



**AFRL-SA-WP-SR-2017-0001**

# **Population Spotting Using “Big Data”: Validating the Human Performance Concept of Operations Analytic Vision**



**Jennifer Geffre**

*The Perduco Group*

**Col Anthony Tvaryanas**

*711<sup>th</sup> Human Performance Wing*



**January 2017**

**Final Report**

**for April 2015 to July 2016**

**DISTRIBUTION STATEMENT A. Approved  
for public release. Distribution is unlimited.**

STINFO COPY

**Air Force Research Laboratory  
711<sup>th</sup> Human Performance Wing  
U.S. Air Force School of Aerospace Medicine  
Aeromedical Research Department  
2510 Fifth St., Bldg. 840  
Wright-Patterson AFB, OH 45433-7913**

# NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

Qualified requestors may obtain copies of this report from the Defense Technical Information Center (DTIC) (<http://www.dtic.mil>).

AFRL-SA-WP-SR-2017-0001 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNATURE//

---

MS. JULIA PARAKKAT  
Chief, Oper Health & Performance Res Div

//SIGNATURE//

---

DR. RICHARD A. HERSACK  
Chair, Aeromedical Research Department

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

<b>REPORT DOCUMENTATION PAGE</b>			<i>Form Approved</i> <i>OMB No. 0704-0188</i>		
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. <b>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</b>					
<b>1. REPORT DATE (DD-MM-YYYY)</b> 5 Jan 2017		<b>2. REPORT TYPE</b> Special Report		<b>3. DATES COVERED (From – To)</b> April 2015 – July 2016	
<b>4. TITLE AND SUBTITLE</b>  Population Spotting Using “Big Data”: Validating the Human Performance Concept of Operations Analytic Vision			<b>5a. CONTRACT NUMBER</b> FA8650-12-D-6280		
			<b>5b. GRANT NUMBER</b>		
			<b>5c. PROGRAM ELEMENT NUMBER</b>		
<b>6. AUTHOR(S)</b> Jennifer Geffre, Col Anthony Tvaryanas			<b>5d. PROJECT NUMBER</b>		
			<b>5e. TASK NUMBER</b> 0031		
			<b>5f. WORK UNIT NUMBER</b>		
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> Prime Contractor: Wyle Laboratories, Inc. 1290 Hercules Houston, TX 77508			Subcontractor: The Perduco Group 3610 Pentagon Blvd, Suite 110 Beavercreek, OH 45432		
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> USAF School of Aerospace Medicine Aeromedical Research Dept/FHO 2510 Fifth St., Bldg. 840 Wright-Patterson AFB, OH 45433-7913			<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>		
			<b>10. SPONSORING/MONITOR'S ACRONYM(S)</b>		
<b>12. DISTRIBUTION / AVAILABILITY STATEMENT</b>  DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited.			<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b> AFRL-SA-WP-SR-2017-0001		
			<b>13. SUPPLEMENTARY NOTES</b> Cleared, 88PA, Case # 2017-0083, 10 Jan 2017.		
<b>14. ABSTRACT</b> This report describes the research into a methodology for determining patient segments consistent with the Air Force Surgeon General's vision for patient care. Patients' health history focuses on the conditions and frequency of visits required to meet their healthcare needs. The methodology consists of a two-part process where patients are first clustered based on the likeness to other patients. Then clusters are aggregated based on similarity and constrained by the availability of providers and resources. The result is patient segments, which can be better serviced to meet healthcare needs.					
<b>15. SUBJECT TERMS</b> Patient segmentation, data mining, cluster analysis, patient hot spotting					
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>	<b>18. NUMBER OF PAGES</b>	<b>19a. NAME OF RESPONSIBLE PERSON</b> William Bell
<b>a. REPORT</b> U	<b>b. ABSTRACT</b> U	<b>c. THIS PAGE</b> U			SAR

*This page intentionally left blank.*

# TABLE OF CONTENTS

Section	Page
LIST OF FIGURES .....	ii
LIST OF TABLES .....	ii
1.0 SUMMARY .....	1
2.0 INTRODUCTION .....	1
3.0 METHODS .....	2
3.1 Study Data .....	2
3.1.1 Patient Data .....	2
3.1.2 Clinic Selection .....	2
3.1.3 Patient Selection .....	2
3.1.4 Diagnoses Codes .....	3
3.1.5 Data Preparation for Modeling .....	3
3.2 Patient Segmentation Modeling .....	3
3.2.1 Cluster Analysis .....	3
3.2.2 Staffing Model .....	4
4.0 RESULTS .....	7
4.1 Clinic Demographics .....	7
4.2 Patient Segmentation Modeling Results .....	7
4.2.1 Large Clinics .....	7
4.2.2 Small Clinics .....	8
4.3 Model Testing .....	8
5.0 CONCLUSIONS .....	9
6.0 REFERENCES .....	10
APPENDIX – Supporting Materials .....	11
LIST OF ABBREVIATIONS AND ACRONYMS .....	14

## LIST OF FIGURES

	<b>Page</b>
Figure 1. Diagnoses distribution (all clinics).....	3
Figure 2. CLARA process.....	4
Figure 3. Staffing model process. ....	5
Figure 4. Pseudo code for staffing model. ....	6

## LIST OF TABLES

	<b>Page</b>
Table 1. Summary Clinic Demographics (2012) .....	7
Table 2. Clinic A Summary Results .....	8
Table 3. Clinic B Summary Results.....	8
Table A-1. ICD Codes Used in Analysis .....	11
Table A-2. Cluster and Staffing Model Results (Base A, 3 years, Tolerance = 0.05).....	12
Table A-3. Cluster and Staffing Model Results (Base B, 3 years, Tolerance = 0.05).....	13

## 1.0 SUMMARY

The Air Force Surgeon General established the strategic goal for the Air Force Medical Service (AFMS) that its supported population become the healthiest and highest performing segment of the United States. Achieving this strategic goal necessitates that the AFMS accurately segment its supported population into groupings that translate into differences in healthcare service delivery needs. These groups are the basis for building a more patient-centered healthcare system and determining what types of providers and services to place in different parts of the AFMS. The purpose of this study was to investigate the reuse of existing medical data to both explore modeling approaches and demonstrate the feasibility of segmenting patient populations around healthcare needs and formulating corresponding provider empanelment. This study used a preexisting dataset consisting of healthcare services delivered in permanent, continuously operating U.S. Air Force Flight and Operational Medicine Clinics between 2003 and 2012. Four representative clinics were chosen based on two considerations: size (large versus small medical treatment facilities) and location (within the contiguous United States versus abroad). Patients actively seen in 2012 at one of the four representative clinics were included in the study. A two-step methodology, combining a cluster model with a constrained staffing model, provided a tractable approach with utility at larger clinics where visit demand exceeded the capacity of a single provider. The approach worked best when data were aggregated at the individual versus family level and when the most recent 3 years of patient data were used for segmentation. Before operationalizing patient segmentation, subsequent studies must address the following issues: segment stability, generalizability, scalability, and new patient population members.

## 2.0 INTRODUCTION

The Air Force Surgeon General established the strategic goal for the Air Force Medical Service (AFMS) that its supported population become the healthiest and highest performing segment of the United States. To achieve this strategic goal, the AFMS is drawing on evidence that it can leverage the Patient Centered Medical Home to improve population health in addition to individual patient health [1,2]. A crucial step in enabling patient care teams to address population health is to give them “meaningful subpopulations” to manage. Targeted base-level populations are far from uniform. Rather, they are composed of distinct groups of individuals or subpopulations with similar healthcare needs that reflect individuals’ risk factors, conditions, the severity of those conditions, and access requirements. Critical to strategy execution is accurate segmentation of populations served into groupings that translate into differences in healthcare service delivery needs. Dividing populations into subpopulations not only enables patient care teams to better meet individuals’ needs, it also enables focused resource planning and delivery of appropriate preventive care services.

After segmenting supported populations at medical treatment facilities (MTFs), individual Patient Centered Medical Home teams can be tailored to more efficiently and effectively meet the healthcare needs of assigned patient segments. By grouping patients into segments, economics of scale are achievable that may justify provision of frequently needed specialty services within the patient care team. Additionally, higher patient volumes for certain medical conditions enable patient care teams to operate farther to the right on the classic learning curve, thereby improving both efficiency and quality [3]. Conceptually this approach dramatically alters the way in which the AFMS builds patient care teams – rather than matching

the patients to the care team based on provider capacity and business rules, the AFMS designs the care team as a potentially unique microsystem that best meets its patients' prevention, clinical care, and access needs.

The purpose of this study was to investigate the reuse of existing medical data to both explore modeling approaches and demonstrate the feasibility of segmenting patient populations around healthcare needs and formulating corresponding provider empanelment.

## 3.0 METHODS

### 3.1 Study Data

**3.1.1 Patient Data.** This study reused a preexisting dataset consisting of healthcare services delivered in permanent, continuously operating U.S. Air Force (USAF) Flight and Operational Medicine Clinics (FOMCs) between 2003 and 2012 [4]. The prior study obtained data from the Military Health System Data Mart on September 10, 2013. Variables used in this study included the following:

- Year of visit
- Diagnosis codes (first four) as defined by the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)
- Unique identifiers for:
  - Patient
  - Patient family
- FOMC identifier

Notably, specific health data (variables like weight, blood pressure, and pain rating) were not available in this dataset.

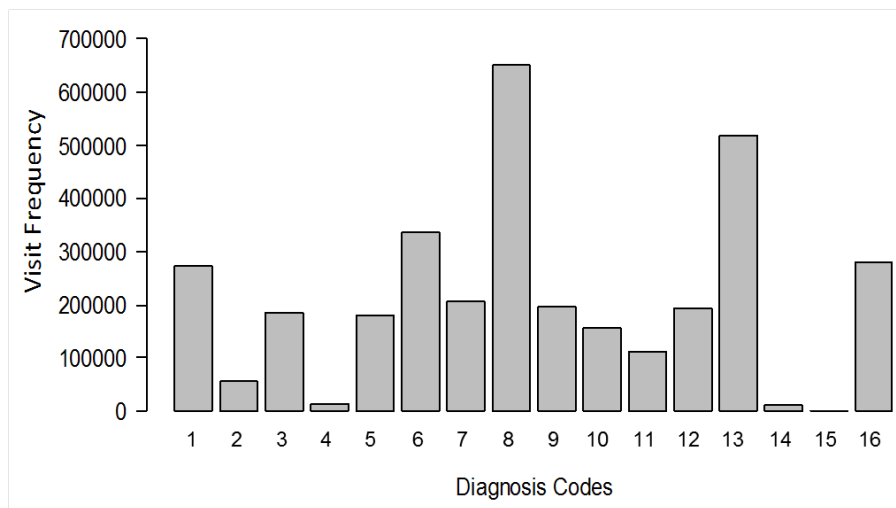
**3.1.2 Clinic Selection.** Four representative clinics were chosen from 79 existing USAF clinics to ensure the methodology was feasible to implement across all FOMCs. Clinic selection was based on two considerations: 1) *size* to account for small and large MTFs<sup>1</sup> and 2) *location* to recognize differences between MTFs within the contiguous United States (CONUS: lower 48 states) and those abroad (OCONUS: outside the United States, to include clinics in Alaska and Hawaii). The four MTFs consisted of (one each): large/CONUS, large/OCONUS, small/CONUS, and small/OCONUS.

**3.1.3 Patient Selection.** Patients actively seen in 2012 at one of the four representative clinics were included in the study. This selection criterion allowed the creation of a patient information baseline and testing various aspects of the amount of retrospective data included. Ultimately, we used the full dataset when aggregating patient data to gain the patients' history regardless of the clinic where they were treated.

---

<sup>1</sup>MTFs providing only outpatient services were classified as small compared to MTFs providing both outpatient and inpatient services (i.e., a hospital), which were classified as large.

**3.1.4 Diagnoses Codes.** Diagnosis codes were recoded using the Clinical Classification Software (CCS) for ICD-9-CM, which aids analysts in collapsing diagnostic data from over 14,000 diagnosis codes that make up the ICD-9-CM standardized coding system into clinically meaningful categories [5]. CCS level 1 categories consists of 16 codes for the major anatomic systems. Figure 1 shows the frequency of all patient visits distributed across the 16 CCS level 1 codes. Since several of these level 1 codes contained a high frequency of patient visits, these were decomposed into the level 2 codes to provide greater granularity about constituent diagnoses. Appendix Table A-1 provides all CCS-derived diagnosis (Dx) codes used in this study. Respiratory (Dx 8), musculoskeletal (Dx 13), nervous system (Dx 6), and infectious diseases (Dx 1) make up the four most prevalent conditions or injuries among the patient visits.



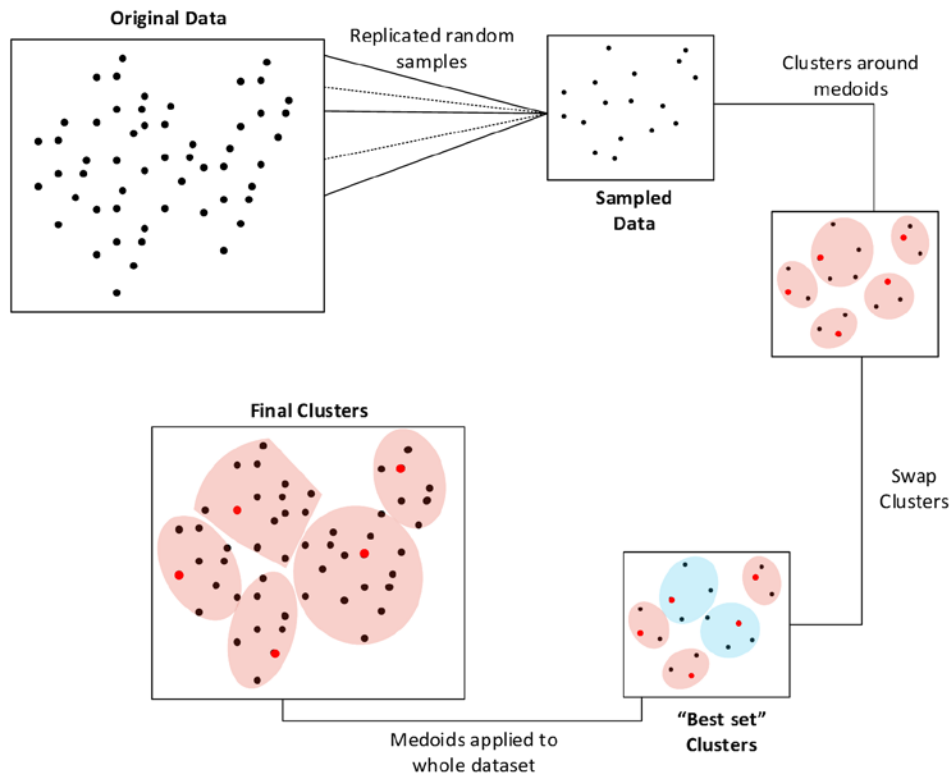
**Figure 1. Diagnoses distribution (all clinics).**

**3.1.5 Data Preparation for Modeling.** The first four Dx codes in each patient visit were translated into indicator variables (0/1) corresponding to the CCS Dx codes used in the study. In the study database, each visit was structured as the complete set of CCS Dx codes with each code having a value of zero or one. This approach allowed both tracking of the number of visits related to each Dx code and multiple codes if used in any single visit. For each patient, data were aggregated by summing each CCS Dx code indicator variable across visits. The resulting values reflect the patient’s visit frequency for each CCD Dx code.

## 3.2 Patient Segmentation Modeling

**3.2.1 Cluster Analysis.** This study used the CLARA (clustering large applications) algorithm [6], available in the cluster package [7] for the statistical software R [8]. The original dataset was randomly sampled, and the sampled data were split into clusters using the PAM (partitioning around medoids) algorithm [6]. The clusters were determined using medoids rather than centroids, as the former represents real patients in the dataset; these patients/medoids have the

smallest average dissimilarity from all other patients in their respective clusters. Patients furthest from medoids were swapped between clusters until a minimum total distance among all sampled members was identified. This process was replicated five times to ensure minimal sampling bias. The resulting lowest total distance was used to identify the “best set” of clusters. CLARA then assigned other patients outside the sample to the nearest medoid based on similarity (distance). Figure 2 depicts the CLARA process.



**Figure 2. CLARA process.**

**3.2.2 Staffing Model.** The staffing model combined like-patient segments while accounting for provider capacity. Figure 3 provides an illustrative representation of the process associated with the staffing model. Cluster results from the patient segmentation model were used as an input to the staffing model. An aggregate visit utilization demand was calculated for each cluster by summing visits across patients forming the cluster. The medoids of the clusters were used to calculate the dissimilarity (i.e., distance) between the other clusters. The minimum distance was identified, and only those cluster pairs with distances nearest the minimum were considered for aggregation; this threshold can be changed to allow more or fewer candidate pairs. The total dissimilarity was calculated to determine which aggregation should be made, thereby allowing for matching clusters that made the most improvement to the data as a whole. Once the aggregation was made, a new medoid was identified for the consolidated cluster and the cluster utilization was updated. Only aggregations with a combined visit utilization within the provider threshold were considered. This process was iterated until there were no other feasible

aggregations to be made. The resulting clusters were then analyzed to determine the performance success of the model (see section 3.4). The model assumed a provider annual capacity of 3,960 visits per provider with an allowed variance of 5% (up to 4,158 annual visits). Provider teams (versus single providers) were addressed by allowing for integer values of the provider threshold. Figure 4 shows the staffing model’s pseudo code.

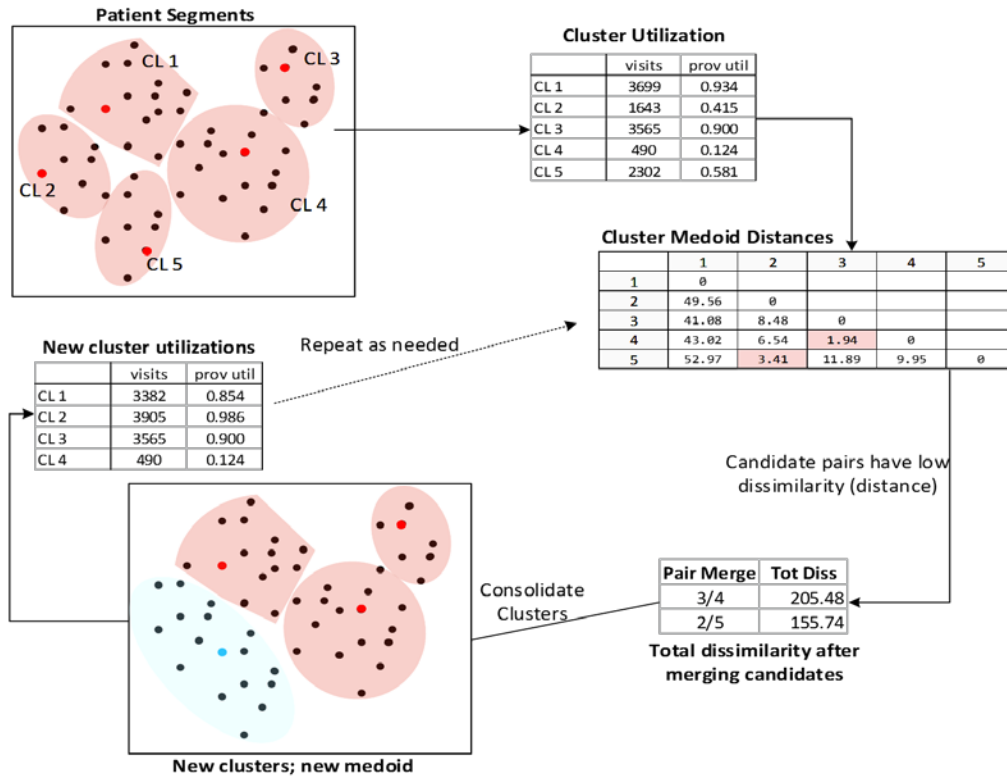


Figure 3. Staffing model process.

```

# Psuedo Code for Staffing Model

#Inputs: patitent assignments from cluster analysis
#      maximum annual visits (provider capacity)

#Variables: annualized visit tolerance

#Objective: aggregate patient clusters based on closeness,
# while keeping cluster utilization under the annualized visit threshold
(maximum annual visits * annualized visit tolerance)

#> Find cluster medoids
#> Calculate the average cluster utilization
#> Calculated the distance (Euclidean) between cluster medoids
#> Create and order a paired list (based on the distance pairs)
#> Add pair distance and combined pair cluster utilization
#> Remove cluster pairs where combined utilization > annualized visits
#> Loop to aggregate clusters
# > Loop: to find the best cluster replacement
# > Find potential cluster medoids and the distances between them
# > Total distance is the sum of all medoid distances
# > Keep the cluster replacement with the maximum total distance
# > Make best cluster replacement permanent, update all patient assignments
# > Recalculate new cluster medoids
# > Create new ordered pair list, with distance and combined utilization
# > Remove cluster pairs where combined utilization > annualized visits
# > Repeat until the ordered pair list is empty
#> Calculate segment summary statistics

#outputs: patient segment assignments
#      count of patients in each segments
#      segment utilization
#      number of patient visits by diagnosis code, for each segment

```

**Figure 4. Pseudo code for staffing model.**

**3.2.3 Model Testing.** The Family Practice medical specialty advocates assigning families rather than individual patients to providers. To understand the impact of this approach, the patient segmentation model was run using families as compared to patients as the unit of analysis.

Another consideration was the relative importance of recent versus more remote patient data in identifying healthcare needs. Keeping too much remote patient data tends to make patients appear more alike and obfuscates patient segmentation. However, keeping too little data biases segmentation to acute health needs. Presently, it is unclear how much patient history to include when performing patient segmentation. Accordingly, several timeframes (1, 3, and 5 years) were explored to determine the impact on clustering.

Finally, the staffing model assumed a provider annualized visit capacity of 3,960. Forcing the model to adhere strictly to this constraint does not reflect the flexible nature of provider availability. Flexibility was addressed in the model by using an adjustable tolerance factor that was set at 5%. A sensitivity analysis of the tolerance factor was performed by varying tolerance at five levels: 1%, 2.5%, 5%, 7.5%, and 10%.

## 4.0 RESULTS

### 4.1 Clinic Demographics

Table 1 provides descriptive data for the four representative FOMCs and associated patient populations examined in this study. Clinic identities were masked to preserve data de-identification.

**Table 1. Summary Clinic Demographics (2012)**

Clinic	Size, Location	Patients	Families	Ages (yr)	Active Duty (%)	Total Visits	Provider Utilization
A	Large, CONUS	7830	7455	10-81	48	14465	3.65
B	Large, OCONUS	5284	4045	Infant-76	45	13621	3.44
C	Small, CONUS	823	750	Infant-75	55	1906	0.48
D	Small, OCONUS	1030	847	2-62	47	3107	0.78

### 4.2 Patient Segmentation Modeling Results

This segmentation analysis used retrospective data for 2012 and the prior 2 years (total of 3 years of data). The annualized visit tolerance was set at 0.05.

**4.2.1 Large Clinics.** The supported patient populations for Clinics A and B were large enough to apply cluster and staffing models. Tables 2 and 3 provide summary results; complete tables of results are available in the Appendix (Tables A-2 and A-3). For each segment, “segment size” is the number of patients in the segment, “providers” is the minimum number of healthcare providers required to cover the annual visit demand of those patients, and “provider utilization” is the number of healthcare provider full-time equivalents needed to meet demand. “# Dx Match” is the number of CCS diagnosis codes for which  $\geq 50\%$  of the patients in the population with that diagnosis are included in the segment.

The patient population for Clinic A partitioned into three segments. The largest segment (A1) had a visit demand necessitating a two-provider team, and it comprised relatively healthy patients who visited the clinic approximately once annually. The next segment (A2) had a higher frequency of visits (3.5 visits per patient per year), but these patients were seen for only a few diagnoses. The last segment (A3) had both multiple medical conditions and the highest visit utilization rate (5.3 visits per patient per year).

**Table 2. Clinic A Summary Results**

Segment	Segment Size	Providers	Provider Utilization	# Dx Match
A1	6062	2	1.898	6
A2	988	1	0.881	9
A3	780	1	1.035	46

The patient population for Clinic B partitioned into four segments. Segment B1 had sufficient visit demand to necessitate a two-provider team; patients in this segment tended to have multiple conditions, indicating a greater likelihood for the need for care planning and coordination. The largest segment (B2) comprised healthy patients with few diagnoses and low visit demand. Segment B3 comprised patients with a few common conditions that drove an average visit demand of three visits per patient per year. Segment B4 comprised a small set of high utilizing patients (average 6.5 visits per patient per year) who aligned to a single diagnosis.

**Table 3. Clinic B Summary Results**

Segment	Segment Size	Providers	Provider Utilization	# Dx Match
B1	1055	2	1.68	44
B2	2712	1	0.884	2
B3	1358	1	1.095	7
B4	159	1	0.262	1

**4.2.2 Small Clinics.** The supported populations for Clinics C and D were too small to apply clustering and staffing models, as a single provider could meet visit demand. Nonetheless, patient segmentation could still be used for cohort management within the single empanelment.

### 4.3 Model Testing

Data aggregation at the family versus the individual level tended to make the units of analysis for clustering more similar, thereby decreasing the effectiveness of segmentation. Accordingly, data aggregated at the individual patient level are preferred to data aggregation by family.

Data aggregation at the individual patient level using 5 years of retrospective data produced the best segmentation, while using 3 years of data yielded almost equivalent results. Thus, use of 3 years of retrospective data is a potential tradeoff for decreased data storage and reduced computational capability relative to using 5 years of retrospective data.

The results of the staffing tolerance sensitivity analysis showed that tolerance levels of 5%, 7.5%, and 10% produced the same clusters. The 1% tolerance, and to a lesser extent the 2.5% tolerance, produced clusters that did not facilitate the efficient use of providers. Accordingly, the 5% tolerance level is recommended.

## 5.0 CONCLUSIONS

This study utilized a preexisting longitudinal dataset to begin exploration of approaches to and the feasibility of segmenting patient populations around healthcare needs and formulating corresponding provider empanelment. Although ascertainment of healthcare needs using this dataset was limited to diagnosis codes, it was nonetheless possible to identify patient segments. A two-step methodology, combining a cluster model with a constrained staffing model, provided a tractable approach with utility at larger clinics where visit demand exceeds the capacity of a single provider.

Before operationalizing patient segmentation, subsequent studies must address the following issues:

1. Segment Stability: A primary reason for segmenting patients is to design tailored microsystems optimized to addressing a specific set of health needs. In theory, a change in a patient's health status should drive a relook at segment assignment. However, continuity of care is also purported to be a factor in healthcare quality and patient satisfaction. Further research needs to determine the periodicity of segment assignments and the proper balancing with other considerations such as continuity of care.
2. Generalizability: It is unclear if results from this study are generalizable to the AFMS at large. The FOMC patient population is not necessarily representative of the larger AFMS population, which is more heterogeneous in terms of demographic factors and disease burden.
3. Scalability: Again, this study focused only on the FOMC, which serves a relatively small patient population as compared to other clinics, such as the Family Health Clinic. Consequently, it is uncertain how the two-step methodology used in this study scales to handle larger patient populations. There is little doubt the required computational resources will increase with a larger patient population. However, the modeling approach should be robust enough to handle larger populations, given clusters are derived from a sample of the total data.
4. New Patient Population Members: Introducing new patients (i.e., new military accessions, new spouses, and new children) without any health history creates a minor problem. The simplest solution to this problem is assigning new patients to a healthy patient segment until there is sufficient health history data that would drive assignment to another segment.

To improve the quality of patient care, we must look for opportunities for change to include better leveraging existing healthcare data. The AFMS has a lot of patient information, but it is not using it to understand patients. Patient segmentation is an important first step because it documents the fact that not all patients are homogeneous. It allows the building of a more patient-centered healthcare system and helps determine what types of providers and services to place in different parts of the AFMS. And as the AFMS becomes smarter about its patient segments, that understanding will inform strategy about how it makes its supported population the healthiest and highest performing population in the United States.

## 6.0 REFERENCES

1. Garg A, Jack B, Zuckerman B. Addressing the social determinants of health within the patient-centered medical home: lessons from pediatrics. *JAMA*. 2013; 309(19):2001–2002.
2. Garg A, Sandel M., Dworkin PH, Kahn RS, Zuckerman B. From medical home to health neighborhood: transforming the medical home into a community-based health neighborhood. *J Pediatr*. 2012; 160(4):535–536.e1.
3. Yelle LE. The learning curve: historical review and comprehensive survey. *Decision Sciences*. 1979; 10(2):302–328.
4. Tvaryanas AP, Maupin GM, Fouts BL. A 10-year cross-sectional analysis of Air Force flight and operational medicine clinic health care services. *Mil Med*. 2016; 181(5):424–433.
5. Agency for Healthcare Research and Quality, Healthcare Cost and Utilization Project. HCUP tools and software. Rockwell (MD): Agency for Healthcare Research and Quality; 2016. [Accessed 13 Nov 2016]. Available from [http://www.hcup-us.ahrq.gov/tools\\_software.jsp](http://www.hcup-us.ahrq.gov/tools_software.jsp).
6. Kaufman L, Rousseeuw PJ. Finding groups in data: an introduction to cluster analysis. New York (NY): Wiley; 1990.
7. Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. Cluster analysis basics and extensions. R package version 2.0.1. 2015.
8. R Core Team. Clustering large applications. In: *R: a language and environment for statistical computing*, version 3.3.1. 2016:2385–2387. [Accessed 12 Sep 2016]. Available from <https://cran.r-project.org/doc/manuals/r-release/fullrefman.pdf>.

## APPENDIX – Supporting Materials

**Table A-1. ICD Codes Used in Analysis**

Dx.1	Infectious Diseases			Dx.9	Diseases of Digestive System		
	Dx.1.1	Bacterial Infection			Dx.9.1	Intestinal Infection	
	Dx.1.2	Mycoses			Dx.9.2	Disorders of Teeth/Jaw	
	Dx.1.3	Viral Infection			Dx.9.3	Diseases of Mouth	
	Dx.1.4	Other Infections			Dx.9.4	Upper Gastrointestinal Disorders	
	Dx.1.5	Immunization & Screening			Dx.9.5	Abdominal Hernia	
Dx.2	Neoplasms				Dx.9.6	Lower Gastrointestinal Disorders	
Dx.3	Endocrine				Dx.9.7	Biliary Tract Disease	
	Dx.3.1	Thyroid Disorder			Dx.9.8	Liver Disease	
	Dx.3.2	Diabetes Mellitus w/o Complication			Dx.9.9	Pancreatic Disorder (not diabetes)	
	Dx.3.3	Diabetes Mellitus w/ Complication			Dx.9.10	Gastrointestinal Hemorrhage	
	Dx.3.4	Other Endocrine Disorders			Dx.9.11	Noninfectious Gastroenteritis	
	Dx.3.5	Nutritional Deficiencies			Dx.9.12	Other Gastrointestinal	
	Dx.3.6	Disorders of Lipid Metabolism		Dx.10	Diseases of Genitourinary System		
	Dx.3.7	Gout		Dx.11	Complications of pregnancy		
	Dx.3.8	Fluid & Electrolyte Disorders		Dx.12	Diseases of Skin		
	Dx.3.9	Cystic Fibrosis		Dx.13	Diseases of Musculoskeletal System		
	Dx.3.10	Immunity Disorders			Dx.13.1	Infective Arthritis	
	Dx.3.11	Other Nutritional			Dx.13.2	Non-traumatic Joint Disorder	
Dx.4	Disease of Blood				Dx.13.3	Spondylosis	
Dx.5	Mental Illness				Dx.13.4	Osteoporosis	
Dx.6	Diseases of Nervous System				Dx.13.5	Pathological Fracture	
	Dx.6.1	Infection of Central Nervous System			Dx.13.6	Acquired Deformities	
	Dx.6.2	Hereditary Nervous System Condition			Dx.13.7	Systemic Lupus	
	Dx.6.3	Paralysis			Dx.13.8	Other Connective Tissue	
	Dx.6.4	Epilepsy			Dx.13.9	Other Bone/Musculoskeletal	
	Dx.6.5	Headache/Migraine		Dx.14	Congenital anomalies		
	Dx.6.6	Coma/Brain Damage		Dx.15	Conditions during Perinatal Period		
	Dx.6.7	Eye Disorder		Dx.16	Injury & Poisoning		
	Dx.6.8	Ear Condition			Dx.16.1	Joint disorder/Dislocations (trauma)	
	Dx.6.9	Other Nervous System			Dx.16.2	Fractures	
Dx.7	Diseases of Circulatory System				Dx.16.3	Spinal Cord Injury	
	Dx.7.1	Hypertension			Dx.16.4	Intracranial Injury	
	Dx.7.2	Diseases of the Heart			Dx.16.5	Crushing/Internal Injury	
	Dx.7.3	Cerebrovascular			Dx.16.6	Open wounds	
	Dx.7.4	Diseases of the Arteries			Dx.16.7	Sprains/Strains	
	Dx.7.5	Diseases of Veins/Lymphatic			Dx.16.8	Superficial Injury/Contusion	
Dx.8	Diseases of Respiratory System				Dx.16.9	Burns	
	Dx.8.1	Respiratory Infection			Dx.16.10	Complications	
	Dx.8.2	COPD/Bronchiectasis			Dx.16.11	Poisoning	
	Dx.8.3	Asthma			Dx.16.12	Other Injuries (external causes)	
	Dx.8.4	Aspiration Pneumonitis					
	Dx.8.5	Pleurisy/Pulmonary Collapse					
	Dx.8.6	Respiratory Failure					
	Dx.8.7	Lung Disease (external agents)					
	Dx.8.8	Other Lower Respiratory Disease					
	Dx.8.9	Other Upper Respiratory Disease					

**Table A-2. Cluster and Staffing Model Results (Base A, 3 years, Tolerance = 0.05)**

Raw Results Data																					
	size	util	prov_util	Dx.1.1	Dx.1.2	Dx.1.3	Dx.1.4	Dx.1.5	Dx.2	Dx.3.1	Dx.3.2	Dx.3.3	Dx.3.4	Dx.3.5	Dx.3.6	Dx.3.7	Dx.3.8	Dx.3.9	Dx.3.10	Dx.3.11	
1	6062	7518	1.898	1	51	42	4	43	48	4	13	0	6	0	85	1	5	0	0	45	
2	988	3490	0.881	17	82	87	4	33	146	13	7	7	9	3	147	16	9	2	0	70	
7	780	4100	1.035	21	112	296	4	28	200	170	29	8	52	11	292	24	14	0	0	145	
	Dx.4	Dx.5	Dx.6.1	Dx.6.2	Dx.6.3	Dx.6.4	Dx.6.5	Dx.6.6	Dx.6.7	Dx.6.8	Dx.6.9	Dx.7.1	Dx.7.2	Dx.7.3	Dx.7.4	Dx.7.5	Dx.8.1	Dx.8.2	Dx.8.3	Dx.8.4	Dx.8.5
1	43	210	1	0	1	9	34	1	87	0	17	63	41	3	87	20	0	10	0	0	2
2	11	183	2	6	1	0	63	0	133	161	60	36	45	0	102	38	924	32	0	0	0
7	63	583	0	11	0	3	143	3	192	377	71	315	114	0	132	75	738	46	0	0	2
	Dx.8.6	Dx.8.7	Dx.8.8	Dx.8.9	Dx.9.1	Dx.9.2	Dx.9.3	Dx.9.4	Dx.9.5	Dx.9.6	Dx.9.7	Dx.9.8	Dx.9.9	Dx.9.10	Dx.9.11	Dx.9.12	Dx.10	Dx.11	Dx.12	Dx.13.1	Dx.13.2
1	0	0	26	106	9	6	1	32	4	10	1	2	0	4	27	28	86	83	0	0	0
2	0	0	81	267	18	23	4	92	13	20	0	2	0	20	80	60	96	163	444	0	745
7	0	1	172	475	20	18	17	126	8	28	6	20	5	15	75	152	693	234	555	3	525
	Dx.13.3	Dx.13.4	Dx.13.5	Dx.13.6	Dx.13.7	Dx.13.8	Dx.13.9	Dx.14	Dx.15	Dx.16.1	Dx.16.2	Dx.16.3	Dx.16.4	Dx.16.5	Dx.16.6	Dx.16.7	Dx.16.8	Dx.16.9	Dx.16.10	Dx.16.11	Dx.16.12
1	106	0	0	14	0	99	23	8	0	8	38	1	11	1	10	111	30	1	0	1	29
2	199	2	0	13	0	261	31	12	0	50	57	0	5	1	19	189	58	5	2	1	96
7	811	3	0	22	3	360	94	29	0	60	45	0	0	4	20	335	43	1	9	6	67
Proportion across Segments (by Diagnosis code)																					
	util	prov_util	Dx.1.1	Dx.1.2	Dx.1.3	Dx.1.4	Dx.1.5	Dx.2	Dx.3.1	Dx.3.2	Dx.3.3	Dx.3.4	Dx.3.5	Dx.3.6	Dx.3.7	Dx.3.8	Dx.3.9	Dx.3.10	Dx.3.11	Dx.4	Dx.5
1	7518	1.898	0.025641	0.208163	0.098824	0.333333	0.413462	0.121827	0.02139	0.265306	0	0.089552	0	0.162214	0.02439	0.178571	0	0	0.173077	0.367521	0.215164
2	3490	0.881	0.435897	0.334694	0.204706	0.333333	0.317308	0.370558	0.069519	0.142857	0.466667	0.134328	0.214286	0.280534	0.390244	0.321429	1	0	0.269231	0.094017	0.1875
7	4100	1.035	0.538462	0.457143	0.696471	0.333333	0.269231	0.507614	0.909091	0.591837	0.533333	0.776119	0.785714	0.557252	0.589366	0.5	0	0	0.557692	0.538462	0.597336
	Dx.6.1	Dx.6.2	Dx.6.3	Dx.6.4	Dx.6.5	Dx.6.6	Dx.6.7	Dx.6.8	Dx.6.9	Dx.7.1	Dx.7.2	Dx.7.3	Dx.7.4	Dx.7.5	Dx.8.1	Dx.8.2	Dx.8.3	Dx.8.4	Dx.8.5	Dx.8.6	Dx.8.7
1	0.333333	0	0.5	0.75	0.141667	0.25	0.211165	0	0.114865	0.152174	0.205	1	0.271028	0.150376	0	0.113636	0	0	0.5	0	0
2	0.666667	0.352941	0.5	0	0.2625	0	0.322816	0.299257	0.405405	0.086957	0.225	0	0.317757	0.285714	0.559597	0.363636	0	0	0	0	0
7	0	0.647059	0	0.25	0.595833	0.75	0.466019	0.700743	0.47973	0.76087	0.57	0	0.411215	0.56391	0.444043	0.522727	0	0	0.5	0	1
	Dx.8.8	Dx.8.9	Dx.9.1	Dx.9.2	Dx.9.3	Dx.9.4	Dx.9.5	Dx.9.6	Dx.9.7	Dx.9.8	Dx.9.9	Dx.9.10	Dx.9.11	Dx.9.12	Dx.10	Dx.11	Dx.12	Dx.13.1	Dx.13.2	Dx.13.3	Dx.13.4
1	0.09319	0.125	0.191489	0.12766	0.045455	0.128	0.16	0.172414	0.142857	0.083333	0	0.102564	0.148352	0.116667	0.098286	0.172917	0	0	0	0.094982	0
2	0.290323	0.314858	0.382979	0.489362	0.181818	0.368	0.52	0.344828	0	0.083333	0	0.512821	0.43956	0.25	0.109714	0.339583	0.444444	0	0.586614	0.178315	0.4
7	0.616487	0.560142	0.425532	0.382979	0.772727	0.504	0.32	0.482759	0.857143	0.833333	1	0.384615	0.412088	0.633333	0.792	0.4875	0.555556	1	0.413386	0.726703	0.6
	Dx.13.5	Dx.13.6	Dx.13.7	Dx.13.8	Dx.13.9	Dx.14	Dx.15	Dx.16.1	Dx.16.2	Dx.16.3	Dx.16.4	Dx.16.5	Dx.16.6	Dx.16.7	Dx.16.8	Dx.16.9	Dx.16.10	Dx.16.11	Dx.16.12	Matches>0.5	
1	0	0.285714	0	0.1375	0.155405	0.163265	0	0.067797	0.271429	1	0.6875	0.166667	0.204082	0.174803	0.229008	0	0	0.125	0.151042	6	
2	0	0.265306	0	0.3625	0.209459	0.244898	0	0.423729	0.407143	0	0.3125	0.166667	0.387755	0.297638	0.442748	0.833333	0.181818	0.125	0.5	9	
7	0	0.44898	1	0.5	0.635135	0.591837	0	0.508475	0.321429	0	0	0.666667	0.408163	0.527559	0.328244	0.166667	0.818182	0.75	0.348958	46	
Proportion across Diagnoses Codes (by Segment)																					
	util	prov_util	Dx.1.1	Dx.1.2	Dx.1.3	Dx.1.4	Dx.1.5	Dx.2	Dx.3.1	Dx.3.2	Dx.3.3	Dx.3.4	Dx.3.5	Dx.3.6	Dx.3.7	Dx.3.8	Dx.3.9	Dx.3.10	Dx.3.11	Dx.4	Dx.5
1	7518	1.898	0.000529	0.026956	0.022199	0.002114	0.022727	0.02537	0.002114	0.006871	0	0.003171	0	0.044926	0.000529	0.002643	0	0	0.023784	0.022727	0.110994
2	3490	0.881	0.003048	0.014701	0.015597	0.000717	0.005916	0.026174	0.002331	0.001255	0.001255	0.001613	0.000538	0.026354	0.002868	0.001613	0.000359	0	0.012549	0.001972	0.032807
7	4100	1.035	0.00225	0.011999	0.031712	0.000429	0.003	0.021427	0.018213	0.003107	0.000857	0.005571	0.001178	0.031283	0.002571	0.0015	0	0	0.015535	0.00675	0.06246
	Dx.6.1	Dx.6.2	Dx.6.3	Dx.6.4	Dx.6.5	Dx.6.6	Dx.6.7	Dx.6.8	Dx.6.9	Dx.7.1	Dx.7.2	Dx.7.3	Dx.7.4	Dx.7.5	Dx.8.1	Dx.8.2	Dx.8.3	Dx.8.4	Dx.8.5	Dx.8.6	Dx.8.7
1	0.000529	0	0.000529	0.004757	0.01797	0.000529	0.045983	0	0.008985	0.033298	0.02167	0.001586	0.045983	0.010571	0	0.005285	0	0	0.001057	0	0
2	0.00359	0.001076	0.000179	0	0.011294	0	0.023844	0.028863	0.010757	0.006454	0.008067	0	0.018286	0.006812	0.165651	0.005737	0	0	0	0	0
7	0	0.001178	0	0.000321	0.01532	0.000321	0.02057	0.04039	0.007607	0.033748	0.012213	0	0.014142	0.008035	0.079066	0.004928	0	0	0.000214	0	0.000107
	Dx.8.8	Dx.8.9	Dx.9.1	Dx.9.2	Dx.9.3	Dx.9.4	Dx.9.5	Dx.9.6	Dx.9.7	Dx.9.8	Dx.9.9	Dx.9.10	Dx.9.11	Dx.9.12	Dx.10	Dx.11	Dx.12	Dx.13.1	Dx.13.2	Dx.13.3	Dx.13.4
1	0.013742	0.056025	0.004757	0.003171	0.000529	0.016913	0.002114	0.005285	0.000529	0.001057	0	0.002114	0.014271	0.014799	0.045455	0.043869	0	0	0	0.056025	0
2	0.014521	0.047867	0.003227	0.004123	0.000717	0.016493	0.002331	0.003586	0	0.000359	0	0.003586	0.014342	0.010757	0.01721	0.029222	0.079598	0	0.13356	0.035676	0.000359
7	0.018427	0.050889	0.002143	0.001928	0.001821	0.013499	0.000857	0.003	0.000643	0.002143	0.000536	0.001607	0.008035	0.016285	0.074245	0.02507	0.05946	0.000321	0.056246	0.086887	0.000321
	Dx.13.5	Dx.13.6	Dx.13.7	Dx.13.8	Dx.13.9	Dx.14	Dx.15	Dx.16.1	Dx.16.2	Dx.16.3	Dx.16.4	Dx.16.5	Dx.16.6	Dx.16.7	Dx.16.8	Dx.16.9	Dx.16.10	Dx.16.11	Dx.16.12		
1	0	0.0074	0	0.052326	0.012156	0.004228	0	0.004228	0.020085	0.000529	0.005814	0.000529	0.005285	0.058668	0.015856	0	0	0.000529	0.015328		
2	0	0.002331	0	0.046791	0.005558	0.002151	0	0.008964	0.010219	0	0.000896	0.00179	0.003406	0.033883	0.010398	0.000896	0.000359	0.000179	0.01721		
7	0	0.002357	0.000321	0.038569	0.010071	0.003107	0	0.006428	0.004821	0	0	0.000429	0.002143	0.03589	0.004607	0.000107	0.000964	0.000643	0.007178		

**Table A-3. Cluster and Staffing Model Results (Base B, 3 years, Tolerance = 0.05)**

Raw Results Data																					
	size	util	prov_util	Dx.1.1	Dx.1.2	Dx.1.3	Dx.1.4	Dx.1.5	Dx.2	Dx.3.1	Dx.3.2	Dx.3.3	Dx.3.4	Dx.3.5	Dx.3.6	Dx.3.7	Dx.3.8	Dx.3.9	Dx.3.10	Dx.3.11	
1	1055	6652	1.68	22	141	302	18	197	204	230	28	4	40	17	354	6	22	0	4	98	
2	2712	3500	0.884	3	33	74	2	46	64	24	12	0	5	2	98	2	2	0	0	23	
3	1358	4337	1.095	11	99	205	8	105	147	95	23	2	14	1	233	13	6	0	1	62	
4	159	1036	0.262	0	26	61	2	16	29	24	12	1	6	1	57	2	0	0	0	12	
	Dx.4	Dx.5	Dx.6.1	Dx.6.2	Dx.6.3	Dx.6.4	Dx.6.5	Dx.6.6	Dx.6.7	Dx.6.8	Dx.6.9	Dx.7.1	Dx.7.2	Dx.7.3	Dx.7.4	Dx.7.5	Dx.8.1	Dx.8.2	Dx.8.3	Dx.8.4	Dx.8.5
1	43	1131	0	21	1	11	223	2	266	367	120	282	157	1	230	52	1316	62	0	2	1
2	14	229	1	4	2	3	41	0	89	143	25	56	33	0	96	19	232	17	0	0	0
3	36	317	0	16	2	9	88	1	189	226	75	193	73	2	145	47	520	38	0	0	2
4	3	132	0	0	0	0	28	0	38	50	31	54	16	1	36	18	119	3	0	0	0
	Dx.8.6	Dx.8.7	Dx.8.8	Dx.8.9	Dx.9.1	Dx.9.2	Dx.9.3	Dx.9.4	Dx.9.5	Dx.9.6	Dx.9.7	Dx.9.8	Dx.9.9	Dx.9.10	Dx.9.11	Dx.9.12	Dx.10	Dx.11	Dx.12	Dx.13.1	Dx.13.2
1	0	4	166	721	23	19	33	226	26	29	12	27	1	19	126	148	559	364	768	1	1125
2	0	0	37	186	3	6	9	31	6	2	1	2	0	5	18	27	99	95	246	0	166
3	0	0	105	422	11	10	9	120	17	37	3	9	2	7	42	77	320	251	583	0	367
4	0	0	30	90	2	2	4	36	1	2	1	1	0	5	15	21	90	24	108	0	115
	Dx.13.3	Dx.13.4	Dx.13.5	Dx.13.6	Dx.13.7	Dx.13.8	Dx.13.9	Dx.14	Dx.15	Dx.16.1	Dx.16.2	Dx.16.3	Dx.16.4	Dx.16.5	Dx.16.6	Dx.16.7	Dx.16.8	Dx.16.9	Dx.16.10	Dx.16.11	Dx.16.12
1	411	8	0	51	0	590	113	42	2	96	47	0	5	5	38	389	91	12	4	6	81
2	124	1	0	17	2	144	28	19	2	7	11	0	1	0	5	78	21	1	2	2	11
3	251	5	1	29	1	328	71	24	1	24	37	0	16	2	17	202	50	4	1	3	32
4	898	0	0	12	0	111	116	6	0	5	5	0	3	2	3	63	9	0	1	0	11
Proportion across Segments (by Diagnosis code)																					
	util	prov_util	Dx.1.1	Dx.1.2	Dx.1.3	Dx.1.4	Dx.1.5	Dx.2	Dx.3.1	Dx.3.2	Dx.3.3	Dx.3.4	Dx.3.5	Dx.3.6	Dx.3.7	Dx.3.8	Dx.3.9	Dx.3.10	Dx.3.11	Dx.4	Dx.5
1	6652	1.68	0.611111	0.471572	0.470405	0.6	0.541209	0.459459	0.616622	0.373333	0.571429	0.615385	0.809524	0.477089	0.26087	0.733333	0	0.8	0.502564	0.447917	0.625207
2	3500	0.884	0.083333	0.110368	0.115265	0.066667	0.126374	0.144144	0.064343	0.16	0	0.076923	0.095238	0.132075	0.086957	0.066667	0	0	0.117949	0.145833	0.126589
3	4337	1.095	0.305556	0.331104	0.319315	0.266667	0.288462	0.331081	0.254692	0.306667	0.285714	0.215385	0.047619	0.314016	0.565217	0.2	0	0.2	0.317949	0.375	0.175235
4	1036	0.262	0	0.086957	0.095016	0.066667	0.043956	0.065315	0.064343	0.16	0.142857	0.092308	0.047619	0.076819	0.086957	0	0	0	0.061538	0.03125	0.072968
	Dx.6.1	Dx.6.2	Dx.6.3	Dx.6.4	Dx.6.5	Dx.6.6	Dx.6.7	Dx.6.8	Dx.6.9	Dx.7.1	Dx.7.2	Dx.7.3	Dx.7.4	Dx.7.5	Dx.8.1	Dx.8.2	Dx.8.3	Dx.8.4	Dx.8.5	Dx.8.6	Dx.8.7
1	0	0.512195	0.2	0.478261	0.586842	0.666667	0.457045	0.466921	0.478088	0.482051	0.562724	0.25	0.453649	0.382353	0.601738	0.516667	0	1	0.333333	0	1
2	1	0.097561	0.4	0.130435	0.107895	0	0.152921	0.181934	0.099602	0.095726	0.11828	0	0.189349	0.139706	0.106081	0.141667	0	0	0	0	0
3	0	0.390244	0.4	0.391304	0.231579	0.333333	0.324742	0.287532	0.298805	0.329915	0.261649	0.5	0.285996	0.345588	0.237769	0.316667	0	0	0.666667	0	0
4	0	0	0	0.073684	0	0.065292	0.063613	0.123506	0.092308	0.057348	0.25	0.071006	0.132353	0.054412	0.025	0	0	0	0	0	0
	Dx.8.8	Dx.8.9	Dx.9.1	Dx.9.2	Dx.9.3	Dx.9.4	Dx.9.5	Dx.9.6	Dx.9.7	Dx.9.8	Dx.9.9	Dx.9.10	Dx.9.11	Dx.9.12	Dx.10	Dx.11	Dx.12	Dx.13.1	Dx.13.2	Dx.13.3	Dx.13.4
1	0.491124	0.508104	0.589744	0.513514	0.6	0.547215	0.52	0.414286	0.705882	0.692308	0.333333	0.527778	0.628866	0.542125	0.523408	0.495913	0.45044	1	0.634518	0.244062	0.571429
2	0.109467	0.131078	0.076923	0.162162	0.163636	0.075061	0.12	0.028571	0.058824	0.051282	0	0.138889	0.089552	0.098901	0.092697	0.129428	0.144282	0	0.093627	0.073634	0.071429
3	0.310651	0.297393	0.282051	0.27027	0.163636	0.290557	0.34	0.538571	0.176471	0.230769	0.666667	0.194444	0.208955	0.282051	0.299625	0.341962	0.341935	0	0.206994	0.14905	0.375143
4	0.088757	0.063425	0.051282	0.054054	0.072727	0.087167	0.02	0.028571	0.058824	0.025641	0	0.138889	0.074627	0.076923	0.08427	0.032698	0.063343	0	0.064862	0.533254	0
	Dx.13.5	Dx.13.6	Dx.13.7	Dx.13.8	Dx.13.9	Dx.14	Dx.15	Dx.16.1	Dx.16.2	Dx.16.3	Dx.16.4	Dx.16.5	Dx.16.6	Dx.16.7	Dx.16.8	Dx.16.9	Dx.16.10	Dx.16.11	Dx.16.12	Matches>0.5	
1	0	0.46789	0	0.502984	0.344512	0.461538	0.4	0.727273	0.47	0	0.2	0.555556	0.603175	0.531421	0.532164	0.705882	0.5	0.545455	0.6	4	
2	0	0.155963	0.666667	0.122762	0.085366	0.208791	0.4	0.05303	0.11	0.04	0	0.079365	0.106557	0.122807	0.058824	0.25	0.181818	0	0.081481	0.081524	2
3	1	0.266055	0.333333	0.279625	0.216463	0.263736	0.2	0.181818	0.37	0	0.64	0.222222	0.269841	0.275956	0.292398	0.235294	0.125	0.272727	0.237037	7	
4	0	0.110092	0	0.094629	0.353659	0.065934	0	0.037879	0.05	0	0.12	0.222222	0.047619	0.086066	0.052632	0	0.125	0	0.081481	1	
Proportion across Diagnoses Codes (by Segment)																					
	util	prov_util	Dx.1.1	Dx.1.2	Dx.1.3	Dx.1.4	Dx.1.5	Dx.2	Dx.3.1	Dx.3.2	Dx.3.3	Dx.3.4	Dx.3.5	Dx.3.6	Dx.3.7	Dx.3.8	Dx.3.9	Dx.3.10	Dx.3.11	Dx.4	Dx.5
1	6652	1.68	0.00178	0.011405	0.024428	0.001456	0.015935	0.016501	0.018604	0.002265	0.000324	0.003235	0.001375	0.028634	0.000485	0.00178	0	0.000324	0.007927	0.003478	0.091483
2	3500	0.884	0.001068	0.011748	0.026344	0.000712	0.016376	0.022784	0.008544	0.004272	0	0.00178	0.000712	0.034888	0.000712	0.000712	0	0	0.008188	0.004984	0.081524
3	4337	1.095	0.001694	0.015242	0.031563	0.001232	0.016166	0.022633	0.014627	0.003541	0.000308	0.002156	0.000154	0.035874	0.000202	0.000924	0	0.000154	0.0009546	0.005543	0.048807
4	1036	0.262	0	0.010117	0.023735	0.000778	0.006226	0.011284	0.009339	0.004669	0.000389	0.002335	0.000389	0.022179	0.000778	0	0	0	0.004669	0.001167	0.051362
	Dx.6.1	Dx.6.2	Dx.6.3	Dx.6.4	Dx.6.5	Dx.6.6	Dx.6.7	Dx.6.8	Dx.6.9	Dx.7.1	Dx.7.2	Dx.7.3	Dx.7.4	Dx.7.5	Dx.8.1	Dx.8.2	Dx.8.3	Dx.8.4	Dx.8.5	Dx.8.6	Dx.8.7
1	0	0.001699	8.09E-05	0.00089	0.018038	0.000162	0.021516	0.029685	0.009706	0.02281	0.012699	8.09E-05	0.018604	0.004206	0.106447	0.005015	0	0.000162	8.09E-05	0	0.000324
2	0.000356	0.001424	0.000712	0.001068	0.014596	0	0.031684	0.050908	0.0089	0.019936	0.011748	0	0.034176	0.006764	0.082592	0.006052	0	0	0	0	0
3	0	0.002463	0.000308	0.001386	0.013549	0.000154	0.029099	0.034796	0.011547	0.029715	0.011239	0.000308	0.022325	0.007236	0.080062	0.005851	0	0	0.000308	0	0
4	0	0	0	0.010895	0	0.014786	0.019455	0.012062	0.021012	0.006226	0.000389	0.014008	0.007004	0.046304	0.001167	0	0	0	0	0	0
	Dx.8.8	Dx.8.9	Dx.9.1	Dx.9.2	Dx.9.3	Dx.9.4	Dx.9.5	Dx.9.6	Dx.9.7	Dx.9.8	Dx.9.9	Dx.9.10	Dx.9.11	Dx.9.12	Dx.10	Dx.11	Dx.12	Dx.13.1	Dx.13.2	Dx.13.3	Dx.13.4
1	0.013427	0.058319	0.00186	0.001537	0.002669	0.01828	0.002103	0.002346	0.000971	0.002184	8.09E-05	0.001537	0.010192	0.011971	0.045216	0.029443	0.062121	8.			

## **LIST OF ABBREVIATIONS AND ACRONYMS**

<b>AFMS</b>	Air Force Medical Service
<b>CLARA</b>	clustering large applications
<b>CONUS</b>	contiguous United States
<b>Dx</b>	diagnosis
<b>FOMC</b>	Flight and Operational Medicine Clinic
<b>ICD-9-CM</b>	International Classification of Diseases, Ninth Revision, Clinical Modification
<b>OCONUS</b>	outside the contiguous United States
<b>USAF</b>	U.S. Air Force