Dataset Integrity Check for the GpCRC Data Files

Prepared by Anne Taylor
IMS Inc.
3901 Calverton Blvd, Suite 200 Calverton MD 20705
November 16, 2016

Contents

1 Standard Disclaimer

The intent of this DSIC is to provide confidence that the data distributed by the NIDDK repository is a true copy of the study data. Our intent is not to assess the integrity of the statistical analyses reported by study investigators. As with all statistical analyses of complex datasets, complete replication of a set of statistical results should not be expected in secondary analysis. This occurs for a number of reasons including differences in the handling of missing data, restrictions on cases included in samples for a particular analysis, software coding used to define complex variables, etc. Experience suggests that most discrepancies can ordinarily be resolved by consultation with the study data coordinating center (DCC), however this process is labor-intensive for both DCC and Repository staff. It is thus not our policy to resolve every discrepancy that is observed in an integrity check. Specifically, we do not attempt to resolve minor or inconsequential discrepancies with published results or discrepancies that involve complex analyses, unless NIDDK Repository staff suspect that the observed discrepancy suggests that the dataset may have been corrupted in storage, transmission, or processing by repository staff. We do, however, document in footnotes to the integrity check those instances in which our secondary analyses produced results that were not fully consistent with those reported in the target publication.

2 Study Background

Gastroparesis, a syndrome in which patients experience delayed emptying of the stomach, is characterized by symptoms such as nausea, vomiting, bloating, abdominal pain, and early satiety. While the incidence and prevalence of gastroparesis are not well-defined, the condition is estimated to affect up to 5 million individuals in the US. Research progress on this syndrome has been hindered by challenges such as single centers, small population base, and limited range of clinical and research techniques available to study. In response to these limitations, the Gastroparesis Clinical Research Consortium (GpCRC) was established by the NIDDK to advance research on the etiology, natural history, and therapy of gastroparesis. Specific aims of the consortium include providing an infrastructure for the efficient design and conduct of multicenter clinical studies; performing clinical trials to investigate clinical, diagnostic, and therapeutic interventions for gastroparesis; and creating a collection of patient samples that may be used for ancillary studies of etiology and pathogenesis.

3 Archived Datasets

All SAS data files, as provided by the Data Coordinating Center (DCC), are located in the Datasets folder in the GpCRC data package. For all tables, variables were taken from the 'bd1', 'bh2', 'gd1', 'ge2', 'pe1', 'rg1', 'se1', and 'ug1' datasets in the GpCRC data package.

4 Statistical Methods

Analyses were performed to duplicate results for the data published by Diabetes Prevention Program Research Group Dig Dis Sci, April 2013. To verify the integrity of the datasets, descriptive statistics were computed.

5 Results

For Table 1 in the publication [1], Baseline demographics and symptom characteristics of patients by cholecystectomy prior to enrollment, Table A lists the variables that were used in the replication and Table B compares the results calculated from the archived data files to the results published in Table 1. The results of the replication are almost an exact match to the published results, with only a few discrepancies due to rounding.

For Table 2 in the publication [1], Comorbidities, psychological function inventories and quality of life assessment in gastroparesis patients by prior cholecystectomy, Table C lists the variables that were used in the replication and Table D compares the results calculated from the archived data files to the results published in Table 2. The results of the replication are almost an exact match to the published results, with only a few discrepancies due to rounding.

6 Conclusions

The NIDDK repository is confident that the GpCRC data files to be distributed are a true copy of the study data.

7 References

[1]Henry P. Parkman • Katherine Yates • William L. Hasler • Linda Nguyen • Pankaj J. Pasricha • William J. Snape • Gianrico Farrugia • Kenneth L. Koch • Jorge Calles • Thomas L. Abell • Irene Sarosiek • Richard W. McCallum • Linda Lee • Aynur Unalp-Arida • James Tonascia • Frank Hamilton. Cholecystectomy and Clinical Presentations of Gastroparesis. Dig Dis Sci (2013) 58:1062–1073

Table A: Variables used to replicate Table 1: Baseline demographics and symptom characteristics of patients by cholecystectomy prior to enrollment

Characteristic	File.Variable(s)
Gender	RG1.RG112
Age at enrollment	RG1.RG110
Ethnicity (any Hispanic)	RG1.RG113
White (yes vs. no)	RG1.RG114E
Marital status (yes vs. no)	RG1.RG121
College degree or higher (yes vs. no)	RG1.RG115
Body mass index (kg/m2)	PE1.PE108A PE1.PE108B PE1.PE109A PE1.PE109B
Etiology: (Idiopathic, DM type 1, DM type 2)	BH2.BH209C BH2.BH231A
Age gastroparesis symptoms started (years)	RG1.RG109 BH2.BH208
Duration (years) of symptoms from enrollment	BH2.BH208
Initial infectious prodrome (yes vs. no)	BH2.BH210
Predominant symptom prompting evaluation: (Nausea, Vomiting, Abdominal pain, Other)	BH2.BH213
Type of symptom onset (insidious vs. acute)	BH2.BH214
Nature of gastroparesis symptoms: (Chronic, but stable, Chronic, but progressive worsening, Chronic with periodic exacerbations, Cyclic	BH2.BH215
Gastroparesis severity: (Mild (grade 1) Compensated (grade 2) Gastric failure (grade 3))	BH2.BH216
Gastric electric stimulator (current use)	BH2.BH244
Proton pump inhibitor or other GI medication	BH2.BH246
Prokinetic	BH2.BH247
Antiemetic	N/A*
Pain relieving analgesic, NSAI, or aspirin in the past 6 months	BH2.BH253
Narcotic	BH2.BH254
Any hormones (HRT, estrogen, progestin) in the past 6 months	BH2.BH261
Any pain modulators	N/A*
Any antidepressants	N/A*
Anxiolytics	N/A*
Percent gastric retention at 2 h	GE2.GE213D
Percent gastric retention at 4 h	GE2.GE213F
Severity of delayed gastric emptying at 4 h: (Mild (B20 %), Moderate (21–35 %), Severe (C35 %))	GE2.GE213F
Nausea severity	GD1.GD110
Retching severity	GD1.GD111
Vomiting severity	GD1.GD112
Stomach fullness severity	GD1.GD113
Inability to finish a meal severity	GD1.GD114
Feeling excessively full after meals	GD1.GD115

Characteristic	File.Variable(s)
Appetite loss severity	GD1.GD116
Bloating severity	GD1.GD117
Stomach visibly larger severity	GD1.GD118
Cardinal symptom index (GCSI)c	GD1.GD110 GD1.GD115 GD1.GD117
Upper abdominal pain severity	GD1.GD119
Upper abdominal discomfort severity	GD1.GD120
Lower abdominal pain severity	GD1.GD121
Lower abdominal discomfort severity	GD1.GD122
Gastrointestinal reflux (GERD) sub-scale	GD1.GD123-GD1.GD129
Constipation severity	GD1.GD130
Diarrhea severity	GD1.GD131

* Variables listed as "N/A" were not included in the data package.

Table B: Comparison of values computed in integrity check to reference article Table 1 values

	Manuscript No (n = 249)	DSIC No (n = 249)	DIFF No (n = 0)	Manuscript Yes (n = 142)	DSIC Yes (n = 142)	DIFF Yes (n = 0)
Demographic, lifestyle, anthropomorphic						
Gender (females)	200(80.3 %)	200(80.3 %)	0(0.0 %)	125(88.0 %)	125(88.0 %)	0(0.0 %)
Age at enrollment	249(40.9 ± 13.5)	249(40.3 ± 13.6)	0(0.6 ± -0.1)	142(45.0 ± 14.2)	142(44.5 ± 14.2)	0(0.5 ± 0.0)
Ethnicity (any Hispanic)	9(3.6 %)	9(3.6 %)	0(0.0%)	10(7.0 %)	10(7.0 %)	0(0.0 %)
White (yes vs. no)	210(84.3 %)	215(86.3 %)	-5(-2.0 %)	124(87.3 %)	126(88.7 %)	-2(-1.4 %)
Marital status (yes vs. no)	128(51.4 %)	128(51.4 %)	0(0.0 %)	100(70.4 %)	100(70.4 %)	0(0.0 %)
College degree or higher (yes vs. no)	73(29.3 %)	73(29.3 %)	0(0.0 %)	38(26.8 %)	38(26.8 %)	0(0.0 %)
Body mass index (kg/m2)	249(26.2 ± 7.3)	249(26.2 ± 7.3)	0(0.0 ± 0.0)	142(28.1 ± 7.2)	142(28.1 ± 7.2)	0(0.0 ± 0.0)
BMI category:						
Underweight/normal	126(50.6 %)	126(50.6 %)	0(0.0 %)	57(40.1 %)	57(40.1 %)	0(0.0 %)
Overweight/obese	123(49.4 %)	123(49.4 %)	0(0.0 %)	85(59.9 %)	85(59.9 %)	0(0.0 %)
Etiology						
Etiology						
Idiopathic	158(63.5 %)	158(63.7 %)	0(-0.2 %)	96(67.6 %)	96(67.6 %)	0(0.0 %)
DM type 1	59(23.7 %)	59(23.8 %)	0(-0.1 %)	19(13.4 %)	19(13.4 %)	0(0.0 %)
DM type 2	32(12.9 %)	31(12.5 %)	1(0.4 %)	27(19.0 %)	27(19.0 %)	0(0.0 %)
Medical history						
Age gastroparesis symptoms started (years)	249(36.2 ± 13.9)	249(36.2 ± 13.9)	0(0.0 ± 0.0)	142(39.6 ± 14.8)	142(39.6 ± 14.8)	0(0.0 ± 0.0)

	Manuscript No (n = 249)	DSIC No (n = 249)	DIFF No (n = 0)	Manuscript Yes (n = 142)	DSIC Yes (n = 142)	DIFF Yes (n = 0)
Duration (years) of symptoms from enrollment	249(4.8 ± 5.7)	249(4.8 ± 5.7)	0(0.0 ± 0.0)	142(5.5 ± 6.8)	142(5.5 ± 6.8)	0(0.0 ± 0.0)
Initial infectious prodrome (yes vs. no)	50(20.1 %)	50(20.1 %)	0(0.0 %)	16(11.3 %)	16(11.3 %)	0(0.0 %)
Predominant symptom prompting evaluation						
Nausea	86(34.5 %)	86(34.5 %)	0(0.0 %)	51(35.9 %)	51(35.9 %)	0(0.0 %)
Vomiting	56(22.5 %)	56(22.5 %)	0(0.0 %)	36(25.4 %)	36(25.4 %)	0(0.0 %)
Abdominal pain	46(18.5 %)	46(18.5 %)	0(0.0 %)	27(19.0 %)	27(19.0 %)	0(0.0 %)
Other	61(24.5 %)	61(24.5 %)	0(0.0 %)	28(19.7 %)	28(19.7 %)	0(0.0 %)
Type of symptom onset (insidious vs. acute)	107 (43.0 %)	105 (42.2 %)	2(0.8 %)	82(57.8 %)	82(57.7 %)	0(0.1 %)
Nature of gastroparesis symptoms						
Chronic, but stable	58(23.5 %)	58(23.5 %)	0(0.0 %)	26(18.3 %)	26(18.3 %)	0(0.0 %)
Chronic, but progressive worsening	73(29.6 %)	73(29.6 %)	0(0.0 %)	58(40.9 %)	58(40.8 %)	0(0.1 %)
Chronic with periodic exacerbations	83(33.6 %)	83(33.6 %)	0(0.0 %)	50(35.2 %)	50(35.2 %)	0(0.0 %)
Cyclic pattern	33(13.4 %)	33(13.4 %)	0(0.0 %)	8(5.6 %)	8(5.6 %)	0(0.0 %)
Gastroparesis severity						
Mild (grade 1)	34(13.8 %)	34(13.8 %)	0(0.0 %)	15(10.6 %)	15(10.6 %)	0(0.0 %)
Compensated (grade 2)	135(54.7 %)	135(54.7 %)	0(0.0 %)	71(50.0 %)	71(50.0 %)	0(0.0 %)
Gastric failure (grade 3)	78(31.6 %)	78(31.6 %)	0(0.0 %)	56(39.4 %)	56(39.4 %)	0(0.0 %)
Gastric electric stimulator (current use)	16(6.4 %)	16(6.4 %)	0(0.0 %)	13(9.2 %)	13(9.2 %)	0(0.0 %)
Medications (current use)						
Proton pump inhibitor or other GI medication	192(77.1 %)	192(77.1 %)	0(0.0 %)	112(78.9 %)	112(78.9 %)	0(0.0 %)
Prokinetic	140(56.2 %)	140(56.2 %)	0(0.0 %)	77(54.2 %)	77(54.2 %)	0(0.0 %)
Antiemetic	148(59.4 %)			100(70.4 %)		
Pain relieving analgesic, NSAI, or aspirin in the past 6 months	144(57.8 %)	144(57.8 %)	0(0.0 %)	91(64.1 %)	91(64.1 %)	0(0.0 %)
Narcotic	99(39.8 %)	98(39.5 %)	1(0.3 %)	74(52.1 %)	74(52.1 %)	0(0.0 %)
Any hormones (HRT, estrogen, progestin) in the past 6 months	53(21.3 %)	53(21.3 %)	0(0.0 %)	31(21.8 %)	31(21.8 %)	0(0.0 %)
Any pain modulators	47(18.9 %)	N/A	N/A	33(23.2 %)	N/A	N/A
Any antidepressants	78(31.3 %)	N/A	N/A	55(38.7 %)	N/A	N/A
Anxiolytics	37(14.9 %)	N/A	N/A	20(14.1 %)	N/A	N/A
Gastric emptying (scintigraphy)						
Percent gastric retention at 2 h	249(65.3 ± 18.6)	249(65.3 ± 18.6)	0(0.0 ± 0.0)	142(62.3 ± 18.1)	142(62.3 ± 18.1)	0(0.0 ± 0.0)

	Manuscript No (n = 249)	DSIC No (n = 249)	DIFF No (n = 0)	Manuscript Yes (n = 142)	DSIC Yes (n = 142)	DIFF Yes (n = 0)
Percent gastric retention at 4 h	249(33.1 ± 22.7)	249(33.1 ± 22.7)	0(0.0 ± 0.0)	141(32.1 ± 22.4)	141(32.1 ± 22.4)	0(0.0 ± 0.0)
Severity of delayed gastric emptying at 4 h						
Mild (B20 %)	88(35.3 %)	88(35.3 %)	0(0.0 %)	61(43.3 %)	61(43.3 %)	0(0.0 %)
Moderate (21–35 %)	78(31.3 %)	78(31.3 %)	0(0.0 %)	28(19.9 %)	28(19.9 %)	0(0.0 %)
Severe (C35 %)	83(33.3 %)	83(33.3 %)	0(0.0 %)	52(36.9 %)	52(36.9 %)	0(0.0 %)
PAGI-SYM severities						
Nausea severity	249(3.3 ± 1.4)	249(3.3 ± 1.4)	0(0.0 ± 0.0)	142(3.6 ± 1.3)	142(3.6 ± 1.3)	0(0.0 ± 0.0)
Retching severity	249(1.8 ± 1.7)	249(1.8 ± 1.7)	0(0.0 ± 0.0)	142(2.5 ± 1.8)	142(2.5 ± 1.8)	0(0.0 ± 0.0)
Vomiting severity	249(2.0 ± 1.8)	249(2.0 ± 1.8)	0(0.0 ± 0.0)	142(2.5 ± 1.9)	142(2.5 ± 1.9)	0(0.0 ± 0.0)
Stomach fullness severity	249(3.5 ± 1.4)	247(3.5 ± 1.3)	2(0.0 ± 0.1)	142(3.7 ± 1.2)	142(3.7 ± 1.1)	0(0.0 ± 0.1)
Inability to finish a meal severity	249(3.4 ± 1.4)	249(3.4 ± 1.4)	0(0.0 ± 0.0)	142(3.5 ± 1.4)	142(3.5 ± 1.4)	0(0.0 ± 0.0)
Feeling excessively full after meals	249(3.6 ± 1.4)	249(3.6 ± 1.4)	0(0.0 ± 0.0)	142(3.6 ± 1.3)	142(3.6 ± 1.3)	0(0.0 ± 0.0)
Appetite loss severity	249(2.9 ± 1.5)	249(2.9 ± 1.5)	0(0.0 ± 0.0)	142(3.1 ± 1.4)	142(3.1 ± 1.4)	0(0.0 ± 0.0)
Bloating severity	249(3.1 ± 1.6)	249(3.1 ± 1.6)	0(0.0 ± 0.0)	142(3.4 ± 1.4)	142(3.4 ± 1.4)	0(0.0 ± 0.0)
Stomach visibly larger severity	249(2.6 ± 1.8)	249(2.6 ± 1.8)	0(0.0 ± 0.0)	142(3.1 ± 1.7)	142(3.1 ± 1.7)	0(0.0 ± 0.0)
Cardinal symptom index (GCSI)c	249(2.9 ± 1.0)	249(3.3 ± 1.1)	0(-0.4 ± -0.1)	142(3.2 ± 0.9)	142(3.6 ± 1.0)	0(-0.4 ± -0.1)
Upper abdominal pain severity	249(2.9 ± 1.8)	249(2.9 ± 1.8)	0(0.0 ± 0.0)	142(3.3 ± 1.7)	142(3.3 ± 1.7)	0(0.0 ± 0.0)
Upper abdominal discomfort severity	249(3.1 ± 1.6)	249(3.1 ± 1.6)	0(0.0 ± 0.0)	142(3.4 ± 1.5)	142(3.4 ± 1.5)	0(0.0 ± 0.0)
Lower abdominal pain severity	249(2.1 ± 1.6)	249(2.1 ± 1.6)	0(0.0 ± 0.0)	142(2.2 ± 1.7)	142(2.2 ± 1.7)	0(0.0 ± 0.0)
Lower abdominal discomfort severity	249(2.2 ± 1.6)	249(2.2 ± 1.6)	0(0.0 ± 0.0)	142(2.3 ± 1.7)	142(2.3 ± 1.7)	0(0.0 ± 0.0)
Gastrointestinal reflux (GERD) sub-scale	248(1.9 ± 1.3)	248(1.9 ± 1.3)	0(0.0 ± 0.0)	142(2.2 ± 1.5)	142(2.2 ± 1.5)	0(0.0 ± 0.0)
Constipation severity	249(2.4 ± 1.7)	249(2.4 ± 1.7)	0(0.0 ± 0.0)	142(2.4 ± 1.9)	141(2.4 ± 1.8)	1(0.0 ± 0.1)
Diarrhea severity	249(1.7 ± 1.7)	249(1.7 ± 1.7)	0(0.0 ± 0.0)	142(2.1 ± 1.8)	142(2.1 ± 1.8)	0(0.0 ± 0.0)

Table C: Variables used to replicate Table 2. Comorbidities, psychological function inventories and quality of life assessment in gastroparesis patients by prior cholecystectomy

Characteristic	File.Variable(s)
Number of comorbid conditions	BH2.BH232A-BH2.BH232Z BH2.BH232AA-BH2.BH232AY
Any comorbidities	BH2.BH232A-BH2.BH232Z BH2.BH232AA-BH2.BH232AY
Migraine	BH2.BH232X
Chronic fatigue syndrome	BH2.BH232Z
History of CVD	BH2.BH232AB BH2.BH232AC
Fibromyalgia	BH2.BH232AL
Eating disorder	BH2.BH232AP
Major depression	BH2.BH232AQ
Severe anxiety	BH2.BH232AU
Number of hospitalizations in the past year	BH2.BH236
Any hospitalization in the past year	BH2.BH235
State anxiety score	SE1.SE108A
Trait anxiety score	SE1.SE108B
BDI Inventory score	BD1.BD108
BDI>28 (severe depression)	BD1.BD109
Activity sub-score	UG1.UG110-UG1.UG119
Clothing sub-score	UG1.UG120 UG1.UG121
Diet sub-score	UG1.UG122-UG1.UG128
Relationship sub-score	UG1.UG129-UG1.UG131
Psychological sub-score	UG1.UG132-UG1.UG139
Total score	UG1.UG110-UG1.UG139
Physical health component sub-score	N/A*
Mental health component sub-score	N/A*

* Variables listed as "N/A" were not included in the data package.

Table D: Comparison of values computed in integrity check to reference article Table 2 values

	Manuscript No (n = 249)	DSIC No (n = 249)	DIFF No (n = 0)	Manuscript Yes (n = 142)	DSIC Yes (n = 142)	DIFF Yes (n = 0)
Comorbidities and hospitalizations						
Number of comorbid conditions	249(3.3 ± 2.5)	249(3.3 ± 2.5)	0(0.0 ± 0.0)	142(5.2 ± 3.3)	142(5.2 ± 3.3)	0(0.0 ± 0.0)
Any comorbidities	228(91.6 %)	228(91.6 %)	0(0.0 %)	140(98.6 %)	140(98.6 %)	0(0.0 %)
Comorbidities (ever diagnosed)						
Migraine	87(34.9 %)	87(34.9 %)	0(0.0 %)	57(40.1 %)	57(40.1 %)	0(0.0 %)
Chronic fatigue syndrome	12(4.8 %)	12(4.8 %)	0(0.0 %)	18(12.7 %)	18(12.7 %)	0(0.0 %)
History of CVD	15(6.0 %)	15(6.0 %)	0(0.0 %)	11(7.8 %)	11(7.7 %)	0(0.1 %)
Fibromyalgia	20(8.0 %)	20(8.0 %)	0(0.0 %)	24(16.9 %)	24(16.9 %)	0(0.0 %)
Eating disorder	9(3.6 %)	9(3.6 %)	0(0.0 %)	3(2.1 %)	3(2.1 %)	0(0.0 %)
Major depression	48(19.3 %)	48(19.3 %)	0(0.0 %)	47(33.1 %)	47(33.1 %)	0(0.0 %)
Severe anxiety	22(8.8 %)	22(8.8 %)	0(0.0 %)	22(15.5 %)	22(15.5 %)	0(0.0 %)
Number of hospitalizations in the past year	249(2.1 ± 4.6)	249(2.1 ± 4.6)	0(0.0 ± 0.0)	142(3.3 ± 4.7)	142(3.3 ± 4.7)	0(0.0 ± 0.0)
Any hospitalization in the past year	112(45.0 %)	112(45.0 %)	0(0.0 %)	84(59.2 %)	84(59.2 %)	0(0.0 %)
State-Trait inventory (STAI)						
State anxiety score	249(44.5 ± 13.0)	249(44.5 ± 13.0)	0(0.0 ± 0.0)	142(47.6 ± 14.1)	142(47.6 ± 14.1)	0(0.0 ± 0.0)
Trait anxiety score	249(43.5 ± 11.7)	249(43.5 ± 11.7)	0(0.0 ± 0.0)	142(46.3 ± 13.4)	142(46.3 ± 13.4)	0(0.0 ± 0.0)
Beck Depression Inventory (BDI)						
BDI Inventory score	249(18.5 ± 11.1)	249(18.5 ± 11.1)	0(0.0 ± 0.0)	142(20.6 ± 11.4)	142(20.6 ± 11.4)	0(0.0 ± 0.0)
BDI>28 (severe depression)	45(18.1 %)	45(18.1 %)	0(0.0 %)	32(22.5 %)	32(22.5 %)	0(0.0 %)
Quality of life (PAGI- QOL) (past 2 weeks)						
Activity sub-score	249(2.4 ± 1.2)	249(2.4 ± 1.2)	0(0.0 ± 0.0)	142(2.0 ± 1.2)	142(2.0 ± 1.2)	0(0.0 ± 0.0)
Clothing sub-score	249(3.1 ± 1.7)	249(3.1 ± 1.7)	0(0.0 ± 0.0)	142(2.6 ± 1.7)	142(2.6 ± 1.7)	0(0.0 ± 0.0)
Diet sub-score	249(1.6 ± 1.2)	249(1.6 ± 1.2)	0(0.0 ± 0.0)	142(1.4 ± 1.2)	142(1.4 ± 1.2)	0(0.0 ± 0.0)
Relationship sub-score	249(3.1 ± 1.5)	249(3.1 ± 1.5)	0(0.0 ± 0.0)	142(2.7 ± 1.6)	142(2.7 ± 1.6)	0(0.0 ± 0.0)
Psychological sub-score	249(2.7 ± 1.4)	249(2.7 ± 1.4)	0(0.0 ± 0.0)	142(2.5 ± 1.4)	142(2.5 ± 1.4)	0(0.0 ± 0.0)
Total score	249(2.6 ± 1.1)	249(2.4 ± 1.1)	0(0.2 ± 0.0)	142(2.2 ± 1.1)	142(2.1 ± 1.1)	0(0.1 ± 0.0)
SF-36v2 health survey (past 4 weeks)						
Physical health component sub-score	248(33.8 ± 10.1)	N/A	N/A	141(31.8 ± 10.1)	N/A	N/A
Mental health component sub-score	248(37.7 ± 12.4)	N/A	N/A	141(35.6 ± 12.9)	N/A	N/A

Attachment A: SAS Code

```
*********************************************************************************************************************

***DSIC of GpCRC.
***Programmer: Michael Spriggs and Anne Taylor
***Date Created: 02Sep2016

********************************************************************************************************************;

title1 'DSIC of GpCRC';
title2 "%sysfunc(getoption(sysin))";

options nofmterr ls=255 mprint source2 symbolgen;

libname sasfile '/prj/niddk/ims_analysis/GpCRC/private_created_data/datasets/';

proc format;
    value ge213f
        0-20='<=20%'
        20<-35='>20-35%'
        35<-high='>35%'
        ;

%macro procdsn(dsn=);
    data &dsn;
        set sasfile.&dsn;

    proc contents data=&dsn;
        title4 "Contents of &dsn";

    proc freq data=&dsn;
        tables visit/missing list;
        title4 "Visits of &dsn";
%mend;

%procdsn(dsn=bh2);
%procdsn(dsn=ge2);
%procdsn(dsn=rg1);
%procdsn(dsn=pe1);
%procdsn(dsn=gd1);
%procdsn(dsn=se1);
%procdsn(dsn=ug1);
%procdsn(dsn=bd1);

%macro shortdsn(dsn);
    data &dsn._short;
```

```
        set &dsn(rename=(visit=old_visit));
        length visit $1;
        if strip(old_visit)='b' then do;
            visit='b';
            output;
        end;
%mend shortdsn;

%shortdsn(ge2);
%shortdsn(pe1);
%shortdsn(gd1);
%shortdsn(se1);
%shortdsn(ug1);
%shortdsn(bd1);

data analysis;
    merge bh2 (in=in_bh)
          ge2_short (in=in_ge)
          rg1 (in=in_rg)
          pe1_short (in=in_pe)
          gd1_short (in=in_gd)
          se1_short (in=in_se)
          ug1_short (in=in_ug)
          bd1_short (in=in_bd)
          ;
    by id visit;
    if not first.visit then abort;
    bh_ok=in_bh;
    ge_ok=in_ge;
    rg_ok=in_rg;
    pe_ok=in_pe;
    gd_ok=in_gd;
    se_ok=in_se;
    ug_ok=in_ug;
    bd_ok=in_bd;

proc freq data=analysis;
    tables bh_ok*ge_ok*rg_ok*pe_ok*gd_ok*se_ok*ug_ok*bd_ok/missing list;
    title4 'Check Datasets';

proc freq data=analysis;
    tables PE108A*PE108B PE109A*PE109B/missing list;
    title4 'Check Variables';


proc freq data=analysis;
    tables bh209a*bh209c/missing list;

data analysis;
```

```
set analysis;
length Cholecystectomy $5. AGE 8.;

array charvars(*) ge213d ge213f gd110-gd131 se108a se108b bd108;
array numvars(*) ge213d_n ge213f_n n_gd110-n_gd131 se108a_n se108b_n bd108_n;

array comorbid(*) bh232a bh232b bh232c bh232d bh232e bh232f bh232g bh232h bh232i bh232j bh232k bh232l bh232m bh232n bh232o bh232p
                  bh232q bh232r bh232s bh232t bh232u bh232v bh232w bh232x bh232y bh232z bh232aa bh232ab bh232ac bh232ad bh232ae
                  bh232af bh232ag bh232ah bh232ai bh232aj bh232ak bh232al bh232am bh232an bh232ao bh232ap bh232aq bh232ar bh232as
                  bh232at bh232au bh232av bh232aw bh232ax bh232ay
                  ;

array pagi_orig(*) ug110-ug139;
array pagi_new(*) score110-score139;

if bh233f='1' then Cholecystectomy="Yes";
else Cholecystectomy='No';
AGE=input(RG110,8.);

if PE108B='1' then HEIGHT_M=PE108A*(0.0254);
else if PE108B='2' then HEIGHT_M=PE108A*0.01;

if PE109B='1' then WEIGHT_KG=PE109A*(0.453592);
else if PE109B='2' then WEIGHT_KG=PE109A;

BMI=round(WEIGHT_KG/(HEIGHT_M*HEIGHT_M),.1);

if 0<=BMI<25 then BMI_CAT=1;
else if BMI>=25 then BMI_CAT=2;

if BH209C='1' then DIAB_TYPE=0;
else if BH231A='1' then DIAB_TYPE=1;
else if BH231A='2' then DIAB_TYPE=2;

birthdays=put(strip(RG109),8.);
gastsymdays=put(BH208,8.);

Age_g_start=(gastsymdays-birthdays)/365.25;
SYMPT_DUR=abs(gastsymdays)/365.25;

if BH213='a' then SYMPTOM_TYPE=1;
else if BH213='b' then SYMPTOM_TYPE=2;
else if BH213='f' then SYMPTOM_TYPE=3;
else SYMPTOM_TYPE=4;

***change missings to true missing to exclude from %s in freq***;
if bh215 in ('5','m') then bh215=' ';  ***5=other***;
if bh216 in ('4','m') then bh216=' ';  ***5=other***;
if bh254='m' then bh254=' ';
```

13

```
do i=1 to dim(charvars);
   if charvars(i)='m' then numvars(i)=.m;
   else if charvars(i)='n' then numvars(i)=.n;
   else numvars(i)=input(charvars(i),8.);
   end;

cardinal_sx=(n_gd110+n_gd115+n_gd117)/3;  ***cardinal symptom index - see footnote c in Table 1***;

***gastrointestinal reflux sub-scale - see footnote c in Table 1***;
gi_reflux=(n_gd123+n_gd124+n_gd125+n_gd126+n_gd127+n_gd128+n_gd129)/7;

if bh232ab='1' or bh232ac='1' then CVD=1;  ***cardiovascular disease (coronary artery disease or cerebrovascular disease)***;

***# and any comorbidities***;
num_comorbid=0;
any_comorbid=0;
do i=1 to dim(comorbid);
   if comorbid(i)='1' then do;
      num_comorbid=num_comorbid+1;
      any_comorbid=1;
      end;
   end;

***# hospitalizations***;
if bh235='1' then num_hosp=input(bh236,2.);
else if bh235='2' then num_hosp=0;
else abort;

***reverse PAGI scores (see 4th paragraph of Methods)***;
do i=1 to dim(pagi_orig);
   if pagi_orig(i)='m' then pagi_new(i)=.m;
   else if pagi_orig(i)='n' then pagi_new(i)=.n;
   else pagi_new(i)=5-input(pagi_orig(i),1.);
   end;

***PAGI-QOL sub-scores***;
qol_activity=sum(of score110-score119)/n(of score110-score119);
qol_clothing=sum(score120,score121)/n(score120,score121);
qol_diet=sum(of score122-score128)/n(of score122-score128);
qol_relationship=sum(of score129-score131)/n(of score129-score131);
qol_psych=sum(of score132-score139)/n(of score132-score139);
qol_total=sum(of score110-score139)/n(of score110-score139);

***3 subjects excluded due to idiopathic gastroparesis and history of diabetes (per Laura Wilson e-mail)***;
if id in ('2237','8132','9875') then delete;

if (BH209A='1' or BH209C='1') and (ge213d_n>60 or ge213f_n>10);
```

```
    label ge213d_n='13d 2 hours: percent retention'
          ge213f_n='13f 4 hours: percent retention'
          n_gd110='10 Nausea'
          n_gd111='11 Retching'
          n_gd112='12 Vomiting'
          n_gd113='13 Stomach fullness'
          n_gd114='14 Unable to finish normal-size meal'
          n_gd115='15 Feel full after meals'
          n_gd116='16 Loss of appetite'
          n_gd117='17 Bloating'
          n_gd118='18 Stomach or belly visibly larger'
          n_gd119='19 Upper abdominal pain'
          n_gd120='20 Upper abdominal discomfort'
          n_gd121='21 Lower abdominal pain'
          n_gd122='22 Lower abdominal discomfort'
          n_gd130='30 Constipation'
          n_gd131='31 Diarrhea'
          se108a_n='8a Form Y-1 score'
          se108b_n='8b Form Y-2 score'
          bd108_n='Score of all 21 items'
          ;
/*
proc freq data=analysis;
   tables birthdays*RG109*gastsymdays*BH208*Age_g_start/missing list;
*/

ODS PATH RESET;
ODS PATH (PREPEND) WORK.Templat(UPDATE) ;

PROC TEMPLATE;
  EDIT Base.Freq.crosstabfreqs;
    EDIT Percent;
      FORMAT = 5.1;
    END;
    EDIT ColPercent;
      FORMAT = 5.1;
    END;
  END;
RUN;

proc freq data=analysis;
   tables bh233f/missing list;

proc freq data=analysis;
   tables rg110*AGE
          rg114c*rg114e*rg114f
          bmi_cat*bmi
          ge213d*ge213d_n
          ge213f*ge213f_n
```

15

```
            gd110*n_gd110
            gd111*n_gd111
            gd112*n_gd112
            gd113*n_gd113
            gd114*n_gd114
            gd115*n_gd115
            gd116*n_gd116
            gd117*n_gd117
            gd118*n_gd118
            gd119*n_gd119
            gd120*n_gd120
            gd121*n_gd121
            gd122*n_gd122
            gd130*n_gd130
            gd131*n_gd131
            cardinal_sx*gd110*gd115*gd117
            gi_reflux*gd123*gd124*gd125*gd126*gd127*gd128*gd129
            cvd*bh232ab*bh232ac
            bh235*num_hosp*BH236
            se108a*se108a_n
            se108b*se108b_n
            qol_activity*ug110*ug111*ug112*ug113*ug114*ug115*ug116*ug117*ug118*ug119
            qol_clothing*ug120*ug121
            qol_diet*ug122*ug123*ug124*ug125*ug126*ug127*ug128
            qol_relationship*ug129*ug130*ug131
            qol_psych*ug132*ug133*ug134*ug135*ug136*ug137*ug138*ug139
            qol_total*ug110*ug111*ug112*ug113*ug114*ug115*ug116*ug117*ug118*ug119*ug120*ug121*ug122*ug123*ug124*
                    ug125*ug126*ug127*ug128*ug129*ug130*ug131*ug132*ug133*ug134*ug135*ug136*ug137*ug138*ug139
            bd108*bd108_n/missing list;

    proc freq data=analysis;
        table (RG112
            RG113
            RG114E)*cholecystectomy/missing norow nocum nopercent;
        table (RG121
            RG115
            BMI_CAT
            DIAB_TYPE
            BH210
            SYMPTOM_TYPE
            BH214
            BH215
            BH216
            BH244
            BH246
            BH247
            /*BH249 BH250*/
            BH253
            BH254
```

```
            BH261
            /*BH256 BH263*/
            ge213f_n)*Cholecystectomy/missprint norow nocum nopercent;
      title4 'Table 1';

proc means data=analysis maxdec=1;
      var age BMI Age_g_start SYMPT_DUR ge213d_n ge213f_n n_gd110-n_gd118 cardinal_sx n_gd119-n_gd122 gi_reflux n_gd130 n_gd131;
      class Cholecystectomy;

proc freq data=analysis;
      table (any_comorbid
            BH232X
            BH232Z
            cvd
            BH232AL
            BH232AP
            BH232AQ
            BH232AU
            BH235
            bd109)*cholecystectomy/missing norow nocum nopercent;
      format ge213f_n ge213f.;
      title4 'Table 2';

proc means data=analysis maxdec=1;
      var num_comorbid num_hosp se108a_n se108b_n bd108_n qol_activity qol_clothing qol_diet qol_relationship qol_psych qol_total;
      class cholecystectomy;
```