Research, Innovation and Development Strategy

February 2022

Professor Dominic Harrington

Purpose of this document

This document provides an overview of the strategic direction and research, innovation and development (RI&D) priorities for Synnovis (2022-27). The three pillars of the strategy are: Mass spectrometry, Real-time Genome Sequencing and Pre- and Post-Analytics.

Our research, innovation and development priorities have been reviewed at this time because the new partnership with SYNLAB creates an exciting opportunity for the Joint Venture partners and other organisations across the wider south east London (SEL) research system. The intention is to position SEL at the forefront of pathology RI&D.

The delivery of our strategy will be in-part supported by the 'Innovation Accelerator Fund' the 'Scientific Learning and Development Fund' and the 'Future Leaders in Innovation group'.

Strategic approach

Synnovis in collaboration with academic and trust colleagues, and with partners from industry, will undertake translational research that supports patient care and brings commercial benefit. This approach builds on our existing strengths and uses established networks that bring early access to opportunities to gain competitive advantage in the provision of pathology services.

Clinical and scientific outcomes

- Earlier, timely diagnoses to support optimal clinical interventions
- Novel approaches to patient monitoring
- Development and application of biomarkers of prognosis
- Efficiency through multiplexing
- Synergy across the three strategic pillars
- Being closer to the patient
- Optimising patient pathways through remote testing
- Use of automation and robotics to improve quality and maximise operational efficiency
- The development of our workforce
- Brand enhancement

Business development outcomes

The RI&D strategy has been designed to align seamlessly with the Synnovis Business Development strategy. There are three components to that strategy which are enhanced by the RI&D pillars:

Hospital and clinic specialist test referrals

Across the initiatives within this strategy there are many examples of translational research which can then be proposed to our existing external referral partners. For example: In the long read real time genome sequencing programme approximately 25% of microbiology samples could potentially undergo same day sequencing.

Healthcare industries

By combining initiatives within the mass spectrometry and pre analytics pillars there are innovative new diagnostics which are in high demand within healthcare industries.
For example: In the pharmaceutical industry Novartis are seeking a way to test patients from their homes for companion diagnostics needed for their drug therapies. The drive for this is to improve the patient experience, reduce risk for the vulnerable and reduce the carbon footprint of their travelling patient population. They have 6,000 tacrolimus tests undertaken each year and a whole pipeline of other companion diagnostics which they want to develop with Synnovis.

Direct to consumer

An integrated approach across the pillars will also provide a rich source of innovative wellbeing panels which would provide a unique offer direct to consumer and to the occupational health industry particularly in pre chronic condition predictive diagnostics. Multiplexing will also provide opportunities in the direct to consumer market

Acknowledgements

With thanks to all contributors. Special thanks are due to Richard Traish, Rachel Carling, Penny Cliff, and Denise Oblein. Text for the ‘PRE-ANALYTICS’ section of this document has been repurposed from ‘Initial Request to Synnovis for Capillary Blood Home Testing’, authors Vanessa Downes, Daniella Chilton, Michaela Cashman and Bharti Malhotra (7/12/2021) with figures from Royal Brompton Hospital.

Science Technology Platform: Mass Spectrometry including Separation Science

Primary five year objective: Synnovis Mass Spectrometry Science Technology Platform designated as biomarker and assay development hub for SYNLAB Group, where research is rapidly translated into the routine laboratory service

Introduction

The breadth of potential collaboration between Synnovis staff, industrial partners and academia is a key strength today. However the service is fragmented and in some areas lacks robustness. Investment is required to prevent a deleterious effect on the specialist services we provide and on innovation. The range and nature of specialist service opportunities continues to increase and presents significant commercial opportunities. The diagnostic landscape is also evolving rapidly and we are well placed to capitalise on this through the identification of new biomarkers for monitoring and stratifying disease. New developments, especially in the field of small molecule analysis and peptides, are opening opportunities for faster and more accurate diagnoses.

Future State

The current application of mass spectrometry in the clinical laboratory is almost exclusively applied to the quantification of small molecules. However, recent developments in clinical proteomics and lipidomics, combined with the increasing availability of new technologies, reagents and automated workflows, mean that laboratory medicine not only has access to a multitude of new biomarkers, it has the capability to successfully translate them into routine use. By introducing mass spectrometry based targeted lipidomics and proteomics, laboratories have the potential to significantly expand their existing test repertoire, the benefits of which will be felt across pathology; mass spectrometry should no longer be considered as a technique with application confined to clinical chemistry.

Advances in genomics mean we can define underlying susceptibility to a particular disease but only phenotypic biomarkers can determine whether a diseased state is developing, likely prognosis and impact of therapeutic interventions. With the rollout of whole genome sequencing (WGS) across the NHS, and the proposed pilot new born screening WGS pilot, there will be an increasing demand to characterise patients who are reported to have genetic changes of uncertain significance. This is a large, and currently unmet need, but the

multiplexing capabilities of mass spectrometry and the ability to measure both small molecules and peptides simultaneously make it the ideal partner technology. One of our priorities is to introduce a multiplex test which could be used for the rapid diagnosis of acutely presenting inherited metabolic disease.

Another rapidly developing area is mass spectrometry imaging (MSI). This is a mass spectrometry based technique that provides a 'chemical map' of a surface, for example, a thin section of tissue, by combining visualisation of the spatial distribution with the chemical information obtained from mass spectrometry. Utilising sample-preparation-free techniques such as Desorption Electrospray Ionisation (DESI) and Rapid Evaporative Ionisation Mass Spectrometry (REIMS), current technology now allows for an image resolution between 10-50 µm to be obtained in minutes; the application of MSI as a tool in the clinical laboratory is imminent. Through our industry partnership and by collaborating with histopathologists at our host Trusts, we have already initiated a project to investigate the application of MSI to distinguish between different tissue types in samples from the gastrointestinal tract and to compare the results with those obtained using light microscopy.

The introduction of MALDI-TOF MS in clinical microbiology has proved to be an accurate and robust tool, allowing rapid and reliable microbial identification. However, despite its success, the technique has yet to be routinely applied to other applications in this area, specifically the determination of antimicrobial resistance. With the spread of broad antimicrobial resistance widely acknowledged to be a serious public health problem, the potential of MALDI-TOF MS to detect drug resistance is something that we will explore.

Objectives

1. Expansion of small molecule test repertoire

a. use automation to improve operational efficiency, quality and safety whilst releasing capacity of existing staff and equipment e.g. application of Cascadion, Andrews robotics, Otto solid phase extraction
b. introduction of new tests e.g. uracil, vitamin K metabolites, anti-fungals and antibiotics

c. transfer of novel tests from the Evelina London Well Child laboratory operated R&D laboratory (earmarked for closure) to Synnovis
d. transfer of existing tests from obsolete technologies to future proof critical services e.g. Biochemical Genetics Unit

2. Introduction of targeted proteomics and lipidomics

a. introduction of new tests e.g. thyroglobulin, phosphatidylethanols (PEth), hepcidin, uromodulin, glycosphingolipids, glycosaminoglycans
b. establish expertise in delivering an automated peptide workflow which can easily be adapted for new applications
c. formalise relationship with academic partner laboratory and scope out next generation of applications

3. Accessing novel technologies

a. partnership with Industry partner A to support the biomarker translational pipeline (Waters Corporation)
b. partnership with Industry partner B to support the automation of existing applications/deliver operational efficiencies (Thermo Fisher Scientific)
c. partnership with Waters Corporation to access novel mass spectrometry technologies

Example application: the analysis of immunosuppressant drugs

Annual transfer of 22,000 tests to the Cascadion – an automated mass spectrometry solution

	Current cost	Incremental costs to move to Cascadion
Staff	£291,000	£125,109
Space	£17,953	
Reagents	£72,000	£118,800
Equipment	£124,000	
Total	£504,953	£243,909

4. **Sophistication of the end-to-end pathway**

    a. introduction of alternative sample collection devices e.g. Mitra for immunosuppressant drug monitoring and Capitainer for PEth
    b. instrument connectivity (LIMS/remote access)
    c. application of robotics and automation
    d. application of machine learning & clinical decision support tools e.g. current collaboration with Amazon Web Services
    e. data archiving solutions

5. **Creation of a local and national mass spectrometry training facility**

    a. develop a pool of trained scientists in south east London with the skill set to deliver over the next five to ten years
    b. provide a structured programme of training through a combination of practical teaching, hands-on experience and e-learning (internal)
    c. consolidate links with academia (Surrey university – visiting positions), contribute to mass spectrometry based teaching and supervision of research students (BSc, MSc, PhD – basic scientists and intercalated medics), provide sandwich placements for students
    d. provision of structured training programme to external candidates

Milestones

Project start April 2022. Intentionally ambitious but realistic and necessary to successfully deliver the re-location of the existing services to a central facility, whilst maintaining and expanding the routine clinical services and embedding a new operational structure.

Short term (year 1)

- Development of a serum uracil assay to complement our existing service for DPYD testing in patients being started on fluoropyrimidine drugs (capecitabine, 5-fluorouracil). DPYD testing, which is commissioned by NHSE, looks for the four common DPYD variants which explain approximately 25% of cases of grade 3-5 fluoropyrimidine related toxicity. Predicted workload = 900 samples per month, based on DPYD testing service at Synnovis. Potential to offer nationally, no service currently exists.

- Urinary metabolites of vitamin K. The only non-invasive marker of vitamin K status (developed by the Nutristasis Unit at Synnovis). This marker of vitamin K dietary exposure and status reflects the turnover and excretion of all K vitamins. At the present time, the only available indicator of vitamin K status that may be said to directly reflect storage and transport is serum phylloquinone – which is invasive and only reflects one K vitamin.

- Blood PEth. Measurement of blood PEth has been shown to be a more sensitive and specific marker of recent alcohol intake than either carbohydrate deficient transferrin (CDT) or urinary markers such as ethyl-glucuronide (EtG) and ethyl-sulfate (EtS). Clinical applications include monitoring of exposure in the paediatric population and alcohol abstinence pre- and post-liver transplant. It also has application in occupational monitoring programmes. The test will be developed on a Capitainer DBS collection device to enable remote monitoring.

- Serum thyroglobulin. We will use SISCAPA assay technologies to measure thryoglobulin - a protein synthesised by the follicular cells of the thyroid gland and used in the production of thyroxine (further information provided in appendix 1).

Medium term (years 2 and 3)

Our primary goals in the medium term will be to take advantage of the economies of scale that co-location of the mass spectrometry service will bring and to maximise the benefits of relationships with our Industry Partners. In parallel to this we will continue to expand both the small molecule and proteomic and lipidomic test repertoires, with a particular focus on introducing tests identified by our host trusts as strategic priorities (metabolic multiplex, GFR, biotinidase, succinylacetone, NTBC) and on the transfer of existing tests from obsolete equipment to mass spectrometry based techniques (glycosaminoglycans, oligosaccharides, purines, pyrimidines). With strategic guidance from our medical colleagues, we will also continue to expand the number of tests that can be accessed remotely via self-sampling, for example, antifungals and antimycotics to support solid organ transplantation and the merger of the Royal Brompton Hospital (RBH).

Long term (year 5)

- Translational mass spectrometry facility up and running at the Blackfriars pathology hub
- Training centre of excellence – offered internally and externally (national/international)
- Synnovis Mass Spectrometry Science Technology Platform designated as biomarker and assay development hub for SYNLAB Group. Technology transfer agreements for novel applications established.
-

Required investment

Infrastructure – acute space constraints. Whilst the service is transitioning to a central facility over the next two years, adequate space will be required.

Resource – additional capacity (staff and equipment) required to support the introduction of new tests and to establish research capability through an academic partner laboratory. Key co-appointments would be required quickly to underpin this strategy including Chair and Research Fellow for academic partner laboratory (these positions provide a succession solution for the Evelina London Well Child laboratory).

Enabler - consolidation of south east London immunosuppressant workload onto the two existing Cascadion platforms to release capacity of an additional two, possibly three MS. Importantly, in the short term whilst the Blackfriars facility is being built (years one and two), this would release three of the key requirements needed to deliver this strategy in years one/two: space (Liver unit laboratory), additional equipment and experienced staff.

Staffing - The existing mass spectrometry service is fragmented across a number of different laboratories, each of which operates independently and also provides the analysis and interpretation of a range of other non-mass spectrometry based tests. The creation of the mass spectrometry Science Technology Platform (STP) will require a new team to be established, with its own operational management structure. This structure, although operated by a dedicated group of mass spectrometry STP scientists, will integrate directly

with each of the specialist services it serves (i.e. Nutristasis, Inherited Metabolic Disease etc.). This integration will be facilitated by clinical scientist posts which will have shared responsibilities between the mass spectrometry STP and the specialist services. These posts will be critical to maintaining application specific expertise and ensuring the clinical integrity of each and every result. Whilst the team will primarily be created from the current pool of scientists within Reference Services, a bespoke structure will be required and several key roles will need to be established (*new posts).

Director
Deputy Director
Technical Lead
Laboratory Manager
Quality Manager
*KHP Chair, MS
*Clinical Academic research fellow
Research Assistant

In addition, the academic R&D partner group will need to be established. It is envisaged that the posts of Chair, Research Fellow and Research Assistant will be co-funded by SYNLAB with the Chair maintaining responsibility for grant applications to fund additional posts. The Research Assistant (band 6) would be a role offered on a secondment basis to SYNLAB Group staff to support the development of the next generation of the laboratory workforce.

A research grant has been submitted to SYNLAB Group to support the introduction of targeted proteomic and lipidomics into the clinical laboratory. Serum thryoglobulin and blood PEth will be developed, validated and accredited within the next 12 months (**APPENDIX 1**).

Science Technology Platform: Long Read Real Time Genome Sequencing

Primary five year objective: Creation of a replicable real-time pathogen genome sequencing facility for the NHS

Introduction

For infectious diseases high volume, low complexity tests (urines, swabs, screening samples and community/GP samples) are likely to continue being cultured manually or by 24/7 automation (Kiestra/Copan), for the foreseeable future. High value, complex tests (blood cultures, respiratory samples and sterile fluids and tissues) have a clear benefit for faster and more comprehensive analysis covering a broader range of pathogens with often more unpredictable antimicrobial resistance and wider significance for infection control, public health and more vulnerable and immunosuppressed patient groups. Synnovis has supported the development of a translational pathway for real-time pathogen genome sequencing. The nanopore sequencing facility opened in late 2017 with an NIHR infrastructure investment award. With funding from BRC, CIDR and Synnovis Innovation awards we conducted development and step-wise evaluation of 16S respiratory sample testing over two years through to provision as a pilot pneumonia diagnostics service for ventilated patients on ICU between December and March 2020 (SYNLAB Global award winner 2021).

Future State

The aim now is to move nanopore sequencing through into a clinical service within five years. The commercialisation opportunity is significant since microbiology is still a culture-based service but there is potential for up to 25% of samples to undergo initial same day sequencing before reflex culture of only a minority. Synnovis are at the forefront of this transformation**.** Target volumes each year at GSTT are 20,000 blood, 5,000 lower-respiratory tract and 2,000 fluid/tissue samples [numbers are mean of actual 2019 and 2020 GSTT activity – and can be multiplied two or three fold for an ICS]. In addition, pure isolate sequencing from automated culture based on infection control or public health significance could add an additional 3,000 samples per annum.

Milestones

Project start April 2022. Intentionally ambitious but realistic based on progress and with continued partnership working between GSTT, Synnovis, KCL and Oxford Nanopore technologies

Short Term
Year 1

- CIDR laboratory refurbishment, recruitment, embedding first-stage increase in daily testing capacity
- Stable capability evaluating technology upgrades
- Commencement of first routine targeted and fully costed nanopore infection test (ICU respiratory metagenomics)

Medium term
Year 2

- Produce 2-3 outline business cases for all-sample processing (respiratory, blood culture, tissue & fluid sterile site, tuberculosis, outbreak & surveillance) with one moving into service

- Complete the integrated clinical and laboratory strategic business case for a Synnovis managed nanopore molecular facility

- Stakeholder engagement with commissioners, the NHS, industry, national and international

Year 3

- Produce full business case for implementing nanopore molecular facility into Synnovis routine diagnostic hub laboratory setting

Long term
Year 5

- First accredited nanopore pathogen sequencing facility up and running in the Blackfriars pathology hub

Long-read (LR) sequencing – application in cancer

Long-read sequencing enables 1,000s to 10,000s of bases to be sequenced, allowing: DNA analysis of genes that have known pseudocopies; detection of CNV more accurately; and better detection of expansion repeats. In cancers, LR sequencing can be used to detect fusion gene products by RNA analysis. At the moment the major disadvantage of LR sequencing is its relative higher sequencing cost and lower base calling accuracy compared

to Illumina sequencing. However, as this improves LR sequencing will be used more generally.

Required investment

A research grant has been submitted to NIHR through the GSTT Biomedical Research Centre Infection theme to support more research focussed early stage laboratory projects, along with research into the data-analysis, decision support and health economic modelling to demonstrate service benefits, which will be important for sustainable investment by the NHS.

Discussions are taking place with Oxford Nanopore technologies for creation of a formal partnership through investment in the laboratory infrastructure to enable iterative high throughput evaluation of their new technology advances and to provide information they require for regulatory approvals and filing.

The resource requirements for continued productive engagement of Synnovis to create the operational model and business case to maintain current involvement but with a graded increase in leadership and delivery through to taking on routine service provision in a fully managed Synnovis facility are described below.

Synnovis resource requirements

1) Continued provision of Synnovis staff comprising one 0.1 WTE consultant clinical scientist and one 0.4 WTE clinical scientist undertaking a Synnovis funded PhD, plus ad-hoc BMS secondment for nanopore testing in CIDRU during periods of pilot service (four Synnovis BMS have been trained and competency assessed for independent laboratory nanopore testing).

2) New funding for 1WTE Band 7 post that could be filled in rotation from 2-3 Synnovis BMS staff as part of a service rota.

3) Engagement of operational, finance, IT and Quality team staff to enable transfer of nanopore tests validated in CIDR through to full delivery by Synnovis as part of a routine accredited billable service.

A research grant has been submitted to SYNLAB Group to support the Development and evaluation of a metagenomics pipeline for identification of pathogens in sterile-site clinical samples (**APPENDIX 2**)

PRE-ANALYTICS

Primary five year objective: Creation of a virtual pathology service for the benefit of patients and business to business and business to consumer users

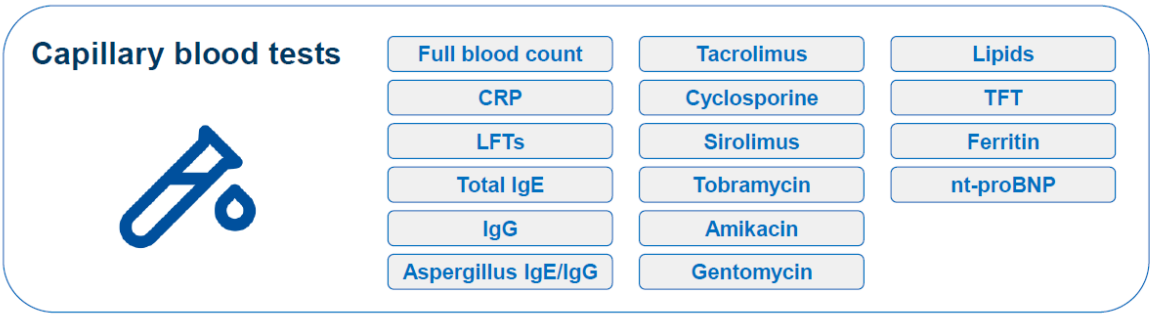
Introduction

Pathology services are heavily dependent on blood samples collected by phlebotomists bringing significant footfall to our hospital sites. During 2019 384,294 adult and 28,166 paediatric outpatients were bled by Synnovis phlebotomists. It is estimated that 65% of all adult outpatients and all paediatric outpatients attend with at least one companion. Travelling into the hospital may be inconvenient for the patient causing missed working or schooling. There is also a carbon footprint consideration. Advancements in blood collection devices present an opportunity to reduce visits to a phlebotomist. Synnovis has used blood collection devices as options for users of the Nutris Service. The approach was extended further during the pandemic for patients being treated by the Evelina London Children's Hospital. The success of these services has attracted interest from other services and hospitals. Synnovis are now in discussions with the RBH heart and lung and GSTT renal transplantation teams about expanding the Mitra tacrolimus/creatinine service for remote monitoring.

Milestones

Short term
Year 1

Build on preliminary discussions with the RBH heart and lung and GSTT renal transplantation teams about expanding the Mitra tacrolimus/creatinine service for remote monitoring. Analytes of particular interest for year one based on pre work from RBUH are shown below:



Infographic- from RBH home testing presentation

Virtual Appointments and Patient Initiated Follow Ups (PIFU)

Both virtual appointments and Patient Initiated Follow Ups (PIFU) are encouraged as part of COVID-19 recovery and NHS England has set targets as part of the H2 planning guidance.

	Target (as part of NHS England H2 Planning Guidance)
Virtual Appointments	• 30% of total consultant led appointments
PIFU pathways – as a % of all OP attendances	• 1.5% of OPD appt. by Dec 21 • 2% of OPD appt. by Mar 22.

Patient Initiated Follow Up (PIFU)

Patient Initiated Follow Up (PIFU), is a clinical model which empowers patients to manage their condition and supports self-management. It is a nationally driven initiative, which is part of the NHS Long Term Plan and in line with Personalised Care agenda. PIFU requires shared decision-making (SDM) between clinician and patient to support self-management.

There are three types of PIFU pathways:

- Discharge to PIFU - Low chance of relapse, full recovery expected
- PIFU to clinical review - Fixed regular review necessary but in between, appointments could be on demand to meet flaring condition
- PIFU to clinical review with remote monitoring – (new pathway) Regular monitoring necessary but follow up could be initiated either by symptomatic patient or by clinician on review of monitoring outcomes

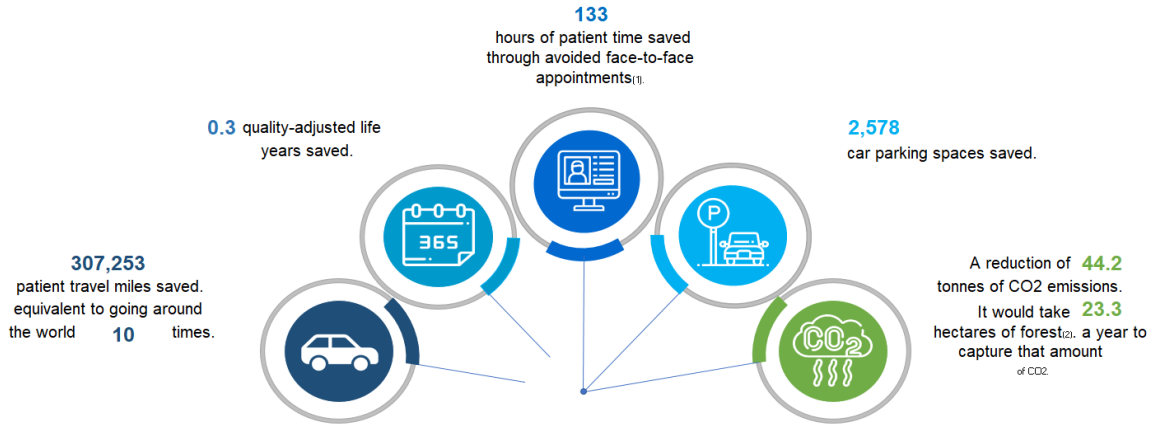
Capillary blood home testing will be essential for those patients that are on a PIFU to clinical review with remote monitoring pathway, and require regular monitoring of bloods, but cannot readily access any of the GSTT hospital or community sites.

Virtual Appointments

During the first wave of the COVID-19 pandemic, many services managed to continue care by delivering it by either video or telephone. Delivering care in this way has a number of advantages for both patients and hospitals, outlined below. The national target is for services to deliver 30% or more of their consultant led clinical activity virtually, however virtual delivery has shrunk to 20% since the first wave. An inability to deliver blood tests at

community sites or in patient's homes has been one of a few major blockers. We would like to support services in delivering remote blood monitoring to enable virtual consultations.

The following is just an example of benefits if we were to save 8,000 patient contacts in a hospital setting.



Estimated Volumes of Capillary Blood Home Testing

Based upon direct returns from services, we expect initially approx. 2,500 capillary tests per week. This number is conservative (there is an assumption that 50% of blood tests requests are suitable for capillary testing service)

PIFU to clinical review (Remote Monitoring)

In this new PIFU pathway, we aim to offer a suite of remote monitoring tools to safely monitor patients without the need for a hospital visit. Capillary blood testing, alongside phlebotomy at community sites locally, ePROM delivery and other devices e.g. O2 sats monitoring, BP, weight, ECG monitoring, will support the PIFU programme. In this pathway, patients will have the option to make contact if they feel unwell, whilst the clinical team monitor clinical data and can initiate follow up based on monitoring outcomes. This combination of PIFU and clinician initiated follow up (CIFU) will reduce unnecessary follow up for patients with long term conditions. A rough estimate of volumes of potential capillary bloods required in the next 4-6 months, is approximately 1,500, with room for further growth as more services start to roll out PIFU to Clinical Review (with Remote monitoring).

Longer term

- Adapt laboratories so that they are remote monitoring compatible. Develop equipment, workflow, IT and automation – without a robust and efficient infrastructure, scalability is not feasible

- Increase the use of remote monitoring to change patient pathways. Gain understanding of the drivers towards remote monitoring – whether that be sustainability and/or reduced clinic footfall

- Gain a better understanding of how remote monitoring fits with point of care testing. Do samples need to come to the laboratory at all?

- Increase repertoire of tests with commercial opportunities

POST-ANALYTICS

Primary five year objective: Pioneers in the use of machine learning in a clinical decision support system for the routine diagnosis of rare diseases

Introduction

The interpretation of data by scientists working in specialist pathology laboratories is predominately performed manually. Machine learning has the potential to improve patient care and increase productivity if it can be developed and used as a clinical support tool. We will use the diagnosis of rare diseases as an example of the potential utility of machine learning at Synnovis. In particular we will focus on plasma amino acid profiles that are used routinely for the diagnosis and monitoring of inherited disorders of amino acid metabolism, organic acidemias, and urea cycle defects as a proof of principle. Machine learning is not however limited to this application and is likely to have utility across Synnovis – for example, in the analysis of variants of uncertain significance in the genetics laboratory where machine learning will predict the impact on the folding of proteins and estimate the impact on function.

Future state

A multi-centre collaboration to validate the use of machine learning in a clinical decision support system for the routine diagnosis of inherited metabolic disorders.

Successful introduction of machine learning as a clinical decision support tool for the routine interpretation of plasma amino acids profiles would be a ground breaking step in the field of metabolic biochemistry. It would be the first true application of machine learning in this field of laboratory medicine and achieving this through a multi-centre collaboration would add to the impact. The immediate benefits to be realised would be savings on time and cost, improved quality and efficiency, a consistent approach to interpretation, a reduction in the error rate and improved resilience/reduced reliance on what is widely acknowledged to be a very limited pool of expertise.

Effectively, implementation of the classifier will fundamentally change the workflow in a metabolic biochemistry laboratory whilst increasing the capacity.

In a field based entirely in the study of rare disease, the advantages of utilising global data to make a patient diagnosis are significant. The machine learning platform will effectively act as a “magnifying lens” on inherited errors of metabolism and will catalyse collaboration between subject matter experts through data sharing. We will use standardised data schemas to drive harmonisation in clinical practice which could potentially accelerate the drive for analytical standardisation.

From a technology perspective, the machine learning platform will be built to ensure it complies with the strict security and data governance requirements dictated by a multi-centre collaboration spanning different hospital, countries and even continents. The machine learning platform will scale the computational resources robustly and automatically to allow the analysis of large volumes of data and to serve large numbers of real-time inference requests.

Table 1: Laboratories and named contacts for multi-centre collaboration.

Contact Name	Laboratory
Stuart Moat	UHW, Cardiff, Wales
Cécile Acquaviva	Hospices Civils de Lyon, Lyon, France
Piero Rinaldo	Mayo Clinic, USA
Denis Dietzen	St Louis Children’s Hospital, St Louis, USA
Alessio Cremonesi	Kinderspital, Zurich, Switzerland
Déborah Mathis	University Hospital Bern, Bern, Switzerland
Claus Dieter Langhans	Heidelberg, Germany
Lars Mokrid	Oslo, Norway
Olivier Braissant	Centre hospitalier universitaire vaudois, Lausaune, Switzerland
Jorgen Bierau	Maastricht, Netherlands
Sabine Scholl Burgi	Innsbruck, Austria
Ronda Greaves	The Royal Children’s Hospital, Melbourne, Australia

Outcomes

Short Term

Year 1

- Data governance and data sovereignty: collate amino acid data from laboratories in Europe, US and Australia
- Standardise a data schema to facilitate collaboration between the laboratories.
- Gain access to an increased number and variety of diagnostic amino acid profiles.
- Use the collated data set to further train and test the three previously developed classifiers
- Assess the precision and recall of each classifier
- Establish Synnovis access to the shared environment. This will be globally accessible, secure and only accessible to those with authorised access.
- Trial the use of the classifier in the routine clinical environment (Inherited Metabolic Disease laboratory, Synnovis) and review performance
- Accredit the algorithm for routine use in the clinical laboratory
- Assess success: a reduction in time spent interpreting amino acid profiles by clinical scientists; performance of classifier as applied to EQA samples

A multi-centre collaboration between an international group of laboratories with expertise in inherited errors of metabolism who are currently accredited providers of a routine clinical plasma amino acid service by either IEC or LC-MS/MS. Plasma amino acid results will be collated in an agreed, defined data scheme. Results will be both numerical and categorical. Age and sex of each patient will be included. Interpretation (ground truth) will be assigned to each profile by the analysing laboratory and will be included in an agreed data schema.

Materials and Methods

The multi-centre data will be collated by method (IEC vs LC-MS/MS) and pooled. Data will be evaluated using feature selection. Both a binary and multi-class classifier will be trained and tested using the k-fold cross validation approach.

Data analysis: precision, recall and F score will be determined for each classifier. EQA data (ERNDIM Cognitive amino acid scheme) will be used to evaluate the performance of each classifier against an independent 'ground truth'.

Synnovis Role

Synnovis will provide clinical, scientific and technical expertise in plasma amino acid analysis and interpretation and other pathology related issues. Synnovis will act as Product Owner and will be responsible for communication and coordination of the multi-centre collaboration, agreeing a common data schema and evaluation of the final classifier(s). Amazon Web Services (AWS) will utilise its Healthcare and Life Science Professional Services team who are a multi-skilled solution builders with expertise across the health and science domain and in developing solutions utilising cloud technology. They will leverage the infrastructure, security, data engineering and analytics, and Machine Learning expertise to work alongside the multi-centre participants to deliver this project.

Milestones

<u>Workstream 1: Gather data governance requirements</u>

- Identify data governance requirements within Synnovis and with collaborating laboratories.
- Centralise the data in Synnovis' AWS account
- Assumptions:
    - Product owner from Synnovis will lead the engagement and ongoing communication with collaborating labs.
    - AWS consultants will support the sessions with technical expertise about best practice and technical feasibility.

<u>Workstream 2: Standardise data schema across collaborating laboratories. Collect data from the collaborating laboratories.</u>

- Agree template for data collection fields from external labs and split of case data.
- Data collection from collaborating centres.
- Assumptions:
    - Product owner from Synnovis will lead the engagement and ongoing communication with collaborating labs.
    - Collaborators will be willing to provide their data early on.

- This is a hard dependency for the data science workstreams below.

Workstream 3: Build foundations of production cloud platform
Estimate: 20 FTE days

- Produce design for cloud environment that establishes secure data transfer, storage and computation.
- Implement environment with infrastructure as code.
- Enablement sessions with technical teams from Synnovis and collaborators.
- Assumptions:
  - Synnovis will identify in-house IT team that will be responsible to take over and manage the platform in the long term.

Workstream 4: Data Science - Data exploration

There will be two potential sources of variability: different laboratories and two methodologies

- Consistent distributions of feature values?
- Consistent data types?
- Establish methodology to handle discrepancies between schema, data type or distribution
- Assumptions:
  - Hard dependency on WS-2: For this WS to start the schema must be standardised and some data from other labs must be available.

Workstream 5: Data Science – Model development / binary classifier

- Develop a binary classifier that can be trained on data from multiple laboratories and on both methodologies.
- Work from the current project phase will be leveraged where we show generalisation across IEC and MS on Synnovis' data.
- Establish whether the classification algorithms can be trained on data from multiple labs.
- Assumptions: Hard dependency on WS-2

Workstream 6: Data Science – Model development / multi-class classifier

- Starting with Synnovis data
- Once feasibility with Synnovis' data is established, feasibility with data from other laboratories will be established.

- Assumptions: Depends on WS-2

Workstream 7: Front End

- Development of user interface: Web portal to upload sample and get inference in real time.

Medium term

Year 2 - 3

- Following successful introduction of the plasma amino acid classifier into the routine clinical environment, use of machine learning-aided diagnostics will be expanded to include other applications, both within the metabolic biochemistry laboratory and more widely across laboratory medicine.

Longer term

- Although we have much to learn and test we know that in other areas of medicine machine learning has been shown as capable of extracting insights from unexpected sources and drawing connections that humans would not normally anticipate, which opens a broad array of possible applications.

A research grant has been submitted to SYNLAB Group to support a multi-centre collaboration to validate the use of machine learning in a clinical decision support system for the routine diagnosis of inherited metabolic disorders (**APPENDIX 3**).