

Protocol

Epigenetic Profiling for Early Detection and Treatment Response Monitoring in Non–Small Cell Lung Cancer: Protocol for a Prospective Translational Biomarker Study

Rajiv Kumar^{1,2,3}, BMedSci, MBChB, MD(Res); Safia May Farry¹, BBiomedSc(Hons); Mark Ezegbogu¹, MSc; Glen Reid¹, PhD; Euan James Rodger¹, PhD; Ben Brockway^{2,4*}, MBBS; Aniruddha Chatterjee^{1*}, PhD

¹Department of Pathology and Molecular Medicine, Faculty of Medicine - Dunedin, University of Otago, Dunedin, New Zealand

²Department of Medicine, Faculty of Medicine - Dunedin, University of Otago, Dunedin, New Zealand

³Mercy Cancer Care, Mercy Hospital, Dunedin, New Zealand

⁴Te Whatu Ora – Southern, Health New Zealand, Dunedin, New Zealand

*these authors contributed equally

Corresponding Author:

Aniruddha Chatterjee, PhD

Department of Pathology and Molecular Medicine, Faculty of Medicine - Dunedin

University of Otago

58 Hanover Street, Hercus Building

Dunedin 9016

New Zealand

Phone: 64 210701558

Email: aniruddha.chatterjee@otago.ac.nz

Abstract

Background: Non–small cell lung cancer (NSCLC) is the leading cause of cancer-related mortality worldwide and continues to have poor survival outcomes, with most patients diagnosed at advanced stages of disease. In New Zealand, NSCLC contributes substantially to cancer inequities, with Māori communities experiencing disproportionately high incidence and mortality rates. Although low-dose computed tomography screening can improve early detection, major limitations remain, including false-positive findings, overdiagnosis, high infrastructure costs, and limited accessibility for rural and underserved populations. Liquid biopsy approaches using circulating tumor DNA (ctDNA), particularly DNA methylation profiling, have emerged as promising, minimally invasive strategies for improving cancer detection, treatment monitoring, and precision oncology.

Objective: This study aims to establish integrated genomic and epigenomic predictive and prognostic biomarkers using ctDNA, tumor tissue, and transcriptomic profiling to improve early detection, risk stratification, treatment selection and response prediction, and longitudinal monitoring, with particular emphasis on identifying molecular mechanisms associated with treatment resistance and disease progression.

Methods: This prospective observational translational biomarker study is being conducted through the University of Otago and associated respiratory and oncology services in New Zealand. The study will recruit participants with NSCLC (including squamous and nonsquamous subtypes), individuals referred to fast-track lung nodule assessment clinics, and nonmalignant respiratory controls. Serial peripheral blood sampling will be performed in selected participants at predefined clinical follow-up time points to evaluate treatment response and disease progression. The availability of formalin-fixed paraffin-embedded archival tissues will be recorded, but will not be mandatory for enrollment. Genome-scale DNA methylation profiling will be performed using cell-free reduced representation bisulfite sequencing (cfRRBS), while targeted genomic profiling and transcriptomic analyses will be conducted using targeted sequencing panels and RNA sequencing. Integrative bioinformatic analyses will be used to identify molecular biomarkers associated with early-stage disease, advanced disease, treatment response, and therapeutic resistance.

Results: Ethics approval for the study has been obtained from the New Zealand Health and Disability Ethics Committee (2022 EXP 12566). This study commenced in 2022, and recruitment and biospecimen collection are ongoing. The study aims to recruit approximately 450 participants, including patients with NSCLC, individuals referred through respiratory diagnostic pathways, and nonmalignant controls. As of July 31, 2026, 205 participants have been recruited, with recruitment continuing

until the target sample size is reached. Molecular and data analyses are ongoing, with additional publications expected as the cohort matures.

Conclusions: This study will generate one of the first integrated genomic, epigenomic, and transcriptomic liquid biopsy datasets for NSCLC in New Zealand. The findings are expected to support the development of sensitive, accessible, and equitable blood-based biomarkers for NSCLC detection and treatment monitoring while also contributing to improved precision oncology approaches and reducing NSCLC inequities among Māori populations.

International Registered Report Identifier (IRRID): DERR1-10.2196/105338

JMIR Res Protoc 2026;15:e105338; doi: [10.2196/105338](https://doi.org/10.2196/105338)

Keywords: lung cancer; liquid biopsy; circulating tumor DNA; DNA methylation; epigenetics; biomarkers; transcriptomics; reduced representation bisulfite sequencing; RRBS; treatment response; Māori health; early detection

Introduction

Background

Non-small cell lung cancer (NSCLC) is the leading cause of cancer-related mortality worldwide, accounting for approximately 1 in 5 cancer deaths. In New Zealand, more than 2000 people die from NSCLC every year. Due to its high incidence and relatively poor survival rate, it continues to have the greatest impact on overall cancer morbidity and mortality in New Zealand. Furthermore, NSCLC is associated with significant outcome inequities [1]. Therefore, a better understanding of NSCLC progression and the development of genetic and epigenetic markers for early detection and treatment guidance are important steps toward improving patient outcomes and contributing to health equity.

NSCLC is frequently diagnosed at an advanced stage, with a 5-year survival rate of approximately 15%. Low-dose computed tomography (LDCT) lung cancer screening programs represent an important opportunity to support earlier detection, improve survival outcomes, and address inequities in health care delivery [2-4]. Although LDCT has demonstrated substantial clinical value, challenges remain regarding specificity and the management of indeterminate nodules, which can lead to additional follow-up investigations [5,6]. Furthermore, distinguishing indolent disease from clinically significant disease remains difficult in some cases [7]. In addition, LDCT screening requires considerable infrastructure, specialist workforce capacity, and equitable access across geographically diverse communities [8-10]. These factors highlight the importance of developing complementary biomarkers and precision-based approaches that could enhance the effectiveness and efficiency of NSCLC screening programs.

To mitigate access, cost, and screening-accuracy issues, we propose developing a sensitive, accessible, and cost-effective blood-based test (a molecular liquid biopsy) that can detect NSCLC at an early stage, monitor patients during therapy, and predict treatment response. These molecular liquid biopsy approaches use the circulating tumor DNA (ctDNA) fraction of cell-free DNA (cfDNA) or circulating tumor cells (CTCs) isolated from plasma. Among these approaches, ctDNA has shown great promise as a dynamic molecular monitoring tool for cancer. The ctDNA test is a minimally invasive, highly specific technology that identifies

the presence of a tumor from plasma-derived DNA. Its development represents a major step toward improving the diagnosis and management of cancer [11,12], with applications including earlier-stage diagnosis, detection of molecular relapses to enable more salvage therapies, monitoring treatment response, and identification of molecular targets to enable direct targeted therapies [13-19].

Tumor-derived DNA is distinguished from background normal DNA in plasma by the presence of distinct somatic DNA alterations, typically mutations and DNA methylation changes [19]. Although ctDNA analysis panels comprising common pathogenic mutations display high sensitivity for late-stage disease [20], they have only modest sensitivity for detecting early-stage disease [21]. However, recent data demonstrate that high sensitivity can be achieved using comprehensive panels that include both DNA methylation and mutation markers [22-24]. We and others have shown that alterations in DNA methylation are not only tissue- and cancer-type specific but, because of their abundance, also enable earlier tumor detection with >95% sensitivity and specificity across all stages of cancer [22-28]. In a large study (involving 123,000 samples), a tumor-specific methylation signature detected lung and other cancers at least 4 years before detection by the standard of care [29].

Early vs Late Cancer Epigenetics

DNA methylation profiles are significantly altered in primary (early-stage) cancer compared with those of normal cells [28]. Primary tumors are globally hypomethylated and exhibit site-specific hypermethylation relative to normal cells. Late-stage tumors (metastases) generally show greater hypomethylation than primary tumors and exhibit large site-specific changes. The effect of a specific mutation on the DNA methylation profile is still a subject of research and can vary depending on tumor type or stage. We have demonstrated that it is possible to identify large and discriminatory methylation changes irrespective of driver mutation status in cancer samples [30]. Further work is needed in this area of genome biology to better understand these relationships.

Treatment Effects on Epigenetics and the Potential for Epigenetic Analysis

In NSCLC and other oncogene-driven tumors, relapse can be driven by pre-existing (genetic) mutations or adaptive (nongenetic) resistance [31]. Adaptive resistance is a recently

identified and poorly understood phenomenon in which a drug-tolerant subpopulation survives initial drug treatment. This population can eventually drive relapse via subsequent mutations or epigenetically fixed nongenetic resistance and is a recognized clinical challenge [32,33]. However, drug-tolerant cells are typically identified after prolonged drug exposure and do not represent the pre-existing population. Before drug exposure begins, it remains unknown whether cells with the tolerant phenotype are pre-existing or induced by therapy [34-36].

Our research aims to identify an epigenetic signature related to the drug-tolerant state. The drug-tolerant cells surviving initial targeted therapy consist of subpopulations of slow-cycling, stem cell-like cells. Recently, drug tolerance has been linked to distinct epigenetically controlled transcriptional states [37]. As DNA methylation provides a somatically heritable, dynamic, and reversible epigenetic mechanism that controls gene expression in cancer cells [28], it is highly likely that DNA methylation changes contribute to the drug-tolerant state. Importantly, DNA methylation changes associated with drug tolerance have the potential to identify this subpopulation if it is present before treatment and to inform clinical management of patients with NSCLC. Furthermore, analysis of methylation changes in ctDNA provides a feasible way to monitor patients during treatment [19].

Aims of the Study

This study has the following objectives:

1. We will establish a blood-based DNA signature (combining methylation and mutation markers) to identify NSCLC with high sensitivity and specificity. To achieve this, we will carry out sequencing-based

analysis of critical gene mutations and DNA methylation landscapes in matched blood and tissue samples from patients with NSCLC, blood samples from patients referred to fast-track nodule follow-up clinics or chest X-ray clinics, and blood samples from healthy or nonmalignant controls.

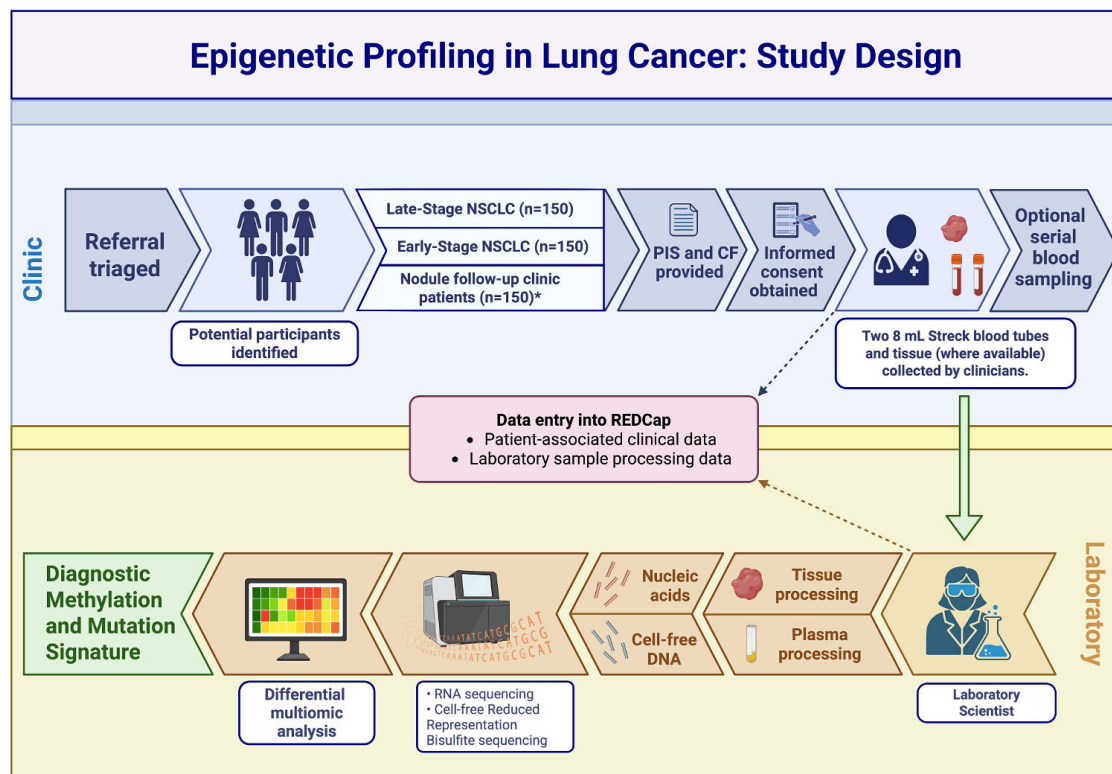
2. We aim to establish a panel of epigenetic and genetic analyses to advance the prediction of treatment response in patients with NSCLC. For this, we aim to identify differences in DNA methylation, gene expression, and targeted mutations between responders and nonresponders with NSCLC receiving targeted therapy or immunotherapy.

Methods

Study Design and Setting

This prospective observational translational biomarker study investigates genomic and epigenomic biomarkers in patients with NSCLC and control participants. The study integrates longitudinal liquid biopsy sampling, tissue-based molecular analyses, and clinical data collection (Figure 1). The study is sponsored through the Department of Pathology and Molecular Medicine and the Department of Medicine, Faculty of Medicine (Dunedin), at the University of Otago, in collaboration with Dunedin Hospital and associated respiratory clinics, including fast-track lung nodule assessment pathways. In addition, we have expanded the local-ity approval to include the Cancer and Chronic Conditions Research Unit in the Waikato District (North Island of New Zealand) for sample collection.

Figure 1. Schematic overview of patient sample collection, processing, and analysis. Patient referrals are triaged and assessed for suitability for participation. Potentially eligible participants are provided with a patient information sheet (PIS) and consent form (CF), followed by the provision of informed consent if they agree to participate. Two 8 mL Streck blood tubes and tissue (where available) are collected from eligible participants by collaborating clinicians from multiple localities, where suitability for serial blood sampling is assessed for future follow-up visits. Patient samples are processed by laboratory scientists for the extraction of nucleic acids. Collaborating clinicians enter patient-associated clinical data into REDCap, and laboratory scientists enter sample processing details into REDCap. Extracted nucleic acids are then further processed in preparation for sequencing. Following sequencing, the data are analyzed and integrated with the corresponding clinical information for each patient to provide multimodal insights. Participants include individuals referred to nodule follow-up clinics and self-declared nonmalignant controls.



Study Population

The study will include nonmalignant controls and patients with precancerous symptoms or early- and late-stage NSCLC

(treated or untreated). The complete eligibility criteria are reported in [Textbox 1](#).

Textbox 1. Eligibility criteria for participation in the study.

Inclusion criteria

- Aged ≥ 18 years
- Histological or cytological diagnosis of non-small cell lung cancer, including nonsquamous or squamous cell carcinoma
- One of the following participant groups:
 - Early- or advanced-stage (stage IIIB or IV) disease, according to the American Joint Committee on Cancer Staging Manual, eighth edition
 - Patients referred to nodule follow-up clinics or chest X-ray clinics
 - Nonmalignant control participants

Exclusion criteria

- Unable to provide informed consent

Recruitment

Potential participants are identified through respiratory services, fast-track lung cancer assessment clinics, chest X-ray clinics, oncology services, and respiratory research databases. Recruitment occurs in both Dunedin and Invercargill. Patients referred for respiratory assessment undergo routine clinical evaluation, including imaging and diagnostic procedures as clinically indicated. Participants are approached regarding study participation during routine clinical care.

Ethical Considerations

Patients identified as meeting the inclusion criteria will be approached for recruitment into the study and asked to provide informed consent. Informed consent will be obtained for the use of their archival tissue samples and blood samples for genomic and epigenomic analyses. This study will be conducted in full conformance with the principles of the Declaration of Helsinki, Good Clinical Practice (GCP), the

Code of Rights of the Health and Disability Commissioner, and the laws and regulations of New Zealand.

All research involving human biological samples is conducted in a manner that respects indigenous values, including the preservation of *mana* (prestige or spiritual power) and adherence to *tikanga* (cultural practices). Both biological samples and the data associated with each participant are handled within a culturally competent framework, recognizing that biological samples are *taonga* (treasured) and that data generated from these samples are *tapu* (sacred or restricted) and must be handled accordingly. *Tikanga* is upheld throughout the study, including culturally appropriate sample storage and disposal procedures. Where requested by participants through the consent process, a *karakia* (blessing) may also be performed during sample disposal.

Ethics approval for the study has been obtained from the New Zealand Health and Disability Ethics Committee (2022 EXP 12566).

Biospecimen Collection and Sample Size

Blood samples are collected from participants at baseline and at longitudinal follow-up time points during clinical management, as applicable. Peripheral blood samples are collected in 8 mL Streck tubes. Serial blood sampling may occur approximately every 6 to 12 months for up to 3 years during routine clinical follow-up, as feasible. Optional archival formalin-fixed, paraffin-embedded (FFPE) tissue samples or fresh-frozen tumor tissue samples are obtained from diagnostic procedures. Tumor-rich regions are identified by pathological review. The study aims to recruit approximately 450 participants, including (1) patients with early-stage NSCLC, (2) patients with advanced-stage NSCLC, (3) participants referred to nodule follow-up clinics, and (4) nonmalignant controls. Nonmalignant controls will comprise self-declared healthy individuals with no known medical comorbidities. These participants will be recruited opportunistically rather than through targeted matching. For all participant groups, relevant patient, tumor, and treatment characteristics will be recorded where available and considered in downstream analyses, where appropriate, for this discovery cohort study.

Plasma and Tissue Sample Processing

Plasma is isolated within 1 week of blood collection to minimize cellular lysis. Blood samples are centrifuged, plasma is retained and stored at -80°C , and buffy coat fractions are preserved for matched nucleic acid analyses. cfDNA is extracted using standardized protocols and subjected to downstream genomic and epigenomic analyses [38-40]. Tumor tissues undergo macrodissection, where required, to enrich tumor cellularity. DNA and RNA are isolated using commercially available extraction kits. Quality control and nucleic acid quantification are performed using fluorometric and electrophoretic methods.

DNA Methylation Analysis

We will use cell-free reduced representation bisulfite sequencing (cfRRBS) to obtain genome-scale DNA methylomes and identify genomic regions with the largest methylation differences between the 2 groups. Our team has extensive experience and is internationally recognized for analyzing epigenomic data [30,41-48]. We will use our well-established laboratory and computational methods [43-46,49-51] to perform library preparation, data processing, and statistical testing to identify differential methylation. We will compare ctDNA methylomes from nonmalignant controls and patients with NSCLC using our established analysis pipeline [45, 46]. We will use ANOVA to determine statistical significance and subsequently apply multiple-testing correction, as appropriate, to detect differentially methylated regions (DMRs) associated with tumors. In addition, we will compare our reduced representation bisulfite sequencing (RRBS) methylomes with The Cancer Genome Atlas (TCGA; array-based methylation data from 585 primary lung tumors are available) and independent ctDNA RRBS datasets from 29 patients with NSCLC and 75 healthy controls (13 million reads per sample; GEO accession GSE79279) [22].

At 100X coverage, the predicted sensitivity for detecting 10 driver methylation changes is 99% with a 1% abundance of ctDNA, and at 100X coverage, the sensitivity for detecting 100 driver methylation changes is 99% with a 0.1% abundance of ctDNA [25]. Furthermore, the observed and expected numbers of DMRs and the detected signals showed a near-perfect linear association when ctDNA was diluted to 0.001% of total cfDNA ($r^2=0.99$) [25]. We will compare ctDNA methylomes from participants without malignancy and patients with NSCLC using our established analysis pipeline [45,46] to detect DMRs that indicate the presence of a tumor. The inclusion of participants without malignancy will allow the identification of methylation changes specific to lung tumors rather than to other respiratory illnesses. We will not sequence germ-line DNA in this work. Our primary aim is to identify large-scale DNA methylation changes specific to NSCLC.

Mutation Profiling

After DNA methylation profiling, when additional DNA remains from a sample, we aim to perform mutation profiling on a subset of samples. Targeted mutation profiling will be performed using clinically relevant next-generation sequencing (NGS) panels designed to detect key genomic alterations in NSCLC, including actionable driver mutations. These assays will enable comprehensive characterization of molecular features associated with NSCLC detection, tumor biology, treatment response, and mechanisms of therapeutic resistance [52,53]. Particular focus will be placed on clinically relevant alterations in genes commonly implicated in NSCLC pathogenesis and targeted therapy response, including epidermal growth factor receptor (EGFR), Kirsten rat sarcoma (KRAS), V-Raf murine sarcoma viral oncogene homolog B (BRAF), anaplastic lymphoma kinase (ALK), c-ros oncogene 1 (ROS1), mesenchymal epithelial transition (MET), and tumor protein p53 (TP53). Mutation profiling

will also be integrated with epigenetic analyses to explore combined genomic and methylomic signatures associated with disease progression, immunotherapy response, and resistance to targeted therapies. Where longitudinal samples are available, temporal changes in mutation profiles will be evaluated to investigate tumor evolution and treatment-associated molecular adaptations.

Transcriptomic Analysis

RNA sequencing (RNA-Seq) data will be generated using our well-established pipelines for clinical tissue samples (where tissue samples are available), including those with low RNA input [30,44,54-58]. Briefly, TruSeq (Illumina Inc) total RNA-Seq libraries will be prepared, yielding 60 to 80 million paired-end reads per sample. Sequence reads will be processed and aligned to the reference genome as previously described [58], and gene expression will be quantified using the transcripts per million (TPM) method after correcting for multiple mapping and fragment bias [59]. Differential expression will be determined with *DESeq2* [39], with *P* values adjusted for multiple testing using the Benjamini-Hochberg method (false discovery rate [FDR] <0.05). Transcriptomic profiles will be analyzed using the gene set enrichment analysis (GSEA) against specific molecular signatures (MSigDB) and integrated with the methylome data, as we have done previously to identify epigenetic drivers associated with transcriptional changes [28,42,56].

RRBS and RNA-Seq Analysis of FFPE Tissues

From pretreatment tumor samples, tumor DNA and RNA will be used to perform RRBS and RNA-Seq. We have developed these methods for both cell lines, tissues, and small amounts of FFPE material. We have demonstrated high reproducibility of RRBS methylation data from FFPE tissues [48,60] and have also validated methylomes using independent methods [42]. In addition, we have demonstrated robust expression profiling using RNA isolated from small amounts of FFPE tissue in multiple tumor types [58]. Our expertise in the regular use of these methods will enable us to successfully profile lung tumors and identify markers of therapy resistance.

Performance Assessment

A central challenge in this study is the selection and validation of candidate markers. In the discovery phase, candidate markers will be chosen from DMRs reaching statistical significance (FDR<0.05) and demonstrating a methylation difference between groups of >20%.

Clinical performance will then be assessed by testing the correlation with treatment response, using receiver operating characteristic (ROC) curve analysis. As a benchmark, for our second aim, the programmed death-ligand 1 (PD-L1) tumor proportion score has variable and modest predictive value, with an area under the curve (AUC) of approximately 0.68 for predicting immunotherapy response [61]. An AUC exceeding this threshold would therefore be considered clinically meaningful in our cohorts, indicating superior predictive

performance relative to PD-L1 alone, and would provide the basis for future translational work. Biomarkers meeting this criterion will proceed to validation in independent external cohorts before prospective evaluation within a clinical trial to establish their utility in guiding treatment selection.

Clinical Data Collection and Management

Clinical and pathological information collected includes the following:

- Demographic variables
- Ethnicity
- Histological subtype
- Disease stage according to the American Joint Committee on Cancer (AJCC)–Union for International Cancer Control (UICC) TNM Staging System (eighth or ninth edition, as applicable)
- Treatment information
- Treatment response
- Longitudinal clinical outcomes

Clinical data are linked to molecular data using deidentified participant identifiers and are collected and managed in REDCap (Vanderbilt University), provided by the University of Otago for staff and students. REDCap is a secure, web-based application designed to support data capture for research studies and provides role-based permissions, secure institutional hosting, audit logging, and controlled data export functionality [62]. The REDCap database, as designed for the Lung Epigenetics Study, is overseen by designated project administrators who are responsible for maintaining database integrity, managing user permissions, and ensuring compliance with institutional data governance policies.

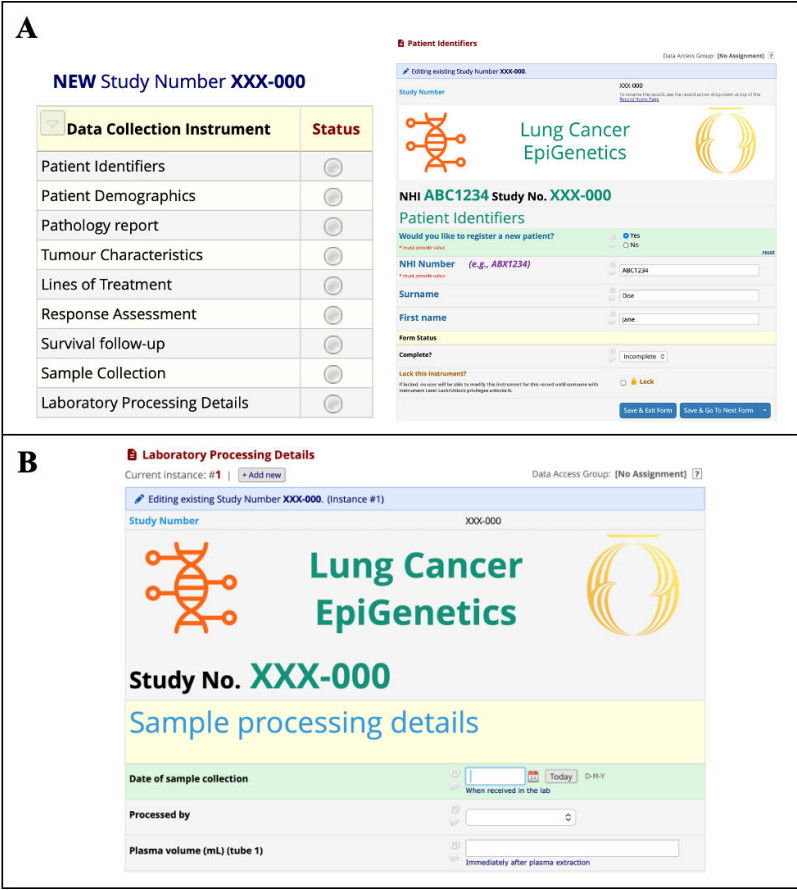
Access to the REDCap database is controlled using predefined user roles, which determine the actions that users are permitted to perform within the database and help ensure that users have access only to the data required for their responsibilities within the project. Users assigned to the admin-level role have full administrative access to the database. Data entry roles are assigned to clinical collaborators responsible for entering patient and clinical data (Figure 2A) from specific contributing sites and updating records as new clinical information becomes available. Users in these roles are restricted to their assigned data access group (DAG), meaning that they cannot view or edit records from other participating localities. Users assigned to the laboratory scientist role are responsible for entering and managing laboratory-related research data. This includes editing sample collection and laboratory processing details (Figure 2B) associated with each deidentified participant identifier. Laboratory scientists do not have permission to view real patient identifiers (National Health Index [NHI] number, name, or date of birth). The auditor role provides read-only access to the database for users who require oversight of the database, excluding patient identifiers, without the ability to modify records.

An additional functionality of REDCap is the ability for users with the admin-level role to “lock” records, meaning that records can no longer be modified. This ensures data integrity by preventing further edits once records have been

finalized, thereby reducing the risk of accidental or unauthorized changes and maintaining data accuracy for downstream analysis and reporting. To further ensure accuracy, exports of associated clinical data are conducted on a quarterly basis

so that downstream analyses are consistently linked to a defined, time-stamped version of the dataset, accounting for the possibility that clinical data may be updated over time.

Figure 2. Example screenshots of clinical data input of a new patient in the Lung Cancer Epigenetics REDCap database. (A) Summary of the data collection instruments and overview of the collection fields in the patient identifiers form to be completed by collaborating clinicians (data entry user role). (B) Overview of the laboratory processing details form to be completed by laboratory scientists (laboratory scientist user role).



Results

Participant recruitment and biospecimen collection are ongoing. The study commenced in 2022 following funding, with recruitment and data collection continuing until the target sample size is reached. As of July 31, 2026, 205 participants have been recruited. Molecular analyses, including DNA methylation profiling, targeted sequencing, and transcriptomic analyses, have commenced, with data analysis progressing alongside recruitment. The first publication describing the technical methodology and feasibility of the study has been published, and analyses addressing the study objectives are ongoing as the dataset matures. Additional findings will be reported through future publications, with further outputs anticipated from 2027 onwards. The study is expected to generate one of the first integrated NSCLC methylome datasets in New Zealand, including valuable liquid biopsy profiling. Continued nationwide recruitment will maximize the value of this collaborative lung cancer resource by enabling robust analyses across multiple research objectives and supporting future translational research.

To support these analyses, it was also necessary to develop low-input analysis methods and perform an assessment to choose an appropriate omics platform for genome-scale DNA methylation analysis. Our comparative analysis demonstrated that both enzymatic and bisulfite-based approaches have distinct advantages for cfDNA methylation profiling in liquid biopsy applications [63]. Enzymatic methyl sequencing (EM-Seq) provided higher mapping efficiency, broader genomic coverage, and improved CpG detection at lower coverage thresholds, whereas bisulfite-based methods showed higher conversion efficiency, superior reproducibility, lower cost, and better coverage of biologically relevant regions such as promoters and exons. Overall, cfRRBS emerged as the most balanced approach in terms of cost-effectiveness, accuracy, and reproducibility, supporting its strong potential for translational cfDNA methylation studies and clinical liquid biopsy applications. Therefore, we will use cfRRBS to analyze NSCLC samples in this study.

Discussion

Blood-Based Biomarkers for Improved NSCLC Outcomes

Against the backdrop of the disproportionately high burden of NSCLC in New Zealand, improving the accuracy of tumor detection and individual risk stratification could inform decisions about personalized investigation, treatment, and monitoring strategies. This could result in improved clinical management and enhanced outcomes for patients with NSCLC. To mitigate the issues of access, cost, and accuracy of screening, the development of blood-based DNA methylation, DNA mutation, and gene expression-based assays that are sensitive, accessible, and cost-effective could improve the detection of lung tumors, assist in treatment management, and contribute to reducing the health care burden and improving NSCLC outcomes.

Molecular Risk Stratification to Support Future Lung Cancer Screening

Although LDCT remains an important and clinically valuable approach for lung cancer screening, integrating complementary molecular biomarkers may further enhance screening efficiency, accessibility, and precision. A New Zealand-based analysis estimated the cost-effectiveness of biennial national LDCT screening at approximately NZ \$145,000 (NZ \$1=US \$0.59, as of July 31, 2026) per quality-adjusted life-year gained [10], highlighting the importance of developing additional strategies that optimize screening pathways and patient selection within resource-constrained health systems. Current international LDCT eligibility criteria typically focus on individuals aged 55 to 75 years with a substantial smoking history [7]. However, NSCLC is increasingly being identified in younger individuals and in never smokers [64,65], populations that may fall outside existing screening frameworks. In this context, ctDNA-based methylation assays could serve as a complementary tool to support risk stratification and prioritization of individuals, particularly those who may not meet conventional criteria but remain at elevated risk. Importantly, ctDNA-based testing has the potential to be relatively cost-effective and scalable as technologies continue to mature.

Epigenetics as an Accessible Tool for Precision NSCLC Investigation

A molecular epigenetic test that can distinguish between nonmalignant disease, nonspecific computed tomography (CT) findings, and clinically significant disease would reduce the number of people requiring follow-up, leading to reduced anxiety and lower costs to the health care system. As ctDNA analysis only requires a 10 mL peripheral blood sample, it enables sampling to be carried out within the community by rural nurses, marae-based clinics, and general practices at any time, greatly improving accessibility and reducing geographic and ethnic inequities.

Epigenetic and Molecular Biomarkers for NSCLC Progression and Therapeutic Resistance

In addition to its use as an adjunct diagnostic tool, a customized ctDNA panel could be used to monitor treatment response (by providing molecular evidence of relapsed disease [18]) and guide treatment decisions. Identification of these markers at the time of diagnosis would be cost-effective and provide the necessary foundation for patients with NSCLC in New Zealand to receive care consistent with international best practice. The objective response rate for targeted therapy in NSCLC varies from 50% to 70% [66-69]. However, most patients develop progressive disease after 18 months of treatment. The development of acquired resistance not only shortens survival but also limits the long-term efficacy of targeted therapies. For patients with NSCLC harboring EGFR mutations who develop resistance to targeted therapy, a second-line therapy is often not available. In New Zealand, osimertinib is funded as the first-line treatment for patients with advanced NSCLC harboring EGFR mutations [68]. Similarly, for patients whose tumors harbor ALK mutations, there is no targeted second-line therapy after they develop resistance to alectinib. The development of markers of tolerance to targeted therapies could enable better clinical management of patients with NSCLC, contributing significantly to reducing the health burden in New Zealand, particularly for patients who lack access to specialized health care, thereby addressing the disparities in outcomes. Furthermore, the results of this project might also provide a basis and opportunity to treat patients with NSCLC with new, lower-toxicity epigenetic drugs (such as guadecitabine, currently being evaluated in 4 active clinical trials for NSCLC) to prevent future tumor progression. Integrative analysis of DNA methylome, expression patterns, and mutation patterns will provide information on the drivers of NSCLC progression and provide a basis for future studies. Using these profiles, epigenetic drivers and pathways that play a role in tumor progression and metastasis in NSCLC can be identified. A plethora of tools is available in the laboratory to study these experimentally and demonstrate how they contribute to cancer progression and metastasis.

As an observational biomarker study, findings from this discovery cohort will require external validation in larger independent cohorts. This study is designed to generate important preliminary insights and identify promising epigenetic biomarkers in NSCLC. Substantially larger independent validation cohorts will enable us to assess the robustness, reproducibility, and generalizability of the identified markers across diverse patient populations and clinical settings [70]. Larger biomarker-driven clinical studies will also be necessary to more confidently evaluate associations with translational outcomes and measures of efficacy, such as treatment response and survival.

Another limitation concerns the paired analysis of tissue- and blood-derived analytes from the same patients. In NSCLC, diagnostic biopsies are often small and clinically

prioritized for routine histopathological and molecular testing, leaving limited residual material available for research purposes. Consequently, matched tissue and blood samples will not always be comprehensively analyzed for all participants, which may reduce the ability to directly correlate tissue-derived and ctDNA epigenomic profiles.

Furthermore, variability in cfDNA yield between patient plasma samples represents a technical challenge. It is well established that cfDNA concentration can vary substantially depending on tumor burden, disease stage, biological variability, and preanalytical factors [71]. Samples with very low cfDNA input may not provide sufficient material for robust genomic or epigenomic profiling, potentially affecting assay performance, data completeness, and downstream analyses. However, we have now developed sensitive, low-input methods for methylation analysis to address this challenge [63].

Funding

The study is primarily supported by funding from the Health Research Council of New Zealand (HRC-NZ; grant 21/989A) and the Sir Charles Hercus Fellowship (awarded to AC; grant HRC 24/002). The authors are also thankful for additional funding from the University of Otago and the Maurice Wilkins Centre, New Zealand.

Data Availability

The epigenomic datasets generated or analyzed during this study (fully anonymized) may be available from the principal investigator (AC) on reasonable request once the study is completed.

Authors' Contributions

All authors contributed to the study conception and design. The first draft of the manuscript was written by AC, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript. AC and BB are joint first authors. BB can be reached via email at Ben.brockway@otago.ac.nz or phone at 64 276 222 475.

Conflicts of Interest

None declared.

References

1. Gurney J, Campbell S, Jackson C, Sarfati D. Equity by 2030: achieving equity in survival for Māori cancer patients. *N Z Med J*. Nov 8, 2019;132(1506):66-76. [Medline: [31778374](#)]
2. National Lung Screening Trial Research Team, Aberle DR, Berg CD, et al. The National Lung Screening Trial: overview and study design. *Radiology*. Jan 2011;258(1):243-253. [doi: [10.1148/radiol.10091808](#)] [Medline: [21045183](#)]
3. de Koning HJ, van der Aalst CM, de Jong PA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. *N Engl J Med*. Feb 6, 2020;382(6):503-513. [doi: [10.1056/NEJMoa1911793](#)] [Medline: [31995683](#)]
4. McLeod M, Sandiford P, Kvizhinadze G, Bartholomew K, Crengle S. Impact of low-dose CT screening for lung cancer on ethnic health inequities in New Zealand: a cost-effectiveness analysis. *BMJ Open*. Sep 24, 2020;10(9):e037145. [doi: [10.1136/bmjopen-2020-037145](#)] [Medline: [32973060](#)]
5. Sozzi G, Boeri M. Potential biomarkers for lung cancer screening. *Transl Lung Cancer Res*. Jun 2014;3(3):139-148. [doi: [10.3978/j.issn.2218-6751.2014.06.04](#)] [Medline: [25806293](#)]
6. Horeweg N, van der Aalst CM, Vliegenthart R, et al. Volumetric computed tomography screening for lung cancer: three rounds of the NELSON trial. *Eur Respir J*. Dec 2013;42(6):1659-1667. [doi: [10.1183/09031936.00197712](#)] [Medline: [23845716](#)]
7. The National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*. Aug 4, 2011;365:395-409. [doi: [10.1056/NEJMoa1102873](#)]
8. Jia Q, Chen H, Chen X, Tang Q. Barriers to low-dose CT lung cancer screening among middle-aged Chinese. *Int J Environ Res Public Health*. Sep 28, 2020;17(19):7107. [doi: [10.3390/ijerph17197107](#)] [Medline: [32998298](#)]
9. Carter-Harris L, Gould MK. Multilevel barriers to the successful implementation of lung cancer screening: why does it have to be so hard? *Ann Am Thorac Soc*. Aug 2017;14(8):1261-1265. [doi: [10.1513/AnnalsATS.201703-204PS](#)] [Medline: [28541749](#)]

Conclusions

Strengths of this study include prospective longitudinal sampling; the integration of genomic, epigenomic, and transcriptomic analyses; the inclusion of participants across the diagnostic and treatment pathway; and the incorporation of equity-focused approaches relevant to Māori health outcomes in New Zealand. This prospective translational study will provide important insights into the genomic and epigenomic landscape of NSCLC in New Zealand. By integrating ctDNA methylation profiling, mutation analysis, and transcriptomic data, the study seeks to establish clinically relevant biomarkers for early detection and treatment monitoring while supporting equitable precision oncology approaches.

10. Jaine R, Kvizhinadze G, Nair N, Blakely T. Cost-effectiveness of a low-dose computed tomography screening programme for lung cancer in New Zealand. *Lung Cancer*. Oct 2018;124:233-240. [doi: [10.1016/j.lungcan.2018.08.004](https://doi.org/10.1016/j.lungcan.2018.08.004)] [Medline: [30268467](https://pubmed.ncbi.nlm.nih.gov/30268467/)]
11. Shohdy KS, West HJ. Circulating tumor DNA testing-liquid biopsy of a cancer. *JAMA Oncol*. May 1, 2020;6(5):792. [doi: [10.1001/jamaoncol.2020.0346](https://doi.org/10.1001/jamaoncol.2020.0346)] [Medline: [32215604](https://pubmed.ncbi.nlm.nih.gov/32215604/)]
12. Wan JC, Massie C, Garcia-Corbacho J, et al. Liquid biopsies come of age: towards implementation of circulating tumour DNA. *Nat Rev Cancer*. Apr 2017;17(4):223-238. [doi: [10.1038/nrc.2017.7](https://doi.org/10.1038/nrc.2017.7)] [Medline: [28233803](https://pubmed.ncbi.nlm.nih.gov/28233803/)]
13. Kunnath AP, Priyashini T. Potential applications of circulating tumor DNA technology as a cancer diagnostic tool. *Cureus*. Jun 16, 2019;11(6):e4907. [doi: [10.7759/cureus.4907](https://doi.org/10.7759/cureus.4907)] [Medline: [31423385](https://pubmed.ncbi.nlm.nih.gov/31423385/)]
14. Fiala C, Diamandis EP. Utility of circulating tumor DNA in cancer diagnostics with emphasis on early detection. *BMC Med*. Oct 2, 2018;16(1):166. [doi: [10.1186/s12916-018-1157-9](https://doi.org/10.1186/s12916-018-1157-9)] [Medline: [30285732](https://pubmed.ncbi.nlm.nih.gov/30285732/)]
15. Hyun MH, Lee ES, Eo JS, et al. Clinical implications of circulating cell-free DNA quantification and metabolic tumor burden in advanced non-small cell lung cancer. *Lung Cancer*. Aug 2019;134:158-166. [doi: [10.1016/j.lungcan.2019.06.014](https://doi.org/10.1016/j.lungcan.2019.06.014)] [Medline: [31319975](https://pubmed.ncbi.nlm.nih.gov/31319975/)]
16. Herbreteau G, Vallée A, Charpentier S, Normanno N, Hofman P, Denis MG. Circulating free tumor DNA in non-small cell lung cancer (NSCLC): clinical application and future perspectives. *J Thorac Dis*. Jan 2019;11(Suppl 1):S113-S126. [doi: [10.21037/jtd.2018.12.18](https://doi.org/10.21037/jtd.2018.12.18)] [Medline: [30775034](https://pubmed.ncbi.nlm.nih.gov/30775034/)]
17. Alama A, Coco S, Genova C, et al. Prognostic relevance of circulating tumor cells and circulating cell-free DNA association in metastatic non-small cell lung cancer treated with nivolumab. *J Clin Med*. Jul 10, 2019;8(7):1011. [doi: [10.3390/jcm8071011](https://doi.org/10.3390/jcm8071011)] [Medline: [31295929](https://pubmed.ncbi.nlm.nih.gov/31295929/)]
18. Guibert N, Pradines A, Favre G, Mazieres J. Current and future applications of liquid biopsy in nonsmall cell lung cancer from early to advanced stages. *Eur Respir Rev*. 2020;29(155):190052. [doi: [10.1183/16000617.0052-2019](https://doi.org/10.1183/16000617.0052-2019)] [Medline: [32051167](https://pubmed.ncbi.nlm.nih.gov/32051167/)]
19. Kilgour E, Rothwell DG, Brady G, Dive C. Liquid biopsy-based biomarkers of treatment response and resistance. *Cancer Cell*. Apr 13, 2020;37(4):485-495. [doi: [10.1016/j.ccell.2020.03.012](https://doi.org/10.1016/j.ccell.2020.03.012)] [Medline: [32289272](https://pubmed.ncbi.nlm.nih.gov/32289272/)]
20. Cohen JD, Li L, Wang Y, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science*. Feb 23, 2018;359(6378):926-930. [doi: [10.1126/science.aar3247](https://doi.org/10.1126/science.aar3247)] [Medline: [29348365](https://pubmed.ncbi.nlm.nih.gov/29348365/)]
21. Bettgowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med*. Feb 19, 2014;6(224):224ra24. [doi: [10.1126/scitranslmed.3007094](https://doi.org/10.1126/scitranslmed.3007094)] [Medline: [24553385](https://pubmed.ncbi.nlm.nih.gov/24553385/)]
22. Guo S, Diep D, Plongthongkum N, Fung HL, Zhang K, Zhang K. Identification of methylation haplotype blocks aids in deconvolution of heterogeneous tissue samples and tumor tissue-of-origin mapping from plasma DNA. *Nat Genet*. Apr 2017;49(4):635-642. [doi: [10.1038/ng.3805](https://doi.org/10.1038/ng.3805)] [Medline: [28263317](https://pubmed.ncbi.nlm.nih.gov/28263317/)]
23. Li J, Zhao S, Lee M, et al. Reliable tumor detection by whole-genome methylation sequencing of cell-free DNA in cerebrospinal fluid of pediatric medulloblastoma. *Sci Adv*. Oct 16, 2020;6(42):eabb5427. [doi: [10.1126/sciadv.abb5427](https://doi.org/10.1126/sciadv.abb5427)] [Medline: [33067228](https://pubmed.ncbi.nlm.nih.gov/33067228/)]
24. Nuzzo PV, Berchuck JE, Korthauer K, et al. Detection of renal cell carcinoma using plasma and urine cell-free DNA methylomes. *Nat Med*. Jul 2020;26(7):1041-1043. [doi: [10.1038/s41591-020-0933-1](https://doi.org/10.1038/s41591-020-0933-1)] [Medline: [32572266](https://pubmed.ncbi.nlm.nih.gov/32572266/)]
25. Shen SY, Singhanian R, Fehringer G, et al. Sensitive tumour detection and classification using plasma cell-free DNA methylomes. *Nature*. Nov 2018;563(7732):579-583. [doi: [10.1038/s41586-018-0703-0](https://doi.org/10.1038/s41586-018-0703-0)] [Medline: [30429608](https://pubmed.ncbi.nlm.nih.gov/30429608/)]
26. Xu RH, Wei W, Krawczyk M, et al. Circulating tumour DNA methylation markers for diagnosis and prognosis of hepatocellular carcinoma. *Nat Mater*. Nov 2017;16(11):1155-1161. [doi: [10.1038/nmat4997](https://doi.org/10.1038/nmat4997)] [Medline: [29035356](https://pubmed.ncbi.nlm.nih.gov/29035356/)]
27. Moss J, Magenheimer J, Neiman D, et al. Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease. *Nat Commun*. Nov 29, 2018;9(1):5068. [doi: [10.1038/s41467-018-07466-6](https://doi.org/10.1038/s41467-018-07466-6)] [Medline: [30498206](https://pubmed.ncbi.nlm.nih.gov/30498206/)]
28. Chatterjee A, Rodger EJ, Eccles MR. Epigenetic drivers of tumorigenesis and cancer metastasis. *Semin Cancer Biol*. Aug 2018;51:149-159. [doi: [10.1016/j.semcancer.2017.08.004](https://doi.org/10.1016/j.semcancer.2017.08.004)] [Medline: [28807546](https://pubmed.ncbi.nlm.nih.gov/28807546/)]
29. Chen X, Gole J, Gore A, et al. Non-invasive early detection of cancer four years before conventional diagnosis using a blood test. *Nat Commun*. Jul 21, 2020;11(1):3475. [doi: [10.1038/s41467-020-17316-z](https://doi.org/10.1038/s41467-020-17316-z)] [Medline: [32694610](https://pubmed.ncbi.nlm.nih.gov/32694610/)]
30. Chatterjee A, Rodger EJ, Ahn A, et al. Marked global DNA hypomethylation is associated with constitutive PD-L1 expression in melanoma. *iScience*. Jun 29, 2018;4:312-325. [doi: [10.1016/j.isci.2018.05.021](https://doi.org/10.1016/j.isci.2018.05.021)] [Medline: [30240750](https://pubmed.ncbi.nlm.nih.gov/30240750/)]
31. Salgia R, Kulkarni P. The genetic/non-genetic duality of drug “resistance” in cancer. *Trends Cancer*. Feb 2018;4(2):110-118. [doi: [10.1016/j.trecan.2018.01.001](https://doi.org/10.1016/j.trecan.2018.01.001)] [Medline: [29458961](https://pubmed.ncbi.nlm.nih.gov/29458961/)]
32. Shen S, Vagner S, Robert C. Persistent cancer cells: the deadly survivors. *Cell*. Nov 12, 2020;183(4):860-874. [doi: [10.1016/j.cell.2020.10.027](https://doi.org/10.1016/j.cell.2020.10.027)] [Medline: [33186528](https://pubmed.ncbi.nlm.nih.gov/33186528/)]

33. Maynard A, McCoach CE, Rotow JK, et al. Heterogeneity and targeted therapy-induced adaptations in lung cancer revealed by longitudinal single-cell RNA sequencing. *bioRxiv*. Preprint posted online on Dec 13, 2019. [doi: [10.1101/2019.12.08.868828](https://doi.org/10.1101/2019.12.08.868828)]
34. Sharma SV, Lee DY, Li B, et al. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell*. Apr 2, 2010;141(1):69-80. [doi: [10.1016/j.cell.2010.02.027](https://doi.org/10.1016/j.cell.2010.02.027)] [Medline: [20371346](https://pubmed.ncbi.nlm.nih.gov/20371346/)]
35. Shaffer SM, Dunagin MC, Torborg SR, et al. Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature*. Jun 15, 2017;546(7658):431-435. [doi: [10.1038/nature22794](https://doi.org/10.1038/nature22794)] [Medline: [28607484](https://pubmed.ncbi.nlm.nih.gov/28607484/)]
36. Oren Y, Tsabar M, Cuoco MS, et al. Cycling cancer persister cells arise from lineages with distinct programs. *Nature*. Aug 2021;596(7873):576-582. [doi: [10.1038/s41586-021-03796-6](https://doi.org/10.1038/s41586-021-03796-6)] [Medline: [34381210](https://pubmed.ncbi.nlm.nih.gov/34381210/)]
37. Boumahdi S, de Sauvage FJ. The great escape: tumour cell plasticity in resistance to targeted therapy. *Nat Rev Drug Discov*. Jan 2020;19(1):39-56. [doi: [10.1038/s41573-019-0044-1](https://doi.org/10.1038/s41573-019-0044-1)] [Medline: [31601994](https://pubmed.ncbi.nlm.nih.gov/31601994/)]
38. Stockwell PA, Rodger EJ, Gimenez G, Morison IM, Chatterjee A. DMAP2: a pipeline for analysis of whole-genome-scale DNA methylation sequencing data. *Curr Protoc*. Sep 2024;4(9):e70003. [doi: [10.1002/cpz1.70003](https://doi.org/10.1002/cpz1.70003)] [Medline: [39258384](https://pubmed.ncbi.nlm.nih.gov/39258384/)]
39. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. [doi: [10.1186/s13059-014-0550-8](https://doi.org/10.1186/s13059-014-0550-8)] [Medline: [25516281](https://pubmed.ncbi.nlm.nih.gov/25516281/)]
40. Rodger EJ, Stockwell PA, Almomani S, Eccles MR, Chatterjee A. Protocol for generating high-quality genome-scale DNA methylation sequencing data from human cancer biospecimens. *STAR Protoc*. Dec 15, 2023;4(4):102714. [doi: [10.1016/j.xpro.2023.102714](https://doi.org/10.1016/j.xpro.2023.102714)] [Medline: [37950864](https://pubmed.ncbi.nlm.nih.gov/37950864/)]
41. Chatterjee A, Macaulay EC, Ahn A, et al. Comparative assessment of DNA methylation patterns between reduced representation bisulfite sequencing and Sequenom EpiTyper methylation analysis. *Epigenomics*. Jun 2017;9(6):823-832. [doi: [10.2217/epi-2016-0176](https://doi.org/10.2217/epi-2016-0176)] [Medline: [28523967](https://pubmed.ncbi.nlm.nih.gov/28523967/)]
42. Chatterjee A, Stockwell PA, Ahn A, Rodger EJ, Leichter AL, Eccles MR. Genome-wide methylation sequencing of paired primary and metastatic cell lines identifies common DNA methylation changes and a role for EBF3 as a candidate epigenetic driver of melanoma metastasis. *Oncotarget*. Jan 24, 2017;8(4):6085-6101. [doi: [10.18632/oncotarget.14042](https://doi.org/10.18632/oncotarget.14042)] [Medline: [28030832](https://pubmed.ncbi.nlm.nih.gov/28030832/)]
43. Chatterjee A, Stockwell PA, Rodger EJ, Parry MF, Eccles MR. scan_tcga tools for integrated epigenomic and transcriptomic analysis of tumor subgroups. *Epigenomics*. Oct 2016;8(10):1315-1330. [doi: [10.2217/epi-2016-0063](https://doi.org/10.2217/epi-2016-0063)] [Medline: [27625317](https://pubmed.ncbi.nlm.nih.gov/27625317/)]
44. Chatterjee A, Stockwell PA, Rodger EJ, et al. Genome-wide DNA methylation map of human neutrophils reveals widespread inter-individual epigenetic variation. *Sci Rep*. Nov 27, 2015;5:17328. [doi: [10.1038/srep17328](https://doi.org/10.1038/srep17328)] [Medline: [26612583](https://pubmed.ncbi.nlm.nih.gov/26612583/)]
45. Stockwell PA, Chatterjee A, Rodger EJ, Morison IM. DMAP: differential methylation analysis package for RRBS and WGBS data. *Bioinformatics*. Jul 1, 2014;30(13):1814-1822. [doi: [10.1093/bioinformatics/btu126](https://doi.org/10.1093/bioinformatics/btu126)] [Medline: [24608764](https://pubmed.ncbi.nlm.nih.gov/24608764/)]
46. Chatterjee A, Stockwell PA, Rodger EJ, Morison IM. Comparison of alignment software for genome-wide bisulphite sequence data. *Nucleic Acids Res*. May 2012;40(10):e79. [doi: [10.1093/nar/gks150](https://doi.org/10.1093/nar/gks150)] [Medline: [22344695](https://pubmed.ncbi.nlm.nih.gov/22344695/)]
47. Chatterjee A, Rodger EJ, Stockwell PA, Le Mée G, Morison IM. Generating multiple base-resolution DNA methylomes using reduced representation bisulfite sequencing. *Methods Mol Biol*. 2017;1537:279-298. [doi: [10.1007/978-1-4939-6685-1_16](https://doi.org/10.1007/978-1-4939-6685-1_16)] [Medline: [27924600](https://pubmed.ncbi.nlm.nih.gov/27924600/)]
48. Ludgate JL, Wright J, Stockwell PA, Morison IM, Eccles MR, Chatterjee A. A streamlined method for analysing genome-wide DNA methylation patterns from low amounts of FFPE DNA. *BMC Med Genomics*. Aug 31, 2017;10(1):54. [doi: [10.1186/s12920-017-0290-1](https://doi.org/10.1186/s12920-017-0290-1)] [Medline: [28859641](https://pubmed.ncbi.nlm.nih.gov/28859641/)]
49. Chatterjee A, Macaulay EC, Rodger EJ, et al. Placental hypomethylation is more pronounced in genomic loci devoid of retroelements. *G3 (Bethesda)*. Jul 7, 2016;6(7):1911-1921. [doi: [10.1534/g3.116.030379](https://doi.org/10.1534/g3.116.030379)] [Medline: [27172225](https://pubmed.ncbi.nlm.nih.gov/27172225/)]
50. Chatterjee A, Lagisz M, Rodger EJ, et al. Sex differences in DNA methylation and expression in zebrafish brain: a test of an extended “male sex drive” hypothesis. *Gene*. Sep 30, 2016;590(2):307-316. [doi: [10.1016/j.gene.2016.05.042](https://doi.org/10.1016/j.gene.2016.05.042)] [Medline: [27259666](https://pubmed.ncbi.nlm.nih.gov/27259666/)]
51. Chatterjee A, Ozaki Y, Stockwell PA, Horsfield JA, Morison IM, Nakagawa S. Mapping the zebrafish brain methylome using reduced representation bisulfite sequencing. *Epigenetics*. Sep 2013;8(9):979-989. [doi: [10.4161/epi.25797](https://doi.org/10.4161/epi.25797)] [Medline: [23975027](https://pubmed.ncbi.nlm.nih.gov/23975027/)]
52. Skoulidis F, Heymach JV. Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy. *Nat Rev Cancer*. Sep 2019;19(9):495-509. [doi: [10.1038/s41568-019-0179-8](https://doi.org/10.1038/s41568-019-0179-8)] [Medline: [31406302](https://pubmed.ncbi.nlm.nih.gov/31406302/)]
53. Wang Z, Xing Y, Li B, Li X, Liu B, Wang Y. Molecular pathways, resistance mechanisms and targeted interventions in non-small-cell lung cancer. *Mol Biomed*. Dec 12, 2022;3(1):42. [doi: [10.1186/s43556-022-00107-x](https://doi.org/10.1186/s43556-022-00107-x)] [Medline: [36508072](https://pubmed.ncbi.nlm.nih.gov/36508072/)]

54. Chatterjee A, Ahn A, Rodger EJ, Stockwell PA, Eccles MR. A guide for designing and analyzing RNA-Seq data. *Methods Mol Biol.* 2018;1783:35-80. [doi: [10.1007/978-1-4939-7834-2_3](https://doi.org/10.1007/978-1-4939-7834-2_3)] [Medline: [29767357](https://pubmed.ncbi.nlm.nih.gov/29767357/)]
55. Gimenez G, Stockwell PA, Rodger EJ, Chatterjee A. Strategy for RNA-Seq experimental design and data analysis. *Methods Mol Biol.* 2023;2588:249-278. [doi: [10.1007/978-1-0716-2780-8_16](https://doi.org/10.1007/978-1-0716-2780-8_16)] [Medline: [36418693](https://pubmed.ncbi.nlm.nih.gov/36418693/)]
56. Rodger EJ, Gimenez G, Ajithkumar P, et al. An epigenetic signature of advanced colorectal cancer metastasis. *iScience.* Jun 16, 2023;26(6):106986. [doi: [10.1016/j.isci.2023.106986](https://doi.org/10.1016/j.isci.2023.106986)] [Medline: [37378317](https://pubmed.ncbi.nlm.nih.gov/37378317/)]
57. Leichter AL, Purcell RV, Sullivan MJ, Eccles MR, Chatterjee A. Multi-platform microRNA profiling of hepatoblastoma patients using formalin fixed paraffin embedded archival samples. *Gigascience.* 2015;4:54. [doi: [10.1186/s13742-015-0099-9](https://doi.org/10.1186/s13742-015-0099-9)] [Medline: [26613016](https://pubmed.ncbi.nlm.nih.gov/26613016/)]
58. Chatterjee A, Leichter AL, Fan V, et al. A cross comparison of technologies for the detection of microRNAs in clinical FFPE samples of hepatoblastoma patients. *Sci Rep.* Jun 3, 2015;5:10438. [doi: [10.1038/srep10438](https://doi.org/10.1038/srep10438)] [Medline: [26039282](https://pubmed.ncbi.nlm.nih.gov/26039282/)]
59. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* Jan 2009;10(1):57-63. [doi: [10.1038/nrg2484](https://doi.org/10.1038/nrg2484)] [Medline: [19015660](https://pubmed.ncbi.nlm.nih.gov/19015660/)]
60. Al Momani S, Rodger EJ, Stockwell PA, Eccles MR, Chatterjee A. Generating sequencing-based DNA methylation maps from low DNA input samples. *Methods Mol Biol.* 2022;2458:3-21. [doi: [10.1007/978-1-0716-2140-0_1](https://doi.org/10.1007/978-1-0716-2140-0_1)] [Medline: [35103959](https://pubmed.ncbi.nlm.nih.gov/35103959/)]
61. Lu S, Stein JE, Rimm DL, et al. Comparison of biomarker modalities for predicting response to PD-1/PD-L1 checkpoint blockade: a systematic review and meta-analysis. *JAMA Oncol.* Aug 1, 2019;5(8):1195-1204. [doi: [10.1001/jamaoncol.2019.1549](https://doi.org/10.1001/jamaoncol.2019.1549)] [Medline: [31318407](https://pubmed.ncbi.nlm.nih.gov/31318407/)]
62. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* Apr 2009;42(2):377-381. [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
63. Ezegbogu M, Stockwell PA, Reid G, et al. Comparison of enzymatic and bisulfite-based methods for sequencing-based cell-free DNA methylation profiling. *Front Epigenet Epigenom.* 2026;3. [doi: [10.3389/freac.2025.1693925](https://doi.org/10.3389/freac.2025.1693925)]
64. Shi J, Li D, Liang D, He Y. Epidemiology and prognosis in young lung cancer patients aged under 45 years old in northern China. *Sci Rep.* 2021;11(1):6817. [doi: [10.1038/s41598-021-86203-4](https://doi.org/10.1038/s41598-021-86203-4)]
65. Miller KD, Fidler-Benaoudia M, Keegan TH, Hipp HS, Jemal A, Siegel RL. Cancer statistics for adolescents and young adults, 2020. *CA Cancer J Clin.* Nov 2020;70(6):443-459. [doi: [10.3322/caac.21637](https://doi.org/10.3322/caac.21637)] [Medline: [32940362](https://pubmed.ncbi.nlm.nih.gov/32940362/)]
66. Wu SG, Shih JY. Management of acquired resistance to EGFR TKI-targeted therapy in advanced non-small cell lung cancer. *Mol Cancer.* Feb 19, 2018;17(1):38. [doi: [10.1186/s12943-018-0777-1](https://doi.org/10.1186/s12943-018-0777-1)] [Medline: [29455650](https://pubmed.ncbi.nlm.nih.gov/29455650/)]
67. Suzuki H, Hirashima T, Okamoto N, et al. The relationship between tyrosine kinase inhibitor therapy and overall survival in patients with non-small cell lung cancer carrying EGFR mutations. *Chin J Cancer.* Mar 2013;32(3):136-140. [doi: [10.5732/cjc.012.10160](https://doi.org/10.5732/cjc.012.10160)] [Medline: [23237215](https://pubmed.ncbi.nlm.nih.gov/23237215/)]
68. So YJ, Fraser A, Rivalland G, McKeage M, Sullivan R, Cameron L. Osimertinib in NSCLC: real-world data from New Zealand. *JTO Clin Res Rep.* 2020;1(2):100022. [doi: [10.1016/j.jtocrr.2020.100022](https://doi.org/10.1016/j.jtocrr.2020.100022)] [Medline: [34589929](https://pubmed.ncbi.nlm.nih.gov/34589929/)]
69. Kobayashi S, Boggon TJ, Dayaram T, et al. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N Engl J Med.* Feb 24, 2005;352(8):786-792. [doi: [10.1056/NEJMoa044238](https://doi.org/10.1056/NEJMoa044238)] [Medline: [15728811](https://pubmed.ncbi.nlm.nih.gov/15728811/)]
70. Robbins HA, Alcalá K, Moez EK, et al. Design and methodological considerations for biomarker discovery and validation in the Integrative Analysis of Lung Cancer Etiology and Risk (INTEGRAL) program. *Ann Epidemiol.* Jan 2023;77:1-12. [doi: [10.1016/j.annepidem.2022.10.014](https://doi.org/10.1016/j.annepidem.2022.10.014)] [Medline: [36404465](https://pubmed.ncbi.nlm.nih.gov/36404465/)]
71. van der Leest P, Schuurin E. Critical factors in the analytical work flow of circulating tumor DNA-based molecular profiling. *Clin Chem.* Jan 4, 2024;70(1):220-233. [doi: [10.1093/clinchem/hvad194](https://doi.org/10.1093/clinchem/hvad194)] [Medline: [38175597](https://pubmed.ncbi.nlm.nih.gov/38175597/)]

Abbreviations

AJCC: American Joint Committee on Cancer
ALK: anaplastic lymphoma kinase
AUC: area under the curve
BRAF: V-Raf murine sarcoma viral oncogene homolog B
cfDNA: cell-free DNA
cfRRBS: cell-free reduced representation bisulfite sequencing
CT: computed tomography
CTC: circulating tumor cell
ctDNA: circulating tumor DNA
DAG: data access group
DMR: differentially methylated region
EGFR: epidermal growth factor receptor

EM-Seq: enzymatic methyl sequencing
FDR: false discovery rate
FFPE: formalin-fixed, paraffin-embedded
GCP: Good Clinical Practice
GSEA: gene set enrichment analysis
KRAS: Kirsten rat sarcoma
LDCT: low-dose computed tomography
MET: mesenchymal epithelial transition
NGS: next-generation sequencing
NHI: National Health Index
NSCLC: non-small cell lung cancer
PD-L1: programmed death-ligand 1
RNA-Seq: RNA sequencing
ROC: receiver operating characteristic
ROS1: c-ros oncogene 1
RRBS: reduced representation bisulfite sequencing
TCGA: The Cancer Genome Atlas
TP53: tumor protein p53
TPM: transcripts per million
UICC: Union for International Cancer Control

Edited by Javad Sarvestan; The proposal for this study was externally peer-reviewed by: the Health Research Council of New Zealand; submitted 24.Jun.2026; final revised version received 09.Jul.2026; accepted 22.Jul.2026; published 27.Aug.2026

Please cite as:

Kumar R, Farry SM, Ezegbogu M, Reid G, Rodger EJ, Brockway B, Chatterjee A
Epigenetic Profiling for Early Detection and Treatment Response Monitoring in Non-Small Cell Lung Cancer: Protocol for a Prospective Translational Biomarker Study
JMIR Res Protoc 2026;15:e105338
URL: <https://www.researchprotocols.org/2026/1/e105338>
doi: [10.2196/105338](https://doi.org/10.2196/105338)

© Rajiv Kumar, Safia May Farry, Mark Ezegbogu, Glen Reid, Euan James Rodger, Ben Brockway, Aniruddha Chatterjee. Originally published in JMIR Research Protocols (<https://www.researchprotocols.org>), 27.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.researchprotocols.org>, as well as this copyright and license information must be included.