

試料・情報利用研究計画書(概要)						
審査委員会 受付番号	2020-2003	利用形態	内部研究	利用する 試料・情報	対象: 三世代コホート調査参加者で、症例(低出生体重)約2,000名の児とその父母(計6,000名)、対照18,000名の児とその父母(計54,000名) 試料: なし 情報: 調査票情報、カルテ転記情報、検体検査情報、ゲノム情報	
主たる研究機関	東北メディカル・メガバンク機構			分担 研究機関	—	
研究題目	低出生体重の発症リスク予測のための機械学習モデルの開発				研究期間	2020年7月～2025年3月
実施責任者	荻島 創一	所属	東北メディカル・メガバンク機構			職位 教授
研究目的と意義	産科における主要な疾患の一つである低出生体重の発症リスク予測のため、調査票の生活習慣情報および検体検査情報、時系列の胎児超音波情報を統合し、組み合わせの特徴を考慮する機械学習モデルの開発を目的とします。					
研究計画概要	低出生体重は全新生児の約10%に見られ、幼児期の発育・発達の遅れだけでなく、若年青年期の心疾患、糖尿病、高血圧など広範囲の疾患リスクを向上させる、健康を守る安心社会の実現のために対策が急務な疾患です。低出生体重はこれまでの研究で、母親の栄養状態、合併症、喫煙歴などに加え、両親および児の遺伝因子が複雑に関係する多因子疾患であることが明らかになる一方で、有効な予測方法の開発が困難である状況が続いています。そのため、これまで検討されてきた少数の因子の組み合わせでなく、患者の生活習慣情報やゲノム情報、時系列の臨床情報を統合した機械学習による、低出生体重の早期リスク予測への応用が期待されています。そこで、本研究では、低出生体重の早期診断支援を目標に、三世代コホート調査で収集された母親・父親の調査票による生活習慣情報および検体検査情報、時系列の胎児超音波情報、ゲノム情報の多様な情報を統合し、組み合わせの特徴を考慮する機械学習モデルの検討を行います。 これまでに妊娠高血圧症候群の早期診断支援のための機械学習の開発を進めており、この経験と、測定者誤差などのノイズを含む時系列データである胎児超音波情報や、超高次元データであるゲノム情報の、前処理などのデータ取り扱い手法を検討し、情報種別ごとの機械学習モデルを開発する。開発した機械学習モデルを統合し、多様な情報の組み合わせの特徴を考慮する機械学習モデルの検討を行う。					
期待される成果	構築された学習モデルは、情報種別ごとの深層学習では捉えることができなかった、低出生体重の発症リスク予測のための、ゲノム情報と生活習慣情報などの組み合わせの特徴を考慮し、より高精度な学習モデルとなることが期待される。					
これまでの倫理 審査等の経過	2020年5月 東北メディカル・メガバンク機構倫理委員会承認					
倫理面、セキュリ ティー面への配慮	ヒトを対象とする医学系研究の倫理指針、ヒトゲノム・遺伝子解析研究倫理指針を遵守します。 また、利用する情報は、ToMMoスーパーコンピュータ内で利用し、外部へは持ち出しません。					
その他特記事項	奨学寄附金					
* 公開日	令和2年8月7日					