

RESEARCH

Open Access



An open-source clinical bioinformatics pipeline for real-world NGS implementation: translating genomic variants into actionable treatment strategies in oncology

Grete Francesca Privitera^{1*}, Salvatore Alaimo^{1*}, Giovanni Micale¹, Luca Giaimi², Marzia Mare², Sofia Paola Lombardo², Emanuele Martorana², Riccardo Villa², Alfredo Ferro¹, Stefano Forte^{2†} and Alfredo Pulvirenti^{1*†} 

Abstract

Background Next-Generation Sequencing (NGS) has become a cornerstone technology in clinical practice, yet its adoption presents significant challenges. Physicians and oncologists must manage vast amounts of genome-scale data and transform it into actionable insights for complex decision-making. While commercial systems exist to synthesize data from NGS experiments into clinical reports, many are hindered by limitations such as closed-source designs that restrict transparency and customization. Additionally, some fail to leverage publicly available genomic databases, missing opportunities to integrate valuable external data. Furthermore, the rigidity of many tools in accommodating diverse NGS panels limits their applicability across varied clinical scenarios.

Methods To address these limitations, we developed OncoReport, an open-source tool that generates comprehensive reports from NGS analyses. By integrating publicly accessible databases, OncoReport provides a robust, user-friendly environment equipped with essential tools for NGS analysis. This design aims to enhance data interpretation and support informed clinical decision-making.

Results Rigorous testing has demonstrated OncoReport's effectiveness in producing detailed, actionable reports that are clear and easy to use. By automating key aspects of the workflow, the tool significantly reduces manual effort and expedites the synthesis and interpretation of NGS results, making genomic insights more accessible to clinicians.

Conclusion OncoReport offers a transparent, flexible, and efficient framework for clinicians to analyze and apply genomic data in patient care. By streamlining workflows and leveraging open-source principles, it empowers

[†]Stefano Forte and Alfredo Pulvirenti contributed equally to this work.

*Correspondence:

Grete Francesca Privitera
grete.privitera@unict.it
Salvatore Alaimo
salvatore.alaimo@unict.it
Alfredo Pulvirenti
alfredo.pulvirenti@unict.it

Full list of author information is available at the end of the article



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

healthcare professionals to make informed, data-driven decisions. OncoReport is freely available at <https://oncoreport.atlas.dmi.unict.it>, with source code and issue tracking on GitHub: <https://github.com/knowmics-lab/oncoreport>.

Keywords Next generation sequencing, Personalized medicine, Cancer, Decision support systems, Drugs

Background

Personalized oncology strives to customize cancer treatment according to an individual's distinctive attributes and genetic profile. Medical experts can employ tailored treatment approaches by evaluating genetic modifications and biomarkers, like mutations and gene expressions. Notable successes encompass focused therapies and immunotherapy, leveraging the immune system in the battle against cancer. This approach enhances patient outcomes, minimizes adverse effects, and potentially propels cancer treatment forward. Next-Generation Sequencing (NGS) technological progress has democratized personalized medicine, enabling broader integration within clinical routines. However, these technological strides have concurrently given rise to pressing concerns linked to the interpretation of molecular data. The prevailing form of NGS sequencing employs short reads (50–500 bp). When applied to patient samples, these experiments generate substantial datasets, necessitating labor-intensive efforts to gather insights into genetic variations. Users must undergo specialized training to install and navigate precise tools for analyzing and interpreting sequencing data. The bioinformatics pipelines for scientific and research purposes demand intricate installation protocols and thorough dependency handling. Additionally, the execution of these software packages requires the retrieval and processing of numerous databases.

Related works

Many pipelines have been developed for NGS analysis in the last few years. In 2019, Joo T. et al. introduced SEQprocess [1], a tool implemented in R to analyze patient NGS data in FASTQ format. Six pre-customized pipelines can be applied to DNaseq and RNA-seq data. The user can analyze Whole Genome Sequencing (WGS) and Whole Exome Sequencing (WES) on both tissue and cell-free (liquid biopsy) DNaseq samples. Notably, SEQprocess facilitates accurate allele frequency estimation and RNAseq analysis, identifying read depths in both mutated and wild-type sequences. However, the tool exclusively supports Linux-based Operating Systems (OS).

SEQVitA [2] is an open-source platform for identifying SNVs and small INDELs in WGS, WES, and custom NGS data. It supports multiple BAM or mpileup files for analysis. The system can also analyze VCF outputs provided by the user. According to the number and type of samples, the detection of mutations is divided into Germline,

Population, and Somatic. For the latter kind of mutations, paired tumor/control samples are needed. In addition, variant annotation is performed for coding and non-coding variants associated with a functional impact score to enable prioritization.

DNAscanv2 [3] can annotate and visualize variants in a NGS sample. The user can analyze a sub-region of the genome. The annotation step is performed by Annovar [4] involving Clinvar [5, 6], Exac [7], dbSNP [8], and dbNSFP [9] databases. Quality control and a results report are generated at the end of the analysis with a tab-delimited list of the variants. The pipeline is available as a Singularity or Docker image. As of December 2025, the tool has not received any updates for two years.

VARIFI [10] is a web-based variant identification, filtering, and annotation pipeline. The user can load a sample file up to 400MB. The tool combines different aligners and variant callers to improve accuracy. It is an easy-to-use pipeline that does not require local computational resources since it runs on the VARIFI servers. Users can load two input files in BAM, FASTQ, or BED format. Reads are mapped using bowtie [11], BWA [12], and NextGenMap [13]. Then, the Genome Analysis Toolkit (GATK) is used for realignment. Variants are called using the UnifiedGenotyper and bcftools. Finally, annotation is done with Annovar, removing potential false positives. A final report contains variants sorted by a confidence score and a plot with amplicon coverage information. Unfortunately, the results produced by such systems are often hard to obtain, and their interpretation in a clinical context could be challenging. Based on variant identification and automated annotation, the principal limit of these approaches is represented by the lack of appropriate filtering and prioritization of actionable variants in the specific context of each medical case. Therefore, applicability in the clinical context is limited. To fill this gap, several solutions have been proposed. As of December 2025, the VARIFI web portal is not operational.

LUSH DNaseq pipeline [14] is a command-line pipeline for variant calling of WGS analysis. It consists of four modules: (i) the aligner; (ii) the BQSR (BaseRecalibrator); (iii) the HC (HaplotypeCaller); (iv) the GenotypeGVCFs. The tool has been claimed to be faster than GATK. The analysis can start from any module according to the user needs.

TGex [15], currently named Genexy Genomex, is a commercial online knowledge-driven platform for clinical genetics analysis. It combines variant annotation and filtering capabilities with an user interface, allowing

interactive interpretation and filtering of the variants by scientists without any bioinformatics skill. To perform the analysis, the user should upload a patient VCF. The output is a report file in PDF or Word with a detailed variant annotation Excel file.

IMPACT (Integrating Molecular Profiles with Actionable Therapeutics) [16] is a pipeline that links somatic variants detected in WES analysis to actionable therapies. It uses a specific drug-target database developed by the same authors [17]. They also developed a web portal to query the variants and find the appropriate therapy [18]. As of December 2025, the IMPACT web portal is not operational.

PCGR (Personal Cancer Genome Report) [19] is an open-source pipeline that creates HTML reports of variants starting from VCF files. It allows users to choose between GRCh37 and GRCh38 and uses several cancer variant databases. The final report provides evidence level for each variant found with the information about associated drugs. Although its report is functional and easy to analyze, its usage could be limited in the clinical setting since it mainly focuses on variant interpretation rather than personalized drug therapy decision support (side effects, drug-food interaction, etc.). Furthermore, it does not allow neither a complete analysis starting from raw reads, nor the filtering of the specific information concerning the patient's disease.

MOAlmanac (Molecular Oncology Almanac) [20] is a clinical interpretation algorithm equipped with its own comprehensive alteration database. It integrates information from somatic and germline mutations and has the capability to analyze transcriptome data. The algorithm prioritizes mutations by leveraging multiple databases to suggest the most effective treatment responses. Its pipeline begins with a mutation list, provided in either TXT or MAF format, and generates an HTML report. This report includes detailed information about predictive implications, mutation types, and potential therapies. MOAlmanac is accessible through Docker and Terra platforms, ensuring flexibility and broad usability.

MTBP (Molecular Tumor Board Portal) [21] is an algorithm developed as part of Cancer Core Europe (CCE) to transform NGS data into actionable therapeutic recommendations for patients. It combines several bioinformatic tools and databases to streamline the identification of cases eligible for clinical trials across CCE centers. MTBP ensures patient data is de-identified and follows a structured process to generate a comprehensive report. While the full version of MTBP is reserved for internal use, a lightweight version is available for online access by external users.

In this paper, we present OncoReport, an open-source application that generates comprehensive reports to support the clinical interpretation of genetic variants in all

cancer types and their implications for prognosis and therapy. It can be easily installed and deployed through a Docker container or a Desktop app with an user-friendly interface. It integrates data from several external databases that provide curated and reliable information on cancer variants and their clinical relevance, such as Clinical Interpretation of Variants in Cancer (CIViC) [22], Cancer Genome Interpreter (CGI) [23], Pharmacogenomics KnowledgeBase (PharmGKB) [24], The Catalogue Of Somatic Mutations In Cancer (COSMIC) [25], Clinvar [5, 6], Refseq [26] and DrugBank [27, 28]. Our methodology is versatile and applicable to liquid biopsy and tissue NGS data. Physicians can utilize this approach to conduct various types of analyses, including but not limited to WGS, WES, and targeted panel analyses.

Results

The OncoReport pipeline

The pipeline illustrated in Fig. 1 consists of four sequential stages: (i) Preprocessing, (ii) Variant calling and filtration, (iii) Variant annotation, and (iv) Report generation. However, our adaptable procedure allows users to initiate the analysis at any of these steps. During preprocessing, OncoReport removes sequencing adapters and low-quality reads via TrimGalore [29, 30]. Subsequently, alignment is carried out using bwa [12]. This process culminates in producing BAM files, sorted based on genomic coordinates, realigned to match the reference contig ordering, and subjected to duplicate removal using Picard [31] tools. It is important to note that the step of duplicate removal is omitted for samples from liquid biopsy experiments. Next, variant calling is performed with GATK Mutect2 [32], VarScan2 [33], and LoFreq [34]. The choice to use one or all of these variant callers is left to the user. After variant calling, VCF files are filtered to remove malformed and artifact calls. Then, a depth (DP) filtration step is executed. In the context of liquid biopsy samples, an Allele Fraction (AF) filter is additionally applied through the GATK VariantFiltration tool. The DP filter retains high-quality bases, while the AF filter estimates the likelihood that a given variant is either germline or somatic. We recommend an upper threshold of 0.3 for somatic variants in liquid biopsy samples and 0.4 for tissue samples. It should be noted that this threshold can be adjusted based on histological evaluations, such as the percentage of tumor cells present in the sample. Following variant calling and filtration, an annotation script is used against an internal variant database. The internal database contains data from CIViC, CGI, Refgene, PharmGKB, and Cosmic, as detailed in the "Databases" section. The annotation script parses the VCF file, conducting a positional search for each variant using the "genome join" function of the "fuzzyjoin" R package. This function assists in linking variants to specific

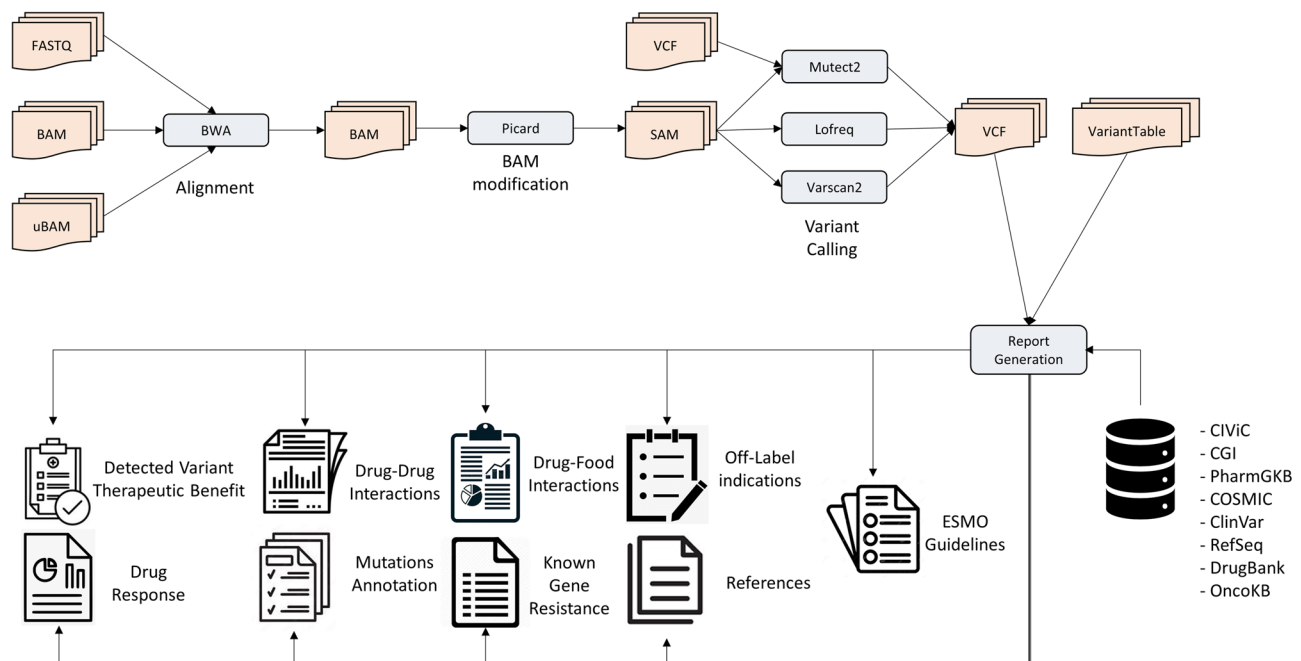


Fig. 1 The OncoReport pipeline. The pipeline can be run using FASTQ, BAM, or Illumina variant table files. The user can run the pipeline with default parameters or choose the most suitable parameters for his analysis. At the end of the analysis, the pipeline automatically generates the report, which will contain all the information retrieved from the databases and integrated with the DNA-seq data of the patient

genomic locations, thereby enriching the annotation process. Finally, all variants that map to at least one database record are stored to build the final report. The custom report consists of a self-contained archive with an HTML page for each section. Our report's structure and contents have been organized for clinical usage. Notably, mutations annotated within the CIViC or CGI databases undergo a scoring procedure that evaluates numerous clinically significant factors. This approach streamlines the decision-making process.

Within OncoReport, the databases, as mentioned earlier, undergo tailored modifications aimed at retaining only essential data for clinical utilization. This strategic refinement serves the purpose of annotating variants with their corresponding drugs. Notably, the enrichment of PharmGKB is achieved by leveraging Ensembl, addressing the information gaps in the original PharmGKB version (varphenoanno.tsv), which lacks variant position and alternative base details. The integration of literature insights is facilitated by using the efetch API [35]. In our pursuit of comprehensive data, two distinct files from COSMIC have been obtained: one encompassing all documented cancer-related mutations and another featuring resistance mutations. A strategic overlap of these files enables the identification of variants correlated with drug resistance. Each drug's relevance is quantified using a three-fold metric: the publication year of the drug-variant association, the pathogenicity of the associated mutation, and the frequency of appearance

in our report. Diverse factors shape the scoring system. The temporal relevance of the association influences the score: 3 for associations published within the last three years, 2 for the last six years, 1 for the last nine years, and 0.5 for the last 12 years. Variants with earlier associations receive a score of 0. Pathogenicity scoring relies on the dbnsfp41c [36] database, where each of the eight pathogenicity predictors assigns a score between 1 and 0 to the variant. Additionally, within the report, drug scores are augmented by 1 for every variant mutation linked to them. Ultimately, these provisional scores are aggregated to yield the final cumulative score. This score considers temporal and pathogenicity aspects and acknowledges the number of associated variant mutations—a comprehensive approach to aid in decision-making. The underlying code for this study is available on GitHub and can be accessed via this link: <https://github.com/knownmics-lab/oncoreport>.

The user interface

The user gains authentication access to the system by utilizing a Personal Access Token created through Laravel Sanctum's scaffold [37]. This token is generated by an administrator within the web application and is initially configured into the client application. These tokens are centrally stored within the server-side database. When a client seeks to make a query, they include this token within the authorization header of their request (see the user manual for more details). We developed the desktop

application using the Electron framework (with HTML, CSS, and JavaScript), which makes OncoReport a cross-platform tool. The graphical user interface and the input information the oncologist can provide for each patient are illustrates in Fig. 2a, b and c and include: Personal information, such as name, surname, age, and email. Fig. 2d shows the summary of patients personal information. Disease history (Fig. 2e) lists the patient's previous or current diseases, with their start and end dates. A full star next to a disease (e.g., colorectal cancer) indicates that it is the condition of interest for the analysis, and the results will be tailored to the suggested therapies and guidelines for that disease. The user can add a new disease by clicking on the "+" button, which will open a form to enter the details of the disease, such as diagnosis and remission dates, primary or secondary status, and TNM staging for cancer cases. Drug history (Fig. 2f) shows the drugs administered to the patient, with their start and

end dates. The user can add a new drug by clicking on the "+" button, which will open a form to enter the details of the drug, such as name, dosage, frequency, and route of administration. When adding a new drug, the user can specify the disease (s) that the drug is prescribed for and the duration of the treatment. If applicable, the user can also indicate the reason for stopping the drug. The last tab, from which it is possible to run a new analysis, shows the list of all the analyses performed on the patient (Fig. 2g). Next to each analysis record, buttons allow the user to view the analysis within the application, download the analysis as a zip archive that contains all the HTML and CSS files, remove the analysis from the system. Fig. 3a displays the analysis summary dashboard, where users can initiate a new analysis by clicking the "plus" button. Fig. 3b–d illustrate the three sequential steps required to launch an analysis: (i) entering the analysis metadata, and

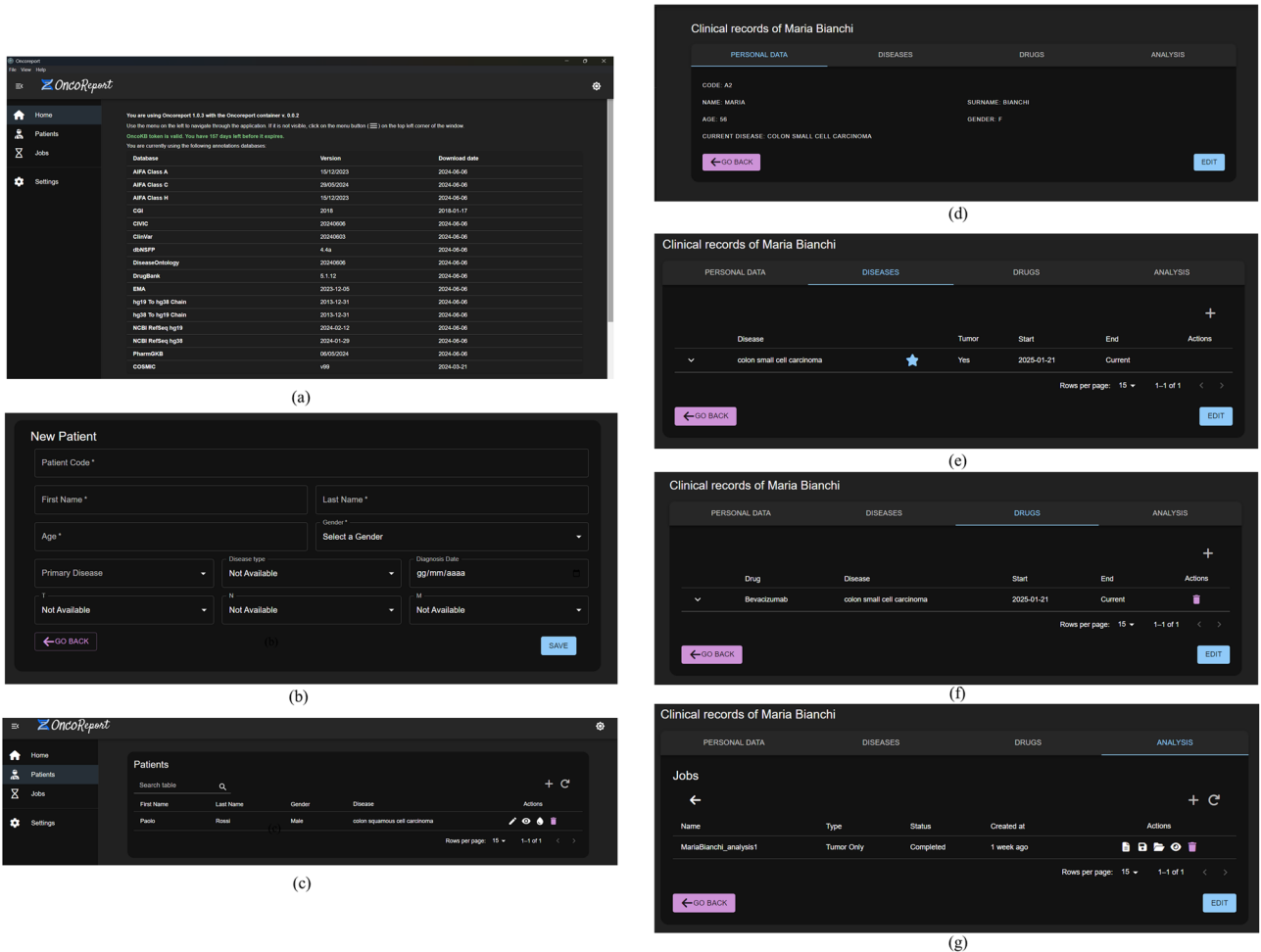


Fig. 2 OncoReport main interface: (a) General app dashboard, on the left, the user has a menu with all the available choices; (b) by selecting the entry patient in the menu, the user gets the patient list stored in the database; (c) patient personal data with the list of past and current patient diseases, with their start and end dates, primary or secondary status, and TNM staging for cancer cases; (f) patient personal data with the list of current and past drug treatments, with their start and end dates, dosage, frequency, route of administration, and reason for discontinuation. *The information displayed in the figure is for demonstration purposes only and does not reflect actual patient data

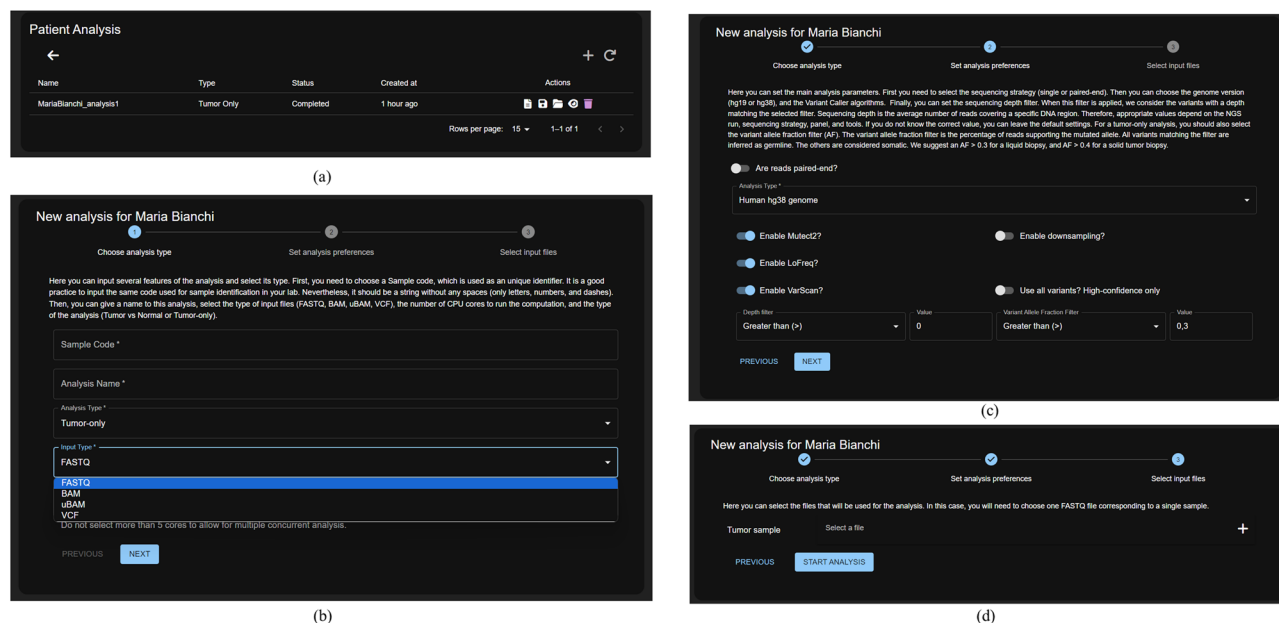


Fig. 3 Steps that the user must follow to perform the analysis of patient data. (a) The jobs entry of the menu allows to see the current and the completed analysis and to start a new analysis; (b) in step 1 it is possible to select sample type providing the sample code and the name of the analysis; in the second step select the custom analysis she/he wants to run and in the third step select the input files. *The information displayed in the figure is for demonstration purposes only and does not reflect actual patient data

selecting the input type, (ii) filtering parameters, and (iii) uploading the sample files.

The HTML report

The report (see Fig. 4) consists of different sections that can be accessed from a NavBar:

- Patient information: This section displays the personal and medical information of the patient that the user entered when uploading the NGS files (see Fig. 4a).
- Therapeutic indications: This section highlights the essential variants found in the patient's tumor and their associated drugs, clinical trials, confidence scores, and publication years. The drugs are sorted by evidence level, starting from the ones approved by the FDA or NCCN, followed by the ones used only in trials, and finally, tested only in vitro. The section also shows the approval information for each drug by different institutions, such as FDA, EMA, and AIFA, to help the user understand which drug is feasible in their country. The section has two tables: one for the variant-drug pairs with clinical evidence or approval and one for those found in case studies or in vitro experiments (see Fig. 4b).
- Drug-drug interactions: This section lists the possible interactions between the drugs recommended by OncoReport and the drugs already taken by the patient in one table and between the drugs

recommended by OncoReport and all the DrugBank drugs in another table (see Fig. 4c).

- Drug-food interactions: This section lists the possible interactions between the drugs recommended by OncoReport and food items, using DrugBank as a source (see Fig. 4d).
- ESMO guidelines: This section shows the European Society for Medical Oncology (ESMO) guidelines related to the patient's cancer type. ESMO is the leading entity in Europe for disseminating best practices for cancer prevention, diagnosis, treatment, and follow-up. This section shows a scrollable list of all clinical practice guidelines relevant to the patient's disease on the left and the details of each guideline in the form of text or dynamic algorithm at the center of the page (see Fig. 4e).
- Drug response: This section lists the variants found in the PharmGKB database that affect the efficacy or toxicity of a drug, regardless of the disease specification.
- Mutation annotations: This section presents a table with all the mutation annotations, using RefGene and ClinVar as sources. It focuses on the function and clinical significance of each mutation.
- Off-label drugs: This section includes the variants that are associated with drugs in tumors different from the patient's, according to current knowledge.
- Known resistance: This section contains annotations from the COSMIC database that indicate drug-resistant mutations.

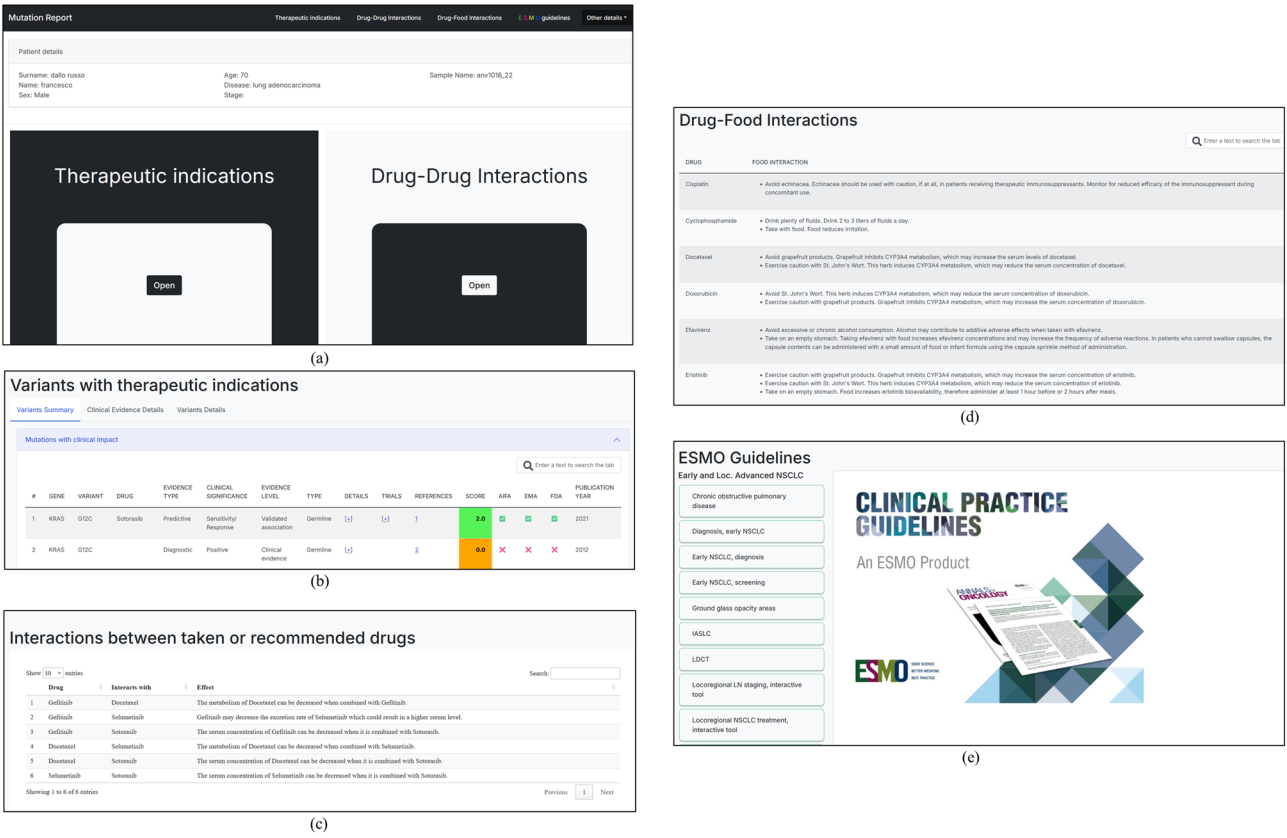


Fig. 4 The report generated by the system consists of a series of sections. Here we can see: **(a)** the report’s main interface comes with a list of options allowing inspecting the details of the generated clinical information. Among the sections, we cite **(b)** the list of all variants with the therapeutic indications. For each variant, we report the name of the gene, the detailed variant, the associated drug with a score, the prescription agencies (European, US, and Italian); **(c)** the interaction between the recommended drugs. Other sections include: the drug-food interactions; the ESMO (European Society for medical oncology) guidelines associated to the case and other details. *The information displayed in the figure is for demonstration purposes only and does not reflect actual patient data

- **References:** This section provides the literature references for each feature in the report, which can be accessed through clickable PMIDs.

Validation and case study
Insights from OncoReport analysis

Of the 40 cases analyzed, we have found 7 featured variants in 7 cases listed in the therapeutic indications section, which aligned with findings in expert-generated reports (true positives). Conversely, 33 cases revealed no variants in the therapeutic indications section, and no variants were reported in the corresponding clinical reports, resulting in no false positives or false negatives. This demonstrates 100% sensitivity and specificity (see Additional file I and Additional file II). OncoReport often provides valuable off-label therapeutic indications even for true-negative cases, offering oncologists alternative options for patient care. For instance, in a case involving a 69-year-old patient with colon cancer, OncoReport identified Imatinib as a therapeutic option linked to the KIT-M541L mutation. This insight extended beyond the

oncologist’s original focus on MMR, enhancing diagnostic possibilities. Similarly, for a 77-year-old female patient with lung carcinoma, OncoReport highlighted associations between EGFR mutations and various drugs used in non-small cell lung cancer, despite the absence of diseasespecific therapeutic indications. Moreover, four representative cases were selected and described with the primary aim of providing a clear overview of the potential applications of the OncoReport tool in typical clinical scenarios. Rather than constituting a systematic evaluation, these cases serve as illustrative examples to explore the appropriateness of advanced indications—such as off-label therapies, variants included in the “other evidence” section, and clinical trial eligibility. The selected cases include two lung cancer cases, one ovarian cancer case, and one breast cancer case.

1. **70-Year-Old Male with Non-Small Cell Lung Cancer.** The patient did not exhibit typical oncogene-addicted mutations suitable for routine targeted therapy. However, the tumor presented

an R130Q mutation in the PTEN gene, linked to sensitivity to PTEN inhibitors. Based on this, an active phase 2 clinical trial (NCT06183736) was identified as a potential treatment option. Additional findings included a PIK3CA mutation for potential off-label treatment and STK11 variants, potentially linked to response to Pembrolizumab/Bemcentinib combination therapy. However, the latter evidence was discarded due to a lack of available interventional options.

2. **58-Year-Old Patient with Non-Small Cell Lung Cancer.** No ESCAT TIER 1 mutations were found. The tumor displayed a targetable HER2 mutation, suggesting potential off-label treatment with neratinib. OncoReport identified an active phase 2 clinical trial (NCT06519110) for which the patient might be eligible.
3. **Woman under the age of 50 with breast cancer and a family history of ovarian and prostate cancers.** The clinical question concerned the oncological risk associated with pathogenic variants in BRCA1 and BRCA2, both for potential therapeutic implications for the patient and as a possible criterion for including her children in targeted screening programs. The test performed was a targeted NGS panel for BRCA1 and BRCA2. The analysis did not reveal any pathogenic variants relevant to the diagnostic question. Therefore, the patient may be referred for second-tier testing to investigate the presence of potential variants in other high or moderate-penetrance genes.
4. **Young woman with ovarian cancer.** An assessment of the presence of pathogenic variants in BRCA1 and BRCA2 is required for a woman diagnosed with high-grade serous ovarian carcinoma, for therapeutic purposes. The identification of such variants is, in fact, a criterion for eligibility for treatment with PARP inhibitors. In this specific case, the analysis did not identify any class 4 or 5 variants in the BRCA1 or BRCA2 genes. Therefore, the patient may be referred for further investigation through HRD testing, in order to define an appropriate therapeutic strategy.

Detailed case study

Stage IV unresectable colon adenocarcinoma A patient with stage IV unresectable colon adenocarcinoma, treated at the IOM in Viagrande, Italy, was analyzed using OncoReport. The VCF file was derived from a routine NGS pipeline using the AmoyDx Classic Handle Panel. The tumor exhibited a BRAF V600E mutation, and OncoReport provided eight clinical pieces of evidence related to drug response or resistance, along with 13 additional findings. The most supported treatment recom-

mendation was a combination of Binimetinib, Cetuximab and Encorafenib. This combination is approved by AIFA, EMA, and FDA.

Alternative options Cetuximab, Irinotecan, and Vemurafenib: One drug lacks EMA approval. However, three recruiting clinical trials offer access to this combination. Dabrafenib, Panitumumab, and Trametinib: This combination is fully approved by AIFA, EMA, and FDA.

Additional considerations The drug-food interactions section advises avoiding grapefruit products and St. John's Wort with Encorafenib. Drug-drug interactions suggest avoiding combinations such as Cetuximab with Bevacizumab or Panitumumab and Binimetinib with Sorafenib, Erlotinib, Cobimetinib, or Vemurafenib. Similarly, Encorafenib should not be combined with drugs such as Gefitinib or Oxaliplatin.

COSMIC database insights

Cetuximab monotherapy is contraindicated for large intestine malignancies due to resistance associated with the BRAF V600E mutation. In contrast to conventional machine learning approaches, OncoReport is not a predictive model. Rather, it functions as a knowledge-driven system that seamlessly integrates heterogeneous sources of information. By aggregating and connecting insights from multiple curated databases, it enables the discovery of clinically relevant associations that may not be immediately evident to the user.

Benchmark

We benchmarked OncoReport against five existing DNA sequencing (DNaseq) pipelines designed for variant calling and annotation. An overview of their primary features and specifications is detailed in the Introduction and summarized in Table 1. We evaluated each tool using a representative TCGA colon cancer sample (TCGA-AY-A54L). All experiments were conducted on a Windows 11 laptop running Docker version 4.47, equipped with an AMD Ryzen 7 PRO 7840HS processor (Radeon 780 M Graphics) and 32 GB of RAM.

The comparison included the following pipelines: (i) SEQprocess, last updated in 2020, could not be tested as several requisite packages are now deprecated. (ii) SEQVita, also last updated in 2020, requires installation via GitHub and relies on pre-aligned, sorted BAM files as input. It is compatible only with the hg19 genome assembly and demands manual installation of dependencies—many of which are undocumented—necessitating advanced technical expertise. (iii) DNAscanv2, last updated in 2023, is available as a standalone installation or a Docker container. However, the Docker container

Table 1 Comparison of OncoReport with five free-use software. *Analysis failed

	OncoReport	SEQprocess	SEQVita	DNAscanv2	PCGR	Lush-DNAseq-pipeline	MTBP
Accepted data types	FASTQ, BAM, uBAM, SAM, VCF, VariantTable	FASTQ	BAM	FASTQ, SAM, BAM, CRAM, VCF	VCF, TXT, TSV	FASTQ	VCF, Variant list
Language	PHP, Bash, HTML, R	R	C++, R, bash	HTML, Python, Bash	Python, R	Bash	Online tool
Output of clinical interpretation	HTML Report	Text	VCF	Text	HTML Report	PDF/Word Report, variant annotation Excel	HTML Report
Pipeline	DNA/RNA	DNA/RNA	DNA	DNA	DNA/RNA	DNA	DNA
Annotation	RefSeq, dbNFSP, ClinVar, COSMIC, CIViC, CGI, OncoKB, PharmGKB	VEP, ANNOVAR	SIFT, Polyphen2, MutationTaster, PhyloP, LRT, ClinVar, OMIM, COSMIC, DECIPHER, PharmGKB	refGene, ClinVar, EXAC, dbSNP, CADD, gnomAD, 1000g	VEP, CIViC, CGI, dbNFSP, gnomAD, dbSNP, Cancer Hotspot, ClinVar, UniProt, Pfam, CancerMine, Open Targets Platform	None	OncoKB, Clinvar, BRCA-Exchange, CIViC
Graphical User Interface	Yes	No	No	Yes	No	No	No
Running time	19h	NA	20h18m*	42 m*	2 m	37 m*	NA
Last Release	01/2024	Unmantained	2020	2023	09/2025	2024	2025
Analysis Type	Tumor vs Normal, Tumor Only	Tumor vs Normal, Tumor Only	Tumor vs Normal, Tumor Only	Tumor Only	Tumor vs Normal, Tumor Only	Tumor only	Tumor vs Normal, Tumor Only

failed to execute, and the standalone version required significant code modifications to run. Even after patching, the pipeline aborted post-alignment due to environment conflicts between Strelka (requiring Python 2) and the recommended Python 3 environment. (iv) LUSH-DNAseq-pipeline, last updated in 2024, operates exclusively in tumor-only mode. It requires users to generate annotation indices independently; however, index generation failed on our system, preventing the alignment step from completing. (v) MTBP offers a free web-based interface accepting VCF files or variant lists of up to 2000 entries. As an online tool, it presents data privacy challenges for clinical applications involving sensitive patient information.

Given the operational failure of three pipelines, our comparative analysis focused primarily on annotation accuracy and the ease of clinical interpretation for the remaining functional tools.

Materials and methods

OncoReport represents a comprehensive pipeline designed to analyze and interpret NGS data, emphasizing gene annotations, identifying potential therapeutic indications thus supporting clinical actions that are tailored on patient specific clinical and molecular profile. In the upcoming sections, we provide an in-depth description of the annotation process and analysis on the databases accessible within our toolkit.

Databases and knowledge bases

OncoReport integrates data from several external databases (see Table 2) to provide comprehensive and reliable information on cancer variants and their clinical implications. These databases include:

- CIViC [22] is an open-source database that provides curated and evidence-based information on the clinical relevance of inherited and somatic variants in cancer. As of February 2025, it contains 3793 variants in 872 features, with 10,692 evidence items that describe their therapeutic, prognostic, diagnostic, or predisposing implications.
- CGI [23] is a free platform that interprets the molecular alterations found in tumors and identifies their potential therapeutic implications based on clinical evidence. It covers 5601 oncogenic alterations in 765 cancer genes and 1631 drug response biomarkers linked to approved or experimental drugs.
- PharmGKB [24] is a resource that integrates and curates knowledge on how human genetic variations affect drug responses. It offers clinically relevant information such as dosing guidelines that recommend optimal drug doses based on genotype; annotated drug labels that highlight pharmacogenetic information from FDA-approved labels; potentially actionable gene-drug associations that suggest possible therapeutic options based on

Table 2 The complete list of databases used in OncoReport. For each database, we report the name, a description of the content, the URL, and its reference. We do not report the version used in our pipeline since databases are automatically updated to the latest version

Database	Description
CIViC	Clinical Interpretation of Variant in Cancer —"CIViC is an open access, open source, community-driven web resource for Clinical Interpretation of Variants in Cancer" [22]
CGI	Cancer Genome Interpreter —"CGI identifies potentially oncogenic alterations, and it flags genomic biomarkers of drug response with different levels of clinical relevance" [23]
COSMIC	Catalogue of Somatic Mutations —"COSMIC is the world's largest and most comprehensive resource for exploring the impact of somatic mutations in human cancer" [25]
PharmGKB	Pharmacogenomics Knowledge Base —"PharmGKB is a comprehensive resource that curates knowledge about the impact of genetic variation on drug response for clinicians and researchers" [24]
ClinVar	(No full name provided) —"ClinVar aggregates information about genomic variation and its relationship to human health" [6]
RefSeq	NCBI Reference Sequence Database [26]
ESMO	European Society for Medical Oncology —"ESMO is the leading European professional medical oncology organization. Its Clinical Practice Guidelines (CPG) are intended to provide the user with a set of recommendations for the best standards of cancer care, based on the findings of evidence-based medicine."
DrugBank	"DrugBank is a unique bioinformatics/cheminformatics resource that combines detailed drug (i.e., chemical) data with comprehensive drug target (i.e., protein) information" [28]
OncoKB	Oncology Knowledge Base —"OncoKB™ is a precision oncology knowledge base developed by Memorial Sloan Kettering Cancer Center. It provides comprehensive biological and clinical insights into genomic alterations associated with cancer, serving as a vital resource for advancing cancer diagnostics and treatment." [38, 39]

genetic evidence; genotype-phenotype relationships that show how genetic variants influence drug efficacy, toxicity, or metabolism.

- COSMIC [25] is a resource that catalogs somatic mutations' impact on human cancers. It contains 9,215,470 curated variants from 29,682 papers as of version 101. It covers various types of mutations, such as non-coding mutations that affect regulatory

regions or non-coding RNAs, gene fusions that result from chromosomal rearrangements or translocations, copy-number variants, and drug-resistance mutations that confer resistance to anti-cancer drugs.

- ClinVar [5, 6] is a resource that archives and interprets the clinical significance of genetic variants and their associated phenotypes. It reports how mutations affect human health and the evidence supporting each interpretation from various sources.
- RefSeq [26] is a project that provides curated and annotated genomic sequences for genes and transcripts. It is maintained by the National Center for Biotechnology Information (NCBI), and it includes information on the nucleotide sequences and their protein products, as well as their functional and evolutionary features.
- OncoKB [38, 39] is a comprehensive database that annotates somatic variants with their biological and oncogenic effects, as well as their predictive and prognostic significance. Drug-variant associations in OncoKB are categorized by evidence levels, based on FDA labeling and other authoritative sources. As of February 2025, the database encompasses 905 genes, 7909 alterations, 142 cancer types, and 148 drugs, making it a critical resource for precision oncology.
- DrugBank [27, 28] is a web resource that provides comprehensive and reliable information on drugs and their interactions. It was launched in 2006 and has evolved over the years to include various types of data, such as pharmacogenomic data that describe how genetic variations affect drug responses; molecular data detailing the chemical structures, properties, and mechanisms of action of drugs; and few more. As of February 2025, DrugBank contains information on 17,449 drugs, including approved, experimental, and natural products. DrugBank aims to support advancing medical research and practice by providing a comprehensive and accessible drug knowledgebase.

Validation method

The validation of OncoReport was conducted to assess its ability to provide clinically relevant results in a real-world setting. A retrospective study was performed using NGS data from genomic profiling activities carried out for diagnostic purposes, alongside the corresponding clinical reports generated by diagnostic staff. The validation criteria focused on the following metrics:

- True Positives (TP): Cases where a variant listed in OncoReport's therapeutic indications matched a clinically relevant variant in the diagnostic reports.

- True Negatives (TN): Cases where no therapeutic variant was identified by OncoReport, and the corresponding clinical report did not mention any clinically relevant variants.
- False Positives (FP): Cases where a variant listed in the therapeutic indications was not included in the clinical report and was deemed irrelevant to the clinical case.
- False Negatives (FN): Cases where no variants were listed in OncoReport's therapeutic indications, but clinically relevant variants were reported in the diagnostic findings.
- Sensitivity: Value that represents the ability of OncoReport to recognize a mutations for its real role with a correct suggestion for the clinician.
- Specificity: This value refers to the OncoReport ability to correctly recognize the patients without actionable mutation, without having false associations.

A total of 40 cases, selected from a patient cohort at the Mediterranean Institute of Oncology (IOM), were evaluated. Various NGS panels were utilized in the diagnostic routine, including BRCA1 and BRCA2 targeted panels, a 40-gene pan-cancer somatic mutation panel, and a custom 37-gene panel for hereditary cancer. Given the retrospective nature of the study and its focus on routine oncology management—which excludes indications for experimental or off-label therapies—the validation primarily examined therapeutic indications. For specific cases where criteria for experimental or off-label therapies were met, experts re-evaluated the results to determine the appropriateness of OncoReport's recommendations. Drug–mutation associations classified by biological databases as 'Case Study,' 'Preclinical Evidence,' or 'Inferential Association' are considered experimental. Off-label treatments refer to the use of targeted therapies beyond their approved indications; specifically, in this context, they involve applying a clinically validated drug–mutation association from one cancer type to a different cancer type relevant to the case under consideration. Validation relied on de-identified patient clinical and pathological data retrieved from the centralized electronic case report forms system at the IOM. Sequencing data files were also obtained from the same institution. Data transfer, storage, and access adhered to robust technical measures aligned with the European legal framework, ensuring compliance with data protection regulations. Ethical approval declarations This study was approved by the Ethics Committee of Catania 2 (protocol no. 116/C.E.). All methods were carried out in accordance with the Declaration of Helsinki and with relevant national and institutional guidelines and regulations. Anonymized NGS data were obtained at the

Istituto Oncologico del Mediterraneo (IOM) from routine diagnostic procedures. Written informed consent was obtained from all participants prior to sample collection and data use for research purposes.

Discussion

OncoReport is a flexible tool that generates comprehensive reports from NGS data to support clinical decision-making in cancer (see Fig. 4). OncoReport is a user-oriented platform developed to support a broad spectrum of professionals—clinicians, biologists, and laboratory technicians—regardless of their bioinformatics expertise. It is adaptable to diverse analytical contexts, enabling its application to both host and tumor genomic data according to the specific needs of the user. OncoReport positively impacts precision-medicine by enabling rapid identification of actionable genomic alterations. Its capacity for off-label drug identification makes it particularly valuable for clinical programs that rely on drug repurposing strategies, such as the Drug Rediscovery Protocol [40] and IMPRESS-Norway [41], where matching existing therapeutics to novel molecular targets is essential for expanding treatment options. Drawing the recommendations of several research groups, we developed a report format that is both easy to interpret and rich in references, thereby supporting clinical decision-making as well as future research. These groups have investigated how to design genetic reports that could be comprehensible to both clinicians and patients [42–44].

In particular, in line with Farmer et al. [44], we engaged clinicians from IOM Ricerca to iteratively refine the interpretability of the report. Our primary objective was to present next-generation sequencing (NGS) results in a clear, unambiguous manner to aid clinicians in selecting the most appropriate therapy for each individual patient. Following best practices suggested in the literature, we prioritized the presentation of clinically actionable information—specifically, mutation–drug associations—at the forefront of the report. To further enhance usability, we assigned scores to each association, providing clinicians with a ranked, urgency- and relevance-based overview. This scoring system facilitates communication with patients and helps prioritize therapeutic options, as advocated by Cutting et al. [42].

To improve readability, we organized results into schematized tables and structured the content to directly address three critical questions: What does the result mean? What action should be taken? Where can further information and support be found? Finally, as emphasized by Deans et al. [43], patient-specific identifiers and clinical context—entered by the OncoReport user—are embedded within the report to ensure unambiguous linkage to the individual case, an essential element for effective care delivery. OncoReport is compatible with both

liquid biopsy and tissue NGS data, supporting diverse analysis modalities including Whole Genome Sequencing (WGS), Whole Exome Sequencing (WES), and targeted panels. System requirements vary by application: WGS analysis mandates a minimum of 64 GB of RAM, eight cores, and 1 TB of storage, whereas WES analysis requires at least 16 GB of RAM with a runtime of up to 24 hours. For custom panels, computational resources and execution time scale with the panel size, while analysis is significantly faster when initiated from VCF inputs. We evaluated OncoReport's performance on a laptop equipped with 32 GB of RAM, utilizing 5 threads. Starting from FASTQ files, the comprehensive analysis of tumor-normal WES samples (combined input size: 48 GB) required 19 hours and consumed approximately 150 GB of disk space. Conversely, when processing pre-computed VCF files for the same samples, annotation and report generation were completed within 1 hour. Execution time can be further reduced by increasing thread allocation to leverage parallel processing across multiple cores.

An user manual explaining how to use OncoReport is provided at <https://github.com/knowmics-lab/oncoreport/wiki>.

Limitations and future perspective

Several limitations of the validation process warrant acknowledgment. First, relying on diagnostic reports as ground truth, while reflective of real-world clinical decision-making, may introduce variability stemming from subjective interpretation and institution-specific protocols. It is important to note, however, that these reports were generated by experienced clinicians utilizing standardized, validated workflows aligned with rigorous national (AIOM, SIAPEC, SIGU) and international (ESMO, ACMG) guidelines. Consequently, the criteria adopted at the IOM for variant classification and therapeutic interpretation, serving as the reference framework for OncoReport, are consistent with established clinical standards and routine practices in oncology centers throughout Italy and Europe. Although the single-center design inherently limits the breadth of validation, it nonetheless provides a representative foundation for evaluating the tool's clinical utility in a real-world context. We acknowledge that multi-center validation is essential to confirm reproducibility across diverse institutional settings and patient populations; accordingly, this has been established as a primary objective for future research.

Second, the validation was conducted on a limited set of retrospective cases. Although this dataset was adequate to demonstrate the practical applicability of OncoReport and its alignment with clinical reasoning, its generalizability remains to be further assessed. As with any clinical decision support system, the tool's

performance requires ongoing evaluation, particularly as it is integrated into broader clinical settings.

OncoReport, as an open-source platform, is designed with adaptability in mind. It is expected to evolve through iterative updates informed by contributions and feedback from the clinical and scientific community.

Finally, the current version of OncoReport does not support the analysis of copy number variants (CNVs), representing a limitation to be addressed in future developments. Additionally, integration of GWAS and eQTL data is not yet available but is planned for the next OncoReport release.

Conclusion

We present OncoReport, a comprehensive clinical bioinformatics framework designed for seamless integration as a microservice. OncoReport accelerates sample analysis and annotation in clinical settings, providing an efficient solution for oncology. It enables automated, end-to-end analysis of samples from custom panels or whole-exome sequencing (WES).

The generated reports deliver actionable clinical insights with high efficiency, enabling timely and informed therapeutic decision-making. By streamlining the analytical workflow and alleviating the technical burden on clinical users, OncoReport is positioned to accelerate the integration of next-generation sequencing (NGS) into routine practice, particularly for the detection of emerging and non-conventional biomarkers.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-026-07718-w>.

Supplementary Material 1: Additional file 1 - .xls - Validation table of 40 patients analyzed with OncoReport and by a clinician searching for variation with therapeutic evidence

Supplementary Material 2: Additional file 2 - .xls - Validation table of 40 patients analyzed with OncoReport and by a clinician

Supplementary Material 3: Additional file 3 - .xls - Off-label variants of 40 patients analyzed with OncoReport. Some patients are missing due to the lack of actionable off-label variants

Acknowledgements

We sincerely thank Dr Santina Cristina Gorgone (Istituto Oncologico del Mediterraneo) for accurate testing the user interface with real clinical cases.

Author contributions

AP, SF, SA, LG, and AF conceived the project. GFP and SA developed the system. EM and RV deeply tested the tool and developed the APIs. MM deeply analyzed and tested the tool within the ward at Istituto Oncologico del Mediterraneo. SPL validated the system. GFP wrote the first draft of the paper. All authors read and approved the final version of the manuscript.

Funding

This research was partially funded by the department for productive activities of the Sicilian Region, project title: "DiOncoGen Diagnostica innovativa", PO FESR (G89J18000700007) and by the project "OMICANCER: Modelli computazionali per l'identificazione di marcatori in oncologia tramite

analisi Multi-Omica" (CUP E63C24001410001) Funded by the European Union – Next Generation EU, Missione 4 Componente 2 Inv. 1.5 (CUP Master B63C22000680007). This study was also funded by the 2024/2026 Research Plan of University of Catania Pia. ce.ri (IMAGINE project).

Data availability

The datasets supporting the conclusions of this article is available in zenodo <https://doi.org/10.5281/zenodo0.14704552>.

Declarations

Ethics approval and consent to participate

The study is conducted in accordance with the Declaration of Helsinki, complied with national ethical guidelines and was approved by the Ethical Committee "Catania 2" under protocol number 116, dated February 22, 2022. Written informed consent was obtained from all participating patients.

Competing interests

The authors declare that a patent regarding the primary technology described in the manuscript has been granted (Italian Patent Office, patent number: 102,023,000,012,882 of 16/06/2025). All authors declare no other financial or non-financial competing interests.

Author details

¹Department of Clinical and Experimental Medicine, Bioinformatics Unit, University of Catania, Via Santa Sofia 89, 95123 Catania, Italy

²Istituto Oncologico del Mediterraneo, Via Penninazzo 7, 95029 Viagrande, Italy

Received: 30 September 2025 / Accepted: 31 December 2025

Published online: 21 January 2026

References

1. Joo T, Choi JH, Lee JH, Park SE, Jeon Y, Jung SH, et al. Seqprocess: a modularized and customizable pipeline framework for NGS processing in R package. *BMC Bioinf.* 2019 Feb;20(1):90.
2. Dharanipragada P, Seelam SR, Parekh N. SeqVItA: sequence variant identification and annotation platform for next generation sequencing data. *Front Genet.* 2018 Nov;9:537.
3. Iacoangeli A, Al Khleifat A, Sproviero W, Shatunov A, Jones AR, Morgan SL, et al. Dnascan: personal computer compatible NGS analysis, annotation and visualisation. *BMC Bioinf.* 2019 Apr;20(1):213.
4. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010 Sep;38(16):e164.
5. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* 2014 Jan;42(Database issue):D980–5.
6. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* 2016 Jan;44(D1):D862–8.
7. Karczewski KJ, Weisburd B, Thomas B, Solomonson M, Ruderfer DM, Kavanagh D, et al. The ExAC browser: displaying reference data information from over 60000 exomes.
8. Sherry ST. dbSNP: the NCBI database of genetic variation.
9. Liu X, Jian X, Boerwinkle E. dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. *Hum Mutat.* 2011 Aug;32(8):894–99.
10. Kronic M, Venhuizen P, Müllerauer L, Kaserer B, von Haeseler A. VARIFI-Web-based automatic variant identification, filtering and annotation of Amplicon sequencing data. *J Pers Med.* 2019 Feb;9(1).
11. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nat Methods.* 2012 Mar;9(4):357–59.
12. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009 Jul;25(14):1754–60. <https://doi.org/10.1093/bioinformatics/btp324>.
13. Sedlazeck FJ, Rescheneder P, von Haeseler A. NextGenMap: fast and accurate read mapping in highly polymorphic genomes. *Bioinformatics.* 2013 Nov;29(21):2790–91.
14. Wang T, Zhang Y, Wang H, Zheng Q, Yang J, Zhang T, et al. Fast and accurate DNaseq variant calling workflow composed of LUSH toolkit. *Hum Genomics.* 2024;18:114. <https://doi.org/10.1186/s40246-024-00666-w>.
15. Dahary D, Golan Y, Mazor Y, Zelig O, Barshir R, Twik M, et al. Genome analysis and knowledge-driven variant interpretation with TGex. *BMC Med Genomics.* 2019 Dec;12(1):200.
16. Hintzsche J, Kim J, Yadav V, Amato C, Robinson SE, Seelenfreund E, et al. IMPACT: a whole-exome sequencing analysis pipeline for integrating molecular profiles with actionable therapeutics in clinical samples. *J Am Med Assoc.* 2016 Jul;32(4):721–30.
17. Home. Accessed: 2021-6-7. <https://impactdatabase.eu/>.
18. Hintzsche JD, Yoo M, Kim J, Amato CM, Robinson WA, Tan AC. IMPACT web portal: oncology database integrating molecular profiles with actionable therapeutics.
19. Nakken S, Fournous G, Vod'ak D, Aasheim LB, Myklebost O, Hovig E. Personal cancer Genome Reporter: variant interpretation report for precision oncology. *Bioinformatics.* 2018 May;34(10):1778–80.
20. Reardon B, Moore ND, Moore NS, Kofman E, AlDubayan SH, Cheung ATM, et al. Integrating molecular profiles into clinical frameworks through the molecular oncology Almanac to prospectively guide precision oncology. *Nat Cancer.* 2021 Oct;2(10):1102–12.
21. Tamborero D, Dienstmann R, Rachid MH, Boekel J, Lopez-Fernandez A, Jonsson M, et al. The molecular tumor board portal supports clinical decisions and automated reporting for precision oncology. *Nat Cancer.* 2022 Feb;3(2):251–61.
22. Griffith M, Spies NC, Krysiak K, McMichael JF, Coffman AC, Danos AM, et al. CIVIC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. *Nat Genet.* 2017 Jan;49(2):170–74.
23. Tamborero D, Rubio-Perez C, Deu-Pons J, Schroeder MP, Vivancos A, Rovira A, et al. Cancer genome interpreter annotates the biological and clinical relevance of tumor alterations. *Genome Med.* 2018 Mar;10(1):25.
24. Whirl-Carrillo M, McDonagh EM, Hebert JM, Gong L, Sangkuhl K, Thorn CF, et al. Pharmacogenomics knowledge for personalized medicine. *Clin Pharmacol Ther.* 2012 Oct;92(4):414–17.
25. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res.* 2019 Jan;47(D1):D941–47.
26. O'Leary NA, Wright MW, Brister JR, Ciufio S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 2016 Jan;44(D1):D733–45.
27. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* 2006 Jan;34(Database issue):D668–72.
28. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018.
29. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads.
30. FelixKrueger: FelixKrueger/TrimGalore. Accessed: 2021-6-1. <https://github.com/FelixKrueger/TrimGalore>.
31. Picard. Accessed: 2021-6-1. <http://broadinstitute.github.io/picard/>.
32. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010 Sep;20(9):1297–303.
33. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 2012 Mar;22(3):568–76.
34. Wilm A, Aw PPK, Bertrand D, Yeo GHT, Ong SH, Wong CH, et al. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cellpopulation heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res.* 2012 Dec;40(22):1189–201.
35. O'Leary NA, Cox E, Holmes JB, et al. Exploring and retrieving sequence and metadata for species across the tree of life with NCBI datasets. *Sci Data.* 2024;5(11(1):732. <https://doi.org/10.1038/s41597-024-03571-y>.
36. Liu X, Li C, Mou C, Dong Y, Tu Y. dbNSFP v4: a comprehensive database of transcript-specific functional predictions and annotations for human nonsynonymous and splice-site SNVs. *Genome Med.* 2020;12(1):103. <https://doi.org/10.1186/s13073-020-00803-9>.
37. Laravel. Laravel Sanctum. Available from: <https://github.com/laravel/sanctum>.
38. Suehnholz SP, Nissan MH, Zhang H, Kundra R, Nandakumar S, Lu C, et al. Quantifying the expanding landscape of clinical actionability for patients with cancer. *Cancer Discov.* 2024 Jan;14(1):49–65.

39. Chakravarty D, Gao J, Phillips SM, Kundra R, Zhang H, Wang J, et al. OncoKB: a precision oncology knowledge Base. *JCO Precis Oncol*. 2017; 2017 Jul. <https://doi.org/10.1200/PO.17.00011>.
40. van der Velden D, Hoes L, van der Wijngaart H, van Berge Henegouwen J, van Werkhoven E, Roepman P, et al. The drug Rediscovery protocol facilitates the expanded use of existing anticancer drugs. *Nature*. 2019;574:7776. <https://doi.org/10.1038/s41586-019-1600-x>.
41. Helland *A, Russnes H, Fagereng G, Al-Shibli K, Andersson Y, Berg T, et al. Improving public cancer care by implementing precision medicine in Norway: IMPRESS-Norway. *J Transl Med*. 2022;20:317. <https://doi.org/10.1186/s12967-022-03432-5>.
42. Cutting E, Banchemo M, Beitelshes AL, Cimino JJ, Fiol GD, Gurses AP, et al. User-centered design of multi-gene sequencing panel reports for clinicians. *J Biomed Inf*. 2016;63:1–10. <https://doi.org/10.1016/j.jbi.2016.07.014>.
43. Deans ZC, Ahn JW, Carreira IM, Dequeker E, Henderson M, Lovrecic L, et al. Recommendations for reporting results of diagnostic genomic testing. *Eur J Hum Genet EJHG*. 2022;30(9):1011–16. <https://doi.org/10.1038/s41431-022-01091-0>.
44. Farmer GD, Gray H, Chandratillake G, Raymond FL, Freeman ALJ. Recommendations for designing genetic test reports to be understood by patients and non-specialists. *Eur J Hum Genet EJHG*. 2020;28(7):885–95. <https://doi.org/10.1038/s41431-020-0579-y>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.