CoDiet

COMBATTING DIET RELATED NON-COMMUNICABLE DISEASE THROUGH ENHANCED SURVEILLANCE

D4.1 Legal and ethical aspects of personalised nutrition

Deliverable number D4.1

Work Package WP4	Causally-Informed Machine Learning Models and Personalized Intervention Recommendation
Task 4.4	Legal and Ethical Consideration
Task Leader	Czech Technical University
Prepared by	Czech Technical University
Contributors	
Version	1.1
Delivery Date	06/10/2024

Co-funded by the European Union
UK Research and Innovation

Foreword

The work described in this report was developed under the project **CoDiet - Combatting Diet related non-communicable disease through enhanced surveillance** (Grant Agreement number: 101084642; Call: HORIZON-CL6-2022-FARM2FORK-01; Topic: HORIZON-CL6-2022-FARM2FORK-01-10). Any additional information, if needed, should be required to:

Project Coordinator:
Itziar Tueros – itueros@azti.es | AZTI |

WP4 Leader: Technion
Shie Mannor – shie@technion.ac.il | TEC |

Task Leader: Czech Technical University
Jakub Marecek – jakub.marecek@fel.cvut.cz| CTU |

Dissemination Level		
PU	Public, fully open	X
SEN	Sensitive, limited under the conditions of the Grant Agreement	
Classified R-UE/EU-R	EU RESTRICTED under the Commission Decision No2015/444	
Classified C-UE/EU-C	EU CONFIDENTIAL under the Commission Decision No2015/444	
Classified S-UE/EU-S	EU SECRET under the Commission Decision No2015/444	

Co-funded by the European Union

CoDiet is part of the Horizon Europe programme supported by the European Union.

Document History

- (June 29th, 2024) Version 1.0 Submitted to the EC and uploaded to CoDiet website.
- (October 6th, 2024) Version 1.1 Submitted to the EC and uploaded to CoDiet website. In response to a comment asking about AI Explainability, we have added a discussion on pages 8 and 12. We have also updated the description of inclusion criteria on page 6.

Executive Summary

Here, we document our work on ethical and legal aspects of the work, capturing joint work of teams at CTU (Jakub Marecek), Imperial College Dublin (esp. the "protocol team" incl. Gary Frost, Aygul Dagbasi, and Monica Hill), NKUA (Dimitrios Gunopulos, Vana Kalogeraki, Kleopatra Markou), and the Technion (Shie Mannor, Mark Kozdoba).

1 Introduction

While our initial study has undergone a rigorous review of their ethical aspects by multiple ethics review boards, and we discuss the ethical and legal aspects in deliverables:

- Deliverable 3.1: Data management for data collected within the project, `https://github.com/codiet-eu/d31/blob/main/CoDiet_DMP_deliverable_3.1_version_1.0.pdf` ("Data Management Plan"), and its updated versions including Deliverable number D3.3 Data Management Plan Update RP1,
- Deliverables 9.1-9.4 in the Work Package 9 ("Ethics Requirements"),

there are several aspects that may be worth commenting on.

The Multiple Perspectives First, it may be worth explaining the multiple perspectives on the ethical aspects. Even within the regulation of the European commission, there are several, including:

- a very abstract perspective of the Regulation (EU) 2021/695 of the European Parliament and of the Council of 28 April 2021 establishing Horizon Europe – the Framework Programme for Research and Innovation, laying down its rules for participation and dissemination, and repealing Regulations (EU) No 1290/2013 and (EU) No 1291/2013 ("Horizon Europe regulation" or "HE regulation"). The HE regulation requires ethical principles to be upheld in Article 19. Article 19, among others, requires the research to be non-disciminatory.
- a very specific rules of the Regulation (EU) No 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC Text with EEA relevance ("Clinical trials directive"). See [7, 18] for the earlier directives. There, Article 18 delegates the regulation of ethics committees to the member states, but Article 29 gives universities rights to perform research utilizing data from controlled trials beyond the stated protocol, whenever an explicit written consent is obtained.
- ethics of the use of AI, governed by Regulation (EU) of the European Parliament and of the Council laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 ("Artificial Intelligence Act", "AI Act"). There, Articles 7 and following refer to the Principles for the Digital Decade and the Ethics guidelines for trustworthy AI of the High-Level Expert Group on Artificial Intelligence (AI HLEG).
- data protection perspective of the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC ("General Data Protection Regulation", "GDPR"). There, Article 22 requires explicit consent for automated individual decision-making, including profiling, which could be relevant in the case of diet recommendations, among others.
- general healthcare data perspective of Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 ("Data Governance Act")
- Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 ("Data Act")

while these are all part of the harmonized, there is a slightly different focus in each. There are also perspectives of the European Commission Ethics Self-Assessment Guidelines, referenced in the Consortium Agreement and Deliverable D3.1, and perspectives introduced by the external ethics advisor in Work Package 9.

Second, it may be worth providing a holistic perspective on the three key steps within the CODIET project: the initial study, the randomized controlled trial, and the exploitation of the results. This holistic perspective is elaborated upon in the following three sections.

2 Legal and Ethical Aspects of the Initial Study

The initial study is observational, which reduces the risks involved. Considering the study uses an enhanced surveillance method, we we utilise wearable micro-cameras and activity monitors to record dietary intake, physical activity, and sleep, there are a number of ethical questions related to the data collection, transfer, and reanalysis.

2.1 Study Population

An interesting ethical aspect involves inclusion criteria. Although it seems desirable (and mandated by the "Horizon Europe" regulation in Article 19) to be non-disciminatory, the inclusion criteria need to be discriminatory, to an extent, in order to obtain meaningful results. In particular, in the initial study, we are recruiting healthy adults at high risk of developing non-communicable disease. Participants eligible for inclusion in this study must be between 18 and 65 years of age (inclusive) and have a Body Mass Index (BMI) greater than or equal to 25 kg/m^2. In addition, they must meet at least one of the following five criteria (no priority in any factor):

- triglycerides levels of $\geq$ 150 mg/dL (1.7 mmol/L),
- HDL cholesterol levels less than 40 mg/dL (1.03 mmol/L) for males or less than 50 mg/dL (1.29 mmol/L) for females,
- blood pressure with a systolic reading of $\geq$ 130 mm Hg or a diastolic reading of $\geq$ 85 mm Hg,
- fasting plasma glucose levels of $\geq$ 90 mg/dL (5.0 mmol/L), or
- being a current smoker.

Individuals with controlled hypertension or dyslipidaemia with medication will be included in the trial. Moreover, potential participants must demonstrate a willingness and ability to sign the informed consent form, along with an understanding and compliance with the study participation requirements. Conversely, individuals will be excluded from the study if they have type 2 diabetes, chronic gastrointestinal conditions such as Crohn's disease, irritable bowel syndrome, or ulcerative colitis, acute infectious diseases, cancer, cardiovascular diseases, , autoimmune conditions, have undergone antibiotic treatment in the 12 weeks prior to enrolment, are pregnant or currently breastfeeding, are participating in another clinical trial, have participated in another clinical trial within the past 12 weeks, or are undergoing any medical intervention during the study period.

This raises several questions:

- could this be seen as discriminatory to populations that do not meet the inclusion criteria? This may include, for instance, persons with BMI strictly less than 25 kg/m^2.
- will the results of the initial study be representative of the general population?
- will the results of the initial study be useful to patients suffering from, e.g., cardiovascular diseases, which may be keen on utilizing the personalized diet and more likely to comply with the recommender diet?

The brief answers are "most likely not". The inclusion criteria have been developed with the hope of collecting sufficient number of samples to present useful answers at least to a certain at-risk population.

2.2 Data Collection

Data collection and data processing are governed by the provisions of the AI Act, where provisions of Article 55 and following establish the so-called 'AI regulatory sandbox', which may override GDPR and other data-protection measures for the purposes of research and development. In particular, the AI regulatory sandbox "shall provide a controlled environment that facilitates the development, testing and validation of innovative AI systems for a limited time". We refer to Deliverables D3.1 and D3.3 for further details.

At the same time, in processing personal data pursuant to the Consortium Agreement, each member shall:

- process, or permit to be processed, personal data only for the purposes of the performance of this Consortium Agreement and under a solid legal basis (Data Sharing Agreement);

- ensure that personnel are subject to an obligation of confidentiality in respect of the processing of personal data under this Consortium Agreement;
- ensure that appropriate technical and organisational measures shall be taken against unauthorised or unlawful processing of personal data and against accidental loss or destruction of, or damage to, personal data;
- not disclose or transfer personal data to any third-party other than where strictly necessary for the purposes of the performance of this Consortium Agreement; notify the other members of any security incidents, events, weaknesses, data breaches or suspected data breaches impacting or possibly impacting the security of personal data;
- be individually responsible for its own processing of personal data pursuant to and in connection with this Consortium Agreement.

Privacy and Reidentification Risks In the initial study, participants wear micro-cameras and activity monitors for 3 separate one-week periods at home (with a 2-3 week break in between each period). At the end of the first and last period of wearing the cameras and activity monitors each week, participants attend a study visit at a clinical research facility to provide blood, urine, stool, and breath samples, as well as undergo body composition, CVD and autonomic nervous system health analysis tests using a range of state-of-the-art technologies.

All potential participants must demonstrate a willingness and ability to sign the informed consent form, along with an understanding and compliance with the study participation requirements.

The collection of the samples of blood, urine, stool, and breath is rather standard and does not pose a major reidentification risk.

The collection of data from the microcameras poses a significant reidentification risk. Imperial College London has developed a toolkit for pseudonymization of the video data (recorded at 0.5 Hz). While there are sophisticated methodologies to estimate reidentification risks [20], it seems very hard to quantify these risks following pseudonymization.

2.3 Data Transfers

In accordance with the rules on transfers of health data in the light of GDPR [11], we transfer only the pseudonymized and encrypted data from the sites collecting the data to a data center hosted physically in the Czech Republic. Subsequently, data can be accessed by the consortium. Notice that from the point of the GDPR, this also considers a data transfer. In particular, "you do not actually have to 'send' the data to a non-EU country for these provisions to apply; if one of your partners or service providers is located outside the EU and is able to access the personal data you have collected, this amounts to a 'data transfer' in the context of the GDPR."

The involvement of non-EU countries poses a non-trivial complication. Since some of the CoDiet consortium members originate from countries outside the EU, the design principle of the procedures for data processing has carefully addressed this information in order to be compliant with the GDPR regulations as well as to carefully address how data are transferred (or not) to non-EU countries. In order to be compliant and lawful, we hereby state that our partners and beneficiaries (ICL and Technion, which originate from United Kingdom and Israel, respectively) have an 'adequacy determination' as defined by the European Commission and ensure the same level of data protection as is required under the EU law. Participants which are established in a non-EU country (if any) undertake to comply with their obligations under the Agreement and:

- to respect general principles (including fundamental rights, values and ethical principles, environmental and labour standards, rules on classified information, intellectual property rights, visibility of funding and protection of personal data)
- for the submission of certificates under Article 24: to use qualified external auditors which are independent and comply with comparable standards as those set out in EU Directive 2006/43/EC (Directive 2006/43/EC of the European Parliament and of the Council of 17 May 2006 on statutory audits of annual accounts and consolidated accounts or similar national regulations (OJ L 157, 9.6.2006, p. 87))
- for the controls under Article 25: to allow for checks, reviews, audits and investigations (including on-the-spot checks, visits and inspections) by the bodies mentioned in that Article (e.g., granting authority, OLAF, Court

of Auditors (ECA), etc.). Special rules on dispute settlement apply (See Consortium's Agreement Data Sheet Point 5).

2.4 Data Analysis

The data is only being processed within the justified interests of the project, fairly and lawfully, for a specific and legitimate purpose and only the data necessary to achieve this purpose will be processed, with personal data safety and security measures considered and upon authorization. The EU's Ethical Guidelines for AI systems, as references in the AI Act, state seven key dimensions to be evaluated and audited by a cross-disciplinary team, namely (a) human agency and oversight, (b) technical robustness and safety, (c) privacy and data governance, (d) transparency, (e) diversity, non-discrimination and fairness, (f) environmental and societal well-being and, (f) accountability. CoDiet's work directly addresses challenges in oversight, safety, transparency, fairness, and accountability.

- Human Agency and oversight: AI systems must support human autonomy and decision-making, enabling users to make informed autonomous decisions regarding the AI systems : The system and techniques developed through CoDiet are not aiming at replacing the human decision-making processes, but to enhance them, allowing for human agency and oversight of the system. CoDiet aims to ensure that the tools that are developed to enhance understanding of NCD risk, dietary monitoring and enhance personalised nutrition will be relevant to all sessions of society across Europe. It is recognised that there are health inequalities based on a range of determinant factors such as socioeconomic status, ethnic and cultural background, access to education, and sex/gender identities.

- Privacy and data governance: AI systems must guarantee privacy and data protection throughout the system's lifecycle. The system and techniques developed through CoDiet will use state-of-the-art privacy and security techniques to ensure data security and integrity.

- Transparency: All datasets and processes associated with AI decisions must be well communicated and appropriately documented. Any benchmarking datasets along with the respective code of the developed fairness-enhancing AI models will be thoroughly documented.

- Explainability: In our uses of AI, such as in three papers (labelled 2, 3, and 4), presented in Deliverable D4.2, we develop approaches that are *a priori* explainable:

  - the structure of the estimator, in the form of dynamic Bayesian network on dozens or hundreds of stochastic processes, is open to visual inspection and human overisight.
  - the approach to learning dynamic Bayesian networks, which we employ, makes it possible to incorporate additional constraints easily. Indeed, if there is human oversight, it needs to be able to easily provide additional constraints suggesting "what had been wrong", when the solutions fall short of the expectations.
  - we have guarantees that we obtain the (constrained) maximum likelihood estimate, which can be seen as the best possible estimate given the data. This is in sharp contrast to deep-learning approaches, where short of achieving training error of 0, such guarantees are not available.

- Fairness, diversity and non-discrimination: Best possible efforts should be made to avoid unfair bias. The AI system built within CoDiet should provide transparency in obtaining any reliable information on the fairness impact of different design measures, in order to empower the user to make decisions appropriately.

- Societal and environmental well-being: The impact of the developed and/or used AI system/technique on the individual, society and environment must be carefully evaluated and any possible risk of harm must be avoided : CoDiet project has established a set of precaution measures such as data minimization, data anonymization and purpose limitation steps, to ensure that the data will be used only for the purpose it is intended to. Producing targeted individualized recommendations that are sensitive to drivers of dietary change in vulnerable populations, will lead to improved awareness of the need for a healthier diet.

- Accountability: Requires that the actors involved in their development or operation take responsibility for the way that these applications function and for the resulting consequences : CoDiet project has considered a role-and-permission system to ensure that the actors involved in the development process of those systems,

as well as, their operation are accountable for the way these applications function and for their resulting consequences.

- Technical Robustness and Safety: AI systems must maintain their robustness, security and safety throughout their entire existence. The system and techniques developed through CoDiet, apart from the privacy and security measures considered, will also focus on the robustness of the results provided through the developed system.

We collaborate closely with Human-compatible AI with guarantees, the Horizon Europe project (humancompatible.org), who develop multiple open-source toolkits (incl. the AI Fairness 360 and AI Explainability 360) to address these Ethical Guidelines.

2.5 Data Reanalysis

Ethical and legal questions of data reanalysis are non-trivial. Within the Clinical trials regulation, Article 29 gives universities and other research institutions the right to perform research utilizing data from controlled trials beyond the stated protocol, whenever an explicit written consent is obtained. In particular, it suggests: "It is appropriate that universities and other research institutions, under certain circumstances that are in accordance with the applicable law on data protection, be able to collect data from clinical trials to be used for future scientific research, for example for medical, natural or social sciences research purposes. In order to collect data for such purposes it is necessary that the subject gives consent to use his or her data outside the protocol of the clinical trial and has the right to withdraw that consent at any time. It is also necessary that research projects based on such data are subject to reviews that are appropriate for research on human data, for example on ethical aspects, before being conducted." This allows for reanalysis by universities and other research institutions.

From the ethics perspective, there is a clear trade-off between the principles of open science, aiming to maximise the return on public investment in research, and data privacy. While there are sophisticated methodologies to estimate reidentification risks [20], we imagine that the video data may be too sensitive to share with third-party researchers, unless a strong legal framework for sharing the data is established.

3 Legal and Ethical Aspects of the Randomized Controlled Trial

There are a wide variety of philosophical [19] and ethical [9] questions to be answered in relation to randomised controlled trials in medical artificial intelligence (AI). Their importance is highlighted by recent recalls of software components of clinical applications of AI, such as the Class I recall for the t:connect mobile app by Tandem Diabetes Care, Inc., which controls the t:slim X2 insulin pump. There, US Food and Drug Administration (FDA) has documented hundreds of instances, where the software malfunctioned. While our randomized controlled trial within personalized nutrition arguably poses a lesser risk, there are several aspects worth commenting on.

3.1 Risks of Narrowly Personalized Diets are not Fully Knowable Ahead of Time

A key concern stems from the fact that the risks and benefits of narrowly personalized diets are not fully known ahead of time. Indeed, a reasonable performance in one participant, or one hundred thousand participants need not guarantee the same performance in the next participant. In the randomized controlled trial, there will not be hundred thousand prior participants. Instead, we offer

- the treatment group a personalized diet that satisfies a variety of a priori constraints in the spirit of safe reinforcement learning [10]
- the control group the mediterranean diet, a well understood diet [16], widely accepted as one of the healthiest options available. Other studies [6] have used US diet recommendations, for instance.

There may be risks associated with a personalized diet. For example, an allergy unknown to the participant prior to the randomized controlled trial could trigger a reaction to a complex meal recommended in the personalized diet. With complex meals consumed outside of the home, such an allergy reaction is near impossible to rule out.

More likely, the personalized diet recommended in the randomized controlled trial may be imperfect in other ways. Either the participant does not enjoy the suggestions, and eats less or does not comply with the recommendations.

Alternatively, the participant is enjoying the recommendations too much (or misunderstands the personalized diet for a certainty of weight loss) and does not follow the recommendations regarding the meal sizes. As in most clinical trials, the opportunities and risks involved are not fully known ahead of time.

3.2 Segmentation and the Degree of Personalization in the Training

In personalized medicine and nutrition, we also face the question as to how personalized should training data be. In general, some form of data-driven population segmentation analysis [21] is known to be beneficial, but the exact nature is changing. In the "foundational models" (e.g., large language models such as GPT4) literature [2, e.g.], the segmentation is typically obscured by the fact that one trains a foundational model on all available data, and then personalizes the model in a process known as fine-tuning. It is well known [13, e.g.] that fine-tuning works reasonably well for in-distribution data and rather poorly for out-of-distribution data.

The personalized approach requires training or fine-tuning a diet recommender on data from a relevant subpopulation. There:

- when one trains or fine-tunes the model on too large a subpopulation, it makes it less targeted and may work less well for very small subpopulations. In the extreme, one could consider, for example, one subpopulation per country. While this would allow for considerations related to the food environment and common complex meals, it would surely obscure gender differences, differences among phenotypes, etc.

- when one trains or fine-tunes the model on too small a subpopulation, one may overfit the model and limit the recommendations to too small a set. In the extreme, one could consider data from one participant in the initial study, and never recommend food items in the randomized controlled trial outside of those eaten by the same participant within the initial study. If the a priori rules discarded some of the food items from the initial study, this could lead to a very monotone diet.

In the CoDiet setting, this is further complicated by the fact that there are many subpopulations to consider, and the subpopulations to use may need to be learned [17, cf.] at the same time as one learns the models, so as to minimize the maximum empirical risk across the models for each subpopulation. In this setting, the computational complexity of such joint problems is non-trivial and many joint problems can be shown to be NP-Hard. This, in turn, may require reasoning about polynomial-time algorithms without strong guarantees on their statistical performance.

3.3 Exploitation and Exploration Tradeoff

Across personalized medicine and nutrition, we also face an exploitation-exploration tradeoff [1], or perhaps some form of exploitation-exploration-compliance tradeoff. In the literature on exploitation-exploration tradeoff, exploitation utilizes (an often small) empirical distribution to pick the best action to take. In our case, this could be the recommendation satisfying a variety of a priori constraints in the spirit of safe reinforcement learning, which maximizes the expected compliance with the recommendation, or the recommendation that maximizes some function of the expected compliance with the recommendation and expected risks and benefits of the recommendation. For instance, once the recommender learns that the user is willing and able to eat watermelon in large quantities, it may keep recommending it very often.

In contrast, exploration aims to expand the support of the empirical distribution of what has been observed so far. For example, if we have never recommended raw kale, the system may recommend eating raw kale to learn whether the user would comply with such a recommendation and what the health risks and benefits would have been. In this particular case, there are many well-documented health benefits of eating kale for the majority population. There are also rare cases where a compound called goitrin in raw kale can affect thyroid function, which poses a risk in a very small subpopulation. Even a well-designed safe reinforcement learning system may be oblivious to such a risk of raw kale. Perhaps more importantly, if the user is asked a couple of times to eat raw kale to estimate their compliance with raw kale recommendations, they may discontinue the randomized controlled trial.

The balancing of the exploitation-exploration tradeoff is a challenging problem. It is well known that one cannot stop expanding the support of the empirical distribution (for a given user or subpopulation), but one also cannot stop making use of what has been collected so far, in order to ensure (asymptotic) optimality of the algorithm. Especially in the case when the feedback obtained on various recommendations is not independent (also known as graph-structured feedback), balancing of the exploitation-exploration tradeoff is an active area of

research [15, e.g.]. Having said that, the Technion team conduct world-leading research [15, 8, e.g.] in handling the exploitation-exploration tradeoff, and the methods developed at Technion will inform our work in this respect.

3.4 Data Collection

Data collection in the randomized controlled trial [11] faces similar challenges as in the initial study. See Section 2.2. In addition, we collect data on the recommendations made, and compliance with the suggestions. We will update this document once the protocol for the randomized controlled trial has been developed in Work Package 5.

3.5 Data Transfers

Data transfers in the randomized controlled trial face similar challenges as in the initial study. See Section 2.3. As above, in accordance with the rules on transfers of health data in the light of GDPR [11], we transfer only the pseudonymized and encrypted data from the sites collecting the data to a data center hosted physically in the Czech Republic. The use of the data by researchers at Imperial College London and the Technion constitutes a data transfer.

3.6 Data Analysis

Data analysis in the randomized controlled trial face similar challenges as in the initial study. See Section 2.4. We will update this document once the protocol for the randomized controlled trial has been developed in Work Package 5.

3.7 Data Reanalysis

Data reanalysis in the randomized controlled trial face similar challenges as in the initial study. See Section 2.5. As above, while there are sophisticated methodologies to estimate reidentification risks [20], we imagine that the video data may be too sensitive to share with third-party researchers, unless a strong legal framework for sharing the data is established. We will update this document once the protocol for the randomized controlled trial has been developed in Work Package 5.

4 Legal and Ethical Aspects of the Exploitation of the Results

In multi-modal AI in medical applications [3], including personalized medicine, digital clinical trials, remote monitoring and care, pandemic surveillance, digital twin technology, and virtual health assistants, one utilizes large datasets of data donated by participants, either to governments, national health systems, biobanks, or individual research projects run by universities or their consortia, and only in a very few cases, donated to companies or purchased by companies. Still, companies affiliated to the univerities (or not) may be keen on exploiting the data for commercial product development, and to utilize publicly funded research in establishing their credibility. This opens up a variety of important questions.

4.1 Communicating the Recommendations

In the case of a clinical trial, it is relatively easy to communicate the recommendations as the work of a thus far untested (in a randomized controlled trial) AI system, and there are wide-ranging exceptions for research across both the AI Act and GDPR.

In contrast, in commerical markets, there is commercial pressure to establish the recommendations as authoritative, e.g., by suggesting high success rate of the personalized diet in earlier studies, or claims about the safety of the recommendations, which can hardly be construed as perfect. In any exploitation activities, we aim to be transparent about the statistical performance of the system, and to present a warning label similar to the "Past performance is no guarantee of future results" known from financial services.

AI systems placed in commercial markets also need to follow the explainability requirements of the GDPR and AI Act, which can be quite non-trivial [12]. While the teams at NKUA and CTU contribute to the leading open-source toolkit in AI explainability (Linux Software Foundations's AIX 360 `https://github.com/Trusted-AI/AIX360`), some of the challenges are inherently difficult, unless one utilizes causal forecasting. For instance, if a recommendation

is made using a long history of compliance with previous recommendations, using a deep-learning model, it may be hard to explain succinctly. One could either truncate the explanation, rendering it less valid, or one could present the full explanation, making it unlikely that a human being would be able and willing to read and understand the explanation in full. In contrast, as we explain on page 8, we strive to develop (WP3) and use (WP4) causal models, from the initial study onwards. This is also reflected in the name of WP4 (Causally-informed ML models).

4.2 Monetization of Data Donated by Participants

While on the legal side, participants donate the data for well-specified uses [14] and possibly research purposes beyond, ethical questions related to data ownership [4] are of particular importance in exploitation of the results of the project:

- what is the legal status of AI models or systems derived from data donated for research purposes? Can the data be the basis for commercial product development? The usual notion of derivative work of copyright law does not seem to answer these questions fully.
- can the donation of biomedical data for commercial product development be seen as one-sided contract [5]? Even if there is less implicit pressure within personalized nutrition than within personalized medicine, where patients may be reluctant to refuse informed consent to data use to medical professionals in charge of their care, contracts can be one-sided due to information asymmetry [5], market failure, or similar.
- can one provide an informed consent to a commercial use of an object with an unknown value? If, for example, the data subject is told their genome is to become an open-science data collection, but then a private corporation sells a model trained on this collection for ten billion dollars, the data subject may be aggrieved.
- what if the value is mis-estimated? If, for example, the data subject is told that the market value of his genome is $ 1000, and thus he would be paid $ 1000, but then the company is able to resell a million genomes for $ 10000 each, or to sell its own shares to investors for ten billion dollars, having collected a million genomes, does this violate the informed consent?

4.3 Monetization of Publicly-Funded Research

A broader question concerns the monetization of publicly-funded research. Some researchers choose to set up start-ups to monetize the research, such as Professor Tim Spector of King's College London and team [6] starting Zoe Ltd. (`https://zoe.com/`). In some ways, this is laudable: Zoe Ltd. claims to have 125,000 members as of June 2024. These members seem to benefit from the research undertaken at King's College London, with funding from the UK government. On the other hand, it is not clear, whether a sufficient share of any future earnings goes to the participants of the original twin study, King's College London, the Department of Twin Research & Genetic Epidemiology therein, or the UK taxpayers. Other researchers prefer to work with large corporations (such as Nestle, `https://www.nestlenutrition-institute.org/`) in a variety of forms, which poses similar, if not the same questions.

Ethically, this presents a conundrum that goes well beyond personalized medicine. If the results of publicly-funded research were not used, ever, one may argue that the funds were wasted. On the other hand, if the profits from the use of publicly-funded research go primarily to a single private corporation, no matter the size, one may wonder if there is equal access to the results to all corporations.

References

[1] Naoki Abe and Philip M Long. Associative reinforcement learning using linear probabilistic concepts. In *ICML*, pages 3–11. Citeseer, 1999.

[2] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.

[3] Julián N Acosta, Guido J Falcone, Pranav Rajpurkar, and Eric J Topol. Multimodal biomedical ai. *Nature Medicine*, 28(9):1773–1784, 2022.

[4] Angela Ballantyne. How should we think about clinical data ownership? *Journal of medical ethics*, 46(5):289–294, 2020.

[5] Lucian A Bebchuk and Richard A Posner. One-sided contracts in competitive consumer markets. *Mich. L. Rev.*, 104:827, 2005.

[6] Kate M. Bermingham, Inbar Linenberg, Lorenzo Polidori, Francesco Asnicar, Alberto Arrè, Jonathan Wolf, Fatema Badri, Hannah Bernard, Joan Capdevila, William J. Bulsiewicz, Christopher D. Gardner, Jose M. Ordovas, Richard Davies, George Hadjigeorgiou, Wendy L. Hall, Linda M. Delahanty, Ana M. Valdes, Nicola Segata, Tim D. Spector, and Sarah E. Berry. Effects of a personalized nutrition program on cardiometabolic health: a randomized controlled trial. *Nature Medicine*, 2024.

[7] European Commission. Consolidated text: Regulation (eu) no 536/2014 of the european parliament and of the council of 16 april 2014 on clinical trials on medicinal products for human use, and repealing directive 2001/20/ec. *Official Journal*, L 294:5, 2022.

[8] Yonathan Efroni, Shie Mannor, and Matteo Pirotta. Exploration-exploitation in constrained mdps. *arXiv preprint arXiv:2003.02189*, 2020.

[9] Thomas Grote. Randomised controlled trials in medical ai: ethical considerations. *Journal of Medical Ethics*, 48(11):899–906, 2022.

[10] Shangding Gu, Long Yang, Yali Du, Guang Chen, Florian Walter, Jun Wang, Yaodong Yang, and Alois Knoll. A review of safe reinforcement learning: Methods, theory and applications. *arXiv preprint arXiv:2205.10330*, 2022.

[11] Johan Hansen, Petra Wilson, Eline Verhoeven, Madelon Kroneman, Mary Kirwan, Robert Verheij, and Evert-Ben van Veen. Assessment of the eu member states' rules on health data in the light of gdpr. 2021.

[12] Joshua Hatherley, Robert Sparrow, and Mark Howard. The virtues of interpretable medical artificial intelligence. *Cambridge Quarterly of Healthcare Ethics*, pages 1–10, 2022.

[13] Ananya Kumar, Aditi Raghunathan, Robbie Jones, Tengyu Ma, and Percy Liang. Fine-tuning can distort pretrained features and underperform out-of-distribution. In *International Conference on Learning Representations*, 2022.

[14] Kathleen Liddell, David A Simon, and Anneke Lucassen. Patient data ownership: who owns your health? *Journal of Law and the Biosciences*, 8(2):lsab023, 2021.

[15] Shie Mannor and Ohad Shamir. From bandits to experts: On the value of side-observations. *Advances in Neural Information Processing Systems*, 24, 2011.

[16] Miguel A Martínez-González, Jordi Salas-Salvadó, Ramón Estruch, Dolores Corella, Montse Fitó, Emilio Ros, Predimed Investigators, et al. Benefits of the mediterranean diet: insights from the predimed study. *Progress in cardiovascular diseases*, 58(1):50–60, 2015.

[17] Mengjia Niu, Xiaoyu He, Petr Rysavy, Quan Zhou, and Jakub Marecek. Joint problems in learning multiple dynamical systems. *arXiv preprint arXiv:2311.02181*, 2023.

[18] Carlo Petrini. Regulation (eu) no 536/2014 on clinical trials on medicinal products for human use: an overview. *Annali dell'Istituto superiore di sanità*, 50:317–321, 2014.

[19] Sridhar Venkatapuram and Alex Broadbent. *The Routledge handbook of philosophy of public health*. Routledge, 2023.

[20] Weiyi Xia, Melissa Basford, Robert Carroll, Ellen Wright Clayton, Paul Harris, Murat Kantacioglu, Yongtai Liu, Steve Nyemba, Yevgeniy Vorobeychik, Zhiyu Wan, and Bradley A Malin. Managing re-identification risks while providing access to the All of Us research program. *Journal of the American Medical Informatics Association*, 30(5):907–914, 02 2023.

[21] Shi Yan, Yu Heng Kwan, Chuen Seng Tan, Julian Thumboo, and Lian Leng Low. A systematic review of the clinical application of data-driven population segmentation analysis. *BMC medical research methodology*, 18:1–12, 2018.