

医学 + 数据科学，必将大有可为

——访清华大学生物医学影像研究中心研究员、
数据科学研究院 RONG 兼职教授赵锡海

◆ 本刊记者 杨璐

在数据时代，相比于传统的统计学分析方法，使用数据科学领域的算法模型或图像识别等技术将能辅助医学专家提高医学研究效率和精准度。作为从始至终支持数据院在医疗健康大数据方向研究工作的医学专家，清华大学生物医学影像研究中心研究员赵锡海早就认识到数据科学对于医学研究的重要性，也一直在践行和推动“医工结合”的创新性合作。

“我在几年前就接触到数据院了，在一场医疗主题的会议上听到韩亦舜老师关于医疗大数据的思考和数据院在相关领域的实践，正巧我也早就有意结合大数据技术来做医学研究，所以我和韩老师一拍即合。”赵锡海老师在回顾与数据院的合作初衷时这样说道：“一般来讲，医学上用的传统统计学分析方法只适用于样本量小、数据变量不多的研究，反之则需要大数据和人工智能等技术从另一个维度挖掘数据的价值。从临床医学的角度，我可以提出一些可能有价值的研究方向，并且知道做什么分析对临床有用，数据院会帮助选择可结合数据技术的课题以及相应的技术合作伙伴和资金支持。我们合作的研究项目都非常成功，有了数据技术的支持，大大提高了研究的准确度和效率。”

“医工结合”带来的不只是数据技术对医学研究的影响，同样会反过来推进数据技术的优化，这是一个互促互进的良性循环。韩亦舜老师邀请

赵锡海老师参与医疗健康大数据方向的研究，不仅是在具体项目上的合作，还有一些战略性、方向性的深层次讨论。赵锡海老师可以提供医学专业的知识、讯息，并积极推荐合适的临床医学专家参与课题设计和开展合作，这些对数据院在医疗大数据方向的研究、规划和未来发展提供了很大的助力。

“在进行医疗数据研究时，我会从临床医学的角度设定研究方向，具体的技术方法和实现途



清华大学生物医学影像研究中心研究员、数据科学研究院 RONG 兼职教授赵锡海

径则需要和数据科学领域的专家合作，共同完成课题。”针对这种跨学科、跨领域的合作，赵锡海老师认为决定合作能否成功的要素有以下三点：

首先，这种合作一定要基于共同的兴趣。这种兴趣点一般是大家在互相交流的过程中讨论出来的，兴趣是做研究的一个原始动力。合作双方如果没有共同的兴趣，则无法开展实质性合作性研究。

其次，合作的双方都要学习对方领域的专业知识，以增强凝练更高水平的科学假说和关键问题的能力。这种学习不仅是简单的聊天，更需要互相提供一些培训的机会。当合作双方对彼此的需求和能力有一定的了解和思考后，才能进一步探讨深入合作的研究方向和合作方式。这种互相交流学习的机会不能限于举办学术讲座或论坛，更为高效的做法是不断地开展小范围的学术讨论，大家需要坐到一起，进行面对面的交流和头脑风暴。

最后，赵锡海老师提到很重要的一个要素是在各方合作过程中，要共同遵循一些规则和规范，具体体现在合作项目的“科研合作协议”，该协议会对合作各方在科研活动中起到一定的约束作用。事实上，清华大学各个研究机构都有自己的科研协议，其中包括合作双方或者多方人员各自的责任和义务；知识产权、成果分享的原则；一旦出现纠纷的仲裁办法等等。然而，以前的协议中有些条款已经不太适应现阶段数据科学研究中产生的新问题。“我认为清华数据院需要牵头推出一系列的适应数据科学时代特点的跨学科跨领域的科研合作协议模板。相信这也会是全国领先的壮举。”赵锡海老师对此也表达了对数据院的期待。

在数据院与清华大学临床医学学院合作搭建清华临床医学科研数据平台（以下简称：医学数据平台）期间，赵锡海老师也全程深度参与。医学

数据平台是一个医疗数据共享的平台，希望将医院的医学专家和校内的工科专家凝聚在一起，共同打造一个医疗数据共享资源池，实现医疗数据的汇集、整理和储存，并为高校和医学界的研究者提供研究素材，更好地发挥医疗数据价值。

“没有人做这件事，也没有人敢做这件事，这绝对开创了国内的先河。”赵锡海老师这样评价数据院带头搭建医学数据平台。从平台建设之初，他就被韩亦舜老师邀请参与到平台建设中来，他非常了解建设这一公共平台的意义：“现在大量的医疗数据堆积在医院或医疗机构，他们不能丢，也不敢用；学校的老师需要真实的数据用于医学研究，他们有需求，但是要不到数据。必须有一个人勇敢地站出来去打破这一僵局，数据院在做的这件事真的够大胆，如能做成，将为社会创造极大的价值。”

赵老师冠以数据院“大胆”二字，是因为他清楚平台搭建过程中的每个环节都面临挑战：数据来源环节，医疗数据涉及大量隐私，没有人敢拿出来；同时，医疗数据没有明确的归属权，想要使用数据是否需要跟每一个患者签订相应的协议，这又是一个具有挑战性的问题；数据处理环节，需要与拥有数据的部门逐一沟通，怎样打通壁垒实现数据融合、怎样给医疗数据脱敏、怎样识别复杂的医疗数据都需要付出大量的人力和精力；研究成果环节，还涉及一个现实的问题，由医学、工科等专家跨学科、跨领域合作，所得出的研究成果如何分配。医疗大数据现在是一个热门的研究方向，我很高兴越来越多的人开始重视医工结合的合作模式。在表达自己的疑虑后，他同样也给关注医疗大数据的人们打了一针强心剂：“虽然机器完全替代医生是难以实现的，但并不能否定数据科学技术对医学将起到变革性的影响。我们合理地使用数据科学技术，一定能在医疗健康领域发挥重要价值。”