Increasing Clinical Trial Accrual via Automated Matching of Biomarker Criteria*

Jessica W. Chen
Departments of Biomedical Data Science and of Pathology, Stanford University School of Medicine, Stanford, CA 94305, USA
Email: jwrchen@stanford.edu

Christian A. Kunder
Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305, USA
Email: ckunder@stanford.edu

Nam Bui
Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA
Email: nambui@stanford.edu

James L. Zehnder
Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305, USA
Email: zehdner@stanford.edu

Helio A. Costa
Departments of Biomedical Data Science and of Pathology, Stanford University School of Medicine, Stanford, CA 94305, USA
Email: hcosta@stanford.edu

Henning Stehr
Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305, USA
Email: stehr@stanford.edu

Successful implementation of precision oncology requires both the deployment of nucleic acid sequencing panels to identify clinically actionable biomarkers, and the efficient screening of patient biomarker eligibility to on-going clinical trials and therapies. This process is typically performed manually by biocurators, geneticists, pathologists, and oncologists; however, this is a time-intensive, and inconsistent process amongst healthcare providers. We present the development of a feature matching algorithmic pipeline that identifies patients who meet eligibility criteria of precision medicine clinical trials via genetic biomarkers and apply it to patients undergoing treatment at the Stanford Cancer Center. This study demonstrates, through our patient eligibility screening algorithm that leverages clinical sequencing derived biomarkers with precision medicine clinical trials, the successful use of an automated algorithmic pipeline as a feasible, accurate and effective alternative to the traditional manual clinical trial curation.

Keywords: Clinical trials, patient recruitment, clinical decision support

1. INTRODUCTION

* This work is supported by an award from Stanford University Department of Pathology to HAC.

Cancer is the second leading cause of death in the US with 163.5 deaths per 100,000 individuals in 2017 as reported by the National Vital Statistics System[1]. Clinical trials with newly developed cancer therapies are an important avenue for patients with otherwise limited treatment alternatives. It is therefore crucial to identify the most relevant clinical trials for which a patient is eligible. Traditionally, patient eligibility screening is a labor- and time-intensive manual process that is susceptible to errors and missed enrollment as the volume of patients and clinical trials increases[2]. Through the development of bioinformatic tools, we sought to improve the speed and accuracy of patient eligibility screening for precision medicine clinical trials. Precision medicine is an approach to patient care that aims to tailor therapies to the molecular abnormalities rather than the organ site of the tumor[3]. For example, Entrectinib, an inhibitor of tyrosine kinases TRKA/B/C, ROS1 and ALK, is used to treat patients diagnosed with solid tumors that have NTRK1/2/3, ROS1 or ALK gene fusions[4]. There currently does not exist any standardized workflow to perform automated eligibility screen at Stanford Hospital that leverages the existing structured genotyping data from the Solid Tumor Actionable Mutation Panel (STAMP) assay, a targeted next-generation sequencing (NGS) assay for tumor biopsy specimens. In this study, we developed an in-house feature matching algorithmic pipeline that identifies patients who meet eligibility criteria of precision medicine clinical trials.

2. MATERIALS AND METHODS

2.1. *Specimens and Retrospective Analysis*

The patient tissue specimens described in this study were obtained from formalin-fixed paraffin-embedded (FFPE) tissue blocks from Stanford Health Care under institutional review board-approved protocol (IRB-36084). An anatomical pathologist reviewed, diagnosed, and estimated tumor purity from hematoxylin and eosin (H&E) slides of each specimen. These samples underwent targeted NGS using the Stanford Health Care STAMP assay that targets clinically actionable somatic mutations.

STAMP test orders were exported from an internal centralized patient database utilized by Stanford Medicine. For each test order, the pathogenicity statuses of the STAMP-identified variants are annotated by the Molecular Genetic Pathology clinical fellows at Stanford Medicine and merged with patient diagnosis data from the Stanford Anatomic Pathology laboratory. For test orders that identified single nucleotide variations (SNVs) and insertions/deletions (Indels), the following fields were extracted per test order - test order (test order identifier, report created date, report date received, amendment note), patient (gender, date of birth, histological diagnosis), primary tumor site, and biomarker (gene name, pathogenicity status, NM accession identifier, sequence variant HGVS (Human Genome Variation Society) nomenclature - i.e. protein, coding, and genomic). Entries were classified into synonymous, SNV, frameshift (i.e. insertions, deletions, deletions/insertions, duplications), and in-frame (i.e. insertions, deletions, deletions/insertions, duplications) mutations based on the protein and coding sequences; exon location was determined based on the genomic sequence. The entries were manually curated to be consistent with standardized HGVS nomenclature - protein sequence and exon number does not exist for intergenic and intronic mutations. Removed from analysis were variants that are either classified as synonymous mutations or have a pathogenicity status that is missing or classified as benign. For

test orders that identified fusions, the following fields were extracted per test order - test order (test order identifier, report created date, report date received, amendment note), patient (gender, date of birth, histological diagnosis), primary tumor site, and biomarker (gene names). For test orders that identified copy number variations (CNVs), the following fields were extracted per test order - test order (test order identifier, report created date, report date received, amendment note), patient (gender, date of birth, histological diagnosis), primary tumor site, and biomarker (gene name, variant type i.e. amplification or deletion). Entries were classified into disease group categories based on the primary tumor site as manually mapped in (Table S5). Entries with missing primary tumor sites were designated the disease group category of unknown. Patient age, rounded down to the nearest integer, was determined using date of birth and report date received, which if missing, was substituted with the report created date. In addition, entries with missing patient gender or patient date of birth, were removed from analysis. For test orders that are amended (i.e. identical report created dates), the most recent version was used for analysis.

The algorithm incorporates the concept of modularity into its design such that the pipeline may be successfully applied to datasets that are not specific to this study. For example, the patient sample data that is used as input for the clinical trial matching component of the pipeline is a data frame whose features are biomarkers that are commonly used as inclusion and exclusion criteria in precision medicine clinical trials. The modularity provides a flexible algorithmic pipeline that can automate clinical trial matching using data not generated from the STAMP assay.

2.2. *Real-time Analysis*

The raw input data used does not contain patient age, primary tumor site, histological diagnosis, or variant pathogenicity status. For test orders that identified SNV/Indels, the gene, genomic sequence, and call status fields. For test orders that identified CNVs, the gene and call status fields were extracted. For test orders that identified fusions, the gene pair field was extracted.

2.3. *Source of Biomarker-based Clinical Trial Data*

Cancer treatment clinical trials open to accrual within Stanford University School of Medicine are tracked on OnCore Enterprise, the institution-wide clinical research management system. All interventional biomarker-based clinical trials from OnCore Enterprise were manually curated onto an Excel file that contains the following fields per trial - OnCore protocol identifier, National Clinical Trial (NCT) identifier, trial title, age group, biomarker gene, biomarker condition (mutation i.e. SNV/Indels, amplification, deletion, or fusion), biomarker detail (i.e. all mutations, specific amino acid change, or translocation partner gene), disease group, disease site, and contact information of the Principal Investigator and Primary Clinical Research Coordinator. Trials were classified into disease group categories based on the disease group as mapped in (Table S5). The input file is restructured into a data frame where every permutation of a unique combination of the individual biomarker criteria, disease group category, and disease site per trial is a separate row.

Cancer treatment clinical trials open to accrual within a multi-institutional precision medicine basket trial (referred to as PMB in this study) is tracked on an Excel file and amendments (i.e. addition and/or removal of new arms) are made on average twice a month. For each trial arm in the PMB trial, the criteria provided are a list of histologic disease exclusion codes, inclusion

non-hotspot rules, exclusion non-hotspot rules, exclusion variants, inclusion variants, immunohistochemistry (IHC) results, and comments. Entries labeled as "Indel" in the variant inclusion and exclusion criteria were re-classified into frameshift and in-frame (i.e. insertions, deletions, indels, duplications) mutations based on the protein sequences.

3. RESULTS

3.1. *STAMP assay identifies somatic mutations*

Molecular testing was performed on solid tumor biopsies using the STAMP targeted NGS assay, which has been offered by the Stanford Molecular Pathology laboratory since 2014. The STAMP v2 assay identifies somatic mutations, specifically, SNVs, Indels, CNVs, and/or fusions, in 130 genes that have been implicated in cancer[5]. The STAMP database, exported for analysis on 04/30/2019, contains 2028 unique test orders after performing quality control measures (Fig. 1, S1A, S2A-C). For this study, a unique patient is defined by a unique test order identifier, and we focused only on test orders associated with the STAMP v2 assay.

3.2. *Algorithmic pipeline flags eligible patients for precision medicine clinical trials*

For this study, we focused on the PMB trial and the internal Stanford Hospital biomarker-based clinical trials (referred to as OnCore in this study). Within the time window of this study, on average, the PMB trial comprised of 18 arms and the OnCore trials comprised of 12 arms. The pipeline we developed uses a hierarchical decision tree-based algorithm to determine whether any of the STAMP-identified mutations per patient satisfy the criteria of the clinical trials of interest and thereby, render the patient potentially eligible for at least one clinical trial. Specifically, each step examines whether a feature of the clinical trials matches the corresponding feature of the STAMP entries being queried, where each branch represents a decision and the leaves are the potential outcomes in the diagnostic report generated per patient.

This study utilizes different data sources derived from the STAMP pipeline - raw data files for the real-time analysis and annotated downstream files for the retrospective analysis. To design a pipeline that can accommodate the multiple data types, age (i.e. select for patients at least 18 years of age), pathogenicity (i.e. select for "likely pathogenic" and "pathogenic" SNV/Indels), and disease (i.e. match by trial disease group and disease site for the OnCore trials and exclude entries matching disease exclusion codes of the PMB trial) are designed as optional filters. If the optional filters associated with the criteria are not applied, the criteria will not be examined.

3.2.1. *Automation of Feature Matching*

The pipeline is designed to examine per class of clinical trials (i.e. OnCore, PMB inclusion variants, and PMB inclusion non-hotspot rules) the different variant types (i.e. SNV/Indels, CNVs, and fusions) via an iteration-based method. The algorithmic workflow for identification of candidate biomarker-based OnCore trials for STAMP-identified mutations are as follows - for SNV/Indels the feature order is gene name, variant type, variant detail (accepts "all mutations" and specific amino acid changes), pathogenicity, age, primary tumor site category, and primary tumor site (Fig. S3A); for CNVs the feature order is gene name, variant detail (i.e. amplification

or deletion), age, primary tumor site category, and primary tumor site (Fig. S3B); for fusions the feature order is gene name, variant type, age, primary tumor site category, and primary tumor site (Fig. S3C). The algorithmic workflow for identification of candidate PMB trial arms for STAMP-identified mutations based on inclusion variants are as follows - for SNV/Indels the feature order is gene name, variant type, genomic region, pathogenicity, and age (Fig. S4A); for CNVs the feature order is gene name, variant type, variant detail, and age (Fig. S4B); for fusions the feature order is gene name, variant type, and age (Fig. S4C). For the non-hotspot rules, we conservatively defined the criteria of deleterious to include any SNV and Indel, which will result in a preferred high false positive rate as it ensures that the algorithm will flag the majority, if not all, of the candidate cases for subsequent manual review. The algorithmic workflow for identification of candidate PMB trial arms for STAMP-identified mutations based on non-hotspot rules are as follows - for SNV/Indels the feature order is gene name, variant type, exon number, pathogenicity, and age (Fig. S5A); for CNVs the feature order is gene name, variant type (matches deletions only), and age (Fig. S5B). If candidate PMB trial arms are identified, the corresponding exclusion criteria and disease exclusions are assessed. The manually curated mapping of the primary tumor sites and histological diagnoses of the STAMP entries to the PMB trial disease classifications are indicated in (Table S1).

3.2.2. *Manual Review of Matching Output*

Upon completion of the multiple feature matching algorithms for each test order, the pipeline aggregates the results to generate summary files and a diagnostic report that is intended for manual review. For each STAMP test order, if candidate trials are identified for any of the identified mutations, then information about the mutation(s) and clinical trial(s) of interest will be exported to a tab-delimited file (Table S2) and formatted into a reader-friendly report (Fig. S6). In addition, a diagnostic report is generated to indicate for each of the STAMP-identified mutations, the leaves of the decision tree algorithm (Fig. S7).

3.3. *Validation of algorithmic pipeline*

We next evaluated the performance of the algorithm to identify eligible patients for the PMB trial based on somatic mutations identified by the STAMP v2 assay. We compared the results derived from our algorithmic pipeline to results derived from manual assessment. The manual patient eligibility screening based on STAMP v2 results that was conducted comprised of test orders from 07/01/2017 to 12/31/2017, inclusively (Fig. 1, 2A-C, S1B). Our pipeline identified a PMB match rate of 25.4% (Fig. 2D). Using the manual assessment results as the ground truth, our algorithm has a recall of 93.8%, precision of 36.6%, and specificity of 91.0%. Of the 78 false positives that the algorithm flagged, 75 entries were excluded in the manual assessment due to either “lung” as a primary tumor site or “lung adenocarcinoma” as the disease type (i.e. primary tumor site and histological diagnosis). While these criteria are not listed in the disease exclusions of the associated PMB trial arms, the compounds of interest (i.e. Afatinib, AZD9291, and Crizotinib) are FDA-approved for use in subsets of lung cancer[6-8]. Exclusion of these cases improved the recall to 93.8% and specificity to 99.7%.

3.4. *Match rate analysis of STAMP-identified mutations*

To predict the rate at which the STAMP assay identifies eligible patients for the PMB and OnCore trials, we performed a retrospective analysis using the STAMP database (Fig. 1, S1C) with the hypothesis that past test orders are representative of future test orders in terms of patient characteristics (Fig. 3A), tumor types (Fig. 3B), and variant types (Fig. 3C). We applied all the optional filters and identified a total match rate of 44.2%, an OnCore match rate of 38.3%, and a PMB match rate of 8.9% (Fig. 4A-C; Table S3). The effect of the optional filters on the match rate (Fig. S8A-I; Table S3) underscores the contributions of filtering by variant pathogenicity and disease grouping on the improvement of the match rate. Examination of the distribution of the OnCore trials and biomarker criteria that contributed to a match rate of 38.3% demonstrated that EGFR mutations for a EGFR/HER2/HER3-focused study of Neratinib therapy (OnCore #VAR0160) and KRAS mutations for a non-small cell lung cancer (NSCLC)-focused study of Regorafenib combination therapy (OnCore #LUN0097) contributed to 79.5% (n=497/625) of the matches (Table S4). Neratinib is a pan-HER inhibitor that binds and inhibits the activity of EGFR[11], a protein that is mutated in more than 60% of non-small cell lung carcinomas[12]. Regorafenib is FDA-approved for use in subsets of lung cancer[13]. Given the status quo of the compounds of interest, we speculate that test orders associated with a primary tumor site of "lung", 54.49% of the STAMP data (Fig. 3B), would be excluded from the trials.

To more accurately predict the rate at which eligible patients for OnCore trials may be identified using the STAMP assay, we excluded all entries where the primary tumor site is "lung" from the analysis (Fig. 1, S1D, S9A-C). This modification identified a total match rate of 17.8%, an OnCore match rate of 9.4%, and a PMB match rate of 9.9% (Fig. 4D-F; Table S3). The continued effect of the optional filters on the match rate (Fig. S10A-I; Table S3) further supports the contributions of filtering by variant pathogenicity and disease grouping on the improvement of the match rate.

4. DISCUSSION

4.1. *Incorporation of informatics into clinical workflows*

Successful and efficient patient eligibility screening is a lynchpin of the clinical trial accrual process. With less than 3% of eligible oncology patients enrolled in clinical trials[14] and the low prevalence of the biomarkers of interest, this is an ideal platform to investigate the benefits that workflow automation may provide. The latter situation is exemplified by the finding that several treatment arms of the largest precision medicine cancer trial to date, the National Cancer Institute Molecular Analysis for Therapy Choice (NCI-MATCH) clinical trials[15], did not accrue 35 patients during the initial screening cohort of approximately 6,000 patients[16]. On average, the total time spent on patient recruitment, from the initial identification to the eventual enrollment, is estimated to be 3.4. to 8.8. hours and $129 to $336, respectively, per patient[17]. Another hurdle is the necessity for physicians to be informed about the current status of all clinical trials, which is subject to frequent amendments. With both the quantity of clinical trials and the rate of genetic data accumulation increasing, the impracticality of the task will become amplified in the coming years. In this vein, automation via bioinformatic tools offers an opportunity to automatically identify

candidates for clinical trials and focus the attention of physicians on suitable candidates and trials. This may also lead to cost improvements in the drug development process, of which clinical trials are the most expensive component[18]. The few alternative solutions for automated clinical trial matching are either commercial (e.g. Molecular Match[9]) or use proprietary analysis engines[10] and hence, there is limited information available about the exact matching algorithm making a direct comparison difficult.

4.2. *Limitations of algorithmic pipelines*

The development of bioinformatic tools may significantly streamline clinical workflows. However, the ability of computational pipelines to successfully complete the designated tasks is dependent on the data sources used as input. In this study, we worked with multiple types of free text - those that were easily parsed by the algorithm (e.g. biomarker descriptions, variant pathogenicity, and age), those that required manual reclassification (e.g. disease groups), and those that lacked any consistent structure to be parsed by the algorithm (e.g. trial comments and IHC results). The challenges in working with text that require modification prior to use is exemplified by the disease mapping of the disease codes of the PMB trial with the primary tumor site and histological diagnoses of the STAMP-identified mutations, a time- and labor-intensive task whose complexity is compounded by the need for domain expertise. The benefits in utilizing the context of the fields is demonstrated in the significant improvement in the match rate upon applying the disease filter to the pipeline. In summary, the ability to accurately automate clinical workflows is dependent on the existence of structured data.

5. CONCLUSION

This study demonstrates through our patient eligibility screening algorithm to feature match STAMP-identified mutations with precision medicine clinical trial (i.e. OnCore and PMB trials) the use of an automated algorithmic pipeline as a feasible, accurate and effective alternative to the traditional manual process. This automation is dependent on the data modalities having features that are easily extractable with defined classifications. The integration of such bioinformatic tools into the existing clinical workflow is advantageous for translational research as it redistributes the limited resources currently allocated to tasks that may be automated to tasks that requires active physician engagement.

6. AUTHOR CONTRIBUTIONS

HAC, HS, and NB conceived and initiated the project. CAK, HAC, HS, NB, and JWC designed and performed the study. HAC, HS, and JWC analyzed and interpreted the data. HAC, HS, and JLZ supervised the work. HAC, HS, and JWC wrote and edited the manuscript.

7. ACKNOWLEDGEMENTS

No competing interests declared. We are grateful to Shivaani Kummar and Agnes Nika for their feedback and assistance during this study. We would like to thank Carlos Suarez for helpful discussions.

8. REFERENCES

1 *Cancer Statistics*, <https://www.cancer.gov/about-cancer/understanding/statistics> (

2 Ni, Y. *et al.* Increasing the efficiency of trial-patient matching: automated clinical trial eligibility pre-screening for pediatric oncology patients. *BMC Med Inform Decis Mak* **15**, 28, doi:10.1186/s12911-015-0149-3 (2015).

3 Carrasco-Ramiro, F., Peiro-Pastor, R. & Aguado, B. Human genomics projects and precision medicine. *Gene Ther* **24**, 551-561, doi:10.1038/gt.2017.77 (2017).

4 Drilon, A. *et al.* Safety and Antitumor Activity of the Multitargeted Pan-TRK, ROS1, and ALK Inhibitor Entrectinib: Combined Results from Two Phase I Trials (ALKA-372-001 and STARTRK-1). *Cancer Discov* **7**, 400-409, doi:10.1158/2159-8290.CD-16-1237 (2017).

5 *Stanford Actionable Mutation Panel for Solid Tumors*, <https://www.stanfordlab.com/esoteric/test-stanford-actionable-mutation-panel-for-solid-tumors.html> (

6 Solomon, B. J. *et al.* First-line crizotinib versus chemotherapy in ALK-positive lung cancer. *N Engl J Med* **371**, 2167-2177, doi:10.1056/NEJMoa1408440 (2014).

7 Soria, J. C. *et al.* Osimertinib in Untreated EGFR-Mutated Advanced Non-Small-Cell Lung Cancer. *N Engl J Med* **378**, 113-125, doi:10.1056/NEJMoa1713137 (2018).

8 Yang, J. C. *et al.* Afatinib versus cisplatin-based chemotherapy for EGFR mutation-positive lung adenocarcinoma (LUX-Lung 3 and LUX-Lung 6): analysis of overall survival data from two randomised, phase 3 trials. *Lancet Oncol* **16**, 141-151, doi:10.1016/S1470-2045(14)71173-8 (2015).

9 Welsh, J. *et al. The MolecularMatch genotype based therapeutics and clinical trials search engine* (The Cancer Genome - Keystone Symposia, 2016).

10 Patel, N. M. *et al.* Enhancing Next-Generation Sequencing-Guided Cancer Care Through Cognitive Computing. *Oncologist* **23**, 179-185, doi:10.1634/theoncologist.2017-0170 (2018).

11 Feldinger, K. & Kong, A. Profile of neratinib and its potential in the treatment of breast cancer. *Breast Cancer (Dove Med Press)* **7**, 147-162, doi:10.2147/BCTT.S54414 (2015).

12 da Cunha Santos, G., Shepherd, F. A. & Tsao, M. S. EGFR mutations and lung cancer. *Annu Rev Pathol* **6**, 49-69, doi:10.1146/annurev-pathol-011110-130206 (2011).

13 Bruix, J. *et al.* Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* **389**, 56-66, doi:10.1016/S0140-6736(16)32453-9 (2017).

14 Somkin, C. P. *et al.* Organizational barriers to physician participation in cancer clinical trials. *Am J Manag Care* **11**, 413-421 (2005).

15 Colwell, J. NCI-MATCH Trial Draws Strong Interest. *Cancer Discov* **6**, 334, doi:10.1158/2159-8290.CD-NB2016-018 (2016).

16 Harris, L. *et al.* Abstract B080: Update on the NCI-Molecular Analysis for Therapy Choice (NCI-MATCH/EAY131) precision medicine trial. *American Association for Cancer Research* **17**, B080, doi:10.1158/1535-7163.TARG-17-B080 (2018).

17 Penberthy, L. T., Dahman, B. A., Petkov, V. I. & DeShazo, J. P. Effort required in eligibility screening for clinical trials. *J Oncol Pract* **8**, 365-370, doi:10.1200/JOP.2012.000646 (2012).

18 Moore, T. J., Zhang, H., Anderson, G. & Alexander, G. C. Estimated Costs of Pivotal Trials for Novel Therapeutic Agents Approved by the US Food and Drug Administration,

2015-2016. *JAMA Intern Med* **178**, 1451-1457, doi:10.1001/jamainternmed.2018.3931 (2018).

9. FIGURES

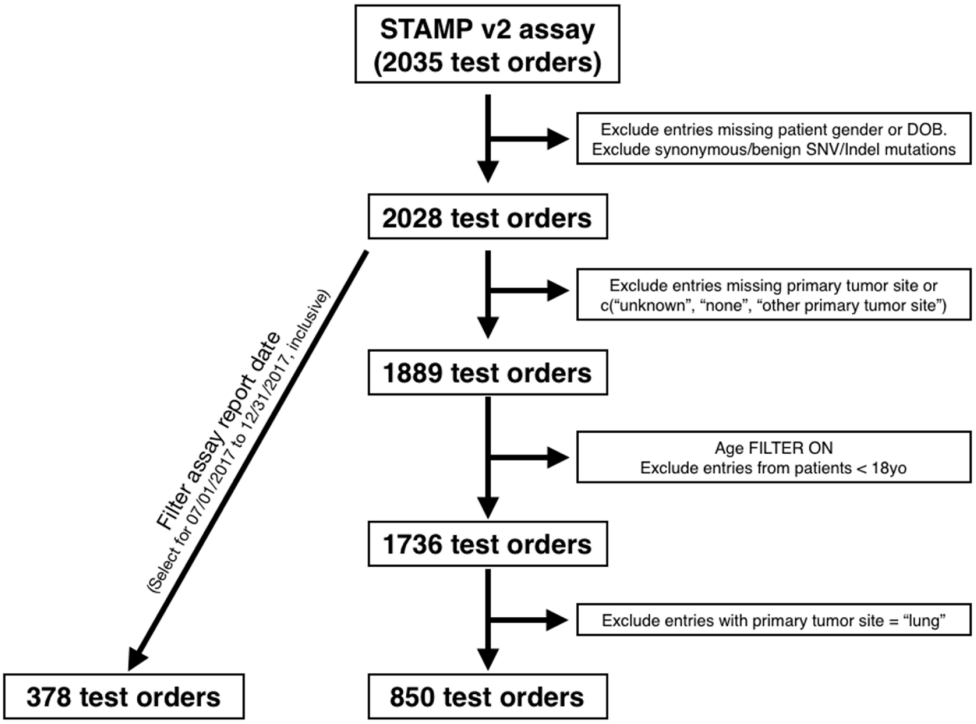

Fig 1. Flow diagram of STAMP database quality control procedure. The exclusion criteria applied to each group analyzed as derived from the STAMP v2 database are indicated. Sample sizes indicated include test orders that did not identify any mutations – refer to Fig. S1 for breakdown.

Fig. 2. Summary-level overview of STAMP database used in manual patient eligibility screening. (A) Age and gender distribution of patients, as determined using unique test order identifiers as proxies. (B) Distribution of primary tumor sites of patients. (C) Distribution of variant types – i.e. SNVs, in-frame indels, frameshift indels, CNVs, and fusions. The SNVs and indels are further partitioned by pathogenicity status of the variants. (D) Distribution of PMB trial matched by the algorithmic pipeline. The algorithm identified a PMB match rate of 25.4%. Using the manual assessment results as the ground truth, our algorithm has a 93.8% recall, 36.6% precision, and 91.0% specificity.

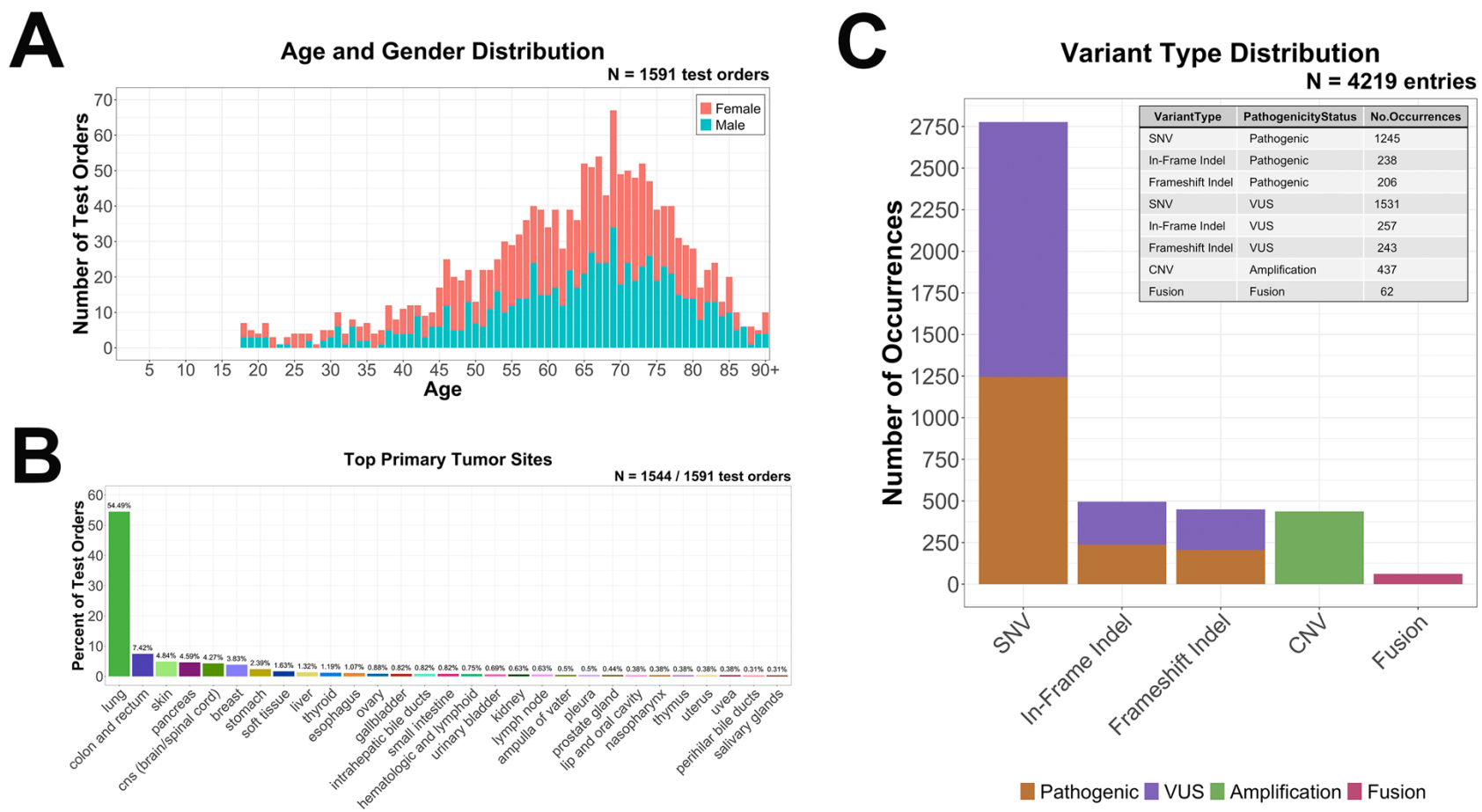

Fig 3. Summary-level overview of post-QC STAMP database. (A) Age and gender distribution of patients, as determined using unique test order identifiers as proxies. (B) Distribution of primary tumor sites of patients. (C) Distribution of variant types – i.e. SNVs, in-frame indels, frameshift indels, CNVs, and fusions. The SNVs and indels are further partitioned by pathogenicity status of the variants.

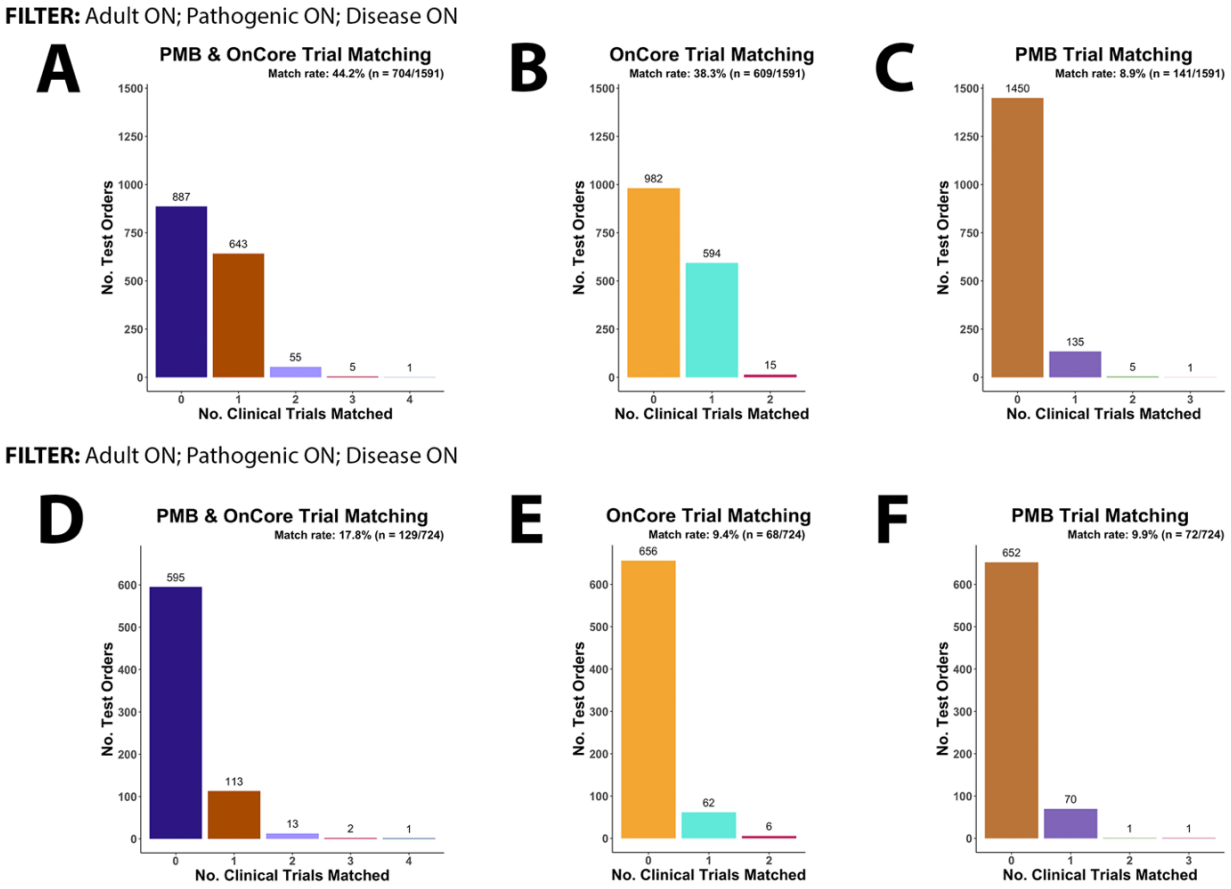

Fig 4. Distribution of PMB and OnCore trials matched by the algorithmic pipeline. (A-C) Match rate of the clinical trials with the input being the entire post-QC STAMP database. Match distribution of (A) combined trials, (B) OnCore trials, and (C) PMB trial. (D-F) Match rate of the clinical trials with input being the post-QC STAMP database entries where the primary tumor site is not lung. Match distribution of (D) combined trials, (E) OnCore trials, and (F) PMB trial. All the optional filters i.e. selection for adult patients, pathogenic variants, and trial-specific disease groups have been applied in this pipeline.

10. SUPPLEMENTARY TABLES AND FIGURES

Refer to the following URL: https://stanfordmedicine.box.com/v/ChenSuppTablesFigs