Universiteit
Leiden
The Netherlands

Multi-omics in research: epidemiology, methodology, and advanced data analysis

Faquih, T.O.

Citation
Faquih, T. O. (2023, March 28). *Multi-omics in research: epidemiology, methodology, and advanced data analysis*. Retrieved from https://hdl.handle.net/1887/3589838

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3589838

Note: To cite this publication please use the final published version (if applicable).

ملخص النتائج

توفر علوم تحليل منتجات الأيض (الميتابولوميكس / Metabolomics) وتحليل البروتينات (البروتيوميكس / proteomics) و تحليل الجينات (الجينومكس / Genomics) رؤى عميقة في علم الأحياء البشري وعلم وظائف الأعضاء المتعلقة بالأمراض وأسبابها في الفيزيولوجيا المرضية البيولوجية (Pathophysiology) على نطاق غير مسبوق. وقد استكشفنا في هذه الأطروحة التحديات البحثية التي تواجه تطبيق تقنيات الـ OMICs[1] (أوميكس) بنظرة تكاملية عميقة وأوضحنا تطبيقاتها العملية في الدراسات الوبائية (Epidemiological Studies).

وفي الجزء الأول من الأطروحة المكون من ثلاثة فصول، ركزنا على بعض تحديات طرق البحث المنهجية التي تواجه أبحاث الاوميكس. و**يتفحص الفصل الأول من الجزء الأول (Chapter 2)** من هذه الرسالة العلمية العلاقة ما بين نتائج قياسات البروتينات باستخدام "الأبتامير" (aptamer-based proteomics) للمؤشرات الحيوية للتخثر الوريدي (VTE[2]) مع طرق القياس المعتمدة التي تتم عادة في المراكز الصحية والممارسات الاكلينيكية المتعارف عليها. وقد وجدت دراستنا البحثية أن المؤشرات الحيوية التسعة التي تم قياسها للتخثر الوريدي أظهرت توافقا ضعيفا مع نتائج القياسات المعتمدة. وكانت العلاقة التوافقية ضعيفة بشكل خاص بالنسبة لـ"ثنائيات **د"** (D-dimer)، وهو مؤشر يستخدم لتشخيص داء التخثر الوريدي.

في **الفصل الثاني من الجزء الأول (Chapter 3)** من هذه الرساله ، تم دراسة التحديات المتمثلة في التعامل مع البيانات المفقودة في تحليل منتجات الأيض ، لا سيما من منصات بيانات تحليل منتجات الأيض الـ"غير مستهدفة".حيث قمنا بتقييم أداء طريقتين لاحتساب القيم المفقودة في بيانات تحليل منتجات الأيض، وهما طريقة "K-nearest neighbor" (kNN) و طريقة"Multiple imputation using chained equations" (التضمين المتعدد)، ومن خلال محاكاة الأنماط المختلفه لفقدان البيانات من الدراسة البحثية الهولندية عن السمنة [3] (NEO). أظهرت نتائجنا أن طريقة "التضمين المتعدد" كان لها تحيز أقل من طريقة kNN في معظم الحالات ، الا في الحالات ذات حجم عينة صغيره مع نسبه عاليه من البيانات المفقوده. كما قدمنا أيضا برنامجاً نصي حاسوبي بلغة برمجة R متاح للعامة لتبسيط عملية استخدام الطريقتين لاحتساب القيم المفقودة في بيانات قياسات منتجات الأيض.

بامكان منتجات الأيض (الميتابولايتس/ Metabolites) توفير نظرة ثاقبة وفريدة في الفيزيولوجيا المرضية البيولوجية لأنها تشمل تأثير العوامل الوراثية و تأثير نمط الحياة و تأثير العوامل البيئية. فلذلك تم استخدام علم تحليل منتجات الأيض لدراسة العملية الحيويه للتقدم بالعمر و الهرم. وقد استخدمت بعض الدراسات منتجات الأيض لتقدير "العمر الأيضي" ، وهو انعكاس للتقدم بالعمر على مستوى منتجات الأيض ، باستخدام طرق نمذجة التنبؤ الإحصائي. في **الفصل الثالث من الجزء الأول (Chapter 4)**، تناولنا التحديات الشائعة لتطوير نماذج التنبؤ بالعمر القائمة على منتجات الأيض

وطورنا نموذجا للتنبؤ بالعمر المبنى على ذلك باستخدام أجهزة منصة Metabolon لتحليل منتجات الأيض. بعد ذلك قمنا بتطوير نموذج في الدراسة البحثية INTERVAL[4] باستخدام أساليب الإحصاء "Ridge regression" و "Bootstrapping" .وقد اظهرت المجموعة المتضمنه في دراسة الـ INTERVAL ان عينة كبيرة من المجتمع كانت في صحة جيده نسبيا (n = 11,977) و مع نطاق عمري واسع ولذلك فان هذه المجموعه كانت مناسبه لتطوير نموذج تنبؤ العمر الأيضي. وبعد تطوير النموذج والتحقق الداخلي من صحته، أظهرت النتائج ان أداء التنبؤ الإحصائي للعمر الأيضي كان جيداً جداً (R^2 المعدل = 0.83).

اما الجزء الثاني من الأطروحة ، والمكون من ثلاثة فصول ايضا، فيتناول دراسة عدة أسئلة بحثية في علم الـepidemiology من خلال استخدام البيانات الجينومية و بيانات تحليل الأيض (باستخدام منصتين Metabolon و Nightingale) بالإضافة إلى طرق تحليل البيانات المتقدمة.

في **الفصل الأول من الجزء الثاني (Chapter 5)،** قمنا بفحص ارتباطات أكثر من 1300 منتج أيضي، تم قياسها بواسطة منصة Metabolon، و محتوى الدهون الثلاثية الكبدية (HTGC). وقد تضمنت هذه الارتباطات مجموعات من منتجات الأيض تابعه لمسارات حيويه مختلفة (خاصة من مجموعات الأحماض الأمينية والدهون) ، بالإضافة إلى منتجات أيض غير معهوده. بعد ذلك، قمنا ببناء شبكتين شموليتين متكاملتين، واحدة قائمة على أساس حيوي والأخرى قائمة على أساس إحصائي ، لتعيين وربط منتجات الأيض مع المسارات الحيويه المناسبه. وقد كشفت هذه الشبكات التفاعلية عن مسارات حيويه غير مألوفه يمكن ان تكون مرتبطة بالـ HTGC وربما أمراض الكبد الدهنية ككل هذا وقد تم توفير هذه الشبكات على صفحة الويب المتاحة على الإنترنت للمزيد من الاستكشاف من قبل المجتمع البحثي العلمي.

اما في الفصل الثاني من الجزء الثاني (Chapter 6)، فاستكشفنا الارتباطات المحتملة لأحجام تباين تكرار CAG الأقل حجما من عتبة تسبب مرض هنتنغتون (اقل من 36 تكرار) في جين *HTT* مع مستويات الأيض المقاسة بواسطة منصة Nightingale. وتبين أن أحد أنواع الخلل الجينية في الجينوم هو خلل ناتج عن اختلاف عدد تسلسل نيوكليوتيدات (Nucleotide) السيتوزين-الأدينين-الجوانين (تكرار تسلسل CAG[5]). وعندما يتجاوز حجم هذا التكرار العتبة الغير مرضيه في جينات معينة (أو بالقرب من هذه الجينات) ، فإنها تؤدي إلى العديد من الأمراض التنكسية العصبية الشديدة. إحدى هذه الحالات هي حجم تكرار تسلسل CAG الكبيرة في جين هنتنغتين[6] *(HTT)* مما يؤدي إلى مرض هنتنغتون[7]. من والمثير للاهتمام ، أن الدراسات الحديثة وجدت أن حجم تكرار CAG حتى اذا كان داخل النطاق الغير مرضي فان له تاثير على وظائف صحية مثل الوظيفة الإدراكية والاكتئاب وانخفاض وزن الجسم وبناء على ذلك تم إجراء تحليل البيانات بطريقه تحليل "التأثيرات المختلطة"[8] و تحليل "الوساطة"[9] في بيانات تم جمعها من ثلاث دراسات أوروبية ضخمه (n = 10,275). حيث أظهرت نتائجنا أن أحجام تكرار CAG الكبيرة الغير مسببة للأمراض كانت مرتبطة بنمط أيضي غير ملائم للصحة، على غرار النمط الأيضي المتصل بزيادة خطر الإصابة بأمراض القلب والأوعية الدموية.

وتتمثل إحدى مزايا تحليل منتجات الأيض في قدرة الغطاء على قياس منتجات الأيض "الخارجية الغريبة"، والتي يشار إليها غالبا باسم Xenobitoics. وهذه المنتجات الأيضيه يكون منشاها من استخدام الأدوية والنظام الغذائي والتعرض للتلوث البيئي. وإحدى هذه الكائنات الخارجية المقاسة هي مواد "Per- and Polyfluoroalkyl substances (PFAS)" (بيفاس) الصناعيه - والمعروفة باسم "الكيماويات الأبديه" بسبب تشبثها في البيئة وجسم الكائنات الحية، و بالإضافة إلى طبيعتها التشبثيه، فقد اوضحت الدراسات ان التركيزات العالية من الPFAS مرتبطة بمشاكل صحيه شديده في الانسان. وانه على الرغم من الجهود المبذولة لتنظيم إنتاجها، لا تزال مستويات PFAS قابلة للاكتشاف في التربة ومياه الشرب وعينات الدم البشري في كل من هولندا وألمانيا.

فلذلك قمنا في **الفصل الثالث من الجزء الثاني (Chapter 7)** ، بدراسه العلاقة ما بين مستويات PFAS في الدم ، التى تم قياسها بواسطة منصة Metabolon ، و مستويات المنتجات الأيضيه و البروتين الدهني التي تم قياسها بواسطة منصة Nightingale في مجموعة من عامة سكان هولندا (دراسة NEO) وألمانيا (دراسة راينلاند[10]). ووجدنا ان مستوى المنتجات الأيضيه للPFAS قابل للقياس في جميع الافراد تقريبا في كلتا الدراستين، علاوة على ذلك وكما أكد تحليل الميتا[11] فقد ارتبطت مستويات PFAS بنمط أيضي غير صحي يتميز بزيادة مستوى "الأحماض الدهنية" و"البروتينات الدهنية منخفضة الكثافة" ومستويات "البروتين الشحمي (**ب**)"[12]. وكانت هذه الارتباطات متشابهة بالنمط الأيضي الذي يزيد خطر الإصابة بأمراض القلب والأوعية الدموية .ووجدنا ايضا أن هذا النمط شديد بشكل خاص لدى الأفراد الأصغر سنا. هذه النتائج تشير الى أن الPFAS يمثل خطرًا صحيًا حتى في الأفراد الذين لم يتعرضوا على ما يبدو لمستويات عالية من PFAS.

[1] "OMICs" هو مجال في العلم الحديث الذي يتطلع إلى الفهم التام للحياة والعمليات الحيوية عبر التحليل الجزيئي للعناصر المختلفة التي تشكل الحياة، مثل الجينات والبروتينات والأملاح والمركبات الكيميائية الأخرى حيث يتم تحليل هذه العناصر عبر عدة انماط "أوميكس" مختلفة، مثل "جينوميكس" (تحليل الجينات) و"بروتوميكس" (تحليل البروتينات). يستخدم هذا المجال في التحقيقات العلمية لفهم أفضل للعمليات الحيوية وللعثور على حلول للمشاكل الصحية المختلفة.

[2] Venous Thrombosis

[3] The Netherlands Epidemiology of Obesity Study (NEO)

[4] تم إنشاء دراسة INTERVAL بالتعاون بين جامعتي كامبريدج وأوكسفورد و التعاون ايضا مع دائرة الدم والتبرعات الصحية في المملكة المتحدة

[5] Cytosine-adenine-guanine nucleotide sequences

[6] Huntingtin gene

[7] Huntington's disease

[8] Multilevel mixed-effects model

[9] Mediation analysis

[10] Rhineland study

[11] Meta analysis

[12] Apolipoprotein B