



**International
Cancer Genome
Consortium**

ARGO



**International Cancer Genome Consortium
The ARGO Project**

International Cancer Genome Consortium The ARGO Project

Accelerating Research in Genomic Oncology

INTRODUCTION	3
THE GENOMICS REVOLUTION	3
THE ERA OF PRECISION MEDICINE	3
THE INTERNATIONAL CANCER GENOME CONSORTIUM	4
THE ARGO PROJECT (ACCELERATING RESEARCH IN GENOMIC ONCOLOGY)	7
ICGC-ARGO SPECIFIC AIMS	10
MEMBERSHIP GUIDELINES	10
STRUCTURE OF ICGC-ARGO	13
GOVERNANCE & COORDINATION	15
TECHNOLOGY DEVELOPMENT	16
<i>Technology Advances</i>	19
<i>Liquid Biopsy</i>	19
CONSORTIUM DEVELOPMENT	20
TISSUE & CLINICAL ANNOTATION	21
<i>Trials & Cohorts</i>	25
<i>Germline Implications</i>	28
COMMUNICATION AND OUTREACH	33
<i>Patient Initiated Participation</i>	36
ETHICS & GOVERNANCE	37
<i>International Data Access Committee</i>	<i>Error! Bookmark not defined.</i>
<i>Data Access Compliance Office</i>	<i>Error! Bookmark not defined.</i>
PHENOMICS	43
DATA COORDINATION AND MANAGEMENT WORKING GROUP (DCMWG)	48
SCIENTIFIC WORKING GROUPS	51
<i>Predisposition WG</i>	51
<i>Precursor lesion WG</i>	51
<i>Therapeutic Development WG</i>	51

Introduction

The Genomics Revolution

Few research fields have witnessed the level of dramatic advancement that has occurred in genomics. The last decade has seen rapid progress from individual gene sequencing to highly parallel whole exome and genome sequencing supported by enhanced computational infrastructure and innovative analytical algorithms. These advances have created unprecedented opportunities to characterise the molecular attributes of disease. Whilst DNA sequencing leads the way, other technologies are interrogating gene expression, genome methylation, and many other attributes of the genome. Moreover, advances in other fields such as proteomics and metabolomics are capable of measuring vast detail, and together with genomics, generating data of unprecedented size and complexity. The ability to rapidly generate massive 'omics datasets, at an ever-decreasing cost, has driven a new era in the world of medicine which promises to optimally manage disease, leading to improvements in outcome and efficiency.

The Era of Precision Medicine

Interpreting accumulating data and generating robust and impactful insights has become the focus of research and investment in many fields. This strategy underpins a transition towards what is variably termed personalised, individualised, stratified or precision medicine. However, rather than this approach being a "Revolution", it is simply the natural "Evolution" of healthcare; where an enhanced understanding of the fundamental basis of disease, as it always has, informs the way we research, diagnose and treat. Central is the use of molecular pathology to complement clinical data that characterise the patient and their disease in greater and greater detail. This, in turn, informs the development and application of selective granular approaches to prevention, diagnosis and treatment.

In oncology, genome sequencing has revealed a subset of cancer-associated aberrations called "driver" mutations. Accumulating experimental data demonstrates that these are responsible for many of the properties that transform normal cells to cancer. "Hallmarks of Cancer" include uncontrolled growth, metastasis, evasion of apoptosis, angiogenesis, and the ability to escape immune surveillance, and many of these mutations result in the acquisition of these attributes. The identification of these aberrations has already defined cancer vulnerabilities and has shown great promise in many cancer types through new drugs and diagnostic tests that are treated based on individual patients selected based on their molecular attributes and that of their cancer.

ICGC-ARGO

Goals, Structure, Policies & Guidelines

Although the rate of scientific discovery in the 'omics era has opened the door to exciting new possibilities, much work is needed before these can be implemented in existing healthcare delivery systems. Healthcare systems need to develop the pipelines, policies and computational capacity to handle the generation and interpretation of 'omics data. More importantly, most current clinical approaches are limited to detect discrete genomic aberrations in a small number of genes (which is usually less than 0.01% of the genome), rather than providing a broad view of the oncogenic events in the context of the clinical attributes of an individual patient. We need to move beyond these narrow approaches to generate valuable insights into cancer, and to pioneer the application of these more advanced methodologies into the clinic.

The International Cancer Genome Consortium

The International Cancer Genome Consortium (ICGC) currently supports over 88 projects across 17 jurisdictions (16 countries & the European Union). To date, this has produced >20,000 tumour genomes for 26 cancer types. The results of the analyses of these data are available through the Data Coordination Centre (DCC) via the ICGC website. ICGC members agree to make the data as broadly available as possible under appropriate governance with minimal restrictions to expedite cancer research. Data is available through a tiered access system. Controlled Access Datasets that contain more sensitive information are only available to scientists and institutions with formal authorisation by the Data Access Compliance office (DACO). To facilitate data access around the globe, ICGC adopted cloud-based solutions to enable users to perform compute-intensive analyses. Initiatives such as ICGC in the Cloud involve non-commercial (the Cancer Genome Collaboratory managed by OICR and the University of Chicago) as well as commercial solutions.

Over the past ten years, the International Cancer Genome Consortium (ICGC) has established itself as an effective and efficient organization, managed by a Secretariat located at the Ontario Institute for Cancer Research (OICR) in Toronto, Canada. Under the guidance of the ICGC Executive Committee and International Steering Committee, the various ICGC Working Groups have addressed emerging issues in ethics, policy, data access, privacy, technology, analysis, data coordination and data management.

The ICGC was the first step to broadly and comprehensively map the structural aberrations of genomes and begin to understand the molecular basis of cancer. It was appropriately focused on cancer that had not yet been treated, and on the tumour at its origin (primary cancer). It achieved these goals with over 20,000 primary cancers of many organs already available, and the remainder sequenced and in process for upload and sharing. The ICGC also extended beyond its initial mandate into the transcriptome and

ICGC-ARGO

Goals, Structure, Policies & Guidelines

the epigenome. Data generated through the ICGC has transformed research strategies in academia and industry alike, with hundreds of seminal works published directly using ICGC data, with landmark articles in the world's elite scientific journals. No therapeutic is developed today without, in some way, applying the knowledge that ICGC has provided the world. Although the ICGC has achieved much, there is a long way to go. Pivotal outstanding challenges remain to be addressed; unanswered questions that are vital in our quest to defeat cancer.

The ICGC has evolved significantly since its inception in 2007. At its heart, ICGC is a consortium of experts in genomics and informatics. Its initial project was to define the genomes of 25,000 primary untreated cancers (the **25K Project**); the second ICGC project, the Pan Cancer Analysis of Whole Genomes (The **PCAWG Project**), commenced in 2013 and continues to analyse ~3,000 of the highest quality whole cancer genomes of multiple cancer types. In 2015, the ICGC, in response to the realization of the potential of genomics in healthcare, released a position "white paper" on the evolution of ICGC into more directly impacting on human health. Emanating from the **ICGC for Medicine (ICGCmed)** white paper is ICGC's next project which aims to Accelerate Research in Genomic Oncology (The **ARGO Project**), where key clinical questions and patient clinical data drive the interrogation of cancer genomes.

The International Cancer Genome Consortium for Medicine (ICGCmed)

The ICGC for Medicine white paper identified central needs and challenges in Genomic Health for Cancer and defined the strategic direction for the ICGC for advancing genomics towards the clinic to improve outcomes for people affected by cancer. The pivotal outstanding questions in cancer molecular pathology discovery that were defined, and that the ICGC aims to address through **The ARGO Project** are:

1. How do we use current treatments better?

Although significant progress has been made, the reality is that most cancer drugs are not very effective in most individuals. The majority of people who receive current treatment do not benefit, they experience toxicity, sometime life threatening, and receipt of the optimal therapy for that individual is delayed, often too late for many. This trial and error approach has a cost, both human and financial. We need to do better with currently available therapies by identifying which treatment will work and which will not for an individual ahead of time; identifying those that will and directing patients to the best therapeutic option; and avoiding treatments that will not work, with the offer of clinical trials of novel therapies when appropriate.

ICGC-ARGO Goals, Structure, Policies & Guidelines

2. How does a cancer change with time and treatment?

Because of the ICGC, we now know the molecular underpinnings of many (primary) cancer types; yet recent evidence tells us that treatment changes the cancer, forcing it to evolve. Even if the treatment is effective, most advanced cancers recur, and their new genetic make-up is the result of what they started with, and where the treatment drove the individual cancer's molecular evolution. We need to know how treatments affect cancer and their genetic makeup after treatment. These "evolved" cancers are the ones that kill most people, and the time in a patient's journey where new drugs are tested in clinical trials. The drugs we make for untreated cancer are unlikely to be appropriate after a cancer has evolved in response to treatment.

The first ICGC 25K project included cancers that are cured by surgery, and those that recur. ICGC did not assess advanced cancers where a high proportion are lethal. Understanding which cancers are lethal as compared to those that are curable, and differentiating those that do, and those that do not require early aggressive treatment.

3. How do we practically implement these approaches in healthcare and drug development?

Implementing these strategies will need adjustments in healthcare systems, with the ultimate goal of making them learning environments that continue to advance knowledge and improve outcomes. Technical innovation will drive much of the advances with critical challenges such as non-invasive diagnostics, monitoring therapeutic response and surveillance after therapy. Moreover, overcoming the challenges of data and information sharing on a global scale that is feasible and efficient to accelerate research and therapeutic development. The ultimate goal is to utilize "real world" data to inform research and development, since over 95% of treated patients are not part of clinical trials. Moreover, clinical trials are selected participants that need to fit in to certain criteria, and may not represent the overall population at large.

4. How do we advance early detection and ultimately prevent cancer?

We only know who has the highest risk of developing cancer based on inherited genomic events in a relatively small number of genes. We need to better determine a person's individual risk for developing cancer based on their genetic predisposition and their environmental and lifestyle exposure. Our greatest advances in early detection have been the identification and treatment of precancerous lesions. Examples include cervical cancer and colon cancer screening. An emerging issue is that although only a small number of such precursor lesions usually progress to cancer, for most we have no way of

ICGC-ARGO

Goals, Structure, Policies & Guidelines

determining which will and which won't, and the treatment of such precursor lesions can be associated with significant detriment. As a consequence, understanding the contribution of the germ line genome to cancer development and progression, impact on drug response and toxicity, and the molecular pathology of precursor lesions is an important area ICGC is well positioned to address.

An important aspect is the difference between germ line variants predisposing to cancer, and germ line variants that are not associated with increased susceptibility, but may confer therapeutic vulnerability. Metrics that distinguish these, and better characterize germ line therapeutically actionable variants in the context of the tumour (loss of second copy of gene in tumour; mutational signature etc.) will become increasingly more relevant.

The ICGC and its membership have established global networks, capacity and expertise in genomics. Tens of thousands of scientists, clinicians, technologists, patients and people at large have built the methodologies and mechanisms to advance genomic health. This expertise, governance, ethical understanding, communication; and importantly the ability to analyze, compile, interpret and share genomic data across the globe are significant. The ICGC forms a strong foundation to build on and advance cancer research through integrating the expertise and capacity of the ICGC with clinical research and cancer care to accelerate research in genomic oncology through ***The ARGO Project***.

The ARGO Project (Accelerating Research in Genomic Oncology)

ICGC-ARGO will link genomic data that is already amassed, new genomic data generated through the 10 years of ARGO, to clinical and health information. This will include information concerning lifestyle, co-morbidity, diagnostics, toxicity and response to therapy and survival. Using this large-scale integrated data, researchers, scientists, policymakers and clinicians will be able to work with patients, health care providers, industry, and others to advance therapeutic development with interventions based on matching the patient's disease molecular subtype with the most effective treatment; develop preventative strategies; markers for early detection of disease; and more specific criteria and methods for diagnosis and prognostication.

As a worldwide consortium, ICGC has the research and organizational expertise to implement the ambitious goal of analyzing the genomes of more than 100,000 patients by 2028, and linking these data to high-quality clinical information including treatment and outcome. The ARGO Project will investigate all samples by extensive sequencing for mutation detection (by exome and/or whole genome or equivalent approaches), and undergo analysis of copy number alterations and rearrangements. Transcriptome will be a common assay across all donors since it is less impacted by cellularity as genomic

ICGC-ARGO

Goals, Structure, Policies & Guidelines

mutation detection is, and incorporates the tumour microenvironment. Methylation analyses will be conducted where possible. The project size and scope will enable an understanding of the regional differences in disease around the world, the heterogeneity of cancer, the diversity of environmental risk factors, as well as describing new cancers with a common genomic background, and common outcomes; and the many different combinations of therapeutic interventions. By pooling large numbers of samples together with clinical information, the consortium will provide the statistical power necessary to enable researchers to achieve a number of goals, and subsequent adoption of evidence-based health care interventions, strategies and policies. A key element is that individual cohorts will have focused clinical parameters, enhancing discovery for specific contexts whilst building the overall dataset for pooled analytical approaches.

The ARGO Project leverages the enormous learnings and infrastructure of ICGC to pursue key aspects of cancer that are required to improve outcomes. The inescapable core of ARGO, which is detailed clinical and treatment annotation, together with cutting edge molecular analyses, will aim to inform:

1. Therapeutic development

- a. Current treatment selection for standard of care therapies**
- b. Understanding lethal cancers versus those that are curable**
- c. Treatment induced tumour evolution**
- d. Pharmacogenomics**
- e. Mechanisms that integrate with contemporary drug development**
- f. “Real World” therapeutic testing**

2. Technology and systems

- a. Liquid Biopsies**
- b. Data-sharing knowledge banks**
- c. Innovative analytics (eg: AI)**

3. Early detection and prevention

- a. Germ line predisposition and contribution to carcinogenesis**
- b. Characterisation of precursor lesions**

ICGC-ARGO

Goals, Structure, Policies & Guidelines

Numerous platforms and for various cancers have already been established, and more are emerging to address these vital questions in many countries around the world. Whilst these build on our knowledge base, there are currently no mechanisms to standardize the complex analyses that are used, or efficient mechanisms for data sharing for cancer that will enable composite and pooled analyses of accumulated data from around the world. Based on the 10-year ICGC experience and resultant infrastructure, ARGO stands poised to accelerate cancer research for the international community through its established infrastructure, expertise and workflows.

The launch of ICGC-ARGO presents an opportunity for countries around the world to combine their efforts to reduce the global burden of cancer, and for all sectors to contribute through shared knowledge. Preceded by the ten highly successful years of ICGC, ICGC-ARGO will “hit the ground running”, immediately positioned to build on the existing core infrastructure, policies and guidelines of ICGC. The support structures for ICGC is both extensive and amenable to adaptation to suit the specific needs of ICGC-ARGO, and are underpinned by an unprecedented level of commitment from clinical and basic research communities, as well as an impressive cadre of international supporters.

Membership and Associate Membership of ICGC-ARGO offers:

1. *Standardised, uniform and evolving genomic analyses – built on the foundation of ICGC 25K Genomes and the PCAWG (Pan Cancer Analysis of Whole Genomes) Projects.*
2. *Data co-ordination and distribution under appropriate ethical and legal frameworks.*
3. *Equitable and fair data sharing arrangements.*
4. *Opportunities for pooled data analyses.*
5. *Network of NGS laboratory partnerships facilitated through ICGC-ARGO for clinical consortia to enable advanced molecular analyses.*
6. *Knowledge exchange through the ICGC-ARGO community with regular meetings, teleconferences and opportunities to contribute to leadership and Working Groups.*
7. *Support for project development (open to non-members as well).*

ICGC has laid the foundations and framework to enable the translation of a wide range of ‘omics data into tangible clinical benefit for cancer patients. It is only through initiatives such as ICGC-ARGO that the mountains of genomic data being generated internationally can be applied to generate novel insights that will underpin advances in the clinical

ICGC-ARGO

Goals, Structure, Policies & Guidelines

management of people with cancer. ICGC-ARGO is the next step for international large-scale efforts and with appropriate governance, will link the wealth of genomic data already amassed with clinical and health information, including lifestyle, patient history, response to therapies and cancer diagnostic data for the international community. ICGC-ARGO will build on the strong foundation of previous ICGC projects, extending the network into the arena of clinical medicine and drug development to expand the expertise and learnings of ICGC into healthcare.

ICGC-ARGO Specific Aims

The central goal of ICGC-ARGO is to analyze biospecimens from at least 100,000 cancer patients with high quality clinical data to address current key outstanding questions that are vital to our quest to defeat cancer. To achieve this goal, the acquisition of accompanying high quality clinical information is of utmost importance.

The sources of cohorts of patients that would constitute ICGC-ARGO projects may include:

- Biospecimens from participants enrolled in active clinical trials;
- Analyses of banked samples from past clinical trials;
- Analyses of samples from clinically well-annotated cohorts that satisfy ICGC-ARGO clinical data requirements.
- Longitudinal cohort studies.
- Autopsy studies with detailed clinical data

Membership Guidelines

1. ICGC-ARGO Membership is based on a Programme of work, and may contain any number of specific “cohorts” within that umbrella. It is encouraged, that when developing these programmes that the ICGC-ARGO Steering Committee is consulted for more detailed guidance. Each programme will be reviewed by the ICGC-ARGO Steering Committee to facilitate an interactive process.
2. An ICGC-ARGO Programme would generally address a specific cancer type, or a clinically relevant grouping of cancer. Examples include cancers of the Head and Neck, which includes many cancer types and is a reflection of the clinical care pathways that are in place to manage patients. This may include specific clinical trial indications. Cohorts may be a reflection of centralised healthcare facilities such as for paediatric cancer. Broad-based general cancer molecular profiling platforms may also be members, with provision of a proportionate breakdown of cancer types expected.

ICGC-ARGO

Goals, Structure, Policies & Guidelines

3. An ICGC-ARGO Programme must address a key clinical and/or biological question of relevance to the specific cancer type on which the programme is focused. This will be different for different cancer types and will impact on sample sizes and analyses performed for individual projects within a programme.
4. Clinical annotation is of utmost importance. Whilst clinical trial data is the Gold Standard, clinically well-annotated cohorts of patients that include the mandatory ICGC-ARGO clinical dataset are eligible.
5. The goal of ICGC-ARGO is to advance discovery, as a consequence, nucleic acid analysis must go significantly beyond the assessment of selected gene sets using panels such as those currently performed in most clinical diagnostic laboratories. Cognizant of the tractability of biospecimen quantity, quality and fixation methods, minimum requirements are either one of the approaches detailed in (a) below; complemented with methodologies listed in (b). The exact composition per project will be defined through discussion with the ICGC-ARGO steering Committee, however, each case **should have a transcriptome** to allow broad pooled data analyses, and a discovery genome sequencing approach. WGS is ideal, however is not often tractable in clinical trials, and whole exome or a “Clinical Genome” that captures attributes beyond point mutations in genes that are clinically relevant may be used.

a) Expected

- **Whole Genome Sequencing OR**
- **Whole Exome Sequencing OR**
- **Clinical Genome** - an assay that captures equivalent data readouts to a WGS. This approach is included as for solid tumours, the above may be intractable for WGS, or less informative (WES). Most clinical trial biospecimens are fixed in formalin or similar fixatives in uncontrolled environments, and up to 50% have low epithelial cellularity which impacts on WGS in particular. This would make over 95% of clinical trials and eligible population based cohorts unable to be used. These composite diagnostics must go well-beyond current diagnostic “panels” and interrogate novel features of clinical and biological relevance in a specific cancer type.
- **ICGC-ARGO Clinical Genome** – ICGC-ARGO is developing a clinical genome diagnostic that may be used. This will become available in 2019.

b) Expected

- **Whole Transcriptome Sequencing**

c) Encouraged

ICGC-ARGO

Goals, Structure, Policies & Guidelines

- **Epigenetic Analyses**
 - **Additional analyses** eg: proteomic, metabolomic ... are encouraged.
6. Financial commitment of \$10 million US equivalent per programme. There are no restrictions as to how a programme comes together, or the timing of the investment; however, the investment must be current and active for a minimum of 3 years from the time of membership. Essential to membership is a commitment and a plan for delivery or access of format and content compliant data to the ICGC-ARGO DCC and/or tangible mechanism for data sharing as per ICGC procedures and policies. A clear formal structure and governance mechanism for consortia that come together to join ICGC-ARGO needs to be articulated and responsible leads identified. Established consortia, networks and co-operative groups are ideally positioned to become ICGC-ARGO Members.
 7. It is anticipated that the number of donors committed by a programme will be a minimum of 2000 depending on the proposed assays and analyses, and will be arrived at in discussion with the ICGC-ARGO Steering Committee on a project-by-project basis if required.
 8. Emerging and smaller scale programmes may become Associate Members. Associate members are defined as those that have a commitment equivalent to members but aim to contribute between 500 and 2000 donors with the same commitment as members with regard to data sharing and an investment of \$5 million with the same conditions as for full members. Again, specific project related parameters can be defined in discussion with the ICGC-ARGO Steering Committee.
 9. Associate and Full Membership will also be considered, on a case by case basis, for groups willing to contribute \$5 or \$10 million respectively in in-kind services, including genome sequencing, clinical data harmonisation, or data analysis compute capacity.
 10. There are no restrictions with regard to Members as to their institution, jurisdiction, or corporate status.
 11. Data access is tiered, and aimed not to disadvantage Members or Associate Member Data producers, with a framework that encourages data sharing, yet provides data generators with sufficient time to perform analyses:
 - Up to 12 months from completion of standardised analyses: Access to Programme submitting data only.
 - 12 months: Access to Members.
 - 18 months: Access to Associate Members.
 - 24 months: Accessible by external parties.

ICGC-ARGO

Goals, Structure, Policies & Guidelines

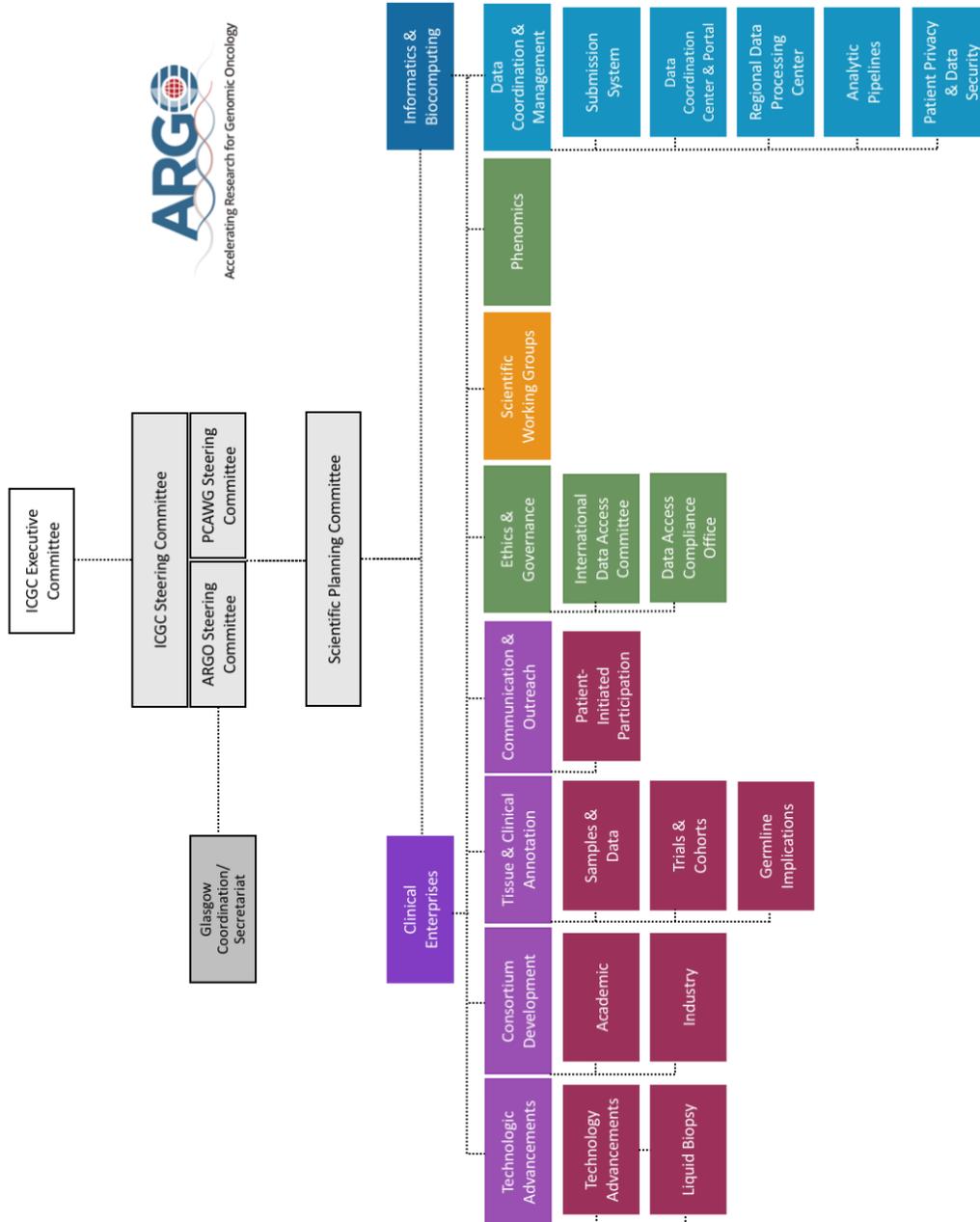
Structure of ICGC-ARGO

The ICGC ethos of partner members sharing common goals and principles whilst working in a coordinated and collaborative manner within a defined structure continues with ICGC-ARGO (Figure 1). The ICGC-ARGO Board, Steering and Scientific Planning Committees provide oversight. These committees oversee 2 broad domains: Clinical Enterprises and Informatics & Biocomputing.

Figure 1.

ICGC-ARGO

Goals, Structure, Policies & Guidelines



ICGC-ARGO

Goals, Structure, Policies & Guidelines

Governance & Coordination

ICGC Executive

A Board, constituted by individuals nominated by ICGC-ARGO Member organisations, will provide oversight of ICGC-ARGO. The ICGC Executive will continue as is with new members added but will be adjusted in May 2020 to reflect advice and emerging ICGC-ARGO Projects. (This allows PCAWG and previous ICGC projects to complete).

The Executive will:

- Review and accept nominations of new Members;
- Work closely with the Steering Committee;
- Revise or adopt new recommendations related to ICGC-ARGO policies;
- Monitor progress, data quality, and data accessibility across projects;
- Periodically report progress to funding agencies;
- Provide a forum to discuss potential overlaps that may arise between projects and negotiate solutions;
- Provide a forum to resolve issues that may arise;
- Decide about recruitment of consultants or establish expert committees on issues related to science, law, intellectual property, ethics, funding, communications, etc.;
- Develop a communications strategy, designate communication leader(s), and assure active consultation of all ICGC-ARGO stakeholders.

A Scientific Planning Committee (SPC) will be composed of the principal investigators of programmes, the Data Coordination Center, working group leads, expert pathologists, oncologists and ethicists, and representatives of funding agencies. This group will interact frequently, through phone conferences, e-mail and regular meetings, to:

- Act as a science coordinating body;
- Evaluate progress;
- Address arising issues of a scientific nature, including those related to samples, consent, ethics, quality standards, evolving technologies;
- Exchange protocols, standard operating procedures;
- Establish temporary or permanent subcommittees that would be assigned focused tasks;
- Establish QC standards.

Coordination Support

Staffing will be committed to help manage the operations of the ICGC committees.