Toward rapid learning in cancer treatment selection: An analytical engine for practice-based clinical data

Samuel G. Finlayson, Mia Levy, Sunil Reddy, Daniel L. Rubin

Harvard Medical School, Boston, MA, United States
Vanderbilt University School of Medicine, Nashville, TN, United States
Stanford University School of Medicine, Stanford, CA, United States

A B S T R A C T

Objective: Wide-scale adoption of electronic medical records (EMRs) has created an unprecedented opportunity for the implementation of Rapid Learning Systems (RLSs) that leverage primary clinical data for real-time decision support. In cancer, where large variations among patient features leave gaps in traditional forms of medical evidence, the potential impact of a RLS is particularly promising. We developed the Melanoma Rapid Learning Utility (MRLU), a component of the RLS, providing an analytical engine and user interface that enables physicians to gain clinical insights by rapidly identifying and analyzing cohorts of patients similar to their own.

Materials and methods: A new approach for clinical decision support in Melanoma was developed and implemented, in which patient-centered cohorts are generated from practice-based evidence and used to power on-the-fly stratified survival analyses. A database to underlie the system was generated from clinical, pharmaceutical, and molecular data from 237 patients with metastatic melanoma from two academic medical centers. The system was assessed in two ways: (1) ability to rediscover known knowledge and (2) potential clinical utility and usability through a user study of 13 practicing oncologists.

Results: The MRLU enables physician-driven cohort selection and stratified survival analysis. The system successfully identified several known clinical trends in melanoma, including frequency of BRAF mutations, survival rate of patients with BRAF mutant tumors in response to BRAF inhibitor therapy, and sex-based trends in prevalence and survival. Surveyed physician users expressed great interest in using such on-the-fly evidence systems in practice (mean response from relevant survey questions 4.54/5.0), and generally found the MRLU in particular to be both useful (mean score 4.2/5.0) and useable (4.42/5.0).

Discussion: The MRLU is an RLS analytical engine and user interface for Melanoma treatment planning that presents design principles useful in building RLSs. Further research is necessary to evaluate when and how to best use this functionality within the EMR clinical workflow for guiding clinical decision making.

Conclusion: The MRLU is an important component in building a RLS for data driven precision medicine in Melanoma treatment that could be generalized to other clinical disorders.

1. Introduction

The promise of leveraging vast medical record data to guide clinical decision making has created growing support for the development of “Rapid Learning Systems” (RLSs) that gather and leverage practice-based clinical evidence for real-time clinical decision support [1]. The need for such systems is particularly evident within the field of oncology, where controlled clinical trial evidence is only available to guide therapy in a minority of patients [2–5]. We developed an analytical engine component of the RLS for Melanoma, called the Melanoma Rapid Learning Utility (MRLU). The analytic engine and graphical user interface (GUI) enables users to quickly build and analyze cohorts of patients through an interactive web application. The MRLU could be useful as a key component of future systems aimed at providing evidence-based practice in the era of electronic medical records (EHRs).

In 2012, the Institute of Medicine (IoM) released a landmark report calling for an immediate and fundamental shift in U.S. healthcare. Noting that traditional pathways of knowledge generation and transmission “can no longer keep pace,” the IoM called for
“computing capabilities and analytic approaches to develop real-time insights from routine patient care” [6]. The IoM report highlights a rapidly growing interest in the community to develop and implement rapid learning healthcare systems (RLSs) that are capable of gathering and leveraging clinical evidence to enable real-time, precision decision support in the clinic [1]. The RLS is an example of the repurposing of primary clinical data to improve healthcare, which falls under the broader term of the “learning health system” [7,2].

Electronic health record (EHR) data has long been recognized as being a potential source of “practice-based evidence” that can supplement traditional forms of medical evidence in guiding clinical decision making [8]. Rapid learning systems represent a modern paradigm for precision clinical practice, in which knowledge mined from electronic medical records is seamlessly integrated into the clinical workflow of physicians [9–12]. A functional RLS therefore requires several components, including clinical databases supplied with EHR data, information pipelines that facilitate rapid transformation and filtration of clinical data to identify cohorts of interest, analytic platforms, and clinical decision support (CDS) utilities that provide physicians with relevant clinical insights at point of care (Fig. 1). While CDS for precision medicine is a primary goal of rapid learning, the implementation of RLS will also provide infrastructure that will support and be enriched by nearly all areas of clinical informatics.

The need for rapid learning systems is particularly evident in oncology [2–5]. Tumor molecular profiling offers great opportunity to identify those patients most likely to benefit from targeted therapies, but it results in small sub-populations of patients that can be used for evidence generation. With many tumor molecular alterations occurring in as few as 1% of patients, it is challenging to learn from the clinical outcomes of the mere 5% of cancer patients who participate in clinical trials [13]. While randomized prospective clinical trials and clinical practice guidelines remain the predominant evidence base for clinical decision making in oncology, in order to realize the promise of precision cancer medicine, it will be important to be able to learn from the experiences of all cancer patients. Oncology is thus an ideal environment in which to deploy RLS, with the prospect of informing decision making about treatments based on observational evidence when clinical trial data is not available.

Though much has been done in describing the opportunities and requirements of RLS, few efforts have yet been undertaken to actually produce a RLS for cancer. The Athena and CancerLinQ projects have recently been undertaken to aggregate data and implement clinical practice guidelines and initial analyses [14,15]. However, these and other related efforts remain limited with respect to providing an analytical engine with physician-facing analytic tools for real-time decision support; in fact, to our knowledge, efforts to build analytical engines and physician-facing tools that enable real-time exploration of clinical data have been limited to date [16–18].

In this paper, we present the Melanoma Rapid Learning Utility (MRLU), which we believe is the first implementation of an analytical engine and physician-facing tool that could be used at the core of a RLS for Melanoma. We chose Melanoma as the focus of our RLS engine because Melanoma is one of the few cancers whose incidence rates and total deaths continue to increase [19], unlike most other cancers, and thus efforts to improve decision making for this disease could be particularly impactful. The large heterogeneity both within and between patients with respect to tumor features and tumor response is particularly challenging in Melanoma treatment planning [20,21]. This inherent complexity adds to the existing challenge of finding and evaluating clinical evidence to guide patient care, and makes Melanoma an excellent candidate for rapid learning.

The scope of work necessary to produce and deploy a functioning RLS in a hospital setting is large. In building the MRLU, we focused our efforts specifically on developing an interactive analytic interface that could provide easy comparisons of treatment efficacy for patients with user-specified characteristics. By making latent clinical knowledge physician-accessible, such tools may foster a virtuous “learning” cycle, in which data-driven decision-making is informed by clinical encounters from the past, the results of which will be captured in the same EMR system and therefore be used to guide care in the future. The resultant pool of continuously expanding and improving health data is central to the notion of a learning health system.

To deploy a fully operational RLS, however, the other components of the RLS (Fig. 1) would be required, including the incorporation of practice guidelines and other clinical knowledge bases. Fully operational RLS infrastructures will therefore be capable of

---

**Fig. 1.** Key components of a rapid learning system. (A) Patient data and outcomes extracted from electronic medical records and used to build up large clinical databases. (B) Data from clinical databases is consolidated and transformed. (C) Patient cohorts of interest are defined and extracted from featureized clinical databases. (D) Data analysis conducted to identify or confirm clinical trends. (E) Results may be shared, motivating new basic and public health research or facilitate the design of clinical trials. (F/G) Clinical knowledge bases developed from a combination of the scientific literature, clinical guidelines, and dynamically generated results from rapid learning data analysis. (H) Clinical decision support utilities employing leverage clinical knowledge bases for personalized medicine. (I) Patient outcomes may be impacted by CDS utilities. The MRLU focuses primarily on (C) and (D), though our work also involved components (A) and (B).
more robust rule-based and statistical learning than the MRLU. Our work represents an essential step toward such systems, and, when integrated with clinical information systems and additional components, could provide physicians with access to real-time, practice-based insights for precision medical practice.

2. Materials and methods

2.1. Data

We acquired clinical, demographic, tumor molecular profiling, and pharmacy data from clinical databases for 200 patients at the Vanderbilt-Ingram Cancer Center and 37 patients from Stanford Hospitals and Clinics with metastatic melanoma. All patient data were de-identified through the creation of new, random patient IDs and the offsetting of all dates. The data collected included:

- Clinical data (~240 rows, 1 per patient): sex, age, days to death from first drug treatment, institution of origin.
- Pharmaceutical data (~2000 rows, 1 per prescription per patient): drug class, drug name, and ordering date.
- Tumor molecular profiling data (~240 rows, 1 per patient): BRAF status, NRAS status, and a Boolean indicating whether other mutations were reported.

Pharmaceutical data were normalized by generic, trade, and development names and mapped onto their drug class according to the MyCancerGenome index of anticancer agents [22]. Combination therapies were identified as consisting of two or more drugs of the same class that were administered on the same date. Unique combinations of treatments were identified and represented within the MRLU as distinct treatment regimens. For example, we separated those patients receiving Carboplatin and Paclitaxel together from those patients receiving only Paclitaxel. Eight patients were identified who received combination therapies in which there were drugs of two or more classes. Since no more than any 2 of these patients received the same treatment protocol, and multi-class protocols could introduce potential biases in outcomes analyses stratified by drug class, we excluded these 8 patients from the cohort. All other patients were included.

The timestamp of the first anti-cancer therapy was labeled as the baseline time point for each patient, and each subsequent treatment time point was considered a follow up. The baseline time point served as a reference point used in conjunction with other timestamps in the medical records to calculate the patient age at first treatment, time to second treatment, and days to death relative to the baseline time point. While calculating the time to second treatment and the days to death relative to the initiation of first treatment, we captured the date of right censoring for all patients (i.e., those who either left the database or whose treatments were ongoing at the time of our data analysis). These data were added to the database as well as genetic results compiled from the EHR.

2.2. System implementation

An overview of the MRLU and its architecture are shown in Fig. 2. The MRLU was implemented using a MySQL database and RStudio’s Shiny package [23]. Shiny enables the rapid prototyping of interactive JavaScript web pages back-ended by the R statistical computing language. The MySQL database contains all the patient data and serves as a data cache for accessing the analyzing the data by Shiny applications. The MRLU accesses the data in its MySQL database through the DBI and RMySQL packages and displays them using Hadley Wickham’s ggplot2 [24–26]. We deployed Shiny Server to host the MRLU on a GNU/Linux server [27].

We defined the requirements for the MRLU and its interface during several consultations with practicing clinical oncologists at the Stanford and Vanderbilt cancer centers. During this process, we determined that the core functionality required by clinician users of the tool could be partitioned into three basic tasks: (1)
(1) **Cohort selection** (Fig. 3) is the process of identifying a group of patients matching particular characteristics of interest. In the RLS scenario, a clinician would describe to the MRLU the characteristics of a patient for whom decision support about treatment is sought. The cohort is specified by means of a series of drop-down, slider, and checkbox filters. As inclusion criteria are tightened and/or relaxed via the user interface, data are fetched from the MySQL database, and an R data frame is constructed on the server-side that consists of those patients who match the user-specified criteria. The data frame is manipulated and displayed using an R plotting routine (ggplot2) in a series of five plots that depict the sex, age, institution, and treatment distributions of the current cohort.

Features to enable expressivity of the user in defining cohorts and analyzing their treatments and outcomes were heavily emphasized in design decisions of the MRLU; even excluding age-based inclusion criteria, the simple menu options allow for more than 8 million combinations of inclusion criteria, outcome variables, and stratification variables. Nevertheless, we sought to maintain a very simple and intuitive user interface. As such, most of the menu options are hidden from view by default and are only dynamically generated when made relevant by other query options or opened explicitly. Furthermore, much of the content of the filters are generated from the contents of the data itself, which allows the engine to be robust to changes in the content of the database. For example, the addition of pharmacy data with previously unlisted drugs or drug classes will not necessitate any adjustment to the interface or server code.

(2) **Outcomes analysis** (Fig. 4) of a given selected patient cohort is conducted using the R survival package [28,29]. The user selects an outcome variable in the MRLU system (survival or time to next treatment) as well as a stratification variable (drug class, drug name, mutation type, BRAF status, NRAS status, or sex) to be used in conjunction with the filters defined in the cohort selection step described previously. This input is then used to construct a Cox proportional hazards model [30]. The MRLU builds a model that is stratified based on the user-defined stratification variable, that controls for all other variables available, and that right censors all patients for whom the data were cut off (due to end of observation in the case of survival, and end of observation or death in the case of time to next treatment). The system displays a Kaplan–Meier plot of the cohort, as well as the specific R code used to generate the model. In addition, summary statistics for each included class including median survival and 95% confidence intervals are provided. The MRLU also invokes the survival package’s implementation of the Mantel–Haenszel test to evaluate the equality of the survival curves, and it reports the chi-square and p-value results of the statistical test, as well as the R code used to generate the result. To lower the risk of inadvertent data dredging (see Discussion), the MRLU also displays a warning that encourages users to conduct only one test at a time, and to avoid adjusting parameters post hoc in search of statistical significance.

(3) **Examination and download of raw data**, as with cohort selection above, is handled through a combination of R and Shiny functionality. The user is presented with menus to toggle
between clinical and pharmaceutical data for those patients that match the inclusion criteria established in the cohort selection step. If viewing pharmaceutical data, the user can choose between displaying all drug entries in the table or merely the first entry for each unique drug used on a patient. By default, the tables are sorted according to patient identification number, though the data can be sorted or searched on the content of any column. Finally, if the user chooses to download the cohort to a file, the underlying R data frame is converted to CSV format before being sent to the user.

2.3. System evaluation

We conducted a preliminary evaluation of the MRLU focusing on two aspects of the system: (1) computational integrity, evidenced by the ability of the MRLU to recapitulate known biomedical knowledge about Melanoma treatment effectiveness for particular genetic mutations, and (2) clinician users’ experience using the system to glean information about Melanoma patients in their clinical practice. For the first evaluation, we compared the output of the MRLU with known statistics from the biomedical literature. Specifically, we used MRLU to (1) evaluate the frequency of BRAF positivity in Melanoma patients and to compare it with published facts, (2) to evaluate the response rate of BRAF+ Melanoma patients to BRAF inhibitor therapy such as Vemurafenib, comparing that with published facts, and (3) to evaluate the differences of frequency of Melanoma deaths according to age and gender [31–33].

For the second evaluation, 13 practicing physicians at the Stanford and Vanderbilt Cancer Centers used the system in the context of a hypothetical patient. Each user completed a survey that explained how to use the system to build a cohort of similar patients to the one under consideration and how to run analyses to compare potential therapies. The physicians were also asked a series of questions regarding their views on Rapid Learning Systems in general and their assessments of the utility of the MRLU in particular. They assessed the usability of the MRLU interface, and provided opinions on limitations of the system and needs for future work using Likert scale ratings. The survey itself is available in the Supplemental Data. To assess inter-user agreement among survey respondents, pairwise agreement percentages were computed using both raw Likert scores, as well as scores binned as negative (Likert scores 1–2), neutral (Likert score 3), or positive (Likert scores 4–5). Pairwise agreement rates were computed over all combinations of reviewers, as well as mean agreement rates for each reviewers versus all other reviewers, mean agreement rates for each question, and mean agreement rate overall.

3. Results

3.1. System Implementation

A fully functional implementation of the MRLU was completed, loaded with data, and deployed on a live server, which is publicly available at http://mrlu.stanford.edu:3838/mrlu/. We outline the workflow of the system and results from the process of implementation.
Upon opening the browser, the user is presented with a set of summary plots that show the makeup of the entire dataset with respect to sex, age, BRAF status, NRAS status, drug name and class (of first drug used), and hospital. A corresponding set of filters enable the user to adjust the cohort to set specific patient characteristics. As filters are adjusted, the summary plots update in real-time. This allows the system, before any analysis is conducted, to leverage the judgment of the physician-user in selecting a cohort that matches the patient as closely as possible while being cognizant of the number of patients in the cohort (so that judgment is based on sufficient cohort size) as well as expected distributions for the variables not used for matching.

With cohort inclusion criteria established via the filters, the user selects a stratification variable (drug name, drug class, mutation type, BRAF status, NRAS status, or sex) as well as the outcome variable (survival or time to next treatment), over which the Kaplan–Meier analysis will be performed as described in Section 2 (Fig. 4). As noted above, summary statistics, confidence intervals, chi-square statistics, and p-values are all displayed in conjunction with the plot itself.

Once the preliminary analysis is completed, the cohort data is made immediately available for browsing, searching, sorting, and downloading.

3.2. Evaluation 1: Confirmation of Known Clinical Trends

An initial assessment of the Melanoma Rapid Learning Utility confirmed that the MRLU successfully identified several known clinical trends. This includes the frequency of BRAF mutations in the Melanoma population [34], as well as a survival rate of approximately 50% in response to BRAF inhibitor therapy such as Vemurafenib among those patients whose tumors had a BRAF mutation [32]. Similarly, in accordance with the literature, the system shows that males account for roughly 56% of cases and 57% of deaths among patients under 65, whereas males account for 68% of cases and 67% of deaths among those patients >65 years of age [35]. Furthermore, the physicians who used our system reported that overall survival outcomes are close to what would be expected for a metastatic Melanoma patient receiving current therapy with immunotherapy and/or BRAF inhibitor therapy [32,33].

3.3. Evaluation 2: Physician survey

The physician feedback on both the vision for and the implementation of the MRLU system was very favorable (Figs. 5–7, Supplemental Data).

When asked various questions regarding the benefit that would be provided by Rapid Learning Systems in both clinical and research settings, the mean response was a score of 4.54 (CI: 4.38–4.70) on a Likert scale from 1 to 5, 5 being the most positive (Fig. 5). The greatest enthusiasm (4.62/5, CI: 4.31–4.92) was for the benefit that would be provided to non-academic centers that could leverage data from academic centers for clinical decision support. Assessment of the utility of the MRLU was also positive, with a mean rating of 4.21/5 (CI: 4.00–4.41) (Fig. 6a). Enthusiasm was greatest for the MRLU in context of clinical research, with average ratings of 4.38/5 (CI: 3.99–4.78) in its current state, and 4.62/5 (CI: 4.22–5.01) if only additional variables were made available. While responses were also generally positive for the utility of the system in a clinical setting, the mean ratings were lower in this context, 4/5 (CI: 3.34–4.66) “as is,” and 4.23/5 (CI: 3.79–4.67) with additional variables. When asked separately if the system currently implements the analytic capabilities necessary for clinical usage, the mean score was 3.75/5 (CI: 3.27–4.23), which could indicate that further capabilities would be requested by many clinicians.

were the MRLU to be rolled out into clinical practice. Clinicians did, however, find the system usable, with a mean rating of 4.42/5 (CI: 4.14–4.71) for interpretability and navigability (Fig. 6b). Taken together, these data (overall mean score 4.37/5, CI: 4.25–4.49) indicate that clinicians are generally eager to see Rapid Learning Systems in both the clinical and research settings, and view the progress made toward that aim by the MRLU as valuable in its current state and potentially even more valuable with simple additions to its underlying database.

We were particularly interested to gauge physician attitudes toward what would be required from a practice-based evidence tool like the MRLU in order for it to be trusted in practice (Fig. 7). When asked how many patients matching their own would be expected in order for their outcomes to be clinically useful, there was wide discordance: 23% of respondents indicated that they would use outcomes data from even one similar patient, whereas 30% indicated that they would expect more than 500 patients. Strikingly, only one clinician felt that (s)he could only answer this question if able to perform a power analysis on the data. In contrast, expectations for confounding variables represented in the database were much more uniform, with a majority of clinicians indicated that they would want to see primary tumor characteristics, co-occurring medical conditions, location of metastases, and additional genetic features. However, the lack of features was generally not listed a major hindrance to the future clinical adoption of these systems, with the strongest concern being an efficient workflow in using the system and, potential distrust of the data by clinicians.

Inter-rater reliability results were mixed, though generally positive (Supplemental Data). While the rates of exact agreement were moderately low for the raw 1–5 scale Likert data (overall mean 0.41, CI 0.37–0.44), agreement rates were much higher when the scores were binned into negative, neutral, and positive ratings (mean agreement 0.80, CI: 0.75–0.85). This indicates that discordance among the respondents was likely due to the subjectivity inherent, for example, in distinguishing a positive rating of 4/5 from an extremely positive rating of 5/5 — two responses which, taken together, comprise 87.3% of all recorded ratings. Nevertheless, we do note that, unsurprisingly, the two questions with the lowest average rating also demonstrated the lowest rate of agreement. These questions, which asked respondents to rate the clinical utility of the MRLU in its current implementation, each had only 47% agreement of respondents on the 1–3 scale of binned
scores. Notably, however, this discordance largely vanished (mean pairwise agreement 72%) when respondents were asked to assess the clinical utility of the system if additional patient variables were added. Similarly, the distribution of one-versus-all reviewer agreement rates was somewhat left-skewed, with two reviewers having average agreement rates more than two standard deviations below the mean. Given the presence of these apparent outliers, evaluations of future systems could be performed to study whether reviewer factors such as age or computer-experience, other reviewer factors such as practice attitudes, or mere chance drives discordance.

4. Discussion

The MRLU we developed is a prototype of the analytical engine component of a RLS for Melanoma (Fig. 1). This analytical engine enables physician users, who in many cases lack statistical programming skills or user-friendly tools that access clinical data, to use primary clinical data to interactively build cohorts and run retrospective outcomes analyses tailored to patients like their own. Our results are promising in that our system was able to recapitulate known facts about Melanoma in an observation cohort. This raises confidence that the MRLU may be useful clinically in providing insight into the disease for patient characteristics for which there is not yet published evidence. In addition, the results of our survey suggest that our interface is usable and potentially clinically useful. However, additional studies with a larger amount of patient data and more detailed evaluation of potential clinical benefit will be needed. In particular, the utility of the MRLU may also be limited by the absence of a number of important clinical features in its database. Our clinical experts identified a number of variables that should be included in MRLU before it would be contemplated for clinical use. Since our system is extensible, adding such variables would not be expected to be difficult if available in the EMR. Further, our user evaluation indicates that many clinical practitioners desired to use such systems.
4.1. MRLU in context of existing systems

The MRLU is complementary to recently developed enterprise-level platforms such as i2b2, STRIDE, RedCAP, and tranSMART [36–39]. These are data warehouse systems that also provide tools for integrating and querying medical record and research data at a large scale. They can fulfill an important role as one of the components of the RLS (Component A in Fig. 1); however, while some of these systems do provide analytics features, they have been generally designed and optimized for research rather than real-time use by physicians for rapid learning in a clinical setting [40]. Similarly, several industry-sponsored visual analytic tools have also been designed to facilitate the transformation and navigation of electronic medical record data, as well as for visualizing “care pathways” among similar patients in order to facilitate intervention planning (Component B in Fig. 1) [16–18]. These latter tools lack analyses needed by a RLS analytical engine (e.g., treatment-stratified Kaplan–Meier, etc.).

Another effort related to our work is the CancerLinQ project, which has created a system based on aggregating primary clinical data from a group of community oncology practices [15]. Their system provides many features of the RLS (Components A, B, and C in Fig. 1), but is still in the early stages of development. Its analytical engine only permits rudimentary data exploration, such as providing histograms of patient characteristics, and it lacks the types of analytical capabilities we have developed in MRLU, such as comparative survival analysis in user-defined patient cohorts. Prototype efforts such as our MRLU could inform further development of efforts such as CancerLinQ. Similar, disease-specific RLS tools like the MRLU have also been reported for other diseases, including lung cancer [41].

4.2. Statistical concerns

Interactive data analysis for real-time clinical decision support is complicated by a range of statistical challenges. As highlighted above, we designed the MRLU analytical engine for flexibility—the user can easily adjust almost all aspects of the analysis. In making this design decision, the MRLU can accommodate a wide range of clinical questions that can be quickly posed by the user and answered in near real-time. However, this flexibility raises the risk of inadvertent errors in judgment by users lacking statistical expertise. Furthermore, as discussed above, our survey corroborates previous work suggesting that many clinicians may not be extremely concerned by statistical limitations while using these systems [12]. As such, before deploying tools such as ours in the clinic, careful considerations should be made to ensure that statistical models are appropriately designed and interpreted.

The first statistical challenge faced by the MRLU arises during cohort selection in the form of the bias-variance tradeoff. In order for models built with the MRLU to be relevant to a given patient, selection criteria must be applied to avoid high bias. Conversely, excessive selection criteria will lead to overfit models of high variance. This tradeoff is particularly complicated to manage in the case of Melanoma decision support, as even the largest institutions will have relatively small, yet still widely heterogeneous, datasets. Ultimately this problem can only be addressed by integrating data from many participating institutions, which is the vision for all learning health systems [7].

In order to assist the user in managing the bias-variance tradeoff as effectively as possible, the MRLU’s summary plots on the cohort selection page (Fig. 3) convey descriptive information. As prospective users apply filters to narrow in on patients of a target demographic, plots depict the makeup of the cohort they have built. A note encourages the physician to confirm that the cohort both reflects the characteristics of the patient and fits with his or her professional judgment. Nevertheless, it is impossible to provide sufficient automation to completely preclude the inadvertent misuse of the system by inexperienced users.

Similarly, the system provides no advanced statistical analyses for managing potential pitfalls such as outlier effects and violations of the proportional hazards assumption. As such, these considerations can only be managed through the manual inspection and/or analysis of the cohort data, and the MRLU thus provides this capability. Though our interface is designed to make this as easy as possible, physicians will likely not have the time nor the expertise to conduct such an analysis in a clinical setting. In addition, while our analytical engine seeks to limit confounders by controlling for all variables other than the stratification variable, the inherent complexity of data gathering in routine care leaves open the possibility of other forms of latent confounding [42].

Another risk inherent within any interactive statistical analysis platform is that of inadvertent data dredging. A user of the MRLU can quickly examine many clinical hypotheses in a single sitting. Due to the principle of multiple hypothesis testing, this ease of use can be dangerous, since it could enable a user running sufficient queries to find statistically significant trends by chance [43]. While the risk of data dredging can never be fully eliminated, the physical separation of the cohort selection page from that of the outcomes analysis was designed to encourage users to carefully formulate a cohort for a single question of interest before seeing whether or not significant results will be obtained. The MRLU produces a warning to users about multiple testing, encouraging them to utilize the tool for one test at a time, and warning them of the false-positive risk in rapid and repeated adjustments of the query in search of statistical significance. Nevertheless, probably the best way to prevent statistical misuse of analytical engines such as MRLU is to engage experts in interpreting the analyses.

To further limit the statistical risks inherent in rapid learning, we will present, in a future work, a hybrid system in which a rapid MRLU analytical engine will be linked to a series of interfaces, some of which will allow expert statisticians to design statistical models, others of which will allow physicians to apply these models to their patients without the capacity to define patient cohorts or new models themselves.

4.3. Looking forward

Broader philosophical and legal questions that surround RLS pertain to where exactly the output of these tools should stand within the existing hierarchy of evidence-based practice [11,12,44,45]. The results from controlled clinical trials provide the highest, most reliable evidence, while retrospective, observational evidence can be less reliable. Rapid learning tools based on engines such as the MRLU are no substitute for traditional forms of medical evidence when they exist and are relevant to the patient of interest. In the future, we plan to implement in MRLU a mechanism to identify and present clinical guidelines and prominent articles from the literature that relate to the user’s query. However, in the (common) circumstance of cancer patients for which there is no clinical trial evidence available, the MRLU could be particularly useful, by leveraging evidence from clinical practice to provide information to physicians that may help in making treatment decisions. While our survey indicates that RLS analytical engines such as the MRLU would likely see clinical use, prospective research studies that relate their use to clinical outcomes will ultimately be needed to show clinical value.

As highlighted in Section 4.2, an increasingly important component of systems such as MRLU will be the ability to expand the number of patients and data types used for analyses across many institutions. We are currently expanding and generalizing MRLU to address this challenge by building a scalable, decentralized...
system. In this system, a front-end interface derived from the MRLU will be used to execute distributed statistical models using data from participating institutions. A server at each institution will compute intermediate statistical results on local data, and computational results will be sent back to the coordinating server. This approach will allow centers to participate in inter-institutional computations without sharing any granular patient data. Each site would do a one-time mapping of certain key patient data fields to those used by the system, and this could expand over time to include new data types in future. The menus and utilities in the system that use these fields would dynamically update based on the data types available from the connected institutions. This system could scale up to including many patients as more sites participate, and these institutions would have the freedom to withdraw at any time.

Finally, while the MRLU was developed specifically for use in Melanoma, the key functionality – integrating genetic variants, treatments, and survival outcomes – is relevant to many types of cancer (and other disease). As such, small adaptations to the covariates stored in and analyzed by the system would allow it to scale across cancer types. Since the menus and model can easily be adapted to fit the data at hand, the rate-limiting steps in such adaptation would almost certainly be data acquisition and clinician interest.

Our MRLU is just a portion of the complete RLS (Components C and D in Fig. 1). Clearly, the other components are needed, and the MRLU must be combined with the other infrastructure in order to realize the RLS. On the other hand, we believe our results provide useful insights into design considerations, feasibility and potential utility of the analytical engine component of the RLS.

5. Conclusion

The MRLU is an analytical engine and user interface that represents a component of the RLS. It can provide real-time, data-driven clinical decision support for Melanoma treatment planning. In a preliminary evaluation, the MRLU successfully recapitulated known biomedical knowledge about Melanoma treatment, and it showed promise for clinical utility when used by oncologists. Given its flexible architecture, it is extensible to other types of cancer and to incorporating more and richer data for greater future clinical utility in the future. We plan to incorporate the MRLU into the rest of the learning system infrastructure and may ultimately enable EHR-driven evidence to be incorporated into medical practice.

Conflict of Interest

None.

Acknowledgments

This project has been funded from National Cancer Institute, National Institutes of Health, under Grant U01CA142555 and seed grant from the Big Data for Human Health Stanford University and Oxford University. This project was also supported by award number T32GM007753 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

Philip Lavori, PhD and Balasubramanian Narasimhan, PhD from Stanford University provided consultation in the development of the MRLU. Vanessa Sochat, Linda Szabo, and Luke Yancy Jr. from Stanford University also provided valuable collaboration on a previous rapid-learning prototype. Finally, Denis Agniel provided helpful feedback as we designed the inter-rater reliability analysis.

Appendix A. Supplemental material

Supplementary data associated with this article can be found in the online version, at http://dx.doi.org/10.1016/j.jbi.2016.01.005.

References


