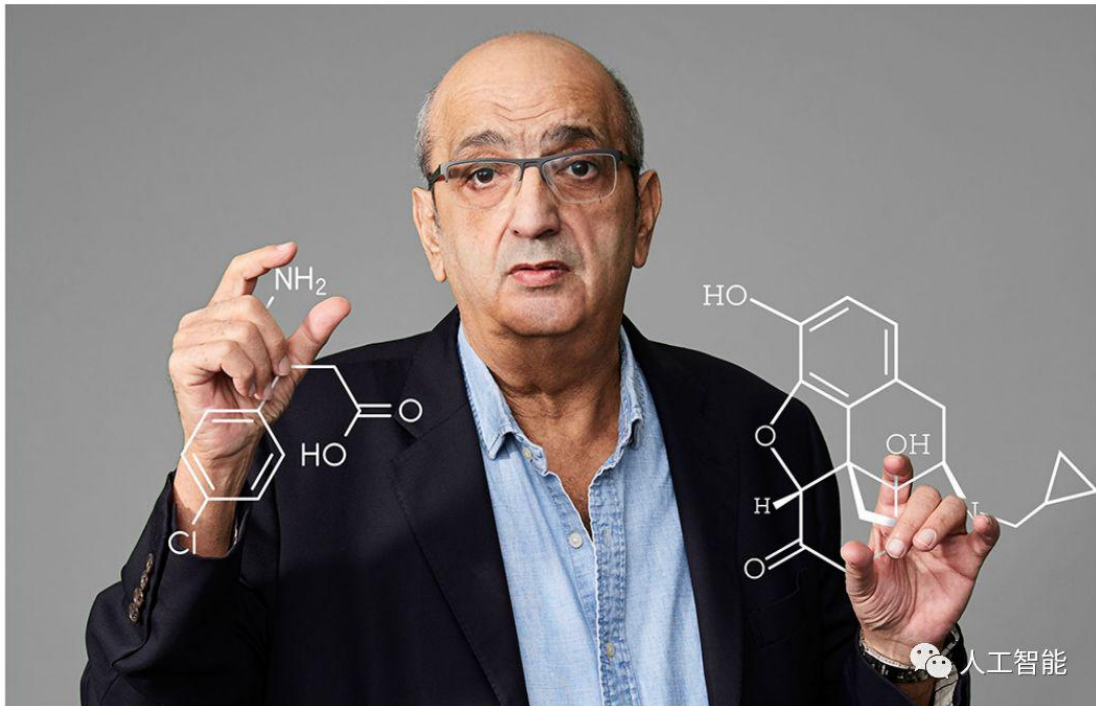


只需 50 种药就可以治疗所有疾病？人工智能显威力

人工智能3月27日

Pharnext 等医药初创公司正借助机器学习来研究“专利过期”药物的新医疗用途。



开处方：Pharnext 的首席执行官丹尼尔·科恩和巴氯芬及纳曲酮分子结构图——这两种药物包括在该公司的创新疗法中。图片来源：Photograph by Roberto Frankenberg for Fortune

在曼哈顿中城瑞典教会狭长的哥特式建筑中，丹尼尔·科恩在雅致而安静的咖啡馆里介绍了一番仿制药后稍事休息。他走到靠近前门的老旧钢琴前坐了下来，然后流畅而完美地弹奏了电影《绿野仙踪》中的插曲《飞越彩虹》。

如果说人体生物学是科学领域里的复杂琴谱，那么科恩就是懂得如何弹奏它的行家。他曾经是法国非营利实验机构 Généthon 的主导人物，后者在 1993 年 12 月绘制了历史上的第一张人类基因组图谱。可以说他是大数据和自动化进入基因研究的“引路人”，因为他和他的团队率先证明了有可能用超高速计算机来加快 DNA 样本的处理速度。

全世界科学家的工作都以科恩提出的思路为基础，而作为免疫学专业医学博士，他本人已经在研究者和医药公司高管的位置上取得了成功。然而，25 年过去了，基因学几乎没能像许多初期创新者期望的那样实现一些改变规则的医学突破。今天，作为巴黎药物初创企业 Pharnext 的首席执行官及创始人，科恩正在全力探究为何这道彩虹的末端不是一罐金子。

科恩离开钢琴走回来，然后对我说：“人体内任何一种蛋白质都有很多功能，而不只是一种。就像你，作为个体你在人群中有很多种角色，而不只是一个。”科恩描述的现象就是“多效性”，即单个基因可以发挥多种作用，而且这些作用看起来都没有关联。这也是疾病复杂性的一个体现，而且它已经多次挫败为最顽固疾病寻找治疗方法的医学研究者。

科恩不光了解多效性的意义，他还相信 Pharnext 等制药公司也许很快就能对其加以利用，而且是在人工智能的有力协助下。科恩希望通过参悟人体的复杂性，并通过用

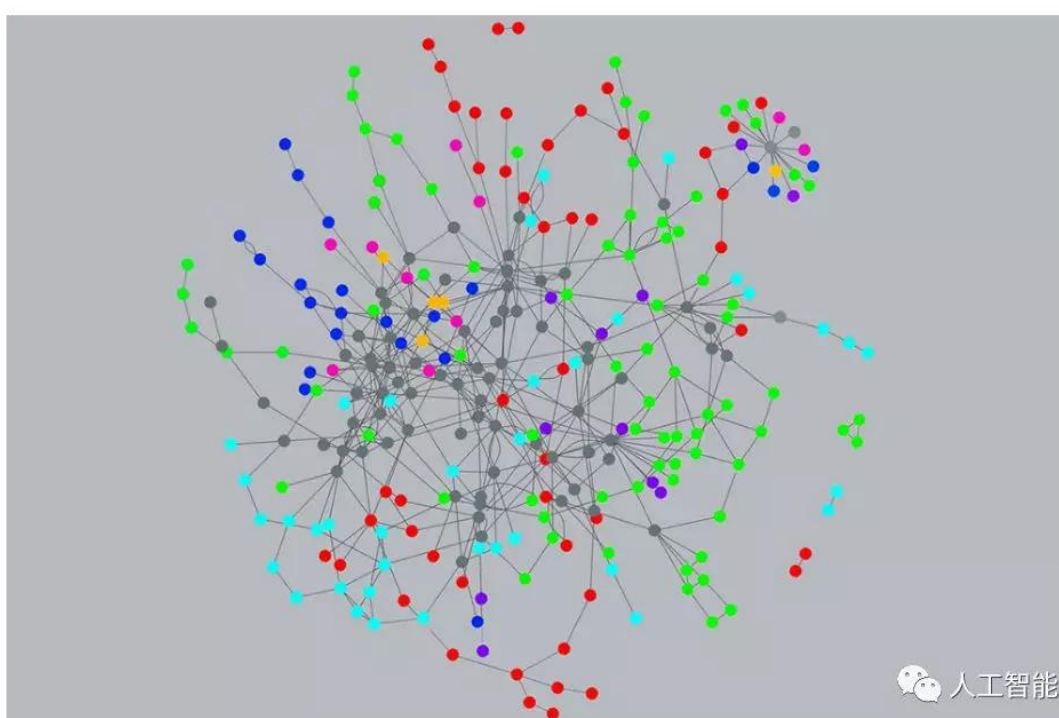
人工智能来更有条理地分析并绘制疾病横扫人体时的连锁反应路径来开发出适于众多疾病的药物组合。

科恩和他的团队还在用人工智能来搜寻“老药新用”带来的治疗方法，也就是将已有药物组合起来，从而产生单一药物不具备的疗效。他们的长期目标是用机器学习来理顺自己的后备药品，使其效率远远超过大型制药公司以缓慢著称的研发部门。说到如何实现这一切，科恩的瞌睡眼就会因为激动而睁大。他说：“这很棒，而且很经济。”

谷歌和 IBM 这样的大企业以及 Insilico Medicine、Recursion Pharmaceuticals 和 BenevolentAI 等初创公司都在与 Pharnext 赛跑。它们都为人工智能工具投入重金，并用这些工具来分析数以百万计的药物样本和患者数据，然后从中找出主要规律。但 2007 年成立的 Pharnext 领先大多数竞争对手好几年，要是考虑到过去几十年科恩率先在基因学和多效性方面进行的研究，它的优势就更明显了。

此外，十几年来 Pharnext 在解决医学难题时对人工智能的应用终于出现了关键拐点，这也许是最重要的。去年 10 月，该公司称一组药物的三期临床试验取得了积极成果。该组药物代号为 PXT3003，用于治疗腓骨肌萎缩症，这是一种罕见的神经退行性疾病，尚未找到治疗方法。该病症的主要诱因是某个基因的复制，而这样的突变会引发一系列“下游”问题，从而使保护神经的施万细胞退化为没有保护作用的干细胞，最终导致神经突触消失。这样人就无法控制肌肉，进而出现肌肉萎缩。

据 Pharnext 介绍，上述三期临床试验结果（尚未得到同行审核）表明 PXT3003 不光能稳定病情，而且还有逆转作用，证据是细胞开始再生。科恩指出，此前的疗法只能延缓病情恶化的速度。在服用 PXT3003 后，统计数据显示患者的两项残疾指标都有了明显好转。基于这些结果，美国食品与药品监督管理局（FDA）在今年 2 月将 FXT3003 纳入“快速通道”。这是一种加速审批程序，只有在 FDA 认为某种药品在治疗某种严重疾病方面展现出“卓越疗效”时才会启动。



通路：为找到治疗方法，人工智能加持的初创公司描绘了基因、蛋白质和人体在患病过程中的互动情况。上图是 Pharnext 对腓骨肌萎缩症的描绘。

当然，这只是针对某一罕见病迈出的有希望的一步。但科技让长期前景变得光明起来，它把药物设计过程缩短了好几年，让 Pharnext 走上了捷径。预临床测试和临床试验

通常需要 8 至 10 年，而从零开始开发新药可能让这个过程再延长 7 年，甚至更长时间。与之相反，就 PXT3003 而言，人工智能帮助 Pharnext 挑选了三种现有药物进行再利用，包括治疗肌肉痉挛和肌肉紧张的巴氯芬、解除阿片类药物依赖的纳曲酮以及用于降低血糖水平的缓泻剂山梨糖醇。由于这些药物都已经得到使用，Pharnext 就可以跳过通常需要保证药品安全性的一期临床试验，也就是那个“从零开始”的阶段。

FDA 的“快速通道”提高了 PXT3003 最早在 2020 年上市的可能性，而这只是 Pharnext 众多项目中的一个。该公司很快就会启动阿尔茨海默病治疗药物的第二次二期临床试验，以及肌萎缩侧索硬化治疗药物的首次二期临床试验，这两个项目也都采用了类似的老药新用方法。

同样重要的是，如果这些试验取得成功，资金实力更强的效仿者将接踵而至。曾经在奥巴马政府中担任卫生及公共服务部部长的凯瑟琳·西贝利厄斯如今是多家医疗保健公司的顾问和董事。她认为这些工作是不断增强的老药新用趋势的一部分。西贝利厄斯说：“所有这些都可能让投资周期大幅缩短，而且有可能形成非常不同的定价水平，并为经济动力一直不足的罕见病带来许多可能性。这非常有吸引力。”哥伦比亚大学 Kavli 脑科学研究所的主任埃里克·坎德尔曾经获诺贝尔生理学或医学奖，他也是 Pharnext 的顾问。坎德尔认为 Pharnext 走在了这种趋势的最前端，其方法“既有原创性，又很强大”。至于这样的方法能否流行起来，坎德尔说：“我们很快就会知道。”

在现代基因研究的“黎明”中，几乎没有人预见到疾病生物学的极端复杂性。当时，许多研究者都把基因视为人体的某种指导手册。Celera Genomics 的首席执行官及创始人克雷格·文特尔和美国国立卫生研究院的院长弗朗西斯·柯林斯等先驱人物被喻为“基因捕手”，这个称号鼓励着改革者们在全球各地搜寻那个“银子弹”基因，也就是可以解释并有助于治疗某种疾病的基因。

从某个角度来说，这些研究者发现了真正的宝藏。举例来说，基因学家南希·韦克斯勒花了好几年时间在委内瑞拉拼凑亨廷顿病患者的族谱。这是一种罕见的遗传疾病。她的工作让人们发现可以通过某个基因的突变来预测一个人会不会患上这种疾病。

但科学家很快意识到基因图谱不太像一本指导手册，而更像是宜家家具随附的那种配件目录。此外，研究人员发现，其他“目录”给基因和疾病的关系增添了复杂的变量，比如蛋白质组、DNA 编码蛋白、转录组以及所有将 DNA 转换成蛋白质的核酸。

在“黎明”过后的“早晨”，失望带来了真正的痛苦，因为研究人员认识到癌症和阿尔茨海默病等复杂病症并非由单个基因引起（就连已经确认诱发基因的亨廷顿病也依然无法治愈）。现在，科恩等人发现了简化执念和药物发现减少之间的联系。药物发现减少的具体表现包括 FDA 批准的新疗法只有十分之一的成功率、螺旋式上升的药物开发成本（塔夫茨大学最近的一项研究被称为“26 亿美元的药丸”）以及少量突破性新药直线上升的价格，比如诺华制药的白血病治疗药物 Kymriah，它的每个疗程的费用为 47.5 万美元。

最近，研究人员开始在网络科学的帮助下应付生物的复杂性。这门科学的首席传播者是美国东北大学的教授艾伯特·拉斯洛·巴拉巴西。他在 2014 年出版的《链接》(Linked) 一书让一种观念流行开来，那就是网络理论可以对很多领域做出解释，比如流行趋势，比如性关系，再比如疾病。巴拉巴西等人意识到疾病就像一个坏信号，它通过基因-蛋白质-细胞-组织这样的关系网络传播，直到所有这些“微小变化”展现出与疾病类似的症状。

复杂疾病是多种因素的共同作用，原因是多效性意味着任何蛋白质都可以在人体的不同部位发挥作用。Pharnext 等初创公司认为药物也有多效性，它们能同时对人体内一种以上的蛋白质产生影响，并进行一种以上的互动。为找到能够应对复杂疾病的药物组合，可以在数据中发现规律而且蕴含着巨大能力的机器学习必须与病理结构概念结合起来。

这就要求计算机科学家和生物学家之间的关系出现“进化”。较新的机器学习方法容纳的数据远超以往，它们还可以将信息分级，从而突破关联这个框架。不过，让这些“深度学习”神经网络获得预测能力需要建立一些漂亮的算法。

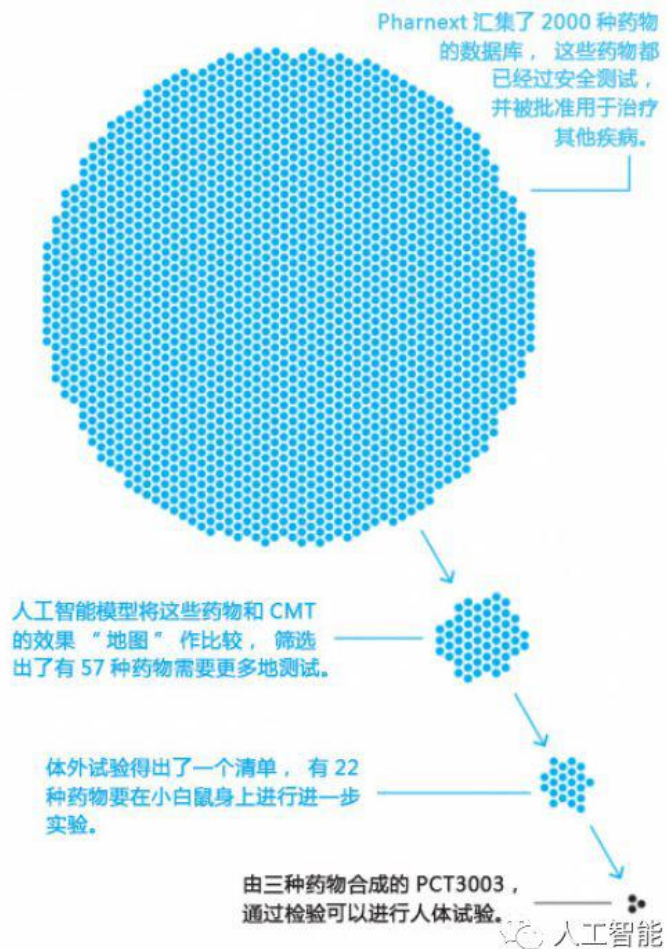
GNS Healthcare 的 CEO 及创始人科林·希尔就是这种算法的塑造者之一。这家设在马萨诸塞州坎布里奇市的公司用了 18 年时间开发出名为 REFS 的计算机系统，意思是“逆向工程、正向模拟”(reverse engineering, forward simulation)。近年来，GNS Healthcare 共筹集资金 3800 万美元，投资方包括制药巨头安进公司旗下的安进风投和赛尔基因等诸多机构。GNS Healthcare 的目标是构建并调整其疾病模型。

在最近的一系列试验中，该公司详细介绍了 REFS 在治疗帕金森病等方面的潜力——多效性因素让现有的帕金森病治疗手段变成了撞大运。GNS Healthcare 的首次试验报告发表在 2017 年的《柳叶刀》医学杂志上。

就帕金森病而言，基因缺陷触发的互动网络具有特定的形状，而运动机能的分解是病情发展的最可靠指标。将帕金森病患者和控制组成员的基因数据输入 REFS 帮助 GNS Healthcare 生成了 100 多个计算机模型，其作用是描述运动机能退化过程中可能发生的情况。这些模型可以发现此前未被人知晓的基因变异，而后者可能就是运动机能加速退化的原因。

THE BIG SORT

Pharnext 利用人工智能模型来搜索现有药物的数据库以寻找新的治疗组合开发过程，这一过程可以缩短几年的药物开发时间。 以下就是 Pharnext 如何帮助公司开发 PTX3003（一种治疗 CMT 的化合物）的过程。



但这只是第一部分。GNS Healthcare 已经用这些发现在计算机上模拟了 5000 次随机控制试验，这些模拟都是为了预测采用不同治疗方法时病情发展的速度。和寻求同样结果的控制人体实验相比，此类“速度测试”的经济性要高的多。同时，已经和其他药厂联手的 GNS Healthcare 目前正在申请用类似的方法来对付糖尿病、肌萎缩侧索硬化、多发性骨髓瘤和乳腺癌等等疾病。

作为 CEO，科林·希尔说：“我们现在可以在电脑上创建、进行并展示患者及其病情，这样就可以用不同的药物和不同的护理管理干预方法逐一进行检验，然后找出哪种治疗方法对哪些患者起作用。”

换句话说，这样的模拟不光是要找到关联，它还回答了如果.....会怎样的问题。如果我们用这种药替代那种药，那么这个病人的情况会如何？这种模拟以及就反向操作得出结论的能力是人工智能应用的一项近期成果。它正在变得越来越重要，而这在很大程度上要归功于该公司的技术顾问朱迪亚·珀尔，后者长期从事人工智能研究，而且是加州大学洛杉矶分校的计算机科学教授。在去年出版的热门书籍《为什么》（The Book of Why）中，珀尔描述了怎样从简单的模式感知中形成真正的智能，机器学习可以辨别出大量的模式，而真正的智能可以基于这些模式进行反向推理并得出结论。脱离了机制观念，光是数据无法做到真正的洞悉。珀尔认为：“数据在因果关系方面非常笨。” 希尔则更直接地说：“深度学习并没有那么深。”

今年 67 岁的丹尼尔·科恩在突尼斯长大，那时他处在犹太人、基督教徒和穆斯林构成的多元化社会中，“他们用非常雅致而且和平的方式生活在一起”。他说这样的经历让他形成了“事物是复合而非复杂”的概念。9 岁时科恩随家人移民到了巴黎，并开始热情地学习弹钢琴。后来他意识到，自己作为科学家的影响可能比音乐家大，这让他立刻转投医学，但保持着同样的热情。他曾经是伦敦皇家交响乐团的客座指挥，梦想着率领该乐团演奏柴科夫斯基的《悲怆交响乐》。他开玩笑说：“交响乐团指挥、CEO 和科学家体质都由同样的基因控制着。”

此前科恩就曾经把基因学和技术转化为药物方面的成功。他同别人一起创立了 Millennium Pharmaceuticals，这是一家肿瘤药物公司，参与开发了多发性骨髓瘤治疗药物 Velcade。

科恩乐观地认为 Pharnext 能通过人工智能取得成功，但他也清楚这项技术的局限。谷歌的人工智能程序 AlphaZero 可以击败最顶尖的人类围棋选手，而且没有使用此前人类的任何围棋知识。但就像科恩所说，围棋有固定规则，而 AlphaZero 也完全明白这些规则。在生物学中，人们还没有完全掌握相关规则，而且或许永远也掌握不了，这在一定程度上要归咎于多效性。



图片来源：Photograph by Roberto Frankenberg for Fortune

但设计周密的人工智能已经让 Pharnext 能够围绕已知规则建立模型并做出相应的选择。在 10000 种已知药物中，该公司的发现模型用其中的 2000 种进行搭配，这些药物的专利已经过期，而且已经得到推广，或者说其疗效和安全性都足以向公众销售。

为开发腓骨肌萎缩症治疗药物，Pharnext 首先用了约一年时间来构建这种疾病的网络模型，或者说和 GNS Healthcare 帕金森病图谱类似的框架，从而说明相关基因突变如何引发了一连串神经和肌肉问题。基于这样的机理，计算机模型挑选出了 57 种

候选药物，分别针对这一连串问题中的各个点。Pharnext 对这些药物逐一进行体外测试，再从中选出 22 种在老鼠身上试验的药物，最终形成由三种药物构成的组合进行临床试验。三期临床最近公布的积极结果证明 PXT3003 正在这一连串问题的多个点上发挥作用。

科恩说，如果没有人工智能模型，预临床测试需要的时间就会多出好几年，而不是 Pharnext 所花的三年，“从 2000 种药物[中着手]，我就可以得出所有可能的组合，即 10 亿种可能性”来进行体外测试。这个“处方”中有无数的错误测试结果和死胡同，意味着好几年的挫折——现在这些都得以避免。

Pharnext 在巴黎上市，去年 10 月公布上述三期临床结果后其股价以上涨了一倍多。10 年来，该公司已经在研发上投入约 1.2 亿欧元（约 1.35 亿美元）。就制药行业的标准而言，这个数字着实不多。Pharnext 尚未盈利，但分析师估算，如果 PXT3003 上市，该公司的收入将从 2020 年开始飙升。2018 年 Pharnext 实现收入 900 万欧元（GNS Healthcare 未上市，没有披露过开支或收入）。

除了投资者可能获得成功，Pharnext 和 GNS Healthcare 的进步还为人工智能以及与之共同成长的药物学指明了发展方向。理解因果关系以及反向探索问题的能力是人工智能用户长期以来一直想跨越的门槛。随着它们掌握并控制众多让人眼花缭乱的变量，这些初创公司的计算机模型正在朝着这个方向迈进。

就连疾病的基本定义都可能发生改变。随着科学家不断加深了解，这些定义已经变得过于简单。学术期刊《Bioinformatics》在去年的一项研究指出，癌症中的基因变异存在“根本性差异”，而这对癌症治疗产生了不利影响——看似是同一种疾病，或者同一类疾病，而实际上患者之间的相似之处寥寥无几，区别却有很多。人体内的任何变化都绝不等同于单个基因的“开/关”引发的一系列症状，弄清楚了这一点，科技就可以帮助从业者克服药物开发的复杂性。

最后但同样重要的是，这些人工智能推动的工作从经济角度带来了一丝希望。在这个时代，药物开发成本已经成为令人畏惧的障碍，智能算法则有可能在某一天让医药利益相关方从投入药物研究的数万亿美元中获得更多的价值。科恩判断，理论上，通过老药新用，“人们就不需要再设计[新]药物。我的感觉是可以用 50 种药物来治疗所有的疾病。”这就意味着另一个词的定义也要变了，那就是“发现”。（财富中文网）

本文另一版本刊登在 2019 年 4 月出版的《财富》杂志上，题为《在老药中找到新疗法》。

译者：Charlie

- END -