Paper CT16

Decimal calculation on BDS parameter

Rupal Rani Bhanja, Eli Lilly, Bangalore, India

ABSTRACT

Decimal calculation on BDS parameter. This paper describes the methodology used to determine and standardize decimal precision for lab parameters. The objective is to ensure consistency and accuracy in the presentation of summary statistics.

INTRODUCTION

The presentation titled Decimal calculation on BDS parameter focuses on the importance of accurate decimal precision in BDS datasets. Its primary purpose is to highlight how maintaining precise decimal values is essential for:

- **Data Integrity:** Ensuring that the dataset remains reliable and trustworthy for analysis.
- **Reproducible Analysis:** Supporting consistent results when analyses are repeated, which is critical for scientific and regulatory standards.
- **Regulatory Expectations:** Meeting the requirements set by regulatory bodies, which often demand specific levels of decimal accuracy.

The scope of the presentation covers how different parameters within the dataset may require varying levels of decimal accuracy. This depends on the nature of each parameter, the sensitivity of its measurement, and its intended use in statistical outputs. The presentation examines practical approaches to decimal calculations within basic dataset structures, providing guidance on best practices for statistical programmers and analysts working with BDS datasets

DECIMAL PRECISION HANDLING IN LB DATASETS.

The analysis focuses on a specific subject, laboratory category, and Lab test code (LBTESTCD). Laboratory measurements from multiple time points, including screening, visit are evaluated. Across these visits, the same subject has lab results reported with varying decimal precision, such as values recorded with two decimal places, three decimal places and one decimal place or no decimal places.

This variability demonstrates that, within a single LBTESTCD, different levels of decimal precision can occur. To address this inconsistency, a predefined rule is applied to align decimal precision across all relevant observations for that Lab test. The selected precision is then consistently used for calculation and presentation of summary statistics.

This paper discusses the statistical approaches outlined in SAP.

- Calculated means and medians will be tabulated to one or more decimal places than the original data.
- Minimum and maximum values will be tabulated to the same number of decimal places as the original data.
- Standard deviation will be tabulated to two or more decimal places than the original.

USUBJID	LBCAT	LBTESTCD	VISIT	LBDTC	LBORRES
101	HEMATOLOGY	WBC	SCREENING	2024-01-01T09...	9
101	HEMATOLOGY	WBC	DAY1	2024-01-03T09...	9.1
101	HEMATOLOGY	WBC	WEEK1	2024-01-08T09...	3.63
101	HEMATOLOGY	WBC	WEEK2	2024-01-15T09...	4.13
101	HEMATOLOGY	WBC	WEEK4	2024-01-29T09...	4.223
101	HEMATOLOGY	BASO	SCREENING	2024-01-01T09...	1.22
101	HEMATOLOGY	BASO	DAY1	2024-01-03T09...	1.1
101	HEMATOLOGY	BASO	WEEK1	2024-01-08T09...	1
101	HEMATOLOGY	BASO	WEEK2	2024-01-15T09...	1.11
101	HEMATOLOGY	BASO	WEEK4	2024-01-29T09...	1
102	HEMATOLOGY	WBC	SCREENING	2024-01-01T09...	6.02
102	HEMATOLOGY	WBC	DAY1	2024-01-03T09...	3.6
102	HEMATOLOGY	WBC	WEEK1	2024-01-08T09...	3.6
102	HEMATOLOGY	WBC	WEEK2	2024-01-15T09...	4
102	HEMATOLOGY	WBC	WEEK4	2024-01-29T09...	4
102	HEMATOLOGY	BASO	SCREENING	2024-01-01T09...	1
102	HEMATOLOGY	BASO	DAY1	2024-01-03T09...	1
102	HEMATOLOGY	BASO	WEEK1	2024-01-08T09...	1
102	HEMATOLOGY	BASO	WEEK2	2024-01-15T09...	1
102	HEMATOLOGY	BASO	WEEK4	2024-01-29T09...	1
101	CHEMISTRY	ALT	SCREENING	2024-01-01T09...	19
101	CHEMISTRY	ALT	DAY1	2024-01-03T09...	12
101	CHEMISTRY	ALT	WEEK1	2024-01-08T09...	15
101	CHEMISTRY	ALT	WEEK2	2024-01-15T09...	14
101	CHEMISTRY	ALT	WEEK4	2024-01-29T09...	19

Step 1:
Find the maximum no. of places past the decimal for each parameter.

```
data lb1;
set lb0;
length cval $200.;
if LBORRES > .z then do;
     cval = trim(left(compress(put(LBORRES, best.))));
end;
if int(LBORRES) = LBORRES then decimal = 0;
     else if indexc(cval, '.') >= 0 then do;
     decimal = length(trim(left(scan(cval, 2, '.'))));
end;
run;
```

This process ensured that the recorded precision of each data point was captured accurately.

USUBJID	LBCAT	LBTESTCD	VISIT	LBDTC	LBORRES	cval	decimal
101	HEMATOLOGY	WBC	SCREENING	2024-01-01T09...	9.1	9.1	1
101	HEMATOLOGY	WBC	DAY1	2024-01-03T09...	3.62	3.62	2
101	HEMATOLOGY	WBC	WEEK1	2024-01-08T09...	4.13	4.13	2
101	HEMATOLOGY	WBC	WEEK2	2024-01-15T09...	4.223	4.223	3
101	HEMATOLOGY	WBC	WEEK4	2024-01-29T09...	6.02	6.02	2
101	HEMATOLOGY	BASO	SCREENING	2024-01-01T09...	1.22	1.22	2
101	HEMATOLOGY	BASO	DAY1	2024-01-03T09...	1.1	1.1	1
101	HEMATOLOGY	BASO	WEEK1	2024-01-08T09...	1	1	0
101	HEMATOLOGY	BASO	WEEK2	2024-01-15T09...	1.11	1.11	2
101	HEMATOLOGY	BASO	WEEK4	2024-01-29T09...	1	1	0
102	HEMATOLOGY	WBC	SCREENING	2024-01-01T09...	6.02	6.02	2
102	HEMATOLOGY	WBC	DAY1	2024-01-03T09...	3.6	3.6	1
102	HEMATOLOGY	WBC	WEEK1	2024-01-08T09...	3.6	3.6	1
102	HEMATOLOGY	WBC	WEEK2	2024-01-15T09...	4	4	0
102	HEMATOLOGY	WBC	WEEK4	2024-01-29T09...	4	4	0
102	HEMATOLOGY	BASO	SCREENING	2024-01-01T09...	1	1	0
102	HEMATOLOGY	BASO	DAY1	2024-01-03T09...	1	1	0
102	HEMATOLOGY	BASO	WEEK1	2024-01-08T09...	1	1	0
102	HEMATOLOGY	BASO	WEEK2	2024-01-15T09...	1	1	0
102	HEMATOLOGY	BASO	WEEK4	2024-01-29T09...	1	1	0
101	CHEMISTRY	ALT	SCREENING	2024-01-01T09...	19	19	0
101	CHEMISTRY	ALT	DAY1	2024-01-03T09...	12	12	0
101	CHEMISTRY	ALT	WEEK1	2024-01-08T09...	15	15	0
101	CHEMISTRY	ALT	WEEK2	2024-01-15T09...	14	14	0
101	CHEMISTRY	ALT	WEEK4	2024-01-29T09...	19	19	0

Step 2-
Identify the maximum precision across all parameters.

/* In this step the dataset is sorted by the parameter (LBTESTCD) and PROC UNIVARIATE to get the maximum decimal count per parameter */

```
proc sort data=lb1;
  by LBTESTCD decimal;
run;

proc univariate data=lb1 noprint;
  by LBTESTCD;
  var decimal;
  output out=lb2 max=decimal;
```

```
run;
```

LBTESTCD	decimal
ALT	0
BASO	2
WBC	3

Step 3-
Assessing Overall Length and Decimal Detail.

In this step “decimal” represents the number of decimal places required by a lab value.
form1 calculates the total display width for numeric formatting.
form2 simply retains the decimal precision.

```
data lb3;
 set lb2;
  /* form1 = 8 + (1 if decimal > 0) + decimal */
  form1 = 8 + (decimal > 0) + decimal;
  /* form2 = just the decimal value */
  form2 = decimal;
run;

/* View the result */
proc print data=lb3 noobs;
run;
```

LBTESTCD	decimal	form1	form2
ALT	0	8	0
BASO	2	11	2
WBC	3	12	3

Step 4-
Count +1 automatically initializes count to 0 and increments it by 1 for each record.
Labnum produces value like 1_1, 1_2, 1_3 to create a stable, ordered identifier for each lab parameter.

Call symput creates a macro variable named 1_1, 1_2 etc to store the corresponding LBTESTCD value.

```
data LB4;
  length labnum $200;
  set LB3 end=eof;

  /* Initialize and increment count */
  count + 1;

  *** RECORD-COUNT LABEL TO EACH PARAMETER NAME ***;
  labnum = "l_" || compress(count);
  call symput(labnum, trim(left(LBTESTCD)));

  *** TOTAL NUMBER OF PARAMETERS INTO A MACRO VARIABLE ***;
  if eof then call symput('toptest', compress(count));
run;
```

labnum	LBTESTCD	decimal	form1	form2	count
l_1	ALT	0	8	0	1
l_2	BASO	2	11	2	2
l_3	WBC	3	12	3	3

Step 5
The step below repeats the same logic ensuring consistent parameter numbering across datasets and create macro variable mapping for numeric formats (forward and backward references) based on previously calculated numeric display attributes.
totfnum identifies a numeric format width
backnum identifies the number of decimal places.

```
data LB5;
  length totfnum backnum $10;
  set LB4 end=eof;
       totfnum = "f_" || compress(count);
       backnum = "b_" || compress(count);
  call symput(totfnum, trim(left(form1)));
  call symput(backnum, trim(left(form2)));
run;
```

totfnum	backnum	labnum	LBTESTCD	decimal	form1	form2	count
f_1	b_1	l_1	ALT	0	8	0	1
f_2	b_2	l_2	BASO	2	11	2	2
f_3	b_3	l_3	WBC	3	12	3	3

```
proc sort data=lb0;by lbcat lbtestcd visit;
run;
```

Step 6
To calculate descriptive statistics.

```
proc univariate data=lb0 noprint;
  by lbcat LBTESTCD visit;
  var LBORRES;
  output out=LB6 n=n
  mean=mean
  std=std
  median=median
  min=min
  max=max;
run;
```

LBCAT	LBTESTCD	VISIT	n	mean	std	max	median	min
CHEMISTRY	ALT	DAY1	1	12	.	12	12	12
CHEMISTRY	ALT	SCREENING	1	19	.	19	19	19
CHEMISTRY	ALT	WEEK1	1	15	.	15	15	15
CHEMISTRY	ALT	WEEK2	1	14	.	14	14	14
CHEMISTRY	ALT	WEEK4	1	19	.	19	19	19
HEMATOLOGY	BASO	DAY1	2	1.05	0.0707106781	1.1	1.05	1
HEMATOLOGY	BASO	SCREENING	2	1.11	0.1555634919	1.22	1.11	1
HEMATOLOGY	BASO	WEEK1	2	1	0	1	1	1
HEMATOLOGY	BASO	WEEK2	2	1.055	0.0777817459	1.11	1.055	1
HEMATOLOGY	BASO	WEEK4	2	1	0	1	1	1
HEMATOLOGY	WBC	DAY1	2	3.61	0.0141421356	3.62	3.61	3.6
HEMATOLOGY	WBC	SCREENING	2	7.56	2.1778888861	9.1	7.56	6.02
HEMATOLOGY	WBC	WEEK1	2	3.865	0.374766594	4.13	3.865	3.6
HEMATOLOGY	WBC	WEEK2	2	4.1115	0.1576848122	4.223	4.1115	4
HEMATOLOGY	WBC	WEEK4	2	5.01	1.428355698	6.02	5.01	4

Step 7

```
proc sql noprint;
    select distinct LBTESTCD
    into :code_1-, :code_9999
    from lb6;
```

```
quit;

%macro lab();

data labs_out;
  length _col $50 _col2 $10;
  set lb6;

  /* Assuming these variables exist in LABS */
  array values (6) n mean std median min max;

  do stat = 1 to 6;
    %do w = 1 %to &toptest;

      /* Replace "code" with the base name of your macro variable family */
      if strip(LBTESTCD) = "&&code_&w" then do;

        select (stat);

          when (1) do; /* N */
            %if %eval(&&b_&w) = 0 %then %do;
              _col  = put(values(stat), %eval(&&f_&w - 1).);      /* integer */
            %end;
            %else %do;
              _col  = put(values(stat), %eval(&&f_&w - 1 - &&b_&w).); /* integer with
reduced width */
            %end;
            _col2 = "N";
          end;

          when (2) do; /* Mean */
            _col  = put(mean,   %eval(&&f_&w + 1).%eval(&&b_&w + 1));
            _col2 = "Mean";
          end;

          when (3) do; /* SD */
            _col  = put(std,    %eval(&&f_&w + 2).%eval(&&b_&w + 2));
            _col2 = "SD";
          end;

          when (4) do; /* Median */
            _col  = put(median, %eval(&&f_&w + 1).%eval(&&b_&w + 1));
            _col2 = "Median";
          end;

          when (5) do; /* Min */
            %if %eval(&&b_&w) = 0 %then %do;
              _col = put(min, %eval(&&f_&w - 1).);                 /* integer */
            %end;
            %else %do;
              _col = put(min, %eval(&&f_&w).%eval(&&b_&w));        /* decimals */
            %end;
            _col2 = "Min";
          end;

          when (6) do; /* Max */
            %if %eval(&&b_&w) = 0 %then %do;
              _col = put(max, %eval(&&f_&w - 1).);                 /* integer */
            %end;
            %else %do;
              _col = put(max, %eval(&&f_&w).%eval(&&b_&w));        /* decimals */
            %end;
            _col2 = "Max";
          end;
```

```
          otherwise; /* do nothing */

        end; /* select */

        output;
      end; /* LBTESTCD match */

    %end; /* w-loop */
  end; /* stat-loop */
run;

%mend lab;
%lab;
```

Final dataset

LBCAT	LBTESTCD	VISIT	n	mean	std	max	median	min
HEMATOLOGY	BASO	WEEK4	2	1	0	1	1	1
HEMATOLOGY	WBC	DAY1	2	6.35	3.8890872965	9.1	6.35	3.6
HEMATOLOGY	WBC	DAY1	2	6.35	3.8890872965	9.1	6.35	3.6
HEMATOLOGY	WBC	DAY1	2	6.35	3.8890872965	9.1	6.35	3.6
HEMATOLOGY	WBC	DAY1	2	6.35	3.8890872965	9.1	6.35	3.6
HEMATOLOGY	WBC	DAY1	2	6.35	3.8890872965	9.1	6.35	3.6
HEMATOLOGY	WBC	DAY1	2	6.35	3.8890872965	9.1	6.35	3.6
HEMATOLOGY	WBC	SCREENING	2	7.51	2.1071782079	9	7.51	6.02
HEMATOLOGY	WBC	SCREENING	2	7.51	2.1071782079	9	7.51	6.02
HEMATOLOGY	WBC	SCREENING	2	7.51	2.1071782079	9	7.51	6.02
HEMATOLOGY	WBC	SCREENING	2	7.51	2.1071782079	9	7.51	6.02
HEMATOLOGY	WBC	SCREENING	2	7.51	2.1071782079	9	7.51	6.02
HEMATOLOGY	WBC	SCREENING	2	7.51	2.1071782079	9	7.51	6.02
HEMATOLOGY	WBC	WEEK1	2	3.615	0.0212132034	3.63	3.615	3.6
HEMATOLOGY	WBC	WEEK1	2	3.615	0.0212132034	3.63	3.615	3.6
HEMATOLOGY	WBC	WEEK1	2	3.615	0.0212132034	3.63	3.615	3.6
HEMATOLOGY	WBC	WEEK1	2	3.615	0.0212132034	3.63	3.615	3.6
HEMATOLOGY	WBC	WEEK1	2	3.615	0.0212132034	3.63	3.615	3.6
HEMATOLOGY	WBC	WEEK1	2	3.615	0.0212132034	3.63	3.615	3.6

CONCLUSION

Managing decimal precision in LB datasets with multiple parameters is essential for ensuring data accuracy, regulatory compliance, and meaningful statistical analysis. By systematically mapping each laboratory parameter to its appropriate decimal format and applying consistent rounding rules, you maintain the integrity and clarity of clinical trial results.

REFERENCES

Archive Page | PHUSE
D is for Dynamic, putting Dynamic back into CDISC
Different Decimal Places For Different Lab Tests
PHUSE APAC Connect 2026

ACKNOWLEDGMENTS

I want to thank Arwa Topiwalla Taher, who read the paper and made very good comments.

CONTACT INFORMATION

Author Name: Rupal Rani Bhanja
Company: Eli lilly
Address: Embassy tech village, Devarabeesanahalli, Bengaluru, Karnataka, India
Email: Bhanja_rupal@lilly.com