**CUPID** – Transforming **C**ancer of **U**nknown **P**rimary with **I**ntelligent **D**iagnostics

**Unique Protocol Identification Number:**

**Australian New Zealand Clinical Trials Registry (ANZCTR) Identified Number:**

**ACTRN12623001221640**

**Principal Investigator: Prof Chris Karapetis**

**Sponsor: SALHN - Southern Adelaide Local Health Network.**

**Funded by: MRFF - EPCDR 2020**

**Version: 2.1**

**15 October 2024**

# 1 Protocol Summary

## 1.1 Synopsis

|  |  |
| --- | --- |
| **Short Title:** | CUPID study |
| **Title:** | Transforming **C**ancer of **U**nknown **P**rimary with **I**ntelligent **D**iagnostics |
| **Study Description:** | This study evaluates the benefits of implementing a uniform care pathway along with the use of a tissue of origin classifier and comprehensive somatic mutation profiling for patients with suspected or confirmed cancer of unknown primary (CUP). This study includes a prospective trial (consented) and a comparator retrospective cohort of patients with CUP (waiver of consent). For the trial cohort, data collection includes the molecular characteristics of CUP tissue and peripheral blood, as well as sociodemographic, clinical, and health service data from participating clinic medical records and linked government registry and administrative datasets. For the retrospective cohort, data collection includes sociodemographic, clinical, and health service data from participating clinic medical records and linked government registry and administrative datasets. |

# 2 Introduction

## 2.1 Background on Cancer of unknown primary

Cancer of Unknown Primary (CUP) is defined as histologically confirmed metastatic cancer for which there is no histologically or clinically confirmed primary site. It is a diagnosis of exclusion for which a standardized diagnostic work-up including comprehensive clinical examination, imaging, tumour markers, morphological and a panel of immunohistochemical (IHC) assessment of tumour sample fails to identify the site of tumour origin at the time of diagnosis1–3.

CUP accounts for 3-5% of all malignancies worldwide3 and prognosis remains poor4–6. Patients with CUP have a median overall survival (OS) of 8-11 months2. In Australia, CUP constitutes 1.6% of all cancers and its incidence has significantly decreased between 1982 and 2019 (from 16 to 8.1 per 100,000 persons), however, like the global statistics, the prognosis is still poor. The 5-year survival of a patient with CUP remains low at 13%7.

Research indicates that significant disparities in the incidence and survival of people with CUP diagnosis exist within Australia. In addition, **the health service delivery, and diagnostic pathways of people with CUP differ markedly compared to people with other cancers**. Although the diagnosis of CUP should only be made in people with metastatic cancer after thorough investigations have failed to identify the site of the primary tumour, analysis of population-based cancer registry data demonstrates that many people with a CUP diagnosis failed to receive adequate investigations8–11. A retrospective US study reported that only 35% of CUP patients reported to the cancer registry received a timely and complete diagnostic evaluation12. These findings were similar in Australia11. Cancer registry-based studies showed that people with CUP were more likely to be elderly and living in aged care facilities, attend an emergency department prior to diagnosis, and undergo less invasive diagnostic tests, leading to missed opportunities for diagnosis and management8, 11,13. Thus, people with suspected CUP may experience inappropriate or underutilisation of recommended tests, leading to delayed or inaccurate diagnosis and missed opportunities to identify a primary site, thus leading to poorer health outcomes. It is also likely that in a proportion of such cases with widely disseminated metastatic disease, additional diagnostic tests were not warranted given the limited ability to change the course of the disease.

Guidelines indicate that if the primary can be identified based on presentation, site-specific therapies are initiated3. However, if a primary site cannot be identified, the CUP is classified into one of two clinic-pathological sub-groups (**favourable and unfavourable risk**). Favourable risk CUP (15-20% of all CUPs) have survival similar to those with advanced cancer with a known primary site. However, the large majority (80-85%) of CUP belong to the unfavourable risk group with no effective treatment options and poor survival from broad-spectrum empirical chemotherapy14. There is a need to develop novel ways to identify the primary site so that effective therapies can be applied and to develop new therapeutic strategies for those without an identifiable primary site.

**2.1.1 Management of Cancer of Unknown Primary**

**2.1.1.1 Cancer of Unknown Primary Optimal Care Pathway**

It is important that the diagnosis of CUP is only made after a comprehensive and standardised diagnostic workup based on the OCP has been done. To further evaluate the potential tissue of origin of CUP, as well as to exclude chemo-sensitive and potentially curable tumours (e.g. germ-cell tumours), one of the most important steps is the ascertainment of tumour histopathological subtype based on the cell of origin using morphology and IHC techniques by an experienced pathologist.

It is anticipated that having a uniform diagnostic pathway will identify those with a true CUP accurately and aid in avoiding incorrect classification arising from inadequate investigations as well as reduce disparities that exist within the CUP population. The Australian optimal care pathways (OCPs) were developed with the goal of reducing variations in cancer care and benchmarking against quality indicators. OCPs set out a standardised approach to delivering consistent and best cancer care. In the year 2020, the Australian OCP for CUP was released15. Elsewhere, the US National Comprehensive Cancer Network (NCCN) and the European Society for Medical Oncology (ESMO) have CUP-specific guidelines3. The OCPs were developed in collaboration with a wide range of clinicians, epidemiologists, consumers and carers by the Cancer Council Victoria, and the National Cancer Expert Reference Group and supported by the Australian Health Minsters’ Advisory Council. The Australian OCP for CUPmaps the whole patient journey from diagnosis of suspected CUP to death. However, the OCP for CUP has not yet been implemented in most states and territories. This study will support the implementation of the OCP for CUP in SA.

**2.1.1.2 Intelligent diagnostics - Assessment of Tissue of Origin and somatic mutations**

Given the poor survival for those with a CUP diagnosis, newer approaches are urgently required. In other cancers, precision medicine approaches with somatic mutation profiling have been increasingly studied. Moreover, tissue of origin studies incorporating molecular changes within tumour tissues was also explored.

*2.1.1.2.1 Precision Medicine*

The National Institutes of Health defines precision medicine as "an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person"16.

Given the therapeutic challenges imposed when the tissue of origin is unknown, patients with CUP would be well suited for a targeted therapy approach. Patients with unfavourable risk CUP have a poor prognosis with currently available chemotherapy. Like in other cancers, somatic mutation molecular profiling techniques have been employed with varying success to identify matched therapies for people with CUP17–22.

We also have evidence to suggest that molecular diagnostic testing methods can assist in identifying the primary site of origin (**tissue of origin**) for patients with CUP. Some methods appear to identify the primary site of origin accurately in more than 90% of patients tested23. However, none of these tests are routinely used in the current clinical practice in Australia either due to lack of proven benefits in clinical trials, difficulty in accessing these tests, lack of national funding or lack of external validation24. There is an urgent need to establish locally available testing methods that are validated and approved by the Therapeutic Goods Administration (TGA). With widespread use and access to somatic mutation profiling in Australia, it is anticipated that such testing in patients with CUP may provide clues towards the primary site of origin and this may lead to improvement in survival outcomes, by enabling therapies that are matched to the identified molecular mutation profile of the cancer and by enabling access to clinical trials when a primary site of the cancer is determined.

# 3 Objectives and Endpoints

|  |  |
| --- | --- |
| **Objectives** | **Corresponding Endpoints** |
| To identify the proportion of patients initially diagnosed with CUP, or suspected to have CUP, who then get assigned a primary cancer using tissue of origin analysis. | Proportion of patients with an initial diagnosis of CUP who then have a primary cancer site identified by tissue of origin analysis. |
| To identify the proportion of patients initially diagnosed with CUP who get a diagnosis other than CUP following somatic molecular profiling. | Proportion of patients with an initial diagnosis of CUP who then have a primary cancer site identified by somatic molecular profiling. |
| To identify the proportion of patients who have an actionable mutation with a recognised matched therapy (or therapies) following somatic molecular profiling. | Proportion of patients with CUP with an actionable mutation following somatic molecular profiling. |
| To implement/increase awareness of the nationally recognised Optimal Care Pathway (OCP) for all patients with CUP. | * Proportion of patients discussed at MDT after implementation of OCP compared to a historical comparator cohort. * Completion rate of a standardised diagnostic workup (measured as time from initial consultation to time of completion of diagnostic work-up). |
| To identify patients’ health care pathway before and after a CUP diagnosis by analysing linked clinic records, Australian government health records, and state and territory government hospital and emergency department records. | Number, type, cost, and order (sequence) of healthcare visits and tests after implementation of the CUP OCP compared to a historical CUP cohort. |
| To identify the proportion of patients whose treatment was changed after a diagnosis other than CUP was made using a tissue of origin test. | Proportion of patients with a change of treatment from results following tissue of origin test. |
| To identify the proportion of patients who were able to receive optimal therapy based on their somatic mutation profiling results. | Proportion of patients that received a therapy guided by somatic mutation profiling. |
| To determine the overall survival of patients with a CUP diagnosis following the implementation of the OCP and to compare it with a historical comparator cohort. | Overall survival (OS) is defined as the time from diagnosis to death from any cause. The date of diagnosis will be taken as the date of histological confirmation of cancer diagnosis, e.g., the date of biopsy. |
| **Exploratory Objectives** | **Exploratory Endpoint** |
| To characterise metabolomic signatures of CUP. | To identify metabolomic signatures. |

# 4 Study design

## 4.1 Description of the Study

This is a multicentre trial evaluating the benefits of implementing a uniform care pathway along with the use of a tissue of origin classifier and comprehensive somatic mutation profiling for people with suspected or confirmed CUP. The study will include a prospective trial and a retrospective comparator cohort, and linkage of these two cohorts to the National Health Data Hub (NHDH). No treatment is being provided to the participants through this study however, the results from the molecular tests may guide treatment.

**Prospective trial**

The study design for the prospective trial cohort is presented in Figure 1. All consecutive patients diagnosed with CUP at the study sites during the study period will be invited to participate. With informed consent, tissue and blood specimens will be collected, or existing specimens utilised, and data such as demographics (age at diagnosis, gender, and ethnicity), dates and types of investigations performed, dates and types of treatment received, number of lines of treatment, progression-free survival and overall survival will be extracted from case notes, electronic medical records, and pharmacy records. Record linkage to cancer registry and administrative datasets will also be performed to ascertain Medicare-subsidised health care utilisation, all hospitalisations and emergency department presentations, Medicare-subsidised tests, costs, and deaths outside of the hospital setting.

**Retrospective comparator cohort**

Data for all patients diagnosed with CUP at the study sites from 1 January 2000 to 31 December 2022 will be retrospectively collected under a waiver of individual consent. Once patients are identified from medical records, information such as demographics (age of diagnosis, gender, and ethnicity), dates and types of investigations done, dates and types of treatment received, number of lines of treatment, progression-free survival and overall survival will be extracted. Information will be obtained from case notes, electronic medical records and pharmacy records. Linkage to registry and administrative health datasets will also be established, but only for patients diagnosed with CUP from 01 January 2011

## 4.1.1 SCHEMA/Study Design:

Figure 1: Workflow of patients with a suspected CUP diagnosis **(prospective trial)**

Change in treatment?

Influence on subsequent. treatment?

* Patients with suspected CUP.
* Local diagnosis (clinical/radiological/histological) compatible with CUP.
* Histological diagnosis of carcinoma or poorly differentiated cancera
* May have commenced treatment for CUP.
* May have commenced treatment for CUP May have commenced treatment for CUP

Tissue of origin, WGTS, CUPPA and somatic mutation profiling results.

Monthly CUPID Multi-Disciplinary team (MDT) discussion and treatment recommendations.

CUP co-ordinator notified.

4-8 weeks

MDT, multi-disciplinary team; OCP, Optimal care pathway; CUP, Cancer of unknown primary;

a. Tumour tissue will be examined further to generate a comprehensive genomic profile. If adequate tumour tissue is not available from the original biopsy, patients will be asked to consent to a biopsy so that more cancer tissue can be obtained for testing. All patients will remain on study to assess compliance with the OCP, even for those where genomic profiling of the cancer cannot be performed.

Figure 2: Workflow to identify and collect data for historical comparator CUP cohort.

Review clinical site electronic medical records for the years 2011-2022.

Identify patients with a diagnosis of suspected or confirmed CUP with:

* A first date of diagnosis of CUP during the study period **OR**
* A first medical oncology clinic visit date for CUP during the study period for CUP, and no prior systematic therapy.

Exclusion criteria:

* patients aged less than 18 years at diagnosis.
* patients with metastatic disease from a known primary site.
* patients with histology and immunohistology profiles (as per 2015 ESMO guidelines) that are compatible with extragonadal germ-cell tumours, neuroendocrine tumours\*, sarcoma, melanoma, mesothelioma, haematological malignancies.

Extract:

1. identifying information about eligible patients for data linkage.
2. socio-demographic and clinical data about eligible patients for analysis.

\*Patients with carcinoma with neuro-endocrine features are eligible. Grade 1 to 3 neuro-endocrine tumours (NETs) are not eligible.

**4.1.2 Screening Period for Prospective Study (trial)**

The population to be included in this study corresponds to patients first diagnosed with CUP for whom a likely tissue of origin cannot be posited from routine diagnostic investigations.

To be eligible, patients must have a histological diagnosis of cancer without pathological findings that confirm a definite primary site, as determined by the study site’s local laboratory on a contemporaneous tissue sample.

The diagnosis of CUP must be ascertained prior to entering the screening period of the study.

**4.1.2.1 During Screening**

As this study aims to direct treatment according to genomic profiles, participants will be asked to agree to access the following:

1. A tumour tissue sample to be analysed for generation of a comprehensive genomic profile.

Archival tumour FFPE block prior to screening will be acceptable for these central analyses. However, if an acceptable archival tumour FFPE block is not available or is not suitable (in quantity and quality) at screening, an FFPE block from a freshly obtained biopsy sample may be obtained if the patient consents to the biopsy.

1. Blood samples; 1 x 5ml EDTA tube and 1 x 9ml lithium heparin peripheral blood sample for WGTS and metabolomic testing.
2. Data from medical records about their health care and outcomes related to the CUP diagnosis.
3. De-identified linked medical and administrative data on their health service use and health outcomes via the Australian Institute of Health and Welfare (AIHW).

**4.1.3 Implementation of the Optimal Care Pathway (OCP)**

1. A survey will be sent out to medical professionals to determine the level of awareness about the OCP at the start and the end of the study.
2. Participants fulfilling eligibility criteria will undergo investigations coordinated by their treating oncologist/treating team as per the OCP while the molecular tests are performed.
3. The assessment based on the CUP OCP includes:
   * Minimal basic workup which will need to be conducted **within 2 weeks** of a medical oncologist/specialist review:
     + Thorough medical history and physical examination (including, e.g., head and neck and breast examination).
     + Basic blood and biochemical analysis.
     + Contrast-enhanced computed tomography (CT) scans of chest, abdomen and pelvis.
   * Further work-up (including, but not limited to):
     + Mammography in women.
     + Breast MRI in women with axillary disease.
     + Endoscopies guided by sign-, symptom- or laboratory abnormalities.
     + Tumour markers:
       - Serum assessment of α-fetoprotein, human chorionic gonadotropin, plasma chromogranin A and PSA in male patients to exclude potentially curable extragonadal germ-cell tumors, neuroendocrine tumors, and prostate cancers amenable to hormonal treatment.
     + Whole-body FDG–PET/CT is suggested for patients with cervical adenopathy from CUP and those with a single CUP metastasis.
   * In situations where clinical and histopathological pictures are suggestive of a specific primary origin for cancer, an appropriate work-up to exclude this specific cancer should also be performed. These situations include (but are not limited to) the following:
     + Breast MRI in women with lymph nodes in the breast drainage areas in the context of IHC results suggestive of breast cancer.
     + MRI liver/MRCP in patients with one or few liver lesions presenting with no extra-hepatic disease or with extra-hepatic disease limited to lung metastases and/or lesions in the upper abdomen, in the context of IHC suggestive of cholangiocarcinoma or pancreatobiliary or upper gastrointestinal disease.
     + Adequate abdominal and pelvic imaging in the presence of high-grade serous carcinoma.
     + Bosniak classification of any kidney lesion in the context of IHC suggestive of renal cell carcinoma or other kidney malignancy.
     + ENT (Ear-Nose-Throat) examination and MRI and/or PET imaging in patients with IHC and/or a clinical picture suggestive of salivary gland carcinoma (this includes poorly defined masses in the neck or predominance of neck lymph nodes in the clinical presentation).

**4.1.4 Evaluation of Treatment Phase for Prospective Study**

4.1.4.1 Initiating treatment:

1. As per the CUP OCP, treatment should start **within 2 weeks** of the decision to treat. Treatment may have commenced prior to genomic profile results.
2. While the molecular tests are performed, the standard care for patients with suspected CUP will proceed as per the OCP. The treating clinician will discuss all cases of suspected CUP in the local hospital or specialist group multi-disciplinary team (MDT) meeting.
3. The treating clinician will proceed with treatment planning according to the OCP guide.

4.1.4.2 Multi-Disciplinary team (CUPID-MDT):

1. Results from somatic mutation profiling will be discussed at the recently established state-wide molecular MDT meetings held monthly.
2. A recommendation on the possible primary site and potentially identified matched therapies including standard therapies and available clinical trials will be provided to the treating clinician/team.
3. The results from the MDT will support the treating team in adopting/modifying treatment after the initial response assessment with the previously chosen therapies. If there is disease progression identified on first-line therapy, the treating clinician and the patient may or may not choose to adopt therapies based on the molecular analysis for the primary site and matched targeted therapies.

## 4.2 Scientific Rationale for TRIAL Design

**4.2.1 Rationale for Population Selection**

The inclusion and exclusion criteria for the study are aimed at selecting patients with confirmed or suspected CUP, for whom a tissue of origin cannot be definitively posited. We will include patients with both favourable and unfavourable prognosis CUP, but we anticipate that a therapeutic approach based on molecular profiling is more likely to change the outcome for the group of patients with unfavourable prognosis CUP, as the disease is inherently heterogeneous in this situation and biomarkers that guide treatment selection are lacking.

**4.2.2 Rationale for Suggested Time Frames for Prospective Study**

The approximate turnaround time for the tissue of origin TSO500, WGTS, CUPPA and somatic mutation profiling results is between 4 and 8 weeks. Results from these tests will then be discussed at the state-wide molecular MDT meetings held monthly. A recommendation on the possible primary site and potentially identified matched therapies including standard therapies and available clinical trials will be provided to the treating clinician/team. Novel treatments may also be recommended to the treating physician. As image assessment for response is usually performed approximately 6-8 weeks after the start of first-line therapy, this timeline for the release of molecular results is appropriate. These results will support the treating team in adopting/modifying treatment after initial response assessment with the previously chosen therapies. If there is disease progression identified on first-line therapy, the treating clinician and the patient may or may not choose to adopt therapies based on the molecular analysis for the primary site and matched targeted therapies.

## 4.4 End of TRIAL Definition

All participants will be followed for survival outcomes until voluntary withdrawal from the study or death.

# 5. TRIAL Population

## 5.1 Inclusion Criteria

Eligible trial participants must fulfill all the following criteria:

1. Age 18 years or more.
2. Suspected or confirmed CUP diagnosis.
3. Willing to provide informed consent.
4. Able to understand English or understand study involvement through interpreter services.

## 5.2 Exclusion Criteria

Patients who meet any of the following criteria will be excluded from trial entry:

1. Metastatic disease from a known primary site.
2. Patients with histology and immunohistology profiles (per 2015 ESMO guidelines) that are compatible with extragonadal germ-cell tumours, neuroendocrine tumours\*, sarcoma, melanoma, mesothelioma, haematological malignancies.

(\*Patients with carcinoma with neuroendocrine features are eligible. Grade 1 to 3 neuroendocrine tumours (NETs) are not eligible)

## 5.4 Screen Failures

Participants who do not meet all the inclusion criteria, or do meet any one of the exclusion criteria, will not be eligible for study participation.

## 5.5 Strategies for Recruitment and Retention

Patients attending cancer clinics with a new diagnosis of CUP will be informed of the study by their treating oncologists. If the patient is interested, the treating clinician will document this and contact the CUPID project coordinator/CUPID study research nurse to provide patient details, including contact details. The CUPID project coordinator/CUPID research nurse will engage with the potential participant. The participant will be fully informed of study requirements, procedures, and any associated risks.

If the patient provides consent, the CUPID team will advise the treating doctor of participation and request details of tissue location for retrieval and testing. The study flow chart is followed, as per Figure 1. After recruitment, the referring medical practitioner may be asked to provide additional information such as favourable/unfavourable prognosis, and investigations/tests undertaken to establish the diagnosis of CUP.

# 6. TRIAL Intervention

## 6.1 TRIAL Intervention(s) Administration

### 6.1.1 TRIAL Intervention Description

Enrolled participants will be offered genomic testing by targeted gene panel testing and Whole Genome and Transcription Sequencing (WGTS). These tests will be performed using archival formalin-fixed paraffin-embedded (FFPE), or fresh tumour tissue (if available).

**6.1.2 Biospecimen collection & management**

6.1.2.1 Archival Tissue:

With the consent of participants, the relevant pathology laboratory will be contacted for access to their archival FFPE tumour tissue specimen (paraffin blocks or unstained slides). Retrieval of all archival material will be coordinated by SA Pathology (Adelaide) for shipping directly to their facility for nucleic acid extraction and gene panel testing.

6.1.2.2 Fresh Tissue:

Some CUP patients will undergo tissue biopsies as part of standard disease workup, and this will be collected whenever feasible. The decision-making process regarding whether to proceed with biopsy collection and the selection of sampling sites will be determined by the treating clinician(s). Participation in biopsy procedures will be entirely voluntary, and patients retain the right to refuse at any point, regardless of prior consent.

WHERE FRESH TISSUE IS AVAILABLE, Tissue biopsies will be collected in RNAlater® solution at the time of the biopsy procedure and transported directly to the University of Melbourne Centre for Cancer Research (UMCCR) for nucleic acid extraction and WGTS. Coordination of fresh tissue collection, including notification to UMCCR for dispatch of Fresh Tissue Collection Kit (containing RNAlater® solution), will be coordinated by SA Pathology. In cases where both FFPE and fresh tissue biopsy are available, fresh tissue will be prioritised over FFPE for utilisation.

6.1.2.3 Blood:

A small blood sample (1 x 5ml EDTA tube) will be collected at recruitment site and shipped to UMCCR for WGTS. This will be timed with blood collection for standard clinical workup where possible, so blood will be drawn from participants from one needle regardless of the number of tubes needed (for genomic testing and standard workup). This will be coordinated by SA Pathology.

**6.1.3 Genomics**

6.1.3.1 Gene panel testing (tumour TSO500):

Somatic gene panel testing will be performed using the Illumina TruSight™ Oncology 500 (TSO500) assay at SA Pathology (Adelaide). TSO500 is a next-generation sequencing (NGS) assay which targets 523 genes across DNA and RNA for assessment of small variants, tumour mutational burden, microsatellite instability, splice variants, gene fusions and copy number variations.

6.1.3.2 Whole Genome & Transcriptome Sequencing (WGTS):

Where there is sufficient high-quality nucleic acid left over after TSO500 panel testing, all remaining tumour DNA/RNA extracts will be shipped to UMCCR (Melbourne) from Adelaide for WGTS. WGS Libraries will be prepared using the Illumina TruSeq Nano Library Preparation Kit. Libraries will be sequenced to 60-100x coverage on tumour and 30-40x coverage on normal blood on Illumina NovaSeq at UMCCR. The bioinformatic analysis will be performed using ICA (Illumina Connected Analytics) and AWS (Sydney Data Centre) using established clinical WGS pipelines, and results will be curated by a dedicated UMCCR curation team.

6.1.3.3 Tissue of Origin Prediction (CUPPA):

WGTS analyses will provide additional actionable genomic features not detected by TSO500 panel testing, including diagnostic and therapeutic mutational signatures. CUPPA is a tissue of origin classifier developed by Hartwig Medical Foundation that utilises both DNA and RNA features derived from WGTS26 and CUPPA will provide a likelihood score of tissue of origin for a CUP case across 36 cancer types and will be reported separately to curated WGTS findings at UMCCR.

**6.1.4 Metabolomic profiling**

Metabolomic profiling using small volumes of peripheral blood analysed on high resolution mass spectrometers will also be performed to identify novel mutation-specific or site-of-origin specific biomarkers. This data will be integrated with the other molecular analyses.

**6.1.5 Project output**

The approximate turn-round time for TSO500, WGTS and CUPPA reports is between 4 and 8 weeks. All molecular data will be curated and reported with recommendations for molecularly targeted therapy, immunotherapy, or site-specific chemotherapy (where applicable) in consultation with genetic pathologists. These will be presented at the monthly CUPID - MDT meetings involving medical oncologists, pathologists, and researchers. Treating clinicians of study participants will be encouraged to attend the MDT, providing them an opportunity to engage directly with scientists and curators and discuss any queries they may have, thereby aiding in the clinical decision-making process for their patients. Final molecular reports for TSO500, WGTS and CUPPA will be signed off by the scientist in charge after TBM and disseminated to treating clinicians via secured file transfer.

**6.1.6 Storage and use of biospecimens for future research**

With patient consent, any residual tissues, blood samples or nucleic acids collected for the study but not consumed for testing will be stored indefinitely at the University of Melbourne Centre for Cancer Research (UMCCR) (Tothill Laboratory). These samples may be accessed for future research through an application to the research study team to ensure such proposed research activities are covered under an approved human research ethics protocol. Research activities may include, but are not limited to, research discovery in CUP patients and the application of new techniques to improve cancer type classification from patient tissue or blood samples.

**6.1.7 Genomics data storage and data use for future research**

All whole-genome and transcriptome data used for primary analysis and clinical reporting will be stored within password-protected accounts on ICA (Illumina) housed within the Australian AWS environment. Data includes individual-level genomic and transcriptomic data (FASTQ, BAM, VCF), and basic associated metadata (laboratory information, sequencing methodologies, quality metrics etc). Data will be archived after reporting on AWS (AWS S3 Glacier). User access and permissions are controlled by Prof Oliver Hofmann (Genomics Platform Group) at the University of Melbourne Centre for Cancer Research (UMCCR). All genomic data is identified by the participant’s unique study identifier, with minimal clinical information.

Collaborative data access requests will be assessed by the study team who will manage and track all requests and will ensure data requestors are approved for access and that the purposes of access are compliant with the original study ethics. Material transfer agreements will be made between institutions and logged with the HREC.

In support of data re-analysis and sharing, the primary data and any associated metadata will be kept accurate, verifiable, unbiased, and annotated with providence and date. Data integrity checks will be performed on every genome sequencing file (FASTQ, BAM, VCF). To support the re-use of the clinical genomic data, for research and secondary purposes, the meta-data associated with genomic sequencing will be collected. This meta-data includes quality metrics, laboratory information and sequencing methodologies.

It is anticipated that following the publication of the study raw genomics data and associated metadata (non-identifying clinical data) will be made available for reuse in a deidentified form by uploading protected data to the European Genome and Phenome Archive (EGA). The data will be managed by a Data Access Committee (DAC) at the UMCCR under the governance of UMCCR DAC (chaired by A/Prof Richard Tothill) to restrict access to only named and authorised medical researchers bound by terms and conditions described in the data access agreement (DAA). The DAA stipulates the use of the data only for agreed purposes by named persons on the DAA and that the requesting institution agrees to preserve patient confidentiality and not to redistribute the data. The DAA will be signed by a legal representative from the requesting institution and all data then be managed through the EGA online portal.

### 6.1.2 Dosing and Administration

Not applicable.

## 6.2 Measures to Minimize Bias: Randomization and Blinding

There is no randomization or blinding.

## 6.3 TRIAL Intervention Compliance

Participant's involvement is minimal as the standard of care biopsies will be utilised for genomic studies. For the majority of patients, no interventions beyond standard of care procedures are planned. In situations where a fresh tumour biopsy is desirable for molecular testing, patients will be informed of the justification for a biopsy and the risks of the procedure and full informed consent will be obtained before the biopsy is performed. Biopsy to obtain tumour material that will enable the molecular characterisation of cancer is, in the current era of molecularly targeted therapies, considered optimal and standard care rather than research-based or experimental intervention. The findings from the genomic analysis of the tumour may influence treatment, and this management effect is a primary outcome measure of the study.

# 7 TRIAL Intervention Discontinuation and Participant Discontinuation/Withdrawal

## 7.1 Discontinuation of Study Intervention

The genomic studies and blood samples for metabolomic studies are one-off testing.

Patient intervention is in keeping with standard practice and optimal care. Survival outcomes will be collected. If participants withdrawal from the study, clarity regarding withdraw from study participation (trial-related activity) versus full study withdrawal will be obtained. Patients will be asked to allow the collection of data, such as survival status, using access to medical records. If medical record access is not permitted, information will be sought through publicly available sources, such as newspaper death notices. If a patient decides to fully withdraw from the study, all study data related to that patient will be withdrawn from the research.

## 7.2 Participant Discontinuation/Withdrawal from the Study

Participants can withdraw from the study at any time without any impact on their care from the treating team.

## 7.3 Lost to Follow-Up

All consented participants will be included for analysis.

# 8 Study Assessments and Procedures

## 8.1 Efficacy Assessments

**8.1.1 Schedule of assessments**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Screening period | Investigation Period- Implementing OCP by treating team | Treatment Period (pre-MDT) | Multi-Disciplinary team (MDT) | Response assessment | Treatment Period  (Treatment as per MDT suggestion) |
| Informed consent | X |  |  |  |  |  |
| Eligibility criteria | X |  |  |  |  |  |
| Complete physical examination | X | X |  |  | X |  |
| Limited physical examination |  |  | X |  |  | X |
| ECOG performance status | X | X | X |  |  |  |
| Blood tests (Haematology, Biochemistry) | X | X | X |  |  | X |
| Tumour assessment  - medical imaging | X | X | X |  | X | X |
| Tumour assessment  - tumour markers | X | X | X |  | X | X |
| Notify CUP co-ordinator if eligibility confirmed | X |  |  |  |  |  |
| Archival tumour tissue | After eligibility confirmation |  |  |  |  |  |
| Tumour tissue biopsy (if insufficient or not suitable archival tissue is available) | After eligibility confirmation |  |  |  |  |  |
| Initiate treatment |  | X |  |  |  |  |
| Discuss tissue of origin and somatic mutation profiling results |  |  |  | X |  |  |
| Molecularly guided therapy in patients with a CR, PR, or SD |  |  |  |  |  | X |

Once potential participants are identified and they have shown interest in the study, a consent and screening visit will be organised.

**8.1.2 Screening visit activities:**

Consent for study participation, including the application of genomic studies, will be obtained. Baseline clinical and demographic data will be collected using a standardised data extraction tool in REDCap; The clinical cancer information including s site of biopsy, tumour features and medical history will all be collected.

Either archival or fresh tumour samples (if available) will be sent for genomic studies (tissue of origin classifier and comprehensive somatic mutation profiling). The participant will then return to the usual treating team for standard of care treatment. If they are started on systemic therapy (with or without local therapy) as per the CUP standards, results of the treatment will be collected (i.e. cancer response) using imaging and cancer markers (if any elevated) as routinely done by the treating teams. In the interim, a molecular MDT meeting will be held to discuss the results from the genomic studies and provide recommendations based on the Mi-ONCOSEQ genomic alteration tiers25. Tier 1 and 2 alterations are considered potentially clinically actionable. The recommendations based on sequencing studies will be provided to the treating team who may choose to use the results for a sequencing-directed therapy through other clinical trials, off-label, or on-label medications to be provided to the participants. This information along with the type of treatment and cancer outcomes from the new second-line treatment will be collected.

The data extraction tool will also be used to capture the dates and types of diagnostic tests performed, the MDT and molecular tumour board meeting dates, the treatments received, and as relevant the date of response to therapy, date of progression, and date and cause(s) of death for each trial participant.

## 8.2 Safety Assessments

As the study does not provide any intervention to the treating physician/treating team, safety assessments are not needed.

## 8.3 HISTORICAL comparator cohort

We will follow the flow-chart depicted in Figure 2 to identify patients eligible for the historical comparator cohort at each study site.

We will perform standardised data extraction and data entry into REDCap using the same data collection tool as for the trial. Data collection will include patient demographics, clinical data, tumour characteristics, the dates and types of diagnostic tests performed, the MDT meeting dates, the treatments received, and as relevant the date of response to therapy, date of progression, and date and cause(s) of death for each eligible patient.

## 8.4 DATA LINKAGE

**8.4.1 Overview**

This study will link both the prospective trial cohort and the historical comparator cohort to the National Health Data Hub (NHDH). The (NHDH) is an enduring linked data asset that includes individual-level data from state and territory hospitals, the Medical Benefits Schedule (MBS), the Pharmaceutical Benefits Scheme (PBS), and the National Death Index (NDI). The hospital data is available via the National Hospital Datasets which is compiled by AIHW from data supplied by states and territories. The NHDH does not include any personally identifying information.

The AIHW linked these datasets using personally identifying information and has overarching ethics approval (AIHW EO/2018/1/438) for it to be used for a broad range of health research statistical analyses. The NHDH was approved by the Australian Health Ministers Advisory Council and is owned by the AIHW, the Commonwealth Department of Health, and state and territory health authorities. It is managed under the custodianship of the AIHW, in consultation with representatives from the state and territory health departments, and the Commonwealth Department of Health and Aged Care.

**8.4.2 Datasets**

Linkage to the NHDH constitutes secondary data collection. The following datasets and indicative variables will be accessed via the NHDH:

1. National Hospital Morbidity Database – Public hospitals and private hospitals. Variables: demographic, diagnosis codes, dates of admission and separation, length of stay, ICU stay, procedures, hospital ID, costs.
2. National Non-admitted Patient Emergency Department Care Database. Variables: presentation date, arrival mode, type of visit, diagnosis codes, hospital ID.
3. National Non-admitted Patient Episode Database. Variables: service delivery mode, service delivery setting, service date, sector,
4. Medicare Benefits Scheme – Medicare Benefits, Speciality, Item Map. Variables: item code, date of processing, broad type of service, service provider ID, in-hospital flag, scheduled fee.
5. Pharmaceutical Benefits Scheme – Pharmaceutical Benefits, Major Specialty, Item Map. Variables: ATC codes, date of dispensing.
6. National Death Index. Variables: date of death, underlying cause, additional causes.

The following datasets are not currently included in the NHDH but will be accessed when available:

1. Australian Cancer Database
2. South Australian Cancer Registry

**8.4.3 Data flow**

The flow of data for the data linkage component of the study is shown in Figure 3.

The AIHW is the data custodian of the NHDH. After all required approvals are received, the AIHW Data Custodian will review and approve linkage to the NHDH and researcher access via the secure access environment (SEAD). The AIHW will provide access in the secure environment to all minimum data items in each of the approved datasets.

Figure 3: Data linkage workflow for prospective trial and historical comparator cohorts

Cohorts:

Study IDs and

content data

Flinders Medical Centre

Cohorts:

Study IDs   
and PII

CUPID Coordinating centre

CGP Laboratory in SA

AIHW

Replace study IDs with PPNs

Replace study IDs with PPNs

Link PIIs to NHDH spine

AIHW

AIHW

Upload PPN-NHDH mapping file and content data to NHDH project workspace in SEAD.

PII, personally identifying information; PPN, project-specific person number; NHDH , National Health Data Hub ; SEAD, secure environment for analysing data.

**8.4.4 Data storage, access, and security**

All data and the analysis datasets will be in an electronic, deidentified format and stored in the Secure Environment for Analysing Data (SEAD) operated by the Australian Bureau of Statistics. Only the approved researchers listed will have access to SEAD. SEAD provides a series of secure, self-contained environments known as “SEADpods” within the ABS Data Lab’s cloud-based infrastructure.

The SEAD system is hosted in Microsoft Azure and meets 'PROTECTED' level security standards as prescribed in the Australian Government [Information Security Manual](https://www.cyber.gov.au/resources-business-and-government/essential-cyber-security/ism) (ISM). The technology underpinning the SEAD system includes:

* Data encryption at rest to mitigate against unauthorised access to microdata.
* Azure Storage Accounts to securely hold individual research products and allow querying from authorised users.
* Cloud servers (including backup servers) are hosted exclusively onshore, with access only authorised for use in Australia unless approved by the ABS.
* Closed network virtual machines to provide secure, isolated research spaces for the analysis of microdata.
* Guarded access through multi-factor authentication and workspace segmentation inhibiting data sharing between projects.
* The SEAD Product Storage Account is protected with Microsoft Defender providing threat detection against malicious/unusual behaviour.

**8.4.5 Data use and disclosure**

Confidentiality will be upheld based on NHDH governance practices:

• The NHDH will not include individual name and address information used to create it. Name and address information will be stored securely and separately from the NHDH and only used for data integration purposes.

• Personal project numbers will be assigned for each individual in lieu of name.

• Individual addresses will be replaced with the Australian Statistical Geography Standard Statistical Area level 2 (SA2) and postcode.

• Date of birth will be replaced with age at service which is calculated in years for all individuals.

• All event dates, including the date of death, will be replaced with dates in month and year only.

• Derived variables will be included comprising number of days since event zero (date of first event) to enable analyses of the order of events and number of days between events.

• The National Health Data and Information Standards Committee (NHDISC) Guidelines for the Disclosure of Secondary Use Health Information for Statistical Reporting, Research and Analysis (NHDISC Guidelines) will be complied with.

Further confidentialisation of outputs from the NHDH will apply as detailed below:

• Unit record data cannot be removed from the host environment.

• An analyst using the NHDH is not to make copies of data from the host environment, neither digital nor handwritten, e.g.by a screenshot, screen share or other digital image, or writing down results etc.

• Project-specific person numbers cannot be removed from the host environment.

• Aggregate outputs cannot be taken from the host environment without AIHW Data Custodian approval.

• Outputs must comply with the confidentiality and privacy requirements of the AIHW Act 1987 and the Privacy Act 1988. This will be assessed by the AIHW Data Custodian in line with requirements for confidentialisation of aggregate outputs.

• Outputs must adhere to the confidentialisation and privacy requirements for the ‘parent’ databases i.e., AIHW Hospital Data Collections, MBS, PBS, and NDI data.

• Following approval from the AIHW Data Custodian aggregate outputs from the NHDH may be integrated with locally held data (for example, population estimates) for the purposes of report development.

• Input and output files must conform to AIHW specifications that may include restrictions on the length of files, the number of files included in a compressed file, and the format of the file.

**8.4.6 Data outputs and dissemination**

NHDH Advisory Committee members will review and provide comments on reports and publications referred to as a third-party release. All third-party releases will be cleared by the NHDH Data Custodian prior to further circulation.

We plan to present the findings from our NHDH analyses and research in peer-reviewed publications in international and national oncology and general medical journals, and present results at international and national conferences to ensure the results of this study reach all relevant decision-makers.

**8.4.7 Data retention**

For the purposes of verification of project findings, files held in project workspaces within NHDH host environments will be securely archived for seven years. After this period, all files will be destroyed. Access to archived files will require the approval of the NHDH AC and will be controlled by the AIHW Data Custodian.

**8.4.8 Data disposal**

Upon completion of the data retention period, we will request personnel from SEAD IT departments to permanently delete files and any relevant workspace, files, or backup tapes.

# 9. Statistical Considerations

## 9.1 Sample Size Determination

The expectation is that the prospective component of this study will recruit 50 patients per year over a 4-year period (total 200 patients). The study will determine the overall survival of patients with a CUP diagnosis following the implementation of the OCP and to compare it with the historical comparator cohort. The expected median survival in the historical cohort is 9 months. Our sample size will have an 80% power of detecting a 33% improvement in survival, from a median of 9 to a median of 12 months, with a sample size of 200.

The study aims to recruit as many participants as possible over a 4-years. Many of the study endpoints are defined as proportions of the total study population. Specifically, these proportions are defined as follows.

1. Proportion of patients with an initial diagnosis of CUP who then have a primary cancer site identified by tissue of origin analysis.
2. Proportion of patients with an initial diagnosis of CUP who then have a primary cancer site identified by somatic molecular profiling.
3. Proportion of patients with suspected CUP with an actionable mutation following somatic molecular profiling.
4. Proportion of patients discussed at MDT after implementation of OCP compared to a historical comparator cohort.
5. Completion rate of a standardised diagnostic workup (measured as time from initial consultation to time of completion of diagnostic workup).
6. Proportion of patients with a change of treatment from results following tissue of origin test.

As no formal comparative tests are planned for the above proportional assessments, pre-specified sample size is not applicable to these endpoints.

The retrospective component of the study will include a review of 20 years of experience managing patients with CUP from each centre. The sample size of the retrospective component will be determined by retrospective record capture.

## 9.2 OUtcomes/exposures and covariates

**9.2.1 Exposures**

The exposure is the implementation of the Optimal Care Pathway (OCP). We will compare outcomes for the trial cohort (OCP exposed) and the historical comparator cohort (non-OCP exposed). Within the historical cohort, we will compare patterns of care and outcomes across 3-year epochs, including the 3 years of 2020 to 2022, as this represents the first 3 years of practice following the publication of the Optimal Care Pathway in Australia. We will take into account the potential impact of the COVID-19 pandemic on the findings.

**9.2.2 Outcomes**

* Health outcomes: fact of and time to cancer progression, death, new onset health conditions (for example, mental health conditions), new cancer diagnoses, and cause(s) of death.
* Health service use: case reviewed by MDT, fact of/timing/sequence of diagnostic tests and procedures, surgery including emergency surgery, chemotherapy/immunotherapy, and radiotherapy, time to diagnosis, time to treatment decision, numbers and timing of admissions to hospital, intensive care units, emergency departments and outpatient clinics, numbers and timing of consultations with GPs, specialists, allied health practitioners, hospital length(s) of stay, timing of supportive/palliative care, and medicines dispensed.
* Health system costs: hospital costs will be estimated using average cost data from the Australian Refined Diagnosis Related Groups (AR-DRG), outpatient and general practice, specialist, allied health service costs, and diagnostic including pathology costs estimated from the Medicare Benefit Schedule (MBS) and medication costs from the Pharmaceutical Benefits Scheme (PBS).
* Trial participants only: proportion of participants with sufficient tissue for successful genomic profiling, proportion of participants who have a primary cancer site identified by tissue of origin analysis, proportion of participants who have a primary cancer site identified by somatic molecular profiling, proportion of participants with an actionable mutation following somatic molecular profiling, proportion of participants with a change of treatment as a result of tissue of origin test, proportion of participants receiving therapy guided by somatic mutation profiling, and the metabolomic signatures of CUP.

**9.2.3 Covariates**

* Patient socio-demographic characteristics: age, gender, ethnicity, remoteness (derived from residential postcode: ARIA index), area-level socio-economic status (derived from residential postcode: SEIFA index), private health insurance status, living arrangements (aged-care resident or other), and marital status.
* Patient co-morbid health conditions, including ECOG performance status, major health conditions present at CUP diagnosis, and Charlson comorbidity index extracted from the hospital data between one year prior and 30 days post CUP diagnosis.
* Clinical and tumour characteristics, including year of diagnosis centre/site.
* Diagnostic pathway, clinical management, and treatment clinical centre/site.

## 9.3 StATISTICAL ANALYSES

All extracted and linked records will undergo quality review and checks for data entry errors and implausible data (e.g., record linkage errors).

We will apply a range of descriptive statistics to quantify the cohorts in terms of patient and disease characteristics, diagnostic work-up, treatments planned and received, health outcomes, other health service use, and health system costs, as listed above in section 9.2. We will calculate the number of days between the first relevant clinical presentation to the completion of diagnostic investigations (the diagnostic interval), and the number of days between the decision to treat and receipt of treatment (the treatment interval) and we will describe the extent of alignment with the CUP OCP recommendations. Patients will be considered aligned with the OCP when evidence of alignment with all recommendations was present. We will evaluate the distribution of the diagnostic interval and treatment interval variables and generate appropriate statistics such as medians and 90th centiles.

For each cohort, we will generate Sankey diagrams to map the movement of patients from one care state to another (e.g. presentation, diagnosis, treatment, maintenance, end-of-life). These diagrams will allow the visualisation of timeliness of care and cancer management/treatment patterns.

Crude overall survival and 95% confidence intervals will be estimated for each cohort and major subgroups (i.e., CUP subtypes, treatment approaches, OCP concordant care) using the Kaplan-Meier method and defined as weeks from histological diagnosis of CUP to death, censored at last follow-up date. Cohort/subgroup differences will be compared using the log-rank test.

Regression models will also be used to examine the association between receipt of key health services, including OCP-aligned care, with survival outcomes, using time to event and relevant censoring (death or end-of-follow-up). All models will consider confounding by relevant sociodemographic and other characteristics (as determined using directed acyclic graphs, DAGs) using approaches to minimise imbalances in confounders (e.g. propensity scores modelling, stratification). Sensitivity analyses will be performed as indicated by the initial descriptive statistics and DAGs, and as justified based on the available sample size.

Health service costs will be determined using National Independent Hospital Pricing Authority costs associated with Australian-Related Diagnosis Related Groups (AR-DRGs) and National Weighted Activity Unit (NWUA) recorded for each occasion of health service in hospital, emergency department, or outpatient service. Rebates for medical services provided by Medicare, and costs of each medicine dispensed, will also be determined. All costs will be aggregated for each patient and stratified in relation to the cohort and period of care (e.g., diagnosis, treatment).

# 10 Supporting Documentation and Operational Considerations

## 10.1 Regulatory, Ethical, and Study Oversight Considerations

### 10.1.1 Ethics and regulatory compliance

This trial will be conducted according to the Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95) annotated with TGA comments (Therapeutic Goods Administration DSEB July 2000) and in compliance with applicable laws and regulations. The study will be performed in accordance with the NHMRC Statement on Ethical Conduct in Research Involving Humans (©Commonwealth of Australia 2007), the NHMRC Australian Code for the Responsible Conduct of Research (©Australian Government 2007), and the principles laid down by the World Medical Assembly in the Declaration of Helsinki 2008.

The study will be reviewed by the Southern Adelaide Clinical Human Research Ethics Committee and the South Australian Department for Health and Wellbeing HREC, in accordance with National Mutual Acceptance (NMA) arrangements for multi-jurisdiction data linkage projects (<https://www.clinicaltrialsandresearch.vic.gov.au/national-mutual-acceptance>).

### 10.1.2 Informed consent **for trial participants**

An Informed Consent Form will be used at each site for the prospective trial. If applicable, it will be provided in a certified translation of the local language. Patients will be told that they are free to refuse to participate and may withdraw their consent at any time for any reason. The patient information sheet and consent form will include check boxes for participants to agree to participate in optional procedures.

The Consent Forms must be signed and dated by the patient before his or her participation in the study. The case history or clinical records for each patient shall document the informed consent process and that written informed consent was obtained prior to participation in the study.

The Consent Forms will be revised whenever there are changes to study procedures or when new information becomes available that may affect the willingness of the patient to participate.

If the Consent Forms are revised (through an amendment or an addendum) to communicate information that might affect a patient's willingness to continue in the study, the patient or a legally authorized representative must re-consent by signing the most current version of the Consent Forms or the addendum in accordance with applicable laws and IRB/EC policy. For any updated or revised Consent Forms, the case history, or clinical records for each patient shall document the informed consent process and that written informed consent was obtained using the updated/revised Consent Forms for continued participation in the study.

A copy of each signed Consent Form must be provided to the patient. All signed and dated Consent Forms must remain in each patient’s study file or the site file and must be available for verification by study monitors at any time.

**Ethical considerations relating to the sharing of genetic test results with participants.**

Genetic testing has benefits and risks that are different from those associated with other pathology tests. This is because of the predictive nature of certain genetic tests, and the shared familial implications and ownership of genetic information. As genetic test results may have implications for relatives in addition to the person being tested, it is important to mention this during the consent process. Therefore, we will take the following steps:

* + When undertaking consent, inform patients of the possible results of genetic testing and their implications.
  + Provide participants with relevant SA Health associated Genetic Testing Insurance Information leaflet if they have more questions or concerns about genetic testing and insurance.
  + Briefly discuss the role of information sharing in the family and notification of at risk family members if a disease-causing variant is identified.
  + Provide access to genetic and clinical advice through relevant clinicians or clearly recommend to participants to seek their services. When a patient has a medical test, they usually expect to receive definitive results and often think that the test has a high chance of producing meaningful results. However, with genetic testing: The pick-up rate varies for each test and clinical indication, and in many contexts is low, ranging from 10% upwards. It is possible to receive complex or uncertain results.
  + Setting realistic expectations about the potential outcomes of a genetic test will help to minimise disappointment and confusion when a patient receives their results.

**Secondary or incidental findings that arise from the analysis of personal information/data.**

The consenting/treating doctor will make it clear to participants when describing the study and obtaining informed consent whether there will be potential actionable or therapeutic benefit to them from the trial, informed by the results of their tumour testing. They will set realistic expectations about the potential outcomes of a genetic test to help minimise disappointment and confusion when a patient receives their results.

**Ethical considerations relating to any dissemination of outcomes to the participants.**

The interpretation of tumour genetic test results can be complex. We will provide clinician training and support for discussing results of tumour genetic testing with participants. We will provide access to genetic counselling for patients and their families including assistance with decision making about genetic testing, reproductive risk counselling and support for adjustment to genetic test results. Additionally, we will co-ordinate family risk notification and predictive genetic testing if a disease-causing variant is identified.

### 10.1.3 **Waiver of consent for historical comparator cohort**

Under the National Statement on Ethical Conduct in Human Research (2023) section 2.3.9, we seek a waiver of consent on the basis that:

**(a) involvement in the research carries no more than low risk to participants.** As this is a retrospective cohort study there is no direct harm or discomfort to participants. There is a small risk of re-identification of individuals during the research process because record linkage requires personally identifying information. This risk will be minimised by strict data governance procedures (see (f) below). Personally identifying information will be kept separate from content data; individuals will be assigned unique person IDs that will be used for data collection and data analysis. Neither individuals nor institutions will be identified in the study findings.

**(b) the benefits from the research justify any risks of harm associated with not seeking consent.** Survival after a diagnosis of CUP is poor and we do not know whether implementation of the CUP OCP improves survival or quality of life. The health service use and health outcomes for the comparator cohort are crucial to understanding these outcomes for the trial cohort.

**(c) it is impracticable to obtain consent.** Obtaining individual consent is impracticable given the very low 5-year survival rate after CUP diagnosis. Most patients in the comparator cohort will have died; obtaining consent from only those who are still living would generate a biased subset.

**(d) there is no known or likely reason for thinking that participants would not have consented if they had been asked.** The benefits of this research outweigh the minimal risk associated with disclosure of personal information, and this research cannot be undertaken using an alternative study design due to the low survival rate and the nature of the data needed to meet the Aims. Improving outcomes for people diagnosed with CUP is rated highly by patients, their carers, and their clinicians.

**(e) there is sufficient protection of their privacy.** Risk mitigation strategies will be employed to minimise the risk to personal privacy throughout the research process i.e. during data extraction, record linkage, data analysis, and reporting. There is a small risk to the privacy of individuals during data extraction and record linkage because personally identifying information is used. However, this risk is minimised by applying the separation principle. Under this principle, personally identifying information is never seen in conjunction with content data (i.e. clinical information). An authorised third party (the AIHW) will undertake the record linkage using best-practice, privacy-preserving protocols. The process of record linkage is largely automated, requiring minimal viewing of individual records. The flow of data has been designed to maximise privacy.

Access to the identified data will be restricted to the smallest number of (i) study staff required to extract the data from clinical records and (ii) AIHW staff required to undertake the linkage. The staff responsible for data extraction will not have access to the linked health data in the secure access environment. Furthermore, all such staff will sign confidentiality agreements. All researchers will sign any necessary confidentiality undertakings, be familiar with the NHMRC Code for the Conduct of Responsible Research, abide by the UNSW Research Code of Conduct, and undertake any necessary training relating to data access/use.

Linked content data accessed by the researchers in a secure access environment will not contain personal identifiers, and all outputs will undergo independent review to ensure that the confidentiality and privacy of individuals and facilities are maintained. All minimum cell size restrictions will be adhered to. Draft presentations, manuscripts, and reports will be shared with all data custodians for review prior to public release.

**(f) there is an adequate plan to protect the confidentiality of data.** The use and access of data by researchers will follow agreed processes to ensure that anonymity is preserved and risks to privacy is minimised.

Data will be extracted from medical records and entered into a REDCap database which will use a 2-factor authentication process. The data will be stored at Flinders Medical Centre. Information and communication technology (ICT) security at Flinders Medical Center will ensure controlled access to this confidential information.

As this project involves cross-jurisdictional linkage, the linked data can only be accessed by researchers in a secure access environment. The environment uses purpose-built technology to automate and manage information flows and provides tight control over data access that overcomes the major privacy challenges, and reputational and legal liability risks associated with releasing data to external parties. The data cannot be copied, downloaded, or transmitted by email or other means. Researchers can take their analytical output from the environment but not the original data. All inputs and outputs are vetted through a unique “curated gateway” for compliance and the system records and archives all transactions for future reference. Access is strongly authenticated requiring three different factors of authentication. All users are required to undertake training on issues of privacy, ethics, information security, and statistical disclosure control prior to gaining access and they sign a deed outlining the terms and conditions of use. The system is hosted in two tiers 3+ (i.e., best available) data centres. Regular on-site and off-site backups of data are made. All off-site backups and archival data are encrypted prior to being transferred to secure off-site storage. An audit record will be maintained, and it will allow the destruction of data to be completed when necessary.

The number of researchers with access to the linked data will be limited at any one time. Other named investigators will have access only to aggregate, tabular summaries of the data. We will submit a change in personnel forms (or equivalent) to the Southern Adelaide Clinical HREC for approval prior to providing data access to new analysts. Likewise, we will notify the HREC when personnel no longer require access to the data.

Results will be published in a form that will not allow individuals or facilities to be identified, that is, in aggregate, tabular form only. All outputs will be reviewed for potentially identifiable information (including small cell sizes) and statistical disclosure control will be used to protect confidentiality and privacy, where needed.

**(g) in case the results have significance for the participants' welfare there is, where practicable, a plan for making information arising from the research available to them.** All data obtained under a waiver of consent is administrative in nature and of no significance to individual patient welfare. There is a clear plan for dissemination of the aggregate research results to all stakeholders. We have established a multidisciplinary Project Advisory Group involving oncologists, molecular pathologists, clinicians, and community representatives. The Group will assist in protocol development, interpretation of results, and the dissemination and translation of findings. Face-to-face meetings will be scheduled quarterly. Findings will be published in academic journals, policy and lay summary briefs, media releases, and relevant websites. They will be presented at major conferences and consumer forums. Our established connections with the Clinical Oncology Society of Australasia (COSA) will ensure the communication of findings to oncologists.

We plan to develop a separate study that specifically involves Aboriginal and Torres Strait Islander peoples, and for that study, an appropriate Indigenous governance process will be set up, and Indigenous researchers will lead analyses and report findings about Indigenous people. We will seek approval from the **Aboriginal Human Research Ethics Committee (AHREC).**

**(h) the possibility of commercial exploitation of derivatives of the data will not deprive the participants of any financial benefits to which they would be entitled.** There is no envisaged commercial exploitation of the administrative data obtained under a waiver of consent.

**(i) the waiver is not prohibited by State, federal, or international law.**

#### 10.1.4 **Confidentiality and Privacy**

Confidentiality standards will be maintained by coding each patient through assignment of a unique patient identification number.

Patient medical information obtained including genetic information by the trial is confidential and may be disclosed to patients’ family/first-degree relatives or third parties only as permitted by the Informed Consent Form (or separate authorization for use and disclosure of personal health information) signed by the patient unless permitted or required by law. Medical information may be given to a patient’s personal physician or other appropriate medical personnel responsible for the patient’s welfare, for treatment purposes.

Given the complexity and exploratory nature of exploratory biomarker analyses, data derived from these analyses can be provided to patients if requested or unless required by law upon patients’ agreement.

Data generated by this study must be available for inspection upon request by representatives of national and local health authorities, representatives, and collaborators, and the IRB/EC for each study site, as appropriate. Trial data, which may include imaging data and data on genomic variants, may be submitted to government or other health research databases or shared with researchers, government agencies, companies, or other groups that are not participating in this study.

As per the Informed Consent Form, de-identified individual and aggregate data collected during the trial will be used for future research purposes, to advance science and public health, or for analysis, development, and commercialization of products to treat and diagnose disease. In addition, redacted Clinical Study Reports and other summary reports will be provided upon request.

**10.1.5 Trial Record Retention**

Trial electronic records will be stored on secure Flinders Medical Centre servers accessible only by approved research staff upon completion of two-factor authentication.

Records and documents pertaining to the conduct of this trial, including electronic and paper medical records informed consent forms, laboratory test results, medication records, genomic results, and images, will be retained by the Principal Investigator for 15 years after completion or discontinuation of the study or for the length of time required by relevant national or local health authorities, whichever is longer.

After that period, the documents will be irreversibly destroyed according to local regulations.

### 10.1.5 Publication and Data Sharing Policy

Regardless of the outcome of a trial, the researchers are dedicated to providing aggregate research findings to healthcare professionals and to the public, at scientific congresses, and in peer-reviewed journals.

The researchers will comply with all requirements for publication of study results. Except for linked health records that will never be held by the researchers, only accessed in a secure, remote analysis environment, trial data may be shared with others who are not participating in this study (see Section 10.1.3 for details), and redacted Clinical Study Reports and other summary reports will be made available upon request. No personally identifying information will be shared with any other party, other than for the purpose of data linkage.

Authorship will be determined by mutual agreement and in line with International Committee of Medical Journal Editors authorship requirements.

# 11 References

1. Pavlidis N, Fizazi K. Carcinoma of unknown primary (CUP). *Crit Rev Oncol Hematol*. 2009;69(3):271-278. doi:10.1016/j.critrevonc.2008.09.005

2. Massard C, Loriot Y, Fizazi K. Carcinomas of an unknown primary origin--diagnosis and treatment. *Nat Rev Clin Oncol*. 2011;8(12):701-710. doi:10.1038/nrclinonc.2011.158

3. Fizazi K, Greco FA, Pavlidis N, Daugaard G, Oien K, Pentheroudakis G. Cancers of unknown primary site: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol Off J Eur Soc Med Oncol*. 2015;26 Suppl 5:v133-8. doi:10.1093/annonc/mdv305

4. Brewster DH, Lang J, Bhatti LA, Thomson CS, Oien KA. Descriptive epidemiology of cancer of unknown primary site in Scotland, 1961-2010. *Cancer Epidemiol*. 2014;38(3):227-234. doi:10.1016/j.canep.2014.03.010

5. Schroten-Loef C, Verhoeven RHA, de Hingh IHJT, van de Wouw AJ, van Laarhoven HWM, Lemmens VEPP. Unknown primary carcinoma in the Netherlands: decrease in incidence and survival times remain poor between 2000 and 2012. *Eur J Cancer*. 2018;101:77-86. doi:10.1016/j.ejca.2018.06.032

6. Kim CS, Hannouf MB, Sarma S, Rodrigues GB, Rogan PK, Mahmud SM, et al. Survival outcome differences based on treatments used and knowledge of the primary tumour site for patients with cancer of unknown and known primary in Ontario. *Curr Oncol*. 2018;25(5):307-316. doi:10.3747/co.25.4003

7. AIHW. Cancer in Australia 2019, Summary. *Cancer Aust 2019*. 2019;Cancer ser(Cat. no. CAN 123):7-18.

8. Schaffer AL, Pearson SA, Dobbins TA, Er CC, Ward RL, Vajdic CM. Patterns of care and survival after a cancer of unknown primary (CUP) diagnosis: A population-based nested cohort study in Australian Government Department of Veterans’ Affairs clients. *Cancer Epidemiol*. 2015;39(4):578-584. doi:10.1016/j.canep.2015.02.007

9. Crawford SM, Skinner J, Coombes E, Jones AP. Cancer of Unknown Primary: a Cancer Registry Study of Factors Affecting Access to Diagnosis. *Clin Oncol*. 2017;29(1):e39-e46. doi:https://doi.org/10.1016/j.clon.2016.09.011

10. Jones W, Allardice G, Scott I, Oien K, Brewster D, Morrison DS. Cancers of unknown primary diagnosed during hospitalization: a population-based study. *BMC Cancer*. 2017;17(1):85. doi:10.1186/s12885-017-3083-1

11. Schaffer AL, Pearson SA, Perez-Concha O, Dobbins T, Ward RL, van Leeuwen MT, et al. Diagnostic and health service pathways to diagnosis of cancer-registry notified cancer of unknown primary site (CUP). *PLoS One*. 2020;15(3):e0230373. doi:10.1371/journal.pone.0230373

12. Smith-Gagen J, Drake CM, White LL, Pinheiro PS. Extent of diagnostic inquiry among a population-based cohort of patients with cancer of unknown primary. *Cancer reports Rev*. 2019;3(3). doi:10.15761/CRR.1000187

13. Vajdic CM, Schaffer AL, Dobbins TA, Ward RL, Er CC, Pearson SA. Health service utilisation and investigations before diagnosis of cancer of unknown primary (CUP): A population-based nested case–control study in Australian Government Department of Veterans’ Affairs clients. *Cancer Epidemiol*. 2015;39(4):585-592. doi:https://doi.org/10.1016/j.canep.2015.02.006

14. Pavlidis N, Pentheroudakis G. Cancer of unknown primary site. *Lancet*. 2012;379(9824):1428-1435. doi:https://doi.org/10.1016/S0140-6736(11)61178-1

15. Overview | Metastatic malignant disease of unknown primary origin in adults: diagnosis and management | Guidance | NICE. https://www.cancer.org.au/assets/pdf/cancer-of-unknown-primary-quick-reference-guide

16. What is precision medicine?: MedlinePlus Genetics. Accessed May 3, 2022. https://medlineplus.gov/genetics/understanding/precisionmedicine/definition/

17. Yoon HH, Foster NR, Meyers JP, Steen PD, Visscher DW, Pillai R, et al. Gene expression profiling identifies responsive patients with cancer of unknown primary treated with carboplatin, paclitaxel, and everolimus: NCCTG N0871 (alliance). *Ann Oncol Off J Eur Soc Med Oncol*. 2016;27(2):339-344. doi:10.1093/annonc/mdv543

18. Varghese AM, Arora A, Capanu M, Camacho N, Won HH, Zehir A, et al. Clinical and molecular characterization of patients with cancer of unknown primary in the modern era. *Ann Oncol Off J Eur Soc Med Oncol*. 2017;28(12):3015-3021. doi:10.1093/annonc/mdx545

19. Kato S, Krishnamurthy N, Banks KC, De P, Williams K, Williams C, et al. Utility of Genomic Analysis In Circulating Tumor DNA from Patients with Carcinoma of Unknown Primary. *Cancer Res*. 2017;77(16):4238-4246. doi:10.1158/0008-5472.CAN-17-0628

20. Clynick B, Dessauvagie B, Strerrett G, Harvey N, Subramaniam S, Herron J, et al. Detection of therapeutic targets in carcinomas of unknown primary. *Ann Oncol*. 2017;28:v597. doi:10.1093/annonc/mdx391.005

21. Krämer A, Losa F, Gay LM, Sokol E, Page DR, Frampton GM, et al. Genomic profiling of carcinomas of unknown primary (CUP) to support clinical decisions. *J Clin Oncol*. 2018;36(15\_suppl):e24162-e24162. doi:10.1200/JCO.2018.36.15\_suppl.e24162

22. Hayashi H, Kurata T, Takiguchi Y, Arai M, Takeda K, Akiyoshi K, et al. Randomized Phase II Trial Comparing Site-Specific Treatment Based on Gene Expression Profiling With Carboplatin and Paclitaxel for Patients With Cancer of Unknown Primary Site. *J Clin Oncol Off J Am Soc Clin Oncol*. 2019;37(7):570-579. doi:10.1200/JCO.18.00771

23. Inocencio RT, Kesari K, Smith S. Diagnostic accuracy of molecular tumor profiling in determining the tissue of origin in adult patients with cancer of unknown primary: A meta-analysis. *J Clin Oncol*. 2017;35(7\_suppl):20. doi:10.1200/JCO.2017.35.7\_suppl.20

24. Hainsworth JD, Rubin MS, Spigel DR, Boccia R V, Raby S, Quinn R, et al. Molecular gene expression profiling to predict the tissue of origin and direct site-specific therapy in patients with carcinoma of unknown primary site: a prospective trial of the Sarah Cannon research institute. *J Clin Oncol Off J Am Soc Clin Oncol*. 2013;31(2):217-223. doi:10.1200/JCO.2012.43.3755

25. Cobain EF, Wu YM, Vats P, Chugh R, Worden F, Smith DC, et al. Assessment of Clinical Benefit of Integrative Genomic Profiling in Advanced Solid Tumors. *JAMA Oncol*. 2021;7(4):525-533. doi:10.1001/jamaoncol.2020.7987

26. Schipper LJ, Samsom KG, Snaebjornsson P, Battaglia T, Bosch LJW, Lalezari F, et al. Complete genomic characterization in patients with cancer of unknown primary origin in routine diagnostics. ESMO Open. 2022;7(6):100611.