Analysis of the Referral Mechanism between Family Doctors and Specialists using Social Network Analysis

João Manuel Caixinha Casal da Veiga

Work project carried out under the supervision of:

Co Advisor: Qiwei Han

Co Advisor: Miguel Castro Neto

07/01/2020

ABSTRACT

The following document represents a master thesis dissertation, that consists in the development of a research project in which its main objective is to understand the referral mechanism between primary care physicians and specialists using network science and other tools. The referral network of a healthcare provider denotes an important source of costs and revenue, as it can affect directly the management of its clients and employees, namely through the quality of the care being provided and the level of satisfaction being achieved. The data provided for the development of the study was given by a European industry leader in private healthcare. It is important to highlight that this research study attempts maps the relationships between general practitioners and specialists using a large dataset of doctor's appointments during 2012 and 2017. These relationships were mapped under the assumptions that two doctors had to share at least one patient, and the period between the two appointments could not be longer than 30 days. Moreover, the impact of the dynamics of the relationships between the two types of doctors in the primary-specialty referral mechanism is done by analyzing the referral patterns exhibited in network, and the performance of the physicians in terms of the centrality measures degree, betweenness and closeness. Additionally, two regression analysis are performed with the objective of identifying potential characteristics that might be affecting the referral rates of doctors. These characteristics include the social network metrics and the physician's backgrounds.

KEYWORDS

Social Network Analysis; Referral Mechanism; Informal Networks; Private Healthcare

INDEX

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

WHO   World Healthcare Organization

OCDE   Organization for Economic Cooperation and Development

CAGR   Compounded Annual Growth Rate

NHS   National Healthcare Service

PCP   Primary Care Physician

GP   General Practitioner

HIT   Healthcare Information Technology

PCMH   Patient Centered Medical Homes

ACO   Accountable Care Organizations

ACSS   Administração Central do Sistema de Saúde

SNA   Social Network Analysis

1. INTRODUCTION

The relationship between a primary care physician1 (pcp) and a specialist (sp) represents a relevant communication channel for the healthcare system, as it can deeply influence the outcome of a patient's treatment quality and the level of satisfaction regarding the care provided by a healthcare institution (Kinchen, Cooper, Levine, Wang, & Powe, 2004). Additionally, it can act as a significant driver of costs (Barnett, Landon, O'Malley, Keating, & Christakis, 2011).

Having knowledge of the relationships established in a referral mechanism, between pcp and sp, can represent an important source of information to manipulate the overall network for the benefit of healthcare providers, as doctors are influenced by their respective networks (B.E. et al., 2012). An efficient healthcare referral system is proven to improve the ability of both types of physicians to provide better quality treatment to their patients (Care, Referral, Means, & Problems, 2019).

This dissertation presents an approach in which the main objective is to better understand the relationship between doctors in primary care and specialists. Specifically, what might be the potential impact of its dynamics in the informal structure, that is the primary-specialty referral network mechanism, using network science and other tools. From an extensive dataset with 9,173,891 patient consultations records, more than 1.3 million unique patients and 2171 unique doctors it was generated a bipartite social network with 459 connected general practitioners, 14 isolated family doctors, 1487 connected specialist and 211 isolated doctors certified in differentiated care. Additionally, the 1946 connected doctors are linked between each other by no more than 59,065 edges. This network was further augmented with the information regarding doctors' backgrounds, their age, the number of years of medical practice, gender and others.

A large percentage of the existent literature focus on the use of descriptive methods to identify and describe patterns based on the informal relationships between primary care doctors and specialists, such as doctor specialties, the patient diagnosis, and disease severity (An, O'Malley, Rockmore, & Stock, 2018). In that sense, this dissertation act as a complement to other research studies, as it allows to validate, or not, patterns that have been recognized to affect the referral decision process between the two types of doctors. Additionally, it is investigated if certain network science metrics have an impact or not in the primary-specialty referral mechanism.

1.1. RESEARCH OBJECTIVES

The research purposes of this dissertation were to add to the current literature more information about the referral patterns of a primary-specialty referral system modelled as a bipartite network and explore its potential inefficiencies. Moreover, it was presented the opportunity to test the hypothesis that the different characteristics of a doctor's background and social network centrality measures have an impact or not in the referral rate of physician. From the perspective of the one making the referrals (pcp) and from the perspective of the one receiving them (sp).

2. LITERATURE REVIEW

In this section it is provided a contextualization of the healthcare industry, by making a helicopter approach at the global landscape (zoom out) and focusing on the Portuguese healthcare system (zoom in). Furthermore, an overview of the referral mechanism between Primary Care Doctors and Specialists will be made, along with its impact to the different type of individuals involved in the process: patients, doctors and respective healthcare providers. Additionally, a synopsis of the patterns so far studied will be referred. To conclude, an overview on the methodology of social network analysis is written for further understanding on how it can be applied to study of the referral mechanism between PCP and Specialist doctors, along with the description of its characteristics.

2.1. HEALTHCARE INDUSTRY

According to the World Health Care Organization (WHO), health can be seen as a state physical, mental and social well-being and not only the absence of sickness or a disease (Larson, 1996). In 2017 it was estimated that In terms of access to essential healthcare half of the worldwide population did not have access (WHO, 2017).

The amount spent in healthcare by households varies according to the country. In 2017 was estimated that the healthcare spending per capita in India to be 238 dollars while in the United States of America was approximately 34 times higher (8047 dollars). In the particular case of Portugal, the household spending in health is roughly 2000 dollars, 1925 to be more precisely (OCDE, 2019).

Between the years of 2013 and 2017 it was registered an increase in healthcare spending (2.9%) and is still expected to grow by even larger numbers, as until 2022 is expected for healthcare spending to have an annual growth of 5.4% (CAGR). This growth is associated with the growing needs of healthcare of the elderly population, the increase of the labour costs in the industry, the advances in treatments, the development of new healthcare technologies, and the expansion of healthcare coverage broaden up to new markets (Deloitte, 2019).

The healthcare stakeholders are investing in a new vision for the healthcare industry in which, instead of their goal being to treat their clients when they are sick already, is to create a system that promotes the well-being, prevention and early intervention on minor or serious diseases. The objective of this to develop of the concept of smart health in order to drive innovation, to reduce costs both to the service provider and for the patient while increasing the quality of the treatments (Deloitte, 2019).

Healthcare organizations are experiencing a state of uncertainty in the health economy as profits have been shrinking, new competitors are entering new markets, personalized medicine is becoming a certainty, disruptive technologies and clinical advances are constantly entering in the market. The healthcare sector is expected to invest 280.25 billion dollars by 2021 in health technologies, related products and services (Deloitte, 2019). Therefore, this reality is creating a sense of urgency in healthcare providers to decide now whether they want to invest to be a market leader, a follower or a niche player (Deloitte, 2019). The factors that dictate the competition within private operators can be summarized to reputation, clinical excellence, technology, price and client satisfaction (Deloitte, 2019).

2.2. HEALTHCARE INDUSTRY IN PORTUGAL

When Portugal was under the government of a dictatorship the healthcare system in Portugal was fragmented in, private healthcare and state hospitals which were usually located in important urban centers. In 1958, the Portuguese health ministry was created. Until 1970 Portugal was one of the worst countries with access to essential care in the European Union. After 1971, it was created a new system where each districted had two different infrastructures implemented, healthcare centers and hospitals. The current national healthcare (NHS) was then created in 1974 with the democratic purpose of making healthcare available to everyone (Baganha, Ribeiro, & Pires, 2002). The coexistence of public and private healthcare is due to the fact the national healthcare service is not able to cover the entire Portuguese population (Baganha et al., 2002).

Currently the Portuguese NHS administration is divided into five major health regions: Norte, Centro, Lisboa and Vale do Tejo, Alentejo and Algarve. Their headquarters are respectively, Porto, Coimbra, Lisboa, Évora, and Faro. Each region is then divided into sub-regions corresponding to districts. In each district is expected for the healthcare services to be ensure by healthcare centers and hospitals. Nowadays, healthcare centers are in charge of providing the primary healthcare to the Portuguese population, with respect to the national public health sector (Baganha et al., 2002).

When compared with other countries worldwide, the number of doctors per 1000 inhabitants in Portugal is estimated to be above the average by registering a number of 4,26 doctors against the average estimation of 3,2 of the OCDE and its 34 state members. In Portugal, is has been forecasted that approximately 600 000 Portuguese still do not have a family doctor (Notícias, 2019). In addition, the number of patients that that a General Practitioner has to see has increase from 1550 in 2011 to 1900 in 2018 (Notícias, 2019), and there are regions that register extreme numbers like 2500 (Teixeira, 2018).

2.3. PRIMARY CARE DOCTORS (PCP) AND SPECIALISTS

Primary care denotes a connection for individuals that requests hospital services in a world where healthcare is more specialized than ever, therefore more fragmented. Thus, Primary Care doctors, also known as General Practitioners or family doctors, often represent the first point of access of the healthcare system. Additionally, they are the ones that are more likely to develop long term relationships, as they regularly follow patients for long periods (University, 2018).

The training of a PCP is particularly focused on giving practitioners the necessary tools to ensure that they have the best first contact and provide quality continuing care of patients undiagnosed signs. Furthermore, the medical expertise's that are comprised in this category of doctors include family medicine, internal medicine and pediatrics[1] (American Academy of Family Physicians, 2016). Therefore, the remaining specialties such as cardiovascular surgery, ophthalmology along with more

---

[1] The term Primary Care does not necessarily characterize exactly the activities or practices a General Practitioner performs or has to perform.

45 areas of expertise available in Table 1 - Doctor Specialties[2] (D. da República, 2018b), will determine who are the doctors classified as Specialists.

Doctor Specialties

Anesthesiology	Physical Medicine and Rehabilitation	
Pathologic anatomy	General and Family Medicine	
Angiology and Vascular Surgery	Intensive Care Medicine	
Cardiology	Internal medicine	
Pediatric Cardiology	legal Medicine	
Cardiac surgery	Nuclear medicine	
Cardiothoracic Surgery	Tropical Medicine	
General surgery	Nephrology	
Maxillofacial Surgery	Neurosurgery	
Pediatric surgery	Neurology	
Reconstructive and Aesthetic Plastic Surgery	Neuroradiology	
Thoracic surgery	Ophthalmology	
Dermato-Venereologia	Medical Oncology	
Infectious diseases	Orthopedics	
Endocrinology and Nutrition	Otolaryngology	
Stomatology	Clinical pathology	
Gastroenterology	Pediatrics	
Medical Genetics	Pneumology	
Gynecology / Obstetrics	Psychiatry	
Immunoallergology	Psychiatry of Childhood and Adolescence	
Immunohemotherapy	Radiology	
Clinical Pharmacology	Radioncology	
Clinical Hematology	Rheumatology	
Sports Medicine	Public health	
Work Medicine	Urology	

Table 1- Doctor Specialties

2.4. THE REFERRAL MECHANISM BETWEEN PRIMARY CARE DOCTORS AND SPECIALISTS

In a significant number of countries, it is required for a patient to see a Specialist or have access to other medical resources such as laboratory tests require first to have an appointment with a PCP (Liddy, Arbab-Tafti, Moroz, & Keely, 2017). However, in the irregular situation, of an emergency, a patient can ask to be seen by a Specialist if it visits a hospital through the emergency room. But further access to other healthcare resources requires the recommendation of the respective Specialists. In addition, nowadays it is possible to schedule an appointment with a Specialist without having to first see a Primary Care Doctor in the private sector.

---

[2] There are two specialties in this table named Tropical Medicine and Cardiothoracic surgery we are not in the aforementioned source of the information, because in 2015 they were recognized as a specialty (D. A. República, 2015).

The practice of referring patients to Specialists is considered to be a common event in the United States of America (Chen & Glover, 2016). Interestingly, two different studies found similar results regarding the percentage of visits to PCP's that include a referral to a Specialist events in the United States of America, 5% (Forrest, Nutting, Von Schrader, Rohde, & Starfield, 2006) and 4.5% (Kinchen et al., 2004). Moreover, in the latter case, 45% of new patients a physician receives is assigned to him or her by referrals. In the particular case of Portugal, it registers a referral rate of 5.56% between primary care and secondary care (Ponte et al., 2006).

Inefficient communication within a referral network can have a negative impact for the referral network, as it can result in ineffective care coordination by the healthcare provider (An, O'Malley, Rockmore, et al., 2018). The level of effectiveness of communication between members of a network is affected by how information and resources flow around an healthcare provider referral network (An, O'Malley, & Rockmore, 2018).

In addition, as good patient experience is related with higher hospital profitability (Deloitte, 2019). Thus, healthcare providers are once more affected by the referral system between the two types of doctors.

Furthermore, in this relationship the Primary Care Doctors are not the only ones that refer their patients to see other doctors if they think that is in the best interest of its patients. Specialists physicians also have that possibility.

In terms of the appropriateness, referrals of patients having a condition that has a low degree of severity by Primary Care Doctors to Specialists, is considered by physicians themselves a waste of "expert" consultations, that ends up adding unnecessarily costs, both to the healthcare provider as to the patient(Jauhar, 2019). However, this kind of behaviour is welcomed by doctors. Moreover, is common for doctors to guarantee that the referrals sources remain happy with them specially major cities and in particular metropolitan areas (Jauhar, 2019). A doctor should only make a referral if it is justified. In that case the ideal is for him or her to refer a doctor or doctors to handle the particular situation of a patient, and not because it is in the best interest of the doctor (Schroeder, 2016).

2.5. PATTERNS IN THE REFERRAL MECHANISM BETWEEN PCP AND SPECIALISTS

It is in the interest of policy makers and healthcare organizations to understand the patterns behind the referral networks as they attempt to control healthcare costs by manipulating how PCPs or Specialists' physicians establish their relationships. Another reason is to cultivate the level of awareness regarding the patterns of referral is to maintain the referrals within a specific institution or organization (Barnet, Song, & Landon, 2012).

The existent referral mechanism between a Primary Care Doctor and Specialists is affected by a variety of complex factors as you might see in table 2 - Factors and Reasons for referral between Doctors [3]. These factors can go from the previous experience a patient had with a certain doctor, as the probability of a referral can depend on the gender of the physician or patient.

---

[3] These table contains factors from multiple sources.

Factors & Reasons for referral between Doctors

Medical skill	Anxiety regarding outcomes
Perceived clinical expertise	Perceived pressure to control referral costs
Patients past experience with a physician	Restricted ability to obtain surgical referrals
Timely availability of a doctor for appointments	Hours of patient care per week
Sharing medical records	Visits per day
Being referred by another doctor	Physician income
Work in the same hospital	Physician income structure
Type of doctor (Primary Care Physician or Specialist)	Ownership practice
Doctor specialty	Level generalization/expertise of a PCP
Patient diagnose	Convenience
Disease severity	Risk aversion
Future impact of a referral on a patient's health care	Race
Getting advice on diagnosis or treatment	Adequate patient time visit
Overall shortage of physicians	Nurse involvement in the process
Costs of tests and procedures	Managed care contracts
Age over 17	Use of HITS
Patient gender	Patient Has chronic conditions
Physician gender	Specialist refers the same PCP
Fear of lawsuit of not consulting an expert	PCP refers the same Specialist
Patient health insurance	Years of practice
Anxiety due to clinical uncertainty scale	

Table 2 - Factors and Reasons for referral between Doctors

A few papers have studied the referral relationship between PCPs and Specialists, the reasons of their choices and their patterns which can extend beyond professional motives (Barnett et al., 2011). For example, according to the study Trends in Physician Referrals in the United States 1999-2009 the most influencing factors affecting referrals were: clinical report, Physician gender, Years of practice, Specialty, Herfindahl Index[4](Barnett, Song, et al., 2012).

When identifying the professional network of each doctor in the overall network, 66% of the referrals of PCPs were made to colleagues inside the network. On the other hand, medical Specialists referred 49% and surgical Specialists referred 52% of colleagues belonging to their respective professional network (Barnett, Keating, Christakis, O'Malley, & Landon, 2012). In the case of Specialists 10% of their referrals were made to Primary Care Physicians. The remaining of the referrals was made to doctors of the same nature (Barnett, Keating, et al., 2012).

Specialists physicians are more likely than PCPs to refer a patient to another doctor depending on the previous experience of his or her patient with the doctor being referred. This can be explained by the fact that they believe to be of less importance if a physician has or does not have a good patient rapport. Conversely PCPs give higher relevance to reasons related to a physician access to one another. For instance: the level of availability of doctor for appointments, the sharing or not of the same medical

---

[4] In this particular case the Herfindahl Index attempts to define the level of expertise of a Primary Care Physician. A doctor who scores one in this index means that it is a Specialist in a specific area. However, if the result is zero it means that he or she a generalist (has a general knowledge about different areas).

record system, or the fact of working in the same hospital. Furthermore, if a patient has an insurance or not, and what type of insurance, are key aspects of the referral process as physicians can be restricted or not to refer patients for further health care (Kinchen et al., 2004).

Another common rules for a Specialist to refer a patient to another doctor depends on the fact of being referred or not by that same doctor (Barnett et al., 2011). The immediately aforementioned article also states that physicians with less capabilities of handling clinical uncertainty were more likely to refer a patient to other doctors. In addition, if the care that a patient requires exceeds the expertise that a Primary Care Physician can give increases the probabilities of referral of a same patient (Franks, Williams, Zwanziger, Mooney, & Sorbero, 2000). In addition, physicians have less encouragement to increase volume of referrals if payments are bundled[5] rather than discriminating for every service (Jauhar, 2019).

Other sources have found out evidence that doctors when referring patients to other colleagues, consider the professionals with who they are familiar with. However , they first consider the ones working at the same hospital (An, O'Malley, & Rockmore, 2018).

It is common for studies to register variations according geographic regions, it is not abnormal when Specialists are found to have a more central role in smaller networks than larger. Despite physicians being assigned to one or more hospitals, the closer proximity of different cluster suggests the existence of multiple ties across different hospitals. In addition, the individual network of a doctor is more likely to have more and stronger ties within the hospital which the professional is affiliated to, despite not always being the verified scenario. Male doctors are more likely to have ties with other male doctors. The same happens with the female gender. Doctors with connections between them reveal a closer proximity in age when compared with unconnected individuals. Furthermore, this study found that not only connected physicians are more likely to be in closer geographical proximity than unconnected doctors as well to have patients with similar medical complexity (An, O'Malley, Rockmore, et al., 2018). Another study supports that doctors professional networks vary across geographic locations (An, O'Malley, Rockmore, et al., 2018). Additionally, it is valid to expect that doctors share patients with colleagues that have personality similarities with them, therefore they demonstrate to have the homophily concept present in their networks. However, the rapid adoption medical electronic records can lead to different findings (B.E. et al., 2012).

A patient diagnostic has a significant amount of explanatory power over the variation of referral rates of PCP as the degree of severity of a patient condition and the possible impact of a referral on the future of his or her health. The availability of Specialist doctors also influences their referral rates. The more Specialists a Primary Care Physician has at his or her disposable for referral the greater are the chances of making one, due to the higher degree of access to specialty care, the patient expectations and demands for seeing a Specialist also increase. Furthermore, patients presenting common pain problems or depression symptoms have lower likelihood of referral. The costs of procedures and surgical interventions (specialist care) represent a constraint towards the referral rates as it is common for policy makers on healthcare organizations implement barriers against inappropriate referrals. Moreover, so far it has not been found an association between the rate of referrals and their quality (Forrest et al., 2006).

In terms of psychological measure they do not reveal to be a determinant aspect in the referral likelihood variability (Franks et al., 2000).On the other hand ethnicity has a stronger explanatory power

---

[5] Payments that covers all physicians' services and hospital care for each patient.

as communities where the degree of concentration of black residents is high have a lower tendency to get treatment when compared with white individuals from the same population (Ghomrawi, Funk, Parks, Owen-Smith, & Hollingsworth, 2018).Additionally, black physician report to have a perception of racial discrimination by white doctors in the referral process (Kinchen et al., 2004).

Sometimes a certain degree of financial transparency from healthcare providers can be demanded to their employees when seeing patients to build stronger relationships between the two parties. In addition, it is common for countries to have laws that prevent doctors to pay other doctors directly for referrals (Schroeder, 2016). Moreover, the availability to a PCP of healthcare information technology, nurse care manager, adequate visit time and quality reports regarding patients with chronic conditions would be expected to increase the referral rates by from 63.9% to 82.7%. The increased administrative burden Primary Care Physician face along with diminish of reimbursements are creating heat on those doctors to see more patients in less time. Having this kind of structure decreases the effectiveness of the communication in a referral network and the quality of the care provided. Moreover, systematic structures, tools and processes for information creation, sharing, receipt and recognition from both parties the sending and receiving individual are necessary to support medical care practices (O'Malley, Ann S., MD & Reschovsky, 2011). An important consequence to bear in mind when increasing the number of referrals, is that, that implies also an additional increase in the number of ambulatory visits for the average person, either in the primary care services and Specialist care (Barnett, Song, et al., 2012).

It has been registered an increase in the referral rates. The study Trends in Physician Referrals in the United States, 2012 it attempts to disclose the justifications behind that event. The study hypothesizes that it might be due to the increasing complexity of care itself, for instance common diseases with general or viral symptoms did not register such increase. In addition, more than before, physicians are asked to do more while having the same amount of consultation time. Patients of today on average require attention as they need more medication and they have one or more chronic condition (Barnett, Song, et al., 2012). In Portugal the average consultation time per patient is approximately 15.2%, compared with other 67 studies from the International variations in Primary Care Physician consultation time: a systematic review of 67 countries, Portugal the 10th position (Irving et al., 2017). In addition, a General Practitioner is estimated to see 19.1 patients per day (Granja, Ponte, & Cavadas, 2014).

3. METHODOLOGY

The current section represents the process of the development of the dissertation since the data pre-processing until all the tools and analysis made in order to better understand the referral mechanism between general practitioners. Additionally, it is presented the reasoning, assumptions and limitations that arose during the construction of the primary-specialty referral network and other frameworks.

3.1. INTRODUCTION

As previously mentioned, the objective of this study is to better understand the referral mechanism between primary care doctors and specialists using network science, and other tools. To do so a referral network between the two types of doctors is estimated. Moreover, in this network the referral patterns and certain social network metrics such as degree centrality will be explored. Finally, a regression analysis is produced with the aim of understanding if the different types of information available, that mean respect to the background of a doctor, including their centrality measures, if they have any impact in the referral rate associated to each doctor.

The database used for the development of the research study was provided by a European healthcare provider in 2017. The originated from Portugal. The database management software used for the storage of the data was Dbeaver. The Programming Languages used for the cleaning, visualization of the data, networks construction and analysis were python16 and postgresql. The development of the dataset necessary for the progress of this report can be divided into two main stages. The first phase consisted in extracting, cleaning and visualizing all the relevant information regarding a doctors' profile. The second stage (explained in the chapter "Schema Evolution") involved adding and validating new information to complement the initial dataset.

3.2. DATA PRE-PROCESSING

This stage represents the selection, integration, description, cleaning, and visualization of final dataset which is used for the development of the respective project. Its main objective is to describe the education background and current places of work.

The development of the final dataset departed from a table called initial_doctor_table. This table represents 3500 observations, each corresponding to a unique doctor, being described by 13 variables. From those variable 7 represent text type variables, 3 bigint type tables, 1 numeric type variable and 2 timestamp with time zone variable. In addition, each variable has a role. There is 1 ID variable, 6 variables defined as interval, 4 variables defined as Nominal, 1 variable defined as Binary, and 1 variable declared as Ordinal (see table 3 – Initial_Doctor_table).

Variable	Variable Type	Role	Description
Doctor_id	Text	ID	Unique identifier of a doctor
Birth_year	Bigint	Interval	Birth of the corresponding doctor
Gender	Text	Binary	Sex of each doctor
Specialty	Text	Nominal	Area of expertise of a doctor
Specialty_2	Text	Nominal	Second area of expertise of a doctor

Academic_Degree	Text	Nominal	Level of education and the training area of a doctor
Years_Of_Practice	Bigint	Interval	Number of years of how long ago a doctor received a school certificate
School_Certificate	Text	Ordinal	Certificate for the conclusion of a certain stage of medical education
Grade	Numeric	Interval	Final grade of doctor school certificate
Learning_Location	Text	Nominal	Higher education institution, location, or both at which an individual concluded a certain stage of his or her medical education
Number_Of_Workshops	Bigint	Interval	Number of workshops a doctor took until 2017
Initial_Date	Timestamp with time zone	Interval	Initiation date of the school certificate
End_Date	Timestamp with time zone	Interval	End date of the school certificate

Table 3 - Initial Doctor Table

3.2.1. Missing Values

The first step to treat the initial doctor table, was to handle the existence of missing values in it. A particular problem was identified related with this issue. Some values in some variables were in blank (not null or erroneous values). Those variables were: specialty, specialty_2, Academic_Degree, School_Certificate, and Learning_Location. This is relevant because the meaning of a blank value is assumed to be the same as a null value. Therefore, to correct this situation the blank values were updated to null values. The output of such action produced the following amount of missing values that is possible to observe in table 4 – Initial Doctor Table Missing Values.

Variable	Variable Type	Percentage of Null Values	
doctor_id	0		0,0%
birth_year	0		0,0%
gender	0		0,0%
specialty	2096		59,9%
specialty_2	3488		99,7%
Academic_Degree	226		6,5%
years_of_practice	0		0,0%
school_certificate	28		0,8%
grade	3136		89,6%
learning_location	28		0,8%
number_of_workshops	0		0,0%
initial_date	28		0,8%
end_date	1559		44,5%

Table 4 – Initial Doctor Table Missing Values

Considering the amount of null values in some variables, some had to be dropped because they could not add any value or information to the dataset. Therefore, the variables specialty_2 and grade, were dropped due to the fact that they had 99.7% and 89.6% of missing values.

Additionally, the variables initial_date and end_date were removed also due to the significant amount of missing values in the latter case. In addition, Academic_Degree is considered to be redundant because the information that is possible to extract from it, is provided by other variables. The different values that the variable can take can be seen in table 5 – Academic Degree Values. Furthermore, the variable number_of_workshops do not carry relevant information for the development of the dissertation. Therefore, 6 different variables were excluded from the study.

Academic Degree
L. Med. Mestr. Medicina Legal
L. Enfermagem
L. Medicina
L. Med. PG Medicina Legal
L. Medicina Dentária
L. Fisioterapia
L. Med. Dout. Medicina
L. Med. PG Saúde Pública
L. Desconhecida / Não especificada

Table 5 - Academic Degree Values

3.2.2. Outliers

Concerning the process of identifying outliers, the distribution of the quantitative variables and their respective descriptive measures were assessed, namely: Birth_Year, Years_of_Practice. The distribution of the variable birth_year, presents to have a slightly skewed distribution to the left. The youngest individual in this sample was born in 1994, while the oldest person was born in 1900, making it to have 117 in 2017. Approximately 75% of the population was born before or in 1980 (37 years old in 2017) (figure 1 – Birth Year Distribution). The average age of the sample is set to be 48, which means that the average birth year was 1969. Additionally, every individual claiming to be born before 1935 is considered to be an extreme value. By applying such rule, we are excluding only 2 observations.
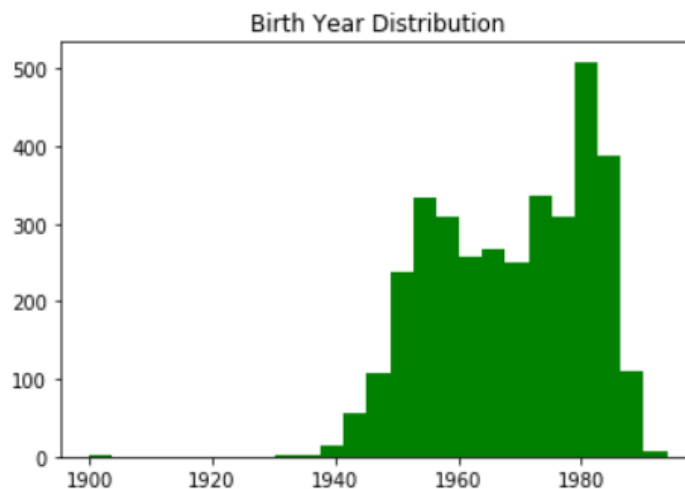

Figure 1 - Birth Year Distribution

The variable years_of_practice presents also a skewed distribution, but in this case is skewed to the right. Meaning the more years of practice, a doctor has the, the fewer are the number of doctors that are able to represent those years of practice. In this sample, on average, doctors finished their studies six years ago. Approximately 75% of the population finished their degrees no more than nine years ago. The range of this variables goes from zero to 39. The value zero can have two different meanings in our sample: It can either mean that a doctor has not finished his or her studies, or it has just finished them in 2017. In terms of outliers, it does not demonstrate to have extreme values (figure 2 – Years of Practice Distribution).
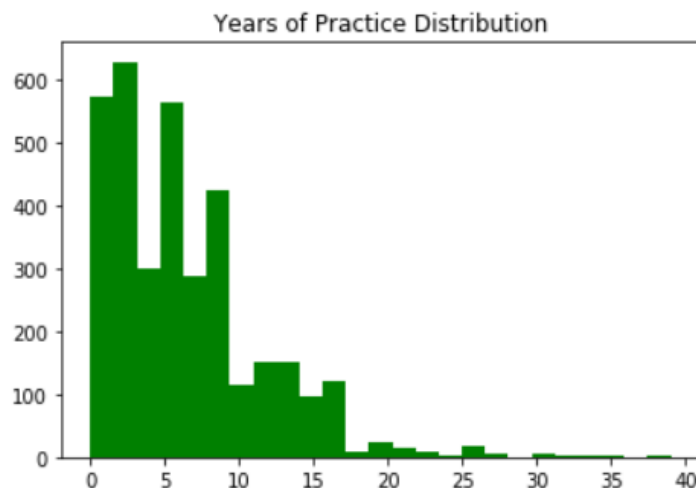

Figure 2 - Years of Practice Distribution

After applying the necessary rules to exclude the outliers from the sample only 1 observation is being removed, which represent 0.06% of the entire dataset. It is possible to observe it more detail in (table 6 – excluded outliers).

Variable	Excluded Observations	Excluded Observations %
Birth_year	2	0,06%
Total	2	0,06%

Table 6 - Excluded Outliers

3.2.3. Data Validation

To become a doctor is necessary to do at least six years of medical school (university), and a minimum of 5 to 7 years to conclude a medical post-graduation known as Internato Médico, that is divided in 1 or 2 years of Internato Geral and a minimum of 4 years to conclude the Internato Complementar, which can reach a maximum of 6 years depending on the specialty. Therefore, it is plausible to assume that the minimum age at which a doctor initiates its medical career in 2017 is at 29 years old. Before 2004, included, what is known today as Ano Comum of the Internato Médico, was in fact recognized as Internato Geral and it had a duration of two years. In the year of 2004 the Internato Geral started to be recognized as Ano Comum and its duration was modified to 1 year (Diário da República, 2004). However, it is important to know, that after concluding the general training of the Internato, a doctor is recognized as an autonomous individual, becoming able to start giving prescriptions to patients and scheduling appointments (D. da República, 2018a). In addition, since 2012 it is now possible for the private healthcare sector to start training medicine students, in order for them to get their post-graduation certificate. However a private healthcare provider has to be first certified by the institution

Ordem dos Médicos (Saúde, 2017). The Healthcare Ministry is the official entity responsible for overseeing the practice of medicine in Portugal, ensuring its quality, security, and ethics (Governo da República Portuguesa, 2017). Additionally, the Administração Central do Sistema de Saúde (ACSS) has established protocols with other entities (Governo da República Portuguesa, 2019) such as Ordem dos Médicos and Conselho Nacional do Médico Interno that helps the Healthcare Ministry to manage the structure and oversee the post-graduation Internato Médico (Ordem Dos Médicos, 2018).

Thus, it should not be appropriate to include individuals that have less than 26 years old in 2017. Because assuming that a person did not make any detour in its academic studies, a doctor has finished its high school degree with 18, finished its college degree with 24 and start its internato with 25 years old. Given that it's only possible to make referrals after concluding the first year of the internato it is not correct to have individuals in the sample of doctors that have 25 or less in 2017. Furthermore, as previously mentioned the objective of the study is to understand the referral mechanism between general practitioners and specialists, therefore, despite being valid to have individuals with 25 years old, they cannot make referrals, because they are in their first year of internato. They have to have at least initiated their second phase of the Internato Médico in 2017. Only then a person can schedule appointments and make prescriptions. By applying these rules only 1 observation is being excluded.

It is important to point out a limitation to the aforementioned rule which is the fact that it is not applicable for cases in which a doctor has more than 26 but it is in fact in its first year of internato. If considered the observations excluded plus the observations being left out, due to inconsistencies issues, the number of records not being considered represents approximately 0.08% of the entire dataset. The values include 2 observation from the exclusion of outliers and 1 observation which do not comply with the consistency rules.

3.3. PRIMARY-SPECIALTY REFERRAL NETWORK

3.3.1. Primary-Specialty Network Mapping

The primary-specialty referral network was built under the purpose of mapping existent relationships between doctors that follow under the category of general practitioners, and physicians that practice specialties that belong to the differentiated care category. In order to do such, it was necessary to import 4 distinct variables: the date of each episode It's unique per visit (represents a doctor appointment with a patient). the unique ids associated with each patient and doctor, regarding each episode, and classification the of each doctor.

The connections estimated to existed in this network, are based on four main assumptions: It cannot be longer than 30 days that between two appointments with different doctors. Only primary care doctors can make referrals. In addition, two distinct doctors have to share a patient. Additionally, in the case of a doctor having one specialty from each category, primary and differentiated care, only the former case was kept. However, if a doctor had two specialties from the same category only the one with the highest represent among the different specialties would remain. Finally, it is not possible to exiting referrals within the same group of doctors.

Initially, the data of the referral network had to be filtered according to the doctors in the final_doctor_table, which contains all the relevant information regarding the doctor's background. Moreover, the data was restricted to episodes that did not have urgency character, that were not

cancelled and in which the date of the episode was available. The final output was a significant large sample with 9 133 477 observations. Afterwards, the dataset was filtered between the years of 2012 and 2017 to ensure a better quality of the data. By doing such, the main dataframe was reduced by 444 observations. Interestingly enough, 137 doctors where identified has not giving a single appointment between those years. Therefore, these doctors will not be taken into account in this research project, leaving us with 2171 doctors in total. However, it would be prudent to verify if such doctors are still active or not. Additionally, in the remaining 9 133 033, it was identified 1 305 361 unique patient ID's. Furthermore, it was registered 1 913 674 unique doctor appointments on behalf of primary care doctors and 4 112 885 on behalf of specialists. Given that the objective is to study the referrals of primary care doctors to specialists, the sample of specialist should only represent doctors who are associated with patients present in the sample of general practitioners.

Unfortunately, the output of the relationships estimated is not valid, as there are multiple doctors which are associated with more referrals than appointments. This means they were estimated to give more appointments than the ones they actually gave. The next best solution found, was to only allow one referral for each unique primary care doctor appointments. The criteria to choose the link that would be assumed as the estimated relationship would the one with the shortest period. This too represents an important limitation to study. Not only, a general practitioner can refer a patient to more than one doctor of the same or different specialties, as well the link with the shortest period is not necessarily the referral that happened in reality. This resulted the estimation of 59 065 unique relationship between doctors.

The next step towards the exploration of the relationship between the two types of doctors was to build the actual network which was done using a library in python called Networkx. Given that it is only possible to have connections between doctors of distinct groups, the network was defined as a bipartite network (figure 3 – Bipartite Network).
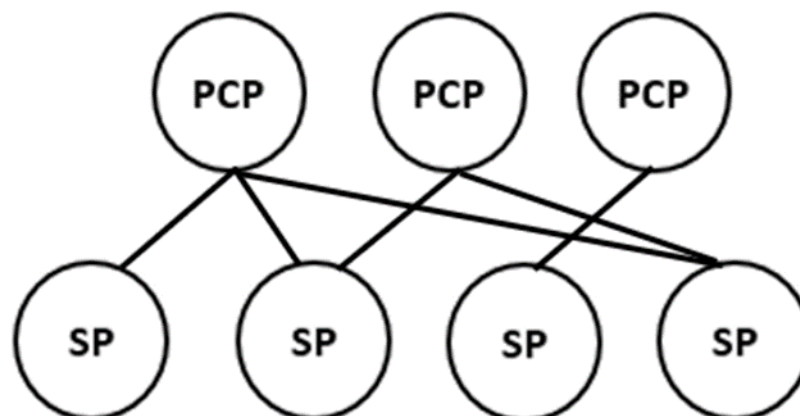

Figure 3 - Bipartite Network

Modelling a social network as a Unipartite network (figure 4 – Unipartite Network) would make it harder to understand the dynamics between the two types of doctors as the analysis and calculations of social network metrics are done are not specific to each type of doctor or do not have in consideration that condition.
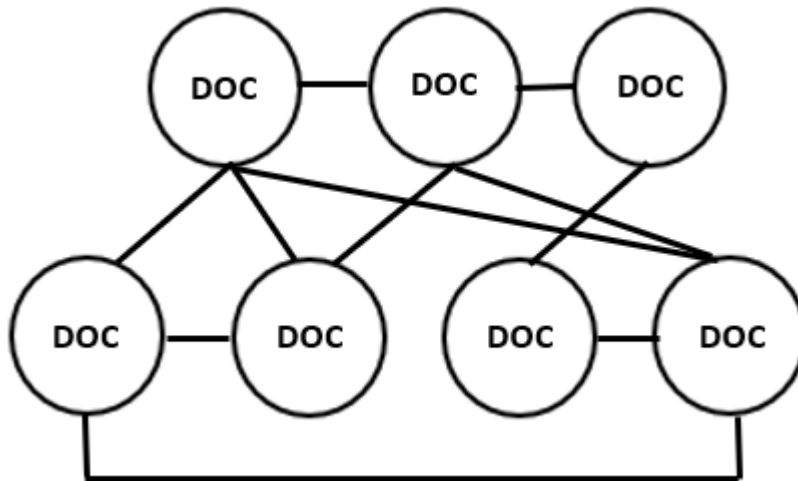
Figure 4 - Unipartite Network

Defining a complex network as two-mode imposes some restrictions and limitation both in the creation and analysis processes. Bipartite networks do not allow to have links within each of the groups in the network, only connections between doctors of different groups are permitted.

Furthermore, the primary-specialty network was computed as a weighted network. In the particular case of this dissertation the weight assigned to an estimated link represents the number of patients each two doctors are expected to share. The nodes with whom there were no connections identified were included in the network grid as isolated nodes. Additionally, this network can be classified as socio-centered, implicit and undirected. To conclude this primary-specialty network does not necessarily corresponds to the formal structure of the private healthcare organization.

3.3.2. Bipartite Random Network

A Bipartite Random network was necessary to build in order to compare the different characteristics of a random network against the macro and micro-level metrics found in the primary-specialty network. It was generated has a Bipartite random graph, where the size of the nodes is the same has the ones in the primary-specialty network. In addition, when creating a network of this nature it is necessary to assign a specific value to a specific parameter called p, which represents the probability of an edge between two nodes being computed. In the equation represented below, μ embodies the weighted average degree of the primary-specialty referral network while symbolizes the number of nodes inserted in the network.

$$p \ = \ \frac{u}{(n-1)}$$

3.3.3. Centrality Measures and Others

3.3.3.1. Degree Centrality

The degree centrality measure in a bipartite network of a doctor is represented by the number of direct ties he or she has, divided by the total possibilities of the opposite partition.

$$\frac{\#\ Direct\ Ties}{\#\ Possible\ Ties\ From\ Opposite\ Patition}$$

On the other hand, if the structure of the network was unipartite, the numerator would remain the same, but in this case the denominator always embodies the same value, which is the total number of nodes in the network.

$$\frac{\#\ Direct\ Ties}{\#\ Possible\ Ties}$$

3.3.3.2. Betweenness Centrality

Betweenness centrality is calculated similarly in both types of networks, bipartite and unipartite. For each node it is the number of occurrences on the shortest paths of each pair of nodes in the network. The main difference when computed in a two-mode network is that this metric is still normalized by the maximum possible betweenness centrality6 values that can be found for each node partition in a bipartite network. The equation for the calculation of such values can be seen in Appendix 9.3 Betweenness Centrality Normalization.

$$CB(v) = \Sigma s, t \in v(\frac{\sigma(s, t\mid v)}{\sigma(s, t)})$$

3.3.3.3. Closeness Centrality

The last centrality measure calculated in this study was the metric Closeness7. The purpose of such metric is to measure the distance of each node to the remaining ones in this network. In addition, given that this metric is being applied in a bipartite network it is normalized by the minimum distance possible, which in the case of a general physician to a specialist is 1 and in the case of doctors within the same partition is 2.

$$Cv = \frac{m + 2(n - 1)}{d}, for\ v \in U,$$

$$Cv = \frac{n + 2(m - 1)}{d}, for\ v \in V$$

3.3.3.4. Global Clustering Coefficient

The global clustering coefficient is a measure that has the purpose of measuring the probability of individuals to cluster together. In a unipartite network, what this metric does is to find all closed triangles in a network (Appendix 7.2 – Closed Triangle) over all triangles that are possible to be formed in a certain network.

However, this is situation is not possible to be observed in a bipartite network as it is not possible to exist closed triangle (Appendix 7.3 – Bipartite Cycles). The alternative solution suggested is to find the

total number of existing squares in the network over all possible squares that could exist in the network given its number of nodes per partition.

$$C4, mn(i) \ = \ \frac{qimn}{(km-\eta imn)+(kn-\eta imn)+qimn}$$

3.3.3.5. Community Detection

The algorithm used to detect possible existent communities in the network was the Louvain algorithm which is an optimization modularity-based algorithm. Communities in social network are no more than groups of nodes that are more densely connected when compared with other community. The Louvain algorithm measures the density between nodes through a metric called modularity. The algorithm starts by defining small communities through the optimization of the aforementioned metric, locally on every node in the sample. Then the communities found are assembled into a single node respectively. The process repeats itself until the optimal number of communities is found, according the algorithm. Applying this algorithm represents a limitation on the current study as the algorithm is not the most appropriate for community detection in bipartite networks.

3.4. PRIMARY-SPECIALTY REGRESSION ANALYSIS

Two distinct regression analysis, representing both types of doctors were performed with the objective of testing the hypothesis that the variables available from the doctor's background and the centrality measures computed from the primary-specialty referral network had any statistical significance in the referral rate that each doctor makes or receives, depending on its specialty. Therefore, the dependent variable (referral rate) is defined as being the number of referrals a doctor has made or received, divided by the number of doctor appointments he or she actually gave between 2012 and 2017. The isolated doctors present were included, with the dependent variable, and centrality measures equal to 0. The initial data can be observed in table 7 – Regression analysis variables. The model used was a multiple linear regression model.

4. RESULTS AND DISCUSSION

The current section has the purpose of critically analyzing the empirical results found. Throughout the current chapter the aforementioned primary-specialty referral network, and regression analysis are used to better understand the existing dynamics between primary care physicans and specialists, and possibly how the relationship between the two of them might be affecting the healthcare provider. Firstly, the primary-specialty referral network is analyzed with the objective of identifying potential referral patterns in the overall network, and then in the communities identified within the network. Finally, the objective of performing a regression analysis was to check if variables like the age, social network centrality measures and other are affecting the number of referrals being made and received by doctors.

4.1. PRIMARY-SPECIALTY NETWORK

A large-scale patient doctor appointments records was obtained from a private European healthcare provider between the years of 2012 and 2017, with 9,173,891 entries, 1.3 million unique patients, 2,171 unique doctors associated to 7 distinct hospitals.

As previously mentioned, a referral between two doctors is estimated to happen when a certain patient goes to see a family doctor and then no later than 30 days he or she meets a specialist for an appointment. The application of this methodology resulted in a weighted, undirected, explicit bipartite network in which 459 general physicians 1,487 specialists are connected. In total 1946 share at least a connection with another doctor. In addition, those doctors are linked by no more than 457,495 links which were converted into 59,065 unique edges.

The actual number of referrals made between each two doctors is measured through the weight associated to that same relationship, as that value represents the number of patients they share. In addition, given that only 1,946 of the 2,171 doctors are estimated to be connected this implies that 225 doctors, approximately 11% of the sample of physicians is expected to be isolated. In these 225 doctors that are estimated as not sharing connections, 14 of them belong to primary care and 211 to differentiated care. Moreover, 89 doctors, 16 general practitioners, and 73 specialists are peripheral nodes in this primary-specialty referral network. This means that they only share a link with one doctor in the entire network. However, despite not being possible to ensure that this is actually the case, it still raises the possibility that those resources are not being optimized.

In addition, it was necessary to prove that the primary-specialty network does not exhibit characteristics of randomness, meaning that it does not have a similar structure of a random network. For that, it was tested if the weighted degree distribution of both groups of doctors followed or not a Poisson distribution. The results of both analyses returned a p-value of 0, allowing us to reject the hypothesis that both samples followed a Poisson distribution. Additionally, the average clustering coefficient was computed on both networks, on the Bipartite random network and on the primary-specialty. The average clustering coefficient computed on the doctor's network (0.0206364) revealed to be approximately 21 times higher than the one computed on the Bipartite random network (0.0009752).

Furthermore, the human resources department of the healthcare provider was able to provide several important information's regarding doctor's background. It was possible to obtain their age, their level

of medical expertise (number of years of practice), the specialty or specialties in which they are certified, their education, and the institution at which they did their post-graduation also known as "internato médico".

4.1.1. Primary-Specialty Referral Patterns

Gender

In terms of gender distribution, the doctors belonging to the specialist group are more or less evenly distributed with a ratio of 44% compared to 56% male physicians. The scenario changes when considering the primary care doctors as approximately 69% are female and only 31% are male (Appendix 7.4 – Gender Distribution per Doctor Classification).

The distribution of the referrals per gender shows a preference for the male gender as 60% of the total of referrals was received by male specialists. From the total amount of referrals, the female gender has referred its own 26% of the times. But when referring the opposite gender this one represented 38% of the total referrals. In the scenario where the male gender is the one making referrals, women are referred 14% the times. The remaining 22% of the total referrals are made between men.

In what concerns the 225 isolated nodes, on the general practitioner's side, the female-male ratio does not change, as it is 71%-29% respectively. Finally, the isolated specialists have a female-male ratio of 42%-58% respectively (Appendix 7.5 – Isolated Gender Distribution per Doctor Classification).

Specialty

From the sample of doctors under analysis 44 distinct specialties are identified. As it turns out all of them are referred at least one time. Meaning they are all represented in the primary-specialty referral network. In the particular group of doctors representing the primary care, the specialty Medicina Geral e Familiar is the most active one (48%), while Medicina Interna (28%) and Pediatria (24%) are responsible for approximately the same number of referrals (Appendix 7.6 - Primary Care Referrals Distribution).

According to the Portuguese national healthcare service the most common diseases that are currently affecting the Portuguese population are: Skin related diseases, depression, obesity related, and diseases such as low back and neck pain (Retrato da Saúde 2018, 2018). It is interesting to observe that most of the conditions aforementioned can be treated by at least one of the specialties in the of the most referred specialties (f Appendix 7.7 – Differentiated Care Referrals Distribution). For example, the conditions of neck and low back can be treated by resorting to the specialties Ortopedia and Radiologia. Skin related conditions can be treated by the specialty Dermato-Venorologia. Morever, in terms of hearing and visual capabilities according to the portuguese national program "Para a Saúde da Visão ", half of population in Portugal does not see properly (Público, 2005). Additionaly, according to a study called "Coping with noise ", Portugal is the second country with worst hearing capabilities (Amplifon, 2017). These two conditions can be treated by the specialties Oftamologia and Otorrinolaringologia. Furthermore, the number of births in private healthcare as duplicated from 2000 to 2017, which is interesting given that it the fourthest most referred specialty, Ginecologia/Obstetrícia.

32

In what concerns the isolated nodes the distribution of the specialties is interesting to see how the picture inverts (Appendix 7.8 – Isolated Primary Care Specialties Distribution). The specialty with the highest representation in the group of isolated general practitioners, Pediatria, is the one with the lowest number of referrals. Additionally, the specialty with the highest number of referrals made, has the lowest representation in this group of isolated doctors.

The specialties represented in the group of doctors belonging to differentiated care (Appendix 7.9 – Isolated Differentiated Care Specialties Distribution), are all of them represent in the class "Top Specialties" of Appendix 7.7 – Differentiated Care Referrals Distribution, with exception of the specialties "Angiologia e Cirurgia Vascular", "Cirurgia Plástica Reconstrutiva e Estética", and "Neurocirurgia". Once more, this raises the question if the healthcare provider is being efficient in its decision making, it might not need that many professionals representing those specialties. However, it is not possible to ensure that this is the case in reality, therefore further analyses are required.

Hospital

The distribution of doctors amongst the different hospitals can produce valuable insights as it is important to understand if the referrals between the two groups of professionals are being kept inside each healthcare institutions not. In the primary-specialty referral network approximately 76% of the referrals estimated only have one hospital in common. Therefore, it is plausible to assume that the vast majority of the referrals is kept inside each of the healthcare institutions. However, it is only possible to ensure that 37% of the distinct belong to the same hospital. Additionally, the data shows that in approximately 19% of the mapped relationships, physicians did not have hospitals in common (Appendix 7.10 - Number of Hospital In Common).

Age

In terms of age difference, it is possible to conclude that close to 80% of the referrals found amongst the two type share an age difference inferior or equal to 20. Additionally, more than 40% of them do not have more than 10 years of difference. This might suggest a certain level of homophily. From Appendix 7.11 – Age Difference Distribution ECDF it appears that higher the age difference the less likely is that a specialist will be referred.

Perceived Clinical Expertise

By observing the distribution of the specialists being referred in Appendix 7.12 – Specialists Perceived Clinical Expertise, the number of specialists in with more than 20 years of medical practice are practically none, in fact they represent less than 10% of the doctor's being referred. In addition, nearly 80% of them does not have more that 15-years practicing medicine. Given that more or less 85% of the entire sample of doctors does not have more than 15-years practising their profession it is hard to state a pattern in terms of referring specialists due to their level of perceived experience.

In addition, it was not associated any justification that would further help comprehend the situation of the isolated nodes in the sample (Appendix 7.13 – Isolated Age & Perceived Clinical Expertise ECDF). However, there are two individuals that distiante themselves from the overall sample of isolated doctors in terms of age, and one in terms of years of practice, which might be a contributive factor for being left out from the primary-specialty referral network.

4.1.2. Primary-Specialty Centrality Measures

Three distinct social network centrality measures were computed after instantiating the primary-specialty referral network. Those measures were: Degree centrality, Betweenness Centrality and Closeness Centrality. In terms of the centrality measure degree (f Appendix 7.14 – Degree Centrality Distribution), it is possible to observe that both samples have a skewed distribution to the right. Furthermore, as the number of referrals increases the number of primary care doctors representing them decreases. A similar situation happens in the specialist distribution. However, the situation is not so much intense, meaning this distribution has a smaller tail. In addition, it was tested the possibility that the weighted degree distribution of both groups could be represented by power-law distribution (null-hypothesis). In the case of general practitioners this would mean that in a certain distribution only a few nodes are responsible for a large proportion of the referrals being made. For specialist doctors it would mean that only a few doctors are receiving a significant percentage of the referrals being made. It was only possible to prove that the primary care physicians group follows a power-law distribution, as the null-hypothesis cannot be rejected, given its p-value of 0.23.

Furthermore, in terms of the betweenness (Appendix 7.15 – Betweenness Centrality Distribution) there is a significant number of individuals on both distributions with values of zero, which means that they aren't on the shortest path of any pair doctors. In addition, the values registered in this metric are extremely low. In fact, the highest values detected in the specialists and general practitioners are 0.007 and 0.035 respectively. It is also, interesting to notice that in general, family doctors have higher values in this metric when compared to specialist. For example, whereas 63% of the general practitioners have values superior to 0.0001, specialists only have 36% of their doctors represented.

36

In what concerns the metric closeness the distribution of both type of doctors is significantly different from the other centrality measures (Appendix 7.16 – Closeness Centrality Distribution8). The only doctors with closeness values equal to zero are the isolated nodes. The higher the values of closeness the more central nodes are in the network. It is interesting to see how the picture has inverted when comparing with the metric degree and betweenness centrality. According to the distributions of the metric closeness, specialists have a more central role in the network, meaning that in general they have more access to the rest of the doctors in the network. This might be happening, because only a few primary care doctors are responsible for most of the referrals being made, and thus, those doctors might be referring the same specialists, to some extension.

It has also been quantified the number of referrals made and received by the individuals that represents top and bottom 25% highest values for all the three centrality metrics (degree, betweenness, and closeness), regarding the two type of doctors. In case of the metric degree, the top 25% individuals are held responsible for 78.9% of the referrals, while the bottom 25% made only 0.33% of the referrals. The scenario, similar for the specialist doctors as the 25% individuals for having more direct ties received 77.8% of the referrals, while the bottom 25% received only 0.08% of the referrals. Moreover, the situation does not change for the remaining centrality measures. Betweenness centrality shows a 25% top-bottom ratio for primary care doctors of 77.8% - 0.34% and a top-bottom ratio for specialists of 73.3% - 0.1%. The metric closeness reveals a top-bottom 25% ratio for general practitioners of 79.2% - 1.4% and a top-bottom 25% ratio of 68.6% - 2.62%. Once again, the issue that

the primary-specialty referral network is not being efficient is raised as the referral process appears to be ensured by only a few doctors, both on the making and receiving side of the referrals.

4.1.3. Primary-Specialty Community Analysis

Communities in network science represent groups of nodes, which in this case are doctors who have a denser network amongst them. Additionally, the connections between groups are expected to be less condensed. They are important structures to be analyzed, as they help to understand the dynamics of complex networks. The algorithm used for the detection of communities in this dissertation was the Louvain, a modularity-based optimization algorithm, which considered the weights of each relationship.

In total, the algorithm estimated the existence of 7 distinct communities. This is interesting as 7, corresponds to the number of unique hospitals that are in our sample (Appendix 7.17– Primary-Specialty Communities). In addition, it is possible to observe the communities individually in Appendix 9.10 – Individual Communities. It is also possible to observe the communities' structure in Appendix 9.11 – Communities Structure.

From community 1 to community 7 the respective sizes of each of them is: 634, 80, 191, 198, 384, 322, 137. It is interesting to observe that in each community, there is one hospital that stands out in terms of representation. This might suggest that each community is representing a unique hospital of the healthcare provider. In the case of community 3, the hospital with the highest representation is 8, but the concentration of individuals working in hospital 8 is significantly higher in community 5. Therefore, group 5 is associated with hospital 8. Thus community 3 will be associated with hospital 15 (Appendix 10 – Communities Structure). Although it is important to bear in mind that the particular situations of communities 3 and 4 are not so clear, when having to assign a hospital to them. It assumed that community 4 represents hospital 12.

Furthermore, the gender distribution amongst all communities is similar and more or less evenly distributed. Only community in community 3 it is verified a big disparity between the two genders as the female-male ratio for that group is 74%-26%. It also the only community where the number general practitioners and specialist nearly the same, which might be a contributing factor, as the female population is significantly higher in doctor certified in primary care.

In addition, the top 5 specialities with the highest representation in all communities produces a list of 13 unique specialties in which, Ginecologia/Obstetrícia, Ortopedia are represented in five communities and Oftamologia, Medicina Geral e Familiar and Anestesiologia are represented amongst 4 different communities.

In what concerns the distribution of the ages of doctors along the distinct communities this shows a very low variance for the metrics calculated (minimum age, average of age, and maximum age). Meaning that all groups estimated by the Louvain algorithm present a similar age structure. Moreover, in terms of perceived clinical expertise the situation is equivalent with exception of the maximum clinical expertise registered in the different communities. For example, in community 2 the doctor with most experience has 16 years on the job, while the most experienced doctor in community 6 has 39 years of medical practice. Even so, on average all communities reveal similar values.

5. CONCLUSIONS

The development of this dissertation is expected to add more insights and new methodologies to the existent literature concern about the referral mechanism between general practitioners and specialists. From a large-scale patient doctor appointments dataset filtered between the years of 2012 and 2017, provided, by a private European healthcare institution, it was obtained a sample with more than 9 million entries, 1,305,361 unique patients, 2171 unique physicians associated to 7 different hospitals. From the overall sample of physicians in the sample 22% are general practitioners and 78% are specialists

The output of the estimated primary-specialty referral network accounted for all the 2171 doctors but only 1946 are expected to share at least one connection with another doctor from these, 459 are primary care doctors and 1487 belong to differentiated care. The remaining 225 of the sample represent isolated doctors, 14 are family doctors and the other 211 doctors are specialists. For these doctors it was not found a doctor with whom they might share a connection. Additionally, within the referral network there are 89 doctors who are identified as having unique relationships with whom they only have one patient in common. Therefore, without making any conclusions that group of doctors may represent underutilized resources for the healthcare provider. Thus, it would be in its best interest to make further analysis to settle if such statements are factual or not. In total the primary-specialty network is comprised of 59,065 distinct relationships that embody 457,495 referrals. The weight of each unique link characterizes the actual number of patients that two doctors share.

In addition, besides the doctor appointments records, it was provided the information regarding doctors' background by the Human Resources department. Such information as doctors' age, years of practice, hospital in which they are currently working in, and the higher learning institution where they got their university degree were given.

The gender composition for primary care doctors is 69% of female and 31% male, and 44% of female and 56% of male for specialists. Women show to be more likely in pursuing a career in the primary care than men. Additionally, it is also interesting to notice that in general the percentage of women in specialties related with childcare is higher. Whereas men, show higher concentration in areas related with surgical fields. In fact, in all the 9 surgical related fields recognized as a specialty in Portugal, including cirurgia pediátrica women are outnumbered. In terms of the 4 specialties including only childcare only, women have a stronger presence in 3.

From the total referrals estimated, approximately 60% of them were to the male. In addition, the distribution of the references per gender shows higher propensity for women to refer men and not the other way around. From the total of referrals that are originated from women (63.9%), 59.9% of the total referrals made are to men. However, from the total of referrals made by men (36.1%) only 38.8% of them were to women. Additionally, it is detected a certain level of homophily in the referrals established. Homophily is a phenomenon that states that individuals have higher tendency to be connected with individuals with similar characteristics. Approximately 48% of the total referrals to the same gender. Moreover, close 80% of the references made are made between doctors with an age difference smaller or equal to 20 and more than 40% were between individuals with an age difference smaller or equal to 10. More interestingly, the majority of the referrals made appear to be kept inside

each of the 7 hospitals, as 76% of the total referrals were made amongst doctors that only have 1 hospital in common. Additionally, approximately 18% of the referrals do not have a single hospital in common. However, only close to 37% of the referrals made between two doctors can be ensured that they represent a unique hospital.

Furthermore, from all the referrals made to specialists more than 60% were made to doctors that claim to have no more than 10 years of medical experienced. However, given that approximately only 4.37% of the doctors who are connected specialists have more than 10 years of experience it might be plausible to infer but not conclude that experience of a doctor is relevant for the referral decision process as 40% of the total referrals made were to doctors with more 10 years of experience.

In what concerns the specific metrics of network science three metrics were computed based on the primary-specialty referral network: Degree Centrality, Betweenness Centrality and Closeness Centrality. In the case of general practitioners, it was possible to prove that only a handful of the primary care doctors is responsible for a significant percentage of the referrals mapped throughout the network. In fact, when selecting the 25% individuals with highest values of degree centrality they are responsible for approximately 79% of all the estimated referrals. Moreover, when quantifying the number of referrals made by the 25% with the highest values in terms of betweenness and closeness centrality they are responsible for 78% and 79% of all the referrals respectively. Therefore, a high percentage of doctors is represented on top 25% of the three centrality measures. In terms of the specialists, despite not having been proven that only small group of the doctors belonging to differentiated care, by selecting the top 25% of each network centrality measure, degree, betweenness and closeness centrality, those doctors represent 78%, 73% and 69% of the total referrals respectively. Therefore, according to these results the primary-specialty referrals network, once more, appears to not be efficient. This suggest that only a few primary care doctors and specialist are being responsible for the referral system.

In terms of communities, through the application of the Louvain algorithm, 7 distinct communities were identified. In fact, 7 is the number of hospitals present in the sample. In what concerns the structure of the communities they show to be similar in age, perceived clinical expertise, specialty distribution, gender and doctor classification. However, community 3 represents the exception as it verified an unbalanced gender distribution with a female-male ratio of 74%-26%. Moreover, in case the different communities are actually representing each unique hospital in the sample, it is possible to support once more that the majority of the referrals is kept inside each community. Considering that on average approximately 30% of the referrals made by a community is to outside, community 2 (58%) and 7 (38%) stand out. Moreover, it is interesting to notice that the only hospital that operates in a different region from the other six hospitals is the most isolated one as only 0.34% of the referral with origin in that community were not kept inside it. Even so, from the 101,274 referrals they received, 22.14% were made from other communities. This might indicate that geographic distance between hospitals plays an important role in the referral decision process. Moreover, is common for doctors to guarantee that the referrals sources remain happy with them especially major cities and in particular metropolitan areas (Jauhar, 2019).

It is also interesting to noticed that in the top 25% of doctors that make more referrals outside their community and the doctors that receive more referrals from outside their community, is partially represented by the doctors that have the most relevant positions in terms of degree, betweenness

and closeness centrality. This might suggest that the number of direct ties a doctor has in the network, the number of times it appears on the shortest paths of other nodes and the position it holds in the network might have an impact in the referral rate of a doctor. Moreover, the statement that the male gender in the specialists has represents a higher percentage of the referrals is evidenced once more.

Finally, by performing a regression analysis on both types of doctors' classification it was possible to conclude within our sample if the background and network metrics have or do not have an impact in the referrals rates associated with each doctor. In the case of primary care from the 11 variables included in that regression 7 were determined as having a statistical impact in the referral rate of general practitioner. The results concluded that the variables age, gender, closeness centrality, working in 2 hospital, having only 1 specialty, being certified in primary care influences the referral rate of a doctor. It is interesting the fact that variable closeness centrality can be the second variable with more impact in the referral rate of a doctor, which depending on its values, it signals the position of the doctor in the network, how much access it has to other doctors. However, the one with the most significant impact is being certified in "Medicina Interna".

In the case of the specialists care the results were not so promising. However, the results within our sample show that the age and closeness centrality have impact in the referral rate of doctor belonging to differentiated care. In addition, from the 40 specialties taken into consideration in the specialists' group, only cirurgia pediátrica revealed to have an impact (positive) in the referral rate of a doctor.

In both regression it was not found evidence that the number of connections a doctor has, or the number of times it appears on the shortest paths of other two doctors, to have any influence in the referral rate of a doctor.

6. BIBLIOGRAPHY

American Academy of Family Physicians. (2016). Primary Care -- AAFP Policies, 1. Retrieved from http://www.aafp.org/about/policies/all/primary-care.html#1

An, C., O'Malley, A. J., & Rockmore, D. N. (2018). Referral paths in the U.S. physician network. *Applied Network Science*, *3*(1). https://doi.org/10.1007/s41109-018-0081-4

An, C., O'Malley, A. J., Rockmore, D. N., & Stock, C. D. (2018). Analysis of the U.S. patient referral network. *Statistics in Medicine*, *37*(5), 847–866. https://doi.org/10.1002/sim.7565

B.E., L., N.L., K., M.L., B., J.-P., O., S., P., A.J., O., … N.A., C. (2012). Variation in patient-sharing networks of physicians across the United States. *JAMA - Journal of the American Medical Association*, *308*(3), 265–273. https://doi.org/10.1001/jama.2012.7615

Baganha, M. I., Ribeiro, J. S., & Pires, S. (2002). O Sector da Saúde em Portugal: Funcionamento do Sistema e Caracterização Sócio-Profissional. Retrieved from https://ces.uc.pt/publicacoes/oficina/ficheiros/182.pdf

Barnett, M. L., Keating, N. L., Christakis, N. A., O'Malley, A. J., & Landon, B. E. (2012). Reasons for choice of referral physician among primary care and specialist physicians. *Journal of General Internal Medicine*, *27*(5), 506–512. https://doi.org/10.1007/s11606-011-1861-z

Barnett, M. L., Landon, B. E., O'Malley, A. J., Keating, N. L., & Christakis, N. A. (2011). Mapping physician networks with self-reported and administrative data. *Health Services Research*, *46*(5), 1592–1609. https://doi.org/10.1111/j.1475-6773.2011.01262.x

Barnett, M. L., Song, Z., & Landon, B. E. (2012). Trends in Physician Referrals in the US, 1999–2009. *Archives of Internal Medicine*, *172*(2), 163–170. https://doi.org/10.1001/archinternmed.2011.722

Bocian, A., Baker, A. E., Forrest, C. B., von Schrader, S., Starfield, B., & Glade, G. B. (2013). Coordination of Specialty Referrals and Physician Satisfaction With Referral Care. *Archives of Pediatrics & Adolescent Medicine*, *154*(5), 499. https://doi.org/10.1001/archpedi.154.5.499

Care, P., Referral, P., Means, G., & Problems, M. (2019). Primary Care Physicians and Specialists Referral Gap Means Marketing Problems, 1–6.

Center, I. K. (2017a). Betweenness. Retrieved from https://www.ibm.com/support/knowledgecenter/SS3J58_9.0.6/com.ibm.i2.anb.doc/about_betweenness.html

Center, I. K. (2017b). Eigenvector. Retrieved from https://www.ibm.com/support/knowledgecenter/SS3J58_9.0.6/com.ibm.i2.anb.doc/about_eigenvector.html

Chen, J. E., & Glover, G. H. (2016). HHS Public Access, *25*(3), 289–313. https://doi.org/10.1007/s11065-015-9294-9.Functional

Clark, L. (2006). *Network Mapping as a Diagnostic Tool Manual*.

Cross, R. (2017). Introduction to Organizational Network Analysis, (July 2011), 1–6.

Cross, R., Parker, A., & Borgatti, S. (2000). A bird's-eye view: Using social network analysis to improve knowledge creation and sharing. *Knowledge Directions*, *2*, 48–61.

https://doi.org/10.2307/1315064

Deloitte. (2019). *2019 Global Health Care outlook: Shaping the Future*. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Life-Sciences-Health-Care/gx-lshc-2016-health-care-outlook.pdf%5Cnhttp://www.deloitte.com/healthcareoutlook.

Diário da República. Portaria n.o 1498/2004, 2002 § (2004). Retrieved from http://data.dre.pt/eli/port/1499/2004/12/28/p/dre/pt/html

Directions, F. (n.d.). SOCIAL NETWORK ANALYSIS. https://doi.org/10.3366/saj.2004.26.1-2.vi

Faust, K. (2006). Comparing Social Networks : Size , Density , and Local Structure, *3*(2), 185–216. https://doi.org/10.1002/smj.444

Forrest, C. B., Nutting, P. A., Von Schrader, S., Rohde, C., & Starfield, B. (2006). Primary care physician specialty referral decision making: Patient, physician, and health care system determinants. *Medical Decision Making*, *26*(1), 76–85. https://doi.org/10.1177/0272989X05284110

Franks, P., Williams, G. C., Zwanziger, J., Mooney, C., & Sorbero, M. (2000). Why do physicians vary so widely in their referral rates? *Journal of General Internal Medicine*, *15*(3), 163–168. https://doi.org/10.1046/j.1525-1497.2000.04079.x

Frey, D., Jégou, A., & Kermarrec, A. M. (2011). Social market: Combining explicit and implicit social networks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6976 LNCS*, 193–207. https://doi.org/10.1007/978-3-642-24550-3_16

Ghomrawi, H. M. K., Funk, R. J., Parks, M. L., Owen-Smith, J., & Hollingsworth, J. M. (2018). Physician referral patterns and racial disparities in total hip replacement: A network analysis approach. *PLoS ONE*, *13*(2), 1–10. https://doi.org/10.1371/journal.pone.0193014

Governo da República Portuguesa, M. da S. (2017). Ministério da Saúde – SNS. Retrieved from https://www.sns.gov.pt/institucional/ministerio-da-saude/

Governo da República Portuguesa, M. da S. (2019). Internato Médico - ACSS. Retrieved from http://www.acss.min-saude.pt/2016/09/12/internato-medico/

Granja, M., Ponte, C., & Cavadas, L. F. (2014). What keeps family physicians busy in Portugal? A multicentre observational study of work other than direct patient contacts. *BMJ Open*, *4*(6), 1–10. https://doi.org/10.1136/bmjopen-2014-005026

IBM Knowledge Center. (2017). Closeness. Retrieved from https://www.ibm.com/support/knowledgecenter/SS3J58_9.0.6/com.ibm.i2.anb.doc/about_closeness.html

Intelligence, C. (2015). EigenCentrality & PageRank. Retrieved from https://cambridge-intelligence.com/eigencentrality-pagerank/

Irving, G., Neves, A. L., Dambha-Miller, H., Oishi, A., Tagashira, H., Verho, A., & Holden, J. (2017). International variations in primary care physician consultation time: A systematic review of 67 countries. *BMJ Open*, *7*(10), 1–15. https://doi.org/10.1136/bmjopen-2017-017902

Jauhar, S. (2019). Referral System Turns Patients Into Commodities, 3–5.

Kenis, P., & Oerlemans, L. (2009). The Social Network Perspective: Understanding the Structure of Cooperation. *The Oxford Handbook of Inter-Organizational Relations*, (April 2015).

https://doi.org/10.1093/oxfordhb/9780199282944.003.0011

Kinchen, K. S., Cooper, L. A., Levine, D., Wang, N. Y., & Powe, N. R. (2004). Referral of patients to specialists: factors affecting choice of specialist...: EBSCOhost. *Annals of Family Medicine*, *2*(3), 245. https://doi.org/10.1370/afm.68

Larson, J. S. (1996). The World Health Organization's Definition of Health: Social versus Spiritual Health. *Springer*, *38*, 181–192. Retrieved from https://www.jstor.org/stable/27522925?seq=1#page_scan_tab_contents

Liddy, C., Arbab-Tafti, S., Moroz, I., & Keely, E. (2017). Primary care physician referral patterns in Ontario, Canada: A descriptive analysis of self-reported referral data. *BMC Family Practice*, *18*(1), 1–8. https://doi.org/10.1186/s12875-017-0654-9

Notícias, D. de. (2019). Ministra assume que deve falhar meta de dar médico de família a todos. *Diário de Notícias*. Retrieved from https://www.dn.pt/pais/interior/ministra-assume-que-deve-falhar-meta-de-dar-medico-de-familia-a-todos-os-portugueses-10730984.html

O'Malley, Ann S., MD, M., & Reschovsky, J. D. (2011). Referral and Consultation Communication Between Primary Care and Specialist Physicians. *Archives of Internal Medicine*, *171*(1), 56–65.

Ordem Dos Médicos. (2018). Ordem dos Médicos – Portal Oficial. Retrieved from https://ordemdosmedicos.pt/

Ponte, C., Moura, B., Cerejo, A., Braga, R., Marques, I., Teixeira, A., … Trancoso, J. (2006). Referenciação aos cuidados de saúde secundários. *Revista Portuguesa de Clínica Geral*, *22*(5), 555–568. https://doi.org/10.32385/RPMGF.V22I5.10282

República, D. A. Lei n.º 117/2015 (2015).

República, D. da. Decreto-Lei n.º 13/2018 (2018). Governo de Portugal. Retrieved from https://dre.pt/web/guest/home/-/dre/114766032/details/maximized?serie=I&print_preview=print-preview&day=2018-02-09&date=2018-02-01

República, D. da. Portaria n.º 78/2018 (2018). Retrieved from https://dre.pt/application/file/a/114880161

Saúde, J. de M. (2017). CUF forma os primeiros médicos internos do setor privado _ CUF. Retrieved from https://www.saudecuf.pt/cuf/noticias/cuf-forma-os-primeiros-medicos-internos-do-setor-privado

Schroeder, M. (2016). What's Behind That Medical Referral?, 1–8. Retrieved from http://health.usnews.com/health-news/patient-advice/articles/2016-04-07/whats-behind-that-medical-referral

Song, Y., B. Keller, F., & Zheng, L. (2017). *Social Network Analysis Methods and Examples*. *Sage* (Vol. 0249). Retrieved from https://us.sagepub.com/sites/default/files/upm-binaries/78651_Chapter_1.pdf

Teixeira, F. (2018). Médicos de família fazem em média 21 consultas por dia. *Público*. Retrieved from https://www.publico.pt/2018/07/13/sociedade/noticia/medicos-de-familia-fazem-21-consultas-por-dia-em-media-1837427

University, S. G. (2018). What Does a Primary Care Physician Do? Exploring This High-Demand Medical

Career. *MEDICAL PRACTICE*. Retrieved from https://www.sgu.edu/blog/medical/what-does-a-primary-care-physician-do/

Valente, T. W. (2010). *Social Networks and Health: Models, Methods, and Applications*. Retrieved from https://books.google.pt/books?id=xnMzd1-7iGgC&pg=PA135&lpg=PA135&dq=network+diameter+The+maximum+distance+between+all+pairs+of+nodes&source=bl&ots=YrsS0oM6YG&sig=ACfU3U14FMJwZ5b32xMrbFHbHMkE8y2bSw&hl=pt-PT&sa=X&ved=2ahUKEwj_r-Pm-KzhAhWt1eAKHXGfDZ0Q6AEwB3o

WHO. (2017). World Bank and WHO: Half the world lacks access to essential health services, 100 million still pushed into extreme poverty because of health expenses. Retrieved from https://www.malariaworldcongress.org/in-the-news/2018/1/11/world-bank-and-who-half-the-world-lacks-access-to-essential-health-services-100-million-still-pushed-into-extreme-poverty-because-of-health-expenses%0Ahttp://www.who.int/news-room/detail/13-12-201

7. APPENDIX

7.1. SPECIALTIES FREQUENCY

Specialty	Frequency
Ginecologia/Obstetrícia	263
Pediatria	215
Anestesiologia	180
Ortopedia	161
Oftalmologia	130
Medicina Geral e Familiar	118
Cirurgia Geral	118
Medicina Interna	116
Otorrinolaringologia	96
Radiologia	82
Urologia	75
Cardiologia	62
Gastrenterologia	59
Neurocirurgia	58
Estomatologia	55
Cirurgia Plástica Reconstrutiva e Estética	51
Neurologia	50
Dermato-Venereologia	45
Angiologia e Cirurgia Vascular	43
Psiquiatria	32
Pneumologia	32
Imunoalergologia	32
Endocrinologia e Nutrição	23
Oncologia Médica	23
Cirurgia Pediátrica	21
Reumatologia	20
Cirurgia Maxilo-Facial	16
Anatomia Patológica	15
Nefrologia	10
Others[6]	62

---

[6] Every Category with a frequency lower or equal to 9 was inserted in the category others. The specialties included in this segment are: Radioncologia; Psiquiatria da Infância e da Adolescência; Cirurgia Cardiotorácica; Medicina Nuclear; Cardiologia Pediátrica; Neurorradiologia; Medicina do Trabalho; Cirurgia Cardíaca; Medicina Desportiva; Cirurgia Torácica; Hematologia Clínica; Genética Médica; Patologia Clínica; Doenças Infecciosas and Imunohemoterapia.

7.2. CLOSED TRIANGLE
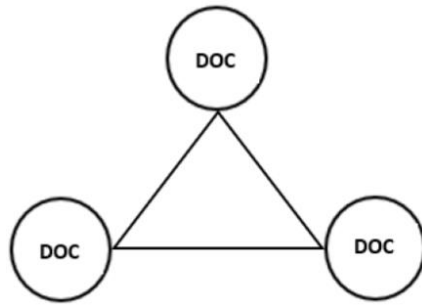
7.3. BIPARTITE CYCLES
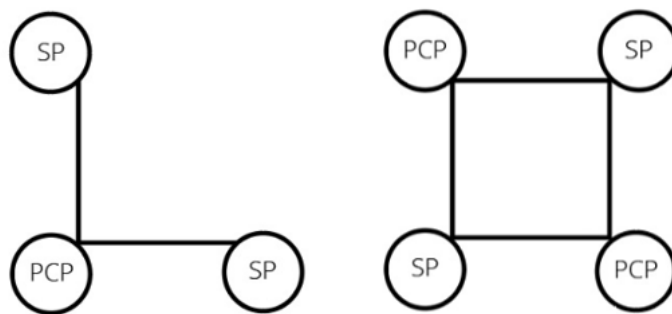
7.4. GENDER DISTRIBUTION PER DOCTOR CLASSIFICATION
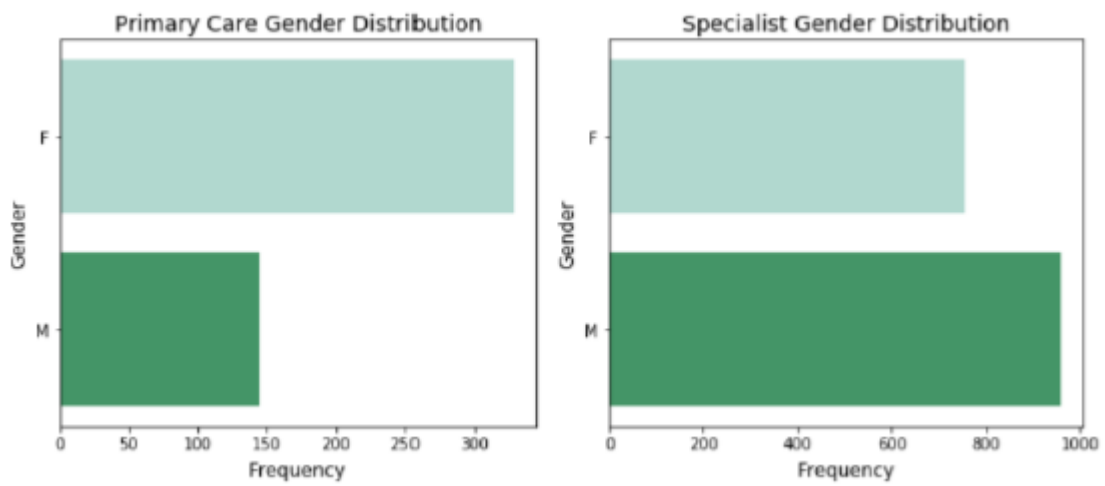
7.5. GENDER DISTRIBUTION PER ISOLATED DOCTOR CLASSIFICATION
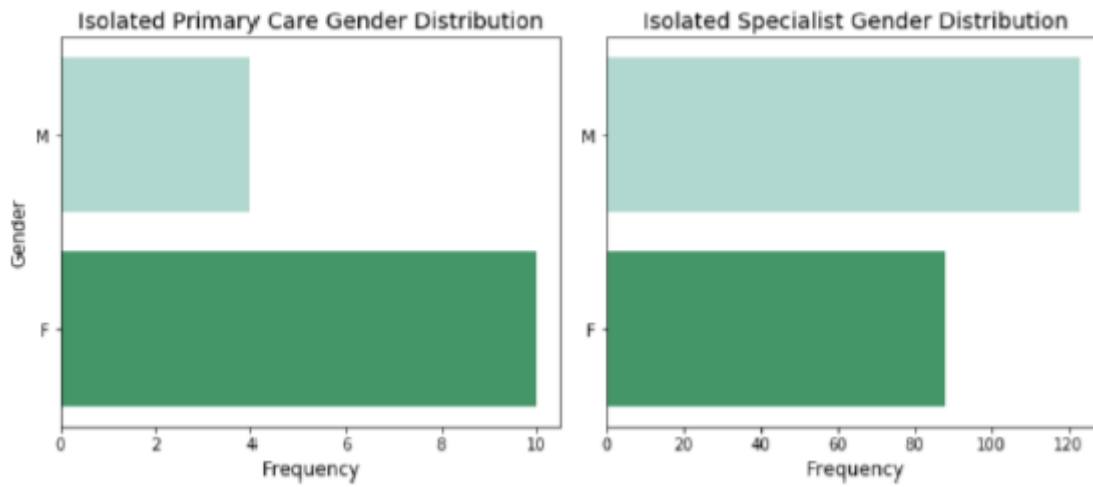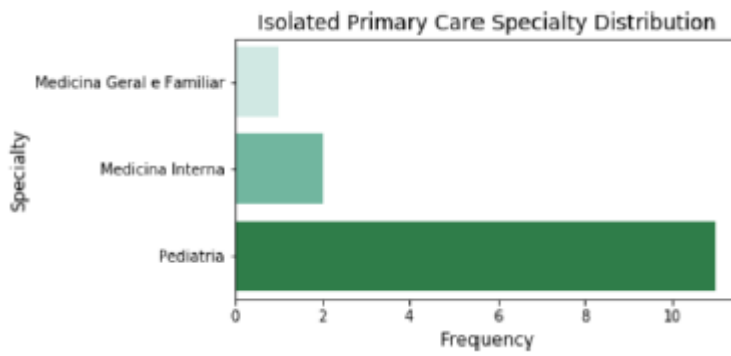

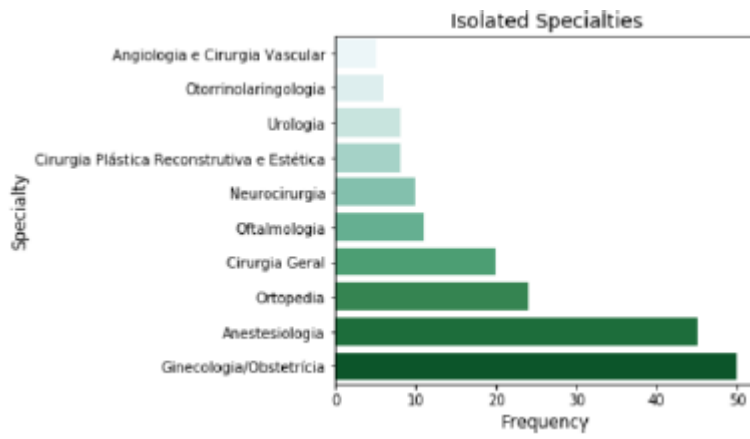## 7.6. PRIMARY CARE REFERRALS DISTRIBUTION
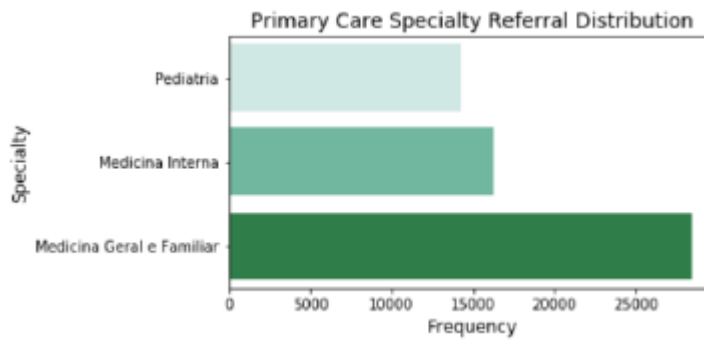
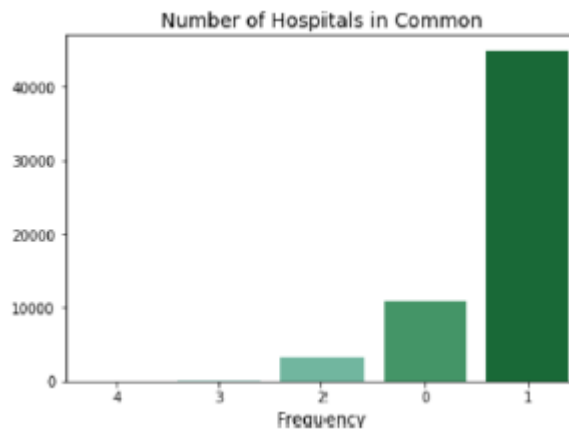## 7.7. DIFFERENTIATED CARE REFERRALS DISTRIBUTION
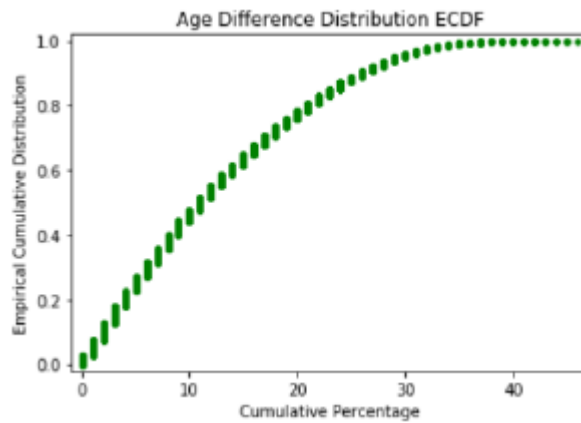
7.8. ISOLATED PRIMARY CARE REFERRALS DISTRIBUTION
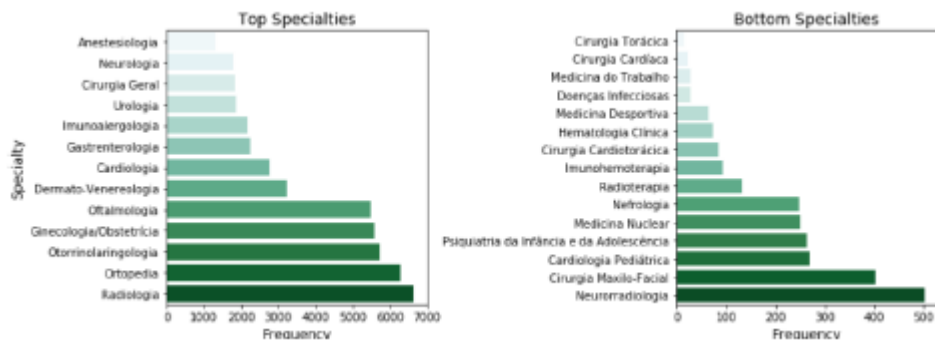
Isolated Primary Care Specialty Distribution

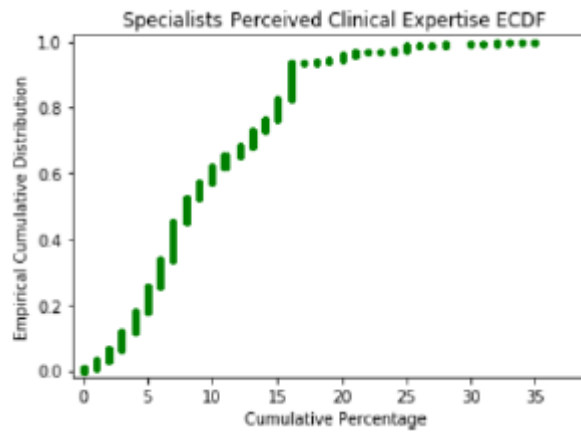## 7.9. ISOLATED DIFFERENTIATED CARE REFERRALS DISTRIBUTION

Isolated Specialties

7.10. NUMBER OF HOSPITALS IN COMMON

Number of Hospitals in Common

7.11. AGE DIFFERENCE DISTRIBUTION ECDF

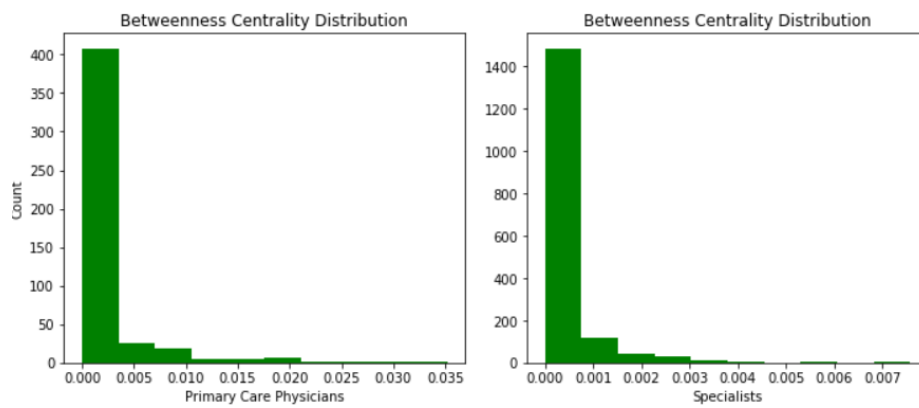## 7.12. SPECIALISTS PERCEIVED CLINICAL EXPERTISE

7.13. ISOLATED AGE & PERCEIVED CLINICAL EXPERTISE
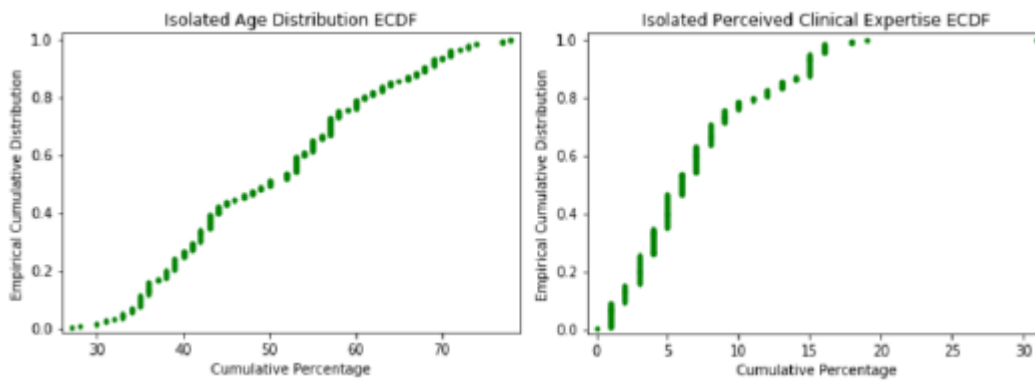
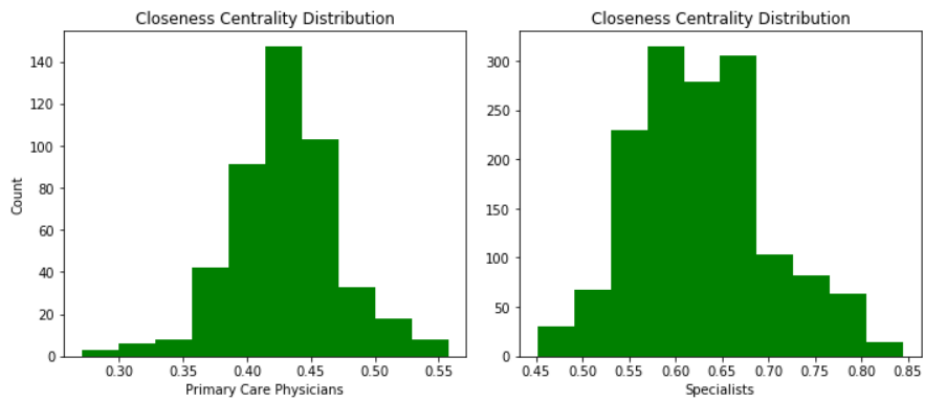## 7.14. DEGREE CENTRALITY DISTRIBUTION
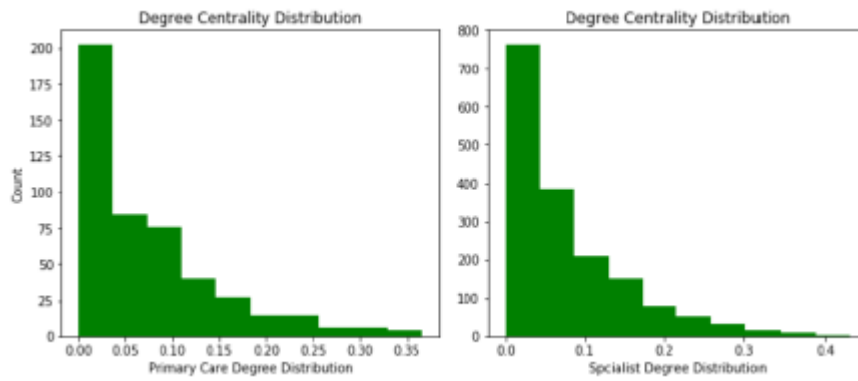
7.15. BETWEENNESS CENTRALITY DISTRIBUTION

7.16. CLOSENESS CENTRALITY DISTRIBUTION

7.17. PRIMARY-SPECIALTY COMMUNITIES