opportunity for our interns to develop sophisticated SAS programming and independent research and problem solving skills. We hope the presentation of this paper will inspire the Chinese SAS professionals to proactively identify business needs from their companies and find solutions to them. And by sharing our experience, we are contributing to the development of Chinese SAS users' community.



Team members:

Cindy (Shengying) Cheng; Mina Chen (SPA intern); Jessie Zhao (SPA intern)

Acknowledge:

Vianna Li; Marek Dorofiejczyk; Susan O'Connor; Annay Mistry;

Thanks for your help!

If you have any comments or questions, please feel free to let us know. Thanks in advance!