



OPEN ACCESS

EDITED BY

Roland Roller,
German Research Center for Artificial
Intelligence (DFKI), Germany

REVIEWED BY

Benjamin Maughan,
The University of Utah, United States
Philippe Thomas,
German Research Center for Artificial
Intelligence (DFKI), Germany

*CORRESPONDENCE

Channing J. Paller
cpaller1@jhmi.edu

SPECIALTY SECTION

This article was submitted to
Natural Language Processing,
a section of the journal
Frontiers in Artificial Intelligence

RECEIVED 02 July 2022

ACCEPTED 17 August 2022

PUBLISHED 12 September 2022

CITATION

Luechtefeld T, Bozada T Jr, Goel R,
Wang L and Paller CJ (2022)
Applications for open access
normalized synthesis in metastatic
prostate cancer trials.
Front. Artif. Intell. 5:984836.
doi: 10.3389/frai.2022.984836

COPYRIGHT

© 2022 Luechtefeld, Bozada, Goel,
Wang and Paller. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Applications for open access normalized synthesis in metastatic prostate cancer trials

Thomas Luechtefeld¹, Thomas Bozada Jr¹, Rahul Goel²,
Lin Wang³ and Channing J. Paller^{4*}

¹Insilica LLC, Bethesda, MD, United States, ²Independent Researcher, San Francisco, CA, United States, ³Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD, United States, ⁴The Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine, Baltimore, MD, United States

Recent metastatic castration-resistant prostate cancer (mCRPC) clinical trials have integrated homologous recombination and DNA repair deficiency (HRD/DRD) biomarkers into eligibility criteria and secondary objectives. These trials led to the approval of some PARP inhibitors for mCRPC with HRD/DRD indications. Unfortunately, biomarker-trial outcome data is only discovered by reviewing publications, a process that is error-prone, time-consuming, and laborious. While prostate cancer researchers have written systematic evidence reviews (SERs) on this topic, given the time involved from the last search to publication, an SER is often outdated even before publication. The difficulty in reusing previous review data has resulted in multiple reviews of the same trials. Thus, it will be useful to create a normalized evidence base from recently published/presented biomarker-trial outcome data that one can quickly update. We present a new approach to semi-automating normalized, open-access data tables from published clinical trials of metastatic prostate cancer using a data curation and SER platform. [Clinicaltrials.gov](https://clinicaltrials.gov) and [Pubmed.gov](https://pubmed.gov) were used to collect mCRPC clinical trial publications with HRD/DRD biomarkers. We extracted data from 13 publications covering ten trials that started before 22nd Apr 2021. We extracted 585 hazard ratios, response rates, duration metrics, and 543 adverse events. Across 334 patients, we also extracted 8,180 patient-level survival and biomarker values. Data tables were populated with survival metrics, raw patient data, eligibility criteria, adverse events, and timelines. A repeated strong association between HRD and improved PARP inhibitor response was observed. Several use cases for the extracted data are demonstrated *via* analyses of trial methods, comparison of treatment hazard ratios, and association of treatments with adverse events. Machine learning models are also built on combined and normalized patient data to demonstrate automated discovery of therapy/biomarker relationships. Overall, we demonstrate the value of systematically extracted and normalized data. We have also made our code open-source with simple instructions on updating the analyses as new data becomes available, which anyone can use even with limited programming knowledge. Finally, while we present a novel method of SER for mCRPC trials, one can also implement such semi-automated methods in other clinical trial domains to advance precision medicine.

KEYWORDS

data curation, prostate cancer, biomarker, text mining, natural language processing, machine learning

Introduction

Interventional biomarker clinical trials report associations between biomarkers, interventions, and patient outcomes. Biomarker-based outcomes and eligibility criteria are now common in cancer clinical trials. Most of the biomarker analyses reported for these trials are only accessible in difficult-to-access publications, with no controlled data access. Published biomarker outcomes are used in clinical guidelines, drug labels, and research. For example, the olaparib drug label references the PROfound trial data (de Bono et al., 2020; Hussain et al., 2020; FDA, 2020), and multiple publications have reviewed PARP inhibitors in mCRPC (Luo and Antonarakis, 2019; Antonarakis et al., 2020a; Jang et al., 2020). These activities involve a redundant, error-prone, time-consuming, and laborious data aggregation step.

Normalizing reported outcomes into an open-access database that can be easily updated as new data becomes available can help address this redundancy and improve the reusability of currently difficult to access biomarker outcome data. We present an approach to extracting the metastatic castration-resistant prostate cancer (mCRPC) trials data across different studies and normalizing it. The aggregated data tables and figures are also made available in an open-access format with a codebase that can be used to update them as new data becomes available. Biomarker, therapy, and outcome data were normalized into common tabular formats to analyze homologous recombination and DNA repair deficiency (HRD/DRD) biomarkers. The analyses herein can be fully reproduced in a short R script without the need for access to local files. The reviewer(s), time of review, and extracted data for each document can be looked up by readers and updated easily. We used Sysrev (Bozada et al., 2021), a web application designed to facilitate data curation and systematic evidence reviews (SER). It has an integrated machine learning (ML) platform with natural language processing (NLP) tools for searching, tagging, and extracting data from database sources. At a fundamental level, Sysrev helps to curate secondary data from biomedical research publications, and analyze them.

The potential value of the extracted and normalized data is demonstrated *via* several visualizations, which help to evaluate associations between therapies and DRD/HRD biomarkers in mCRPC trials. Computational access to these data points has the potential to accelerate precision medicine by enabling programmatic exploration of trial outcomes. This will become even more relevant in the future as results from more clinical trials get published and it will become even more difficult to search and screen for appropriate papers to include in a review, if done manually, to synthesize results across studies. Access to databases that can provide this data easily can facilitate clinical trial development and provide more transparency for guidelines, drug labels, and research.

Five analyses are done: biomarker usage counts, biomarker conditional treatment hazard ratios, biomarker response rates, adverse event comparisons, and automated discovery of therapy/biomarker relationships.

Many of these analyses focus on comparisons between the AR-directed (ARD) therapies abiraterone acetate + prednisone/enzalutamide and PARP inhibitors (PARPi) olaparib, rucaparib, veliparib, niraparib, and talazoparib. HRD specifically is widely recognized to mediate PARPi survival benefit, potentially *via* synthetic lethality (Jang et al., 2020).

Our approach demonstrates that normalizable biomarker/outcome data exist in prostate cancer clinical trial publications and identifies an opportunity to accelerate precision medicine by making this data computationally accessible. Precision oncology could benefit from efforts to collect this normalized data from clinical trials rather than re-collecting it in a systematic literature review (SLR), which is often outdated by the time it is published ((Beller et al., 2013)). We do not intend to promote SLR as the ideal solution for collecting this data. In *The Future Of Annotation/Biocuration*, Amos Bairoch states, “It is quite depressive to think that we are spending millions in grants [clinical trials] for people to perform experiments, produce new knowledge, hide this knowledge often in badly written text and then spend some more million trying to second guess what the authors really did and found.” (Bairoch, 2009).

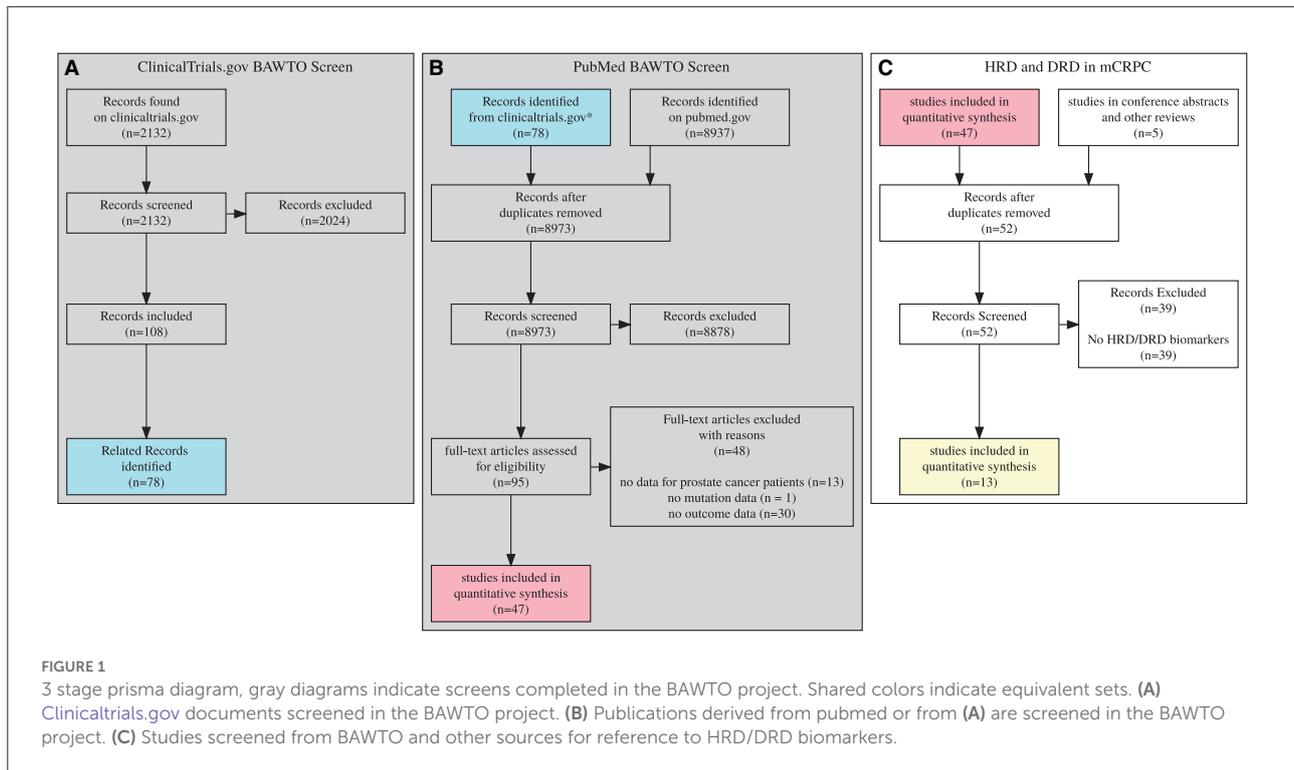
Materials and methods

Search strategy

The first step in the process was to carry out a review of biomarkers associated with therapeutic outcomes (BAWTO), from clinical trial publications. We extracted biomarkers matching one or more HRD/DRD biomarkers in the list (BRCA1, BRCA2, ATM, FANCA, PALB2, CHEK2, BRIP1, HDAC2, ATR, MLH1, MRE11A, NBN, RAD51C, RAD51, MLH3, ERCC3, MRE11, RAD51B, BARD1, CDK12, CHEK1, FANCL, PPP2R2A, RAD51D, RAD54L) (Luechtefeld et al., 2021).

A multi-stage Prisma diagram (Figure 1) describes the screening and search process for included documents (Page et al., 2021). Columns A and B describe the BAWTO review, which screens documents from [PubMed.gov](https://pubmed.gov) and [Clinicaltrials.gov](https://clinicaltrials.gov).

Briefly, the BAWTO review performs three screening steps. In each screening stage, documents were reviewed in duplicate, with conflicts resolved by a project administrator not involved in the original document review. Documents were included during each screening process when they contained three primary features:



1. A therapeutic intervention such as a drug, combination of drugs, surgery, or radiation.
2. A prostate cancer population with a well-defined mutation.
3. A measured patient outcome such as overall survival, adverse events, and progression-free survival.

In column A, the BAWTO review screened 2,132 documents from [Clinicaltrials.gov](#) matching the query for text matching “prostate”, study type “interventional,” and the overall status of “Completed” or “Active, not recruiting” as of April 22, 2021. These 2,132 documents were reviewed by five researchers with varying levels of research experience (three PhDs, BS, current university student); there were 58 conflicts which were resolved. In column B, the BAWTO review used a PubMed query for “prostate cancer” with the PubMed clinical trial filter. On April 29, 2021, this query generated 8,937 articles. The 8,937 were reviewed by the same five researchers. There were 52 conflicts which were resolved. The articles resulting from the two screens were deduplicated and overall 95 articles underwent a full text screen and initial data extraction phase, performed by four researchers (three PhDs, BS). Forty-seven articles were found to have the primary features stated above for any biomarker.

Column C in [Figure 1](#) describes the screening process completed in this HRD/DRD biomarker trials review. Forty-seven articles were sourced from BAWTO, and five reports were collected by searching conference abstracts and identifying relevant citations from other reviews of HRD/DRD in mCRPC. The data already extracted from these documents was reviewed

to identify articles with HRD/DRD biomarkers. Of the 52 resulting documents, 39 exclusions occurred due to a lack of HRD/DRD biomarkers. In total, we collected 13 documents for evidence synthesis over the course of 3 months.

Evidence synthesis

[Sysrev.com](#) was used to extract, transform, and load data from documents into controlled “labels”. Labels are boolean, string, categorical, or tabular forms that constrain user input in order to build consistent, structured data. Data were accessed and analyzed in R using the *rsr* package ([Bozada et al., 2021](#)).

Label design

During evidence synthesis, we extracted five primary tables from each document: Biomarker measurement, Relative population outcomes, Single population outcomes, Single population response rates, and Adverse Events. We also created an uncontrolled sixth table to contain patient-level data, including specific days to death, days to last follow-up, day of progression, biomarker status, and trial arm membership data. Future iterations of this project will attempt to control patient-level data, but the heterogeneity of patient-level data within and between publications made creating a normalized patient outcomes table too challenging for the current scope of this work.

TABLE 1 Ten mCRPC trials [0mm][8mm]measuring HRD or DRD biomarkers were included with results reported in 13 publications.

Trial	Study	Intervention(s)	Assignment	Brief observation(s)	Patient	Study type
NCT02987543	1. PROfound-1 2. PROfound-2	1. A1 - Olaparib 2. A2 - Enz or AAP	Randomized Parallel	1. Stratified population with HRD to AAP/Enza or Olaparib. 2. Identified strong evidence of PARPi survival benefit for HRD	None	Complete
NCT01682772	1. TOPARP-A 2. TOPARP-B	Olaparib	Single Group	1. First mCRPC specific HRD Olaparib trial. 2. Identifies strong response rates in patients with BRCA2 mutations	Limited	Complete
NCT02952534	1. TRITON2-1 2. TRITON2-2	Rucaparib	Single Group	1. Only Rucaparib (a PARPi) trial 2. Stage 1 gives BRCA1/2 associations stage 2 focuses on non BRCA HRD.	Full	Complete
NCT01576172	Hussain et al. (2018)	1. A1 - AAP 2. A2 - AAP+Veliparib	Randomized Parallel	1. ETS Fusion stratification factor 2. identifies no significant difference between arms 3. identifies HRD survival association in both arms	Full	Complete
NCT01972217	Clarke et al. (2018)	1. A1 - Olaparib + AAP 2. A2 - Placebo + AAP	Randomized Parallel	1. Precursor to PROfound trial with the same biomarkers measured 2. weak HRD - OS relationship, but stronger HRD - rPFS benefit	None	Complete
NCT02125357	Annala et al. (2018)	1. A1 - AAP 2. A2 - Enzalutamide	Randomized Crossover	1. No difference in survival between arms in any subgroup 2. Large number of tested biomarkers identified in liquid biopsy 3. Circulating Tumor DNA strong association with survival	Full	Complete
NCT01078662	Kaufman et al. (2015)	Olaparib	Single Group	1. The only trial with diseases in addition to mCRPC. 2. The first trial applying olaparib to mCRPC	None	Complete
NCT02312557	Graff et al. (2020)	Enz + Pembro	Single Group	1. The only trial involving immunotherapy - Pembrolizumab 2. Only 16 patients with baseline biopsies 3. Potentially weak psa response in patients with HRD 16% (3–50%)	None	Complete
NCT03148795	TALAPRO-1	Talazoparib	Single Group	1. Only talazoparib (a PARPi) trial. 2. Shows a strong composite response 76.1% in BRCA1/2 patients 3. Much weaker response in non BRCA1/2 HRD (0–40%)	None	Preliminary
NCT02854436	GALAHAD	Niraparib	Single Group	1. Only Niraparib (a PARPi) trial 2. Strong composite response in BRCA1/2 (65%) 3. Weaker composite response in non BRCA1/2 HRD (31%)	None	Preliminary

Publications were collected on April 22, 2021. The left identifier is the [Clinicaltrials.gov](https://clinicaltrials.gov) identifier for each trial. The study column provides a trial acronym (if available), followed by –1 for the first publication and a –2 for the second. The last author's name and year of publication are provided when a trial acronym was not available. The availability of patient data is broken down into none—no data available, limited—only biomarker data available, and full—both biomarker and clinical outcome data available. The study type indicates whether the collected data were a complete peer reviewed publication(s) or a preliminary conference abstract result.

Biomarker measurement

A simple label was constructed to associate gene names with their use as biomarkers in the respective study. Genes were categorized as used for inclusion, used for exclusion, measured, or not measured. Genes that were both measured and used for inclusion/exclusion, were only marked as used for inclusion/exclusion.

Relative population outcome

The “relative population outcome” table records hazard ratios as computed by univariate and multivariate Cox proportional hazards models for overall survival, prostate specific antigen (PSA) progression-free survival, and radiographic progression-free survival (Cox, 1972). The table captures treatment and biomarker descriptors for an experimental population and a control population. In total, 105 hazard ratios were extracted, with 47 hazard ratios comparing treatment outcomes for populations with the same biomarker status, and 58 hazard ratios comparing outcomes between populations with the same treatment but different biomarker statuses.

Single population outcome

The “single population outcome” table captures overall survival and progression-free survival median duration and confidence intervals. These tables define the population by known prior treatment, biomarker status, and treatment type. Two hundred and ten single population outcomes were recorded, 78 median overall survival measurements, and 132 median progression-free survival-based measurements.

Single population response rates

The “single population response rate” table defines a population by biomarker status, treatment, and known prior treatment. Two hundred and seventy single population response rates were recorded, including measurements of radiographic response (134), PSA response (94), circulating tumor cell reduction (10), and death (32).

Patient-level data

Four studies associated patient identifiers with biomarker status and survival/response rate metrics. These data were often provided in heatmap form and some supplementary

tables. Patient-level data were collected but have not yet been transformed into a normalized table.

Adverse event data

A simple table associates adverse event names with grade and event counts. This table suffers from uncontrolled vocabulary, but many trials used consistent naming. Some of these data were also available on [Clinicaltrials.gov](https://clinicaltrials.gov), but many trials did not have AE data available on [Clinicaltrials.gov](https://clinicaltrials.gov).

Clinicaltrials.gov trial protocols and metadata

Study protocols were downloaded from [Clinicaltrials.gov](https://clinicaltrials.gov) and included with their associated publications for reference during document review. Trial metadata describing phase, eligibility criteria, design, primary and secondary objectives were accessed programmatically *via* the [Tasneem et al. \(2012\)](#). Clinical trial events (start, end, etc.) were manually extracted from historical [Clinicaltrials.gov](https://clinicaltrials.gov) updates.

Pathway databases and genomic data commons access

Normalized biomarker data can be compared to other public databases. We found homologous recombination repair (HRR) genes that were missing in the reviewed papers by accessing KEGG, wikipathways, and reactome pathway databases with the `msigdb` package in R ([Dolgalev, 2020](#)). The referenced pathways included:

1. KEGG_HOMOLOGOUS_RECOMBINATION
2. WP_HOMOLOGOUS_RECOMBINATION
3. REACTOME_HDR_THROUGH_HOMOLOGOUS_RECOMBINATION_HRR

Missing gene relevance was measured by counting occurrences in the genomic data commons prostate cancer cohort with the `TCGABiolinks` package ([Colaprico et al., 2015](#)) and code: “`GDCquery_Maf (“PRAD”, pipelines=“muse”)`”.

Model-based recursive partitioning

Model-based recursive partitioning (MOB) was used to construct decision trees that optimally separate populations on treatment hazard ratios. MOB trees and random forests were created, using the R package MOB for recursive partitioning and the R survival package for Cox modeling

(Therneau and Grambsch, 2000; Zeileis et al., 2010; Garge et al., 2013; Therneau, 2021).

All patient-level data were collected into a single table to build MOB trees. Because different publications had different methods for annotating gene alterations, we chose to mark genes as either “altered” or “normal”. This simplifies model building at the cost of losing features like zygosity, which is an opportunity for future research.

The MOB tree for patient-level data from the Chinnaiyan study (Hussain et al., 2017) was built using all the Chinnaiyan genes. Parameters included a minimum of 20 patients in leaf nodes and no Bonferroni or confidence constraints.

The MOB tree for combined Chi/Chinnaiyan data used genes (1) measured in both trials and (2) had at least 10 patients for every combination of alteration status and treatment status. This tree was restricted to 15 children in each leaf node, a maximum depth of 5, and had no Bonferroni or confidence constraints.

The MOB random forest combined TRITON2/Chi/Chinnaiyan data. In this forest, we assumed that unmeasured genes were not altered. The forest was built from 10 trees and required a minimum of 10 patients in every leaf node. Trees used binary features for BRCA1, BRCA2, ATM, TP53, PTEN, AR, and engineered binary pathway features set to true if any gene in the pathway was altered - PI3K pathway, AR associated pathway, DNA repair. A final feature, “multiple alterations,” was true if the patient had more than two alterations in any gene.

Plotting

Base R, meta and metafor packages were used to generate confidence intervals, response rates, and forest plots (Viechtbauer, 2010; Balduzzi et al., 2019). The treemap visualization uses the plot_ly package with treemap layout (Inc. PT, 2015). Other visualizations involved custom code, ggplot, ggparty, patchwork and ggkm (Wickham and Wickham, 2007; Pedersen, 2017; Borkovec and Madin, 2019; Sachs, 2020).

Results

Trial characterization

Ten mCRPC clinical trials were identified that measured biomarkers for HRD or DRD (Table 1). Three trials had multiple publications, some having preliminary/final results (PROfound-1 vs. PROfound-2) and some having reports on different stages and subgroup analyses (TOPARP-A vs. TOPARP-B and TRITON2-1 and TRITON2-2).

Biomarkers were used as part of eligibility criteria, part of the primary or secondary objectives, and as part of the

trial design. In addition to HRD/DRD, several trials measured biomarkers related to other mechanisms in prostate cancer, including androgen receptor, ETS fusions, WNT, PI3K, and Cell cycle mechanisms. However, this publication primarily focuses on the outcomes associated with HRD/DRD.

The PROfound study was the only phase 3 trial identified. All other trials were phase 2 trials. There were four multi-arm trials comparing treatments in patients with and sometimes without specific HRD mutations. Three of these trials compared a PARP inhibitor against next-generation androgen receptor (AR) directed (ARD) therapies (Abiraterone Acetate + Prednisone or Enzalutamide), which are routinely used in mCRPC treatment (Cabot et al., 2012; Fizazi et al., 2012, 2017; Beer et al., 2014; Ryan et al., 2015). The remaining trials all had single group assignments and used PARP inhibitors (Talazoparib, Niraparib, Olaparib, Rucaparib, Veliparib) and, in one case, a combination of Enzalutamide and Pembrolizumab. None of the trials used radium-223 or sipuleucel-T. None of the trials used taxane-based therapies, although several had prior treatment conditions and subgroup analysis based on taxane use.

AR-directed therapies AAP and enzalutamide

The most common treatment class used in five trials was one of the AR-directed (ARD) therapies Abiraterone Acetate + Prednisone (AAP) or Enzalutamide. ARD was frequently used as a control (three trials) and/or in combination with another therapy (three trials), and one trial compared AAP to Enzalutamide. ARD works by targeting the androgen signaling pathway, which is critical in mCRPC progression (Sumanasuriya and Bono, 2018). Several ARD resistance mechanisms, including AR modifications, can reduce therapeutic benefits. These mechanisms may impact differential outcomes, particularly when ARD progression is an eligibility requirement.

PARP inhibitors

Many of the evaluated studies used PARP inhibitors like Olaparib (four trials), Rucaparib (TRITON), Veliparib (Chinnaiyan), Talazoparib (TALAPRO-1), and Niraparib (GALAHAD). PARP enzymes have an important role in single-strand DNA repair. One of the mechanisms for the therapeutic effect of PARP enzyme inhibitors is their ability to disrupt single-strand DNA repair (Dziadkowiec et al., 2016). Disruption of ssDNA repair results in the aggregation of double-strand breaks. When dsDNA repair is simultaneously disrupted *via* homologous recombination defects, like a BRCA2 alteration, the cell has no way to recover, and synthetic lethality occurs. Synthetic lethality is the primary reason for the high prevalence of PARP inhibitors as the experimental arm or therapy in the single group studies in these trials. The hypothesis is often that patients with a BRCA2 or other HRD mutation will see greater

benefit from PARP inhibitors relative to existing therapies like ARD.

Pembrolizumab

Pembrolizumab, immunotherapy, was tested in combination with Enzalutamide in the Moran, 2020 study (Graff et al., 2020). The resulting PSA response was low relative to other trials, even in the HRD group. Still, the authors saw other metrics to conclude that Pembrolizumab “Responses were deep and durable and did not require tumor PD-L1 expression or DNA-repair defects” (Graff et al., 2020).

Taxanes, like docetaxel and cabazitaxel, were not used as an intervention or control in any of the trials. Still, many trials had conditions or subgroup analysis based on prior treatment with taxanes.

Timelines

Seven of the ten reviewed studies have estimated completion dates in the future, with all trials estimated to complete before 2023. All included trials have already completed recruiting. Several trials started as recently as 2017 (Figure 2). We should expect the number of publications from these trials to double or more in the next 2 years relative to the reviewed publications found and included.

Biomarker eligibility, measurement and GDC prevalence

We recorded the measurement status and use of biomarkers as eligibility criteria for each study. Figure 3 provides a heatmap with rows ordered from the study with the most HRD/DRD biomarkers to the study using the least. The columns are ordered from the biomarkers used in the most studies (BRCA1/2) to the HRD/DRD biomarkers used in the least. BRCA1/2 was used in every identified study except TRITON2-1. TRITON2-2 and Kaufman et al. (2015) were the only studies that did not measure ATM, and TRITON2-2 is a special case where ATM results were published in TRITON2-1. Some of the heterogeneity in measured genes can be explained by choices to measure HRD and DRD vs. HRD alone. For instance, the FANCA gene is not considered part of the homologous recombination repair pathway in KEGG, Wikipathway, and Reactome pathway databases.

Unstudied HRR genes

KEGG, Wikipathway, and Reactome pathway databases identify 59 HRR genes that aren't measured in any trials. These

genes may be excluded for technical reasons, low prevalence, the low likelihood of having a deleterious effect, or other reasons. By using TCGAbiolinks (Colaprico et al., 2015; Silva et al., 2016; Mounir et al., 2019) we found that these unmeasured genes do have variants in Genomic Data Commons prostate adenocarcinoma population. Variants in the measured genes occurred in 63 patients, and variants in the unmeasured genes occurred in 30 GDC prostate cancer patients with 39 Moderate and four High impact variants identified by the IMPACT algorithm (McLaren et al., 2016).

The only study that excluded participants based on a biomarker result was TRITON2-1 which focused on the relationship of non-BRCA1/2 biomarkers of HRD with rucaparib (a PARP inhibitor). TRITON2-2 complemented this analysis with a focus on populations with a BRCA1/2 mutation.

Many biomarkers were measured in addition to HRD and DRD biomarkers. Other categories of biomarkers covered androgen receptor biomarkers, cell cycle, WNT pathway, chromatin modifier, and ETS fusions.

Treatment hazard ratios

Overall survival (OS) and progression-free survival (PFS) treatment hazard ratios (HRs) were collected from four trials and calculated from combined patient data in the Chinnaiyan, Chi, and TRITON studies.

In Figure 4A, PROfound OS HRs and patient counts are used to build a treemap. The overall HR is reported at the top (0.79), and subgroup hazard ratios are placed within parent groups with size proportional to prevalence and color from green to red indicating HR (Olaparib vs. AAP). The BRCA2|BRCA1|ATM group covers patients with BRCA1/2 or ATM mutations. It has five subgroups, a group with *only* BRCA2 alterations (@BRCA2), only BRCA1 (@BRCA1), prior taxane + only ATM, no prior-taxane + ATM alterations, and an “Other” group with HR found *via* pooling. These subgroups' pooled HR and patient count are equal to the BRCA2|BRCA1|ATM group.

Prior taxane therapy and ATM

In PROfound, the overall ATM HR was 0.93 (0.53–1.75), but prior-taxane status drops to 0.45, below even the BRCA2 HR, and the absence of prior taxane raises ATM HR to 2.82. This prior-taxane/OS relation is less significant in BRCA1/2 and CDK12 subgroups (Figure 3B PROfound). CDK12 may not be a robust biomarker for HRD (Antonarakis et al., 2020b).



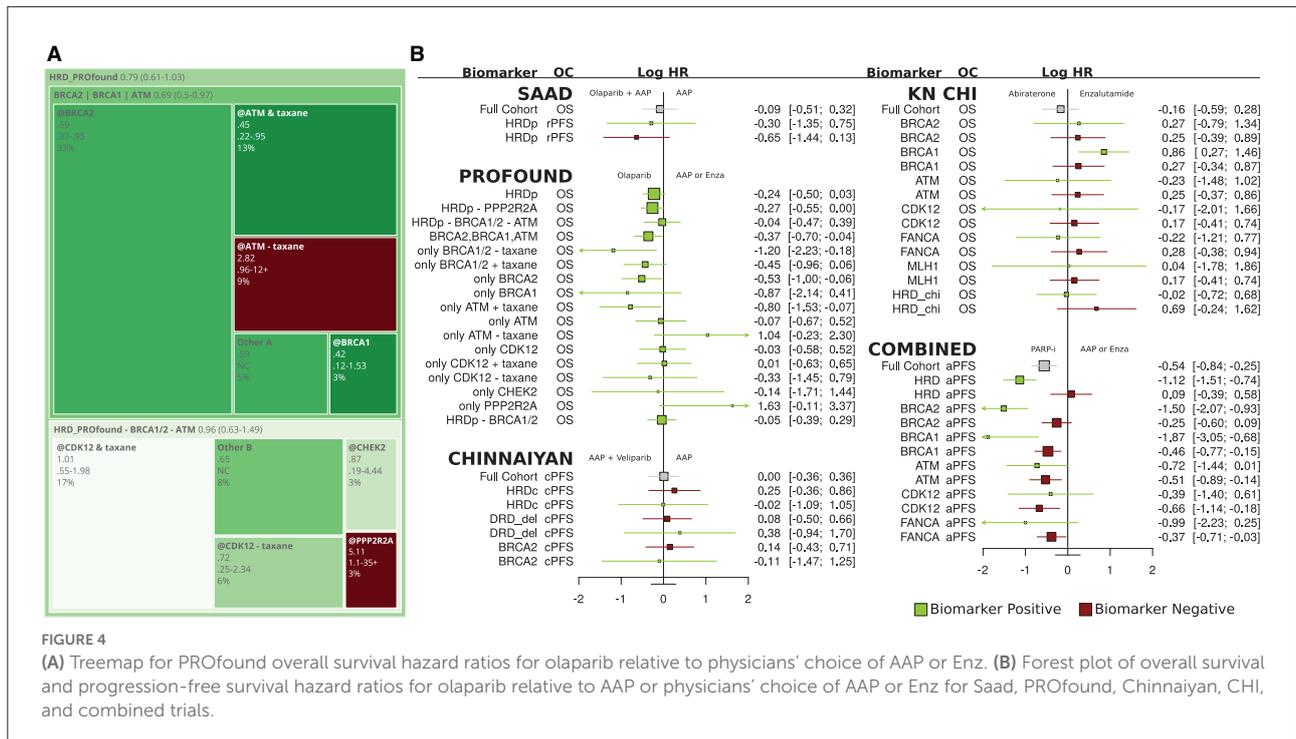
Significant results and BRCA2

The PROfound study was the only study to report significant treatment HRs at the 95% confidence interval level. The Chinnaiyan, Chi, and SAAD studies did not report significant differences at the 95% level (Figure 4B). BRCA2 had the strongest reported treatment HRs of.21 (0.13–0.32) for rPFS in PROfound. Chinnaiyan did not duplicate this, perhaps due

to the sample size, Veliparib specific factors, or differences in eligibility criteria.

Combined group

The Figure 4B COMBINED group reports HR between PARPi in Chinnaiyan/TRITON2 and ARD in Chi which were



chosen due to overlapping survival curves. PARPi HR was strong for the HRD group (green, row 2) relative to the full cohort and better still relative to patients who did not have HRD alterations (red).

Response rates

Single population response rates were extracted from all thirteen studies. In Figure 5, simple PSA response rates (PSA drop > 50%) and radiographic (RECIST any version) response rates are given. These results are filtered to only the reported HRD/DRD response rates. Abbreviations starting with HRD or DRD correspond to groups with mutations in any HRD/DRD biomarkers recorded in the respective trial (see Figure 3).

BRCA1/2 zygosity

Zygosity is an important consideration for DNA repair mutations. Monoallelic tumors may still have functional DNA repair pathways, and it is often hypothesized that homozygous loss/alteration is a better predictor of PARP-i response than alteration alone. In the TRITON2-2 study (2nd group from bottom), different response rates are given for BRCA1/2 based on zygosity. The homozygous loss group saw the best response rate, but the monoallelic group also saw a very strong response rate, perhaps indicating some weakness in this hypothesis.

ATM CDK12, and non-BRCA1/2 response rate

The ATM, CDK12, and other non-BRCA1/2 groups saw poor relative response rates in the TRITON2-1, GALAHAD, and TOPARP-B studies, providing more evidence of BRCA1/2 dominance in the observed HRD/PARP-i relation.

AAP differences

The PROfound trial ARD arm had poor response rates, possibly due to inclusion criteria requiring patients to have progressed on prior ARD therapy. The Chi study also reports this relationship and writes that HRD may prime tumors for ARD resistance (Annala et al., 2018). In the Chinnaiyan study, which did not have an ARD failure eligibility criteria, HRD positive AAP patients (green) had better aggregate response rates than those without (red).

Adverse events

In total, 113 adverse events (AEs) types were reported in at least one study, 20 were reported in 3 or more studies and are shown in Figures 6A,B. Similar data are stored on Clinical trials.gov, but with less coverage (Appendix A1). In Figure 6A,

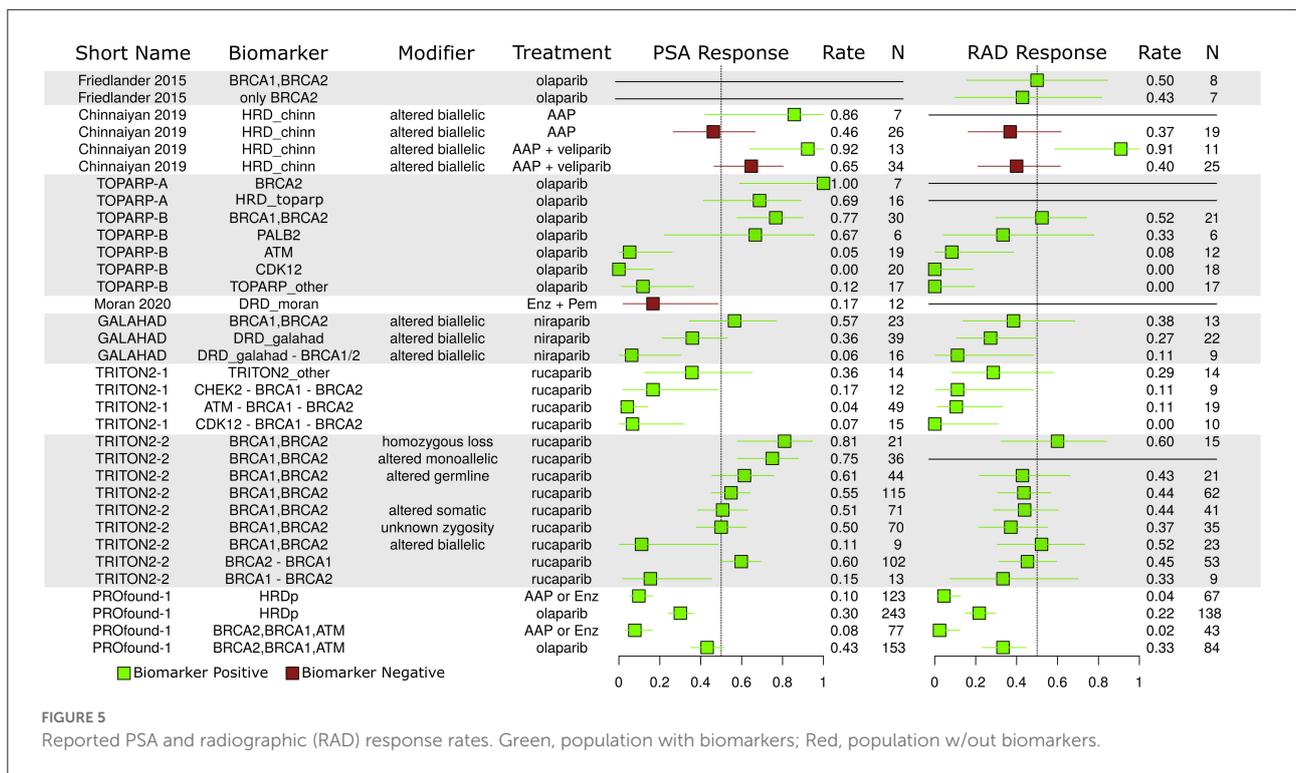


FIGURE 5

Reported PSA and radiographic (RAD) response rates. Green, population with biomarkers; Red, population w/out biomarkers.

median AE rates across grade and study are stacked and colored by grade. The median rate of serious anemia events was 20% in PARPi populations and ~20% for grade 1/2.

ARD events

The rate of severe grade 3+ AEs was <11% in all ARD populations for all event types. Hyperglycemia in the Chinnaiyan study and anemia/asthenia in the PROfound studies were the only reported serious AEs that impacted 5+ patients at >5%. Three out of 28 patients in the Moran ARD arm had a severe hypertension event. Asthenia was the most common event in all ARD populations except Moran.

PARP-I events

Anemia was the most common serious AE (>15%) in all PARP-i populations except the Chinnaiyan population, where only 2 (2.5%) of patients had a severe anemia event. Serious platelet count decreases and asthenia were also frequently reported at lower median rates.

AE dose dependence is seen in TOPARP-A-300 (300 mg olaparib) vs. TOPARP-A-400 (400 mg). Every serious AE affecting 2+ patients had a higher rate in the 400 mg group

(except back pain). Abdominal pain, spinal cord compression, arthralgia, and hypertension affected 8%+ of the 400 mg and 0% of the 300 mg group.

Comparison

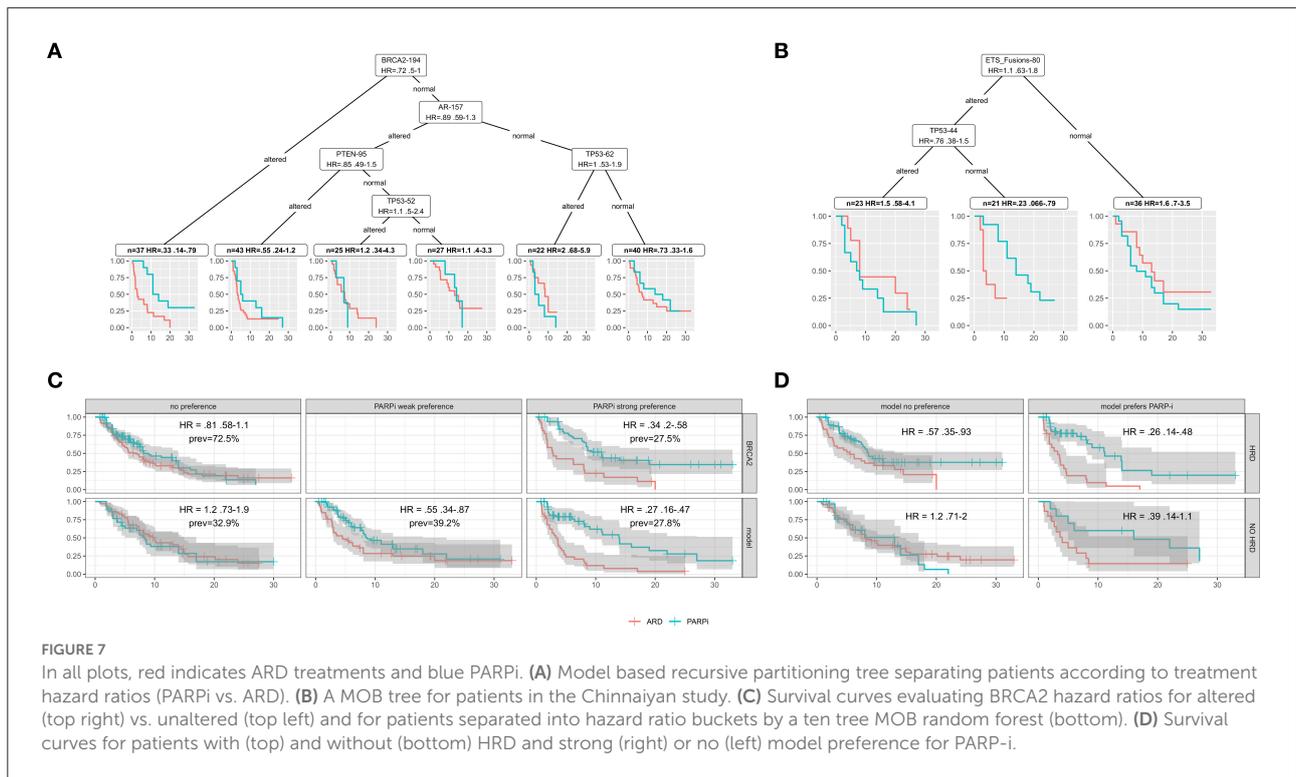
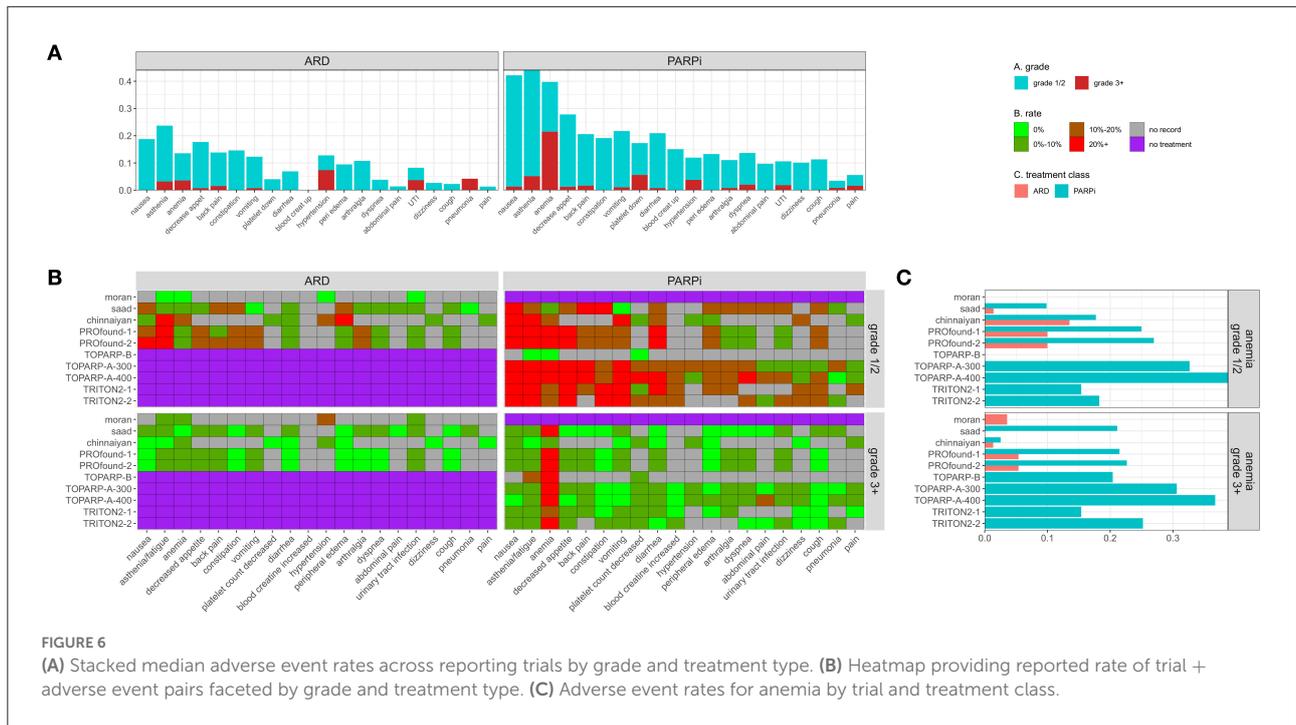
In studies with ARD and PARPi arms, the rate of most adverse events was higher for PARPi, with significance in some studies for anemia (Figure 6C), nausea, lymphocyte decrease, pyrexia, and asthenia. The only significant risk difference in grade 3+ AE was for anemia in the PROfound-2 and Saad studies (Appendix A2).

Patient-level data and simple models

In Figure 7, models are constructed from combined patient-level data from 2 to 4 trials. There are issues using combined data to build models, and here they are only meant to describe use cases for the extracted data.

Models as research tools

In Figure 7A, combined Chinnaiyan/Chi patient data is partitioned into groups that differ the most in Cox model



treatment hazard ratios. The model-based partitioning (MOB) tree (Zeileis et al., 2010) identifies BRCA2 as the strongest treatment differentiator. It prioritizes AR next, perhaps due to ARD resistance in the altered subgroup.

A tree built on all the patients in the Chinnaiyan paper (Figure 7B) identifies an ETS fusion subgroup negative for TP53 alteration with a strong PARPi preference. This subgroup provides evidence for molecular theories. ETS fusions are a

factor in PARPi activity, and p53 modifications are known to impact PARP1 PARylation (Fischbach et al., 2017). Clinical researchers could use models/visualizations like these in a research tool to explore treatment hypotheses.

Models to extend/refine drug indications

HRD can be thought of as a simple model - HRD is true if any HRR gene is altered. More complex models can add predictive power to biomarker indications. Figure 7C uses a random forest of MOB trees (Garge et al., 2013) to group patients from 4 trials by those with no, weak, or strong (top 15%) PARPi preference. In this combined cohort, BRCA2 alteration has an HR of 0.34 (upper right), and the strong PARPi preference group for the model has a lower HR of 0.21. Better stratification of patients can be useful when therapy is weighed against factors like AE risk.

Figure 7D uses the random forest to find an HRD negative subgroup that may still benefit from PARPi (lower right). The model also identifies a subgroup of HRD positive patients that may not benefit from PARPi (upper left). These examples demonstrate that complex models could extend or refine drug indications with the potential to improve trial success.

Queryable data

While machine learning models can provide powerful aids in developing precision medicine, often what is needed are answers to simple questions. Normalized, queryable data can answer questions like:

What was the longest reported median overall survival time in each trial?

One line of code can answer this question, and it allows us to generate ideas about the mCRPC trial's overall survival. We should not use this question to compare results. Median overall survival in months in increasing order:

1. 13.8 - BRCA1/2 negative HRD patients receiving olaparib (TOPARP-A)
2. 17.7 - BRCA1/2 altered patients receiving olaparib (TOPARP-B)
3. 22.7 - AAP + Olaparib without specified biomarker (Saad)
4. 24.8 - BRCA2 altered patients receiving olaparib (PROfound)
5. 32.3 - AAP + veliparib, without specified biomarker (Chinnaiyan)

Which PARP-I had the lowest and highest rate of grade 3+ anemia?

A single line of code can answer this question. In increasing order, the reported rates PARPi of serious anemia:

1. 2.53% - AAP + Veliparib in Chinnaiyan
2. 15.3% - Rucaparib in TRITON2-1
3. 20.4% - TOPARP-B
4. 21.1% - AAP + Olaparib Saad
5. 22.7% - Olaparib PROfound
6. 25.2% - Rucaparib in TRITON2-2
7. 30.6% - 300 mg group TOPARP-A
8. 36.7% - 400 mg group TOPARP-A

Open access, queryable data, will help providers and clinical researchers ask and answer questions like these in real-time. These queries can be viewed at insilica.co/bawto/figure-scripts/statements.html.

Discussion

Today, biomarker-based clinical trial outcome data are not readily accessible by oncologists, researchers, and policy makers in publicly available databases. Instead, the research, treatment, and regulatory decisions that rely on these data are limited by the need to search, read and recall literature manually. Accessibility in published literature indicates that the data are not confidential. An opportunity exists to improve data access, save time, reduce errors, and enable better therapy and research. Here, we present an approach to explore some of those possibilities in HRD/DRD biomarkers mCRPC clinical trials. Additionally, all of our work is open-source, with step-by-step instructions, and anyone with limited programming knowledge can build upon it.

The publication of trial results in closed-access journals makes it hard to access by oncologists, researchers, and policymakers. Also, the absence of computationally accessible data prevents the use of modern informatics methods on this data. When normalized, clinical trial results can be combined, which has its challenges, but can enable greater statistical power and improve our ability to understand the disease. In this review of mCRPC trials using our approach, we extracted thousands of data points from only 13 studies in a normalizable manner.

The data points analyzed in this trial are difficult and time-consuming to find. They can't be computationally accessed without human labor. They are not stored in an interoperable format so that results can be jointly analyzed. Even when the manual data extraction process is done, the results are not reusable by future researchers. In short, biomarker outcome data is not findable, accessible, interoperable, or reusable (FAIR).

In this publication, we show that for a set of 13 mCRPC HRD/DRD trial publications, outcome data for hazard ratios

could be extracted, compared, and visualized, thus identifying (1) high impact biomarkers like BRCA2, (2) possible biomarker interactions with prior treatments, (3) surprising results in specific trials, like the lack of an HRD survival benefit for Veliparib in the Chinnaiyan study. The combination of semi-normalized patient data was used to build simple models to automate the discovery of interesting patient subgroups in combined data.

With access, any internet-connected machine can run this entire analysis from a simple RScript. These scripts can be viewed on insilica.co/bawto or at <https://github.com/insilica/bawto>.

Models

Machine learning models like those demonstrated here can identify biomarker-outcome patterns that are too complex for humans or remain unobserved due to human bias (Maertens et al., 2020). Machine learning models can identify causal relationships, an important aspect when considering disease progression and treatment effects (Luechtefeld et al., 2020). Guidelines for using algorithms in trial designs indicate that models could be used to refine eligibility criteria or stratify patients based on model outputs from biomarker data. Guidelines like SPIRIT-AI provide suggestions and checklists to navigate the possible issues with clinical trial model use (Rivera et al., 2020).

Limitations

Our approach was restricted to clinical trial publications and, as a result, fails to identify secondary research identifying gene - outcome relationships. Many publications reference large amounts of retrospective molecular data derived from tissue banks and other sources. Some of these data are accessible in repositories like the genomic data commons and NCBI-GEO and would make an excellent addition to analyses performed for this study.

Conclusion

There exists a rich set of currently inaccessible but normalizable data in biomarker-based clinical trial publications. In a subset of 13 mCRPC clinical trial publications reporting the results of trials measuring homologous recombination and DNA repair deficiency biomarkers, our approach demonstrates that biomarker clinical trials have sufficient normalizable data. This will help inform biomarker selection in clinical trial design, compare and rank survival metrics, investigate the effects of biomarker-outcome relationships, evaluate adverse events, and

automate the discovery of therapeutic mechanisms. Today, oncologists, regulatory agencies, and researchers use many of these same data points, but they must search, read, and recall data from publications to retrieve them. The approach presented in this work to collect and store data allows for data to be reused, combined with other data sets, and extended in future analyses. While we showed its utility for metastatic prostate cancer, our approach is domain-agnostic and can be easily extended to other clinical domains and advance precision medicine.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://insilica.co/bawto/>; <https://github.com/insilica/bawto>.

Author contributions

This project was directed by CP, with critical review, editing and feedback from LW. RG provided editorial and administrative support. Technical development and analysis was performed by TL. Project administration was provided by TB. All authors contributed to the article and approved the submitted version.

Conflict of interest

Authors TL and TB were employed by Insilica LLC.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frai.2022.984836/full#supplementary-material>

References

- Annala, M., Vandekerckhove, G., Khalaf, D., Taavitsainen, S., Beja, K., Warner, E. W., et al. (2018). Circulating tumor DNA genomics correlate with resistance to abiraterone and enzalutamide in prostate cancer. *Cancer Discov.* 8, 444–457. doi: 10.1158/2159-8290.CD-17-0937
- Antonarakis, E. S., Gomella, L. G., and Petrylak, D. P. (2020a). When and how to use PARP inhibitors in prostate cancer: a systematic review of the literature with an update on on-going trials. *Eur. Urol. Oncol.* 3, 594–611. doi: 10.1016/j.euo.2020.07.005
- Antonarakis, E. S., Velho, P. I., Fu, W., Wang, H., Agarwal, N., Santos, V. S., et al. (2020b). CDK12 -altered prostate cancer: clinical features and therapeutic outcomes to standard systemic therapies, poly (ADP-ribose) polymerase inhibitors, and PD-1 inhibitors. *Jco Precis. Oncol.* 4, 370–381. doi: 10.1200/PO.19.00399
- Bairoch, A. (2009). The future of annotation/biocuration. *Nat. Preced.* 2009, 1. doi: 10.1038/npre.2009.3092.1
- Balduzzi, S., Rücker, G., and Schwarzer, G. (2019). How to perform a meta-analysis with R: a practical tutorial. *Evid. Based Ment. Health* 2019, 153–160. doi: 10.1136/ebmental-2019-300117
- Beer, T. M., Armstrong, A. J., Rathkopf, D. E., Loriot, Y., Sternberg, C. N., Higano, C. S., et al. (2014). Enzalutamide in metastatic prostate cancer before chemotherapy. *N. Engl. J. Med.* 371, 424–433. doi: 10.1056/NEJMoa1405095
- Beller, E. M., Chen, J. K., Wang, U. L., and Glasziou, P. P. (2013). Are systematic reviews up-to-date at the time of publication? *Syst Rev.* 2, 36. doi: 10.1186/2046-4053-2-36
- Borkovec, M., and Madin, N. (2019). *ggparty: “ggplot” Visualizations for the “partykit” Package*. R package version 1. Available online at: <https://CRAN.R-project.org/package=ggparty>
- Bozada, T., Borden, J., Workman, J., Cid, M. D., Malinowski, J., and Luechtefeld, T. (2021). Sysrev: a FAIR platform for data curation and systematic evidence review. *Front. Artif. Intell.* 4, 685298. doi: 10.3389/frai.2021.685298
- Cabot, R. C., Harris, N. L., Rosenberg, E. S., Shepard, J.-A. O., Cort, A. M., Ebeling, S. H., et al. (2012). Increased survival with enzalutamide in prostate cancer after chemotherapy. *N. Engl. J. Med.* 367, 1187–1197. doi: 10.1056/NEJMoa1207506
- Clarke, N., Wiechno, P., Alekseev, B., Sala, N., Jones, R., Kocak, I., et al. (2018). Olaparib combined with abiraterone in patients with metastatic castration-resistant prostate cancer: a randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet Oncol.* 19, 975–986. doi: 10.1016/S1470-2045(18)30365-6
- Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., et al. (2015). TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 44, e71–e71. doi: 10.1093/nar/gkv1507
- Cox, D. R. (1972). Regression models and life-tables. *J. Royal Statist. Soc. Ser. B* 34, 187–202. Available online at: <https://www.jstor.org/stable/2985181>
- de Bono, J., Mateo, J., Fizazi, K., Saad, F., Shore, N., Sandhu, S., et al. (2020). Olaparib for metastatic castration-resistant prostate cancer. *N. Engl. J. Med.* 382, 2091–2102. doi: 10.1056/NEJMoa1911440
- Dolgalev, I. (2020). *msigdb: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format*. R package version 7. Available online at: <https://CRAN.R-project.org/package=msigdb>
- Dziadkowiec, K. N., Gasiorowska, E., Nowak-Markwitz, E., and Jankowska, A. (2016). PARP inhibitors: review of mechanisms of action and BRCA1/2 mutation targeting. *Menopausal Rev.* 4, 215–219. doi: 10.5114/pm.2016.65667
- FDA (2020). *LYNPARZA FDA Drug Label*. Available online at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2020/208558s014lbl.pdf (accessed October 10, 2021).
- Fischbach, A., Krüger, A., Hampp, S., Assmann, G., Rank, L., Hufnagel, M., et al. (2017). The C-terminal domain of p53 orchestrates the interplay between non-covalent and covalent poly(ADP-ribosylation) of p53 by PARP1. *Nucleic Acids Res.* 46, gkx1205. doi: 10.1093/nar/gkx1205
- Fizazi, K., Scher, H. I., Molina, A., Logothetis, C. J., Chi, K. N., Jones, R. J., et al. (2012). Abiraterone acetate for treatment of metastatic castration-resistant prostate cancer: final overall survival analysis of the COU-AA-301 randomised, double-blind, placebo-controlled phase 3 study. *Lancet Oncol.* 13, 983–992. doi: 10.1016/S1470-2045(12)70379-0
- Fizazi, K., Tran, N., Fein, L., Matsubara, N., Rodriguez-Antolin, A., Alekseev, B. Y., et al. (2017). Abiraterone plus prednisone in metastatic, castration-sensitive prostate cancer. *N. Engl. J. Med.* 377, 352–360. doi: 10.1056/NEJMoa1704174
- Garge, N. R., Bobashev, G., and Eggleston, B. (2013). Random forest methodology for model-based recursive partitioning: the mobForest package for R. *BMC Bioinform.* 14:1, 5. doi: 10.1186/1471-2105-14-125
- Graff, J. N., Beer, T. M., Alumkal, J. J., Slottke, R. E., Redmond, W. L., Thomas, G. V., et al. (2020). A phase II single-arm study of pembrolizumab with enzalutamide in men with metastatic castration-resistant prostate cancer progressing on enzalutamide alone. *J. Immunother. Cancer* 8, e000642. doi: 10.1136/jitc-2020-000642
- Hussain, M., Daignault-Newton, S., Twardowski, P. W., Albany, C., Stein, M. N., Kunju, L. P., et al. (2018). Targeting androgen receptor and DNA repair in metastatic castration-resistant prostate cancer: Results from NCI 9012. *J Clin Oncol.* 36, 991–999. doi: 10.1200/JCO.2017.75.7310
- Hussain, M., Daignault, S., Twardowski, P., Albany, C., Stein, M. N., Kunji, L. P., et al. (2017). Abiraterone+ prednisone (Abi)+/-veliparib (Vel) for patients (pts) with metastatic castration-resistant prostate cancer (CRPC): NCI 9012 updated clinical and genomics data. *J. Clin. Oncol.* 35, 5001. doi: 10.1200/JCO.2017.35.15_suppl.5001
- Hussain, M., Mateo, J., Fizazi, K., Saad, F., Shore, N., Sandhu, S., et al. (2020). Survival with olaparib in metastatic castration-resistant prostate cancer. *N. Engl. J. Med.* 383, 2345–2357. doi: 10.1056/NEJMoa2022485
- Jang, A., Sartor, O., Barata, P. C., and Paller, C. J. (2020). Therapeutic potential of PARP inhibitors in the treatment of metastatic castration-resistant prostate cancer. *Cancers* 12, 3467. doi: 10.3390/cancers12113467
- Kaufman, B., Shapira-Frommer, R., Schmutzler, R. K., Audeh, M. W., Friedlander, M., Balmaña, J., et al. (2015). Olaparib monotherapy in patients with advanced cancer and a germline BRCA1/2 mutation. *J. Clin. Oncol.* 33, 244–250. doi: 10.1200/JCO.2014.56.2728
- Luechtefeld, T., Bozada, T., Cid, M. D., Chang, L., and Paller, C. (2021). *Biomarkers Associated With Therapeutic Outcomes*. PROSPERO 2021 CRD42021245432. Available online at: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021245432 (accessed October 10, 2021).
- Luechtefeld, T., Lin, N., Paller, C., Kuhns, K., Lartera, J. J., and Bressler, J. P. (2020). Measuring cancer hallmark mediation of the TET1 glioma survival effect with linked neural-network based mediation experiments. *Sci. Rep.* 10, 1–11. doi: 10.1038/s41598-020-65369-3
- Luo, J., and Antonarakis, E. S. (2019). PARP inhibition — not all gene mutations are created equal. *Nat. Rev. Urol.* 16, 4–6. doi: 10.1038/s41585-018-0129-3
- Maertens, A., Tran, V. P., Maertens, M., Kleensang, A., Luechtefeld, T. H., Hartung, T., et al. (2020). Functionally enigmatic genes in cancer: using TCGA data to map the limitations of annotations. *Sci. Rep.* 10, 1–11. doi: 10.1038/s41598-020-60456-x
- McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., et al. (2016). The ensembl variant effect predictor. *Genome Biol.* 17, 122. doi: 10.1186/s13059-016-0974-4
- Mounir, M., Lucchetta, M., Silva, T. C., Olsen, C., Bontempi, G., and Chen, X. (2019). New functionalities in the TCGAbiolinks package for the study and integration of cancer data from GDC and GTEx. *PLoS Comput. Biol.* 15, e1006701. doi: 10.1371/journal.pcbi.1006701
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst. Rev.* 10, 89. doi: 10.1186/s13643-021-01626-4
- Pedersen, T. L. (2017). *Patchwork: The Composer of ggplots*. R package version 00.1. Available online at: <https://CRAN.R-project.org/package=patchwork>
- Rivera, S. C., Liu, X., Chan, A.-W., Denniston, A. K., Calvert, M. J., Group TS-A, and, C.-A. W., et al. (2020). Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Lancet Digital Heal.* 2, e549–e560. doi: 10.1136/bmj.m3210
- Ryan, C. J., Smith, M. R., Fizazi, K., Saad, F., Mulders, P. F. A., Sternberg, C. N., et al. (2015). Abiraterone acetate plus prednisone versus placebo plus prednisone in chemotherapy-naïve men with metastatic castration-resistant prostate cancer (COU-AA-302): final overall survival analysis of a randomised, double-blind, placebo-controlled phase 3 study. *Lancet Oncol.* 16, 152–160. doi: 10.1016/S1470-2045(14)71205-7
- Sachs, M. (2020). *ggkm: Survival Plots for ggplot2*. R package version 0.7.0.
- Sievert, C. (2020). *Interactive Web-Based Data Visualization with R, Plotty, and Shiny*. Boca Raton, FL: Chapman and Hall/CRC.

- Silva, T. C., Colaprico, A., Olsen, C., D'Angelo, F., Bontempi, G., and Ceccarelli, M. (2016). TCGA Workflow: Analyze cancer genomics and epigenomics data using Bioconductor packages. *F1000Research* 5, 4–5. doi: 10.12688/f1000research.8923.1
- Sumanasuriya, S., and Bono, J. D. (2018). Treatment of advanced prostate cancer—A review of current therapies and future promise. *Csh Perspect. Med.* 8, a030635. doi: 10.1101/cshperspect.a030635
- Tasneem, A., Aberie, L., Ananth, H., and Chakraborty, S. (2012). The database for aggregate analysis of ClinicalTrials.gov (AACT) and subsequent regrouping by clinical specialty. *PLoS ONE* 7, e33677. doi: 10.1371/journal.pone.0033677
- Therneau, T. M. (2021). *A Package for Survival Analysis in R*. R package version 3.2-13. Available online at: <https://CRAN.R-project.org/package=survival>
- Therneau, T. M., and Grambsch, P. M. (2000). *Modeling Survival Data: Extending the Cox Model*. New York: Springer.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* 36, 1–48 doi: 10.18637/jss.v036.i03
- Wickham, H., and Wickham, M. H. (2007). *The ggplot Package*. Available online at: <https://cranr-project.org/web/packages/ggplot2/index.html>
- Zeileis, A., Hothorn, T., and Hornik, K. (2010). *Party With the Mob: Model-Based Recursive Partitioning in R*. R package version 09–9999.