



TECHNICAL REPORT

No. CCLS-13-01

Title: Data pre-processing for the preterm prediction study
MFMU dataset

Authors: Clinical Informatics Group (CING)
Ilia Vovsha, Ansaf Salleb-Aouissi, Axinia Radeva,
Anita Raja, Hatim Diab, Ashish Tomar, Ashwath Rajan

Copyright Center for Computational Learning Systems - CCLS
Columbia University
<http://www.ccls.columbia.edu>

Contents

Summary	2
Bibliography	3
Dictionary and Notation	4
Timeline	5
Group Template.....	6
Group Synopsis	7
Group Tables	8-20
CLASS	8
DATES, GA.....	9
DMG, JOB, OBST.....	10
PPH, PPHD	11
SAD, SAD3.....	12
INFEC, INFEC3, MEDT, MEDT3.....	13
CPH, CPH3	14
CPM, SPEC, CRVM, SYMP	15
CPM3, SPEC3, CRVM3, SYMP3.....	16
FFN, FFN2, FFN3, FFN4, VISIT2, VISIT4	17
BUA, BUA3, PSYCH.....	18
OUTM, OUTN, OUTS, IPRE.....	19
CHORS, PDELM, PDELC.....	20
Fill-In Steps	21
Processing Steps	22

1. Summary

The United States spends over 26 billion dollars per annum on the delivery and care of the 12-13% of infants who are born preterm (Behrman and Butler, 2007). As preterm birth is a major public health problem with profound implications on society, there would be extreme value in being able to identify women at risk of preterm birth during the course of their pregnancy. Previous research has largely focused on individual risk factors correlated with preterm birth (e.g. prior preterm birth, race, and infection) and less on combining these factors in a way to understand the complex etiologies of preterm birth. Today, there is no widely tested prediction system that combines well-known factors (Davey et al., 2011) with a good prediction technique to provide actionable decisions within a clinical environment (Mercer et al., 1996).

We attempt to address this gap by conducting a deeper analysis of the preterm prediction study data collected by the NICHD Maternal Fetal Medicine Units (MFMU) Network, a high-quality data for over 3,000 singleton pregnancies having detailed study visits and biospecimen collection at 24, 26, 28 and 30 weeks gestation. Reports from this dataset used relatively straightforward biostatistical methodologies such as relative risk assessments to measure associations between risk factors and PTB (Maternal Fetal Medicine Units Network. Biostatistical Coordinating Center NICHD Networks, 1995). These methods include descriptive statistics, Pearson correlation, Fisher’s exact tests and linear/logistic regression where risk factors are studied independent of each other. In order to perform detailed experiments on this data using non-linear Support Vector Machines and other machine learning (ML) methodologies, it is necessary to complete several pre-processing steps.

1.1 Data Pre-Processing

The MFMU dataset was compiled from a set of tests and a detailed questionnaire administered to mothers (patients) over the entire duration of the pregnancy. As a result, multiple processing steps are required to harness this very rich and highly structured data. We face three specific challenges,

1. *Missing Data:* A substantial number of features is missing, both randomly (a particular question was not completed) and structurally (dependency between questions).
2. *Varying Sample Size:* The number of examples changes over time as patients give birth, withdraw, or miss a visit.
3. *Skewed Class Distributions:* By the nature of the preterm birth problem, class sizes are significantly uneven. This fact is relevant to subsets of the data as well (e.g. when we consider nulliparous mothers only).

We handle the complexity of the data by organizing features into groups (according to the original questionnaire and the definition). In addition, we keep track of the number and type of missing values, and the processing steps taken to complete or delete features. Our main objective is to retain as many features and examples as possible, while converting the data into a standard format suitable for off-the-shelf ML algorithms. We believe that the MFMU data (if properly processed) is an exciting application from a machine learning perspective.

1.2 Notes

- The blood and urine analysis tables are incomplete.

References

R.E. Behrman and A.S. Butler, editors. *Preterm Birth: Causes, Consequences, and Prevention*. The National Academies Press, 2007. Committee on Understanding Premature Birth and Assuring Healthy Outcomes.

M. Davey, L. Watson, J.A. Rayner, and S. Rowlands. Risk scoring systems for predicting preterm birth with the aim of reducing associated adverse outcomes. *Cochrane Database Syst Rev*, 11, 2011.

Maternal Fetal Medicine Units Network. Biostatistical Coordinating Center NICHD Networks. Final Report Screening for Risk Factors for Spontaneous Preterm Delivery. Singleton Study. Technical report, 09 1995.

B.M. Mercer, R.L. Goldenberg, A. Das, A.H. Moawad, J.D. Iams, P.J. Meis, R.L. Copper, F. Johnson, E. Thom, D. McNellis, M. Miodovnik, M.K. Menard, S.N. Caritis, G.R. Thurnau, S.F. Bottoms, and J. Roberts. The preterm prediction study: A clinical risk assessment system. *American Journal of Obstetrics and Gynecology*, 174(6):1885 – 1895, 1996.

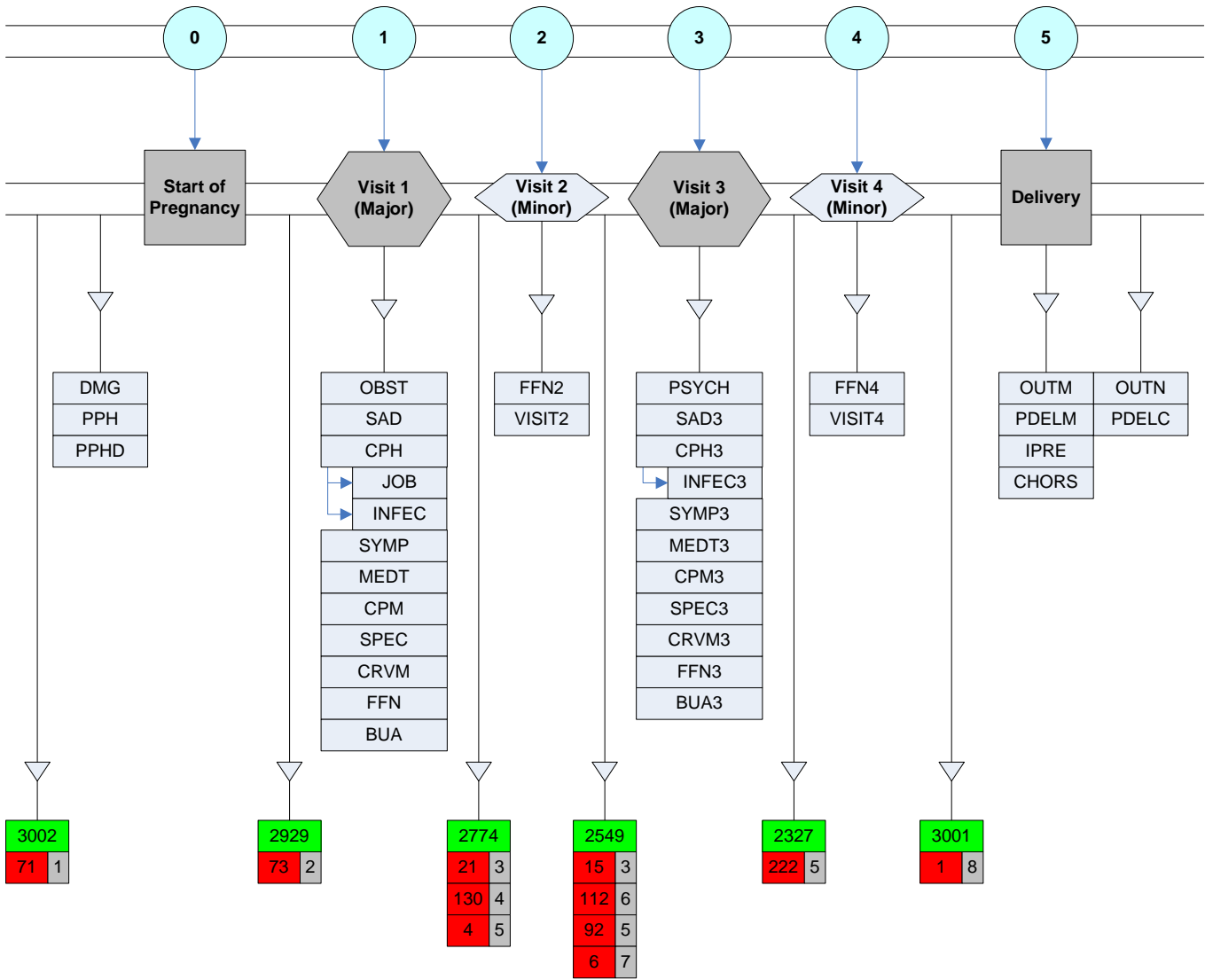
Acronym Dictionary

GA:	Gestational Age
DMG:	Demographics and Home Life
PPH:	Previous Pregnancy History
PPHD:	Previous Pregnancy History Detail
OBST:	Obstetrical & Medical Complications
SAD:	Substance Use V1
SAD3:	Substance Use V3
CPH:	Current Pregnancy History V1
CPH3:	Current Pregnancy History V3
JOB:	Current or Last Job
INFEC:	Infections During This Pregnancy V1
INFEC3:	Infections During This Pregnancy V3
MEDT:	Medications and Treatments V1
MEDT3:	Medications and Treatments V3
SYMP:	Symptoms During Previous Week V1
SYMP3:	Symptoms During Previous Week V3
CPM:	Current Pregnancy Measurements V1
CPM3:	Current Pregnancy Measurements V3
SPEC:	Specimen Collection V1
SPEC3:	Specimen Collection V3
CRVM:	Cervical Measurements V1
CRVM3:	Cervical Measurements V3
FFN:	Fetal Fibronectin Analysis V1
FFN2:	Fetal Fibronectin Analysis V2
FFN3:	Fetal Fibronectin Analysis V3
FFN4:	Fetal Fibronectin Analysis V4
BUA:	Blood & Urine Analysis V1
BUA3:	Blood & Urine Analysis V3
PSYCH:	Psychological Questionnaire
VISIT2:	Yeast & Intercourse Variables V2
VISIT4:	Yeast & Intercourse Variables V4
OUTM:	Pregnancy Outcome, Maternal Data
OUTN:	Pregnancy Outcome, Neonatal Data
OUTS:	Pregnancy Outcome Status
IPRE:	Indicated Preterm Birth Reasons
CHORS:	Chorioamnionitis Suspected
PDELM:	Preterm Delivery, Maternal Data
PDELC:	Preterm Delivery, Clinical Data

Notation

INT	Integer variable
BIN	Binary variable
CAT	Categorical variable
GRP	Group
d.	During
Prev.	Previous
hr / wk / yr	hour / week / year
Spont.	Spontaneous
Pre.	Preterm
Preg.	Pregnancy

Timeline



Reasons for removing examples:

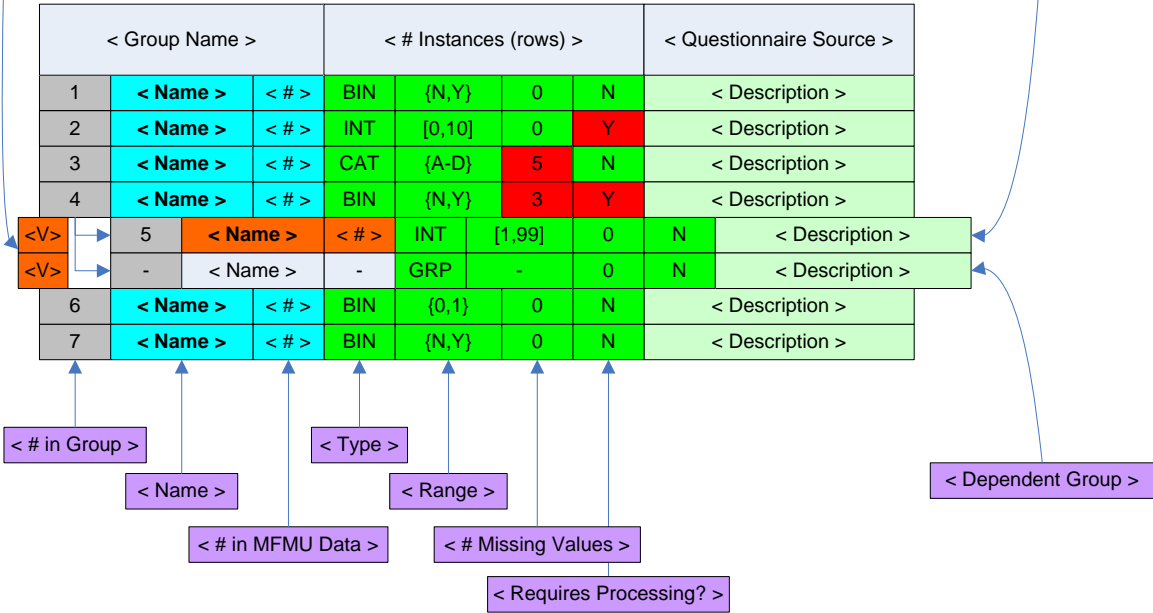
- 1) Lost to follow-up (no outcome data).
- 2) Visit 1 outside of GA window [154,174] allowed by protocol.
- 3) Patient delivered.
- 4) Patient withdrew.
- 5) No more visit data (skipped / withdrew / delivered).
- 6) Visit 3 outside of GA window [182, 209] allowed by protocol.
- 7) Skipped major visit (visit 3).
- 8) No preterm delivery information

Group Template

Colors:
 Red: variable requires fill-in or processing steps.
 Orange: dependency.
 Light Gray: group.

< Value Dependency >

< Dependent Variable >



Name	T	#	Miss	Dep	MFMU Features	Source
CLASS	-	3002	-	-	1,6, 255, 265, 284, 291-293, 295, 306-314, 359	-
DATES	-	3002	-	-	2-5, 40-41, 108-111, 174-176, 220-221, 248-253, 256, 266, 282, 304, 349, 352	-
GA	-	3002	-	-	7-8, 354-368	-
DMG	0	3002	-	-	23-24,42-45,48-51,416-418	(MP02A.1) Q.2-8,13-17
PPH	0	3002	-	-	6, 52-59, 71-76, 434	(MP02A.2) Q.35-43
PPHD	0	551	-	-	77-89, 415	(MP02B)
OBST	1	2929	-	-	29-39, 432-433	(MP02A.2) Q.22-34
SAD	1	2929	-	-	25-28, 60-70, 419-422, 424-428	(MP02A.2) Q.18-21
CPH	1	2929	-	-	9,10, 90-106, 120-129, 143-144, 154-155, 440	(MP02C.1) Q.8-15
JOB	1	1346	-	10	11-22, 46-47, 361	(MP02A.1) Q.10-12
INFEC	1	1068	-	95	145-153, 335	(MP02C.1) Q.12
MEDT	1	2929	-	-	96-102, 154	(MP02C.1) Q.13
SYMP	1	2929	-	-	123-128	(MP02C.1) Q.10
CPM	1	2929	-	-	112-119, 156, 437, 439	(MP02C.1) Q.1-7, 16-17
SPEC	1	2929	2	-	130-132, 279, 319	(MP02C.2) Q.21-24
CRVM	1	2929	2	-	107, 133-142	(MP02C.2) Q.25-34
FFN	1	2929	3	-	323-324, 331, 336-337, 344	(MP02C.2) Q.20

VISIT2	2	2774	247	-	348, 350	(MP03) Q.3,5
FFN2	2	2774	259	-	325-326, 332, 338-339, 345	(MP03) Q.4
PSYCH	3	2549	33	-	267-278	(MP08)
SAD3	3	2549	-	-	157-159, 206-208, 423, 429-430	(MP04.1) Q.3-5
CPH3	3	2549	-	-	160-165, 177-180, 187, 441	(MP04.1) Q.2-11
INFEC3	3	321	-	164	209-216, 431	(MP04.1) Q.10
MEDT3	3	2549	-	-	166-172, 217	(MP04.1) Q.12
SYMP3	3	2549	-	-	181-186	(MP04.1) Q.8
CPM3	3	2549	-	-	188-192, 438	(MP04) Q.13-16
SPEC3	3	2549	5	-	193-195, 280, 320	(MP04.2) Q.20-23
CRVM3	3	2549	7	-	173, 196-205	(MP04.2) Q.24-33
FFN3	3	2549	5	-	327-328, 333, 340-341, 346	(MP04.2) Q.19

VISIT4	4	2327	3	-	351, 353	(MP03) Q.3,5
FFN4	4	2327	10	-	329-330, 334, 342-343, 347	(MP03) Q.4

OUTM	5	3001	3	-	218-219, 222-228, 436	(MP05) Q.A
PDELM	5	433	-	307	229-234, 241-247, 254, 257, 315	(MP06) Q.A
IPRE	5	124	-	307, 314	258-264	(MP06) Q.A7
CHORS	5	33	-	234	235-240	(MP06) Q.A8
OUTN	5	3001	1	-	281, 283-290, 321-322	(MP05) Q.B
PDELC	5	420	?	-	296-303, 305, 316-318	(MP06) Q.B

BUA	1	?	?	-	369, 371, 373-379, 387-388, 391, 393-394, 397, 399, 401-406, 413	(MP02C.2) Q.18-19
BUA3	3	?	?	-	370, 372, 380-386, 389-390, 392, 395-396, 398, 400, 407-412, 414	(MP04.2) Q.17-18

CLASS			3002				-
1	PRE34	306	BIN	{0,1}	0	N	Spont. Pre. birth d. wks 23-34
2	PRE36	307	BIN	{0,1}	1	N	Spont. Pre. birth d. wks 23-36
3	PRE27	308	BIN	{0,1}	0	N	Spont. Pre. birth d. wks 23-27
4	PRE30	309	BIN	{0,1}	0	N	Spont. Pre. birth d. wks 23-29
5	PRE32	310	BIN	{0,1}	0	N	Spont. Pre. birth d. wks 23-31
6	PRE28_34	311	BIN	{0,1}	0	N	Spont. Pre. birth d. wks 28-34
7	PRE35_36	312	BIN	{0,1}	1	N	Spont. Pre. birth d. wks 35-36
8	SPRE34	313	BIN	{0,1}	0	N	Spont. & Ind. Pre. birth d. wks 23-34
9	SPRE36	314	BIN	{0,1}	0	N	Spont. & Ind. Pre. birth d. wks 23-36
10	GADEL_D	359	INT	[163,314]	0	N	GA at delivery in days
11	EPSTATUS	291	CAT	{1-5}	7	N	Infant status at <= 30 days
12	STLBRN	292	BIN	{0,1}	0	N	Stillbirth
13	INFDTH	293	BIN	{0,1}	0	N	Infant death
14	FPCLASS	255	CAT	{1-5}	1	N	Preterm delivery classification
15	FPIUFD	295	BIN	{N,Y}	-	N	Stillbirth among pre. deaths
16	APPARITY	6	BIN	{1,2}	0	N	Parity
17	EPSEX	284	BIN	{1,2}	2	N	Gender of infant
18	APLMPSUR	1	BIN	{N,Y}	-	N	LMP date sure
19	GPLOST	265	BIN	{N,Y}	-	N	Lost to follow-up

DATES			-				-
1	BRVISDT	40	DATE	-	-	N	Visit 1 date
2	VIS2DT	349	DATE	-	-	N	Visit 2 date
3	DRVISDT	174	DATE	-	-	N	Visit 3 date
4	VIS4DT	352	DATE	-	-	N	Visit 4 date
5	ARUSDATE	2	DATE	-	-	N	Date of 1 st dating U/S
6	ARLMPDT	3	DATE	-	-	N	LMP date
7	ARPROEDC	4	DATE	-	-	N	Project EDC
8	ARENUSDT	5	DATE	-	-	N	Date of enrollment U/S
9	BRSTOPDT	41	DATE	-	-	N	Date patient stopped working
10	BRPRENAT	108	DATE	-	-	N	Date of 1 st prenatal visit
11	BRBLPRDT	109	DATE	-	-	N	Most recent BP reading V1
12	BRHEMAT	110	DATE	-	-	N	1 st Hematocrit / Hemoglobin
13	BRALPHDT	111	DATE	-	-	N	Date of serum alpha-fetoprotein
14	DRBLPRDT	175	DATE	-	-	N	Most recent BP reading V3
15	DRHEMAT	176	DATE	-	-	N	Last Hematocrit / Hemoglobin
16	HRDATE	266	DATE	-	-	N	Psychological survey date
17	ER1STVIS	220	DATE	-	-	N	Date of 1 st L&D / hospital visit
18	ERTOCODT	221	DATE	-	-	N	Date of 1 st IV tocolysis
19	FRADMDT	248	DATE	-	-	N	PTD admission date
20	FPRUPTDT	249	DATE	-	-	N	Date membranes ruptured
21	FPRUPTTM	252	TIME	-	-	N	Time membranes ruptured
22	FPCHORDT	250	DATE	-	-	N	Date chorio was confirmed
23	FPCHORTM	253	TIME	-	-	N	Time chorio was confirmed
24	FRDISCDT	251	DATE	-	-	N	Maternal discharge date
25	ERBRDT	282	DATE	-	-	N	Date of birth
26	ERBRDTM	256	TIME	-	-	N	Time of birth
27	FRDTHDT	304	DATE	-	-	N	Date of death

GA			-				-
1	VISGA1_D	354	INT	[146,196]	-	N	GA in days at visit 1
2	VISGA2_D	355	INT	[161,208]	-	N	GA in days at visit 2
3	VISGA3_D	356	INT	[175,231]	-	N	GA in days at visit 3
4	VISGA4_D	357	INT	[189,237]	-	N	GA in days at visit 4
5	VISGAD24	365	INT	[146,174]	-	N	GA in days at fFN (wk 23-24)
6	VISGAD26	366	INT	[175,188]	-	N	GA in days at fFN (wk 25-26)
7	VISGAD28	367	INT	[189,202]	-	N	GA in days at fFN (wk 27-28)
8	VISGAD30	368	INT	[203,223]	-	N	GA in days at fFN (wk 29-30)
9	APUSWKS	7	INT	[4,32]	-	N	GA in weeks by first U/S
10	APUSDAYS	8	INT	[0,6]	-	N	GA in days by first U/S
11	GAPREN_D	358	INT	[0,192]	-	N	GA in days at 1 st prenatal visit
12	GASTOP_D	361	INT	[4,171]	-	N	GA in days at stop-work date
13	GATOCO_D	362	INT	[141,257]	-	N	GA in days at 1 st tocolysis
14	GAALPH_D	363	INT	[50,172]	-	N	GA in days at 1 st serum AFP
15	GARUP_D	360	INT	[163,258]	-	N	GA in days at membrane rupture
16	GAADM_D	364	INT	[162,258]	-	N	GA in days at admission for PTD
17	GADEL_D	359	INT	[163,314]	-	N	GA at delivery in days

DMG			3002				(MP02A.1) Q.2-8,13-17
1	BPPHONE	23	BIN	{N,Y}	0	N	Has home phone
2	BPCAR	24	BIN	{N,Y}	0	N	Use of car
3	BPMARITL	42	CAT	{1-4}	0	N	Marital status
4	BPINSUR	43	CAT	{1-3}	0	N	Source of medical care payment
5	BPINCOME	44	INT	[1-4]	1	Y	Family unit total income
6	BPDEPEND	45	INT	[1,14]	0	N	# People supported by income
7	BPKIDS	48	INT	[0,6]	0	N	# Preschool children at home
8	BPLASTYR	49	INT	[0,20]	1	N	# Times changed address last yr
9	BPLST5YR	50	INT	[0,40]	2	N	# Times changed addr. last 5 yrs
10	BPCOND	51	INT	[1,4]	0	N	Patient description of residence
11	AGEMOM	416	INT	[17,40]	0	Y	Age of mother in yrs
12	RACE	417	CAT	{1-3}	0	N	Predominant race
13	SCHOOLYR	418	INT	[8,17]	0	Y	Total yrs of schooling

JOB			1346				(MP02A.1) Q.10-12	
1	BPWORK	11	BIN	{N,Y}	0	N	Currently working at paying job	
N	2	GASTOP_D	361	INT	[4,171]	0	N	GA at stop date (days)
3	BPLOAD	12	BIN	{N,Y}	0	N	Carry more than 20lb	
4	BPSTREN	13	BIN	{N,Y}	0	N	Do strenuous activity	
5	BPCLIMB	14	BIN	{N,Y}	0	N	Climb stairs	
Y	6	BPSTRNUM	47	INT	[1,99]	0	N	# Stairs climbed at work
7	BPSTAND	15	BIN	{N,Y}	0	N	Stand more than 3 hours	
8	BPBREAK	16	BIN	{N,Y}	0	N	Breaks when tired	
9	BPASSEMB	17	BIN	{N,Y}	0	N	Works on assembly line	
10	BPVIBES	18	BIN	{N,Y}	0	N	Use vibrating machine	
11	BPTASKS	19	BIN	{N,Y}	0	N	Do repetitive tasks	
12	BPCOLD	20	BIN	{N,Y}	0	N	Works in cold/wet area	
13	BPNOISE	21	BIN	{N,Y}	0	N	Encounter much noise	
14	BPBORING	22	BIN	{N,Y}	0	N	Finds work boring	
15	BPHRS	46	INT	[1,105]	0	N	Ave. # hrs worked per week	

OBST			2929				(MP02A.2) Q.22-34	
1	BPFIBR	29	BIN	{N,Y}	0	N	Uterine Fibroids	
2	BPDIAB	30	BIN	{N,Y}	0	N	Diabetes	
Y	3	DIABCLS	432	CAT	{1-4}	0	N	Class of diabetes
4	BPHYPER	31	BIN	{N,Y}	0	N	Chronic hypertension	
Y	5	BPMEDIC	32	BIN	{N,Y}	0	N	Medication required for hyp.
6	BPCARD	33	BIN	{N,Y}	0	N	Cardiac disease w/limited activity	
7	BPENDO	34	BIN	{N,Y}	0	N	Endocrine disorder req. meds	
8	BPHEMO	35	BIN	{N,Y}	0	N	Hemoglobinopathy	
9	BPPULMO	36	BIN	{N,Y}	0	N	Acute pulmonary disorder req. m.	
10	BPLUNG	37	BIN	{N,Y}	0	N	Chronic lung disease	
11	BPRENAL	38	BIN	{N,Y}	0	N	Acute or chronic renal disease	
12	BPELVIC	39	BIN	{N,Y}	0	N	Pelvic inflammatory disease	
13	OBRISK	433	BIN	{0,1}	0	N	Obstetrical complications	

PPH			3002				(MP02A.2) Q.35-43
1	APPARITY	6	BIN	{1,2}	0	N	Parity
2	BPTOTAL	52	INT	[1,12]	0	N	Total Pregnancies
3	BPTERM	53	INT	[0,7]	0	N	# Prev. term deliveries
4	BPPRETERM	54	INT	[0,5]	0	N	# Prev. preterm deliveries
5	BP13TO19	55	INT	[0,2]	0	N	# Prev. spon. deliveries 13-19 wks
6	BPINDUCE	56	INT	[0,6]	0	N	# Prev. induced abortions
7	BPSTILL	57	INT	[0,3]	0	N	# Prev. stillbirths over 20 wks
8	BPDEATHS	58	INT	[0,2]	0	N	# Prev. neonatal deaths
9	BPCHILDREN	59	INT	[0,8]	0	N	# Prev. children still living
10	TERM	71	BIN	{0,1}	0	N	Any prev. term delivery
11	PRETERM	72	BIN	{0,1}	0	N	Any prev. preterm delivery
12	SPON	73	BIN	{0,1}	0	N	Any prev. spon. deliv. 13-19 wks
13	FRST	74	BIN	{0,1}	0	N	Any prev. spon. miscarriage
14	INDUCE	75	BIN	{0,1}	0	N	Any prev. induced abortions
15	STILL	76	BIN	{0,1}	0	N	Any prev. stillbirths over 20 wks.
16	MISC1ST	434	INT	[0,4]	0	Y	# Prev. 1 st trimester miscarriages

PPHD			551				(MP02B)
1	PCLASS1	77	INT	[0,2]	0	N	# Prev. spon. PTL W/deliv. 13-19 wks
2	PCLASS2	78	INT	[0,2]	0	N	# Prev. PROM W/spon. PTD 13-19 wks
3	PCLASS3	79	INT	[0,1]	0	N	# Prev. PROM W/induced. PTD 13-19 wks
4	PCLASS4	80	INT	[0]	0	N	# Prev. Indicated PTD-Fetal 13-19 wks
5	PCLASS5	81	INT	[0]	0	N	# Prev. Indicated PTD-Mat. 13-19 wks
6	PCLASS6	82	INT	[0,2]	0	N	# Prev. elective abortions 13-19 wks
7	PCLASSA1	83	INT	[0,4]	0	N	# Prev. spon. PTL W/deliv. 20-36 wks
8	PCLASSA2	84	INT	[0,3]	0	N	# Prev. PROM W/spon. PTD 20-36 wks
9	PCLASSA3	85	INT	[0,1]	0	N	# Prev. PROM W/induced. PTD 20-36 wks
10	PCLASSA4	86	INT	[0,2]	0	N	# Prev. Indicated PTD-Fetal 20-36 wks
11	PCLASSA5	87	INT	[0,3]	0	N	# Prev. Indicated PTD-Mat. 20-36 wks
12	PCLASSA6	88	INT	[0,1]	0	N	# Prev. elective abortions 20-36 wks
13	SECAB	89	INT	[0,2]	0	N	# Prev. 2 nd trimester abortions
14	LASTPREG	415	INT	[0,23]	0	N	Calendar yrs since last (13-36 wks) preg.

SAD			2929				(MP02A.2) Q.18-21	
1	BPSMOKE	25	BIN	{N,Y}	0	N	Ever smoked cigarettes	
Y	2	CIGSPRE	419	INT	[0,40]	1	Y	# CIG / day pre-pregnancy
Y	3	CIG12	420	INT	[0,30]	1	Y	# CIG / day <= 12 wks preg.
Y	4	CIG24	421	INT	[0,25]	0	Y	# CIG / day 13-24 wks preg.
Y	5	CIGSNOW	422	INT	[0,25]	0	Y	# CIG / day last week
6	SMOKPREG	60	BIN	{0,1}	0	N	Smoked during pregnancy	
7	BPALCHOL	26	BIN	{N,Y}	0	N	Ever drunk alcoholic beverages	
Y	8	DRINKPRE	424	INT	[0,6]	6	Y	# Drinks / wk pre-pregnancy
Y	9	DRK12	425	INT	[0,4]	4	Y	# Drinks / wk <= 12 wks preg.
Y	10	DRK24	426	INT	[0,2]	4	Y	# Drinks / wk 13-24 wks preg.
Y	11	BPALCNOW	27	BIN	{N,Y}	2	N	Drink alcoholic beverages now
Y	12	TOTALC	61	INT	[0,29]	0	N	Total drinks during pregnancy
Y	13	DRNKMOST	427	INT	[0,4]	0	Y	Most drinks in 1 day last month
Y	14	DRKLSTDY	428	INT	[0,2]	0	Y	# Drinks in last 24 hrs
15	BPDRUGS	28	BIN	{N,Y}	0	N	Ever used street drugs	
Y	16	BHEROIN	62	BIN	{N,Y}	3	N	Heroin pre-preg.
Y	17	BCOCAINE	63	BIN	{N,Y}	3	N	Cocaine pre-preg.
Y	18	BMARI	64	BIN	{N,Y}	3	N	Marijuana pre-preg.
Y	19	BLSD	65	BIN	{N,Y}	3	N	LSD pre-preg.
Y	20	BSPEED	66	BIN	{N,Y}	3	N	Speed pre-preg.
Y	21	BOTHER	67	BIN	{N,Y}	3	N	Other drugs pre-preg.
Y	22	DCOCAINE	68	BIN	{N,Y}	510	N	Cocaine during preg.
Y	23	DMARI	69	BIN	{N,Y}	510	N	Marijuana during preg.
Y	24	DOTHER	70	BIN	{N,Y}	510	N	Other drugs during preg.

SAD3			2549				(MP04.1) Q.3-5	
1	DPSMOKED	157	BIN	{N,Y}	0	N	Cigarettes smoked since visit 1	
Y	2	VCIGDAY	423	INT	[0,25]	1	Y	# CIG / day since visit 1
3	DPALCHOL	158	BIN	{N,Y}	0	N	Alcohol since visit 1	
Y	4	VALCDY	429	INT	[1,4]	0	Y	Max # Drinks / day since visit 1
Y	5	VALCWK	430	INT	[0,3]	3	Y	# Drinks / wk since visit 1
6	DPDRUGS	159	BIN	{N,Y}	0	N	Street drugs since visit 1	
Y	7	VCOCAINE	206	BIN	{N,Y}	0	N	Cocaine since visit 1
Y	8	VMARI	207	BIN	{N,Y}	0	N	Marijuana since visit 1
Y	9	VOTHER	208	BIN	{N,Y}	0	N	Other drugs since visit 1

INFECC			1068				(MP02C.1) Q.12
1	CHLAM	145	BIN	{N,Y}	0	N	Chlamydia
2	GONORR	146	BIN	{N,Y}	0	N	Gonorrhoea
3	SYPH	147	BIN	{N,Y}	0	N	Syphilis
4	TRICH	148	BIN	{N,Y}	0	N	Trichomonas
5	BV	149	BIN	{N,Y}	0	N	Bacterial vaginosis
6	HERPES	150	BIN	{N,Y}	0	N	Herpes
7	BACTER	151	BIN	{N,Y}	0	N	Asymptomatic bacteriuria
8	CYS	152	BIN	{N,Y}	0	N	Cystitis
9	GBS	153	BIN	{N,Y}	0	N	Group B strep
10	PYELO	335	BIN	{N,Y}	0	N	Pyelonephritis

INFECC3			321				(MP04.1) Q.10
1	VCHLAM	209	BIN	{N,Y}	0	N	Chlamydia
2	VTRICH	210	BIN	{N,Y}	0	N	Trichomonas
3	VBV	211	BIN	{N,Y}	0	N	Bacterial vaginosis
4	VHERPES	212	BIN	{N,Y}	0	N	Herpes
5	VBACTER	213	BIN	{N,Y}	0	N	Asymptomatic bacteriuria
6	VCYS	214	BIN	{N,Y}	0	N	Cystitis
7	VPYELO	215	BIN	{N,Y}	0	N	Pyelonephritis
8	VGBS	216	BIN	{N,Y}	0	N	Group B strep
9	VSYPHON	431	BIN	{N,Y}	0	N	Gonorrhoea or Syphilis

MEDT			2929				(MP02C.1) Q.13
1	BPIRON	96	BIN	{N,Y}	0	N	Iron taken
2	BPVITA	97	BIN	{N,Y}	0	N	Vitamins taken
3	BPASPIRIN	98	BIN	{N,Y}	0	N	Aspirin taken
4	BPANTIB	99	BIN	{N,Y}	0	N	Antibiotics
5	BPTOCOL	100	BIN	{N,Y}	1	N	Tocolytics
6	BPBRONCH	101	BIN	{N,Y}	0	N	Bronchodilators
7	BPACTIV	102	BIN	{N,Y}	0	N	Activity restriction
8	MEDTRT	154	BIN	{0,1}	0	N	Any meds/treatments

MEDT3			2549				(MP04.1) Q.12
1	DPIRON	166	BIN	{N,Y}	0	N	Iron taken
2	DPVITA	167	BIN	{N,Y}	0	N	Vitamins taken
3	DPASPIRIN	168	BIN	{N,Y}	0	N	Aspirin taken
4	DPANTIB	169	BIN	{N,Y}	0	N	Antibiotics
5	DPTOCOL	170	BIN	{N,Y}	0	N	Tocolytics
6	DPBRONCH	171	BIN	{N,Y}	0	N	Bronchodilators
7	DPACTIV	172	BIN	{N,Y}	0	N	Activity restriction
8	VMEDTRT	217	BIN	{0,1}	0	N	Any meds/treatments

CPH			2929				(MP02C.1) Q.?	
1	APPLACEN	9	CAT	{1-7}	12	N	Placental location at enroll. U/S	
2	BPJOB	10	BIN	{N,Y}	0	N	Worked at paying job d. preg.	
Y	-	JOB	-	GRP	-	0	N	Job details
3	BPVAG1ST	90	BIN	{N,Y}	0	N	Vaginal bleeding/spot 1 st trimester	
4	BPVAG2ND	91	BIN	{N,Y}	0	N	Vaginal bleeding/spot 2 nd trimester	
5	VAGBLD	143	BIN	{0,1}	0	N	Vaginal bleeding, 1 st or 2 nd trim.	
6	PERBLD	144	BIN	{0,1}	0	N	Persistent vaginal bleeding	
7	BPCONTRA	92	BIN	{N,Y}	0	N	Uterine contractions in last 2 wks	
Y	8	BPCONHR	120	INT	[0,30]	1	N	Max. uterine contractions / hr
Y	9	BPCONDAY	121	INT	[0,168]	2	N	Max. uterine contractions / day
Y	10	BPCONWK	122	INT	[0,504]	2	N	Max. uterine contractions / wk
11	BPINTER	93	BIN	{N,Y}	0	N	Intercourse w/in last 24 hrs	
12	BPLIMIT	94	BIN	{N,Y}	5	N	Advised to limit intercourse	
13	BPINTNUM	129	INT	[0,30]	29	N	# Times had Intercourse per wk	
14	BPINFEC	95	BIN	{N,Y}	0	N	Infections occurred d. preg.	
Y	-	INFEC	-	GRP	-	0	N	Infections d. preg.
15	BPCVS	103	BIN	{N,Y}	0	N	Performed CVS	
16	BPAMNIO	104	BIN	{N,Y}	0	N	Performed Amniocentesis	
17	BPABD	105	BIN	{N,Y}	0	N	Performed ABD surgery	
18	PROCED	155	BIN	{0,1}	0	N	Performed any procedures	
19	BPHOSPIT	106	BIN	{N,Y}	0	N	Hospitalized for any reason	
Y	20	HOSPREAS	440	CAT	{A-J}	0	N	Primary reason for hospitalization

CPH3			2549				(MP04.1) Q.?	
1	DPWORK	177	INT	[1,4]	0	N	Work status	
2	DPBLEED	160	BIN	{N,Y}	0	N	Vaginal bleeding/spot since V1	
3	DPCONTRA	161	BIN	{N,Y}	0	N	Uterine contractions since V1	
Y	4	DPCONHR	178	INT	[0,30]	8	N	Max. uterine contractions / hr
Y	5	DPCONDAY	179	INT	[0,144]	10	N	Max. uterine contractions / day
Y	6	DPCONWK	180	INT	[0,672]	9	N	Max. uterine contractions / wk
7	DPINTER	162	BIN	{N,Y}	0	N	Intercourse w/in last 24 hrs	
8	DPLIMIT	163	BIN	{N,Y}	13	N	Advised to limit intercourse	
9	DPINTNUM	187	INT	[0,30]	74	N	# Times had Intercourse per wk	
10	DPINFEC	164	BIN	{N,Y}	0	N	Infections occurred since visit 1	
Y	-	INFEC3	-	GRP	-	0	N	Infections since visit 1
11	DPCOMPL	165	BIN	{N,Y}	0	N	Obst. / surg. / med. complications	
Y	12	COMPLIC3	441	CAT	{A-J}	0	N	Major complications since V1

CPM			2929				(MP02C.1) Q.1-7, 16-17
1	BP1STSYS	112	INT	[70,180]	29	N	Systolic BP at 1 st prenatal
2	BP1STDIA	113	INT	[10,110]	29	N	Diastolic BP at 1 st prenatal
3	BPLSTSYS	114	INT	[72,168]	4	N	Systolic BP at last reading
4	BPLSTDIA	115	INT	[30,100]	5	N	Diastolic BP at last reading
5	BPHEMVAL	116	REAL	[18.5,47]	578	N	Most recent hematocrit value
6	BPGLOVAL	117	REAL	[5.8,17]	393	N	Most recent hemoglobin value
7	BPALPVAL	118	REAL	[0.1,7]	923	N	Alpha-fetoprotein value (mom)
8	BPURINE	119	INT	[0,4]	6	N	Highest urine protein by dip
9	HEIGHT	156	REAL	[48,78]	5	N	Maternal height in inches
10	WEIGHTV1	437	REAL	[40,191]	5	N	Weight in kgs, visit 1
11	WGTPRE	439	REAL	[36,170]	87	N	Weight in kgs, pre-preg.

SPEC			2929				(MP02C.2) Q.21-24
1	BPVAGPH	130	REAL	[2.4,8.4]	12	N	Vaginal pH
2	BPTRICHO	131	BIN	{1,2}	2	N	Wet prep trichomonas results
3	BPYEAST	132	BIN	{1,2}	2	N	KOH yeast results
4	GSTOT1	279	INT	[0,10]	19	N	Gram stain result
5	BV1	319	BIN	{0,1}	29	N	Bacterial vaginosis

CRVM			2929				(MP02C.2) Q.25-34		
1	BPCERLNG	133	REAL	[0.5,5.5]	13	N	Cervical length in cm (digital)		
2	BPEXTDIL	134	REAL	[0,3]	13	N	External OS dilation (cm)		
10	3	BPINTDIL	135	REAL	[0,3]	0	N	Internal OS dilation (cm)	
	10	4	BPMEMB	136	INT	[1,4]	0	N	Membranes palpated (digital)
5	BPSTAT	137	INT	[-3,3]	6	N	Fetal station		
6	BPPOS	138	INT	[1,3]	5	N	Cervical position		
7	BPCONS	139	INT	[1,3]	6	N	Cervical consistency		
8	BPLOWER	140	BIN	{1,2}	5	N	Lower uterine segment		
9	BPCRVLT	141	INT	[0,70]	14	N	Cervical length by U/S (mm)		
10	BPMEMPRO	107	BIN	{N,Y}	14	N	Membranes protruding into cervix		
Y	11	BPMAXPEN	142	INT	[3,52]	1	N	Max membrane penetration (mm)	

SYMP			2929				(MP02C.1) Q.10
1	BPCPELV	123	INT	[0,3]	0	N	Level of pelvic pressure last wk
2	BPCRAMP	124	INT	[0,3]	0	N	Level of abdominal cramp. last wk
3	BPDISCH	125	INT	[0,3]	0	N	Level of vaginal discharge last wk
4	BPDIAAR	126	INT	[0,3]	0	N	Level of diarrhea last wk
5	BPBKPAIN	127	INT	[0,3]	0	N	Level of low back pain last wk
6	BPSPOT	128	INT	[0,3]	0	N	Level of bleed/spot last wk

CPM3			2549				(MP04) Q.13-16
1	DPLSTSYS	188	INT	[60,160]	2	N	Systolic BP at last reading
2	DPLSTDIA	189	INT	[0,120]	2	N	Diastolic BP at last reading
3	DPHEMVAL	190	REAL	[21.3,46.9]	164	N	Most recent hematocrit value
4	DPGLOVAL	191	REAL	[7.1,17]	796	N	Most recent hemoglobin value
5	DPURINE	192	INT	[0,4]	4	N	Highest urine protein by dip
6	WEIGHTV3	438	REAL	[43,194.4]	4	N	Weight in kgs, visit 3

SPEC3			2549				(MP04.2) Q.20-23
1	DPVAGPH	193	REAL	[2.3,10.5]	20	N	Vaginal pH
2	DPTRICHO	194	BIN	{1,2}	8	N	Wet prep trichomonas results
3	DPYEAST	195	BIN	{1,2}	7	N	KOH yeast results
4	GSTOT3	280	INT	[0,10]	25	N	Gram stain result
5	BV3	320	BIN	{0,1}	38	N	Bacterial vaginosis

CRVM3			2549				(MP04.2) Q.24-33		
1	DPCERLNG	196	REAL	[0,4.5]	11	N	Cervical length in cm (digital)		
2	DPEXTDIL	197	REAL	[0,4]	11	N	External OS dilation (cm)		
10	3	DPINTDIL	198	REAL	[0,3]	2	N	Internal OS dilation (cm)	
	10	4	DPMEMB	199	INT	[1,4]	0	N	Membranes palpated (digital)
5	DPSTAT	200	INT	[-3,3]	9	N	Fetal station		
6	DPOPOS	201	INT	[1,3]	10	N	Cervical position		
7	DPCONS	202	INT	[1,3]	9	N	Cervical consistency		
8	DPLOWER	203	BIN	{1,2}	9	N	Lower uterine segment		
9	DPCVLTH	204	INT	[0,68]	18	N	Cervical length by U/S (mm)		
10	DPMEMPRO	173	BIN	{N,Y}	18	N	Membranes protruding into cervix		
Y	11	DPMAXPEN	205	INT	[2,48]	3	N	Max membrane penetration (mm)	

SYMP3			2549				(MP04.1) Q.8
1	DPCPELV	181	INT	[0,3]	0	N	Level of pelvic pressure last wk
2	DPCRAMP	182	INT	[0,3]	0	N	Level of abdominal cramp. last wk
3	DPDISCH	183	INT	[0,3]	0	N	Level of vaginal discharge last wk
4	DPDIAR	184	INT	[0,3]	0	N	Level of diarrhea last wk
5	DPBKPAIN	185	INT	[0,3]	0	N	Level of low back pain last wk
6	DPSPOT	186	INT	[0,3]	0	N	Level of bleed/spot last wk

FFN			2926				(MP02C.2) Q.20
1	FFNC1	323	REAL	[0,2418]	13	N	Cervical FFN value (wk 23-24)
2	FFNV1	324	REAL	[0,1835]	5	N	Vaginal FFN value (wk 23-24)
3	FFCEA24	336	BIN	{0,1}	11	N	Cervical FFN +/- (wk 23-24)
4	FFVAA24	337	BIN	{0,1}	4	N	Vaginal FFN +/- (wk 23-24)
5	FF1	331	BIN	{0,1}	14	N	Visit 1 FFN (wk 23-24)
6	FFA24	344	BIN	{0,1}	0	N	Summary FFN (wk 23-24)

FFN2			2515				(MP03) Q.4
1	FFNC2	325	REAL	[0,2079]	22	N	Cervical FFN value (wk 25-26)
2	FFNV2	326	REAL	[0,1937]	13	N	Vaginal FFN value (wk 25-26)
3	FFCEA26	338	BIN	{0,1}	85	N	Cervical FFN +/- (wk 25-26)
4	FFVAA26	339	BIN	{0,1}	79	N	Vaginal FFN +/- (wk 25-26)
5	FF2	332	BIN	{0,1}	21	N	Visit 2 FFN (wk 25-26)
6	FFA26	345	BIN	{0,1}	73	N	Summary FFN (wk 25-26)

FFN3			2544				(MP04.2) Q.19
1	FFNC3	327	REAL	[0,2363]	17	N	Cervical FFN value (wk 27-28)
2	FFNV3	328	REAL	[0,2593]	12	N	Vaginal FFN value (wk 27-28)
3	FFCEA28	340	BIN	{0,1}	247	N	Cervical FFN +/- (wk 27-28)
4	FFVAA28	341	BIN	{0,1}	243	N	Vaginal FFN +/- (wk 27-28)
5	FF3	333	BIN	{0,1}	17	N	Visit 3 FFN (wk 27-28)
6	FFA28	346	BIN	{0,1}	237	N	Summary FFN (wk 27-28)

FFN4			2317				(MP03) Q.4
1	FFNC4	329	REAL	[0,2758]	15	N	Cervical FFN value (wk 29-30)
2	FFNV4	330	REAL	[0,2226]	12	N	Vaginal FFN value (wk 29-30)
3	FFCEA30	342	BIN	{0,1}	63	N	Cervical FFN +/- (wk 29-30)
4	FFVAA30	343	BIN	{0,1}	62	N	Vaginal FFN +/- (wk 29-30)
5	FF4	334	BIN	{0,1}	14	N	Visit 4 FFN (wk 29-30)
6	FFA30	347	BIN	{0,1}	56	N	Summary FFN (wk 29-30)

VISIT2			2527				(MP03) Q.3,5
1	INTER2	348	BIN	{N,Y}	2	N	Intercourse w/in last 24 hours
2	YEAST2	350	BIN	{1,2}	12	N	Yeast results from KOH prep

VISIT4			2324				(MP03) Q.3,5
1	INTER4	351	BIN	{N,Y}	2	N	Intercourse w/in last 24 hours
2	YEAST4	353	BIN	{1,2}	6	N	Yeast results from KOH prep

Blood & Urine Analysis			V1				(MP02C.2) Q.18-19
1	LCR1	369	?	?	?	?	Chlamydia
2	CRF1	371	?	?	?	?	Cortico releasing factor
3	AFP1	373	?	?	?	?	Alpha feto protein
4	ALP1	374	?	?	?	?	Alkaline phosphatase
5	B2M1	375	?	?	?	?	Beta2 Microglobulin
6	FER1	376	?	?	?	?	Ferritin
7	ICAM1A	377	?	?	?	?	ICAM
8	IL6A	378	?	?	?	?	IL6
9	LCRP1	379	?	?	?	?	C reactive protein
10	LACTO1	387	?	?	?	?	Lactoferrin
11	DEFEN1	388	?	?	?	?	Defensin
12	REL1	391	?	?	?	?	Relaxin
1	IL101	393	?	?	?	?	IL10
2	GCSF1	394	?	?	?	?	GCSF
3	CORT1	397	?	?	?	?	Cortisol
4	ACT1	399	?	?	?	?	Activin
5	SELECT1	401	?	?	?	?	Beta2 Microglobulin
6	PLGF1	402	?	?	?	?	Placental growth factor
7	TNF_RA1	403	?	?	?	?	TNF receptor type 1
8	TNF_RB1	404	?	?	?	?	TNF receptor type 2
9	IL_2SRA1	405	?	?	?	?	IL2 receptor alpha
10	ANGIO1	406	?	?	?	?	Angiogenin
11	TAT1	413	?	?	?	?	Thrombin anti-thrombin

Blood & Urine Analysis			V3				(MP04.2) Q.17-18
1	LCR3	370	?	?	?	?	
2	CRF3	372	?	?	?	?	
3	AFP3	380	?	?	?	?	
4	ALP3	381	?	?	?	?	
5	B2M3	382	?	?	?	?	
6	FER3	383	?	?	?	?	
7	ICAM1B	384	?	?	?	?	
8	IL6B	385	?	?	?	?	
9	LCRP3	386	?	?	?	?	
10	LACTO3	389	?	?	?	?	
11	DEFEN3	390	?	?	?	?	
12	REL3	392	?	?	?	?	
1	IL103	395	?	?	?	?	
2	GCSF3	396	?	?	?	?	
3	CORT3	398	?	?	?	?	
4	ACT3	400	?	?	?	?	
5	SELECT3	407	?	?	?	?	
6	PLGF3	408	?	?	?	?	
7	TNF_RA3	409	?	?	?	?	
8	TNF_RB3	410	?	?	?	?	
9	IL_2SRA3	411	?	?	?	?	
10	ANGIO3	412	?	?	?	?	
11	TAT3	414	?	?	?	?	

PSYCH			2516				(MP08)
1	TOTLQUES	267	INT	[26,28]	0	N	Total questions answered
2	PSYSCORE	268	INT	[1,140]	0	N	Total psychological score
3	ANXIETY	269	INT	[1,40]	0	N	Anxiety subscale
4	TOTANX	270	INT	[7,8]	0	N	Total q. ans. anxiety
5	ESTEEM	271	INT	[1,30]	0	N	Self esteem subscale
6	TOTEST	272	INT	[5,6]	0	N	Total q. ans. self esteem
7	MASTERY	273	INT	[4,20]	0	N	Mastery subscale
8	TOTMAS	274	INT	[3,4]	0	N	Total q. ans. mastery
9	DEPRESS	275	INT	[1,40]	0	N	Depression subscale
10	TOTDEP	276	INT	[6,8]	0	N	Total q. ans. depression
11	STRESS	277	INT	[2,10]	0	N	Stress subscale
12	TOTSTR	278	INT	[2]	0	N	Total q. ans. stress

OUTM			3001				(MP05) Q.A	
1	EPLABOR	222	CAT	{1-4}	4	N	How labor started	
2	EPADMIT	218	BIN	{N,Y}	3	N	Admitted or L+D for PTL	
Y	3	EPINTVEN	219	BIN	{N,Y}	0	N	Other interventions
Y	4	ORALTOC	226	BIN	{N,Y}	0	N	Received oral tocolysis
Y	5	SUBQTOC	227	BIN	{N,Y}	0	N	Received Sub Q tocolysis
Y	6	HYDRA	228	BIN	{N,Y}	0	N	Hydration during admission
Y	7	EPDILVS	223	REAL	[0,4]	13	Y	Internal OS dilation at 1 st visit
Y	8	IVTOC	225	BIN	{0,1}	0	N	IV tocolysis administered
1	9	EPDILTOC	224	REAL	[0,5]	5	N	Internal OS dilation at IV tocolysis
Y	10	CONTRNUM	436	INT	[1,5]	0	Y	# Visits with observed contractions

OUTN			3001				(MP05) Q.B	
1	EPCongen	281	BIN	{N,Y}	6	N	Congenital Malformations	
2	EPDELIV	283	CAT	{1-4}	4	N	Type of delivery	
3	EPSEX	284	BIN	{1,2}	2	N	Gender of infant	
4	EPAPGAR1	285	INT	[0,10]	30	N	1-minute APGAR	
5	EPAPGAR5	286	INT	[0,10]	29	N	5-minute APGAR	
6	EPLNGTH	287	REAL	[21,61]	99	N	Length	
7	EPWEIGHT	288	INT	[195,6390]	5	N	Birth weight	
8	LBW	321	BIN	{0,1}	5	N	Low birth weight < 2500 gms	
9	FGR	322	BIN	{0,1}	5	N	Fetal growth retardation (brenner)	
10	EPHC	289	REAL	[18,55]	145	N	Head circumference	
11	EPABDOM	290	REAL	[19,50]	2583	N	Abdominal circumference	

OUTS			3001				(MP05) Q.B11	
1	EPSTATUS	291	CAT	{1-5}	7	N	Infant status at <= 30 days	
2	STLBRN	292	BIN	{0,1}	0	N	Stillbirth	
3	INFDTN	293	BIN	{0,1}	0	N	Infant death	
1	4	DEATHDAYS	444	INT	[0,6]	0	Y	Neonate age at death in days
5	HOSPDAYS	294	INT	[0,230]	8	N	Infant days in hospital	
6	DISAGE	435	INT	[0,16]	8	Y	Neonatal age at discharge	

IPRE			124				(MP06) Q.A7	
1	CLAMNION	258	BIN	{N,Y}	1	N	Clinical amnionitis	
2	ABRUPT	259	BIN	{N,Y}	0	N	Abruptio placenta	
3	DISTRESS	260	BIN	{N,Y}	0	N	Fetal distress	
4	IUGR	261	BIN	{N,Y}	0	N	Intra-uterine growth restriction	
5	FDEATH	262	BIN	{N,Y}	0	N	Fetal death	
6	PREEC	263	BIN	{N,Y}	0	N	Preeclampsia	
7	FOTHR	264	BIN	{N,Y}	0	N	Other reason	
8	OLIGO	442	BIN	{N,Y}	0	N	Oligohydramnios	
9	PREV_CS	443	BIN	{N,Y}	0	N	Previous C-section	

CHORS			33				(MP06) Q.A8
1	FPTEMP	235	BIN	{N,Y}	0	N	Temp > 100.4 F or 38.0 C
2	FPTENDER	236	BIN	{N,Y}	0	N	Unexpl. UT tenderness at hosp adm.
3	FPFOUL	237	BIN	{N,Y}	0	N	Foul vaginal discharge
4	FPWBC	238	BIN	{N,Y}	0	N	WBC > 20,000
5	FPTACHY	239	BIN	{N,Y}	0	N	Maternal tachycardia (HR > 100)
6	FPFETACH	240	BIN	{N,Y}	0	N	Fetal tachycardia (HR > 160)

PDELM			433				(MP06) Q.A	
1	FPCONTRA	229	BIN	{N,Y}	1	N	Recurrent contractions at hosp. adm.	
2	FPHYDRA	230	BIN	{N,Y}	0	N	Hydration during adm.	
3	FPTOCOL	231	BIN	{N,Y}	0	N	Tocolysis during adm.	
4	FPANTIB	232	BIN	{N,Y}	1	N	Antibiotics during adm.	
5	FPOTHER	233	BIN	{N,Y}	1	N	Other treatments during adm.	
6	TRTADM	257	BIN	{0,1}	0	N	Any treatment during PTD adm.	
7	FPCHORIO	234	BIN	{N,Y}	0	N	Chorioamnionitis suspected	
Y	-	CHORS	-	GRP	-	0	N	Reasons for suspected chorio
8	FPDIAG	241	BIN	{N,Y}	407	N	Diagnosis of chorio made before PROM	
9	FPENDO	242	BIN	{N,Y}	0	N	Endometritis	
10	FPCSECT	243	BIN	{N,Y}	0	N	C-section wound infection at hosp. adm.	
11	FPTHROMB	244	BIN	{N,Y}	0	N	Spetic pelvic thrombophlebitis	
12	FPEPISIO	245	BIN	{N,Y}	0	N	Episiotomy wound infection	
13	FPUTI	246	BIN	{N,Y}	0	N	UTI diagnosed post-partum	
14	FPPYELON	247	BIN	{N,Y}	0	N	Pyelonephritis	
15	FPDILADM	254	REAL	[0,10]	19	N	Dilation on adm.	
16	TIMEADDG	315	INT	[1,39]	1	N	Time from admission to discharge	

PDELC			420				(MP06) Q.B	
1	FPNICADM	296	BIN	{N,Y}	0	N	Admitted to the NICU	
2	FPRDS	297	BIN	{N,Y}	5	N	Respiratory distress syndrome	
3	FPNEC	298	BIN	{N,Y}	5	N	Necrotizing enterocolitis	
4	FPVENT	299	BIN	{N,Y}	9	N	Mechanical ventilation	
5	FPSEPSIS	300	BIN	{N,Y}	5	N	Newborn sepsis suspected	
Y	6	SEPSIS	316	INT	[0,10]	0	N	Neonate age at suspected sepsis
7	FPCULTUR	301	BIN	{N,Y}	5	N	Newborn sepsis confirmed by culture	
Y	8	CULT	317	INT	[0,41]	0	N	Neonate age at confirmed sepsis
9	FPNEUMO	302	BIN	{N,Y}	5	N	Pneumonia confirmed by X-ray or cult.	
Y	10	NEUMO	318	INT	[0,6]	0	N	Neonate age at confirmed pneumonia
11	FPCSF	303	BIN	{N,Y}	5	N	Meningitis confirmed by culture	
12	FPIVH	305	?	{0-4,9}	30	N	IVH grade	

Fill-in Steps (Structurally Absent Variables)

GRP	#	Fill-in Rule
PPHD	1...13	{? → 0}
PPHD	14	{? → AGEMOM}
JOB	1, 3..5	{? → N}
JOB	7, 9...14	{? → N}
JOB	8	{? → Y}
JOB	6,15	{? → 0}
INFEC	1...10	{? → N}
OBST	3	{? → 0}
OBST	5	{? → N}
SAD	2...5	{? → 0}
SAD	8...10	{? → 0}
SAD	12...14	{? → 0}
SAD	11, 16...24	{? → N}
CPH	8...10	{? → 0}
CPH	20	{? → K}
CRVM	3,4,11	{? → 0}
PSYCH	1...12	{? → mean(variable)}
VISIT2	1	{? → U}
VISIT2	2	{? → 0}
VISIT4	1	{? → U}
VISIT4	2	{? → 0}
INFEC3	1...9	{? → N}
SAD3	2, 4, 5	{? → 0}
SAD3	7...9	{? → N}
CPH3	4...6	{? → 0}
CPH3	12	{? → K}
CRVM3	3,4,11	{? → 0}
IPRE	1...9	{? → N}
CHORS	1...6	{? → N}

Fill-in Steps (Randomly Missing Variables)

GRP	#	Fill-in Rule
DMG	5	{? → 4}
SAD	2,3	{? → CIG24}
SAD	11	{? → N}
SAD	16,19,20	{? → N}
SAD	17	{? → DCOCAINE}
SAD	18	{? → DMARI}
SAD	21	{? → DOTHER}
SAD	22...24	{? → N}
MEDT	5	{? → N}
CPH	1	{? → 0}
CPH	12	{? → N}
SPEC	2,3	{? → 0}
CRVM	2,6,7,11	{? → mean}
CRVM	5	{? → -3}
CRVM	8	{? → 0}
CRVM	10	{? → U}
CPM	1...4	{? → mean(variable)}
CPM	8	{? → 0}
VISIT2	1	{? → U}
VISIT2	2	{? → 0}
VISIT4	1	{? → U}
VISIT4	2	{? → 0}
SAD3	2	{? → CIGSNOW}
CPH3	8	{? → N}
SPEC3	2,3	{? → 0}
CRVM3	2,6,7,11	{? → mean}
CRVM3	5	{? → -3}
CRVM3	8	{? → 0}
CRVM3	10	{? → U}
CPM3	1,2	{? → mean(variable)}
CPM3	5	{? → 0}
IPRE	1	{? → N}

Simple Processing Steps

GRP	#	Name	#	Processing Rule
DMG	11	AGEMOM	416	{'<=17' → 17, '>=40' → 40}
DMG	13	SCHOOLYR	418	{'<9' → 8, '>16' → 17}
PPH	16	MISC1ST	434	{'>3' → 4}
SAD	2	CIGSPRE	419	{'>=40' → 40}
SAD	3	CIG12	420	{'20-29' → 25, '>=30' → 30}
SAD	4	CIG24	421	{'>20' → 25}
SAD	5	CIGSNOW	422	{'>20' → 25}
SAD	8	DRINKPRE	424	{'>=6' → 6, '4-5' → 4.5}
SAD	9	DRK12	425	{'>=4' → 4, '2-3' → 2.5}
SAD	10	DRK24	426	{'>=2' → 2}
SAD	13	DRNKMOST	427	{'>=4' → 4}
SAD	14	DRKLSTDY	428	{'>1' → 2}
SAD3	2	VCIGDAY	423	{'>20' → 25}
SAD3	4	VALCDY	429	{'>3' → 4}
SAD3	5	VALCWK	430	{'>2' → 3}

OUTS	4	DEATHDAYS	444	{'>=6' → 6}
------	---	-----------	-----	-------------

Nontrivial Processing Steps:

Group: CRVM

Case: 10.BPMEMPRO = Y & 11.BPMAXPEN = ?
Rule: Set F11 to {16.00}.

Case: 1.BPCERLNG = ? & 9.BPCRVLT = ?
Rule: Set F1,F9 to {2.71,35.24}.

Group: CRVM3

Case: 10.DPMEMPRO = Y & 11.DPMAXPEN = ?
Rule: Set F11 to {14.04}.

Case: 1.DPCERLNG = ? & 9.DPCVLTH = ?
Rule: Set F1,F9 to {2.55,33.64}.

Group: CPM

Case: Too many missing values
Rule: Remove 5.BPHEMVAL, 6.BPGLOVAL, 7.BPALPVAL

Case: 11.WGTPRE = ?
Rule: Set F11 to {F10 - mean(F10 - F11)}.

Case: 10.WEIGHTV1 = ?
Rule: Set F10 to {F11 + mean(F10 - F11)}.

Case: 9.HEIGHT = ?
Rule: Set F9 to {mean(F9 s.t. F11 is within a range of one)}.

Group: CPM3

Case: Too many missing values
Rule: Remove 3.DPHEMVAL, 4.DPGLOVAL

Case: 6.WEIGHTV3 = ?
Rule: Set F6 to {WEIGHTV1 + mean(F6 - WEIGHTV1)}.

Nontrivial Processing Steps:

Group: DMG

Case: 8.BPLASTYR & 9.BPLST5YR = ?
Rule: Set F8,F9 to {1,2}

Case: 9.BPLST5YR = ?
Rule: Set F9 to {8}

Case: 5.BPINCOME = 4
Rule: Set F5 to the average F5 value of all examples with identical F4 (BPINSUR) value.

Group: JOB

Case: 2.GASTOP_D = ? & 1.BPWORK = ?
Rule: Set F2 to {0}.

Case: 2.GASTOP_D = ? & 1.BPWORK = Y
Rule: Set F2 to {VISGA1_D}.

Group: FFN / FFN2 / FFN3 / FFN4

Case: All features are missing
Rule: Set F1...F6 to {mean(F1),mean(F2),0,0,0,0}

Case: 1.FFNC# & 2.FFNV# = ?
Rule.a: If 3.FFCEA = 0, Set F1 to {mean(F1)}.
Rule.b: If 3.FFCEA = 1, Set F1 to {mean(F1 s.t. F3 = 1)}.
Rule.c: If 4.FFVAA = 0, Set F2 to {mean(F2)}.
Rule.d: If 4.FFVAA = 1, Set F2 to {mean(F2 s.t. F4 = 1)}.

Case: 1.FFNC# <OR> 2.FFNV# = ?
Rule: If F1 = ?, Set F1 to {F2}. If F2 = ?, Set F2 to {F1}.

Case: 3.FFNCEA = ?
Rule: If F1 < 50, Set F3 to {0}, else, Set F3 to {1}.

Case: 4.FFNVA = ?
Rule: If F2 < 50, Set F4 to {0}, else, Set F4 to {1}.

Case: 5.FF# = ?
Rule: If F3 == 0 & F4 == 0, Set F5 to {0}, else, Set F5 to {1}.

Case: 6.FFA = ?
Rule: If F3 == 0 & F4 == 0, Set F6 to {0}, else, Set F6 to {1}.

Group: SAD

Case: 7.BPALCHOL = Y
Rule: Set F8...F10 to {mean(F8...F10 s.t. F7 = Y)}.

Group: SAD3

Case: 3.DPALCHOL = Y & 5.VALCWK = ?
Rule.a: If F4 = 1, Set F5 to {mean(F5 s.t. F4 = 1)}.
Rule.b: If F4 = 2, Set F5 to {mean(F5 s.t. F4 = 2)}.

Group: CPH

Case: 7.BPCONTRA = Y & F8...F10 = (?, ?, ?)
Rule: Set F8...F10 to {2,4,12}.

Case: 7.BPCONTRA = Y & F9...F10 = (?, ?)
Rule.a: Set F9 to {F8 + median(F9 - F8)}.
Rule.b: Set F10 to {F8 + median(F10 - F8)}.

Case: 13.BPINTNUM = ?
Rule.a: If F11 = Y, Set F13 to {median(F13 s.t. F11 = Y)}.
Rule.b: If F11 = N, Set F13 to {median(F13 s.t. F11 = N)}.

Group: CPH3

Case: 3.DPCONTRA = Y & F4...F6 = (?, ?, ?)
Rule: Set F4...F6 to {2,6,17.5}.

Case: 3.BPCONTRA = Y & 4.BPCONHR = ?
Rule: Set F4 to {F5}.

Case: 3.BPCONTRA = Y & 5.BPCONDAY = ?
Rule: Set F5 to {F4 + median(F5 - F4)}.

Case: 3.BPCONTRA = Y & 6.BPCONWK = ?
Rule: Set F6 to {F5 + median(F6 - F5)}.

Case: 9.DPINTNUM = ?
Rule.a: If F7,F8 = (N,N), Set F9 to {median(F9 s.t. F7,F8 = N,N)}.
Rule.b: If F7,F8 = (N,Y), Set F9 to {median(F9 s.t. F7,F8 = N,Y)}.
Rule.c: If F7,F8 = (Y,N), Set F9 to {median(F9 s.t. F7,F8 = Y,N)}.

Group: SPEC / SPEC3

Case: 1.VAGPH = ? & 4. GSTOT = ?
Rule: Set F1,F4 to {4.5,3}.

Case: 1.VAGPH = ?
Rule.a: If F4 < 7, Set F1 to {4.5}.
Rule.b: If F4 >= 7, Set F1 to {mean(F1 s.t. F4 >= 7)}.

Case: 4.GSTOT = ?
Rule.a: If F1 > 4.5, Set F4 to {mean(F4 s.t. F1 > 4.5)}.
Rule.b: If F1 <= 4.5, Set F4 to {mean(F4 s.t. F1 <= 4.5)}.

Case: 5.BV = ?
Rule: If F1 > 4.5 & F4 >= 7 Set F5 to {1}, else, Set F5 to {0}.

Group: CRVM / CRVM3

Case: 1.CERLNG = ?
Rule: Set F1 to {F9 / mean(F9 / F1)}.

Case: 9.CRVLT = ?
Rule: Set F9 to {F1 * mean(F9 / F1)}.