

循证管理中“证据”的内涵及测量

颜士梅 梅丽珍

2013-05-22 13:50:00 来源: 《软科学》2012年11期

摘要: 基于以往文献,系统地分析了“证据”的内涵及测量的不同观点,即研究证据观、研究证据与个体经验证据结合观、以及多重证据观。最后指明了循证管理的未来研究方向。

关键词: 循证管理; 证据; 证据强度

作者简介: 颜士梅,浙江大学管理学院副教授,博士生导师,研究方向为人力资源管理与创业管理;梅丽珍,浙江大学管理学院硕士研究生,研究方向为人力资源管理。

管理者往往要在任务和时间的压力下利用有限信息做出重要决策。许多管理者在决策时通常会参照两类依据:一是过时信息或者自身经验,二是管理“妙方”或者其他公司的“最佳实践”。这两类依据都可能导致错误决策。而循证管理(Evidence-based management)则为寻找可靠的管理决策依据提供了新思路。循证管理的思想来源于20世纪80年代兴起的循证医学(Evidence-based medicine)和循证理论(Evidence-based theory),是循证理论在管理领域的应用和推广。循证管理具体是指将建立在最佳证据之上的管理原则转化为组织的管理实践的过程;其使管理决策从基于个人偏好和不系统的经验转变为基于最佳的科学证据。

近年来,循证理论及循证管理越来越受到管理领域学术界和实业界的关注。纵览以往相关文献发现,以往研究一般聚焦在推行循证实践(evidence-based practice)的障碍因素以及实现途径上;而明显欠缺对证据本身的内涵及测量的具体分析;另外,要在循证管理领域进行深入地实证研究,必须明确其中的核心概念——“证据(evidence)”的内涵及测量。同时,对证据内涵的明晰离不开对证据特征的探究,而证据强度(strength of evidence)是表征证据特征的重要概念,其在以往研究中被反复提及却相对缺少系统梳理和分析。鉴于此,本文基于以往文献,对证据的内涵、强度以及测量进行了系统总结和分析,从而为进一步地深入研究奠定基础。

一、证据的内涵

自20世纪80年代循证理论在临床决策领域的兴起,有关循证理论中证据研究已有近30年的历史,研究者对证据内涵的认识在不断延伸。概括这些认识和观点,可以发现关于证据的内涵主要存在三种观点:研究证据观、研究证据与个体经验证据结合观和多重证据观。

(一)研究证据观

英国流行病学家Cochrane(1972)最早阐述了研究证据观。他在其专著《疗效与效益:医疗保健中的随机对照试验》中首次提出将随机对照试验的研究结果作为临床决策的依据,并建议不断收集新的研究结果,从而为临床治疗实践提供可靠依据。该观点引起了学界对于研究证据尤其是随机对照试验的极大关注,也促成了研究证据观及循证医学的蓬勃发展。

循证医学作为概念本身是由加拿大麦克玛斯特大学(McMaster University)的临床流行病学专家们正式提出的,作为医学实践的新范式,循证医学强调临床决策不应当过分强调直觉、无章可循的临床经验以及病理生理,而应当基于临床研究的证据。可见,循证医学从最初提出到内涵界定都非常强调研究证据。

借鉴于循证理论在医学领域的发展,Hewison(1997)将它推广到了管理领域。针对医疗组织中的管理问题,他首次提出了循证管理的概念,即“管理者依照循证实践的原则来指导自身行动”;还进一步指出,医疗护理组织中的管理者在鼓励临床医生采取循证实践的同时,缺乏对自身管理实践的反思。后来,Stewart(1998)强调,管理者需要在组织内营造注重研究的氛围,鼓励质疑,培养对研究结果的兴趣,并尝试通过其找寻解决问题的方案。Young(2002)、Walshe和Rundall(2001)也都指出了在战略决策中使用研究证据的益处。Reay等(2009)则明确概括道:在过去50多年中,利用研究证据指导管理实践已经渗透到管理的研究中。另外,循证管理协会也坚持研究证据观点。循证管理协会是一个旨在缩小管理研究以及管理者决策之间鸿沟的组织,他们认为循证管理需要通过慎重地使用相关的最佳的科学证据以提高组织决策和实践质量。他们强调,研究证据虽然不是唯一依据,但却是最重要的。

由上可见,无论医学领域还是管理领域,许多学者都非常强调循证决策与循证管理的研究证据观点。这一观点强调实践者在决策中应该从基于个人偏好和不系统的经验转变为基于最佳的科学研究证据。

(二) 研究证据与经验证据结合观

随着循证理论的发展,证据概念的外延也在不断延伸,研究证据不再是行动的唯一指南。循证理论中有关证据内涵的观点的一个重要分支,即是研究证据与经验证据相结合的观点。这一观点是由著名临床流行病学家Sackett(1996)率先提出。他指出,临床医生应当在实践中结合个体的临床经验以及由系统研究得到的最佳外部临床证据。其中,个体临床经验指的是临床医生自身通过临床实践经历所获得的熟练度和判断力。最佳的外部临床证据指的是与临床问题相关的研究,通常来自于基础医学,尤其是以病人为导向的临床研究。鉴于此,证据的内涵已经由单一的研究证据扩展到了临床医生的个体经验。在管理领域,也有部分循证管理研究者更注重个体经验这一类证据,例如,Yates(1990)提倡使用更为有效的决策工具、技术,利用管理者的经验以及判断力来指导管理活动中的每项决策。

显然,在循证管理中,使用研究证据就具体情境做出决策时,个体的经验判断、对研究证据的批判性评价都是关键的。正如我国学者吕力(2011)认为,循证管理的中心内容就是评价,由于研究证据通常不具有普适性,需要主观参与的评价,因此个体经验在循证管理中的证据占据了重要的地位。

然而,在循证理论的发展历程中,结合个体经验与研究证据来指导决策和实践的观点也受到了越来越多的质疑。在决策情境的高复杂性以及客户期望的独特性的双重压力下,多重证据观随之诞生。

(三) 多重证据观

研究者们通过对循证理论的探究,对证据类型进行了进一步的扩充,使证据的内容逐渐拓宽至多重。多重证据观最早是由Sackett(1997)在其专著《循证医学:如何实践和教育》中提出并加以阐述的。他将循证医学定义为“慎重、准确和明智地应用所能获得的最好研究依据来确定患者的治疗措施”,他具体提出,临床医生必须将他们的个体经验运用到评价病人所面临的问题上,并在决策时结合研究证据以及病人的偏好、价值观。其中,病人的偏好和价值观不仅仅指的是病人对生命和健康的观点、信念、期望以及目标,还包括个体对不同方案的利弊、成本、便利程度的权衡。根据这一观点,证据的种类再次得到扩充,临床决策不仅仅应当基于研究证据、个体经验,也应当基于相关利益者(病人)的偏好。

在有关多重证据观的探究中,具有开拓性的是Haynes等(2002)的研究。他们提出一个循证临床决策的新模型,它基于病人所处的具体情境、病人的偏好以及最佳研究证据,而个体的临床经验作为核心贯穿于决策中。这就意味着,临床医生应当运用自身经验,获悉病人的个体信息,例如人口特征、病史等情境证据,并通过搜寻文献选择、记录评价最佳研究证据,最终获得符合病人期望的治疗方案。可见,Haynes等(2002)对证据的界定扩展至病人所处的具体情境。他们认为,病人所处的临床状态或者具体背景和情境对于临床决策是极其关键的,甚至占据了主导地位,具体的临床方案应当随着个体的不同而变化。

而在管理学领域，研究者们提出了符合管理学学科特点的多重证据观。在这一观点下，不同研究者强调的证据类型也不尽相同。Pfeffer和Sutton(2006)强调的是局部情境证据。他们认为，管理者收集、分析组织的局部证据，在决策时更多地关注具体事实信息，而较少地强调对外部研究证据的使用。有些学者则强调紧密结合研究证据与局部情境证据，认为两者在循证管理中同等重要，缺一不可。例如，Rousseau(2006)提出，循证管理中的“最佳证据”可分为大证据和小证据。“大证据”指的是通过科学方法获得的有关因果联系的普适性知识，即系统性的科学研究证据。“小证据”是局部性质的、有关特定组织的事实性的证据，指的是在特定情境下，管理者系统收集的用以指导局部决策的数据，即管理者听取的事实真相、获取的事实数据。

随着美国学界对循证管理讨论的激烈化，Briner等(2009)对循证管理的内涵进行了更加明确的界定，这使循证管理中“证据”的外沿也随之得到了延伸。他们提出了包含四个证据类型的多重证据观，即管理者的经验或判断、局部情境中的证据、对外部的最佳研究证据的系统评价以及可能会被这个决策所影响的人们的观点等四种证据。尤其强调恰当进行的系统综述(systematic review)能够明确总结与某个实践问题有关的已知信息和未知信息，这种研究证据是循证管理的基石。他们还强调循证管理是在管理者同时获得以上四种信息下才得以发生的。但至关重要的一点是，每一个证据元素的影响力大小将随着决策问题的不同而变化。在某些情况下，利益相关者或决策者的道德考虑可能比外部研究证据更重要，从而在决策中占据更大的权重。在其他情况下，由于可获得的局部证据很少，所以它对这项决策的影响相对较小。总之，在所有的案例中，应谨慎小心地确定对不同证据种类的重视程度。

从上述文献中可以看出，在近年的循证管理研究中，学者们对证据的界定范畴愈来愈广，由起初的仅仅包含“研究证据”延伸至包含“个体经验”、“利益相关者的偏好”乃至“具体情境证据”，从而形成了证据的多重观。

在证据内涵的不断扩展中，研究者及实践者面对的一个重要问题便是如何评价纷繁多样的证据，于是衍生出了对证据特征的探讨，这对于理解证据内涵具有重要的作用。由于证据强度是证据特征的一个重要表征，因此，有必要对有关证据强度的文献进行梳理。

二、证据强度

证据的特征与证据的内涵具有密切的关系，为了更好地理解证据内涵，必须理清表征证据特征的一个关键概念的含义，该概念为证据强度。

证据强度指研究证据“接近真实的程度”、“影响力的大小”和“在实践中的适用性”。它体现了实践者对研究质量的判断、对研究结论真实性的自信程度，以及研究结论是否与其他研究者的研究结论或者与面向不同人群展开的研究结论相吻合。从对该概念的界定和陈述中，不难发现证据强度主要用来表征研究证据的特征。

判断证据强度的标准包括：①内部效度(internal validity)，即研究本身的质量，具体是指该研究根据其特定的研究对象所得的结论是否有效；②外部效度(external validity)，即研究结论能够推广到更广泛群体的程度，具体是指指导解决问题的可行性和可推广性；③连贯性或一致性(coherence or consistency)，即在不同研究中所得出结论的一致性。

很多研究按照证据的强弱程度对研究证据进行划分。1979年，加拿大定期体检特别工作组(Canadian Task Force on the Periodic Health Examination, CTFPHE)首次对研究证据进行系统分级并提出推荐意见，由最强到最弱依次有三个水平。其中第一个等级是至少一项设计良好的随机对照试验。第二个等级是设计良好的队列或病例对照研究，尤其是来自多个中心或研究组；或者是比较了不同时间、地点的研究证据，或有重大结果的非对照研究；第三个是基于临床研究、描述性研究或专家委员会的报告，或权威专家的意见。

1986年Sackett教授在该分级基础上又提出了证据的五个等级：第一个等级是具有肯定结论(较低错误概率)的大样本随机控制试验；第二个等级是具有不肯定结论(有中度乃至高度的错误风险)的小样本随机控制试验；第三个等级是非随机同期对照试验(Nonrandomized, contemporaneous controls)；第四个等级是非随机历史对照试验(Nonrandomized, historical controls)；第五个等级是无对照的病例系列(No controls, case-series only)。2004年，包括WHO在内的19个国家和国际组织的67名专家组成GRADE工作组，又推出了国际统一的证据强度评级系统。

在管理学领域, Reay等(2009)依照D' Agostino和Kwan(1995)提出来的循证医学中证据的六等级思路, 提出了循证管理中证据的六等级, 即证据按照强度由强到弱一共有六个水平: 第一个水平是大样本随机控制试验或者元分析(meta-analysis)结论; 第二个水平是一个高质量的文献综述或系统综述(systematic review)结论; 第三个水平是有比较的、多点的案例研究或者大样本的定量研究结论; 第四个水平是小样本、单案例的定性或定量研究结论; 第五个水平是描述性研究和/或自我报告的故事; 最弱的第六个水平指权威意见机构或专家的观点, 非基于数据的证据。

由上可见, 研究证据具有强弱特征, 根据强度可以将证据分为不同等级或者水平, 高强度的研究证据在实践中往往具有更高可信度和更强的指导意义。

三、证据的测量

在科学研究中, 对一个概念的清晰界定最终要落实到对该概念的测量上, 因此, 为了进一步明确证据的内涵, 本部分将梳理和讨论证据的测量问题。

证据量表的开发自20世纪90年代起便得到相关研究者的普遍关注。纵览以往文献可以发现, 证据测量根据证据内涵的差异而有所不同, 具体可以分为三类: 基于研究证据的测量、基于研究证据和个体经验证据结合的测量、多重证据测量。

(一) 基于研究证据的测量

对研究证据进行测量的量表受到研究者广泛而持久的关注。这些根据不同情境开发出的量表, 在测量视角上也有所不同, 具体涉及研究证据的运用能力视角、研究证据的运用障碍视角和研究证据的运用行为视角。

就运用能力视角的研究证据测量而言, McColl等(1998)在对基于研究证据的医疗决策的相关研究中, 具体测量医生“对相关期刊、综述类等出版物以及数据库的意识和认可度”以及“进入相关数据库或者网络的能力”。Fritsche等(2002)也开发并验证了基于能力视角的研究证据量表。该量表测量了医生所掌握的有助于解读临床研究证据的基本知识、从临床难题中提取问题, 并做出解决该问题的最佳设计的技能以及利用已发表研究文献中的定量信息处理各种具体问题的能力。这些量表都强调从运用研究证据的能力来对研究证据进行测量。

就有关运用研究证据的障碍的测量而言, 被普遍认可的是Funk等(1991)构建的研究运用障碍量表(Barriers to Research Utilization Questionnaire)。该量表包括29个项目, 共分为四个维度: 对研究证据的价值判断、搜索并采用研究证据的技巧和意识、研究质量以及研究证据在组织内的分享程度。该量表最早强调了研究证据的运用障碍视角。这份针对美国情境的量表在随后13年间被依次开发于适应不同的国家情境, 其有效性被广泛验证。

就研究证据的运用行为量表而言, 较为成熟的是Johnston等(2003)开发的以大学生为研究对象的研究证据量表。该量表具体测量了大学生们通过教材、网络、文献以及特定专业数据库和特定专业期刊获取证据的行为频率。该量表已被运用于不同的研究对象, 其可靠性和有效性均得到了验证。

(二) 基于研究证据和个体经验证据结合的测量

随着证据内涵的发展, 研究者开始逐渐重视个体经验的测量。Aarons(2004)较早关注了循证管理和实践中的个体经验证据及其测量问题。他开发了循证实践的态度量表(The Evidence-Based Practice Attitude Scale, EBPA), 具体包括四个维度: 循证实践的吸引力、外界强制使用循证实践的程度、对创新的接受度以及循证实践与传统实践的差异。在循证实践与传统实践差异维度中, 他强调循证实践是基于研究证据的, 而传统实践则是基于个体经验的, 进而分别对其进行了测量。后来, Upton和Upton(2006)在开发和验证循证实践认知、态度和履行行为自我报告量表时, 又再次强调和完善了研究证据与个体经验证据结合的测量。其中的履行行为维度考察了实践者搜索、记录研究证据的频率以及结合自身经验专长的频率。而Gerrish等(2007)构建了推进循证实践(Developing Evidence-Based Practice, DEBP)的障碍及途径量表, 测量了有关研究证据和个体经验运用的发展条件及瓶颈, 又一次强调了研究证据与个体经验证据的结合。

在这类研究证据和个体经验证据结合的量表中, 研究者们广泛认可的是Estabooks(1998)开发的实践知识来源量表(Sources

of Practice Knowledge Questions), 它测量了护士护理实践的依据。这一量表相比以往量表, 其不同之处在于: 以往对证据的测量都局限于研究证据的运用程度, 该量表首次测量了实践者对个体经验证据的使用程度。后来该测量被Gerrish等(2007)扩充后, 又进行了信度、效度检验, 其有关个体经验证据的测量项目被广泛引用。

这类量表在证据测量中并非占据主流, 因为证据的内涵在随后获得了迅猛地发展, 许多研究者往往基于更宽范畴的多重证据观对其进行测量。

(三) 多重证据测量

多重证据观下的证据一般从循证实践的流程视角进行测量。循证实践流程描述实践者运用不同证据的行为步骤, 这样, 量表开发者可以从流程角度测量实践者获取、利用不同证据的情况。然而, 由于每个量表开发者所依照的循证实践流程不尽相同, 这类量表在项目的具体表述上就会存在诸多差异。

Rubin和Parrish(2010)开发了循证实践流程的评估量表(The EBP Process Assessment Scale), 旨在调查和评估医生对参与循证实践流程的自我效能感、态度、熟悉度、倾向以及参与度。该量表列出的循证实践步骤包括: 形成一个循证实践问题; 搜寻有助于解答该问题的研究证据; 批判性地评价该证据; 将最佳证据与实践专长、实践情境以及顾客特质和偏好相结合, 从而做出最佳决策; 监控决策的实施。他们的检验结果也显示这一量表有良好的内部一致性信度、校标效度以及灵敏度。从该量表所依据的流程可以发现, 其测量了研究证据、经验判断、情境证据以及利益相关者偏好这四类证据。

Nancy和Susan(2010)开发了循证实践信心量表(evidence-based practice confidence, EPIC), 该量表评价了专业卫生保健人员达成循证实践的自我效能感。他们通过逐步地修订完善, 获得了包括11个项目的自评量表, 通过描述循证实践的步骤, 从而让卫生保健人员为自己能够完成这些活动的自信程度打分。量表所考察的循证实践的步骤是由Rappolt(2003)提出的, 具体包括: 识别一个临床问题、形成一个问题来引导文献搜索、有效地搜索相关文献、批判性地评价证据信度和效度、恰当地将证据应用于临床问题、理解病人的需要和偏好、在恰当的行动步骤中与病人合作决策、持续评价实践效果。可见, 该量表的测量涵盖了证据的三大内容: 研究证据、病人偏好以及个体经验, 是多重证据观下的一个量表。

Tilson(2010)开发和评价了一个改编版的Fresno检测(FT)来测量医生掌握的有关循证实践的知识和技能。Fresno检测是一个普遍使用的测量循证实践知识和技能的方法, Tilson(2010)按照Dawes等(2005)描述的五步骤模型对原始的Fresno检测版本加以修订。原始的Fresno检测只评价了循证实践模型的第一至第三个步骤(提出问题, 获取研究证据, 评价研究证据), 没有评价第四个步骤——将病人的观点以及临床经验与可得的研究证据结合起来。这份研究改编了原始FT中有关特定学科的内容, 并添加了两个新问题来更全面地评测循证决策的流程。由此, 这份量表测量的证据对象也在原始的FT量表上扩充至多重证据。

纵览以往文献, 对研究证据、个体经验证据、利益相关者偏好证据以及具体情境证据的测量均存在可靠、有效的量表。虽然这些量表更多针对的是医学和教育学领域的相关研究问题, 但其测量视角、测量维度和具体内容, 都特别值得管理领域相关问题研究的借鉴。这是因为无论医学、教育学领域还是管理领域, 循证决策、循证管理和循证实践都无外乎涉猎研究证据、个体经验证据、利益相关者偏好证据以及情境证据这样几大类。

四、总结及有待进一步研究的问题

“证据”是循证决策、循证管理及实践的核心概念, 只有明晰了其界定、特征和测量, 才可能对该概念的内涵有清晰了解, 也才能在管理领域进行更深入地研究。

本文基于以往相关研究文献, 首先, 明确了证据内涵的不同观点, 即研究证据观、研究证据与个体经验证据结合观以及多重证据观。研究证据观强调在决策及实践中以最佳的科学研究证据为基础; 而研究证据与个体经验证据结合观则强调利用个体经验判断对研究证据进行搜寻、记录、批判性评价以指导决策; 多重证据观在以上观点基础上又将具体情境证据和利益相关者的偏好纳入证据中。其次, 本文明确了研究证据特征的一个重要表征——“强度”的内涵。证据强度主要用来表征研究证据的特征, 具体指研究证据的真实性和普适性。根据证据强度的不同, 研究证据可以分为6个等级, 并且强度越强的研究证据越具有可信度和指导意义。另外, 本研究还明确了不同观点下的证据测量。分析发现以往研究者基于研究证据、个体经验与研究证据结合以及多

表 1 三种证据观下的代表研究者及量表

	代表研究者	代表量表	
		名称	视角
研究证据观	<ul style="list-style-type: none"> • EBM Working Group(1992) • Stewart (1998) 	• 柏林问卷(Berlin questionnaire) (Fritsche 等,2002)	• 证据运用能力视角
		• 研究运用障碍量表(Barriers to Research Utilization Questionnaire)(Funk 等,1991)	• 证据运用障碍视角
个体经验与研究证据结合观	<ul style="list-style-type: none"> • Sackett(1996) 	• 循证实践态度量表(Evidence - Based Practice Attitude Scale) (Aarons,2004)	• 证据运用态度视角
		• 实践知识来源量表(Sources of Practice Knowledge Questions) (Estabrooks,1998)	• 证据运用行为视角
多重证据观	<ul style="list-style-type: none"> • Sackett(1997) • Haynes 等(2002) • Briner 等(2009) 	• 循证实践流程的评估量表(The EBP Process Assessment Scale) (Rubin 和 Parrish,2010)	• 证据运用流程视角
		• 循证实践信心量表(evidence - based practice confidence) (Nancy 和 Susan,2010)	• 证据运用流程视角

资料来源:根据相关文献整理

上述这些分析结论为后续实证研究奠定了坚实基础。但是，概括以往文献，也可以发现以往研究仍存在一些不足之处：(1) 以往关于“证据”内涵及测量问题尽管在不少研究中都有涉及，但不同研究者有不同的观点，那么在以后的管理研究中，应该运用何种证据观？强调研究证据还是个体经验证据？还是多种证据？循证理论和循证管理到底应该坚持何种证据观？(2) 通过上述分析，也可以发现关于“证据”测量的开发在医学研究领域更加成熟，在管理研究领域还仅是一个开端，那么当将医学研究领域的“证据”内涵和测量应用到管理研究领域时，是否应该进行修正和调整？以往研究并未给出答案。(3) 就循证理论在管理领域的研究而言，以往文献还比较少，并且这些研究多聚焦在描述性研究上，针对特定问题的深入理论和实证分析明显欠缺，关于“证据”的研究应用也随之缺乏。这为以后的研究提供了契机。

鉴于此，循证理论在管理研究和实践领域都会有广阔的应用和发展空间，未来研究可以在如下几个方面展开：(1) 将循证理论应用到管理决策的研究中。管理决策包括人事甄选决策、薪酬决策、培训方案的选择、激励策略选择以及人力资源管理策略选择都涉及基于证据的决策问题，那么，管理者决策主要依据何种类型的证据？基于不同类型的证据，其决策机制是否不同？这些问题都值得进一步探讨。(2) 将循证理论应用到创业决策的研究中。创业决策问题一直是创业领域的一个研究热点，那么，创业者在对创业机会的把握时，到底依据研究证据还是个人经验？还是两者的结合？哪些因素使他们在决策时会依据不同类型的依据？(3) 将循证理论应用到中国背景下的管理问题的研究中。如中国的民营企业的管理者在管理过程中更多倚重研究证据还是个人经验？为什么？总之，将循证理论引入管理研究中，会在新的视角下探索出非常有理论和实践意义的相关结论，从而可以丰富和延伸管理问题的研究。

责任编辑：毕雁

文档附件：

隐藏评论

用户昵称: (您填写的昵称将出现在评论列表中) ∈ 匿名

请遵纪守法并注意语言文明。发言最多为2000字符（每个汉字相当于两个字符）

1236

中国社会科学网电话：010-84177688；84177878；84177879 Email: skw01@cass.org.cn
投稿信箱: skw01@cass.org.cn 网友之声信箱: skw02@cass.org.cn 地址: 北京市朝阳区望京中环南路1号
版权所有: 中国社会科学院 [版权声明](#) 京ICP备05072735号