Adding Individual Patient Case Data to The Melanoma Targeted Therapy Advisor

Jovan Stevovic¹
CRG - Centro Ricerche GPI,
Trento, Italy
jovan.stevovi@cr-gpi.it

Jeff Shrager
Stanford and CommerceNet,
Palo Alto, California, USA
jshrager@stanford.edu

Alban Maxhuni¹
Ubiquitous Interaction Group,
CREATE-NET, Trento, Italy
alban.maxhuni@create-net.org

Gregorio Convertino
Xerox Research Centre Europe
Grenoble, France
gregorio.convertino@xrce.xerox.com

Iman Khaghanifar
¹DISI - University of Trento,
Trento, Italy
khaghanifar@disi.unitn.it

Randy Gobbel
CollabRx, Inc.
San Francisco, California, USA
gobbel@acm.org

Abstract—The emergence of genomically targeted cancer treatments has spurred the development of methods that correlate genomic information with treatments and outcomes. Because this information is usually pulled from published literature, such methods are limited to summarizing only the data generated through the slow and narrow publication pipeline. However, many thousands of patients are treated each year whose data does not make it into publications. Each of these is, in effect, a case study whose capture would add to our overall knowledge of cancer treatment, and could speed up the search for treatments and cures. Our work extends one such literature-based system, the Melanoma Targeted Therapy Advisor (TTA), by adding direct patient profiling. This extension of the TTA, or PTTA (Personalized or Patient TTA), both enriches the TTA knowledge base by collecting case reports directly from patients, and gives patients and/or physicians immediate feedback by ranking the best-performing treatments for genomic profiles of interest to them. The PTTA will permit patients to register their test and treatment results and then to obtain rankings for additional potentially useful tests and treatments. It will also provide a report with statistical and literature evidence that justifies the rankings. These functionalities can aid physicians in treating patients in the most effective manner.

Keywords—targeted therapies, melanoma, therapy efficacy, personalized medicine, genomics, cancer, biomedical knowledge.

I. INTRODUCTION

The diagnosis and treatment of cancer are evolving rapidly, and patients are acquiring an ever-increasing amount of information about treatments from the Internet [3]. One value of web-based medical information is the efficiency with which it can be updated, so that it is effectively “always-up-to-date”. One such example is the Targeted Therapy Analyzer (TTA), developed by the Melanoma Molecular Map Project (MMMP [10]) and CollabRx, Inc. Cancer is generally driven by genetic mutations in tumor cells. So the tumor’s genomic profile can be used to personalize treatments to individual mutations. The TTA algorithm matches a patient’s molecular profile, obtained through tests, with a manually-curated literature database, the Targeted Therapy Database (TTD) [11]. Unfortunately,
II. BACKGROUND AND RELATED WORK

Evidence-Based Medicine (EBM), introduced in the early 1990s [5], grounds medicine strongly in data. Clinical decision support systems are active, point-of-care knowledge systems that bring this grounding directly into physician decisions regarding a particular patient’s care. One version of this marriage, so-called “personalized genomic medicine”, utilizes the patient’s molecular profile to guide treatment decisions, and is expected to play an increasingly significant role in patient care [7, 15, 16]. Precise diagnosis and effective treatment will increasingly involve access both to existing knowledge, and to rapid analysis of new data in order to recognize disease patterns and treatment response relations. A good example of this is the DEBM clinical decision support system [8], which collects patient data from Electronic Health Records (EHRs) and uses data mining to identify various disease patterns, even very rare ones. These patterns are extracted automatically and help physicians in making informed treatment decisions.

It is widely accepted that personal health management or self-management is an effective way of managing chronic disease and avoiding illness in healthy people [2]. Online communities and social-networks are known to be effective in helping individuals achieve health outcomes; patients and caregivers have created many personal health information settings (services, forums, etc.) to help manage chronic illnesses such as “Follow Me” (www.followme.com) and “PatientsLikeMe” (www.patientslikeme.com) [4]. In addition, mobile devices are playing a rapidly increasing role in facilitating data collection and therapy guidance, as well as early disease detection [9]. For example, the MelApp application [9] can detect early melanoma by analyzing pictures taken by personal cameras. Many similar personal health apps are rapidly appearing. These tools are gathering enormous amounts of data on treatments and outcomes, which via data mining methods can be used to improve personalized treatment guidance.

Our work focuses on personalized data collection and molecularly-based therapy guidance in late stage melanoma. Other applications, such as [14], have informative purpose providing news about cancers and therapies. Through [13] users can query for specific information in Q&A style, while [12] provides news on the research on cancer and patient care stories. However these applications do not provide patients with personalized therapy guidance by exploiting a patient-populated database such as the platform presented in this paper.

III. USE CASE

Figure 1 depicts a typical use case for the PTTA:

1. A patient registers and is guided to create a molecular profile – a collection of test results. This can be updated as new test results become available. The patient can also record the qualitative state of her/his disease, and updates on her/his general health status.
2. The treatment definition process records the patient’s current and historical treatment information.
3. The profile and treatment content is used to create a report ranking potential treatments and tests through the TTA algorithm [11]. The report includes relevant literature and other statistical information that may be useful to a physician in choosing further tests and possible treatments.
4. As treatment takes place, the patient updates her or his profile. This directly updates the TTD database, and so in turn, updates the rankings for this patient. Importantly, this data gathering also updates the results for other patients with similar profiles. Patients might opt to receive push updates that suggest changes to their treatment.
5. The patient, and his or her physicians, can also use the database to track the patient’s progress, or to explore alternative treatments in a “sandbox” mode.

IV. THE PTTA PLATFORM

The PTTA platform extends the web-based TTA to include this per-case, individualized data, and a new user interface. The three layers of the PTTA architecture, Data (PTTD), Algorithm, and Presentation, are described next.

A. Data Layer (Patient Profile, PTTD)

The Data Layer (PTTD database, or PTTD) extends the TTD [11] to include patient profiling and treatment outcomes (Table 1, Figure 1 (points 1 and 5), see [11] for more details).

<table>
<thead>
<tr>
<th>Molecule</th>
<th>State</th>
<th>Relationship</th>
<th>Drug</th>
<th>Model</th>
<th>Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRAF</td>
<td>mut V600E</td>
<td>sensitivity</td>
<td>Sorafenib</td>
<td>in vitro</td>
<td>14</td>
</tr>
<tr>
<td>BRAF</td>
<td>mut V600E</td>
<td>no rel. with</td>
<td>Sorafenib</td>
<td>human in vitro</td>
<td>10</td>
</tr>
<tr>
<td>BRAF</td>
<td>mut V600E</td>
<td>resistance</td>
<td>Sorafenib</td>
<td>human/clinical</td>
<td>1</td>
</tr>
</tbody>
</table>

In Table 1, Molecule represents the gene or gene product under investigation. State refers to the molecular state of that molecule (e.g., mutated, overexpressed, phosphorylated) by virtue of which it exerts the biological activity related to the disease process and targeted therapy hypothesis. Relationship reports the hypothesized relationship between the molecule and the corresponding treatment/drug (the Drug column): possible values are sensitivity to (positive effect), no relationship with (neither positive nor negative effect), or resistance to (negative) the effect with the drug. Model types can be animal in vitro, human in vitro, human/clinical. Finally, Cases refers to the number of cases reported in literature or by patients.
Surprisingly, implementing the PTTA requires no change to the TTA itself. A large clinical trial is indicated as Model = “human/clinical” and Cases >> 1. All that we need do is store for each patient the current treatment and its effect as a new single row in the table, with Model = “human/clinical” and Cases = 1. The TTA algorithm already accounts for the weight of this sort of observation simply by virtue of already being able to interpret clinical trials; in a sense, cases are interpreted as “n-of-1” clinical trials. This small change – adding case studies – simultaneously (a) constitutes a contribution to research by providing data for mining, (b) aids the provider obtaining the latest reports regarding their tumor, and (c) contributes statistically to the report that other patients obtain through the TTA algorithm.

The individual patient case data added to the PTTD may come from patients, physicians, or clinical researchers; each has different information to offer. Information such as disease state, tumor load, and treatments could be imported from an EMR system (e.g., see [8]), but patients are in a unique position to report on their own well-being. Physicians could provide treatment suggestions and semi-quantitative descriptions of melanoma progress and therefore treatment efficiency for their patients. Importantly, only physicians (or patients’ families or caregivers) can report on severe negative results of a treatment. This bias towards positive outcomes is common in biomedical statistics data (i.e., “incidence-prevalence”, “survivorship”, or “Neyman” bias [6]), and can be dealt with statistically. Finally, clinical researchers can contribute data based on the results of clinical trials and meta-analyses, as well as results for potential drugs from in vitro or in vivo studies.

Although the data provided to the PTTA directly and immediately affects treatment rankings, data quality should, at a large scale, be buffered against normal noise. However, large-scale fake data could seriously skew the results, so spam protection is critical. To ensure data quality and security it may be necessary to assign a private and unique ID to each patient that can be given by physician in order to confirm the identity.
2. Whereas in TTA the user needed to provide a value for hypothesis strength cutoff, PTTA sets this value to the recommended value of 0.5 (explorable in sandbox mode).
3. Whereas in TTA the set of hypotheses was explicitly reported, in PTTA these interim hypotheses are hidden (except in a special research mode).
4. Whereas in TTA the user had to manually enter the results for each test that TTA had proposed in the previous step, PTTA automatically fills in this information using the patient profile, as specified initially, or when it is updated.

Currently, PTTA is under development and the proposed user interface is a proof-of-concept prototype intended to test the PTTD and underlying algorithms. Before going public, it will be revised based on a usability study with real users.

V. PRIVACY AND SECURITY

A key set of functions for healthcare applications are those enabling secure and privacy-aware data management (collection, storage, and sharing). Privacy and security policies are usually dictated by regulations defined by governments. As for other applications in this area, PTTA access control is ensured by usernames and passwords, with the possible addition of unique IDs assigned by physicians, as mentioned above. Also, as suggested in [1], we will evaluate the extent to which exact values of Personal Identifiable Information (PII) are needed to provide useful targeted information. When the results do not depend upon exact values we will employ input generalization [1]. We encrypt the PII with the Advanced Encryption Standard (AES) algorithm and store it separately from non-PII (literature-based data). These are decrypted just at the point of each computation, thereby reducing the possibility of PII leakage.

We do not intend to share patient data with any external entity. If the collected data will be used for research and statistical purposes, as is already done by some patient-centered online communities [4], appropriate privacy-preserving strategies such as data anonymization will be undertaken to avoid PII disclosure. The expected uses of the data will be clearly described in the Terms of Service and the users will be explicitly asked to authorize any new proposed uses.

VI. CONCLUSIONS AND FUTURE WORK

The Personal Targeted Therapy Analyzer adds individualized case data into the Targeted Therapy Database and Analyzer (TTD/TTA) [11], thus significantly extending their utility by enabling them to be kept up-to-date much more rapidly than is possible if one waited for cases to appear in the literature (many of which will never make it there).

This approach can be applied to diseases other than melanoma. In considering the breadth of applicably, one can distinguish "diagnostic" tests from "treatment selective" tests. The present approach is applicable to any disease for which there is a body of evidence linking particular tests to the differential efficacy of particular treatments. Many diseases and tests are in this category, particularly cancers and related genomic tests, but this is not true for some classes of cancers/tests and for some others classes of diseases. For example, whereas HER2-positive breast cancer directly suggests treatment with trastuzumab, elevated PSA levels (even combined with histological results) may be considered "diagnostic" for prostate cancer, but are not "treatment selective" as these observations do not differentially suggest particular treatments. The case is similar for many non-cancers.

In our future work we will evaluate the PTTA by making it available to volunteers from an existing patient community. We will also extend the platform to include social features. For example, one planned feature will match patients who have similar test and/or treatment profiles, allowing them to engage in mutual support. Our design goal is to extend current web-based tools for self-support communities (e.g., www.PatientsLikeMe.com [4]) so that the selective sharing and discussion can motivate entering test and outcomes data.

REFERENCES

[4] Frost, J.H., et al., Social uses of personal health information within PatientsLikeMe, an online patient community: what can happen when patients have access to one another’s data”, JMIR. 10 (3) (2008).