POLYCYTHEMIA VERA: AN INVESTIGATION OF INFLUENTIAL FACTORS FOR DEATH RATES BETWEEN 1962 AND 2009

by

Fangfang Chen

BS, Wuhan University, China, 2007

Submitted to the Graduate Faculty of

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This thesis was presented

by

Fangfang Chen

It was defended on

Jul 31th, 2013

and approved by

Gary Marsh, PhD, Professor, Biostatistics, Graduate School of Public Health, University of Pittsburgh

Marnie Bertolet, PhD, Assistant Professor, Epidemiology, Graduate School of Public Health, University of Pittsburgh

Thesis Advisor: Jeanine Buchanich, PhD, Research Assistant Professor, Biostatistics, Graduate School of Public Health, University of Pittsburgh

Jeanine Buchanich, PhD

POLYCYTHEMIA VERA: AN INVESTIGATION OF INFLUENTIAL FACTORS FOR DEATH RATES BETWEEN 1962 AND 2009

Fangfang Chen, M.S.

University of Pittsburgh, 2013

ABSTRACT

Introduction: Polycythemia vera (PV) is a bone marrow disorder that causes overproduction of red blood cells. While some studies have documented increased local incidence rates of PV, no study has explored the mortality rates of PV over time. Public health significance of this study is firstly examining PV mortality rates in detail; a better understanding of these rates may help contextualize changes.

Methods: We used age-race-sex-and time period specific death counts from the Mortality and Population Data System (MPDS) from 1962 to 2009 for Pennsylvania (PA), California (CA), and the United States (US). Age-adjusted and age-specific mortality rates were examined for the whole US, as well as PA, California (CA), and the remaining US. We also accounted for changes in diagnostic criteria, International Classification of Diseases (ICD) mortality coding, and treatment. Negative binomial models were fit to account for age, sex, region, and critical time point changes in criteria, separated into seven sections to match events in PV medical history: 1962-1967, 1968-1974, 1975-1977, 1978-1986, 1987-1998, 1999-2000, and 2001-2009.

Results: We identified 18,743 PV death cases from 1962 to 2009; 95% of the deaths occurred in whites, and almost half of the deaths were in people 75 and older. Age-adjusted PV mortality of males decreased after 1967, while that of females decreased after 1978. The rate among males was always higher than that of females. Our time period analysis, corresponding to changes in

ICD coding, diagnostic criteria, and implementation of Hydroxyurea as standard treatment, found that each significantly influenced PV mortality.

Conclusion: This study is the first comprehensive examination of PV mortality rates over time. We determined that acceptance of the PVSG diagnostic criteria in 1975 decreased the magnitude of the difference in PV mortality between males and females seen prior to 1975, apparently due to the introduction of sex-specific red cell mass values. The transition from the 8th to 9th ICD revision in 1979, in which PV was changed from malignant to "benign and/or uncertain and unspecified behavior", decreased PV mortality significantly in both males and females. PV researchers should be aware of these changes when interpreting their data.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

1.1 BACKGROUND

Polycythemia vera (PV), a bone marrow disorder that causes overproduction of red blood cells as well as a slight increase in white blood cells and platelets, is one of the myeloproliferative disorders (MPD), which include chronic myeloid leukemia (CML), essential thrombocythemia, and primary myelofibrosis. [1] People with PV have a higher risk of thrombosis, which phlebotomy, chemotherapy (P^{32}) and myelosuppressive drugs like Hydroxyurea and aspirin can control. [2] PV can lead to acute myeloid leukemia, myelofibrosis, and thrombosis, which is the formation of a blood clot inside a blood vessel and is the leading causes of death among PV patients.

1.1.1 Mortality from PV

Mortality from PV is based on information from death certificates. Cause of death is coded using the International Classification of Diseases, or ICD, codes published by the World Health Organization (WHO) and used worldwide for morbidity and mortality statistics, reimbursement systems, and automated decision support in health care. Because the ICD codes underwent several revisions between 1962 and 2009, the ICD codes of PV have changed as follows: 6th and 7th revision, 294 polycythemia (1950-1967); 8th revision, 208 polycythemia vera (1968-1978); 9th revision, 238.4 polycythemia vera (1979-1998); and 10th revision, D45 polycythemia vera (1999-2013). PV was classified generally as polycythemia (code 294) in the 6th and 7th ICD

revisions. In the 8th revision, PV was differentiated from secondary polycythemia. This differentiation is still present in the current 10th revision. From 1950 to 2009, the classification of PV altered, from malignant neoplasm to unspecified neoplasm in the 8th to 9th revision.

1.1.2 Diagnosis of PV

PV was initially described by Dameshek in 1951 as one spectrum of myeloproliferative diseases. [3] According to Dameshek, these interrelated disorders were due to some undiscovered stimulus acting on the bone marrow. In 1967, the Polycythemia Vera Study Group developed diagnostic criteria to recruit PV patients into clinical trials. In 1975, after a series of clinical trials from the late 1960s to early 1970s, these clinical diagnostic criteria, based on the concentration of hemoglobin and hematocrit (volume percentage of red blood cells in blood), were accepted as the first ones for PV. As shown in Table 1, the major criteria included a raised red cell mass (RCM), oxygen saturations, and splenomegaly [4]. The patients were diagnosed if they fit all the major criteria, or had a high red cell mass and oxygen saturation with two of the minor criteria. The PVSG criteria were the gold standard for PV clinical diagnosis in the 1970s and 1980s. Before 1975, US hematologists diagnosed PV patients using only one major criterion, raised red cell mass, and two of four minor criteria: thrombocytosis, leukocytosis, oxygen saturation (without overt lung disease or renal tumor), and splenomegaly.[4,22] During that time, the polycythemia cut-off value for males was over 55% hematocrit or 18g/100 ml of hemoglobin. The cut-off value for females was over 52% hematocrit or 17g/100ml hemoglobin. In 1975, the PVSG criteria were thought to be stringent enough to exclude a false positive. The 1975 criteria ruled out the common etiology of secondary polycythemia and required evidence of myeloproliferative state.

Table 1 : 1975 PVSG diagnostic criteria

A1 Raised red cell mass, male ≥ 36 ml/kg, female ≥ 32 ml/kg
A2 Normal arterial oxygen saturation >92%
A3. Splenomegaly
B1. Thrombocytosis platelet count >400,000µl
B2. Leukocytosis >12,000µl
B3. Leukocyte alkaline phosphatase score >100
B4. Serum B_{12} (>900pg/ml) or unsaturated B_{12} binding capacity (>2200pg/ml)
Note: Diagnosis of PV if one of these combinations is present: A1+A2+A3;A1+A2+any two from B category

With the passage of time, more recently developed tests superseded the first diagnostic criteria. [5] In 1996, Pearson and Messinezy produced a new set of diagnostic criteria that encompassed burst forming unit erythroid (also known as erythroid progenitor cell, which gives rise to colonies that only contain erythrocytes) growth and other advancements to problematic diagnostic individuals [6] (Table 2). Another criteria change was that the red cell mass (RCM) was measured in relation to surface area rather than ml/kg; a significant diagnostic reading was 25% above the normal predicted value of red cell mass. In the 1975 PVSG criteria, the red cell mass value of obese people was underestimated due to being measured and then divided by weight. Also, these criteria differentiated secondary polycythemia from polycythemia vera, because secondary polycythemia is an acquired form of a rare disorder characterized by an abnormal increase in the number of mature red cells in the blood. Those people who smoke, are obese, or have high blood pressure are at greater risk for secondary polycythemia. [7] In this version of the criteria, the absence of secondary polycythemia was explicitly noted, along with presence of high red cell mass, palpable splenomegaly, abnormal marrow karyotype, or any two of the minor criteria.

Table 2: : 1996 polycythemia vera clinical diagnostic criteria

A1. Raised red cell mass (>25% above mean normal predicted value)
A2. Absence of cause of secondary polycythemia
A3. Palpable splenomegaly
A4. Clonality marker (abnormal marrow karyotype)
B1. Thrombocytosis platelet count >400,000 / μl
B2. Neutrophil leukocytosis count > 10,000 / μl
B3. Splenomegaly demonstrated on isotope or ultrasound scanning
B4. Characteristic BFU-E growth or reduced serum erythropoietin
Note: Diagnosis of PV if one of these combinations is present: A1+A2+A3 or A4; A1+A2 + two of B category

As shown in Table 3, in 2001, the World Health Organization (WHO) published new criteria for PV. These criteria showed significant changes in the measurement of RCM (adding new diagnostic criteria considering age, gender, and regional factors based on the population average value) and endogenous erythroid colony (EEC) formation. [6] The bone marrow produces endogenous erythroid colonies (EED) in the absence of exogenous erythropoietin (drugs have structure similar to erythropoietin that can control blood cell production). These criteria differed from the 1996 criteria in three main aspects: 1) the major criterion of red cell mass was defined less stringently; 2) the elevation of EEC formation went from a minor criterion to a major one; and 3) the WHO criteria included bone marrow biopsy as a minor criterion. People without secondary polycythemia but with a high total blood cell mass and any one of the other major criteria or any two of the minor criteria would receive a diagnosis of PV.

Table 3 : 2001 WHO diagnostic criteria

A1. Elevated total blood cell mass >25% above mean normal predicted value, or hemoglobin >18.5g/dl in men, 16.5 g/dl in women or >99th percentile of method-specific reference range for age, sex and altitude of residence
A2. No cause of secondary erythrocytosis
A3. Splenomegaly
A4. Clonal genetic abnormality other than Philadelphia chromosome or Bcr-Abl fusion gene in marrow cells
A5. Endogenous erythroid colony formation in vitro
B1. Thrombocytosis platelet count >400,000µl
B2. Leukocytosis >12,000µl
B3. Bone marrow biopsy showing panmyelosis with prominent erythroid and megakaryocytic proliferation
B4.Low serum erythropoietin levels
Note: Diagnosis of PV if one of these combinations is present: A1+A2 and any other category of A; A1+A2 and any two of category B

In 2005, multiple research studies identified a mutation in *JAK2* V617F on the *JAK2* gene in large numbers of MPDs patients. [8-10]. This mutation enhances kinase activity, which causes hyper-activation of erythropoietin, or overproduction of the red cells. [11] Scientists discovered that somatic *JAK2* mutations can be observed in 65-97% of PV patients, in 23–57% of ET patients, as well as in 35–57% of patients with other myeloid neoplasms [12]. Although this genetic mutation cannot differentiate MPDs from one another, it can be used to rule out the possibility of reactive thrombocytosis or myelofibrosis. [13] When the typical blood profile is obscured, the *JAK2* mutation becomes the only way to establish the diagnosis. [14] In a genetic analysis of PV patients, no association was observed between mutational status and current age, diagnostic age, gender, race, or self-reported family history; hence, it was suggested that the mutation is acquired. [6]

In 2008, WHO added this genetic mutation to the PV diagnostic criteria, so that molecular tools could more effectively help pinpoint the disease (Table 4). The most significant change of this version of criteria was adding the *JAK2* mutation to the major criteria. People who met both major criteria would receive a diagnosis of PV. Those who met the first major criteria (high total blood cell mass) would need any two of the minor criteria to be diagnosed with PV.

Table 4 : 2008 WHO diagnostic criteria

A1. Elevated total blood cell mass >25% above mean normal predicted value, or hemoglobin >18.5g/dl in men, 16.5 g/dl in women or >99th percentile of method-specific reference range for age, sex and altitude of residence
A2. Presence of JAK2V617F or similar mutation
B1. Bone marrow trilineage myeloproliferation
B2. Subnormal serum Epo level
B3. EEC growth
Diagnosis of PV requires meeting either both A criteria and one B criterion or the first A and 2 B criteria.

1.1.3 Treatment of PV

Although there is no cure for PV, hematologists have found ways to extend the survival time of PV patients. Because thrombosis is the main cause of PV death, hematologists targeted treatment of thrombocytosis. Eighty percent of the clinicians used phlebotomy alone, phlebotomy combined with Hydroxyurea, or Hydroxyurea alone to treat PV patients. Aspirin, Interferon, ^{32}P, and Busulfan were chosen only in special situations in clinical practice. . Hydroxyurea plays a significant role among the drugs to treat PV because it has fewer side-effects compared to other drugs. Since 1987, Hydroxyurea has been widely used to treat high thrombosis risk patients because it has less risk of causing acute leukemia than other drugs, like Busulfan.[11]

1.1.4 Epidemiology of PV

PV is rarely found in children or adults younger than thirty years of age. Only 5% of patients diagnosed with PV are younger than forty-years-old. The incidence of PV increases with age; the average diagnostic age is seventy-years-old. Reports from Europe indicate that the annual incidence rates for PV are variable, ranging from 0.02 to 2.8 per 100,000 inhabitants. [15] Most reports demonstrated that the incidence rates do not differ between genders, although some studies report that males have a higher occurrence. [5, 12] The highest reported annual incidence was 23.5/100,000 among 70-79-year-old males. [16] The prevalence of PV was 30 cases/100,000 inhabitants in the U.S [17].

The median survival time for untreated PV patients is 18 months, compared to 15 years for those treated.[2] It is recommended that all patients undergo phlebotomy, with the goal of keeping the hematocrit value below 0.45; stable patients who are at low risk for thrombosis do not require additional therapy. For those at high risk for thrombosis, the choice of a myelosuppressive agent depends on the patient's age or history of thrombosis. However, these myelosuppressive drugs have potential side effects, including leukemia. [2]

In 2004, four cases of PV found in people living on the same rural road near Tamaqua, in Northeastern Pennsylvania (PA), triggered the Pennsylvania Department of Health (PADOH) to review cancer cases reported to the Pennsylvania Cancer Registry (PCR) from the three counties (Carbon, Luzerne, and Schuylkill) surrounding the Tamaqua areas. The PADOH found that the overall cancer rate in the areas was similar to that in other parts of the state, but there were more PV cases than expected.[18] The PADOH invited the Agency for Toxic Substances and Disease Registry (ATSDR) to perform an extensive case-finding effort in the tri-county areas. The ATSDR used PCR records, outreach to hematologists, and case self-identification to determine cases diagnosed from 2001 to 2005 in the tri-county areas. The ATSDR identified a statistically significant ($p < 0.001$) cluster of 15 PV cases in the Tamaqua community. [18]The investigators also found that almost half of the confirmed PV cases had not been reported to the PCR, while a third of the participants who reported to the registry with a hybrid clinical diagnosis of PV did not meet PV diagnostic criteria before 2004. [18] Most of the population-based PV studies have focused on PV incidence. In 1994, a retrospective study based on the local clinical medical records found that the sex- and age-adjusted incidence of PV in Olmsted County, Minnesota, was

1.9 per 100,000 population per year from 1935 to 1989; during this period, this rate hovered around the average incidence. The age and sex adjusted incidence was 2.0, 1.1, 2.8, and 1.9 per 100,000 population per year in 1950-1959, 1960-1969, 1970-1979, and 1980-1989, respectively. [16]During this study, potential cases of PV were observed than 55 cases were diagnosed by 1975 diagnostic criteria. Additionally, in Malmo, Sweden, this incidence rate was 1.0 to 2.6 per 100,000 population per year with an increasing pattern from 1950 to 1984.This retrospective study also based on the local hospital records, and they were identified 177 PV patients by PVSG criteria. No examination of PV mortality in the US can be found.

While studies, such as the one focusing on PV in Northeastern PA, have documented increased local incidence rates of PV, no study has explored the mortality rates of PV over time. This study will be the first attempt to examine PV mortality rates in detail; a better understanding of these rates will help contextualize the changes of PV-associated rates.

2.0 METHODS

2.1.1 Objectives

As no epidemiological study has focused on examining PV morality changes on a large scale, this research will be the first attempt to describe and analyze PV mortality trends in different gender-race groups from 1962 to 2009 in Pennsylvania and in the United States. The recent findings of increased PV incidence in Northeastern PA led us to investigate whether PV mortality rates in PA are significantly different from those in other parts of the US. In the medical history of PV from 1962 to 2009,.diagnostic criteria has changed two times. At first, there were only clinical measurements; after several PVSG clinical trials, PVSG criteria in 1975 was accepted for diagnosing PV more specifically; as cellular and genetic methods developed, WHO published in 2001 new PV criteria. These changes may have affected mortality rates by improving case ascertainment. As the clinical trials progressed, researchers found that Hydroxyurea reduced the risk of death from PV-related thrombosis with few side effects. Wide usage of Hydroxyurea may have affected PV mortality rates by increasing survival time for PV patients. From 1962 to 2009, ICD codes have changed three times (in 1968, 1978, and 1999 respectively); these changes may cause discontinuity in the cause-specific mortality trends. This research will analyze the influences of the three above factors by modeling and comparing the results with graphical study.

2.1.2 Research design

In this study, data were evaluated for the Pennsylvania, California, and the United States. Our target population for analysis was PA mortality. We first examined descriptive mortality graphs

to show how PV mortality rate altered from 1962 to 2009. Then, a generalized linear model was fit to find what influenced the rate of PV mortality. During the model building, we used the US population to compare with the PA population to determine if PA had a different mortality pattern from the US. Also, we compared PA and CA because California has no known incidence of increased risk of PV, and California also has a larger population than PA, which could produce more stable estimates of mortality. With these comparisons, we want to detect if PA had different mortality patterns from other areas or if the high incidence pattern only happened in that cluster in PA.

In the model, we had six predictors: age, region, sex, ICD code revision (four segments), diagnostic criteria changes (three segments), and treatment of Hydroxyurea (two segments). During the model building process, we ran univariate models to first show the direction of the predictors' influence; we then modeled the ICD code revision changes, diagnostic criteria changes, and usage of treatment combined with other predictors. After that, we built region-specific models to test the effects of these changes to the three regions. Finally, we built the model with predictors' region, age, sex, and a categorical time variable with seven segments to include all the changes in time in order to reduce the influence of the overlapping time periods if we had included all three time variables in one model.

2.1.3 Data source

The Mortality and Population Data System (MPDS) was used to generate the cancer mortality rates for the priori cause of interest mortality from PV International Classification of Disease (ICD) 7th revision codes 294; (ICD) 8^{th} revision 208; (ICD) 9^{th} revision 238.4; and (ICD) 10^{th} revision D45. MPDS, a data repository and retrieval system for detailed mortality data provided by the National Center for Health Statistics and the US Census Bureau, has been maintained in the Biostatistics Department of the University of Pittsburgh since the 1980s. The MPDS contains the underlying cause of death codes (ICD four-digit codes) for all persons who died in the United States from 1950 to 2009 (limited to deaths from malignant neoplasms during the 1950-1961

period). Individual death records include codes for gender, race, age of death, year of death, and geographic location (at time of death).

In this thesis, the exact count data of death and population were used to calculate the mortality rates and build the model. Table 5 shows the specific coded causes of death for PV.

PV was classified generally as polycythemia (code 294) in the 6th and 7th ICD revisions. In the 8th revision, PV was differentiated from secondary polycythemia. This differentiation is still present in the current 10th revision. From 1950 to 2009, the classification of PV altered, from malignant neoplasm to unspecified neoplasm in the 8th to 9th revision.

Table 5: ICD codes used in mortality analysis of PV

6th & 7th revision(1950-1967)	8th revision (1968-1978)	9th revision (1979-1998)	10th revision (1999-)
294 polycythemia	208 polycythemia vera	238.4 polycythemia vera	D45 polycythemia vera

2.1.4 Summary of age covariates

Initially, we combined the age group under 45 because we knew that 95% of PV patients are older than 40. After calculating the age-specific mortality of different age strata with the date of white males and females in the three populations, we found that ages 45-54, 55-64, and 65-74 had a similar mortality change trend, while the 75-84 and the 85+ age groups showed a different trend compared with the younger age groups . The PA group of both white males and females 75 and older varied around average. PV mortality of CA white males 75 and older varied and declined, and that of corresponding females showed an inverse bathtub pattern. Although the mortality patterns differed according to regions and genders in the over 75 group, the 45-74 subgroups in these six region-gender groups showed similar declining patterns of PV mortality (Figure 1). For the less than 45 age group the mortality is close to

zero; it appeared as a stable line. Therefore, we summarized the data according to the following three age groups: less than 45, 45 to 74, over 75. After summarizing the data, we found that the mortality trends in the three age categories differed. The mortality rate in US white males ages 75 and older in areas of the US outside PA and CA increased from 1962 to 1967, and then decreased from 1968 to 2009. The mortality rate in US white females ages 75 and older in areas of the US outside PA and CA increased from 1962 to 1977, and then decreased from 1978 to 2009. varied (Age-specific mortality tables also appear in the Appendix.)

When calculating the mortality rates, we separated the time period into seven sections to match the events in PV medical history: 1962-1967, 1968-1974, 1975-1977, 1978-1986, 1987-1998, 1999-2000, and 2001-2009.

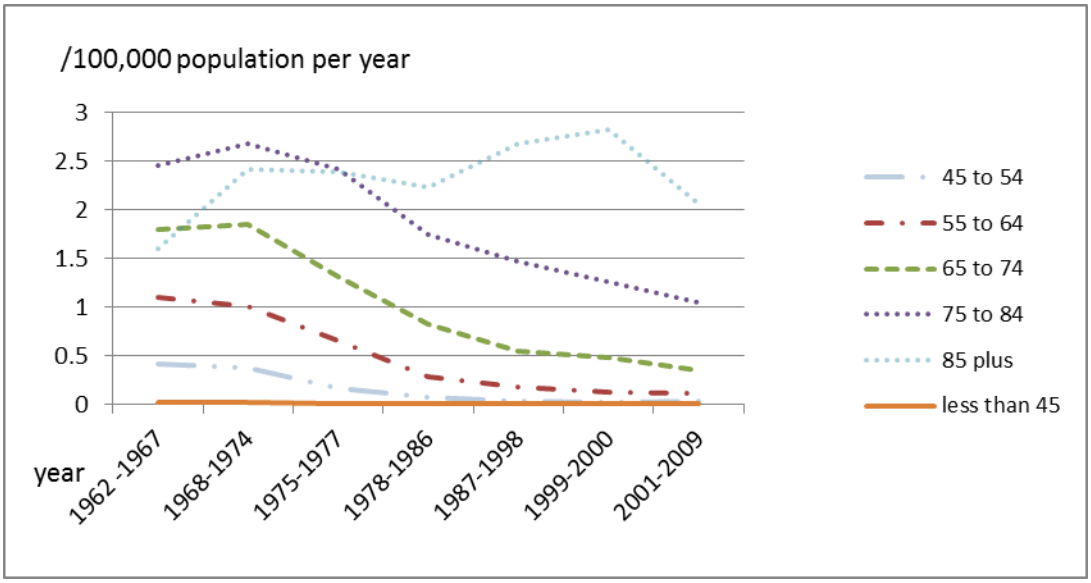

Figure 1: Age-specific mortality for ungrouped data of Other US white males

In white males in other areas of the US, whose age in 65 to 74, 55 to 64, 45 to 54 groups, PV mortality decreased after 1967. In the 75 to 84 age group, PV mortality increased from 1962 to 1967 then decreased. In the over 85 age group, PV mortality increased from 1962 to 2001 then decreased. (Figure 1)

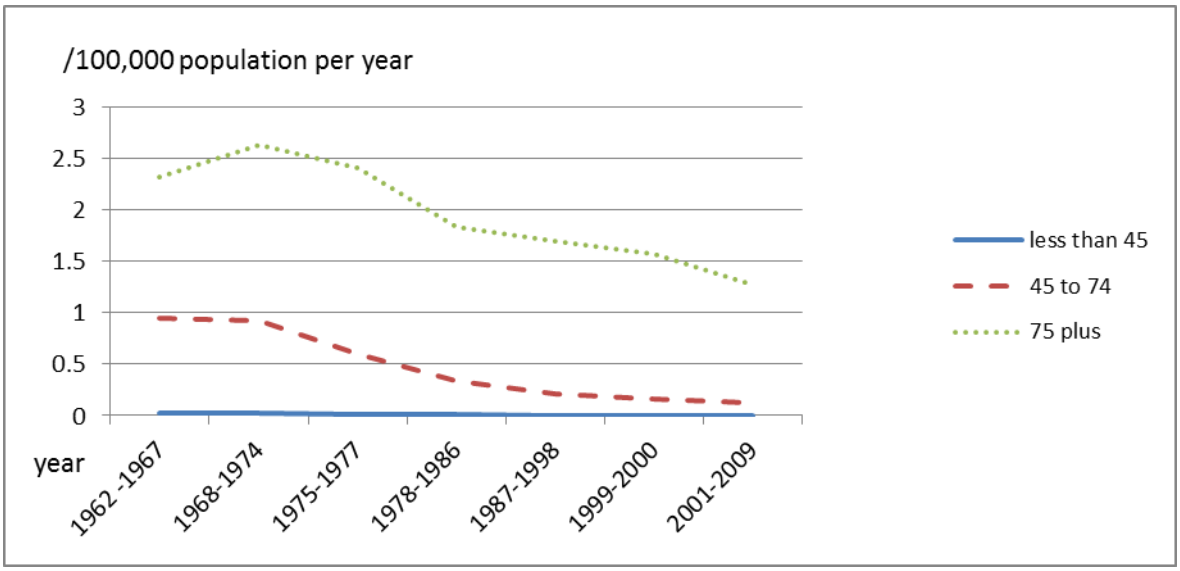

Figure 2: Age-specific mortality for grouped data of Other US white males

PV mortality of US males 75 and older increased from 1962 to 1974, and then decreased; PV mortality of US males 45-74 continued to decline. (Figure 2)

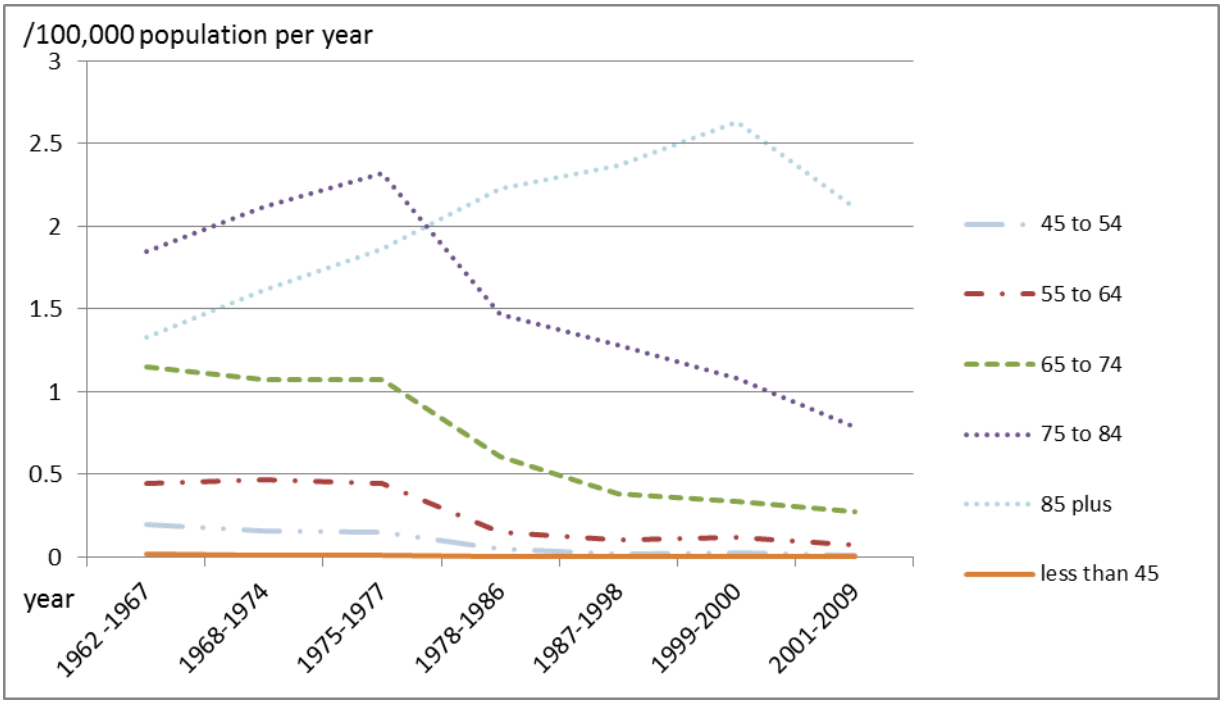

Figure 3: Age-specific mortality for ungrouped data of Other US white females

In white females in other areas of the US, whose age in 65 to 74, 55 to 64, 45 to 54 groups, PV mortality decreased after 1975. In the 75 to 84 age group, PV mortality increased from 1962 to 1975 then decreased. In the over 85 age group, PV mortality increased from 1962 to 2001 then decreased. (Figure 3)

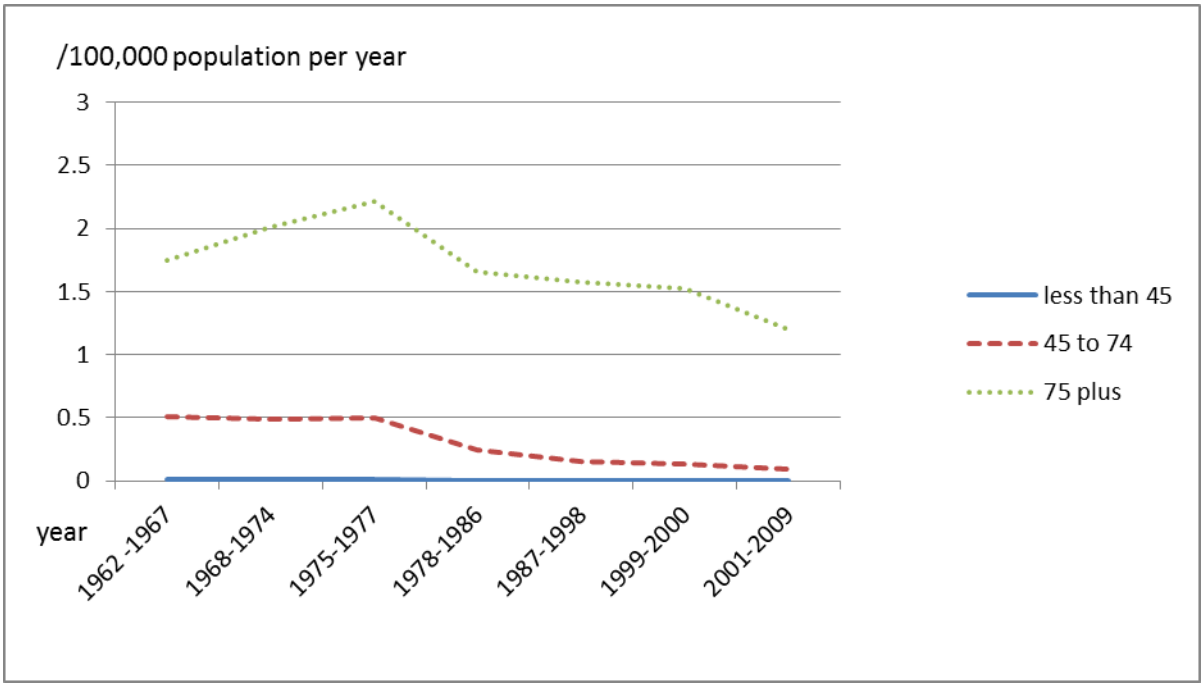

Figure 4: Age-specific mortality for grouped data of Other US white females

PV mortality of US females 75 and older increased from 1962 to 1977, and then decreased; again, PV mortality of US females 45-74 kept decreasing. (Figure 4)

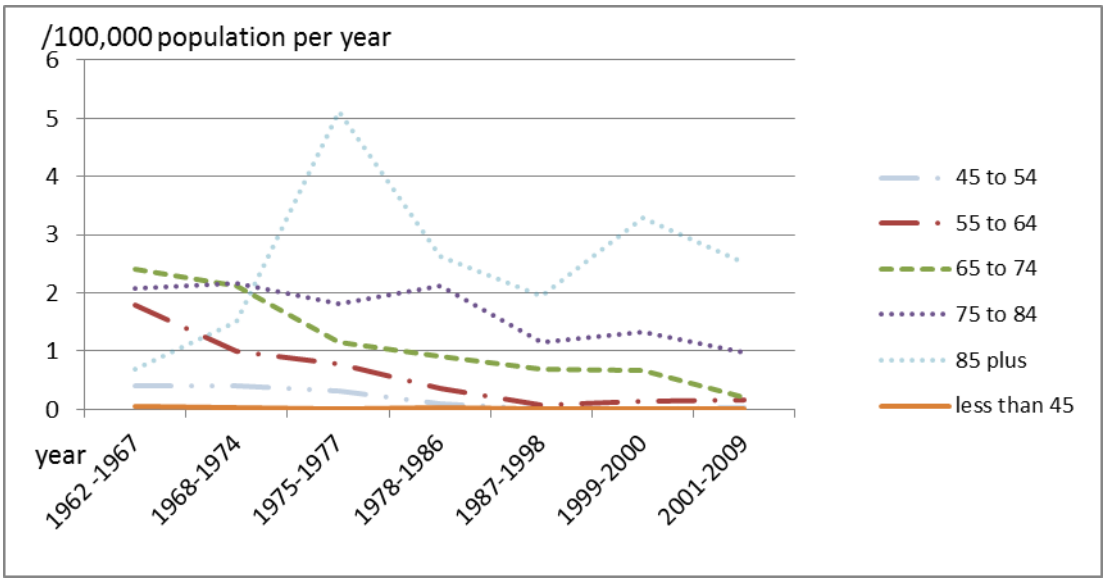

Figure 5: Age-specific mortality for ungrouped data of PA white males

In PA white males, whose age in 75 to 84, 65 to 74, 55 to 64, 45 to 54 groups, PV mortality decreased after 1968. In the 75 to 84 age group, PV mortality varied and decreased. In the over 85 age group, PV mortality increased from 1962 to 197, then varied and decreased. (Figure 5)

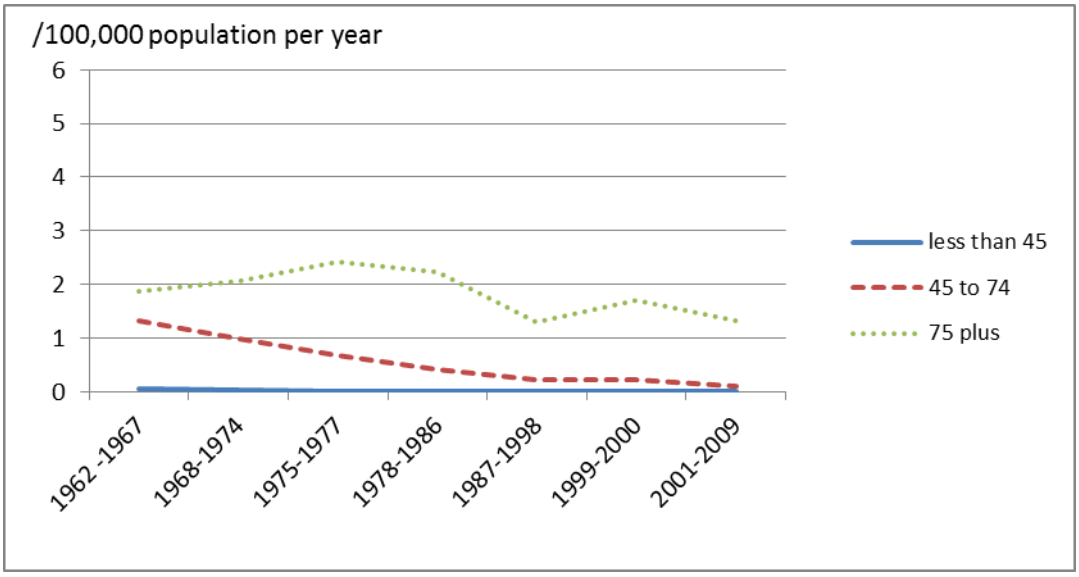

Figure 6 : Age-specific mortality for grouped data of PA white males

PV mortality of PA males 75 and older did not have an obvious trend; PV mortality of PA males 45-74 continued to decline.(Figure 6)

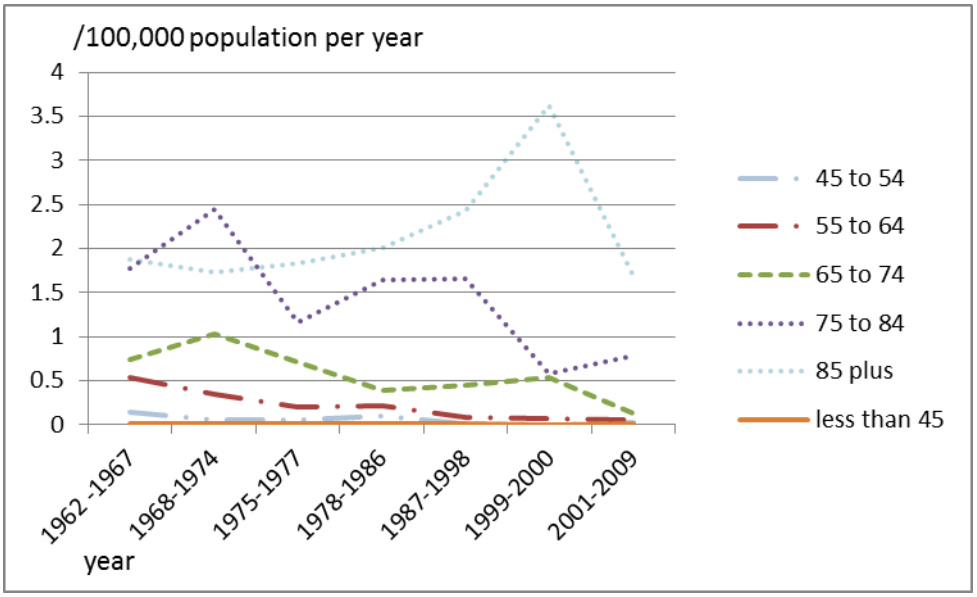

Figure 7: Age-specific mortality for ungrouped data of PA white females

In PA white females, whose age in 55 to 64, 45 to 54 groups, PV mortality decreased after 1962. In the 65 to 74 and 75 to 84 age group, PV mortality increased from 1962 to 1968 then decreased and varied. In the over 85 age group, PV mortality increased from 1962 to 1999 then decreased. (Figure 7)

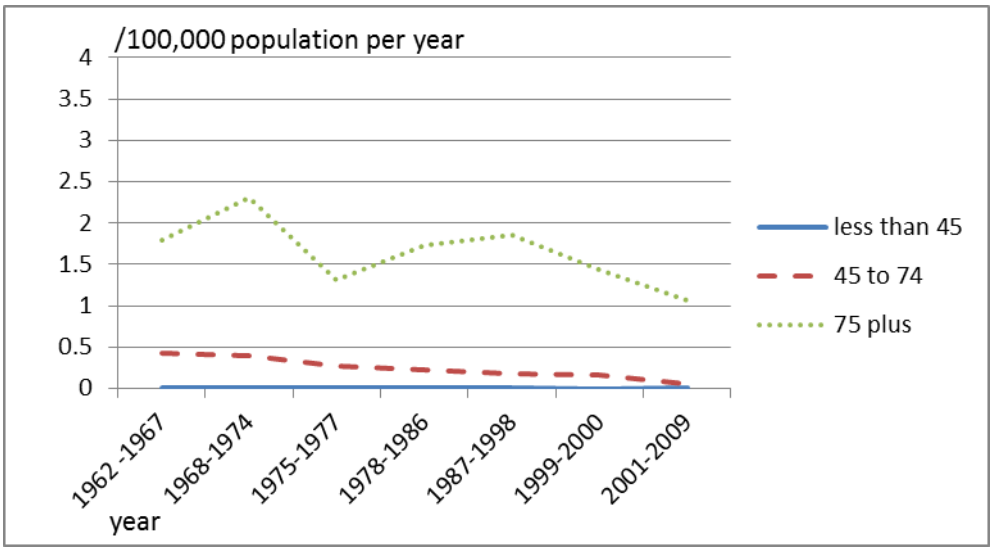

Figure 8: Age-specific mortality for grouped data of PA white females

PV mortality of PA females 75 and older did not have an obvious trend; PV mortality of PA females 45-74 continued to decline.(Figure 8)

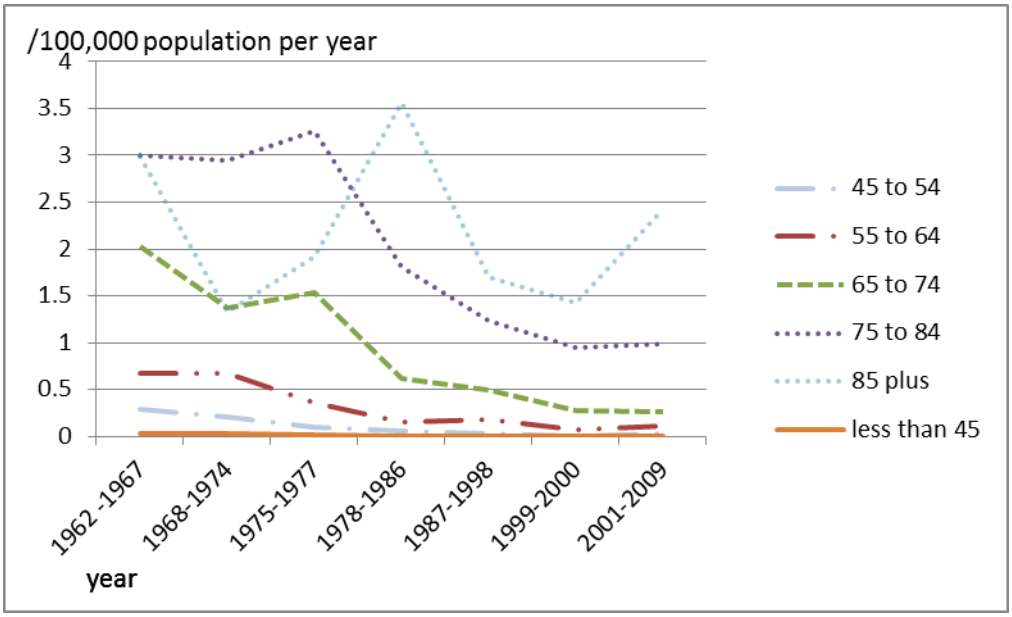

Figure 9: Age-specific mortality for ungrouped data of CA white males

In CA white males, whose age in 65 to74, 55 to 64, 45 to 54 groups, PV mortality decreased after 1962. In the 65 to 75 to 84 age group, PV mortality increased from 1962 to 1975 then decreased. In the over 85 age group, PV mortality varied from 1962 to 2009. (Figure 9)

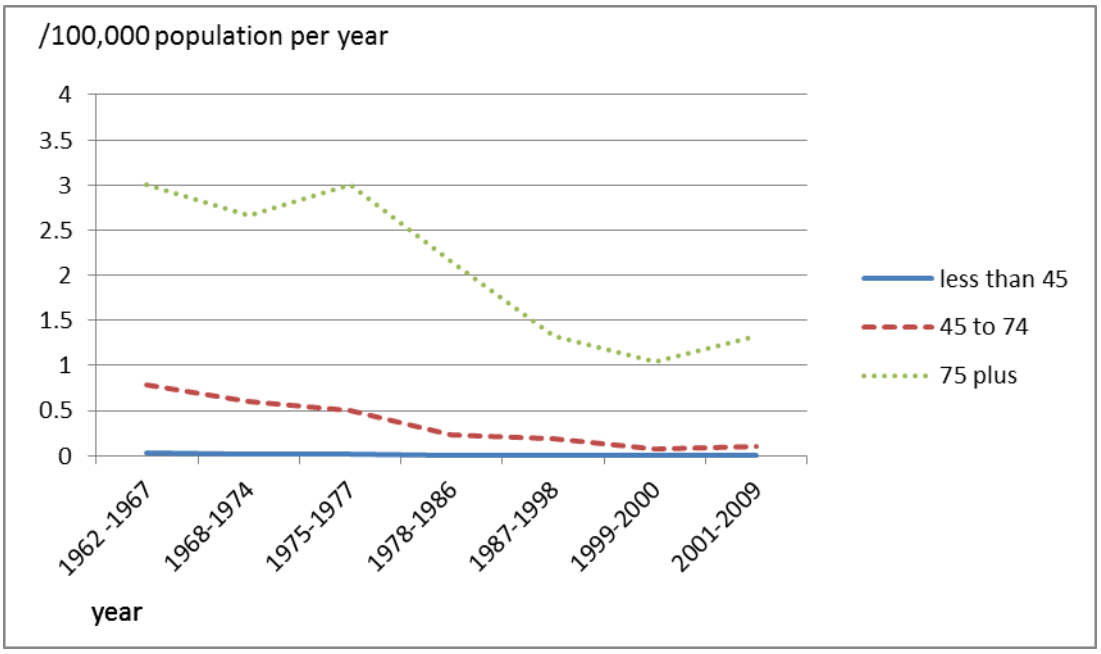

Figure 10: Age-specific mortality for grouped data of CA white males

PV mortality of CA males 75 and older decreased after 1975; PV mortality of CA males 45-74 continued to decline.(Figure 10)

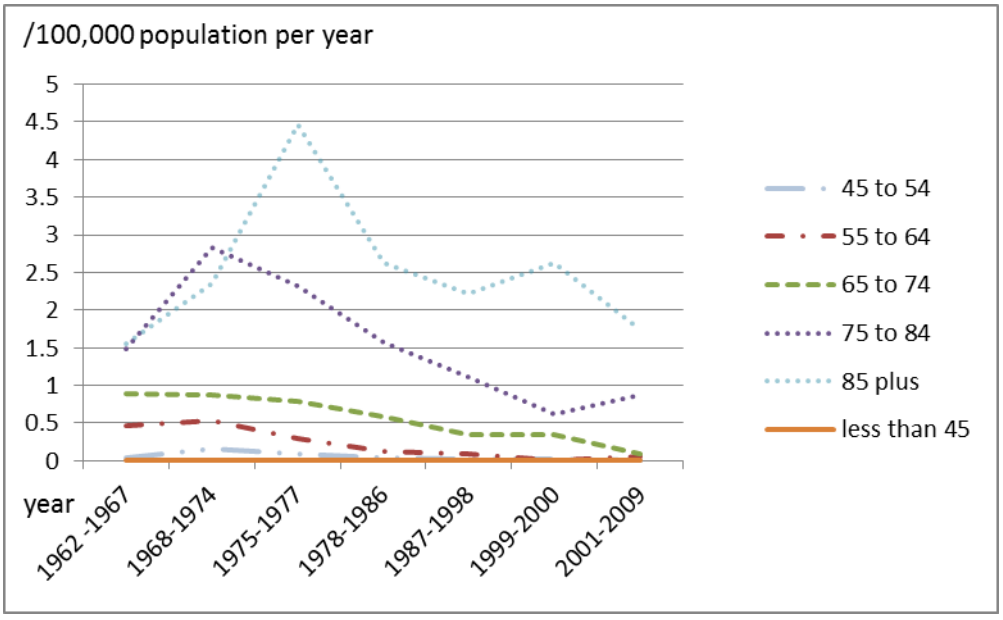

Figure 11: Age-specific mortality for ungrouped data of CA white females

In CA white females, whose age in 65 to 74, 55 to 64, 45 to 54 groups, PV mortality decreased after 1962. In the 75 to 84 age group, PV mortality increased from 1962 to 1968 then decreased and varied. In the over 85 age group, PV mortality increased from 1962 to 19 then decreased and varied. (Figure 11)

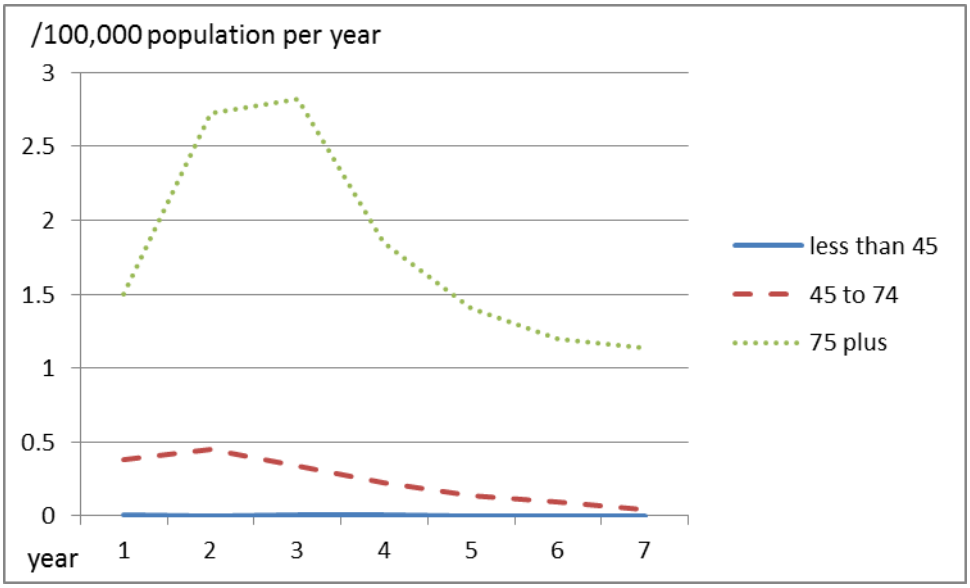

Figure 12: Age-specific mortality for grouped data of CA white females

PV mortality of CA females 75 and older increased from 1962 to 1977, and then decreased; again, PV mortality of CA females 45-74 increased from 1962 to 1974, and then decreased. (Figure 12)

2.1.5 Death count summary

Death counts were summarized in different categorical variables (Table 6). In the 18,743 PV death cases from 1962 to 2009, 95% of the deaths occurred in white males and females, 97% of the deaths affected people 45 and older, and almost half of the deaths were concentrated in people 75 and older. Males and females had a very similar death count. The counts of death were evenly distributed in the time period divided by the ICD revision. Two-thirds of the deaths were distributed before the wide usage of Hydroxyurea (1987), and almost 90% of deaths were

distributed before the publication of the 2001 WHO criteria. The three regions had similar distributions of death counts in these covariates.

Table 6: Descriptive characteristics of covariates

		CA		PA		OTHER		TOTAL
		No.	%	No.	%	No.	%	
Death Total		1623		1172		15948		18743
Race	**W**	1539	*94.8*	1118	*95.4*	15102	*94.7*	17759
	NW	84	*5.2*	54	*4.6*	810	*5.3*	948
Gender	**Male**	787	*48.5*	609	*52*	7945	*49.8*	9341
	Female	836	*51.5*	563	*48*	8003	*50.2*	9402
Age group	**<45**	51	*3.1*	39	*3.3*	520	*3.3*	610
	45-74	706	*43.5*	597	*50.9*	7913	*49.6*	9216
	75+	866	*53.4*	536	*45.7*	7515	*47.1*	8917
ICD changes	**6th-7th(1962-1967)**	262	*16.1*	255	*21.8*	2934	*18.4*	3451
	8th(1968-1978)	485	*29.9*	323	*27.6*	4537	*28.4*	5345
	9th(1979-1998)	586	*36.1*	422	*36*	5641	*35.4*	6649
	10th (1999-)	290	*17.9*	172	*14.7*	2836	*17.8*	3298
Diagnostic changes	**BEFORE PVSG(1962-1974)**	608	*37.5*	489	*41.7*	6199	*38.9*	7296
	PVSG(1975-2000)	803	*49.5*	561	*47.9*	7775	*48.8*	9139
	WHO(2001-)	212	*13.1*	122	*10.4*	1974	*12.4*	2308
Treatment changes	**BEFORE HU(1962-1986)**	1015	*62.5*	759	*64.8*	9960	*62.5*	11734
	HU(1987-)	608	*37.5*	413	*35.2*	5988	*37.5*	7009

The mortality graphs of non-white people and those individual in the <45 age group are not shown as there are too many zero values in the data. The age-adjusted mortality for all race-gender-age groups is shown in Appendix I.

Age-adjusted mortality (0.35 /100,000 population per year) is much lower than that of unadjusted mortality (1~2/100,000 population per year). This situation is due to the concentration of deaths in a small proportion of people: less than 5% of people are 75+ in the population, but this group represents approximately half of the PV deaths overall (Table 7). The complete death count data summary is shown in Appendix II.

Table 7: Death distribution in age group

Region	Gender	Age	death count	Population percent in the age (%)
US	FEMALE	<45	173	65.3
		45-74	3300	28.5
		>75	4530	6.2
	MALE	<45	347	69.4
		45-74	4613	26.8
		>75	2985	3.8
PA	FEMALE	<45	13	61.5
		45-74	211	31.2
		>75	339	7.3
	MALE	<45	26	66.2
		45-74	386	29.5
		>75	197	4.4
CA	FEMALE	<45	13	67.8
		45-74	301	26.7
		>75	522	5.5
	MALE	<45	38	71.9
		45-74	405	24.7
		>75	344	3.4

Because we observed too many zero values in the non-white population and detected 95% of the deaths among whites, we decided to only use white population data to build our model. From the background information, we knew that 95% of the PV patients are over 40-years-old; therefore, we combined all age groups under 45. Even when combined, these age groups demonstrated a very small proportion of the overall death count. (Table 7)

2.1.6 Distribution of death count

After we summarized the death count, we had 126 observations in the dataset. (We summarized the death count by gender in 2 category, age in 3 categories, region in 3 categories, 3 time variables in 7 categories.) Because the test for the normal distribution of the count data rejects the null ($p<0.001$) (Fig 13), and the dependent variable is count data, but not dichotomized classified data, we rejected the possibility of using a linear regression or logistic regression to model the data. Usually, the Poisson model would be used to model count data, but in our dataset, the mean of the grouped death count is 148.75, while the variance is 70,901. The data are extremely over-dispersed, thereby violating the null hypothesis of the Poisson distribution. These data had five zero values in the 126 observations; thus, zero-truncated and zero-inflated problems did not occur.

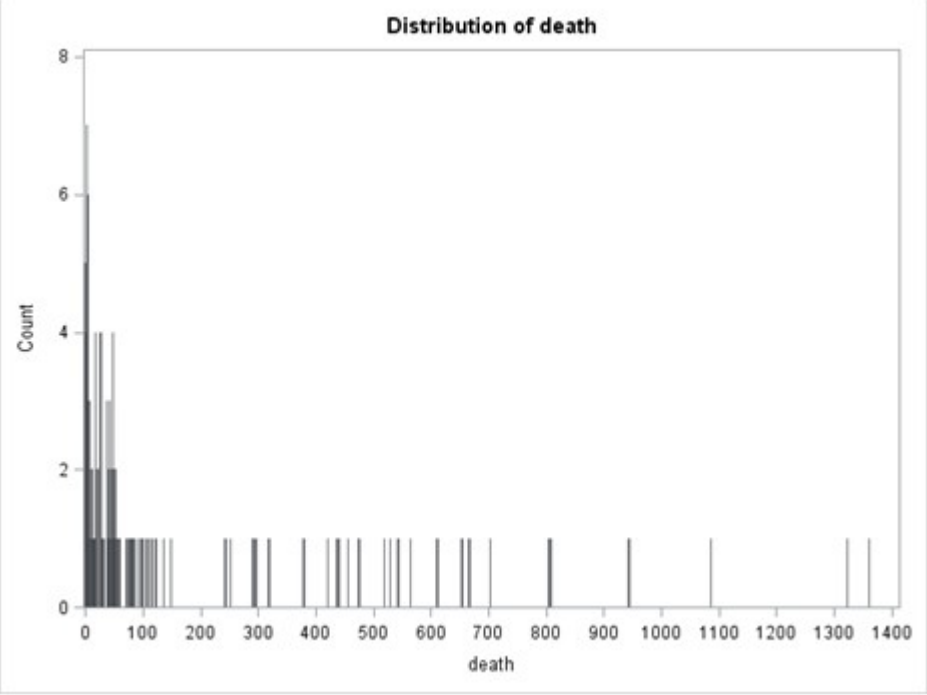

Figure 13: Death count distribution of PV

2.1.7 Descriptive analyses

The descriptive statistics analysis was primarily used to investigate age-adjusted mortality in different age-gender-race groups by year. The whole population was separated into 12 age groups: less than 5, 5-9, 10-14, 15-19, 20-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-84, and 85 and older. Four race-gender groups were used: white males, white females, non-white males, and non-white females.

The graphs are drawn in two parts: the first one is for the whole US, while the second and third ones show PA, CA, and the US minus PA and CA by gender. The graphs separately show age groups less than 45-years-old and over 45-years-old as most PV deaths were observed among those over 40-years-old.

The age-adjusted mortality was calculated by the formula below:

$$R_i = (n / N) * w,$$

$$R = \Sigma R_i \ (i = 12),$$

In the formula, R_i is the weighted mortality in the i^{th} age group, R is the age-adjusted mortality, n is the exact count of death of PV in this age-gender-race subgroup in that year, N is the number of population in this age-gender-race subgroup in that year, and w is the weight function that was calculated using the US age-gender-race specific population in the matching year as the standard population. In the graph of US white males and females, mortality is adjusted using the 1980 population.

The descriptive statistics also include a summary of the exact count of PV deaths by race-gender groups in the PA, US, and CA population in each year. In this part, age groups are separated into three groups: less than 45, 45-74, and 75 and older. This step helps identify the zero inflated issue for the Poisson model building in the next step.

2.1.8 Data analysis

Poisson regression is usually used in mortality studies. This technique is a branch of the generalized linear model, in which the outcome (number of deaths) is count data following the Poisson distribution. Epidemiological studies since the 1980s, especially cancer mortality studies, have used this method. [19] The assumption is when the number of deaths (d) is small relative to the total cohort size (n), the Poisson approximation to the exact sampling distribution of $\lambda = d/n$ will be sufficient. That means when d is small, withdrawing the deaths from future person-years is acceptable and the number of deaths in different cells can be considered to be statistically independent.

The following model is used:

$$\text{Log } E(d) = \log(n) + xb$$

Whereas d is the summarized death from a specified group, n is the person-year at risk, and b represents a vector of regression coefficients describing the effects of primary interest.[18]

However, from Figure 13, we see that a significant number of small count values existed; at the same time, extreme large values were distributed sparsely. Among these small counts, there were only five zero value counts; as a result, the issue of zero-inflated or zero-truncated problem did not exist, but overdispersion did exist. To fix the overdispersion problem, we decided to determine whether the use of the overdispersion Poisson regression or negative binomial regression would be more appropriate to model these data.

An assumption of over-dispersed Poisson regression is that the variance is proportional to the mean:

$\text{Var}(Y) = \vartheta E(Y) = \vartheta\mu$ ($\vartheta > 1$, we have overdispersion).

An assumption of negative binomial regression is that:

$\text{Var}(Y) = \mu(1 + \alpha\mu)$. (If $\alpha > 0$, overdispersion exists.)

To determine whether the over-dispersed Poisson model or negative binomial model would be better, we fit three multivariate models, including age, gender, and region, and one of the time-period categorical variables. After that, we compared the AIC and BIC values of the

two series of models. After comparing the results, we chose to use negative binomial regression as it had smaller AIC and BIC values in this situation. (Table 8)

Table 8: Comparison of AIC/BIC value of overdispersed Poisson model and negative binomial model

	Negative binomial		**Over-dispersed Poisson**	
	AIC	BIC	AIC	BIC
R+A+S+T1	1046.21	1074.6	2465.52	2491.07
R+A+S+T2	1098.93	1124.46	3044.58	3067.27
R+A+S+T3	1084.23	1106.92	3659.75	3679.6

R =REGION A= AGE S = SEX T1 = ICD CODES T2 = Diagnostic criteria change T3 = Treatment

Since our research goal is to determine whether PV deaths in PA differed from PV deaths in other places in the US, we have a predictor for region and several potential confounders (age, gender, times of ICD codes changes, times of diagnostic criteria changes e, time when Hydroxyurea became widely used). Table 9 lists how these independent variables were coded. In the model building, we first used univariate testing to see the direction of the influences of these predictors. Then, we built multivariate models to see how these predictors affect the overall mortality. In this step, we included interaction terms to detect the different influences of these time categorical variables given different regions or genders. Finally, we built region-specific models to see the different appearance if the influences of the other areas were removed. In the region-specific models, only the US, PA, or CA data were used (42 observations in each model). These region-specific models were also built separately by three time categorical variables with age and gender in the model.

Table 9: Variable codes for analysis

variable	**different level**			
sex	female = 0	male = 1		
age	<45 = 0	45-74 = 1	>75 = 2	
region	other = 0	PA = 1	CA = 2	
ICD	6&7th = 0	8th = 1	9th = 2	10th = 3
diagnostic criteria	before PVSG = 0	PVSG = 1	WHO = 2	
treatment	before HU = 0	HU = 1		

3.0 RESULT

3.1 GEOGRAPHICAL STUDIES

PV mortality in the US (in this section, 'US' means the whole US) decreased from 1962 to 2009. As the graphs indicate, several different time segments occurred. The graphs also illustrate when Hydroxyurea was widely used. These graphs indicate that PV mortality of white males was always higher than that of white females. However, the difference of the rates between genders dropped quickly after 1975, the time when the PVSG criteria were accepted. Also, the rates of PV mortality in white males and females dropped sharply after 1978, the time when the ICD revision changed from the 8th to the 9th (Figure 14). Age-adjusted mortality was adjusted to the 1980 US population here.

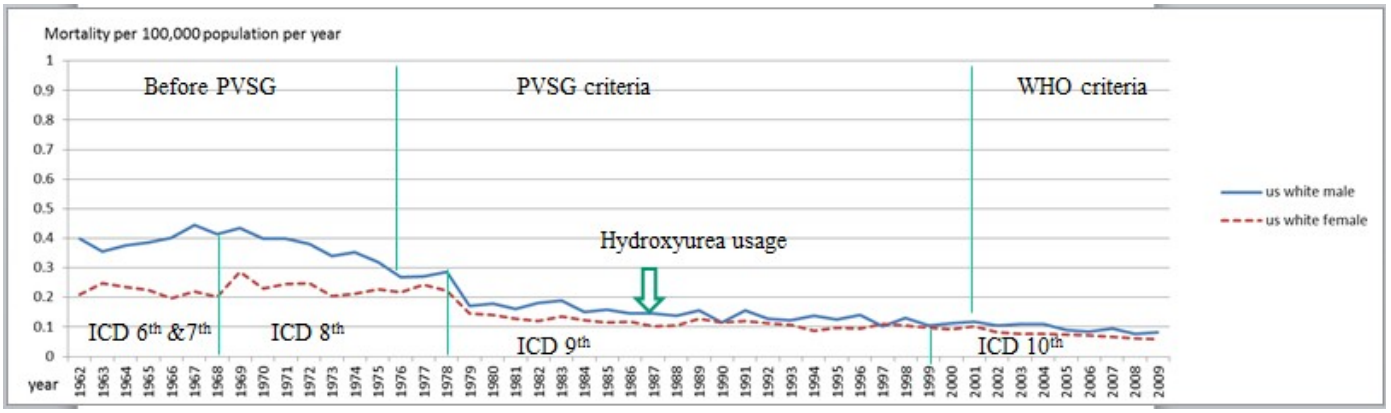

Figure 14: Comparison of Age-adjusted mortality of US white male and female

A comparison of PV mortality of white males in three different regions (PA, CA, other areas of the US) shows that mortality in PA is higher than that of the other two areas before 1980. After 1986, the mortalities of the three regions were very similar. PV mortality of white males in the three regions decreased from 1962 to 2009 (Fig 15).

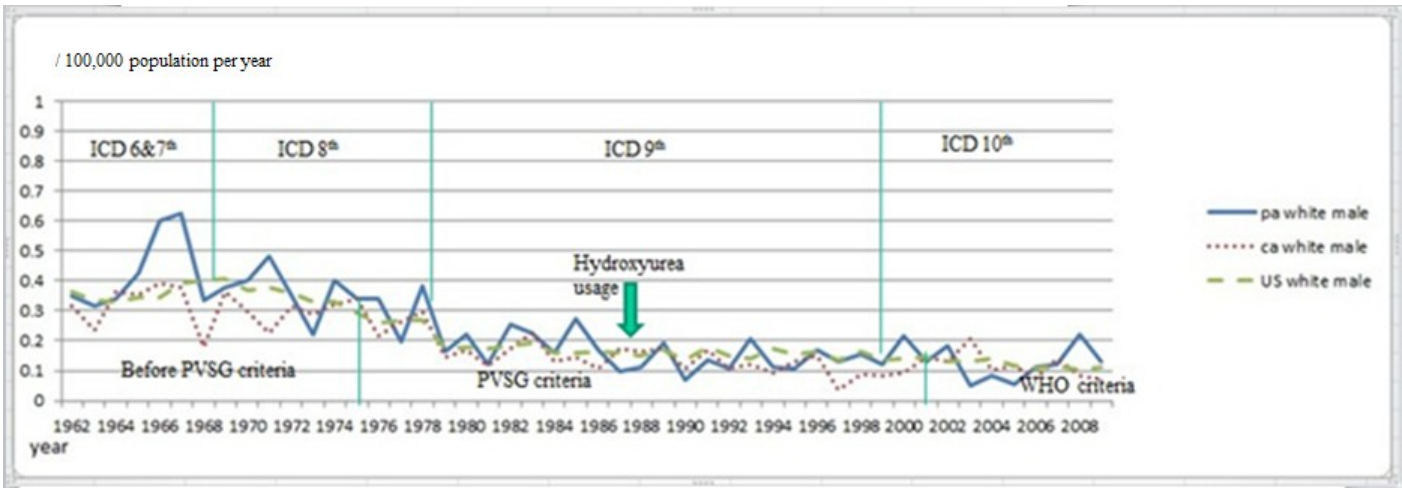

Figure 15: Comparison of Age-adjusted mortality of white males in PA, CA, and other areas in the US

A comparison of PV mortality of white females in three different regions (PA, CA, other areas of the US) shows that the mortality rates of the three groups were also similar. These rates of white females in the three regions slightly decreased from 1962-2009 (Fig 16).

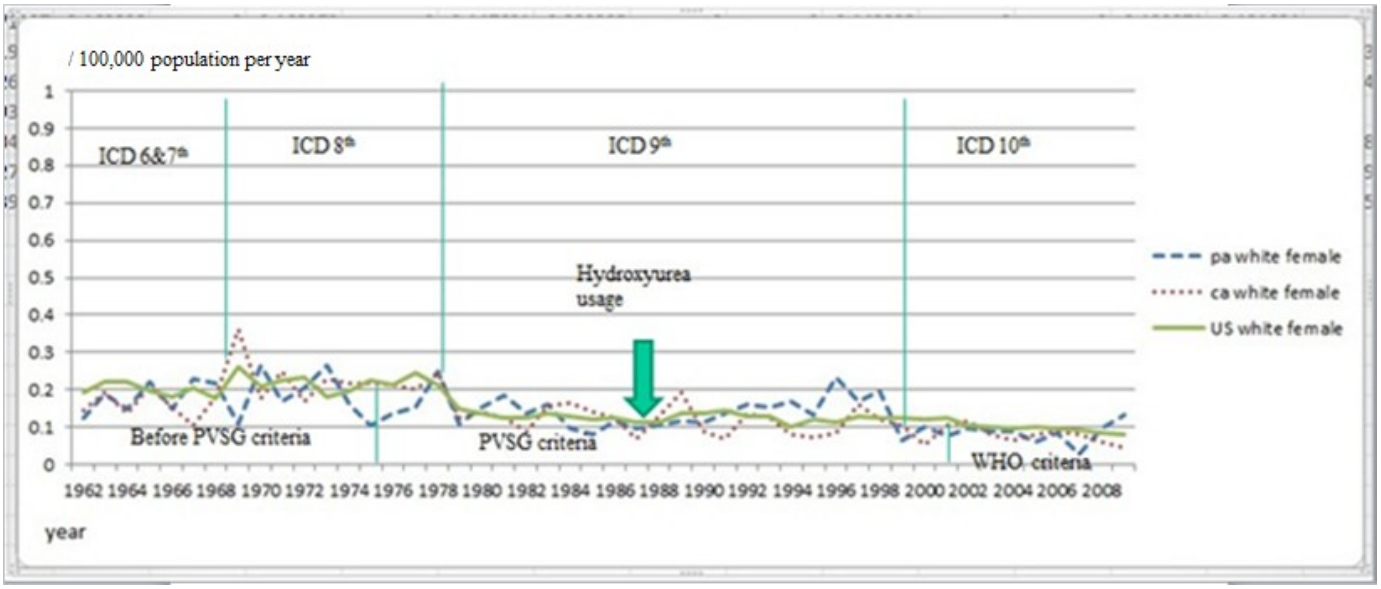

Figure 16: Comparison of Age-adjusted mortality of white females in PA, CA, and other areas in the US

3.2 MODEL BUILDING

3.2.1 Univariate negative binomial models

Tables 16 to 21 show the univariate models of predictors and confounders. We placed more emphasis on the strength and direction of the influence of these predictors than on the statistical significance. In this section, the baseline 'US' population excludes both the PA and CA populations.

Age has a large positive influence on PV mortality; the people in the 75+ age group (coefficient = 5.24) were more likely to die due to PV than people whose ages ranged from 45 to 74 (coefficient = 3.61). The P-value of age demonstrates that age is a significant predictor of PV mortality. The trend test was statistically significant, indicating that PV mortality increases as people age (Table 10).

Table 10: Negative binomial regression on single variable: age

		β	p-value	global test	trend test
age	<45	0	-	<0.001	<0.001
	45-74	3.61	<0.001		
	75+	5.24	<0.001		
constant		-16.2			

Compared to other areas in the US, PA had a lower mortality from PV (coefficient = -0.01); similarly, CA also had a lower mortality when compared to the remainder of the US (coefficient = -0.01). However, this predictor was not statistically significant (Table 11).

Table 11: Negative binomial regression on single variable: region

		β	p-value	global test
region	US	0	-	0.997
	PA	-0.01	0.982	
	CA	-0.01	0.982	
constant		-11.87		

Compared to females, males had a higher mortality from PV although the difference was not statistically significant (coefficient = 0.21). (Table 12)

Table 12: Negative binomial regression on single variable: sex

		β	p-value
sex	female	0	-
	male	0.21	0.475
constant		-11.99	

After the first ICD code revision, PV mortality slightly increased, and then decreased after the other two revisions (Table 13). This pattern of change matches what we observed in the age-adjusted PV mortality graph of US white males(Table 13). ICD code revision is not a statistically significant predictor in this model. The trend test is statistically significant, indicating that PV mortality decreased as ICD code changed.

Table 13: Negative binomial regression on single variable: ICD code revision

		β	p-value	global test	trend test
ICD	6&7th	0	-	0.1684	0.029
	8th(1968)	0.07	0.875		
	9th(1979)	-0.38	0.402		
	10th(1999)	-0.72	0.124		
constant		-11.64			

After every change in PV diagnostic criteria, the mortality decreased (Table 14). Diagnostic criteria changes were not statistically significant.

Table 14: Negative binomial regression on single variable: diagnostic criteria

		β	p-value	global test	trend test
diagnostic criteria	before PVSG	0	-	0.2048	0.142
	PVSG(1975)	-0.36	0.275		
	WHO(2001)	-0.81	0.079		
constant		-11.59			

Furthermore, after wide usage of Hydroxyurea by clinicians, PV mortality decreased (Table 15).. This predictor is statistically significant.

Table 15: Negative binomial regression on single variable: treatment

		β	p-value
treatment	before HU	0	-
	HU(1987)	-0.6	0.037
constant		-11.66	

3.2.2 Time categorical negative binomial models

Our final negative binomial regression models are listed in Tables 16-18. The model selection method was backward selection. Although region and the three time variables were not statistically significant in the univariate model global test, they were our research targets. So we included these factors in the beginning then did backward selection. The results of the test of α were all statistically significant, which indicated that the variance is a quadratic form of the mean. The age, gender, and time categorical variables were statistically significant in the three models. Region was only significant in the ICD code revision model.

In the modeling of ICD codes revision, PV mortality in PA is not statistically significant different from that in other areas of the US; the ICD codes revision from 7^{th} to 8^{th} did not statistically significantly reduce the PV mortality. We also tested the difference of different

levels in the categorical variables and found that after the ICD code 9th revision and 10th revision, PV mortality differed. PV mortality in PA and CA were statistically significant different. (Table 16)

We also tested the interaction terms between time variables and region in the following models, however, none of them was statistically significant.

Table 16: Negative binomial regression model for ICD code revision

		β	sig level	test of previous category	global test	trend test
age	<45	0	-		<0.001	<0.001
	45-74	3.63	0.048			
	75+	5.55	<0.001	<0.001		
sex	female	0	-			
	male	0.37	<0.001			
ICD code	6&7th	0	-		<0.001	0.029
	8th (1968)	-0.155	0.158			
	9th (1978)	-0.96	<0.001	<0.001		
	10th (1999)	-1.42	<0.001	<0.001		
region	OTHER	0	-		0.0338	
	PA	0.04	0.553			
	CA	-0.23	0.04	0.016		
constant		-15.91	<0.001			

In the modeling of diagnostic criteria, neither PA nor CA had statistically significant difference in PV mortality compared to other areas of the US. We also tested the difference of different levels in the categorical variables and found that after the WHO criteria acceptance, the PV mortality differed. PV mortality in PA and CA were not statistically significant different. The interaction between gender and criteria changes here indicated that PV mortality of white males decreased more than those of white females after acceptance of PVSG criteria (Table 17).

Table 17: Negative binomial regression model for diagnostic criteria change

		β	sig level	test of previous category	global test	trend test
age	<45	0	-		<0.001	<0.001
	45-74	3.63	<0.001			
	75+	5.51	<0.001	<0.001		
sex	female	0	-			
	male	0.39	<0.001			
diagnostic criteria	BEFORE PVSG	0	-		<0.001	0.142
	PVSG (1975)	-0.81	<0.001			
	WHO (2001)	-1.47	<0.001	<0.001		
region	OTHER	0	-		0.13	
	PA	0.05	0.997			
	CA	-0.16	0.29	0.06		
sex#criteria	female#before PVSG	0	-			
	male#PVSG	-0.47	0.01			
	male#WHO	-0.41	0.12			
constant		-15.95	<0.001			

In the modeling of treatment, neither PA nor CA had statistically significant differences in PV mortality as compared with other areas of the US. We also tested the difference of different levels in the categorical variables and found that PV mortality in PA and CA were statistically significant different (Table 18).

Table 18: Negative binomial regression model for treatment

		β	sig level	test of previous category	global test	trend test
age	<45	0	-		<0.001	<0.001
	45-74	3.62	<0.001			
	75+	5.51	<0.001	<0.001		
sex	female	0	-			
	male	0.39	<0.001			
Treatment	before HU	0	-			
	HU (1987)	-1.04	<0.001			
region	OTHER	0	-			
	PA	0.06	0.554		0.06	
	CA	-0.18	0.074	0.027		
constant		-16.15	<0.001			

3.2.3 Region-specific models

From Table 19, we found that all of the predictors (age, gender, and ICD codes revision), significantly affected PV mortality. The influence of the first revision of ICD codes in PA was greater than that of CA and other US areas. In the three populations, the first revision of ICD codes did not significantly affect PV mortality. Comparing the three models, we observed that PA white males had a higher ratio of PV mortality than that of white males in CA and other areas of the US. In CA, people 75 and older had a higher mortality ratio than in PA and other areas of the US. Tests between the three times of the ICD code revision showed that after each change, PV mortality statistically significant decreased.

Table 19: Region- specific model for ICD code revision

US		**β**	**sig level**	**test of previous category**	**global test**	**trend test**
age	<45	0	-		<0.001	<0.001
	45-74	3.69	<0.001	<0.001		
	75+	5.54	<0.001			
sex	female	0	-			
	male	0.36	0.001			
ICD code	6&7th				<0.001	0.16
	8th(1968)	-0.16	0.335	<0.001		
	9th(1979)	-1.04	<0.001	<0.001		
	10th(1999)	-1.38	<0.001			
constant		-15.91	<0.001			
PA		**β**	**sig level**	**test of previous category**	**global test**	**trend test**
age	<45	0	-		<0.001	<0.001
	45-74	3.46	<0.001			
	75+	5.24	<0.001	<0.001		
sex	female	0	-			
	male	0.44	0.001			
ICD code	6&7th	0	-		<0.001	0.103
	8th(1968)	-0.25	0.237			
	9th(1979)	-0.84	<0.001	<0.001		
	10th(1999)	-1.33	<0.001	<0.001		
constant		-15.72				
CA		**β**	**sig level**	**test of previous category**	**global test**	**trend test**
age	<45	0	-		<0.001	<0.001
	45-74	3.64	<0.001			
	75+	5.77	<0.001	<0.001		
sex	female	0	-			
	male	0.32	0.004			
ICD code	6&7th	0	-		<0.001	0.21
	8th(1968)	-0.06	0.733			
	9th(1979)	-0.94	<0.001	<0.001		
	10th(1999)	-1.51	<0.001	<0.001		
constant		-16.18				

From Table 20, we found that after the acceptance of the PVSG diagnostic criteria, PV mortality in all of the regions decreased, although the mortality in PA decreased less than that of CA and the rest of the US. Comparing the three models, we observed that PA white males had a higher ratio of PV mortality than white males in CA and other area of the US. In CA, people 75 and older had a higher mortality ratio than those in the same age group in PA and other areas of the US. Tests between the two times of diagnostic criteria change showed that after the acceptance of the WHO criteria, PV mortality statistically significant decreased.

Table 20: Region- specific model for diagnostic criteria

US		**β**	**sig level**	**test of previous category**	**global test**	**trend test**
age	<45	0	-		<0.001	<0.001
	45-74	3.68	<0.001	<0.001		
	75+	5.47	<0.001			
sex	female	0	-			
	male	0.36	0.007			
diagnostic criteria	before PVSG				<0.001	0.32
	PVSG (1975)	-0.8	<0.001	<0.001		
	WHO (2001)	-1.46	<0.001			
constant		-15.9				
PA		**β**	**sig level**	**test of previous category**	**global test**	**trend test**
age	<45	0	-		<0.001	<0.001
	45-74	3.47	<0.001			
	75+	5.21	<0.001	<0.001		
sex	female	0	-			
	male	0.46	0.001			
diagnostic criteria	before PVSG		0	-	<0.001	0.253
	PVSG (1975)	-0.68	<0.001			
	WHO (2001)	-1.33	<0.001	<0.001		
constant		-15.8				
CA		**β**	**sig level**	**test of previous category**	**global test**	**trend test**
age	<45	0	-		<0.001	<0.001
	45-74	3.67	<0.001			
	75+	5.78	<0.001	<0.001		
sex	female	0	-			
	male	0.36	0.026			
diagnostic criteria	before PVSG	0	-		<0.001	0.478
	PVSG (1975)	-0.8	<0.001			
	WHO (2001)	-1.52	<0.001	<0.001		
constant		-16.2				

From Table 21, we found after wide usage of Hydroxyurea for treating PV, PV mortality decreased significantly. The decrease in mortality in PA was less than that in CA and other area in the US. Also, from these three tables, in comparison to CA and other areas in the US, PA has a higher risk ratio of PV mortality among white males. In CA, people 75 and older had a higher mortality ratio than the same age group in PA and other areas of the US.

Table 21: Region- specific model for treatment

US		**β**	**sig level**	**test of previous category**	**global test**	**trend test**
age	<45	0	-		<0.001	<0.001
	45-74	3.67	<0.001	<0.001		
	75+	5.49	<0.001			
sex	female	0	-			
	male	0.37	0.004			
Treatment	before HU	0	-			
	HU (1987)	-1.05	<0.001			
constant		-16.16				
PA		**β**	**sig level**	**test of previous category**	**global test**	**trend test**
age	<45	0	-		<0.001	<0.001
	45-74	3.48	<0.001			
	75+	5.2	<0.001	<0.001		
sex	female	0	-			
	male	0.44	0.005			
Treatment	before HU	0	-			
	HU (1987)	-0.87	<0.001			
constant		-15.9				
CA		**β**	**sig level**	**test of previous category**	**global test**	**trend test**
age	<45	0	-		<0.001	<0.001
	45-74	3.65	<0.001			
	75+	5.74	<0.001	<0.001		
sex	female	0	-			
	male	0.36	0.01			
Treatment	before HU	0	-			
	HU (1987)	-1.16	<0.001			
constant		-16.39				

3.2.4 Non-overlapping time period model

After analyzing each time variable separately, we found that each was statistically significant with region, gender, and age in the model. We created a new time variable (seven segments) to include all these events in one model. Results are shown below (Table 22): PV mortality in PA was not statistically significant different from that in other parts of the US, while CA mortality was statistically significantly different from that in PA or other areas in the US. As people aged, PV mortality increased. White males have statistically significant higher rates than white females. The coefficients in the model are all negative, which indicates that PV mortality decreased from 1962 to 2009. By comparing with the previous time segment, we found that the acceptance of the PVSG criteria in 1975, the change from the ICD 8^{th} to 9^{th} revision in 1978, and the wide usage of Hydroxyurea in 1987 statistically significantly decreased PV mortality. The trend test was statistically significant, indicating that PV mortality decreased from 1962 to 2009.

Table 22: Non-overlapping time period model

		β	sig level	test of previous category	global test	trend test
age	**<45**	0	-		<0.001	<0.001
	45-74	3.62	<0.001			
	75+	5.55	<0.001	<0.001		
sex	female	0	-		<0.001	
	male	0.37	<.001			
region	US	0	-		0.03	
	PA	0.058	0.489			
	CA	-0.16	0.04	<0.001		
Time	1962	0	-		<0.001	0.044
	1968(ICD 7th-8th)	-0.07	0.541			
	1975(PVSG criteria)	-0.25	0.044	<0.001		
	1978(ICD 8th-9th)	-0.83	<0.001	<0.001		
	1987(HU used)	-1.1	<0.001	0.027		
	1999(ICD 9th-10th)	-1.3	<0.001	0.102		
	2001(WHO criteria)	-1.5	<0.001	0.25		
constant		-15.91				

4.0 DISCUSSION

In the US population, age-adjusted PV mortality of males decreased after 1967, while that of females decreased after 1978. The rate among males was always higher than that of females. The age-specific PV mortality of males ages 45-74 continuously decreased over time, but that of males over 75-years-old increased from 1962 to 1974 and then decreased. The mortality rate of females ages 45-74 was stable from 1962 to 1978 and then quickly decreased, while that of females over 75-years-old increased from 1962 to 1978 and then decreased, similar to the pattern in males over 75. Ania reported in Olmsted County, Minnesota, that the PV age-adjusted incidence was 2.0, 1.1, 2.8, and 1.9 per 100,000 population per year in 1950-1959, 1960-1969, 1970-1979, and 1980-1989, respectively. [16]This rate varied around the average incidence 1.9.[16] On the contrary, Berglund found the PV incidence rose in Malmo, Sweden, from 1950-1984, with decade specific rates of: 1.0, 1.6, 2.2, and 2.6 in 1950-1959, 1960-1969, 1970-1979, and 1980-1984, respectively.[20] Similarly, Silverstein observed an increasing PV incidence (average is 1.6 per 100,000 population per year) in Rochester, Minnesota, from 1935 to 1969, though the number of cases (only 19) were too small to draw any conclusions about secular trend.[21] The average of age-adjusted mortality in the US is 0.21 per 100,000 population per year in males, and 0.15 in females. Our result found PV had a decreasing mortality trend. Why our trend is different from above result possibly caused by the difference of methods. Both Ania and Berglund got their cases from the local medical records to diagnosed PV patients, but our data sources were drawn from the death certificates by MPDS system. Ania had 55 cases and Berglund had 177 cases, but we had 18,743 cases. Obviously, difference in the sample size can lead to different statistical result. Berglund reported that the highest age-specific incidence occurred in males over 80 years old and was 18.3/100,000 population per year, and the peak incidence in 70-79 year old females was 14.6/ 100,000 population per year. In our age-specific mortality analysis, we found the highest age-specific mortality rate similarly occurred in over 75-84 year old males and

females after 1977 but before that time in over 85 year old group. Additionally, Ania found in the Olmsted County analysis that the PV incidence of males was higher than that of females. Our model also showed a statistically significantly higher mortality rate in males. Berglund drew the conclusion that the increasing PV incidence was caused by increasing case ascertainment. It could be possible that in a similar manner, PV mortality is decreasing due to better treatment, as the survival time of treated patients is almost 10 times higher than those who are untreated.

The table of death counts shows that 95% of deaths due to PV occur in whites. In the age-adjusted mortality tables, we observed that the rates in non-whites are much lower than those in whites in the US population; there were many zero values in non-whites in the PA and CA populations. In an epidemiologic study of PV in Baltimore from 1950 to 1959, the PV incidence in African-Americans was 1.6 per million residents in the 10-year period, while that of whites was 4.3 per million. In this study, we found that PV deaths among non-whites are rare. In his paper, Modan said "The reported low frequency of polycythemia in Negroes is reflected in our findings, but the observed incidence, although only a third of that in whites is not as low as reported previously. In addition, as will be demonstrated below, polycythemia vera is more prevalent in the older age groups, where there are relatively less Negroes than whites." Although Modan said these words in 1965, life-expectancies of whites and non-whites are still statistically significantly different today. Additionally, differences in medical care can lead to differences in mortality rates between whites and non-whites. In short, the small numbers of PV deaths among non-whites found in this study may reflect a complex interaction of factors.

In the non-overlapping models, we found no statistically significant difference in PV mortality between PA and other parts of the US, and the coefficients of PA in the model were positive. This indicates that the pattern of PV deaths in PA is not different from that of other parts of the US. However, we observed that PA males had higher PV age-adjusted mortality compared with CA and other areas of the US before the ICD code revision change in 1978.

Our time period analysis, corresponding to changes in ICD coding, diagnostic criteria, and implementation of Hydroxyurea as standard treatment, found that all of these variables significantly influenced PV mortality. The overlapping time periods made these models difficult to fit. If we included all three time variables in one model, the overlapping time periods could lead to poor statistical significance for some variables; on the contrary, if we included only one time variable in the model, we risked ignoring important predictors. To solve this problem, we

built a model including all of the time period changes by dividing the timeline into seven segments, and found that ICD 8th to 9th revision, acceptance of PVSG criteria, and Hydroxyurea usage were time points when PV mortality significantly decreased. Then we compared these results with the three time specific variables, and found the model including all of the time periods had fewer statistically significant changes. In the one-time-factor models, except for the ICD 6&7th to 8th revision change, other time categories were statistically significant when compared with earlier time periods. We also compared the model with the adjusted mortality table of the US (Fig 2). Based on all of the findings, we concluded that the 8th to the 9th ICD codes revision (in 1978) significantly reduced PV mortality, and the acceptance of the PVSG diagnostic criteria (in 1975) significantly reduced the difference of PV mortality between males and females. Moreover, the widespread usage of Hydroxyurea in 1987 helped to statistically significantly reduce PV mortality.

From the age-adjusted mortality table of US males and females, we observed a drop in 1978 when the ICD code changed from the 8th to 9th revision. As it influenced both genders in PV was classified as a "malignant neoplasm" before the 9th revision ICD was introduced in 1979 and in the 9th and 10th revisions it was classified as a "benign neoplasm and neoplasm of uncertain behavior and of unspecified nature". In 1978, the age-adjusted PV mortality of US males was 0.28 per 100,000 population per year, and the rate in females was 0.22 per 100,000 population per year. However in 1979, the rates in males and females were 0.17 and 0.14 per 100,000 population per year, respectively. The change in ICD revision does appear to have influenced PV mortality directly. From the time period 1975-1977 to the period 1978-1987, it is observed that for people under 75 there is a 50% decrease in age-specific mortality in both males and females. Between the same two time periods, there was a 25% decrease in the mortality rate seen in 75+ males and a 30% decrease in 75+ females. The 8th to 9th ICD revision change influenced the mortality rates in younger people more than older people. . Identifying this discontinuity may help other PV researchers avoid bias when analyzing long-term mortality trends.

After the 1975 PVSG diagnostic criteria were introduced, the major diagnostic criteria required for diagnosis increased from one to three. Etiology of other polycythemia was ruled out, and evidence of myeloproliferative disorder was needed. After these criteria were accepted, the difference in PV mortality between males and females decreased quickly. From the age-specific

mortality tables, we found that this decrease was caused by a large decrease in the mortality rate among 45-74 year old males and a small increase in the rate in 75+ year old females. In the 1968-1974 time period, age-adjusted rate in white male was 0.92 per 100,000 population per year, while in the 1975-1977 time period, this value was 0.59 per 100,000 population per year. At the same time, the rate in 75+ year old females changed from 2.01 to 2.21 per 100,000 population per year. The rate in 75+ year old males slightly decreased while that in other female groups did not change. We also know that the PVSG began in 1967 and did a series of clinical trials to determine the most appropriate diagnostic criteria and treatment for PV. Because we know the diagnostic criteria were designed to reduce false positive diagnoses after 1975, and we observed a decline in the difference between PV mortality of two genders, it is possible that males had more false-positive PV diagnoses than females before 1975. It was reported that secondary polycythemia, caused mainly by smoking or high blood pressure, can lead to false-positive diagnosis of PV. Because males were more likely to be smokers in this time period, the declining difference in PV mortality between the genders may have been caused by more false-positive cases in males being excluded.

Limitations of this research include those that existed in the source data. The data in this study were generated using the MPDS system, which was based on death certificates in the NCHS repository. Because PV patients mainly die from thrombosis if untreated, some PV patients' deaths may be recorded in other underlying causes of death categories, such as cardiovascular disease.[15] Furthermore, PV patients could develop myelofibrosis or acute myeloid leukemia (AML) as their disease progresses. In these situations, the underlying causes of deaths would not be PV despite its strong association with those diseases. These limitations in cause of death coding and reporting may have underestimated PV mortality.

In the other similar epidemiology research about cause-specific mortality trends, especially in cancers, Poisson models were widely used to model the count of death with an offset of population in the year. However, our research found that the Poisson model did not fit the data because of serious over dispersion. After examining the possible use of the over-dispersed Poisson model and negative binomial model, we finally picked the second one as it had relative smaller AIC and BIC values. Although no direct method to check the fit exists, we found that the modeling results matched the age-adjusted graphs well. We found a lessening of the difference in PV mortality between males and females after the acceptance of the 1975 PVSG

diagnostic criteria, a decrease in mortality from the ICD 8th to 9th ICD revision, and a slow, steady decline in mortality after Hydroxyurea became widely used in 1987, in both the non-overlapping model and the age-adjusted graphs. Furthermore, males had statistically significantly higher mortality than females. In short, our model fits these data well.

Although some limitations exist, this study is the first attempt to examine PV mortality in PA and the US over an almost 50-year period. Previous population-based studies of PV mortality have used only local populations with small sample sizes. The present study included 18,743 cases and offered a more complete and comprehensive examination of PV mortality than any other to date.

5.0 CONCLUSION

This study is the first comprehensive examination of PV mortality rates over time in the US and PA. We found that PV age-adjusted mortality in US males increased from 1962 to 1967, then decreased after 1967 while of the mortality rate in US females variedwas relatively stable at first then decreased after 1978. PV death patterns in PA and the US were not statistically significantly different. In the US, PV mortality of males was statistically significantly higher than that of females, while in PA there was higher mortality in males compared with females. After consideration of possible influences on PV mortality rates, we determined that the acceptance of the PVSG diagnostic criteria in 1975 decreased the magnitude of the difference in PV mortality between males and females seen prior to 1975, apparently due to the introduction of sex-specific red cell mass values. The transition from the 8^{th} to the 9^{th} ICD death code revision in 1979 decreased PV mortality significantly in both males and females. While the specificity of the code remained the same from the 8^{th} to 9^{th} revision, PV was one of several malignancies changed to “benign” and/or “uncertain and unspecified behavior” in the 9^{th} revision. This may have had an effect on mortality coding and subsequent effect on mortality rates.

APPENDIX A

SUMMARY OF DEATH COUNT

Table 23 Summary of death count

	WM < 45			WM 45-74			WM >75			WF < 45			WF 45-74			WF >75			NWM < 45			NWM 45-74			NWM >75			NWF< 45			NWF 45-74			NWF >75		
	US	PA	CA	US	PA	CA	US	PA	CA	US	PA	CA	US	PA	CA	US	PA	CA	US	PA	CA	US	PA	CA	US	PA	CA	US	PA	CA	US	PA	CA	US	PA	CA
1962	19	1	2	202	17	13	49	1	6	14	0	0	105	5	9	44	3	3	0	0	0	5	0	0	1	0	0	0	0	0	5	0	0	0	0	0
1963	16	2	0	189	13	14	42	2	2	7	0	0	131	10	12	58	2	4	4	0	0	12	0	1	1	0	0	0	0	0	8	1	0	1	0	0
1964	17	3	2	184	11	15	62	4	8	7	0	0	121	6	7	71	4	5	1	0	0	6	1	0	0	0	0	2	0	0	4	0	0	0	0	0
1965	19	2	3	198	19	15	61	2	7	4	0	0	124	8	11	67	7	8	8	2	0	22	2	1	1	0	0	1	0	0	11	0	1	0	0	0
1966	18	0	1	219	28	20	58	5	7	9	0	2	111	6	7	53	4	5	3	0	0	8	1	0	2	0	1	0	0	0	11	0	1	2	0	0
1967	10	1	2	249	29	17	68	4	8	10	2	1	112	7	4	76	6	5	1	0	0	10	1	1	3	0	0	1	0	0	5	0	0	2	0	0
1968	23	1	1	203	15	9	77	2	3	8	0	0	108	9	8	69	6	10	2	0	0	7	0	0	2	0	0	0	0	0	6	0	0	1	0	0
1969	15	2	0	235	14	23	75	4	5	8	0	1	158	2	22	104	6	12	2	0	0	14	0	0	2	0	0	0	0	0	5	0	1	1	0	0
1970	10	0	2	218	15	12	69	6	8	9	1	0	134	13	6	77	4	13	3	0	1	14	2	2	2	1	0	0	0	0	4	0	0	1	0	0
1971	19	2	1	223	22	10	66	2	6	7	1	0	128	5	10	107	6	16	1	0	0	9	0	0	1	0	0	0	0	0	8	0	1	0	0	0
1972	19	1	1	210	16	13	67	2	10	1	0	0	141	6	12	108	9	5	3	0	1	16	0	1	0	0	0	0	0	0	6	1	0	2	0	1
1973	12	1	2	184	9	13	69	2	7	9	1	0	113	10	14	89	8	10	2	0	0	11	0	1	4	0	0	1	0	0	7	0	0	1	0	0
1974	18	0	3	182	14	15	75	7	7	5	0	0	111	4	8	108	8	16	4	0	1	3	0	1	3	0	1	1	0	0	9	1	0	5	0	0
1975	12	1	1	149	9	14	85	7	11	3	0	1	128	3	10	118	5	13	2	0	0	9	1	0	2	0	0	2	0	0	3	0	0	3	0	1
1976	6	0	0	155	14	12	56	4	5	9	0	0	119	6	8	111	4	17	4	0	0	6	0	0	1	0	0	1	0	0	4	0	0	3	1	0
1977	7	0	1	130	8	10	78	2	9	9	1	1	143	6	9	122	4	13	1	0	0	10	0	2	5	0	0	0	0	0	6	1	0	2	0	0
1978	12	1	0	144	14	9	78	5	13	9	2	3	114	7	12	137	9	13	1	0	1	8	0	0	3	0	0	1	0	0	2	0	1	4	0	0
1979	5	0	0	79	2	4	55	5	7	4	0	1	83	5	10	89	3	3	0	0	0	3	2	0	2	1	0	0	0	0	4	0	0	3	0	0
1980	9	0	2	88	6	7	54	5	5	1	0	0	74	4	4	100	8	14	3	0	0	5	0	2	4	0	0	1	0	0	9	0	2	3	0	0
1981	7	1	0	82	2	5	49	3	4	2	0	0	53	4	6	109	11	9	0	0	0	0	0	0	2	0	1	0	0	0	5	0	0	4	0	2
1982	2	2	0	85	4	9	66	6	5	3	0	1	57	3	2	98	8	8	2	0	0	7	0	0	2	0	0	0	0	0	4	0	0	4	0	0
1983	5	0	0	84	8	6	75	4	11	1	0	0	71	5	6	110	7	14	1	0	1	6	0	2	2	0	0	0	0	0	2	1	0	2	1	0
1984	3	0	0	67	4	4	63	4	6	2	0	0	56	2	8	113	6	13	2	0	0	14	0	1	3	1	0	0	0	0	4	0	0	3	0	1
1985	4	1	0	70	8	4	65	5	7	4	1	0	60	1	7	96	4	11	0	0	0	3	0	1	1	0	1	0	0	0	4	0	0	3	0	0
1986	1	0	0	56	6	4	75	3	5	2	0	0	60	5	2	103	4	15	0	0	0	6	0	0	6	0	0	0	0	0	7	1	0	6	1	0
1987	4	0	0	65	4	12	68	1	4	3	0	0	53	4	4	90	4	5	0	0	0	5	0	0	2	0	0	0	0	0	3	0	0	3	0	0
1988	1	0	0	63	4	6	66	2	8	2	1	1	58	2	7	93	5	9	0	0	0	3	0	0	5	0	0	0	0	0	9	0	1	5	0	0
1989	7	1	1	85	3	8	58	6	6	1	0	0	70	2	10	120	8	15	1	0	0	4	0	0	5	0	1	0	0	0	4	0	0	4	0	0
1990	1	0	0	46	3	3	63	1	6	3	0	0	55	6	2	123	3	10	0	0	0	7	0	0	1	0	0	0	0	0	12	2	0	3	0	0
1991	1	0	1	73	4	9	81	3	6	0	0	0	58	2	6	133	10	3	0	0	0	2	0	0	4	0	0	1	0	0	4	3	0	7	2	0
1992	3	0	0	58	4	3	67	2	6	1	0	0	39	3	3	151	11	17	0	0	0	9	0	2	1	0	0	0	0	0	3	0	1	5	0	0
1993	3	0	0	54	4	7	69	6	4	2	0	0	38	0	3	149	14	15	2	0	0	8	0	2	1	0	1	0	0	0	7	1	1	7	1	0
1994	1	0	0	55	4	2	88	2	7	1	0	1	39	4	0	110	10	11	2	0	0	6	1	1	6	2	0	0	0	0	6	1	0	3	1	0
1995	0	0	0	48	3	3	85	3	9	1	0	0	38	1	2	133	11	8	1	0	1	6	0	0	2	0	0	0	0	0	2	0	0	8	0	0
1996	1	0	0	64	3	5	88	5	9	1	0	0	28	4	4	148	17	8	1	0	1	5	0	0	4	0	2	1	0	0	5	1	0	11	1	3
1997	2	0	0	39	1	3	76	6	1	2	1	0	44	5	5	153	7	18	0	0	0	6	0	1	2	0	0	1	0	0	4	0	0	4	0	0
1998	2	0	0	65	3	4	82	5	5	1	0	0	49	7	4	145	10	14	2	0	1	3	1	0	4	1	0	1	0	0	4	1	1	3	0	0
1999	2	0	0	39	4	2	81	3	6	1	0	0	45	1	3	132	5	13	1	0	0	5	0	0	3	0	1	0	0	0	6	0	0	5	2	0
2000	4	0	0	43	3	2	87	8	7	1	0	0	40	1	3	136	9	5	0	0	0	3	0	0	4	0	1	1	0	0	3	0	0	4	0	0
2001	6	0	0	51	2	7	86	5	8	2	0	0	56	2	4	128	5	12	0	0	0	3	0	0	3	0	1	0	0	0	5	0	2	4	0	1
2002	3	1	0	46	2	2	80	6	10	3	0	0	30	3	0	134	5	19	1	0	0	3	0	0	5	0	0	1	0	0	4	1	0	7	1	0
2003	3	0	0	59	1	9	75	2	12	0	0	0	36	2	1	117	6	12	1	0	1	5	0	0	2	0	1	1	0	0	3	0	2	7	0	0
2004	3	0	0	51	2	1	85	3	9	3	0	0	33	0	1	114	9	9	0	0	0	3	0	1	2	0	1	1	0	0	1	0	0	8	0	1
2005	1	0	0	37	1	3	80	2	9	0	0	0	27	0	2	131	6	10	0	0	0	3	0	1	5	0	2	0	0	0	1	0	0	5	1	2
2006	1	0	0	44	3	4	66	3	4	2	1	0	32	1	3	119	5	10	1	0	0	1	1	0	3	1	0	0	0	0	2	1	0	10	1	4
2007	3	0	1	45	3	3	80	4	11	0	0	0	30	0	3	111	3	9	1	0	0	3	0	0	4	1	0	0	0	0	2	0	1	4	0	1
2008	2	0	1	32	3	2	77	9	6	1	0	0	31	1	0	103	8	10	2	0	0	4	0	0	2	0	1	0	0	0	4	0	0	6	0	2
2009	2	0	1	42	1	4	73	6	3	2	1	0	30	1	1	100	9	6	1	0	0	4	1	0	3	0	0	0	0	0	1	0	0	2	0	0

APPENDIX B

AGE SPECIFIC MORTALITY

Table 24 Age-specific mortality of US white males

	1962 -1967	1968-1974	1975-1977	1978-1986	1987-1998	1999-2000	2001-2009
45 to 54	0.42	0.38	0.16	0.08	0.04	0.03	0.04
55 to 64	1.09	1	0.65	0.28	0.18	0.12	0.11
65 to 74	1.79	1.85	1.31	0.83	0.54	0.48	0.35
75 to 84	2.46	2.68	2.41	1.74	1.47	1.25	1.05
85 plus	1.6	2.41	2.38	2.23	2.68	2.82	2.06
less than 45	0.03	0.03	0.01	0.01	0	0	0
45 to 74	0.95	0.92	0.6	0.34	0.21	0.16	0.12
75 plus	2.32	2.63	2.41	1.84	1.7	1.57	1.27

Table 25 Age-specific mortality of US white females

	1962 -1967	1968-1974	1975-1977	1978-1986	1987-1998	1999-2000	2001-2009
45 to 54	0.2	0.16	0.15	0.05	0.02	0.03	0.01
55 to 64	0.44	0.47	0.45	0.15	0.1	0.12	0.08
65 to 74	1.15	1.07	1.07	0.61	0.39	0.33	0.28
75 to 84	1.85	2.12	2.32	1.47	1.28	1.08	0.79
85 plus	1.33	1.61	1.86	2.23	2.37	2.63	2.11
less than 45	0.02	0.01	0.01	0	0	0	0
45 to 74	0.51	0.49	0.5	0.25	0.15	0.13	0.09
75 plus	1.75	2.02	2.22	1.66	1.58	1.53	1.2

Table 26 Age-specific mortality of PA white males

	1962 -1967	1968-1974	1975-1977	1978-1986	1987-1998	1999-2000	2001-2009
45 to 54	0.41	0.41	0.32	0.08	0.02	0	0.03
55 to 64	1.79	1	0.77	0.35	0.08	0.14	0.17
65 to 74	2.41	2.12	1.15	0.91	0.68	0.67	0.2
75 to 84	2.07	2.16	1.82	2.13	1.15	1.32	0.98
85 plus	0.68	1.51	5.11	2.64	1.94	3.3	2.52
less than 45	0.04	0.03	0.01	0.02	0	0	0
45 to 74	1.32	0.98	0.67	0.4	0.23	0.21	0.11
75 plus	1.86	2.06	2.41	2.22	1.29	1.7	1.32

Table 27 Age-specific mortality of PA white females

	1962 -1967	1968-1974	1975-1977	1978-1986	1987-1998	1999-2000	2001-2009
45 to 54	0.14	0.06	0.05	0.1	0.01	0	0.03
55 to 64	0.54	0.35	0.21	0.22	0.08	0.06	0.05
65 to 74	0.74	1.04	0.71	0.38	0.44	0.53	0.12
75 to 84	1.78	2.45	1.17	1.64	1.65	0.58	0.78
85 plus	1.87	1.73	1.83	2.01	2.43	3.62	1.69
less than 45	0.01	0.01	0.01	0.01	0.01	0	0.01
45 to 74	0.42	0.4	0.28	0.23	0.18	0.17	0.06
75 plus	1.8	2.31	1.32	1.73	1.85	1.43	1.06

Table 28 Age-specific mortality of CA white males

	1962 - 1967	1968-1974	1975-1977	1978-1986	1987-1998	1999-2000	2001-2009
45 to 54	0.29	0.21	0.1	0.06	0.04	0	0.03
55 to 64	0.68	0.67	0.35	0.16	0.18	0.07	0.11
65 to 74	2.03	1.37	1.54	0.62	0.49	0.28	0.26
75 to 84	3	2.94	3.26	1.81	1.24	0.94	0.98
85 plus	2.99	1.33	1.92	3.56	1.7	1.43	2.43
less than 45	0.03	0.02	0.01	0	0	0	0
45 to 74	0.78	0.6	0.5	0.23	0.19	0.08	0.1
75 plus	3	2.66	3.01	2.15	1.33	1.04	1.32

Table 29 Age-specific mortality of CA white females

	1962 - 1967	1968-1974	1975-1977	1978-1986	1987-1998	1999-2000	2001-2009
45 to 54	0.03	0.16	0.09	0.04	0.03	0.02	0.01
55 to 64	0.46	0.53	0.29	0.13	0.09	0	0.05
65 to 74	0.88	0.88	0.78	0.59	0.34	0.35	0.09
75 to 84	1.49	2.82	2.31	1.57	1.1	0.62	0.87
85 plus	1.55	2.36	4.47	2.62	2.22	2.63	1.73
less than 45	0.01	0	0.01	0.01	0	0	0
45 to 74	0.38	0.45	0.33	0.22	0.13	0.09	0.04
75 plus	1.5	2.73	2.82	1.85	1.41	1.19	1.14

BIBLIOGRAPHY

1 A.D.A.M. Medical Encyclopedia: Polycythemia vera [On line] Available from: http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0001615/

2 Levine, R.L., et al., *Activating Mutation in the Tyrosine Kinase JAK2 in Polycythemia Vera, Essential Thrombocythemia, and Myeloid Metaplasia with Myelofibrosis*. Cancer Cell. 2005; 7(4):387–97.

3 Dameshek, W., *Some speculations on the Myeloproliferative Syndromes.* Blood 1951; 6(4):372–375

4 Turkington, R.C., *Comparison of Diagnostic Criteria for Polycythaemia Vera.* Hematology, Volume 12, Issue 2, pp. 123 - 130

5 Green, A.R. and Vassiliou, G.S., *Management of the Myeloproliferative Disorders: Distinguishing Data from Dogma.* The Hematology Journal, 2004, Volume 5 Suppl 3, pp. 126 - 132

6 Heike, L.P., *Diagnostic Approaches to Polycythemia Vera in 2004.* Expert Review of Molecular Diagnostics, 07/2004, Volume 4, Issue 4, p. 495

7 Laurie, E. and Fundukian, J. *Secondary Polycythemia*. The Gale Encyclopedia of Medicine, 2011, Vol.5, p3884-3885.

8 Polycythemia Vera [Online] Available from: http://emedicine.medscape.com/article/205114-overview

9 Pearson, T.C., *A Polycythemia Vera Update: Diagnosis, Pathobiology, and Treatment.* Hematology, 2000, pp. 51-68

10 Vannucchi, A.M., *Advances in Understanding and Management of Polycythemia Vera.* Current Opinion in Oncology, 11/2010, Volume 22, Issue 6, pp. 636 - 641

11 Zhao R., *Identification of an Acquired JAK2 Mutation in Polycythemia Vera.* Journal of Biological Chemistry, 06/2005, Volume 280, Issue 24, pp. 22788 - 22792

12 Ayalew T., *The Diagnosis of Polycythemia Vera: New Tests and Old Dictums.* Best Practice & Research Clinical Haematology, Volume 19, Issue 3, pp. 455 – 469

13 Tefferi, A. and Vardiman, J.W., *Classification and Diagnosis of Myeloproliferative Neoplasms: the 2008 World Health Organization Criteria and Point-of-care Diagnostic Algorithms.* Leukemia, 01/2008, Volume 22, Issue 1, pp. 14 - 22

14 Mannucci, P.M., *Polycythaemia Vera and JAK2 Mutation.* Internal and Emergency Medicine, 06/2007, Volume 2, Issue 2, p. 147

15 Johansson, P., *Epidemiology of the Myeloproliferative Disorders Polycythemia Vera and Essential Thrombocythemia.* Seminars in Thrombosis and Hemostasis, 04/2006, Volume 32, Issue 3, pp. 171 - 173

16 Ania B.J., et al, *Trends in the Incidence of Polycythemia Vera Among Olmsted County, Minnesota Residents, 1935-1989*. Am J Hematology, 1994, Vol. 47, pp.89–93

17 Seaman, V. and Marchetto, D., *A Multidisciplinary Investigation of a Polycythemia Vera Cancer Cluster of Unknown Origin.* Int J Environ Res Public Health, 03/2010, Vol.7, Issue 3, pp. 1139–1152.

18 Frome, E.L., *The Analysis of Rates Using Poisson Regression Models.* Biometrics, 09/1983, Volume 39, Issue 3, pp. 665 - 674

19 Berglund, S., *Incidence of Polycythemia Vera in a Defined Population.* Eur J Haematology, 1992, Vol. 48, pp. 20-26

20 Modan, B., and Lilienfeld, A. M., *Polycythemia Vera and Leukemia-The Role of Radiation Treatment*. Medicine, 1965, Vol.44, pp. 305.

21 Modan B., *An Epidemiological Study of Polycythemia Vera*. Blood, 11/1965, Vol.26, No.5.