MHS GENESIS Person File
for the
MHS Data Repository (MDR)
(Version 2.00.01)

Future Specification

Revision History

Version	Date	Originator	Para/Tbl/Fig	Description of Change
1.00.00	01/15/2018	Wendy Funk	• Initial Document	• Initial Document
1.01.00	02/06/2019	Tracy Comer for Wendy Funk	• Section 3 • Table 2	• Adjusted wording regarding restricted file • Added wording to place person_key field in the restricted file
2.00.00	09/21/2023	Wendy Funk	• Entire Document	• Incorporated BDE 3.0
2.00.01	07/05/2024	Keith Hofmann	• Section VI, Table 2	• Added merge to MPI to get PARC and enhanced sponsor SSN

MDR GENESIS Person File

I.   SOURCE:

The source system is the Cerner Millennium database. BDE 3.0 based records are from the person table. BDE 2.4 records are from the WH_CLN_PERSON_REF table.

II.   TRANSMISSION (FORMAT AND FREQUENCY)

TBD

III.   ORGANIZATION AND BATCHING

Source data: The first step in MDR processing is to batch records received from MHS Genesis. Raw data batches are stored in /mdr/raw/genesis according to routine MDR operating procedures.

Output products:

Two SAS datasets will be produced from MHS Genesis person data. These datasets are described below.

File Name	Description	Location
Person File	File with most patient information and ordered by the MHS Genesis Person Key	/mdr/pub/genesis/person/ person.sas7bdat – BDE 2.4 records only. This file stopped being updated in October 2023. /mdr/pub/genesis3/person/ person.sas7bdat – BDE 3.0 records combined with BDE 2.4
Person File Restricted	File with patient addresses, phone numbers, and email addresses and ordered by the MHS Genesis Person Key	/mdr/restricted/genesis/person/ person.sas7bdat – BDE 2.4 records only. This file stopped being updated in October 2023. mdr/restricted/genesis3/person/ person.sas7bdat – BDE 3.0 records combined with BDE 2.4

IV.   RECEIVING FILTERS

All records were provided with the initial batch of data. Thereafter, new and changed records are sent each week.

V.   UPDATE PROCESS

After sending the initial batch of person data (11/16/2017), the raw feeds for the person file represent either insert and update records. These records shall be used to maintain the master MDR person dataset.

The primary key for the person table is the PERSON_KEY field (BDE 2.4) and PERSON_ID field (BDE 3.0). During the extraction of the raw person records, de-duplication of records, or anytime a person key collision occurs between incoming data and existing master data, the processor de-duplicates data by selecting the record with the most recent value of the update date/time (UPDT_DT_TM) for any multiple of records with the same primary key.

When combining BDE 2.4 and BDE 3.0 data, BDE 3.0 data is prioritized. If a record exists with data from BDE 2.4 (by PERSON_SK) and a new record is obtained from BDE 3.0 (by PERSON_ID), the information from BDE 3.0 will

overwrite the information from BDE 2.4. When BDE 3.0 is complete, it is expected that BDE 2.4 will only contribute to a minority of records. This may not be broadly true due to changes in data syndication rules over time.

Once the dataset has been updated, the processor assigns many other internally-derived variables as described in Table 2.

VI.   FIELD TRANSFORMATIONS AND DELETIONS FOR MDR CORE DATABASE

This section of this functional specification describes the data merges that are necessary to append fields in the MDR Genesis Person file. Table 1 describes the reference files that are used in processing.

Merge	Date Matching	Additional Matching
Person Alias Table (person_alias)	Latest update for each type of alias	BDE 2.4: person_sk BDE 3.0: person_id
Code Value table (code_value)	N/A	BDE 2.4: code value ref, health_system_id; active_ind BDE 3.0: code_value, code_set, active_ind = 1
Person Name Table (person_name)	Latest update	BDE 2.4: person_sk, health_system_source_id where src_active_ind=1 BDE 3.0: person_id and name_type_cd = 766 and active_ind = 1
Phone Table (phone)	Latest update for each type of phone	BDE 2.4: person_sk= src_parent_entity_sk (and src_parent_entity_name=person, src_active_ind=1, health_system_source_id=18635 BDE 3.0: person_id=parent_entity_id, parent_entity_name = PERSON, phone_type_cd, active_ind = 1
Address (BDE 3.0 only) (address)	Latest update for each type of address	BDE 3.0: person_id = parent_entity_id, parent_entity_name = PERSON, address_type_cd, active_ind = 1
Person Patient Table (BDE 3.0 only) (person_patient)	Latest update	BDE 3.0: person_id
DEERS VM-6 Beneficiary File	Most recent file	EDIPN, where primary record is 1 and eligibility indicator is 1. Only match where src_active_ind
Encounter Table (encounter)	N/A	BDE 2.4: person_sk BDE 3.0: person_id
Person to Person Relationship Table (person_person_reltn)	N/A	BDE 2.4: person_sk BDE 3.0: person_id
Master Person Index (MPI)	Most recent file	edipn, ssn, sponssn, gender, last_name, first_name, birth_dt
Person Info Table (person_info)	end_effective_dt_tm > the current date	BDE 3.0: person_id, info_sub_type_cd = 2678703703, active_ind = 1

Business rules for each of the appended fields are described in the body of the format table in Section VII.

VII.   FILE LAYOUT

The MDR Genesis Person file is stored in a SAS data set. Table 2 provides the file layout and processing rules.

Table 2 File Layout for MDR GENESIS Person File

Field	Format	SAS Name	BDE 2.4 Field	BDE 3.0Field	Transformation
MHS Genesis Unique Person ID	$100	person_sk	person_sk	person_id	No transformation
MHS Genesis Person Key	8	person_key	person_key	null	BDE 2.4: No transformation BDE 3.0: Null
Health System ID	8	hs_id	health_system_id	null	BDE 2.4: No transformation BDE 3.0: Null
Gestational Age	8	gest_age	gestational_age_days_at_birth	person_patient.gest_age_at_birth	BDE 2.4: No transformation BDE 3.0: Merge to person_patient table by person_id
Sex Change Flag	$1	sex_chg	sex_change_ind	sex_age_change_ind	No transformation
Birth Weight	8	birth_weight	birth_weight	person_patient.birth_weight	BDE 2.4: No transformation BDE 3.0: Merge to person_patient table by person_id
First Name	$100	first_name	first_name	name_first	No transformation
Middle Name	$100	mid_name	middle_name	name_middle	No transformation
Last Name	$100	last_name	last_name	name_last	No transformation
Full Name	$100	full_name	full_name	name_full_formatted	No transformation
Rank	$100	rank	employee_title	null	BDE 2.4: Set to employee_title BDE 3.0: null
Death Date/Time	Date/Time	death_dt	deceased_dt_tm	deceased_dt_tm	BDE 2.4: Convert to local time using deceased_tm_zn BDE 3.0: Convert to local time using deceased_tz. The numeric code might need to be converted to the time zone name (e.g. 35 = America/Los_Angeles)
Begin Effective Date/Time	Date/Time	beg_dt	beg_effective_dt_tm	beg_effective_dt_tm	No transformation
End Effective Date/Time	Date/Time	end_dt	end_effective_dt_tm	end_effective_dt_tm	No transformation.
Extract Date to MDR	Date/Time	extract_dt_tm	extract_dt_tm	dl_eff_dt	No transformation

Field	Format	SAS Name	BDE 2.4 Field	BDE 3.0Field	Transformation
Update Date/Time	Date/time	updt_dt_tm	updt_dt_tm	updt_dt_tm	No transformation.
Fields from the Code Value Reference Table					
Gender	$10	gender	alias	display	BDE 2.4: Set to alias where gender_ref matches code_value_ref and code_set=gender (57) BDE 3.0: Set to display where sex_cd matches code_value and code_set = gender (57)
Ethnicity	$22	ethnic	alias	display	BDE 2.4: Set to alias where ethnic_group_ref matches code_value_ref and code_set=ethnicity (27) and contributor_source_ref 105499759 BDE 3.0: Set to display where ethnic_grp_cd matches code_value and code_set = ethnicity (27)
Race	$41	race	alias	display	BDE 2.4: Set to alias where race_ref matches code_value_ref and code_set= race (282) and contributor_source_ref = '105099617' BDE 3.0: Set to display where race_cd matches code_value and code_set = race (282)

Field	Format	SAS Name	BDE 2.4 Field	BDE 3.0Field	Transformation
Marital Status	$20	marital	alias	display	BDE 2.4: Set to alias where marital_status_ref matches code_value_ref and code_set= marital status (38) BDE 3.0: Set to display where marital_type_cd matches code_value and code_set = marital status (38)
Language	$20	language	alias	display	BDE 2.4: Set to alias where language_ref matches code_value_ref and code_set= language (36) BDE 3.0: Set to display where language_cd matches code_value and code_set = language (36)
Birth Weight Units	$10	birth_weight_units	desc_meaning	display	BDE 2.4: Set to desc_meaning where birth_weight_units matches code_value_ref and the code set = birthweight (14659) BDE 3.0: Merge to person_patient, retrieve birth_weight_units_cd, then set to display where birth_weight_units_cd matches code_value and code_set = weight units (14659)
Religion	$40	religion	desc_display	display	BDE 2.4: Set to desc_display where religion_ref matches code_value_ref and code_set= religion (49) BDE 3.0: Set to display where religion_cd matches code_value and code_set = religion (49)
Fields from the Person Name Table					
Name suffix	$100	suffix	name_suffix	name_suffix	No Transformation

Field	Format	SAS Name	BDE 2.4 Field	BDE 3.0Field	Transformation
Fields from the Phone Number Table					
Home Phone Number	$20	home_phone	phone_nbr	phone_num	BDE 2.4: Most recently reported phone_nbr (based on begin and end effective dates), where phone_type_ref=home (170). Only available in /restricted. BDE 3.0: Set to phone_num where phone_type_cd = home (170) and active_ind = 1 and parent_entity_name = 'PERSON'
Work Phone Number	$20	work_phone	phone_nbr	phone_num	BDE 2.4: Most recently reported phone_nbr (based on begin and end effective dates), where phone_type_ref=work (163). Only available in /restricted. BDE 3.0: Set to phone_num where phone_type_cd = work (163) and active_ind = 1 and parent_entity_name = 'PERSON'
Cell Phone Number	$20	cell_phone	phone_nbr	phone_num	BDE 2.4: Most recently reported phone_nbr (based on begin and end effective dates), where phone_type_ref=cell (4149712). Only available in /restricted. BDE 3.0: Set to phone_num where phone_type_cd = cell (4149712) and active_ind = 1 and parent_entity_name = 'PERSON'

Field	Format	SAS Name	BDE 2.4 Field	BDE 3.0Field	Transformation
Fields from the Person-Person Relationship File					
Relationship Flag	8	reltn_fl	N/A	null	If person_sk has a record in the Person-Person Relationship table (person_person_reltn), set to 1.
Fields from the Encounter File					
Workload Flag	8	wkld_fl	N/A	null	If person_sk has a record in the Encounter table, and the encounter type is not in('History','Lifetime Pharmacy') (20058643 , 19962609), then set to 1.
Fields from the Person Alias File					
MRN	$40	mrn	alias	alias	BDE 2.4: Set to Alias, where alias_type_ref=MRN (10) BDE 3.0: Set to MRN where person_alias_type_cd = MRN (10) and active_ind = 1
EDIPN	$10	edipn	alias	alias	BDE 2.4: Fille with alias, where alias_type_ref=EDIPN (22) BDE 3.0: Set to alias where person_alias_type_cd =EDIPN (22) and active_ind = 1
SSN	$9	ssn	alias	alias	BDE 2.4: Set to alias, where alias_type_ref=SSN (18). If length is not 9, then set to null. BDE 3.0: fille with alias where person_alias_type_cd = SSN (18) and active_ind = 1
Interim Person Identifier	$40	ipi	alias	null	BDE 2.4: Set to alias, where alias_type_ref=IPI (106127789) BDE 3.0: null

Field	Format	SAS Name	BDE 2.4 Field	BDE 3.0Field	Transformation
Foreign Identifier	$40	foreign_id	alias	null	BDE 2.4: Set to alias, where alias_type_ref=Foreign Identifer (106127795) BDE 3.0: null
Fields from the Address File (Code value ref table applied also for some fields, noted in business rule)					
Address Line 1	$100	addr_line1	address_line_1	street_addr	BDE 2.4: No Transformation BDE 3.0: Set to street_addr if address_type_cd = home (756) and active_ind = 1 and parent_entity_name = 'PERSON'
Address Line 2	$100	addr_line2	address_line_2	street_addr2	BDE 2.4: No Transformation BDE 3.0: Set to street_addr2 if address_type_cd = home (756) and active_ind = 1 and parent_entity_name = 'PERSON'
Address Line 3	$100	addr_line3	address_line_3	street_addr3	BDE 2.4: No Transformation BDE 3.0: Set to street_addr3 if address_type_cd = home (756) and active_ind = 1 and parent_entity_name = 'PERSON'
Address Line 4	$100	addr_line4	address_line_4	street_addr4	BDE 2.4: No Transformation BDE 3.0: Set to street_addr4 if address_type_cd = home (756) and active_ind = 1 and parent_entity_name = 'PERSON'

Field	Format	SAS Name	BDE 2.4 Field	BDE 3.0Field	Transformation
City	$100	city	city	city	BDE 2.4: No Transformation BDE 3.0: Set to city where address_type_cd = home (756) and active_ind = 1 and parent_entity_name = 'PERSON'
State	$30	state	alias	state	BDE 2.4: Set to the alias where state_ref matches code_value_ref and code_set= state (62) BDE 3.0: fille with state where address_type_cd = home (756) and active_ind = 1 and parent_entity_name = 'PERSON'
Postal Code	$11	zip	postal_code	zipcode	BDE 2.4: No Transformation BDE 3.0: Set to zipcode where address_type_cd = home (756) and active_ind = 1 and parent_entity_name = 'PERSON'
Country	$30	country	alias	country	BDE 2.4: Set to alias where state_ref matches code_value_ref and code_set=country (15) BDE 3.0: Set to country where address_type_cd = home (756) and active_ind = 1 and parent_entity_name = 'PERSON'
Email Address	$100	email	email_address	street_addr	BDE 2.4: No Transformation BDE 3.0: Set to street_addr where address_type_cd = email (756) and active_ind = 1 and parent_entity_name = 'PERSON'

Field	Format	SAS Name	BDE 2.4 Field	BDE 3.0Field	Transformation
Fields from the DEERS VM6 (no difference between BDE 2.4 and BDE 3.0)					
DEERS Eligibility Indicator	$1	eligibility_ind	d_mhs_elig_indic	d_mhs_elig_indic	No transformation.
Sponsor SSN	$9	sponssn	spn_pn_id	spn_pn_id	After merging from DEERS, enhance with the Master Person Index (see below).
Beneficiary Category	$3	bencat	r_ben_cat_cd	r_ben_cat_cd	No transformation.
HCDP - Enrolled	$3	hcdp_enr	d_mi_hcdp_pln_cvg_cd	d_mi_hcdp_pln_cvg_cd	Set to d_mi_hcdp_pln_cvg_cd when enrollment group is P, L or U.
HCDP - Assigned	$3	hcdp_asgn	asg_hcdp_ pln_cvg_cd	asg_hcdp_ pln_cvg_cd	Set to asg_hcdp_pln_cvg_cd
PCM ID	$10	pcm_id	d_mi_pcm_id	d_mi_pcm_id	Set to d_mi_pcm_id when enrollment group is P, L or U.
Enrollment MTF	$4	enr_dmisid	d_mi_pcm_edvsn_dmis_id	d_mi_pcm_edvsn_dmis_id	Set to d_mi_pcm_edvsn_dmis_id when enrollment group is P, L or U.
Enrollment Group	$1	enr_grp	d_enr_grp_cd	d_enr_grp_cd	No transformation.
Eligibility Group	$1	elg_grp	d_elg_grp_cd	d_elg_grp_cd	No transformation.
Enrollment PCM Type	$1	pcm_type	d_pcm_typ_cd	d_pcm_typ_cd	No transformation.
DEERS File Date	YYMM	deers_month	d_fde_ext_month	d_fde_ext_month	No transformation.

Field	Format	SAS Name	BDE 2.4 Field	BDE 3.0Field	Transformation
Fields from the Master Person Index (MPI) (no difference between BDE 2.4 and BDE 3.0)					
Person Association Reason Code	$2	parc	parc	parc	Merge to MPI and return the PARC
Internally Derived Fields					
Birth Date/Time	Date/Time	dob	birth_dt_tm	birth_dt_tm	BDE 2.4: No transformation BDE 3.0: Convert to local time using birth_tz. The numeric code might need to be converted to the time zone name (e.g. 35 = America/Los_Angeles). Some time zones are off by one hour, resulting in the wrong date of birth. If the birth_prec_flag = 0 (precise to the day) and the time of birth is exactly 11:00:00 pm, add one hour to the raw birth_dt_tm after converting to local time.
Test Record Indicator	$1	test_record_ind	N/A	N/A	See Appendix A

VIII.   REFRESH FREQUENCY

Daily

IX.   DATA MARTS

N/A.

X.   SPECIAL OUTPUTS

N/A

Appendix A: Test Person Logic

Test patients are included in the raw data from GENESIS with no simple way to extract them. This logic uses a "wide net" approach to flag patients that may be test patients and should be excluded from any analysis.

Test patients can be flagged two ways:
1) By the Test Patient Identifier (info_sub_type_cd = 2678703703) in the person_info table
2) By the patient's name

Some patients will be flagged by one of the rules above but identified post hoc as NOT test patients and should not be considered test patients. These will be addressed at the bottom of this appendix.

To identify test patients via the Test Patient Identifier, merge to the person_info table by person_id where info_sub_type_cd = 2678703703 and active_ind = 1 and end_effective_dt_tm > the current date. If the identifier has any non-zero value (value_cd > 0), that counts as a test patient (except for the exclusions detailed below).

To identify test patients via their name, they must have person_type_cd = 903 (PERSON), active_ind = 1 and end_effective_dt_tm > the current date. Unless otherwise specified, when a name "contains" the strings below it should be used as a "matches pattern". In other words, the indicated name contains that string anywhere within it. For example, if the string indicated is "ZZZ" that means that the string "ZZZ" appears anywhere in the name, the beginning, the end, the middle, or it's the entire name. All the following name patterns count as test patients (except for the exclusions detailed below).
- Last name or first name starts with 'QQ'
- Last name starts with 'CHDR', 'COLLECTIVE', 'DIALYS', 'DONOTUSE', 'GENERIC', 'MANAGER', 'MASCAL', 'MHV', 'UNKNOWN'
- Middle name is 'TEST', PRACTICE' (exact match)
- Last name or first name contains 'PATIENT', 'ZZZ', 'NMCSD', 'SURVEY'
- Last name contains 'ZZTEST', 'GREASEBOARD', 'PRODUCT', 'PROVIDER', 'TRANSFUSION'
- Last name, first name, or middle name contains 'MAMC', 'BAMC', 'CERNER'
- Last name, first name, or middle name contains numeric digits
- Last name is 'APPTEST', 'SMURF', 'USER', 'VAVBASCHEDULEONLY' (exact match)
- First name is 'MICKEY' and last name is 'MOUSE'
- First name contains 'PHERESIS', 'TESTCONFIRM'
- First name starts with 'TRAUMA'
- Last name is 'TEST' (exact match) and first name is any of the following: 'ACCOUNT', 'ALEXIS', 'AMBULATORYONCOLOGYRN', 'ANOTHERKID', 'AUTOMATE', 'BABY', 'CHILD', 'COURTNEY', 'CSR', 'DALOOP', 'DUMMY', 'EDGREASEBOARD', 'EDPHARMACIST', 'EVERETTE', 'EXERCISE', 'EXERCISETHERAP', 'FELLOW', 'GALE', 'HANA', 'HOMECAREOT', 'HOMECAREPT', 'HOMECARERN', 'HOMECARESOCIALWORKERMSW', 'HOMECAREST', 'HOMEHEALTHAIDE', 'HOSPICEVOLUNTEER', 'INTERN', 'IQHCONFIGADMIN', 'IQHEALTHIDENTIFICATIONADMINSTRATOR', 'JOHNNY', 'JULIET', 'LONGTERMCARE', 'LONGTERMCARERECREATIONALTHERAPIST', 'OBRNMANAGER', 'OFFUTT', 'PATIENT', 'PHYSICALTHERAPYSTUDENT', 'PHYSICIANAMBULATORY', 'PHYSICIANAMBULATORY3', 'PHYSICIANONCOLOGIST', 'REHABINTAKECOORDINATOR', 'REHABOCCUPATIONALTHERAPISTSTUDENT', 'REHABPHYSICALTHERAPISTASSISTANT', 'REHABPHYSICALTHERAPISTSTUDENT', 'REMOTEUSER', 'RNONCOLOGY', 'ROLE1PROVIDER', 'ROLE1SECONDARY', 'ROLE1SYSADMIN', 'RT', 'SSN', 'SUPPLYCHAINRECEIVINGCLERK', 'TEST', 'TESTTT', 'TESTUSER', 'VICTOR'

Exclusions: Even if a patient qualifies by any of the above, they are not flagged as a test patient if:
- The person_id is 12609416, 4600665, 9018287, 11528352, 11180977, 7419750, 11667143, 22468914, 5093018, 6213110, 7072696, 5908362, 14894624, or 5601993
- The merge to obtain test patient identifier was described above. Look up value_cd from the person_info table in the code_value table using code set 100451. If the display for value_cd is 'Not a Test Patient' (value_cd = 2678703509) or if the display is null, do not flag this person as a test patient.