**Distribution Agreement**

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

______________________________ _______________

Thi Minh Hieu Nguyen Date

**Physician social networks and variation in outcomes for traditionally fatal cancers**

By

Thi Minh Hieu Nguyen
Master of Public Health

Department of Epidemiology

_________________________________________

Kevin C. Ward, PhD, MPH
Committee Chair

_________________________________________

Vicki Hertzberg, PhD
Committee Member

**Physician social networks and variation in outcomes for traditionally fatal cancers**

By

Thi Minh Hieu Nguyen

Bachelor of Engineering
National University of Singapore
2009

Thesis Committee Chair: Kevin C. Ward, PhD, MPH

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Department of Epidemiology
2017

**Abstract**

**Physician social networks and variation in outcomes for traditionally fatal cancers**

By Thi Minh Hieu Nguyen

This study aims to explore the relationship between physician social networks and the odds of survival from liver, stomach and pancreatic cancer. Little information is known about why some patients diagnosed with these deadliest cancers survive longer than expected, especially with respect to provider-specific factors. In this study, we constructed physician social networks based on their shared cancer patients using SEER-Medicare data from 2004 to 2010. Network clusters, comprised of providers who are more densely connected with one another, were identified from the network. We then examined whether the network clusters' structure and characteristics were associated with variation in survival outcome and the odds of long-term survival using linear mixed effects and generalized linear mixed effects models, controlling for patients' age, sex, race, comorbidity, tumor stage, geographic region and median household income. Only the network for pancreatic cancer exhibited significant network cluster structure, hence further analyses on the relationship with survival were limited to pancreatic cancer. There were 3133 physicians (medical oncologists, radiation oncologists and surgeons) and 25 large network clusters (with 20 or more patients per cluster) in this network. Significant variation in survival outcome across the network clusters was identified (Mixture test, test statistic = 11.10, p = 0.002). Patients whose providers belonged to the large clusters (with 20 or more patients per cluster) had significantly higher odds of 1-year survival compared to the small clusters (OR = 1.34, 95% CI 1.15-1.56, p <0.001). Among the 25 large network clusters, there was no significant difference in survival outcome of the patients. This study, although largely exploratory in nature, showed that survival outcomes were superior for patients involved in the large compared to small network clusters.

**Length**: 277 words

**Physician social networks and variation in outcomes for traditionally fatal cancers**

By

Thi Minh Hieu Nguyen

Bachelor of Engineering
National University of Singapore
2009

Faculty Thesis Advisors: Kevin C. Ward, PhD, MPH
Vicki S. Hertzberg, PhD

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Department of Epidemiology

2017

# Acknowledgements

I would like to dedicate this work to the late professor William McClellan, who patiently listened to my ideas and helped me to connect the dots. This thesis will not be a reality without him. I will forever remember his smile and his kindness.

I also would like to express my gratitude to Dr. Kevin Ward and Dr. Vicki Hertzberg for their ongoing support and guidance despite their busy schedule. Without Dr. Ward's introduction to the SEER – Medicare database during our very first meeting and Dr. Hertzberg's course on network science in public health, I could not come up with this idea.

In addition, I want to thank Mr. Renjian Jiang at the Georgia Center for Cancer Statistics for helping to establish the cohort from the SEER – Medicare database and continue giving me advice regarding the data. I also appreciated the assistance of Mr. Zach Binney, Department of Epidemiology, Ms. Lisa Elon and Mr. Liao Peizhou, Department of Biostatistics and Bioinformatics to help me figure out how to specify the mixed effects models.

Finally, I want to thank my friends and family for their tireless support during this journey.

# Table of Contents

Page number

# I. Background

### *A. Recalcitrant cancers: pancreatic cancer, liver cancer and stomach cancer*

The American Cancer Society estimated half of the 595,690 cancer deaths expected in 2016 will be caused by eight of the deadliest site specific cancers: pancreas, lung, liver, esophagus, stomach, brain, ovary and myeloma (1). These cancers have 5-year relative survival rates below 50% and are called recalcitrant cancers, as defined by the 2012 Recalcitrant Cancer Research Act.

Pancreatic cancer, liver cancer and stomach cancer are among the recalcitrant cancers with both higher incidence and very low prognosis. Moreover, all three organs in which these cancers originate are part of the digestive system. Since these organs are not easily accessible for clinical observation and there is no effective screening program in place, patients with these types of cancer tend to be diagnosed with more advanced stage disease. According to Surveillance, Epidemiology, and End Results (SEER) data (Appendix Table 1) from 2004 to 2010, the proportion of patients who survived more than 3 years was 8.4% for pancreatic cancer, 20.6% for liver cancer and 28.6% for stomach cancer. For the age group over 65, the numbers were 5.9%, 15.2% and 25.1% respectively.

- **Pancreatic cancer**

Pancreatic cancer is the deadliest cancer in the United States (1). Therapeutic options for the management of pancreatic cancer vary by stage and include surgical resection, neoadjuvant therapy, adjuvant chemotherapy, and radiation therapy (2). According to the National Cancer Institute's treatment options guideline for pancreatic cancer, patients at stage I-III are recommended a multimodal treatment approach of surgery, chemotherapy and chemoradiation therapy. For stage IV patients, only chemotherapy combined with palliative therapy is recommended. Early and complete tumor resection is presumed necessary to enable higher survival in pancreatic cancer patients (3). According to a study using the SEER database, receipt of curative intent surgery (CIS), early stage disease and white race were the strongest predictors of better survival. In addition, younger age, female gender, and geographic location were also predictors of receiving CIS, which in turn affects survival (4). A systematic review of the volume-outcome relationship in pancreatic surgery showed that improved survival was consistently observed in high patient-volume centers (5). Last but not least, close cooperation within a multidisciplinary team of physicians is strongly recommended for the management of pancreatic cancer to maximize survival outcomes (3, 6, 7). As early as the time of primary diagnosis of pancreatic cancer, a comprehensive and coordinated evaluation by a multidisciplinary clinic could change the therapeutic recommendations for patients (5). A SEER-Medicare study of 4105 elderly pancreatic cancer patients showed that post-operative outcomes were better for patients treated with surgery and adjuvant therapy than for patients that underwent pancreatectomy alone (8). Patients with borderline resectable pancreatic cancers who underwent multimodality therapy also had comparable survival outcomes to those who initially presented with resectable tumors (6).

- **Liver cancer**

Liver cancer is also a complex disease with a very poor prognosis. It is the third deadliest cancer in the United States (1). The therapeutic options for liver cancer depend on the cancer stage, liver function and performance status of the patient (9). Curative therapy such as hepatic resection, liver transplantation, percutaneous ethanol injection and radiofrequency ablation have shown to improve survival in the early stages of liver cancer (10). For intermediate stages, transarterial chemoembolization (TACE) is recommended as the first-line treatment. For unresectable and advance staged patients, only systemic therapies such as systemic chemotherapy, hormonal therapy and immunotherapy are indicated (9). With the availability of several treatment options and clinical trials, a comprehensive management approach with a team of multidisciplinary healthcare professionals is recommended for better patient outcomes (11, 12). Such an approach is beneficial in both diagnosis and treatment of liver cancer due to its complex pathological and oncological features (13). Earlier evidence showed that overall survival of stage II hepatocellular cancer patients was significantly improved after the establishment of a multidisciplinary treatment team at a Veteran Affair hospital (14).

- **Stomach cancer**

Stomach cancer is the fifth deadliest cancer in the United States (1). Most patients are diagnosed at advanced stages hence prognosis is very low (15). For localized gastric cancer, surgical resection is the treatment of choice. However, only 20-30% of patients with advanced-stage disease survive long-term with surgery alone (16). The recent development of a multimodal strategy, which combines curative intent surgery with post-operative chemoradiation and perioperative chemotherapy, has improved the prognosis of gastric cancer (17). For patient suffering from metastatic disease, chemotherapy and targeted therapy can improve overall survival (15). Palliative chemotherapy is also indicated to prolong survival (18). Similar to pancreatic cancer and liver cancer, the management of gastric cancer also involved a multidisciplinary and multimodal approach to improve long-term outcomes (17, 18).

### *B. Physician social networks*

A social network consists of one or more sets of units together with the relationships among them. The units are usually individual people, e.g. patients or clinicians, but can also be organizations such as hospitals. Relationships often represent increased opportunities for communication, influence, and trust among the units (19). Naturally occurring social networks of physicians and hospitals may contribute to variation in evidence-based health care.

Physicians are often embedded in formal and informal networks that result in their sharing patients, information, and behaviors. In physician social networks, physicians are considered connected to one another if they provide care to the same patient. By reflecting both formal (due to practice structure or hospital affiliation) and informal (referral patterns and advice seeking) connections that may shape clinical practice, physician patient-sharing networks may provide insight into variations in care and care outcomes (20).

Social networks can also shape health care delivery. Physicians' locations within networks of colleagues may serve to make some physicians aware of innovations in medicine sooner than others. Local opinion leaders occupying strategic, central network positions may disseminate influential assessments of both established and innovative medical regimens. Physicians may also look to nearby role models within their social networks for guidance in treating their patients (21).

### *C. Relationship between physician social networks and variation in cancer care*

Recent studies have shown that physician social networks are associated with variation in treatment for men with localized prostate cancer as well as rates of complications following radical prostatectomy. Using data from the linked SEER-Medicare database, researchers found that specific network subgroups of physicians involved in the care of prostate cancer patients (physicians who are more densely connected) were significantly associated with the likelihood that patients undergo prostatectomy in three different cities (22). Specifically, four among fourteen large subgroups (more than 50 patients per subgroup) in city A had significantly lower odds of prostatectomy compared with the baseline after adjusting for patient clinical and sociodemographic characteristics. The respective numbers for cities B and C were two among eight subgroups and five among eight subgroups. Network subgroup characteristics such as urologist centrality and patient racial compositions were also significantly associated with rates of surgical complications following prostatectomy (20). Average urologist degree was significantly associated with 30-day surgical complications and long term incontinence; while the proportion of non-white patients in a network subgroup was significantly associated with long term incontinence. Previously, by comparing physician referral and advice relationships measured by web-based surveys and patient-sharing relationships measured by Medicare data, a study validated the “Medicare” method as an informative ‘‘diagnostic test’’ for predicting the existence of relationships between physicians (23).

Patient-sharing may reflect an important aspect of provider relationships that facilitates timely, coordinated, consistent, state-of-the-art care for cancer patients, thus resulting in better outcomes. Collaboration among cancer specialists, measured by the number of patients shared among them, was shown to be associated with lower mortality among patients with stage III colon cancer (24). In this study, the authors reported an approximate 20% improvement in survival benefits from all-cause and colon cancer-specific mortality when the number of patients shared between specialists increased from one to five. For each additional patient, both all-cause mortality and colon cancer-specific mortality improved by 5%. Another study using SEER-Medicare data also found that cancer survivors treated by physicians who shared more patients with one another had higher quality and lower cost care (25). Thus we hypothesize that there is an association between physician social networks defined by shared patients and variation in the odds of survival among recalcitrant cancer patients with pancreatic, liver or stomach cancers.

## II. Manuscript

### A. Title, Authors, Abstract

***Physician social networks and variation in outcomes for traditionally fatal cancers***

Thi Minh Hieu Nguyen (1), Kevin C. Ward, PhD, MPH (1), Vicki Hertzberg, PhD (2)

1. Department of Epidemiology, Rollins School of Public Health, Emory University
2. Neil Hodgson Woodruff School of Nursing, Emory University

This study aims to explore the relationship between physician social networks and the odds of survival from liver, stomach and pancreatic cancer. Little information is known about why some patients diagnosed with these deadliest cancers survive longer than expected, especially with respect to provider-specific factors. In this study, we constructed physician social networks based on their shared cancer patients using SEER-Medicare data from 2004 to 2010. Network clusters, comprised of providers who are more densely connected with one another, were identified from the network. We then examined whether the network clusters' structure and characteristics were associated with variation in survival outcome and the odds of long-term survival using linear mixed effects and generalized linear mixed effects models, controlling for patients' age, sex, race, comorbidity, tumor stage, geographic region and median household income. Only the network for pancreatic cancer exhibited significant network cluster structure, hence further analyses on the relationship with survival were limited to pancreatic cancer. There were 3133 physicians (medical oncologists, radiation oncologists and surgeons) and 25 large

network clusters (with 20 or more patients per cluster) in this network. Significant variation in survival outcome across the network clusters was identified (Mixture test, test statistic = 11.10, p = 0.002). Patients whose providers belonged to the large clusters (with 20 or more patients per cluster) had significantly higher odds of 1-year survival compared to the small clusters (OR = 1.34, 95% CI 1.15-1.56, p <0.001). Among the 25 large network clusters, there was no significant difference in survival outcome of the patients. This study, although largely exploratory in nature, showed that survival outcomes were superior for patients involved in the large compared to small network clusters.

### *B. Introduction*

Pancreatic, liver and stomach cancer are among the deadliest cancers in the United States, with moderate incidence and very low prognosis (Deadliest Cancer Coalition, 2016). According to the Surveillance, Epidemiology, and End Results (SEER) data (Appendix Table 1) from 2004 to 2010, the proportion of patients who survived more than 3 years following diagnosis were 8.4% for pancreatic cancer, 20.6% for liver cancer and 28.6% for stomach cancer. For the over 65 age group, the numbers were 5.9%, 15.2% and 25.1% respectively. Due to the complexity and severity of these cancers generally due to late stage at diagnosis, multimodal approaches with the combination of surgical, chemotherapy and radiation therapy options are recommended to maximize the survival outcomes of the patients (3, 6, 9, 12, 17, 18, 26). These approaches require close cooperation within a multidisciplinary team of physicians, including but not limited to medical oncologists, surgeons and radiation oncologists.

Naturally occurring social networks of physicians is a new concept that attempts to represent the relationships and cooperation between physicians. In physician social networks, physicians are considered connected to one another if they provide care to the same patient. By reflecting both formal (due to practice structure or hospital affiliation) and informal (referral patterns and advice seeking) connections that may shape clinical practice, physician patient-sharing networks may provide insight into variation in care and care outcomes (20).

Recent studies have shown that physician social networks are associated with variation in treatment for men with localized prostate cancer as well as rates of complications following radical prostatectomy. Using data from the linked SEER-Medicare database, researchers found that specific network subgroups of physicians involved in the care of prostate cancer patients (physicians who are more densely connected) were significantly associated with the likelihood that patients undergo prostatectomy in three different cities (22). Specifically, four among fourteen large subgroups (more than 50 patients per subgroup) in city A had significantly lower odds of prostatectomy compared with the baseline after adjusting for patient clinical and sociodemographic characteristics. The respective numbers for cities B and C were two among eight subgroups and five among eight subgroups. Network subgroup characteristics such as urologist centrality and patient racial compositions were also significantly associated with rates of surgical complications following prostatectomy (20). Average urologist degree was significantly associated with 30-day surgical complications and long term incontinence; while the proportion of non-white patients in a network subgroup was significantly associated with long term incontinence. Previously, by comparing physician referral and advice relationships measured by web-based surveys and patient-sharing relationships measured by Medicare data, a study validated the "Medicare" method as an informative ''diagnostic test'' for predicting the existence of relationships between physicians (23).

Patient-sharing may reflect an important aspect of provider relationships that helps enable timely, coordinated, state-of-the-art care for cancer patients, thus resulting in better outcomes. Collaboration among cancer specialists, measured by the number of patients shared among them, was shown to be associated with lower mortality among patients with stage III colon cancer (24). In this study, the authors reported an approximate 20% improvement in survival benefits from all-cause and colon cancer-specific mortality when the number of patients shared between specialists increased from one to five. For each additional patient, both all-cause mortality and colon cancer-specific mortality improved by 5%. Another study using SEER-Medicare data also found that cancer survivors treated by physicians who shared more patients with one another had higher quality and lower cost care (25).

To our knowledge, there has been no study that examines the relationship of physician social networks in the multimodal treatment approaches for some of the deadliest cancers and the association between these relationships with patient outcomes. In an effort to explore if patient sharing between physicians plays a role in explaining better survival outcomes for fatal cancers, we hypothesize that there is an association between physician social networks defined by shared patients and variation in outcomes among recalcitrant cancer patients, namely pancreatic, liver and stomach cancer.

## *C. Methods*

### Study design

This is a retrospective population-based cohort study utilizing the linked Surveillance, Epidemiology, and End Results (SEER) - Medicare database. The study was approved by the Institutional Review Board at Emory University.

### Data Source

The SEER-Medicare database links patient demographic and tumor-specific information collected by SEER cancer registries to longitudinal health care claims for Medicare enrollees in the fee-for-service program. Medicare claims include inpatient, outpatient and physician files covering all procedures and corresponding diagnoses billed to Medicare. While physician specialty is recorded in Medicare, there has been some concern regarding the accuracy of this information (27). For this study, physician specialty was decided based on the treatment that the physician provided as documented in the claims data (Appendix Table 3 and 4).

**Study Population**

Patients age 66 years and older identified in the SEER-Medicare database with a diagnosis of pancreatic, liver or stomach cancer between 01/01/2004 and 12/31/2010 were included in the study. From those, we further excluded patients with incomplete Medicare records (i.e., those enrolled in health maintenance organizations), and patients who did not start their first course of treatment within 9 months from the initial diagnosis. Patients without identifiable provider NPIs/UPINs were also excluded. The final sample includes 1375 liver cancer patients, 5540 pancreatic cancer patients and 3622 stomach cancer patients (Appendix Table 2).

Physicians who were most likely to be involved in the care of those patients and thus most likely to directly or indirectly interact were included to form the network. For each patient in the study, we attempted to identify a surgical, medical and radiation oncologist involved in the patent's care. Since it is possible that a patient can have more than one physician of a given specialty involved in their care, we selected the first provider of each specialty who billed for the service in the Medicare claims.

**Network construction**

The network was constructed using data from the cancer patients and all of their providers as described above. In this network, physicians (represented by nodes) were connected to one another via shared patients (called edges).

Network construction was performed in Gephi 0.9.1.

Network clusters were identified from this network. Clusters define physicians who are more densely connected with one another via shared patients than to physicians outside the cluster. Each physician belonged to a single, mutually exclusive cluster. Because each patient has as many as three physicians, patients may have doctors who were assigned to multiple clusters. As a result, there was a correlation structure in which patients and providers were nested within network clusters. In addition, to emphasize the potential importance of larger clusters, we focus to study network clusters with at least 20 patients. Each cluster with 20 patients or more was assigned a separate cluster ID. All of the small network clusters with fewer than 20 patients were combined into a single cluster ID.

**Variable definitions**

***Dependent variables***

The primary outcome of this study was survival (in months). The secondary outcome, long-term survival, was defined as survival time (time until death or censoring) greater than 12 months.

***Independent variables***

The main predictors of interest were network cluster and network cluster characteristics, including the size of the cluster as well as the average degree of all the providers in the cluster. The size of the network cluster is defined as the total number of patients in that cluster. The degree of a provider is the number of other providers with whom he/she shared patients. These two measures were dichotomized. Network clusters were defined as having high average degree if they were in the top 75% of the distribution. For the size of the clusters, we were interested to compare the largest cluster with the rest of the clusters, as well as large clusters (more than 20 patients per cluster) with small clusters (less than 20 patients per cluster).

Other predictors included patient-level covariates such as age, sex, race, (white, non-white), geographic region (West, Midwest, South, Northeast), area-level median income, comorbidity, and tumor stage. Since our time window for assessing comorbidities is from 13 months prior to diagnosis to 1 month prior to diagnosis (the diagnosis month was not included), the Prior Charlson comorbidity score was used in the analysis (28, 29). Area-level U.S. Census information was used as a proxy for individual measures of socioeconomic status. Patients were linked to their census tract and, when not available, ZIP Code to determine median income, which was categorized into quartiles based on the sample distribution.

**Statistical analyses**

Descriptive statistics and univariate analyses were conducted to examine patient and network characteristics. Subsequently, the association of survival and network clusters was studied in multivariable analyses. To account for the correlation structure of patients and providers being nested within network clusters, we used linear mixed effects model analysis where survival time in months was the dependent variable and patient-level covariates were independent variables. Network cluster was included as a random effect. Due to a computational limitation, we could not include the random effects due to providers and patients.

For the second analysis, we examined the association between network cluster characteristics and odds of long-term survival, using generalized linear mixed effects models for binary outcomes. We included fixed effects for patient-level covariates, and for network cluster characteristic, and random effects for network clusters. Random effects due to providers and patients were not included due to the reason mentioned above. Analyses were performed separately for each type of network cluster characteristics, i.e. largest subgroup or high average degree of physicians.

Data analysis was performed using PROC MIXED and PROC GLIMMIX procedure in SAS (version 9.4, Cary, NC).

### D. Results

#### 1) Construction of physician social networks for pancreatic cancer, liver cancer and stomach cancer

Three different physician social networks were constructed to represent the patient-sharing relationships of all the providers (i.e medical oncologists, radiation oncologists, surgeons) involved in the care of liver, pancreatic and stomach cancer patients from 2004-2010. Table 1 presents some important characteristics of these networks. Pancreatic cancer had the largest network with 3133 nodes (physicians) and 3760 edges (patient-sharing relationships) that connect them. Since only the network for pancreatic cancer exhibited the network structure of interest, i.e. formation of large network clusters (Figure 1, Appendix Figure 1-3), we decided to focus our further analysis on this network only.

According to network theory, degree centrality is a measure of the direct relationships (edges) a physician has with other physician in the network. A physician with higher degree centrality means that he or she is more well-connected, thus occupying an important position in the network. The highest degree of the pancreatic cancer network was 33, meaning that this physician shared patients with 33 other physicians in this network. On average, each provider connected with 1.2 other providers in the network. It took on average 8.2 steps to get from one provider to another provider in the network (average path length). The shortest distance between two most distant providers in the network, i.e. network diameter, was 23. In other words, a provider had to communicate with 8.2 providers on average and 23 providers at most to contact a total stranger in the network. There were 1741 connected components (i.e. appearance of densely connected groups of nodes and sparser connections between groups) in the pancreatic cancer network and these components formed the basis for our network clusters. There were 25 large network clusters (cluster ID 1-25), which were defined as components having at least 20 patients. Each large network cluster had 33 providers and 85 patients on average (Table 2.)

There were 5540 pancreatic cancer patients (53.07% female, 85.16% white, aged 66 and above) in the network. More than half of the patients had one or more comorbidities. The majority of patients were diagnosed at stage II (36.16%) and stage IV (38.32%) disease. The mean survival time was 11.88 months with a standard deviation of 12.24 months. Other sociodemographic characteristics are presented in Table 3.

Figure 1 depicts the pancreatic cancer network where the network clusters are colored coded. For a clearer resolution, only 1755 providers with a degree larger than 0 (i.e. share at least 1 patient with other providers in the network) are shown in the network diagram. Among these providers, 42% are medical oncologists, 35% are radiation oncologists and 23% are surgeons (Appendix Figure 1).

### 2) *Variation in survival outcomes in pancreatic cancer across network clusters*

The significance of network clusters random effect was tested using the Mixture Test method, which is believed to be more statistically desirable compared to the Wald test (30). According to the test, there was an overall significant variation in survival outcomes across the network clusters (Likelihood ratio test statistic = 11.10, p = 0.002). However, this significant variation was largely due to the random effect of all of the clusters with less than 20 patients (i.e. cluster ID 26, Table 4). There was no significant difference in survival of the patients in the large clusters, i.e. cluster ID 1-25 compared to the population average. As explained previously, cluster IDs 1-25 were the large network clusters with 20 patients or more per cluster. All of the small network clusters with fewer than 20 patients were combined into a single cluster ID 26. To further understand the effect of this cluster on survival time, we decided to run another linear mixed effects model, this time including cluster ID 26 as a fixed effect. The result showed that patients belonging to this cluster ID survived less than 1.32 months on average compared to patients not in this cluster (95% CI 0.52 – 2.12, p = 0.001, data not shown).

### 3) *Association of network cluster characteristics with long-term survival in pancreatic cancer*

In models that examined characteristics of network clusters, controlling for patients' age, sex, race, comorbidity, tumor stage, geographic region and median household income, the largest cluster (OR = 1.15, 95% CI 0.96 – 1.38, p

= 0.12) and average degree of providers (OR = 1.04, 95% CI 0.88 - 1.24, $p = 0.61$) were not significantly associated with long-term survival (Table 5). The odds of 1-year survival for patients in the large clusters was 34% higher than the same odds of those in the small clusters (OR = 1.34, 95% CI 1.15 – 1.56, $p < 0.001$). This result showed that patients whose physicians belong to a group of providers that share many patients with each other have significantly higher odds of 1-year survival.

## *E. Discussion*

The findings of this study raise a number of important issues to be discussed. Although the literature recommends the use of a multimodal approach in the treatment of liver, stomach and pancreatic cancer, we were only able to detect significant network and network cluster structure in pancreatic cancer. According to the National Cancer Institute's Physician Data Query guideline for cancer treatment, a treatment approach combining multiple therapies is recommended for stage I-III stomach cancer and pancreatic cancer. For liver cancer, surgical options are recommended for early stages whereas chemotherapy and radiation therapy are recommended for late stages. This partly explains what was observed in the network for liver cancer. However, it is interesting that we only observed significant network clusters structure in pancreatic cancer but not stomach cancer, although both samples contain 55-60% patients in stages I-III (Figure 2).

Overall, there is significant variation in survival outcomes across the network clusters for pancreatic cancer. However, when we look at the individual clusters, only patients from the small clusters, i.e. less than 20 patients per cluster, significantly differed from the rest in terms of survival. It should be noted that about 50% of the patients in these small clusters are stage IV, compared with 33% stage IV patients in the large clusters. Stage IV patients usually have worse outcomes and are less likely to be involved in multimodal treatment. Although our models controlled for tumor stage, we suspect that other unmeasured factors related to tumor stage may have played some role in the results we observed.

It should be noted that we did not observe a significant association between average degree of the providers and the outcome. Physicians with high degrees, i.e. those who have many connections with other physicians, may exert higher influence in the network, but this influence may not necessarily translate to better outcomes for patients. In fact, being in a large cluster with a small degree, i.e. sharing lots of patients with a single other provider, may be more optimal for outcomes. Nevertheless, we believe these findings suggest that some aspect of patient sharing between physicians results in higher odds of long-term survival for their patients.

This study presents a novel approach in using network science to visualize the naturally occurring network of physicians and examine the association of these network structures with patient outcomes. However, it also suffered from a number of limitations. First, there is currently no optimal approach to construct the patient-sharing network of physicians and identify the clusters within this network. Hence our work is largely exploratory in nature. Second, we identified provider specialty based on the type of procedures that provider performed for the patient. Ideally, this information should be obtained from the American Medical Association Physician Masterfile link with SEER-Medicare. For this reason, we also failed to include Primary Care Providers (PCPs) in our network. We suspect that the referral role that the PCPs play may help shape the network structure. Third, we did not control for provider volume in our analysis due to time constraints. Fourth, we could not run the full models, which included the random effects of not only clusters, but also providers and patients due to a computational limitation (although we did try to run them in a high performance computing cluster). Hence our final models only included the random effects due to network clusters.

In summary, by using network analysis, we were able to construct a network of physicians involved in the care of pancreatic cancer patients based on their patient-sharing relationships. From this network, we identified different network clusters, which were defined as groups of physicians that were "tightly connected" to each other through shared patients. There was significant variation in survival outcome across the network clusters. Patients whose providers belonged to the large clusters (with 20 or more patients per cluster) had significantly higher odds of 1-year survival compared to the small clusters. Among the large network clusters, there was no significant difference in survival outcome of the patients. Further research needs to be done in order to understand the mechanism behind these observations.

This is the first study, to our knowledge, that applies network analysis to examine the association between physician relationships and pancreatic cancer outcomes. This method can be used to study other public health outcomes which are believed to be associated with the cooperation and relationship between care providers. If validated, this kind of study could help to improve quality of care and target interventions to specific groups of physicians identified from such studies.

## *F. References*

1. Society AC. Cancer Facts & Figures 2016. *Atlanta: American Cancer Society* 2016.
2. Rossi ML, Rehman AA, Gondi CS. Therapeutic options for the management of pancreatic cancer. *World J Gastroenterol* 2014;20(32):11142-59.
3. Buanes TA. Pancreatic cancer-improved care achievable. *World J Gastroenterol* 2014;20(30):10405-18.
4. Shaib Y, Davila J, Naumann C, et al. The impact of curative intent surgery on the survival of pancreatic cancer patients: a U.S. Population-based study. *Am J Gastroenterol* 2007;102(7):1377-82.
5. Gooiker GA, van Gijn W, Wouters MW, et al. Systematic review and meta-analysis of the volume-outcome relationship in pancreatic surgery. *Br J Surg* 2011;98(4):485-94.
6. Pawlik TM, Laheru D, Hruban RH, et al. Evaluating the impact of a single-day multidisciplinary clinic on the management of pancreatic cancer. *Ann Surg Oncol* 2008;15(8):2081-8.
7. Cooper AB, Tzeng CW, Katz MH. Treatment of borderline resectable pancreatic cancer. *Curr Treat Options Oncol* 2013;14(3):293-310.
8. Nayar P, Chandak A, Gupta N, et al. Postoperative mortality following multi-modality therapy for pancreatic cancer: Analysis of the SEER-Medicare data. *Journal of surgical oncology* 2017;115(2):158-63.

9. Colombo M, Raoul JL, Lencioni R, et al. Multidisciplinary strategies to improve treatment outcomes in hepatocellular carcinoma: a European perspective. *Eur J Gastroenterol Hepatol* 2013;25(6):639-51.

10. Chen KW, Ou TM, Hsu CW, et al. Current systemic treatment of hepatocellular carcinoma: A review of the literature. *World J Hepatol* 2015;7(10):1412-20.

11. Gish RG, Lencioni R, Di Bisceglie AM, et al. Role of the multidisciplinary team in the diagnosis and treatment of hepatocellular carcinoma. *Expert Rev Gastroenterol Hepatol* 2012;6(2):173-85.

12. Choi C, Choi GH, Kim TH, et al. Multimodality Management for Barcelona Clinic Liver Cancer Stage C Hepatocellular Carcinoma. *Liver Cancer* 2014;3(3-4):405-16.

13. Siddique O, Yoo ER, Perumpail RB, et al. The importance of a multidisciplinary approach to hepatocellular carcinoma. *Journal of multidisciplinary healthcare* 2017;10:95-100.

14. Chang TT, Sawhney R, Monto A, et al. Implementation of a multidisciplinary treatment team for hepatocellular cancer at a Veterans Affairs Medical Center improves survival. *HPB : the official journal of the International Hepato Pancreato Biliary Association* 2008;10(6):405-11.

15. Delaunoit T. Latest developments and emerging treatment options in the management of stomach cancer. *Cancer Manag Res* 2011;3:257-66.

16. Moehler M, Galle PR, Gockel I, et al. The multidisciplinary management of gastrointestinal cancer. Multimodal treatment of gastric cancer. *Best Pract Res Clin Gastroenterol* 2007;21(6):965-81.

17. Das P, Jiang Y, Lee JH, et al. Multimodality approaches to localized gastric cancer. *J Natl Compr Canc Netw* 2010;8(4):417-25.

18. Misleh JG, Santoro P, Strasser JF, et al. Multidisciplinary management of gastric cancer. *Surg Oncol Clin N Am* 2013;22(2):247-64.

19. O'Malley AJ, Marsden PV. The Analysis of Social Networks. *Health Serv Outcomes Res Methodol* 2008;8(4):222-69.

20. Evan Pollack C, Wang H, Bekelman JE, et al. Physician social networks and variation in rates of complications after radical prostatectomy. *Value Health* 2014;17(5):611-8.

21. Christakis NA, Fowler JH. Social Network Visualization in Epidemiology. *Nor Epidemiol* 2009;19(1):5-16.

22. Pollack CE, Weissman G, Bekelman J, et al. Physician social networks and variation in prostate cancer treatment in three cities. *Health Serv Res* 2012;47(1 Pt 2):380-403.

23. Barnett ML, Landon BE, O'Malley AJ, et al. Mapping physician networks with self-reported and administrative data. *Health Serv Res* 2011;46(5):1592-609.

24. Hussain T, Chang HY, Veenstra CM, et al. Collaboration Between Surgeons and Medical Oncologists and Outcomes for Patients With Stage III Colon Cancer. *J Oncol Pract* 2015;11(3):e388-97.

25. Pollack CE, Frick KD, Herbert RJ, et al. It's who you know: patient-sharing, quality, and costs of cancer survivorship care. *J Cancer Surviv* 2014;8(2):156-66.

26. Yoon HI, Seong J. Multimodality treatment involving radiotherapy for advanced liver-confined hepatocellular carcinoma. *Oncology* 2014;87 Suppl 1:90-8.

27. Pollack LA, Adamache W, Eheman CR, et al. Enhancement of identifying cancer specialists through the linkage of Medicare claims to additional sources of physician specialty. *Health Serv Res* 2009;44(2 Pt 1):562-76.

28. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40(5):373-83.

29. Klabunde CN, Potosky AL, Legler JM, et al. Development of a comorbidity index using physician claims data. *J Clin Epidemiol* 2000;53(12):1258-67.

30. Bardenheier, BH. Proc Mixed: Macro to assess fix and random effect for significance using the Likelihood Ratio test and the approximate Mixture Method.

## *G. Tables*

Table 1. Characteristics of the physician social networks for liver cancer, pancreatic cancer and stomach cancer patients

| Type of cancer | Liver cancer | Pancreatic cancer | Stomach cancer |
|---|---|---|---|
| Size | 1083 nodes, 226 edges | 3133 nodes, 3760 edges | 2998 nodes, 2316 edges |
| Average degree | 0.209 | 1.2 | 0.773 |
| Average path length | 1.067 | 8.194 | 1.806 |
| Diameter | 3 | 23 | 7 |
| Average clustering coefficient | 0.008 | 0.112 | 0.109 |
| Number of connected components | 973 | 1741 | 2064 |

Table 2. Descriptive characteristics of the pancreatic cancer network structure

| **Variables** | **Total** |
|---|---|
| Number of patients | 5540 |
| Number of providers | 3133 |
| Number of Radiation oncologists | 903 |
| Number of Medical oncologists | 1519 |
| Number of Surgeons | 711 |
| Number in large network cluster* | 25 |
| Average number of patients per large network cluster (range) | 85 (20-905) |
| Average number of providers per large network cluster (range) | 33 (4-328) |

* Large network cluster is defined as having at least 20 patients

Table 3. Sociodemographic and clinical characteristics of pancreatic cancer patients cohort in the network

| **Variables** | **Total** |
|---|---|
| No. (%) | 5540 (100.00) |
| Age | |
| 65-69 years old | 1216 (21.95) |
| 70-74 years old | 1559 (28.14) |
| 75-79 years old | 1465 (26.44) |
| 80+ years old | 1300 (23.47) |
| Sex | |
| Female | 2940 (53.07) |
| Male | 2600 (46.93) |
| Race | |
| White | 4718 (85.16) |
| Non-white and Unknown race * | 822 (14.84) |
| Geographic region | |
| Midwest | 659 (11.90) |
| Northeast | 1323 (23.88) |
| South | 1279 (23.09) |
| West | 2279 (41.14) |
| Median household income in dollars | |
| Less than 36381 | 1384 (24.98) |
| 36381 - 48975 | 1385 (25.00) |
| 48975 - 65853 | 1384 (24.98) |
| Greater than 65853 | 1387 (25.04) |
| Comorbidity | |
| 0 | 2729 (49.26) |
| 1 | 1614 (29.13) |
| 2+ | 1197 (21.61) |
| AJCC-6 Stage Group | |
| Stage 0 and Stage I* | 369 (6.66) |
| Stage II | 2003 (36.16) |
| Stage III | 558 (10.07) |
| Stage IV | 2123 (38.32) |
| Unknown | 487 (8.79) |
| Survival time in months, Mean(Std) | 11.88 (12.24) |
| Less than 12 months, N (%) | 3562 (64.30) |
| 12 months and greater, N(%) | 1978 (35.70) |

* Data were merged for confidentiality reasons due to small cell size

Table 4. Variation in outcome due to random effect of network clusters

| Cluster ID | Intercept estimate (95% CI) | p value |
|---|---|---|
| 1 | 0.52 (-0.41 - 1.45) | 0.27 |
| 2 | -0.23(-1.32 - 0.85) | 0.67 |
| 3 | 0.04 (-1.30 - 1.37) | 0.96 |
| 4 | 0.42 (-0.92 - 1.76) | 0.54 |
| 5 | 0.03 (-1.35 - 1.42) | 0.96 |
| 6 | -0.16 (-1.56 - 1.24) | 0.82 |
| 7 | -0.40 (-1.82 - 1.01) | 0.58 |
| 8 | -0.30 (-1.75 - 1.15) | 0.69 |
| 9 | 0.06 (-1.38 - 1.51) | 0.93 |
| 10 | -0.09 (-1.56 - 1.38) | 0.91 |
| 11 | -0.47 (-1.94 - 1.01) | 0.54 |
| 12 | 0.35 (-1.12 - 1.83) | 0.64 |
| 13 | 0.08 (-1.40 - 1.55) | 0.92 |
| 14 | 0.04 (-1.45 - 1.53) | 0.96 |
| 15 | 0.58 (-0.92 - 2.07) | 0.45 |
| 16 | 0.17 (-1.33 - 1.67) | 0.83 |
| 17 | 0.14 (-1.36 - 1.64) | 0.86 |
| 18 | 0.11 (-1.41 - 1.63) | 0.89 |
| 19 | 0.32 (-1.18 - 1.82) | 0.67 |
| 20 | 0.04 (-1.46 - 1.54) | 0.96 |
| 21 | -0.17 (-1.67 - 1.33) | 0.83 |
| 22 | -0.13 (-1.62 - 1.37) | 0.87 |
| 23 | -0.29 (-1.80 - 1.21) | 0.70 |
| 24 | 0.60 (-0.91 - 2.10) | 0.44 |
| 25 | -0.27 (-1.79 - 1.25) | 0.72 |
| 26 | -0.98 (-1.72 - (-0.23)) | 0.01 |

Table 5. Association of network characteristics and long-term survival

| | Odds of Long-term Survival | |
|---|---|---|
| Variables | ORs (95% CI) | p-value |
| Largest size | 1.15 (0.96 - 1.38) | 0.12 |
| Large size (>= 20 patients/cluster) | 1.34 (1.15 - 1.56) | < 0.001 |
| High average degree | 1.04 (0.88 - 12.4) | 0.61 |

***H. Figures/Figure legends***

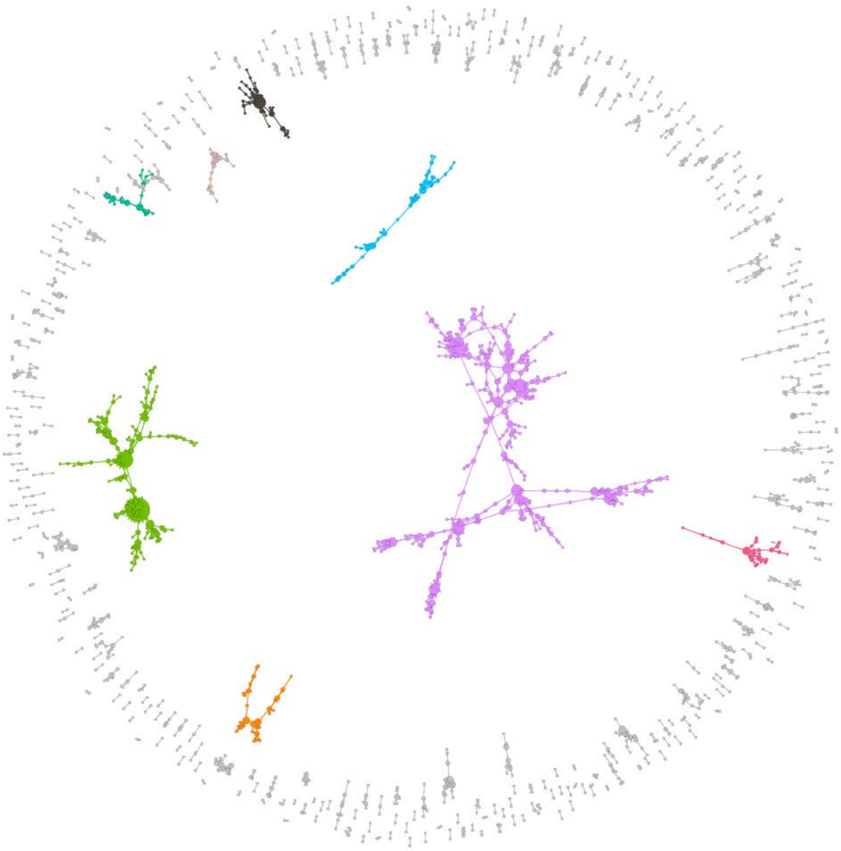
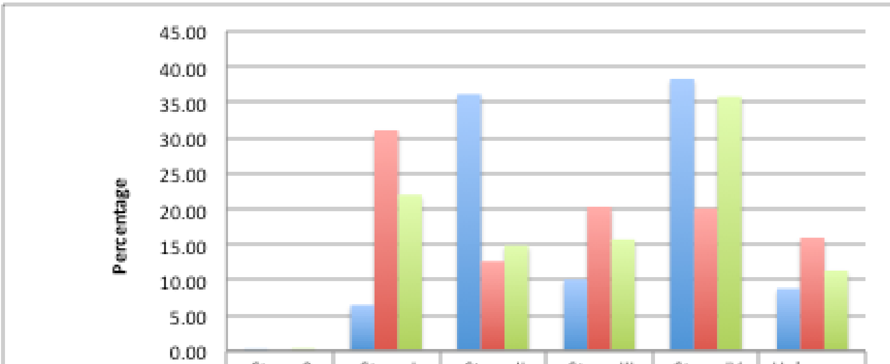


Figure 1. The physician social network for pancreatic cancer. 1755 providers are represented by circles. They are connected by shared patients represented by lines. Providers with larger circles have higher degree, i.e. more connections with other providers. Network clusters are shaded different colors.




| | Stage 0 | Stage I | Stage II | Stage III | Stage IV | Unknown |
|---|---|---|---|---|---|---|
| Pancreatic cancer | 0.16 | 6.50 | 36.16 | 10.07 | 38.32 | 8.79 |
| Liver cancer | 0.00 | 31.05 | 12.65 | 20.29 | 20.07 | 15.93 |
| Stomach cancer | 0.36 | 21.98 | 14.88 | 15.68 | 35.81 | 11.29 |

Figure 2: Proportion of patients diagnosed by stage for pancreatic cancer, liver cancer and stomach cancer, SEER 18 Regs Research Data, Year 2004-2010

## III. Summary, Public Health Implication, Possible Future Directions

In summary, by using network analysis, we were able to construct a network of physicians involved in the care of pancreatic cancer patients based on their patient-sharing relationships. From this network, we identified different network clusters, which were defined as groups of physicians that were “tightly connected” to each other through shared patients. There was significant variation in survival outcome across the network clusters. Patients whose providers belonged to the large clusters (with 20 or more patients per cluster) had significantly higher odds of 1-year survival compared to the small clusters. Among the large network clusters, there was no significant difference in survival outcome of the patients. Further research needs to be done in order to understand the mechanism behind these observations.

This is the first study, to our knowledge, that applies network analysis to examine the association between physician relationships and pancreatic cancer outcomes. This method can be used to study other public health outcomes which are believed to be associated with the cooperation and relationship between care providers. If validated, this kind of study could help to improve quality of care and target interventions to specific groups of physicians identified from such studies.

# IV. Appendices

*Appendix Table 1. Distribution of lethal cancers according to 3 years survival, SEER data 2004-2010 **

| ***Survival*** | ***Liver cancer*** | | | |
|---|---|---|---|---|
| | All age groups | | **65+ age groups** | |
| | n | % | **n** | **%** |
| ***<= 3 yrs*** | 31,453 | 79.40% | 14,948 | 84.80% |
| ***> 3 yrs*** | 8,169 | 20.60% | 2,669 | 15.20% |
| ***Total*** | 39,622 | 100.00% | 17,617 | 100.00% |
| ***Survival*** | ***Stomach cancer*** | | | |
| | All age groups | | **65+ age groups** | |
| | n | % | **n** | **%** |
| ***<= 3 yrs*** | 31,538 | 71.40% | 20,602 | 74.90% |
| ***> 3 yrs*** | 12,626 | 28.60% | 6,913 | 25.10% |
| ***Total*** | 44,164 | 100.00% | 27,515 | 100.00% |
| ***Survival*** | ***Pancreatic cancer*** | | | |
| | All age groups | | **65+ age groups** | |
| | n | % | **n** | **%** |
| ***<= 3 yrs*** | 62,653 | 91.60% | 42,737 | 94.10% |
| ***> 3 yrs*** | 5,776 | 8.40% | 2,684 | 5.90% |
| ***Total*** | 68,429 | 100.00% | 45,421 | 100.00% |

* Incidence - SEER 18 Regs Research Data + Hurricane Katrina Impacted Louisiana Cases, Nov 2015 Sub (1973-2013 varying), exclude death certificate only case

Appendix Table 2. Inclusion/exclusion criteria and sample size

| | **Liver cancer** | | **Pancreatic cancer** | | **Stomach cancer** | |
|---|---|---|---|---|---|---|
| | N | % | N | % | N | % |
| ***Total Eligilbe Patients*** | ***8441*** | ***100%*** | ***21523*** | ***100%*** | ***9612*** | ***100%*** |
| **Radiation therapy** | | | | | | |
| Within 9 month | 445 | 5.27% | 2758 | 12.81% | 1795 | 18.67% |
| Has NPI/ Can Match NPI by UPIN | 390 | 4.62% | 1986 | 9.23% | 1334 | 13.88% |
| **Chemotherapy** | | | | | | |
| Within 9 month | 470 | 5.57% | 5103 | 23.71% | 1882 | 19.58% |
| Has NPI/ Can Match NPI by UPIN | 412 | 4.88% | 3684 | 17.12% | 1485 | 15.45% |
| **Surgery** | | | | | | |
| Within 9 month | 1020 | 12.08% | 2095 | 9.73% | 2424 | 25.22% |
| Has NPI/ Can Match NPI by UPIN | 683 | 8.09% | 1504 | 6.99% | 1808 | 18.81% |
| ***Final sample size (Had any of the treatment modality)*** | ***1375*** | ***16.29%*** | ***5540*** | ***25.74%*** | ***3622*** | ***37.68%*** |

Appendix Table 3. Cancer surgical procedure codes

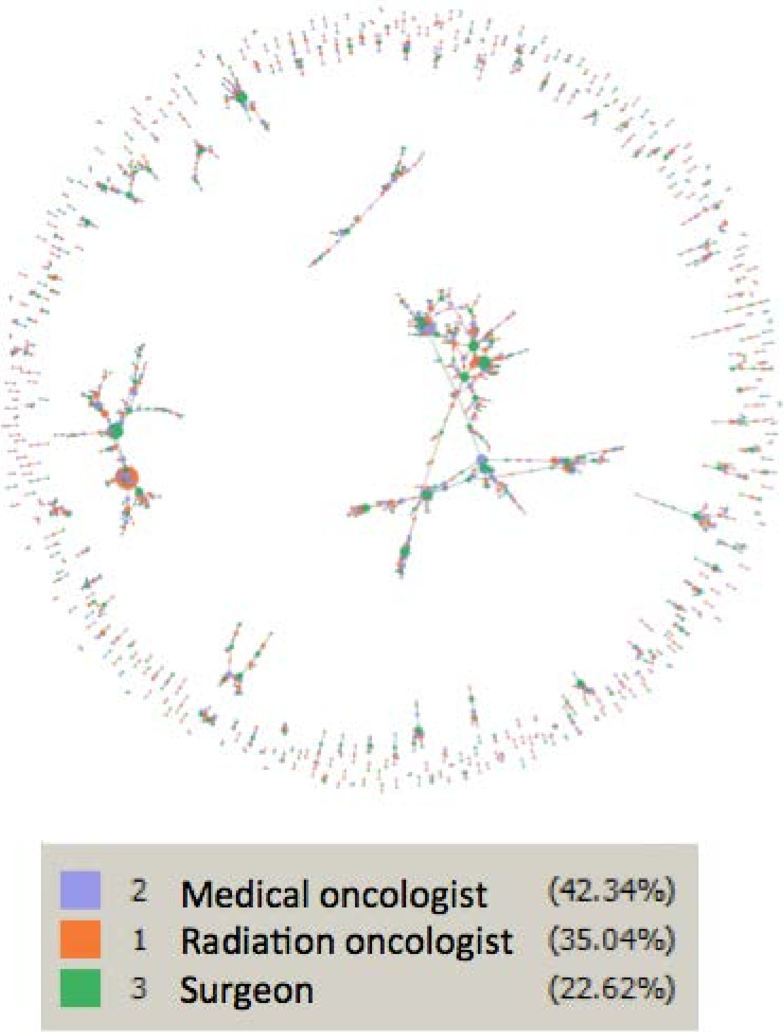
| Cancer | Code | Brief Text |
|---|---|---|
| **Stomach** | | |
| Procedure (ICD9) | 43.42 | Local excision of other lesion or tissue of stomach |
| | 43.5 | Partial gastrectomy with anastomosis to esophagus |
| | 43.6 | Partial gastrectomy with anastomosis to duodenum |
| | 43.7 | Partial gastrectomy with anastomosis to jejunum |
| | 43.81 | Partial gastrectomy with jejunal transposition |
| | 43.82 | Laparoscopic vertical (sleeve) gastrectomy |
| | 43.89 | Open and other partial gastrectomy |
| | 43.91 | Total gastrectomy with intestinal interposition |
| | 43.99 | Other total gastrectomy |
| CPT | *43610* | Excision, local; ulcer or benign tumor of stomach |
| | 43611 | malignant tumor of stomach |
| | 43620 | Gastrectomy, total; with esophagoenterostomy |
| | 43621 | with Roux-en-Y reconstruction |
| | 43622 | with formation of intestinal pouch, any type |
| | 43631 | Gastrectomy, partial, distal; with gastroduodenostomy |
| | 43632 | with gastrojejunostomy |
| | 43633 | with Roux-en-Y reconstruction |
| | 43634 | with formation of intestinal pouch |
| | 43635 | Vagotomy when performed with partial distal gastrectomy (List separately in addition to code(s) for primary procedure) (Use 43635 in conjunction with 43631, 43632, 43633, 43634) |
| **Liver** | | |
| Procedure (ICD9) | 50.22 | Partial hepatectomy |
| | 50.3 | Lobectomy of liver |
| | 50.4 | Total hepatectomy |
| CPT | 47120 | Hepatectomy, resection of liver; partial lobectomy |
| | 47122 | trisegmentectomy |
| | 47125 | total left lobectomy |
| | 47130 | total right lobectomy |
| | 47370 | Laparoscopy, surgical, ablation of 1 or more liver tumor(s) |
| | 47371 | Radiofrequency cryosurgical |
| | 47379 | Unlisted laparoscopic, liver |

Appendix Table 3 (cont). Cancer surgical procedure codes

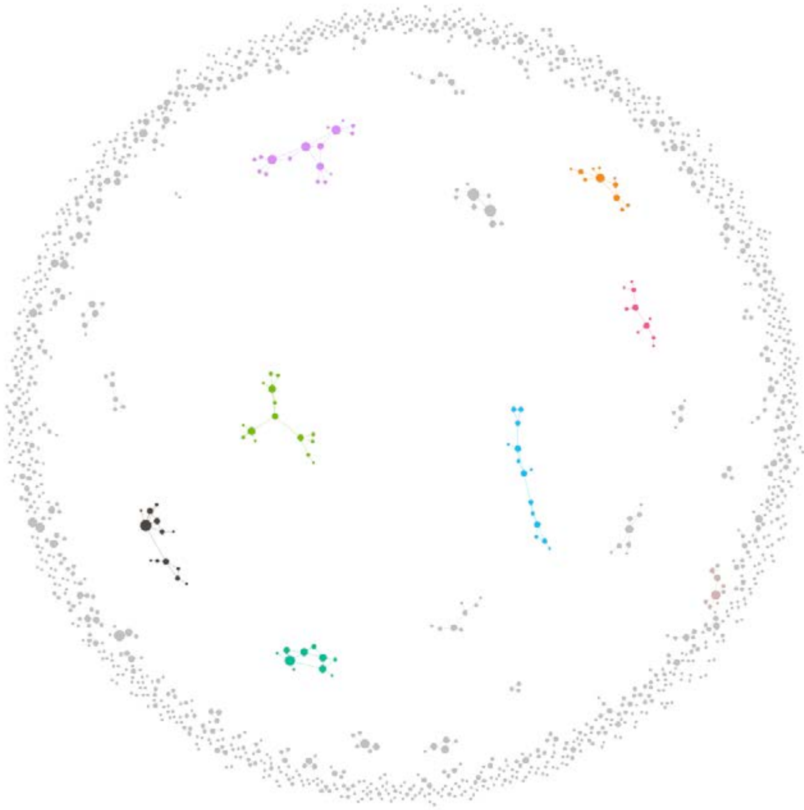
| Cancer | Code | Brief Text |
|---|---|---|
| Pancreas | | |
| Procedure | 57.71 | Pelvic exeneration |
| (ICD9) | 60.4 | Retropubic prostatectomy |
| | 60.5 | Radical prostatectomy |
| | 60.69 | Other prostatectomy |
| CPT | 48120 | Excision of lesion of pancreas (eg, cyst, adenoma) |
| | 48140 | Pancreatectomy, distal subtotal, with or without splenectomy; without pancreaticojejunostomy |
| | 48145 | with pancreaticojejunostomy |
| | 48146 | Pancreatectomy, distal, near-total with preservation of duodenum (Child -type procedure) |
| | 48148 | Excision of ampulla of Vater |
| | 48150 | Pancreatectomy, proximal subtotal with total duodenectomy, partial gastrectomy, cholecystoenterostomy and gastrojejunostomy (Whipple-type procedure); with pancreatojejunostomy |
| | 48152 | without pancreatojejunostomy |
| | 48153 | Pancreatectomy, proximal subtotal with near-total duodenectomy, cholecystoenterostomy and duodenojejunostomy (pylorus-sparing, Whipple-type procedure); with pancreatojejunostomy |
| | 48154 | without pancreatojejunostomy |
| | 48155 | Pancreatectomy, total |
| | 48160 | Pancreatectomy, total or subtotal, with autologous transplantation of pancreas or pancreatic islet cells |

Appendix Table 4. Chemotherapy and radiation therapy codes

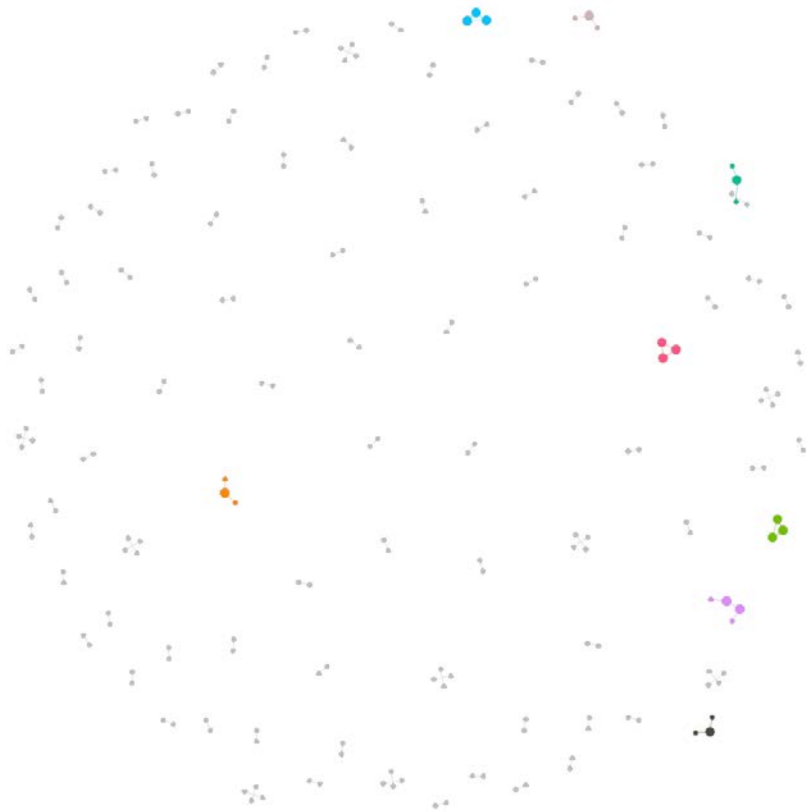
| Therapy | Codes | Note |
|---|---|---|
| Chemotherapy | J9000-J9999 | Level II HCPCS only |
| Radiation therapy | 77401-77499, 77520, 77523, 77750- 77799, G0256, G0261 | CPT codes |




Appendix Figure 1. The physician social network for pancreatic cancer. 1755 providers are represented by circles. They are connected by shared patients represented by lines. Providers with larger circles have higher degree, i.e. more connections with other providers. Providers specialty are shaded different colors.



Appendix Figure 2. The physician social network for stomach cancer. 1413 providers are represented by circles. They are connected by shared patients represented by lines. Providers with larger circles have higher degree, i.e. more connections with other providers. Network clusters are shaded different colors.



Appendix Figure 3. The physician social network for stomach cancer. 211 providers are represented by circles. They are connected by shared patients represented by lines. Providers with larger circles have higher degree, i.e. more connections with other providers. Network clusters are shaded different colors.

**SAS codes for linear mixed model and generalized linear mixed model**

```
***************************************************************;
* Linear mixed effect model                                    ;
***************************************************************;

*Linear mixed model, random effect for cluster and provider
--> to see if cluster associated with variation in survival time;
proc mixed data=hannah.network_pedsf_pancr_medonly method=reml
covtest;

      class cluster_id (ref = '26') provider_id (ref = 'NoMedicalO')
                  patient_id income (ref = '1') comorbidity (ref =
'0')
                  agecat (ref = '1')stage (ref = '0') raceeth (ref =
'1')
                  region_cat (ref = '1') sex (ref = '0');
      model survmonth = income comorbidity agecat stage raceeth
region_cat sex/
                  s ddfm=contain;
      *survmonth is survival by month;
      random intercept /type = vc subject = cluster_id s cl;
      *random intercept/ type = vc subject = provider_id
(cluster_id);
      *random intercept/ type = vc subject = patient_id
(cluster_id);

run;


*Misture test for significance of random effect;
%include
"Z:\medicare\Nguyen\SharedFolder\Hannah\SAS\Pancreas\New\mixture_mac
ro.sas";

*Full model;
proc mixed data=hannah.network_pedsf_pancr_medonly method=reml
covtest;

      class cluster_id (ref = '26') provider_id (ref = 'NoMedicalO')
                  patient_id income (ref = '1') comorbidity (ref =
'0')
                  agecat (ref = '1')stage (ref = '0') raceeth (ref =
'1')
                  region_cat (ref = '1') sex (ref = '0');
      model survmonth = income comorbidity agecat stage raceeth
region_cat sex/
                  s ddfm=contain;
      *survmonth is survival by month;
      random intercept /type = vc subject = cluster_id ;
      *random intercept/ type = vc subject = provider_id
(cluster_id);
      *random intercept/ type = vc subject = patient_id
(cluster_id);
      ods output FitStatistics = fullmodel;

run;

*Reduce model;
proc mixed data=hannah.network_pedsf_pancr_medonly method=reml
covtest;
```

```
      class cluster_id (ref = '26') provider_id (ref = 'NoMedicalO')
            patient_id income (ref = '1') comorbidity (ref = '0')
            agecat (ref = '1')stage (ref = '0') raceeth (ref = '1')
            region_cat (ref = '1') sex (ref = '0');
      model survmonth = income comorbidity agecat stage raceeth
region_cat sex/
            s ddfm=contain;
      *survmonth is survival by month;
      *random intercept /type = vc subject = cluster_id ;
      *random intercept/ type = vc subject = provider_id
(cluster_id);
      *random intercept/ type = vc subject = patient_id
(cluster_id);
      ods output FitStatistics = redmodel;

run;

%LRMixtureTest(fullmodel=fullmodel,redmodel=redmodel);

proc mixed data=hannah.network_pedsf_pancr_medonly method=reml
covtest;

      class cluster_id (ref = '26') provider_id (ref = 'NoMedicalO')
                  patient_id income (ref = '1') comorbidity (ref =
'0')
                  agecat (ref = '1')stage (ref = '0') raceeth (ref =
'1')
                  region_cat (ref = '1') sex (ref = '0');
      model survmonth =  network_1 income comorbidity agecat stage
raceeth region_cat sex/
                  s ddfm=contain cl;
      *survmonth is survival by month;
      *network_1 = 1 if small cluster, i.e. cluster_id = 26, = 0
else;
      random intercept/type = vc subject = cluster_id s cl;
      *random intercept/ type = vc subject = provider_id
(cluster_id);
      *random intercept/ type = vc subject = patient_id
(cluster_id);

run;


*************************************************************;
* Generalized linear mixed effect model                     ;
*************************************************************;

* Generalized linear mixed model for binary outcome, fixed effect
for largest cluster,
random effect for cluster and provider -->
to see if largest cluster associated with long-term survival,
i.e. survival more than 12 months;
proc glimmix data=hannah.network_pedsf_pancr_medonly empirical;

      class cluster_id (ref = '26') provider_id (ref = 'NoMedicalO')
                  income (ref = '1') comorbidity (ref = '0')
                  agecat (ref = '1')stage (ref = '0') raceeth (ref =
'1')
                  region_cat (ref = '1') sex (ref = '0');
      model survival (ref = '0') = network income comorbidity agecat
```

```
stage raceeth region_cat sex/
                  s dist=bin link=logit ddfm =bw;
      * network = 1 if largest cluster, = 0 else;
      random intercept /subject = cluster_id ;
      *random intercept/ subject = provider_id (cluster_id) s;
      *random intercept/ subject = id (cluster_id) s;
      estimate "Largest cluster" network 1/ exp cl;

run;

* Generalized linear mixed model for binary outcome, fixed effect
for largest cluster,
random effect for cluster and provider -->
to see if small clusters associated with long-term survival,
i.e. survival more than 12 months;
proc glimmix data=hannah.network_pedsf_pancr_medonly empirical;

      class cluster_id(ref = '26') provider_id (ref = 'NoMedicalO')
                  income (ref = '1') comorbidity (ref = '0')
                  agecat (ref = '1')stage (ref = '0') raceeth (ref =
'1')
                  region_cat (ref = '1') sex (ref = '0');
      model survival (ref = '0') = network_1 income comorbidity
agecat stage raceeth region_cat sex/
                  dist=bin link=logit ddfm =bw;
      random intercept /subject = cluster_id ;
      *network_1 = 1 if small cluster, i.e. cluster_id = 26, = 0
else;
      *random intercept/ subject = provider_id (cluster_id) s;
      estimate "Large cluster" network_1 1/ exp cl;

run;

* Generalized linear mixed model for binary outcome, fixed effect
for
network characteristic (i.e. high average degree),
random effect for cluster and provider -->
to see if high average degree associated with long term survival;
proc glimmix data=hannah.network_pedsf_pancr_medonly empirical;

      class cluster_id(ref = '26') provider_id (ref = 'NoMedicalO')
                  income (ref = '1') comorbidity (ref = '0')
                  agecat (ref = '1')stage (ref = '0') raceeth (ref =
'1')
                  region_cat (ref = '1') sex (ref = '0');
      model survival (ref = '0') = ave_degree_cat income comorbidity
agecat stage raceeth region_cat sex/
                  dist=bin link=logit ddfm =bw;
      random intercept /subject = cluster_id ;
      *ave_degree_cat = 1 if average degree is in the first Q (top
25%), = 0 else;
      *random intercept/ subject = provider_id (cluster_id) s;
      estimate "High average degree" ave_degree_cat 1/ exp cl;

run;
```

EMORY UNIVERSITY | Institutional Review Board

TO: Thi Minh Hieu Nguyen
Principal Investigator
*SPH: Epidemiology

DATE: August 24, 2016

RE: **Expedited Approval**
IRB00089231
Physician social network and variation in the odds of survival longer than three years of pancreatic cancer, liver cancer and stomach cancer patients

Thank you for submitting a new application for this protocol. This research is eligible for expedited review under 45 CFR.46.110 and/or 21 CFR 56.110 because it poses minimal risk and fits the regulatory category F[5] as set forth in the Federal Register. The Emory IRB reviewed it by expedited process on **August 24, 2016** and granted approval effective from **August 24, 2016** through **August 23, 2017**. Thereafter, continuation of human subjects research activities requires the submission of a renewal application, which must be reviewed and approved by the IRB prior to the expiration date noted above. Please note carefully the following items with respect to this approval:

- Scientific Protocol_Recalcitrant cancers ver. 8/5/2016
- Complete HIPAA Waiver granted

Any reportable events (e.g., unanticipated problems involving risk to subjects or others, noncompliance, breaches of confidentiality, HIPAA violations, protocol deviations) must be reported to the IRB according to our Policies & Procedures at www.irb.emory.edu, immediately, promptly, or periodically. Be sure to check the reporting guidance and contact us if you have questions. Terms and conditions of sponsors, if any, also apply to reporting.

Before implementing any change to this protocol (including but not limited to sample size, informed consent, study design, you must submit an amendment request and secure IRB approval.

In future correspondence about this matter, please refer to the IRB file ID, name of the Principal Investigator, and study title.

Thank you,

Emilie Scheffer
IRB Analyst Assistant

*This letter has been digitally signed*

| CC: | Hertzberg | Vicki | Surg Serv Admin |
|---|---|---|---|
| | McClellan | William | *SPH: Epidemiology |

---

Emory University
1599 Clifton Road, 5th Floor - Atlanta, Georgia 30322
Tel: 404.712.0720 - Fax: 404.727.1358 - Email: irb@emory.edu - Web: http://www.irb.emory.edu/
*An equal opportunity, affirmative action university*