A Guide to Identifying Non-Veteran Records in the Inpatient and Outpatient Databases

Samuel King, M.S., M.Div.
Andrea Shane, B.A.
Mark W. Smith, Ph.D.

September 2006

Technical Report #20

Table of Contents

1. Overview

The majority of the individuals who receive medical attention from the Department of Veterans Affairs health care system are individuals who have completed military service and are considered to hold veteran status. However, a small number of patients who are treated within the VA health care system are not veteran individuals. This non-veteran population consists of individuals such as VA employees, the widows of veterans, and the family members of veterans.

In this report, we provide a workable approach for identifying non-veteran records. Thus, researchers who need to identify, analyze or exclude non-veteran records in their studies may find our approach helpful. The identification of non-veteran records is especially important in studies of the healthcare use of female veterans.

2. Databases Used to Identify Non-Veteran Records

The "PM04" SAS inpatient file (i.e., the PTF main file for non-extended care) and the "SF04" SAS outpatient file (i.e., an NPCD outpatient visit dataset) were used to identify non-veteran records in FY2004. Within both files, two variables can be used to identify non-veteran records with high consistency[1]. A third variable can also be used to identify non-veterans records, but yields less consistent results than the other two variables.

3. Variables Used to Identify Non-Veteran Records

Compensation and Pension Status/Eligibility

The compensation and pension status (CP) variable in the inpatient file and the eligibility (ELIG) variable in the outpatient file identify the degree to which the medical condition for which an individual is seeking medical assistance is service connected. In the inpatient file, a CP value of 8 indicates non-veteran status while values of 1-7 signify veteran status. In the outpatient file, ELIG values of 11-17 and 19 indicate non-veteran status while all other values identify veteran status.[2] Table 3.1 provides statistics on the levels of the CP variable while Table 3.2 provides statistics on the levels of the ELIG variable.

[1] In this report, we define the word consistency to mean a lack of fluctuation between veteran and non-veteran status.

[2] While the majority of the records containing values of 11-17 or 19 of the ELIG variable indicate non-veteran status, there may exist some individuals in this group who are indeed veterans, but elect not to use their veteran status to obtain healthcare benefits.

Table 3.1:
Compensation and Pension Status (CP) Values in the FY04 PTF Inpatient Discharge File

Internal Value	Formatted Value	N	%	Indicates Vet or Non-Vet Status
1	Service Connected>10%	54,318	9.51	Vet
2	Service Connected<10%	4,504	0.79	Vet
3	Non-Service Connected + Service Connected>10%	138,318	24.21	Vet
4	Non-Service Connected + Pension + Service Connected<10%	4,154	0.73	Vet
5	Non-Service Connected + Pension	59,429	10.40	Vet
6	Non-Service Connected + Service Connected<10%	17,822	3.12	Vet
7	Non-Service Connected	287,277	50.28	Vet
8	Non-Veteran	5,567	0.97	Non-Vet
	TOTAL	571,389	100.00	

Table 3.2: Eligibility (ELIG) Values in the FY04 NPCD Outpatient File

Internal Value	Formatted Value	N	%	Indicates Vet or Non-Vet Status
-1	Purple Heart	200,055	0.40	Vet
0	Service Connected 0%	1,879,188	3.76	Vet
1	Service Connected 50%	10,762,878	21.54	Vet
2	Authorized Absence/Housebound	822,523	1.65	Vet
3	Service Connected 40%	1,518,559	3.04	Vet
4	Non-Service Connected, Pension	2,662,193	5.33	Vet
5	Non-Service Connected	24,513,942	49.06	Vet
7	Service Connected 30%	1,804,239	3.61	Vet
8	Service Connected 20%	1,832,481	3.67	Vet
9	Service Connected 10%	2,880,621	5.77	Vet
11	CHAMPVA	213,650	0.43	Non-Vet
12	Collateral	30,974	0.06	Non-Vet
13	Employee	509,762	1.02	Non-Vet
14	Other Federal	37,086	0.07	Non-Vet
15	Allied Veteran	3,820	0.01	Non-Vet
16	Humanitarian	42,999	0.09	Non-Vet
17	Sharing	179,884	0.36	Non-Vet
18	Reimbursable	193	0.00	Vet
19	TRICARE/CHAMPUS	71,221	0.14	Non-Vet
	TOTAL	49,966,268	100.00	

Means Test Indicator

The means test indicator, or the MEANS variable, determines if a veteran is required to make co-payments as part of his/her VA healthcare benefits. The MEANS variable exists in both the inpatient and outpatient files. This variable contains a specific value (N) to indicate non-veteran status while values AN, AS, G, and C of the MEANS variable signify veteran status. Caution must be taken with two MEANS values: U (Not Done) and X (Not Applicable.) These values are ambiguous and may

signify either veteran or non-veteran status. However, only 3.56% of the total records in the SF04 file and 1.26% of the records in the PM04 file contain these values. Tables 3.3 and 3.4 provide statistics on the levels of the MEANS variable for the inpatient and outpatient files, respectively.

Table 3.3: Means Test Indicator (MEANS) Values in the FY04 PTF Inpatient Discharge File

Internal Value	Formatted Value	N	%	Indicates Vet or Non-Vet Status
AN	Co-Pay Exempt Non-Service Connected	297,129	52.00	Vet
AS	Co-Pay Exempt Service Connected/Special	213,773	37.41	Vet
C	Means Test Co-Pay Required	37,295	6.53	Vet
G	Geographic Means Test Co-Pay Required	10,478	1.83	Vet
N	Non-Veteran	5,514	0.97	Non-Vet
U	Not Done	3,419	0.60	Ambiguous
X	Not Applicable	3,781	0.66	Ambiguous
	TOTAL	571,389	100.00	

Table 3.4: Means Test Indicator (MEANS) Values in the FY04 NPCD Outpatient File

Internal Value	Formatted Value	N	%	Indicates Vet or Non-Vet Status
AN	Co-Pay Exempt Non-Service Connected	19,659,576	39.35	Vet
AS	Co-Pay Exempt Service Connected/Special	19,398,724	38.82	Vet
C	Means Test Co-Pay Required	6,845,932	13.70	Vet
G	Geographic Means Test Co-Pay Required	1,190,485	2.38	Vet
N	Non-Veteran	1,091,990	2.19	Non-Vet
U	Not Done	450,030	0.90	Ambiguous
X	Not Applicable	1,329,531	2.66	Ambiguous
	TOTAL	49,966,268	100.00	

Period of Service

Period of Service is indicated by the PSX variable in the inpatient file and the POS variable in the outpatient file. The Period of Service variable identifies the war or conflict during which a veteran served. However, some values of the Period of Service variable do not identify a war or conflict, but rather indicate how an individual is associated with the VA. For example, values A, B, and C indicate individuals who are Army, Navy, or Air Force active duty participants, respectively, but have received medical attention from a VA hospital.

As indicated in Tables 3.5 and 3.6, values A-V of the PSX and POS variables signify non-veteran status while values 0-8 and W-Z indicate veteran status. Special caution must be taken with value 9 (Other or None.) In many cases, multiple records pertaining to a single individual fluctuate between value 9 and other values of the PSX and POS variables. However, only 0.56% of the total records in the inpatient file and 2.69% of the records in the outpatient file contain a value of 9. We discuss this issue in greater detail in Section 5 of this technical report.

Table 3.5:
Period of Service (PSX) Values in the FY04 PTF Inpatient Discharge File [3]

Internal Value	Formatted Value	N	%	Indicates Vet or Non-Vet Status
A	Army	602	0.11	Non-Vet
B	Navy or Marine	539	0.09	Non-Vet
C	Air Force	358	0.06	Non-Vet
D	Coast Guard	26	0.00	Non-Vet
E	Retired Military	45	0.01	Non-Vet
H	Public Health Service	1	0.00	Non-Vet
I	Observation & Examination	2	0.00	Non-Vet
J	Workers Compensation	1	0.00	Non-Vet
M	Foreign Government	3	0.00	Non-Vet
N	Emergency	173	0.03	Non-Vet
P	Contract	7	0.00	Non-Vet
Q	Other Federal	2	0.00	Non-Vet
R	Donors	4	0.00	Non-Vet
S	Special Study	172	0.03	Non-Vet
T	Other Non-Veteran	1,402	0.25	Non-Vet
U	Survivor-CHAMPVA	1,341	0.23	Non-Vet
V	CHAMPUS	161	0.03	Non-Vet
X	Desert Storm (Veteran)	21,140	3.70	Vet
Y	Philippine Veterans	27	0.00	Vet
Z	Merchant Marine	355	0.06	Vet
0	Spanish-American	8	0.00	Vet
1	World War I	9	0.00	Vet
2	World War II	112,273	19.65	Vet
3	Pre-Korea	3,743	0.66	Vet
4	Korean	81,767	14.31	Vet
5	Post-Korean	44,596	7.80	Vet
6	Vietnam	236,734	41.43	Vet
7	Post-Vietnam	62,662	10.97	Vet
8	Desert Storm	27	0.00	Vet
9	Other or None	3,209	0.56	Ambiguous
	TOTAL	571,389	100.00	

[3] Values F, G, K, L, O, and W of the PSX variable do not exist.

Table 3.6: Period of Service (POS) Values in the FY04 NPCD Outpatient File [4]

Internal Value	Formatted Value	N	%	Indicates Vet or Non-Vet Status
	Missing Value	1,142	0.00	Missing
A	Army—Active Duty	6,741	0.01	Non-Vet
B	Navy, Marine—Active Duty	2,562	0.01	Non-Vet
C	Air Force—Active Duty	2,468	0.00	Non-Vet
D	Coast Guard—Active Duty	704	0.00	Non-Vet
E	Retired, Uniformed Forces	104	0.00	Non-Vet
F	Medical Remedial List	2	0.00	Non-Vet
G	Merchant Seamen— U.S. Public Health Service	14	0.00	Non-Vet
H	Other U.S. Public Health Service Beneficiaries	65	0.00	Non-Vet
I	Observation/Examination	342	0.00	Non-Vet
J	Office of Workers Compensation	13	0.00	Non-Vet
M	Beneficiaries-Foreign Government	113	0.00	Non-Vet
N	Humanitarian (Non-Veteran)	13,957	0.03	Non-Vet
P	Others Reimbursable (Non-Veteran)	541	0.00	Non-Vet
Q	Other Federal-Dependent	175	0.00	Non-Vet
R	Donors (Non-Veteran)	126	0.00	Non-Vet
S	Special Studies (Non-Veteran)	12,369	0.02	Non-Vet
T	Other Non-Veteran	44,255	0.09	Non-Vet
U	CHAMPVA—Spouse, Child	488	0.00	Non-Vet
V	CHAMPUS	31,575	0.06	Non-Vet
W	Czechoslovakia/Poland Service	8	0.00	Vet
X	Persian Gulf War	2,705,591	5.41	Vet
Y	CAV/NPS	390	0.00	Vet
Z	Merchant Marine	35,863	0.07	Vet
0	Korean	6,998,241	14.01	Vet
1	World War I	2,322	0.00	Vet
2	World War II	9,055,895	18.12	Vet
3	Spanish American	4,483	0.01	Vet
4	Pre-Korean	345,987	0.69	Vet
5	Post-Korean	3,933,118	7.87	Vet
6	Operation Desert Shield	6,152	0.01	Vet
7	Vietnam Era	20,163,018	40.35	Vet
8	Post-Vietnam	5,255,622	10.52	Vet
9	Other or None	1,341,822	2.69	Ambiguous
	TOTAL	49,966,268	100.00	

[4] Values K, L, and O of the POS variable do not exist.

4. Identifying Non-Veteran Records

Programming Approach for Assigning Non-Veteran Values to the Patient Records
In order to assign non-veteran status to records, we first prepared SAS code to assign an indicator variable for each of the variables in Tables 3.1 – 3.6 (Means Test Indicator, Period of Service, and Compensation and Pension Status [in the inpatient file] or Eligibility [in the outpatient file].) Each of the indicator variables has a value of either 1 (for non-veteran status) or 0 (for other than non-veteran status), based upon the level of the variables in Tables 3.1 – 3.6. We then created another variable (NON_VET_COUNTER) that counted (for each record) the number of indicator variables that were assigned 1 (for non-veteran status.) Thus, the possible values of the NON_VET_COUNTER variable (either 0, 1, 2 or 3) on each record represent the number of non-veteran values for that record.

The 0 Value
A NON_VET_COUNTER value of 0 identifies a record that does not contain any non-veteran indicator variables. For example, an inpatient record that contains the following values would be considered a veteran record since the values of all three variables indicate veteran status:

- MEANS value of AN (Co-Pay Exempt Non-Service Connected, a veteran value),
- CP value of 6 (Non-Service Connected + Service Connected<10%, a veteran value), and
- PSX value of 2 (World War II Veteran, a veteran value.)

The 3 Value
A NON_VET_COUNTER value of 3 indicates that a record contains non-veteran values for all three variables and is therefore considered to be a non-veteran record. For example, we would consider an inpatient record that contains the following values to be a non-veteran record since the values of all three variables indicate non-veteran status:

- MEANS value of N (Non-Veteran, a non-veteran value),
- CP value of 8 (Non-Veteran, a non-veteran value), and
- PSX value of T (Other Non-Veteran, a non-veteran value.)

The 1 and 2 Values
Records with NON_VET_COUNTER values of 0 or 3 present no inconsistencies between any values of the relevant variables (MEANS, POS/PSX, and CP/ELIG.) However, records with a NON_VET_COUNTER value of 1 or 2 indicate an inconsistency in non-veteran status among the three variables. For example, a NON_VET_COUNTER value of 1 indicates a record with only one non-veteran value and two veteran (or ambiguous) values. Similarly, a record with a NON_VET_COUNTER value of 2 contains two non-veteran values and one veteran (or ambiguous) value. An example of an outpatient record with a NON_VET_COUNTER value of 2 is as follows:

- MEANS value of “N” (Non-Veteran, a non-veteran value),
- ELIG value of “16” (Humanitarian, a non-veteran value), and
- POS value of “6” (Operation Desert Shield, a veteran value.)

Frequencies of the NON_VET_COUNTER Values
Hereafter in this report, “non-vet value” (or “non-veteran value”) refers to the value of the NON_VET_COUNTER variable.

Table 4.1 presents the frequency of non-veteran values using all three of the previously discussed variables in the inpatient and outpatient file records. For example, in the inpatient file, 99.00% of the

total records do not contain any non-veteran values while 0.82% of the records contain three non-veteran values. In the outpatient file, 97.76% of the total records do not contain any non-veteran values while 0.19% contain three non-veteran values.

Table 4.1: Non-Veteran Values Contained in the FY04 PTF Inpatient Discharge File and NPCD Outpatient File Records

	Inpatient Admissions		Outpatient Visits	
Non-Vet Values	N	%	N	%
0	565,698	99.0	48,848,169	97.76
1	144	0.03	34,000	0.07
2	865	0.15	988,297	1.98
3	4,682	0.82	95,802	0.19
Total Records	571,389		49,966,268	

5. Discrepancies between Variables

Inconsistencies between the Variables

Upon assigning a NON_VET_COUNTER value of 0, 1, 2, or 3 to all inpatient and outpatient records based upon the number of non-veteran values contained in each record, we noticed that a large number of records in the outpatient file had a NON_VET_COUNTER a value of 2. This indicates a discrepancy (on many records) between the three variables that identify non-veteran status. After further analyzing the patient records, we determined that the variable that most frequently differed from the other two variables with regard to the identification of non-veteran status was the Period of Service variable.

In order to better understand this discrepancy, we compared the values of multiple records that pertain to individual patients. As previously noted in Section 3 (Variables Used) of this technical report, the values of the PSX/POS variable frequently fluctuated between value 9 (Other or None) and other values. This suggests that the 9 value is sometimes used as a catch-all category when the Period of Service value to which an individual should be assigned is unclear.

Elimination of the Period of Service Variable from the Analyses

Due to inconsistencies between the Period of Service values and the values of the other two variables (MEANS and CP/ELIG) used to identify non-veteran patient records, we excluded the PSX and POS variables from our analyses. Tables 5.1 – 5.3 reflect this change. Note that the phrase "Revised Non-Veteran Values" in the table titles refers to the use of the MEANS and CP/ELIG (not POS/PSX) values to determine the non-veteran status of medical records.

In addition, we edited our program to analyze male and female patient records both separately and collectively. Table 5.1 presents the frequencies of non-veteran values for both the inpatient and outpatient file records using only the MEANS and CP/ELIG variables. Table 5.2 stratifies this information by sex for the inpatient file while Table 5.3 stratifies the information by sex for the outpatient file.

Exclusion from Analyses of Outpatient Records with Unknown Sex

A very small number (8) of outpatient records have sex variable value of "U" (unknown). As these records represent only a very small percentage of the total number of outpatient records, we have excluded records with an unknown sex from our analysis.

Table 5.1: Frequency of Revised Non-Veteran Values in the FY04 PTF Inpatient Discharge File and FY04 NPCD Outpatient File for Females and Males Combined

	Inpatient Admissions		Outpatient Visits	
Non-Vet Values	N	%	N	%
0	565,822	99.03	48,866,675	97.80
1	53	0.01	17,800	0.04
2	5,514	0.97	1,081,793	2.17
Total Records	571,389		49,966,268	

Table 5.2: Frequency of Revised Non-Veteran Values in the FY04 PTF Inpatient Discharge File Stratified by Sex

	Female Admissions		Male Admissions	
Non-Vet Values	N	%	N	%
0	21,962	88.19	543,860	99.52
1	10	0.04	43	0.01
2	2,930	11.77	2,584	0.47
Total Records	24,902		546,487	

Table 5.3: Frequency of Revised Non-Veteran Values in the FY04 NPCD Outpatient File Stratified by Sex

	Female Outpatient Visits		Male Outpatient Visits	
Non-Vet Values	N	%	N	%
0	2,500,740	77.83	46,365,931	99.17
1	2,510	0.08	15,289	0.03
2	709,676	22.09	372,114	0.80
Total Records	3,212,926		46,753,334	

Inconsistencies between NON_VET_COUNTER Variable Values

Table 5.1 reveals that the outpatient file has a slightly larger percentage (0.04%) of records with a NON_VET_COUNTER value of 1 compared with the inpatient file (0.01%.) However, these very small percentages suggest that very little inconsistency exists between the MEANS and CP/ELIG variables with respect to their corresponding indicator variables (that define non-veteran status.) In other words, the MEANS and CP/ELIG variables are generally reliable indicators of non-veteran status in the outpatient and inpatient records.

It is also important to note that both the inpatient and outpatient files contain a relatively small number of non-veteran records (records with a NON_VET_COUNTER value of 2.) Such records comprise only 0.97% of the total records in the inpatient file and 2.17% of the total records in the outpatient file.

However, upon reviewing the number of non-veteran values stratified by sex, a key result is that a much larger percentage of non-veteran records exist for female patients than for male patients. In the inpatient file, 11.77% of the female patient records belong to non-veteran individuals while only 0.47% of the male patient records pertain to non-veteran individuals. With regard to the outpatient file, 22.09% of the female patient records belong to non-veteran individuals while only 0.80% of the male patient records pertain to non-veteran individuals.

Exploring Inconsistencies in the Non-Veteran Values of Multiple Records of Individual Patients

We examined multiple records belonging to individual patients to determine if the values of the MEANS or CP/ELIG variables fluctuated. Recall that non-veteran values are stored in the NON_VET_COUNTER variable, which counts (for each record) the number of times that a non-veteran status was reported by the MEANS or CP/ELIG indicator variables. In Table 5.4 – 5.6, we explored whether the NON_VET_COUNTER values fluctuate between different records for individual patients. Table 5.4 presents the frequency of patients who had a least one change in non-veteran values in the inpatient or outpatient files. Tables 5.5 and 5.6 stratify this information by sex for the inpatient and outpatient files, respectively.

Table 5.4: Consistency of Revised Non-Veteran Values in the FY04 PTF Inpatient Discharge and NPCD Outpatient Files for Individual Patients

	Inpatient Individuals		Outpatient Individuals	
	N	%	N	%
No Change	366,823	99.97	4,831,951	99.16
At Least One Change	125	0.03	40,886	0.84
Total Patients	366,948		4,872,837	

Table 5.5: Consistency of Revised Non-Veteran Values in the FY04 PTF Inpatient Discharge File Stratified by Sex for Individual Patients

	Females		Males	
	N	%	N	%
No Change	17,125	99.94	349,698	99.97
At Least One Change	10	0.06	115	0.03
Total Patients	17,135		349,813	

Table 5.6: Consistency of Revised Non-Veteran Values in the FY04 NPCD Outpatient File Stratified by Sex for Individual Patients

	Females		Males	
	N	%	N	%
No Change	392,663	98.40	4,439,288	99.23
At Least One Change	6,387	1.60	34,499	0.77
Total Patients	399,050		4,473,787	

Discussion

A change in non-veteran status among individual patient records occurred for less than 1% of inpatients and outpatients. This result reveals that the non-veteran status (as reported by the MEANS and CP/ELIG variables) among different records for individual patients was remarkably consistent. This indicates that the MEANS and CP/ELIG variables are generally reliable indicators of non-veteran status in both the outpatient and inpatient records.

6. Conclusion and Suggestions for Researchers

Significance of the Analysis

The significance of our analyses of the FY2004 inpatient and outpatient files appears to be that we have a workable approach for identifying non-veteran records. Should researchers wish to duplicate our approach, we recommend that they create a variable (e.g., NON_VET_COUNTER) that cumulates (for each record) the number of times that non-veteran status was reported by the MEANS and CP/ELIG indicator variables. Non-veteran records can then be identified as those that have a NON_VET_COUNTER value of 2. Researchers may also find it beneficial to eliminate records containing "ambiguous" values from their analyses of veteran populations as it is unclear whether such records pertain to veteran or non-veteran individuals.

Variables Used to Identify Non-Veteran Records

Although the Period of Service variable (PSX in the inpatient file and POS in the outpatient file) can be used to identify non-veteran records, it yields much less consistent results compared with the MEANS and CP/ELIG variables. If researchers who want to identify or analyze non-veteran records are concerned about the reliability of the Period of Service variable as an indicator of non-veteran status, they may want to consider either (1) eliminating the PSX/POS variable from their analysis or (2) eliminating all records that contain a PSX/POS value of 9.

Identifying Non-Veteran Female Records

Although the percentage of non-veteran records (having a NON_VET_COUNTER value of 2) in the FY2004 inpatient and outpatient files is 0.97% and 2.17% respectively, it may be beneficial for researchers to identify or eliminate these non-veteran records from their analyses of veteran populations. This may be especially true when researching female veteran populations as the percentage of non-veteran records for female patients is significantly larger than for male patients. A key finding of our analyses is that the percent of non-veteran female records in the inpatient and outpatient files is 11.77% and 22.09%, respectively.

Veteran Indicator Field

Researchers might be interested to know that (effective FY2006) a new field, Veteran Indicator, has been added to the medical SAS outpatient datasets. The SAS variable is VETIND, the label is Veteran Indicator, the format is $XXXYB, and the variable has possible values of YES, NO or space. Further information is available on the VA Intranet. As the Veteran Indicator field is not on the (2004) outpatient dataset used in our analysis, we have not compared our analysis of non-veteran records with an analysis of the VETIND variable in the medical SAS outpatient datasets.