

摘要

大脑年龄范式被证明越来越有助于探索与衰老相关的疾病，并可以预测未来重要的健康结果。大多数脑时代研究使用结构神经成像来索引脑容量。然而，衰老影响大脑结构和功能的多个方面，这可以使用多模态神经成像进行检查。使用英国生物银行，用 T1 加权磁共振成像、T2-FLAIR、T2*、扩散磁共振成像、任务功能磁共振成像和静息态功能磁共振成像对 $n = 2205$ 名健康人的大脑年龄进行建模。在一个持续健康验证组($n = 520$)，使用 LASSO 回归准确预测了实足年龄($r = 0.78$ ，平均绝对误差 = 3.55 岁)，高于单独使用任何一种方法。34 个神经影像表型通过回归(引导后)被认为是具有信息的；主要是灰质体积和白质微结构测量。当应用于英国生物库的新个体($n = 14,701$)时，发现中风史、糖尿病诊断、吸烟、酒精摄入和一些(但不是全部)认知测量与多模态脑预测年龄差(brain-PAD)显著相关 $p < 0.05$ 。多模态神经成像可以改善脑年龄预测，并且导出的脑 PAD 值对负面影响脑和认知健康的生物医学和生活方式因素敏感。

关键词:大脑老化，神经成像，多模态磁共振成像，英国生物银行，生物医学测量

1. 介绍

衰老对人类大脑有明显的影响，导致认知能力下降，增加神经退行性疾病和痴呆症的风险。尽管大脑老化无处不在，但个体之间的差异可能是巨大的。一些人在中年经历认知衰退、神经退化和与年龄相关的脑部疾病，而另一些人在 10 多岁时仍保留大部分认知功能(Deary 等人，2009 年;门德斯，2017;Wyss-Coray，2016)。

衰老过程中较差的大脑健康对个人、他们的家庭和社会都有明显的负面影响(Wasay 等人，2016 年)。因此，识别衰老过程中大脑健康状况不佳的人的方法已经成为一个重要的研究目标。一种这样的方法，所谓的“大脑时代范式”(科尔和弗兰克，2017;弗兰克和加斯，2019)，旨在定义大脑的“生物年龄”(杰克逊等人，2003 年;路德维希和烟雾，1980 年)。这里的观点是，遗传和环境因素可以影响与年龄相关的生物学变化的累积速度，而一个人的生物年龄可能比实际年龄更能预测疾病风险、功能能力和剩余寿命。通过定义健康大脑老化的统计模型，使用神经成像数据预测实际年龄，然后可以评估来自新个体的神经成像数据，并考虑他们的“大脑年龄”是否比他们的实际年龄更年轻或更老。

拥有一个看起来更老的大脑之前已经与生理老化(例如握力、肺功能、行走速度)和认知老化(科尔等人，2018 年)，说明大脑和身体老化是有关联的(科尔等人，2019 年 c)。大脑年龄较大也与不良的未来结果有关，包括从轻度认知障碍进展到痴呆(弗兰克和加塞，2012 年;Gaser 等人，2013 年)和死亡率(科尔等人，2018 年)。大脑老化也可以通过一系列不同的神经和精神疾病来缓解(科尔等人，2019 年 b)。

使用神经影像建立大脑年龄模型的统计方法通常是高度准确的，大多数方法能够解释实足年龄变化的 80%。预测的平均绝对误差(MAE)通常在 4-5 年左右(当预测 18-90 年之间时)，尽管最近一次竞赛的获胜者使用集成深度学习方法实现了 MAE < 3 年(<https://www.photon-ai.com/pac2019>)。然而，进一步提高模型的准确性是在临床环境中应用的重要一步，其中预测将在个体水平上进行。潜在的临床应用包括在认知正常的中年人中筛查较差的脑健康，在临床试验招募中分层或作为神经保护治疗的替代结果测量。

大多数脑年龄模型只使用 T1 加权结构磁共振成像，反映大脑体积。鉴于衰老会影响大脑结构和功能的许多方面，而且这些方面可以通过其他神经成像方式进行测量，人们可以通过纳

入大脑连通性、白质高密度、铁沉积以及任务和静息时大脑活动的补充数据来提高准确性。多模态方法的好处是为大脑老化的个体差异提供了更丰富和更全面的机制解释。例如，脑白质微结构的变化可能先于脑容量的变化，或者对年龄相关的大脑变化的易感性可能存在区域差异。包括更多的测量值应该捕获更多的与年龄相关的变化，从而提高模型的准确性。多模态成像也可以用来生成一系列特定模态的大脑“年龄”。例如，一个人可能有结构性脑龄、扩散性脑龄和功能连接性脑龄。根据生物年龄的概念，在健康人群中，这些独立的“年龄”应该是密切相关的，但对于患者群体，可能会出现异常大脑老化的不同模式。

先前的几项大脑年龄研究使用了 2 或 3 种模式(布朗等人, 2012 年,切鲁比尼等人, 2016 年,Groves 等人, 2012 年,Liem 等人, 2017 年,牛等, 2019,理查德等人, 2018 年).在第一个大脑年龄研究中, Brown 和他的同事(2012 年)使用 T1 加权的区域体积和皮质厚度测量, T2 加权的标准化信号强度和扩散 MRI 测量来预测神经发育期间的年龄(3-20 岁)。由此产生的多模态模型在这个年龄范围内高度准确(平均误差= 1.1 年), 尽管他们没有明确地将这个模型与其单峰成分的表现进行比较。Liem 及其同事(2017)研究了 $n = 2354$ 人, 发现结合 T1 加权结构 MRI 和静息态 fMRI (MAE = 4.29 年)时准确性提高。当独立使用时, 结构磁共振成像比功能磁共振成像具有更高的准确性。与此同时, Groves 及其同事(2012 年)使用关联独立成分分析来合并 $n = 484$ 名参与者的 T1 加权和扩散 MRI, 发现由此产生的最大成分可以高精度地预测年龄(实足年龄和大脑年龄之间的相关性 $r = 0.95$)。同样, Richard 和他的同事(2018)在他们对 877 名参与者的研究中使用了 T1 加权和扩散磁共振成像。此外, 各种形式的组合实现了最佳性能($r = 0.86$, 平均寿命= 6.14 年), 而 T1 衍生值比扩散磁共振成像的白质微结构测量值具有更高的年龄预测准确性。切鲁比尼及其同事(2016)使用了 3 种模式: T1 加权, T2 *弛豫和扩散磁共振成像数据。这项小型研究($n = 140$)取得了优异的成绩($r = 0.96$), 尽管单个模态的准确度仍然普遍较高。最近, 牛及其同事(2019)结合 T1 加权结构磁共振成像、扩散磁共振成像和静息态功能磁共振成像来预测年龄在 8-21 岁的 $n = 839$ 名参与者的年龄, 评估 36 种不同特征集和统计模型组合的性能。这项综合研究发现, 深度神经网络模型结合了所有 3 种模式(MAE = 1.38 年)的 MAE 最低, 尽管其他统计模型和 3 种模式中 2 种模式的组合达到了类似的性能水平(MAE 范围为 1.38-1.55 年)。

这些多模态脑年龄研究的结果表明, T1 加权 MRI 是最好的单一预测指标, 但是尽管各模态之间存在明显的共线性, 但加入其他模态比任何单一模态都具有更高的准确性。这些结果是有希望的, 但仍然局限于 2 或 3 种模式。这在很大程度上是由于数据的可用性; 英国生物银行成像研究(米勒等人, 2016 年)为克服这一限制提供了新的机会。随着使用 4 台相同的专用 MRI 扫描仪对总共 $n = 100,000$ 名个体进行标准化神经成像, 大量的多模态 MRI 数据变得可用。这些方法包括 T1 加权磁共振成像、T2-FLAIR、磁敏感加权成像(SWI)、扩散磁共振成像、任务功能磁共振成像和静息态功能磁共振成像。除了神经成像, 参与者还提供当前健康状况、生活方式和病史的详细信息, 以及参与认知测试和提供血液进行生物评估。

在这里, 我测试了健康人的大脑年龄预测是否可以通过结合这 6 种模式的数据以及评估它们的独立表现来改善。然后, 我将由此产生的多模态大脑年龄模型应用于英国生物库参与者的保留样本, 以测试大脑预测年龄与健康、生活方式和认知功能相关因素之间的关系。我假设 i)最高的准确性(即最低的 MAE, 最高的年龄差异)将通过结合所有模态的数据来实现; ii)T1 加权 MRI 将提供最高的独立准确性; iii)所有模式都可以用于独立地显著预测实足年龄; iv)吸烟、饮酒、主要的生理健康状况和较差的认知能力将与大脑老化有关。

2.材料和方法

2.1.参与者

分析了来自英国生物银行的 22,392 名参与者的数据。其中, $n = 20,237$ 人完成了脑部 MRI 评估; 该研究包括 $n = 2776$ 名缺少一种或多种神经影像方式的参与者。这使得 $n = 17,461$ 名参与者(年龄 45-80 岁, $n = 9274$ 名女性, $n = 8187$ 名男性)拥有完整的神经影像数据。然后定义了两个子集: 一个用于训练统计模型(训练集), 一个用于测试与其他英国生物库指标的关联(测试集)。训练数据只包括在扫描时符合健康标准的人。排除标准是 ICD-10 诊断, 自我报告的长期疾病残疾或虚弱(英国生物库数据字段#2188), 自我报告的糖尿病(字段#2443), 中风史(字段#4056), 自我报告的健康状况不佳或优秀(字段#2178)。这得出 $n = 2725$ (平均年龄 = 61.47 ± 7.2 岁, 1343 名女性和 1382 名男性)。测试集包括剩余的 $n = 14,701$ 名参与者(平均年龄 = 62.64 ± 7.5 , 7914 名女性, 6787 名男性)。所有英国生物银行的数据都是用 R package ukbtools(汉斯科姆等人, 2019 年)。

所有参与者都提供了知情同意书。英国生物银行获得了西北多中心研究伦理委员会(MREC)的伦理批准。有关英国生物银行道德和治理框架的更多详情, 请点击此处:<https://www.ukbiobank.AC.uk/the-ethics-and-governance-Council/>。本分析是在数据申请号 40933 下进行的, 优化大脑老化的神经成像生物标志物, 以识别不良大脑健康的遗传和环境风险因素。

2.2.数据采集

这里提供了英国生物银行神经影像数据的全部细节:https://biobank.ctsu.ox.AC.uk/crystal/crystal/docs/brain_MRI.pdf。简而言之, T1 加权磁共振成像使用 1 毫米各向同性分辨率的 MPRAGE 序列。T2 协议使用了液体衰减反转恢复(FLAIR)对比和 3D 空间优化读出, 分辨率为 $1.05 \times 1 \times 1$ 毫米。SWI 协议使用双回波 3D 梯度回波采集, 分辨率为 $0.8 \times 0.8 \times 3$ 毫米($TEs = 9.4, 20$ 毫秒)。T2 ρ 值(即信号衰减时间)是从 2 个回波时间的幅度图像中估算的。扩散磁共振成像数据通过 2 个 b 值获得($b = 1, 000, 2000 \text{ s/mm}^2$) 在 2 毫米各向同性分辨率下, 多波段加速系数为 3。对于两个扩散加权壳, 获得 50 个扩散编码方向(覆盖 2 个 b 值的 100 个不同方向)。任务和静息态 fMRI 使用以下采集参数, 空间分辨率为 2.4 毫米, $TR = 0.735 \text{ s}$ 秒, 多波段加速系数为 8。每个参与者的总神经成像采集方案持续了 32 分钟。

2.3.数据处理

当前分析中使用的数据是由英国生物库(米勒等人, 2016 年)并通过数据展示(<http://biobank.ctsu.ox.ac.uk/crystal/index.cgi>)。这些数据是 T1 加权 MRI、扩散 MRI、T2-FLAIR、SWI(即 T2*)和任务 fMRI 的可用汇总指标。来自静息状态 fMRI 的数据是所谓的批量数据, 在这种情况下是 25 维偏相关矩阵(数据字段#25752)。这些矩阵已被转换成长度为 210 的向量, 表示来自 21 个独立的静息状态网络的粗体时间序列之间的成对相关性(详情请参见第 16 页:https://biobank.ctsu.ox.AC.uk/crystal/crystal/docs/brain_MRI.pdf)。25 个网络中的 4 个被英国生物银行的研究人员归类为噪音, 并在发布前被删除。210 个值的数字是通过 21 个网络乘以 20(即, 排除身份相关性), 然后除以 2(即, 因为矩阵是对角对称的)得到的。使用偏相关的目的是提供比使用全相关更好的网络间直接“连接”强度的估计。

最终的一组神经影像表型编号如下: T1 加权 = 165, T2-FLAIR = 1(白质高信号的总体积), T2* = 14, 扩散-MRI = 675, 任务 fMRI = 14, 静息态 fMRI = 210。这总共给出了 1079 种表型(参见表 A1 在附录中)。

2.4.统计分析

所有统计分析均使用 R 版本 3.5.2(r 核心团队, 2015).包含所有统计代码的 R Markdown 笔记本可在此处找到:<https://james-cole.github.io/UKBiobank-Brain-Age/>。

为了定义一个多模态健康老化模型, 训练数据被随机分成单独的训练集(80%, $n = 2205$)和验证集(20%, $n = 520$), 以确保模型的准确性能够以一种无偏见的方式进行评估。所有的神经影像表型都被标准化(即, 通过标准偏差和均值居中来衡量)以说明不同模式使用的不同测量尺度。

为了从神经影像学数据预测年龄, 进行了最小绝对收缩和选择算子(LASSO)回归, 以年龄作为结果变量, 神经影像学表型作为预测因子。套索回归施加了 L1 范数惩罚, 其目标是最小化模型中 β 系数的绝对值。收缩到阈值(λ)以下的系数自动设置为零。这导致稀疏解, 这减少了模型中的方差(即, 正则化); 实际上是努力选择有用的特征, 去掉无用的特征。为了获得最佳值 λ , 首先使用一系列 λ 进行 10 倍交叉验证, 在随后的分析中使用最小值的一个标准误差内的最大值。

由于 LASSO 易受多重共线性导致的有偏解的影响, 因此使用了自举(即替换重采样)来生成每个预测变量的系数分布。这里, 运行了 1000 次引导, 并计算了平均系数和 95%置信区间(使用基本引导方法计算)。具有不与零重叠的自举 95%置信区间的神经影像表型被认为是预测年龄的有用信息。

然后对每个模式(T1 加权、T2-FLAIR、扩散-MRI、SWI、任务 fMRI、静息态 fMRI)进行套索回归程序, 其中仅包括该模式的表型。这一次没有进行引导, 因为目标是跨模态比较模型级性能, 而不是识别每个模型中的重要特性。接下来, 使用留一个模态出来的方法重复套索, 与单一模态分析相反, 以测试包括一些模态是否会降低模型性能。

为了测试英国生物库中大脑年龄和健康相关变量之间的联系, LASSO 模型被用于预测测试集中人们的年龄。最近的研究强调了大脑年龄计算中的比例偏差, 由此实际年龄和大脑预测年龄之间的差异与实际年龄负相关(乐等, 2018;梁等, 2019;史密斯等人, 2019 年), 使用了年龄偏差校正程序。这需要计算训练集中年龄(预测值)和大脑预测年龄(结果)之间的回归线, 然后使用该线的斜率(即系数)和截距来调整测试集中大脑预测的年龄值(通过减去截距, 然后除以斜率)。在应用年龄偏差校正后, 计算脑预测年龄差(脑垫);从大脑预测的年龄中减去实际年龄。这给出了一个以单位年为单位的结果值, 正值代表看起来更老的大脑, 负值代表看起来更年轻的大脑。

然后使用线性回归模型测试了脑垫值与人口统计学、生物医学、认知和生活方式测量值之间的关联。在这些模型中, 脑垫是结果测量, 兴趣变量是预测因子, 年龄, 年龄²性别、身高、从 T1 加权 MRI 到标准(数据字段#25000)的体积缩放, 以及平均任务 fMRI 头部运动(跨空间和时间点平均; 数据字段#25742)作为协变量。选择这些协变量是由于面部有效性(即, 理论上可能与大脑结构相关)和与脑垫的统计相关性。

测试的生物医学测量值(来自第 2 次随访, 成像随访)为收缩压、舒张压、体重、体重指数(身体质量指数)、臀围、糖尿病诊断; 中风诊断和面部衰老。生活方式测量包括吸烟状况、饮酒频率、每天适度活动或剧烈活动的持续时间。认知表现测量是流体智力分数、路径制造任务:

完成数字路径路径 1 的持续时间、完成字母数字路径路径 2 的持续时间、矩阵模式完成:正确解决的谜题数量、回答每个谜题和塔重排花费的持续时间:正确的谜题数量。关于这些变量编码的全部细节在英国生物银行数据展示中报告(<http://biobank.ctsu.ox.ac.uk/crystal/search.cgi>)。特别值得注意的是,酒精摄入频率的编码值较低,反映了较高的摄入量(1 = 每天或几乎每天; 2 = 每周 3、4 次; 3 = 一周一到两次; 4 = 一个月一到三次; 5 = 仅特殊场合; 6 = 从不; 3 = 最好不要回答)。对这 18 个单独的测量进行了多次测试校正,使用了错误发现率校正(本杰明和霍克伯格, 1995 年)。

3. 结果

3.1. 训练和测试集中健康参与者的统计人口数据

在将英国生物库细分为健康训练集和独立测试集后,这两组具有广泛的可比性(表 1)。测试组的年龄略高,女性比例也较高。这些组在血压、身体质量指数、体重和臀围方面是模棱两可的。通过设计,测试组的中风史和糖尿病诊断率较高(分别为 1.4% 和 5.7%),因为这些是健康训练组的排除标准。

3.2. 多模式神经成像可以预测健康人的实际年龄

当将脑模型模型应用于验证数据集时,实足年龄和脑预测年龄之间的相关性为 $r = 0.786$, $R^2 = 0.618$, $MAE = 3.515$, 年龄偏差 $r = 0.65$ (图一)。

从最初输入的 1079 个神经影像表型中,有 221 个在 LASSO 回归模型中设置为非零。在引导之后,34 个神经影像表型具有不重叠零的 95% 置信区间(表 2)。尽管 1079 个神经影像表型中有 886 个与年龄显著相关 $p < 0.05$; 即使使用 Bonferroni correction 704,神经影像学也与年龄显著相关(调整 $p = 4.6 \times 10^{-5}$)。

这 34 个“信息性”神经影像表型主要来自 T1 加权磁共振成像或扩散磁共振成像,尽管包括 T2、任务功能磁共振成像和静息功能磁共振成像变量。假设变量在各种模式间进行了缩放,则可以考虑模型中的相对贝塔系数。具有最大绝对系数的变量是灰质体积(根据头部尺寸标准化) $\beta = 1.466$, 95% 置信区间(1.85, 1.24),表明较低的灰质体积与较高的大脑年龄估计值相关。小镊子中较低的加权平均细胞内体积分数(ICVF)也与较高的脑年龄估计值相关,这表明胼胝体前部和相关区域的神经突起密度在较老的大脑中较低。第四脑室体积的增加,相当于脑脊液空间的扩大,也与较高的脑年龄估计值相关。其他值得注意的关联包括体积的减少和小脑扩散指标的变化。

3.3. 单一模态脑年龄预测

来自每种神经成像方式的变量分别用于训练模型以预测验证数据集中的年龄(表 3)。表现最好的模式(即,与年龄的相关性最高,MAE 最低)是 T1 加权磁共振成像和扩散磁共振成像。其他 4 种模式只能解释有限的年龄差异,尤其是功能性磁共振成像。每种模式预测的不同“年龄”之间的相关性模式如所示图 2。2 个“年龄”之间的最高相关性在 T1 加权和扩散 MRI 之间($r = 0.73$)。

3.4. 排除一种形态的大脑年龄预测

为了进一步探索不同模态对大脑年龄预测的影响,我进行了一个留一模态分析,本质上是上述单一模态方法的逆。这种分析背后的动机是测试是否包括一些模态实际上降低了模型性

能。这里，每个模态依次被省略，LASSO 回归在剩余的 5 个模态上训练，然后在验证数据集上测试(表 4)。与使用所有 6 种模式相比，去除任务功能磁共振成像或静息态功能磁共振成像对表现没有明显影响。尽管排除 T1 加权表型也降低了模型的准确性，但是排除扩散 MRI 数据时，表现下降最大。值得注意的是，当排除单一模态时，模型性能从未得到改善。

3.5. 大脑预测-年龄、健康、生活方式和认知表现

当应用于来自英国生物银行的 $n = 14,701$ 测试数据集参与者时，性能类似于验证集中的性能： $r = 0.803$ ， $R^2 = 0.644$ ，平均寿命 = 3.555 年(图 3a)。一个明显的年龄偏差是明显的($r = 0.640$)，使用大脑预测年龄关系的年龄斜率(0.59)和截距(24.7)进行了调整。调整后，年龄偏差最小(图 3b)。

使用偏差调整的脑垫值，评估与所选结果测量的关系(表 5)。每项分析都使用线性回归模型，根据年龄、年龄²、性别、身高、体积 T1 加权 MRI 标度和头部运动。致使 p-数值错误-发现率得到纠正(18 次测试)。经过多重比较校正后，增加 brain-PAD 与以下因素相关：高血压、糖尿病诊断、中风史、过去或现在吸烟、饮酒频率较高、液体智力较低、完成字母数字路径轨迹 2 的持续时间较长、完成正确的矩阵图案拼图较少以及正确的塔重排拼图较少。与人体测量、面部衰老、身体活动水平和一些认知任务没有关联。

4. 讨论

使用英国生物库，通过结合来自 6 种不同神经成像模式(T1 加权 MRI、T2-FLAIR、T2*、扩散-MRI、任务 fMRI、静息态 fMRI)的数据，可以准确预测健康人的实足年龄。在这个联合模型中，T1 加权 MRI 和扩散 MRI 表型通常是最能提供年龄预测信息的。Bootstrapping 强调了 34 个变量的信息，正如假设的那样，多模态模型优于任何独立模态的预测模型。这表明许多与年龄相关的变异可以被这 34 种表型单独捕获。在独立的模式预测因子中，只有 T1 加权磁共振成像和扩散磁共振成像达到了合理的预测准确性。与 T1 加权磁共振成像将提供最准确的独立模型的预期相反，扩散磁共振成像的性能稍高。当将多模态模型应用于保留的测试数据时，异常脑老化(通过脑垫量化)与高血压、中风史、糖尿病、吸烟、频繁饮酒和认知能力差相关。

大多数现存的脑年龄研究仅使用 T1 加权 MRI(科尔等人, 2019 年 b)，尽管以前的多模态研究使用了 2 或 3 种模态(布朗等人, 2012 年, 切鲁比尼等人, 2016 年, Groves 等人, 2012 年, Liem 等人, 2017 年, 牛等, 2019, 理查德等人, 2018 年)。多亏了英国生物银行，我能够组合和比较 6 种不同的模式。正如预期的那样，T1 加权 MRI 被证明对脑年龄预测很重要，标准化灰质体积是最具信息性的神经影像表型。LASSO 模型中的扩散 MRI 表型也提供了信息，壳核中的 T2，面孔和形状任务中的 BOLD 反应，以及来自静息态 fMRI 的独立成分也包括在内。脑白质高信号的体积(来自 T2-FLAIR)没有被模型保留。从生物学、技术和统计学的角度，有几个原因可以解释 T1 加权和扩散 MRI 表型的重要性。首先，表型反映了神经病理学研究显示的随年龄变化的生物学过程，特别是脑容量损失和髓鞘变化(彼得斯, 2002 年, Savva 等人, 2009 年);因此，地面-真相关系是强有力的。其次，这些模态具有相对较高的信噪比，尤其是 T1 加权 MRI(陆等, 2005, Perthen 等人, 2008 年, Polzehl 和 Tabelow, 2016 年)，这意味着与年龄相关的信号不太可能被与数据采集相关的因素混淆。第三，T1 加权和扩散 MRI 表型在数据集中比其他模式多得多(例如， $n = 675$ 个扩散 MRI 表型，而 T2-FLAIR 表型 $n = 1$ 个)。这意味着不仅有更大的解剖学特异性(因为许多表型是局部 ROI)，而且通过简单的概率，这些表型有更多的机会检测数据中的年龄相关信号。

使用 bootstrap($n = 1000$)得出置信区间为所选特征的稳健性提供了保证, 尽管 34 种信息表型中每一种的贡献效应相对较小。有趣的是, 多模态模型的预测性能($r = 0.79$)低于一些使用多种模态的先前报告(例如, 切鲁比尼等人, $r = 0.96$), 尽管训练样本量大得多($n = 2205$ 对 $n = 140$)。目前使用的衍生汇总测量(即, 影像衍生表型)可能会对脑年龄预测性能产生负面影响, 因为这些汇总测量将复杂的高维信息压缩为单个值, 潜在地消除了大量与年龄相关的差异。事实上, 单独使用体素方式 T1 加权的 MRI 数据的脑年龄预测性能往往非常准确(例如, $r = 0.96$ 科尔等人, 2017 年 a)。未来的工作将使用来自每个英国生物库模态的原始神经成像数据, 以更高维度的方式融合多模态数据。

当前研究的一个重要发现是, 尽管 1079 个神经影像表型中的许多与年龄显著相关, 但这一大样本($n = 2725$)与年龄相关 $r \geq 0.05$ 将是重要的 $p < 0.05$ 。因此, 许多这些统计学上显著的效应是如此之小以至于可以忽略不计, 只有 34 种表型被选择作为年龄预测的信息。这意味着, 尽管大脑结构和功能的多种模式在统计学上与年龄相关, 但只有少数测量可能适合于获得具有足够临床实用性精确度的个体化大脑年龄预测。与其他研究一样, 对于大脑时代来说, 重要的是要注意新数据中统计显著性、效应大小和预测能力之间的差异。

另一个值得注意的问题是获得神经影像表型的处理方法的有效性。例如, 在当前的研究中, 皮质下灰质体积表型以两种方式产生; 首先使用 FSL(帕特诺德等人, 2011 年)和 FSL 快(张等, 2001), 后者具有应用于组织分割图像的皮层下区域掩模。尽管理论上测量了相同的潜在结构, 例如, 第一丘脑体积和快速丘脑体积之间的相关性仅为 $r = 0.44$ 且 $r = 0.41$, 分别用于左侧和右侧。这种内部一致性的缺乏凸显了使用英国生物银行等“大数据”资源的研究人员需要仔细考虑数据处理的细微差别, 并明确选择表型, 而不是盲目分析所有可用数据。当使用这些数据解释当前和其他作品时, 这是一个重要的警告。正确的验证不仅来自于使用真正独立的数据集, 还来自于不同的处理方法, 以收敛到相同的构造(即, 大脑区域分割), 从而克服每个单独方法的假设和误差。

通过仅使用健康的人进行训练, 新个体与这种健康的大脑老化模型的偏差可以被索引。这一关键点使得在广泛的包容性人群样本上训练的模型和在健康人群上专门训练的模型之间有所不同。对于后者, 目标总是改进模型预测, 而对于前者, 改进的模型预测可能意味着可以准确预测任何人的年龄, 但偏差的概念失去了价值。以前的研究并不总是在模型训练期间筛选健康的人, 或者至少没有报道过这种情况(史密斯等人, 2019 年, 孙等, 2019)。这意味着潜在的脑-年龄差异度量(即, 脑-年龄差距、脑-年龄增量)不反映与健康脑老化的实际偏差, 因此可能对随后与其他测量的关系不太敏感, 或者更容易出现假阳性。目前的研究有一个详细的生物医学数据的大样本, 允许对健康训练数据集进行彻底的筛选。虽然这些健康参与者中确实有一些可能有未被发现的或前驱病变, 但这是所有病例对照研究的情况, 只能通过长期随访来正确评估。由于英国生物银行的研究设计, 这在未来也将成为可能。

看起来更老的大脑与更高的舒张压和收缩压、中风史、糖尿病诊断、吸烟、饮酒和认知表现的某些方面相关。身体质量指数、体重、臀围、面部衰老和认知表现的其他方面与脑垫无关。这些调查结果与以前的报告部分一致。弗兰克及其同事(2014 年)报道的脑容量(相当于脑垫)与舒张压和身体质量指数的关系; 这里只复制了前者。肥胖以前被认为与大脑老化有关(Kolenic 等人, 2018 年, 罗南等人, 2016 年), 而身体质量指数已被证明能影响大脑结构(科尔等人, 2013 年), 虽然这里没有发现身体质量指数和脑垫的关联。锻炼持续时间与脑垫无关,

这与 Steffener 及其同事(2016 年)。然而,许多不同的方法来衡量身体活动是可能的,目前的自我报告措施不一定是最有效的。

关于病史, Franke 及其同事(2013 年)发现糖尿病与脑容量增加= 4.6 年相关,而这里糖尿病使脑容量增加 2 年(协变量调整后)。中风也与脑垫增加相关= 2.6 年;有趣的是,在 1936 年洛锡安出生队列中,糖尿病和中风都与脑垫无关(科尔等人, 2018 年)。当前和以前的工作之间的差异可能是由于采样偏差,如群组效应或招募策略,尽管统计能力肯定有利于当前的研究,这要归功于英国生物库的大样本量。另一个原因可能是所使用的分析策略,尤其是协变量的使用。这里,多个协变量用于增加对脑垫关系的敏感性,最显著的是头部运动(在任务 fMRI 会话中平均),它占脑垫变化的 4.7%。虽然几个生物学、生活方式和认知因素与脑垫有关,但影响范围通常很小。例如,糖尿病、中风、当前吸烟(+2.1 年脑垫)和每日酒精摄入(+0.97 年)对表观脑年龄(科尔等人, 2020 年, Franke 等人, 2010 年)。尽管直觉上认为这些因素与衰老相关的大脑健康状况不佳有关,但这种影响的程度不应被夸大。未来的工作将调查基因-环境的相互作用,以更好地了解衰老过程中病史和生活方式对大脑健康影响的个体差异。

流体智力先前已经与脑垫(科尔等人, 2018 年),因为在追踪任务中的表现(科尔等人, 2015 年,科尔等人, 2017 年 b),其他研究报告了大脑老化和认知表现之间的中度显著关系(Liem 等人, 2017 年,理查德等人, 2018 年)。在这里,我能够复制这种关系,尽管对大脑老化如何与认知老化相关的更详细的分析需要更全面的认知测试。面部衰老数据的出现提供了一个有趣的机会,可以测试脑垫和参与者自我报告之间的关系,即参与者是否觉得人们说他们看起来“比你年轻”、“比你老”或“年龄差不多”。虽然有趣的相关工作已经报道了主观“年龄”和大脑年龄之间的关系,但是没有检测到关系(Kwak 等人, 2018 年)。在目前的研究中,只有 1% 的受访者表示,人们认为他们看起来更老,这可能会质疑这一主观衡量标准的有效性。

目前的研究受益于英国生物银行的极大样本量和深入的生物学数据。这使得对健康个体的详细筛选能够包括在训练数据集中,以最小化疾病相关效应对大脑的影响,这种影响可能会混淆健康大脑老化的模型。尽管本文使用的来自英国生物库的神经影像样本中有 84% 具有至少一项 ICD-10 诊断、自我报告的长期疾病、不良的自我评估健康状况或中风或糖尿病病史,但剩余的 16% 仍留有超过 2500 名参与者进行训练和验证。此处的另一个好处是将 LASSO 回归与 bootstrapping 结合使用,将 LASSO 识别重要预测要素的能力(通过将无信息变量收缩为零)与 bootstrapping 的稳健性相结合,从而克服了使用 LASSO 与高度相关的预测变量的局限性。这项研究的一个弱点是使用了概要水平的神经影像表型。如上所述,模型精度大大低于体素数据通常达到的精度。另一个潜在的限制是目前对生物学、生活方式和认知变量的有限使用。潜在的更多与生物、心理和行为参数的关联可以被评估;然而,增加更多的测试会增加多重测试的负担,并降低对实际效果的敏感度。因此,决定只测试具有先前研究证据和强表面效度的变量。另一个弱点是,目前的研究是横向的;因此,大脑老化的长期后果无法测试,也无法进行因果推断。计划对英国生物银行参与者进行的长期医学随访使其成为未来研究的一个明显途径。

5. 结论

总之,可以通过结合 6 种不同的 MRI 模式来预测脑年龄,最强的预测因子是 T1 加权和扩散 MRI 表型。健康大脑老化的偏差与病史、吸烟和酒精摄入有关。较差的认知表现也与大脑老化有关,这表明生物学和生活方式因素会对大脑老化和认知老化产生负面影响。大脑年龄

模式提供了一种潜在的筛查工具,用于检测医学和生活方式因素对无症状人群大脑健康的负面影响,英国生物银行计划的纵向性质提供了验证这种潜力的机会。