



# **HCHS/SOL GOLD Ancillary Study Codebook**

**February 2019  
INV Version 2.0**

**Prepared by  
HCHS/SOL Coordinating Center  
Collaborative Studies Coordinating Center  
UNC Department of Biostatistics**

Daniela Sotres-Alvarez  
Yanping Teng

# HCHS/SOL GOLD Ancillary Study

## Investigator Use Database INV2 (Feb. 2019)

### Codebook

#### Table of Contents

<b>Section A: Questionnaire Battery</b> .....	3
GOLD AS Enrollment and Tracking Form (GOL) .....	3
GOLD Informed Consent Tracking Form (GIC) .....	5
GOLD AS Questionnaire (GLE).....	7
<b>Section B: Participant Derive Files</b> .....	12
GOLD Participant Derive File (GOLD_PART_DERV) .....	12

**Section A: Questionnaire Battery**

GOLD AS Enrollment and Tracking Form: FORM CODE=GOL VERSION=1

<i>VERS</i>		<i>Vers</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	1	

<i>OCCURRENCE</i>		<i>Occurrence</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	1	

<i>GOL0A</i>		<i>Completion Date</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3488	Range	11/10/2016 - 05/19/2018
1		Missing

<i>GOL0C</i>		<i>Lab ID GOLD Site</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	Present	Text suppressed

<i>GOL0C1</i>		<i>Lab ID number</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	Present	Text suppressed

<i>GOL1</i>		<i>GOLD enrollment status</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
1166	2	2=Agreed to participate, enrolled during V2 contact
2323	3	3=Agreed to participate, enrolled at a contact other than V2

<i>GOL2</i>		<i>Kit distribution date</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	Range	11/10/2016 - 05/19/2018

<i>GOL2Y</i>		<i>Year of Kit distribution date</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	Range	2016 - 2018 ( median=2017 mean=2017.0 std=0.4 )

**Section A: Questionnaire Battery**

<i>GOL2M</i>		<i>Month of Kit distribution date</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	Range	1 - 12 ( median=6 mean=6.2 std=3.6 )

<i>GOL2D</i>		<i>Day of Kit distribution date</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	Range	1 - 31 ( median=15 mean=15.6 std=8.6 )

<i>GOL3</i>		<i>Kit return date</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
292	Range	11/17/2016 - 01/12/2018
3197		Missing

<i>GOL3Y</i>		<i>Year of Kit return date</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
292	Range	2016 - 2018 ( median=2017 mean=2016.7 std=0.5 )
3197		Missing

<i>GOL3M</i>		<i>Month of Kit return date</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
292	Range	1 - 12 ( median=2 mean=5.0 std=4.8 )
3197		Missing

<i>GOL3D</i>		<i>Day of Kit return date</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
292	Range	1 - 31 ( median=21 mean=18.6 std=9.1 )
3197		Missing

**Section A: Questionnaire Battery**

GOLD Informed Consent Tracking Form: FORM CODE=GIC VERSION=1

<i>VERS</i>		<i>Vers</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	1	

<i>OCCURRENCE</i>		<i>Occurrence</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	1	

<i>GIC0A</i>		<i>Completion Date</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	Range	11/10/2016 - 05/19/2018

<i>GIC1</i>		<i>Agrees to GOLD to collect data and stool specimen</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	1	1=Yes

<i>GIC2</i>		<i>Agrees use of DNA from SOL and GOLD</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
9	0	0=No
3480	1	1=Yes

<i>GIC3</i>		<i>Agrees use of data by SOL and GOLD investigators</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
1	0	0=No
3488	1	1=Yes

<i>GIC4</i>		<i>Agrees use samples in future by SOL and GOLD investigators</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
12	0	0=No
3477	1	1=Yes

**Section A: Questionnaire Battery**

<i>GIC5</i>		<i>Agrees future contact</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
14	0	0=No
3472	1	1=Yes
3		Missing

**Section A: Questionnaire Battery**

GOLD AS Questionnaire: FORM CODE=GLE VERSION=1

<i>VERS</i>		<i>Vers</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3488	1	

<i>OCCURRENCE</i>		<i>Occurrence</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3486	1	
2	2	

<i>GLE0A</i>		<i>Completion Date</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3485	Range	11/10/2016 - 05/19/2018
3		Missing

<i>GLE0C</i>		<i>Comment:</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
1811	Present	Text suppressed
1677		Missing

<i>GLE1</i>		<i>Method of interview</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3473	1	1=In person interview
4	2	2=Telephone interview
5	3	3=Self administered by participant
6		Missing

<i>GLE2</i>		<i>Antibiotic treatment</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
970	1	1=Yes
2511	2	2=No
7		Missing

**Section A: Questionnaire Battery**

<i>GLE2A</i>		<i>Date for Antibiotic treatment</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
9	Range	04/01/2017 - 01/17/2018
456	.S	Legitimately skipped
3023		Missing

<i>GLE2AY</i>		<i>Year of Date for Antibiotic treatment</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
963	Range	2016 - 2018 ( median=2017 mean=2016.8 std=0.5 )
2525		Missing

<i>GLE2AM</i>		<i>Month of Date for Antibiotic treatment</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
963	Range	1 - 12 ( median=6 mean=6.3 std=3.7 )
2525		Missing

<i>GLE2AD</i>		<i>Day of Date for Antibiotic treatment</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
9	Range	1 - 28 ( median=16 mean=14.2 std=8.7 )
3479		Missing

<i>GLE3</i>		<i>Probiotic intake</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
280	1	1=Yes
3199	2	2=No
9		Missing

<i>GLE3A</i>		<i>IF YES, when was the last month you took probiotics? (mm/yyyy):</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
2	Range	05/01/2017 - 10/16/2017
571	.S	Legitimately skipped
2915		Missing

**Section A: Questionnaire Battery**

<i>GLE3AY</i>		<i>Year of IF YES, when was the last month you took probiotics? (mm/yyyy):</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
280	Range	2016 - 2018 ( median=2017 mean=2016.9 std=0.4 )
3208		Missing

<i>GLE3AM</i>		<i>Month of IF YES, when was the last month you took probiotics? (mm/yyyy):</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
280	Range	1 - 12 ( median=6 mean=6.3 std=3.5 )
3208		Missing

<i>GLE3AD</i>		<i>Day of IF YES, when was the last month you took probiotics? (mm/yyyy):</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
2	Range	1 - 16 ( median=8.5 mean=8.50 std=10.61 )
3486		Missing

<i>GLE4</i>		<i>Medication for indigestion, heart burn, or stomach problems</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
1061	1	1=Yes
2417	2	2=No
10		Missing

<i>GLE4A</i>		<i>Date for stomach problems medication</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
40	Range	02/02/2017 - 03/15/2018
460	.S	Legitimately skipped
2988		Missing

<i>GLE4AY</i>		<i>Year of Date for stomach problems medication</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
1057	Range	2016 - 2018 ( median=2017 mean=2016.9 std=0.4 )
2431		Missing

<i>GLE4AM</i>		<i>Month of Date for stomach problems medication</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
1057	Range	1 - 12 ( median=6 mean=6.4 std=3.5 )
2431		Missing

**Section A: Questionnaire Battery**

<i>GLE4AD</i>		<i>Day of Date for stomach problems medication</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
40	Range	1 - 31 ( median=15 mean=13.6 std=8.5 )
3448		Missing

<i>GLE5</i>		<i>Yogurt consumption in past 6 mo</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
1542	1	1=Yes
1937	2	2=No
9		Missing

<i>GLE6</i>		<i>Meat dietary preference</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
2849	0	0=I eat all kinds of meat
597	1	1=I eat poultry and/or fish, but not red meat
30	2	2=I am vegetarian (no meat)
5	3	3=I am vegan (no meat, dairy or animal products)
7		Missing

<i>GLE7A</i>		<i>Special diet for Weight loss</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
321	1	1=Yes
3161	2	2=No
6		Missing

<i>GLE7B</i>		<i>Special diet: Low fat</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
838	1	1=Yes
2644	2	2=No
6		Missing

<i>GLE7C</i>		<i>Special diet: High protein / low carbohydrate</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
532	1	1=Yes
2950	2	2=No
6		Missing

**Section A: Questionnaire Battery**

<i>GLE7D</i>		<i>Special diet: Macrobiotic</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
64	1	1=Yes
3418	2	2=No
6		Missing

<i>GLE7E</i>		<i>Special diet: Gluten-free</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
138	1	1=Yes
3344	2	2=No
6		Missing

<i>GLE7F</i>		<i>Special diet: Low sodium</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
853	1	1=Yes
2629	2	2=No
6		Missing

<i>GLE7G</i>		<i>Special diet: Lactose free</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
374	1	1=Yes
3108	2	2=No
6		Missing

<i>GLE7H</i>		<i>Special diet: Other</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
106	1	1=Yes
3376	2	2=No
6		Missing

<i>GLE7H1</i>		<i>Special diet: other specify</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
101	Present	Text suppressed
595	~	Legitimately skipped
2792		Missing

**Section B: Participant Derive Files**

GOLD Participant Derive File: FORM CODE=GOLD\_PART\_DERV VERSION=1

**Design**

<i>WEIGHT_EXPAND ED_GOLD</i>		<i>GOLD Expanded Census 2010 Calibrated, Trimmed, Nonresponse Adjusted Weights (frozen as of 01/22/2019)</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	Range	10.284450741 - 3428.7511391 ( median=78.494171322 mean=187.8607050731 std=303.9803368672 )

<i>WEIGHT_NORM_C ENTER_GOLD</i>		<i>GOLD Center Normalized, Census 2010 Calibrated, Trimmed, NonResponse Adjusted Weights (frozen as of 01/22/2019)</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	Range	0.052352939 - 16.216641452 ( median=0.4107235532 mean=1.00000000000 std=1.52044383089 )

<i>WEIGHT_NORM_O VERALL_GOLD</i>		<i>GOLD Overall Normalized, Census 2010 Calibrated, Trimmed, NonResponse Adjusted Weights (frozen as of 01/22/2019)</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	Range	0.0547450875 - 18.251561111 ( median=0.4178317722 mean=1.00000000000 std=1.61811559660 )

**Administrative**

<i>KIT_ID</i>		<i>GOLD Lab ID (GOL0c &amp; GOL0c1)</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	Present	Text suppressed

<i>BURK_ID</i>		<i>Lab use BURK ID (G#####)</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3057	Present	Text suppressed
432		Missing

<i>CLINDATE_GOLD</i>		<i>GOLD AS Clinic Visit Date</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	Range	11/10/2016 - 05/19/2018

<i>GOLD_ENROLL</i>		<i>Final GOLD AS Enrolled Participant with GOLD AS Consent</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3489	1	Yes

**Section B: Participant Derive Files**

<i>SAMPLEYN</i>		<i>GOLD Sample received by LAB</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
432	0	No
3057	1	Yes

<i>REC_DATE</i>		<i>Rec. Date, from Burk Receiving File</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
3057	Range	11/17/2016 - 06/18/2018
432		Missing

<i>SURVEY_DATE</i>		<i>Survey Date, from Burk Receiving File</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
2449	Range	12/30/1899 - 06/20/2027
1040		Missing

<i>S1_FREQ</i>		<i>S.1 (Frequency) Response, from Burk Receiving File</i>
<i>N</i>	<i>Value</i>	<i>Description</i>
5	0.5	
2	0.66666666 666666663	
1309	1	
1	1&3	
1	1, 2	
3	1/2	
1	1/3	
1119	2	
1	2, 3	
1	2/3	
322	3	
3	4	
1	5	
1	7	
7	Blank	
712		Missing

**Section B: Participant Derive Files**

S2_TYPE		S.2 (Type) Response, from Burk Receiving File
N	Value	Description
1	0	
112	1	
1	1&6	
1	1,2,5,6	
15	1/2	
9	1/3	
7	1/4	
1	1/4/6	
8	1/5	
2	1/7	
602	2	
60	2/3	
2	2/3/4	
1	2/3/6	
28	2/4	
7	2/5	
11	2/6	
5	2/7	
443	3	
42	3/4	
1	3/4/6	
5	3/5	
1	3/5/6/7	
20	3/6	
4	3/7	
838	4	
2	4/2	
1	4/3	
1	4/4	
20	4/5	
10	4/6	
8	4/7	
102	5	
15	5/6	
8	5/7	

**Section B: Participant Derive Files**

S2_TYPE		S.2 (Type) Response, from Burk Receiving File
N	Value	Description
336	6	
1	6/3	
1	6/5	
33	6/7	
138	7	
1	7/4	
2	7/6	
1	other	
582		Missing

**Socio-demographic**

AGE_GOLD		Age of GOLD Participant in years at the time of GOLD clinic visit
N	Value	Description
3489	Range	23 - 83 ( median=59 mean=57.6 std=11.6 )

AGEGROUP_C6_G OLD		Age Groups (6-level) at GOLD AS Visit
N	Value	Description
11	1	Ages 18-24
194	2	Ages 25-34
240	3	Ages 35-44
714	4	Ages 45-54
1333	5	Ages 55-64
997	6	Ages 65+