

5-2019

Data Visualization of Treatment Outcomes for Tuberculosis Patients

Joy Jenkins

Follow this and additional works at: <https://scholarworks.uark.edu/ineguht>

Part of the [Industrial Engineering Commons](#), [Investigative Techniques Commons](#), [Other Operations Research, Systems Engineering and Industrial Engineering Commons](#), and the [Respiratory Tract Diseases Commons](#)

Recommended Citation

Jenkins, Joy, "Data Visualization of Treatment Outcomes for Tuberculosis Patients" (2019). *Industrial Engineering Undergraduate Honors Theses*. 64.

<https://scholarworks.uark.edu/ineguht/64>

This Thesis is brought to you for free and open access by the Industrial Engineering at ScholarWorks@UARK. It has been accepted for inclusion in Industrial Engineering Undergraduate Honors Theses by an authorized administrator of ScholarWorks@UARK. For more information, please contact ccmiddle@uark.edu.

Data Visualization of Treatment Outcomes for Tuberculosis Patients

An Honors Thesis Submitted in partial fulfillment of the requirements for Honors studies in the degree of
Bachelor of Science in Industrial Engineering

By

Joy Jenkins

Spring 2019

The University of Arkansas

Thesis Advisor: Dr. Shengfan Zhang

Thesis Reader: Dr. Chase Rainwater

Acknowledgements

This would not have been possible without the help of the Saxena Honors Academy Scholarship, granted by the University of Arkansas and grants sponsored by the Arkansas Academy of Engineers.

I would like to Dr. Don Catanzaro for entrusting me to be a part of this project and Dr. Zhang for introducing me to this topic, consistently guiding me, and being a wonderful advisor, mentor, and friend. She has been a wonderful reminder of what good engineering can do in the world. Maryam Kheirandish was essential to this project, so thank you for your hard work in cleaning the data. I would also like to thank Dr. Rainwater for reviewing my thesis and being a gracious, patient, and understanding professor who was always willing to help.

My parents, Steve and Carla Jenkins, have been my wonderful source of encouragement. My friends Lucie, Lizzy, Alex, and Raquel are the reasons I can find this world bright and beautiful. Finally, Jack- all we have is time.

Abstract

Tuberculosis is an infectious disease, and different treatments have been discovered over the years. However, patients may develop various drug resistance levels that affect the likelihood of becoming cured or dying. In this study, we sought to employ data visualization to explore the relationship between treatment trajectory, as indicated by smear and culture results in the follow-up tests and patient outcomes. A large sample of patients have been broken down by demographics including age, gender, and drug resistance status. Sankey diagrams were used to visualize the pathway progression of the patients over time split between two time periods- months 0-6 and months 6-24. It was determined that the most crucial months of treatment for all drug resistant types were during months 0-2, since there was high variability within that time frame for all demographics. It was also observed that younger patients were much more varying test results. It is thus recommended that the standards be updated to test every month for the first nine months of treatment in order to better track the pathway variety and that younger patients be more closely monitored throughout the treatment process. Future studies may investigate the possibility of creating a prediction diagram for patient pathway progression based on demographic status and past medical information.

Abstract 3

1. Background and Significance..... 6

 1.1 Past-Related Tuberculosis Studies 7

 1.2 Sankey Diagrams..... 8

2. Materials and Methodology 10

 2.1 Research Objectives 11

 2.2 Methodology 11

 2.3 Key Assumptions 12

 2.4 Program Diagraming 12

 2.5 Sankey Diagram Design..... 13

3. Diagram Results 15

 3.1 Male MDR..... 15

 3.1.1. Month 0-6..... 16

 3.1.2. Month 6-24..... 17

 3.2 Female MDR 18

 3.2.1. Month 0-6..... 19

 3.2.2. Month 6-24..... 20

 3.3 Male PanS..... 21

3.3.1. Month 0-6.....	22
3.3.2. Month 6-24.....	23
3.4 Female PanS	24
3.4.1. Month 0-6.....	25
3.4.2. Month 6-24.....	26
3.5 Male Neither	27
3.5.1. Month 0-6.....	28
3.5.2. Month 6-24.....	29
3.6 Female Neither	30
3.6.1. Month 0-6.....	31
3.6.2. Month 6-24.....	32
4. Discussion and Analysis	33
4.1 Gender (Male vs. Female).....	33
4.2 Month Range (0-6 Months vs. 6-24 Months).....	33
4.3 Drug Resistance Type (MDR vs. PanS vs. Neither)	33
4.4 Age (1 vs. 2 vs. 3)	34
6. Conclusion and Future Works	36
Bibliography.....	37

1. Background and Significance

Tuberculosis (TB) is an infection caused by the bacteria called *Mycobacterium tuberculosis*, which may lay dormant in the body or be an active infection. It is most commonly located in the lungs, but the bacteria may also reside within other organs such as the kidney or spine. The bacteria is transmitted through the air whenever an infected individual coughs, sneezes, etc., and only a few of the germs are needed to be inhaled in order for another person to become infected. It should be noted, however, that approximately 25% of the world's population has latent TB, who are not considered sick or able to transmit the disease.

According to the World Health Organization (WHO), diagnosis and treatment of tuberculosis averts millions of death, but there is still improvement to be made in regards to patient treatment (WHO 2017). Since there are varying reasons why some people do not receive treatment, such as lack of healthcare or knowledge of infection, more efficient and effective diagnosis and treatment is necessary. In addition to an effective diagnosis and follow-up testing, patients are in need of proper treatment plans. Since people have varying levels of multidrug resistance (MDR) levels, different drug treatments may not be as effective for certain patients. In previous studies, the status of “multi-drug resistant” (MDR) tends to be the most threatening to patients, as those with MDR tend to suffer from higher death rates (Abdelbary 2017). There also exists a resistant level called “PanS,” which has different resistances than MDR, but is still a significant threat since there is no treatment regimen. It is thus best to detect drug resistance as soon as possible so that another drug treatment may be implemented for the patient, as diagrams of untreated TB generally assume a total duration of only 2 years until self-cure or death (Tiemersma et. al. 2011).

Smear and culture tests are common follow-up tests for TB patients under treatment in order to monitor the progress of recovery. Smear tests study the signs for TB bacteria while culture tests indicate the growth of bacteria. Patient status falls into one of four possible categories, depending on whether the smear reading is positive or negative and whether the culture reading is positive or negative. According to Lab

Tests Online, when both smear and culture read positive, there is significant evidence to determine the presence of the TB micro-bacteria, and the patient is diagnosed as having tuberculosis. If only the smear is positive, the results are questionable, as the bacteria present may not be that of tuberculosis. If only the culture is positive, this is a rare occurrence, and usually indicates an error and that further testing is required. Whenever both smear and culture test as negative, it indicates that no tuberculosis bacteria is present (Chakravorty et. al. 2005). Using this information, patient prognosis can be tracked over time based on which of the four statuses possible.

Although those previously conducted projects are beneficial to understanding how to prevent infection and contamination, this research will help already infected patients by knowing if they are being properly treated. If not, it can be determined earlier so that more patients may start getting healthier sooner.

1.1 Past-Related Tuberculosis Studies

Currently, most tuberculosis studies have been concerned with factors that impact the infection of tuberculosis and how to better prevent the infection from developing (Wallis et. al. 2013). For example, the Tanzania study determined that smoking and HIV positivity are significant risk factors for mortality (Senkoro et. al. 2010, Mollel and Chilongola 2017, Philips et. al. 2017). These preexisting health concerns usually indicate that the patients begin treatment at some drug resistant level, making treatment even more difficult.

It has also been determined that diabetes and multi-drug resistance are risk factors for cavitory disease, usually a result of prolonged untreated TB (Zhang et. al. 2016 and Raznatovska and Khudiakov 2018). This indicates that although untreated TB may not directly result in death, it often leads to further health problems (Chan, Edward D. 2004). Since so many patients may begin treatment with some form of drug resistance level, the development of TB drug treatments is vital in progression of disease treatment (Wallis et. al. 2016).

The idea of determining optimal treatment combinations is a fairly recent development within the medical and engineering community. There have already been projects dedicated to determining a drug combination that results in the highest yield of cured patients with the lowest death yield (Vite 2014 and Ahuja Shama D. 2011). Recently, many studies are being conducted to visualize complex medical process for simpler understanding (Lerner 1996, Loorak et. al. 2016, Laptev, Orlov, and Dragunova 2017, and Alemasoom et. al. 2014).

1.2 Sankey Diagrams

The first draft of a flow diagram based on volume, perhaps the most commonly known diagram resembling the Sankey diagram principles, was completed in 1869 by Charles Joseph Minard illustrating the size of Napoleon's army during the Russian campaign of 1812-1813 (Minard). This graph shows the population of Napoleon's army was much larger at the beginning of the campaign, then slowly diminished over time and distance until Napoleon returned from Russia with considerably depleted forces.

The "Sankey diagram" itself was coined by Captain Matthew Henry Phineas Riall Sankey in 1898 in order to demonstrate the efficiency of a steam engine by visualizing the volume of energy (in this case, steam) moving throughout the system and its destinations. This visual allowed easier comprehension of wasted energy and demonstrating where heat was leaving and entering the engine. Sankey himself intended the diagram to mainly be used for energy flows within a confined system (Unknown, 1926).

Commercially, the concept of visualizing the data in an engaging way makes communicating ideas much more attractive for both engineers and clients. Since the eyes naturally are drawn to arrows with a larger width, this makes drawing attention to the most significant flows much easier, and if the arrow diminishes over time, then that makes the change between the two appear even more drastic.

This mapping of patient volume over time will be beneficial since it will clearly demonstrate the volume of patients that are being cured and the number of patients who die due to incomplete or insufficient

treatment. Khan et. al. (2015) used the Sankey diagram to better understand the trends of breast cancer patients, demonstrating its usage in the medical field.

2. Materials and Methodology

The data was collected for Informational Monitoring and Evaluation System of National Control Program of Tuberculosis in the Republic of Moldova (SYME TB) over the course of 2009 to 2016. Patients were only included in the data set if testing proved susceptibility to infection, preventing any unconfirmed tuberculosis patients from being included. The set of data is composed of the clinical records of 23,160 tuberculosis patients from diagnosis to the end of treatment. The dataset was cleaned and prepared by Maryam Alimohammadi.

For each patient, attributes were collected, and these attributes have been further divided into either a variable or a covariate. The covariate is the independent variable impacting the outcome of the dependent variables. The relevant divisions are as follows:

- a. Covariates: Gender (Male or Female), Age (less than 35 years of age, between ages 35 and 55, and greater than 55 years of age), TB group (MDR, PanS, or Neither)
- b. Variables: Smear, Culture, and Culture-Grade

By the end of the 2-year time study, there is an additional response variable designated “Final_Status,” indicating if the patient’s status classified as “Cured,” “Not Cured,” or “Death.” Based on the NIH definition (WHO 2010), to obtain the “Cured” status for non-MDR classified patients, there must be at least two consecutive smear/culture tests with negative readings. However, for MDR patients, there must be at least three consecutive time periods of negative smear/culture test readings. The “Death” status is designated to any patient who dies for any reason within the 24 month period. Cases that are neither “Cured” or “Death” cases are designated as “Not Cured.” The percentages of each demographic was determined (**Figure 1**).

Variable	Levels	Level Population	Description (if necessary)
Gender	Male	76.715%	
	Female	23.285%	
Age Group	1	39.468%	Less than 35 years of age

	2	44.038%	Between 35 and 55 years of age
	3	16.494%	Greater than 55 years of age
TB Group	MDR	23.246%	Multi-drug Resistance
	PanS	65.782%	Susceptible to both Isoniazid and Rifampicin
	Neither	10.972%	
Final Status	Cured	82.347%	
	Death	5.753%	
	Not Cured	11.9%	

Figure 1: Dataset Attribute Analysis

2.1 Research Objectives

The main objective of this research project was using data visualization in order to demonstrate patient treatment progression within certain demographics. By analyzing the typical trends of the treatment pathways of different demographics, the decision-making process for future TB treatment could be improved. Analyzing trends may also allow a better understanding of what groups are more at risk when diagnosed with a TB infection.

2.2 Methodology

The first main task of this project was developing programming code that would count the number of patients traveling along each possible path according to appropriate measures. Once all the possible paths were properly accounted for, a new Excel data file was created, only listing the pathways with non-zero values. This data set was then be loaded into Power BI for visualization. The visualization was completed using the Sankey Diagram extension on Microsoft's Power BI, available for free download online. The visual was then rearranged in order to correspond with the defined layout of the axis.

2.3 Key Assumptions

Before programming could begin, key assumptions had to be established, as the big data was not perfect and had to be cleansed for proper diagraming. The key assumptions are as follows:

- a. If a patient missed a check-up session, then it is assumed that his or her status remained unchanged until the next check-up attended
- b. This project would only be concerned with the patients who had test readings at time 0. This is to ensure that only patients present for the full two years are analyzed.
- c. Smear would only be classified as “Positive” or “Negative”. Although the dataset had varying degrees of smear positivity, in order to maintain simplicity in the diagram, culture was the only reading accounting for varying degrees.
- d. Since Smear indicates the presence of the TB bacteria, any culture readings would be considered false positives, since culture reads the severity of the infection. Therefore, Negative Smear only has a 0 Culture grade, but Positive Smear has varying grades of Culture.
- e. If a patient were to leave the system before the time study deadline of 24 months with the status “Not Cured,” they would be counted as in the “Death” category. This is due to the fact that if the patient weren’t cured yet, it is assumed that they would continue treatment before the 24 month time period was up. The only time “Death” and “Not Cured” are differentiated is in the final month.

2.4 Program Diagraming

The programming was completed using formulas already installed in Excel. In order for this visualization to form properly, the data was arranged in a format mapping out the quantities from origins to targets. This determined that in addition to the programming separating the data according to different demographics, it had to be recorded in a specific way in order to create a Sankey diagram in accordance with the extension

requirements. The following format required for the Sankey diagram is similar to the example below (**Figure 2**).

Origin	Target	Quantity
A	B	10
A	C	2
B	D	7

Figure 2: Sankey Diagram Sample

2.5 Sankey Diagram Design

For each diagram, the y-axis is a reading of the patient's status, with the positive direction indicating improvement and the negative direction signifying health decline. Since each test reading contains a smear and culture reading, the axis is sorted first by smear reading, with negative readings lying on the positive direction of the y-axis, then by culture, with lower positive readings lying in the positive y-axis direction. The x-axis demonstrates the progression over time in units of months.

In order to keep the diagrams relatively simple, the diagrams are broken up into 2 distinct time periods: Month 1-6 and Month 6-24. Since the first 6 months are the most crucial in the treatment process, it was deemed important that those months be separated from the later months of testing. Each age group can be seen off to the side (A for age group 1, B for age group 2, C for age group 3), and the number of patients included in the first time period of the figure.

In addition to separating the charts into different time periods, they are also divided according to various patient demographics. First, the pathways were constructed when only separated by two categories- gender and drug status. Drug status is the most crucial demographic to study, as this is the most impactful attribute directing patient treatment progression. In addition to drug resistance level, the patients were also separated based on

gender, since this was more insightful than separation based on age groups. Once the Gender/TB Group graphs were completed, they were further broken down based on age groups within each demographic.

It was deemed vital that the diagrams were overall organized with a clear direction for decline and improvement. Since the diagrams are automatically generated by the program, any visual complication is not intentional.

3. Diagram Results

In this section, each general diagram will be analyzed before being broken down according to age groups. This general analysis of a broader group further broken down with the more specific graphs allowed further insight into an initial broad analysis. Within each section, the different month sets can be seen, each further broken down by age group.

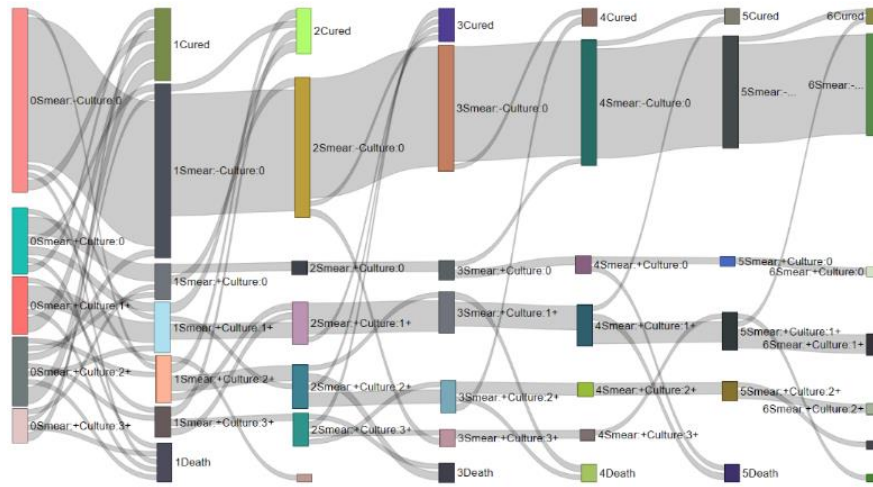
3.1 Male MDR

For the first 6 months within this demographic, most patients don't change test results between intervals (**Figure 3**). In month 1, the only patients who died were previously tested with readings of negative smear results. However, later in month 5, the only patients designated with "Death" had previously had positive smear and culture readings, but this is to be expected. Throughout the 6 month period, there are multiple cases of patients with positive TB readings being designated as "Cured" in the next time period. The majority of cured patients previously tested having no signs of smear, but from the patients who had previously shown signs of smear, there was an even distribution getting cured.

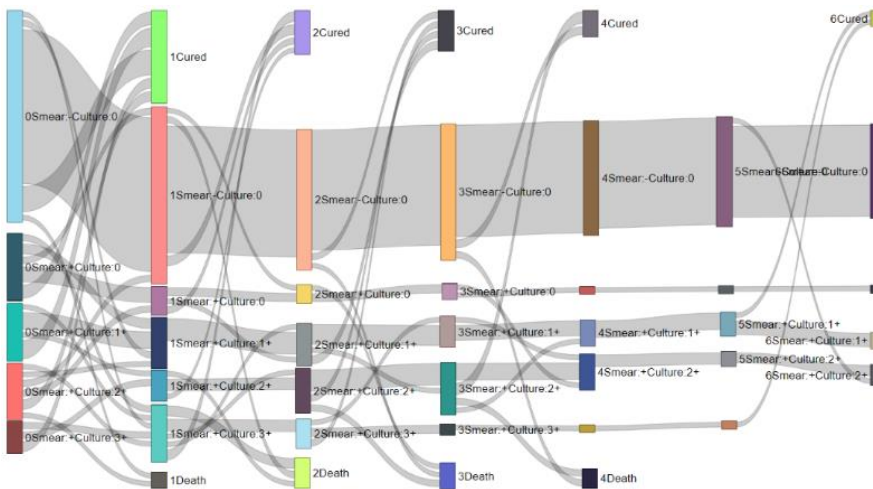
During the remaining months, the patients were much more stagnant in test readings, with relatively few patients classified as "Cured" or "Dead" in each time period. In the last time period, however, there were no deaths occurring, only instances of patients becoming designated as "Cured" or continuing treatment (**Figure 4**).

3.1.1. Month 0-6

a:
145



b:
165



c:
52

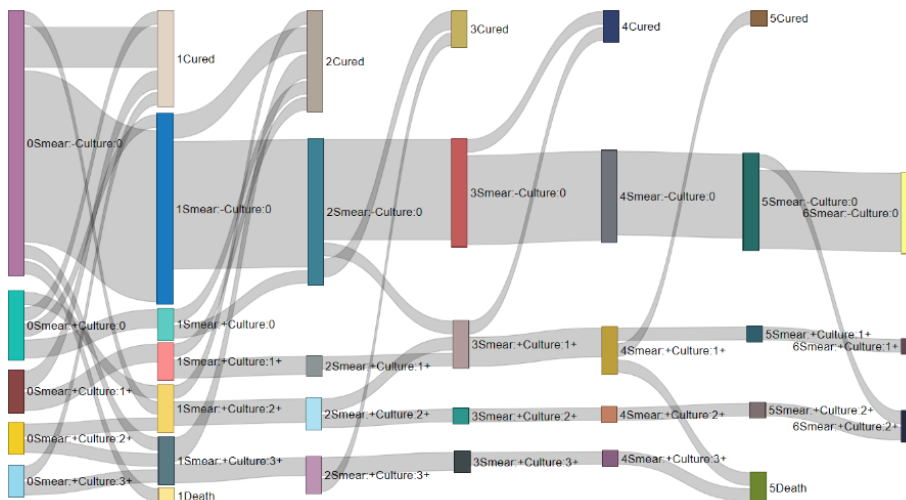
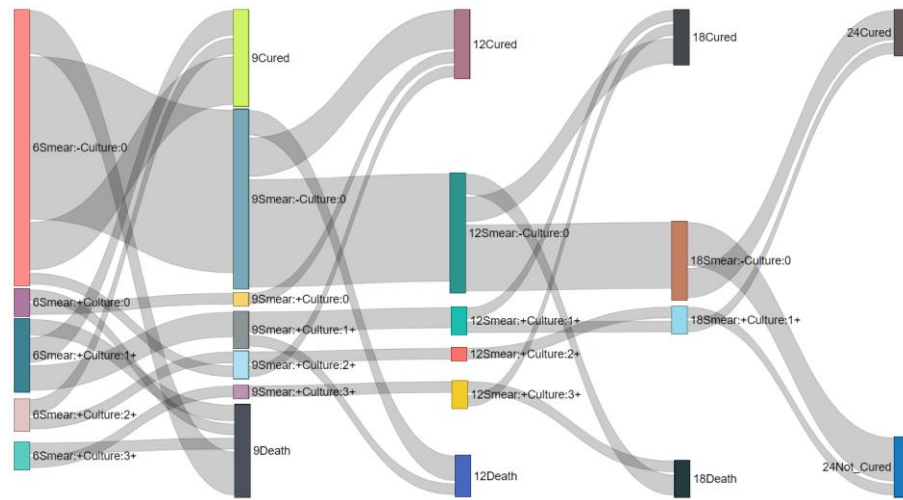


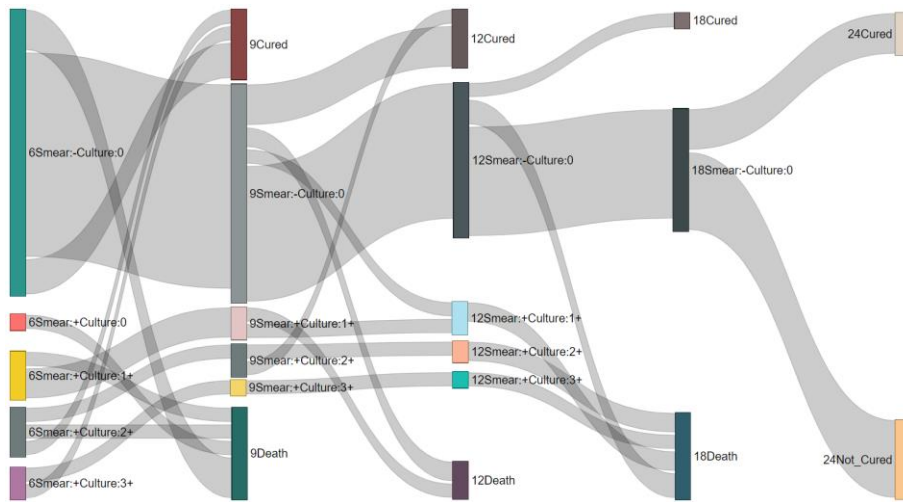
Figure 3: Sankey Diagram for Male MDR Patients Months 0-6

3.1.2. Month 6-24

a:
66



b:
59



c:
14

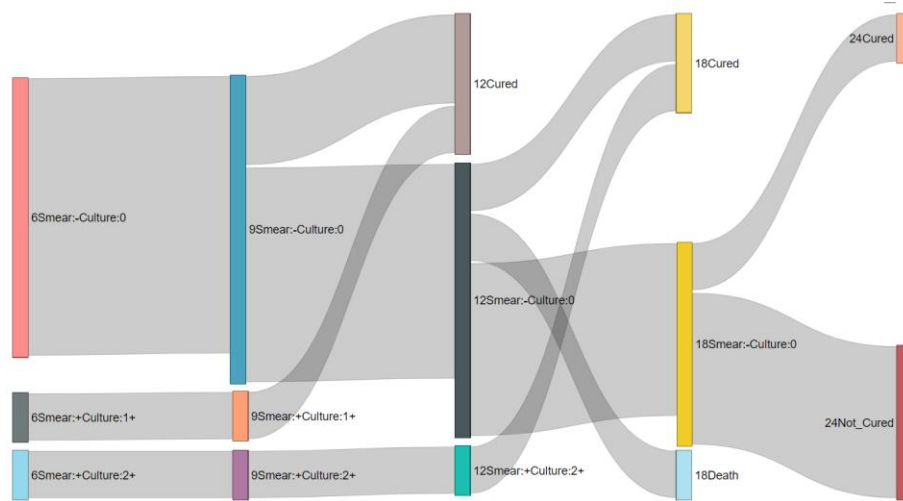


Figure 4: Sankey Diagram for Male MDR Patients Months 6-24

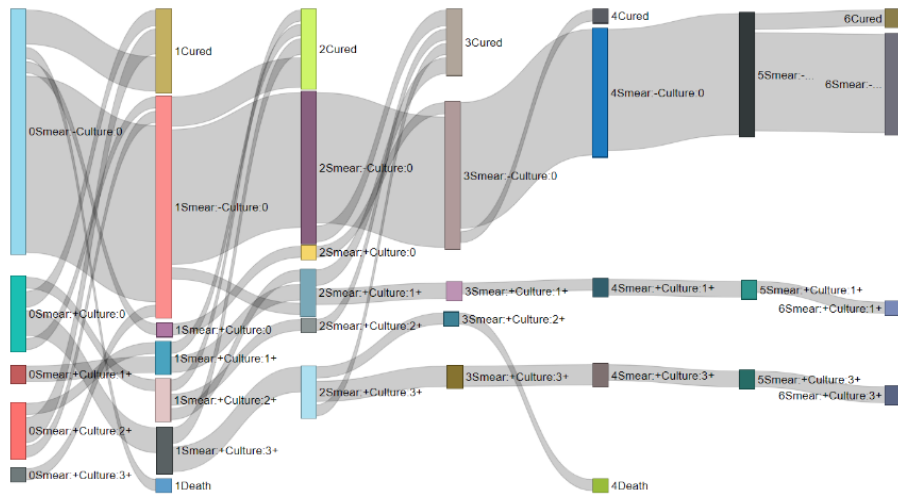
3.2 Female MDR

During the first 6 months, the only occurring deaths were in age groups 1 and 2 (**Figures 5a and 5b**), so this may signify that younger patients are more at risk for death during this first time period. However, the most patients recorded as being cured were also in the first two age groups, so this could also signify that the younger ages work in more extremes. That is to say, younger patients don't tend to stay stagnant in treatment status and are quicker to reaching either a "Cured" or "Death" status within the first six months of treatment.

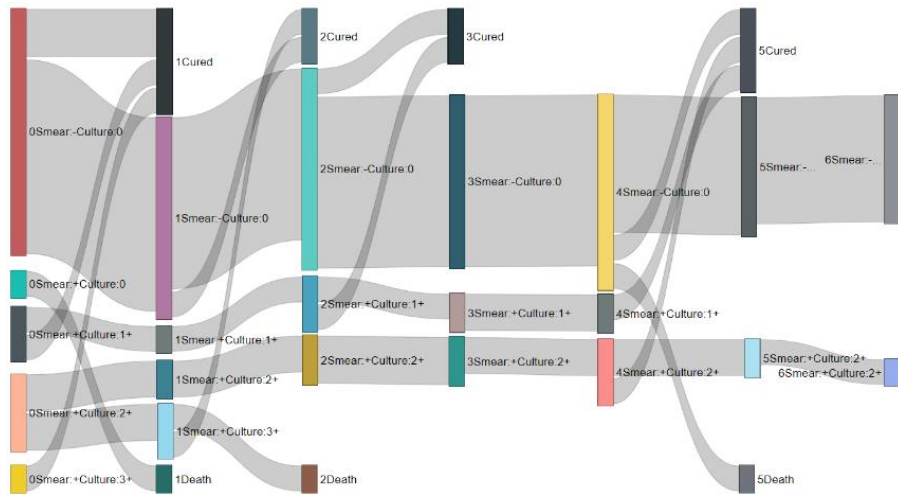
For the remaining 18 months, there were few instances of degrading in test results before dying. Instead, most patients stayed within the past test reading, only moving when exiting the system. Again, there were no deaths recorded in the final time period, but there appeared to be significant portions of patients dying in other months, with the most deaths occurring in either age group 1 or 2 (**Figures 6a and 6b**).

3.2.1. Month 0-6

a:
62



b:
27



c:
13

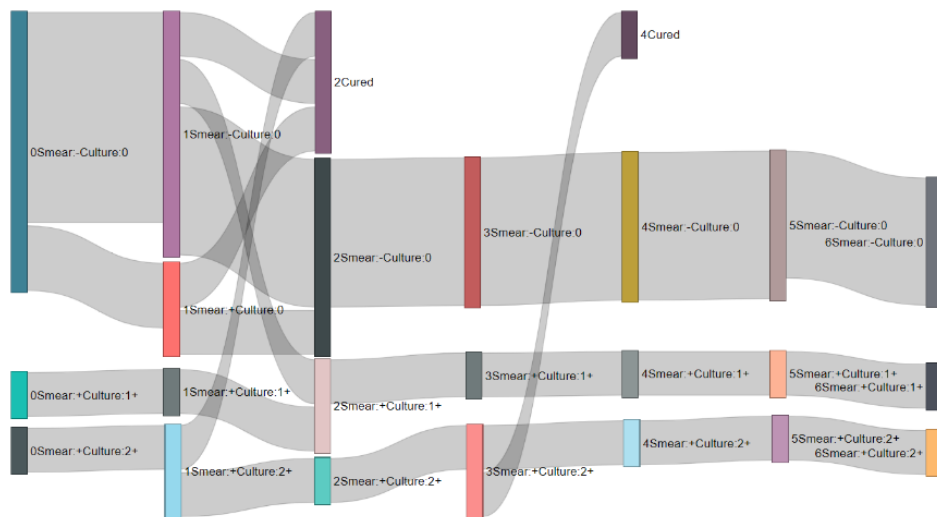
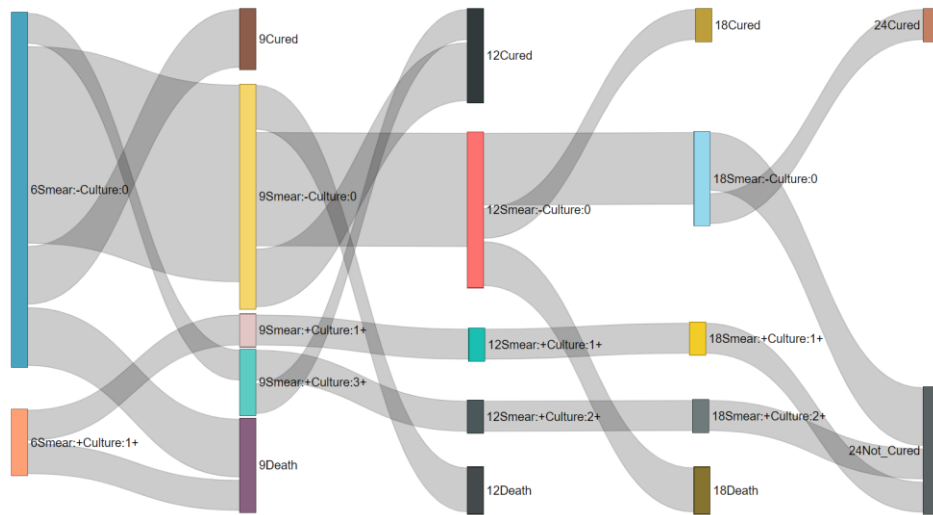


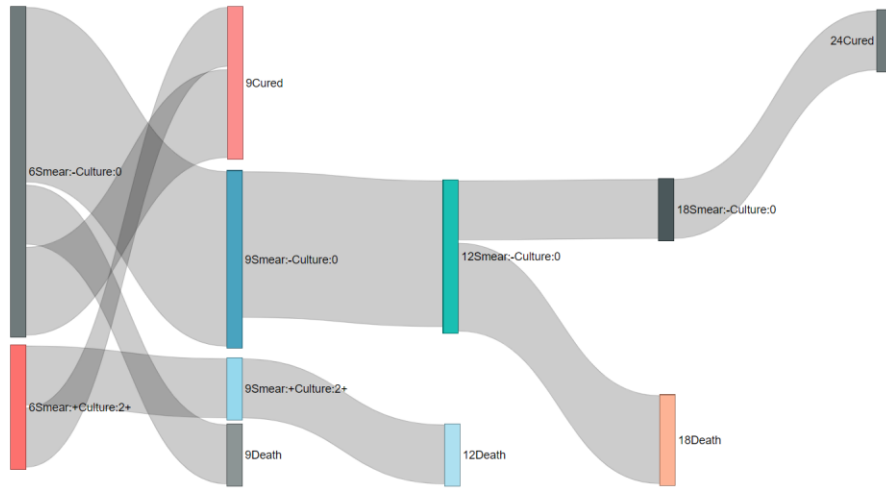
Figure 5: Sankey Diagram of Female MDR Patients Months 0-6

3.2.2. Month 6-24

a:
21



b:
10



c:
6

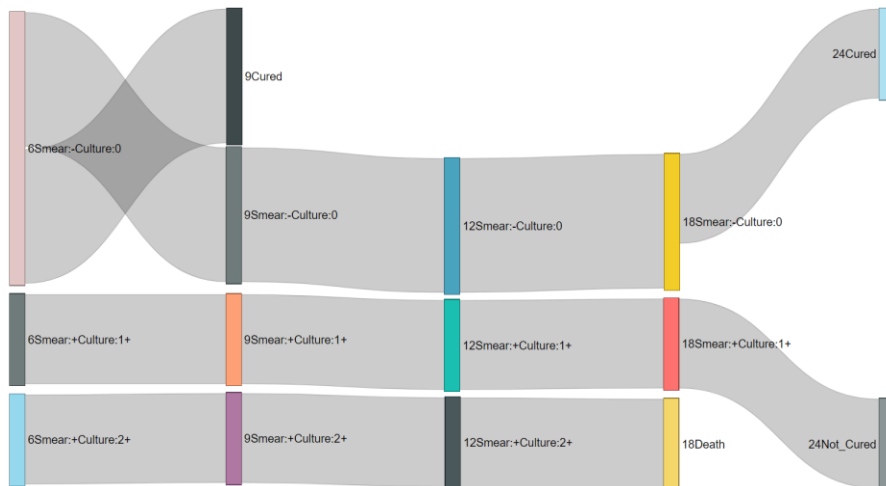


Figure 6: Sankey Diagram of Female MDR Patients Months 6-24

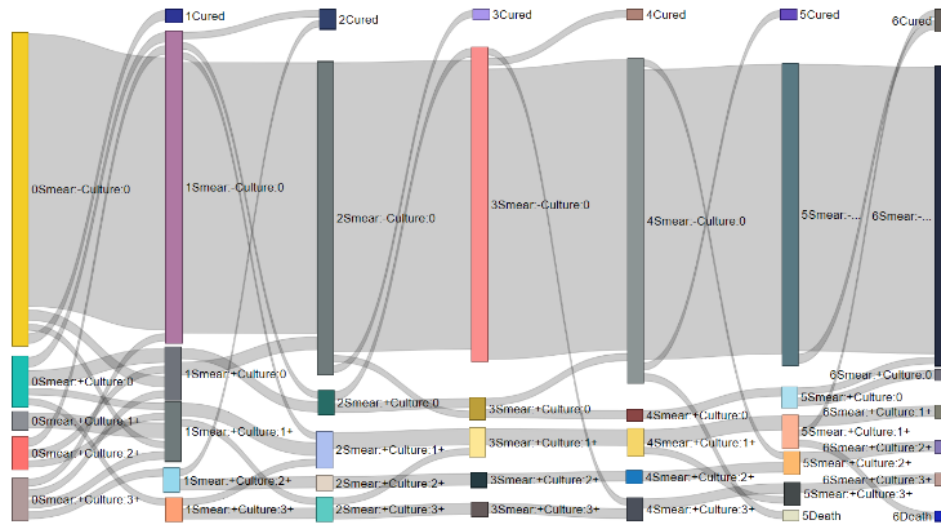
3.3 Male PanS

Within the first 6 months, the largest proportion of patients within this time period were diagnosed as having no evidence of the TB bacteria, implying that these patients were simply waiting in order to be officially designated as “Cured,” and in fact these patients were usually the only ones designated into the “Cured” status. However, in younger ages, there were other patients from different test categories (**Figure 7a and 7b**). It should also be noted that the first two age groups are the only ones suffering from patients dying during the course of treatment, and the third age group had no deaths, but also very little “Cured” patients.

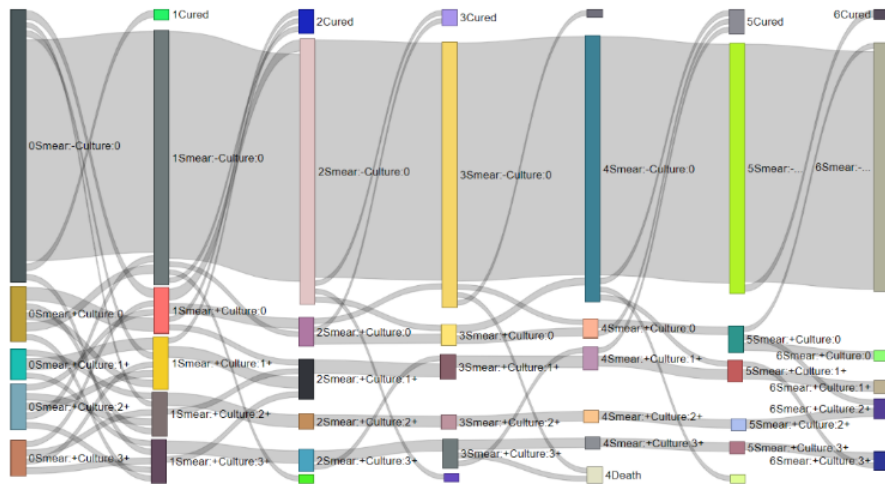
From months 6-24, there were very few patients with positive readings for the TB infection, and many of these patients were eventually classified as “cured.” Month 9 contained the largest proportion of patients becoming cured, but that was also the month with the highest records of patients dying, though it was much smaller than the amount cured (**Figure 8**). Patients in the second age group tended to leave the system faster than the other age groups, with the oldest patients remaining in the treatment system longer (**Figure 8b**).

3.3.1. Month 0-6

a:
486



b:
623



c:
359

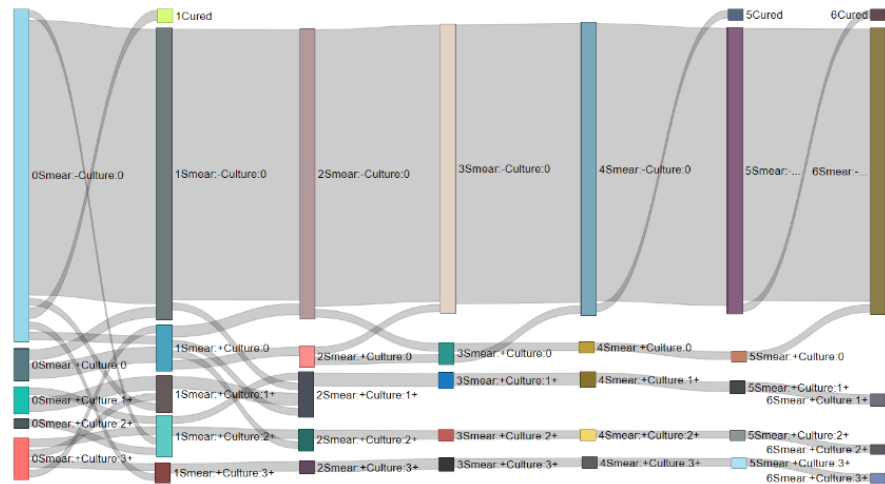
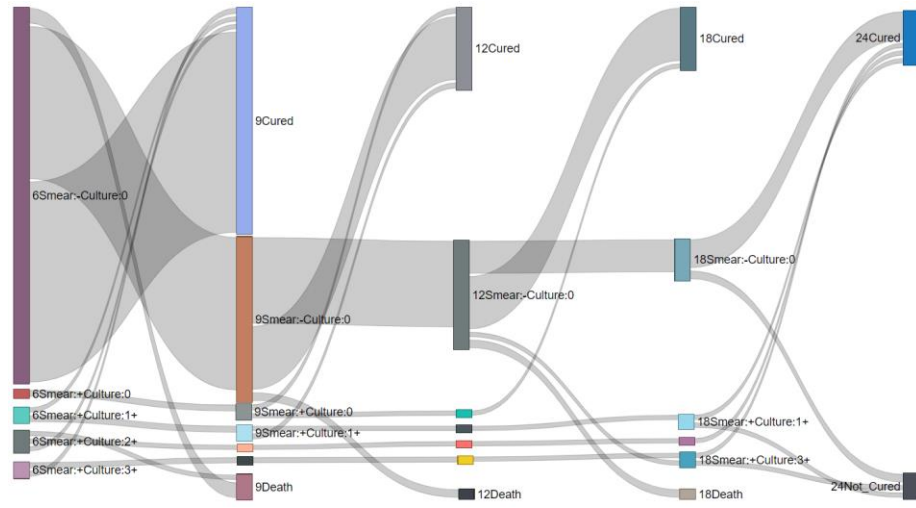


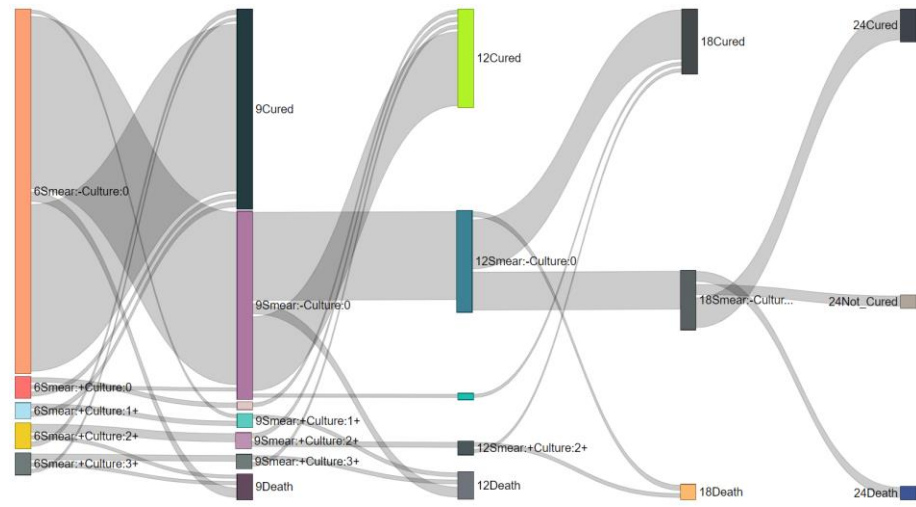
Figure 7: Sankey Diagram of Male PanS Patients Months 0-6

3.3.2. Month 6-24

a:
453



b:
585



c:
320

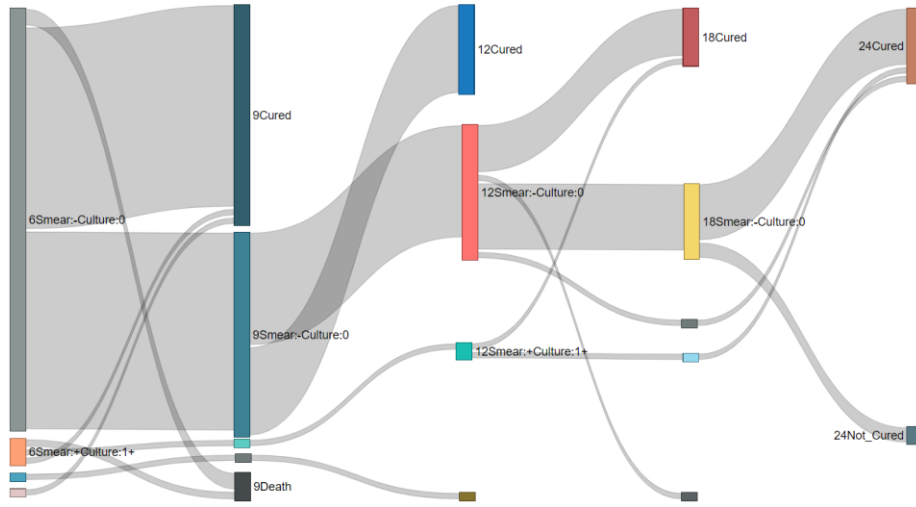


Figure 8: Sankey Diagram of Male PanS Patients Months 6-24

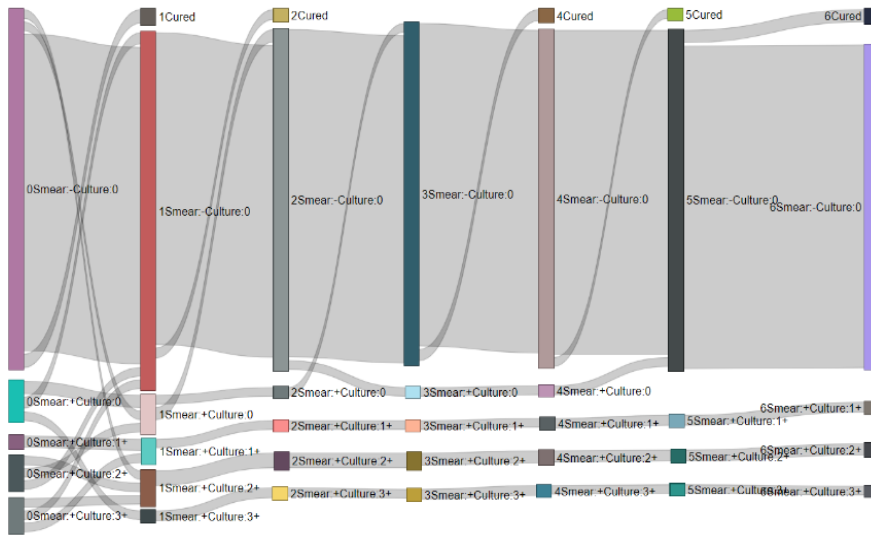
3.4 Female PanS

Within the first six months, the second age group recorded the highest rates of cured patients and there were only occurrences of death at month 5, but there was a more even spread of different test results, especially compared to age groups 1 and 3, which have the majority of patients testing of negative smear and culture results (**Figure 9a and 9c**). Age group 3 also only had death at month 5, but there were sparse instances of patients becoming classified as cured (**Figure 9c**).

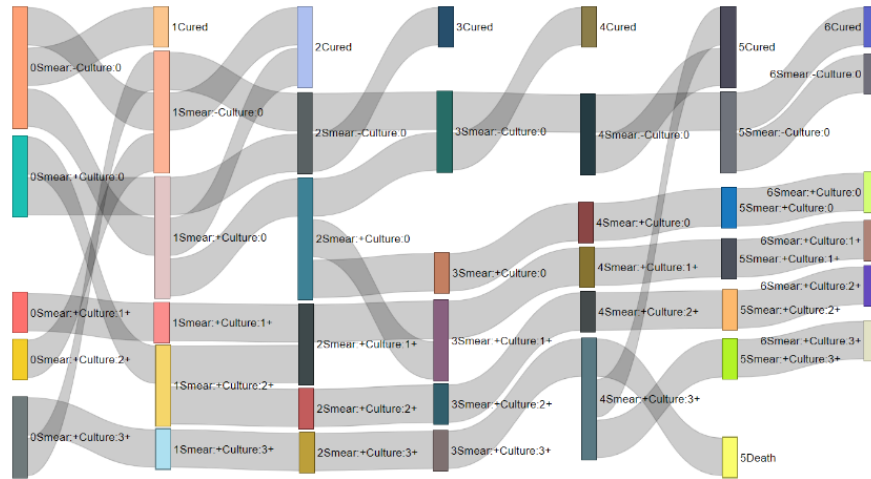
During the remaining months in the treatment process, very few patients remained in the treatment process until month 24. Most patients were cured at month 9, and patients who died were evenly spread between months 9 and 12, with some occurring at month 18 (**Figure 10**). It should be noted that most deaths were patients who had not previously shown signs of the TB infection. Again, the majority of the patients within this demographic were showing no signs of infection, with very few having a positive smear reading.

3.4.1. Month 0-6

a:
284



b:
185



c:
99

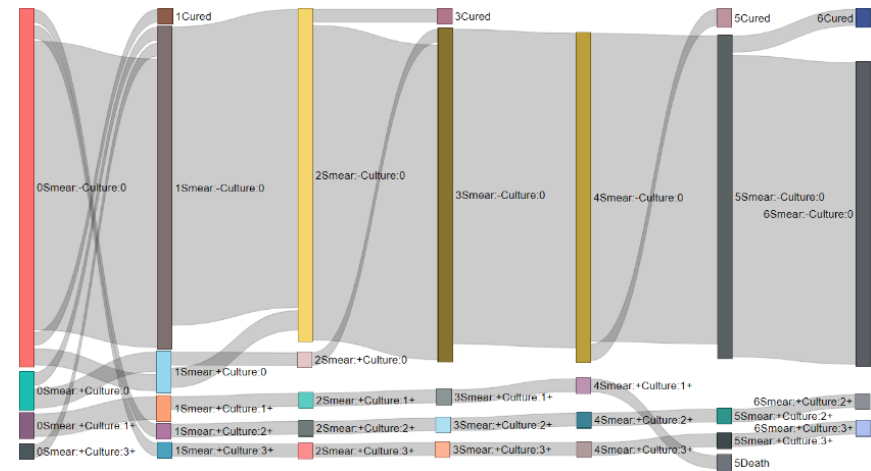
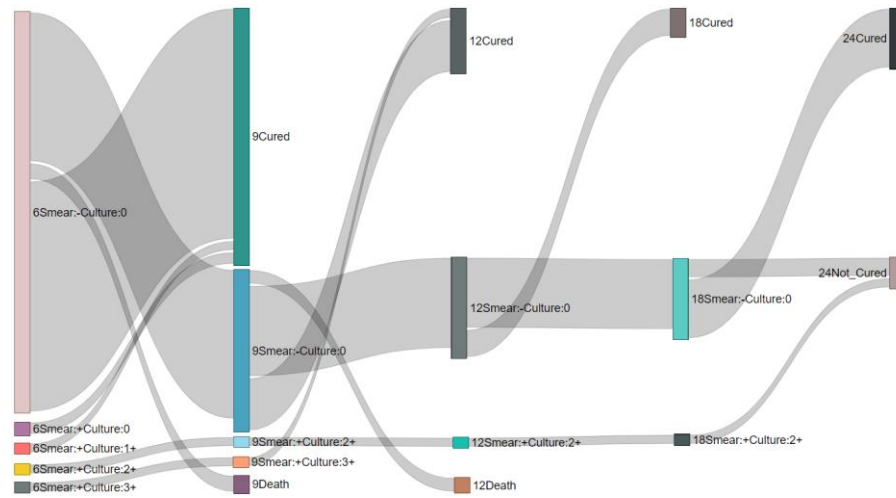


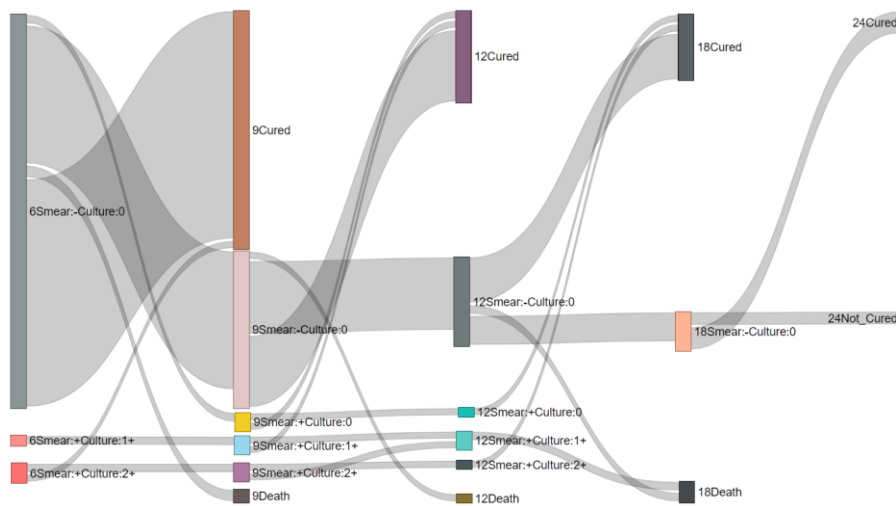
Figure 9: Sankey Diagram of Female PanS Patients Months 0-6

3.4.2. Month 6-24

a:
271



b:
171



c:
92

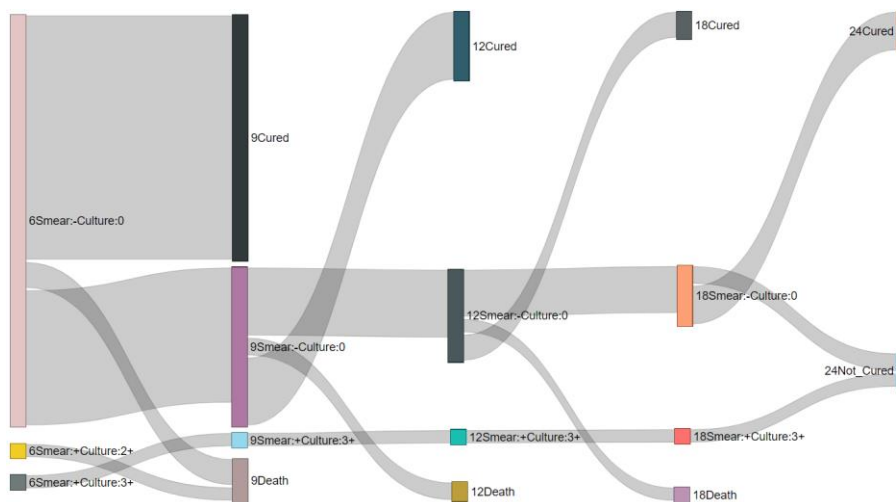


Figure 10: Sankey Diagram of Female PanS Patients Months 6-24

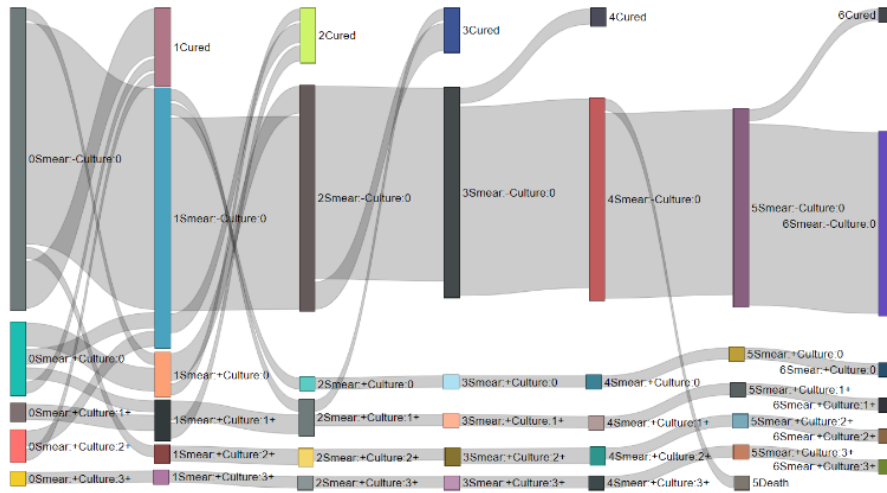
3.5 Male Neither

From the start of the time study to month 6, the highest rates of patients becoming classified as “Cured” or “Death” occurred up to month 3. After month 3, there were much less significant instances of patients either dying or being cured, instead staying stagnant in the current test readings (**Figure 11**). This is better than the patients’ health readings declining, but there is concern that many patients aren’t improving despite treatment.

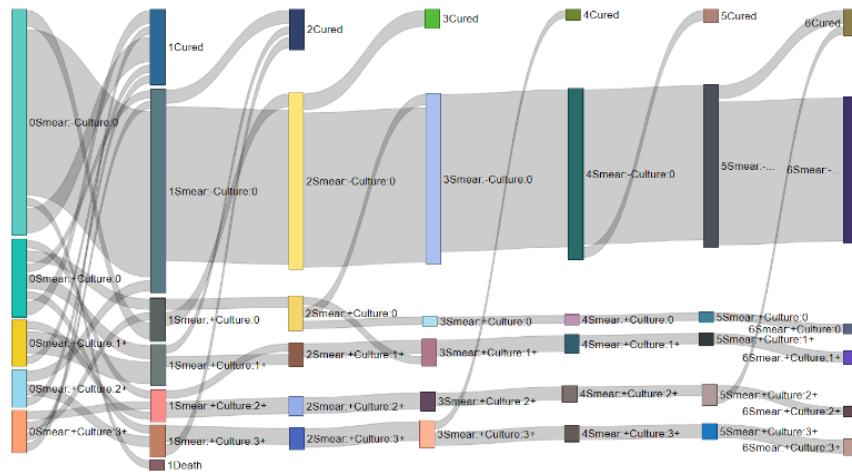
For the remainder of the time study, there were large portions of patients being cured at months 9 and 18, but the highest records of death were also at month 9. No patients died at month 24, and in fact a relatively small amount needed to continue treatment. The youngest age group experienced the highest rates of death, with patients dying at months 9, 12, and 18 (**Figure 12a**). All age groups showed not much test reading variety, instead keeping consistent test readings for each time period (**Figure 12**).

3.5.1. Month 0-6

a:
80



b:
97



c:
29

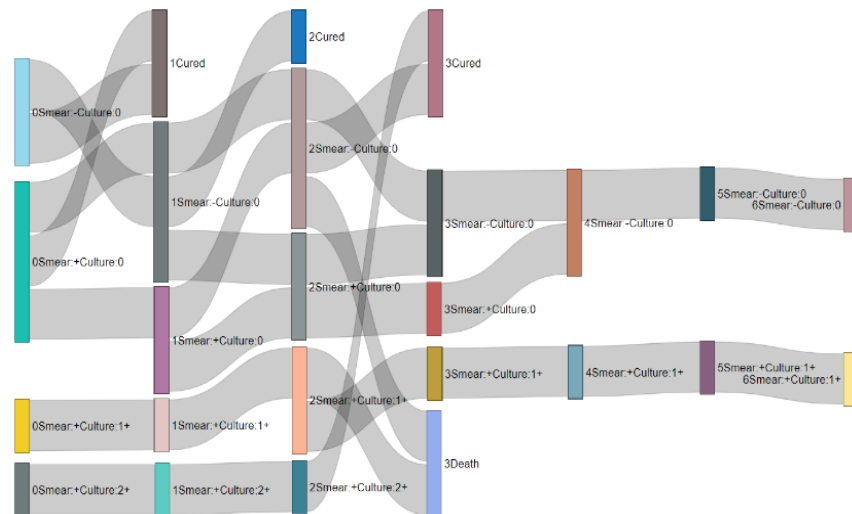
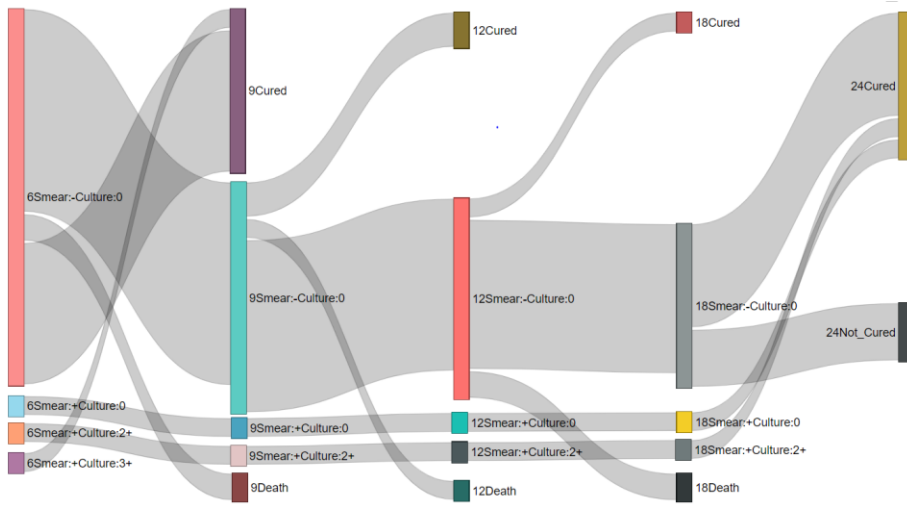


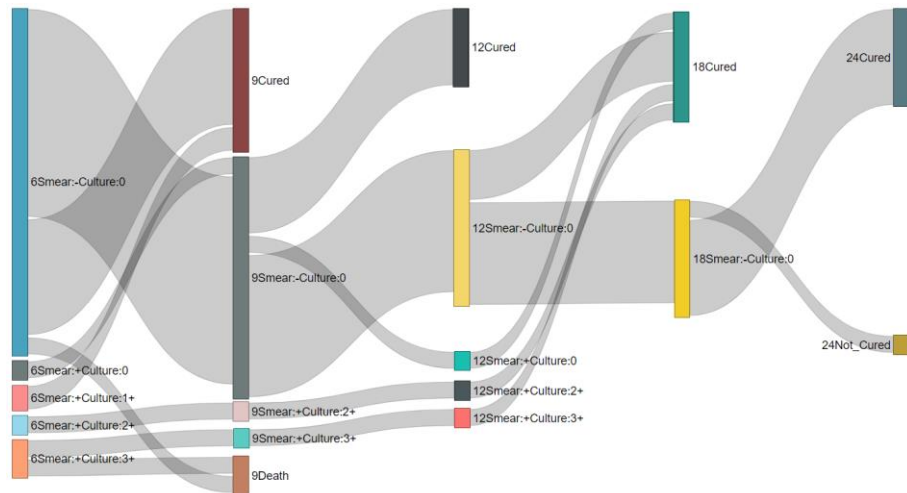
Figure 11: Sankey Diagram of Male Neither Patients Months 0-6

3.5.2. Month 6-24

a:
47



b:
53



c:
19

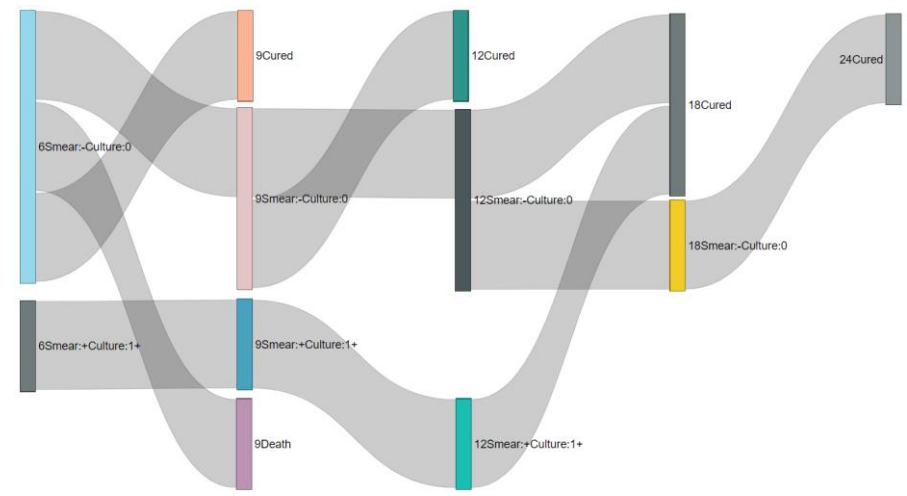


Figure 12: Sankey Diagram of Male Neither Patients Months 6-24

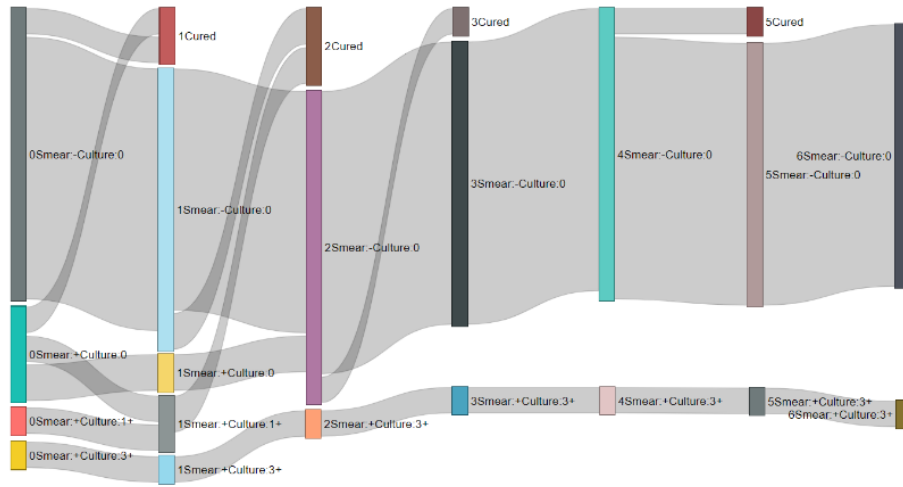
3.6 Female Neither

For the first six months, the majority of patients within this demographic have tested negative for traces of smear, and those who show positive signs either quickly exit the process or improve readings by testing negative for smear. Age group 1 is the only group with a maintained different test type for all six months (**Figure 13a**).

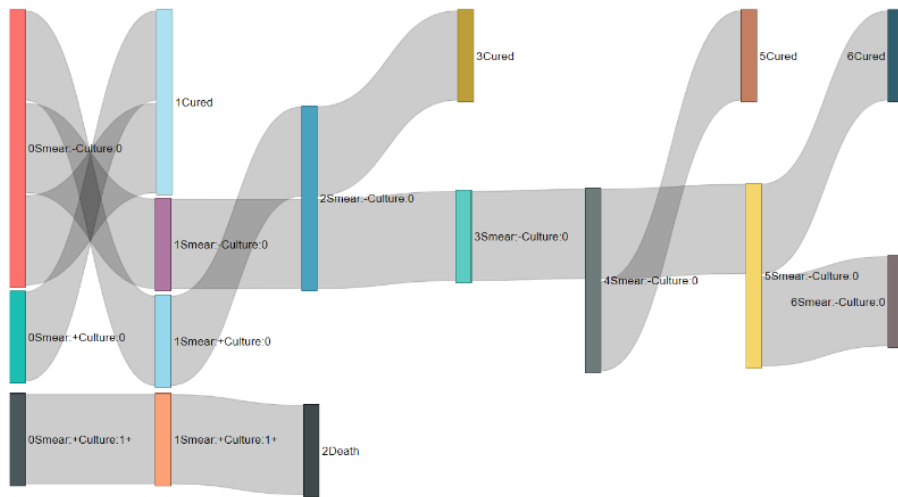
During the remaining 18 months, the only age group with a test reading other than negative smear is the first age group (**Figure 14a**). Death only occurred at months 9 and 12, but the majority of the patients within this demographic were ultimately cured, with a smaller portion continuing treatment. Since most patients had a negative smear reading, most were waiting for the “Cured” status to become official, they didn’t degrade in test readings.

3.6.1. Month 0-6

a:
31



b:
17



c:
18

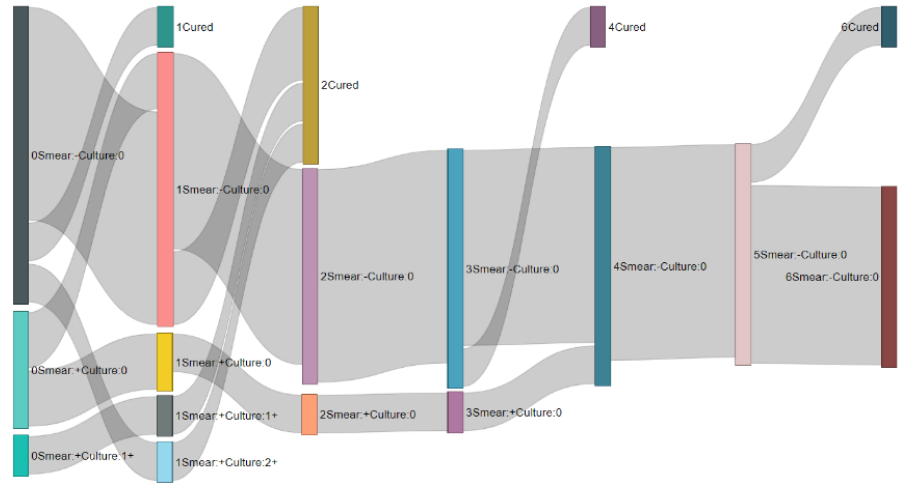
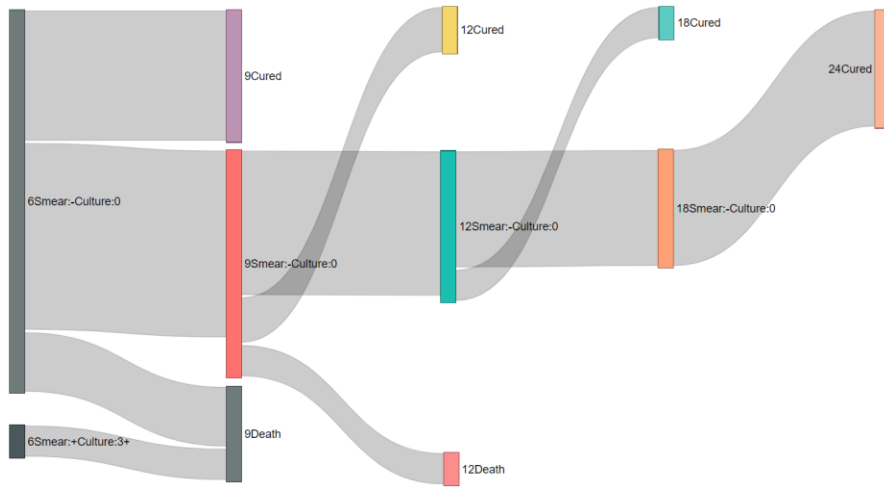


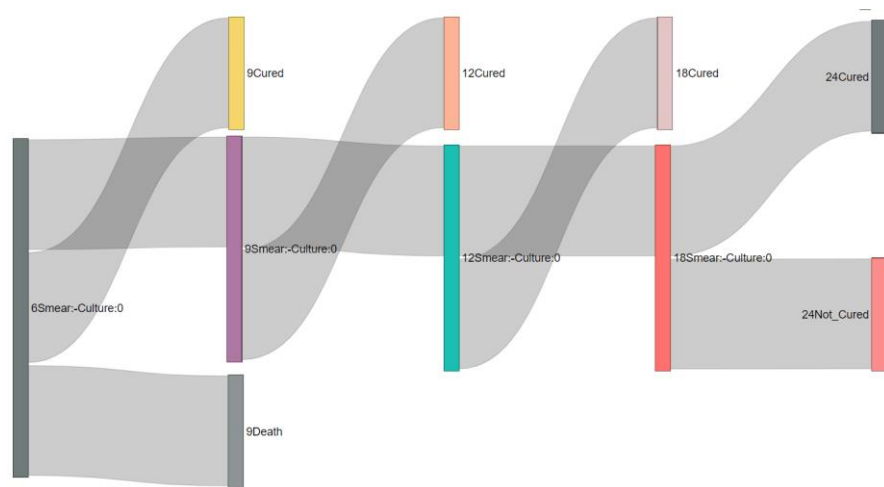
Figure 13: Sankey Diagram of Female Neither Patients Months 0-6

3.6.2. Month 6-24

a:
24



b:
13



c:
9

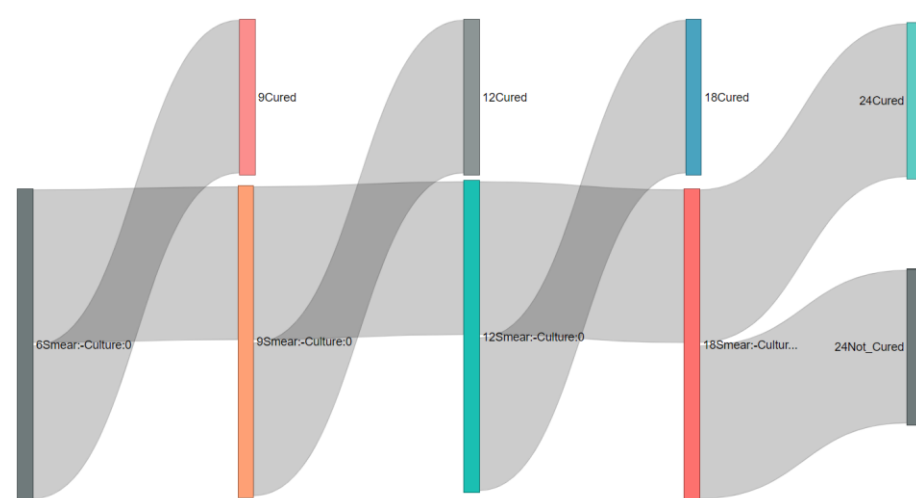


Figure 14: Sankey Diagram of Female Neither Patients Months 6-24

4. Discussion and Analysis

In this section, the trends and comparisons between different demographics were completed. Within all six demographics, there were instances of patients testing positive for smear and/or culture, but then being considered cured in the next time period. Since it was required that patients test negative for smear/culture for at least two consecutive time periods in order to be considered “Cured,” this could be the result of a recording error or patients being clear for 2 months before the final month.

4.1 Gender (Male vs. Female)

When comparing the males and females in this study, it should be noted that there were many more male patients than female, so the male patients may offer a more complex and detailed picture than the female demographic. The male population experiences higher levels of death and cured patients, but since the male population is larger than the female, this outcome is understandable. Overall, there didn’t appear to be a significant difference between the gender demographics within this study.

4.2 Month Range (0-6 Months vs. 6-24 Months)

Since patient readings are described every month for the first six months, there is more variety in test results than in the months following, where the patient is only tested every 3-6 months. Since the first few months of treatment are the most crucial, the variety in most test results is understandable, as different treatments are tested in order to determine if they are effective.

In the later months, month 9 has shown to be a major month, with many records of patients becoming cured or dying. After month 9, most test readings remain stagnant, with some proportions leaving the system.

4.3 Drug Resistance Type (MDR vs. PanS vs. Neither)

The MDR patients experienced the highest proportions of death rates among the three different drug resistant levels. These patients also experienced the most complexity in changing test grades. There were much more

variation in the amount of routes different patient test readings would take, making trend identification more difficult.

The patients diagnosed with PanS resistance status tended to have much lower rates of positive smear and culture readings. That is, the majority of patients would test with no readings for the TB bacteria. This is odd, however, because there are very few reports of cured patients in the first six months, even though the patients tended to stay consistent in testing negative. Once the time study reached month 9, however, there were much greater proportions of patients becoming designated as “Cured.” This aligns with the WHO’s consideration that PanS patients should be cured by month 9, however, there was a decent number still within the treatment process. There were also few reports of patients dying, so it is unclear why it took so long for patients to be designated “Cured.”

The patients neither MDR nor PanS experienced the highest rates of deaths and cured patients during the first 3 months of treatment, but then most patients remained stationary in status readings. There weren’t any particularly large ratings of cured patients until month 9, coinciding with the WHO’s predications of being cured within 9 months.

4.4 Age (1 vs. 2 vs. 3)

The youngest age group (1) was the most responsive over the course of this study. That is, younger patients tended “exit the system” faster (i.e. were cured or died), and over the course of time, these patients had more various test readings before eventually exiting the system. Older patients (age group 3) tended to stay within the system longer with the same test results. Age group 2 results was much more dependent on drug resistant status. For instance, MDR patients had much more variety in test results than PanS patients, who in turn had more variety than patients without a drug resistance.

4.5 Overall Analysis

One of the main takeaways concerns the rate at which the patient reports back for a check-up. The standard is 2 months, but after observing the importance of the first six months, as well as how important Month 9 is, it could be beneficial to schedule monthly appointments for the first 9 months of treatment instead of the first 6. This could allow further insight into when the patients stop drastically exiting the system and settle into a stagnant test result for an unexpected duration of time. However, a great emphasis should be placed on the time period containing the first two months, since these months contained the most variability.

In addition to possible revision to the standard checkup rate, age should also be investigated. Since the younger patients are much more reactive during treatment, these patients should be more closely monitored in the future.

6. Conclusion and Future Works

Tuberculosis is a disease that affects different portions of the world, be it active or latent. In order to better understand the trends of patient treatment progress, diagrams were created to display the quantities of patients over time according to test results. Using these diagrams, trends were identified overall and according to different demographics.

The main shortcoming of this research is that only quantities of patients were tracked throughout the treatment process. Although this allowed general trends to be identified from the different diagrams, this did not track individual patient progression.

In future studies, it would be beneficial to design a new visualization diagramming program in order to track typical patient progression over time according to each demographic whilst including the variation along the path. The inclusion of history along the entire treatment path can allow for future patients to compare his or her own progression path to a standard for treatment.

Bibliography

- Abdelbary, B.E., et al. *Predicting treatment failure, death and drug resistance using a computed risk score among newly diagnosed TB patients in Tamaulipas, Mexico*. *Epidemiol. Infect.* (2017), 145, 3020–3034. Cambridge University Press 2017. Retrieved 08/28/2018 doi:10.1017/S0950268817001911
- Ahuja, Shama D.et. al. *Multidrug Resistant Pulmonary Tuberculosis Treatment Regimens and Patient Outcomes: An Individual Patient Data Meta-analysis of 9,153 Patients*. *PLoS Med* Vol. 9, Issue 8, 2011. doi:10.1371/journal.pmed.1001300
- Alemasoom, Haleh et. al. *Interactive Visualization of Energy Systems*. *The Visual Computer*, 2014. DOI 10.1109/CW.2014.39
- Chakravorty, Soumitesh et. al. *Utility of Universal Sample Processing Methodology, Combining Smear Microscopy, Culture, and PCR, for Diagnosis of Pulmonary Tuberculosis*. *Journal of Clinical Microbiology*, June 2005, pp. 2703-2708. doi:10.1128/JCM.43.6.2703–2708.2004
- Chan, Edward D. et. al. *Treatment and Outcome Analysis of 205 Patients with Multidrug-resistant Tuberculosis*. *American Journal of Respiratory and Critical Care Medicine*. May 2004, Vol. 169, Issue 10, p. 1103. DOI: 10.1164/rccm.200308-1159OC
- Khan, Azam et. al. *Benefits of visualization in the Mammography Problem*. *International Journal of Human-Computer Studies*, Vol. 83, Nov. 2015, pp. 94-113. <https://doi.org/10.1016/j.ijhcs.2015.07.001>
- Lab Tests Online. “Acid-Fast Bacillus (AFB) Testing.” Available at <https://labtestsonline.org/tests/acid-fast-bacillus-afb-testing>. Retrieved 10/04/2018.
- Laptev, V.V., Orlov, P.A., and Dragunova, O.V. *Visualization of Dynamic Data Structures with Flow Charts in Web Analytics*. *St. Petersburg State Polytechnical University Journal. Computer Science. Telecommunications and Control Systems*, 2017, Vol. 10, No. 4, Pp. 7–16. DOI: 10.18721/JCSTCS.10401

- Lerner, Seth E. *Analysis of Risk Factors for Progression in Patients with Pathologically Confined Prostate Cancers after Radical Retropubic Prostatectomy*. The Journal of Urology, Vol. 156, pp. 137-143, 1996.
<https://www.ncbi.nlm.nih.gov/pubmed/8648775>
- Loorak, Mona Hosseinkhani et. al. *TimeSpan: Using Visualization to Explore Temporal Multi-dimensional Data of Stroke Patients*. IEEE Transactions on Visualization and Computer Graphics, Vol. 22, No. 1, Jan. 2016. *Digital Object Identifier no. 10.1109/TVCG.2015.2467325*
- Mollel, Edson W. and Chilongola, Jaffu O. *Predictors for Mortality among Multidrug-Resistant Tuberculosis Patients in Tanzania*. (2017) Volume 2017, Article ID 9241238, 6 pages
<https://doi.org/10.1155/2017/9241238>
- Philips, Lauren et. al. *The association between tuberculosis and the development of insulin resistance in adults with pulmonary tuberculosis in the Western sub-district of the Cape Metropole region, South Africa: a combined cross-sectional, cohort study*. BMC Infectious Diseases, Vol. 17, Issue 570, 2017. DOI 10.1186/s12879-017-2657-5
- Raznatovska, O.M. and Khudiakov, G.V. *Factors of chemoresistant pulmonary tuberculosis progression in patients receiving palliative treatment*. Zaporozhye medical journal 2018; Vol. 20, issue 3, pp. 388-391. DOI: 10.14739/2310-1210. 2018.3.130829
- Senkoro, Mbazi, et. al. *Smear microscopy and culture conversion rates among smear positive pulmonary tuberculosis patients by HIV status in Dar es Salaam, Tanzania*. BMC Infectious Diseases, Vol. 10, Issue 210, 2010. <http://www.biomedcentral.com/1471-2334/10/210>
- Tiemersma EW, van der Werf MJ, Borgdorff MW, Williams BG, Nagelkerke NJD (2011) *Natural History of Tuberculosis: Duration and Fatality of Untreated Pulmonary Tuberculosis in HIV Negative Patients A Systematic Review*. PLoS ONE 6(4): e17601. doi:10.1371/journal.pone.0017601
- Unknown. "OBITUARY: Captain Matthew Henry Phineas Riall Sankey 1853-1925". *Minutes of the Proceedings*, Vol. 221, Issue 1, Jan. 1926, pp. 271-271. E-ISSN 1753-7843

- Vite, Aleidy Marlene Silva. (2014) *Feedback System Control: optimizing drug combinations for tuberculosis treatment*. (Doctoral Dissertation). Retrieved from ProQuest. UMI Number: 3640128
- Wallis, Robert S. et. al. *Tuberculosis-advances in development of new drugs, treatment regimens, host-directed therapies, and biomarkers*. Lancet Infect Dis Vol 16 April 2016, [https://doi.org/10.1016/S1473-3099\(16\)00070-0](https://doi.org/10.1016/S1473-3099(16)00070-0)
- Wallis, Robert S. et. al. *Tuberculosis biomarkers discovery: developments, needs, and challenges*. Lancet Infect Dis, Vol. 13, pp. 362-72, 2013. [http://dx.doi.org/10.1016/S1473-3099\(13\)70034-3](http://dx.doi.org/10.1016/S1473-3099(13)70034-3)
- World Health Organization (WHO). Global Tuberculosis Report 2017. Geneva, Switzerland: World Health Organization (WHO); 2017. Retrieved 10/09/2018
- World Health Organization & Stop TB Initiative (World Health Organization). *Treatment of tuberculosis: guidelines*. World Health Organization (2010).
- Zhang, Liqun et. al. *Risk Factors for Pulmonary Cavitation in Tuberculosis Patients from China*. Emerging Microbes & Infections (2016) 5, e110; doi:10.1038/emi.2016.111