

BIMO SAS® Macros and Programming Tools

Rohit Kamath, Regeneron Pharmaceuticals Inc.;
Mi Young Kwon, Regeneron Pharmaceuticals Inc.;

ABSTRACT

As part of the regulatory review process, FDA conducts site inspections to ensure that clinical investigators, sponsors, and Institutional Review Boards (IRBs) comply with regulations. The current submission format for study data in NDA and BLA packages does not facilitate efficient site selection for FDA because these data are submitted as subject-level data. Therefore, FDA has requested that pharmaceutical companies submit data that describes the characteristics and outcomes of clinical investigations at the site level. FDA uses this data to plan their site inspections. The submission of data for this request is also sometimes referred to a Bioresearch Monitoring Program (BIMO) submission because the data is placed in the BIMO section of Module 5 in the eCTD. FDA BIMO develops guidelines for inspections of clinical investigators, sponsors, and IRBs.

INTRODUCTION

FDA Office of Scientific Investigations (OSI) manages the BIMO program for drugs, and FDA Office of Inspections and Surveillance manages the BIMO program for biologics.

There are three parts to the request to be provided to FDA:

- General study related information and specific Clinical Investigator information
- Subject-level data listings by site
- Site-level dataset in a standardized electronic format.

Submission teams will make a proposal as to which studies are within scope of BIMO and which studies are not in scope for BIMO with their FDA review team during pre-NDA or pre-BLA communications to determine what Summary Level Clinical Site data to provide for their submission. Per FDA guidance document Specifications for Preparing and Submitting Summary Level Clinical Site Data for CDER's Inspection Planning, the site-level dataset should contain data from all major (e.g. pivotal) studies used to support safety and efficacy in the application, including studies with different treatment indications. If a site-level dataset is needed, FDA and sponsor will need to discuss and decide on what studies to include in the dataset and details about the content.

This paper will introduce SAS® macros and programming tools to automate and standardize the BIMO dataset and listing generation per regulatory authorization requirements and standards for electronic submissions.

BIMO CLINSITE SAS DATASET AND LISTINGS GENERATION

PROGRAMMING PROCESS FLOW CHART

Figure 1 is a flowchart to show the process of the generation BIMO Clinsite SAS® dataset and Listings.

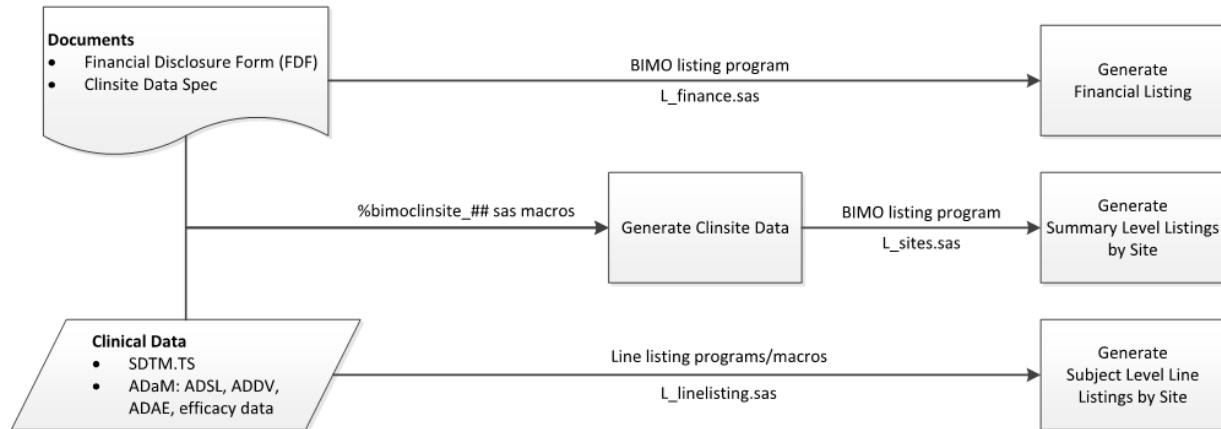


Figure 1. Flowchart

REQUIRED COMPONENTS

Clinsite dataset

The following components are required to generate the Clinsite dataset:

- Clinical Data
 - CDISC conformity datasets for submission will be used to generate Clinsite dataset.
 - SDTM TS domain
 - ADaM Datasets
 - ADSL
 - ADDV
 - ADAE
 - Efficacy dataset (that contains the primary endpoint)
- Financial Disclosure Form (FDF) in excel format (xlsx)
- Clinsite dataset specification in excel format
- SAS® macro(s) (referenced in Table 1)

Subject-Level Data Line Listing

The following components are required to generate Subject-Level Data Line Listings by Clinical Site.

- SAS® macro (referenced in Table 1)

Site Summary Listing

For every site, Site Summary listing summarizes the number of subjects that are screened, treated, discontinued early from treatment or discontinued early from study in a tabular format. The following components are required to generate this listing:

- Clinsite dataset
- SAS® macro (referenced in Table 1)

Finance Summary Listing

Finance Summary Listing is to summarize the information of investigators and sub-investigators and financial interests at each site. The following components are required to generate this listing:

- Financial Disclosure Form in excel format (xlsx)
- SAS® macro (referenced in Table 1)

List of SAS® macros

Table 1 shows SAS® macros to generate the components.

Clinsite Dataset	Line Listing	Site Summary Listing	Finance Summary Listing
%bimoclinsite_section1	l_linelisting.sas	l_sites.sas	l_finance.sas
%bimoclinsite_section2			
%bimoclinsite_section3			
%bimoclinsite_section4			
%bimoclinsite_section5			
%bimoclinsite_final			

Table 1. SAS® macros to generate the components

INPUT FILES AND CONVENTION

Financial Disclosure Form (FDF) Excel File

The FDF Excel file contains information of investigators (Last Name, First Name, Role), sites (Site address, phone number, email), and financial disclosure amount. This file usually is released from Project management team after they collect all information from different functions. The macro expects the FDF excel file to have a standard format with the required columns for SAS macro inputs

Clinsite Data Specification (clinsite_spec_Study ID)

The Clinsite Data Specification contains all variables that should be populated in the data according to BIMO guidance_Technical Conformance Guide (<https://www.fda.gov/media/85061/download>, July 202).

Table 2 shows the detailed description of all variables included in the data.

Section	Section Description	Variable Index	Variable Name	Variable Label	Type
1	Study & Registration Information	1	STUDYID	Study Identifier	Char
1	Study & Registration Information	2	TITLE	Study Title	Char
1	Study & Registration Information	3	SPONCNT	Sponsor Count	Num
1	Study & Registration Information	4	SPONSOR	Sponsor Name	Char

Section	Section Description	Variable Index	Variable Name	Variable Label	Type
1	Study & Registration Information	5	IND	IND Number	Num
1	Study & Registration Information	6	UNDERIND	Under IND	Char
1	Study & Registration Information	7	NDA	NDA Number	Num
1	Study & Registration Information	8	BLA	BLA Number	Num
1	Study & Registration Information	9	SUPPNUM	Supplement Number	Num
2	Disposition & Population by Site	10	SITEID	Study Site Identifier	Char
2	Disposition & Population by Site	11	ARM	Description of Planned Treatment Arm	Char
2	Disposition & Population by Site	12	COHORT	Description of Planned Cohort	Char
2	Disposition & Population by Site	13	SAFPOP	Number of Subjects in Safety Population	Num
2	Disposition & Population by Site	14	SCREEN	Number of Subjects Screened	Num
2	Disposition & Population by Site	15	DISCSTUD	Number of Subject Discons from Study	Num
2	Disposition & Population by Site	16	DISCRT	Number of Subject Discons from Study Treatment	Num
3	Efficacy by Site	17	ENDPOINT	Primary Endpoint	Char
3	Efficacy by Site	18	ENDPTYPE	Primary Endpoint Type	Char
3	Efficacy by Site	19	TRTEFFR	Treatment Efficacy Result	Num
3	Efficacy by Site	20	TRTEFFS	Treatment Efficacy Result STD	Num
3	Efficacy by Site	21	CENSOR	Number of Censored Observations	Num
4	Safety by Site	22	NSAE	Number of Non-Serious Adverse Events	Num
4	Safety by Site	23	SAE	Number of Serious Adverse Events	Num
4	Safety by Site	24	DEATH	Number of Deaths	Num
4	Safety by Site	25	IMPDEV	Number of Important Protocol Deviations	Num
4	Safety by Site	26	NOIMPDEV	Number of Non-Important Protocol Deviations	Num

Section	Section Description	Variable Index	Variable Name	Variable Label	Type
5	Site & Financial Information	27	FINLDISC	Financial Disclosure Amount	Char
5	Site & Financial Information	28	LASTNAME	Investigator Last Name	Char
5	Site & Financial Information	29	FRSTNAME	Investigator First Name	Char
5	Site & Financial Information	30	INITIAL	Investigator Middle Initial	Char
5	Site & Financial Information	31	PHONE	Investigator Phone Number	Char
5	Site & Financial Information	32	FAX	Investigator Fax Number	Char
5	Site & Financial Information	33	EMAIL	Investigator Email Address	Char
5	Site & Financial Information	34	COUNTRY	Country	Char
5	Site & Financial Information	35	STATE	State	Char
5	Site & Financial Information	36	CITY	City	Char
5	Site & Financial Information	37	POSTAL	Postal Code	Char
5	Site & Financial Information	38	STREET	Street Address	Char
5	Site & Financial Information	39	STREET1	Street Address Continued	Char

Table 2. Clinsite Data Specification

Display 1 is a screenshot of Clinsite Data Specification

Section	Section Description	Variable Index	Variable Name	Variable Label	Type	Controlled Terms or Format	Programming Note
1	Study & Registration Information	1	STUDYID	Study Identifier	Char	String	ADaM.ADSL.STUDYID
1	Study & Registration Information	2	TITLE	Study Title	Char	String	SDTM.TS.TSVAL when TSPARMCD="Title"
1	Study & Registration Information	3	SPONCNT	Sponsor Count	Num	Integer	1
1	Study & Registration Information	4	SPONSOR	Sponsor Name	Char	String	XXXXXX
1	Study & Registration Information	5	IND	IND Number	Num	6 digit identifier	XXXXXX
1	Study & Registration Information	6	UNDERIND	Under IND	Char	String	Y
1	Study & Registration Information	7	NDA	NDA Number	Num	6 digit identifier	
1	Study & Registration Information	8	BLA	BLA Number	Num	6 digit identifier	XXXXXX
1	Study & Registration Information	9	SUPPNUM	Supplement Number	Num	Integer	

Display 1. A screenshot of Clinsite Data Specification

BIMO SAS® MACROS AND PROGRAMS

%bimoclinsite_section1

Function:

Macro, *%bimoclinsite_section1*, is designed to generate a section 1 dataset from clinical database. The section 1 contains Study & Registration Information (clinsite data spec Variable Index 1 - 9). Table 3 shows the description of parameters and the default value at SAS® macro *%bimoclinsite_section1*.

Definition:

```
%bimoclinsite_section1(studyid=, filepath=, filename=, filetype);
```

Parameter	Default	Description
studyid		full study ID
filepath		Path of Clinsite Specs
Filename		Excel File name of clinsite Specs
filetype		XLS or XLSX

Table 3. SAS® macro %bimoclinsite_section1

Sample call:

```
%bimoclinsite_section1(studyid=XXXXX,  
                        filepath = &studypath./eSub/Docs/bimo,  
                        filename = %str(clinsite_spec_XXXXX),  
                        filetype = xlsx);
```

%bimoclinsite_section2

Function:

Macro, *%bimoclinsite_section2*, is designed to generate a section 2 dataset from clinical database. The section 2 contains information about Disposition & Population by Site (clinsite data spec Variable Index 10 - 16). Table 4 shows the description of parameters and the default value at SAS® macro *%bimoclinsite_section2*.

Definition:

```
%bimoclinsite_section2(data=, subjid=, armn=, arm=, cohort=, pops=, popwheres=);
```

Parameter	Default	Description
data		dataset name used for counting population
subjid		unique subject ID for counting
armn		treatment group name based on population
arm		numeric treatment group name
cohort		cohort variable name in the data
pops		variable name in Clinsite specs for each population Note: Multiple populations can be defined by using delimiter.

		delimiter for each population: '
popwheres		algorithm for counting population Note: Multiple algorithms can be defined by using delimiter. delimiter for each population: '

Table 4. SAS® macro %bimoclinsite_section2

%bimoclinsite_section3

Function:

Macro, %bimoclinsite_section3, is designed to generate a section 3 dataset from clinical database. The section 3 contains information about Efficacy by Site (clinsite data spec Variable Index 17 - 21). Table 5 shows the description of parameters and the default value at SAS® macro %bimoclinsite_section3.

Definition:

%bimoclinsite_section3(data=, armn=, arm=, cohort=, endpoint=, endtype=, censor);

Parameter	Default	Description
data		dataset name used for counting population
armn		treatment group name based on population
arm		numeric treatment group name
cohort		cohort variable name in the data
endpoint		plain text label used to describe the primary endpoint
endtype		variable type of the primary endpoint (i.e., Continuous, Binary, Discrete, Time to event, or other).
censor		censored variable name in the data when endtype='Time to event'

Table 5. SAS® macro %bimoclinsite_section3

%bimoclinsite_section4

Function:

Macro, %bimoclinsite_section4, is designed to generate a section 4 dataset from clinical database. The section 4 contains summary information about Safety by Site (clinsite data spec Variable Index 22 - 26). Table 6 shows the description of parameters and the default value at SAS® macro %bimoclinsite_section4.

Definition:

%bimoclinsite_section4(data1=, subjid=, armn=, arm=, cohort=, pops1=, popwheres1=, data2=, pops2=, popwheres2=);

Parameter	Default	Description
data1		AE dataset name used for counting population
subjid		unique subject ID for counting
armn		treatment group name based on population
arm		numeric treatment group name
cohort		cohort variable name in the data

Parameter	Default	Description
pops1		variable name in Clinsite specs for each population in AE dataset Note: Multiple populations can be defined by using delimiter. delimiter for each population: ' '
popwheres1		algorithm for counting population in AE dataset Note: Multiple algorithms can be defined by using delimiter. delimiter for each population: ' '
data2		Protocol Deviation dataset name used for counting population
pops2		variable name in Clinsite specs for each population in Protocol Deviation dataset Note: Multiple populations can be defined by using delimiter. delimiter for each population: ' '
popwheres2		algorithm for counting population in Protocol Deviation dataset Note: Multiple algorithms can be defined by using delimiter. delimiter for each population: ' '

Table 6. SAS® macro %bimoclinsite_section4

%bimoclinsite_section5

Function:

Macro, %bimoclinsite_section5, is designed to generate a section 5 dataset from FDF. The section 5 contains information about Site & Financial Information (clinsite data spec Variable Index 27 - 39). Table 7 shows the description of parameters and the default value at SAS® macro %bimoclinsite_section5.

Definition:

%bimoclinsite_section5(studyid=, filepath=, filename=, filetype=);

Parameter	Default	Description
studyid		full study ID
filepath		Path of FDF
filename		Excel File name of FDF
filetype		XLS or XLSX

Table 7. SAS® macro %bimoclinsite_section5

% bimoclinsite_final

Function:

Macro, %bimoclinsite_final, is designed to generate a final clinsite dataset by combining the individual datasets from Section 1 to 5. Table 8 shows the description of parameters and the default value at SAS® macro %bimoclinsite_final.

Definition:

%bimoclinsite_final(xptpath=);

Parameter	Default	Description
xptpath		Location to save the final Clinsite.xpt file

Table 8. SAS® macro %bimoclinsite_final

Sample output:

STUDYID	TITLE	SPONCNT	SPONSOR	IND	UNDERIND	NDA	BLA	SUPPNUM	SITEID	ARM	COHORT
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124101	Placebo	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124101	Placebo	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124101	Active 1	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124101	Active 1	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124101	Active 2	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124102		
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124103	Active 1	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124103	Active 1	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124103	Active 2	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124103	Active 2	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124104	Placebo	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124104	Placebo	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124104	Active 1	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124104	Active 1	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124104	Active 1	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124106	Active 2	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124107	Active 1	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124107	Active 1	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124107	Active 2	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124107	Active 2	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	124108		
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	233101	Active 1	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	233101	Active 1	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	233101	Active 2	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	233101	Active 2	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	233103	Placebo	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	233103	Placebo	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	233103	Active 1	
XXXX-XX-XXXX	A Phase 3 Ran...	1	Regeneron Pharmaceuticals Inc.	111111	Y	.	222222	.	233103	Active 1	

Display 2. SAS® output from macro %bimoclinsite_final

I_linelisting.sas®

Function:

SAS® program, I_linelisting.sas, is designed to generate the subject-level data line listing per site per input data.

I_sites.sas®

Function:

SAS® program, I_sites.sas, is designed to generate the site summary listing from Clinsite dataset.

Sample output:

Clinical Site Listing				
Site	Number of Patients Screened (N= xxxxxx)	Number of Patients Treated (N= xxxxx)	Number of Patients Who Prematurely Discontinued from Study (N= xxxxx)	Number of Patients Who Prematurely Discontinued from Treatment (N= xxxxx)
100001	xx	xx	x	x
100002	xxx	xx	x	x
100003	xxx	xx	x	x
100004	xxx	xx	x	x
100005	xxx	xx	x	x
100006	xxx	xx	x	x
100007	xxx	xx	x	x
100008	xxx	xx	x	x

Display 3. A screen shot of site summary listing

I_finance.sas®

Function:

Regeneron Pharmaceuticals, Inc. Protocol: XXXX-XX-XXXX		Site: XXXXXXX
1.1 a Listing for each subject consented/enrolled; for subjects who were not randomized to treatment and/or treated with study therapy, include reason not randomized and/or treated		
Patient Identifier	Population	Primary Reason for Screen Failure
XXXX-XX-XXXX -XXXXXXXX	Not Randomized/ Not Treated	INCLUSION CRITERIA NOT MET AND/OR EXCLUSION CRITERIA MET: INCXX
XXXX-XX-XXXX -XXXXXXXX	Not Randomized/ Not Treated	INCLUSION CRITERIA NOT MET AND/OR EXCLUSION CRITERIA MET: EXCXX
XXXX-XX-XXXX -XXXXXXXX	Randomized/ Treated	
XXXX-XX-XXXX -XXXXXXXX	Not Randomized/ Not Treated	INCLUSION CRITERIA NOT MET AND/OR EXCLUSION CRITERIA MET: EXCXX
XXXX-XX-XXXX -XXXXXXXX	Randomized/ Treated	
XXXX-XX-XXXX -XXXXXXXX	Not Randomized/ Not Treated	INCLUSION CRITERIA NOT MET AND/OR EXCLUSION CRITERIA MET: INCXX

Display 5. A screenshot of Subject-Level Line Listing

POOLING CLINSITE DATASETS

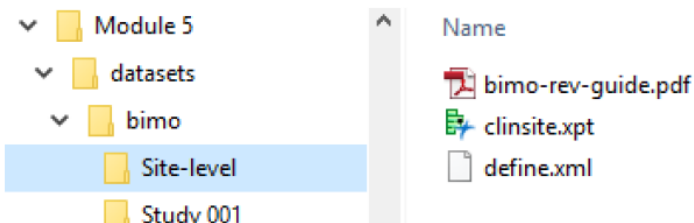
According to BIMO technical conformance guidance, a single file containing data from all major (pivotal) studies used to support safety and efficacy in the application is required for submission. After the Clinsite dataset from each individual study is generated, the final step is to pool individual Clinsite datasets into a single Clinsite dataset. Define.xml is required for the pooled Clinsite dataset. A Reviewer’s Guide is optional for the clinsite dataset. The project team should determine on a case-by-case basis if a Reviewer’s Guide is needed for the clinsite dataset in a submission.

C. eCTD Folder Structure for Summary-Level Clinical Site Dataset

For the site-level dataset, use the filename “clinsite.xpt.” A single file containing data from all major (i.e., pivotal) studies used to support safety and efficacy in the application should be provided.

Within the eCTD folder structure, place the site-level dataset define file and BIMO Reviewer’s Guide, if it is being submitted, in the M5 folder as follows:

Figure 2: Place the Site-Level Dataset Define File and BIMO Reviewer’s Guide in the M5 Folder



Display 6. A screenshot of BIMO technical conformance guidance

CONCLUSION

This paper introduces SAS® macros and programming tools to automate and standardize the BIMO dataset and listing generation per regulatory authorization requirements and standards for electronic submissions:

- Standard Financial Disclosure Form (FDF)
- SAS® macros to automate dataset, title file and listing generation
- Sample programs for training and reference

This process, SAS® macros and standard programs will provide a consistent way across Therapeutic Areas to ensure that all requirements are efficiently performed and conformity with regulatory guidance for electronic submission.

REFERENCES

“BIMO guidance_Technical Conformance guide” released at July 27, 2020:
<https://www.fda.gov/media/85061/download>

Paper RG07 “BIMO Listings – Check That Off Your NDA To-Do List” PhUSE 2018:
<https://www.lexjansen.com/phuse-us/2018/rg/RG07.pdf>

CONTACT INFORMATION

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

Rohit Kamath
Regeneron Pharmaceuticals Inc.
Rohit.kamath@regeneron.com

Mi Young Kwon
Regeneron Pharmaceuticals Inc.
miyoung.kwon@regeneron.com

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