

Using Similarity Metrics on Real World Data and Patient Treatment Pathways to Recommend the Next Treatment

Kyle Haas, B.S.¹, Stuart Morton, Ph.D.², Simone Gupta, Ph.D.², Malika Mahoui, Ph.D.²
¹Indiana University-Purdue University Indianapolis (IUPUI), Indianapolis, IN, USA; ²Eli Lilly and Company, Indianapolis, IN, USA

Abstract

Non-small-cell lung cancer (NSCLC) is one of the most prevalent types of lung cancer and continues to have an ominous five year survival rate. Considerable work has been accomplished in analyzing the viability of the treatments offered to NSCLC patients; however, while many of these treatments have performed better over populations of diagnosed NSCLC patients, a specific treatment may not be the most effective therapy for a given patient. Coupling both patient similarity metrics using the Gower similarity metric and prior treatment knowledge, we were able to demonstrate how patient analytics can complement clinical efforts in recommending the next best treatment. Our retrospective and exploratory results indicate that a majority of patients are not recommended the best surviving therapy once they require a new therapy. This investigation lays the groundwork for treatment recommendation using analytics, but more investigation is required to analyze patient outcomes beyond survival.

Introduction

Lung cancer is one of the most prevalent cancers worldwide for both men and women¹, and accounts for nearly 15% of new cancer diagnoses². Of the various types of lung cancer, non-small cell lung cancer (NSCLC) represents over 80% of all documented lung cancer occurrences. Reductions in smoking have aided in the decline of death rates, but advanced stage diagnoses (IIIB and IV) continue to have an ominous 5-year survival rate² (26% and 1% respectively). While improvements in survival will continue with the development of new treatments and therapies, large-scale patient analytics may provide the means for reinforcing or improving traditional recommendation methods with the currently available therapies.

Traditionally, patients are administered a new line of therapy after either receiving a sufficient number of cycles or an unresponsive outcome to the prior treatment³. Clinicians are faced with the dilemma of selecting this new line. Considerable work has been accomplished in analyzing the viability of the treatments offered to NSCLC patients³; however, while many of these treatments have performed better over populations of diagnosed NSCLC patients, a specific treatment may not be the most effective therapy for a given patient. More recently, and with the growth of patient data digitization, the potential for performing large-scale patient analytics becomes more accessible. Patients are grouped together and viewed as collective target subgroups, where each patient shares the same genetic, demographic, clinical, and treatment profile with patients within the same group. This patient-centric, analytics approach allows for more patient-centric care.

Patient similarity metrics are not a novel concept. They have been used to help quantify the relationship between patients, which provide practical applications and insight for a given patient based on the known outcomes of patients with comparable profiles. Applications utilizing similarity metrics for better patient health and well-being include the following: medical diagnosis⁴, mortality predictions⁵, treatment recommendations^{6,7} and more⁸. These studies however, do not leverage either the patients' prior treatments or the ordering of these treatments. Prior line knowledge is integral in selecting the next therapy⁹, and has shown its value in studies beyond pure analytics applications. Using sequential pattern mining techniques to represent treatment features, Malhotra et al. were able to improve survival prediction models¹⁰. Wright et al. also used sequential pattern mining techniques on patient treatments to develop supervised machine-learning models to predict the next prescribed patient therapy¹¹. These studies further prove that prior treatment knowledge should be considered when recommending the next treatment.

Applying patient analytics in the treatment recommendation domain could potentially provide:

- An analytics perspective and reference on therapy outcomes to complement traditional recommendation methods
- A method for targeting clinical trial participants who may not be receptive to currently available therapies

Perer et al has led the forefront on visualizing treatment outcomes for similar patients with the CareFlow application¹². This data-driven, visual analytics tool recommends an entire care plan to a specific patient based from the outcomes

of similar patients. Using entire therapy lines, initial results for a group of similar congestive heart failure patients pointed to patients having a better outcome by following a care plan where the first line was different from the clinically recommended initial first line. While the approach is similar to our study, this work does not provide a detailed investigation over the impact of each line. We seek to recommend a single line of therapy based on the increased survival time each patient experiences using our proposed *After Stem Patient Survival* method (see Methods equation (1)). To further our investigation, we varied parameters (similarity threshold and which line number is recommended) in the models to analyze their impact as well. Also, another key distinction is that we used genetic biomarker results in our study. Biomarkers are one of the most influential indicators for clinicians in recommending treatments and the CareFlow application provides no clear indication of using any biomarkers in their study.

Other techniques including machine learning have been investigated in an effort to improve patient survival outcomes. In particular, Glioblastoma¹⁰ and NSCLC¹³ are two disease conditions where models were trained to predict overall patient survival using a myriad of patient features including therapy knowledge. Our data-driven approach using similarity metrics does not try to predict survival but rather investigates if better treatment recommendations exist for NSCLC patients using known outcomes from real world data.

This paper proposes a novel method that incorporates not only features that are used in traditional patient profiling (genetic, demographic, clinical, etc...), but we also account for prior treatment history in selecting the next best treatment.

Methods

Data Source

This study was conducted using the FlatIron® Advanced NSCLC proprietary dataset¹⁴, which aggregates genetic, clinical, and demographic information regarding each de-identified patient. All data available in FlatIron® dataset is in a structured and normalized format, including information extracted from medical notes by FlatIron® medical scientists. This dataset includes patients in stage IIIB or stage IV who are administered at least a single therapy on or after December 2010 and for which there is at least two documented EMR (Electronic Medical Record) visits during the same period. The FlatIron® advanced NSCLC dataset includes a geographically diverse sample of patients with a majority of patients treated at Community Oncology Hospitals and additional patients from Academic Centers in the United States. This longitudinal study contains data from December 2010 through May 2018. The dataset includes the following attributes: encompassing demographic data, clinical diagnosis, laboratory data, biomarker tests and results, medication, line of therapy (LOT), at the drug and class level, month and year of death, and patients' clinical characteristics (i.e. stage of diagnosis, tumor histology, smoking status, and ECOG performance status). Note that this study does not qualify as human subjects research in accordance with the U.S. Code of Federal Regulations (CFR), 45 CFR 46.102(f), and is thereby exempt from Institutional Review Board evaluation.

Certain patients and lines of therapy were removed from consideration. Any therapy line begun three months before the official cutoff date (any line administered after March 1, 2018) was removed from the patient profile. This filtering ensures new therapies had a consistent amount of time to take effect. We also removed any patient who had the same dates for the start and end of their therapy. This uncommon occurrence results from only a single administration of the therapy. Our main outcome of this study is to analyze survival relating to each line and these types of administrations do not provide any survival duration, but are still considered a therapy administration in the FlatIron® database. For this reason, we removed any patient with this rare behavior. Following these exclusion criteria, 30,974 unique patients remained.

Feature Space

All features listed in the dataset are partitioned into three categories: Genomic, Demographic/Clinical, and Treatment. The genomic features include results from well-regarded biomarker tests: ALK, EGFR, KRAS, ROS1, and PD-L1¹⁵. The outcomes from these biomarker tests influence on the treatment decision process for clinicians, and these genetic features also play a critical role in our recommendation procedure. The demographic/clinical features include race, age, gender, presence of maintenance therapy, initial performance score, and site of metastasis. These features and their corresponding statistics along with the genetic biomarker tests can be found in Table 1. Altogether, there were 23 features used in creating the patient profiles used for similarity (see non-treatment-based similarity). Although treatment information was used in this study, it was not included in the patient feature profile. These treatments generate the treatment similarity (see Treatment-Based Similarity). This study was performed at a class level analysis,

and each treatment was assigned a class using a hierarchy defined by FlatIron®. The line breakdown and the percentage of class representation across the dataset is shown in Table 2 and Figure 1 respectively.

Table 1. Feature statistics for the genomic biomarkers and the demographic/clinical data

Feature		Feature		Feature	
<i>Gender</i>		<i>Biomarker Results</i>		<i>Site of Metastasis</i>	
-Female	14547 (47.0%)	-ALK Positive	691 (2.2%)	-Liver Positive	1642 (5.3%)
-Male	16426 (53.0%)	-ALK Negative	16882 (54.5%)	-Liver Negative	29332 (94.7%)
-Unknown	1 (~0%)	-Unknown	13401 (43.3%)		
<i>Race</i>		-EGFR Positive	3222 (10.4%)	-Lung Positive	1323 (4.3%)
-White	21673 (70.0%)	-EGFR Negative	16203 (52.3%)	-Lung Negative	29651 (95.7%)
-Asian	799 (2.5%)	-Unknown	11549 (37.3%)		
-Black/African American	2566 (8.3%)	-KRAS Positive	2522 (8.1%)	-Brain Positive	4287 (13.8%)
-Other	2529 (8.2%)	-KRAS Negative	6257 (20.2%)	-Brain Negative	26687 (86.2%)
-Hispanic/Latino	47 (0.2%)	-Unknown	22195 (71.7%)		
-Unknown	3360 (10.8%)			-Adrenal Positive	811 (2.6%)
<i>Time between Diagnosis and Advanced Diagnosis (mean)</i>	5.6 ± 17.7 months	-PDL1 Positive	1421 (4.6%)	-Adrenal Negative	30163 (97.4%)
<i>First Known ECOG</i>		-PDL1 Negative	2778 (9.0%)		
-0	7423 (24.0%)	-Unknown	26775 (86.4%)	-Lymph Node Positive	1611 (5.4%)
-1	8428 (27.2%)			-Lymph Node Negative	29313 (94.6%)
-2	3028 (9.8%)	-ROS1 Positive	123 (0.4%)		
-3	676 (2.2%)	-ROS1 Negative	9360 (30.2%)	-Bone Positive	8001 (25.8%)
-4	31 (0.1%)	-Unknown	21491 (69.4%)	-Bone Negative	22973 (74.2%)
-Unknown	11388 (36.8%)			-CNS Positive	264 (0.9%)
<i>Region</i>		<i>Stage at Diagnosis</i>		-CNS Negative	30710 (99.1%)
-Northeast	6286 (20.3%)	-Occult	1 (~0%)		
-West	4502 (14.5%)	-Stage 0	1 (~0%)	-Other Positive	2539 (8.2%)
-Midwest	5547 (17.9%)	-Stage I	417 (1.3%)	-Other Negative	28435 (91.8%)
-South	12233 (39.5%)	-Stage IA	972 (3.1%)		
-Unknown	2406 (7.8%)	-Stage IB	864 (2.8%)	<i>Histology</i>	
<i>1st Line Maintenance</i>		-Stage II	150 (0.5%)	-Squamous	7618 (24.6%)
-0	27929 (90.3%)	-Stage IIA	672 (2.2%)	-Non-squamous	21780 (70.3%)
-1	2995 (9.7%)	-Stage IIB	577 (1.9%)	-Not Otherwise Specified	1576 (5.1%)
		-Stage III	344 (1.1%)		
		-Stage IIIA	2152 (6.9%)	<i>Smoking History</i>	
		-Stage IIIB	3863 (12.5%)	-Smoker	26476 (85.5%)
		-Stage IV	20029 (64.7%)	-Non-Smoker	4092 (13.2%)
		-Unknown	932 (3.0%)	-Unknown	406 (1.3%)
				<i>Age at First Line (mean)</i>	66.9 ± 9.8 yrs

Table 2. Number of patients having at least the number of lines in their TS

Line	Number of patients
1	30974
2	13579
3	5345
4	2111
5	787
6	290
7	132
8	52
9	19
10	5
11	2
12	2
13	2

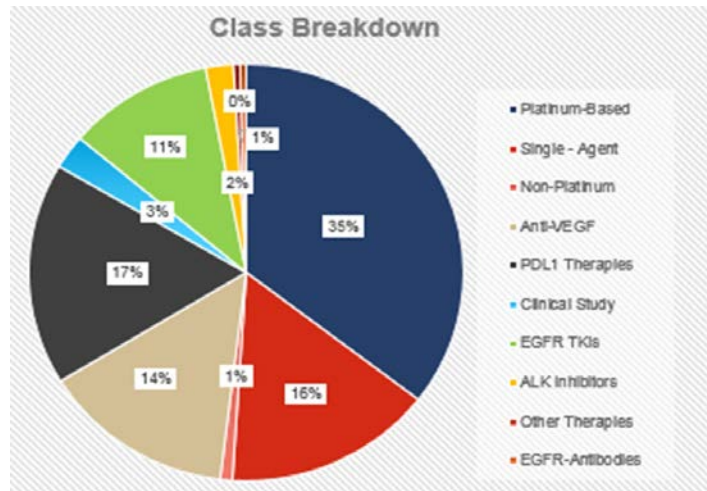


Figure 1. Class percentage breakdown

Definitions

Let D represent the set of drugs received by patients. A treatment, also referred to as a line of therapy (LOT), is an administration received by a patient over a specified time period. A treatment sequence (TS) is a series of LOTs received by a patient in a sequential manner to indicate the order in which the treatments were administered. For

example, a patient P that took a series of three therapies (A, B, and C) in D , patient P 's TS is defined and represented as the following: " $A \rightarrow B \rightarrow C$ " where A is the first LOT, B the second, and C the third. In this representation, all LOT's are separated by arrows.

Primary Outcomes

Given a group of similar patients Sim_p to a patient P , recommend a treatment corresponding to the subgroup of patients having the "best" outcome following the same treatment stem sequence (prior treatments). More formally, the stem is the entire TS before the recommended LOT. For patient P , if we want to recommend a fourth LOT, the TS from the first line A to third line C would be considered the stem.

This study uses patient line survival as the primary outcome. *After Stem Patient Survival (ASPS)* and its corresponding calculation, define the "best" outcome in our study. The following paragraphs provide more details on the *ASPS* definitions and how the method determines the similar patient subgroups and treatment stems.

ASPS: Is defined as the percent time duration from the start of the next LOT after the treatment stem sequence until either death or last known treatment end date (computed in days) divided by the overall survival (also computed in days). More formally, given P and the set of similar patients Sim_p , let " $t_1 \rightarrow t_2 \rightarrow \dots \rightarrow t_k$ " be the TS of patient P used as the stem treatment sequence to find patients with similar previous treatment as P .

In the example of Figure 3, this treatment stem sequence corresponds to $(A \rightarrow B)$. If P_i is a patient in Sim_p , then we define the treatment sequence of P_i as " $t_{i1} \rightarrow t_{i2} \rightarrow \dots \rightarrow t_{ij} \rightarrow \dots \rightarrow t_{ik} \rightarrow t_{ik+1} \rightarrow \dots \rightarrow t_{in}$ ". In this case, the *ASPS* of patient P_i in Sim_p , noted $ASPS(P_i)$ is defined by the following equation, where duration (t_{ij}) is the time duration of the LOT t_{ij} .

$$ASPS(P_i) = \frac{\sum_{j=k+1}^n Duration(t_{ij})}{\sum_{j=1}^n Duration(t_{ij})} \quad (1)$$

In the example of Figure 3, $ASPS(P_i)$ is the ratio of the duration of the sum of the treatment durations of the LOTs D and E (100 days) and the overall survival of patient P_i (250).

Methodology Pipeline

Given a "reference" patient P described by a set of demographic, genetic, clinical characteristics, and a set of already prescribed treatments, our approach is to recommend the next treatment for patient P by finding the treatment sequence administered to similar patients that led to the "best" outcome. In this study, the best outcome is determined using line therapy survival. Figure 2 depicts the main steps involved in our approach.

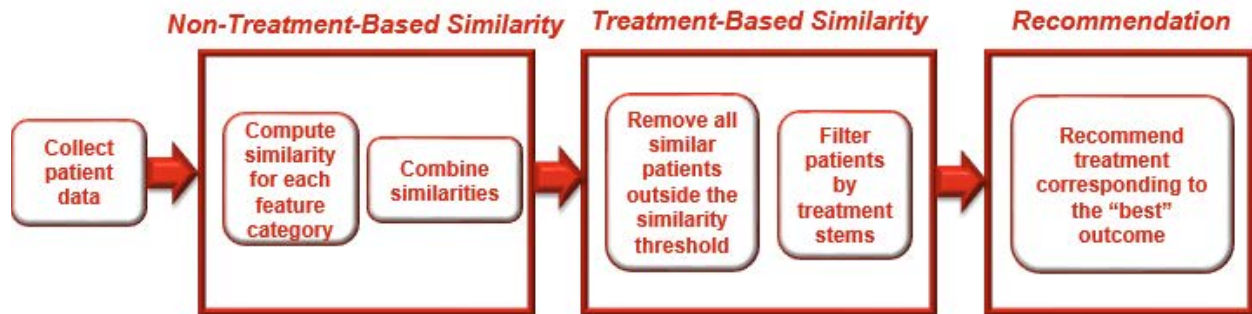


Figure 2. Proposed methodology pipeline

Non-Treatment-Based Similarity

We divide the set of patient features into two categories: genetic and clinical/demographic. We first compute the similarity between patients for each category and then combine the two similarity measures using a geometric mean calculation. A similar approach for combining multiple patient similarity metrics was used in a study by Gottlieb et al⁴. We compute the similarity between patients within each category using the Gower Similarity Coefficient¹⁶ given its design to incorporate both continuous and binary features. The resulting similarity matrix is then used to determine the set of similar patients above a pre-defined similarity threshold w.r.t a single reference patient. When calculating the similarity between patients, only features where both patients have valid entries are considered. If one of the patients has a null feature, this feature is ignored for the purposes of non-treatment-based similarity.

Treatment-Based Similarity

Given a reference patient P , we first compute a set of similar patients Sim_p (determined from the genetic and clinical/demographic features). The next step is to filter from Sim_p the patients that do not share the same stem (prior treatments) as the reference patient. The objective is to consider only patients who share similar previous treatments in order to determine the next treatment for the reference patient. An example illustrating this filtering is depicted in Figure 3. Treatments are represented as letters, while arrows represent the sequencing of treatments. Patients 4 and 6 are filtered from further consideration as they do not share the same treatment stem in their sequence as the reference patient.

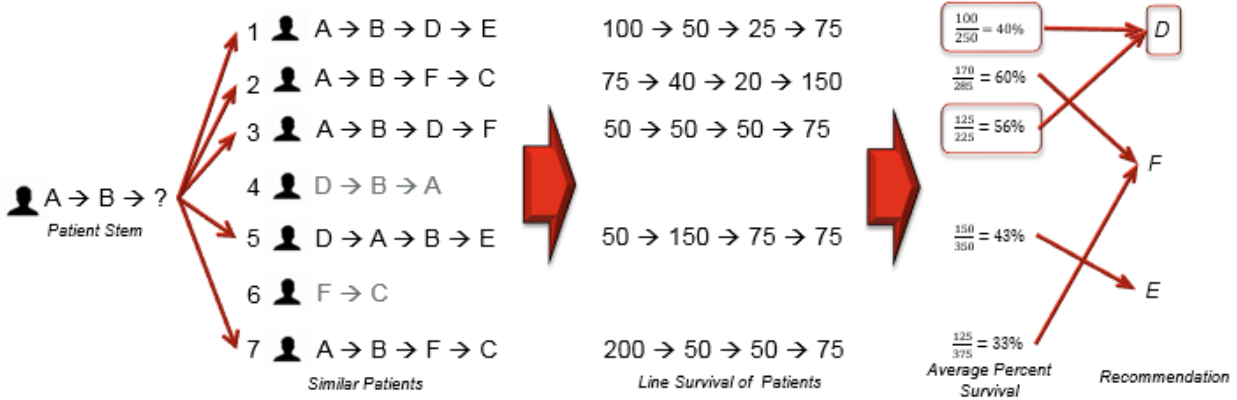


Figure 3. Example illustration of the algorithm recommending treatment D using line survival as the sole outcome

Next Treatment Recommendation

After the two-step filtering process, only patients who share a common treatment stem and have a similarity above a required threshold with the reference patient remain. From this subset of patients, we determine the percentage of each patient's survival following the treatment stem. Once this value is computed, we average the percentage across all patients with the same next treatment following the stem. This process is detailed in Figure 3 and shows that the percentage of survival is only computed for the subgroup (Patients 1,2,3,5, and 7) that share the same treatment stem (A → B). All percentage of patients sharing the same next treatment after the stem are then averaged together (48% for D, 46.5% for F, and 43% for E), which results in the algorithm selecting treatment D as the best next treatment for the patient.

To formally describe how we recommend the next treatment, consider the following definitions. Let $P_{t1}, P_{t2}, \dots, P_{tm}$ be m subsets of patients from Sim_p such that, patients in subset P_{ti} share the same next LOT ti following the stem treatment sequence. In the example of Figure 3, patients P_1 and P_3 belong to the same subset sharing the same next line of therapy D.

The equation below details the method for averaging the ASPS for each m subset in Sim_p :

$$\text{Max}_{i=1 \dots m} \left(\frac{\sum_{j=1}^{|P_{ti}|} ASPS(P_j)}{|P_{ti}|} \right) \quad (2)$$

The next treatment belonging to the subset of patients containing the highest average ASPS is selected as the “best” next treatment.

Evaluation

To evaluate the performance of our approach, we propose to consider a portion of patients in the dataset as a reference patient. Given that the patient survival time is known, we can compute the number of times that the recommended treatment would have led to a longer survival than the reference patient's original survival time. In our proposed implementation, we divide the dataset into ten random subsets (with replacement following the entire selection of a subset), compute the performance of the approach for each sample, and then average the performance across all samples. We also analyze the impact of previous (i.e. stem) treatment knowledge on the performance of the method and the selection of the similarity threshold.

Results

Altogether, there were 30,974 patients considered in this study. First, we separated ten random 300 patient subsets with replacement (a patient may not appear more than once in a same subset but may or may not be represented in another subset). As a baseline, we evaluated each subset of patients using the methodology proposed above, but without any prior treatment knowledge (0-stem). Each patient in this subset was evaluated against the other 30,674 patients. After each subset evaluation was complete, the results were aggregated together. We repeated this process with the 0-stem length five different times considering different minimum similarity thresholds each evaluation (0.65, 0.7, 0.75, 0.8, 0.85). The results are displayed below.

Table 3. Results from evaluating the treatment recommendation algorithm with five different thresholds for zero prior treatments (0-Stem). The similarity threshold is listed along with the percentage representing the likelihood where a reference patient in the random subset was recommended a better different therapy than the one the patient took. The most and least frequent recommendations are also listed.

Similarity Threshold	Percent of Patients Recommended a “better” treatment	Most Common Recommended Treatment Class	Least Recommended Treatment Class
0.85	69.8%	Clinical Study Drug (22.3%)	PDL1 Therapies (1.7%)
0.8	69.3%	Clinical Study Drug (23.3%)	PDL1 Therapies (0.9%)
0.75	72.7%	Other Therapies (27.0%)	PDL1 Therapies (0.3%)
0.7	72.7%	Other Therapies (40.6%)	PDL1 Therapies (0.2%)
0.65	72.8%	Other Therapies (44.3%)	PDL1 Therapies (0.1%)

After the 0-stem baseline, we then repeated the same process with all five similarity thresholds for both 1-stem and 2-stem subsets. 1-Stem ensures all treatment recommendations are from therapies following whichever therapy was the first line of the reference patient. 2-Stem follows a similar format but with two therapies. Figure 3 above illustrates the evaluation for a 2-stem process. The results for both 1-stem and 2-stem are found below in Table 4 and Table 5.

Table 4. Results from evaluating the treatment recommendation algorithm with five different thresholds with a single line prior treatment knowledge (1-Stem). The similarity threshold is listed along with the percentage representing the likelihood where a reference patient in the random subset was recommended a better different therapy than the one the patient took. The most and least frequent recommendations are also listed.

Similarity Threshold	Percent of Patients Recommended a “better” treatment	Most Common Recommended Treatment Class	Least Recommended Treatment Class
0.85	66.5%	Platinum-Based Chemo (17.8%)	Other Therapies (2.6%)
0.8	68.1%	Clinical Study Drug (19.5%)	Other Therapies (4.5%)
0.75	72.1%	Clinical Study Drug (22.3%)	Single Agent Chemo (2.0%)
0.7	71.6%	Clinical Study Drug (26.3%)	Single Agent Chemo (1.0%)
0.65	71.2%	Clinical Study Drug (22.3%)	Single Agent Chemo (0.3%)

Table 5. Results from evaluating the treatment recommendation algorithm with five different thresholds with a two line prior treatment knowledge (2-Stem). The similarity threshold is listed along with the percentage representing the likelihood where a reference patient in the random subset was recommended a better different therapy than the one the patient took. The most and least frequent recommendations are also listed.

Similarity Threshold	Percent of Patients Recommended a “better” treatment	Most Common Recommended Treatment Class	Least Recommended Treatment Class
0.85	62.9%	PDL1 Therapies (21.6%)	Other Therapies (0.9%)
0.8	64.1%	PDL1 Therapies (19.9%)	Other Therapies (0.6%)
0.75	68.5%	Anti-VEGF Therapies (18.6%)	Other Therapies (1.9%)
0.7	74.9%	Clinical Study Drug (19.6%)	Other Therapies (3.9%)
0.65	77.4%	Clinical Study Drug (20.3%)	Other Therapies (4.2%)

Discussion

Discussion of Results

These results indicate a majority of patients are not taking the next best treatment once they need to switch to a different therapy. For all three stem lengths and five similarity thresholds, more than 50% of patients were recommended a better therapy for survival. Moreover, the length of stem and similarity threshold impacted the likelihood in which a patient was recommended a better therapy. In particular, for all the three stems, we found that as the similarity threshold increased (the subgroups of patients providing recommendations collectively are more similar to the reference patient), the likelihood a patient was recommended a better therapy decreased. Only two thresholds at the 1-stem length (0.7 and 0.75) produced an increase in the likelihood of recommending a better therapy. Additionally, larger stem lengths (more prior knowledge), decrease this likelihood. The likelihood of recommending a better therapy decreased for all stem lengths and similarity thresholds except for the 2-stem length at the 0.7 and 0.65 thresholds.

These results are intuitive. We would expect to determine a better next treatment after already considering the outcomes of the prior therapies. In addition, as the similarity threshold decreases, the considerations for what constitutes as a similar patient is reduced, which in turn may produce recommendations that may not be the most appropriate for a given patient. However, there is a tradeoff with high similarity thresholds. Higher thresholds will produce more refined groups of patients to consider, but more refinement means less patients are considered in the analysis. In order to compensate for the lower amounts of patients, larger datasets should be used in determining the “best” treatment recommendation.

The other takeaway from these investigative results is recommendations from the drug class of Clinical Study and Other Therapies reduced as the similarity threshold increased. This also indicates the grouping of patients are becoming more refined and following more consistent standards-of-care, as most of the therapies considered in these two classes are either unique, new, or not yet approved therapies. Moreover, Clinical Study was the most recommended therapy class throughout this investigation (eight of the fifteen studies recommended this class). The identity of the treatments in this class are unknown which hinder any further investigation that may provide insight into which clinical study drugs are providing the largest impact. However, this encouraging result supports many of the promising clinical study results, and offers hope for many of the potential treatments for NSCLC patients in the future.

Limitations

Certain limitations do exist with this study and must be considered. First, many patients did not have complete feature profiles. In particular, 10,190 of the patients used in the study had no record of any biomarker results, and out of the remaining patients, the average patient only had between two and three results of the five biomarkers (2.86 average). For the patients with no biomarker information, this proved challenging as there is no genetic similarity that can be determined with these patients. In these cases, the similarity was determined using only the demographic and clinical

features. Along with the lack of genetic features, 1,927 patients contained gaps in the lines of therapy. For most patients, once the end of the line occurs, the next line begins. For the few patients with gaps, the beginning of the next line may not occur until months after the end of the last line. In this study we accounted for these gaps in the overall survival computation of the *ASPS*, however it is not the same as a typical treatment sequence of many the other patients. Additionally, while certain demographic features such as race, gender, and region were included in the similarity computation, no formal analysis of potential cofounders was investigated.

Second, there is no comprehensive knowledge of why certain patients were administered given therapies. In our model, we use survival time as the sole outcome, whereas with some patients, quality-of-life or even cost of treatment is more important than the potential to live longer. Future applications of this work should investigate collective outcomes. In addition, this is a retrospective analysis. Many patients at the beginning of the study may not have had access to newly approved therapies or promising clinical study treatments. This may be why the model appears to be very inclined to recommend clinical study drugs for all the tested stem lengths.

Lastly, there is no effective way to represent living patients. If the living patients are removed, the data becomes biased to only represent deceased patients. On the other hand, if you keep the living patients, their total survival becomes underestimated, as they are yet to complete their entire treatment pathway. In the case of our study, we included the living patients and represented their ending survival as the last day of the data cutoff. Altogether, there were 9,765 living patients in the study. Any future researcher interested in applying similar work must consider the impact of living patients.

Future Work

Future work will center upon developing a composite outcome that accounts for more than patient survival (quality of life, cost, etc...). This will aid in representing what the clinician believes is the best next treatment for a given patient. In addition, as more genetic biomarkers are discovered and reported, the model will be adjusted in order to account for any new discoveries.

Conclusion

Prior treatment knowledge plays a pivotal role in the selection of the next treatment for a patient. Using prior treatment knowledge along with patient similarity metrics, we proposed a methodology for recommending the next treatment. The “best” outcome for a given patient was determined by selecting the best percentage survival following the same administered treatment from a subgroup of similar patient. Our initial investigative evaluations indicated that while many patients are not taking the best next treatment for survival, the likelihood that a patient took the best next treatment increases as the subgroups are refined closer to the given patient or the target patient has a longer prior treatment history. Further investigation is necessary in order to validate these findings.

References

1. Key statistics for lung cancer. [Internet]. American Cancer Society, 2018 [cited 2018 Jul 2]. Available from: <https://www.cancer.org/cancer/non-small-cell-lung-cancer/about/key-statistics.html>
2. Lung cancer – non-small cell: statistics. [Internet]. American Society of Clinical Oncology, 2018 [cited 2018 Jul 2]. Available from: <https://www.cancer.net/cancer-types/lung-cancer-non-small-cell/statistics>
3. Azzoli CG, Baker S Jr, Temin S, Pao W, Aliff T, Brahmer J, Johnson D, Laskin J, Masters G, Milton D, Nordquist L, Pfister DG, Piantadosi S, Schiller JH, Smith R, Smith TJ, Strawn JR, Trent D, Giaccone G. American society of clinical oncology clinical practice guideline update on chemotherapy for stage IV non-small-cell lung cancer. *J Clin Oncol* 2009;27:6251-6266.
4. Gottlieb A, Stein GY, Ruppin E, Altman RB, Sharan R. A method for inferring medical diagnoses from patient similarities. *BMC Medicine* 2013;11:194
5. Lee J, Maslove DM, Dubin JA. Personalized mortality prediction driven by electronic medical data and a patient similarity metric. *PLoS One*: 2015;10:5
6. Zhang P, Wang F, Hu J, Sorrentino R. Towards personalized medicine: leveraging patient similarity and drug similarity analytics. *AMIA Jt Summits Transl Sci Proc* 2014;2014:132-136.
7. Panahiazar M, Taslimitehrani V, Pereira NL, Pathak J. Using EHRs for heart failure therapy recommendation using multidimensional patient similarity analytics. *Digital Healthcare Empowering Europeans* 2015;210:369-373.

8. Sharafoddini A, Dubin JA, Lee J. Patient Similarity in prediction models based on health data: a scoping review. *JMIR Med Inform* 2017;5:1.
9. Maione P, Rossi A, Bareschino MA, Sacco PC, Schettino C, Falanga M, Barbato V, Ambrosio R, Gridelli C. factors driving the choice of the best second-line treatment of advanced NSCLC. *Reviews on Recent Clinical Trials* 2011;6:44-51.
10. Malhotra K, Navathe SB, Chau DH, Hadjipanayis C, Sun J. Constraint based temporal event sequence mining for glioblastoma survival prediction. *J Biomed Inform* 2016;61:267-275.
11. Wright AP, Wright AT, McCoy AB, Sittig DF. The use of sequential pattern mining to predict next prescribed medications. *J Biomed Inform* 2015;53:73-80.
12. Perer A, Gotz D. Data-driven exploration of care plans for patients. *Extended Abstracts on Human Factors in Computing Systems* 2013;439-444.
13. Haas K, Mahoui M, Gupta S, Morton S. Leveraging treatment patterns to predict survival of patients with advanced non-small-cell lung cancer. *Proceedings of the 18th ACM international conference on bioinformatics, computational biology, and health informatics*. 2018:283-290
14. Flatiron Health, Inc. <http://www.flatiron.com/life-sciences>, 2018.
15. Bernicker EH, Miller RA, Cagle PT. Biomarkers for selection of therapy for adenocarcinoma of the lung. *J Oncol Prac* 2017;13:221-227.
16. Gower JC. A general coefficient of similarity and some of its properties. *Biometrics* 1971;27:857-874.