·综述·

非综合征型唇腭裂的遗传预测模型研究进展

王斯悦[1]　彭和香[1]　薛恩慈[1]　陈曦[1]　王雪珩[1]　范梦[1]　王梦莹[1]　李楠[2]　李静[3]
周治波[2]　朱洪平[2]　胡永华[1]　吴涛[1,4]
[1]北京大学公共卫生学院流行病与卫生统计学系，北京 100191；[2]北京大学口腔医院颌面外科，北京 100081；[3]北京大学口腔医院儿科，北京 100081；[4]北京大学生育健康研究所/国家卫生健康委员会生育健康重点实验室，北京 100191
通信作者：吴涛，Email：twu@bjmu.edu.cn

【摘要】 非综合征型唇腭裂（NSOC）是我国常见的出生缺陷。近年来，随着我国生育政策相继调整两次，与高龄生育伴发的出生缺陷防控形势日益严峻。开展NSOC风险预测将为健全出生缺陷防控链条提供重要证据。近年来，全基因组关联研究和二代测序等发现了多个与NSOC有关的遗传位点，为开展预测提供了有益信息。本文综述了NSOC风险预测，特别是利用遗传信息开展风险预测的常用方法，以期在研究设计、变量筛选、构建策略及评价方法等方面，为进一步开发和完善NSOC等复杂出生缺陷的风险预测模型提供参考。

【关键词】 出生缺陷；　非综合征型唇腭裂；　风险预测；　风险预测模型
基金项目：国家自然科学基金（81102178，81573225）；北京市自然科学基金（7172115）

Progress in research of risk prediction of non-syndromic oral clefts using genetic information

Wang Siyue[1], Peng Hexiang[1], Xue Enci[1], Chen Xi[1], Wang Xueheng[1], Fan Meng[1], Wang Mengying[1], Li Nan[2], Li Jing[3], Zhou Zhibo[2], Zhu Hongping[2], Hu Yonghua[1], Wu Tao[1,4]
[1]Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing 100191, China; [2]Department of Oral and Maxillofacial Surgery, School of Stomatology, Peking University, Beijing 100081, China; [3]Department of Pediatrics, School of Stomatology, Peking University, Beijing 100081, China; [4]Institute of Reproductive and Child Health/Key Laboratory of Reproductive Health, National Health Commission of the People's Republic of China, Beijing 100191, China
Corresponding author: Wu Tao, Email: twu@bjmu.edu.cn

【Abstract】 Non-syndromic oral cleft (NSOC), a common birth defect, remains to be a critical public health problem in China. In the context of adjustment of childbearing policy for two times in China and the increase of pregnancy at older childbearing age, NSOC risk prediction will provide evidence for high-risk population identification and prenatal counseling. Genome-wide association study and second generation sequencing have identified multiple loci associated with NSOC, facilitating the development of genetic risk prediction of NSOC. Despite the marked progress, risk prediction models of NSOC still faces multiple challenges. This paper summarizes the recent progress in research of NSOC risk prediction models based on the results of extensive literature retrieval to provide some insights for the model development regarding research design, variable selection, model-build strategy and evaluation methods.

【Key words】 Birth defect; Non-syndromic oral cleft; Risk prediction; Genetic risk prediction model
Fund programs: National Natural Science Foundation of China (81102178, 81573225); Beijing Municipal Natural Science Foundation (7172115)

DOI：10.3760/cma.j.cn112338-20220624-00556
收稿日期 2022-06-24　　**本文编辑** 万玉立
引用格式：王斯悦，彭和香，薛恩慈，等. 非综合征型唇腭裂的遗传预测模型研究进展[J]. 中华流行病学杂志，2023，44(3)：504-510. DOI：10.3760/cma.j.cn112338-20220624-00556.
Wang SY, Peng HX, Xue EC, et al. Progress in research of risk prediction of non-syndromic oral clefts using genetic information[J]. Chin J Epidemiol, 2023, 44(3):504-510. DOI: 10.3760/cma.j.cn112338-20220624-00556.

非综合征型唇腭裂（non-syndromic oral cleft，NSOC）指单独发生的颌面部畸形，不伴有其他结构畸形或者发育障碍，是我国常见出生缺陷和重要公共卫生问题[1-2]。目前该病病因尚不清楚且缺乏有效的预防措施。2021 年，我国印发《健康儿童行动提升计划（2021-2025 年）》[3]。另外，随着我国近年来两次调整生育政策，女性生育渐有高龄趋势，高龄生育伴发的出生缺陷防控形势愈发严峻。开展 NSOC 等出生缺陷的产前风险预测是完善出生缺陷防控保障的重要环节，可辅助识别高危人群、健全二级预防、促进精准健康管理，具有重要研究价值。

NSOC 不同于成年期发病的复杂疾病，出生时可判断疾病状态，且遗传度较高[4-6]。因此，基于遗传信息预测该病发生风险具备一定的可靠性。既往开展的多个全基因组关联研究（genome-wide association study，GWAS）已发现数十个影响该疾病发病风险的基因或区域[6-8]，不仅具有病因学研究价值，也具有预测价值。

对唇腭裂等出生缺陷而言，目前常见产前预测多可诊断综合征型[9-11]，大致可分为基于超声影像和基于遗传信息预测两类；如超声检查“大排畸”[12]，或孕早期胎儿游离 DNA 无创产前检测，或辅助生殖中胚胎植入前遗传诊断等[13-14]。然而，尽管 NSOC 相比于综合征型更为常见，但截至目前，适用于 NSOC 产前预测的研究证据还较少[15-16]。因此，本文综述了该病各类预测模型，特别是遗传预测模型的研究进展，同时梳理了目前常见的遗传预测方法，以期为进一步开发和完善 NSOC 预测模型提供参考。

1. 常见 NSOC 预测模型分类：NSOC 预测模型通常可根据研究目的分为诊断预测模型和前瞻性预测模型两类。诊断预测模型基于孕期胎儿在娩出前可获取的遗传信息，实现该胎儿的患病风险预测，常见于产前诊断；前瞻性预测模型常基于父母亲代的遗传、环境等信息，在配子尚未形成时，预测未来可能孕育的胎儿的发病风险，常见于遗传咨询。尽管前瞻性预测具有防控出生缺陷关口前置的重要意义，但由于既有研究设计、数据利用方式、预测方法学等方面的原因，非综合征型出生缺陷的预测研究证据较少，尚不能满足该病临床应用的现实需求。截至目前已发表的 NSOC 预测模型研究见表 1，仅有 Li 等[17]的研究为前瞻性预测模型。

2. NSOC 预测模型构建方法：预测模型的构建与常规流行病学研究中的解释型模型不同。前者的主要目标是预测误差的最小化；而后者的目标则是探索潜在危险因素与疾病的关联，估计潜在危险因素的关联强度，为因果推断提供证据。由于两者的目标差异，基于解释型模型获得的关联证据未必能直接用于预测模型的构建。例如，涉及低频罕见遗传变异等遗传预测模型的研究[20]。罕见变异可能具有病因学价值，但由于人群中这类遗传变异频率较低，直接纳入罕见变异预测，未必在人群中获得更优的预测效果[20]。通常，NSOC 预测模型也遵循研究设计、模型构建方法及模型评价等维度的考量，本文拟针对上述内容，综述该病风险预测的研究进展，以期为出生缺陷中复杂性状的预测提供参考。

（1）研究设计：常见出生缺陷的预测模型研究设计包括病例-对照研究、家系设计等。

病例-对照研究通常收集出生缺陷患儿及健康活产儿的父母，以父母的暴露特征为预测因子。Li 等[17]开展了以医院为基础的病例-对照研究，调查对象包括来自中国湖南省 52 家出生缺陷监测医院的 113 例 NSOC 和 226 例健康对

表 1　目前已发表的非综合征型唇腭裂预测模型研究基本信息

作者及发表时间	研究设计	样本量	结局	地点	预测模型构建方法	预测变量池	变量筛选方法	AUC[a]
Wen 等[18]，2015	家系、病例-对照	1 875/3 692	单纯腭裂、单纯唇裂	亚洲、欧洲地区	多类似然比	148 个 SNP	过滤法	唇裂合并腭裂：0.572 唇裂不合并腭裂：0.589 唇裂合并或不合并腭裂：0.604 软型单纯腭裂：0.617 硬型单纯腭裂：0.623 单纯腭裂：0.556
Li 等[17]，2016	病例-对照	113/226	单纯唇裂	中国	logistic 回归、Fisher 判别分析	13 个环境变量	封装法	0.846
Zhang 等[19]，2018	病例-对照	382/205[b] 103/205[d] 279/205[e]	单纯唇裂	中国（汉族、维吾尔族）	PRS、logistic 回归、SVM、NB、KNN、RF、DT、ANN[f]	43 个 SNP	过滤法、嵌入式法	PRS[c]：0.882、0.716 SVM[c]：0.89、0.64 logistic 回归[c]：0.90、0.62 NB[c]：0.87、0.60 RF[c]：0.89、0.54 KNN[c]：0.75、0.57 DT[c]：0.74、0.54 ANN[c]：0.85、0.51

注：[a]受试者工作特征曲线下面积（area under the curve，AUC）；[b]合并后的病例/对照样本数量；[c]分别表示汉族、维吾尔族人群中预测模型结果；[d]汉族病例/对照样本数量；[e]维吾尔族病例/对照样本数量；[f]分别表示加权遗传评分（polygenic risk score，PRS）、支持向量机（supporting vector machine，SVM）、朴素贝叶斯（Naïve Bayes，NB）、K 近邻（K-Nearest neighbor，KNN）、随机森林（random forest，RF）、决策树（decision tree，DT）、人工神经网络（artificial neural network，ANN）

照。研究采用面对面问卷调查法回顾收集双亲在孕前和孕期与NSOC有关的暴露，基于logistic回归模型分析该病的关联影响因素，并采用逐步Fisher判别分析构建预测模型。其中8个亲代危险环境暴露与NSOC的关联有统计学意义，并被进一步纳入判别分析构建预测模型。基于8个预测变量（家族史、家庭收入，父/母职业暴露、婚前检查、豆浆牛奶摄入、父亲饮浓茶、住房改造史）构建的预测模型显示：83.8%的样本可以被正确区分为病例或对照。该预测模型的敏感度为74.3%、特异度为88.5%，该模型的受试者工作特征曲线下面积（area under the curve，AUC）为0.846。然而，上述单独基于环境危险暴露开展预测的研究不仅样本量较小，且缺乏预测模型外部验证，可能存在过度拟合的问题，在外推到其他人群中开展实际应用前仍有待进一步的验证和评估。

此外，由于出生缺陷的特殊性，部分预测模型采用家系研究设计开展。相比于病例对照研究，同时纳入患儿与父代的信息，可探索亲源效应、母亲围孕期暴露与子代基因型的交互作用及宫内环境对出生缺陷的遗传效应[21]。Wen和Lu[18]采用核心家系设计纳入患儿-父母三联体及对照-父母三联体，以148个SNP位点建立的预测模型其AUC在唇裂合并或不合并腭裂为0.604、单纯腭裂为0.556。尽管队列设计在其他出生缺陷预测模型研究中应用较为广泛，但截至目前，在NSOC领域较少见到基于队列设计构建的预测模型。

当前多基因复杂疾病的风险预测亦多有采用复杂的嵌套设计，例如巢式病例-对照设计[22-23]、家系队列[24-25]、基于家系队列的巢式病例-对照[26]等设计，或利用“表型组”等策略探讨“单一模型预测多种疾病表型”的设计[27]。这些研究利用“遗传信息”的特点，暴露和疾病的时间顺序明确，结合纵向设计利用相对较小的样本，实现人群内危险分层。另外，家系研究发现，利用不同家系结构可能具有提高该病预测研究效率的潜力[28-29]。因此探索基于复合研究设计（如不同家系结构、家系及人群等复合设计）的遗传预测模型，是未来NSOC遗传预测进一步发展的方向之一[28]。

（2）模型构建技术：预测模型的构建涉及组织预测因子的方式，这也是影响预测模型效果的重要环节。目前常见的预测模型构建技术包括基于回归模型的方法和机器学习方法两类。回归模型是目前常见的预测模型构建方法。该法可通过较为直观的方式组合已知的危险因素[29]，亦可纳入交互作用，有利于提升预测模型的准确性[30-31]。然而，随着后GWAS时代的到来，传统的回归模型难以处理数据维度极大、数据间存在复杂关联、噪音数据极多的全基因组信息。当前研究常采用两种策略解决上述问题。

第一种策略是构建多基因风险评分（polygenic risk score，PRS）[32-33]。由于复杂疾病的发病风险往往受到多个遗传和环境因素的影响，任何单一的遗传或环境因素变异仅对整体风险贡献微弱的效应[34-35]。PRS通过加权平均多个位点效应的方法，成为可同时纳入多个甚至全基因组位点效应的预测技术，在慢性病、癌症等研究领域获得了较好的预测效果[36-37]。然而，PRS预测应用至今，也引发了较多争议[38]，主要表现为PRS纳入的预测变量过多，可能不具备实践应用价值[34]；同时PRS预测研究在外部验证中多数失灵，对其预测有效性的方法学也提出挑战[39]。总之，PRS日益成为开展人群复杂疾病风险评估、促进精准健康的重要工具。但如何弥合PRS研究与临床、公共卫生实践的鸿沟，仍有待进一步探索。首先，PRS预测研究的方法学及结果解读策略尚有待国际认可的指南规范；其次，可将目前开展的、针对同一目标疾病的PRS预测模型进行汇总和梳理，以便评估和校准当前相对分散的研究方法和结果；最后，也可推进PRS证据在临床、公共卫生预防实践中的阈值研究，推进向应用的转化。第二种策略是改进算法。相较于传统的回归模型，机器学习预测方法能有效拟合数据间的复杂关系、提高预测精度，逐渐被广泛应用于遗传预测研究。同时，由于构建预测模型对模型的解释性需求较低，机器学习模型的黑箱算法恰好并不构成预测模型研究中的劣势[40]。在NSOC应用方面，由于预测表型多为患病/不患病等分类变量，机器学习多以分类器形式出现，通常被分为有监督和无监督式两类[41]。监督式通过建立个体变异特征与复杂疾病表型的关联实现表型的预测[42]。无监督式则仅基于暴露变量，通过探索变量间结构实现个体分类的目的[43]。另外，深度学习亦常应用于复杂疾病预测[44-45]。深度学习是模仿人脑工作机制来解释数据的一种机器学习技术，主要分为卷积神经网络（convolutional neural networks）和深度置信网（deep belief nets）两类。有研究比较了深度学习和经典机器学习算法，结果显示对于非加性遗传效应占比高的复杂表型，深度学习算法预测的效果更优[46]。

数据不平衡处理也是利用机器学习技术构建遗传预测模型的重要又特殊的考量范畴[47]。以病例-对照研究设计为例，由于病例、对照GWAS数据收集的难易程度往往存在较大差异，易发现病例的样本数量远大于对照样本数量的情况；或反之，如罕见病研究。这种病例、对照样本数目相差巨大的情况也称为数据不平衡。如果忽略数据不平衡情况，机器学习可能倾向于以朴素行为的方式预测病例-对照的分类状态[48]，最终表现为预测模型在内部样本中效果较好，在外部验证样本中失效。处理相关问题的策略可以基于数据和算法两类考虑。数据视角下，除了基于一定规则的数据扩充外，还可以考虑过采样（over-sampling）和欠采样（under-sampling）策略[49]。算法视角下，可以设计不同“错误分类”（miss-classification）代价函数（cost function）进而优化机器学习过程，例如代价敏感学习算法（cost-sensitive learning）中的adacost算法等。同时，热门领域也有将不平衡数据转化为异常值检测（novelty detection）问题开展策略研究的探讨，相关算法有One-class SVM等。

另外，小样本的机器学习方法也是遗传预测模型研究的热点问题之一[50-51]。由于机器学习算法具有较强的样本量依赖，尤其是考虑到GWAS信息的预测特征维度远远高

于预测样本量时，如何使基于小样本的机器学习方法也获得相对可靠的预测效果，也是当前遗传预测的重要挑战之一。适用于小样本的机器学习方法大多利用了数据扩充和算法优化两类策略。具体包括属性值随机采样、少数类样本的合成过采样技术（synthetic minority over-sampling technique）及通过弱学习器相互堆叠[52]，递归为boosting算法等。相关进展大多应用于慢性病和癌症等领域[47,53]，NSOC领域未来可借鉴其他复杂疾病预测前沿方法开展相关工作。

同时，变量筛选策略对于预测模型效果具有重要影响。GWAS检测的暴露数量远远超过传统关联研究，海量遗传信息在大多数情况下无法直接用于构建预测模型[54]，因此如何筛选纳入变量成为实际操作中的重要问题。目前，基于GWAS变量筛选策略可分为知识驱动（knowledge driven）和数据驱动（data driven）两类。在知识驱动策略下，研究假设具有病因学价值的位点也具有预测能力，通常选择具有生物学功能的位点或GWAS阳性位点。然而，早期的遗传风险预测研究表明，仅利用GWAS获得的阳性关联位点进行预测并不理想[55-56]，重要原因之一在于GWAS为了控制假阳性关联位点，采取了较为严苛的显著性阈值[57-58]，也从侧面反映了病因学研究与预测模型研究目标的差异。在NSOC方面，Zhang等[19]即采用此策略，纳入既往GWAS提示NSOC阳性的43个位点进行后续NSOC预测模型研究，内部验证结果提示在汉族人群中预测效果尚可（AUC>0.74），在维吾尔族人群中预测效果则较为不足（AUC>0.51）。数据驱动策略下，常见的变量筛选策略包括过滤（filtering）、嵌入式模块（embedded modules）和封装（wrapper）3类[59]。过滤常根据预设的阈值（如$P<5\times10^{-8}$）筛选预测变量；嵌入式模块先基于某些机器学习算法获得各变量的重要性排序，再依据该排序筛选变量；封装法多依据目标函数（通常为预测效果评分）筛选预测变量，并在排除选中变量后的剩余数据中再次循环上述过程，直至所有变量均被遍历。实际预测模型构建过程中，上述3种方法也可串联或并联使用。

另外，随着NSOC病因学研究逐渐在罕见变异、非编码RNA[60]、MicroRNA[61]交互作用网络等方向获得突破性研究进展，利用GWAS与上述新病因发现关联开展预测也可能是一种新型遗传预测模型构建策略。

（3）评价方法：评价模型的预测能力通常基于原始研究人群开展内部测试（即内部测试集）。若该模型预测性能较好，则应在外部研究中验证（即外部测试集）。外部数据验证可防止数据过度拟合时高估模型预测能力，并可提示该模型的外推性。

预测模型准确度的具体评价指标包括校准度（calibration）和区分度（discrimination）[62]。校准度是评价模型实际观察的风险与模型估计的风险的一致程度，常用Hosmer-Lemeshow卡方检验。校准度不佳的预测模型将系统性地低估或高估结局事件概率[63]。区分度是反映模型正确区分对象结局差异的能力，是在个体水平评估一个模型能够把将要发生所关注疾病的人与那些不会发生该疾病的人区分开来的能力。若病例的预测风险高于对照，则该模型的区分度较好。区分度常用AUC或*C*-index、敏感度及特异度等指标评价。

目前NSOC预测模型的评价较多关注区分度，对校准度的评价关注不足。此外，目前已发表的模型均未经过外部验证。在现有的该病预测模型中，Zhang等[19]基于中国人GWAS阳性位点，采用支持向量机等机器学习方法构建的预测模型AUC数值最大。但值得注意的是，基于不同人群构建的不同预测模型难以直接通过AUC大小直接判断不同预测模型的优劣[64]。这是由于不同研究中的AUC与患者在人群中的分布特征有关，若模型中纳入的变量特征在人群中有更大的异质性，则模型的AUC趋向于更高。因此，当前比较不同NSOC预测模型效果时，尚不能单独依靠AUC数值大小推断Zhang等[19]的支持向量机模型为中国人群NSOC的最优预测模型。

值得一提的是，尽管预测模型不以模型解释为终极目标，但试图理解未知的黑箱算法，评估不同预测特征对预测的贡献，长期以来都是热点方向[65]。其中，SHAP值（SHapley Additive exPlanation）及其可视化是相关领域的重要进展[66]。相关应用研究多发表于常见慢性复杂疾病及癌症等[67-68]，研究结果为后续集中预测资源在高预测能力的遗传变量上奠定了基础。

另外，机器学习遗传预测模型面临着临床实践的考验。仅凭借群体特征水平上的AUC、敏感度、特异度等指标去遴选应用于个体风险预测的方法，是难以保障临床防治诊疗现实安全需求的。因此，黑箱算法的“可靠性评估”也是预测应用推广前的重要考量。因此，发展应用可靠性的评估方法，例如新近提出的逐点可靠性（pointwise reliability）评估方法等[69]，也是遗传预测模型的重要发展方向之一。相较之下，NSOC仍有待更多研究丰富相关证据。总之，目前已发表的NSOC预测模型研究数量较少，现有研究样本量较小，预测模型的内部预测效果较好。从预测研究的变量筛选策略上，现有研究主要基于GWAS阳性或环境暴露阳性的知识驱动策略，未系统性采用数据驱动等策略筛选变量，也并未考虑具有潜在生物学功能的位点及基因-环境交互作用等。从模型评价角度而言，现有研究多基于区分度展开评价，对于校准度的评价相对缺乏，且尚未关注机器学习模型的可解释性、评估可靠性等问题。

3. NSOC预测模型研究的机遇及挑战：NSOC作为一种常见的出生缺陷，遗传度较高，因而基于遗传因素构建NSOC预测模型具有重要意义。综合考量预测模型当前的研究现状，其所面临的机遇与挑战如下。

首先，该病病因未明，且不同疾病亚型间存在病因异质性，这给预测模型构建带来诸多挑战。如何在既有病因学研究证据的基础上，精细划分疾病亚型，辅助预测模型的开发，提升预测模型的外推性，仍有待进一步研究。目前，Wen和Lu[18]已在相关领域的方法学研究中开展了有益探

索。该研究开发了多类似然比（multi likelihood ratio，MLR）方法，在不依赖既有分型的基础上，从遗传数据的相似性出发，以数据驱动的方式，逐渐将具有相似遗传病因的个体合并为同质的亚型。模拟研究和基于NSOC的实证分析结果表明，通过MLR在该病常见分型基础上，进一步划分为精细的亚型，再构建各亚型的遗传预测模型，可获得更高的预测区分度。

其次，丰富的高质量GWAS数据为预测模型的构建提供了海量的数据资源，但同时也伴随着数据维度极大、数据间存在复杂关联、噪音数据极多等挑战。如何充分挖掘大数据资源的优势，并探索符合GWAS大数据特征的预测方法成为提升预测模型效果的重要挑战之一。随着复杂疾病预测模型不断发展，例如机器学习方法及高维度变量的压缩筛选方法等，为NSOC预测模型提供了理论和方法学基础，亟待在该病领域利用上述研究进展探索预测模型的构建方法和预测效果。

最后，现有的NSOC预测模型均单独基于遗传或环境变量构建，缺乏基于系统流行病学的多维度、多水平的预测模型研究；也缺乏基于队列研究及外部人群验证等。这些因素均在一定程度上限制了该病预测模型的可外推性。相关问题的解决将有望为孕前咨询、临床产前风险评估提供重要研究基础。

随着NSOC人群研究证据不断丰富，以及多源数据的连接互通成为可能，探索高效利用多层次多水平的综合信息，构建准确可靠的风险预测模型方法，为丰富临床个体化医疗实践、风险评估和疾病预防提供重要科学证据。

利益冲突　所有作者声明无利益冲突

参 考 文 献

[1] Wang MY, Yuan Y, Wang ZF, et al. Prevalence of orofacial clefts among live births in China:a systematic review and meta-analysis[J]. Birth Defects Res, 2017, 109(13): 1011-1019. DOI:10.1002/bdr2.1043.

[2] Machado RA, Popoff DAV, Martelli-Júnior H. Relationship between non-syndromic oral clefts and cancer: a systematic review and meta-analysis[J]. Oral Dis, 2022, 28(5):1369-1386. DOI:10.1111/odi.14179.

[3] 国家卫生健康委. 国家卫生健康委关于印发健康儿童行动提升计划(2021-2025年)的通知:国卫妇幼发〔2021〕33号[EB/OL]. (2021-10-29)[2022-06-19]. http://www.gov.cn/zhengce/zhengceku/2021-11/05/content_5649019.htm.

[4] Grosen D, Bille C, Petersen I, et al. Risk of oral clefts in twins[J]. Epidemiology, 2011, 22(3): 313-319. DOI: 10.1097/EDE.0b013e3182125f9c.

[5] Sivertsen Å, Wilcox AJ, Skjærven R, et al. Familial risk of oral clefts by morphological type and severity:population based cohort study of first degree relatives[J]. BMJ, 2008, 336(7641):432-434. DOI:10.1136/bmj.39458.563611.AE.

[6] Beaty TH, Murray JC, Marazita ML, et al. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near *MAFB* and *ABCA4*[J]. Nat Genet, 2010, 42(6):525-529. DOI:10.1038/ng.580.

[7] Yu YQ, Zuo XB, He M, et al. Genome-wide analyses of non-syndromic cleft lip with palate identify 14 novel loci and genetic heterogeneity[J]. Nat Commun, 2017, 8: 14364. DOI:10.1038/ncomms14364.

[8] Huang LL, Jia ZL, Shi Y, et al. Genetic factors define CPO and CLO subtypes of nonsyndromicorofacial cleft[J]. PLoS Genet, 2019, 15(10): e1008357. DOI: 10.1371/journal.pgen.1008357.

[9] Tian MM, Xiao LX, Jian N, et al. Accurate diagnosis of fetal cleft lip/palate by typical signs of magnetic resonance imaging[J]. Prenat Diagn, 2019, 39(10): 883-889. DOI: 10.1002/pd.5499.

[10] Machado RA, de Oliveira Silva C, Martelli-Junior H, et al. Machine learning in prediction of genetic risk of nonsyndromic oral clefts in the Brazilian population[J]. Clin Oral Invest, 2021, 25(3): 1273-1280. DOI: 10.1007/s00784-020-03433-y.

[11] Jia SS, Zhang Q, Wang Y, et al. PIWI-interacting RNA sequencing profiles in maternal plasma-derived exosomes reveal novel non-invasive prenatal biomarkers for the early diagnosis of nonsyndromic cleft lip and palate[J]. eBioMedicine, 2021, 65:103253. DOI:10.1016/j.ebiom.2021.103253.

[12] Li XL, Xiu GH, Yan F, et al. First-trimester evaluation of cleft lip and palate by a novel two-dimensional sonographic technique: a prospective study[J]. Curr Med Imaging, 2022, 19(3): 278-285. DOI: 10.2174/1573405618666220713103500.

[13] Zaninovic N, Rosenwaks Z. Artificial intelligence in human in vitro fertilization and embryology[J]. Fertil Steril, 2020, 114(5):914-920. DOI:10.1016/j.fertnstert.2020.09.157.

[14] Serafin D, Grabarek BO, Boroń D, et al. Evaluation of the risk of birth defects related to the use of assisted reproductive technology:an updated systematic review[J]. Int J Environ Res Public Health, 2022, 19(8): 4914. DOI: 10.3390/ijerph19084914.

[15] Martinelli M, Palmieri A, Carinci F, et al. Non-syndromic cleft palate: an overview on human genetic and environmental risk factors[J]. Front Cell Dev Biol, 2020, 8: 592271. DOI:10.3389/fcell.2020.592271.

[16] Stanier P, Moore GE. Genetics of cleft lip and palate: syndromic genes contribute to the incidence of non-syndromic clefts[J]. Hum Mol Genet, 2004, 13 Spec No 1:R73-81. DOI:10.1093/hmg/ddh052.

[17] Li HX, Luo MY, Luo JY, et al. A discriminant analysis prediction model of non-syndromic cleft lip with or without cleft palate based on risk factors[J]. BMC Pregnancy Childbirth, 2016, 16(1): 368. DOI: 10.1186/s12884-016-1116-4.

[18] Wen Y, Lu Q. Risk prediction models for oral clefts allowing for phenotypic heterogeneity[J]. Front Genet. 2015, 6:264. DOI:10.3389/fgene.2015.00264.

[19] Zhang SJ, Meng PQ, Zhang JN, et al. Machine learning models for genetic risk assessment of infants with non-syndromic orofacial cleft[J]. Genomics Proteomics Bioinformatics, 2018, 16(5): 354-364. DOI: 10.1016/j.gpb.2018.07.005.

[20] Lali R, Chong M, Omidi A, et al. Calibrated rare variant genetic risk scores for complex disease prediction using large exome sequence repositories[J]. Nat Commun, 2021, 12(1):5852. DOI:10.1038/s41467-021-26114-0.

[21] Ainsworth HF, Unwin J, Jamison DL, et al. Investigation of maternal effects, maternal-fetal interactions and parent-of-origin effects (imprinting), using mothers and their offspring[J]. Genet Epidemiol, 2011, 35(1): 19-45. DOI:10.1002/gepi.20547.

[22] Liu C, Wang YY, Zheng W, et al. Putrescine as a novel biomarker of maternal serum in first trimester for the prediction of gestational diabetes mellitus: a nested case-control study[J]. Front Endocrinol, 2021, 12:759893. DOI:10.3389/fendo.2021.759893.

[23] Xu C, Guo Z, Zhang J, et al. Non-invasive prediction of fetal growth restriction by whole-genome promoter profiling of maternal plasma DNA: a nested case-control study[J]. BJOG, 2021, 128(2): 458-466. DOI: 10.1111/1471-0528.16292.

[24] Tapia G, Suvitaival T, Ahonen L, et al. Prediction of type 1 diabetes at birth: cord blood metabolites vs genetic risk score in the norwegian mother, father, and child cohort[J]. J Clin Endocrinol Metab, 2021, 106(10):e4062-4071. DOI: 10.1210/clinem/dgab400.

[25] Farmer RF, Gau JM, Seeley JR, et al. Family-based predictors of alcohol use disorder (AUD) recurrence and new non-alcohol substance use disorder onset following initial AUD recovery[J]. J Stud Alcohol Drugs, 2022, 83(2): 239-247. DOI:10.15288/jsad.2022.83.239.

[26] Monson KR, Goldberg M, Wu HC, et al. Circulating growth factor concentrations and breast cancer risk: a nested case-control study of IGF-1, IGFBP-3, and breast cancer in a family-based cohort[J]. Breast Cancer Res, 2020, 22(1): 109. DOI:10.1186/s13058-020-01352-0.

[27] Ballout N, Garcia C, Viallon V. Sparse estimation for case-control studies with multiple disease subtypes[J]. Biostatistics, 2021, 22(4):738-755. DOI:10.1093/biostatistics/kxz063.

[28] Lou XY, Hou TT, Liu SY, et al. Innovative approach to identify multigenomic and environmental interactions associated with birth defects in family-based hybrid designs[J]. Genet Epidemiol, 2021, 45(2): 171-189. DOI: 10.1002/gepi.22363.

[29] Wray NR, Goddard ME. Multi-locus models of genetic risk of disease[J]. Genome Med, 2010, 2(2): 10. DOI: 10.1186/gm131.

[30] Wang H, Bennett DA, de Jager PL, et al. Genome-wide epistasis analysis for Alzheimer's disease and implications for genetic risk prediction[J]. Alzheimers Res Ther, 2021, 13(1):55. DOI:10.1186/s13195-021-00794-8.

[31] Moore JH, Williams SM. Epistasis and its implications for personal genetics[J]. Am J Hum Genet, 2009, 85(3): 309-320. DOI:10.1016/j.ajhg.2009.08.006.

[32] Lencz T, Backenroth D, Granot-Hershkovitz E, et al. Utility of polygenic embryo screening for disease depends on the selection strategy[J]. Elife, 2021, 10: e64716. DOI: 10.7554/eLife.64716.

[33] Liu GQ, Peng JJ, Liao ZX, et al. Genome-wide survival study identifies a novel synaptic locus and polygenic score for cognitive progression in Parkinson's disease[J]. Nat Genet, 2021, 53(6): 787-793. DOI: 10.1038/s41588-021-00847-6.

[34] Lewis CM, Vassos E. Polygenic risk scores: from research tools to clinical instruments[J]. Genome Med, 2020, 12(1): 44. DOI:10.1186/s13073-020-00742-5.

[35] Ikeda M, Saito T, Kanazawa T, et al. Polygenic risk score as clinical utility in psychiatry: a clinical viewpoint[J]. J Hum Genet, 2021, 66(1): 53-60. DOI: 10.1038/s10038-020-0814-y.

[36] Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores[J]. Nat Rev Genet, 2018, 19(9):581-590. DOI:10.1038/s41576-018-0018-x.

[37] Marston NA, Kamanu FK, Nordio F, et al. Predicting benefit from evolocumab therapy in patients with atherosclerotic disease using a genetic risk score: results from the FOURIER trial[J]. Circulation, 2020, 141(8): 616-623. DOI:10.1161/CIRCULATIONAHA.119.043805.

[38] Janssens ACJW. Validity of polygenic risk scores: are we measuring what we think we are? [J]. Hum Mol Genet, 2019, 28(R2):R143-150. DOI:10.1093/hmg/ddz205.

[39] Wray NR, Lin T, Austin J, et al. From basic science to clinical application of polygenic risk scores: a primer[J]. JAMA Psychiatry, 2021, 78(1): 101-109. DOI: 10.1001/jamapsychiatry.2020.3049.

[40] Adler ED, Voors AA, Klein L, et al. Improving risk prediction in heart failure using machine learning[J]. Eur J Heart Fail, 2020, 22(1):139-147. DOI:10.1002/ejhf.1628.

[41] Eckardt JN, Wendt K, Bornhäuser M, et al. Reinforcement learning for precision oncology[J]. Cancers (Basel), 2021, 13(18):4624. DOI:10.3390/cancers13184624.

[42] Andaur Navarro CL, Damen JAA, Takada T, et al. Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review[J]. BMJ, 2021, 375: n2281. DOI: 10.1136/bmj.n2281.

[43] Eshaghi A, Young AL, Wijeratne PA, et al. Identifying multiple sclerosis subtypes using unsupervised machine learning and MRI data[J]. Nat Commun, 2021, 12(1): 2078. DOI:10.1038/s41467-021-22265-2.

[44] Lee S, Lee HC, Chu YS, et al. Deep learning models for the prediction of intraoperative hypotension[J]. Br J Anaesth, 2021, 126(4):808-817. DOI:10.1016/j.bja.2020.12.035.

[45] Doppalapudi S, Qiu RG, Badr Y. Lung cancer survival period prediction and understanding: deep learning approaches[J]. Int J Med Inform, 2021, 148: 104371. DOI: 10.1016/j.ijmedinf.2020.104371.

[46] Abdollahi-Arpanahi R, Gianola D, Peñagaricano F. Deep learning versus parametric and ensemble methods for genomic prediction of complex phenotypes[J]. Genet Sel Evol, 2020, 52(1):12. DOI:10.1186/s12711-020-00531-z.

[47] Wu YF, Fang Y. Stroke prediction with machine learning methods among older Chinese[J]. Int J Environ Res Public Health, 2020, 17(6):1828. DOI:10.3390/ijerph17061828.

[48] Qureshi Z, Maqbool A, Mirza A, et al. Efficient prediction of missed clinical appointment using machine learning[J]. Comput Math Methods Med, 2021, 2021: 2376391. DOI: 10.1155/2021/2376391.

[49] Fotouhi S, Asadi S, Kattan MW. A comprehensive data level analysis for cancer diagnosis on imbalanced data[J]. J Biomed Inform, 2019, 90: 103089. DOI: 10.1016/j.jbi.2018.12.003.

[50] Riley RD, Snell KI, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: PART Ⅱ - binary and time-to-event outcomes[J]. Stat Med, 2019, 38(7):1276-1296. DOI:10.1002/sim.7992

[51] Archer L, Snell KIE, Ensor J, et al. Minimum sample size for external validation of a clinical prediction model with

a continuous outcome[J]. Stat Med, 2021, 40(1): 133-146. DOI:10.1002/sim.8766.

[52] Yokose J, Marks WD, Yamamoto N, et al. Entorhinal cortical Island cells regulate temporal association learning with long trace period[J]. Learn Mem, 2021, 28(9):319-328. DOI:10.1101/lm.052589.120.

[53] Vabalas A, Gowen E, Poliakoff E, et al. Machine learning algorithm validation with a limited sample size[J]. PLoS One, 2019, 14(11): e0224365. DOI: 10.1371/journal.pone.0224365.

[54] Wu JC, Pfeiffer RM, Gail MH. Strategies for developing prediction models from genome-wide association studies [J]. Genet Epidemiol, 2013, 37(8):768-777. DOI:10.1002/gepi.21762.

[55] Konuma T, Okada Y. Statistical genetics and polygenic risk score for precision medicine[J]. Inflamm Regen, 2021, 41(1):18. DOI:10.1186/s41232-021-00172-9.

[56] Zeng P, Dai J, Jin SY, et al. Aggregating multiple expression prediction models improves the power of transcriptome-wide association studies[J]. Hum Mol Genet, 2021, 30(10): 939-951. DOI:10.1093/hmg/ddab056.

[57] Visscher PM, Wray NR, Zhang Q, et al. 10 years of GWAS discovery: biology, function, and translation[J]. Am J Hum Genet, 2017, 101(1): 5-22. DOI: 10.1016/j. ajhg. 2017.06.005.

[58] Lu TY, Forgetta V, Richards JB, et al. Capturing additional genetic risk from family history for improved polygenic risk prediction[J]. Commun Biol, 2022, 5(1): 595. DOI: 10.1038/s42003-022-03532-4.

[59] Handorf E, Yin YN, Slifker M, et al. Variable selection in social-environmental data: sparse regression and tree ensemble machine learning approaches[J]. BMC Med Res Methodol, 2020, 20(1): 302. DOI: 10.1186/s12874-020-01183-9.

[60] Wei YY, Tian Y, Yu X, et al. Advances in research regarding the roles of non-coding RNAs in non-syndromic cleft lip with or without cleft palate: a systematic review[J]. Arch Oral Biol, 2022, 134: 105319. DOI: 10.1016/j. archoralbio. 2021.105319.

[61] Fu CY, Lou S, Zhu GR, et al. Identification of new miRNA-mRNA networks in the development of non-syndromic cleft lip with or without cleft palate[J]. Front Cell Dev Biol, 2021, 9: 631057. DOI: 10.3389/fcell.2021.631057.

[62] Fenlon C, O'Grady L, Doherty ML, et al. A discussion of calibration techniques for evaluating binary and categorical predictive models[J]. Prev Vet Med, 2018, 149: 107-114. DOI:10.1016/j.prevetmed.2017.11.018.

[63] Alba AC, Agoritsas T, Walsh M, et al. Discrimination and calibration of clinical prediction models: users' guides to the medical literature[J]. JAMA, 2017, 318(14): 1377-1384. DOI:10.1001/jama.2017.12126.

[64] Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction[J]. Circulation, 2007, 115(7): 928-935. DOI: 10.1161/CIRCULATIONAHA.106.672402.

[65] Musolf AM, Holzinger ER, Malley JD, et al. What makes a good prediction? Feature importance and beginning to open the black box of machine learning in genetics[J]. Hum Genet, 2022, 141(9): 1515-1528. DOI: 10.1007/s00439-021-02402-z.

[66] Wang K, Tian J, Zheng C, et al. Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP[J]. Comput Biol Med, 2021, 137: 104813. DOI:10.1016/j.compbiomed.2021.104813.

[67] Tseng PY, Chen YT, Wang CH, et al. Prediction of the development of acute kidney injury following cardiac surgery by machine learning[J]. Crit Care, 2020, 24(1): 478. DOI:10.1186/s13054-020-03179-9.

[68] Li WY, Song YN, Chen K, et al. Predictive model and risk analysis for diabetic retinopathy using machine learning: a retrospective cohort study in China[J]. BMJ Open, 2021, 11(11):e050989. DOI:10.1136/bmjopen-2021-050989.

[69] Nicora G, Rios M, Abu-Hanna A, et al. Evaluating pointwise reliability of machine learning prediction[J]. J Biomed Inform, 2022, 127: 103996. DOI: 10.1016/j.jbi.2022.103996.