Project_description

PROJECT TITLE

Metabolic Signatures Toward Type 2 Diabetes: Decoding Natural Incretin and Glucagon Pathways in conjunction with metabolic traits in the Progression to Type 2 Diabetes and Related Complications

BRIEF PROJECT DESCRIPTION incl. CHOICE OF DATABASE (maximum 2,000 characters, including spaces, line breaks and special characters)

This project investigates the role of natural incretin and glucagon responses as drivers of progression toward diabetes and diabetes-related complications.

People at high risk of type 2 diabetes are not currently being systematically captured, primarily due to heterogeneity in who progresses to diabetes. Many of the driving and compensatory protective mechanisms involved in the development of type 2 diabetes and its related complications remain to be uncovered. Evidence from cohort studies on the long-term consequences of disrupted natural incretin and glucagon responses remains limited.

This proposal project investigates a missing piece in the puzzle the role of natural incretin and glucagon, aiming to identify the primary drivers of progression toward diabetes and diabetes-related complications. The project will leverage the deep-phenotyped data in the Fenland study representing the UK general population and validate in the Danish ADDITION-PRO and UK biobank. The project is divided into two parts: (1) Using traditional and newer methods of causal inference, including genetic predicted and phenotypic measures, we aim to determine the causal relationship between incretin and glucagon and long-term cardiometabolic outcomes and to identify mediating pathways such as liver health, inflammation and metabolites. This will be investigated using data from the Fenland Study and validated in large population studies such as the Danish ADDITION-PRO and UK Biobank. (2) We will investigate incretin and glucagon responses in conjunction with current components of heterogeneity in the risk of

metabolic dysfunction, such as liver health and inflammation, and aim to decipher the impact of each metabolic trait on long-term cardiometabolic risk.

By using novel epidemiological methods to characterizing peoples' risk of progression towards cardiometabolic disease, and identifying those with a high potential benefit from earlier intensive intervention, this proposal aim to aid prevention of type 2 diabetes.

Choice of database

PROJECT DESCRIPTION (maximum 20,000 characters, including spaces, line breaks and special characters).

Specific aim

This Inter-SUSTAIN Part 2 projects aims to elucidate the role of natural incretin and glucagon responses as drivers of heterogeneity in the long-term risk of diabetes and related complications, and to map clinical and biological traits related to these responses with particular emphasis on liver health, adiposity and inflammation. Thereby, identifying patterns of metabolic dysfunction that may predict progression to type 2 diabetes and inform personalized treatment strategies.

The specific aims are to:

a) To phenotypically and genotypically investigate the causal role of natural incretin and glucagon responses to an oral glucose tolerance test (OGTT) in determining the risk of regression to normoglycemia, progression to type 2 diabetes, and the development of diabetes-related complications. This will be conducted using data from the Fenland study and the Danish ADDITION-PRO cohort, linked to comprehensive outcomes from the Danish National Health Registries.

b) To identify metabolites and biomarkers for liver health and inflammation mediating the association between genetically predicted (Fenland study, ADDITION-PRO, UK Biobank) and phenotypically measured (Fenland study) incretin and glucagon responses, and cardiometabolic outcomes; with a particular emphasis on direct and indirect markers of liver function.

c) To map dimensions of metabolic traits in conjunction with incretin and glucagon responses, in order to decipher their role in the context of deteriorated glucose metabolism and their relationship to long-term cardiometabolic health

Background and significance of

Pre-diabetes is a complex state associated with increased risk of progression to type 2 diabetes and related complications[1–3]. However, not all people with pre-diabetes progress to diabetes; many persist in the prediabetic state or regress to normoglycemia over the course of various years[2]. Recently the idea of metabolic heterogeneity in type 2 diabetes has been extended to pre-diabetes, and traits that distinguish individuals most likely to progress to diabetes, such as insulin-resistant fatty liver and visceral adiposity-related renal dysfunction, have been identified[1]. In recent years, incretins (Glucagon Like Peptide 1 [GLP 1] and Gastric Inhibitory Polypeptide [GIP]) and the hormone glucagon have gained increasing attention as targets in therapeutics for prevention of type 2 diabetes and diabetes-related complications[4–8]. GLP-1 and GIP enhance post-meal insulin secretion, supporting efficient nutrient handling[9]. Glucagon serves as a counter-regulatory hormone to insulin and maintains glucose balance both during fasting and following glucose intake[10]. However, in the context of heterogeneity in pre-diabetes, the role of natural incretin and glucagon responses for the progression to type 2 diabetes and diabetes-related complications remains unexplored. Prior research in the Danish ADDITION-PRO study showed that, individuals with pre-diabetes and type 2 diabetes exhibited up to 25% lower natural GLP-1 responses to oral glucose[11]. Additionally, impaired glucagon suppression and elevated fasting glucagon levels were observed in individuals with insulin resistance and early glucose dysregulation[12]. A limitation of the ADDITION-PRO cohort is the selection of individuals at high risk of diabetes based on diabetes-specific risk scores, which limits the generalizability of the findings to the broader population

The Fenland Study represents a pioneering population-based cohort that has substantially advanced the precision of identifying individuals at high risk of type 2 diabetes[13] and elucidating causal pathways underlying cardiometabolic disease [14]15[16]17. Leveraging extensive phenotypic and genetic data, the study has established causal relationships between liver function markers, alanine aminotransferase (ALT), aspartate aminotransferase (AST), and alkaline phosphatase (ALP), and type 2 diabetes[14], as well as between adiposity-related traits, fat-mass ratios, and metabolic dysfunction-associated liver disease[17]. Moreover, Fenland data have uncovered mechanistic links between interleukin-6–mediated inflammation and type 2 diabetes[18], highlighting the interplay between inflammatory and metabolic pathways. The cohort's precise characterization of adiposity, inflammation, and hepatic health, combined with detailed assessments of incretin, glucagon, insulin, and glucose dynamics in a general population, provides an unparalleled resource for investigating the constellation of metabolic traits and their long-term impact on cardiometabolic health. Their cohort provide an unique timing for studying contellations of metabolic traits relevant to progression to cardiometablic disease and how the play in conjuctions to glucose and hormones.

This project is grounded in the hypothesis that dysregulated incretin secretion and inadequate glucagon suppression, together with insulin resistance, accelerate the progression of metabolic dysfunction by activating hepatic and inflammatory pathways in individuals with pre-diabetes.

Work Package 3 (Heterogeneity) in the new DP-Next

Clinical applicability

This project supports the development of a precision approach to pre-diabetes by identifying a set of easily obtainable biomarkers that can optimally distinguish individuals with a high probability of stable pre-diabetes or remission from those at greatest risk of progressing to diabetes. This work is connected to the objectives of Work Package 3 (Heterogeneity) in the new DP-Next project, which aims to develop new strategies for diabetes prevention[19]. I will work in close connection with the DP-Next team and it is expected that the findings from the Fenland study and ADDITION-PRO will inform the design and structure of the new DP-Next cohort.

The ADDITION-PRO study, in combination with the Danish National Health Registries, offers a unique opportunity through long-term follow-up data to support these efforts. It enables the identification of phenotypic and biological metabolic traits that characterize individuals at high risk of developing type 2 diabetes and related complications, including macrovascular and microvascular conditions, as well as increased mortality. Furthermore, these traits can be linked to longitudinal changes in HbA1c, lipid profiles, and eGFR, allowing for the identification of individuals with accelerated progression toward cardiometabolic complications. These insights support the timely identification of individuals who may benefit from early interventions, such as intensive lifestyle modifications or specific medications.

Work Package 1: Investigate causal relationship of incretin and glucagon and long-term cardiometabolic disease and identify mediating pathways.

Hypothesis for Work Package 1: Natural incretin and glucagon responses play a causal role as drivers of the risk of progression from pre-diabetes to type 2 diabetes and diabetes-related complications. In this pathway, biomarker and metabolites related to liver health play a mediating role.

Natural incretin and glucagon reposes in epidemiological studies have been mostly investigated cross-sectionally. Longitudinally, a higher GLP-1 response during OGTT was associated with a steeper decrease in fasting glucose over time[20]; however, with a limited sample size (n=121). [The ADDITION PRO cohort includes 1657 individuals at high risk of diabetes who underwent measurement of natural incretin and glucagon responses during a 3-point (0, 30, 120 min) oral glucose tolerance test (OGTT) in 2009-2010, i.e. before the introduction of incretin based therapies[11,12,21]. This unique timing provides an opportunity to prospectively analyze the associations between individuals's natural hormonal responses to glucose and progression to diabetes, regression to normoglycemia, or risk of diabetes-related complications.

Preliminary findings from ADDITION PRO indicate that per SD higher glucagon (incidence rate ratio [IRR]: 1.38 [CI: 1.15 to 1.67]) and GLP-1 responses (IRR: 1.34 [CI: 1.05 to 1.71])

at 120 minutes during the OGTT are associated with incident heart failure over an 11-year follow-up period.

Genetic variants, clinical biomarkers and metabolites serve as valuable indicators of individual predisposition for specific phenotypes (genes) and mediating pathways (metabolites and clinical biomarkers) that link phenotypes to disease outcomes. Genetic variants can strengthen causal inference by serving as instrumental variables in Mendelian randomization studies[22]. Identification of metabolic pathways that mediate the effect of genetically determined phenotypes on the incidence of diseases such as type 2 diabetes helps us understand the causal biological mechanisms linking traits to disease. This understanding in turn supports the development of potential new targeted interventions based on metabolic pathways.

Evaluating the risk associated with incretin and glucagon responses to a glucose load in isolation may be limited without considering concurrent insulin and glucose regulation. To address this, in WP1 principal component analysis will be applied to extract the strongest sources of variance in the concurrent OGTT response across all 4 hormones and glucose.

Research design and methods

To quantify the causal relationship between incretins and glucagon and cardiometabolic disease, we will first apply traditional epidemiological time-to-event analysis methods (Poisson and Cox regression) in Fenland study, with adjustments informed by pathways outlined in directed acyclic graphs (DAGs). To strengthen causal inference and account for unmeasured confounding, Mendelian randomization will be conducted using pathway specific polygenic risk scores (PRS) as instruments for incretins and glucagon[22], leveraging data from ADDITION-PRO, the UK Biobank and Fenland Study[23].

To sequence mediating pathways and identify key metabolites linking incretin and glucagon responses to cardiometabolic disease, we will apply a structural causal algorithm (NetCoupler)[24] to identify sequential causal pathways involving metabolites. Based on the identified metabolites, structured causal mediation analyses[25] will be conducted to assess and quantify the direct and indirect effects of individual metabolites seen as most likely to be causal mediators in NetCoupler on cardiometabolic outcomes.

trash can and in a subgroup of around 800 participants, still B-mode ultrasound images of the liver, collected with a standardized protocol[21]. WP3 will apply novel AI-based image processing and analysis methods to extract features and cluster images.

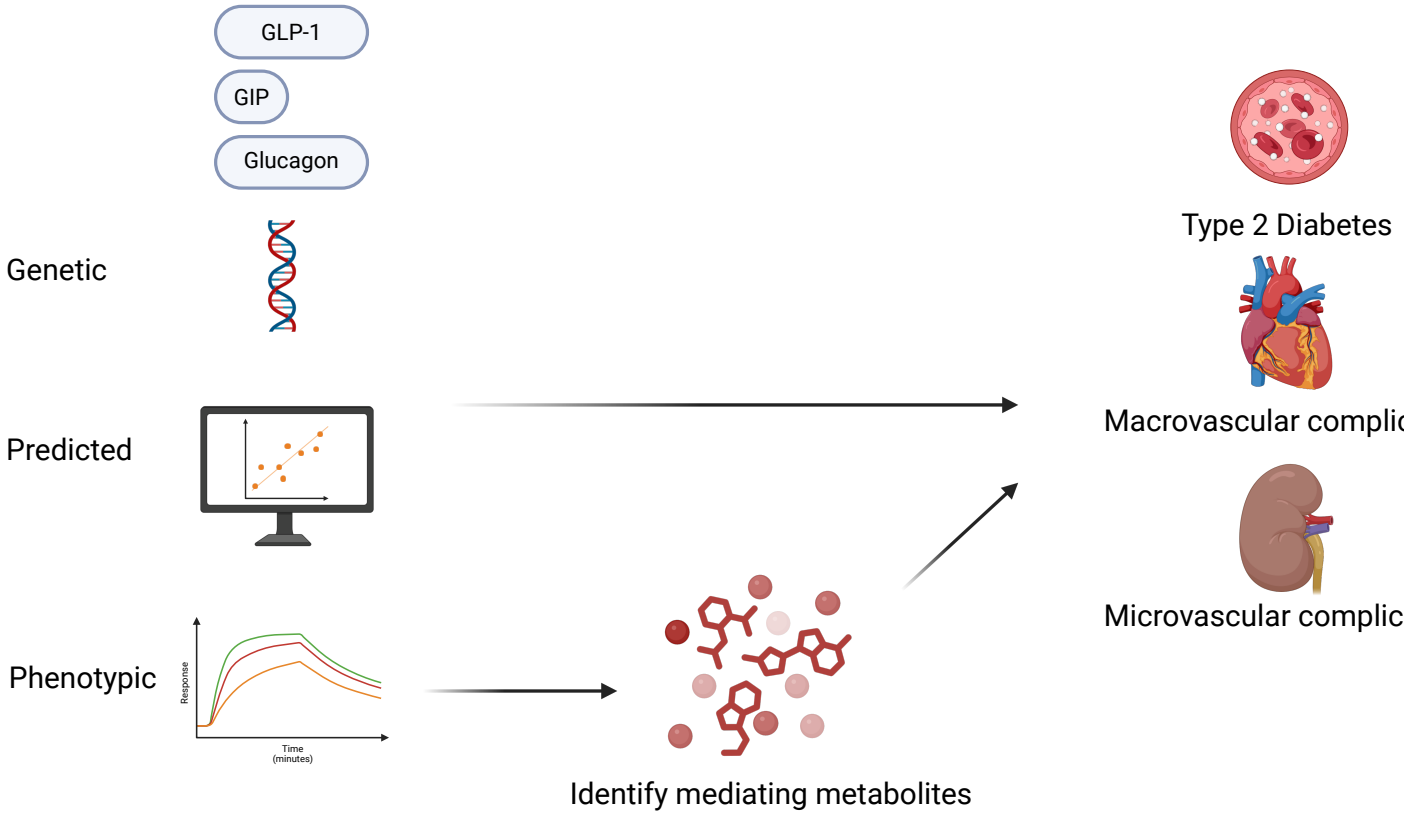

Figure 1: **Figure 1: Structured causal analysis framework of incretin and glucagon responses and their association with the risk of progression to cardiometabolic disease**

Work Package 2: To map dimensions of metabolic traits in conjunction with incretin and glucagon responses, in order to decipher their role in the context of deteriorated glucose metabolism.

Hypothesis for Work Package 2: natural incretin and glucagon responses may exert protective, compensatory effects under conditions of obesity, hyperglycemia, insulin-resistance, and loss of beta-cell function. We further hypothesize that liver health may act as an effect modifier in the associations between incretin-glucagon responses and metabolic outcomes.

Obesity and insulin resistance have long been regarded as central components in the development of type 2 diabetes. Recent clustering analyses have identified liver fat as a main determinant heterogeneity among people with pre-diabetes with regard to their risk of progression to type 2 diabetes[1]. Reduced liver function interacts with insulin resistance, glucose and glucagon responses, and is accompanied by low grade inflammation and morphological changes to the liver, particularly MAFLD (Metabolically Associated Fatty Liver Disease)[10,26,27]. The biological actions of GLP-1 and GIP include reducing hepatic inflammation[28]. Liver function therefore appears to be a central pathway modulating diabetes risk and interacting with or mediating the impact of natural incretin and glucagon responses. The Fenland study includes baseline measurements of -glutamyl transferase and alanine transaminase, and 11,559 participants had Dual Energy X-ray Absorptiometry (DEXA)[29] performed allowing to estimate regional fatmass inclunding liver fat. In conjunction with liver function biomarkers, this will enable the calculation of a general indication of liver health.

We will further integrate dimensions of metabolically relevant low-grade inflammation by integrating markers of macrophage activation (soluble CD163), adiponectin, and C-reactive protein[30] in our multidimensional characterization of liver health. In addition, fasting branched-chain amino acids (BCAA)[31]and lipid-related metabolites during the OGTT[32] were obtained in ADDITION-PRO. These data allow us to characterize aspects of metabolic function, such as liver fat accumulation[33]and tissue-specific insulin resistance[34,35] .

Based on phenotypes, metabolites, and genotypes, we want to identify metabolic traits that, in conjunction with incretin and glucagon responses, contribute to cardiometabolic risk. Machine learning–based dimensionality reduction techniques help identify and characterize distinct constellations of metabolic traits across individuals. From phenotypic clusters, we want to map dimensions of metabolic function that either provide compensatory protection against or contribute to cardiometabolic disease. Then, we will add dimensions of metabolite-based profiles to the existing clusters to enhance their characterization.

To support phenotypic clustering, genetic data will be incorporated through partitioned polygenic risk scores (PRS) for type 2 diabetes. These scores reflect underlying traits such as glucagon and incretin, inflammation, beta-cell function, obesity, liver fat, and insulin resistance. This approach enables the identification of key phenotypes by genetic risk scores that drive cardiometabolic disease.

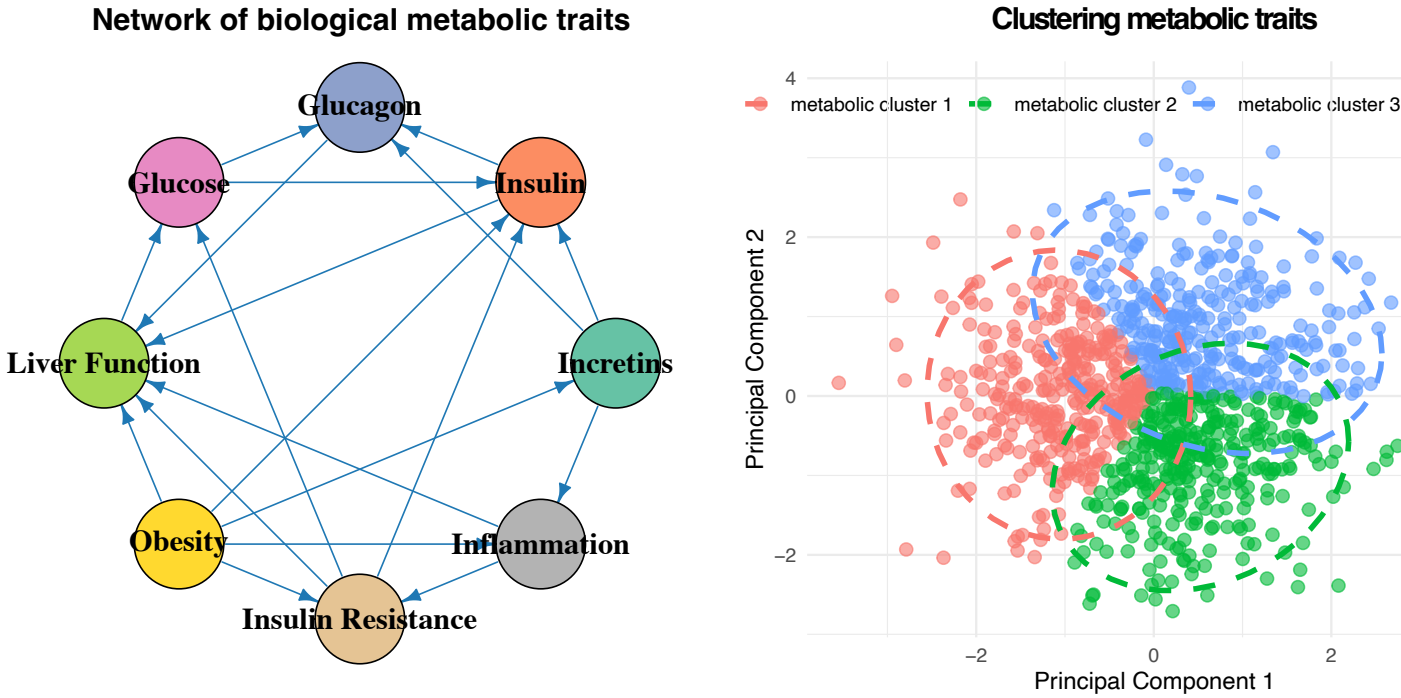

Figure 2: Figure 2: Illustration of dimensionality reduction of metabolic traits into metabolic clusters

Research design and methods

We use data from previous work in The Fenland study and ADDITION-PRO, including genotypes, metabolomic profiles, and markers of metabolism and inflammation[11,12,21,30–32,36].

Combined with liver function biomarkers, this approach will enable us to generate a general, multidimensional indication of liver health.

We will apply principal component analysis to integrate variables and identify a restricted number of metabolic dimensions that reflect the interplay between incretin and glucagon responses along with concurrent glucose-insulin dynamics during the OGTT[37]. We will further add dimensions of obesity, insulin resistance, inflammation, liver health markers, subsequently examine these dimensions in relation to diabetes risk[38]. Finally, metabolite-based profiles and partitioned PRS will be analyzed separately in relation to cardiometabolic outcomes and subsequently incorporated to enrich the dimensional metabolic profiling. All components will be investigated prospectively in relation to cardiometabolic outcomes.

trash can and in a subgroup of around 800 participants, still B-mode ultrasound images of the liver, collected with a standardized protocol[21]. WP3 will apply novel AI-based image processing and analysis methods to extract features and cluster images. We will leverage recent advances in AI-based image analysis by applying transfer learning approaches and open-source feature extraction and image classification algorithms to ultrasound liver images[39].

Work Package 3: To predict incretin and glucagon responses based clinical and biological traits

Hypothesis for Work Package 3: dysfunctional individual incretin and glucagon responses in conjunction with can be predicted to a clinically useful degree based on clinical, genetic and metabolomic biomarkers, and these predictions can stratify individuals by their risk of progression to type 2 diabetes and related complications.

Given the fact that it is impractical to assess a full incretin and glucagon response in clinical practice, WP2 aims to establish a practical set of biomarkers that jointly have the capacity to characterize incretin and glucagon responses. Analyses in the Fenland study will be carried out across three levels: (1) traditional clinical characteristics, (2) genomic data, and (3) metabolomics profiles. The utility of these predicted responses will be investigated in relation to cardiometabolic outcomes in large population-based cohorts that include the same biomarkers but have not measured incretin and glucagon responses, including the UK Biobank and Fenland Study.

Research design and methods

To predict incretin and glucagon responses based on clinical and metabolite data, we will apply a two-layered modeling approach in the Fenland study. Clinical and metabolic profiles will be used separately and in combination to predict responses. For variable selection in both layers, we will apply two strategies: restricting variance and using linear LASSO regression to shrink and select the most predictive variables. To support this, we will employ machine learning models to capture undefined interactions and identify important predictors based on variable importance[40]. Our approach follows established principles in machine learning and statistical modeling, adhering to standardized practices for prediction, reporting, and validation[41–44].

As part of the collaboration with the MRC Epidemiology Unit and Novo Nordisk Foundation Center for Basic Metabolic Research, current efforts are focused on developing PRS for GLP-1, GIP, and glucagon. These PRSs will be incorporated into the aforementioned predictive layers as composite predictors of incretin and glucagon responses. PRS for other relevant traits such as obesity, insulin resistance, beta-cell function, liver function and low-grade inflammation will also be considered where relevant.

Research design and methods

	Fenland study	**ADDITION-PRO Study**
Baseline data collection period	2005-2015 (Phase 1-3)	2009–2011
Included participants	UK general population aged 30-55 years	Individuals at high risk of diabetes
Number of participants	12,435	2,082
Measure from OGTT (t= 0, 30, 120 min):	GLP-1, GIP, glucagon, glucose and insulin	GLP-1, GIP, glucagon, glucose and insulin

	Fenland study	**ADDITION-PRO Study**
Other relevant metabolic measures	DEXA-scan	Liver function: • Ultrasound liver images (still B-mode images with liver protocol) • GGT, Alanine aminotransferase Inflammation: • soluble CD163, adiponectin, and high sensitivity C-reactive protein
Assessment method of metabolites	LC electrospray ionization and flow-injection analysis tandem MS (ref), targeted metabolomics	Proton nuclear magnetic resonance spectroscopy, targeted metabolomics[31]
Sample tissue	Fasting plasma blood samples	Both fasting and during the OGTT[31,32]
Number of metabolites	175 (acylcarnitines, amines, sphingolipids and phospholipids)	231 lipid-related and 3 BCAA (fasting isoleucine, leucine and valine levels)
Genotyping	Yes	Yes
Follow-up period	NA	13-years

	Fenland study	**ADDITION-PRO Study**
Outcome	Metabolic-related outcomes: • Progression to type 2 diabetes Macrovascular complications: • Ischemic-related cardiovascular disease • Heart failure Microvascular complications: • Chronic kidney disease	Metabolic-related outcomes: • Progression to type 2 diabetes • Regression to normoglycemia • HbA1c, lipids and eGFR trajectories Macrovascular complications: • Ischemic-related cardiovascular disease • Heart failure Microvascular complications: • Chronic kidney disease • Neuropathy

To extend findings and predicted characterizations in larger population-based cohorts, will use ADDITION-PRO Study and UK Biobank[45] and as validation cohorts.

Collaboration

Synergy with the MRC Epidemiology Unit

The synergy between this project and the MRC Epidemiology Unit lies in the integration of deep-phenotyped metabolic data with large-scale population health surveillance. While my work in Denmark leverages the specialized ADDITION-PRO cohort, the MRC Epidemiology Unit provides an unparalleled environment for precision epidemiology. By combining the Unit's global leadership in studying the genetic and environmental determinants of obesity and type 2 diabetes with my focus on natural incretin and glucagon responses, we create a powerful framework to move from observational data to causal understanding. This collaboration allows for the cross-pollination of Danish clinical depth and the UK's extensive population-based datasets, ensuring that findings regarding metabolic drivers are both biologically robust and representative of the broader population.

The Role of the Collaborator

The MRC Epidemiology Unit, specifically through the mentorship of experts in the Fenland Study including Professor Nick Wareham and Professor Simon Griffin, will play a critical role as the primary international host and scientific advisor. Their role is three-fold:

Data Provision and Oversight: Providing access to the Fenland Study (n=12,435), which is uniquely suited for this project due to its objective measures of metabolic function and long-term follow-up.

Methodological Mentorship: Facilitating my training in causal inference and Mendelian Randomization, specifically using genetic instruments to determine the causal role of incretin and glucagon in cardiometabolic outcomes.

Validation Expertise: Assisting in the harmonisation of data between the Fenland Study, UK Biobank, and ADDITION-PRO to ensure that the identified metabolic profiles and mediating pathways (such as liver health and inflammation) are validated across different geographical and demographic contexts.

Ongoing Synergy and Knowledge Exchange

The synergy with the MRC Epidemiology Unit is not limited to a single research stay; it is an ongoing institutional bridge. My research is strategically aligned with the Unit's mission to improve population health through better risk stratification.

Continuous Methodological Alignment: We will maintain regular scientific exchange through joint virtual meetings to align the analytical protocols used in the Fenland Study with the DP-Next (Work Package 3) framework.

Future Frameworks: This collaboration serves as a pilot for long-term data integration between Danish and UK cohorts, establishing a permanent channel for investigating the heterogeneity of type 2 diabetes and refining precision prevention strategies across borders.

LAY PROJECT DESCRIPTION

1 Wagner R, Heni M, Tabák AG, *et al.* Pathophysiology-based subphenotyping of individuals at elevated risk for type 2 diabetes. *Nature Medicine* 2021; **27**: 49–57.

2 Tabák AG, Herder C, Rathmann W, Brunner EJ, Kivimäki M. Prediabetes: A high-risk state for diabetes development. *The Lancet* 2012; **379**: 2279–90.

3 Birkenfeld AL, Franks PW, Mohan V. Precision medicine in people at risk for diabetes and atherosclerotic cardiovascular disease: A fresh perspective on prevention. *Circulation* 2024; **150**: 1910–2.

4 Kahn SE, Deanfield JE, Jeppesen OK, *et al.* Effect of semaglutide on regression and progression of glycemia in people with overweight or obesity but without diabetes in the SELECT trial. *Diabetes Care* 2024; **47**: 1350–9.

5 Jastreboff Ania M., Roux Carel W. le, Stefanski Adam, *et al.* Tirzepatide for obesity treatment and diabetes prevention. *New England Journal of Medicine* 2025; **392**: 958–71.

6 Jastreboff Ania M., Kaplan Lee M., Frías Juan P., *et al.* Triple–hormone-receptor agonist retatrutide for obesity — a phase 2 trial. *New England Journal of Medicine* 2023; **389**: 514–26.

7 Sanyal AJ, Kaplan LM, Frias JP, *et al.* Triple hormone receptor agonist retatrutide for metabolic dysfunction-associated steatotic liver disease: A randomized phase 2a trial. *Nature Medicine* 2024; **30**: 2037–48.

8 Lincoff A. Michael, Brown-Frandsen Kirstine, Colhoun Helen M., *et al.* Semaglutide and cardiovascular outcomes in obesity without diabetes. *New England Journal of Medicine* 2023; **389**: 2221–32.

9 NAUCK MA, HOMBERGER E, SIEGEL EG, *et al.* Incretin effects of increasing glucose loads in man calculated from venous insulin and c-peptide responses*. *The Journal of Clinical Endocrinology & Metabolism* 1986; **63**: 492–8.

10 Hædersdal S, Andersen A, Knop FK, Vilsbøll T. Revisiting the role of glucagon in health, diabetes mellitus and other metabolic diseases. *Nature Reviews Endocrinology* 2023; **19**: 321–35.

11 Færch K, Torekov SS, Vistisen D, *et al.* GLP-1 response to oral glucose is reduced in prediabetes, screen-detected type 2 diabetes, and obesity and influenced by sex: The ADDITION-PRO study. *Diabetes* 2015; **64**: 2513–25.

12 Færch K, Vistisen D, Pacini G, *et al.* Insulin resistance is accompanied by increased fasting glucagon and delayed glucagon suppression in individuals with normal and impaired glucose regulation. *Diabetes* 2016; **65**: 3473–81.

13 Carrasco-Zanini J, Pietzner M, Lindbohm JV, *et al.* Proteomic signatures for identification of impaired glucose tolerance. *Nature Medicine* 2022; **28**: 2293–300.

14 De Silva NMG, Borges MC, Hingorani AD, *et al.* Liver function and risk of type 2 diabetes: Bidirectional mendelian randomization study. *Diabetes* 2019; **68**: 1681–91.

15 Wittemans LBL, Lotta LA, Oliver-Williams C, *et al.* Assessing the causal association of glycine with risk of cardio-metabolic diseases. *Nature Communications* 2019; **10**: 1060.

16 Lotta LA, Pietzner M, Stewart ID, *et al.* A cross-platform approach identifies genetic regulators of human metabolism and health. *Nature Genetics* 2021; **53**: 54–64.

17 Agrawal S, Luan J, Cummings BB, Weiss EJ, Wareham NJ, Khera AV. Relationship of fat mass ratio, a biomarker for lipodystrophy, with cardiometabolic traits. *Diabetes* 2024; **73**: 1099–111.

18 Bowker N, Shah RL, Sharp SJ, *et al.* Meta-analysis investigating the role of interleukin-6 mediated inflammation in type 2 diabetes. *EBioMedicine* 2020; **61**: 103062.

19 Witte DR, Johnston L, Røikjer J, Rasmussen N, Juhl CB. Work package 3: heterogeneity. 2025; published online Aug 19. https://dp-next.github.io/wp3.html.

20 Koopman ADM, Rutters F, Rauh SP, *et al.* Incretin responses to oral glucose and mixed meal tests and changes in fasting glucose levels during 7 years of follow-up: The hoorn meal study. *PLOS ONE* 2018; **13**: e0191114.

21 Johansen NB, Hansen AL, Jensen TM, *et al.* Protocol for ADDITION-PRO: a longitudinal cohort study of the cardiovascular experience of individuals at high risk for diabetes recruited from Danish primary care. *BMC Public Health* 2012; **12**: 1078.

22 Davey Smith G, Hemani G. Mendelian randomization: Genetic anchors for causal inference in epidemiological studies. *Human Molecular Genetics* 2014; **23**: R89–98.

23 MRC Epidemiology Unit. Fenland study. https://studies.mrc-epid.cam.ac.uk/fenland.

24 Johnston L, Wittenbecher C. Inference of causal links between metabolomics and disease incidence. 2020. https://github.com/NetCoupler/NetCoupler.

25 VanderWeele T. Explanation in causal inference: Methods for mediation and interaction. Oxford University Press, 2015.

26 Winther-Sørensen M, Galsgaard KD, Santos A, *et al.* Glucagon acutely regulates hepatic amino acid catabolism and the effect may be disturbed by steatosis. *Molecular Metabolism* 2020; **42**: 101080.

27 Stefan N, Yki-Järvinen H, Neuschwander-Tetri BA. Metabolic dysfunction-associated steatotic liver disease: Heterogeneous pathomechanisms and effectiveness of metabolism-based treatment. *The Lancet Diabetes & Endocrinology* 2025; **13**: 134–48.

28 Hammoud R, Drucker DJ. Beyond the pancreas: Contrasting cardiometabolic actions of GIP and GLP1. *Nature Reviews Endocrinology* 2023; **19**: 201–16.

29 Powell R, De Lucia Rolfe E, Day FR, *et al.* Development and validation of total and regional body composition prediction equations from anthropometry and single frequency segmental bioelectrical impedance with DEXA. *medRxiv* 2020; : 2020.12.16.20248330.

30 Deichgræber P, Witte DR, Møller HJ, *et al.* Soluble CD163, adiponectin, c-reactive protein and progression of dysglycaemia in individuals at high risk of type 2 diabetes mellitus: The ADDITION-PRO cohort. *Diabetologia* 2016; **59**: 2467–76.

31 Mahendran Y, Jonsson A, Have CT, *et al.* Genetic evidence of a causal effect of insulin resistance on branched-chain amino acid levels. *Diabetologia* 2017; **60**: 873–8.

32 Buckley MT, Racimo F, Allentoft ME, *et al.* Selection in europeans on fatty acid desaturases associated with dietary changes. *Molecular Biology and Evolution* 2017; **34**: 1307–18.

33 Gnatiuc Friedrichs L, Trichia E, Aguilar-Ramirez D, Preiss D. Metabolic profiling of MRI-measured liver fat in the UK biobank. *Obesity* 2023; **31**: 1121–32.

34 Vogelzangs N, Kallen CJH van der, Greevenbroek MMJ van, *et al.* Metabolic profiling of tissue-specific insulin resistance in human obesity: Results from the diogenes study and the maastricht study. *International Journal of Obesity* 2020; **44**: 1376–86.

35 Beyene HB, Hamley S, Giles C, *et al.* Mapping the associations of the plasma lipidome with insulin resistance and response to an oral glucose tolerance test. *The Journal of Clinical Endocrinology & Metabolism* 2020; **105**: e1041–55.

36 Madsen AL, Bonàs-Guarch S, Gheibi S, *et al.* Genetic architecture of oral glucose-stimulated insulin release provides biological insights into type 2 diabetes aetiology. *Nature Metabolism* 2024; **6**: 1897–912.

37 Zhou Y, Chen H, Iao SI, *et al.* Fdapace: Functional data analysis and empirical dynamics. 2024 https://CRAN.R-project.org/package=fdapace.

38 Healy J, McInnes L. Uniform manifold approximation and projection. *Nature Reviews Methods Primers* 2024; **4**: 82.

39 Cohn R, Holm E. Unsupervised machine learning via transfer learning and k-means clustering to classify materials image data. *Integrating Materials and Manufacturing Innovation* 2021; **10**: 231–44.

40 Dietrich S, Floegel A, Troll M, *et al.* Random survival forest in practice: A method for modelling complex metabolomics data in time to event analysis. *International Journal of Epidemiology* 2016; **45**: 1406–20.

41 Collins GS, Moons KGM, Dhiman P, *et al.* TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* 2024; **385**: e078378.

42 Lopez-Ayala P, Riley RD, Collins GS, Zimmermann T. Dealing with continuous variables and modelling non-linear associations in healthcare data: Practical guide. *BMJ* 2025; **390**: e082440.

43 Collins GS, Dhiman P, Ma J, *et al.* Evaluation of clinical prediction models (part 1): From development to external validation. *BMJ* 2024; **384**: e074819.

44 Riley RD, Archer L, Snell KIE, *et al.* Evaluation of clinical prediction models (part 2): How to undertake an external validation study. *BMJ* 2024; **384**: e074820.

45 Bycroft C, Freeman C, Petkova D, *et al.* The UK biobank resource with deep phenotyping and genomic data. *Nature* 2018; **562**: 203–9.