

Use RTF Control Word in SAS Integration with MS Word: Mechanisms and Applications

Ryan Wang, Zhi Zhi Medicine Technology Co., Ltd.

ABSTRACT

During study, changes in documents like SAP, Shell, or CSR might cause form title numbers to change. This means we often need to update SAS programs or prepare table-of-contents (TOC) files. To handle this, managing document versions and auto-updating numbers became essential.

This paper explains how to use simple Word features (styles and field codes) for better document control. It also describe how to create RTF files with titles and field codes in SAS, and how to connect these steps with later processing workflows to make everything work smoothly.

INTRODUCTION

In SAP, Shell, and CSR documentation, combining title styles with field codes enables auto-updating of table title numbers. This also allows automatic extraction of TFLs (tables, figures, listings) directories, significantly improving document management efficiency and accuracy. Both manual execution and VBA (Visual Basic for Applications) scripting can achieve this automation. When generating TFLs, embedding specific control characters in RTF files ensures consistent numbering and structure during output.

LIST OF FIELD CODES IN WORD

The company template libraries typically include standardized font styles. Heading styles use outline level numbering through the “Home → Paragraph → Multilevel List” , enabling auto-updated heading numbering.

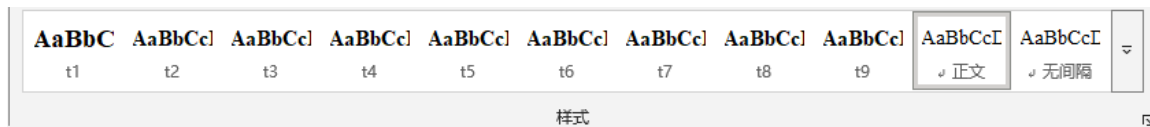


Figure 1. Heading Styles t1-t9



Figure 2. Heading Styles with Multilevel List t1-t9

When using the heading styles described above, the Navigation Pane displays the document's structure. If content changes occur, dragging headings in the pane to new positions automatically moves their subheadings and content, with subsequent heading numbers updating accordingly.

标题	页面	结果
14. 参考但不纳入文本的表格、图示和图表		
14.1. 人口统计学资料及基线特征		
14.1.1. 人口统计学资料		
表 14.1.1 人口学资料总结(FAS)		
表 14.1.2 身高体重(FAS)		
表 14.1.3 血清学检查(FAS)		
表 14.1.4 其他基线特征(FAS)		
14.1.2. 病史		
表 14.1.5 肿瘤病史总结(FAS)		
表 14.1.6 其他肿瘤病史总结(FAS)		
表 14.1.7 其他既往病史/伴随疾病(FAS)		
14.1.3. 既往肿瘤治疗史		
表 14.1.8 既往肿瘤药物治疗史总结(FAS)		
14.1.4. 既往/合并治疗		
14.1.5. 用药依从性		

Figure 3. Navigation Pane

FIELD CODES : INSERT AND HOTKEYS

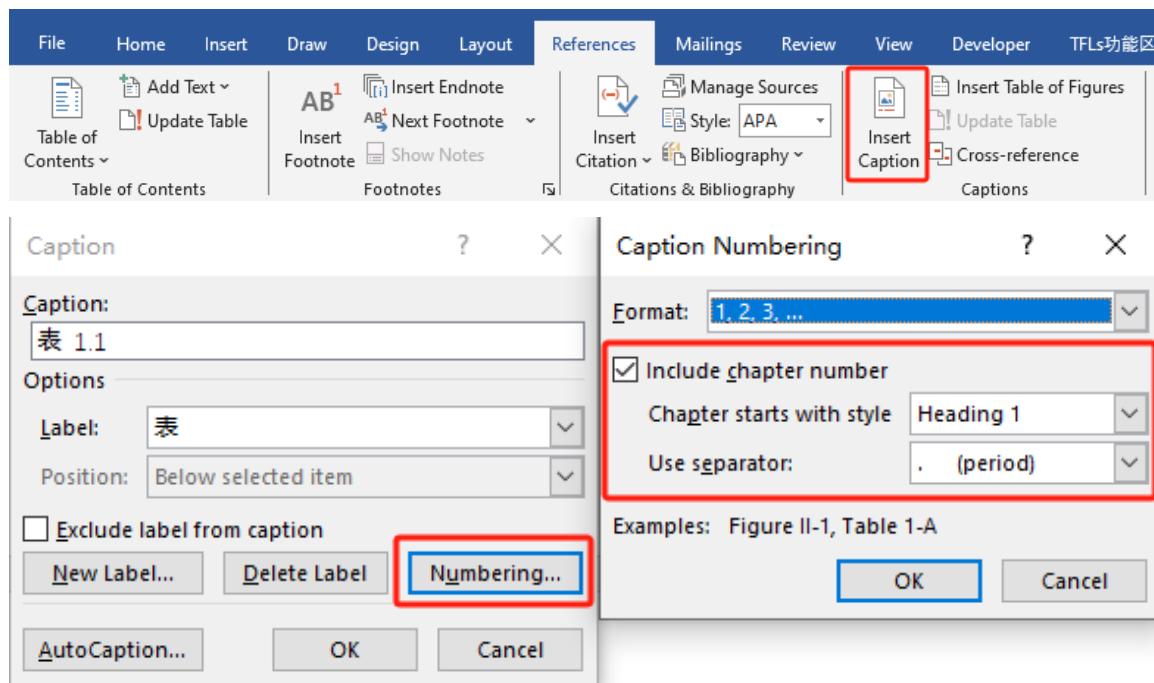


Figure 4. Insert Caption about Table

Inserted content appears directly in the document : 表 1.1 title text .

Switch to field code view to see the underlying code : 表 { STYLEREF 1 \s }. { SEQ 表 * ARABIC \s 1 } title text .

(You could create Label like Table, Figure, or Listing instead of 表)

And , use these keyboard shortcuts for field code operations:

Insert field code: Ctrl+F9

Show / Hide codes : ALT+F9 (all document) or Shift+F9 (single location)

Update fields : F9

Update single field : Click target field → F9

Update all fields : Ctrl+A (select all) → F9。

FIELD CODES : STYLEREF

The StyleRef field inserts text formatted with specified style settings. 。

Syntax :

{ STYLEREF *StyleIdentifier* [Switches] }

Example :

{ STYLEREF *t2* \s }.{ SEQ 表 * ARABIC \s 1 }

Explanation:

- The bold parts switch between field code and display.
- When refreshed, this code retrieves the most recent heading number using the “*t2*” style before the field.
- E.g.: If the qualifying heading is “14.3 SAFETY DATA”, the field will display as “14.3.x” after refresh (where x increments automatically).

FIELD CODES : SEQ

The SEQ (Sequence) field numbers document items like chapters, tables, and figures sequentially using identifiers.

Syntax in documents:

{ SEQ *Identifier* [Bookmark] [Switches] }

Example:

{ STYLEREF *t2* \s }.{ SEQ 表 * ARABIC \s 1 }

Explanation:

- The bold parts switch between field code and content.
- When refreshed, this code continues numbering from the last used “表” identifier.
- E.g.: If the previous table number was “14.3.3”, this field will display “14.3.4” after refresh.
- Of note, to reset numbering to start at “1”, modify “\s 1” to “\r 1” in the code.

METHODOLOGY AND DISCUSSION

Next, we will introduce how to use RTF control words in SAS to create RTF documents containing field codes:

INSERT FIELD CODES IN RTF VIA SAS

In general, when outputting RTF documents in SAS, an escape character such as “ods escapechar='^';” is declared, and the RTF control words following the escape character are input into the document to achieve some text or table formatting.

\tab	Retract	缩进
\i	Italicize	斜体
\ul	Underline	下划线
\dul	Double underline	双下划线
\line	New line	换至新行
\bullet	Bullet	添加项目符号
\embo	Emboss	字体突出
\impr	Engrave	字体凹陷
\sub	Subscript	下标
\super	Superscript	上标
\outl	Outline	文字轮廓线
\shad	Shadow	文字阴影
\strike	Strike	删除线
\strike	Double strike	双删除线
\uld	Dotted underline	下划线-虚线
\ulw	Wave underline	下划线-波浪线
\ulth	Thick underline	下划线-粗线
\cfN	Foreground color	字体颜色，数字选择颜色
\fsN	Font size	字体大小，数字选择字号
\highlightN	Highlight	高亮，数字选择颜色
\outlinelevelN	Outline Level	段落大纲级别，数字选择级别（0~9），0 表示 1 级大纲，9 表示正文级别
\b	Bold	字体加粗
\ql	Left aligned	字体左对齐
\qr	Right aligned	字体右对齐
\qc	Centered	字体居中

In this way, the RTF control words about field code can also be written into the document via SAS, with the key statements :表 \$R'{\$field{\$fldinst{STYLeref t3 \s }}}.{\$field{\$fldinst{ SEQ 表 * ARABIC\s 1 }}} 班级信息。

```
ods escapechar='$';
```

```
title "表 $R'{$field{$fldinst{STYLeref t3 \s }}}.{$field{$fldinst{ SEQ 表 \* ARABIC\s 1 }}} 班级信息";
```

```
ods results off;
```

```
ods rtf file="C:\Users\Anti\Desktop\class.rtf" bodytitle;
```

```
proc report data=sashelp.class;
    column sex name age height weight;
run;
ods rtf close;
ods results on;
```

Output 1 shows this example.

表 . 班级信息

性别	姓名	年龄	身高 (英寸)	体重 (磅)
男	阿尔弗雷德	14	69	112.5
女	爱丽丝	13	56.5	84
女	芭芭拉	13	65.3	98
女	凯露	14	62.8	102.5
男	亨利	14	63.5	102.5

表 *STYLEREF t3* SEQ 表 **ARABIC t1* 班级信息

性别	姓名	年龄	身高 (英寸)	体重 (磅)
男	阿尔弗雷德	14	69	112.5
女	爱丽丝	13	56.5	84
女	芭芭拉	13	65.3	98
女	凯露	14	62.8	102.5
男	亨利	14	63.5	102.5

Output 1. The output document with Field Codes via SAS code above

THE BEST PRACTICE

As mentioned earlier, field codes need to work in conjunction with heading styles for full-text, therefore, the configuration of the title and the style used need to be considered. Additionally, control words for paragraph outline levels can be incorporated to meet the requirements for multi-level bookmarks, namely:

- 1. Output chapter titles with outline levels and apply styles
- 2. Output TFLs titles with outline levels and field codes

After exploring and practicing multiple approaches, the final implementation plan is as follows:

- 1. Output chapter titles with outline levels (SAS) and apply styles (Word)
- 2. Output TFLs titles with outline levels (SAS) and field codes (SAS)

The previous code from the “title” statement has been adjusted to the “pretext” within the proc report statement, and the optimized result is as follows:

```

/*Blank is a temp dataset*/
data blank;
    blank = "";
run;
ods results off;
ods rtf file="C:\Users\Anti\Desktop\class.rtf" style=ThreeLines startpage=no;
*Section title;
proc report data=blank
    style(report)={pretext="\b\q\outlinelevel2 初二年级"};
    /*Use \outlinelevelN to create muti-level bookmark , the level followed Shell Section title*/
    column blank;
    define blank / noprint;
run;
*Table title and body;
proc report data=sashelp.class
    style(report)={pretext="\outlinelevel4 表 {field{\fldinst{STYLEREf t3 \s }}}.{\field{\fldinst{ SEQ 表
\\* ARABIC\s 1 }}} 班级信息"};
    /*For the table , default outline level is 5 class*/
    column sex name age height weight;
run;
ods rtf close;
ods results on;

```

And Output 2 shows this example in MS Word.



性别	姓名	年龄	身高 (英寸)	体重 (磅)
男	阿尔弗雷德	14	69	112.5
女	爱丽丝	13	56.5	84
女	芭芭拉	13	65.3	98
女	凯露	14	62.8	102.5
男	亨利	14	63.5	102.5

Output 2. The output document with Field Codes and Outline Levels via SAS code above

Complete the configuration of chapter title "t3" style in Word, and after refreshing the field codes throughout the document, the document updates entirely:

标题 页面 结果

14.1.1. 初二年级

表 14.1.1.1 班级信息

14.1.1.初二年级

分节符(连续)

表 14.1.1.1 班级信息

性别	姓名	年龄	身高 (英寸)	体重 (磅)
男	阿尔弗雷德	14	69	112.5
女	爱丽丝	13	56.5	84
女	芭芭拉	13	65.3	98
女	凯露	14	62.8	102.5
男	亨利	14	63.5	102.5

Figure 5. The finally output document

There are two noteworthy points in the above code:

1. In the first proc report statement, a temporary dataset with only a blank variable was read in, and "noprint" was declared in the output. The RTF control words "\b", "\ql", and "\outlinelevel2" were used to achieve bold, left alignment, and add a 3-level outline level, respectively. The final presentation effect is that only text is output without a table, which corresponds to chapter titles in Shell;
2. In the above code, the RTF control words within the *pretext (and posttext)* parameter will take effect directly, so there is no need to write a preceding escape character (even though it is already defined before the proc report). However, this situation also leads to special characters being garbled in UTF-8 Unicode environment.

ENCODING ISSUES

As previously discussed, most studies use UTF-8 Unicode environments, where pretext statements experience text corruption due to double encoding by the SAS system. After a lot of testing, confirmed that converting text to decimal Unicode encoding (e.g., &#XXXXX) resolves the issue.

查看字符编码 (UTF-8)

请输入文字 (可以输入多个): 班级信息

查看编码

如: "春眠不觉晓, 处处闻啼鸟。"

请输入编码 (只可输入一个):

查看文字

如: "6HB4BA", "65"

字符	编码10进制	编码16进制	Unicode编码10进制	Unicode编码16进制
班	15175597	E78FAD	29677	73ED
级	15186599	E7BAA7	32423	7EA7
信	14991265	E4BFA1	20449	4FE1
息	15106479	E681AF	24687	606F

Figure 6. Characters in different encoding environment

The re-coding operation can be performed using the UNICODDEC function in SAS, but the encoded text needs to be further adjusted to RTF control word format, which mean replacing "&#" with "\u".

Syntax

STR=UNICODDEC(<instr> (<Unicode type>))

Required Argument

str
data string that has been converted to Unicode encoding.

Optional Arguments

instr
input data string.

Unicode type
Unicode character formats

ESC Unicode Escape (for example, \u0042) ESC is the default format.
NCR Numeric Character Representation (for example, 大 or ± ;)
PAREN Unicode Parenthesis Escape (for example, <u0061>)

Str	UTF	RTF
班级信息	班级信息	\u29677;\u32423;\u20449;\u24687;

Figure 7. The Syntax and Example about UNICODDEC function in SAS

Additionally, this way can also solve special characters encoding issues with SAS statements in Unicode environment, such as using Chinese characters path in the "pipe" statement (in this case, needn't adjust to RTF control word format).

POST-PROCESSING FOR OUTPUT DOCUMENTS

There are some post-processing for TFLs.

Concatenate Individual RTF Files

After generating all TFL files (.rtf or .doc), combine them into a single document or section-based files as required. (Various concatenate methods exist; details omitted.)

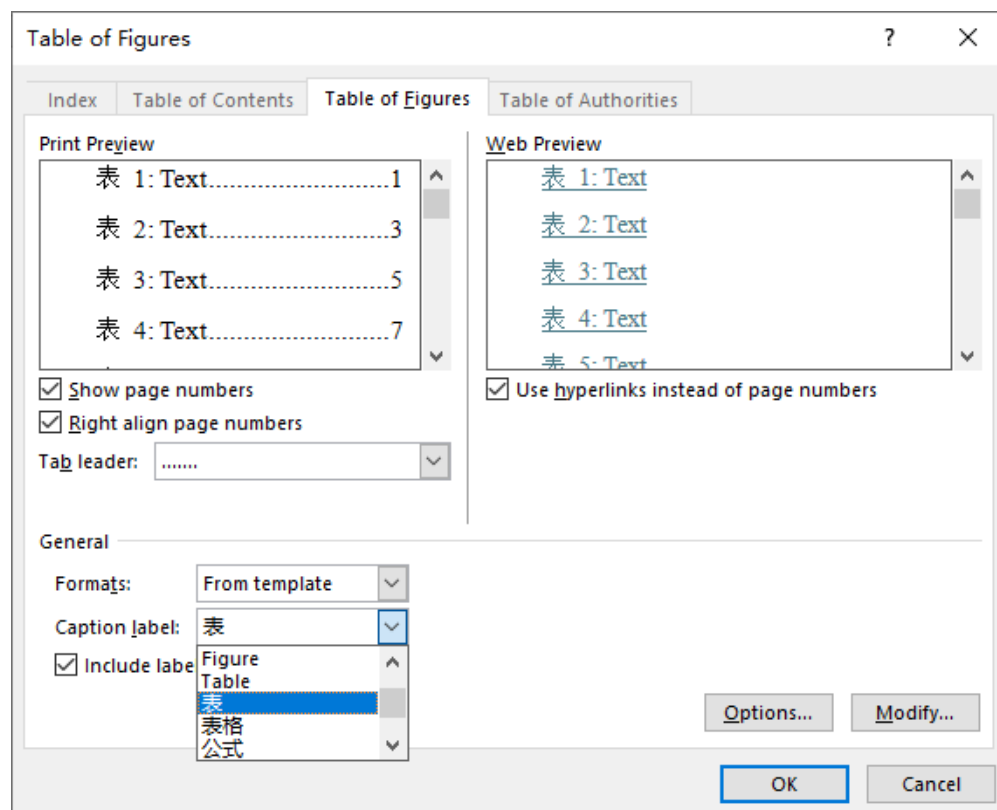
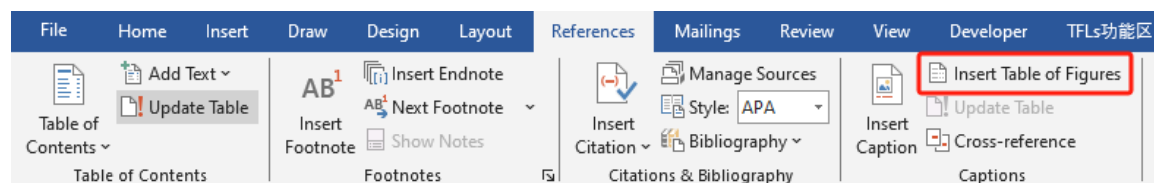
Then add predefined heading styles (e.g., "t1", "t2" or "Heading 1", "Heading 2") to chapter titles across the entire document, and update all field codes to ensure numbering consistency. (And you could remove section breaks if needed.)

The document is now fully processed and ready for review, following your habits or standard procedures for reviewing proposals, just like reviewing SAP, Shell, or CSR files.

Hyper-Linked Table of Contents

For document table/figure/listing directories: Since field codes are already embedded in table titles during output, you can directly use Word's **Insert Table of Figures** feature (under the **References** tab) to automatically extract TOC for figures, tables, and listings based on the identifiers in the field codes.

This eliminates the need for additional coding complexity, and, if there are some TFLs added in or deleted , what you need to do is only to select all story and Press the F9 button.



表目录

表 14.1.1 受试者分布	2
表 14.1.2 入组受试者分布(ITT)	3
表 14.2.1.1 人口学资料连续/分类指标描述(ITT)	5
表 14.2.2.1 用药依从性情况汇总表(ITT)	6
表 14.3.1.1 贝伐珠单抗药物暴露情况汇总表(SS)	7
表 14.3.1.2 厄洛替尼各剂量累积暴露时间(SS)	9
表 14.3.2.1 治疗期不良事件汇总表(SS)	10
表 14.3.3.1 按MedDRA编码的SOC和PT分类汇总治疗期不良事件(SS)	12
表 14.3.3.2 按MedDRA编码的SOC和PT及严重程度分类汇总治疗期不良事件(SS)	26
表 14.3.3.3 按ATC04和PN分类汇总既往合并用药(SS)	59
表 14.3.3.4 按MedDRA编码的PT分类汇总治疗期间常见不良事件(SS)	100
表 14.3.3.5 按MedDRA编码的PT分类汇总治疗期间常见不良事件(SS)	102
表 14.3.3.6 按MedDRA编码的PT及严重程度分类汇总治疗期间常见不良事件(SS)	104
表 14.3.4.1 生命体征治疗后相对基线的历时性变化(SS)	107
表 14.3.5.1 血常规检查治疗后相对基线的正异常变化(SS)-2×2	131
表 14.3.5.2 血常规检查治疗后相对基线的正异常变化(SS)-4×4	161
表 14.3.5.3 血常规检查治疗后相对基线的正异常变化(SS)-5×5	251

Figure 8. Create Table of Contents by Caption Label

Convert Combined RTF File to Other Formats

The final submitted document is in PDF format, which can be converted from Word using the “Save As” function or other methods. When converting, check “Create bookmarks using” and “Headings” to simultaneously create multi-level bookmarks in the PDF document, without the need for additional or complex code operations.

(Well, sometimes this way will spend lots of hours convert to PDF when document is largely.).

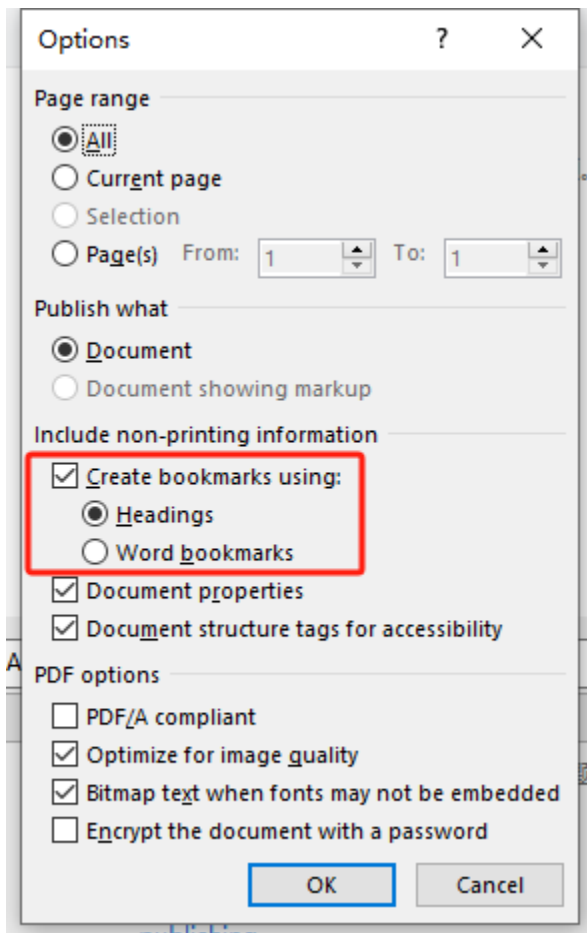


Figure 9. Options when Save as PDF

CONCLUSION

This article provides the best practice for using RTF control words in SAS to insert field codes when outputting TFLs, and introduces how to solve the problem of special character garble; using this method, Table of Contents and multi-level bookmarks in PDF documents can be generated in the document post-processing without additional or complex operations.

Ultimately making the entire document have good integrity and convenience; the document is adaptable to various subsequent application scenarios, such as splitting the forms of baseline, efficacy, safety, and other parts, and handing the entire document to Artificial Intelligence (AI) to let it complete report summaries and analysis according to the overall or each chapter.

REFERENCES

List of field codes in Word, Microsoft Word

Advanced RTF Layout Based on SAS Software, Katja Glaß, Bayer Schering Pharma AG, Berlin, GER
UNICODEC Function, Dictionary of Functions for NLS, SAS 9.4

CONTACT INFORMATION

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

Ryan Wang
WuHan Zhi Zhi Medicine Technology Co., Ltd.
+86 18507131519
1181190880@qq.com