JAMIA Open, 4(4), 2021, 1–6
https://doi.org/10.1093/jamiaopen/ooab090
Application Notes

Application Notes

Follow-up Interactive Long-Term Expert Ranking (FILTER): a crowdsourcing platform to adjudicate risk for survivorship care

Alex C. Cheng [1], Li Wen [2], Yanwei Li [3], Tatsuki Koyama [4], Lynne D. Berry [4], Tuya Pal [5], Debra L. Friedman [6], and Travis J. Osterman [1]

[1]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee, USA, [2]Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, Tennessee, USA, [3]School of Engineering, Vanderbilt University, Nashville, Tennessee, USA, [4]Department of Biostatistics, Vanderbilt University Medical Center, Nashville, Tennessee, USA, [5]Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA, and [6]Department of Pediatrics, Vanderbilt University Medical Center, Nashville, Tennessee, USA

Corresponding Author: Alex Cheng, PhD, MEM, Department of Biomedical Informatics, Vanderbilt University Medical Center, 2525 West End Ave, Suite 1475, Nashville, TN 37203, USA; a.cheng@vumc.org

Received 25 May 2021; Revised 14 September 2021; Editorial Decision 14 October 2021; Accepted 21 October 2021

ABSTRACT

Objectives: To develop an online crowdsourcing platform where oncologists and other survivorship experts can adjudicate risk for complications in follow-up.

Materials and Methods: This platform, called Follow-up Interactive Long-Term Expert Ranking (FILTER), prompts participants to adjudicate risk between each of a series of pairs of synthetic cases. The Elo ranking algorithm is used to assign relative risk to each synthetic case.

Results: The FILTER application is currently live and implemented as a web application deployed on the cloud.

Discussion: While guidelines for following cancer survivors exist, refinement of survivorship care based on risk for complications after active treatment could improve both allocation of resources and individual outcomes in long-term follow-up.

Conclusion: FILTER provides a means for a large number of experts to adjudicate risk for survivorship complications with a low barrier of entry.

Key words: cancer survivors (D000073116), crowdsourcing (D063045), risk factors (D012307), expert systems (D005103)

BACKGROUND AND SIGNIFICANCE

An estimated 5% of the US population (16.9 million people) are cancer survivors.[1] For these individuals, care should focus on disease surveillance and health promotion to prevent or ameliorate chronic health issues and subsequent malignancies. Adverse health outcomes are well documented among cancer survivors, especially in underserved populations due to barriers to obtaining care such as financial toxicity, low income, transportation, and insurance inadequacy.[2–4] Services must therefore address and minimize adverse cancer treatment sequele and decrease risk for recurrent or subsequent malignancies.

For well over a decade, national best practice guidelines have recommended that cancer survivors receive survivorship care. Definitions of levels of survivorship care vary widely, but generally determine the frequency that patients are seen by an oncology team vs a primary care provider.[5,6] A "one-size-fits-all" model is neither feasible nor sustainable. As cancer care is becoming more precise, so too should survivorship care. Care plans should be tailored to the

LAY SUMMARY

With advances in cancer treatment, more people are being cured of cancer, leading to a growing population of cancer survivors. The therapies that have resulted in longer lives for cancer patients can be accompanied by long-term health complications. While guidelines for caring for cancer survivors exist, refinement of survivorship care based on risk for complications would allow doctors to determine if a patient needs to be seen by a cancer specialist at an academic medical center, or a primary care physician at a community clinic. Creating risk models would typically require detailed data from a large group of cancer survivors to better understand what attributes make someone more at-risk, which are currently unavailable. We created a website where survivorship experts can judge a series of pairs of hypothetical patients for which patient has a higher risk for complications. This platform, called Follow-up Interactive Long-Term Expert Ranking (FILTER), uses an algorithm originally designed to rank chess players to rank hypothetical patients by risk. FILTER takes a crowdsourcing approach, where many experts will judge many matchups to capture the wisdom of the crowd. We will use the risk scores obtained by FILTER to attribute the effect of individual attributes on survivorship risk.

level of services required and based on each patient's unique set of risk factors. This concept of stratified survivorship care exists in the United Kingdom through the National Cancer Survivorship Initiative and has been appropriately proposed in the United States.[7]

As defined by Oeffinger and McCabe, an individual's survivorship risk is their chance of premature mortality, serious morbidity, or adverse health status.[5] Currently, risk stratification models for survivorship primarily consider intensity of the cancer treatment and likelihood of adverse health conditions, based on limited expert opinion.[5] Development of a data-driven survivorship risk model would require rigorously collected and sufficiently comprehensive long-term follow-up data—a resource currently lacking in the electronic health record (EHR) or cancer registries. Expert opinion remains the best available resource to assign a survivor into a low, medium, or high-risk group. However, access to these experts is limited in the healthcare settings where they are needed most, such as in community-based rural clinics.

OBJECTIVES

To address this gap, we are seeking to develop a survivorship risk model that calculates a patient's required level of follow-up care based on their disease, treatment, genetic, socioeconomic, and demographic factors using clinical knowledge from a large group of experts. We have developed a risk stratification crowdsourcing platform called follow-up interactive long-term expert ranking (FILTER), which invites oncologists and survivorship care experts to judge survivorship follow-up complexity.

MATERIALS AND METHODS

Design considerations for clinical expert crowdsourcing

Our goal was to leverage the expertise of many oncologists and survivorship experts to create a risk-informed algorithm for survivorship care. Crowdsourcing as a means of codifying clinical knowledge is a relatively new concept in oncology and clinical research.[8] Nevertheless, expert data curation through tools custom-designed for crowdsourcing has been essential for generating datasets for machine learning and artificial intelligence research in healthcare.[9,10]

For success, such tools must address barriers inherent to development and implementation of a crowdsourcing platform. Table 1 summarizes some of these challenges and how our innovative design overcomes each barrier. Because we used a synthetic dataset instead of actual cancer cases, our institutional review board (IRB) determined the adjudication process to be nonhuman subjects research. We were able to avoid many of the regulatory hurdles common in crowdsourced medical research such as waivers of informed consent, privacy controls, and handling of sensitive data.[9] Instead of developing extensive training materials to ensure experts of different backgrounds applied a uniform approach to rating, we developed our interface with simple instructions. Experts can sign up for an account and start adjudicating cases within minutes. Furthermore, each expert may adjudicate any number of cases with no minimum requirement. Even with just a few cases, each expert still provides information for the ranking algorithm to determine a synthetic case's relative risk. This low barrier of entry facilitates knowledge capture from busy experts across a diverse range of expertise.

RESULTS

Ranking interface and algorithm

The FILTER application has been deployed on the cloud and is available for experts to create an account and adjudicate cases. Figure 1 is a screenshot of the FILTER interface with an example matchup that an expert might adjudicate. For each adjudication, the expert is presented with the question, "Which of the following scenarios requires a higher level of survivorship follow up?"

The expert has the option to choose the case to the right, the case to the left, or rank as equal. After each judgment, the scores for the two cases are adjusted using the Elo rating algorithm.[11]

$$R'_A = R_A + 23(1 - E_A)$$
$$\text{where } E_A = \frac{1}{1 + 10^{(R_{Bi} - R_A)/400}}$$

R_A is the current score and R_{Bi} is the score of opponent.

This algorithm, originally developed to rate chess players, sets an expectation that cases with higher scores will likely "win" against cases with lower scores. With each selection, the "winning" case is increased in point value, and the point value of the "losing" case is decreased. The magnitude of point-value change is dependent on how far apart the scores started, with larger changes in the event of an "upset". We chose the Elo ranking algorithm over others primarily because Elo does not require us to predefine the number of cases that we want to adjudicate, which allows us to add new cases as more matchups are adjudicated. Additionally, the resulting Elo score is parametric, which allows us to consider the magnitude of differences in scores rather than just their order.

Table 1. Summary of crowdsourcing platform barriers and how FILTER addresses these barriers

Crowdsourcing platform barrier	FILTER solution
Experts are limited to a single institution's staff and/or external experts under a data use agreement (DUA). IRB approval and oversight can delay data collection.	Use of synthetic cases instead of real patient information allows a large number of adjudicators from many institutions without IRB approvals or DUAs.
Due to experts of different types and levels of expertise, training participants on uniform criteria and approach for adjudication is difficult and can lead to low inter-rater agreement.	Use of a self-explanatory interface with simple instructions for adjudication obviates the need for training.
Each expert must adjudicate a minimum number of cases to ensure sufficient overlap for assessment of inter-rater reliability. This can delay the project as experts tend to be busy.	Experts can adjudicate as many or as few cases as they wish. Even with just a few cases, each expert provides information useful to determine risk.
Due to the complexity of clinical cases, accurate assessment of risk on an absolute scale may present challenges.	Requiring only one-to-one comparison of relative risk is easier to understand and judge.

FILTER: follow-up interactive long-term expert ranking.

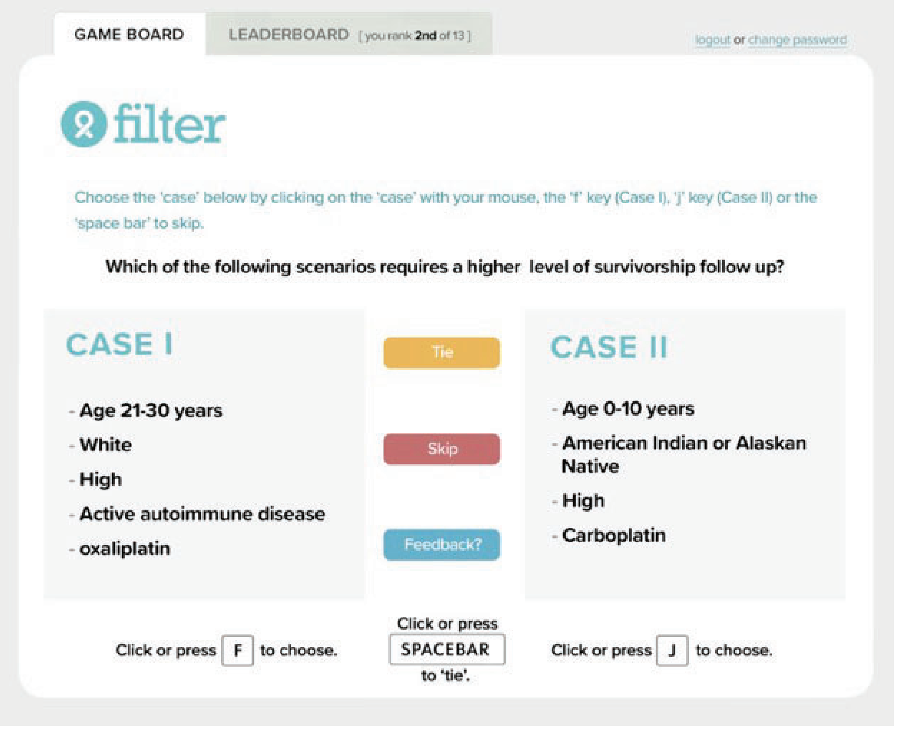

Figure 1. Expert adjudicator FILTER interface.

New case creation and matchup algorithm

Because neither the number of experts nor the number of adjudications per expert is prespecified, we designed FILTER to dynamically generate new synthetic cases whenever a sufficient number of matchups have occurred. Each new case is randomly assigned risk factors from each domain in Table 2. This list was generated by authors (TO, DF, and TP) who are experts in oncology survivorship and genetic risk factors in cancer. Our matchup algorithm ensures new cases matched to enough existing cases to establish a starting rank. Existing cases are also periodically rematched against one another to reconfirm their place in the ranking.

Since risk factors for each synthetic case are selected randomly, it is possible that clinically unlikely combinations may occur. We considered creating a list of such combinations and eliminating them from possible synthetic cases. However, for almost every unlikely combination, we were able to come up with an edge case where that combination might occur. Our solution was to instruct experts to make their best determination of risk based on the synthetic case even if the combination of treatments would be impossible. We included a disclaimer whenever an expert logged in that stated that synthetic cases were generated randomly, and that many combinations would not be realistic. We also informed experts that the goal of the process was to determine the contribution of each factor to overall risk independently.

Figure 2 illustrates the sequence of matchups starting from new case creation. In Phase 1, a new case is matched against roundup [log2(n)] existing cases selected at random, where n is the total number of existing cases. In Phase 2, each case matched to the new case is matched to roundup [log2(n)] other existing cases, selected at random. In each matchup both cases involved are score-adjusted immediately, according to the Elo formula. After all Phase 2 matchups have occurred, the newly added case is considered an existing case. Another new case is added, and the process repeats. We designed this algorithm so that matchups would

Table 2. Survivorship risk factors by domain

Domain	Risk factor
Surgery	
	Breast resection
	Lung resection
	Kidney resection
	Colon resection
	Small intestine resection
	Extremity resection
	Pancreas resection
	Liver resection
	Brain resection
	Larynx resection
	Esophagus resection
	Lymph node resection
	Testicle resection
	Ovary resection
	Uterus resection
	Bladder resection
	Prostate resection
	Breast removal
	Lung removal
	Kidney removal
	Colon removal
	Small intestine removal
	Extremity removal
	Pancreas removal
	Liver removal
	Larynx removal
	Esophagus removal
	Lymph node removal
	Testicle removal
	Ovary removal
	Uterus removal
	Bladder removal
	Prostate removal
	Stomach removal
	Thyroid removal
Immune modulation	
	Allogeneic transplant (CyTBI conditioning)
	Allogeneic transplant (BuCy conditioning)
	Allogeneic transplant (BuFlu conditioning)
	CAR-T cell therapy
Comorbidity	
	Active autoimmune disease
	Traumatic brain injury
	Congestive heart failure (CHF)
	COPD or obstructive airway disease
	Renal failure
	Obesity
	Tobacco use
	Substance abuse
	Developmental delay
	Hepatic impairment
	Hypertension
	Psychiatric illness
	Neuropathy
	Stroke
Radiation	
	Radiation to the breast
	Radiation to the lung
	Radiation to the kidney
	Radiation to the colon
	Radiation to the small intestine
	Radiation to the extremity
	Radiation to the pancreas
	Radiation to the liver
	Radiation to the brain
	Radiation to the larynx
	Radiation to the esophagus
	Radiation to the lymph node
	Radiation to the testicle
	Radiation to the ovary
	Radiation to the uterus
	Radiation to the bladder
	Radiation to the prostate
	Radiation to the neck
	Radiation to the stomach
Systemic drug	
	Anthracyline (like adriamycin)
	Vinca alkaloid (like vincristine)
	Tumor antibiotic (like bleomycin)
	Alkylating agent (like cyclophosphamide)
	Cisplatin
	Carboplatin
	Oxaliplatin
	Microtubule inhibitor (like paclitaxel)
	Immunotherapy (like pembroluzimab)
	Monocloncal antibody (like blinatumomab)
	Tetrahydrofolate reductase inhibitor (like pemetrexed)
	Corticosteroids
	Antimetabolites (like mercaptopurine or cytarabine)
	Topoisomerase I inhibitor (like topotecan)
	Topoisomerase II inhibitor (like etoposide)
Genetic risks	
	Multiple close family members with cancer
	Inherited cancer gene mutation (eg, BRCA, Lynch) identified
	Increased risk of treatment toxicity due to inherited gene mutation
	Multiple primary cancers of paired organs or different organs
Age (years)	
	0–10
	11–20
	21–30
	31–40
	40–65
	65+
Socioeconomic status	
	Low
	Medium
	High

be well distributed among cases. Additionally, we hold one case the same in a series of matchups (the new case in Phase 1 and subsequently matched cases in Phase 2) so that both cases do not change from matchup to matchup. We believe this will assist the expert cognitively to improve the ease and speed of adjudicating cases.

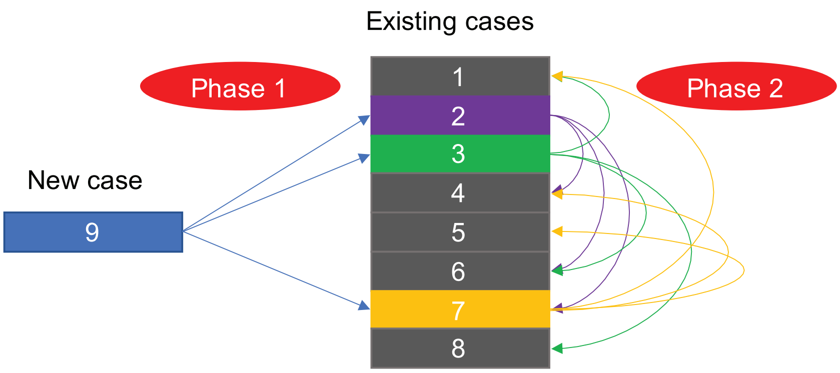

Figure 2. Matchup algorithm following the addition of a new synthetic case "9" with eight existing cases.

Incentives

Participation of experts is incentivized with gift cards for the three experts who adjudicate the most cases. When logged into FILTER, each expert can see his/her own case count, and the counts for the top adjudicators, on the leaderboard. Each expert is required to affiliate with an institution at the time of account creation. This information is used to display an institution leaderboard to encourage friendly competition among groups of experts.

Initial testing

We have done initial testing with a group of 13 Vanderbilt-Ingram Cancer Center oncologists. These oncologists have adjudicated 1174 matchups for 64 cases. In the next phase of FILTER implementation, we plan to invite members of the National Comprehensive Cancer Network (NCCN) Survivorship Guidelines Panel to participate as experts.[12] These individuals will be authenticated through their institutions' email addresses.

DISCUSSION

We have created an application with a low barrier of entry to obtain expert adjudication of risk. Although we have designed FILTER for risk of clinical complications in cancer survivorship, the platform is generalizable to other medical use cases that require risk or severity scores generated through crowdsourcing. As a crowdsourcing platform, FILTER is unique and powerful because it does not use real patient data, it does not require much instruction for experts to use, and it does not prescribe a minimum input for the contribution of each expert to be considered complete.

FILTER has several limitations. Its ranking algorithm is limited to adjudicating a single ordinal or continuous scale. This precludes FILTER's use to identify individual, disease, or treatment phenotypes, a common use case for crowdsourcing in cancer research. As experts start to use FILTER, other limitations may emerge. Given our invitation to a wide range of physician participants, inter-rater disagreement may arise due to differences in opinion based on training or background. In addition, "bad actors" may enter purposefully wrong or random information. We believe that, as with other crowdsourcing platforms, these limitations can be overcome by having a large number of adjudicators so that the wisdom of the crowd is captured.

After adjudication of a sufficient number of cases, our next step is to use the risk scores as outcomes in a regression model that will ascertain the contribution of each factor to survivorship risk. The end result will be an online tool that calculates survivorship risk based on the risk factors in Table 2. One limitation to determining how many experts we must engage is that there is no prior data to determine the extent to which experts will disagree on levels of risk. Part of what we will assess in pilot testing with the NCCN survivorship guidelines panel is expert agreement and risk score variability. We estimate that there will need to be at least 870 synthetic cases to obtain a reliable regression model. Assuming FILTER's matchup algorithm effectively rates those 870 cases, we would need 77 266 matchups adjudicated. Therefore, we anticipate that there must be 772 experts adjudicating an average of 100 matchups each to get a reliable model.

CONCLUSION

The FILTER crowdsourcing platform addresses a critical need for capturing clinical knowledge from experts when real-world data are scarce. Results from data obtained using FILTER will allow oncologists to better assess patient need for cancer survivorship follow-up care, thus allowing healthcare systems to allocate resources and services according to need.

FUNDING

This work was supported by the National Cancer Institute grant numbers R01 CA240093-02S1 and P30 CA068485.

AUTHOR CONTRIBUTIONS

TO conceived of the crowdsourcing platform and use of the Elo rating algorithm. TO, TP, and DF compiled the list of risk factors. LW, AC, and YL developed the platform. TK created the risk model and LB built the risk calculator. All authors contributed to the final manuscript.

CONFLICT OF INTEREST STATEMENT

None declared.

DATA AVAILABILITY

No new data were generated or analyzed in support of this research.

REFERENCES

1. Bluethmann SM, Mariotto AB, Rowland JH. Anticipating the 'Silver Tsunami': prevalence trajectories and comorbidity burden among older cancer survivors in the United States. *Cancer Epidemiol Biomarkers Prev* 2016; 25: 1029–36.
2. Carver JR, Shapiro CL, Ng A, *et al*. American Society of Clinical Oncology clinical evidence review on the ongoing care of adult cancer survivors: cardiac and pulmonary late effects. *J Clin Oncol* 2007; 25: 3991–4008.
3. Lee M, Salloum RG. Racial and ethnic disparities in cost-related medication non-adherence among cancer survivors. *J Cancer Surviv* 2016; 10: 534–44.
4. Pisu M, Kenzik KM, Oster RA, *et al*. Economic hardship of minority and non-minority cancer survivors 1 year after diagnosis: another long-term effect of cancer? Economic hardship in cancer survivors. *Cancer* 2015; 121: 1257–64.
5. Oeffinger KC, McCabe MS. Models for delivering survivorship care. *J Clin Oncol* 2006; 24: 5117–24.
6. Nekhlyudov L, O'malley DM, Hudson SV. Integrating primary care providers in the care of cancer survivors: gaps in evidence and future opportunities. *Lancet Oncol* 2017; 18: e30–8.
7. Jefford M, Rowland J, Grunfeld E, *et al*. Implementing improved post-treatment care for cancer survivors in England, with reflections from Australia, Canada and the USA. *Br J Cancer* 2013; 108: 14–20.
8. Créquit P, Mansouri G, Benchoufi M, *et al*. Mapping of crowdsourcing in health: systematic review. *J Med Internet Res* 2018; 20: e187.
9. Ye C, Coco J, Epishova A, *et al*. A crowdsourcing framework for medical data sets. *AMIA Jt Summits Transl Sci Proc* 2018; 2017: 273–80.
10. He X, Zhang H, Bian J. User-centered design of a web-based crowdsourcing-integrated semantic text annotation tool for building a mental health knowledge base. *J Biomed Inform* 2020; 110: 103571.
11. Lehmann R, Wohlrabe K. Who is the 'Journal Grand Master'? A new ranking based on the Elo rating system. *J Informetr* 2017; 11: 800–9.
12. NCCN Survivorship Guidelines Panel. https://www.nccn.org/guidelines/guidelines-panels-and-disclosure/disclosure-panels-details-page?MeetingId=0&GroupId=652 Accessed September 11, 2021.