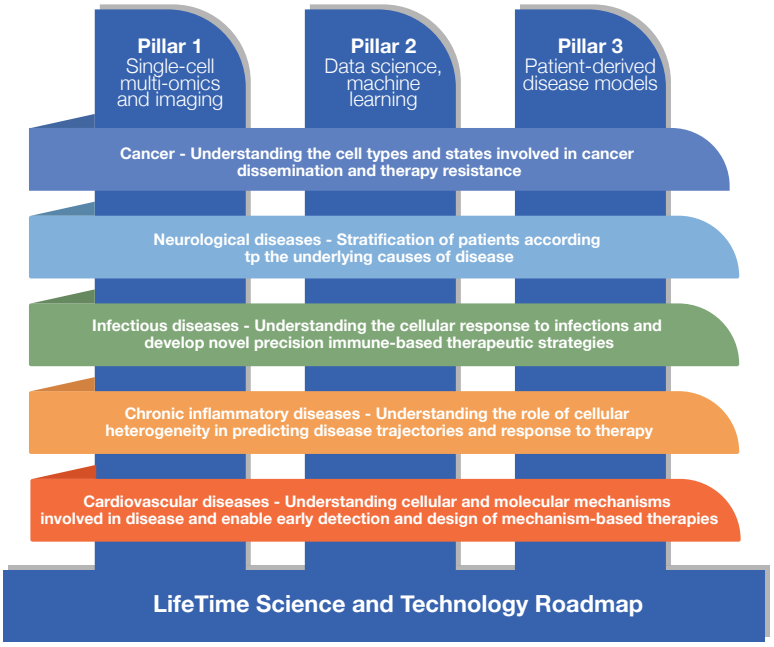
03
LifeTime Science and Technology Roadmap: Towards Cell-based Medicine in Europe

This Science and Technology Roadmap describes the technology and infrastructure developments required to implement cell-based interceptive medicine in Europe in the next decade. It details roadmaps to address key medical challenges in five disease areas and recommends funding of scientific programmes required to transition to a new way of detecting and treating diseases.

At the core of LifeTime's vision are **three technology pillars:** i) single-cell multi-omics and imaging; ii) data science, artificial intelligence (AI) and machine learning (ML); iii) patient-derived experimental disease models. The co-development and integration of these technologies will provide a suite of standardised technologies to address any key medical challenge and be subsequently deployed in the clinics.

LifeTime has identified **five key medical challenges** in different disease areas where the application of these integrated technologies will have a direct impact on patient care and outcome. The five disease areas are: i) cancer, ii) neurological and neuropsychiatric diseases, iii) infectious diseases iv), chronic inflammatory diseases, v) cardiovascular and metabolic diseases.

These medical challenges have been selected considering societal impact, evidence for cellular heterogeneity that limits current clinical avenues, availability of samples from biobanks, relevant pre-clinical models, existence of patient cohorts including those enabling longitudinal studies, clinical feasibility, issues relating to sex-and-gender-in research and ethical aspects, as well as alignment with national and EU funding priorities.



LifeTime recommends the establishment of a working group to ensure the sex-and-gender-in-research dimension and other relevent diversity aspects such as age or ethical background are taken into consideration from the very beginning in planning basic and translational research, including AI, activities.

Solving these medical challenges requires the synergistic combination of integrated technology developments in the three key technological pillars across five medical areas. Technology developments essential for introducing cell-based medicine at a population scale are described in three enabling technology sections (sections 3.1.1 - 3.1.3). These three technology pillars are built on a foundation of principles and infrastructures that will ensure that the technologies and data are fully Findable, Accessible, Interoperable and Reusable (FAIR) and secure when required (section 3.1.2). Through LifeTime, Europe has an opportunity to coordinate research efforts, standardise technologies and data formats, prevent duplication of efforts and ensure that citizens in each country have the opportunities to benefit from these developments.

The five ten-year disease roadmaps contain research programmes required to develop and apply these technologies to address key medical challenges (section 3.2.1 - 3.2.5), each disease roadmap follows the same structure with recommended investments in:

1. **Immediate research into the identified medical challenge using established, scaled single-cell technologies, computational tools and disease models**

2. **Development of new technologies required to address specific medical challenges, including the development of patient-derived model systems for longitudinal analyses**

3. **Applying new, next-generation technologies for longitudinal analyses of patient samples, or patient-derived models, combined with machine learning to generate patient trajectories and predictive models of disease. Identification of biomarkers, drug targets and validation in clinical studies**

In addition to the challenges, aspects common to several diseases have been identified across disease areas and would benefit from coordinated research and use of standardised, interoperable technologies. These include inflammation-based mechanisms and the role of metabolism in disease, as well as links between infection/inflammation and cardiovascular diseases, cancer or neurological disorders.

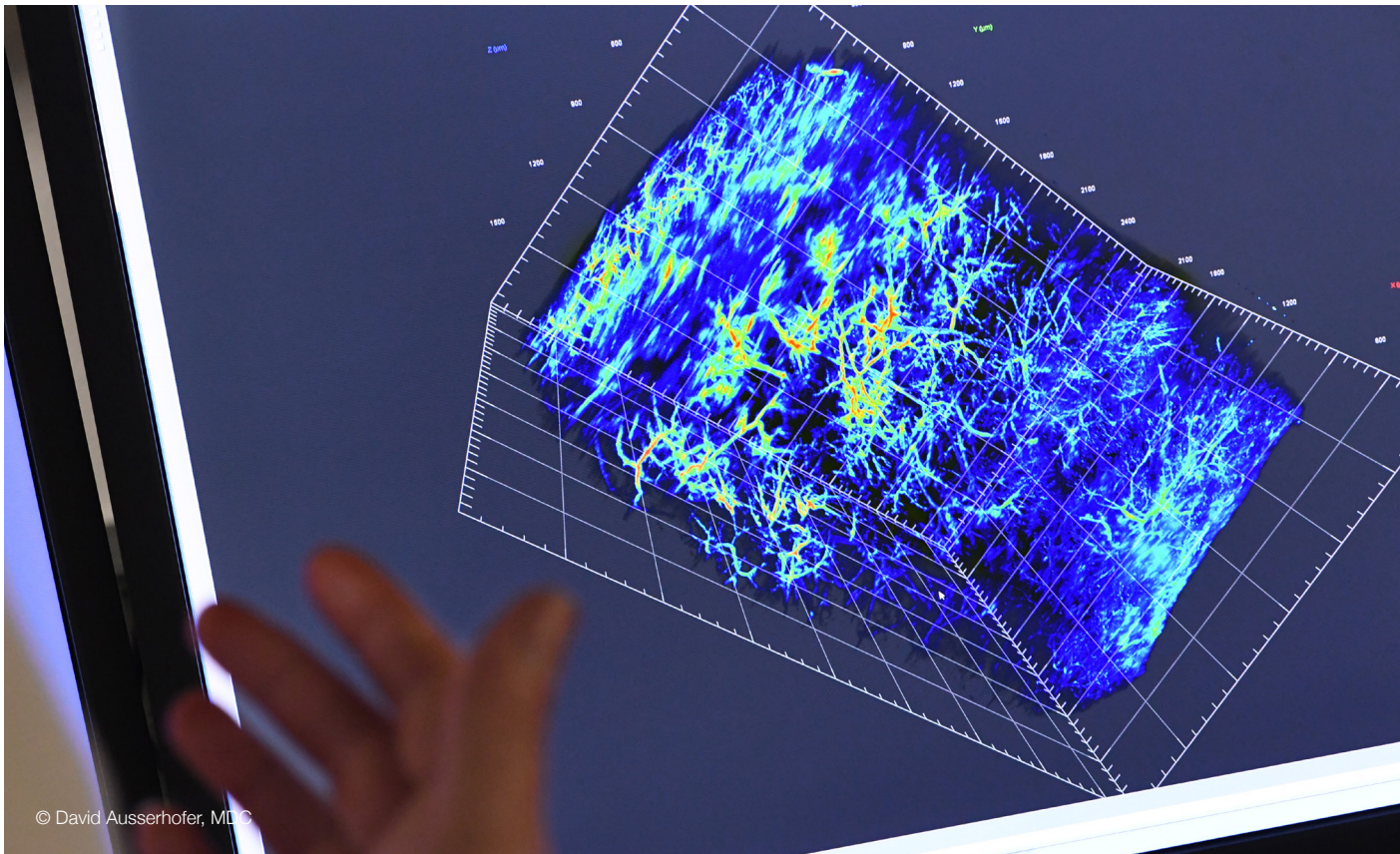
Together, implementation of the research programmes will drive the transition to data-driven and cell-based-medicine in Europe.

3.1 Developing Technologies for Cell-based Medicine

Implementing cell-based interceptive medicine requires the development and integration of single-cell multi-omics and imaging, artificial intelligence in particular machine learning and patient-derived experimental disease models. In the following sections, LifeTime identifies key technology development priorities that should be implemented in research activities addressing key medical challenges described in the disease roadmaps (section 3.2.1 - 3.2.5).

3.1.1 Single-cell Multi-omics and Imaging

Single-cell multi-omics and advanced imaging are technologies that simultaneously detect and enable the analysis of multiple types of macromolecules (DNA, RNA, proteins and metabolites) in individual cells. They are essential for cell-based medicine because disease processes within individual cells can occur on different molecular levels and no individual technology can provide a complete picture of a disease. Integration of single-cell technologies is needed to understand molecular mechanisms by describing the different molecular components, cell types and states in a tissue, the role of cell-cell interactions and tissue microenvironments and how these change through time. Realising their potential requires applying them to medical samples at different times, including during therapy, and patient-derived disease models (section 3.1.3) to reveal the status of hundreds of thousands of individual patient cells through the course of health and disease. Extracting meaningful information from these large, heterogeneous datasets using machine learning analytics (section 3.1.2) will provide a new level of information required to understand the cause of diseases, predict their future course and implement precision medicine.



Single-cell based technological solutions will impact our knowledge of disease in several ways. They will increase our resolution and understanding of disease mechanism diversity in patients diagnosed with a particular disease; enable the identification of single-cell based biomarkers to detect diseases earlier, stratify patients and select optimal therapeutic strategies; provide high-quality large datasets, including the consequences of genetic and molecular perturbations, to derive predictive models of disease; discover molecules and processes that will lead to new drug and therapy concepts based on understanding mechanisms; lead to new clinical trial designs and a new level of digital molecular pathology based on multi-dimensional data.

LifeTime recommends technologies to be developed in three main areas. The state-of-the-art of these areas is briefly described before outlining eight specific development priorities required to obtain mature single-cell multi-omics and imaging technologies.

- **Single-cell multi-omics.** These approaches are typically performed in parallel on individual cells isolated from body fluids or from dissociated tissues. Methods for detecting single classes of macromolecules (DNA, RNA, protein or metabolites) are currently at different stages of development. Transcriptomics and chromatin accessibility are mature technologies that have been scaled and can be applied to hundreds of thousands of individual cells. Detecting mutations and structural variants in the genome, DNA and chromatin modifications or chromosomal contacts are still limited to thousands of cells. Unbiased single-cell proteomics and metabolomics technologies have not routinely reached single-cell resolution. So far, combined multi-omics approaches can simultaneously detect genome and transcriptome, combinations of transcriptome, chromatin accessibility and DNA methylation, as well as DNA-protein interactions and transcriptome or chromosomal contacts and DNA methylation, but do not include information on proteins or metabolites. The sensitivity of these integrated methods and the number of cells that can be analysed typically decrease. When further developed and integrated single-cell multi-omics will provide a complete view of the different molecular layers in a cell to reveal its status and underlying molecular mechanisms.

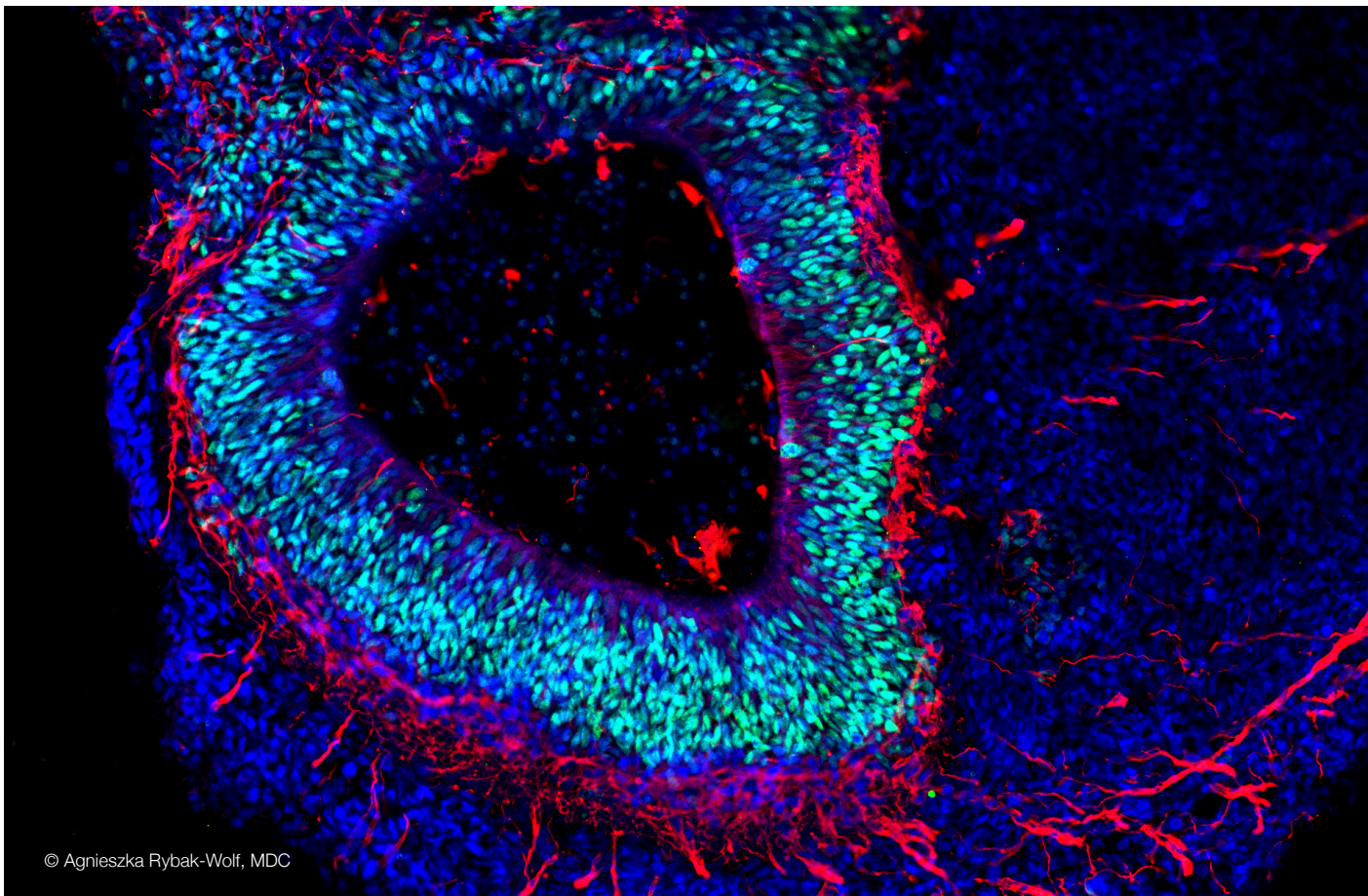
- **Imaging- and sequencing-based spatial omics.** These approaches detect single types of macromolecules (DNA, RNA, protein or metabolites) in the context of tissue sections using sequencing or imaging approaches. They are important because they maintain information on the relative position of a cell within a tissue and its neighbouring cells and microenvironment - which are key to understanding diseases. Current spatial transcriptomics methods are based on either sequencing molecules from particular regions of tissues or imaging of labelled RNA to detect rare molecules with sub-micrometer precision. Similar approaches are being applied to detect DNA and proteins with spatial resolution: spatial genomics enables the tracing of the 3D structure of DNA in single cells, while spatial proteomics and metabolomics uses fluorescence or mass spectrometry-based imaging to detect proteins and lipids. Epitope-based methods that use tagged (fluorophore or metal) antibodies can detect up to a hundred different proteins in clinical tissue samples.

- **Spatial single-cell multi-omics and advanced imaging.** Spatial multi-omics approaches detect multiple different types of macromolecules in tissues with spatial resolution. Imaging-based technologies offer the unique possibility to combine the detection of multiple molecular species in a single acquisition. For instance, 3D DNA structure can be traced while simultaneously detecting tens to hundreds of RNA

species, similarly, proteins and RNAs can also be monitored in the same cell within tissues. These technologies can be scaled and included in clinical workflows. Other advanced imaging approaches that can also detect different molecules in the same cell, such as multi-photon, light-sheet, or super-resolution microscopy can be difficult to transfer to the clinic, but will play a key role to understand disease mechanisms *in situ*. For instance for imaging deep within tissues at cellular resolution, for fast dynamic imaging of organoids and disease models, or for the visualisation of molecules and complexes in cells at sub-cellular resolution.

Building on a Fertile European Landscape

European research groups have pioneered many of these key technologies and are at the forefront of research for developing novel methods and applying them to study disease and therapy response. These include single-cell technologies for studying transcriptomes, genomes, epigenomes, proteomes and metabolomes. European groups are pioneering multi-omics technologies and integration with novel computational methodologies as well as imaging-based spatial genomics and multi-omics. They have established many approaches for spatial transcriptomics, proteomics and metabolomics, as well as super-resolution, multi-photon and light-sheet microscopy. Europe also has a dynamic ecosystem of companies with single-cell biology expertise, key proprietary technologies for single-cell isolation and single-cell analysis including spatial technologies. It also includes world leaders in the development of imaging hardware, components required for imaging, and software dedicated to image analysis. EU companies develop algorithms and computing platforms to analyse single-cell omics data and use AI to discover biomarkers. The expertise that is distributed across Europe is an excellent basis to drive the creation of next-generation single-cell multi-omics and imaging technologies, and to benchmark and standardise them to ensure rapid uptake in the clinics. More needs to be done to ensure that these technologies are taken up and commercialised in Europe to create a new single-cell technology-based ecosystem (section 4).



Priorities for Developing Next-generation Single-cell Multi-omics and Imaging Technologies

LifeTime has identified several technology developments that are required to address the disease challenges and enable large-scale profiling of patient cohorts. Implementation of these recommendations in specific research programmes is described in the individual disease roadmaps (section 3.2.1 - 3.2.5).

Development of robust, high-resolution and single-cell technologies for different macromolecules

Transcriptomic approaches have been scaled, standardised, benchmarked, are commercially available and can be performed on hundreds of thousands of cells. These gene expression data are used to identify different cell types and states in tissues but do not provide the whole picture. Understanding disease mechanisms requires additional knowledge of the genome and chromatin state, signalling pathways involved including post-translational protein modifications and lipids as well as the metabolic status of a cell. These molecules are not routinely assessed at large-scale at the single-cell level. High-resolution proteomics and metabolomics for analysis of small cell numbers, towards single-cell resolution with a useful depth of analysis, will require development of fast and sensitive mass spectrometry instrumentation (such as Orbitrap Fourier transform mass spectrometry and MALDI) and also new sample preparation strategies. For example, cell digestion strategies and methods for coupling of microfuildics platforms to mass-spectrometers should maintain the cell state and minimise the loss of the miniscule amounts of proteins and metabolites in individual cells.

Novel single-cell multi-omics modalities to understand disease mechanisms

Current single-cell multi-omics approaches have focused on combining transcriptomic and chromatin/epigenomic information. To reveal the complexity of cell states, novel multi-omics combinations are required that incorporate additional information on protein levels and modifications, metabolites and genetic variation. These will be applied to profiling of cells acquired from liquid biopsies, dissociated tissue biopsies or patient-derived organoids. They should also be integrated with techniques providing functional information on cells where required. For example, coupling with single-cell electrophysiology to record neuronal activity can inform about changes in the circuitry in neurological and neurodevelopmental diseases. Innovative multiplexing strategies, such as barcoding and cell hashing, using barcoded antibodies, lipids, or click chemistry are required to increase the throughput of approaches to make profiling at the cohort scale affordable and technically feasible. Coordinated efforts are required to develop reagents, protocols, and benchmarks to assist research groups with multiplexing strategies, which would accelerate the adoption of new methods.

Developing sensitive, high-resolution imaging- and sequencing-based spatial omics technologies

LifeTime has identified the development of next-generation spatial omics technologies to be a high priority. These developments will be important to address disease challenges, such as understanding the role of the tumour microenvironment in cancer, the functions of immune cells during homeostasis and immunotherapy, the link between cell types and pathological lesions in the brain, infections as well as specialised structural regions of the heart. The first step is to increase the sensitivity and resolution of technologies that detect specific types of macromolecules (either RNA, DNA, protein or metabolites). Sequencing-based spatial transcriptomics methods need to reach the subcellular resolution of imaging-based methods. In addition, the number of RNA species that can be profiled per spatial unit needs to be increased while acquisition times and costs should be drastically reduced. The number of RNA molecules and DNA loci detectable by imaging-based spatial transcriptomics/genomics is still relatively low (few thousands at most), which needs to be improved. Sample processing should be adapted and/or optimised, so these spatial methods can be applied to existing tissue collections, which are routinely stored as formaldehyde-fixed and paraffin-embedded sections. There is also a need for increased resolution, speed and sensitivity in spatial proteomics and metabolomics approaches. Typically these involve mass spectrometry-based technologies, which require development and improvement in instrumentation and data interpretation. Instrumentation needs to be simplified and commercialised to make it widely applicable, especially in clinical settings. Furthermore, development of new reporter chemistries is needed to increase sensitivity, resolution and the number of proteins and metabolites that can be simultaneously detected.

Novel spatial single-cell multi-omics and advanced imaging technologies

Next-generation spatial transcriptomics, genomics, proteomics and metabolomics approaches need to be integrated to create novel spatial single-cell multi-omics technologies that can detect multiple types of macromolecules in the same cells within tissues. This is required to discover hidden morphological features, facilitating the development of digital molecular pathology by visualising localisation of disease or patient-specific biomarkers, assisting early assessment of treatment, and improving prognosis. This

will require improvements in acquisition and analysis pipelines to achieve robust, fast, reliable and user-friendly imaging-based multi-scale spatial omics. This will enable acquisition of tissue-wide (millimeter-scale areas), spatial information on the presence and abundance of hundreds of thousands of different molecules in conjunction with other molecular features at cellular resolution. These developments should be accompanied by the establishment of standard data formats and validation datasets. Super-resolution and light-sheet microscopy should be further developed to achieve higher spatial resolutions, higher throughput and multi-scale imaging capabilities in tissues, organisms and organoids. Altogether, the coordinated development, optimisation, standardisation and cost-scaling of these technologies will accelerate the introduction of spatial multi-omic applications to medicine.

Computational methodologies for data integration and analysis

Creating new single-cell multi-omics and imaging technologies cannot occur without parallel advances in computation. Integration of multiple modalities requires new computational approaches, including those to register images and omics profiles. Analytic computational approaches are to reduce confounding factors, visualise highly multiplexed data from hundreds of thousands of cells or more, reveal molecular features and subpopulations hidden behind noise, and ultimately discover links between cell types, phenotypes, and states and their molecular networks. Approaches are required to provide molecular and spatial multi-scale models ranging from molecules to organs and humans (section 3.1.2).

Integration of perturbation screens and lineage tracing with single-cell multi-omics

Establishing the causative molecular changes that lead to disease and identifying the cells-of-origin will require integration of CRISPR-based technologies with single-cell multi-omics approaches. CRISPR-Cas technologies can be used to remove specific molecules from patient-derived disease models (section 3.1.3) and the disease-relevant consequences monitored using single-cell multi-omics approaches. Genetic lineage tracing approaches, including those using CRISPR-Cas modifications, provide information about the number of progeny that derive from a cell of interest, as well as their location and their cellular state. Coupling CRISPR-perturbation with spatial multi-omics technologies will provide much needed spatial information in the context of genome-engineered and patient-derived disease models. This will not only be important to understand the early events during disease onset or progression, but also to identify key molecules and cells for drug or therapy development. These approaches will also provide the large datasets required for the development and iterative evaluation of machine learning-based computational models of diseases, which will be critically needed to predict disease outcomes.

Benchmarking and standardisation

Newly developed multi-omics technologies need to be optimised, benchmarked and standardised to ensure that they produce high-quality data required for rapid adoption by the community and downstream analyses including machine learning-based computational approaches. Coordinated efforts will provide

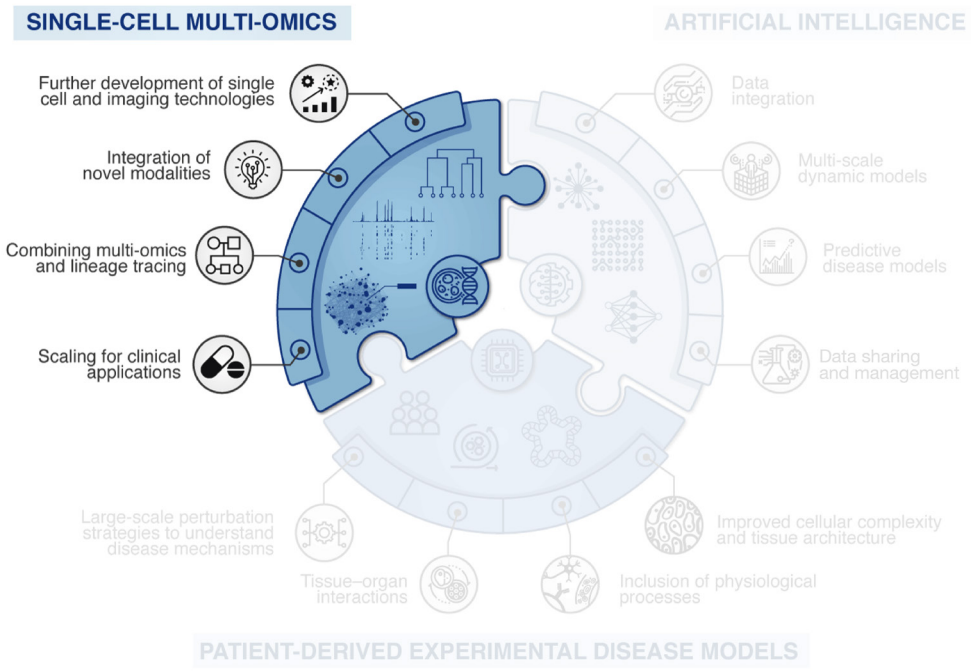
standardised protocols to ensure robustness and reproducibility and assure interoperability of datasets. This will also include reagents, for example establishing antibody standards for spatial proteomic approaches including cell lines for validation and longitudinal normalisation, as well as standardisation of procedures for clinical sample acquisition or recording of the sampling procedure and sample fixation.

Robust clinical sample acquisition and processing

For many diseases, most of the sampled tissue is composed of many "normal" healthy cells, while the cells relevant for pathology can be extremely rare. Major challenges to routinely applying single-cell approaches to clinical samples include the cost and ability to systematically process patient blood samples and tissue biopsies from the site of pathology. Strategies and technologies are required in order to obtain patient material containing high-quality dissociated cells for live processing. This will require development of affordable instrumentation for routine portable/robust acquisition of single-cells in the clinics, for example, new devices for single-cell processing of needle biopsies. Importantly, introduction of single-cell acquisition protocols into clinical procedures will facilitate acquisition of high-quality single-cell samples during routine diagnostic or operation procedures.

Preservation of patient samples (e.g. in single-cell biobanks) maintains the natural state found in a tissue and is required for subsequent sampling and downstream analysis. This will require the development of standardised protocols for cell or nuclei preparation that maintain an unbiased composition and unperturbed state of the tissue.

For patient sample processing, high-throughput microfluidic technologies for enriching specific types of cells will enable sampling of thousands of rare, disease-relevant cells (e.g. circulating tumour cells, circulating immune cells from tissue) routinely from thousands of patients across multiple centres. This will enable longitudinal monitoring of patient disease progression, response to treatment and relapse.



Single-cell multi-omics and imaging technology developments

Expected Scientific Impact

In the short term:

Currently available technologies will be used to study disease processes to identify the cell types and pathways involved in disease and develop biomarkers

New technologies will be developed and scaled, so that they can provide additional levels of information that is crucial to understanding disease mechanisms and discovering ways to target the cells involved

Development of instrumentation from prototypes to successful commercial products will be done in partnerships between academia and industry. This should be supported by joint technology development programmes and adoption platforms across Europe (section 4)

As technologies mature and are integrated, Europe should take a lead in benchmarking and setting standards to ensure interoperability

In the long-term:

Innovative multi-omics technologies will be able to profile and map hundreds of thousands of patients' cells for entire patient cohorts. Their scaling should ensure robustness and affordability, so that they can be applied to in-depth analysis of longitudinal human samples

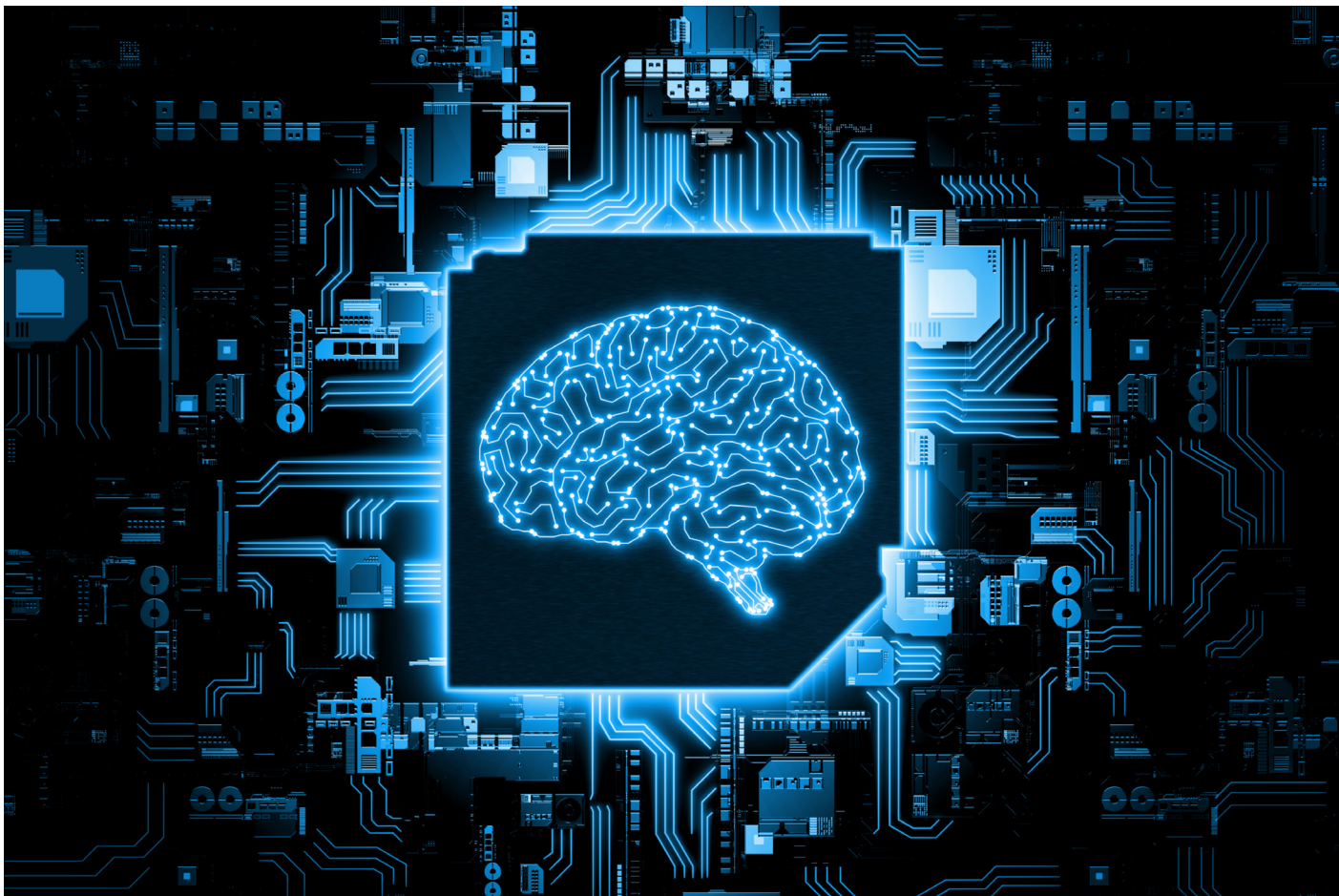
Multi-dimensional descriptors of cell states from patients taken from different stages of disease or therapy will be used to derive new biomarker panels and guide therapeutic strategies. This information for earlier detection of diseases depending on a patient's particular single-cell profile provides a rational way for interceptive medicine

Eventually, these technologies would be incorporated into digital molecular pathology and together with clinical and medical imaging data, the machine learning-based models used to inform physicians

3.1.2 Data Science, Artificial Intelligence and Machine Learning

Data-driven cell-based medicine - the real-time integration of molecular and cellular data generated using cutting-edge technologies with clinical information - will provide critical information for clinicians as they diagnose their patients, transforming healthcare across Europe and globally. Its successful implementation depends on the development of new computational methods - using advances in artificial intelligence, in particular machine learning, to identify complex associations between different molecular features and distinct clinical outcomes. Unlocking the full potential of novel molecular features depends upon access to large amounts of high-quality training data, so that complex patterns can be reliably learnt and their output can be made interpretable for humans. LifeTime's vision is to create the required high-quality big data at unprecedented scale and resolution (section 3.1.1). These will form the basis for the development of novel predictive computational models of diseases based on changes in thousands of molecular and clinical readouts from hundreds of thousands of individual cells at different stages of health and disease from large cohorts of patients. The derived predictive machine learning-based models will be tested and refined in an iterative manner, by performing systematic large-scale perturbations in patient-derived experimental disease models monitored by single-cell multi-omics and imaging approaches (section 3.1.1 and 3.1.3).

Computational and big data solutions will provide unique opportunities to drive early detection and interception of diseases. They will enable the discovery of unanticipated disease relationships, derivation of novel hypotheses and facilitate making predictions for disease prognosis and optimal therapeutic strategies. To fully exploit the large amounts of molecular health data in infrastructures and national health systems, LifeTime will address key issues of securely storing, sharing, accessing and jointly analysing data across national borders.



LifeTime recommends developments in three main areas, including computational approaches and infrastructure. The state-of-the-art of these areas is briefly described before outlining specific development priorities that are required for data sharing and analysis.

- **Large-scale data sharing and management.** Europe has invested in infrastructures for storing and accessing large scale biological and medical data (e.g. ELIXIR, European Open Science Cloud (EOSC), Federated European Genome-phenome Archive (EGA)). These infrastructures and frameworks provide the foundation for scientists to find and share data, exchange expertise, and define best practices on a European level. Differences in European healthcare systems render accessing and integrating data from Electronic Health Records (EHRs) for research across borders particularly challenging. Data are usually not accessible outside a national, regional clinical care system or specified data 'safe haven'. When data are accessible, accredited systems are often required for storing the data and information governance may be at the hospital, federal or international level. Recent international population and disease specific cohorts such as the UK BioBank, have associated molecular omics data generated at scale and represent an essential bridge to healthcare data. In the future, a flexible ecosystem of computational infrastructures will be required to enable data sharing, access and querying across Europe.

- **Computational analytics.** Artificial intelligence and machine learning are already an integral component for biomedical data science. With the first applications being focused on biomedical images, machine learning is increasingly transforming analysis paradigms for virtually all data types and fields, including genomics, proteomics, metabolomics, imaging, EHR, medical imaging as well as clinical data. Machine learning approaches have been successfully applied to medical imaging where they are able to diagnose disease at a similar level as a clinician, as well as predict the onset of some diseases based on EHR. The field of machine learning is evolving rapidly, where, in particular, methods based on deep neural networks are gaining more and more importance. New developments will be essential for integration, scaling and analysis of different types of data to create predictive models that identify actionable disease features.

- **Data visualisation and dissemination.** Extracting knowledge from complex biomedical data critically depends on the deep integration of algorithms, software, models and statistics, deep learning, and artificial intelligence with visualisation tools. Complex datasets are commonly visualised by projecting them into lower-dimensional spaces, for example using dimensionality reduction methods. As the volume of medically relevant data grows physicians increasingly depend on digital decision support systems to quickly identify the patient's most relevant medical data and patterns, apply the latest medical insights and explore the best treatment options. More advanced forms of these tools will need to be driven by sophisticated machine learning-based approaches on diagnostic and therapeutic data from real patients.

Building on a Fertile European Landscape

Europe is strongly placed to build on pan-European investments in computational infrastructure and computing capacity for biomedicine. The European Bioinformatics Institute (EMBL-EBI) provides central databases, tools and software to align, verify and visualise the diverse data and make that information freely available to all. The EGA is a service for permanent archiving and sharing of all types of personally identifiable genetic and phenotypic data resulting from biomedical research projects. EOSC aims to establish a trusted, virtual, federated environment in Europe to store, share and re-use research data across borders and scientific disciplines. Coordinated infrastructures such as ELIXIR for data, Euro-BioImaging for imaging and EuroHPC and Partnership for Advanced Computing in Europe (PRACE) for high performance computing provide the standards, interoperability strategies and computational power necessary to deliver LifeTime's vision.

Complementary to data infrastructure, initiatives such as the European Laboratory for Learning and Intelligent systems (ELLIS), foster excellence in machine learning research. This provides an excellent basis to develop an ecosystem where innovation at the interface of artificial intelligence and biomedicine in academia and industry can flourish (section 4). Investment is required to link biomedical research and clinical data and drive application of knowledge for patient benefit and to improve outcomes. European small and medium enterprises (SMEs) are developing tools for downstream functional interpretation and navigation of LifeTime-related data. They include bioinformatics and data science expertise, image analysis, patient omics data-driven medicine, as well as AI-based decision-aiding systems required to integrate and interpret available molecular, cellular, individual disease trajectory and imaging information. Several companies and EU-initiatives are providing solutions for secure and European General Data Protection Regulation (GDPR) compliant data management platforms. These platforms will make derivatives of data, which have been collected and are stored locally at hospitals and institutes, available for consolidation and subsequent AI-based analytics. Together, the European landscape is well positioned to rapidly provide returns on investment and implement data-driven and cell-based medicine within a decade and form a dynamic healthcare ecosystem.

Priorities for Developing Computational Infrastructure and Tools

LifeTime recommends prioritising the below computational developments to address the disease challenges and to enable the derivation of patient specific computational models that can be used to predict disease outcome and guide treatment. Implementation of these recommendations in specific research programmes is described in the individual disease roadmaps (section 3.2.1 - 3.2.5).

Large-scale Data Sharing and Management

LifeTime has broad infrastructural and data management requirements, encompassing support for methods development, bioinformatics, clinical informatics and imaging and with the assumption that the data conform to the Findable, Accessible, Interoperable and Reusable (FAIR) principles. Consequently, a 'LifeTime ecosystem' of interoperable resources will be required and linked to and expanding on infrastructure that are either available or in development by existing projects (e.g. Human Cell Atlas (HCA), Global Alliance for Genomics and Health (GA4GH)), as well as leveraging investment in European infrastructures (e.g. ELIXIR, EOSC, Federated EGA). LifeTime's data platform will be based on a federated approach with control of the data, especially with respect to clinical data or controlled access data, such as EHR, and is likely to remain within national or federal jurisdictions.

Data storage and coordination

Biomedical data sharing has advanced considerably with the creation of large and complex datasets based on single-cell and imaging technologies. These include the development of the Human Cell Atlas Data Coordination Platform (HCA-DCP) - a unified cloud based modular system for single-cell transcriptomic data offering end to end data management and analysis. While the platform appropriately covers biological data management it requires improved integration of clinical and image data. LifeTime therefore recommends the development of a flexible and extensible ecosystem of resources linking existing infrastructures such as ELIXIR and Euro-BioImaging, Clinical Systems and project specific efforts. This will require significant investment in computing, cloud storage, novel visualisation and query technologies and is an unparalleled opportunity to shape Europe's research data landscape. Optimising cloud usage for analyses, and delivering rapid integration via technologies such as Jupyter NoteBooks will enable access to data in specialised cellular atlases and improve integration with clinical data providing lasting impact by delivery of an extensible ecosystem and by cloud portability of analysis pipelines. As the number of disease atlases increases, these technologies will be extended to a scenario of federated atlases with common semantics and linked to omics datasets.

LifeTime recommends the implementation of common metadata standards to ensure that data are interoperable across technologies and with clinical data such as EHR. The initiative has the potential to generate a resource that has an even greater impact than the Human Protein Atlas (HPA), which provides large scale reference datasets and browsing tools. The HPA has invested significantly in the annotation and presentation of datasets and such investment is required also for LifeTime, to ensure data is accessible by the user community both for analysis and for incorporation into the clinical practice. With the integration

of large and complex datasets LifeTime recommends the inclusion of domain experts to coordinate the acquisition, quality control and curation of data. LifeTime should not only implement but also evolve technical standards for data quality, metadata as well as data sharing, similarly to those recommended by the GA4GH designed to bring omics data and clinical data standards together. These activities will be aligned with the European GDPR.

For controlled access to omics data, LifeTime proposes a federated solution involving local EGA installations using the ELIXIR developed Local EGA infrastructure. A Local EGA can be operated within the GA4GH data security framework and has the potential to be deployed on a secure cloud with GA4GH encryption standards. For LifeTime this will provide the necessary collocated access to the data and to computing, for example at existing high-performance computing facilities or via private clouds. This will address challenges of sharing data across borders by co-locating storage and analysis with data, and enabling federated analyses via common application programme interfaces.

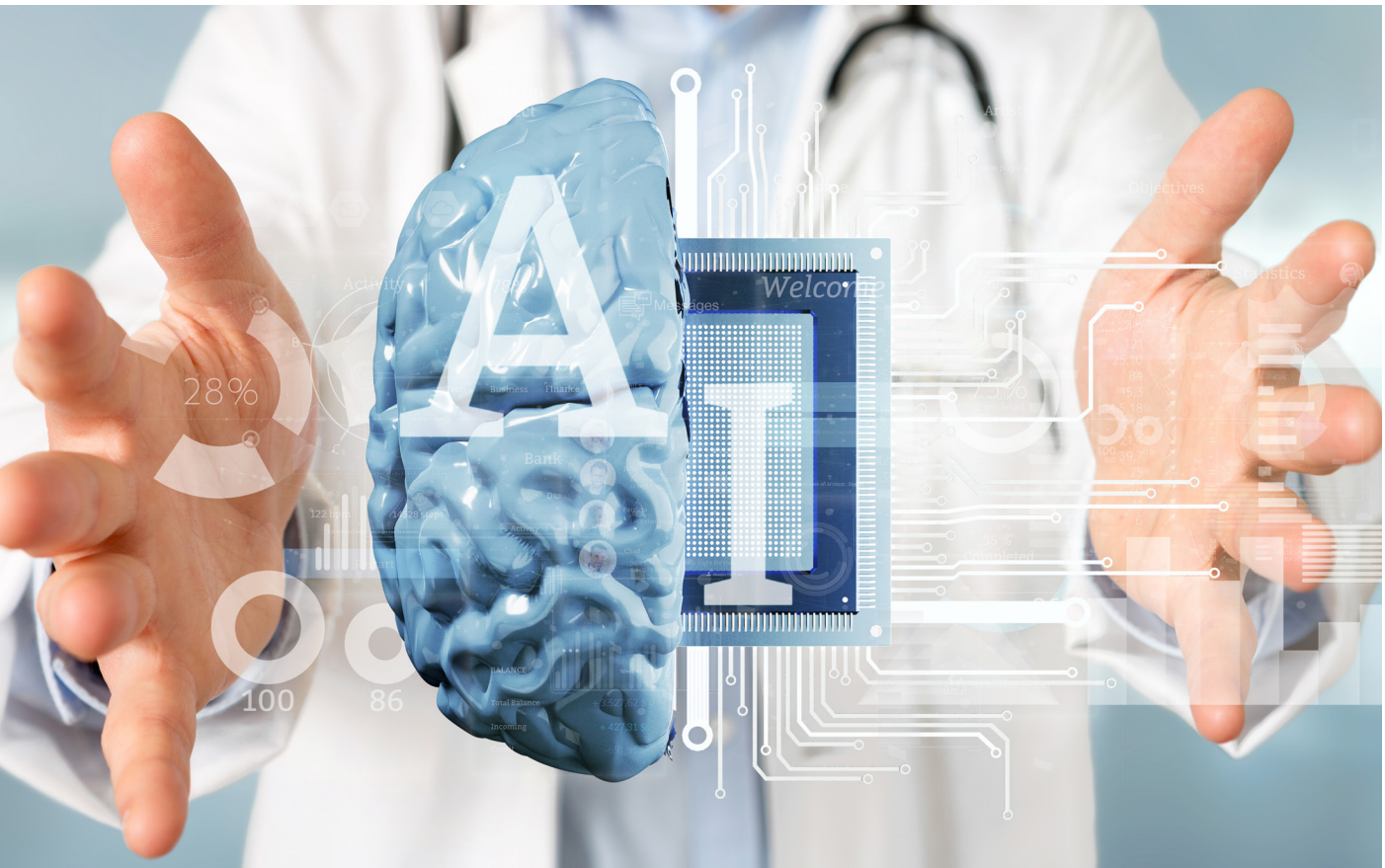
Access to open and controlled data

LifeTime‘s data platform will need to provide access to federated data subject to ethical constraints associated with national and pan-European models. Existing Authorisation, Access and Identification (AAI) protocols such as ELIXIR AAI can be deployed across national platforms and pan-European platforms enabling secure access to health records that typically reside in safe havens, and/or national mandated secured platforms residing in accredited national infrastructures or accredited national or institutional private clouds.

LifeTime provides an exemplar use case for driving European strategy for sharing health data across national borders, as part of the European Commission's Digital Single Market strategy by exploring open EHR exchange standards as a tool for improving portability of data. Such a standard could be driven by LifeTime and data consumed by LifeTime and we identify an opportunity to extend LifeTime's constituency with representatives of the OpenEHR community. Such activities may be considered as part of LifeTime's infrastructure network, though in the short-term access is likely to be within, or transfer of analysis to, pipelines in a secured environment. Where EHR or other health data is accessed within LifeTime, de-identified data can also be supplied for some research questions, using the UK BioBank model of linked clinical data.

Federated queries and analyses

European healthcare systems contain large amounts of rich health data. Today it is feasible to collect longitudinal health data on a national as well as cohort scale, which should be used to improve the quality of delivered healthcare. Analysis of this large-scale data by machine learning algorithms requires large training datasets to make accurate new discoveries. Beyond the technical challenges involved in producing such high-dimensional records, there are concerns about re-identification and associated privacy concerns. Regulations addressing access to clinical data necessarily introduce higher security standards, but anonymisation is intrinsically limited when dealing with high dimensional data such as medical records. LifeTime therefore recommends focusing on developing federated learning approaches for training of machine learning models while keeping records in decentralised trusted data warehouses at the hospital, avoiding single points of failure. Standardisation of federated access using minimal data models for harmonisation has been demonstrated in projects such as OSIRIS in cancer genomics. Both OSIRIS and PhenoPackets, a GA4GH standard for phenome/genomic data sharing have the potential to be adopted for cross border data sharing and interoperability with PhenoPacket already used internationally.



Computational Analytics

Computational approaches for integrating data from different modalities

While integrating data obtained from different experimental modalities will be extremely powerful to understand the molecular basis of disease, there remain significant challenges. These include the indirect relationships between different types of features (genes, chromatin regions, proteins, metabolites), differing coverage of the data as well as the multi-scale integration of complex datasets. Additionally, approaches are required to link single-cell data across scales, to make clinically relevant predictions on patients and their future health trajectories at the level of organs and patients. LifeTime proposes to focus on developing machine learning principles, such as multi-view learning and variational autoencoders, to project data from different modalities, including both multi-omics and spatial data, onto a common coordinate system, and construct a multi-space similarity measure that takes into account the different omics layers. A major avenue of future research is to derive methods that span different scales and have latent representations for individual cells while simultaneously accounting and modelling patient covariates and clinical variables. Concepts such as transfer learning will be important developments to fill in incomplete missing information from the different molecular layers.

Scaling computational analyses to higher dimensional data

Computational scalability and extensibility of computational approaches to cope with unprecedented volumes of single-cell imaging and omics data across large patient cohorts is essential. Currently, most analyses are based on compiling large datasets at a central location where analytical models are trained, which is problematic both for scaling as well as data sharing and privacy reasons. The scale of data that will be generated by LifeTime will foster novel innovations. LifeTime recommends developing modular models and software to support dynamic and online training, directly connected to data influx. The federated nature of European data deposition will require sharing of trained models and training schemes whereby the same models are iteratively refined based on the data available at different locations. Finally, there are important opportunities to compress large-scale datasets into models rather than sharing raw data to promote new modes of collaboration. This model-based dissemination of data proposed by LifeTime, for example using highly parameterised models such as deep neural networks, will permit sharing insights and democratising access to complex insights without the need for users to train these models on local high-performance computing systems and having access to potentially GDPR-protected data.

Integration of multi-omics molecular data into electronic health records

Data accessibility and governance have been, and remain, the primary challenges in accessing and integrating data from EHRs for research purposes. To integrate molecular data into health records and present data in a meaningful way to clinicians, LifeTime will address several challenges. These include: i) representing multi-omics data in clinical records, ii) a lack of reference data structures, iii) rapidly changing omics data and different evolution cycles between omics data and clinical standards, iv) lack of data and EHR

relevant standards as well as analysis pipeline execution standardisation, and v) application programming interfaces amenable to clinical data workflows.

LifeTime recommends prioritising the development of tools to process standardised data, such as single-cell data or summaries for inclusion in EHR, in collaboration with clinical information standards bodies such as Health Level Seven International (HL7). Molecular data from LifeTime will be made available via simple and lightweight data sharing tools aligned with the Human Cell Atlas Data Coordination Platform approach. LifeTime will work towards making EHR data formats consistent and interoperable, with a view to cleaning and standardisation of the data required to describe disease progression using machine learning. To address this, the initiative will explore the different standards pioneered by initiatives such as the PhenoPackets proposed by GA4GH to ensure interoperability of data with EHR. PhenoPackets 'packages' genomic information into a standard format, and adheres to the FAIR principles by the use of ontologies to describe individual phenotypes. A data exchange format compatible with Fast Healthcare interoperability Resources (FIHR) and supporting software libraries could be extended to handling image data and new applications, such as biomarkers based on cellular data. LifeTime recommends this path is followed to leverage the momentum and impact of GA4GH. Projects such as the federated European Health Data and Evidence Network (EHDEN) goes beyond individual records as it provides an effective way of analysing large quantities of data (millions of EHRs). It relies on the Observational Medical Outcomes Partnership (OMOP) model providing a "common data model" for databases to label and structure their data consistently.

Generating predictive models based on patients' cellular trajectories

Cell-based medicine involves the longitudinal sampling of patients. This requires integration of additional dynamic information into computational approaches for inferring continuous changes in cell types and states across disease and time. To infer additional information on cellular trajectories during disease onset and progression LifeTime recommends incorporating extra information such as multiple modalities, experimental time and spatial information. This will provide a more complete picture of cell transitions and lineage relationships. Creating predictive trajectory models requires additional information on directionality and dynamics such as that provided by sampling and temporal information, lineage tracing, and information on the dynamics of biochemical reactions. Such predictive models will infer future dynamics from observed data, combining sparse and incompletely sampled data at multiple time points with adequate models of uncertainty.

With a particular focus on spatial multi-omics approaches (section 3.1.1) LifeTime recommends developing computational tools to combine information from several time points as well as multiple patient samples with lineage tracing and multi-modal, including spatial, readouts. Functional single-cell readouts from spatial omics approaches such as morphometries could be used to determine cell state and transitions across lineages, these would improve lineage reconstruction and produce more accurate trajectories.

To robustly identify and annotate the resulting cell dynamics, it is paramount to be able to map to a reference atlas as for example generated in healthy subjects within the Human Cell Atlas project using machine learning and deep learning-based projection methods. These data integration tasks need to be sufficiently constrained to guarantee that disease states are not regressed out and thus need to be benchmarked accordingly.

Differentially comparing cellular trajectories between cells in a tissue, between organs or between individuals through time will require not only method transfer from the field of time series analysis (such as change point detection) but also the use of a coordinate system, as proposed by the Human Cell Atlas. This will be challenging, since various factors (e.g. age, height, gender) can lead to differences in an organ's morphology, making mapping onto a common template challenging. LifeTime recommends leveraging machine learning methods to account for covariates to robustly describe the underlying integrated latent space.

Establishing causality for molecular mechanisms

Modelling disease progression based on longitudinal cellular readouts will not only enable predictions but also enable discovery of the underlying molecular mechanisms. While the proposed machine and deep learning techniques allow for the scalable and robust incorporation of variation across samples, the resulting network architectures make interpretation of the sources of variation difficult. LifeTime proposes to address this limitation by going beyond pure statistical approaches and constrain the networks towards interpretable models, which will be key for explainable AI-based approaches to be accepted for clinical use. This would include backpropagating variation in classification in supervised learning, or incorporation of mechanistic models in the network and extend these tools to the mostly unsupervised and non-localised/convolutional

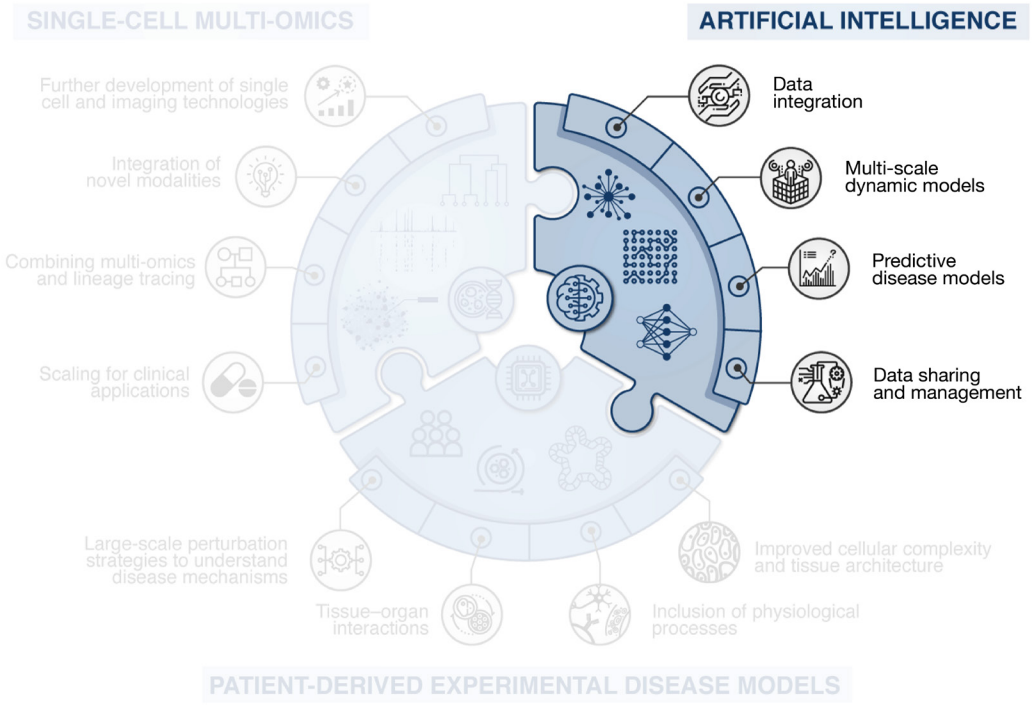
methods employed in single-cell omics. This will reveal sets that jointly explain certain variation and hence fully data-driven pathways, which will be complemented by traditional annotation-based efforts. Integrating learned trajectories, time series information with spatial structure, will identify gene regulatory motifs and causal networks, cell-cell interaction graphs and dynamic transitions affecting disease. Enriching inference methods with epigenomic data will lead to interpretable disease regulatory motifs that can be translated to more effective treatments.

The resulting molecular mechanisms will need to be integrated within a multi-scale model of a patient, to address tissue, organ and organism-level covariates beyond a purely statistical setting. Multi-scale models of organs, generated by integrating information from tissues, cells and molecules, will simulate the dynamic interplay of factors during disease progression and treatment. LifeTime will develop the needed computational frameworks to integrate machine learning and multi-scale modelling to identify causality by modelling molecular mechanisms and predicting their effects at the tissue, organ or system level.

Validation and benchmarking computational tools

Ensuring comparability between multiple analyses will require validation and benchmarking of computational approaches. Through LifeTime, Europe is well placed to implement standardised benchmarks for datasets and methods needed for data-driven medicine. LifeTime recommends defining datasets (where possible with known ground truth) and metrics for systematic benchmarking of the developed methods. This includes defining which aspects of analysis can be performed using common strategies across multiple biological contexts and which require genuine context-specific tools. The initiative will drive the identification of more

accurate metrics for assessing the quality, robustness and impact of different computational strategies on downstream biological inference to be implemented in clinical settings. These metrics will include self-consistency under subsampling, robustness with respect to outliers and noise, performance on gold standard data sets or more broadly identification of independently validated medical knowledge. LifeTime recommends promoting the regular development of such standards as well as open community-based competitions, also involving industry (section 4), which could lead to benchmarks - similarly hackathons and other community efforts could be leveraged to build compatible and "best-of" computational pipelines. LifeTime will ensure that such benchmarking will be done in collaboration with related initiatives such as the GA4GH and Human Cell Atlas.



Computational, artificial intelligence, machine learning developments

Data Visualisation and Dissemination

The outputs from large-scale machine learning and mechanistic models may be difficult to interpret. A key challenge is developing intuitive, comprehensible visualisation and dissemination tools and strategies that are useful to all user groups including computational and experimental researchers as well as clinicians to facilitate better diagnoses and treatment.

Complex datasets are commonly visualised by projecting them into lower-dimensional spaces, for example using dimensionality reduction and latent-space learning methods. These approaches provide accessible entry points for practitioners and non-experts to explore the complexity of data and gain insights. Resources and computation infrastructures that provide such functionality and then allow mapping model results on top of the raw data are critical to enable data and models to be widely used in the community. LifeTime recommends building upon existing efforts and in particular existing portal infrastructures and systems developed in the context of the Human Cell Atlas and other efforts. This approach will ensure that LifeTime visualisation is closely interlinked and connected to existing infrastructures, such as the bio data resources offered by the European Bioinformatics Institute. This will require approaches for the visualisation and connection of spatial omics data together with genomic data and modelling results, thereby connecting mutational, genomic and epigenomic data to tissues and their two-dimensional and three-dimensional makeup.

LifeTime recommends working closely with clinicians to create the next-generation clinical support systems to produce new AI-based decision-aiding systems that will integrate and interpret available molecular, cellular, individual disease trajectory and imaging information. Using interpretable and accountable AI systems will also provide the basis for the predictive models and clinical recommendations. Simple dashboards will enable these clinical teams to exploit data produced during the patient journey, from the patient narrative to rich imaging and genomic data. They should highlight predictive scenarios as well as actionable targets identified by modelling, as well as link these targets to available clinical performance indicators. Alongside dashboards, browsers capable of handling genomes as graphs, such as those used by the 1+ Million Genomes Initiative, would enable population genomics. These approaches will provide new medical knowledge and be used for visual purposes or as the foundation for tools such as decision support systems.

Expected Scientific Impact

In the short term:

Machine learning methods for single-cell data analysis and integration will be accessible to a wider community for broad application.

Newly developed methods will translate disparate European datasets from diverse samples into mechanistic insights of disease.

Standardisation and sharing of data by openly releasing datasets via a federated approach that is FAIR and can be computationally treated from different locations across Europe.

Standardised common formats for incorporating molecular data into EHR.

Implementation of benchmarked methods with different standard datasets and in different scenarios in clinical settings.

Integration of existing infrastructures by promoting an ecosystem of data resources with common standards and cloud portability. Scaling beyond traditional approaches and hybrid architectures will be ideal for data-driven analyses.

In the long-term:

Scaling of newly developed methods and visualisation tools to very large federated cohort datasets to enable the detection of early onset of disease and inform medical doctors of the best possible treatment not only for the specific disease but also for the current status of the disease in the specific patient.

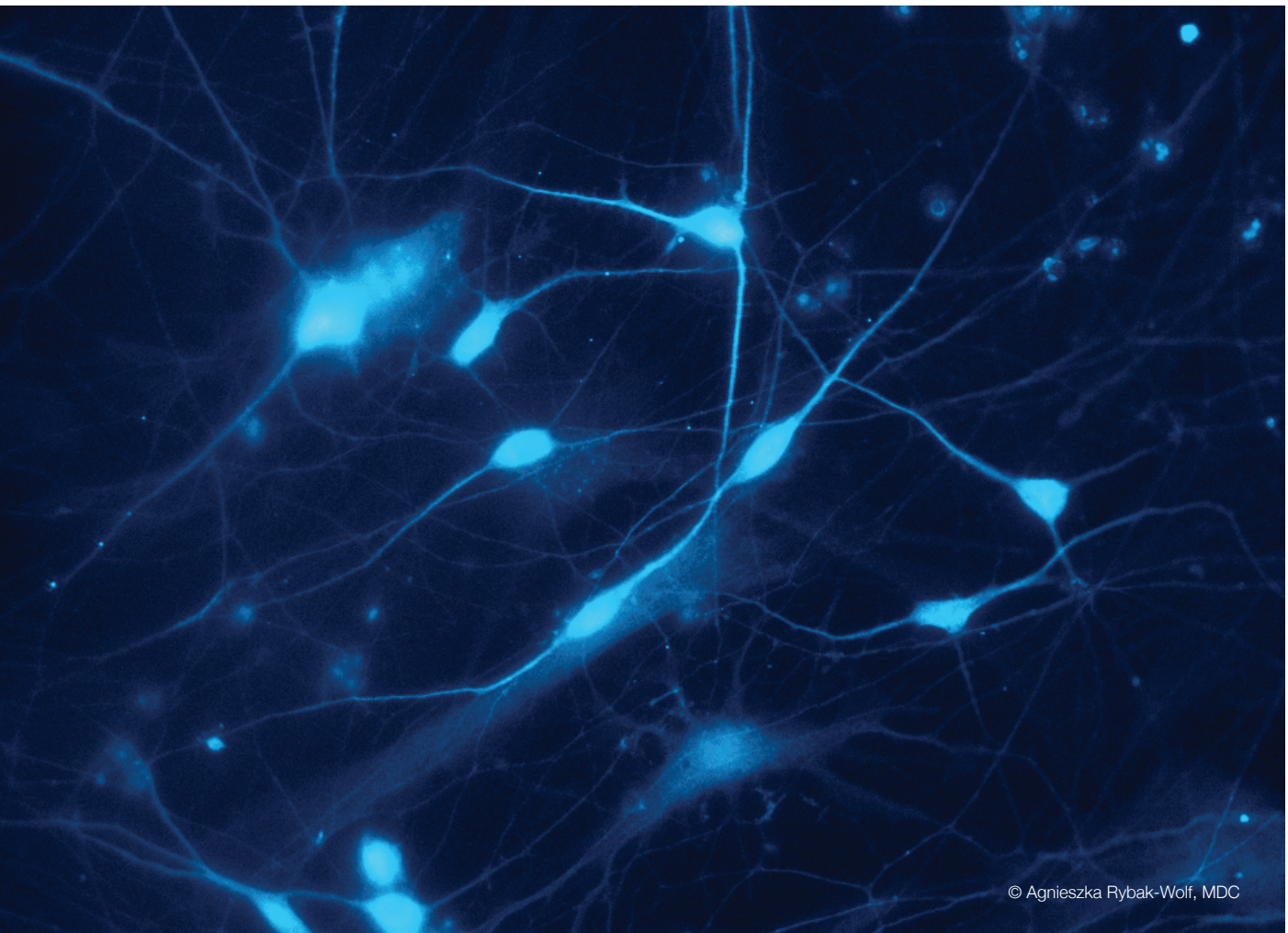
Accurate machine learning approaches to enable multi-scale models of disease in patients to make data-driven decisions on treatment.

Inclusion of novel resources into the infrastructure ecosystem to enable users to access and analyse large datasets relying on federated approaches for the benefit of patients. Such distributed learning models will need to be processed in near real-time.

3.1.3 Patient-derived Experimental Disease Models

Patient-derived experimental disease models are cell or organism-based systems that recapitulate the pathophysiological processes involved in disease. As they are derived from adult, pluripotent human stem cells or tissues, and can capture individual features that are unique to each patient. Importantly, by comparing patient-derived models from diseased and healthy individuals, unique features can be extracted even without knowing the specific genetic cause of the disease. They are especially important for studying diseases where repeated acquisition of patient tissues is limited or impossible, for instance for neurological disorders or cardiovascular disease, or genetic disorders that are caused by combinations of many unknown genes. Developing these key technologies will provide a range of models that reflect human biology and pathophysiology as closely as possible. Together with single-cell multi-omics and imaging (section 3.1.1) and large-scale genetic and molecular perturbations they will be used to understand disease mechanisms and iteratively test machine learning-based predictive computational models (section 3.1.2) that predict outcomes or most effective treatments for an individual patient.

Patient-derived models will have a major impact on the understanding of disease mechanisms. They will help discover new drug targets based on identifying the molecular cause of disease, new concepts for cell-based therapies as well as more accurately predict potential drug toxicity. Developing biobanks of patient-derived models will lead to new concepts for screening platforms that cover a wide range of the worldwide population genetic variance. Functioning as patient ‘avatars’ they will be tools for predicting disease prognosis and testing tailored therapeutic strategies and drug combinations on patients cells before administering them.



LifeTime recommends developments in three main technology areas. The state-of-the-art of these areas are briefly described before outlining specific development priorities required to create the necessary patient-derived disease model technologies and tools for cell-based medicine:

- **Patient-derived organoids.** Organoids are emerging multi-cellular three-dimensional experimental systems that model aspects of organ development, regeneration and pathophysiology. Derived from adult or pluripotent human stem cells, they have been created for most healthy tissues, as well as from diseased tissue e.g. tumours. Organoids have tremendous potential for disease research, especially for organs such as the brain where key developmental and physiological processes can only be studied in a human setting as well as enabling routine genetic analysis in human tissues. They can be used to characterise potential disease risk factors, disease specific cell types, disease-specific molecular changes and mechanistic analysis in human tissues, bridging gaps between biological research and medicine. Therapeutically, patient-derived organoids have been also used to test the likelihood with which a patient will respond to a given therapy. Further development of organoids is required in order to more faithfully recapitulate key physiological features in human tissues and to understand more precisely disease processes and mechanisms.

- **Advanced animal models containing human genes and cells.** *In vivo* model systems such as patient-derived xenografts (PDXs) and genetically engineered mouse models are necessary to translate the science from bench to bedside. They are important for understanding complex temporal relationships that occur in disease such as those involving the vasculature and blood brain barrier, the resident microbiota, the immune system and pathogens as well as neuronal networks in the brain. Additionally, these models capture the cellular heterogeneity required to understand disease progression and therapy response, which is not possible in many *in vitro* systems. Advanced models are required that more closely represent specific disease processes in humans and will be used in parallel with animal-free research until *in vitro* models are sufficiently developed to reduce their use.

- **Large-scale perturbation tools.** Understanding the molecular causes of disease in a personalised manner in human tissues will require the development of tools for systematically perturbing molecules in disease models and analysing their consequences on the single-cell level. These include both genetic (CRISPR-Cas) and pharmacological inhibitors (e.g. small molecules or next-generation therapeutics such as conjugated antibodies or protein degradation agents) combined with single-cell multi-omics approaches and imaging or lineage tracing.

Building on a Fertile European Landscape

Europe is well placed to develop and implement next generation patient-derived disease models in cell-based medicine. The recent surge in the development and use of patient-derived models, especially organoid systems, has been driven by European research groups. This includes the development of these models to study normal physiological and developmental processes, disease mechanisms and pre-clinical applications. Europe has invested in large-scale infrastructure with the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-ERIC) which links over 600 biobanks across Europe and facilitate the use of disease models and standards as well as the EuroPDX initiative which facilitates the exchange of disease models and avoids duplication in pre-clinical cancer research. Europe also has a dynamic ecosystem of SMEs involved in the standardised production and maintenance of organoid collections of different organs and diseases for drug discovery screens, drug and companion diagnostics development and pre-clinical patient stratification. Investment in European expertise and know-how would enable a coordinated effort to further develop these models, standardise and scale them to be routinely used for precision medicine. Research is required into the fundamental biology of these models and how they can be improved, which should happen hand-in-hand with model development and integration with single-cell multi-omics and imaging technologies.

Priorities for Developing Next-generation Patient-derived Experimental Disease Models

LifeTime has identified five priorities for developing next-generation patient-derived disease models. Implementation of these recommendations and integration with other technologies are described in specific research programmes in the individual disease roadmaps (section 3.2.1 - 3.2.5).

Increasing cellular, structural and functional complexity of organoids

Organoids are derived from homogenous cell populations that are directed into specific lineages using a variety of signalling molecules that are present in the *in vivo* tissue environment. Despite their potential, they need to be extended to create tissues from multiple distinct origins and contain the full complement of organ-specific cell types. Currently, they are typically more representative of fetal rather than mature tissues. LifeTime recommends increasing the complexity and maturity of organoids to create improved pre-clinical models and enable the further understanding of disease mechanisms. More precise control of tissue structures will require the development of more complex and manipulatable extracellular matrices, tissue engineering approaches such as biofabrication and synthetic biology. Increasing the reproducibility and cellular complexity of organoids needs more accurate spatial and temporal control of signalling pathways to better mimic the intracellular microenvironment.

Organoids lack many key features of tissues such as the presence of nerves, blood vessels, immune cells and microbes. These are key to understanding various aspects of neurological, cardiovascular, chronic inflammatory and infectious diseases, cancer as well as precision immunotherapies. LifeTime recommends establishing coculture systems where relevant progenitors are incorporated into the organoids at precise times during development. Engineered vasculature systems will need to be developed to facilitate diffusion of oxygen and nutrients. Microfluidic based systems to apply mechanical forces such as fluid shear stress and solid mechanical forces are required for full maturation of certain tissues. Studying disease mechanisms will require establishing protocols and careful benchmarking to ensure that different types of organoids can be studied using single-cell multi-omics, spatial and imaging technologies (section 3.1.1).

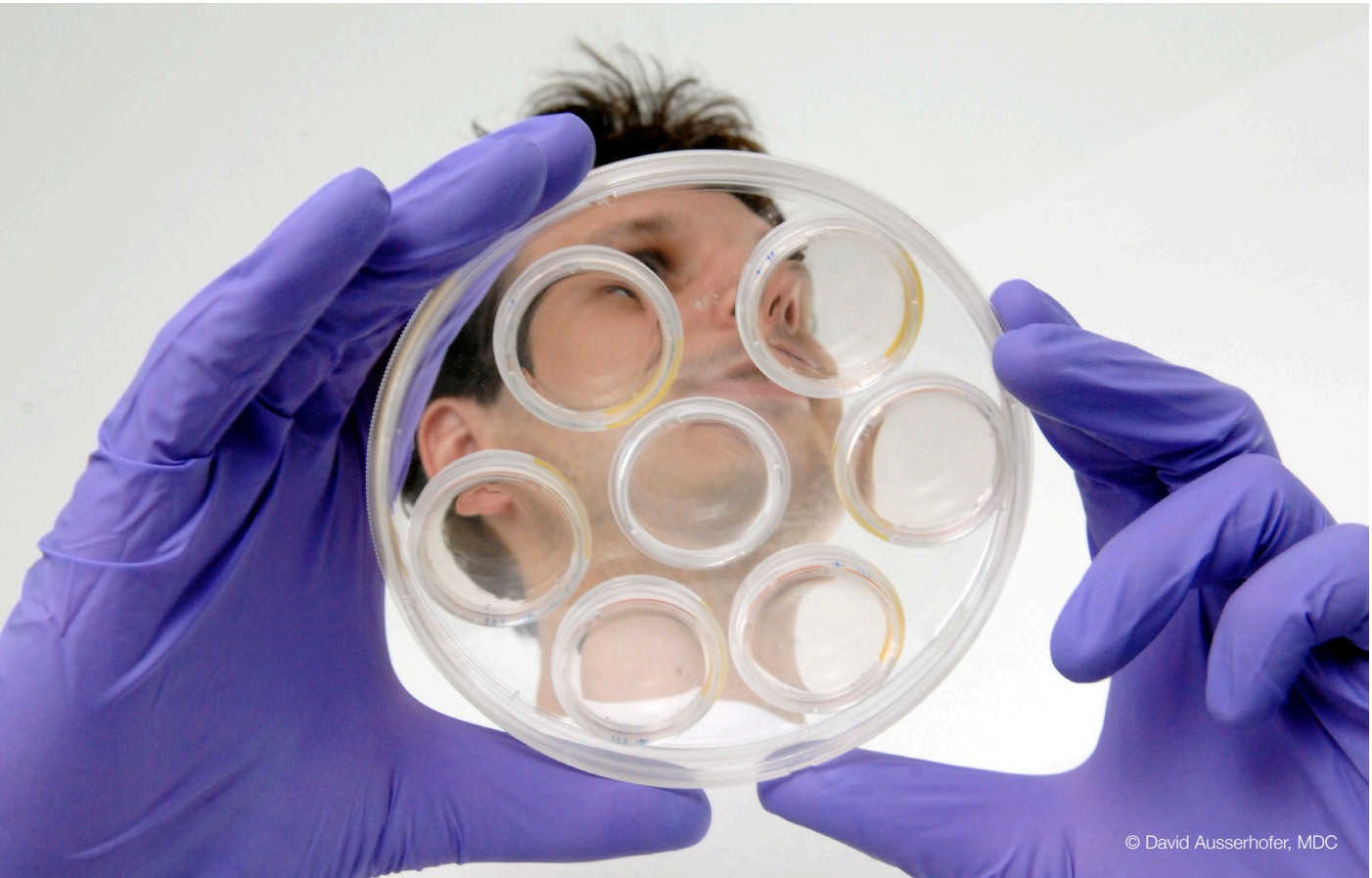
Organoid integration to study tissue-tissue and organ-organ interactions

Many disease relevant biological processes involve interactions between different tissues or organs. Modelling these processes will require the integration of different organoids. For example, the assembly of organoids from different regions of the brain (assembloids) to recapitulate mature functional circuits that may be perturbed in neurological or neurodevelopmental disorders. LifeTime recommends expansion of those technologies towards the reconstruction of major functional brain circuits are required for further insight into neuropsychiatric diseases. Fusion of organoids resembling different interacting organs and engineered tissues will allow the assembly of functional units of organs and create distinct microenvironments. For example, the multi-layered alveolar–capillary unit in the lung, which is composed of closely interacting alveolar epithelial cells and pulmonary microvascular endothelial cells which interact with air and blood to study respiratory infections. Likewise, fusion of organoid systems resembling heart muscle and blood

vessels will allow for unprecedented modelling of cardiovascular diseases. Such bioengineered systems enable the introduction of physiologically important mechanical or shear stress of fluid flow forces required to mature organoids. Since many diseases involve dynamic systemic interactions between different organs, microfluidic-based platforms will connect different organoids to simulate multi-organ interactions. Such structured organoids require novel microchip concepts that account for the stereotypical organisation of organoids and the incorporation of real-time monitoring of single-cells in the tissue context using for example electrodes for electrophysiological recordings, cell-cell interactions and metabolic processes in tissues.

Standardisation, automation, scaling and benchmarking of organoid propagation and banking

Generation of organoids is currently based on manual laboratory procedures that result in extensive inter-organoid variability. Although this allows for remarkable precision in recapitulation of endogenous processes, more reproducible and robust three-dimensional structures are required. In partnership with industry (section 4), LifeTime proposes establishing common guidelines and common practices for different tissues and organs. Scaling of organoidogenesis requires automation, optimisation of protocols and microengineering and miniaturisation of procedures and assays to reduce costs, this will require use of robotic cell handlers, three-dimensional bioprinting and microfluidic chips. To reduce costs of longitudinal single-cell multi-omics profiling of patient-derived organoid cohorts or genome-wide perturbation studies, LifeTime proposes the development of strategies for multiplexing of samples. Standardised protocols will need to be benchmarked against cell type composition of healthy tissues based on current atlas studies of healthy tissues such as the Human Cell Atlas, using several single-cell based technologies. Standardised protocols for biobanking and recovery will be developed in collaboration with BBMRI-ERIC to ensure quality control checks, so samples are comparable across different countries, facilitating reproducible and reliable data.



Advanced animal models for studying human pathophysiology and therapy response

Studying specific disease related processes that cannot be easily recapitulated in organoids will require models in which patient-derived tissues are integrated into *in vivo* models. For example, differences in the immune system between humans and mice means that successful engraftment of PDXs requires the use of immunocompromised organisms, which prohibits studying the tumour microenvironment and human-specific therapy such as immunotherapy. To overcome this limitation, LifeTime proposes developing models with an intact/functional human immune system. Mice with induced pluripotent stem cell-derived xenografted human cells will be important pre-clinical models for studying interactions between these cell types and neurodegenerative pathologies associated with human proteins over longer periods of time. Alternatively, neurodegenerative disorders could be modelled under physiological conditions by engrafting human brain organoids into an adult immunodeficient mouse brain. These would contain key human cell types, functional neuronal networks, blood vessels and have the potential to model sensory inputs. Investment is required in advanced animal models for different disease applications that faithfully reflect human population dynamics, cellular heterogeneity, disease risk factors, age, lifestyle and chronic environmental exposures.

LifeTime also recommends developing models to study early events in disease such as for example mechanisms involved in tumour initiation, early tumour development or early biochemical alterations in Alzheimer's disease when clinical symptoms are not yet present and dementia can still be prevented. Precise modelling of disease progression requires replicating the natural course of events, for example for cancer, ranging from tumour initiation to metastatic outgrowth or the inflammatory response in Alzheimer's disease. Studying the early events in disease onset will require the further development of Genetically Engineered Models (GEMs). CRISPR-Cas systems have great potential incorporating human alleles, loci or pathways to study diseases. These will be particularly important for modelling human disease for personalised medicine to address the great variability between individuals, including response to treatment and clinical trajectories.

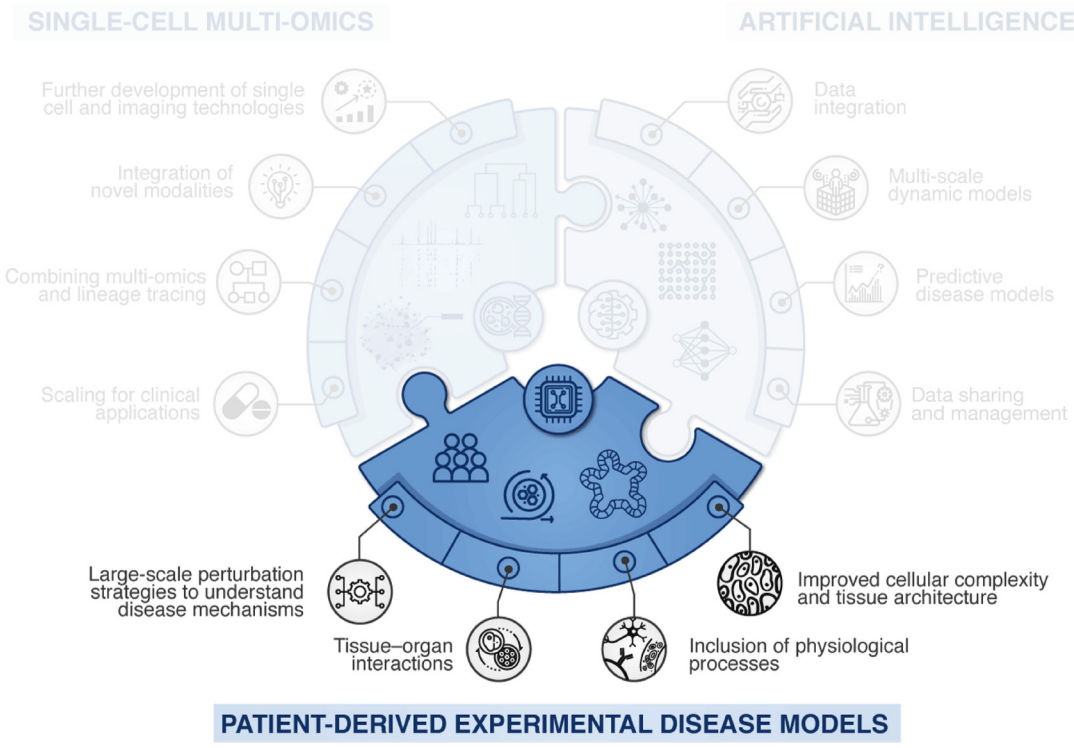
Adaptation of technologies will be required to study disease processes in these next-generation models such as expanding intravital imaging to all tissue sites and techniques that allow for quantitative dynamic information *in vivo*, such as lattice-light sheet microscopy, or single-cell sequencing-based methods (section 3.1.1). Development of imaging-based spatial multi-omics will capture spatial relationships and cellular heterogeneity and cell-cell interactions within intact tissues at single-cell resolution.

Large-scale perturbation tools to establish causality in model organisms

Understanding the causes of disease requires integration of new methods for large-scale perturbation experiments. These involve reducing the levels or activity of specific molecules in advanced disease models and monitoring the consequences on disease processes. Large-scale perturbation screens have been described in using CRISPR-Cas technologies with a single-cell transcriptome readout. LifeTime proposes to develop strategies for combinations of perturbations to reveal biological mechanisms that are controlled by parallel processes. Also tools are required for the targeted perturbation of the human epigenome and of protein signalling pathways to determine cause and effect among these highly interconnected layers. Large-

scale perturbation screens will profit from the development of high-throughput sequencing- and imaging-based multi-omics assays, especially those that can provide single-cell and spatial resolution (section 3.1.1), and testing predictive computational models (section 3.1.2). Moreover, CRISPR-based methods for lineage-tracing and molecular recording of cellular events and environmental exposures (including the storage of transcriptional events into the genome) will make it possible to capture each cell's developmental history, its present regulatory state, and its epigenetic potential for future plasticity.

Large-scale perturbation experiments, in combination with elements of synthetic biology, have the potential to accelerate the development of cell-based therapies, both in cancer immunotherapy and regenerative medicine, for example by enabling the rational, data-driven design of application-specific gene-regulatory networks. Clinical trials provide a rigorous, standardised path for the clinical validation of novel therapies. Unfortunately, a large percentage of clinical trials fail to reach their primary endpoint, often for unclear reasons. To improve insight into the molecular basis for success or failure of clinical trials (especially phase I/II) LifeTime recommends reproducing them in suitable experimental disease models including patient-derived organoids together with single-cell technologies and high-throughput perturbation analyses.



Patient-derived experimental disease model technology developments

Expected Scientific Impact

In the short term:

Currently available disease models will be used to study disease processes with established single-cell technologies, which will provide insights into cell types and pathways involved in disease

Generation of next-generation disease models containing increased complexity of cell types and containing tissue elements e.g. vasculature and immune system components or tissue engineering to study respiratory infections in the lung, that are key to addressing the disease challenges

Standardisation, automation and scaling of patient-derived disease models to create the required large cohort scale model biobanks and for perturbation screens

In the long-term:

Patient-derived disease models will be required at scale to be able to understand the patient-to-patient heterogeneity in mechanisms underlying disease

Disease models and iterative large-scale perturbations will be used to test and train machine learning derived predictive computational models of disease to increase their accuracy

Avatars of patient-derived material disease models will be used to test therapeutic strategies based on predictions from machine learning-based computational models

3.2 Solving Medical Challenges - Disease Roadmaps

3.2.1 Cancer

Medical Challenge

Cancer will kill almost 2 million citizens in Europe every year. It affects everyone regardless of age, gender or social status and represents a tremendous burden for patients, families and societies at large. Almost half of cancers are diagnosed after 65 years and, considering the ageing European population, a tidal wave of cancer cases will sweep across Europe within the next decades. If no further action is taken, the number of people newly diagnosed with cancer every year in Europe will increase from the current 4.2 million to 5.2 million by 2040. Preventing such an increase, as a minimum goal, would require a 0.75% annual reduction in risk and 1% reduction in risk of death. Such challenges call for coordinated European efforts, rational strategies, implementation of new structures and interdisciplinary research initiatives that are able to address major challenges in cancer diagnostics and treatment.

Cancer is a broad medical field with hundreds of individual tumour types, each with their own clinical presentations, molecular portraits, diagnostic and treatment strategies and prognostic outlook. Nevertheless, there are critical knowledge gaps that are common for all cancer entities, including early dissemination and therapy resistance. Metastatic dissemination of a subpopulation of cancer cells is a leading cause of death in almost all cancer types. Successful treatment of advanced and metastasised forms of cancer remain difficult, despite the development of targeted therapies and immunotherapies, due to the emergence of drug or therapy resistance.

Although tumours originate as single-cells, during disease progression some cells accumulate mutations and/or epigenetic changes to evolve into spatially and temporally distinct lineages and subclonal populations. These changes lead to intratumour cellular heterogeneity (the specific cell types and states involved) which contributes to early dissemination with subsequent metastasis and the development of therapy resistance. Until recently there has been a lack of technologies that are able to resolve the underlying mechanisms, the role of cellular heterogeneity as well as the complex interactions between cancer cells and the tumour microenvironment and metastatic niche. This level of complexity can be deconvolved using a combination of single-cell technologies providing cellular spatial resolution, machine learning and patient-derived disease models.

Driven by cancer genomics, technologies to resolve cellular heterogeneity and molecular risk stratification, the cancer field has been a ”first mover” in the precision medicine field. Cancer can, therefore, serve as a template for other diseases. Identifying cancer at the earliest stage possible will enable **interception of the disease** before it becomes too advanced or spreads via dissemination and metastasis. A European strategy for early cancer detection and interception of small tumours curable by surgery +/- radiotherapy as

well as/or tailored drugs will **increase survival of European citizens.** Alongside prevention strategies, this will require implementation of new translational research strategies focused on identification and targeting of the driving cancer cell population(s).

To address the **key challenge of cancer spread,** LifeTime's approach should address the following medical challenges to provide the greatest impact for cancer patients in the next decade:

Understanding the cell types and states - malignant cells and their microenvironment - involved in early stages of cancer dissemination

Understanding the reprogramming of cellular states during disease and their impact on therapy resistance

Cancer Roadmap

LifeTime proposes the following objectives to deliver an understanding of tumour evolution leading to dissemination and therapy resistance of cancer cells:

- **Define the cell types and states involved in early cancer dissemination using patient biomaterial (tissue and liquid biopsies) and pre-clinical models (patient-derived organoids and advanced animal models)**
- **Understand the molecular mechanisms and reprogramming events underlying cellular diversity and acquisition of properties essential for disease progression (including metastatic dissemination) and the development of therapy resistance with temporal and spatial resolution. Use this knowledge to generate new single-cell based molecular and cellular biomarkers for early disease detection and interception**
- **Integrate clinical phenotypes (responders versus non-responders as well as metastasis versus primary tumours) with molecular mechanisms, using machine learning to generate *in silico* models of cancer progression, metastasis and relapse to guide therapy selection and treatment**

Addressing these objectives will enable: i) the stratification of patients based on biomarkers derived from disease mechanisms in single-cells; ii) early disease detection and prediction of disease prognosis for an individual patient; iii) the systematic identification of key molecules/pathways for drug targeting and/or repurposing and development of effective, preventive or early interception treatments. Together, these insights will enable newly designed clinical trials by bridging clinical with molecular/cellular endpoints and therapeutic approaches on the basis of validated patient-specific disease mechanisms. For this roadmap, LifeTime's expected impact on breast cancer is used as a specific showcase example, but similar impact has been identified for colorectal and pediatric cancer, melanoma, glioblastoma and acute leukemia in the broad field of cancer.

Required Investments for Cancer

To address the above objectives LifeTime recommends the implementation of the following multidisciplinary research programmes:

Longitudinal mapping of cancer cell types and states relevant for early dissemination and/or treatment resistance using available single-cell multi-omics approaches

LifeTime proposes to immediately apply currently established single-cell multi-omics approaches (transcriptomic, genomic and epigenomic approaches on frozen or dissociated cells) and imaging- and sequencing-based spatial technologies (spatial transcriptomics, genomics and proteomics) to increase knowledge of the early events during tumour evolution. Longitudinal multi-omics analysis will be applied to tissue and liquid biopsies, patient-derived organoids and patient-derived xenograft models to uncover the relevant cell populations and their molecular characteristics. Data will be linked to ongoing efforts towards generating cancer atlases of primary tumour tissues, in particular the Human Tumor Atlas Network (HTAN), the International Cancer Genome Consortium - Accelerating Research in Genomic Oncology (ICGC-ARGO), The Cancer Genome Atlas (TCGA) and other international consortia.

These studies will identify disease-relevant cell types/states involved in early dissemination and therapy resistance and begin to reveal how these cells interact with each other and the tumour microenvironment or metastatic niche to cause relapse and/or metastasis. Disease-driving cell populations will be used for early detection of cancer and be immediately validated in clinical trials. Consequently, novel strategies for molecular targeted approaches and immunotherapies will be identified for disease interception as well as drug development or drug repurposing and the rational design of clinical trials.

Development and integration of novel multi-omics approaches with machine learning to reveal mechanisms involved in early cancer dissemination and therapy resistance

Understanding underlying molecular mechanisms requires the integration of information from different molecular networks as well as the functional importance of cellular location, cell-cell interactions and the tumour environment. LifeTime proposes prioritising the development and integration of several spatial-based multi-omics single-cell technologies (section 3.1.1) to provide a longitudinal view of the changes involved in early dissemination and resistance mechanisms of aggressive and deadly cancers. The important role of signalling pathways, metabolism and interactions of cancer cells with immune cells, stromal cells and vascular cells in the tumour microenvironment requires advances and integration of imaging and sequencing-based spatial transcriptomics/ genomics with proteomics and metabolomics. This will be done in combination with advanced methodologies for non-invasive live-cell imaging of disease models at subcellular resolution (light-sheet and multi-photon microscopy). Linking these molecular data with clinical data-sets and medical imaging modalities (i.e. radiomics) will provide insight into patient outcomes.

These technologies and new machine learning approaches (section 3.1.2) will go beyond gene expression programmes and substantially enrich mechanistic insights. They will begin to generate new predictive models for disease based on cell states and trajectories for disease interception. Discovery of the relevant biomarkers, cell types and their molecular characteristics for innovative targeted and immunotherapeutic treatment approaches will most likely provide a new level of treatment efficacy compared to current precision oncology approaches. These approaches would be validated immediately in new clinical trials and also initiate a new level of digital molecular pathology, based on spatial multi-omics, as the resulting information and models can be provided to clinicians to help guide decision making.

Development of advanced patient-derived cancer models

Obtaining longitudinal tumour tissue samples is challenging. Therefore, LifeTime proposes a particular focus on deriving advanced patient-derived models for major cancer entities, which can also be used to investigate underlying disease mechanisms (section 3.1.3), using newly developed single-cell and machine learning approaches (section 3.1.1 and 3.1.2). More representative patient-derived organoids or human-on-chip models require increased cellular complexity, to ensure that all key cell types are represented including tumour cells, stromal cells, immune cells and vascular cells with predictable spatial organisation and accurate cell-cell interactions. The need for vascularisation will require the use of microfluidic systems and long-term culture. Protocols that model different tumour microenvironments and metastatic niches (i.e. bone marrow, liver, brain or lung) must be developed to recapture cell-cell interactions in the respective environments. Longitudinal drug response measurements require the development of human-on-a-chip methodologies, in which organoids are grown within 3D-scaffold in proximity to niche cells. Disease models should be developed in conjunction with strategies to implement large-scale perturbation screens (e.g. CRISPR-Cas, small molecule) with single-cell multi-omics readouts and lineage tracing to understand causation in resistance and dissemination mechanisms (section 3.1.1 and 3.1.3).

These advanced models will not only serve as a driving force for understanding cancer metastasis and resistance mechanisms but are required to close the gap between basic research and translation of these findings into the clinics. They will also enable optimisation of therapies, where applicable, by testing tailored strategies and therapy combinations for individual patients.

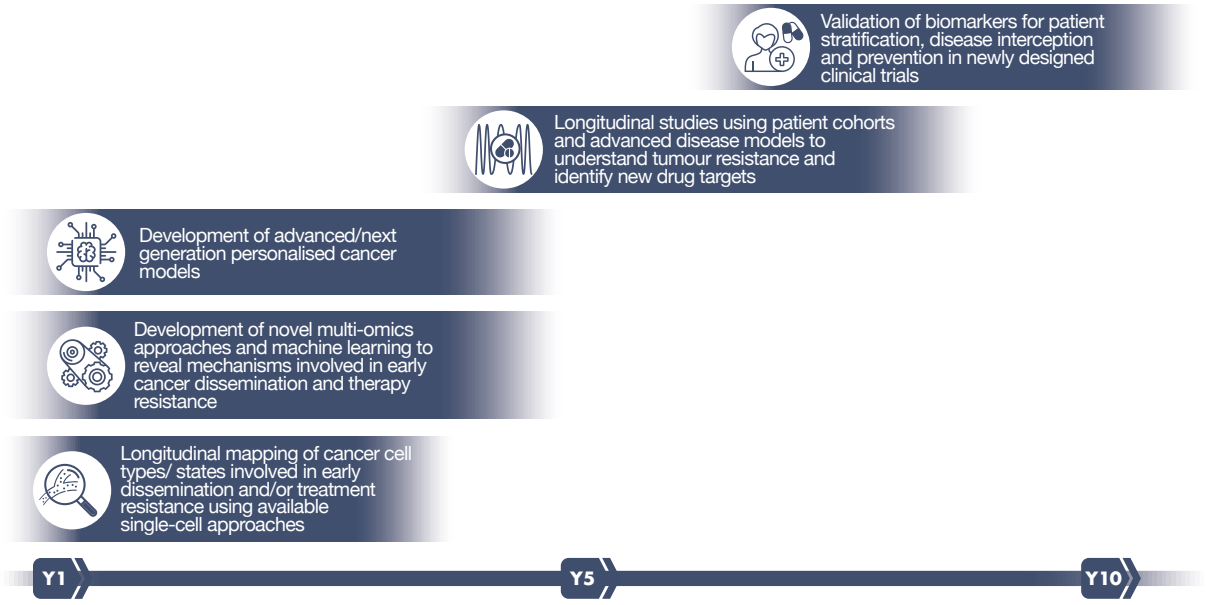
Longitudinal studies using patient cohorts and advanced disease models to understand tumour resistance and identify new drug targets

Understanding diversity in individual patient disease trajectories will require complementary studies using non-invasive liquid biopsies and newly developed patient-derived disease models. These will be generated from samples of patients obtained from clinical trial cohorts, that either respond or not to new treatments. These large-scale cohorts will be investigated using the newly developed, benchmarked and standardised imaging and sequencing-based multi-omics technologies (described above and section 3.1.1), and integrated with clinical and medical imaging data. In the process, the large high-quality datasets will be generated that are required for predictive computational models using machine learning algorithms (section 3.1.2) for early disease detection, interception and identification of novel mechanism-based drug-targets. Disease mechanisms and new potential drug targets will be studied using a combination of classical chemotherapy, molecular and immunotherapeutic approaches as well as novel large-scale CRISPR and small molecule screens in newly developed disease models.

Identified biomarkers will enable the stratification of patients based on the underlying disease mechanism. New modes of digital molecular pathology based on single-cell and machine learning approaches will create predictive models that will facilitate the selection of treatments for individuals. Where the disease course allows these treatments will be first tested in patient-derived model ‘avatars’ to ensure the required therapeutic response. Based on these models the portfolio of druggable targets will be systematically expanded targeting the driving cancer cell populations and subsequently lead to more efficient combination therapies.

Validation of biomarkers for patient stratification, disease interception and prevention in newly designed clinical trials

Promising new therapies would be tested in relevant and accurate models, allowing fast translation to the clinics for stratified groups of patients. Mechanism-based biomarkers for early dissemination and therapy resistance will be validated in patient cohorts and used to stratify patient groups for clinical trials for existing and repurposed drugs, and also for novel therapies that demonstrate efficacy in proof-of-concept studies in advanced pre-clinical models. These biomarkers will provide the basis for secondary disease prevention and would be used to molecularly screen at risk populations for early disease detection. We expect that such sensitive diagnostic and progression biomarkers as well as promising new combination therapies for major aggressive cancer entities would be identified and validated within the five to ten years.



Cancer Roadmap timeline

Expected Scientific Impact

SHORT-TERM (1-7 years):

Identification of cell types and states involved in early metastatic dissemination and treatment resistance

Single-cell spatial methods to understand infiltration/dissemination in the tumour microenvironment and metastatic niche

New pre-clinical cancer models for testing and predicting drug responses in patients

Identification of novel drug targets, biomarkers and molecular signatures suitable for risk stratification for early disease detection and interception

LONG-TERM (8-15 years):

Assessment of biomarkers, risk stratification schemes and targeted treatments in personalised clinical trials. Prospective validation of biomarkers within clinical trials

Introduction of biomarkers and molecular-guided therapeutic interventions into clinical practice

Single-cell based predicted prognosis of cancer entities will enable the development of preventive and interception therapeutic strategies and stratify tumours for appropriate aggressiveness of treatment

Application of machine learning will create new tools for early diagnosis and to define prognostic factors

Technologies to enable physicians to delay/ prevent metastatic spreading and diagnose cancer much earlier and more effectively as well as predict the future course of the disease and select the optimal treatment for an individual cancer patient

Medical Challenge

- Understanding the cell types and states involved in early cancer dissemination and therapy resistance

Outputs

- Increased understanding of the cell types and states responsible for cancer dissemination and therapy resistance and how they cause relapse and/or metastasis
- New technologies for understanding the cellular basis and evolution of tumours in the context of the microenvironment
- Next generation patient derived cancer models for longitudinal cohort studies and testing of new or predicted personalised therapies
- Novel multi-dimensional biomarkers for patient stratification for therapeutic strategies
- Identification of new targets for drug development or repurposing based on cell types and states involved in metastasis and drug resistance
- Development of machine learning enabled systems for new molecular pathology of tumours that can be used for secondary prevention applications

Impact

- Increased likelihood of curing cancer at an early stage, increased patient survival
- Less burden on health systems and less indvidual suffering through early detection of cancer and emergence of drug resistance based on informative multi-dimensional biomarkers to enable disease interception, before metastasis or resistance occurs
- Improved patient outcomes through precision therapeutic strategies to intercept disease, based on predictive and personalised models of disease
- Faster and better validation of therapies based on new technologies and preclinical models as well as newly designed clinical trials using clinical and cellular/molecular endpoints

Example – Impact on Breast Cancer

- Identification of disease-relevant cell types/states involved in early dissemination and therapy resistance, how these cells interact with each other and the tumour microenvironment to cause local relapse and/or metastasis. These disease-driving cell populations would be used to identify relevant molecular targeted approaches and immunotherapies for disease interception.

- Increased knowledge of disease mechanisms and new computational predictive models for disease course based on cell states and trajectories. Novel single-cell multi-omics and imaging technologies will form a basis for a new level of digital molecular pathology.

- Advanced patient-derived disease models will serve as a driving force for understanding cancer metastasis and resistance mechanisms as well as accelerate translation.

- Stratification of patients based on multi-dimensional biomarkers and the underlying disease mechanisms. Together with predictive computational models these biomarkers will facilitate the selection of treatments for individuals. Where disease course allows these treatments will be first tested in patient-derived disease model ‘avatars’ to ensure required therapeutic response to provide optimised treatments. Based on these models the portfolio of druggable targets will be systematically expanded targeting the driving cancer cell populations and subsequently to more efficient combination therapies.

- Mechanism-based biomarkers for early dissemination and therapy resistance will be validated in patient cohorts and used to stratify patient groups for clinical trials for existing and repurposed drugs, and also for novel therapies that demonstrate efficacy in proof-of-concept studies in advanced pre-clinical models. Biomarkers will provide the basis for secondary disease prevention and would be used to molecularly screen at risk populations for early disease detection. It is expected that such sensitive diagnostic and progression biomarkers as well as promising new combination therapies for major aggressive cancer entities could be identified and validated within the next 5 to 10 years.

3.2.2 Neurological and Neuropsychiatric Diseases

Medical Challenge

The majority of chronic neurological diseases are currently incurable. Collectively they are the leading global cause of Disability Adjusted Life Years and the second-leading global cause of deaths[1]. Brain disorders affect the quality of life of **over 160 million European citizens**, constituting an ever-growing social and economic burden that threatens the stability of the EU's healthcare systems.

In cases where treatments are available, they mostly attempt to alleviate the symptoms of the disease but are not able to cure or even slow its progression. This situation coupled with an ageing European population, offers poor future prospects for patients and their families. After more than a decade of failed clinical trials, drug discovery is increasingly limited to academia and SMEs, with the pharmaceutical industry progressively withdrawing from translational brain research programmes due to failing returns on large early investments. As a result, brain diseases are currently among the most understudied clinical conditions, with new strategies and approaches urgently required to tackle this unmet challenge. A key factor explaining the poor translation of research outcomes to the clinic is the complexity and heterogeneity in disease causes, mechanisms and presentation, which could not be adequately addressed using previous approaches. Consequently, clinical trials have so far included patients with both diverse symptoms and underlying disease biology, and who are often at late stages of disease, which complicates developing effective disease-modifying drugs.

A major barrier to implementing personalised medicine is the large knowledge gap in our understanding of the human brain structure and function at the cellular level and how this relates to clinical phenotypes captured through functional brain imaging studies. Due to the cellular complexity of the brain, revealing these mechanisms requires single-cell approaches, including spatial and imaging. These would identify the specific cell subpopulations and corresponding molecular pathways associated with specific functions or disease. Being able to map specific cell (sub)types to specific brain locations is key as specific brain regions are selectively or predominantly affected in several brain diseases. Another major barrier is the difficulty of studying disease biology, especially early events, in humans due to a lack of routine access to tissue from patients to study disease onset and progression. Longitudinal sampling of disease states will require highly innovative *in vitro* models that fully reproduce human biology, including ageing and disease. These include developing human induced pluripotent stem cells (hiPSCs), advanced animal models containing hiPSC-derived specific human cell types that can be exposed to human pathology, for example amyloid plaques composed of human proteins. Alternatively, brain organoids generated from hiPSCs recapitulate *in vitro* salient features of the *in vivo* developing brain, especially the sequential emergence of progenitor and neuronal subpopulations that capture aspects of spatial organisation of several brain regions.

1 the Global Burden of Disease Study 2015

Understanding the cellular and molecular mechanisms that underlie diseases in individual patients will reveal the heterogeneity and diversity of causes of neurological disorders. This will provide new multi-dimensional biomarkers that can be used for **earlier disease detection**, and most importantly, used to **stratify patients according to disease mechanism** to re-test previous, as well as new therapeutic concepts for **disease-modifying drugs and therapies.** Such biological information is also needed to design tailored treatment strategies for specific patient subgroups and disease stages.

Therefore, LifeTime has identified:

The stratification of patients according to the heterogeneity of cell states and trajectories in neurodegenerative and neurodevelopmental diseases

as the key transforming challenge to provide the greatest impact for patients in the next decade.

Neurological and Neuropsychiatric Diseases Roadmap

LifeTime proposes the following objectives to deliver an understanding of the diversity of neurological and neuropsychiatric disease mechanisms in patients:

- **Define the cellular states in disease-relevant contexts and analyse how they are altered due to an individual's complex genetic constitution using combinations of patients' primary samples (biopsies, cerebrospinal fluid and peripheral blood), brain organoids and advanced animal models**
- **Understand the cellular trajectories and mechanistic pathways underlying neurodegenerative and neurodevelopmental diseases with temporal and spatial resolution**
- **Integrate clinical phenotypes with molecular mechanisms, using machine learning to generate *in silico* models of Alzheimer's disease, Parkinson's disease, fronto-temporal dementia, epilepsy and autism spectrum disorders**

Addressing these objectives will enable: i) a rational stratification of patients on the basis of biomarkers empirically linked to pathogenic cascades; ii) the systematic identification of key molecules/pathways for drug targeting and/or repurposing; iii) early disease detection and better prediction of disease prognosis. Together, these insights are poised to transform the design of clinical trials by bridging clinical with molecular/cellular endpoints and therapeutic approaches on the basis of validated patient-specific disease mechanisms. For this roadmap, LifeTime's expected impact on autism spectrum disorders and Alzheimer's disease is used as showcase examples for the broad field of neurodevelopmental and neurodegenerative disorders.

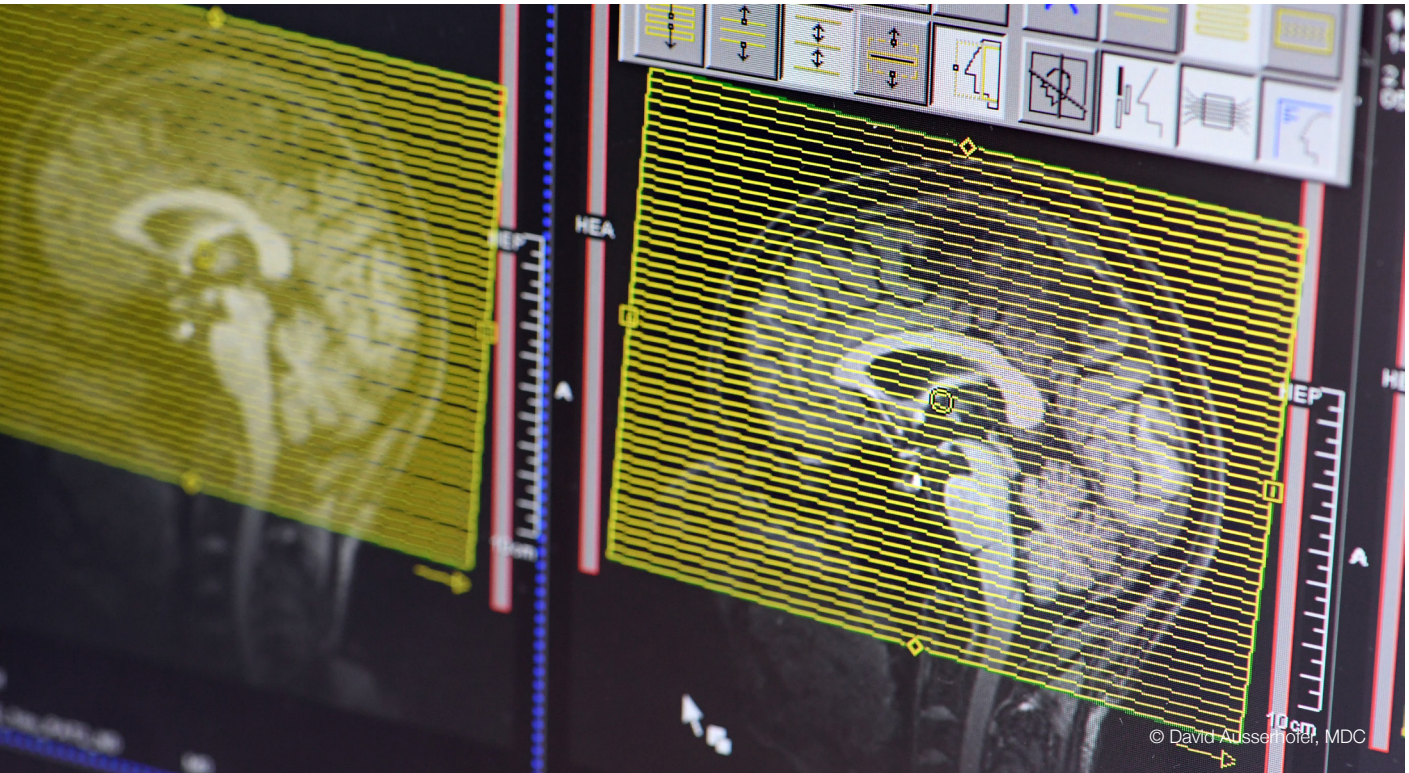
Required Investments for Neurological and Neuropsychiatric Diseases

To address the above objectives LifeTime recommends the implementation of the following multidisciplinary research programmes:

Longitudinal mapping of disease cell states and types in established patient-derived models using available single-cell multi-omics approaches

To understand the early pathophysiological events that occur during neurodevelopmental and neurodegenerative disorders, LifeTime proposes to immediately apply currently established single-cell multi-omics approaches (transcriptomic, chromatin accessibility, epigenetic approaches on dissociated cells) and spatial technologies (spatial transcriptomics and proteomics). For neurodevelopmental disorders longitudinal multi-omics analysis will be pursued in patient-derived brain organoids to uncover the molecular pathways and derangements in cellular trajectories. As organoid models for neurodegenerative diseases still require significant technological development (see below), initial longitudinal analyses will use disease models in which reprogrammed cells from patients with a high risk of developing a specific disease (based on polygenic risk score) are grafted into the mouse brain to uncover major disease-relevant phenotypes. This would be linked to ongoing efforts towards generating brain atlases, in particular the Human Cell Atlas, the Human Brain Project, and the H2020 project BRAINTIME (molecular atlas of the brain across the human lifespan).

These studies will provide insight into disease mechanisms at the single-cell level. For Alzheimer's disease they will reveal the specific cell types (neurons, glia, vascular cells) involved, the early responses to amyloid pathology and how these cells interact with each other to ultimately cause neurodegeneration. Most importantly, they would begin to reveal where and when the many genes associated with risk of Alzheimer's disease are expressed and exert their function. Analysing the cellular reaction in different brain areas will provide insights into regional vulnerability. Comparing cell states of patients with various genetic



© David Ausserhofer, MDC

risks of Alzheimer's disease and especially healthy centenarians will provide information on the cellular mechanisms of resilience. For autism spectrum disorders, longitudinal organoid modelling will reveal the changes in developmental trajectories, the timing of their deviation from normal development and the effect on neuronal-network architectures. These are necessary for the meaningful stratification of patients and the rational design of clinical trials and drug repurposing pipelines.

Development of novel multi-omics approaches to study disease mechanisms in brain models

Based on the functional importance of cellular location, cell-cell interactions and the local environment in the brain, LifeTime proposes prioritising the development of several spatial-based multi-omics single-cell technologies (section 3.1.1) together with machine learning (section 3.1.2). These will provide an integrated view of the molecular layers involved in brain diseases using patient-derived models required for cohort scale analysis below. Increasing cellular resolution and sensitivity as well as the throughput of spatial transcriptomic approaches is required to reveal the roles of different cell types including both neurons and glial cells, the neurovasculature and, in the case of neurodegeneration, protein aggregates. Unravelling the important role of signalling events, proteostasis, cellular inflammation and lipid metabolism requires advances and integration of spatial proteomics and metabolomics as well as advanced methodologies for non-invasive imaging of disease models, including live cell imaging at subcellular resolution (section 3.1.1). Integration of data-sets and other medical imaging modalities will require novel machine learning based approaches for extraction of meaningful features that can be used as biomarkers for disease onset and progression (section 3.1.2). Since pathology will also result in changes in cellular activity it will also be essential to integrate functional assays, such as electrophysiology, within these spatial approaches.

These technologies will substantially enrich the mechanistic insights obtained from current established technologies, since changes in the proteome and especially the lipidome cannot be reliably predicted based on gene expression. For both Alzheimer's disease and autism spectrum disorders this will reveal the cell types responsible for pathophysiological changes. In the process revealing the alterations in proteins, protein modifications and turnover, and metabolites including lipids both in these cells and in the surrounding tissue. The outcome will be a detailed view of disease progression at high spatial and temporal resolution.

Development of advanced patient-derived disease models for neurological disorders

Obtaining longitudinal patient brain samples is particularly challenging for neurological disorders. Therefore, LifeTime proposes a particular focus on deriving advanced patient-derived models for both neurodegenerative and neurodevelopmental diseases (section 3.1.3) to dissect the underlying molecular and cellular mechanisms and identify and validate new drug targets. More representative patient-derived organoids require increased cellular complexity, to ensure that all key cell types are represented, including neurons and glia, with predictable spatial organisation and accurate cell-cell interactions. In addition, there is a need for vascularisation, which will require the use of microfluidic systems, and for long-term culture. Since symptoms in patients originate in distinct functional neural circuits, new research, analysis and

therapy concepts depend on the reconstruction of brain circuitry on a histological and functional level. While assembly of circuits in organoid model systems is currently in its infancy, the fusion of organoids resembling different brain areas is an example of a promising approach to establish more sophisticated brain circuit organoid systems with the correct neuronal connections. Analysing these circuits requires the parallel development of optogenetic tools in organoids, fluorescent readouts for neuronal activities and - most importantly- readouts for neural firing patterns compatible with single-cell multi-omics.

Longitudinal activity measurements require the development of brain-on-a-chip methodologies in which organoids are grown within a 3D-scaffold in proximity to inbuilt electrodes that enable measurement of real time electrical activity. Disease models should be developed in conjunction with strategies to implement large-scale perturbation screens (e.g. CRISPR-Cas, small molecule) with multi-omics readouts and lineage tracing to understand causation in disease mechanisms (section 3.1.3). Profiling of disease trajectories (see below) at a cohort-wide scale will require that organoid production be scaled, standardised, automated and benchmarked to primary tissues as described in detail in section 3.1.3.

These advanced models will not only serve as a driving force for understanding disease mechanisms but they are essential for closing the gap between basic research and translation of these findings into the clinics. For example for Alzheimer's disease current models based on human cells are in their infancy and have the potential to revolutionise how the disease is understood, but also explore the potential of new therapeutic strategies in a human *in vitro* pre-clinical trial platform. For autism spectrum disorders, complex hiPSC-based models have already been developed and used to study monogenic autism spectrum disorder syndromes, and non syndromic autism spectrum disorders would also benefit from these advanced models.

Longitudinal studies of disease onset and progression using advanced patient-derived models to understand disease mechanisms and identify new drug targets

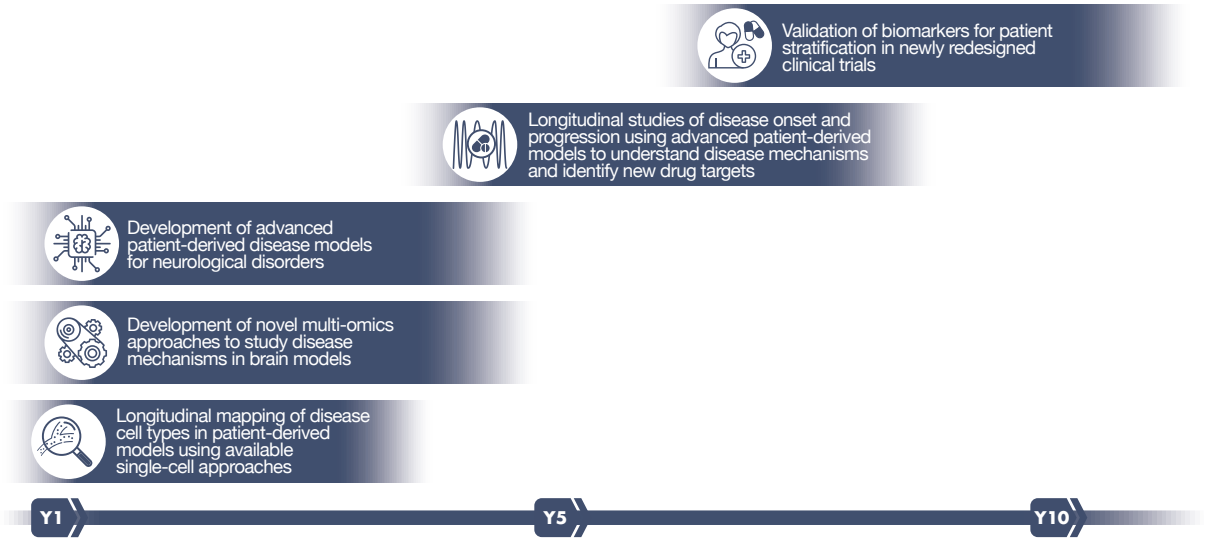
To understand both normal ageing and disease trajectories of individuals and to identify novel meaningful biomarkers for patient stratification will require longitudinal studies of advanced models based on hiPS cells derived from established patient cohorts. For example, the EMIFAD MBD (European Medical Information Framework for Alzheimer's disease Multimodal Biomarker Discovery) study or national cohorts such as the Dementia Platform UK for Alzheimer's disease; and the EU-AIMS (European Autism Interventions - A Multicentre Study for Developing New Medications) study or the Italian autism network (ITAN) Rome cohort and Milan Cohort from the San Paolo Research Hospital for autism spectrum disorders. Newly developed patient-derived disease models described above will be generated from selected patients from cohorts. This will be based on extremes in the polygenic risk scores to identify extreme high-risk and extreme protected individuals (for instance healthy centenarians) with different pathophysiological mechanisms e.g. inflammation or defects in lipid metabolism. These large-scale model cohorts will be investigated using the newly developed, benchmarked and standardised single-cell multi-omics and imaging technologies (described above and section 3.1.1), which will be integrated with clinical and medical imaging data. These approaches will generate the large high-quality datasets required for generating predictive computational models and identify novel mechanism-based drug targets. Machine learning-based models will be developed to generate *in silico* predictions of the most promising therapeutic targets, using mechanistic knowledge.

They will predict which type of intervention will be the most efficacious for a particular disease phase (section 3.2.2). Disease mechanisms and new potential drug targets will be studied using a combination of classical molecular and genetic approaches as well as newly developed large-scale CRISPR and small molecule screens in novel disease models (section 3.1.3).

For both Alzheimer's disease and autism spectrum disorders this will identify biomarkers that enable the stratification of patients based on the underlying disease mechanism and lead to a considerable expansion of the portfolio of druggable targets. This will also allow antisense and genetic therapies to be rapidly tested in clinical trials over the coming years, and encourage the pharmaceutical industry to re-invest much more than today in R&D for brain diseases.

Validation of biomarkers for patient stratification in newly redesigned clinical trials

Promising therapies that have previously failed to prove efficacy in clinical trials will be re-tested in relevant and accurate models, allowing fast translation to the clinics for targeted groups of patients. Mechanism-based biomarkers will be validated in patient cohorts and used to stratify populations for clinical trials for existing and repurposed drugs, and also for therapies that demonstrate efficacy in proof-of-concept studies in advanced pre-clinical models. It is expected that such sensitive diagnostic and progression biomarkers would be identified and validated within the next five to ten years, putting an end to the "one-size-fits-all" approach to treating brain diseases.



Neurological and Neuropsychiatric Diseases Roadmap timeline

Expected Scientific Impact

SHORT-TERM (1-7 years):

For Alzheimer's disease:
Stratification of patients according to the main cellular mechanisms driving the disease (biological pathways, neuroinflammation, vascular defects, lipid metabolism, etc.)

New design of clinical trials to re-evaluate drugs that were tested without patient stratification (eg amyloid directed therapies)

Broaden the drug target portfolio beyond amyloid and tau biology, initially to inflammatory and vascular targets and perform pre-clinical validation in relevant advanced patient-derived models

For neurodevelopmental disorders:
Advanced patient-derived models reflecting specific disease aspects to identify pathways that can be targeted, whenever possible by repurposed drugs, to drive cell states into a healthy trajectory

Pre-clinical testing of new drugs or therapies in relevant and accurate models to drive faster translation to the clinics

LONG-TERM (8-15 years)

These approaches will revolutionise the field of brain diseases research and change the way drugs and therapeutic strategies are developed and diseases diagnosed. Focusing on the initial cellular phases of these disorders will provide opportunities to intercept them at an earlier stage, to halt or even prevent disease manifestations

LifeTime will shift the focus to the early cellular biology of these diseases in the complex context of the brain. This will provide a novel arsenal of therapies addressing pathological cell states in brain disease and replace our traditional and obsolete battery of diagnostic tools and symptomatic treatments

Patients will be treated in a personalised manner, using combinations of therapies that target the right cellular responses at the right time

Medical Challenge

- The stratification of patients according to the heterogeneity of cell states and trajectories

Outputs

- Increased understanding of the molecular mechanisms and cell types and states causing disease
- New technologies for understanding the molecular mechanisms, cell types involved, their activity and spatial interactions in the brain. Organoid systems recapitulating main human brain circuits
- Tools for circuit-level analysis of neurological disorders in organoids and neural firing patterns on the single-cell level compatible with other single-cell level analysis
- Next generation patient-derived organoids to create ‘cohorts’ to study early events in disease and progression and used as a pre-clinical trial platform to test new drugs and therapies
- Identification of biomarkers for patient stratification in clinical trials based on underlying molecular mechanisms and pathways involved
- Systematic identification of new drug targets and therapeutic concepts for disease modifying drugs and a much broader drug target portfolio

Impact

- Rational development of disease modifying drugs and therapies that can slow or cure of diseases to improve patients’ quality of life
- Faster translation of therapies to clinics using advanced pre-clinical models and applying improved clinical trial design according to patient subgroup’s underlying mechanisms
- Improved patient outcomes based on earlier detection of neurological diseases; more precise therapy selection using machine learning enabled systems for predictive computational models of brain diseases

Example - Impact on Alzheimer's Disease

- New insight into disease mechanisms at the single-cell level, including the specific cell types (neurons, glia, vascular cells) involved, the early responses to amyloid pathology and how these cells interact with each other to ultimately cause neurodegeneration. Reveal where and when the many genes associated with risk of Alzheimer's disease are expressed and exert their function and provide insights into brain region vulnerability.

- Enriched mechanistic insights from novel single-cell multi-omics and imaging technologies that incorporate changes in the proteome and the lipidome in addition to gene expression and epigenome. Reveal which cell types are responsible for pathophysiological changes and the alterations in proteins, protein modifications and turnover, and metabolites including lipids in these cells and in the surrounding tissue. Provide a detailed view of disease progression at high spatial and temporal resolution.

- Advanced patient-derived models as a driving force for understanding disease mechanisms and closing the gap between basic research and translation of these findings into the clinics. Revolutionise how the disease is understood and also explore the potential of new therapeutic strategies in a human *in vitro* pre-clinical trial platform.

- Stratification of patients according to single-cell based biomarkers and understanding the underlying disease mechanisms will lead to a considerable expansion of the portfolio of druggable targets. This will encourage the pharmaceutical industry to re-invest much more than today in R&D for brain diseases.

- Re-testing of promising therapies that have previously failed to prove efficacy in clinical trials in relevant and accurate models, allowing fast translation to the clinics for targeted groups of patients. Mechanism-based biomarkers will be validated in patient cohorts and used to stratify populations in clinical trials for existing and repurposed drugs, and also for therapies that demonstrate efficacy in proof-of-concept studies in advanced pre-clinical models.

- Sensitive diagnostic and progression biomarkers could be identified and validated within the next five to ten years, putting an end to the "one-size-fits-all" approach to treating brain diseases.


MEDICAL REPORT
02-08-38 : MALE
DOCTOR
: 02 :43 080
:586 :89 403
:253 :684 :01
:99 :RP_809

3.2.3 Infectious Diseases

Medical Challenge

Infectious diseases remain one of the most challenging areas in medicine and responsible for one in four deaths worldwide. Antimicrobial resistance is increasing, and epidemics and pathogens such as SARS, Influenza H1N1, multi-resistant bacterial pathogens, novel fungi such as Candida auris and very recently SARS-CoV-2 continually emerge. Difficulties in treating persistent chronic infections, increasing resistance and pandemic outbreaks such as COVID-19, call for the rapid development of novel therapeutic strategies.

Despite the need for new treatments, only a handful of new antibiotics have been launched over the past 40 years and many companies have abandoned their R&D programmes in infectious diseases. In addition, antimicrobial therapy is unlikely to provide full protection for a weakened patient and novel, personalised, intervention strategies are required to boost an individual's immune response. Importantly, interception of infections using recent medical breakthroughs such as immunotherapy, which has transformed the therapeutic landscape of autoimmune diseases and cancer, has not yet been successfully implemented for infectious diseases. Previous immunotherapeutic approaches, for example in sepsis, have been unsuccessful due to the "one-size-fits-all" approach, which does not consider the immune status of patients. This has resulted in the failure of important therapeutic concepts in infection biology such as hyperinflammation or immunoparalysis. An important factor explaining the difficulty in translating new molecular insights into novel therapies is the complexity and heterogeneity in the causes, mechanisms and presentation of infectious diseases.

Implementing novel therapeutic approaches such as immunotherapy for infectious diseases requires a more comprehensive understanding of the cellular basis of disease heterogeneity, as well as the underlying mechanisms and dynamics. For example, in pneumonia or sepsis there is a complex cellular immune response that continuously changes during the course of the infection. Selecting the most effective available therapy at a particular time or developing new immunotherapies will require knowledge of the phase of the immune response, e.g. hyperinflammation or immunosuppression, and take into account an individual's preconditions, genetics and environmental factors. Understanding and targeting host-pathogen interactions requires longitudinal single-cell level analyses of the cell types, sub-types, functional cell states and cellular memories involved. Disease mechanisms will require patient-derived disease models of anatomically complex tissues such as the lungs to resolve spatio-temporal interactions within diverse microenvironments that can be exploited by pathogens. Infectious diseases offer a unique opportunity to combine single-cell multi-omics and imaging approaches with both cohort-based studies and human controlled experimental infections to decipher pathophysiology before and after challenge (e.g. with a vaccine) as well as during therapy.

Reducing the burden of infectious diseases requires a coordinated, international, multi-dimensional approach that encompasses increased research, development and innovation in areas such as improved diagnostics, patient stratification, new and precise therapeutic strategies and more effective vaccines. This will provide new multi-dimensional biomarkers that can be used for **earlier disease detection**, and selection of **tailored immunotherapy** and combination strategies to specific patient populations to identify optimal strategies for treating patient subgroups and disease stages.

In order to achieve these goals, LifeTime has identified the following key transformational challenge to be studied in the next decade that could have the greatest impact for patients:

Understanding the cellular response to infections and develop novel precision immune-based therapeutic strategies to combat infectious diseases

Infectious Diseases Roadmap

LifeTime proposes the following objectives to deliver a transformative improvement in the monitoring and treatment of infectious diseases:

- **Identify single-cell multi-omics-based biomarkers for cellular immunopathology and patient stratification**
- **Develop new immunotherapy concepts for infections based on single-cell resolution of patient disease trajectories**
- **Develop novel approaches to improve vaccination efficacy in at risk patients (e.g. elderly, immunosuppressed), including using novel controlled models of human infection**

Addressing these challenges will: i) provide a single-cell resolution spatio-temporal map of infections and the trajectory of the immune response from the initiation of the infection towards the recovery; ii) identify biomarkers to stratify patients for immunotherapy; iii) increase our understanding of disease mechanisms based on single-cell multi-omics analysis, machine learning and patient-derived disease models for earlier detection and therapy selection and monitoring immunotherapy response; iv) develop novel approaches for improving vaccination by modulating host immune pathways and cells. For this roadmap, LifeTime's expected impact on COVID-19 and sepsis are used as showcase examples for the broad field of infectious diseases.

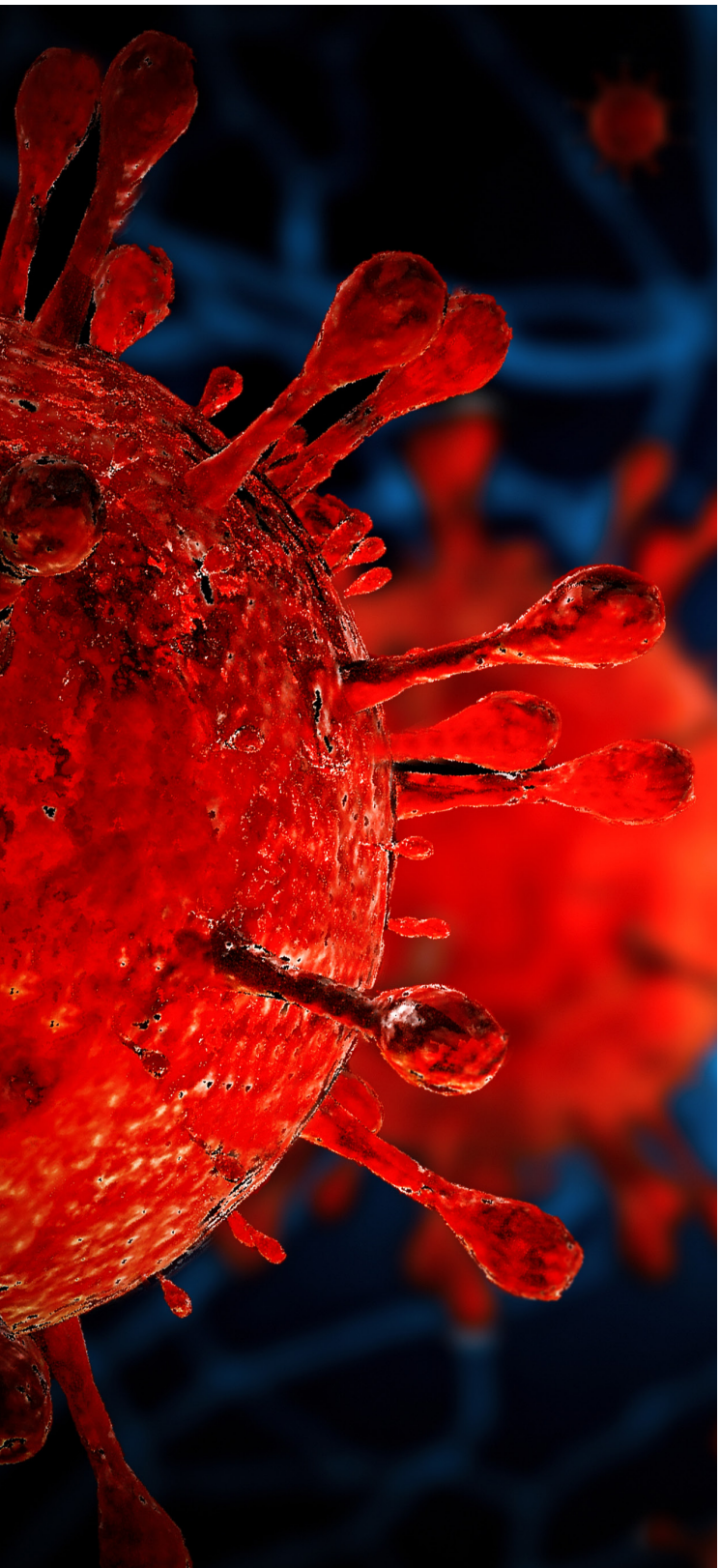
Required Investments for Infectious Diseases

To address the above objectives LifeTime recommends the implementation of the following multidisciplinary research programmes:

Longitudinal single-cell multi-omics analysis of patients' immune response to infection

To understand the immune response to infection and vaccination, LifeTime proposes to immediately apply established and scaled single-cell approaches to body fluids and tissue biopsies from (1) acute lower respiratory tract infections or (2) sepsis patient cohorts or (3) individuals undergoing controlled infections or vaccination programmes. Immediately deciphering the heterogeneity in a patient's immune response

to infection will be initiated using current single-cell technologies such as transcriptomics, epigenomics and proteomics. These will be performed on accessible patient material (peripheral blood mononuclear cells, plasma, nasal scrapings, bronchoalveolar lavage for respiratory diseases and whole blood for sepsis) from patient cohorts with standardised clinical data longitudinally across key timepoints. Immune responses during respiratory diseases will also be characterised at the site of infection using minimally invasive tissue sampling and spatial transcriptomic approaches depending on the pathogen and infection site (i.e. spatial transcriptomics on lungs, bronchial linings). This will be complemented with studies using established patient-derived airway and lung organoid models to model infections and disease progression and combined with medical imaging (computerised tomography scans) to facilitate early diagnostics.



For COVID-19 this will reveal immune dysregulated pathways in patients with disparate symptoms and enable screening of available immunomodulatory drugs in established disease models. These will identify biomarkers of disease severity. These will be used to select the immunotherapies that have the greatest potential to restore body/lung homeostasis for specific patients and facilitate the initiation of rapid clinical trials with particular drugs.

For sepsis this longitudinal analysis will reveal the dynamics of sepsis progression, including the timing of hyperinflammation and immunosuppression phases. It will identify cell types or states that are associated with early events that could be used as biomarkers for early detection and stratify patients based on the type of immune dysregulation and organ dysfunction and stage of sepsis. This will indicate the optimal immunotherapies and timing of administration to ensure they have the greatest potential to restore immune homeostasis in combination with antimicrobial therapy.

For vaccinations this will provide insight of immunological mechanisms and molecular pathways involved at the single-cell level. It will reveal individual heterogeneity in the response to vaccination and how these cells interact with each other to ultimately determine an efficient or defective immune response. Most importantly, these technologies and analyses will reveal which pathways should be modulated for improved vaccine design, and how the individual variability should be approached for maximum efficacy.

Developing advanced patient-derived models that recapitulate pathophysiology and reveal infection/disease mechanisms

Obtaining tissue samples for longitudinal analysis of disease progression is particularly challenging for infections that target particular organs. While iPSC lung organoids and air-liquid interface are key to deciphering the immediate response to respiratory virus infection, they need to be further developed to go beyond the epithelial layer and to integrate components of the basal, stromal and endothelial cells to more accurately reflect complex lung tissue (section 3.1.3). They also require the incorporation of resident and recruited immune cells that play a role during the infection. Advanced models should also mimic age-dependent structures since lower respiratory tract infections depend on age. Establishing causality in dysregulated pathways and host manipulation strategies by the virus and creation of predictive computational models requires integration of organoids, single-cell multi-omics and genome-scale gene perturbation screens e.g. CRISPR based technologies to resolve the complex intertwined host-pathogen interaction (section 3.1.2 and 3.1.3). Creating organoids from population cohorts will provide an opportunity to investigate individual susceptibility and pathology. The respiratory tract is highly structured in successive compartments such as the conductive airway and the respiratory airway. Studying lung-specific processes requires development of novel tissue engineering approaches such as lung-on-a-chip that approximate the whole organ (section 3.1.3). Microfluidics-based approaches will be used to model complex structures that also contain multiple layers of stromal cells and immune cells that play a major role in mounting an efficient immune response and tissue repair.

For COVID-19 these models will provide a means to study temporal aspects of infections and patient-to-patient variability and sensitivity to infections. When coupled with CRISPR-Cas screens they will play a key role in understanding the molecular mechanisms involved in host pathogen interactions and the immune response.

Developing novel multi-omics approaches to study infection disease mechanisms longitudinally

Timely intervention of infectious diseases requires development and optimisation of antimicrobial and immunotherapeutic strategies. LifeTime prioritises understanding the mechanisms that contribute to cellular heterogeneity in the pathogen and the immune response by integrating longitudinal changes in transcriptomes, intracellular signalling pathways, including phosphoproteomes, metabolomes, lipidomes, and cell-cell interactions, using machine learning (section 3.1.1 and 3.1.2). This requires investment in the development of untargeted single-cell proteomics and metabolomic approaches. These are particularly key to understand the states of immune cells, including the regulation of inflammatory responses and immunoparalysis in sepsis and how pathogens manipulate host signalling pathways and metabolism during infection. Interactions between host and pathogen are also influenced by cell-cell interactions and the lung tissue microenvironment. Interception of the early stages of infections requires the identification of cell types/states that are targeted by pathogens, the molecules involved and the temporal and cellular response of the immune system. These require systematic studies including longitudinal approaches using patient-derived samples or models. These molecular data will be integrated together with clinical data collected for lower respiratory tract infections such as computerised tomography scans and X-rays to improve diagnosis

(section 3.1.2). This approach will also be tested in the context of treatment (companion diagnostics) which could potentially accelerate the development of new therapies.

For COVID-19 and sepsis, integration of these approaches and creation of computational models will provide mechanistic insight into the regulation of the various molecular networks in cells. In the process revealing the roles of cell types and subtypes involved in viral infection and host defense mechanisms, especially those that lead to acute respiratory distress syndrome (ARDS) and sepsis.

Apply novel single-cell multi-omics approaches to study infections longitudinally using patient cohorts and patient-derived disease models

To understand the role of heterogeneity of infections and the corresponding immune response, including sepsis, as well as response to treatment within individual patients and across cohorts LifeTime recommends applying the single-cell multi-omics and machine learning approaches as well as patient-derived disease models described above to systematically analyse patient disease trajectories. Next generation multi-omics technologies will be used to analyse samples from multiple patient cohorts and biobanks, including with comorbidities (e.g. cardiovascular or renal). This will include body fluid samples from patients with respiratory infections and sepsis caused by a variety of pathogens, to reveal general and pathogen specific signatures, and also include samples obtained from controlled human infections and vaccination programmes. The later will provide longitudinal data, prior to and during infection through to therapy response and clearance as well as subsequent timepoints. Machine learning approaches will be used to integrate these data with clinical and imaging data, as well as single-cell multi-omics analysis in patient-derived organoid and tissue models and include perturbation studies to create predictive causal mechanistic models of infections for individual patients (section 3.1.2).

For COVID-19 and sepsis, integration of novel multi-omics technologies with machine learning approaches at a cohort/population scale will enable the creation of predictive models for disease progression based on multi-dimensional datasets. Elucidation of disease mechanisms will systematically identify therapies, including combinations, as well as new cellular drug/antiviral targets to be validated in patient-derived models.

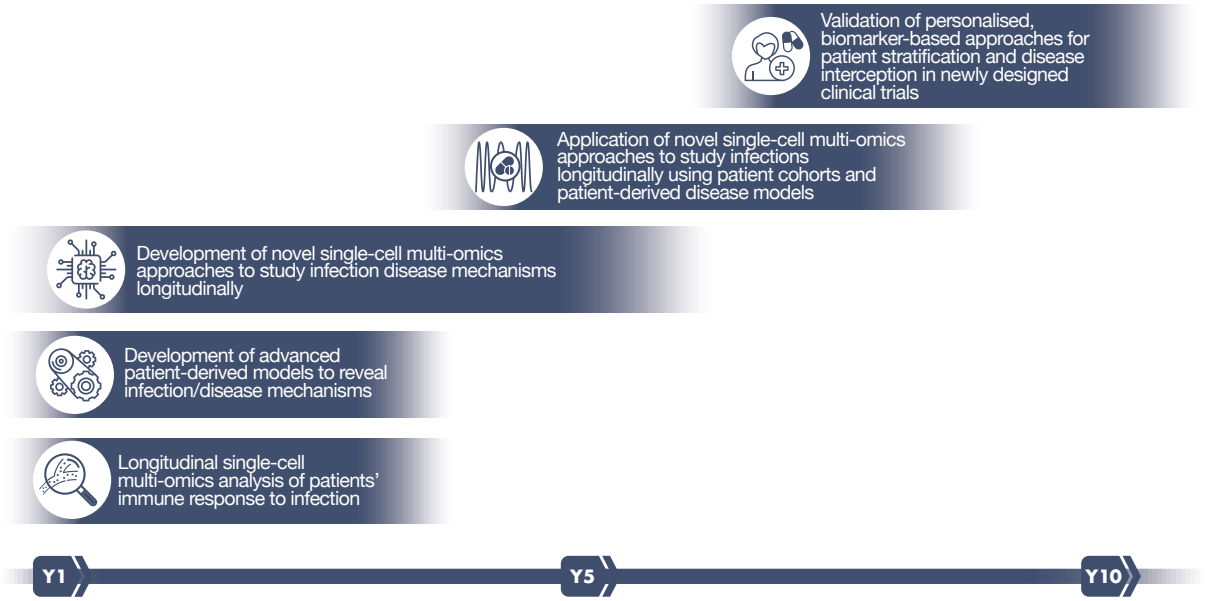
For sepsis this will provide a better classification of disease and enable early detection and accurate diagnosis of sepsis from the very first hours of the patient admission. The informative biomarkers based on enhanced disease classification will be developed into easy-to-perform clinical and laboratory tools to allow tracing the subgroups and provide patient stratification at the bench-side. Together these cellular and molecular biomarkers will help guide the physician in selecting the most effective therapies to intercept the progression of sepsis in individual patients.

For vaccination, new mechanistic insight from longitudinal analyses using next generation multi-omics technologies will reveal the specific cell types and molecular pathways involved in the heterogeneity of the response to vaccination. This will provide new strategies for enhancing vaccine efficacy. Enhanced mechanistic understanding of host-response to infections will lead to new systematic strategies for developing host-directed therapies, new vaccine leads and novel antiviral drugs, which differs from current candidate screening approaches.

Validation of personalised, biomarker-based approaches for patient stratification and disease interception in newly designed clinical trials

For COVID-19 longitudinal studies in patient cohorts will identify therapies and combinations (e.g. anti-IL-1 approaches, anti-IL-6 antibodies, interferon-beta, passive immunisation with hyperimmune serum, anti-SARS-CoV-2 monoclonal antibodies) that will be effective in specific patient subgroups. These will be tested in relevant and accurate patient-derived models, allowing fast translation to the clinics for targeted groups of patients. Furthermore, novel cellular-based therapy concepts will be likely identified in the process of understanding disease mechanisms at the single-cell level and can be tested in advanced pre-clinical models. Mechanism-based biomarkers for disease onset and therapy response will be validated in patient cohorts and used to stratify populations for clinical trials for existing and repurposed drugs, and also for novel therapies that demonstrate efficacy in proof-of-concept studies in advanced pre-clinical models. Multi-dimensional diagnostic and progression biomarkers as well as promising new combination therapies for immunotherapeutic intervention will be validated and implemented in the next seven to ten years.

For sepsis biomarkers based on multi-omics single-cell technologies for patient immune stratification at the bench-side will be validated in large-scale randomised clinical trials. Personalised treatment of infections of the critically ill patients will use an algorithm guided by the sepsis classification tool and deliver specific treatment tailored to each patient's needs. The outcome will demonstrate the treatment's potential to prolong sepsis survival compared to the standard-of-care modalities.



Infectious Diseases Roadmap timeline

Expected Scientific Impact

SHORT-TERM (1-7 years):

Increased understanding of infectious diseases in space and time and classification of the cell types, states and localisation required to predict disease progression based on cellular biomarkers

Quantification of immunotherapy treatment efficiency, and stratification of patients for the appropriate type of immunotherapy

Development of early diagnostics and therapy response based on study of patient cohorts

LONG-TERM (8-15 years):

Introduction of novel biomarkers into clinical practice for disease detection and guided therapeutic intervention

Circulating immune cells will be used as cellular biomarkers to stratify patients and predict the disease progression based on data-driven multi-omics

Design of guided-therapeutic interventions (both therapeutic and prophylactic, such as vaccines) based on single-cell multi-omics profiles

Design of new clinical trials for targeting precisely novel immune cell types and sub-types

Medical Challenge

- To understand the cellular response to infections and develop novel precision immune-based therapeutic strategies to combat infectious diseases

Outputs

- Increased understanding of the cellular basis of the immune response to infections and vaccination
- New technologies for understanding the cell types and molecular mechanisms involved in the early events during infection through to immunotherapy response
- Advanced patient-derived models to recapitulate pathophysiology and understanding infection mechanisms and immune response
- Identification of biomarkers for patient stratification for specific immunotherapy regimens
- New immunotherapy concepts to combat infectious diseases

Impact

- Earlier diagnosis of infectious diseases based on circulating immune cells obtained through minimally-invasive diagnostics
- More effective treatment of patients using precision immunotherapies and response based on multi-dimensional biomarkers used to stratify patients according to underlying pathophysiology
- Enhanced prevention by developing more long-lasting and effective vaccines based on understanding immune response

Example - Impact on COVID-19

- Reveal immune dysregulated pathways in patients with disparate symptoms and enable screening of available immunomodulatory drugs in established disease models. Identify multi-dimensional biomarkers of disease severity and immunotherapies that will have the greatest potential to restore body/lung homeostasis in specific patient subgroups. Facilitate the initiation of rapid clinical trials with selected drugs.

- Advanced disease models will enable the analysis of the temporal aspects of infections and patient-to-patient variability and sensitivity.

- Integration of single-cell multi-omics, advanced disease models and creation of machine learning-based predictive computational models will provide mechanistic insight into the regulation of the various molecular networks in cells and reveal the roles of cell types and subtypes involved in viral infection and host defense mechanisms, especially those that lead to acute respiratory distress syndrome.

- Integration of novel multi-omics technologies with machine learning approaches at a cohort/population scale will enable creation of predictive models for disease progression based on multi-dimensional datasets. Elucidation of disease mechanisms will enable systematic identification of therapies, including combinations as well as new cellular drug/antiviral targets to be validated in new patient-derived models.

- Longitudinal studies in patient cohorts will identify therapies and combinations that could be effective in specific patient subgroups. These will be tested in relevant and accurate patient-derived models, allowing fast translation to the clinics for targeted groups of patients.

- Novel cellular-based therapy concepts will be likely identified in the process of understanding disease mechanisms at the single-cell level and tested in advanced pre-clinical models.

- Mechanism-based biomarkers for disease onset and therapy response will be validated in patient cohorts and used to stratify populations for clinical trials for existing and repurposed drugs, and also for novel therapies that demonstrate efficacy in proof-of-concept studies in advanced pre-clinical models. Multi-dimensional diagnostic and progression biomarkers as well as promising new combination therapies for immunotherapeutic intervention will be validated and implemented in the next seven to ten years.

3.2.4 Chronic Inflammatory Diseases

Medical Challenges

Chronic inflammatory diseases (CIDs) are a group of non-communicable, currently incurable diseases involving the immune system. They are characterised by relapsing inflammatory activity and disease-specific secondary diseases, in which uncontrolled inflammation leads to tissue destruction. Lifetime prevalence of CIDs is estimated at over 10% in Central Europe.

Although CIDs are defined by the symptoms of the inflammatory reaction in the affected organ, they are considered to be systemic diseases with overlapping phenotypes and manifestations. Although disease-onset occurs during specific timeframes, often during early adulthood, they remain difficult to detect and there is a diagnostic delay of 12-24 months in Europe. Diagnosis relies on a combination of clinical symptom and/or organ damage scores together with additional information from endoscopic, radiologic and histological procedures. These methods are rather crude, only describe the magnitude of inflammation and often contain subjective patient reported outcome measures (e.g. intensity of pain, disability) that are difficult to use in clinical trials. The scores are unable to define the disease entity or predict outcome, meaning there is an urgent need for a more granular set of molecular biomarkers for early disease detection and stratification of patients.

The high burden of CIDs is caused by chronicity and long-term debilitating consequences due to structural destruction of the affected organ or tissue. Current therapies only treat the symptoms and do not cure or fully control the chronic inflammatory pathophysiology. While many different targeted therapies exist they are expensive and are limited by high rates of non-response to treatment. For example, while anti-TNF therapy has significantly improved the quality of life for many patients by reducing disease symptoms and slowing disease progression, full therapeutic benefit is achieved in less than 20-30% of the patients with inflammatory bowel disease or rheumatoid arthritis at an economic cost of approximately €9 billion in Europe.

The difficulty in treating CIDs is due to the observed large degree of heterogeneity in disease progression. This includes both disease activity (i.e. mild, moderate and severe) and complexity that can range from easily controlled disease to a severe destructive course, which ultimately leads to permanent impairment of organ function with a severe impact on the quality of life. Understanding the cellular basis for disease heterogeneity and the pathogenic cell-types and states involved is required to impact the future clinical management of the diseases. This requires single-cell level approaches to reveal the cellular complexity and heterogeneity of the immune system that links to disease type and disease behaviour. Single-cell based biomarkers and innovative pre-clinical models are required for early disease detection, to identify new therapeutic mechanisms, extend indications and guide therapy selection. Longitudinal analyses have the potential to go beyond simply guiding the selection of therapy and define the optimal timing of intervention (decision support) during the course of CID, which would result in a higher degree of disease control than is possible today. Future mechanism-based therapies must address the molecular cause of the disease to enable interception and deliver a cure.

Therefore, LifeTime has identified:

Understanding the role of cellular heterogeneity in predicting disease trajectories and response to therapy

as a pivotal challenge of translational research to provide a disruptive impact for CID patients in the next decade.

Chronic Inflammatory Disease Roadmap

LifeTime proposes the following objectives to generate an understanding of the molecular and cellular heterogeneity of CIDs:

- **Define the full functional cellular heterogeneity in disease-relevant context and analyse the impact of inter-patient genetic variation and microbiome on regulatory molecular networks**
- **Develop approaches for predicting disease trajectories based on disease mechanisms and endotypes**
- **Understand the heterogeneity of response to targeted therapies for optimising disease control**

This will lead to: i) a molecular taxonomy of CIDs across individual disease entities empirically linked to pathogenic cascades and individual immune network states to unambiguously detect and diagnose them earlier; ii) a systematic identification of cell populations and molecular pathways that must be targeted (by potentially different therapeutic principles) to overcome the problem of heterogeneity of disease progression; iii) a personalised predictive computational model for therapy selection and optimisation. For this roadmap, the expected impact of LifeTime´s approach on inflammatory bowel disease, rheumatoid arthritis and multiple sclerosis are used as showcase examples for the broad field of CID.

Required Investments for Chronic Inflammatory Diseases

To address the above objectives LifeTime recommends the implementation of the following multidisciplinary research programmes:

Longitudinal mapping of disease cell states and types in CID patients using single-cell based multi-dimensional analyses: novel markers for unambiguous diagnosis

To understand early pathophysiology requires increased knowledge of the roles of different cell types, including immune cells, epithelial (neuronal and glial cells for multiple sclerosis) and stromal populations (endothelial cells, fibroblasts). LifeTime proposes to apply currently established single-cell approaches (transcriptome, epigenome, chromatin accessibility) on purified cell populations (from biopsies, stool, cerebrospinal fluid and peripheral blood) and relevant disease models (organoid, pre-clinical models). Particularly important will be analyses related to life history events (e.g. treatment changes, triggers for flare) in longitudinal datasets. As regional differences of inflammatory responses are known (e.g. skip lesions in inflammatory bowel disease or specific joints being affected in rheumatoid arthritis), spatial transcriptomic analyses will link impaired cellular response states to anatomical structures and compartments to provide a currently unknown dimension to understanding the underlying disease mechanisms. Simultaneous analysis of microbiota and bacterially derived metabolites will provide additional insight into host-microbe crosstalk that shapes heterogeneity of immune cell lineages and subpopulations. These investigations will be linked to high-dimensional well curated clinical data sets and generalised clinical features such as fatigue, depression or metabolic comorbidities.

Ultimately, such studies will resolve a molecular taxonomy of disease beyond the organ-based manifestation and lead to early predictors of disease trajectories for earlier detection. As an example, for multiple sclerosis this will comprise molecular and cellular stratification markers, which reflect specific disease behaviour and comorbidities. Such an approach will enable the meaningful stratification of patients and the rationale design of future clinical trials.

Development of novel patient-derived disease models and spatial multi-omics approaches to study disease: therapy selection and new therapeutic concepts

Emerging contributions of epithelial and mesenchymal cells in CID etiology require that models comprise more than just a singular layer of cellular complexity to recreate the dynamic cell-cell interactions within specific tissue compartments. LifeTime recommends developing novel patient-derived *in vitro* disease models of CIDs (section 3.1.3). This will involve microfluidics and other solutions to model perfusion, tissue oxygenation, the role of vascular cells and immune cell recirculation. iPSC-based technologies will enable also the generation of tissue compartments, which are usually not accessible to routine tissue sampling (e.g. microglia and immune cell reactions in iPSC-derived brain organoids in multiple sclerosis or synovial organoid cultures). Promising organoid models such as the intestinal mucosa in inflammatory bowel disease, from which patient samples are accessible and can be repeatedly taken, will be further developed to capture specific aspects of disease. For example, further miniaturisation and the addition of environmental factors such as microbial communities in trans-well systems will allow a more detailed and robust view on immune-mediated processes and the parallel testing of therapeutic compounds ("clinical trial in a dish"). The integration of defined genetic models (e.g. cells derived from patients with rare monogenic phenocopies of CID) will provide invaluable insight into the role of genetic variants in influencing disease processes in susceptible individuals. Findings and hypotheses from the models should be validated in pre-clinical *in vivo* models and in smaller sets of patients in order to corroborate the relevance of the systems.

Understanding the mechanisms underlying the complex biology of cell-based CID models will require spatial-based multi-omics single-cell technologies, including spatial transcriptomics, proteomics and metabolomics as tissue compartmentalisation is particularly important (section 3.1.1). Physiological spatial gradients such as oxygen and microbiota-dependent metabolites are deregulated in inflammatory bowel disease and understanding and defining pathological microcompartments will provide essential insights into mechanism of disease manifestation and progression. To capture dynamic spatial interactions multi-omics approaches will be complemented and integrated with advanced imaging approaches to decipher the complex multicellular processes in the inflammatory niche. Imaging methodologies will also delineate processes at subcellular resolution or to decipher migratory dynamics and interaction kinetics using live-imaging (section 3.1.1). CID models offer the opportunity to develop strategies for large-scale perturbation screens (e.g. CRISPR-Cas), which can be assessed using these single-cell multi-omics technologies (section 3.1.1 and 3.1.3).

Ultimately, such refined cell-based models will substantially enhance our mechanistic understanding of CIDs and identify cellular compartments which might be specifically accessible by targeted treatment approaches. These systems will be important tools to test therapeutic principles, which will include new drug targets, repurposed available compounds as well as completely new approaches (e.g. site-specific epigenetic reprogramming to overcome differentiation impairments). For example, for inflammatory bowel disease, the current challenge is to mimic the disease process between epithelial, mesodermal (immune cells, vasculature and fibroblast compartment) and neuronal cells to understand the impairment at each layer. Overcoming current limitations (e.g. modelling influx/efflux of inflammatory cells) will enable another crucial layer of understanding and rapid pre-clinical testing of novel targeted therapies.

Development of advanced pre-clinical disease models to study CID progression and therapy response: testing new therapeutic concepts

Current animal models of CID only capture parts of the respective underlying pathophysiology. Genetically altered mice are often complete-loss-of-function models, which do not reflect specific genetic variants that are associated with CID. Therefore, LifeTime proposes a particular focus for advancing pre-clinical models for CID using CRISPR-Cas based technologies to create the required mutations in selected genetic background, introduce heterozygous SNPs and genomic regions genetically associated with CID in humans (section 3.1.3). These novel pre-clinical models will better mimic the effect of human coding variants or the effect of regulatory elements. Coupled to appropriate perturbation (e.g. by gnotobiotics, antigenic stimuli or cellular transfer), such models will be powerful tools for understanding human CID mechanisms and potential targets. Efforts to incorporate human aspects of such models will provide crucial additional insights to understand why translation from mice to humans has often failed in the past. Such models need to include the chronicity and co-morbidity aspects of CID, which reflect aspects of the respective disease e.g. fibrotic scarring and atrophy or metabolic consequences of chronic inflammation. Single-cell multi-omics technologies, including transcriptomic, proteomic and metabolic approaches developed above will enable the delineation of signalling events, cellular interactions, reconstruction of (immune)metabolic events and may also include non-invasive physiological measurements (e.g. of mitochondrial metabolism) at single-cell level (section 3.1.1).

These advanced pre-clinical models will be an important step towards closing the gap between basic research and translation of these findings into the clinics. In proof-of-concept studies, they will enable new cell-specific therapeutic concepts such as sequential immunotherapies, microbiota-based strategies to modulate specific cell types or delivery of targeted epigenetic modification therapies to be tested. This approach will lead to a disruptive advance in generating personalised pre-clinical models that will enable stratified testing of repurposed drugs and new therapies. For example, for rheumatoid arthritis and inflammatory bowel disease, one of the current challenges is to understand the exact impact of immunologically active cells, which do not belong to the classical migratory immune cell compartment, such as fibroblasts and epithelial cells. Advanced pre-clinical models to study human risk gene loci, epigenetic reprogramming of specific loci or the therapeutic transfer of autologous re-programmed cells will enable a radical paradigm shift in the treatment of CIDs.

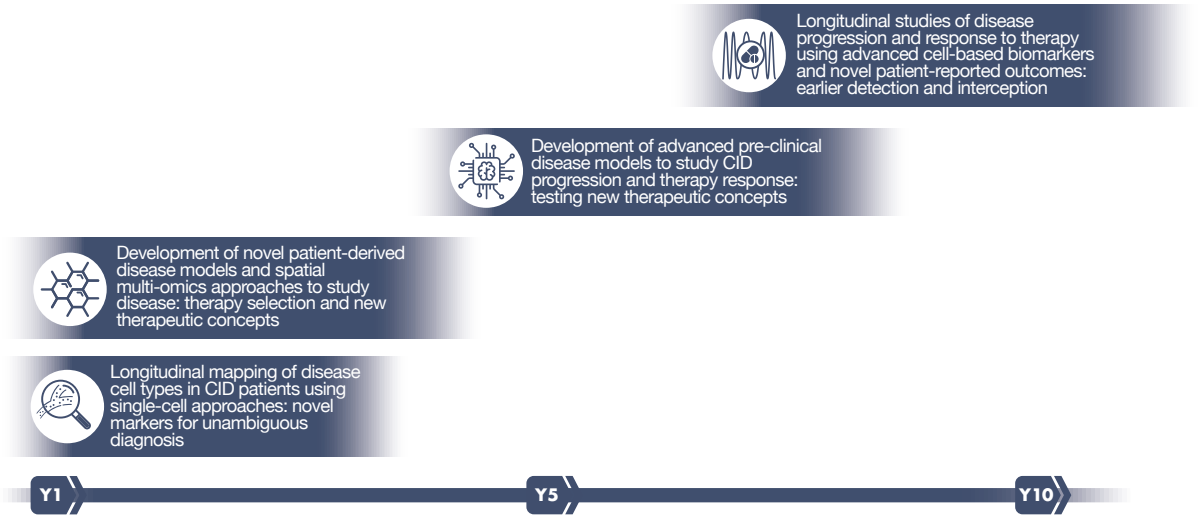
Longitudinal studies of disease progression and response to therapy using advanced cell-based biomarkers and novel patient-reported outcomes to develop innovative clinical endpoints: earlier detection and interception

A major challenge in future CID patient care is the definition of actionable disease sub-phenotypes to reliably predict disease progression and to provide a rationale for disease interception and individual therapy selection. CIDs as a disease group share a common set of genetic variants, signalling pathways and environmental risk factors, however, determinants of disease etiology are distinct from biomarker predictors of disease behaviour or diagnosis. Exploiting longitudinal high-resolution data sets from well-characterised clinical cohorts will define molecular and cellular biomarkers that identify prognostic disease endotypes to clearly predict future disease behaviour. These biomarkers will address the problem of primary and secondary non-response to targeted therapies and tailor specific therapy concepts for individual patients. LifeTime recommends to build on efforts in several EU-funded projects (H2020 SYSCID, IMI2 3TR) that are collating large CID patient cohorts receiving first time targeted therapies in long-term observational studies using harmonised protocols. This includes harmonisation of follow-up schedules, clinical scoring systems and recording of co-morbidities. LifeTime´s approach delivers a unique opportunity to exploit these and future systematic collections in the most meaningful sense by assessing the dynamics of cellular and functional heterogeneity in diseased organs and peripheral blood to extract meaningful predictive biomarkers. The approach will develop spatially-resolved single-cell multi-omics technologies into a disruptive clinical

diagnostics tool-set using machine learning-based algorithms and create predictive disease models (section 3.1.1 and 3.1.2). Successful translation requires benchmarking against state-of-the-art standard tissue-based methods, such as histological scores or routine lab tests. Defining deep remission on the cellular level, e.g by associating the presence or absence of particular cell states with prolonged response to therapy, will have an immediate impact as an innovative molecular endpoint in clinical trials. It will lead to improved and individualised usage of targeted therapies by employing quantifiable multi-modal biomarkers for a better anticipation of individual disease behaviour and definition of disease control.

Such single-cell multi-omics derived insights will provide novel concepts to clinically address disease severity and remission. Including innovative patient-centric outcome parameters will provide crucial insights into disease impact on patients' lives and integrate innovative objective measurements (wearables, smart patient home). This advanced dimensionality of patient-derived data will be important to close the gap between the high granularity and dynamics of molecular assessments and still more than crude ways to clinically describe disease behaviour (e.g. number of stools in inflammatory bowel disease).

Another important aspect of clinical diagnostics relies on innovative imaging-based methods (e.g. functional magnetic resonance imaging, *in vivo* microscopy), which offer the possibility of correlating results from cell-based molecular markers with non-invasive diagnostic principles. In proof-of-concept studies, such new combined diagnostic concepts should be tested longitudinally to demonstrate the predictive power of LifeTime´s cell-based approach and enable earlier interception of disease. This requires interactions with different stakeholders (e.g. regulatory authorities and patient organisations) for communicating and proper benchmarking of innovative clinical and molecular endpoints.



Chronic Inflammatory Disease Roadmap timeline

Expected Scientific Impact

SHORT-TERM (1-7 years):

Identification of immune cell subtypes involved in disease heterogeneity and therapy refractoriness, resolving molecular taxonomy of disease beyond the organ-based manifestation will lead to early predictors of disease trajectories, which are necessary for the meaningful stratification of CID patients and the rationale design of future clinical trials.

Refined patient-derived models will substantially enhance the mechanistic understanding of CIDs and identify cell populations for targeted treatment approaches. These models will be important tools to test therapeutic principles, including new drug targets, repurposed available compounds as well as completely new approaches.

Advanced pre-clinical models will be an important step towards closing the gap between basic research and translation of these findings into the clinics. In proof-of-concept studies, they would enable the testing of new cell-specific therapeutic concepts.

LONG-TERM (8-15 years)

Larger, biomarker-informed basket trial for targeted therapeutic intervention (proof-of principle trial to optimise efficacy), development of targeted strategies for epigenetic reprogramming of immune cells.

Medical Challenge

- Understand the role of cellular heterogeneity in predicting a patient's disease trajectory and response to therapy

Outputs

- Increased understanding of the cellular basis for autoimmune diseases and variability between patients
- New technologies for identifying the cell types and molecular mechanisms involved during the onset of disease and response to immunotherapy
- Creation of advanced patient-derived models to recapitulate pathophysiology, drug repurposing pipelines and to test new immunotherapeutic concepts (sequential immunotherapy or targeting microbiota)
- Identification of new drug targets and novel cell-specific therapeutic concepts
- Identification of biomarkers for patient stratification for specific immunotherapy regimens
- Biomarker-informed clinical trials for targeted strategies of immune cells

Impact

- Better quality of life for patients through improved disease management and control
- Earlier detection and interception of diseases
- Improved translation of new therapies based on next generation pre-clinical models and advanced cell biomarkers and clinical endpoints

Example – Impact on Inflammatory Bowel Disease

- Provide a molecular taxonomy of disease beyond the organ-based manifestation and lead to early predictors of disease trajectories for earlier detection. These will include molecular and cellular markers to enable the meaningful stratification of patients and the rationale design of future clinical trials.

- Advanced patient-derived models to enhance the mechanistic understanding of the IBD disease process between epithelial, mesodermal (immune cells, vasculature and fibroblast compartment) and neuronal cells. These models will be important tools to test therapeutic principles, which may include new drug targets, repurposed available compounds as well as completely new approaches (e.g. site-specific epigenetic reprogramming to overcome differentiation impairments). Advanced pre-clinical models to study human risk gene loci, epigenetic reprogramming of specific loci or the therapeutic transfer of autologous re-programmed cells will enable a radical paradigm shift in the treatment of CIDs.

- Single-cell multi-omics analyses will provide novel concepts to clinically address disease severity and remission. Including innovative patient-centric outcome parameters will provide crucial insights into disease impact on patients' lives and integrate innovative objective measurements (wearables, smart patient home). This advanced dimensionality of patient-derived data will be important to close the gap between the high granularity and dynamics of molecular assessments and the crude ways to clinically describe disease behaviour (e.g. number of stools).

- New combined diagnostic concepts based on combining innovative imaging-based methods (e.g. functional MRI, *in vivo* microscopy) with cell-based molecular markers for non-invasive diagnostic principles will enable earlier interception of disease.

3.2.5 Cardiovascular and Metabolic Diseases

Medical Challenge

Cardiovascular and metabolic diseases (CMD) are the leading cause of death worldwide today. In Europe, 45% of all deaths are due to CMD. With 29% of all deaths under 65 years, CMD is also the leading cause of premature mortality of which many cases are deemed to be preventable through timely and effective treatment. Despite enormous successes of interventional and drug therapies, they remain the world's largest health problem. Currently, over 85 million Europeans suffer from cardiovascular diseases, causing early disability and interfering with healthy ageing - a dramatic situation that represents an enormous economic burden.

Many CMDs lack effective therapies. These include heart failure with reduced and preserved ejection fraction, and inherited, toxic, viral, auto-immune and storage disease related cardiomyopathies, which when untreated also result in heart failure (HF). A major reason is that despite substantial progress in the understanding of epidemiology, physiology and pathophysiology of CMDs, knowledge of their cellular and molecular causes is still in its infancy. Current therapies generally do not take the underlying etiology into consideration and primarily alleviate symptoms accompanying the disease. Furthermore, the relationship between abnormal cardiac cell structure/function and pathophysiology is not sufficiently incorporated into clinical decision-making during HF therapy. A detailed characterisation of changes in cell-type composition and gene expression of individual cells in different types of HF is urgently required for a new molecular classification. This will enable subsequent patient stratification needed for optimisation and individualisation of HF treatments. In addition, the resulting novel insights into HF etiology-specific pathophysiology will trigger **novel therapeutic approaches** based on molecular profiling that will pave the way for **personalised HF treatment.**

New insight into the molecular and cellular basis of disease processes will open up a wealth of opportunities for new therapeutic and regenerative therapies as well as early recognition of CMD. This will require developing new technologies to investigate the complex interactions between different cell types, including immune cells, in the heart and the different autocrine, paracrine and endocrine signalling pathways involved, as well as various pathomechanisms such as fibrosis. Single-cell multi-omics approaches and particularly spatial technologies are needed to understand the complexity of cardiac cells and their relationships in health and disease. Single-cell multi-omics and spatial analysis should be integrated with cardiac medical technologies used to detect arrhythmias, vascular disease, or recognise acute cardiovascular events contributing myocardial infarcts to improve and automate analysis of digitalised data in cardiovascular imaging (going from electrocardiogram (ECG), echo to magnetic resonance) for **earlier disease detection and management.**

Given the emerging role of bone marrow derived inflammatory cells, which partially replace tissue resident immune cells, as drivers of pathophysiological remodelling and cardiac fibrosis, circulating cells could represent **new biomarkers** for identifying patients in whom inflammation may play a critical role in cardiac remodelling and fibrosis. Using this knowledge for early detection of CMD will allow early **interception**, which is especially important as chronification of the response to injury is of particular concern due to

the limited potential of the adult heart to regenerate. Early recognition of CMD will not only enable the prevention or an early treatment of CMD, it will also increase the sustainability of our healthcare systems.

Therefore, LifeTime has identified:

Understanding cellular and molecular mechanisms involved in CMD to enable early diagnosis and to design new mechanism-based therapies for precise clinical treatment

as a major unmet clinical challenge in the field of CMD.

Cardiovascular and Metabolic Diseases Roadmap

To address the key challenges in the field of CMD LifeTime proposes to address the following objectives:

- **Identify the heterogeneous response of parenchymal and non-parenchymal cells (e.g. inflammatory cells, endothelial cells, fibroblasts etc.) and underlying molecular mechanisms involved in CMD in organoids, tissue engineering constructs mimicking disease states, established mouse models, and human samples**
- **Deeply assess the contribution of immune cells to disease progression by determining the heterogeneity of circulating and tissue resident cells in mouse models and human samples**
- **Use the mechanistic knowledge to identify targets for interfering with pathophysiological secondary disorders that lead to cardiac dysfunction and assess their therapeutic benefit in disease models. Use single-cell multi-omics and imaging technologies to guide the use of novel therapeutic interventions to treat patient subgroups that are predicted to respond to treatment**

For this roadmap LifeTime's expected impact for heart failure is used as a showcase example in the field of CMD.

Required Investments for Cardiovascular and Metabolic Diseases

To address the above objectives LifeTime recommends the implementation of the following multidisciplinary research programmes:

Elucidating cell states and types associated with pathomechanisms

To study changes in cell-type composition and gene expression profiles of individual cells in response to different triggers of HF, LifeTime proposes to immediately use established single-cell multi-omics techniques such as transcriptome profiling and measuring chromatin accessibility in individual dissociated cells as well as spatial transcriptomics. These approaches will characterise cardiac cell states and types, their expression networks and cellular circuits, and locate these in 3D space. This will require tissue sections from patients with different forms of heart failure to understand the cellular and molecular drivers that enable functional plasticity in response to different stress responses leading to heart failure. Such tissue is available through numerous tissue banks established in a variety of population studies that exist for cardiovascular diseases. These provide the material from clinically well characterised patients required to correlate cellular and molecular mechanisms underlying structural alterations with clinical, imaging and hemodynamic data. Understanding disease mechanisms requires the systematic integration of human tissue analysis with animal models for longitudinal sampling in the presence of systemic influences such as blood pressure and immune responses.

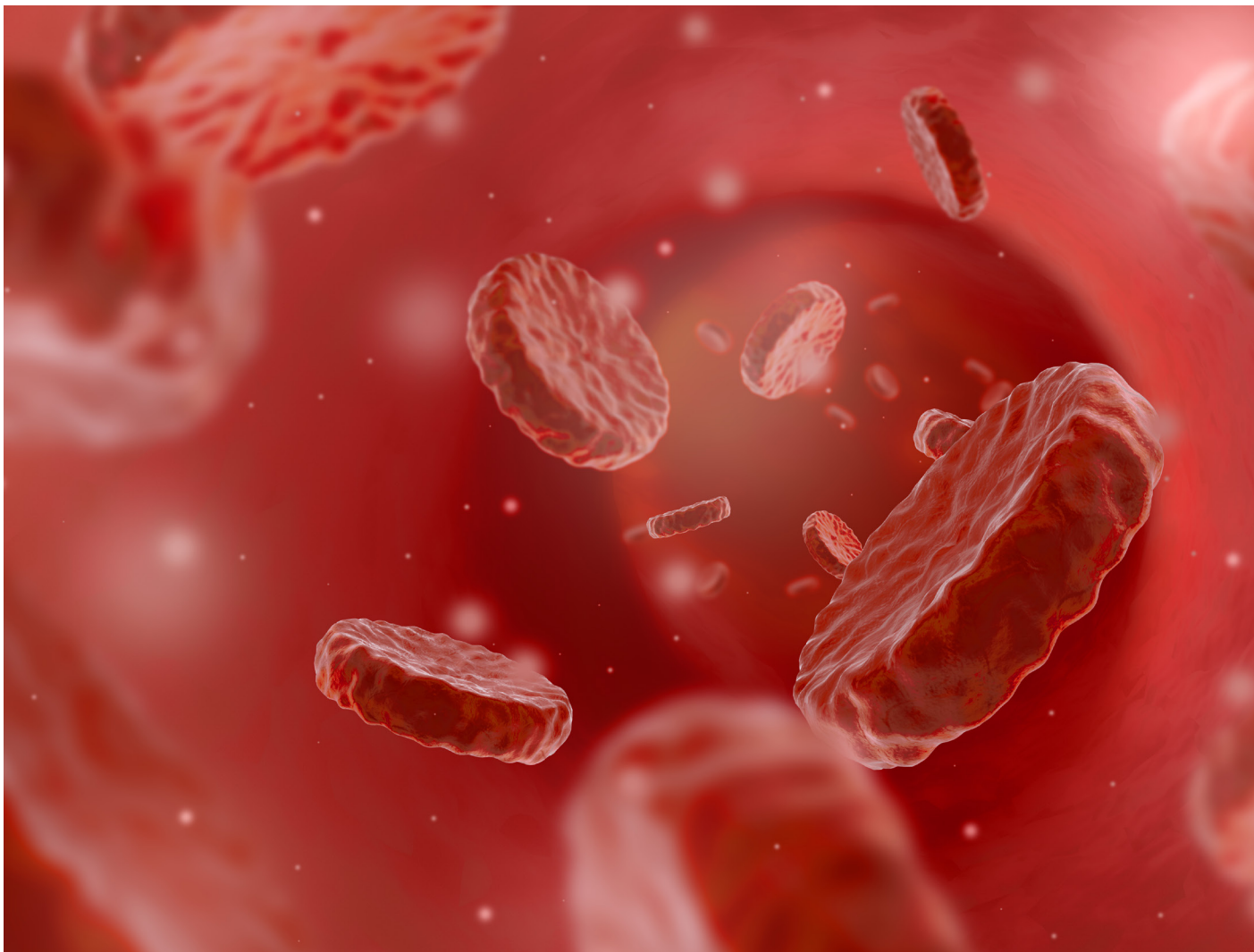
Molecular profiling together with clinical characterisation will serve as a starting point for a new molecular classification of HF and assessment of prognosis to significantly advance patient stratification for treatment and allow for a more informed interpretation of genetic testing results in the clinic. Comparing changes in cell type composition of cardiac tissue at different disease stages will provide new insight into the underlying molecular mechanisms driving the course of disease and identify early disease markers and new therapeutic targets for precision therapy of HF pathologies.

Development of novel multi-omics approaches to study cardiac disease mechanisms

Complexity of the heart and cardiac disease biology requires the development of new technologies to address specific aspects of pathology. For example, integration of single-cell full-length RNA-isoform will address currently unknown heterogeneity in isoform usage between cells, distinct cardiac regions and response to physiological stimuli as well as disease triggers. Its complex physiology means that spatial information is particularly important for the heart, and spatial knowledge of cardiac cell type arrangement in three dimensions is urgently needed to understand the heart at the single-cell level (section 3.1.1). For example, it is fundamentally important to comprehend how the swirling pattern of cardiac muscle tissue contributes to the heart's ability to pump blood effectively and how disease associated changes affect cardiac function. The cardiac conduction system initiates and coordinates the contraction of the heart across the entire organ – from atria down to the interventricular septum and along the ventricles. Dysfunction of the conduction system can cause arrhythmia and knowledge of how the specialised cells

involved are arranged in three-dimensions will lead to a better understanding of how changes predispose a person to complications such as stroke or HF. Studying cardiac cell type arrangement can be initiated using spatial transcriptomics in two dimensions, but requires tracing cells in three dimensions and ultimately in four-dimensional spatial trajectories (section 3.1.1 and 3.1.2). These will be key to elucidate how muscle tissue is built up, how neighbouring cells communicate and interact through autocrine/paracrine signals, electromechanical connectivity and biophysical interactions. Elucidation of immune cell interactions with connective tissue cells is especially important for understanding disease related cardiac fibrosis mechanisms such as cytokine release that can serve as novel therapy targets. Additional aspects of cardiac physiology require the integration of other spatial technologies (section 3.1.1). Spatial proteomics will enable the investigation of the important role of cell-cell communication in the heart based on, for example, receptor and ligand expression. Energy metabolism is of particular importance to the heart as it must contract perpetually which requires a tight coupling of ATP production and myocardial contraction. The heart is capable of remodelling the metabolic pathways in chronic pathophysiological conditions, which modulates myocardial energetics and contractile function. Therapeutic interventions have been clinically tested to target substrate preference, insulin sensitivity, and mitochondrial function to variable success. To fully understand the complexity of the cardiac metabolism in concert with cardiac contraction and cardiac tissue structures requires the development and integration of novel spatial metabolomic technologies.

The development and integration of spatial technologies will provide novel insights into structural and myocardial energetics alterations in three dimensions that will enable a detailed characterisation of contractile and cardiac conduction system dysfunctions. This will significantly expand the understanding of cardiac disease pathomechanisms and lead to the recognition of novel region-specific therapeutic targets.



Patient-derived experimental disease model development to enable longitudinal analysis of disease onset, progression and response to therapy

Cardiac organoids hold great potential in studying the molecular mechanisms driving disease progression (section 3.1.3). However, in addition to increasing the multicellularity and maturation of cardiac organoids to more accurately model the tissue it also needs to be considered that the human heart consists of specialised substructures that acquire different features during development and are physiologically and functionally distinct. Developing these aspects of human *in vitro* models is particularly important for stem cell-based regenerative, therapeutic purposes. For example, developing cardiac "patches" to recapitulate specialised substructures could be introduced into patients and replace damaged tissues. Also in disease modelling differences in cardiac tissue composition in the various heart regions is important since cardiac diseases can show structural alterations in particular regions. For example dilated cardiomyopathy affects the left ventricle to a far greater extent than the right ventricle. This clearly indicates the need for advancing cardiac organoids with the goal to achieve multicellularity including immune cells, maturing organoids to resemble adult cardiac tissue and to model the different cardiac substructures found in an adult human heart.

These new technologies will provide for the first time the possibility to perform pre-clinical screening of novel drugs and for testing of repurposing drugs in sophisticated models to provide a better understanding of their effectiveness and translatability to humans. Furthermore, patient-derived cardiac organoids can serve as personalised disease models to test new therapeutic approaches. The use of such models will significantly accelerate the bench-to-bedside process of novel therapeutic strategies based on basic research findings and will boost the field of engineered heart tissue grafts for regenerating diseased myocardium.

Identification of novel biomarkers for earlier disease detection and interception based on non-invasive liquid biopsies

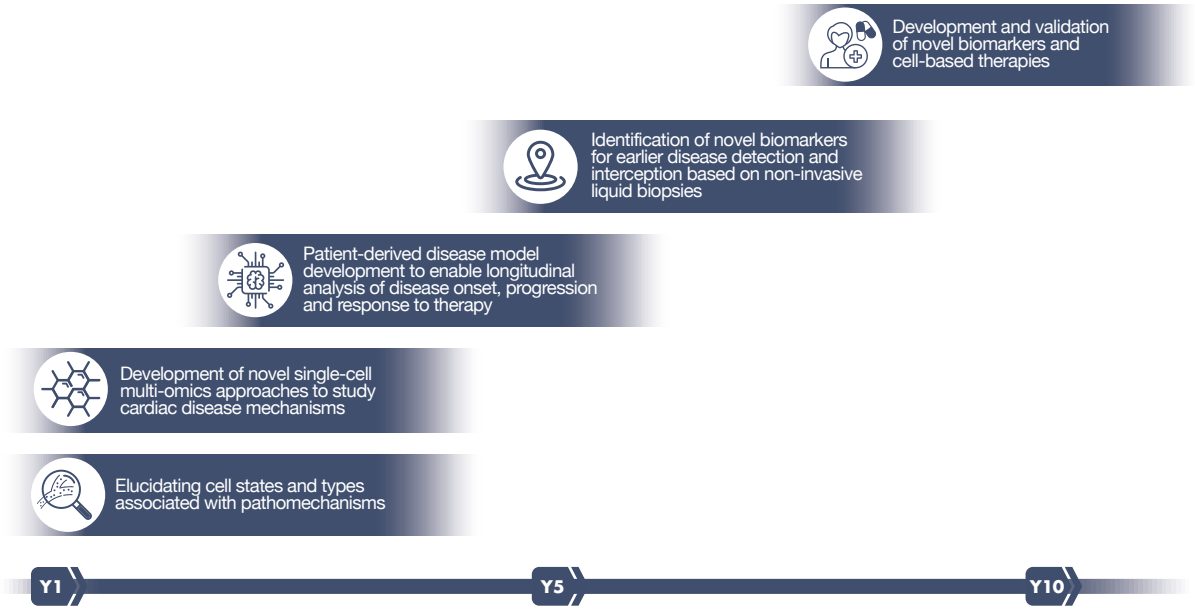
Inflammation, fibrosis and microvascular dysfunction/regression are hallmarks of CMDs and comorbidities, including diabetes, liver disease, kidney disease and cancer. The early recognition and monitoring of these processes requires establishing minimally-invasive detection strategies to identify novel multi-dimensional biomarkers. Immune cells play an important role in not only inflammation but also fibrosis and microvascular dysfunction. To obtain minimally-invasive information on the status of these processes, LifeTime proposes to investigate the potential of circulating immune cells for their correlation with tissue resident immune cell populations. Longitudinally studying the relationship between circulating and resident immune cells in animal models and correlation with results from patient blood and tissue samples will reveal if circulating cell profiles reflect resident immune cells in humans at different stages and can be used in clinical decision making. Furthermore, LifeTime proposes to study clonal hematopoiesis of undetermined potential (CHIP) in peripheral blood cells by profiling single circulating cells to identify the impact of known and unknown CHIP mutations on cardiovascular diseases. Understanding the mechanisms underlying the connection between somatic mutation-driven clonal hematopoiesis and cardiovascular disease will be highly relevant in the context of personalised medicine, as it may provide key information for the design of diagnostic, preventive or therapeutic strategies tailored to the effects of specific somatic mutations. Multiple human cohorts such as the Flemengho study, The European network on database and biobanking in cardiomyopathies and

myocarditis, The LifeLines Cohort Study and The German National Cohort (GNC) exist with banked blood for single-cell profiling of immune cells and CHIP identification. In addition, numerous tissue banks exist for the validation of biomarkers identified in studying cardiovascular disease mechanisms.

Biomarkers based on blood samples will provide, for the first time, a measure to enable detection of early signs of HF risk and thus allow for early intervention. They will facilitate tracking of disease progression and clinical decision making aligned with the course of disease. Furthermore, such biomarkers will greatly aid the stratification of populations for clinical studies of existing and repurposed drugs as well as therapies that demonstrate efficacy in proof-of-concept studies in advanced pre-clinical models.

Development and validation of novel biomarkers and cell-based therapies

We expect that in ten years we will have a significantly increased understanding of the cellular and molecular mechanisms underlying cardiovascular diseases and novel biomarkers for patient stratification. To serve as the foundation of new guidelines relevant for clinical decision-making, biomarkers will need to be validated in clinical studies or patient cohorts. The identification of novel therapeutic targets based on understanding HF mechanisms will lead to the development of new therapy approaches such as cell-based targeted anti-inflammatory and anti-fibrosis therapies as well as therapies addressing expanded somatic blood-cell clones. These approaches are already beginning to emerge, in a mouse model of cardiac injury it was shown that T cell immunotherapy can be used to limit cardiac fibrosis. Moving forward with engineered T-cell therapies in human heart failure is a promising strategy but requires further developments such as generation of T cells that can be turned on and off to limit unwanted activity. Cellular therapies need to be tested in animal models, followed by clinical studies.



Cardiovascular and Metabolic Diseases Roadmap timeline

Expected Scientific Impact

SHORT-TERM (1-7 years):

Increased understanding of underlying HF disease mechanisms, including HF trigger-specific pathophysiology that will result in the recognition of novel targets for precision therapy.

Novel biomarkers will pave the way for a new molecular classification of heart failure and for the optimisation and individualisation of heart failure treatments based on understanding underlying molecular mechanisms.

Novel biomarkers will aid patient stratification for future clinical studies of existing and repurposed drugs as well as of therapies that demonstrate efficacy in proof-of-concept studies in advanced pre-clinical models.

LONG-TERM (8-15 years)

New molecular understanding of HF pathology will impact pathophysiological concepts and clinical decision making for therapy.

Novel therapeutic modalities will be tested in patient stratified clinical studies using molecular biomarkers. Together with the novel ability to track disease course, this strategy will lead to personalised treatments that target the cellular and molecular causes of disease at the right time and will make current symptomatic treatment obsolete.

Medical Challenge

- Understand cellular and molecular mechanisms involved in CMD to enable early diagnosis and to design new mechanism-based therapies for precise clinical treatment

Outputs

- Increased understanding of the cellular basis for pathomechanisms involved in CMD
- Development of new technologies for understanding the roles of cell types and spatial arrangements including cardiac regions
- Creation of advanced patient-derived models to recapitulate cardiac structure as well as pathophysiology and be used as a basis for regenerative therapies
- Identification of biomarkers using minimally-invasive biopsies for early CMD detection
- Biomarker-informed clinical trials for clinical decision making for novel therapeutic concepts

Impact

- Earlier detection and interception of CMD risk and heart failure based on novel biomarkers and minimally-invasive liquid biopsies
- Accelerated translation of novel cell-based therapies and repurposed drugs using next generation pre-clinical disease models

Example – Impact on Heart Failure

- New molecular classification of HF and assessment of prognosis to significantly advance patient stratification for treatment and allow for a more informed interpretation of genetic testing results in the clinic.

- New insight into the underlying cellular and molecular mechanisms driving the course of HF, including specialised contractile and cardiac conduction system dysfunctions. These will identify early disease markers and new therapeutic targets for precision therapy, including new region-specific therapeutic drug targets for HF pathologies.

- New patient-derived model systems will enable pre-clinical screening of novel drugs and for testing of repurposing drugs to provide a better understanding of their effectiveness and translatability for HF in humans. The use of such models will significantly accelerate the bench-to-bedside process for novel therapeutic strategies and developments in the field of engineered heart tissue grafts for regenerating diseased myocardium.

- Biomarkers based on blood samples will provide for the first time a measure to enable detection of early signs of HF risk and thus allow for early disease detection, interception, disease progression and clinical decision making aligned with the course of disease. These biomarkers will be used to stratify populations for clinical studies of existing and repurposed drugs and new therapies.

- Validation of novel biomarkers for patient stratification in clinical studies or patient cohorts. The identification of novel therapeutic targets based on understanding HF mechanisms will lead to the development of new therapy approaches such as cell-based targeted anti-inflammatory and anti-fibrosis therapies (e.g. engineered T-cell therapies) as well as therapies addressing expanded somatic blood-cell clones.