

DOI: 10.13288/j.11-2166/r.2014.19.001

学术探讨

大数据技术与中医现代化

黄欣荣^{*}, 张艳朋

(江西财经大学管理哲学研究中心, 江西省南昌市双港东大街 168 号, 330013)

基金项目: 国家社会科学基金资助项目 (14AZX006)

摘要 正在兴起的大数据技术革命给中医现代化带来了新机遇。大数据通过重视整体性、提倡多样性、凸显相关性的大数据思维为中医现代化提供了切实可行的方法论, 通过量化世界与数据挖掘为中医知识系统化提供了可以操作的技术手段, 通过数据规律的合法化为中医现代化提供了合理的实现路径。

关键词 大数据; 中医现代化; 中医方法论

中医的科学性与现代化问题是中医哲学讨论的永恒话题, 也是中医未来发展的重要问题。中医属于复杂性科学, 从复杂性科学的路径可以将中医纳入现代科学体系^[1]。但是从复杂性科学视角只解决了中医的科学性问题, 究竟怎样将中医纳入现代科学体系? 中医走向现代化的路径在哪里? 这些问题只有通过技术手段才能解决, 而正在兴起的大数据技术革命正好为中医走向现代化提供了有效的技术手段。

1 大数据与中医的发展机遇

中医是中华民族的传统医学, 具有数千年的悠久历史, 为中华民族的健康事业做出了巨大贡献, 其疗效已被医疗实践反复证明。然而随着西方科学的传入以及西医的普及, 中医的科学性在近百年来颇受争议, 甚至被批评为“不科学”而被排斥, 更有极端者提出要彻底废除中医。其原因是简单地照搬近现代科学的标准来判断中医的科学性。值得注意的是, 西方科学近年来同样遇到了极大的困难, 在遇到“南墙”之后已经开始了自我批判和反思, 并提出了复兴整体论、走向复杂性的主张。中医学天然地具有整体性、多样性和实践性等复杂性科学的特征, 是复杂性思想的朴素表现, 并为现今复杂性研究提供了一个活样板, 中医也因复杂性科学的兴起而找到了融入国际科学共同体的道路。

但是试图通过复杂性科学来实现中医的现代

化, 还会遇到不少困难, 复杂性科学目前主要还停留在思想观念的转变上, 即在哲学思想、科学思维上形成了从简单性向复杂性的转变。在技术层面上, 目前还没有具体的实现途径, 特别是面对特别复杂的研究对象, 还缺少解决问题的技术工具。正因如此, 目前复杂性科学被人嘲笑为“混杂学”, 甚至被人说成是不可证伪的“伪科学”^[2]。我们虽然不能接受“伪科学”的论断, 但复杂性科学的不成熟是显而易见的。中医试图用复杂性科学范式来完成其现代化的过程, 还有漫长的道路要走, 其中首先要完成的就是必须找到技术化的实现路径。

数与人类文明紧密相关。人类在古埃及、古巴比伦时期就发明了数, 并且知道用数和量来描述事物的属性。古希腊哲学家和数学家毕达哥拉斯曾提出过“数是万物本原”的观点, 由此把数上升到形而上的层次。西方科学的历史基本上就是一部数据化的历史。最先进入数据化阶段的是天文学和力学, 其他各门学科也紧随其后。正因为有了精确的数据, 才有随后西方科学的证实或证伪的科学检验过程。正如马克思所说, 一门科学只有成功地运用数学时, 才算达到完善的地步。只有实现了数据化, 才能逐渐实现精确化、形式化和标准化等科学化的过程。

令人遗憾的是, 西方近代科学的数据化主要停留在对客观自然世界的数据化, 而对于更复杂的人类思想、行为和人类社会等主观世界, 数据化的过程却举步维艰, 因此就出现了这样的局面: 对自然

^{*} 通讯作者: 32478179@qq.com, (0791) 83816923

界和自然科学,主要运用定量研究方法,而对人类和社会科学,则主要运用定性研究方法。中医一直保持着整体性,停留在模糊、经验的研究阶段,其中主客体没有分开,渗入了行医者的主观性判断、个性化经验,因此,量化过程遇到了极大的困难,经过几千年的发展依然停留在只可意会不可言传的“默会知识”(tacit knowledge)阶段。

不过,人类与人类行为的数据化困境近年来有了突破性的改变,这就是大数据技术的出现和大数据时代的来临。所谓大数据(big data),从字面来说就是数据量特别巨大的数据库,但数据规模特别巨大(volume)只是其特点之一,它还具有数据类型丰富多样(variety)、数据采集处理迅捷(velocity)、数据蕴含无价之宝(value)等特点^[4]。从哲学观点来看,大数据主要是一种世界观,它认为世间万物皆可用数据来表示,世界的本质就是数据^[5]。与传统的小数据相比,大数据不但能将客观世界数据化,借助当代的智能设备、互联网、物联网等技术手段同样能将主观精神世界数据化。原来难于数据化的人类思想、行为、经验以及人类社会的历史等等,如今都可以用大数据来描述。这样,具有模糊性、经验性特点的中医就有了数据化的技术工具,也就有了走向现代化的途径。因此,大数据技术革命给中医的现代化带来了前所未有的机遇。有了大数据技术,中医的数据化和现代化就有了走向现实的可能性。

2 大数据思维与中医方法论

大数据首先在思维方式上发生革命,使得在大数据时代看问题的视角以及分析、解决问题的方法与小数据时代相比有了巨大的不同。大数据时代思维方式变革主要表现出其整体性、模糊性、多样性和相关性等几个方面的特点^[6],并由此奠定了中医走向现代化的方法论基础。

2.1 整体性

系统科学和复杂性科学都一再强调整体性的重要性,而中医一直重视生命的整体性,注重人体各部分之间的联系和协调,所以中医使用望、闻、问、切等整体诊断手段,不会像西医那样通过人体解剖来打开人体黑箱。整体论的愿景十分美好,它希望保持完整,不丢失任何信息,但很快就会发现,没有具体的技术手段来实现这种愿景。还原论看起来有点残忍,将整体剖开,丢失了不少信息,特别是让有生命的东西失去了生命活力,但它是一种技术上可行的路径。大数据将万事万物数据化,

通过收集所有数据,从大数据中挖掘知识,寻找规律,从而使整体论变成了数据化、可操作的整体论^[7]。大数据认为,通过云存储、云计算等手段,我们有能力收集、处理海量数据,而不必从海量数据中只抽取一小部分进行处理,即用部分来代替整体,因此,大数据提出了要全部数据而不要随机抽样的主张^[8-9]。用全部数据来描述整体,那么整体就变成了一个数据化的整体,因而也变成了可进行技术处理和操作的整体,而不是传统整体论那种抽象、不可分的整体。通过数据化的整体论,中医就有了进入科学阵营的入场券。

2.2 模糊性

大数据时代不必像小数据时代那样对数据从一开始就提出精确要求。在小数据时代,数据的采集和处理成本都比较高,因此,在数据收集之前就提前筹划并制定标准,只采集那些精确、标准的数据,而那些不符合要求的数据被提前抛弃以免浪费存储、处理的资源。然而属于默会知识的中医存在大量不精确、不标准的描述,例如中医方剂的所谓“少许”、“适量”、“若干”等都是模糊语言,传统科学对此无法进行科学处理。大数据通过对全部数据进行挖掘处理,能够将模糊数据中反映的信息挖掘出来。大数据认为,在海量的数据中难免有些数据不精确,而且现实生活中的许多现象本身就具有模糊性。传统科学苦于没有合适的处理手段以至于将中医划入非科学或伪科学的范围。如今的大数据认为,在大数据时代并不追求过去的精确性,而是乐于接受那些反映事物真实状态的模糊数据^[10-16]。因此,大数据为中医提供了处理模糊描述的技术手段,为中医的现代化开辟了道路。

2.3 多样性

自古希腊开始,科学就一直在追求统一性。古希腊哲学家和科学家们认为,世界是一个统一的世界,它表面上丰富和多样,但实际上都由某种或某些作为宇宙之砖的“始基”经不同的组合而构成的。所以他们都试图通过还原论来将丰富多彩的世界还原成一个冷冰冰的“始基”世界,而且所有学科都可以还原为最为基础的力学模型。中医理论很难通过还原论还原为力学模型,因此,难于纳入西方科学的统一体系中。然而复杂性科学开始接受多样性,承认世界是一个丰富而多样的世界,因而复杂性科学能够接受中医理论。正在兴起的大数据技术不但继承了复杂性对多样性的兼容态度,而且从技术上能够处理来自不同路径、不同模式、不同格式的各种数据^[17]。因此,大数据为中医这种一

直被西方科学特别是西医视为异端或伪科学的地方性知识纳入世界科学的海洋铺平了技术的道路。

2.4 因果性

西方科学的最大特点是寻求因果关系。从古希腊哲学家寻求万物始基开始,追问“为什么”就成了一种传统和习惯。正因为要寻求现象之间的因果关系,所以西方科学喜欢还原论,喜欢将整体的事物逐渐分解、解剖,发现其中的内部机制,揭示它们之间的因果关系。因此,当人们生病治疗的时候,医生们首先要弄清楚致病的因果关系,找到并排除病灶,因为西医遵循的是简单、线性的因果联系。但中医师只用“望、闻、问、切”几个简单的手段了解患者的外在特征,了解病症的表象,这看似不符合因果关系,也很难得到科学的解释,但大数据的兴起为中医找到了解释的理论依据。大数据认为,事物之间的相关性比因果性更重要^[7]。一个数据的变化会引起另一个数据的变化,我们就说这两个数据具有相关性,也就是说,具有相关关系的变量之间会有伴随的变化,至于其内在的原因我们不得而知或者没必要知道。中医很早并一直就使用如今大数据提倡的相关分析方法来认识、分析人体及疾病,大数据让中医原来看似不科学的方法成了先进的科学方法之一。

3 数据挖掘与中医知识生产

科学知识从何处而来?这是科学哲学一直关心的重大问题,也被称为科学发现的逻辑问题。西方近现代科学,包括西方医学之所以发展迅速,就是因为它们找到了科学知识生产的合适途径和方式,找到了科学发现的逻辑通道。

文艺复兴之后,近代科学走上了快速发展的道路。哲学家们在总结科学知识的生产途径之时,出现了两大学派,即唯理论和经验论。唯理论认为科学知识是从作为大前提的原有正确知识中推导演绎而来,而经验论则认为一切知识来源于人类经验,特别是科学实验。现代科学哲学在近代科学认识论的基础上,出现了逻辑实证主义的“科学始于观察”和证伪主义的“科学始于问题”的分歧^[8]。“科学始于观察”强调了科学实验及其观察结果的重要性,从客观观察中归纳、提炼出科学的结论,而“科学始于问题”则强调了问题意识,需要研究者大胆猜想,充分发挥主观能动性,而猜想正确与否则由科学实验来检验。

科学实验,特别是受控实验,在包括现代医学在内的西方科学发现中起着举足轻重的作用。所谓

受控实验,就是将研究对象与环境隔离,并且在试验之前就制定了明确的目标和实验手段,包括实验过程都提前预设好,甚至理论预设都已经准备好,实验的目的只是验证其设想的结论。中医跟西医的根本差别在于中医基本上没有受控实验,中医数据都是中医师在临床诊疗过程中取得的,可以说科研实验和临床诊疗合二为一。中医师不提前做理论预设,不提前做实验假设,他们只是收集实际诊疗的各种案例。中医师的每一次诊疗都是一次科学实验,一切过程都是顺其自然,没有任何地方有“受控”的成分或过程,因此,“真实世界中医临床科研范式”的提法准确地反映了中医研究的真实过程^[9]。

西医的“受控”和中医的“自然”各有优点。西医的研究和诊疗完全分开,研究是在理想条件下进行的,一切过程都有备而来,取得的各种数据都按预设标准进行了处理,因此,数据标准、精确,有利于后续的数据处理和模型建构。但西医的问题是理想条件和真实临床存在较大的差异,从科研到临床还有很大的一段距离。中医的研究过程反映了真实世界,诊疗和实验合二为一,科研和临床融为一体,从科研到临床没有距离^[9],但因为提前规划和部署,其取得的数据就千差万别,缺乏精确描述,因此,其取得的数据缺少精确性和标准化。

西方近现代科学知识生产的模式虽然有不同,但西方科学都强调理性、还原、实验等要素,而中医则特别强调经验、整体和实践,因此西方科学发现的逻辑路径很难容纳中医,或者说中医知识的生产难以用西方近现代科学解释模型来解释。那么中医知识究竟是如何产生的?它的生产模式是什么?它具有什么特点?正当中医知识生产模式问题陷入困境之时,大数据技术为中医知识提供了另一种可选择的生产模式,因为大数据技术让传统难于数据化的人类精神、思维、行为等,都可以实现数据化,给主观世界的客观化、数据化带来了可行的技术手段。与小数据时代不同,大数据时代的数据也像中医一样,不是从受控实验而来,而是从原先无意中留下的各种数据所构成的海量数据库中去挖掘出预先并未意识到的问题或规律。所以在大数据时代,数据在先,然后才有目的或用途。另外,大数据时代的数据挖掘技术更加先进,它并不要求数据必须结构化,因为它能够处理来自各种途径的非结构化数据。因此,大数据技术提供了“科学始于数据”这一知识发现新模式。从大数据角度来看,一切知识都必须从数据开始,除了上帝,其他

任何人都必须用数据说话^{[7] 20}。通过让数据发声所获取的知识更加客观、可靠,而且从看似无序的数据中找到秩序和规律。借助于大数据技术,可以快捷地处理来自不同中医师无意间留下的各种非结构化数据,并从中挖掘出中医知识或规律。也就是说,只要有海量的原始中医诊疗数据,就能够从中发掘出有用的知识或规律,非结构化的中医诊疗原始数据在大数据面前变成了无价之宝。

大数据是怎样从海量数据中发现规律的呢?答案是数据之间的相关性,即大数据从数据间的相关性中找出数据规律。由于数据量巨大,不可能像小数据时代那样去寻找每个数据的因果关系,只能通过数据挖掘技术去发现其中具有显著相关性的变量。相关关系通过识别有用的关联物来帮助我们分析一个现象,而不是通过揭示其内部的运作机制;通过找到一个现象的良好的关联物,相关关系可以帮助我们捕捉现在和预测未来^{[7] 72}。中医可以利用这种大数据挖掘技术,通过关联物来发现其中的关联关系,从中发现中医知识和规律,从而完成中医经验的科学化表述。

从大数据中挖掘出来的由相关关系所构成的知识或规律可以称之为数据规律。小数据时代通过变量之间内在的逻辑关系而得出的规律被称为因果规律。因果规律具有普遍必然性,而数据规律是一种经验规律,只是具有关联关系,因此,不一定具有普遍必然性。那么这种不具有普遍必然性的数据规律能被看作是真理吗?任何真理都需要经验的检验,只有经过经验的证实或证伪才能被称为真理。从科学哲学,特别是从波普尔的证伪主义观点看,任何科学真理都只是暂时的、相对的真理,未来都有可能被证伪,因此,根本不存在绝对的真理^[8]。从中医大数据中发现的中医规律就是属于数据规律,这些经验规律被中医师的实践反复证实或证伪,因此也可以说具有相对的真理。

4 大数据与中医现代性的路径

通过“科学始于数据”这条科学发现的逻辑新通道,中医找到了知识生产的新模式,中医从此可以从诊疗的原始数据中挖掘、提炼,发掘出数据化的科学规律。大数据成为中医走上科学之路、真正实现现代化的通道,并提供了具体的实现路径。

4.1 树立中医自信

自从西方近代科学技术特别是西医传入我国以来,中医的地位日益衰落。不仅是外行人对中医存在怀疑,而且内行的中医人也有人怀疑中医。随着

复杂性科学的兴起,中医的科学地位获得了充分的肯定。正在兴起的大数据技术不但再次肯定了中医的科学地位,而且为中医纳入未来科技体系提供了切实可行的技术路径。中医不能纳入西方科学体系,并不能说明它不科学,反而说明西方科学体系可能存在某些局限。如今兼容并包的大数据技术能够将存在于各民族的智慧进行技术处理,从而显现出未来科学是全人类集体智慧的集成。在复杂性科学与大数据技术革命的时代,中医获得了新生力量,因此,我们完全应该对中医有足够的自信。

4.2 建立大数据平台

要用大数据技术对中医进行数据化,就必须有中医大数据,因此建立中医大数据平台就成了中医现代化的重要一步^[9]。首先,中医文献资料必须数据化。中医具有数千年的悠久历史,积累了丰富的文献。过去由于这些文献都是先辈们个性化的经验积累,属于非结构化的默会知识,所以无法进行科学处理,也就无法纳入现代科学研究的视野。如今大数据技术可以对这种非结构化数据进行挖掘处理,所以中医文献就成了一座丰富的宝藏。不过要将这些记载在纸质媒体中的文献进行大数据技术挖掘,中医文献的数据化和网络化是当务之急。其次,中医临床方药必须数据化。中医医院每天都在开出大量的方药,诊治各种疾病,记录在纸质媒介上的病历、处方资料难以被数据化并被数据挖掘和共享。医院信息系统的日益普及为中医诊疗数据的积累提供了技术支持。最后,中医名医经验必须数据化。中医名医经验的传承工作日益受到重视,民间也留存了不少独特的诊疗方法,这些宝贵财富目前大部分都未实现数据化,无法实现数据挖掘与共享。如何尽早将中医默会知识数据化,是摆在我们面前的重要任务^[10]。

4.3 联合挖掘大数据

中医药经验性知识经过数据化、网络化,就变成了一个巨大无比的原始宝藏,但这些经验要变成可在不同主体间传递的具有规律性的知识,也就是说要把中医药大数据变成中医药科学知识,还有一项重要的工作要做,那就是数据挖掘。针对中医药大数据,仅仅靠中医药工作者是无法进行全面挖掘的,必须有中医药行业之外的人员参与其中,比如统计工作人员、信息技术人员等其他专业人员。通过中医药大数据的挖掘,发现中医药之间的显著相关性,就能将传统的中医药经验知识逐渐变成不同主体间可以交流的客观知识,从而让中医药逐渐走上现代化之路。

4.4 实现数据共享

传统中医药体系的一大特点是其私有性、隐秘性,疾病诊疗完全凭个人积累的经验,而且每个人都必须从头做起,因此难于继承和发展。特别是有些灵验方药被当作家族遗产秘而不宣,因此难于成为公共知识,并且有些宝贵经验就随着家族的衰落而被埋没。在大数据时代,开放、共享、合作成了国际共同的准则,那种秘而不宣的私有性与大数据精神是相冲突的。因此,为了让传统中医药发扬光大,为了中医药的科学化和现代化,就必须抛弃传统的狭隘观念,让中医药数据成为全世界共知、共享的公共数据。中医药工作者要有开放的胸怀和气度,通过开放、共享,才能实现中医药的国际合作研究和发

5 结语

正在兴起的大数据革命让中医有了走向现代化的技术手段和实现路径。借助于大数据技术,隐秘化、个性化和经验性的中医默会知识可以变成非结构化的数据。通过开放、共享和合作,借助古今中医药实践经验而构成的中医药大数据,从中可以挖掘出中医药的数据规律,并形成中医药科学知识体系。应当及时抓住这次大数据革命的机遇,制定中医

药大数据发展战略,争取早日实现中医药的现代化。

参考文献

- [1] 黄欣荣. 复杂性科学与中医 [J]. 中医杂志, 2013, 54 (19): 1621-1626.
- [2] 约翰·霍根. 科学的终结 [M]. 孙雍君, 译. 呼和浩特: 远方出版社, 1997: 285.
- [3] Zikopoulos PC. Understanding Big Data [M]. New York: Mc Graw Hill, 2013: 5-9.
- [4] 黄欣荣. 从复杂性科学到大数据技术 [J]. 长沙理工大学学报(社科版), 2014, 29(2): 5-9.
- [5] 黄欣荣. 大数据时代的思维变革 [J]. 重庆理工大学学报(社科版), 2014, 28(5): 13-18.
- [6] 黄欣荣. 大数据技术对科学方法论的革命 [J]. 江南大学学报(社科版), 2014, 13(2): 28-33.
- [7] V·M-舍恩伯格, K. 库克耶. 大数据时代 [M]. 盛杨燕, 周涛, 译. 杭州: 浙江人民出版社, 2013.
- [8] Wessler M. Big Data Analytics [M]. New Jersey: John Wiley & Sons, Inc., 2013: 6.
- [9] 邱仁宗. 科学方法与科学动力学 [M]. 2 版. 北京: 高等教育出版社, 2006: 51.
- [10] 刘保延. 真实世界的中医临床科研范式 [J]. 中医杂志, 2013, 54(6): 451-455.
- [11] 刘保延. 大数据绘制当代中医航海图 [N]. 中国中医药报, 2013-6-5(3).

Big Data Technology and Modernization of Traditional Chinese Medicine

HUANG Xinrong, ZHANG Yanpeng

(Jiangxi University of Finance and Economics, Jiangxi Province 330013)

ABSTRACT The emerging big data technical revolution has brought new opportunities for TCM modernization. By focusing on integrity, promoting diversity and highlighting relevance big data provides a practical methodology for TCM modernization. By quantifying the world and data mining big data provides operational techniques for TCM knowledge systematization. By legalizing data law big data provides a reasonable and legitimate path to achieve TCM modernization.

Keywords big data; TCM modernization; TCM methodology

(收稿日期: 2014-06-19; 修回日期: 2014-07-16)

[编辑: 李春梅]

欢迎订阅《国际中医中药杂志》

《国际中医中药杂志》为中华人民共和国国家卫生和计划生育委员会主管,中华医学会、中国中医科学院中医药信息研究所主办的国际医学系列杂志之一,系国内外公开发行的专业学术期刊。本刊以反映国内外中医药学研究的最新进展,促进国内外中医药学的学术交流为己任,以中医药科研、临床、生产、管理工作为主要对象,介绍中医药学领域有关的新进展、新技术和新方法。本刊主要栏目有专题论坛、国外传统医学、论著、短篇论著、综述等,报道内容涉及中医药动态、临床与实验研究、流行病学调查、质量标准研究、制剂与工艺、临床报道、专家经验、临证心得、思路与方法等。

《国际中医中药杂志》为月刊,16开,96页,每册定价12元,全年144元。国内邮发代号:2-611,各地邮局订阅;国外代号:BM6563,中国国际图书贸易集团有限公司(北京399信箱)订阅。也可直接汇款至本刊编辑部订阅。

地址:北京市东直门内南小街16号中国中医科学院中医药信息研究所《国际中医中药杂志》编辑部,邮政编码:100700。电话:010-64014411-3225。投稿网址: <http://gjzy.cintcm.com>, E-mail: guowaiyixue@aliyun.com。