

# AsthmaNet Microbiome Data Documentation for BioLINCC

## Microbiome Study Forms, Data sets and Variables

Each CRF has an abbreviated name located in a shaded box in the lower right-hand corner of the form. Abbreviated CRF names beginning with a 'P3\_' prefix are specific to the Microbiome protocol; data associated with these Microbiome-specific forms are stored in a corresponding data file of the same name with the 'P3\_' prefix removed. A listing of each CRF and data set name is included in Table 1.

Variable names for a particular CRF begin with a prefix assigned to that form followed by an underscore and the number of the field; prefixes for each form can be found in Table 1. Fields within each CRF are usually numbered sequentially and are located (in most cases) in parentheses on the left hand side of the responses to each question.

For example, data regarding adverse events experienced by the Microbiome subjects are collected on the Clinical Adverse Events form (form name AECLIN). Corresponding data are stored in file aeclin.sas7bdat. The variable name corresponding to the ICD9 code for a particular adverse event (field #1010) on the Clinical Adverse Events form is aec\_1010.

Where no form name is given in Table 1, no CRF existed for collection of the relevant information during study implementation. Where no dataset name is given, the corresponding "form" was simply a reference card that was used to explain some element of data collection.

Table 2 below shows the CRFs that were completed at each study visit, either routinely or only on an as-needed basis.

## Standard Data Edits

Each Microbiome CRF includes a set of key variables in the upper right-hand corner. These variables may include: Participant ID (part\_id), participant initials, visit number (vnum), visit/current date (vdate), and interviewer/technician/supervisor/coordinator ID. For the purpose of de-identifying the dataset, the following standard changes were applied to the key variables for all public use datasets:

- Participant initials were omitted.
- All visit dates were converted to intervals representing the number of days from the participant's Visit 0 (initial screen visit).
- All interviewer/technician/supervisor/coordinator IDs were omitted.

In addition, the following changes were made to all public use data sets:

- All dates collected within a given CRF were converted to days from Visit 0. Unless otherwise noted, converted dates were stored in the original variable names given on the CRFs.
- Data collected in source documentation boxes or fields (for example, snq\_1050, snq\_1060, snq\_1070 from the Sinonasal Questionnaire (SNQ) form) were omitted from the public datasets. Variables in the source documentation boxes remain annotated on the forms but they were not included in the datasets.
- Comment fields and text descriptions related to a given question were deleted. These fields contain free text which may include dates or other participant-related details that risk possible identification.

Note the 'vnum\_c' variable is a character variable representing the visit numbers as outlined in the protocol, with the phone contact between Visits 3 and V4 denoted with 'vnum\_c' of 3A. Variable 'vnum' is a corresponding numeric variable that was assigned for ease of sorting the data in chronological order, with 'vnum\_c' of 3A having 'vnum' of 3.01. For the remainder of the visits, 'vnum\_c' and 'vnum' values are identical.

Any additional changes made to a given dataset are summarized in Table 1 below.

## **Background**

One hundred eighty-six (186) individuals had Phadiatop screening at Microbiome Visit 0. Of the 186 screened, 102 were characterized as having asthma. Of those, 27 allergic non-asthmatics, 27 non-allergic non-asthmatics and 67 allergic asthmatics entered the run-in. Among the non-asthmatics, 21 allergic and 21 non-allergic individuals completed Visit 2, ending their participation in Microbiome. Among the asthmatics, 42 were randomized at Visit 2 to receive either ICS (N=28) or matched placebo (N=14) until Visit 5. At Visit 5, participants were placed on open-label ICS until returning for Visit 6.

The primary hypothesis of the Microbiome trial was: The microbiota of the bronchial airways of allergic asthmatic, allergic non-asthmatic, and non-allergic, non-asthmatic healthy subjects differ in diversity, richness, evenness, and/or taxonomic composition.

Visit numbers for the Microbiome data generally match the numbers outlined in the protocol (i.e., Visits 0 through 6).

### **Table 1. Microbiome Forms and Datasets**

Form Name Abbreviation	Dataset Name	Variable Prefix	Form Name or Description	Field Changes/ Comments
ACQ	acq.sas7bdat	acq	Asthma Control Questionnaire	<ul style="list-style-type: none"> <li>• Used with permission from Elizabeth Juniper</li> <li>• 6 question version of questionnaire (excluding FEV<sub>1</sub>) used</li> </ul>
AECLIN	aeclin.sas7bdat	aec	Clinical Adverse Events	<ul style="list-style-type: none"> <li>• aec_1020, aec_1030 (dates) were altered as described above</li> <li>• aec_1000 is a sequential number assigned to each adverse event within a participant (starting from 01 up to 99)</li> <li>• Variables added to this dataset include: <ul style="list-style-type: none"> <li>○ icd9_cat: ICD-9 category</li> <li>○ icd9long: Full description of ICD-9 code</li> <li>○ icd9shrt: abbreviated description of ICD-9 code</li> </ul> </li> <li>• The dataset used to code adverse events and link to descriptions was downloaded from <a href="http://www.cms.gov">www.cms.gov</a>. ICD-9-CM Diagnosis and Procedure Codes: Abbreviated and Full Code Titles effective October 1, 2009 were used.</li> </ul>
ASTHMA_HX_ADULT	asthma_hx_adult.sas7bdat	aha	Adult Asthma and Allergy History	<ul style="list-style-type: none"> <li>• aha_1400D, aha_1410D, aha_1420D, aha_1460D were deleted (description fields)</li> <li>• aha_6000 was deleted (comment field)</li> </ul>
	bmi.sas7bdat		Body Mass Index	<p>File contains the following variables:</p> <ul style="list-style-type: none"> <li>• 'part_id' = participant ID number</li> <li>• 'bmi' = body mass index computed from height and weight measured at Visit 1</li> </ul>

Form Name Abbreviation	Dataset Name	Variable Prefix	Form Name or Description	Field Changes/ Comments
CMED	cmed.sas/bdat	cme	Concomitant Medications for Asthma/Allergy and Adverse Events	<ul style="list-style-type: none"> <li>• cme_1060, cme_1070 (dates) have been altered as described above</li> <li>• cme_1000 is a sequential number assigned to each medication within a participant (starting from 01 up to 99)</li> <li>• Variables added to this dataset include: <ul style="list-style-type: none"> <li>○ class: drug class text</li> <li>○ class_id: class ID number</li> <li>○ gen_name: generic drug name</li> <li>○ Units (cme_1040), frequency (cme_1050) and route (cme_1055) contain formatted values</li> </ul> </li> <li>• Drug codes and classes were obtained from the American Society of Health-System Pharmacists, Inc. drug coding system (AsthmaNet purchased a license to use and update the data annually)</li> </ul>
COLD_HX	cold_hx.sas7bdat	chx	Cold History	<ul style="list-style-type: none"> <li>• chx_1000D was deleted (description field)</li> <li>• chx_6000 was deleted (comment field)</li> </ul>
HEQ	heq.sas7bdat	heq	Home Environment Questionnaire	<ul style="list-style-type: none"> <li>• heq_1710, heq_1720 were deleted (source doc fields)</li> <li>• heq_1000D, heq_1160D, heq_1250D, heq_1340D, heq_1410D, heq_1420D, heq_1460D, heq_1620D were deleted (description fields)</li> <li>• heq_6000 was deleted (comment field)</li> </ul>
HOUSEHOLD_SEI	household_sei.sas7bdat	sei	Household Socio-Economic Information	<ul style="list-style-type: none"> <li>• sei_1000D was deleted (description field)</li> <li>• sei_6000 was deleted (comment field)</li> </ul>
METHA	metha.sas7bdat	mth	Methacholine Challenge Testing	<ul style="list-style-type: none"> <li>• mth_6000 was deleted (comment field)</li> </ul>

Form Name Abbreviation	Dataset Name	Variable Prefix	Form Name or Description	Field Changes/ Comments
METHACHK_ADULT	methachk_adult.sas7bdat	mca	Adult Methacholine Challenge Testing Checklist	<ul style="list-style-type: none"> <li>• mca_1100D was deleted (description field)</li> <li>• mca_6000 was deleted (comment field)</li> </ul>
METHA_ADD_TRT	metha_add_trt.sas7bdat	mad	Additional Treatment Post Methacholine Challenge Testing	<ul style="list-style-type: none"> <li>• mad_1240, mad_1250, mad_1260 were deleted (source doc fields)</li> <li>• mad_1060D, mad_1190D were deleted (description fields)</li> <li>• mad_6000 was deleted (comment field)</li> </ul>
PALB4_SPIRO	palb4_spiro.sas7bdat	pa4	Post-Albuterol (4 puffs) Spirometry Testing	<ul style="list-style-type: none"> <li>• pa4_6000 was deleted (comment field)</li> </ul>
PREG_TEST	preg_test.sas7bdat	prg	Urine Pregnancy Test	<ul style="list-style-type: none"> <li>• prg_1050, prg_1060 were deleted (source doc fields)</li> <li>• prg_6000 was deleted (comment field)</li> </ul>
PRIOR_COND_ADULT	prior_cond_adult.sas7bdat	pad	Prior Conditions for Adult Participants	<ul style="list-style-type: none"> <li>• pad_1000D, pad_1010D, pad_1020D, pad_1030D, pad_1040D, pad_1050D, pad_1060D, pad_1070D, pad_1080D, pad_1090D, pad_1100D, pad_1110D, pad_1120D were deleted (description fields)</li> <li>• pad_6000 was deleted (comment field)</li> </ul>
PRIOR_COND_ALL	prior_cond_all.sas7bdat	pal	Prior Conditions for All Participants	<ul style="list-style-type: none"> <li>• pal_1000D, pal_1010D, pal_1060D, pal_1110D, pal_1130D, pal_1150D, pal_1170D, pal_1180D were deleted (description fields)</li> <li>• pal_6000 was deleted (comment field)</li> </ul>

Form Name Abbreviation	Dataset Name	Variable Prefix	Form Name or Description	Field Changes/ Comments
PRIOR_TR1	prior_trt.sas/bdat	ptr	Prior Asthma/Allergy Treatment	<ul style="list-style-type: none"> <li>• ptr_1000D, ptr_1470D, ptr_1500D, ptr_1600D, ptr_1830D, ptr_1870D, ptr_1910D were deleted (description fields)</li> <li>• ptr_6000 was deleted (comment field)</li> <li>• Date intervals for the medication dates in Questions 2-22 were stored in variables date_q02-date_q22. If day was missing, but month and year were present, we calculated the date using the 1<sup>st</sup> of the month. If only the year was present, no date interval was computed (i.e., missing data).</li> </ul>
REGISTRY	registry.sas7bdat	reg	AsthmaNet Registry information, including demographics	<ul style="list-style-type: none"> <li>• reg_1020, reg_1040, reg_1060 (dates) were altered as described above</li> <li>• reg_1070 (birth date) was deleted; the participant's age at visit 0 was calculated and is stored in variable 'age' (as a whole number)</li> <li>• ethnicity and individual race variables have been deleted due to sparseness (reg_1090, reg_1100, reg_1110, reg_1120, reg_1130, reg_1140)</li> <li>• reg_1150 (primary racial identification) has been re-coded such that those who were in categories 1, 2 and 6 were included in the 'other' category (code 6)</li> <li>• reg_1160 (description field) was deleted</li> </ul>

Form Name Abbreviation	Dataset Name	Variable Prefix	Form Name or Description	Field Changes/ Comments
SERIOUS	serious.sas7bdat	ser	Serious Adverse Event Reporting Form	<ul style="list-style-type: none"> <li>ser_1000, ser_1080, ser_1090 (dates) were altered as described above</li> <li>ser_1010D, ser_1180D, ser_1210D, ser_1220D were deleted (description fields)</li> <li>ser_6000 was deleted (comment field)</li> </ul>
SNQ	snq.sas7bdat	snq	Sinonasal Questionnaire	<ul style="list-style-type: none"> <li>snq_1050, snq_1060, snq_1070 were deleted (source doc fields)</li> <li>Permission was obtained from Anne Dixon, MD, lead author, to use the SNQ (CHEST 2009; 136:1324-1332) for Microbiome in the AsthmaNet format</li> </ul>
SPIRO	spiro.sas7bdat	spi	Spirometry Testing	<ul style="list-style-type: none"> <li>spi_6000 was deleted (comment field)</li> </ul>
SPUTLAB	sputlab.sas7bdat	sla	Sputum Induction Lab Values	<ul style="list-style-type: none"> <li>sla_1000 (date) was altered as described above</li> <li>sla_6000 was deleted (comment field)</li> </ul>
SPUTREAD	sputread.sas7bdat	sre	Sputum Induction Read	<ul style="list-style-type: none"> <li>sre_1000 (date) was altered as described above</li> <li>sre_6000 was deleted (comment field)</li> <li>Cell counts were performed by a central overreader</li> </ul>
SPUTUM	sputum.sas7bdat	spu	Sputum Induction	<ul style="list-style-type: none"> <li>spu_6000 was deleted (comment field)</li> </ul>
SPUTUMCHK	sputumchk.sas7bdat	sch	Sputum Induction Checklist	<ul style="list-style-type: none"> <li>sch_1060D was deleted (description field)</li> <li>sch_6000 was deleted (comment field)</li> </ul>
SPUTUM_ADD_TRT	sputum_add_trt.sas7bdat	sad	Additional Treatment Post Sputum Induction	<ul style="list-style-type: none"> <li>sad_1080, sad_1090, sad_1100 were deleted (source doc fields)</li> </ul>

Form Name Abbreviation	Dataset Name	Variable Prefix	Form Name or Description	Field Changes/ Comments
	allergy.sas7bdat		Blood testing for Allergy results (includes, phadiatop, total IgE and allergen-specific IgE test results)	File contains the following variables: <ul style="list-style-type: none"> <li>• part</li> <li>• phad</li> <li>• ige</li> <li>• re1</li> <li>• re5</li> <li>• re72</li> <li>• rmx1</li> <li>• ri6</li> <li>• rgx2</li> <li>• rtx4</li> <li>• rtx6</li> <li>• rwx1</li> <li>• rw3</li> <li>• rd2</li> <li>• rd1</li> </ul>
P3_ASTHMA_LOG	asthma_log.sas7bdat	al	Microbiome Asthma Monitoring Log	<ul style="list-style-type: none"> <li>• ddate (date) was altered as described above</li> <li>• al_6000 was deleted (comment field)</li> </ul>
P3_BPD	bpd.sas7bdat	bpd	Microbiome Bronchoscopy Procedure Documentation	<ul style="list-style-type: none"> <li>• bpd_1140 bpd_1150 were deleted (source doc fields)</li> <li>• bpd_1360D, bpd_1650D, bpd_1760D were deleted (description fields)</li> <li>• bpd_6000 was deleted (comment field)</li> </ul>
P3_BRONCHCHK_A	bronchchk_a.sas7bdat	bca	Microbiome Bronchoscopy Checklist Asthmatic	<ul style="list-style-type: none"> <li>• bca_1100D was deleted (description field)</li> <li>• bca_6000 was deleted (comment field)</li> </ul>
P3_BRONCHCHK_C	bronchchk_c.sas7bdat	bcc	Microbiome Bronchoscopy Checklist Control	<ul style="list-style-type: none"> <li>• bcc_1100D was deleted (description field)</li> <li>• bcc_6000 was deleted (comment field)</li> </ul>
P3_COMPLY	comply.sas7bdat	com	Microbiome Compliance Checklist	<ul style="list-style-type: none"> <li>• com_6000 was deleted (comment field)</li> </ul>



Form Name Abbreviation	Dataset Name	Variable Prefix	Form Name or Description	Field Changes/ Comments
P3_CTXQX	ctxqx.sas/bdat	ctx	Microbiome Coordinator Study Treatment Questionnaire	<ul style="list-style-type: none"> <li>ctx_1015D was deleted (description field)</li> <li>ctx_1020, ctx_1030 were deleted (source doc fields)</li> </ul>
P3_ELIG0A	elig0a.sas7bdat	e0a	Microbiome Eligibility Checklist 0 Asthmatic	<ul style="list-style-type: none"> <li>e0a_1010 (date) was altered as described above</li> <li>e0a_1050D, e0a_1130D were deleted (description fields)</li> <li>e0a_1360, e0a_1370 were deleted (source doc fields)</li> <li>e0a_6000 was deleted (comment field)</li> </ul>
P3_ELIG0C	elig0c.sas7bdat	e0c	Microbiome Eligibility Checklist 0 Control	<ul style="list-style-type: none"> <li>e0c_1010 (date) was altered as described above</li> <li>e0c_1050D, e0c_1130D were deleted (description fields)</li> <li>e0c_1360, e0c_1370 were deleted (source doc fields)</li> <li>e0c_6000 was deleted (comment field)</li> </ul>
P3_ELIG1A	elig1a.sas7bdat	e1a	Microbiome Eligibility Checklist 1 Asthmatic	<ul style="list-style-type: none"> <li>e1a_1040D, e1a_1050D were deleted (description fields)</li> <li>e1a_1080, e1a_1090 were deleted (source doc fields)</li> <li>e1a_6000 was deleted (comment field)</li> </ul>
P3_ELIG1C	elig1c.sas7bdat	e1c	Microbiome Eligibility Checklist 1 Control	<ul style="list-style-type: none"> <li>e1c_1040D, e1c_1050D were deleted (description fields)</li> <li>e1c_1080, e1c_1090 were deleted (source doc fields)</li> <li>e1c_6000 was deleted (comment field)</li> </ul>

Form Name Abbreviation	Dataset Name	Variable Prefix	Form Name or Description	Field Changes/ Comments
P3_ELIG2A	elig2a.sas7bdat	e2a	Microbiome Eligibility Checklist 2 Asthmatic	<ul style="list-style-type: none"> <li>e2a_1000D, e2a_1070D were deleted (description fields)</li> <li>e2a_1014, e2a_1017, e2a_1090, e2a_1100 were deleted (source doc fields)</li> <li>e2a_6000 was deleted (comment field)</li> </ul>
P3_ELIG2C	elig2c.sas7bdat	e2c	Microbiome Eligibility Checklist 2 Control	<ul style="list-style-type: none"> <li>e2c_1000D, e2c_1070D were deleted (description fields)</li> <li>e2c_1014, e2c_1017, e2c_1090, e2c_1100 were deleted (source doc fields)</li> <li>e2c_6000 was deleted (comment field)</li> </ul>
P3_ELIG3A	elig3a.sas7bdat	e3a	Microbiome Eligibility Checklist 3 Asthmatic	<ul style="list-style-type: none"> <li>e3a_1100D was deleted (description field)</li> <li>e3a_1120, e3a_1130 were deleted (source doc fields)</li> <li>e3a_6000 was deleted (comment field)</li> </ul>
P3_ELIG3C	elig3c.sas7bdat	e3c	Microbiome Eligibility Checklist 3 Control	<ul style="list-style-type: none"> <li>e3c_1100D was deleted (description field)</li> <li>e3c_1120, e3c_1130 were deleted (source doc fields)</li> <li>e3c_6000 was deleted (comment field)</li> </ul>
	group.sas7bdat		Identifies study groups (allergic asthmatic, allergic control and non-allergic control)	File contains the following variables: <ul style="list-style-type: none"> <li>part</li> <li>group</li> </ul>
P3_LAB	lab.sas7bdat	lab	Microbiome Laboratory Results	<ul style="list-style-type: none"> <li>lab_6000 was deleted (comment field)</li> </ul>
P3_LAB_BAL	lab_bal.sas7bdat	bal	Microbiome BAL Laboratory Results	<ul style="list-style-type: none"> <li>bal_6000 was deleted (comment field)</li> </ul>

Form Name Abbreviation	Dataset Name	Variable Prefix	Form Name or Description	Field Changes/ Comments
P3_MEQ	meq.sas7bdat	meq	Microbiome Microbial Exposure Questionnaire	<ul style="list-style-type: none"> <li>• meq_1000D, meq_1050D, meq_1060D, meq_1340D were deleted (description fields)</li> <li>• meq_1360, meq_1370 were deleted (source doc fields)</li> <li>• meq_6000 was deleted (comment field)</li> </ul>
	nutrition.sas7bdat		Block Brief 2000 Food Frequency Questionnaire	<p>Block Brief 2000 Food Frequency Questionnaire was administered online through NutritionQuest. Data set includes raw questionnaire and nutrient analysis data provided by NutritionQuest. Data dictionary is included in Microbiome_contents.pdf; paper copy of questionnaire is in Microbiome_forms.pdf. The following variables were removed:</p> <ul style="list-style-type: none"> <li>• booknum</li> <li>• sex</li> <li>• pregnant</li> <li>• age</li> <li>• weight</li> <li>• heightfeet, heightinches</li> <li>• hispanic, white, blackafrican, asian, nativeindian, Hawaiian</li> <li>• smokenow, howmanycigs</li> </ul>
P3_PARTTXQX	parttxqx.sas7bdat	ptx	Microbiome Participant Study Treatment Questionnaire	<ul style="list-style-type: none"> <li>• ptx_1050D was deleted (description field)</li> <li>• ptx_1060, ptx_1070 were deleted (source doc fields)</li> </ul>
P3_PULMONARYCHK	pulmonarychk.sas7bdat	pch	Microbiome Pulmonary Procedure Checklist	<ul style="list-style-type: none"> <li>• pch_1150D was deleted (description field)</li> <li>• pch_1170 (height) was deleted</li> <li>• pch_6000 was deleted (comment field)</li> </ul>

Form Name Abbreviation	Dataset Name	Variable Prefix	Form Name or Description	Field Changes/ Comments
	regimen.sas7bdat		Randomized treatment assignments	File contains the following variables: <ul style="list-style-type: none"> <li>part_id</li> <li>regimen (either "Placebo" or "ICS")</li> </ul>
P3_SIGEX	sigex.sas7bdat	sig	Microbiome Significant Asthma Exacerbation	<ul style="list-style-type: none"> <li>sig_1060 (date) was altered as described above</li> <li>sig_1230D was deleted (description field)</li> <li>sig_6000 was deleted (comment field)</li> </ul>
P3_TERM_A	term_a.sas7bdat	tma	Microbiome Termination of Study Participation Asthmatics	<ul style="list-style-type: none"> <li>tma_1030D, tma_1160D were deleted (description fields)</li> <li>tma_1060 (date) was altered as described above</li> <li>tma_1180, tma_1190, tma_1200, tma_1210 were deleted (source doc fields)</li> </ul>
P3_TERM_C	term_c.sas7bdat	tmc	Microbiome Termination of Study Participation Control	<ul style="list-style-type: none"> <li>tmc_1030D, tmc_1160D were deleted (description fields)</li> <li>tmc_1060 (date) was altered as described above</li> <li>tmc_1180, tmc_1190, tmc_1200, tmc_1210 were deleted (source doc fields)</li> </ul>

**Table 2. Forms/Procedures Completed at each Study Visit (Visits 0-6)**

(•=regular visit procedure; ○=completed only as needed)

	Visit Number							
	0	1	2	3	3.01	4	5	6
ACQ	•	•	•	•		•	•	•
AECLIN	○	○	○	○		○	○	○
ASTHMA_HX_ADULT		•						
CMED	○	○	○	○		○	○	○
COLD_HX		•						
HEQ		•						
HOUSEHOLD_SEI		•						
METHA		•				•		
METHACHK_ADULT		•				•		
METHA_ADD_TRT		○				○		
PALB4_SPIRO		○	•				•	
PREG_TEST		•	•			•	•	
PRIOR_COND_ADULT		•						
PRIOR_COND_ALL		•						
PRIOR_TRT		•						
Randomization			•					
REGISTRY	•							
SERIOUS	○	○	○	○		○	○	○
SNQ		•						
SPIRO		•	•	•		•	•	•
SPUTLAB		•				•		
SPUTREAD		•				•		
SPUTUM		•				•		

	Visit Number							
	0	1	2	3	3.01	4	5	6
SPUTUMCHK		•				•		
SPUTUM_ADD_TRT		○				○		
P3_ASTHMA_LOG			•	•	○	•	•	•
P3_BPD			•				•	
P3_BRONCHCHK			•				•	
P3_COMPLY				•	•	•	•	
P3_CTXQX			○	○		○	•	
P3_ELIG0	•							
P3_ELIG1		•						
P3_ELIG2		•						
P3_ELIG3			•					
P3_LAB		•	•					
P3_LAB_BAL			•				•	
P3_MEQ		•						
P3_PARTTXQX			○	○		○	•	
P3_PULMONARYCHK		•	•	•		•	•	•
P3_SIGEX			○	○		○	○	○
P3_TERM	○	○	○	○		○	○	•