IDENTIFICATION OF DEMOGRAPHICS AND COMORBIDITIES ASSOCIATED WITH VASCULAR HAMARTOMAS

A THESIS IN
Bioinformatics

Presented to the Faculty of the University
of Missouri-Kansas City in partial fulfillment of
The requirements for the degree

MASTERS OF SCIENCE

by
MICHEL CONN

B.S., Missouri Western State University, 2012

Kansas City, Missouri
2016

IDENTIFICATION OF DEMOGRAPHICS AND COMORBIDITIES ASSOCIATED WITH VASCULAR HAMARTOMAS

Michel Allan Conn, Candidate for the Master of Science Degree

University of Missouri-Kansas City, 2016

ABSTRACT

Vascular hamartomas (VH) are tumor like growths that typically appear in infants and manifest as a blemish on the skin. The goal of this study is to substantiate clinical data on the demographic characteristics and discover comorbid conditions associated with VH. Previous clinical data suggests that VH are associated with preterm birth. This study will be using two data sources: the nationally-representative National Hospital Discharge Survey (NHDS) and Cerner Health Facts (HF). NHDS contained 2,944,459 patient discharges with a weighted total of 386,186,183 and HF contained 46,721,119 unique patient IDs. Survey regressions were run for the NHDS data on vascular hamartoma patients and a series of 55 ICD-9 CM codes with a frequency of 5 or higher in vascular hamartoma patients in the NHDS. Logistic regressions were run on the HF dataset for vascular hamartoma patients and the same set of 55 ICD-9 CM codes. Race, sex, region, and age were evaluated as predictors. The results show age as a significant negative predictor. Blacks and Asian/Pacific Islanders were significantly less likely to have VH than whites. Female subjects were more likely to have VH than males and patients in the

Northeast, Midwest, and West were significantly less likely to have VH than patients in the South. There were 13 significant comorbidities in the NHDS and 35 significant in HF. In both datasets 11 ICD-9 CM codes were found to be significantly associated with the diagnosis of VH. The demographics found in this analysis reflect previous clinical data and offer a population wide view of VH. The comorbidities show that VH may be associated with over development of the fetus.

APPROVAL PAGE

The faculty listed below, appointed by the Dean of the Dean of the School of Medicine have examined a thesis titled “Identification of demographics and comorbidities associated with Vascular Hamartomas,” presented by Michel A. Conn, candidate for the Master of Science degree, and certify that in their opinion it is worthy of acceptance.

Supervisory Committee

Mark Hoffman, Ph.D., Committee Chair
Department of Biomedical and Health Informatics

Jenifer E. Allsworth, Ph.D.
Departments of Biomedical and Health Informatics
And Obstetrics and Gynecology

Daniel F. Pauly, MD Ph.D.
Department of Biomedical and Health Informatics
Department of Medicine

CONTENTS

TABLES

ILLUSTRATIONS

Figures Pages

CHAPTER 1

INTRODUCTION

Vascular hamartomas (VH) are benign, tumor-like growths that typically develop in infancy, are usually dermal, and consist of primitive to well-formed blood vessels. VH consist of three major diagnoses: port-wine stains, strawberry nevi, and birthmarks (World Health Organization, 2015e). VH are relatively common, but despite this, little is known about comorbid associations or potential disease complications that could arise in vascular hamartoma patients.

Port-wine stains are cutaneous capillary malformations and usually involve the head and neck. Port-wine stains occur in 3 out of 1000 newborns. They appear red and darken over the years (Shirley et al., 2013). Port-wine stains are found in a few combined vascular diseases such as Sturge-Weber Syndrome and Klippel-Trenaunay Syndrome (Eerola et al., 2003). Port-wine stains' most prevalent complication is emotional problems related to their appearance. There is an association with glaucoma if the port-wine stain is located on the eyelid (Berman, 2015). Port-wine stains are associated with two genes, but the cause of birthmarks and strawberry nevi are mostly unknown (Cleveland Clinic, n.d.; Darrow, Greene, Mancini, & Nopper, 2015; Shirley et al., 2013). The gene GNAQ has been found as a cause of port-wine stains. A mutated version of RASA1 has also been found in patients with port-wine stains (Eerola et al., 2003; Shirley et al., 2013). Port-wine

stains clinically have been shown to affect males and females equally and be more common in whites than African Americans (Antaya, 2014).

Strawberry nevi, also known as infantile hamartomas, are the most common tumors in children. They tend to resolve without intervention. Strawberry nevi's complications are pain, functional impairment, or disfigurement (Aderibigbe, Treat, & Maguiness, 2015). Darrow, Greene, Mancini, & Nopper (2015) state that strawberry nevi occur in 5% of children, occur 1.4 to 3 times more often in females, and have the risk factors "white race, prematurity, low birth weight, advanced maternal age, multiple gestation pregnancy, placenta previa, and preeclampsia" (p. 787). The pathogenesis is unknown. The cause is thought to be a mix of intrinsic factors such as angiogenic and vasculogenic factors and extrinsic factors such as tissue hypoxia and developmental field disturbance (Darrow et al., 2015).

The generic term birthmark can be classified as a vascular hamartoma, but many birthmarks, such as café au lait spots and Mongolian spots, don't classify as VH. The cause of most birthmarks aren't known but some can be hereditary (Cleveland Clinic, 2013; U.S. National Library of Medicine, 2014).

The databases being used for this analysis are the National Hospital Discharge Survey (NHDS) and Health Facts (HF). Previously, most vascular hamartoma demographic information and comorbidities have been gathered clinically. The NHDS and HF should provide a broader, population basis, for understanding VH. The NHDS has been conducted annually since 1965 by the Center

for Disease Control (CDC) and the National Center for Health Statistics (NCHS). The goal of the NHDS is to meet the need for information on the characteristics of patients discharged from non-federal short-stay hospitals in the U.S. (Center for Disease Control, 2015). The dataset contains a variety of information such as discharge status, marital status, length of stay, and procedure codes (National Center for Health Statistics, 2010).

Health Facts is provided to UMKC from Cerner through a collaboration with Truman Medical Center. The data in HF comes from the electronic medical records of hospitals that have a data use agreement with Cerner. The HF dataset contains a large variety of encounters with each patient receiving an unique identifier. Each encounter or hospital visit is associated with the patient's identifier, sex, race, region, age, ICD-9 CM codes, and more. Cerner Corporation has established Health Insurance Portability and Accountability Act-compliant operating policies to establish de-identification for Health Facts.

This approach was modeled after genome wide association studies (GWAS). GWAS examine hundreds of thousands of single-nucleotide polymorphisms (SNPs) for association with a disease in thousands of patients. The main goal of GWAS are to gather a large number of patients with the disease being examined and sequence their genomes. The genome of the patients are then compared against a reference genome or set of controls. P-values and odds ratios are then found for each SNP. Typically a chi-square or logistic regression is computed for each SNP and an alpha

correction is applied to the analysis to account for the multiple testing issue (Bush & Moore, 2012). Associated SNPs can help localize disease-causing genes. As the proximity of the SNP to a disease-causing gene shortens, the SNP will segregate with the causative gene. This can help lead to a cure or better understanding of a disease. Similarly, a single disease can have an effect on many different parts of the body and interact and be associated with many different diagnoses. Type I and type II polyglandular autoimmune (PGA) syndrome is an example of a disease that is associated with other threatening comorbidities. PGA syndrome is caused by damaged adrenal glands and is associated with hepatitis, juvenile anemia, diabetes, and more (Neufeld, Maclaren, & Blizzard, 1981). While VH are typically thought to be benign, understanding diseases associated with VH could show underlining mechanisms and potentially dangerous diseases that occur with VH.

CHAPTER 2

METHODOLOGY

This analysis included ten years of NHDS data from 2001-2010. A new variable was created to indicate that a patient had a vascular hamartoma diagnosis. The ICD-9 CM Code 757.32 codes for VH and was used to create the vascular hamartoma variable. SAS 9.4, from the SAS institute, was the statistical program used for this analysis. The data was first imported into SAS using the code provided by the NHDS. The ten years of data was combined and then diagnosis variables were cleaned and reformatted with Stata.

In the NHDS, hospitals are selected based on their number of annual beds and discharged patients. Hospitals with the most beds and discharges are selected with certainty. The remaining are sampled using a three-stage stratified design. The data collection is a two-step process starting with hospital staff or U.S. Bureau staff manually abstracting hospital records. The next step involves automatically purchasing hospital data from commercial organizations, state data systems, hospitals, or hospital associations (Dennison & Pokras, 2000). An important note about the NHDS is that the data includes a weight variable. The NHDS is created from a small sampling of all hospitals in the United States. In order to inflate records to national or regional estimates the NHDS requires the use of the weight variable. The NHDS states, “to produce an estimate of the number of discharges, the weights for the desired records must be summed” (12). The CDC calculates weight estimates

based on the 2000 U.S. census. Also, starting in 2010 the number of diagnosis codes was increased from seven to fifteen and due to funding the number of hospitals polled were reduced (National Center for Health Statistics, 2010). The reduction of hospitals resulted in an increase in the variability of the weight variable.

A survey logistic regression was performed on the NHDS demographic variables race, sex, region, and age (table 1). In the NHDS the variable age originally consisted of two variables but was condensed for this analysis. The first variable indicated whether age was in days, months, or years and the second variable was the age of the patient. These two variables were used to create age in years. The regression was modeled for patients having VH with the demographic variables being used as effects. Race, sex, and region were treated categorically and age was treated as a continuous variable. Race consisted of white, black, American Indian/Alaskan Native, Asian, Native Hawaiian/Other Pacific Islander, and other. The census regions are West, South, Midwest, and Northeast. Whites were the reference for race, males were the reference for sex, and South was the reference for region. In the NHDS dataset the variable race included the categories multiple race and not stated. Patients that contained the race categories multiple race and not stated made up 28% of the data and were not used in this regression analysis because of their ambiguity. Results of the demographic regression can be seen on table 2. Unlike HF, the NHDS dataset did not have a Hispanic race category.

An additional analysis involved comparing comorbidities between patients with VH and patients without VH using a survey logistic regression. Vascular hamartoma patients had a total of 345 unique ICD-9 CM codes. Any ICD-9 CM code that did not appear in 5 or more vascular hamartoma patients were not used in the analysis decreasing the total to 55 unique ICD-9 CM codes. Any patient from the main dataset that did not have at least one of the 55 ICD-9 CM codes found in vascular hamartoma patients were removed from the dataset. After removing patients there were 64,573,269 weighted patients in the NHDS dataset. A survey logistic regression was run for each of the 55 ICD-9 CM codes using VH as the outcome, one of the ICD-9 CM codes as the effect, and age, race, and sex as the confounders. These variables were chosen as confounders based on results from the first regression examining demographic characteristics. The p-value used to indicate significance was 0.000909. This p-value was chosen to account for the multiple testing issue and was derived from a standard p-value (0.05) divided by 55.

A logistic regression in HF was performed for the demographics in a similar manner as the NHDS dataset. Only unique patient IDs provided by HF were included in the analysis. The race category included white, black, Asian/Pacific Islander, Hispanic, and other. HF contained Asian and Asian/Pacific islander but these were condensed into a single group for this analysis. The groups Native American, other, biracial, and Mid-Eastern Indian were also condensed into a single group called "other". Patients with the race categories NULL, Not Mapped, Null, and Unknown

were not used for this analysis and made up 5.2% of the data. Patients with the sex categories NULL, Not Mapped, Null, other, and Unknown/Invalid were not used in this analysis and comprised 0.02% of the data. The categories excluded from the analysis were removed because of their ambiguity.

A logistic regression was run for each of the 55 ICD-9 CM codes in HF in the same manner as the NHDS. Any patient that did not have at least one of the 55 ICD-9 CM codes was removed from the dataset and the same alpha (0.000909) was used. After removing patients that did not have at least one of the 55 ICD-9 CM codes there were 3,821,992 patients in the HF dataset used to analyze comorbidities.

CHAPTER 3

RESULTS

The NHDS dataset had a total of 2,944,459 observations with a weighted total of 386,186,183. Within the dataset there was a total of 1188 patients with VH and a weighted total of 162,859. The Health Facts data contained 46,721,119 total patients with 7488 vascular hamartoma patients. Figure 1 shows the flow of the analysis and table 1 contains the descriptive statistics for the demographic variables. An appropriate univariate analysis was run on each demographic.

Figure 1.—Analysis flowchart

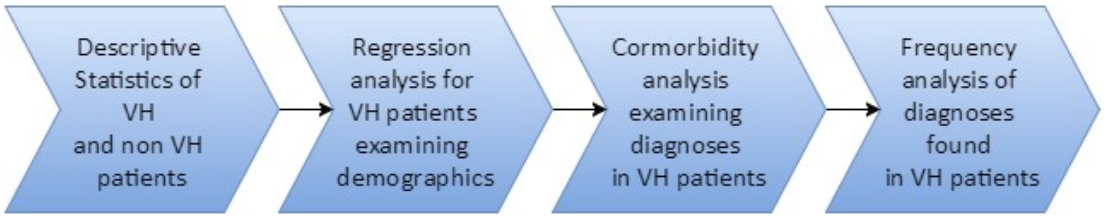

Table 2 contains the significant results from the demographic regressions. For the NHDS data age was significantly associated with VH (OR=0.90, CI=0.89-0.90, P<0.0005). Blacks when compared to whites were almost two times less likely to have VH (OR=0.52, CI=0.38-0.71, P<0.0005). Females were 29% more likely to have VH than males (OR=1.29, CI=1.04-1.62, P=0.024). Patients from the Midwest were

more than two times less likely to have VH than those from the South (OR=0.44, CI=0.29-0.67, P<0.0005). Northeast patients were 64% less likely to have VH (OR=0.61, CI=0.44-0.84, P=0.003) and patients from the West were 89% less likely to have VH than those from the south (OR=0.53, CI=0.38-0.73, P<0.0005).

In Health Facts age was a significant predictor (OR=0.92, CI=0.92-0.92, P<0.0005). Blacks were 67% less likely to have VH (OR=0.60, CI=0.56-0.65, P<0.0005), Asian/Pacific Islanders were 27% more likely to have VH (OR=1.27, CI=1.09-1.48, P=0.003) and the category other was 41% more likely to have VH than whites (OR=1.41, CI=1.31-1.52, P<0.001). Females were 28% more likely to have VH than males (OR=1.28, CI=1.22-1.34, P<0.0005). The regions Midwest (OR=0.65, CI=0.60-0.69, P<0.0005), Northeast (OR=0.70, CI=0.66-0.75, P<0.0005), and West (OR=0.60, CI=0.55-0.65, P<0.0005) were all less likely to have VH than the south.

Table 3 contains the significant comorbidities for the individual logistic regressions using the 55 ICD-9 CM codes. Diagnoses that were significant in both datasets were esophageal reflux, congenital pigmentary anomalies of skin, other hamartoses (not elsewhere classified), other injuries to scalp, respiratory distress syndrome in newborn, septicemia [sepsis] of newborn, cutaneous hemorrhage of fetus or newborn, hemolytic disease of fetus or newborn due to ABO isoimmunization, congenital hydrocele, other specified conditions involving the integument of fetus and newborn, and single liveborn (born in hospital) and delivered without mention of cesarean section. The raw results in both datasets

were similar, but certain diagnoses found only in children and specific genders were different. In table 3 and 4 the bold rows indicate ICD-9 CM codes that were significant in both datasets.

Table 1.—Descriptive statistics and univariate analysis for NHDS (2001-2010) and HF (2000-2014) datasets

Variable	No VH	VH	P-value
National Hospital Discharge Survey			
N (weighted)	2944459 (386186183)	1188 (162859)	
Age (years), mean (SD)	47.4 (323.5)	2.8 (80.0)	<.0005
Race			
White	233021568 (78.0%)	102815 (80.0%)	<.0005
Black	48042728 (16.1%)	14538 (11.3%)	
American Indian/Alaskan Native	1655827 (0.5%)	1219 (0.9%)	
Asian	6118568 (2.0%)	4203 (3.3%)	
Native Hawaiian/Other Pacific Islander	843152 (0.3%)	609 (0.5%)	
Other	9030105 (3.0%))	5268 (4.1%)	
Sex			
Male	159207399 (41.2%)	68444 (42.0%)	<.0005
Female	226815925 (58.8%)	94415 (56.0%)	
Region			
Northeast	80671716 (20.9%)	26827 (16.5%)	<.0005
South	145026956 (37.6%)	83202 (51.1%)	
Midwest	86170975 (22.3%)	22509 (13.8%)	
West	74153677 (19.2%)	30321 (18.6%)	
Health Facts			
N	46721119	7488	
Age (years), mean (SD)	45.4 (25.3)	12.3 (25.6)	<.0005
Race			
White	31945304 (72.2%)	4378 (63.7%)	<.0005
Black	7671539 (17.3%)	1042 (15.2%)	
Hispanic	1445333 (3.3%)	415 (6.0%)	
Asian/Pacific Islander	789567 (1.8%)	168 (2.4%)	
Other	2424068 (5.5%)	874 (12.7%)	
Sex			
Male	18951205 (40.6%)	3114 (41.6%)	0.067
Female	27757030 (59.4%)	4369 (58.4%)	
Region			
Northeast	18122064 (38.8%)	2358 (31.5%)	<.0005
South	11917102 (25.5%)	2511 (33.5%)	

Table 1.—Continued

Midwest	11129540 (23.8%)	1506 (20.2%)	
West	5508195 (11.8%)	1111 (14.8%)	

Table 2.—Logistic regression results for the demographics using NHDS and Health Facts data

Variable	Reference	Odds Ratio	CI
National Hospital Discharge Survey			
Age (years)		0.90	0.89-0.90
Race			
Black	White	0.52	0.38-0.71
Sex			
Female	Male	1.29	1.04-1.62
Region			
Midwest	South	0.44	0.29-0.67
Northeast	South	0.61	0.44-0.84
West	South	0.53	0.38-0.73
Health Facts			
Age (years)		0.92	0.92-0.92
Race			
Black	White	0.60	0.56-0.65
Asian/Pacific Islander	White	1.27	1.09-1.48
Other	White	1.41	1.31-1.52
Sex			
Female	Male	1.28	1.22-1.34
Region			
Midwest	South	0.65	0.60-0.69
Northeast	South	0.70	0.66-0.75
West	South	0.60	0.55-0.65

Table 3.—Regression analysis of comorbidities in vascular hamartoma patients from NHDS and Health Facts datasets

ICD-9 Codes	Label	Odds Ratios HF	Odds Ratios NHDS	P-values HF	P-values NHDS
216.1	Benign neoplasm of eyelid, including canthus	1.67	32.04	0.1468	<.00005
228.01	Hemangioma of skin and subcutaneous tissue	9.69	1.11	<.00005	0.9026
530.81	**Esophageal reflux**	**0.06**	**0.00**	**<.00005**	**<.00005**
745 .5	Ostium secundum type atrial septal defect	0.57	0.21	<.00005	0.0027
752.51	Undescended testis	0.52	0.99	<.00005	0.9871
757.33	**Congenital pigmentary anomalies of skin**	**6.56**	**7.61**	**<.00005**	**<.00005**
757.39	Other specified congenital anomalies of skin	2.32	3.19	<.00005	0.0255
759.6	**Other hamartoses, not elsewhere classified**	**24.32**	**8.15**	**<.00005**	**<.00005**
763.82	Abnormality in fetal heart rate or rhythm during labor	2.11	0.69	<.00005	0.5253
765.29	37 or more completed weeks of gestation	1.48	0.39	<.00005	0.0806
766.1	Other "heavy-for-dates" infants	2.24	1.29	<.00005	0.3505
766.21	Post-term infant	3.04	1.88	<.00005	0.3921
767.19	**Other injuries to scalp**	**3.80**	**5.91**	**<.00005**	**<.00005**
767.3	Other injuries to skeleton due to birth trauma	2.02	9.58	0.1624	0.0002
769	**Respiratory distress syndrome in newborn**	**0.39**	**0.02**	**<.00005**	**<.00005**
770.83	Cyanotic attacks of newborn	1.99	0.41	0.0001	0.2258
771.81	**Septicemia [sepsis] of newborn**	**0.30**	**0.06**	**<.00005**	**<.00005**
772.6	**Cutaneous hemorrhage of fetus or newborn**	**5.38**	**2.93**	**<.00005**	**0.0004**
773.1	**Hemolytic disease of fetus or newborn due to ABO isoimmunization**	**1.68**	**0.14**	**0.0001**	**<.00005**

Table 3.— Continued

774.2	Neonatal jaundice associated with preterm delivery	0.52	0.19	<.00005	0.0308
774.6	Unspecified fetal and neonatal jaundice	1.77	1.18	<.00005	0.3086
775 .0	Syndrome of 'infant of a diabetic mother'	1.62	1.18	<.00005	0.8068
778.6	**Congenital hydrocele**	**4.20**	**7.06**	**<.00005**	**<.00005**
778.8	**Other specified conditions involving the integument of fetus and newborn**	**4.17**	**6.01**	**<.00005**	**<.00005**
779.84	Meconium staining	2.32	0.19	<.00005	0.0033
779.89	Other specified conditions originating in the perinatal period	5.85	1.94	<.00005	0.0105
785 .2	Undiagnosed cardiac murmurs	0.68	1.02	<.00005	0.9588
910.0	Abrasion or friction burn of face, neck, and scalp except eye, without mention of infection	0.20	1.84	<.00005	0.1838
V05.3	Need for prophylactic vaccination and inoculation against viral hepatitis	2.38	0.83	<.00005	0.2206
V05.9	Need for prophylactic vaccination and inoculation against unspecified single disease	1.79	0.95	<.00005	0.9023
V20.1	Other healthy infant or child receiving care	0.92	0.00	0.7821	<.00005
V20.2	Routine infant or child health check	0.30	0.97	<.00005	0.9341
V29.0	Observation for suspected infectious condition	1.78	0.51	<.00005	0.1025
V30.00	**Single liveborn, born in hospital, delivered without mention of cesarean section**	**1.83**	**0.64**	**<.00005**	**0.0001**
V30.01	Single liveborn, born in hospital, delivered by cesarean section	1.47	0.76	<.00005	0.0459
V50.2	Routine or ritual circumcision	1.76	1.97	<.00005	0.1657

Table 3.—Continued

V64.05	Vaccination not carried out because of caregiver refusal	2.07	0.61	<.00005	0.3671
V72.19	Other examination of ears and hearing	0.57	0.74	0.0007	0.6112

Table 4.—Frequency and percent of total for ICD-9 CM codes in NHDS and HF databases

ICD-9 Codes	Label	Frequency HF	Frequency NHDS
216.1	Benign neoplasm of eyelid, including canthus	4161 (0.11%)	9817 (0.02%)
228.01	Hemangioma of skin and subcutaneous tissue	17493 (0.46%)	90712 (0.14%)
530.81	**Esophageal reflux**	**1777978 (46.52%)**	**22130623 (34.27%)**
745 .5	Ostium secundum type atrial septal defect	70822 (1.85%)	888613 (1.38%)
752.51	Undescended testis	24888 (0.65%)	192801 (0.3%)
757.33	**Congenital pigmentary anomalies of skin**	**18782 (0.49%)**	**404283 (0.63%)**
757.39	Other specified congenital anomalies of skin	13197 (0.35%)	106045 (0.16%)
759.6	**Other hamartoses, not elsewhere classified**	**3352 (0.09%)**	**33194 (0.05%)**
763.82	Abnormality in fetal heart rate or rhythm during labor	7928 (0.21%)	119382 (0.18%)
765.29	37 or more completed weeks of gestation	43683 (1.14%)	798554 (1.24%)
766.1	Other "heavy-for-dates" infants	38846 (1.02%)	2070341 (3.21%)
766.21	Post-term infant	26497 (0.69%)	404237 (0.63%)
767.19	**Other injuries to scalp**	**17622 (0.46%)**	**579325 (0.9%)**
767.3	Other injuries to skeleton due to birth trauma	561 (0.01%)	50401 (0.08%)
769	**Respiratory distress syndrome in newborn**	**21031 (0.55%)**	**1104559 (1.71%)**
770.83	Cyanotic attacks of newborn	4582 (0.12%)	130496 (0.2%)
771.81	**Septicemia [sepsis] of newborn**	**19690 (0.52%)**	**792410 (1.23%)**

Table 4.-- Continued

772.6	**Cutaneous hemorrhage of fetus or newborn**	**8768 (0.23%)**	**371994 (0.58%)**
773.1	**Hemolytic disease of fetus or newborn due to ABO isoimmunization**	**9123 (0.24%)**	**566385 (0.88%)**
774.2	Neonatal jaundice associated with preterm delivery	31325 (0.82%)	1787018 (2.77%)
774.6	Unspecified fetal and neonatal jaundice	157544 (4.12%)	5870438 (9.09%)
775 .0	Syndrome of 'infant of a diabetic mother'	11082 (0.29%)	414817 (0.64%)
778.6	**Congenital hydrocele**	**6282 (0.16%)**	**247432 (0.38%)**
778.8	**Other specified conditions involving the integument of fetus and newborn**	**23953 (0.63%)**	**574382 (0.89%)**
779.84	Meconium staining	11504 (0.3%)	197671 (0.31%)
779.89	Other specified conditions originating in the perinatal period	47429 (1.24%)	889323 (1.38%)
785 .2	Undiagnosed cardiac murmurs	175524 (4.59%)	849416 (1.32%)
910.0	Abrasion or friction burn of face, neck, and scalp except eye, without mention of infection	110840 (2.9%)	356546 (0.55%)
V05.3	Need for prophylactic vaccination and inoculation against viral hepatitis	307374 (8.04%)	12203956 (18.9%)
V05.9	Need for prophylactic vaccination and inoculation against unspecified single disease	53597 (1.4%)	880319 (1.36%)
V20.1	Other healthy infant or child receiving care	3381 (0.09%)	13143 (0.02%)
V20.2	Routine infant or child health check	863498 (22.59%)	30545 (0.05%)
V29.0	Observation for suspected infectious condition	65103 (1.7%)	2923442 (4.53%)
V30.00	**Single liveborn, born in hospital, delivered without mention of cesarean section**	**496243 (12.98%)**	**27400589 (42.43%)**
V30.01	Single liveborn, born in hospital, delivered by cesarean section	217060 (5.68%)	10808074 (16.74%)

Table 4.-- Continued

V50.2	Routine or ritual circumcision	38721 (1.01%)	344694 (0.53%)
V64.05	Vaccination not carried out because of caregiver refusal	9821 (0.26%)	191205 (0.3%)
V72.19	Other examination of ears and hearing	22286 (0.58%)	818958 (1.27%)

CHAPTER 4

DISCUSSION

The NHDS and HF data provided a population wide look into vascular hamartoma patients. Previous clinical data has shown occurrences in patients anywhere between 0.3-5% (Darrow et al., 2015; Shirley et al., 2013). The data from the NHDS has an occurrence of ~0.04% and Health Facts has an occurrence of ~0.02%. A possible reason for the low incidence rate is differential diagnoses. Birthmarks cover a wide variety of possible diagnoses with many having a different ICD-9 CM code (American Academy of Pediatrics, 2009). With the physical similarities between the different birthmarks, there is a possibility of misdiagnosis or miscoding. Strawberry nevi tend to resolve on their own and wouldn't be documented in later visits. Also, older patients returning to the hospital are not likely to be re-diagnosed with a port-wine stain.

The demographic characteristics in the NHDS and Health Facts typically reflect previous clinical data. Age was expected to be significantly different since a majority of cases of vascular hamartomas are diagnosed in infancy. Previous clinical data show that females are significantly more likely to have strawberry nevi and that port-wine stains and birthmarks affect both sexes evenly (Aderibigbe et al., 2015; J. Nelson, n.d.). Females are 29% more likely to have VH in the NHDS and are 28% more likely in HF. The increased likelihood of VH in females agrees with previous clinical data but 28-29% is not consistent with the 1.4-3.0 time increase seen

clinically in female strawberry nevi patients. The differences between this study and previous clinical data may be related to port-wine stains affecting both sexes equally.

The regression analysis shows that blacks are almost two times less likely in the NHDS dataset and 67% less likely in the Health Facts dataset to have VH documented compared to white patients. In the Health Facts dataset Asian/Pacific Islanders were 27% more likely to have VH documented in their medical record compared to whites. Further analysis should be done to determine if Asians/Pacific Islanders are more likely to have VH. Previous studies have shown that port-wine stains and strawberry nevi were more common among whites than blacks (Amrock & Weitzman, 2013; Antaya, 2014).

The datasets, NHDS and Health Facts, found similar racial tendencies, but disagreed in some areas. These differences could be attributed to the relative standard error (RSE) in the NHDS for people under 15 years of age. RSE is a measure of the sampling variability in the NHDS. Since the majority of patients with VH were less than 15 years of age, the RSE could be as high as 30% for whites and blacks. With a lack of understanding for the cause of vascular hamartomas, it is difficult to determine why race is significantly associated. With the cause of port-wines stains being traced to the GNAQ gene perhaps there is an underlining genetic disposition (Shirley et al., 2013).

There are few studies on the regional distribution of VH. The NHDS and Health Facts show that Midwest, West, and Northeast patients are all significantly less associated with VH than patients from the south. The NHDS and Health Facts are similar in this statistic with the NHDS showing a more negative association. It is worth noting that all regions in the NHDS had a RSE of at least 20%. The West had RSE of 35%, which is higher than the 30% recommended by the NHDS. If the cause of most vascular hamartomas are genetic, the regional distribution could easily be affected. The quality of health care is lower for southern states which might contribute to the increase of vascular hamartoma patients (U.S. Department of Health and Human Services, 2015).

There were a variety of comorbidities found to be associated with VH. Most comorbidities positively associated with VH in both datasets can be broken down into three main categories: hemorrhaging/vascular, skin issues, and birth trauma. Other hamartomas, not elsewhere classified and cutaneous hemorrhage of fetus or newborn both show possible vascular issues (World Health Organization, 2015a, 2015b). Congenital pigmentary anomalies of skin and other specified conditions involving the integument of fetus and newborn both involve diagnoses that affect the skin (World Health Organization, 2015b, 2015d). Other injuries to skeleton due to birth trauma is an example of birth trauma (World Health Organization, 2015c).

Overdevelopment of the fetus' physiology might connect the positively associated diagnoses. Overdevelopment is backed by the ICD-9 CM codes other

“heavy-for-dates” infants and post-term infants having an increased likelihood in both datasets and being significantly associated in Health Facts. This could mean that VH are linked to a growth factor in early development. One thing to note is that many of the skin conditions are possible differential diagnoses (A. Nelson et al., 2007; Vorvick, 2015).

The negatively associated comorbidities may have under development in common. This is supported by the ICD-9 CM code other preterm infants, 1,750-1,999 grams being less associated in both datasets but not significant. Esophageal reflux, respiratory distress syndrome (RDS), undescended testicle, and ostium secundum type atrial septal defect are diagnoses that are caused by under development of a physiological feature. Esophageal reflux is caused by the lower esophageal sphincter not closing properly and RDS is caused by a lack of surfactant (a slippery substance in the lungs that aids in keeping the air sacs from deflating) (Lee, 2013; Subodh, 2015). Ostium secundum type atrial septal defect is an opening in the wall of tissue that separates the left and right atria and is typically caused by an inadequate formation of the septum secundum (one of two major septal structures involved in partitioning the atrium) (Gessner, 2016). A few negatively associated comorbidities are also associated with infants being premature. Neonatal jaundice and septicemia [sepsis] of newborn are examples of diseases that occur more often in preterm infants (Hansen, Windle, & Carter, 2015; Lee, 2015).

Three results, congenital hydrocele, single liveborn (born in hospital) delivered without mention of cesarean section, and hemolytic disease of fetus or newborn due to ABO isoimmunization, don't follow the pattern seen in this analysis or contain opposite associations for each dataset. Congenital hydrocele (778.6) contains a higher prevalence in preterm infants and is positively associated with VH. This is contrary to most results in this analysis. Congenital hydrocele is believed to be related to the amount of smooth muscle in the processus vaginalis and is six times more common in males (Ortenberg & Roth, 2014). Single liveborn (born in hospital) delivered without mention of cesarean section (V30.00) and hemolytic disease of fetus or newborn due to ABO isoimmunization (773.1) was significant in both datasets but positively associated in HF and negatively associated in the NHDS. The difference could be related to the RSE in the NHDS data. For this dataset the RSE for patients under 15 years of age is abnormally high and thus might affect the results. HF doesn't rely on weighted estimates and is likely more reliable.

The limitations with this research are related to the NHDS dataset. One limitation is that the NHDS does not recommend using data with less than 30 unweighted observations. This was unavoidable during the comorbidity analysis. This would make the RSE for some of the comorbidities very high. The large RSE for the data increases the chance for a type I and type II error. Health Facts on the other hand does not suffer from RSE issues and many of the significant comorbidities found in the NHDS were also found in Health Facts. The use of ICD-9 CM codes are

also a limitation. Many ICD-9 CM codes cover quite a few diseases. This limits the precision of the analysis and what can be learned. The NHDS and HF are also completely reliant on physicians diagnosing correctly with no way to verify accuracy.

Future studies, examining comorbidities in vascular hamartoma patients, should involve a more precise dataset that provides more granularity for the comorbidities. More granularity in the dataset could make connections more obvious and provide more details on possible mechanisms. While the use of ICD-10 codes isn't prevalent in U.S. hospitals, ICD-10 codes could provide a more detailed look at the comorbidities. Also if a large enough hospital system is available the EHR might prove beneficial for this study. Discovering the cause and effect of each comorbidity might also prove beneficial. Ascertaining from the data if VH is the cause of the comorbidity or if the comorbidity caused a vascular hamartoma is difficult. The number of diagnoses in vascular hamartoma patients compared to all other patients and the way physicians typically assign ICD-9 CM codes could be beneficial information. Future studies could also look into other ICD-9 CM codes since the methods in this research could apply to any code.

REFERENCE LIST

Aderibigbe, O., Treat, J., & Maguiness, S. (2015, March 1). Infantile Hemangiomas and Vascular Malformations Basic Dermatology Curriculum. The Society for Pediatric Dermatology. Retrieved from https://www.aad.org/education/basic-derm-curriculum/suggested-order-of-modules/infantile-hemangiomas-and-vascular-malformations

American Academy of Pediatrics. (2009). Birthmarks & Hemangiomas. Retrieved January 19, 2016, from http://www.healthychildren.org/English/health-issues/conditions/skin/pages/Birthmarks-Hemangiomas.aspx

Amrock, S. M., & Weitzman, M. (2013). Diverging Racial Trends in Neonatal Infantile Hemangioma Diagnoses, 1979-2006. *Pediatric Dermatology*, *30*(4), 493–494. http://doi.org/10.1111/pde.12075

Antaya, R. (2014). Capillary Malformation: Background, Pathophysiology, Epidemiology. Retrieved from http://emedicine.medscape.com/article/1084479-overview

Berman, K. (2015, April 14). Port-wine stain: MedlinePlus Medical Encyclopedia. Retrieved January 26, 2016, from https://www.nlm.nih.gov/medlineplus/ency/article/001475.htm

Bush, W. S., & Moore, J. H. (2012). Chapter 11: Genome-Wide Association Studies. *PLoS Comput Biol*, *8*(12), e1002822. http://doi.org/10.1371/journal.pcbi.1002822

Center for Disease Control. (2015, November 6). About the Hospital Care Surveys. Retrieved January 19, 2016, from http://www.cdc.gov/nchs/nhds/about_nhds.htm

Cleveland Clinic. (2013, February 6). Birthmarks. Retrieved January 22, 2016, from http://my.clevelandclinic.org/health/diseases_conditions/hic_Birthmarks

Cleveland Clinic. (n.d.). Klippel-Trenaunay Syndrome: Symptoms, Treatment and Therapy Options. Retrieved September 22, 2014, from http://my.clevelandclinic.org/services/heart/disorders/vs_klippel-trenaunay_syndrome

Darrow, D. H., Greene, A. K., Mancini, A. J., & Nopper, A. J. (2015). Diagnosis and Management of Infantile Hemangioma: Executive Summary. *Pediatrics*, *136*(4), 786–791. http://doi.org/10.1542/peds.2015-2482

Dennison, C., & Pokras, R. (2000). Design and Operation of the National Hospital Discharge Survey: 1988 Redesign.

Eerola, I., Boon, L. M., Mulliken, J. B., Burrows, P. E., Dompmartin, A., Watanabe, S., … Vikkula, M. (2003). Capillary malformation-arteriovenous malformation, a new clinical and genetic disorder caused by RASA1 mutations. *American Journal of Human Genetics*, *73*(6), 1240–1249. http://doi.org/10.1086/379793

Gessner, I. (2016). Ostium Secundum Atrial Septal Defects: Background, Pathophysiology, Prognosis. Retrieved from http://emedicine.medscape.com/article/890991-overview

Hansen, T., Windle, M., & Carter, B. (2015). Neonatal Jaundice: Background, Pathophysiology, Epidemiology. Retrieved from http://emedicine.medscape.com/article/974786-overview

Lee, K. G. (2013, April 26). Neonatal respiratory distress syndrome: MedlinePlus Medical Encyclopedia. Retrieved February 2, 2016, from https://www.nlm.nih.gov/medlineplus/ency/article/001563.htm

Lee, K. G. (2015, April 27). Neonatal sepsis: MedlinePlus Medical Encyclopedia. Retrieved February 2, 2016, from https://www.nlm.nih.gov/medlineplus/ency/article/007303.htm

National Center for Health Statistics. (2010). National Hospital Discharge Survey. U.S. Department of Health and Human Services. Retrieved from ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHDS/NHDS_2010_Documentation.pdf

Nelson, A., Ulfgren, A.-K., Edner, J., Ståbi, B., Brismar, H., Hultenby, K., & Marchini, G. (2007). Urticaria Neonatorum: accumulation of tryptase-expressing mast cells in the skin lesions of newborns with Erythema Toxicum. *Pediatric Allergy and Immunology: Official Publication of the European Society of*

Pediatric Allergy and Immunology, *18*(8), 652–658. http://doi.org/10.1111/j.1399-3038.2007.00568.x

Nelson, J. (n.d.). Port Wine Stain [Organization]. Retrieved February 7, 2015, from https://birthmark.org/node/23

Neufeld, M., Maclaren, N. K., & Blizzard, R. M. (1981). Two types of autoimmune Addison's disease associated with different polyglandular autoimmune (PGA) syndromes. *Medicine*, *60*(5), 355–362.

Ortenberg, J., & Roth, C. (2014). Pediatric Hydrocele and Hernia Surgery: Background, Pathophysiology, Epidemiology. Retrieved from http://emedicine.medscape.com/article/1015147-overview

Shirley, M. D., Tang, H., Gallione, C. J., Baugher, J. D., Frelin, L. P., Cohen, B., … Pevsner, J. (2013). Sturge-Weber syndrome and port-wine stains caused by somatic mutation in GNAQ. *The New England Journal of Medicine*, *368*(21), 1971–1979. http://doi.org/10.1056/NEJMoa1213507

Subodh, L. (2015, January 22). Gastroesophageal reflux disease: MedlinePlus Medical Encyclopedia. Retrieved February 2, 2016, from https://www.nlm.nih.gov/medlineplus/ency/article/000265.htm

U.S. Department of Health and Human Services. (2015, August). State Snapshots [Text]. Retrieved February 8, 2016, from http://www.ahrq.gov/research/data/state-snapshots/index.html

U.S. National Library of Medicine. (2014, February 5). Birthmarks [Text]. Retrieved January 22, 2016, from https://www.nlm.nih.gov/medlineplus/birthmarks.html

Vorvick, L. (2015, April 11). Bleeding into the skin: MedlinePlus Medical Encyclopedia. Retrieved February 9, 2016, from https://www.nlm.nih.gov/medlineplus/ency/article/003235.htm

World Health Organization. (2015a). Cutaneous hemorrhage of fetus or newborn. In *International statistical classification of diseases and related health problems (9th ed.)*. Retrieved from http://www.icd9data.com/2015/Volume1/760-779/764-779/772/772.6.htm

World Health Organization. (2015b). Other hamartoses, not elsewhere classified. In *International statistical classification of diseases and related health problems (9th ed.)*. Retrieved from http://www.icd9data.com/2012/Volume1/740-759/759/759.6.htm

World Health Organization. (2015c). Other injuries to skeleton due to birth trauma. In *International statistical classification of diseases and related health problems (9th ed.)*. Retrieved from http://www.icd9data.com/2015/Volume1/760-779/764-779/767/767.3.htm

World Health Organization. (2015d). Other specified anomalies of skin. In *International statistical classification of diseases and related health problems*

(9th ed.). Retrieved from http://www.icd9data.com/2015/Volume1/740-759/757/757.39.htm

World Health Organization. (2015e). Vascular hamartomas. In *International statistical classification of diseases and related health problems (9th ed.)*. Retrieved from http://www.icd9data.com/2015/Volume1/740-759/757/757.32.htm

VITA

Michel Conn was born on May 26, 1989 in Saint Joseph, Missouri. He was educated at Saint Joseph Christian School and later moved to the Avenue City School district until high school. He graduated from Lafayette High school in the top twenty in his class. He received the Golden Griffon scholarship from Missouri Western State University (MWSU) in Saint Joseph Missouri. The Golden Griffon is the highest scholarship awarded by MWSU. He graduated in 2012 from MWSU magna cum laude with a Bachelor of Science in Biotechnology and a minor in chemistry. At MWSU he was a five time recipient of the President's list and two time recipient of the Dean's list. At MWSU he also participated in research that went on to be published in the Interdisciplinary Bio Central Journal.

After college he worked at Cenetra Scientific as the scientific lead in the quality control laboratory and helped maintain documentation standards.

In 2014 he started attending the University of Missouri-Kansas City with plans to graduate May 2016. After college he plans to pursue a career in database research.