Taylor & Francis
Taylor & Francis Group

Future Oncology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ifon20

Defining Treatment Regimens and Lines of Therapy Using Real-World Data in Oncology

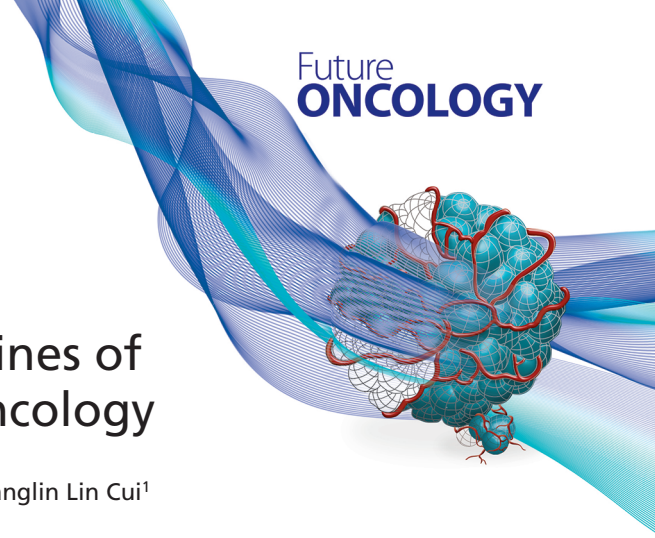
Lisa M Hess, Xiaohong Li, Yixun Wu, Robert J Goodloe & Zhanglin Lin Cui

To cite this article: Lisa M Hess, Xiaohong Li, Yixun Wu, Robert J Goodloe & Zhanglin Lin Cui (2021) Defining Treatment Regimens and Lines of Therapy Using Real-World Data in Oncology, Future Oncology, 17:15, 1865-1877, DOI: 10.2217/fon-2020-1041

To link to this article: https://doi.org/10.2217/fon-2020-1041

Methodology

For reprint orders, please contact: reprints@futuremedicine.com



Defining treatment regimens and lines of therapy using real-world data in oncology

Lisa M Hess*[,1], Xiaohong Li[1], Yixun Wu[1], Robert J Goodloe[1] & Zhanglin Lin Cui[1]

[1]Eli Lilly & Company, Indianapolis, IN 46285, USA

*Author for correspondence: hess_lisa_m@lilly.com

Retrospective observational research relies on databases that do not routinely record lines of therapy or reasons for treatment change. Standardized approaches to estimate lines of therapy were developed and evaluated in this study. A number of rules were developed, assumptions varied and macros developed to apply to large datasets. Results were investigated in an iterative process to refine line of therapy algorithms in three different cancers (lung, colorectal and gastric). Three primary factors were evaluated and included in the estimation of lines of therapy in oncology: defining a treatment regimen, addition/removal of drugs and gap periods. Algorithms and associated Statistical Analysis Software (SAS®) macros for line of therapy identification are provided to facilitate and standardize the use of real-world databases for oncology research.

Lay abstract: Most, if not all, real-world healthcare databases do not contain data explaining treatment changes, requiring that rules be applied to estimate when treatment changes may reflect advancement of underlying disease. This study investigated three tumor types (lung, colorectal and gastric cancer) to develop and provide rules that researchers can apply to real-world databases. The resulting algorithms and associated SAS® macros from this work are provided for use in the Supplementary data.

First draft submitted: 13 October 2020; Accepted for publication: 22 January 2021; Published online: 24 February 2021

Keywords: administrative claims • colorectal cancer • electronic medical records • gastric cancer • line of therapy • lung cancer • retrospective research • SAS macro

Cancer is a global public health problem leading to an estimated over 17.2 million new cases diagnosed and 8.9 million deaths each year [1]. Among patients whose disease has become metastatic, cure is rarely possible. For these patients, care becomes chronic, with the primary goals to manage disease growth, maintain patient quality of life and to extend life for as long as possible. Treatment guidelines, such as those produced by the European Society of Medical Oncology, the National Comprehensive Cancer Network and American Society of Clinical Oncology, summarize evidence-based systemic therapy that are recommended based on patient-specific disease characteristics, progression and response to prior therapies. The sequence or 'line of therapy' is the cornerstone of planning systemic therapy over the course of disease in clinical practice.

The availability of electronic data for billing and reimbursement (e.g., administrative claims data) and for clinical care (e.g., electronic medical record [EMRs]) has led to the availability of large population-based databases for research purposes. These databases typically contain drug names and dates of billing or administration, but rarely characterize the treatment plan or intent, and do not generally specify the treatment regimen used or the line of therapy in which the patient is being treated. Therefore, researchers must evaluate the data to make estimations of the clinical intent of treatments when conducting research. Rules must be derived to allow for programming that will determine the regimen and line of therapy. The rules used by researchers are rarely transparent and are likely inconsistent across studies even within the same disease, which may lead to different conclusions and potentially incorrect findings.

In general, retrospective database studies have operationalized these concepts by defining the first line of therapy as the initial set of drugs used after advanced or metastatic diagnosis during a specified time period and until a specified gap period or addition of different systemic therapies [2–5]. This first step, however, is dependent on the correct

identification of an advanced/metastatic cohort, so as to not combine adjuvant therapy with systemic therapy for metastatic disease. Subsequent lines of therapy in the post-progression setting are more difficult to study, as rules cannot rely simply on the first drugs administered, but must carefully evaluate the order and timing of drugs, in the context of the specific disease under study, to estimate when a new line of therapy may be initiated. As a result, some studies simply limit the cohort being investigated to those with prior exposure to specific drugs to identify a pretreated population, which limits the generalizability of the findings.

While rules can be applied across tumor types, the specific nature of treatment strategies in a given disease must be taken into consideration that may require specific considerations in the general rules, such as if there is a role for maintenance therapy, the prevalence of allowing drug holidays for patient convenience or quality of life, or the practice supplementing chemotherapy with specific biologic or targeted agents in the absence of advancing a line of therapy, as well as diseases that may be usually treated by continuing the same biologic or targeted agents across multiple lines of therapy [6–9]. Unfortunately, the details of the rules applied within individual studies are often lacking in published manuscripts [10,11] As a result, replication of methods is usually not possible, and therefore, limits the consistency among retrospective database studies, as the underlying rules for identifying sequences of treatments may be dissimilar. For cases in which detailed rules are provided, these studies have been very focused and valuable for a specific tumor, research question, and/or database, with limited to no generalizability to other settings [12–15]. The provision of detailed rules for specific tumors in these publications elucidates key consistencies that are to be accounted for in the determination of lines of therapy more broadly in oncology, such as the time period of drug use to determine a regimen, the addition or discontinuation of drugs to a regimen and treatment interruptions (gap periods) where the patient does not receive therapy. Specifically, tumor agnostic rules developed only from melanoma and non-small-cell lung cancer (NSCLC) in one study; such agnostic rules have yet to be investigated across a broader set of solid malignancies [15].

Therefore, this work was initiated to provide standardization and transparency in the application of rules to regimens and lines of therapy in retrospective observational database research in oncology. The primary goal of this work was to provide researchers with SAS® macros to apply such rules consistently, while allowing for variations to accommodate the needs of specific research questions. This will allow for more transparent assumptions within retrospective database research that relies on defining regimens and lines of therapy. The macros were developed to enable them to be applied to administrative claims or EMR database research across several solid tumor types.

Materials & methods

Tumor types

The development of rules for SAS macros for every tumor type is not feasible within a single project. Therefore, three specific tumor types were selected for development due to their unique settings and ability to be adapted to other disease sites that may have comparable treatment patterns. NSCLC is a complex disease, with multiple targeted agents available for patients with actionable mutations, maintenance therapy is common and the treatment of this disease has changed considerably as novel targeted agents have become available for patient care [16,17]. Colorectal cancer (CRC) is a slower growing disease, with long treatment and gap periods [18], whereas gastric cancer remains an aggressive disease with limited treatment options [19,20]. Each of these diseases has been extensively studied by our research team, and rules for lines of therapy have been developed and applied across multiple projects [4,21–26]. Based on our learning and application of line of therapy rules to date as well as on other publications applying rules for lines of therapy, the SAS macros were developed and tested in the settings of NSCLC, CRC and gastric cancer.

Fundamental to the development of rules for lines of therapy is clinical understanding of the specific disease and its treatment. The rules for lines of therapy were developed through an assessment of the National Comprehensive Cancer Network Treatment Guidelines, as well as iterative review and revision by clinical experts who care for these patients.

Systemic therapy

For the determination of lines of therapy in the advanced/metastatic setting, all systemic anticancer drugs were included. Any surgical or radiation therapies were not included in the regimen or line of therapy rules. All systemic anticancer therapies were further categorized as either chemotherapy or as targeted/biologic drugs. Consideration of interchangeable drugs within a drug class was evaluated for each tumor type, where the changing of a drug within a class would not suggest disease progression (hence, advancement of the line of therapy), but rather may reflect convenience, cost or intolerance.

The role of drug interchangeability was evaluated. In gastric and CRC, 5-fluororuacil and capecitabine may be considered interchangeable and were categorized as 'fluoropyrimidines'. For all three cancers, cisplatin and carboplatin may also be considered interchangeable and in this work were categorized as 'platinum agents'. While other platinum drugs exist (e.g., oxaliplatin), this was not included in the categorization for the macro development due to the unique nature of regimens that include this agent (e.g., FOLFOX, fluoropyrimidine plus oxaliplatin). In addition to specific drug names, dates for each drug was used in the development of the rules and macros. Supportive care drugs are not considered in the rules for lines of therapy.

Regimen definition

For all tumor types, a regimen was defined as the single drug or set of drugs used within the initial 28-day period after treatment initiation. This was based on consistency in cycle length in the labels of all drugs across all tumor types in oncology, none of which exceed 28 days. In cases of potential regimen augmentation strategies (the practice of adding a biologic or targeted drug to an existing chemotherapy backbone), the regimen name was updated to include the drug that was added. Additionally, in cases of maintenance therapy (e.g., NSCLC), the agent used to maintain or consolidate response in the first-line setting was included in the regimen (the regimen name was not changed or specified for the maintenance period).

Starting a line of therapy

The first line of therapy began on the date of the first oral or infused drug within the initial 28-day period of systemic therapy. If a line of therapy was stopped as described below, the new line of therapy would similarly start upon the first date of the subsequent oral or infused drug, with the new line of therapy regimen defined again as the drug or the set of drugs within the initial 28-day period starting at the time of the initiation of the new drug or following the end of the previous line of therapy.

Stopping a line of therapy

The line of therapy discontinuation rules differed by tumor type due to the way in which each disease is managed. However, there are consistent factors across all tumor types that were evaluated, such as time to disease progression, adverse events/toxicity and drug holidays; rules were devised to indicate a stopping rule: a specified gap period of a no-treatment interval was reached; a new chemotherapy drug was added to the regimen; a new biologic/targeted agent was added after a prespecified period of time; there was a complete change in the set of drugs being used, with or without a gap period or discontinuation of all systemic anticancer drugs.

The line of therapy date of discontinuation was the last day on which a drug in the regimen was infused or on the final date of the oral drug supply (e.g., if a 30-day fill was made on a certain date, the final date would be +30 days from the last drug fill), whichever was later in the case of combination oral plus infused drug regimens.

Testing & revision

Stopping rules were evaluated for each criterion at different levels: gap periods of 60, 90 and 180 days; new biologic agents added after the initial 28-day period, after 90 and 180 days, or with no time limit for a line of therapy advancement; and new chemotherapy added after 60, 90 or 180 days were evaluated (Table 1). Specific considerations included maintenance therapy (such as in NSCLC), where the continuation of bevacizumab or pemetrexed was considered after a regimen containing that same agent after gap periods of 60, 90 or 180 days, or with no time limit (Table 2). At last, consideration was made for diseases such as CRC, which may supplement a regimen with new drugs, which may not all represent a change in line of therapy (Table 3). Rules were applied to various electronic medical record and claims data sets and the resulting regimens and sequences were visually evaluated to compare with treatment guidelines and were inspected by practicing oncologists to identify the time periods that resulted in sequences of care that most resembled clinical practice patterns.

Results

The rules identified to develop the line of therapy macros are summarized in Tables 1–3 and the final SAS macros are presented in the Supplementary data for analytic use. The macros include options as summarized in Tables 1–3, which an investigator may wish to vary depending on the research question. Therefore, the macros were developed to allow for investigator flexibility in modifying the time periods for each factor. The start and stop logic as presented

Table 1. Line of therapy rules in oncology.

Rule	Option A	Option B	Option C	Option D	Optimal rule and considerations
Line therapy changes when:					
Patient has had no treatment for a gap of greater than ____ and then restarts the same treatment	60 days	90 days	180 days	The line of therapy never changes when this happens	NSCLC – option A Gastric cancer – option B CRC – option C
Any biologic/targeted is added to a regimen greater than ____ of the start of the line of therapy	After initial 28-day period	90 days	180 days	The line of therapy never changes with adding a biologic	NSCLC – option D Gastric cancer – option D CRC – option C (except for panitumumab or cetuximab which follow option B)
Any biologic/targeted agent is started with discontinuation of all agents in the initial regimen	Upon initiation of biologic targeted agent				In general, the start of a new drug with discontinuation of all prior drugs indicates a new line of therapy
Adding one or more chemotherapy drugs in a regimen will constitute a change in line of therapy	After initial 28-day period				In general, a new chemotherapy agent (with or without discontinuing all other agents) that is not considered to be interchangeable indicates a new line of therapy
Line of therapy does not change when:					
Any biologic/targeted agent is added to a regimen less than or equal to ___ days of the start of therapy	28 days	90 days	180 days	The line of therapy never changes with adding a biologic	NSCLC – option D Gastric cancer – option D CRC – option C (except for panitumumab or cetuximab which follow option B)
Exchange of cisplatin or carboplatin (these agents are considered equivalent)	Never changes line of therapy				Any interchangeable agent should never change a line of therapy. The drugs considered interchangeable may vary by tumor type. These two platinums are provided as they apply to all three tumor types
Exchange of 5-fluorouracil or capecitabine (these agents are considered equivalent)	Never changes line of therapy				Any interchangeable agent should never change a line of therapy. The drugs considered interchangeable may vary by tumor type. Fluoropyrimidines are provided as they are used in both CRC and gastric cancer

CRC: Colorectal cancer; NSCLC: Non-small-cell lung cancer.

in the methods above are applicable for each tumor type, however, tumor-specific macros are provided to allow the investigator to select nuances of a specific disease type, should the research question warrant additional precision.

Key components of all macros

Table 1 summarizes the line of therapy advancement rules that are consistent across all tumor types studied. These key factors can be summarized as starting a new drug (chemotherapy agents are handled differently than biologic or targeted agents in estimating a change in line of therapy), gap periods without any treatment (representing the intent to change or a period of disease stability or drug holiday) and drug interchangeability.

For all cancers, if there are new chemotherapy agents added after the initial 28-day period (other than drugs considered interchangeable, as defined in the methods), this does indicate that a new strategy is being implemented to control disease, and the line is advanced. For biologic and targeted agents, these are frequently used to supplement a regimen, perhaps due to delays in reimbursement or insurance approvals or to the time to receive results of biomarker testing. However, the practice of biologic/targeted agent supplementation does vary by disease type, and rules are provided for scenarios in which supplementation may indicate a change in disease status. In all cases, if there are changes to the backbone chemotherapy regimen that occur with the biologic/targeted agent supplementation, the line with advance by one due to the new chemotherapy drug that has been added.

A gap period of 60 days was set as default to allow for re-initiation of the same drug for NSCLC due to the aggressiveness of disease. If disease were progressing it would be unlikely that a long drug holiday would be part of the treatment plan. Therefore, if a gap period of at least 60 days is reached, the line will advance by one. For other cancers that in general progress less rapidly, a longer gap period was identified. CRC tends to have longer nonmedication periods between drug changes, due to the slower progression of disease than lung cancer. Because of

Table 2. Additional considerations for maintenance therapy (example is specific to non-small cell lung cancer).

Rule	Option A	Option B	Option C	Option D	Optimal rule and considerations
Line of therapy changes when:					
Patient receiving a first-line regimen containing pemetrexed has no treatment for a gap of more than ____ and reinitiates single-agent pemetrexed or bevacizumab + pemetrexed therapy	Never changes line of therapy	60 days	90 days	180 days	NSCLC – option C For other cancers for which maintenance therapy is used, 'pemetrexed' and 'bevacizumab' may be replaced by maintenance drugs used in the specific cancer type
Patient receiving a first-line regimen containing bevacizumab and has had no treatment for a gap of more than ____ and reinitiates single-agent bevacizumab or bevacizumab + pemetrexed therapy	Never changes line of therapy	60 days	90 days	180 days	NSCLC – option C For other cancers for which maintenance therapy is used, 'pemetrexed' and 'bevacizumab' may be replaced by maintenance drugs used in the specific cancer type
Line of therapy does not change when:					
Patient receiving a first-line regimen containing pemetrexed and has had no treatment for a gap of less than or equal to ____ and reinitiates single-agent pemetrexed or bevacizumab + pemetrexed therapy	Never changes line of therapy	60 days	90 days	180 days	NSCLC – option C For other cancers for which maintenance therapy is used, 'pemetrexed' and 'bevacizumab' may be replaced by maintenance drugs used in the specific cancer type
Patient receiving a first-line regimen containing bevacizumab and has had no treatment for a gap of less than or equal to ____ and reinitiates single-agent bevacizumab or bevacizumab + pemetrexed therapy	Never changes line of therapy	60 days	90 days	180 days	NSCLC – option C For other cancers for which maintenance therapy is used, 'pemetrexed' and 'bevacizumab' may be replaced by maintenance drugs used in the specific cancer type

NSCLC: Non-small-cell lung cancer.

this biology, a line of therapy will advance by one only after a 180-day treatment gap period. However, biologic and targeted agents are commonly used; the addition of these drugs, without discontinuing the chemotherapy drugs used in the line of therapy, does not advance the line.

Maintenance therapy: NSCLC macro

Table 2 summarizes the specific advancement rules that should be considered for maintenance therapy (in this case, NSCLC). In this setting, patients with metastatic disease are assigned a regimen based on the set of systemic anticancer drugs received within the first 28 days. Maintenance therapy remains a part of the clinical practice for NSCLC in the first-line setting and a treat-to-progression model is in place for all lines of therapy. Therefore, if a patient is receiving a first-line regimen that contains agents included in the guidelines for maintenance therapy (e.g., pemetrexed, bevacizumab), the discontinuation of all other agents and continuation of these maintenance drugs, if the continuation occurs within a 90-day gap will not advance the line of therapy. The maintenance therapy rules only apply in the first-line setting, after this point the general macro rules apply and the maintenance rules are no longer necessary (for second-line and later therapies the rules in Table 1 alone apply). Figure 1 presents the flow chart translation of the rules in Tables 1 & 2 to the logic for macro development for NSCLC.

Drug supplementation: CRC macro

As mentioned above, the practice of biologic/targeted agent supplementation varies by disease type, and Table 3 presents rules (in this case, CRC) for which supplementation may indicate a change in disease status. There are two EGFR inhibitors (cetuximab and panitumumab) used in CRC that deserve specific attention. Other biologic agents can be added to a regimen without advancing the line, but addition of these specific biologic drugs later in the course of treatment likely indicates that there is something about the disease that needs attention and advances the line. The initial 90-day period for these drugs to be added is permitted without line advancement, as there may be early delays, perhaps due to reimbursement or access issues. Additionally, new chemotherapy drugs do not advance the line among patients receiving single-agent capecitabine or 5-fluorouracil within 60 days of initiating the line of therapy as this is another case of common supplementation in CRC; however, if drugs are added to the single agent regimen after that period, the line of therapy advances. Figure 2 presents the flow chart translation of the rules in Tables 1 & 3 to the logic for macro development for CRC.

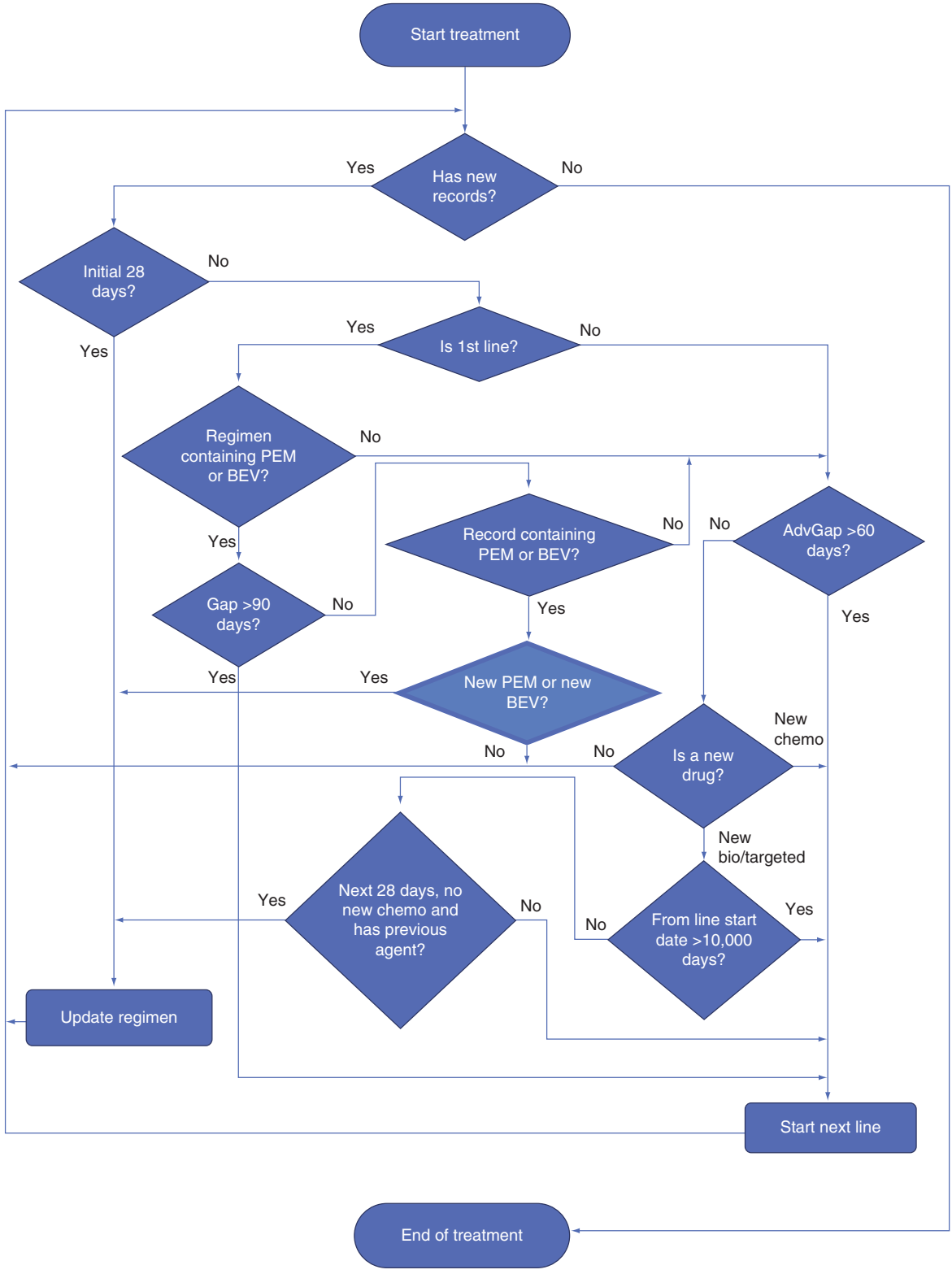


Figure 1. Flow chart summarizing content of Statistical Analysis Software® macro for non-small cell lung cancer.
BEV: Bevacizumab; bio: Biologic or targeted agent; PEM: Pemetrexed.

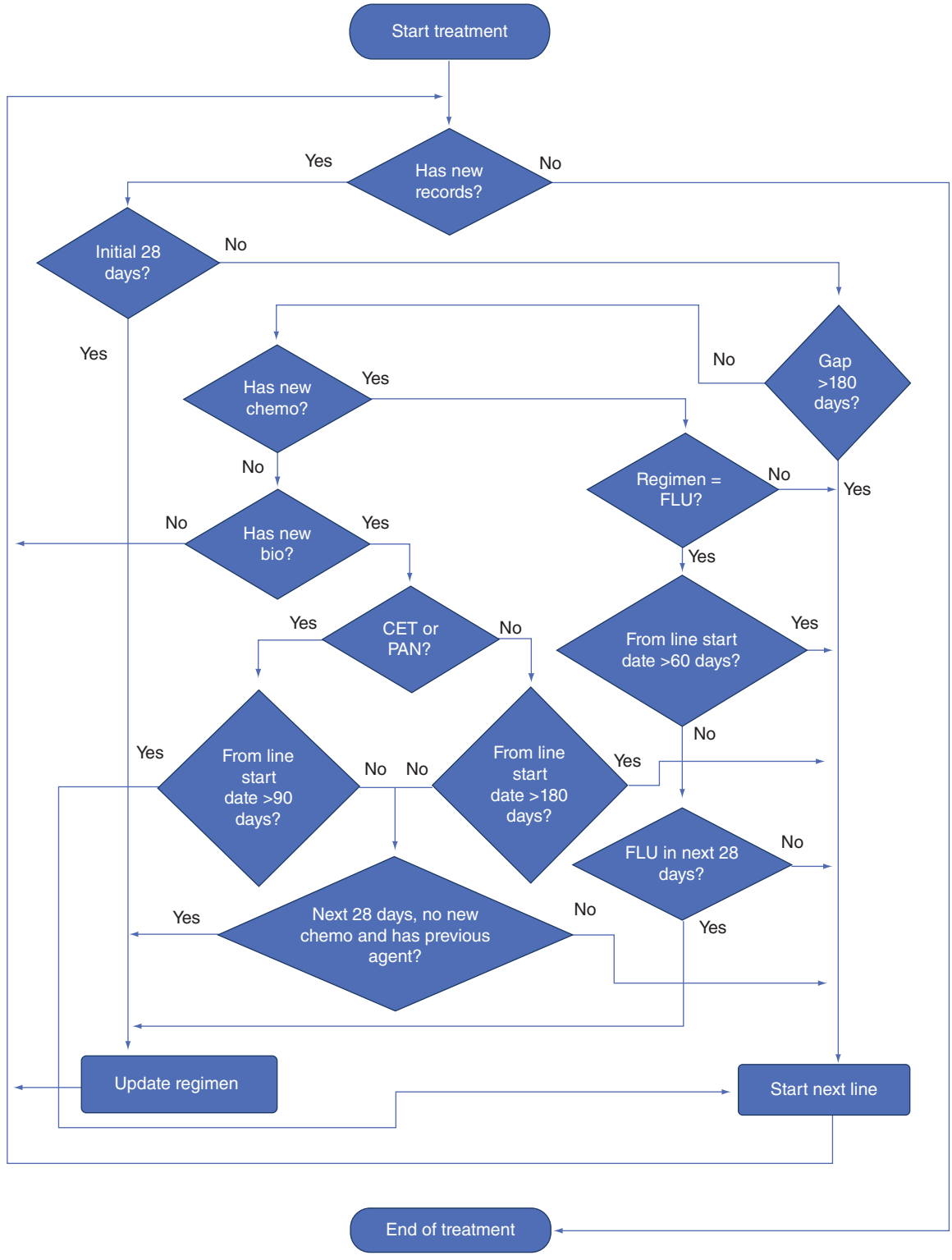


Figure 2. Flow chart summarizing content of Statistical Analysis Software® macro for colorectal cancer.
bio: Biologic or targeted agent; CET: Cetuximab; FLU: Fluoropyrimidine; PAN: Panitumumab.

Table 3. Additional considerations for diseases where drug supplementation is not uncommon (example is specific to colorectal cancer).

Rule	Option A	Option B	Option C	Option D	Optimal rule and considerations
Line of therapy changes when:					
Any biologic/targeted (except cetuximab or panitumumab) is added to a regimen greater than ____ of the start of the line of therapy	After initial 28-day period	90 days	180 days	The line of therapy never changes with adding a biologic	CRC – option C For other cancers that utilize a targeted therapy that could indicate a change in disease status, consider replacing 'cetuximab' and 'panitumumab' with other drugs that are used in this manner
Cetuximab or panitumumab is added to a regimen greater than ____ of the start of the line of therapy	60 days	90 days	180 days	The line of therapy never changes with adding a biologic	CRC – option B For other cancers that utilize a targeted therapy that could indicate a change in disease status, consider replacing 'cetuximab' and 'panitumumab' with other drugs that are used in this manner
A new chemotherapy drug is added to single-agent fluoropyrimidine (5-fluorouracil and capecitabine are considered equivalent) greater than ____ of the start of single-agent therapy	60 days	90 days	180 days	The addition of these agents does not change a line of therapy	CRC – option A For other cancers that commonly supplement a regimen, consider replacing 'fluoropyrimidine' with other drugs that are used in this manner
Adding one or more chemotherapy drugs in a regimen will constitute a change in line of therapy, except for drugs added to the regimen of a single-agent fluoropyrimidine (5-fluorouracil and capecitabine considered equivalent)	After initial 28-day period				CRC – option A For other cancers that commonly supplement a regimen, consider replacing 'fluoropyrimidine' with other drugs that are used in this manner
Line of therapy does not change when:					
Any biologic/targeted agent (except cetuximab or panitumumub) is added to a regimen less than or equal to ___ days of the start of therapy	28 days	90 days	180 days	The line of therapy never changes with adding a biologic	CRC – option C For other cancers that utilize a targeted therapy that could indicate a change in disease status, consider replacing 'cetuximab' and 'panitumumab' with other drugs that are used in this manner
Cetuximab or panitumumab is added to a regimen less than or equal to ____ of the start of the line of therapy	60 days	90 days	180 days	The line of therapy never changes with adding a biologic	CRC – option B For other cancers that utilize a targeted therapy that could indicate a change in disease status, consider replacing 'cetuximab' and 'panitumumab' with other drugs that are used in this manner
A new chemotherapy drug is added to single-agent fluoropyrimidine (5-fluorouracil or capecitabine) less than or equal to ____ of the start of single-agent therapy	60 days	90 days	180 days	The addition of these agents does not change a line of therapy	CRC – option A For other cancers that commonly supplement a regimen, consider replacing 'fluoropyrimidine' with other drugs that are used in this manner
A new chemotherapy drug is added to single-agent fluoropyrimidine (5-fluorouracil and capecitabine are considered equivalent) less than or equal to ____ of the start of single-agent therapy	60 days	90 days	180 days	The addition of these agents does not change a line of therapy	CRC – option A For other cancers that commonly supplement a regimen, consider replacing 'fluoropyrimidine' with other drugs that are used in this manner

CRC: Colorectal cancer.

Base case conditions: gastric cancer macro

The gastric cancer macro may be the most generalizable to other cancers without unique treatment practice patterns, such as maintenance or drug supplementation. Future work should apply this macro to a variety of solid and hematologic malignancies to evaluate its performance. The logic of the gastric cancer macro is based on the scenario of a fairly aggressive disease with limited treatment options. These features are common with many solid malignancies that have become metastatic. After a patient is diagnosed with metastatic disease, the first anticancer systemic therapy drug used is identified. Similar to all macros, the set of drugs administered or received by the patient within the initial 28-day period from the first drug is considered the first-line regimen. The line of therapy advances by one (e.g., to second-line therapy) if the patient has a gap of 90 days or longer with no therapy, before re-initiation of treatment. The line of therapy does not advance if a biologic or targeted drug is added to the regimen.

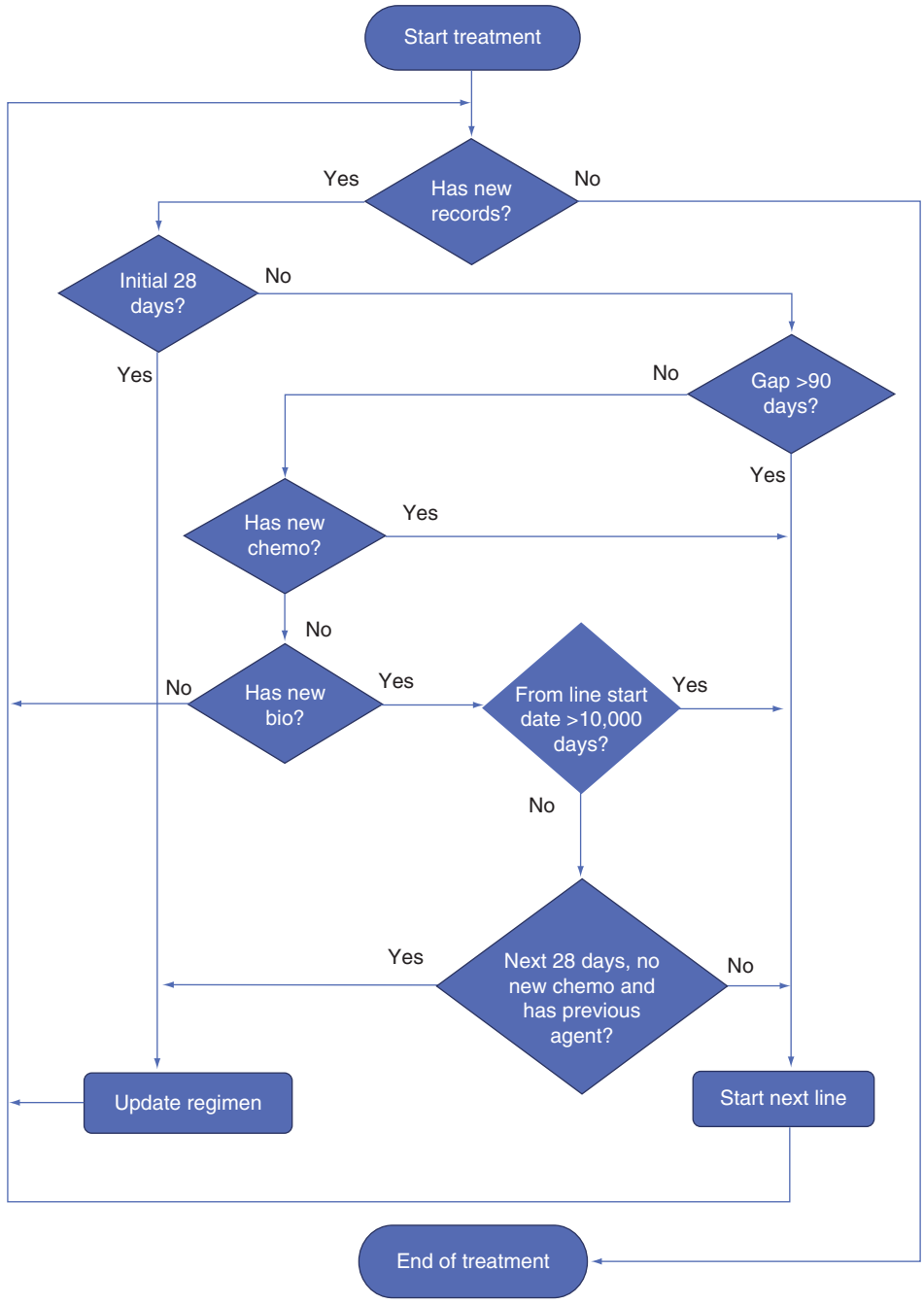


Figure 3. Flow chart summarizing content of Statistical Analysis Software® macro for gastric cancer.
bio: Biologic or targeted agent.

However, if the regimen is discontinued and a biologic or targeted agent is started, this would advance the line of therapy. If new chemotherapy drugs are added any time after the initial 28-day period (other than an exchange of drugs considered equivalent), the line of therapy will advance by one. In this more aggressive disease state, changes in therapy are considered more meaningful and there is less tolerance for gap periods or drug interchanges than CRC as described above, which tends to grow at a slower rate, but is longer than NSCLC, which has the shortest tolerance for a gap period. Therefore, this macro may be adapted to other diseases taking into consideration the aggressiveness of disease when determining the optimal gap period to select. The more aggressive the cancer, the shorter the gap period to be used to estimate a change in line of therapy. Figure 3 presents the flow chart translation

of the rules in Table 1 to the logic for macro development for gastric cancer.

Discussion & conclusion

This research has resulted in a method to estimate lines of therapy in the advanced/metastatic setting of the care of patients with NSCLC, CRC or gastric cancer. Macros are presented in the Supplementary data for use in SAS that require the following variables in a real-world dataset for implementation: patient identifier; drug names and dates of administration or billing. The application of these macros requires that several steps be completed first by the researcher to arrange their data in a manner that will accommodate macro coding. These include cohort development (which is a critical element of research, not discussed in this work) and identifying drug codes to then categorize the type of therapy received (i.e., as chemotherapy or as a biologic/targeted agent). Non-systemic therapies, such as hormonal agents or supportive care drugs, should be excluded from the tumor types we explored in this study. For other tumor types, such as breast or prostate cancer, these agents will play an important role in the care of the patient. The researcher should investigate the role of these drugs before applying or adapting these macros for other tumor types (such as breast or prostate cancer, where hormonal agents play an important role), as it is highly likely that as demonstrated with these three tumor types, each disease is likely to have unique aspects of care that should be incorporated into line of therapy determinations. No single algorithm will fit all tumor types.

The algorithms and SAS code provided to the reader are useful in a variety of research scenarios using real-world EMR or administrative claims data that require some form of definition of line of therapy from limited data. These include comparative effectiveness research, where an investigator may wish to perform a comparative analysis of the survival outcomes of two or more different treatment options in the real-world setting. There are multiple recent examples of this type of work being published [27,28], as well as studies generating a real-world control cohort from databases to compare with a single-arm clinical trial cohort that all rely on some form of line of therapy rule to balance cohorts [29,30]. If patients are not compared with treatment received within the same line of therapy between groups, there may be more patients with further advanced disease in one comparator versus another, and will result in bias in the survival outcomes between groups. Research investigating treatment sequencing and patterns of clinical practice must also rely on researcher-defined line of therapy rules to ensure that the sequences represent changes from the same line to the subsequent line of therapy [31,32]. If the lines of therapy are misrepresented, there will be incorrect treatments, sequences or outcomes reported. The algorithms and SAS code provided are designed to improve standardization across studies to ensure comparability across similar study designs and across those with similar research questions, as most publications to date fail to provide sufficient detail to allow for replication of the research specifically when it relies on determination of line of therapy.

While there is no one set of rules that will fit all research scenarios, the definition of initial regimen within the first 28-day period has been a constant across all tumor types investigated. The other consistent factors to consider are what happens when a new drug is used during care, and if the type of drug could reflect any underlying differences in clinical decision making. Additionally, one must consider periods without drug treatment. These 'gap periods' could have implications in terms of potential drug holidays (that do not advance the line when drug is restarted) or could reflect tumor response and a 'watch and wait' approach for further treatment decision making (which would suggest a line advancement takes place upon treatment initiation). The need for clinical expertise in the development of the rules for any research project cannot be overstated.

There are several assumptions made in the development of these rules to build lines of therapy. Some specific drugs were considered interchangeable. In some cases, the researcher may not wish to consolidate these drugs and should remove that aspect of the algorithms. In other cases, additional drugs may be considered to be interchangeable (such as taxanes). The decision to consolidate drug classes will be dependent on the research question. In the current work, the only drugs that were considered interchangeable were those that clinical experts deemed to least likely reflect a progression in disease when one drug was substituted for another. This may not be ideal for every research project. For example, if one conducts a study evaluating heterogeneity of treatment patterns, there may be less interest in consolidation in order to understand the variability of drugs used in clinical care. On the other hand, a comparative effectiveness study investigating a specific type of treatment versus standard of care may wish to increase the drug consolidation into broader categories or drug classes, particularly if there are known efficacy data showing equivalence of one drug to another. While these macros can be applied as is, it is recommended that the research question be considered in the context of how these algorithms were developed.

There are several limitations to note. While these rules have been applied in multiple studies and have been found to reflect expected treatment patterns, there is no known gold standard against which to test our rules.

Even within in-depth chart review, it is difficult to assign the intended sequence of care in retrospective studies. Therefore, we could not validate these rules against any reliable data source, other than observing the results in the context of appropriate clinical care of patients with these conditions. Additionally, it must be acknowledged that any set of rules applied broadly will not correctly characterize all patients. In any broad rule applied across a diverse patient cohort, there will be some who are incorrectly classified as advancing a line of therapy who did not, and *vice versa*. The algorithms presented here are designed to minimize the error of both over and underestimation of line advancement, but are not expected to be fully accurate for all patients as the reasons for changing therapies is unknown. If data were available that contain these reasons, the algorithm could be improved by including such a variable. The role of radiation therapy was not included in these macros due to limited data in most EMR systems related to concurrent radiation therapy. Therefore, for diseases in which chemoradiation is common and where variables are available for study, these macros must be adapted to account for this. At last, the macros and algorithms must be reviewed periodically as new treatments and novel therapeutic strategies become available, to ensure they reflect current practice patterns.

Future perspective

It is important that particularly EMR or other healthcare delivery systems consider including line of therapy and disease progression information into electronic data capture systems without burdening oncology healthcare providers. This will further enable the conduct of health systems research that can rely on accurate versus estimated information.

Summary points

- The majority of current databases do not contain sufficient fields to allow for the identification of reasons for treatment change or line of therapy change, requiring that rules be applied to estimate these factors in administrative and electronic medical record databases.
- Each tumor site within the field of oncology is unique, and there is no one set of rules that will fit across all diseases.
- The primary factors for consideration when developing rules for lines of therapy include the current treatment patterns, establishing the appropriate study cohort, defining how drugs may be changed within a line of therapy and defining the time period that is typical between lines of therapy that suggest progression versus drug holiday periods.
- This work provides rules and SAS® macros to investigators to provide transparency and consistency in the application of rules that incorporate these factors in the settings of non-small-cell lung cancer, colorectal cancer and gastric cancer.
- Guidelines to investigators are provided about how to adapt these macros to database research in other tumor types.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/suppl/10.2217/fon-2020-1041

Author contributions

LM Hess conceptualized the project and drafted the manuscript; LM Hess and ZL Cui designed the work; X Li, Y Wu and RJ Goodloe analyzed the data; all co-authors evaluated and interpreted the data, critically revised the draft manuscript, and approved the final manuscript for submission.

Financial & competing interests disclosure

All authors are employees of Eli Lilly and Company. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. Fitzmaurice C, Akinyemiju TF, Al Lami FH *et al.* Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2016: a systematic analysis for the global burden of disease study. *JAMA Oncol.* 4(11), 1553–1568 (2018).
2. Bittoni MA, Arunachalam A, Li H *et al.* Real-world treatment patterns, overall survival, and occurrence and costs of adverse events associated with first-line therapies for medicare patients 65 years and older with advanced non-small-cell lung cancer: a retrospective study. *Clin. Lung Cancer* 19(5), e629–e645 (2018).
3. Wang F, Mishina S, Takai S *et al.* Systemic treatment patterns with advanced or recurrent non-small cell lung cancer in Japan: a retrospective Hospital Administrative Database Study. *Clin. Ther.* 39(6), 1146–1160 (2017).
4. Abrams T, Hess LM, Zhu YE, Schelman W, Liepa AM, Fuchs C. Predictors of heterogeneity in the first-line treatment of patients with advanced/metastatic gastric cancer in the US. *Gastric Cancer* 21(5), 738–744 (2018).

• **Utilizes the line of therapy rules in the current study. The reader may review this publication to evaluate their performance**

5. Kennecke H, Berry S, Maroun J *et al.* A retrospective observational study to estimate the attrition of patients across lines of systemic treatment for metastatic colorectal cancer in Canada. *Current Oncol.* 26(6), e748 (2019).
6. Mcmullen S, Hess LM, Kim ES *et al.* Treatment decisions for advanced non-squamous non-small cell lung cancer: patient and physician perspectives on maintenance therapy. *Patient* 12(2), 223–233 (2019).
7. Nagy A, Rahman OA, Abdullah H, Negida A. Do we need maintenance chemotherapy in advanced NSCLC in the era of immune and targeted therapy? *Curr. Cancer Ther. Rev.* 15(1), 50–55 (2019).
8. Milano G. More rationale for optimal sequencing of therapeutic monoclonal antibodies in metastatic colorectal cancer. *Eur. J. Cancer* 117, 131–132 (2019).
9. Sveen A, Kopetz S, Lothe RA. Biomarker-guided therapy for colorectal cancer: strength in complexity. *Nat. Rev. Clin. Oncol.* 17(1), 1–22 (2019).
10. Shah S, Raskin L, Cohan D, Freeman M, Hamid O. Treatment patterns of malignant melanoma in the United States from 2011 to 2016: a retrospective cohort study. *Curr. Med. Res. Opin.* 36(1), 63–72 (2020).

• **Applies the concepts of regimen definition, gap period and addition of new drugs in melanoma to define lines of therapy.**

11. Singal G, Miller PG, Agarwala V *et al.* Association of patient characteristics and tumor genomics with clinical outcomes among patients with non–small cell lung cancer using a clinicogenomic database. *JAMA* 321(14), 1391–1399 (2019).

• **Utilizes an algorithm (unpublished) that prepopulates the line of therapy in the database and unfortuntely is not reproducible.**

12. Bikov KA, Mullins CD, Seal B, Onukwugha E, Hanna N. Algorithm for identifying chemotherapy/biological regimens for metastatic colon cancer in SEER-Medicare. *Med. Care* 53(8), e58–e64 (2015).

•• **Provides a Statistical Analysis Software® macro for lines of therapy using the Surveillance, Epidemiology and End Results (SEER)-Medicare database.**

13. Song X, Cong Z, Wilson K. Real-world treatment patterns, comorbidities, and disease-related complications in patients with multiple myeloma in the United States. *Curr. Med. Res. Opin.* 32(1), 95–103 (2016).
14. Rajkumar SV, Richardson P, San Miguel JF. Guidelines for determination of the number of prior lines of therapy in multiple myeloma. *Blood* 126(7), 921–922 (2015).
15. Meng W, Ou W, Chandwani S, Chen X, Black W, Cai Z. Temporal phenotyping by mining healthcare data to derive lines of therapy for cancer. *J. Biomed. Inform.* 100, 103335 (2019).
16. Planchard D, Popat S, Kerr K *et al.* Metastatic non-small cell lung cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* 30(5), 863–870 (2019).
17. Ettinger DS, Wood DE, Aggarwal C *et al.* NCCN Guidelines insights: non-small cell lung cancer, version 1.2020: featured updates to the NCCN guidelines. *J. Natl Compr. Canc. Netw.* 17(12), 1464–1472 (2019).
18. Messersmith WA. NCCN guidelines updates: management of metastatic colorectal cancer. *J. Natl Compr. Canc. Netw.* 17(5.5), 599–601 (2019).
19. Alsina M, Miquel JM, Diez M, Castro S, Tabernero J. How I treat gastric adenocarcinoma. *ESMO Open* 4(Suppl. 2), e000521 (2019).
20. Ajani JA, D'Amico TA, Bentrem DJ *et al.* Esophageal and esophagogastric junction cancers, version 2.2019, NCCN clinical practice guidelines in oncology. *J. Natl Compr. Canc. Netw.* 17(7), 855–883 (2019).
21. Hess LM, Cui ZL, Mytelka DS, Han Y, Goodloe R, Schelman W. Treatment patterns and survival outcomes for patients receiving second-line treatment for metastatic colorectal cancer in the USA. *Int. J. Colorectal Dis.* 34(4), 581–588 (2019).

• **Utilizes the line of therapy rules in the current study, to identify the second-line cohort.**

22. Zhu Y, Hess L. PCN258 improving the accuracy of measuring duration of therapy by identifying patients who remain on-treatment in a real world database. *Value Health* 22, S105 (2019).

23. Hess LM, Zhu Y, Liepa AM *et al.* Geographic and site variability in the treatment of patients with metastatic gastric cancer. *J. Clin. Oncol.* 3(Suppl. 54), 153 (2017).

24. Hess LM, Cui ZL, Wu Y, Fang Y, Gaynor PJ, Oton AB. Current and projected patient and insurer costs for the care of patients with non-small cell lung cancer in the United States through 2040. *J. Med. Econ.* 20(8), 850–862 (2017).

25. Hess LM, Louder A, Winfree K, Zhu YE, Oton AB, Nair R. Factors associated with adherence to and treatment duration of erlotinib among patients with non-small cell lung cancer. *J. Manag. Care Spec. Pharm.* 23(6), 643–652 (2017).

26. Hess LM, Michael D, Mytelka DS, Beyrer J, Liepa AM, Nicol S. Chemotherapy treatment patterns, costs, and outcomes of patients with gastric cancer in the United States: a retrospective analysis of electronic medical record (EMR) and administrative claims data. *Gastric Cancer* 19(2), 607–615 (2016).

- **Utilizes the line of therapy rules in the current study. The reader may review this publication to evaluate their performance.**

27. Cui P, Li R, Huang Z *et al.* Comparative effectiveness of pembrolizumab vs. nivolumab in patients with recurrent or advanced NSCLC. *Scientific Reports* 10(1), 1–7 (2020).

28. Mudad R, Patel MB, Margunato-Debay S, Garofalo D, Lal LS. Comparative effectiveness and safety of nab-paclitaxel plus carboplatin vs gemcitabine plus carboplatin in first-line treatment of advanced squamous cell non-small cell lung cancer in a US community oncology setting. *Lung Cancer (Auckl.)* 8, 179 (2017).

29. Vander Velde N, Guerin A, Ionescu-Ittu R *et al.* Comparative effectiveness of non-cisplatin first-line therapies for metastatic urothelial carcinoma: Phase II IMvigor210 study versus US patients treated in the Veterans Health Administration. *Eur. Urol. Oncol.* 2(1), 12–20 (2019).

30. Davies J, Martinec M, Delmar P *et al.* Comparative effectiveness from a single-arm trial and real-world data: alectinib versus ceritinib. *J. Comp. Eff. Res.* 7(09), 855–865 (2018).

31. Princic N, Aizer A, Tang DH, Smith DM, Johnson W, Bardia A. Predictors of systemic therapy sequences following a CDK 4/6 inhibitor-based regimen in post-menopausal women with hormone receptor positive, HEGFR-2 negative metastatic breast cancer. *Curr. Med. Res. Opin.* 35(1), 73–80 (2019).

32. Burton T, Byfield SD, Smith GL *et al.* Clinical and economic outcomes by first-line treatment among women with HR+/HER2- metastatic breast cancer in a large US health plan database. *Curr. Med. Res. Opin.* 32(8), 1417–1423 (2016).