# Learnings From Precision Clinical Trial Matching for Oncology Patients Who Received NGS Testing

Neha M. Jain, PhD[1]; Alison Culley, RN, BSN[2]; Christine M. Micheel, PhD[3]; Travis J. Osterman, MS, DO[3,4]; and Mia A. Levy, MD, PhD[4,5,6]

*original reports*

*abstract*

**PURPOSE** Tumor next-generation sequencing reports typically generate trial recommendations for patients based on their diagnosis and genomic profile. However, these require additional refinement and prescreening, which can add to physician burden. We wanted to use human prescreening efforts to efficiently refine these trial options and also elucidate the high-value parameters that have a major impact on efficient trial matching.

**METHODS** Clinical trial recommendations were generated based on diagnosis and biomarker criteria using an informatics platform and were further refined by manual prescreening. The refined results were then compared with the initial trial recommendations and the reasons for false-positive matches were evaluated.

**RESULTS** Manual prescreening significantly reduced the number of false positives from the informatics generated trial recommendations, as expected. We found that trial-specific criteria, especially recruiting status for individual trial arms, were a high value parameter and led to the largest number of automated false-positive matches.

**CONCLUSION** Reflex clinical trial matching approaches that refine trial recommendations based on the clinical details as well as trial-specific criteria have the potential to help alleviate physician burden for selecting the most appropriate trial for their patient. Investing in publicly available resources that capture the recruiting status of a trial at the cohort or arm level would, therefore, allow us to make meaningful contributions to increase the clinical trial enrollments by eliminating false positives.

## INTRODUCTION

Oncology clinicians are using an increasing number of tumor next-generation sequencing (NGS) tests to help determine treatment options for patients with cancer, including for use of standard-of-care targeted therapies, enrollment on clinical trials, and use of off-label treatments. With increased testing and the growing size of panels, more variants are being uncovered. A recent study showed that clinical trial participation in oncology is only about 8%.[1] The impact of low accrual rates on biomarker-driven trials is severe, since the target patient population for these trials is typically small.[2] This, combined with the low accrual rates for adult oncology trials, may delay or prevent targeted drug from being approved and reaching the market.

A typical tumor NGS report may contain recommendations for clinical trials based on the cancer diagnosis and genomic mutation profile of the patient. Many clinical trial matching algorithms used by laboratories that report tumor NGS results favor high sensitivity at the cost of poor positive predictive value. The onus,

therefore, lies on the managing physician or the clinical research staff to review and eliminate false-positive trials to find the best match for the patient. This demands a tremendous amount of time, human effort, and clinical expertise as eligibility criteria may include diagnosis, tumor markers, disease state, prior line of therapies, tumor genomic profile, and several other criteria. According to one estimate, research nurses spend about 4-9 hours evaluating a patient for clinical trial eligibility.[3]

This issue is further complicated by the difficulty in correctly predicting whether a particular trial or arms of a clinical trial is open at a given institution at the particular time. The trial recruiting status reported on publicly available resources such as ClinicalTrials.gov[4] often serve as the source of information for trials listed in NGS reports. Jones et al[5] found that the trial status reflected on the ClinicalTrials.gov could be lagging behind by up to 7 months. Furthermore, the clinical trial status on publicly available resources only reflects the overall open or closed status of the trial and not of all individual sites and/or arms of the trial. This can

**CONTEXT**

**Key Objective**

How accurate are automated algorithms for clinical trial matching based on diagnosis and biomarker criteria?

**Knowledge Generated**

Clinical trial recommendations based on diagnosis and biomarker criteria had a high rate of false positives (approximately 88%). Outdated information about open sites for specific clinical trial arms and trial slots was the biggest factor for false positives.

**Relevance**

Focusing on developing infrastructure that can support curation of trial metadata across various sites and host it on publicly available domains is urgently needed to improve the accuracy of trial matching algorithms.

have a significant impact on trial recommendations listed in NGS reports and could hypothetically increase false positives.

As the cost of sequencing decreases and the clinical focus moves from limited gene-panel testing to whole-exon sequencing, RNA sequencing, and ultimately to whole-genome sequencing, the pool of potentially matching trials—that requires manual review—will expand. Informatics tools that assist in refining the trials that need to be manually reviewed or suggest potential matches based on the patient's clinical and biomarker profile may reduce the manual burden of this process and reduce the barrier to trial enrollment.

There have been efforts led by private and commercial organizations to address this issue. Some of these efforts are MatchMiner,[6] IBM Watson,[7] Personalized Cancer Therapy,[8] MolecularMatch,[9] and Trial Prospector.[10] However, there have been limited studies outlining the real-world implementation of these tools in the setting of a large cancer center. As NGS testing becomes standard of care for several cancers, it is important to incorporate clinical decision support pertaining to clinical trial matching as part of the clinical workflow. Pilot feasibility trials that mimic the real-world challenges of implementing such interventions will go a long way in supporting uptake of these tools across the oncology community.

In this paper, we share our experiences and challenges in evaluating augmented clinical trial matching at the Vanderbilt-Ingram Cancer Center (VICC), an NCI-designated Comprehensive Cancer Center in Nashville, TN. *Augmented* trial matching refers to the addition of human review to refine clinical trial results following the output of the automated clinical trial matching algorithm. The precision clinical trial matching (PCTM) service at VICC was developed in collaboration with our software development partner GenomOncology, LLC, and trial results were augmented by prescreening performed by a research nurse. An iterative design philosophy was followed to rapidly identify issues and modify efforts accordingly to minimize false positives. This paper describes the study design and learnings from a clinical trial matching study at a tertiary cancer center.

## METHODS

### Study Design

Figure 1 shows the schematic for the prospective study. All patients with a solid tumor diagnosis who received a new NGS test result were automatically included in the study. A new tumor NGS result triggered the generation of clinical trial recommendations by the PCTM informatics platform based on the diagnosis and biomarkers reported on the NGS test. A research nurse manually reviewed these clinical trial recommendations and performed an initial prescreen to determine eligibility. The refined results were compared against the initial number for trials to determine the added value of augmented matching, ie, manual prescreening, and to elucidate the reasons for false-positive trial matches. The study was approved by the institutional investigational review board (IRB# 171809), which determined that neither the physicians nor the patients needed to be consented for this pilot feasibility decision support study.

### Data Flow for Study

Extensible markup language (XML) files from NGS tests were directly consumed by PCTM's proprietary software codeveloped with GenomOncology, and relevant demographic and biomarker information was parsed and structured. PCTM was directly interfaced with the VICC OnCore[11] clinical trial management system (CTMS) and received nightly refresh of all active trials that were recruiting patients (Fig 2). PCTM also interfaced with the My Cancer Genome clinical trial model[12] and showed the structured curated data for the relevant trials based on the OnCore results. The structured curated data included trial eligibility groups based on disease and biomarkers as well as the trial treatment contexts, which included the treatment setting (ie, neoadjuvant, adjuvant, and metastatic) as well as the drugs used for each trial cohort or arm. This model has been described in more detail elsewhere.[12]
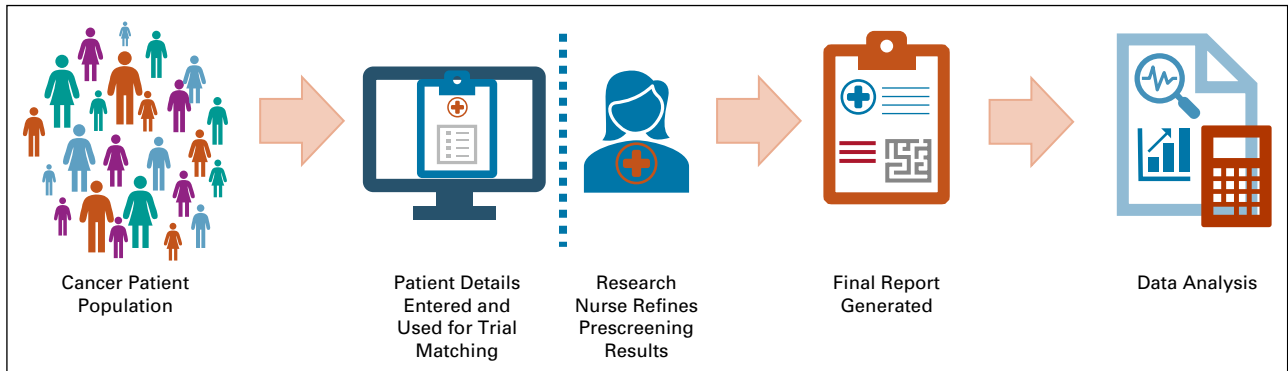
**FIG 1.** Schematic outlining the study workflow. Patients were identified based on the receipt of an NGS report, and their clinical reports were uploaded using a web-based user interface to PCTM. Relevant patient details such as vital status, diagnosis, date of diagnosis, and other biomarkers tested outside of the NGS test (eg, PD-L1 and IHC for ER or PR or HER2) were added manually. PCTM then generated a list of potential trials for each patient based on diagnosis and biomarker profile. Additional refinements to these trial results were applied using built-in multifaceted filtering for treatment setting and trial phase. Furthermore, a research nurse performed an initial manual prescreening to evaluate eligibility. The final list of eligible trials was compared against the initial recommendations by the PCTM software. ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; IHC, immunohistochemistry; NGS, next-generation sequencing; PCTM, precision clinical trial matching; PD-L1, programmed death-ligand 1; PR, progesterone receptor.

Based on these three data streams, PCTM presented a set of potentially eligible institutional as well as national clinical trials for each patient. The research nurse had the ability to augment or edit a patient's case with more granular clinical details—such as protein expression detected through an immunohistochemistry (IHC) report, a report from an outside institution that mentions a separate comorbidity, among other examples.

## Study Protocol

Briefly, PCTM was used to give an initial list of trials for all eligible patients. The following eligibility criteria were used to define the patient population: (1) NGS test resulted from October 2018 to April 2019; (2) was not deceased or on hospice as documented in electronic medical record (EMR); (3) had at least one appointment with a medical oncologist, neuro-oncologist, or gynecologic oncologists. Surgical oncologists were not part of the study. PCTM trial results were refined by a research nurse by using a set of customized filters. These filters allowed trial filtering based on trial phase, recruiting status, and treatment context (ie, neoadjuvant, adjuvant, and metastatic). Finally, an initial prescreening was performed and trials were classified as either (1) a *potential match* (patient is eligible for the trial at the time of review), (2) a *future match* (patient may be eligible for the trial in the future), (3) a *partial match* (patient seems to be eligible for the trial at the time of review but further information is needed to confirm eligibility), or (4) a *no match* (patient is ineligible for a trial and this ineligibility is unlikely to change in the future). We have described these concepts in more detail elsewhere.[13] The final shortlisted trials included the *potential match* and *future match* trials. False-positive rate referred to the number of no match trials compared with the overall trials shortlisted by PCTM.

## Data Collection

A secure REDCap form[14] with repeatable forms was designed for this study. Initial trial matches, refined trial results after expert review and prescreening, and reasons for ineligibility were recorded manually in the REDCap



**FIG 2.** Data flow in PCTM. This schematic shows how the PCTM received data feeds from three different sources: (1) the curated clinical trial data from My Cancer Genome clinical trial data model, (2) the NGS data feed in the form of XML files from VICC, and (3) updated and refreshed clinical trial metadata from the CTMS. CTMS, clinical trial management system; NGS, next-generation sequencing; PCTM, precision clinical trial matching.
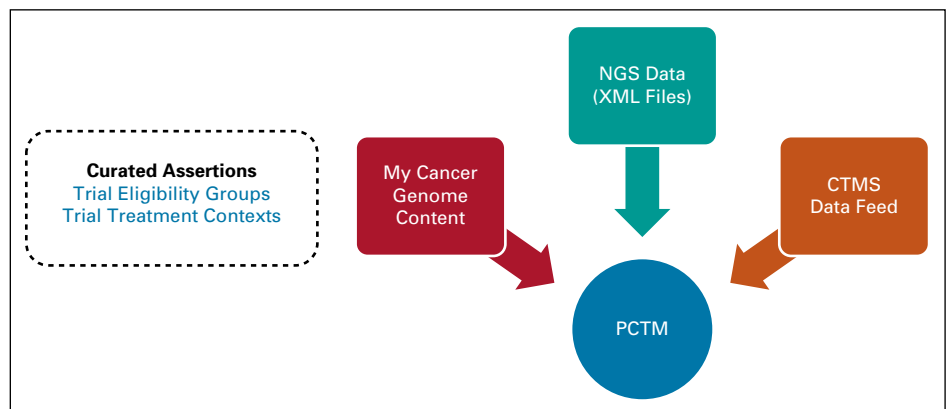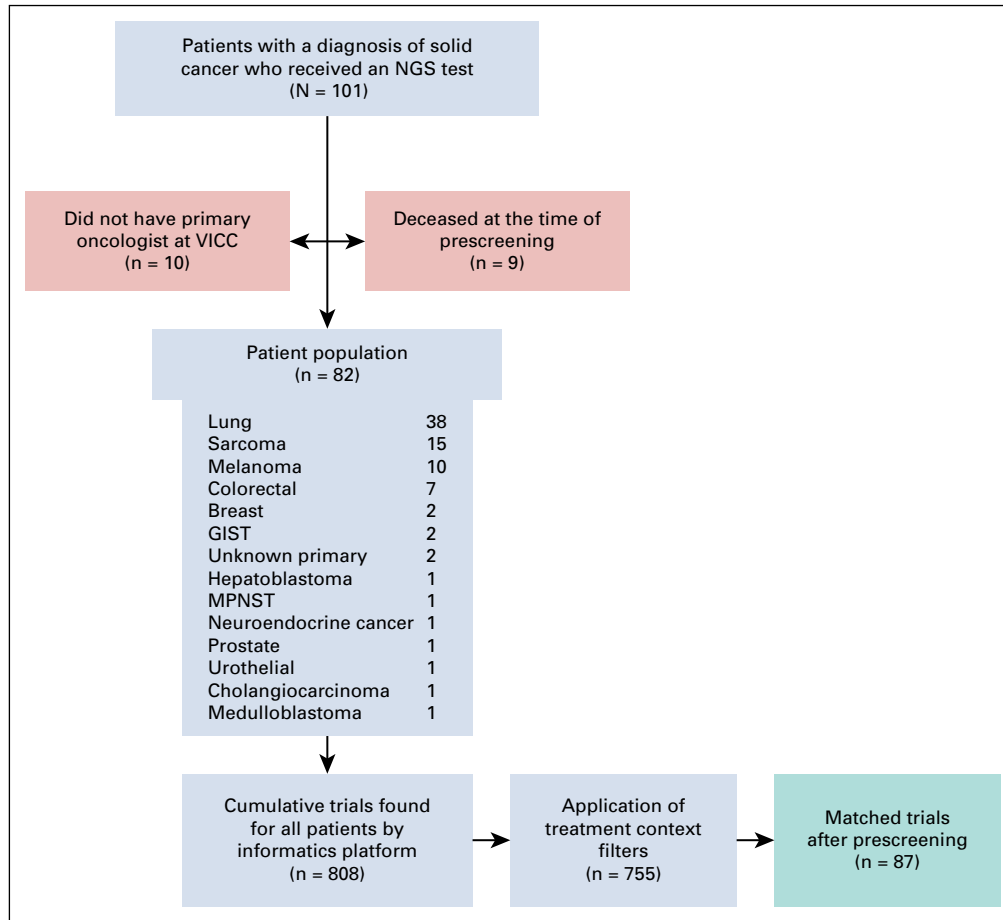
**FIG 3.** Cohort CONSORT diagram. A total of 101 patients had a solid cancer diagnosis and received a qualifying NGS test within the study period (October 2018-April 2019). At the time of prescreening, nine were found to be deceased and 10 patients did not have a primary oncologist at VICC and therefore were removed from the study. The remaining 82 patients had a corresponding oncologist and were entered into the study. Cancer diagnosis distribution can be seen in the gray box. A total of 808 cumulative trials were designated as matched trials by the informatics platform. These were reduced to 755 after the application of treatment context filters and ultimately to 87 matched trials after manual prescreening (shown in green box). All data are cumulative. NGS, next-generation sequencing.

database in a retrospective fashion. Data were exported out of REDCap as a CSV file at the end of study for further analysis.

## RESULTS

### Cohort Characteristics

The characteristics of the patient cohort have been detailed in Figure 3. Clinical trial recommendations were generated for 82 patients. About 9% of the patient population was deceased at the time of prescreening and another 10% received NGS testing here but were receiving primary cancer care elsewhere. Clinical trial recommendations were not generated for these patients (n = 19) and they were excluded from the analysis.

### Trial Recommendations

A total of 808 trials were designated as matches by the PCTM software (cumulative data). After the application of treatment context filters, this number dropped to 755 trials. These 755 trials were manually prescreened by the research nurse based on the clinical details of the patient and real-time clinical trial recruiting status data for individual trial cohorts.

### Prescreening Outcomes

We observed that approximately 88% of the trials that initially matched based on diagnosis and biomarker criteria of a patient were eventually classified as a no match and were therefore false positives (Fig 4). Less than 2% of trials were found to be potential matches for the patient's current clinical state, whereas 10% of trials were found to be a potential match for future lines of therapy.

A deeper analysis into the reasons for no match revealed that 72% of the false-positive matches occurred when a patient matched to a particular arm of a clinical trial that was either closed to accrual, did not have available slots,
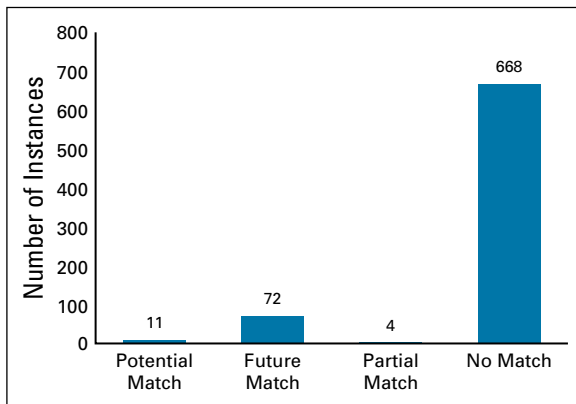
**FIG 4.** Results from the pilot feasibility study. The prescreen outcomes based on the manual review of the suggested matches from PCTM. A high number of trials (88%) that initially matched to the patient based on diagnosis and biomarkers were eventually found to be false matches (cumulative patient data). PCTM, precision clinical trial matching.

was never opened at VICC, or was suspended while other arms of the trial were still open to accrual. The status of individual arms of the clinical trial is not maintained in publicly available resources and is inconsistently maintained in the local CTMS. Figure 5A shows all the reasons for trials qualifying as a false positive despite matching on diagnosis and biomarkers. The blue bars represent failure to match because of clinical criteria, whereas the red bars represent failure to match because of inaccurate or incomplete trial or trial arm statuses. Trial-related clinical criteria include prior treatments, disease histology, biomarker criteria, performance status, treatment setting or context, cardiac abnormalities, presence or history of brain metastases, presence of progressive disease, prior malignancies, HIV status, and age-related criteria, among others. Trial and trial arm-related criteria include trial open or close status, trial arm open or close or suspension status, and slot availability in an arm of interest.

Figure 5B shows the cumulative data for incremental refinement of selected trials with application of selected filters—disease, biomarker, treatment context (ie, neoadjuvant, adjuvant, and metastatic), trial-cohort open or closed status, and trial-cohort slot availability. Only approximately 12% of all the trials that matched initially based on diagnosis and biomarker were eventually found to be accurate matches (a potential or a future match). The blue bars indicate criteria for which a filter exists in PCTM, and the green bars represent an existing gap in the process. It is noteworthy that some trials were a no match based on multiple criteria. In some instances, these were related to both the trial-related eligibility criteria, eg, HIV status and multiple malignancies, as well as administrative details pertaining to the trial, eg, no slots available on the cohort of interest and cohort never opened.

## DISCUSSION

In this study, we aimed to assess the stepwise quantitative refinement of clinical trial recommendations when trial results from an informatics platform were augmented by manual prescreening. While doing so, we also wanted to elucidate factors that have the highest impact on the matching efficiency and thereby inform future clinical trial matching decision support efforts.

Administrative details of clinical trials, both at the institutional and global levels, were the single-most important reason for false-positive trial matches. Institutional-level administrative features included trial recruiting status, trial open or close status, and trial-arm open or close or suspended status. Global administrative features included data that might not be readily available to the local institution in real time; this primarily includes slot availability for phase I clinical trials. The most important result of this study was the impact of these administrative clinical trial features on the false-positive rate for clinical trial matching. Although institutional-level administrative features can be addressed with improved manual upkeep of data or automated efforts, addressing global administrative features would require support from multiple stakeholders and potentially redesigning clinical trial data infrastructures and/or policies.

Interfacing institutional trial matching services with the institution's internal CTMS can give a more accurate picture about the recruiting statuses of clinical trials that are open at local institutions. It is important to note that the recruiting status at the institutional level may be different than that listed on ClinicalTrials.gov or other similar public databases because of update delays.[5] Similarly, information related to an individual arm of a trial having an open, closed, or suspended status can be recorded in a structured fashion in existing CTMS. A more robust approach of assigning a unique identifier to each trial arm in addition to the trial may be pivotal in teasing apart the recruiting statuses among trial arms. We discovered that existing CTMS may already be capable of recording these criteria. Since these fields are not consistently used for downstream uses, these details are rarely recorded. Educating clinical trial administrative staff about the potential downstream uses of the data encoded in the CTMS can enhance the quality of data related to recruiting statuses and thereby improve matching efficiency.

To obtain accurate and updated information about slot availability across multiple sites, an orchestrated effort would be required across trial sponsors and clinical trial research staff from one or more participating sites. Slot availability is an inherently challenging variable to track because of its temporal nature—slots often rotate between all open sites—and the window of enrollment on these slots can be < 24 hours. There are no existing resources that have a consolidated and aggregated view of this information across multiple sites. The key is that information needs to
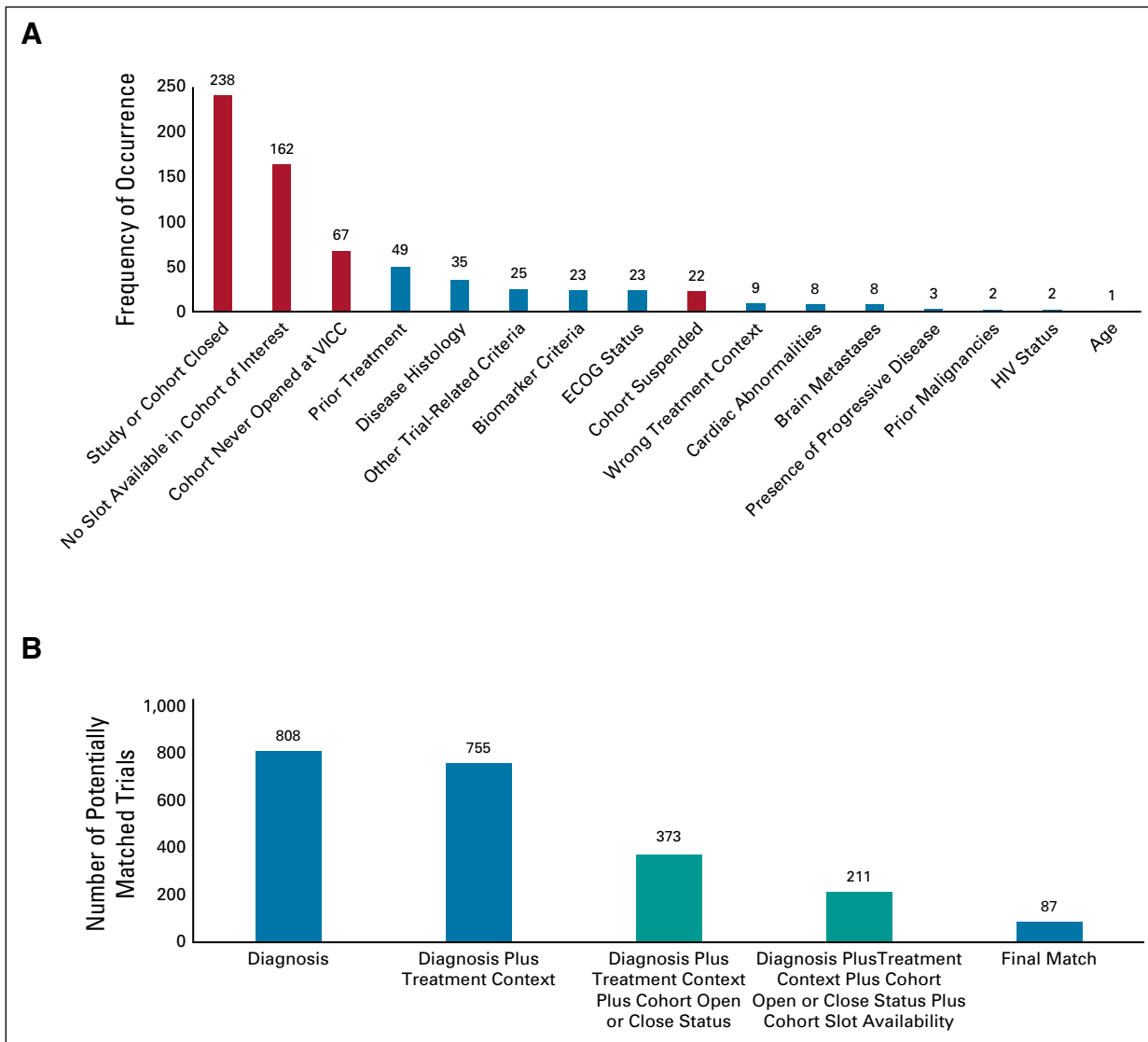
**A**



**B**



**FIG 5.** Results from the pilot feasibility study. (A) The reasons that led trials to qualify as no match. The red bars represent trial-specific criteria that are not currently curated and require manual review, whereas the blue bars represent patient-specific trial criteria that led the trial to be a false positive. (B) The incremental benefit seen in the number of potentially matched trials by refining specific trial criteria. Initial PCTM recommendations without any filters or manual refinement showed a total of 808 matched trials (cumulative data) for all patients. On applying the appropriate built-in treatment context filter, this number dropped to 755. Manual review by a research nurse was then used to further reduce this number as follows: (1) manual evaluation of actual trial arm status (open v closed) reduced this number to 373; (2) manual review to verify if the trials (mostly phase I and multidisease trials) had open slots reduced the number to 211; (3) prescreening of these trials reduced the number to the final 87 trials. These 87 trials were found to be a match (potential, partial, or future). The blue bars indicate criterion for which a filter exists in PCTM, and the green bars represent an existing gap in the process that had to be bridged by manual review. The data are cumulative and across the entire cohort of patients. PCTM, precision clinical trial matching.

be centrally managed and accessible for updates to and use by multiple sites in real time. This needs to be further supported by enabling a seamless data architecture between CTMS, matching services, and central reporting resources. Built-in process checks that can detect data interruptions and other anomalies will support maintenance of a robust and agile system.

This study also confirms previously reported inconsistencies between the actual recruiting status of individual arms of multiarm trials with that published on publicly available resources.[5,15,16] We observed an astonishingly high false-positive rate (88%) of clinical trial matches when they were exclusively based on diagnosis and biomarker criteria. And yet, it is not uncommon to see these two criteria being used by NGS vendors and trial-matching algorithms to predict potential trial matches for patients. Gathering updated trial-arm information from local institutions and using that to filter down trial results on NGS

reports would minimize false positives and reduce pre-screening efforts.

In addition to outdated and incomplete recruiting statuses, the presence of multiple malignancies also significantly reduced patient matching. The two patients in our cohort with multiple malignancies were excluded from all trials. The NCI recently accepted recommendations from ASCO and Friends of Cancer Research that allow patients with multiple malignancies to enroll on trials,[17] but these have not yet been widely adopted.

About 9% of the patients in our patient cohort were found to be deceased at the time of prescreening (1-2 months after the NGS report was obtained). The vital status of some of these patients was outdated in the EMR, which led to futile prescreening efforts. Outdated vital status can negatively affect efforts for clinical trial matching by holding waitlist spots that may result in enrollment delays and subsequent loss of trial enrollment opportunity. This increase in mortality was not surprising since patients who received NGS tests at VICC typically had advanced disease and may already have progressed on multiple lines of therapy. This highlights the importance of timeliness of trial matching efforts. Ideally, these efforts should be performed in real time. Recommendations not sent in real time can still be helpful to guide future lines of therapy, but they run the risk of becoming outdated because of change in patient's disease status or other clinical criteria. We have discussed the various process triggers that can be designed to kickstart reflex clinical trial matching elsewhere.[13] Such efforts take the onus off the clinical trial staff and the physician and could result in increased treatment options for patients as well as higher rates of trial enrollment.

In conclusion, it is important to uncover the factors that result in a high rate of false positives while generating clinical trial recommendations. Such studies can help to drive efficiency and improve the design for future trial matching efforts. Furthermore, such studies should be designed such that trial recommendations are sent to physicians in real time to maximize patient impact and generate enough learnings to inform design of a larger pragmatic study.

It is not possible to make meaningful contributions to increase the clinical trial enrollments without investing in publicly available resources that capture the recruiting status of a trial at the cohort or arm level. This can be achieved by enhanced reporting of statuses of arms on ClinicalTrials.gov and Cancer.gov, as well as by upgrading current CTMS to support curation of recruiting statuses at the level of arms. Efforts—both policy-based improvements and technological advancements—of this magnitude can benefit the entire oncology community and can be taken up by the NLM, NCI, and/or commercial vendors since it requires multi-institutional collaboration.

## AFFILIATIONS
[1]Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, TN
[2]Vanderbilt-Ingram Cancer Center, Clinical Trial Shared Resource, Vanderbilt University Medical Center, Nashville, TN
[3]Department of Medicine, Division of Hematology/Oncology, Vanderbilt University Medical Center, Nashville, TN
[4]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN
[5]Department of Internal Medicine, Division of Hematology/Oncology, Rush University Medical Center, Chicago, IL
[6]Rush University Cancer Center, Rush University Medical Center, Chicago, IL

## CORRESPONDING AUTHOR
Neha M. Jain, PhD, Vanderbilt-Ingram Cancer Center, 2525 West End Ave, Suite 1582, Nashville, TN 37232-6838; Twitter: @nehajain77; @VUMC_Cancer; e-mail: neha.jain@vumc.org.

## AUTHOR CONTRIBUTIONS
**Conception and design:** Neha M. Jain, Christine M. Micheel, Travis J. Osterman, Mia A. Levy
**Financial support:** Mia A. Levy
**Administrative support:** Neha M. Jain, Mia A. Levy
**Provision of study materials or patients:** Mia A. Levy
**Collection and assembly of data:** Neha M. Jain, Alison Culley, Christine M. Micheel, Mia A. Levy
**Data analysis and interpretation:** Neha M. Jain, Christine M. Micheel, Travis J. Osterman, Mia A. Levy
**Manuscript writing:** All authors
**Final approval of manuscript:** All authors
**Accountable for all aspects of the work:** All authors

## AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST
The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/cci/author-center.

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians (Open Payments).

**Neha M. Jain**
**Research Funding:** GE Healthcare

**Christine M. Micheel**
**Consulting or Advisory Role:** Roche
**Research Funding:** GenomOncology, GE Healthcare

**Travis J. Osterman**
**Stock and Other Ownership Interests:** Infostratix

## REFERENCES

1. JM Unger, R Vaidya, DL Hershman, et al: Systematic review and meta-analysis of the magnitude of structural, clinical, and physician and patient barriers to cancer clinical trial participation. J Natl Cancer Inst 111:245–255, 2019
2. SH Shin, AM Bode, Z Dong: Addressing the challenges of applying precision oncology. NPJ Precis Oncol 1:28, 2017
3. LT Penberthy, BA Dahman, VI Petkov, et al: Effort required in eligibility screening for clinical trials. J Oncol Pract 8:365–370, 2012
4. Home—ClinicalTrials.gov: https://clinicaltrials.gov/
5. CW Jones, MR Safferman, AC Adams, et al: Discrepancies between ClinicalTrials.gov recruitment status and actual trial status: A cross-sectional analysis. BMJ Open 7:e017719, 2017
6. MatchMiner—An open source computational platform for matching patient-specific genomic profiles to precision cancer medicine clinical trials. https://matchminer.org/
7. IBM Watson Oncology Clinical Trial Matching—Overview—United States: https://www.ibm.com/us-en/marketplace/clinical-trial-matching-oncology
8. Knowledge Base for Precision Oncology: https://pct.mdanderson.org/
9. Clinical Intelligence: Precision Oncology. Clinical Decision Support. https://www.molecularmatch.com/
10. SS Sahoo, Tao S, Parchman A, et al: Trial prospector: Matching patients with cancer research studies using an automated and scalable approach. Cancer Inform 13:157–166, 2014
11. OnCore—Login: https://oncore.app.vumc.org/forte-platform-web/login
12. N Jain, Mittendorf KF, Holt M, et al: The My Cancer Genome clinical trial data model and trial curation workflow. J Am Med Inform Assoc 27:1057–1066, 2020
13. NM Jain, A Culley, T Knoop, et al: Conceptual framework to support clinical trial optimization and end-to-end enrollment workflow. JCO Clin Cancer Inform 3:1–10, 2019
14. PA Harris, R Taylor, R Thielke, et al: Research Electronic Data Capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 42:377–381, 2009
15. J Fleminger, B Goldacre: Prevalence of clinical trial status discrepancies: A cross-sectional study of 10,492 trials registered on both ClinicalTrials.gov and the European Union Clinical Trials Register. PLoS One 13:e0193088, 2018
16. D Hartung, DA Zarin, JM Guise, et al: Reporting discrepancies between the ClinicalTrials.gov results database and peer reviewed publications. Ann Intern Med 160:477–483, 2014
17. SM Lichtman, Harvey RD, Damiette Smit MA, et al: Modernizing clinical trial eligibility criteria: Recommendations of the American Society of Clinical Oncology–Friends of Cancer Research Organ Dysfunction, Prior or Concurrent Malignancy, and Comorbidities Working Group. J Clin Oncol 35:3753–3759, 2017