

A Standard Data Set Format for the Validation of Tables and Listings in Regulatory Submissions for Clinical Trials

Kevin R. Viel, Ph.D., kevin.viel@navitaslifesciences.com, Navitas Data Sciences

- A standardize **VALDS** (VAlidity DAta SEt) results in **fewer changes** between TFL (Tables, Figure, Listings) programs and between deliveries/studies
- **New Resources** do not need to learn as many of the team conventions to be immediately productive.
- The shell (**Figure 1**), often included in the SAP, is a “blueprint” for the programmers.
- The shell defines the format and structure of the output.
- The programmers *derive the values* in the cells of the shell.
- A cell is *indexed* by Page, Section, Row, and Column (PSRC).
- The labels of the PSRC link to their order variable values (**Table 1**).
- The VALDS (**Figure 3**) looks like the TLF.

Table 14.1.X Demographics
Safety Population

Characteristic	Arm 1 (N=XXX)	Arm 2 (N=XXX)	Total (N=XXX)
Sex, n (%)			
Female	xx (xx.x)	xx (xx.x)	xx (xx.x)
Male	xx (xx.x)	xx (xx.x)	xx (xx.x)
Missing	xx (xx.x)	xx (xx.x)	xx (xx.x)
Age (years)			
n	xx	xx	xx
Mean (SD)	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)
Median	xx.x	xx.x	xx.x
Min, Max	xx, xx	xx, xx	xx, xx
Missing (%)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Age Group, n (%)			
< 65 years	xx (xx.x)	xx (xx.x)	xx (xx.x)
>= 65 years	xx (xx.x)	xx (xx.x)	xx (xx.x)
Missing	xx (xx.x)	xx (xx.x)	xx (xx.x)
Race, n (%)			
White	xx (xx.x)	xx (xx.x)	xx (xx.x)
Black or African American	xx (xx.x)	xx (xx.x)	xx (xx.x)
Asian	xx (xx.x)	xx (xx.x)	xx (xx.x)
American Indian or Alaska Native	xx (xx.x)	xx (xx.x)	xx (xx.x)
Native Hawaiian or Other Pacific Islander	xx (xx.x)	xx (xx.x)	xx (xx.x)
Other	xx (xx.x)	xx (xx.x)	xx (xx.x)
Missing	xx (xx.x)	xx (xx.x)	xx (xx.x)

Figure 1. Example shell for TFL

Table 14.1.X Demographics
Safety Population (Page_1)

Section_1	Section_Order_1	Row_1	Row_order_1	Col_1	Col_2	Col_3	Type
		Characteristic		Arm 1 (N=XXX)	Arm 2 (N=XXX)	Total (N=XXX)	H
Sex, n (%)	1	Sex, n (%)	1				B
Sex, n (%)	1	Female	2	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Sex, n (%)	1	Male	3	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Sex, n (%)	1	Missing	4	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Age (years)	2	Age (years)	1				B
Age (years)	2	n	2	xx	xx	xx	B
Age (years)	2	Mean (SD)	3	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)	B
Age (years)	2	Median	4	xx.x	xx.x	xx.x	B
Age (years)	2	Min, Max	5	xx, xx	xx, xx	xx, xx	B
Age (years)	2	Missing (%)	6	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Age Group, n (%)	3	Age Group, n (%)	1				B
Age Group, n (%)	3	< 65 years	2	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Age Group, n (%)	3	>= 65 years	3	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Age Group, n (%)	3	Missing	4	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Race, n (%)	4	Race, n (%)	1				B
Race, n (%)	4	White	2	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Race, n (%)	4	Black or African American	3	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Race, n (%)	4	Asian	4	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Race, n (%)	4	American Indian or Alaska Native	5	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Race, n (%)	4	Native Hawaiian or Other Pacific Islander	6	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Race, n (%)	4	Other	7	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Race, n (%)	4	Missing	8	xx (xx.x)	xx (xx.x)	xx (xx.x)	B

Figure 2. The VALDS annotated example shell

Table 1. VALDS Variables

Variable Name	Type	Codelist/Controlled Terms	Core	Notes
PAGE_y	Char		Perm	Higher level PAGE variables are nested in the lower level PAGE variables.
PAGE_ORDER_y	Num	Integer	Cond	Key variable(s) for the dataset.
SECTION_y	Char		Perm	Nested in the PAGE dimension. Higher level SECTION variables are nested in the lower level SECTION variables.
SECTION_ORDER_y	Num	Integer	Cond	Key variable(s) for the dataset.
ROW_y	Char		Req	Nested in the SECTION dimension. Higher level ROW variables are nested in the lower level ROW variables.
ROW_ORDER_y	Num	Integer	Req	Key variable(s) for the dataset.
COL_y	Char		Req	Nested in the ROW dimension.
TYPE	Char	B,H,F	Req	Indicator of whether the observation pertains to the Body, Header, or Footnote

Page_1	Page_Order_1	Section_1	Section_Order_1	Row_1	Row_order_1	Col_1	Col_2	Col_3	Type
Safety Population	1			Characteristic		Arm 1 (N=XXX)	Arm 2 (N=XXX)	Total (N=XXX)	H
Safety Population	1	Sex, n (%)	1	Sex, n (%)	1				B
Safety Population	1	Sex, n (%)	1	Female	2	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Safety Population	1	Sex, n (%)	1	Male	3	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Safety Population	1	Sex, n (%)	1	Missing	4	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Safety Population	1	Age (years)	2	Age (years)	1				B
Safety Population	1	Age (years)	2	n	2	xx	xx	xx	B
Safety Population	1	Age (years)	2	Mean (SD)	3	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)	B
Safety Population	1	Age (years)	2	Median	4	xx.x	xx.x	xx.x	B
Safety Population	1	Age (years)	2	Min, Max	5	xx, xx	xx, xx	xx, xx	B
Safety Population	1	Age (years)	2	Missing (%)	6	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Safety Population	1	Age Group, n (%)	3	Age Group, n (%)	1				B
Safety Population	1	Age Group, n (%)	3	< 65 years	2	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Safety Population	1	Age Group, n (%)	3	>= 65 years	3	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Safety Population	1	Age Group, n (%)	3	Missing	4	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Safety Population	1	Race, n (%)	4	Race, n (%)	1				B
Safety Population	1	Race, n (%)	4	White	2	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Safety Population	1	Race, n (%)	4	Black or African American	3	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Safety Population	1	Race, n (%)	4	Asian	4	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Safety Population	1	Race, n (%)	4	American Indian or Alaska Native	5	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Safety Population	1	Race, n (%)	4	Native Hawaiian or Other Pacific Islander	6	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Safety Population	1	Race, n (%)	4	Other	7	xx (xx.x)	xx (xx.x)	xx (xx.x)	B
Safety Population	1	Race, n (%)	4	Missing	8	xx (xx.x)	xx (xx.x)	xx (xx.x)	B

Figure 3. The standardized VALDS for the shell



- Example (**Figure 2**): the Section label of the ADaM variable SEX is “Sex n(%)” (**SECTION_1**) and its value “F” has the Row label “Female” (**ROW_1**)
- By default, all VALDS character variables have a length of 200 bytes, the maximum CDISC length.

