

## Building Traceability for End Points in Analysis Datasets Using SRCDOM, SRCVAR, and SRCSEQ Triplet

Xiangchen Cui, Vertex Pharmaceuticals Incorporated  
Tathabbai Pakalapati, Cytel Inc.  
Qunming Dong, Vertex Pharmaceuticals Incorporated

### ABSTRACT

To be compliant with ADaM Implementation Guide V1.0, traceability feature should be incorporated to possible extent in study analysis datasets. SRCDOM, SRCVAR, and SRCSEQ triplet are used to establish data point traceability in ADaM datasets. It can facilitate the transparency in FDA submission data, build confidence in analysis results, help efficient programming validation, speed up the overall review progress by FDA reviewers, and build a good relationship with FDA reviewers. This paper provides various examples of applying SRCDOM, SRCVAR, and SRCSEQ triplet to establish traceability in efficacy ADaM datasets from Cystic Fibrosis therapeutic area, and shows the art of applying the triplet to different scenarios.

### INTRODUCTION

SDTM DOMAIN variable value, the name of the SDTM source variable, and the relevant SDTM domain --SEQ value serve as primary candidates for establishing data point traceability through SRC--- triplet method. Using SRCDOM, SRCVAR and SRCSEQ triplet, one can link back to the source SDTM record(s) used to derive an analysis value.

Variable Name	Variable Label	Type	CDISC Notes
SRCDOM	Source Domain	Char	The 2-character identifier of the SDTM domain that relates to the derived analysis value
SRCVAR	Source Variable	Char	The name of the column (in the SDTM domain identified by SRCDOM) that relates to the derived analysis value
SRCSEQ	Source Sequence Number	Num	The sequence number SEQ of the row (in the SDTM domain identified by SRCDOM) that relates to the derived analysis value

**Table 1 Definition of SRCDOM, SRCVAR, and SRCSEQ Triplet**

SRC--- triplet in ADaM datasets aids the reviewer (Biostatistician, Agency Reviewer, QC Programmer) to trace back to the specific predecessor record(s) used to derive an analysis value. The definition of SRCDOM, SRCVAR, and SRCSEQ variables in Table 1 as defined ADaM IG V1.0 can be applied only if the derived analysis variable depends only on a single predecessor record and single variable from the source SDTM. But there might be situations where an analysis value is derived from multiple predecessor records from single variable of source SDTM. This paper provides examples of using SRC--- triplet to handle such situations by slightly modifying the usage of SRC--- triplet without losing the actual definition.

Sweat Chloride and Spirometry are often primary efficacy endpoints in a Cystic Fibrosis study. ADSW (Sweat Chloride Analysis Dataset) and ADSP (Spirometry Analysis Dataset) are used to provide examples of applying data point traceability using SRC--- triplet. Firstly, this paper tries to present the structure of an analysis dataset used to build the traceability. Secondly, the paper provides various examples of establishing data point traceability using SRC--- triplet.

### STRUCTURE OF ADSW

Typically, sweat chloride is collected from both left arm and right arm at every visit. The average of both left arm and right arm is often used as the analyses value at each analysis visit. Three important derivations in ADSW would be

- Baseline value
- Average Sweat Chloride (of both left and right arms) at every analysis visit
- Average on-treatment sweat chloride value (Example: Average through Day 28)

To establish proper traceability, the structure of ADSW will be as follows:

- Inherit all the SDTM SW records to ADSW to establish data point traceability as per ADaM implementation guide V1.0. Values of all analysis variables (such as AVAL, CHG, BASE, ABLFL, AVISIT etc.) will be set to missing for inherited records
- Derive an additional record for every nominal visit in SDTM SW with PARAMTYP="DERIVED" and DTYPE='AVERAGE' whose AVAL is equal to average of sweat chloride at left and right arms.

**USAGE OF SRCDOM, SRCVAR AND SRCSEQ TRIPLET IN ESTABLISHING TRACEABILITY FOR AVERAGE OF SWEAT CHLORIDE AT BOTH LEFT AND RIGHT ARMS AT EVERY ANALYSIS VISIT IN ADSW**

Sweat Chloride at an analysis visit (AVISIT) is an average of sweat chloride from both left and right arms collected at a nominal visit. Specification (metadata) of derived records for variables ASWSEQ, AVAL, SWDTC, SWENDTC, and for SRCDOM, SRCVAR, and SRCSEQ which helps in building data point traceability is shown in Table 2 followed by a sample SAS® code that populates these variables. Snapshot of these variables in an analysis dataset is shown in Display 1.

Variable Name	Variable Label	Type	Length	Comments
ASWSEQ	Analysis Sequence Number	Num	8	Equal to SW.SWSEQ for records inherited from SDTM SW. An average of SW_CL_L (left) and SW_CL_R (right) is derived for each time point in SDTM SW. The sequence number for derived records starts from 1001 and incremented by 1 successively.
AVAL	Analysis Value	Num	8	AVAL is equal to the average of SW_CL_L (left) and SW_CL_R (right) for each time point in SDTM SW. The AVAL will be rounded to 1 decimal. Note: There will be no derived records for never dosed subjects or for assessments at which the sample volume is <15 ul at both left and right arms
SWDTC	Start Date/Time of Sweat Collection	Char	20	Equal to SW.SWDTC for records inherited from SDTM SW. Equal to the earliest date/time among SW.SWDTC for sweat chloride tests SW_CL_L, VOLUME_L, SW_CL_R and VOLUME_R corresponding to the same visit for derived records.
SWENDTC	End Date/Time of Sweat Collection	Char	20	Equal to SW.SWENDTC for records inherited from SDTM SW. Equal to the latest date/time among SW.SWENDTC for sweat chloride tests SW_CL_L, VOLUME_L, SW_CL_R and VOLUME_R corresponding to the same visit for derived records.
SRCDOM	Source Domain	Char	4	Equal to "ADSW" for derived records (PARAMTYP="DERIVED") and AVISIT is equal to "Baseline" or "Average through Day 28". Equal to "SW" for all other derived records (PARAMTYP="DERIVED")
SRCVAR	Source Variable	Char	8	Equal to "AVAL" for derived records (PARAMTYP="DERIVED") where AVISIT is equal to "Baseline" or "Average through Day 28". Equal to "SWSTRESN" for all other derived records (PARAMTYP="DERIVED")
SRCSEQ	Source Sequence Number	Char	40	For derived records SRCSEQ lists the sequence numbers (SW.SWSEQ) of records used to derive the analysis value at each analysis visit. Example: if sequence numbers of SW_CL_L and SW_CL_R at DAY 7 are 4 and 5 respectively, then SRCSEQ is equal to 4\$5 for AVISIT=Day 7

**Table 2 Metadata of SRCDOM, SRCVAR, and SRCSEQ Establishing Traceability for Average Sweat Chloride**

```
data sw99;
length srcdom $4 srcvar $8 srcseq $40 swdtc swendtc $20;
set sw9;
sw_cl_lsdt=input(sdtc_sw_cl_l,??is8601dt.); /*Left Arm Start Date/Time*/
sw_cl_rsdt=input(sdtc_sw_cl_r,??is8601dt.); /*Right Arm Start Date/Time*/
sw_cl_ledt=input(edtc_sw_cl_l,??is8601dt.); /*Left Arm End Date/Time*/
sw_cl_redt=input(edtc_sw_cl_r,??is8601dt.); /*Right Arm End Date/Time*/
```

```

/*As per the SAP it is physiologically implausible to have sweat chloride value >160
mmol/L or <10 mmol/L*/
if sw_cl_l>160 then sw_cl_l=.;
else if .<sw_cl_l<10 then sw_cl_l=.;
if sw_cl_r>160 then sw_cl_r=.;
else if .<sw_cl_r<10 then sw_cl_r=.;

if volume_l>=15 and volume_r>=15 then do;
  if sw_cl_l ne . and sw_cl_r ne . then do;
    aval=(sw_cl_l+sw_cl_r)/2;
    srcdom="SW";
    srcvar="SWSTRESN";
    if seq_sw_cl_l<seq_sw_cl_r then
srcseq=strip(put(seq_sw_cl_l,best.)||"$"||strip(put(seq_sw_cl_r,best.)));
    else if seq_sw_cl_r<seq_sw_cl_l then
srcseq=strip(put(seq_sw_cl_r,best.)||"$"||strip(put(seq_sw_cl_l,best.)));
    swdtc=substr(put(min(sw_cl_lsdt,sw_cl_rsdt),is8601dt.),1,16);
    swendt=substr(put(max(sw_cl_ledt,sw_cl_redt),is8601dt.),1,16);
    adt=min(datepart(sw_cl_lsdt),datepart(sw_cl_rsdt));
    adtm=min(sw_cl_lsdt,sw_cl_rsdt);
    atm=timepart(adtm);
  end;
  else if sw_cl_l ne . and sw_cl_r eq . then do;
    aval=sw_cl_l;
    srcdom="SW";
    srcvar="SWSTRESN";
    srcseq=strip(put(seq_sw_cl_l,best.));
    swdtc=substr(put(sw_cl_lsdt,is8601dt.),1,16);
    swendt=substr(put(sw_cl_ledt,is8601dt.),1,16);
    adt=datepart(sw_cl_lsdt);
    adtm=sw_cl_lsdt;
    atm=timepart(adtm);
  end;
  else if sw_cl_l eq . and sw_cl_r ne . then do;
    aval=sw_cl_r;
    srcdom="SW";
    srcvar="SWSTRESN";
    srcseq=strip(put(seq_sw_cl_r,best.));
    swdtc=substr(put(sw_cl_rsdt,is8601dt.),1,16);
    swendt=substr(put(sw_cl_redt,is8601dt.),1,16);
    adt=datepart(sw_cl_rsdt);
    adtm=sw_cl_rsdt;
    atm=timepart(adtm);
  end;
end;
else if volume_l<15 and volume_r>=15 and sw_cl_r ne . then do;
  aval=sw_cl_r;
  srcdom="SW";
  srcvar="SWSTRESN";
  srcseq=strip(put(seq_sw_cl_r,best.));
  swdtc=substr(put(sw_cl_rsdt,is8601dt.),1,16);
  swendt=substr(put(sw_cl_redt,is8601dt.),1,16);
  adt=datepart(sw_cl_rsdt);
  adtm=sw_cl_rsdt;
  atm=timepart(adtm);
end;

```

```

else if volume_l>=15 and volume_r<15 and sw_cl_l ne . then do;
    aval=sw_cl_l;
    srcdom="SW";
    srcvar="SWSTRESN";
    srcseq=strip(put(seq_sw_cl_l,best.));
    swdtc=substr(put(sw_cl_lsdt,is8601dt.),1,16);
    swendtc=substr(put(sw_cl_ledt,is8601dt.),1,16);
    adt=datepart(sw_cl_lsdt);
    atm=sw_cl_lsdt;
    atn=timepart(atm);
end;
else if volume_l<15 and volume_r<15 then aval=.;
run;

```

ASWSEQ	PARAMCD	PARAMTYP	AVAL	DTYPE	AVISIT	ADTM	SRCDOM	SRCVAR	SRCSEQ	VISIT	SWORRES	SWDTC	SWENDTC
17	SW_CL_R		.			.				DAY 7	60	2012-08-22T13:58	2012-08-22T14:28
19	SW_CL_L		.			.				DAY 7		2012-08-22T14:50	2012-08-22T15:20
21	SW_CL_R		.			.				DAY 14		2012-08-29T14:04	2012-08-29T14:34
23	SW_CL_L		.			.				DAY 14	63	2012-08-29T14:14	2012-08-29T14:44
25	SW_CL_L		.			.				DAY 21	74	2012-09-04T14:02	2012-09-04T14:32
27	SW_CL_R		.			.				DAY 21		2012-09-04T14:10	2012-09-04T14:40
29	SW_CL_L		.			.				DAY 28	110	2012-09-12T08:00	2012-09-12T08:30
31	SW_CL_R		.			.				DAY 28	88	2012-09-12T08:07	2012-09-12T08:37

ASWSEQ	PARAMCD	PARAMTYP	AVAL	DTYPE	AVISIT	ADTM	SRCDOM	SRCVAR	SRCSEQ	VISIT	SWORRES	SWDTC	SWENDTC
1005	SW_CL_M	DERIVED	60	AVERAGE	Day 7	22AUG2012:13:58:00	SW	SWSTRESN	17	DAY 7	60	2012-08-22T13:58	2012-08-22T14:28
1006	SW_CL_M	DERIVED	63	AVERAGE	Day 14	29AUG2012:14:14:00	SW	SWSTRESN	23	DAY 14	63	2012-08-29T14:14	2012-08-29T14:44
1007	SW_CL_M	DERIVED	74	AVERAGE	Day 21	04SEP2012:14:02:00	SW	SWSTRESN	25	DAY 21	74	2012-09-04T14:02	2012-09-04T14:32
1008	SW_CL_M	DERIVED	99	AVERAGE	Day 28	12SEP2012:08:00:00	SW	SWSTRESN	29\$31	DAY 28	99	2012-09-12T08:00	2012-09-12T08:37

### Display 1 Snapshot of Average Sweat Chloride Records in an Analysis Dataset

The above example shows that SRCDOM, SRCVAR and SRCSEQ triplet builds a clear path from an ADaM derived record to its predecessor in source SDTM. PARAMTYP='DERIVED' tells the reviewer (Biostatistician and/or FDA reviewer) that these records are not inherited from SDTM instead derived in ADaM for analyses purpose. SRCDOM='SW' and SRCVAR='SWSTRESN' let the reviewer know that the records were derived using values in variable SWSTRESN in SDTM SW domain. SRCSEQ lists sequence number, separated by '\$', of those records in SDTM SW used in deriving AVAL. We can also populate DTYPE='AVERAGE' to let the reviewer know that AVAL is the average of SWSTRESN corresponding to records with sequence numbers listed in SRCSEQ. In the above scenario, SRCSEQ=17 for AVISIT='Day 7' suggests that only sweat chloride at right arm is used to populate AVAL as sweat chloride at left arm is missing. Similarly, SRCSEQ='29\$31' for AVISIT='Day 28' suggests that sweat chloride assessments in SDTM SW with sequence numbers 29 and 31, corresponding to left and right arms respectively, are used to derive AVAL. Note: For the derived analysis records (PARAMTYP='DERIVED') SWORRES is equal to AVAL. SWDTC is equal to minimum of Sweat Chloride Start Date/Time (SWDTC) corresponding to left and right arms. SWENDTC is equal to maximum of Sweat Chloride End Date/Time (SWENDTC) corresponding to left and right arms.

### USAGE OF SRCDOM, SRCVAR AND SRCSEQ TRIPLET IN ESTABLISHING TRACEABILITY FOR BASELINE ANALYSIS VISIT IN ADSW

Baseline sweat chloride is defined as the average of Screening and Day 1 predose values. As discussed earlier, an additional record for every nominal visit in SDTM SW is derived with PARAMTYP="DERIVED" whose AVAL is equal to average of sweat chloride at left and right arms. The sequence number for derived records starts from 1001 and incremented by one. Analysis visit 'Baseline' is then derived whose AVAL is equal to average of analysis values corresponding to analysis visits 'Screening' and 'Day 1' respectively. Specification (metadata) of derived records for SWSEQ, AVAL, at baseline and for SRCDOM, SRCVAR, and SRCSEQ which helps in building data point traceability is shown in Table 3 followed by a sample SAS code that populates these variables. Snapshot of these variables in an analysis dataset in shown in Display 2.

Variable Name	Variable Label	Type	Length	Comments
ASWSEQ	Analysis Sequence Number	Num	8	Average of records from the predose scheduled visits is derived and populated as the analysis baseline with ASWSEQ equal to 0.5 more than the sequence number of the last predose assessment.
AVAL	Analysis Value	Num	8	AVAL is equal to the average of the records from the predose scheduled visits for Baseline analysis visit

SRCDOM	Source Domain	Char	4	Equal to "ADSW" for derived records (PARAMTYP="DERIVED") and AVISIT is equal to "Baseline" or "Average through Day 28". Equal to "SW" for all other derived records (PARAMTYP="DERIVED")
SRCVAR	Source Variable	Char	8	Equal to "AVAL" for derived records (PARAMTYP="DERIVED") where AVISIT is equal to "Baseline" or "Average through Day 28". Equal to "SWSTRESN" for all other derived records (PARAMTYP="DERIVED")
SRCSEQ	Source Sequence Number	Char	40	For AVISIT="Baseline" SRCSEQ lists the sequence numbers (SW.SWSEQ) of records that are used to derive the analysis baseline value. Example: if sequence numbers of Screening and Day 1 are 1 and 2 respectively, then SRCSEQ is equal to 1\$2

**Table 3 Metadata for SWSEQ, AVAL at Baseline and for SRCDOM, SRCVAR, and SRCSEQ Establishing Data Point Traceability**

```

/*APHASE=0 means Pre-treatment assessments. VISITNUM=80001 means Unscheduled visits*/
/*Subsets and transposes all scheduled predose analysis values*/
proc transpose data=sw6(where=(aphasen=0 and visitnum ne 80001)) out=baseline;
  by usubjid;
  id avisitn;
  var aval;
run;
/*Subsets and transposes all sequence numbers corresponding to scheduled predose*/
proc transpose data=sw6(where=(aphasen=0 and visitnum ne 80001)) out=bsrcseq
  prefix=seq_;
  by usubjid;
  id avisitn;
  var swseq;
run;
/*Get the sequence number corresponding to last scheduled predose assessment to derive
SWSEQ for analysis Baseline visit*/
proc sort data=sw6 out=baseseq;
  by usubjid adt atm;
  where aphasen=0 and visitnum ne 80001;
run;
data baseseq;
  set baseseq;
  by usubjid adt atm;
  if last.usubjid;
  keep usubjid swseq swnam swcat;
run;
data baseline;
length srcdom $4 srcvar $8 srcseq aphase avisit $40 dtype $20;
merge baseline(in=in_baseline) bsrcseq(in=in_bsrcseq) baseseq(in=in_baseseq);
  by usubjid;
  if in_baseline;
  aswseq=swseq+0.5;
  dtype="AVERAGE";
  aphase="Pre-Treatment Phase";aphasen=0;
  avisit="Baseline";avisitn=30001.1;anl01fl="Y";
  array y{*} _;
  do l=1 to dim(y);
    if l=1 then do;sumbl=0;nbl=0;end;
    if y{l} ne . then do;sumbl=sumbl+y{l};nbl=nbl+1;end;
  end;
  aval=sumbl/nbl;
  srcdom="ADSW";srcvar="AVAL";

```

```

array z{*} seq_;;
do m=1 to dim(z);
  if m=1 and z{m} ne . then srcseq=strip(put(z{m},best.));
  else if z{m} ne . and srcseq^="" then
srcseq=strip(srcseq)||"$"||strip(put(z{m},best.));
  else if z{m} ne . and srcseq="" then srcseq=strip(put(z{m},best.));
end;
run;

```

ASWSEQ	PARAMCD	PARAMTYP	AVAL	DTYPE	AVISIT	ADTM	SRCDOM	SRCVAR	SRCSEQ	VISIT	SWSTRESN	SWTPT
1	SW_CL_L	.	.	.	.	.	.	.	.	SCREENING	67	.
3	SW_CL_R	.	.	.	.	.	.	.	.	SCREENING	86	.
5	SW_CL_L	.	.	.	.	.	.	.	.	DAY 1	79	PREDOSE
7	SW_CL_R	.	.	.	.	.	.	.	.	DAY 1	67	PREDOSE
9	SW_CL_L	.	.	.	.	.	.	.	.	DAY 1	88	2-4 HOURS POSTDOSE
11	SW_CL_R	.	.	.	.	.	.	.	.	DAY 1	52	2-4 HOURS POSTDOSE
13	SW_CL_R	.	.	.	.	.	.	.	.	DAY 1	71	24 HOURS POSTDOSE
15	SW_CL_L	.	.	.	.	.	.	.	.	DAY 1	87	24 HOURS POSTDOSE
1001	SW_CL_M	DERIVED	76.5	AVERAGE	Screening	08AUG2012:16:46:00	SW	SWSTRESN	1\$3	SCREENING	76.5	.
1002	SW_CL_M	DERIVED	73	AVERAGE	Day 1	15AUG2012:08:22:00	SW	SWSTRESN	5\$7	DAY 1	73	PREDOSE
1002.5	SW_CL_M	DERIVED	74.8	AVERAGE	Baseline	.	ADSW	AVAL	1001\$1002	.	74.8	.
1003	SW_CL_M	DERIVED	70	AVERAGE	Day 1, 2-4 hrs post dose	15AUG2012:11:55:00	SW	SWSTRESN	9\$11	DAY 1	70	2-4 HOURS POSTDOSE
1004	SW_CL_M	DERIVED	79	AVERAGE	Day 1, 24 hrs post dose	16AUG2012:08:47:00	SW	SWSTRESN	13\$15	DAY 1	79	24 HOURS POSTDOSE

**Display 2 Snapshot of Analysis Baseline Visit with Traceability Variables in Sweat Chloride Analysis Dataset**

In the above scenario, SRCDOM='ADSW', SRCVAR='AVAL' and SRCSEQ='1001\$1002' for AVISIT='Baseline' indicates that analysis value (AVAL) is derived, using the AVAL in ADSW, corresponding to sequence numbers 1001 and 1002. We can also populate DTYPE='AVERAGE' to let the reviewer know that AVAL is the average of AVAL corresponding to records with sequence numbers listed in SRCSEQ. Note: ASWSEQ for AVISIT='Baseline' is equal to 0.5 more than the sequence number corresponding to the last predose analysis visit. Also, AVISIT is populated only for those visits that need to be presented in analyses.

**USAGE OF SRCDOM, SRCVAR AND SRCSEQ TRIPLET IN ESTABLISHING TRACEABILITY FOR AVERAGE ON-TREATMENT SWEAT CHLORIDE IN ADSW**

Analysis value for Average On-treatment Sweat Chloride is defined as the average of on-treatment analysis values. Specification (metadata) of derived records for ASWSEQ, AVAL, for Average through Day 28 and for SRCDOM, SRCVAR, and SRCSEQ which helps in building data point traceability is shown in Table 4 followed by a sample SAS code that populates these variables. Snapshot of these variables in an analysis dataset in shown in Display 3.

Variable Name	Variable Label	Type	Length	Comments
ASWSEQ	Analysis Sequence Number	Num	8	Average of on-treatment analysis visits Day 7, Day 14, Day 21 and Day 28 is derived and populated as Average through Day 28 analysis visit with SWSEQ equal to 0.5 more than the sequence number of the last analysis visit on or before Day 28.
AVAL	Analysis Value	Num	8	AVAL is equal to average of the records on-treatment Day 7, Day 14, Day 21 and Day 28 analysis visits for the Average through Day 28 analysis visit. The AVAL will be rounded to 1 decimal
SRCDOM	Source Domain	Char	4	Equal to "ADSW" for derived records (PARAMTYP="DERIVED") and AVISIT is equal to "Baseline" or "Average through Day 28". Equal to "SW" for all other derived records (PARAMTYP="DERIVED")
SRCVAR	Source Variable	Char	8	Equal to "AVAL" for derived records (PARAMTYP="DERIVED") where AVISIT is equal to "Baseline" or "Average through Day 28". Equal to "SWSTRESN" for all other derived records (PARAMTYP="DERIVED")
SRCSEQ	Source Sequence Number	Char	40	For AVISIT="Average through Day 28" SRCSEQ lists the analysis visits (ADSW.AVISIT) of records that are used to derive the analysis value (average of Day 7, Day 14, Day 21, Day 28). Example: if Day 7, Day 14, Day 21, Day 28 are used to drive the average then SRCSEQ is equal to Day 7\$Day 14\$Day 21\$Day 28

**Table 4 Metadata of SWSEQ, AVAL for Average through Day 28 and of SRCDOM, SRCVAR, and SRCSEQ Triplet**

```

/*Subsets and transposes all scheduled on-treatment analysis values*/
proc transpose data=sw6(where=(aphasen>0 and avisitn in (30007,30014,30021,30028)))
out=avg28(drop=_name_);
  by usubjid;
  id visitnum;
  var aval;
run;
proc transpose data=sw6(where=(aphasen>0 and avisitn in (30007,30014,30021,30028)))
out=avg28srcseq(drop=_name_) prefix=seq_;
  by usubjid;
  id visitnum;
  var swseq;
run;
/*Sequence number corresponding to last scheduled on-treatment predose assessment
to derive SWSEQ for analysis Average through Day 28 visit*/
proc sort data=sw6 out=avg28seq;
  by usubjid adt atm;
  where aphasen>0 and avisitn in (30007,30014,30021,30028);
run;
data avg28seq;
  set avg28seq;
  by usubjid adt atm;
  if last.usubjid;
  keep usubjid swseq swnam swcat;
run;
data avg28;
length srcdom $4 srcvar $8 srcseq aphase avisit $40 dtype studyid $20;
merge avg28(in=in_avg28) avg28srcseq(in=in_avg28srcseq)
avg28seq(in=in_avg28seq);
by usubjid;
if in_avg28;

  studyid="&study_lbl.";
  aswseq=swseq+0.5;
  dtype="AVERAGE";
  aphase="On-Treatment Phase";aphasen=1;
  avisit="Average through Day 28";avisitn=30028.3;
  anl01fl="Y";

array a{*} _30007 _30014 _30021 _30028;
  do i=1 to dim(a);
    if i=1 then do;sum28=0;n28=0;end;
    if a{i} ne . then do;sum28=sum28+a{i};n28=n28+1;end;
  end;
  aval=sum28/n28;
  srcdom="ADSW";
  srcvar="AVAL";
  if seq_30007 ne . then srcseq="DAY 7";
  if seq_30014 ne . and srcseq^="" then srcseq=strip(srcseq)||"$DAY 14";
  else if seq_30014 ne . and srcseq="" then srcseq="DAY 14";
  if seq_30021 ne . and srcseq^="" then srcseq=strip(srcseq)||"$DAY 21";
  else if seq_30021 ne . and srcseq="" then srcseq="DAY 21";
  if seq_30028 ne . and srcseq^="" then srcseq=strip(srcseq)||"$DAY 28";
  else if seq_30028 ne . and srcseq="" then srcseq="DAY 28";
run;

```

SUBJID	ASWSEQ	PARAMCD	PARAMTYP	AVAL	DTYPE	AVISIT	ANL01FL	SRCDOM	SRCVAR	SRCSEQ	VISIT	SWSTRESN	SWTPT
999006	17	SW_CL_R		.							DAY 7	60	
999006	19	SW_CL_L		.							DAY 7	.	
999006	21	SW_CL_R		.							DAY 14	.	
999006	23	SW_CL_L		.							DAY 14	63	
999006	25	SW_CL_L		.							DAY 21	74	
999006	27	SW_CL_R		.							DAY 21	.	
999006	29	SW_CL_L		.							DAY 28	110	PREDOSE
999006	31	SW_CL_R		.							DAY 28	88	PREDOSE
999006	1005	SW_CL_M	DERIVED	60	AVERAGE	Day 7	Y	SW	SWSTRESN	17	DAY 7	60	
999006	1006	SW_CL_M	DERIVED	63	AVERAGE	Day 14	Y	SW	SWSTRESN	23	DAY 14	63	
999006	1007	SW_CL_M	DERIVED	74	AVERAGE	Day 21	Y	SW	SWSTRESN	25	DAY 21	74	
999006	1008	SW_CL_M	DERIVED	99	AVERAGE	Day 28	Y	SW	SWSTRESN	29\$31	DAY 28	99	PREDOSE
999006	1008.5	SW_CL_M	DERIVED	74	AVERAGE	Average through Day 28	Y	ADSW	AVAL	DAY 7\$DAY 14\$DAY 21\$DAY 28		74	
999007	17	SW_CL_R		.							DAY 7	95	
999007	19	SW_CL_L		.							DAY 7	50	
999007	21	SW_CL_R		.							DAY 14	89	
999007	23	SW_CL_L		.							DAY 14	.	
999007	25	SW_CL_R		.							DAY 21	.	
999007	27	SW_CL_L		.							DAY 21	.	
999007	29	SW_CL_R		.							DAY 28	.	PREDOSE
999007	31	SW_CL_L		.							DAY 28	.	PREDOSE
999007	1005	SW_CL_M	DERIVED	72.5	AVERAGE	Day 7	Y	SW	SWSTRESN	17\$19	DAY 7	72.5	
999007	1006	SW_CL_M	DERIVED	89	AVERAGE	Day 14	Y	SW	SWSTRESN	21	DAY 14	89	
999007	1006.5	SW_CL_M	DERIVED	80.8	AVERAGE	Average through Day 28	Y	ADSW	AVAL	DAY 7\$DAY 14		80.8	

**Display 3 Snapshot of Average through Day 28 Analysis Visit with Traceability Variables in Sweat Chloride Analysis Dataset**

In the above scenario, for subject 999006, SRCDOM='ADSW', SRCVAR='AVAL' and SRCSEQ='DAY 7\$DAY 14\$DAY 21\$DAY 28' for AVISIT='Average through Day 28' indicates that analysis value (AVAL) is derived, using the AVAL in ADSW, corresponding to analysis visits Day 7, Day 14, Day 21, and Day 28. DTYPE='AVERAGE' will let the reviewer know that AVAL is the average of AVAL corresponding to analysis visits listed in SRCSEQ. Importance of traceability will be more evident when all possible on-treatment values don't contribute while deriving average on-treatment sweat chloride. For subject 999007, SRCSEQ='DAY 7\$DAY 14' for AVISIT='Average through Day 28' indicates that only analysis values corresponding to analysis visits Day 7, and Day 14 were used to derive the AVAL as sweat chloride assessments on DAY 21 and DAY 28 were missing. Note: ASWSEQ for AVISIT='Average through Day 28' is equal to 0.5 more than the sequence number corresponding to the last analysis visit on or before Day 28 predose. Alternatively, sequence numbers of records used in deriving AVAL for Average through Day 28 can be listed in SRCSEQ but listing visits will be more readable and easy for a reviewer.

**STRUCTURE OF ADSP**

Typically, Spirometry SDTM has FEF (FORCED MIDEPIRATORY FLOW RATE (L/sec)), FEV (FORCED EXPIRATORY VOLUME IN 1 SECOND (L)), and FVC (FORCED VITAL CAPACITY (L)) raw values collected at every visit. Important derivations in ADSP would be

- Ratio of FEV1/FVC
- PFEF (PREDICTED FORCED MIDEPIRATORY FLOW RATE (L/sec)), PFEV (PREDICTED FORCED EXPIRATORY VOLUME IN 1 SECOND (L)), PFVC (PREDICTED FORCED VITAL CAPACITY (L)), and PFEVFVC (PREDICTED RATIO OF FEV1/FVC) at every analysis visit. In short trials with adult subjects age and height can be assumed constant and therefore the predict values stay the same at all visits and derived only at baseline; this is not true for longer trials and/or with younger subjects whose height can increase during the trial and hence predicted values are derived at every analysis visit.
- PPFEF (PERCENT PREDICTED FORCED MIDEPIRATORY FLOW RATE (%)), PPFEV (PERCENT PREDICTED FORCED EXPIRATORY VOLUME IN 1 SECOND (%)), PPFVC (PERCENT PREDICTED FORCED VITAL CAPACITY (%)), and PPFEVFVC (PERCENT PREDICTED FEV/FVC (%)) at every analysis visit
- Baseline value for FEF, FEV, FVC, Ratio of FEV1/FVC, PPFEF, PPFEV, PPFVC and PPFEVFVC
- Average on-treatment FEF, FEV, FVC, Ratio of FEV1/FVC, PPFEF, PPFEV, PPFVC, and PPFEVFVC values (Example: Average through Day 28)

To establish proper traceability, the structure of ADSP will be as follows:

- Inherit all the SDTM SP records to ADSP to establish data point traceability as per ADaM implementation guide V1.0.
- Derive an additional record with Ratio of FEV1/FVC for every nominal visit in SDTM SP



- Derive PFEF, PFEV, PFVC, and PFEVFC based on age and height only at baseline or at every analysis visit depending upon the population and length of a study
- Derive an additional record of PPFEF, PPFEV, PPFVC, and PPFEVFC for every nominal visit in SDTM SP
- Derive baseline value for FEF, FEV, FVC, Ratio of FEV1/FVC, PPFEF, PPFEV, PPFVC, and PPFEVFC
- Derive average on-treatment value for FEF, FEV, FVC, Ratio of FEV1/FVC, PPFEF, PPFEV, PPFVC, and PPFEVFC (Example: Average through Day 28)

**USAGE OF SRCDOM, SRCVAR AND SRCSEQ TRIPLET IN ESTABLISHING TRACEABILITY FOR RATIO OF FEV1/FVC AT EVERY ANALYSIS VISIT IN ADSP**

FEV and FVC raw values are collected and mapped to SDTM SP. A record with PARAMTYP="DERIVED" and PARAMCD="FEVFC" will be derived in ADSP for every nominal visit in SDTM SP for Ratio of FEV1/FVC. Specification (metadata) of derived Ratio of FEV1/FVC records for ASPSEQ, AVAL variables and for SRCDOM, SRCVAR, and SRCSEQ which helps in building data point traceability is shown in Table 5 followed by a sample SAS code that populates these variables. Snapshot of these variables in an analysis dataset is shown in Display 4.

Variable Name	Variable Label	Type	Length	Comments
ASPSEQ	Sequence Number	Num	8	Ratio of FEV/FVC is derived for each subject at each time point. These derived records are sorted byUSUBJID ADT ATM and are assigned sequence numbers starting from 301 and incremented by 1 for each subject.
AVAL	Analysis Value	Num	8	Equal to ratio of analysis values corresponding to parameter FEV and FVC at each time point for PARAMCD='FEVFC'. AVAL is rounded to suggested precision as per variable SPPREC, only at the end. All the derivations are done using analysis values before rounding to the suggested precision.
SRCDOM	Source Domain	Char	4	Equal to 'SP' for PARAMTYP='DERIVED' and PARAMCD is equal to 'FEVFC'
SRCVAR	Source Variable	Char	8	Equal to 'SPSTRESN' for PARAMTYP='DERIVED' and PARAMCD is equal to 'FEVFC'
SRCSEQ	Source Sequence Number	Char	40	SRCSEQ lists the sequence numbers of records, separated by '\$', used in deriving the analysis values for records with PARAMTYP='DERIVED'. Example: If sequence numbers of DAY 7 visit for a subject are 4 and 7 for parameters FEV and FVC respectively then SRCSEQ is equal to 4\$7 for PARAMCD equal to 'FEVFC' for analysis visit Day 7

**Table 5 Metadata of SWSEQ, AVAL for Ratio of FEV1/FVC and of SRCDOM, SRCVAR, and SRCSEQ Triplet**

```

/*Subset FEV and FVC assessments*/
proc sort data=spl out=fevfc;
by usubjid aphasen avisitn adt atm adtm ady visit spdtc anl01fl;
  where paramcd in ("FEV","FVC") and spstresn ne .;
run;

/*Transpose FEV and FVC values*/
proc transpose data=fevfc out=fevfc_val(drop=_name_ _label_);
  by usubjid aphasen avisitn adt atm adtm ady visit spdtc anl01fl;
  id paramcd;
  var spstresn;
run;

/*Transpose corresponding FEV and FVC sequence numbers for traceability purpose*/
proc transpose data=fevfc out=fevfc_seq(drop=_name_ _label_) prefix=seq;
  by usubjid aphasen avisitn adt atm adtm ady visit spdtc anl01fl;
  id paramcd;
  var spseq;
run;

```

```

data fevfvc_ratio;
length srcdom $4 srcvar paramcd $8 srcseq $40 param sprptlbl $80 paramtyp $20;
merge fevfvc_val(in=in_ratio) fevfvc_seq(in=in_seq);
by usubjid aphasen avisitn adt atm adtm ady visit spdtc anl01fl;
paramcd="FEVFVC";
param="RATIO OF FEV1/FVC";
paramtyp="DERIVED";
aval=fev/fvc;
srcdom="SP";
srcvar="SPSTRESN";
srcseq=strip(put(seq_fev,best.))||"$"||strip(put(seq_fvc,best.));
sprptlbl="Ratio of FEV1/FVC";
run;
proc sort data=fevfvc_ratio;by usubjid paramcd adt atm;run;
/****Derive sequence numbers for derived FEV1/FVC records****/
data fevfvc_ratio;
retain spseq;
set fevfvc_ratio;
by usubjid paramcd adt atm;
if first.usubjid then aspseq=301;
else aspseq=aspseq+1;
run;

```

SUBJID	ASPSEQ	PARAMCD	PARAMTYP	AVAL	DTYPE	AVISIT	ADTM	ANL01FL	SRCDOM	SRCVAR	SRCSEQ	VISIT	SPSTRESN
999001	8	FEV1		2.96		Screening	31MAY2012:11:56:00	Y				SCREENING	2.96
999001	9	FEV1		2.95		Day 1	07JUN2012:08:09:00	Y				DAY 1	2.95
999001	9.5	FEV1	DERIVED	2.95		Baseline	07JUN2012:08:09:00	Y				DAY 1	2.95
999001	10	FEV1		3.08		Day 7	14JUN2012:08:25:00	Y				DAY 7	3.08
999001	11	FEV1		3.12		Day 14	21JUN2012:08:35:00	Y				DAY 14	3.12
999001	12	FEV1		3.13		Day 21	26JUN2012:09:22:00	Y				DAY 21	3.13
999001	12.5	FEV1	DERIVED	3.11	AVERAGE	Average through Day 28	.	Y	ADSP	AVISIT	DAY 7\$DAY 14\$DAY 21		3.11
999001	13	FEV1		3.27			05JUL2012:09:07:00					EARLY TERMINATION	3.27
999001	14	FEV1		3		Follow-up/Day 56	18JUL2012:08:34:00	Y				FOLLOW-UP/DAY 56	3
999001	15	FVC		3.8		Screening	31MAY2012:11:56:00	Y				SCREENING	3.8
999001	16	FVC		3.98		Day 1	07JUN2012:08:09:00	Y				DAY 1	3.98
999001	16.5	FVC	DERIVED	3.98		Baseline	07JUN2012:08:09:00	Y				DAY 1	3.98
999001	17	FVC		4.06		Day 7	14JUN2012:08:25:00	Y				DAY 7	4.06
999001	18	FVC		4.1		Day 14	21JUN2012:08:35:00	Y				DAY 14	4.1
999001	19	FVC		4.06		Day 21	26JUN2012:09:22:00	Y				DAY 21	4.06
999001	19.5	FVC	DERIVED	4.07	AVERAGE	Average through Day 28	.	Y	ADSP	AVISIT	DAY 7\$DAY 14\$DAY 21		4.07333333
999001	20	FVC		4.31			05JUL2012:09:07:00					EARLY TERMINATION	4.31
999001	21	FVC		4.02		Follow-up/Day 56	18JUL2012:08:34:00	Y				FOLLOW-UP/DAY 56	4.02
999001	301	FEV1FVC	DERIVED	0.779		Screening	31MAY2012:11:56:00	Y	SP	SPSTRESN 8\$15		SCREENING	0.77894737
999001	302	FEV1FVC	DERIVED	0.741		Day 1	07JUN2012:08:09:00	Y	SP	SPSTRESN 9\$16		DAY 1	0.74120603
999001	302.5	FEV1FVC	DERIVED	0.741		Baseline	07JUN2012:08:09:00	Y	SP	SPSTRESN 9\$16		DAY 1	0.74120603
999001	303	FEV1FVC	DERIVED	0.759		Day 7	14JUN2012:08:25:00	Y	SP	SPSTRESN 10\$17		DAY 7	0.75862069
999001	304	FEV1FVC	DERIVED	0.761		Day 14	21JUN2012:08:35:00	Y	SP	SPSTRESN 11\$18		DAY 14	0.76097561
999001	305	FEV1FVC	DERIVED	0.771		Day 21	26JUN2012:09:22:00	Y	SP	SPSTRESN 12\$19		DAY 21	0.77093596
999001	305.5	FEV1FVC	DERIVED	0.764	AVERAGE	Average through Day 28	.	Y	ADSP	AVISIT	DAY 7\$DAY 14\$DAY 21		0.76351075
999001	306	FEV1FVC	DERIVED	0.759			05JUL2012:09:07:00		SP	SPSTRESN 13\$20		EARLY TERMINATION	0.7587007
999001	307	FEV1FVC	DERIVED	0.746		Follow-up/Day 56	18JUL2012:08:34:00	Y	SP	SPSTRESN 14\$21		FOLLOW-UP/DAY 56	0.74626866

**Display 4 Snapshot of Ratio of FEV1/FVC records with Traceability Variables in Spirometry Analysis Dataset**

In above example, PARAMCD='FEV1FVC', SRCDOM='SP', SRCVAR='SPSTRESN', and SRCSEQ='12\$19' at AVISIT='Day 21' indicates that values of SPSTRESN variable from SDTM SP domain with sequence numbers 12 and 19, respectively, are used to derive ratio of FEV1/FVC. We can also populate DTYPE='RATIO' to let the reviewer know that AVAL is the ratio of SPSTRESN corresponding to records with sequence numbers listed in SRCSEQ. Note: AVISIT is populated only for records that are used in analyses. In above example, Early Termination assessments are not analyzed and hence AVISIT has been set to missing. Derived Ratio of FEV1/FVC records are assigned a separate set of sequence numbers. In the above scenario the sequence number for Ratio of FEV1/FVC records starts with 301 and there by incremented by one.

**USAGE OF SRCDOM, SRCVAR AND SRCSEQ TRIPLET IN ESTABLISHING TRACEABILITY FOR PPFEE, PPFEV, AND PPFVC AT EVERY ANALYSIS VISIT IN ADSP**

Percent Predict values (PPFEF, PPFEV, and PPFVC) are derived for every nominal visit in SDTM SP using corresponding SPSTRESN and predicted values. Predicted values are derived using KNV or Hankinson/Wang Equation Coefficients by Sex and Age at baseline or at every analysis visit depending upon the population (pediatrics or adults) and length of a study (short or long). Specification (metadata) of derived Percent Predicted records for SPSEQ, AVAL variables and for SRCDOM, SRCVAR, and SRCSEQ which helps in building data point traceability is

shown in Table 6 followed by a sample SAS code that populates these variables. Snapshot of these variables in an analysis dataset is shown in Display 5.

Variable Name	Variable Label	Type	Length	Comments
ASPSEQ	Analysis Sequence Number	Num	8	Equal to 200 +SP.SPSEQ for Percent Predicted FEF, FEV, and FVC, where SP.SPSEQ is the corresponding sequence number of FEF, FEV, and FVC assessments.
AVAL	Analysis Value	Num	8	Derived per the information in SAP for 'PPFEF', 'PPFEV', and 'PPFVC'. AVAL is rounded to suggested precision as per variable SPPREC, only at the end. All the derivations are done using analysis values before rounding to the suggested precision.
SRCDOM	Source Domain	Char	4	Equal to 'SP' for PARAMTYP='DERIVED' and PARAMCD is equal to 'PPFEV', 'PPFEF', or 'PPFVC'
SRCVAR	Source Variable	Char	8	Equal to 'SPSTRESN' for PARAMTYP='DERIVED' and PARAMCD is equal to 'PPFEV', 'PPFEF', or 'PPFVC'
SRCSEQ	Source Sequence Number	Char	40	SRCSEQ lists the sequence numbers of records, separated by '\$', used in deriving the analysis values for records with PARAMTYP='DERIVED'. Example: If sequence number of Day 7 visit for a subject is 4 for PARAMCD equal to FEV then SRCSEQ is equal to 4 for PARAMCD equal to 'PPFEV' for analysis visit Day 7

**Table 6 Metadata of SWSEQ, AVAL for % Predicted Values and of SRCDOM, SRCVAR, and SRCSEQ Triplet**

```

/*****Derive Percent Predicted FEV, FVC and FEF*****/
data ppfef; /*Similarly, for ppfev, ppfvc*/
length srcdom $4 srcvar $8 srcseq $40;
  set sp(where=(paramcd="FEF" and knvfef ne . and spstresn ne .));
  paramcd="PPFEF";
  param="PERCENT PREDICTED FORCED MIDEXPIRATORY FLOW RATE (%)";
  paramtyp="DERIVED";
  aval=(spstresn/knvfef)*100;
  srcdom="SP";
  srcvar="SPSTRESN";
  srcseq=strip(put(spseq,best.));
  sprptlbl="Percent Predicted Forced Midexpiratory Flow Rate (%)";
run;
data percent_predicted;
  set ppfef(in=in_ppfef) ppfev(in=in_ppfev) ppfvc(in=in_ppfvc);
  drop spseq;
run;
data percent_predicted;
  retain spseq;
  set percent_predicted;
  by usubjid paramcd adt atm;
  if first.usubjid then aspseq=201;
  else aspseq=aspseq+1;
run;

```

SUBJID	ASPSEQ	PARAMCD	PARAMTYP	AVAL	DTYPE	AVISIT	ADTM	ANL01FL	SRCDOM	SRCVAR	SRCSEQ	VISIT	SPSTRESN
999001	8	FEV1		2.96		Screening	31MAY2012:11:56:00	Y				SCREENING	2.96
999001	9	FEV1		2.95		Day 1	07JUN2012:08:09:00	Y				DAY 1	2.95
999001	9.5	FEV1	DERIVED	2.95		Baseline	07JUN2012:08:09:00	Y				DAY 1	2.95
999001	10	FEV1		3.08		Day 7	14JUN2012:08:25:00	Y				DAY 7	3.08
999001	11	FEV1		3.12		Day 14	21JUN2012:08:35:00	Y				DAY 14	3.12
999001	12	FEV1		3.13		Day 21	26JUN2012:09:22:00	Y				DAY 21	3.13
999001	12.5	FEV1	DERIVED	3.11	AVERAGE	Average through Day 28		Y	ADSP	AVISIT	DAY 7\$DAY 14\$DAY 21		3.11
999001	13	FEV1		3.27			05JUL2012:09:07:00					EARLY TERMINATION	3.27
999001	14	FEV1		3		Follow-up/Day 56	18JUL2012:08:34:00	Y				FOLLOW-UP/DAY 56	3
999001	102	PFEV1	DERIVED	4.041									4.0405
999001	208	PPFEV1	DERIVED	73.258		Screening	31MAY2012:11:56:00	Y	SP	SPSTRESN	8	SCREENING	73.25826012
999001	209	PPFEV1	DERIVED	73.011		Day 1	07JUN2012:08:09:00	Y	SP	SPSTRESN	9	DAY 1	73.01076599
999001	209.5	PPFEV1	DERIVED	73.011		Baseline	07JUN2012:08:09:00	Y	SP	SPSTRESN	9	DAY 1	73.01076599
999001	210	PPFEV1	DERIVED	76.228		Day 7	14JUN2012:08:25:00	Y	SP	SPSTRESN	10	DAY 7	76.22818958
999001	211	PPFEV1	DERIVED	77.218		Day 14	21JUN2012:08:35:00	Y	SP	SPSTRESN	11	DAY 14	77.21816607
999001	212	PPFEV1	DERIVED	77.466		Day 21	26JUN2012:09:22:00	Y	SP	SPSTRESN	12	DAY 21	77.46566019
999001	212.5	PPFEV1	DERIVED	76.971	AVERAGE	Average through Day 28		Y	ADSP	AVISIT	DAY 7\$DAY 14\$DAY 21		76.97067195
999001	213	PPFEV1	DERIVED	80.931			05JUL2012:09:07:00		SP	SPSTRESN	13	EARLY TERMINATION	80.9305779
999001	214	PPFEV1	DERIVED	74.248		Follow-up/Day 56	18JUL2012:08:34:00	Y	SP	SPSTRESN	14	FOLLOW-UP/DAY 56	74.2482366

**Display 5 Snapshot of Percent Predicted FEV derived records with Traceability Variables in Spirometry ADaM**

SRCSEQ for percent predicted values indicates the sequence number of the record in SDTM SP that is used to derive the AVAL. In above example, PARAMTYP='DERIVED', SRCDOM='SP', SRCVAR='SPSTRESN', and SRCSEQ='12' at AVISIT=Day 21' for PARAMCD='PPFEV' indicates that value of SPSTRESN variable from SDTM SP domain with sequence number 12 is used to derive Percent Predicted FEV at analysis visit Day 21. Note: Derived Percent Predicted records are assigned a separate set of sequence numbers. In the above scenario the sequence number for Percent Predicted records starts with 201 and there by incremented by one. If predicted values (PFEF, PFEV, PFVC, and PFEVFVC) are derived at every analysis visit then it is recommended to list the sequence numbers of predicted value records along with sequence numbers of raw values (FEF, FEV, FVC, and FEVFVC) in SRCSEQ for percent predicted value records.

## CONCLUSION

Incorporating traceability features in ADaM datasets helps in effective program validation, speeds up the review process, and facilitates transparency in submitted analysis data. This paper provides examples of establishing data point traceability in ADaM datasets by using SRCDOM, SRCVAR, and SRCSEQ triplet where an analysis value is derived either from a single predecessor record or from multiple predecessor records from single variable of source SDTM.

## ACKNOWLEDGMENTS

Appreciation goes to Appreciation goes to Cindy Song for her review and comments.

## CONTACT INFORMATION

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

Name: Xiangchen Cui  
Enterprise: Vertex Pharmaceuticals Incorporated  
Address: 130 Waverly Street  
City, State ZIP: Cambridge, MA, 02139  
Work Phone: 617/444-6069  
E-mail: [xiangchen\\_cui@vrtx.com](mailto:xiangchen_cui@vrtx.com)

Name: Tathabbai Pakalapati  
Enterprise: Cytel Inc.  
Address: 675 Massachusetts Avenue  
City, State ZIP: Cambridge, MA, 02139  
Work Phone: 617/528-7144  
E-mail: [Tathabbai.Pakalapati@cytel.com](mailto:Tathabbai.Pakalapati@cytel.com)

Name: Qunming Dong  
Enterprise: Vertex Pharmaceuticals Incorporated  
Address: 130 Waverly Street  
City, State ZIP: Cambridge, MA, 02139  
Work Phone: 617-961-7514  
E-mail: [qunming\\_dong@vrtx.com](mailto:qunming_dong@vrtx.com)

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. © indicates USA registration.

Other brand and product names are trademarks of their respective companies.