

• 非血管介入 Non-vascular intervention •

食管癌 ^{125}I 粒子支架临床研究中文文献质量评价

朱海东, 滕皋军, 郭金和, 朱光宇

【摘要】目的 评估我国 ^{125}I 内照射支架治疗食管癌文献的质量。**方法** 按照循证医学标准, 分别按不同年份对 2010 年 6 月前的不同设计类型文献的一般情况进行分类。由 2 名评阅者对文献进行质量评价, 对评阅者的评价结果进行 Kappa 检验, 检验结果的一致性。参照 Jadad 评分量表, 计算随机对照试验(RCT)的 Jadad 得分, 按照 CONSORT 声明修订版标准评估其质量。对临床对照试验(CCT)文献中的内容进行描述性分析。**结果** 纳入临床治疗性文献 54 篇, 其中叙述性研究 39 篇、RCT 文献 2 篇、CCT 文献 13 篇。2 名评价者间 Kappa 值为 0.815, 说明评价的一致性较好。2 篇 RCT 文献的 Jadad 评分均为 1 分; 对照 CONSORT 声明修订版标准, “标题和摘要”部分平均得分 0.500 分, “引言”部分平均得分 0.500 分, “方法”部分平均得分 0.325 分, “结果”部分平均得分 0.250 分, “讨论”部分平均得分 0.500 分。13 篇 CCT 中, 纳入与排除标准明确者仅 3 篇(23.08%), 疗效标准明确者 12 篇(92.31%); 仅 2 篇(15.38%)提及“随机”字样, 但没有 1 篇是完全随机, 样本量 ≥ 60 例者 3 篇(23.08%), 有基线可比性者 12 篇(92.31%), 无文献提到盲法, 有统计学方法的 12 篇(92.31%), 随访的 13 篇(100%), 失访的 1 篇(7.69%), 有 11 篇(84.62%)讨论时同时考虑了统计学意义和临床意义。**结论** 我国食管癌 ^{125}I 粒子支架临床研究中文文献总体质量不高, 证据的可信度级别较低, 较循证医学标准质量差距较大。

【关键词】食管癌; 内照射; 碘 125; 支架; 文献计量学

中图分类号: R735.1 文献标志码: A 文章编号: 1008-794X(2011)-06-0455-04

Self-expandable esophageal stent loaded with ^{125}I seeds for the treatment of esophageal cancer: quality assessment of related Chinese literature ZHU Hai-dong, TENG Gao-jun, GUO Jin-he, ZHU Guang-yu.
Department of Radiology, Zhongda Hospital, Southeast University, Nanjing 210009, China

Corresponding author: TENG Gao-jun, E-mail: gjteng@vip.sina.com

【Abstract】Objective To evaluate the quality of medical documents concerning clinical studies about self-expandable esophageal stent loaded with ^{125}I seeds for esophageal cancers in China. **Methods** Medical documents published before June 2010 were collected. According to the standard of evidence-based medicine, the documents were classified. The quality of the academic papers was evaluated by two reviewers separately. The results were assessed with Kappa T-test. Jadad scores of randomized controlled trials (RCT) and CONSORT scores of controlled trials were calculated. **Results** A total of 54 therapeutic articles, including description studies ($n = 39$) and RCT/CCT papers ($n = 15$) were enrolled in this study. The Kappa value was 0.815, indicating a good consistency in evaluation. The Jadad score of 2 RCT papers was 1. According to revised CONSORT statement score scale, the average score for the “title and abstract”, the “introduction”, the “method”, the “result” and the “discussion” was 0.500, 0.500, 0.325, 0.250 and 0.500, respectively. Accurate criteria of case selection and exclusion were designed in 3 CCT papers (23.08%), and therapeutic effect evaluation was mentioned in 12 CCT papers (92.31%). In 3 CCT papers (23.08%) the number of cases was over 60. The baselines were compared in 12 CCT papers (92.31%). No randomization and blind method were employed in all papers (0%). Follow-up check-up were performed in 13 CCT papers (100%) and lost-in-touch of cases was mentioned in one CCT paper (7.69%). Statistical analysis was conducted in 12 CCT papers (92.31%), and both the statistical significance and the clinical significance

基金项目: 国家 863 项目(2009AA02Z402); 国家自然科学基金面上项目(81071238)

作者单位: 210009 南京 东南大学医学院附属中大医院放射科

通信作者: 滕皋军 E-mail: gjteng@vip.sina.com

were discussed in 11 CCT papers (84.62%). **Conclusion** Compared with the standard of evidence-based medicine, the general quality of Chinese medical documents concerning clinical studies about self-expandable esophageal stent loaded with ^{125}I seeds for esophageal cancers is, for the present, not high and poorly credible. (J Intervent Radiol, 2011, 20: 455-458)

【Key words】 esophageal cancer; intraluminal brachytherapy; ^{125}I seed; stent; bibliometrics

食管癌是临床常见的恶性肿瘤,其发病率及死亡率分列恶性肿瘤的第 8 位和第 6 位^[1-3]。 ^{125}I 粒子内照射支架在扩张狭窄食管病变改善患者吞咽困难同时,能一定程度地抑制肿瘤组织生长。郭金和等^[4]进行的一系列动物实验基础研究及小规模的前瞻性临床研究证明 ^{125}I 粒子内照射支架安全、有效。国内亦可见不少相关临床报道。为了全面、系统地了解我国 ^{125}I 粒子内照射支架治疗食管癌临床研究文献现状和存在的问题,我们采用循证医学系统评价的方法对我国相关中文文献进行分析,以便进一步提高临床研究水平和文献质量。

1 材料与方法

1.1 文献来源

根据国际 Cochrane 系统评价手册标准^[5],于 2010 年 7 月 8 日选择万方数据库 (1989 - 2009 年)、中国知网(CNKI,包括中国期刊网,1979 - 2010 年)、中文科技期刊全文数据库 (重庆维普,1989 - 2010 年)、中国科学文献数据库服务系统(CSCD,1989 - 2010 年)、中国生物医学文献数据库(CBM,1978 - 2010 年)和中国医学科普文献数据库(2000 - 2010 年),以“食管癌”、“食道癌”、“ ^{125}I ”、“碘 125”、“内照射”和“支架”为检索词进行文献检索并翻阅。

1.2 纳入与排除标准

1.2.1 纳入标准 ①研究设计:随机对照试验(randomized controlled trials,RCT)和临床对照试验(clinical controlled trials,CCT),为文献中提及“对照”、“分组”的临床研究;无对照研究为无对照的系列病例观察研究和叙述性研究。②纳入患者诊断:食管癌或食管癌伴吞咽困难。③干预措施: ^{125}I 粒子内照射支架置入。

1.2.2 排除标准 ①非治疗性的临床研究,包括护理、器械改进设计、手术并发症研究等。②动物实验和细胞组织研究。

1.3 文献质量评价

按照临床科研设计、衡量、评价和循证医学的原则^[6],由 2 名接受过循证医学培训的介入专业医师阅读全文,并按统一的调查评价表要求独立进行

评阅。独立评价完成后,进行交叉核对,对有分歧者,再次讨论或者提请第三方达成共识。文献的一般情况分别按不同年份对其进行分类。对纳入的所有 RCT 文献,参考 Jadad 评分量表^[9]:如果文章提及随机、盲法以及失访情况,有一项加 1 分,总共为 3 分。在此基础上,如果随机方法描述错误,则减 1 分;若正确,则加 1 分。盲法不正确减 1 分,正确加 1 分,总积分为 0 ~ 5 分;具体按照 CONSORT 声明修订版^[10],若符合该标准,得 1 分,完全未提及该标准,得 0 分;提及但阐述不全面,得 0.5 分。而对于其余所有 CCT 文献,则按 Cochrane 系统评价手册对其中的纳入与排除标准、疗效判定标准、随机法、样本含量、组间均衡性比较、盲法的运用、治疗方法、统计学方法、随访、退出与失访、结论推导等内容进行描述性分析。

1.4 统计学方法

采用 Excel 2003 建立数据库,进行数据的录入和管理。对 2 名评阅者的评价结果采用 SPSS 13.0 软件进行 Kappa 检验行可靠性评价。

2 结果

2.1 文献筛选结果

以“食管癌/食道癌 + ^{125}I /碘 125 + 支架”检索,初步筛选结果为:万方数据库:34 篇,中国知网:23 篇,维普网:47 篇,中国科学文献数据库:3 篇;中国生物医学文献数据库 + 中国医学科普文献数据库:31 篇;共 138 篇,去除重复及不相关文献后余文献 61 篇。以“食管癌/食道癌 + 内照射 + 支架”检索,初步筛选结果为:万方数据库 28 篇,中国知网 28 篇,维普网 12 篇,中国科学文献数据库 1 篇;中国生物医学文献数据库+中国医学科普文献数据库 20 篇,共 89 篇,去除重复及不相关文献后余 29 篇。对上述“61 + 29”篇文献再次整合去除重复及不相关文献后共得文献 68 篇。进一步删除护理相关文献 8 篇、综述类文献 3 篇、支架设计相关文献 1 篇、支架置入并发症文献 1 篇以及支架置入后 CT 表现文献 1 篇,最终纳入临床治疗性文献 54 篇,其中叙述性研究 39 篇,RCT 文献 2 篇、CCT 文献 13 篇。2 名评

价者间 Kappa 值为 0.815,说明评价一致性较好。

我国 ^{125}I 粒子内照射支架治疗食管癌临床研究文献最早见于 2003 年,以后逐渐增加,但非对照性的叙述性研究仍占多数,具体文献的年度分布见表 1。

表 1 ^{125}I 内照射支架治疗食管癌临床研究文献不同年份分布情况

文献 年份	对照试验 随机对照(%)	临床对照(%)	无对照研究 (%)	综述(%)	合计(%)
2003	0(0.00)	0(0.00)	1(1.75)	0(0.00)	1(1.75)
2004	0(0.00)	1(1.75)	2(3.51)	0(0.00)	3(5.26)
2005	0(0.00)	1(1.75)	5(8.77)	0(0.00)	6(10.53)
2006	0(0.00)	3(5.26)	6(10.53)	0(0.00)	9(15.79)
2007	1(1.75)	2(3.51)	12(21.05)	0(0.00)	15(26.32)
2008	0(0.00)	3(5.26)	6(10.53)	1(1.75)	10(17.54)
2009	1(1.75)	3(5.26)	4(7.02)	2(3.51)	10(17.54)
2010	0(0.00)	0(0.00)	3(5.26)	0(0.00)	3(5.26)
合计 (%)	2(3.51)	13(22.81)	39(68.42)	3(5.26)	57(100.00)

2.2 临床试验设计质量评估

2.2.1 RCT 质量评价 在仅有的 2 篇 RCT (文献 1^[11]、2^[12])中,其 Jadad 量表评分均仅得 1 分(进行了随机分组,但未详细阐述,未提及双盲及退出、失访等情况)。对照 CONSORT 声明修订版标准,此 2 篇 RCT 文献中,仅“干预措施”和“定义明确的指标”2 项阐述完全,引言部分中文献 1 未能阐明原理;“方法”部分平均得分 0.325 分,主要在于随机化相关项目得分较低,甚至不得分;“结果”部分仅平均得分 0.250 分,主要因缺乏“受试者流程”以及“分析人数与意向性分析”相关内容,其他几项虽有提及但阐述都不完全;“讨论”部分平均得分 0.500 分,各方面都提及但阐述都不全面。详见表 2

2.2.2 CCT 质量评估 13 篇 CCT 文献中,参考 Cochrane 系统评价手册,描述如下。

2.2.2.1 纳入与排除标准:仅有纳入标准明确的 10 篇(76.92%);只有 3 篇(23.08%)有明确排除标准,且纳入和排除标准均有,其中的 2 篇(15.38%)标准较明确。

2.2.2.2 疗效判定标准情况:其中 11 篇(84.62%)以吞咽困难程度为疗效判定标准,其中 8 篇(61.54%)采用 Stooler 评分标准,2 篇(15.38%)采用 Neuhans 评分标准^[13],1 篇(7.69%)有具体评分但未交待采用何种标准;没有文献涉及到肿瘤大小、淋巴结以及远距离转移等相关方面的标准。另有 1 篇(7.69%)仅考虑患者生存质量,采用的标准为:Spitzer 生存质量总体评分标准^[14];甚至有 1 篇

表 2 按照 CONSORT 声明修订版标准评估两篇 RCT 报告质量

CONSORT 条目	文献1得分	文献2得分	平均分
标题和摘要	0.500	0.500	0.500
引言	0.500	1.000	0.750
方法			0.325
受试者	0.500	0.500	0.500
干预措施	1.000	1.000	1.000
研究目的	0.500	1.000	0.750
测量指标及方法	0.500	0.500	0.500
确定样本量	0.000	0.000	0.000
产生随机分配顺序的方法	0.000	0.000	0.000
分配隐蔽	0.000	0.000	0.000
执行情况	0.000	0.000	0.000
盲法	0.000	0.000	0.000
统计学方法	0.500	0.500	0.500
结果			0.250
受试者流程	0.000	0.000	0.000
受试者招募	0.500	0.500	0.500
基线资料	0.500	0.500	0.500
分析人数和意向性分析	0.000	0.000	0.000
结果和估计	0.500	0.500	0.500
辅助分析	0.000	0.000	0.000
不良事件	0.000	0.500	0.250
讨论			0.500
结果解释	0.500	0.500	0.500
可推广	0.500	0.500	0.500
全部证据	0.500	0.500	0.500

(7.69%)无任何疗效标准。

2.2.2.3 样本量:样本量 < 60 例者 10 篇(76.92%), ≥ 60 例者 3 篇(23.08%)。试验组和对照组样本含量比值 = 1.0 者 1 篇 (7.69%), 比值 > 1.0 者 5 篇(38.46%),比值 < 1.0 者 7 篇(53.85%);所有文献均未说明样本含量估算的依据。

2.2.2.4 随机化的应用:13 篇研究文献中仅有 2 篇(15.38%)提及随机抽取非 ^{125}I 粒子支架多少例为对照组,余未提及“随机对照”、“随机分组”等字样。

2.2.2.5 盲法的应用:无文献采用双盲法,其中 10 篇(76.92%)都是根据患者本人及其家人意愿进入试验组,另 3 篇(23.08%)未提及盲法。

2.2.2.6 组间均衡性比较的应用:所有文献中组间均衡性的比较内容均涉及年龄、性别、肿瘤长度、食管狭窄程度、机体一般情况、支架长度、肿瘤病理类型、肿瘤临床分期等其中的 1 ~ 7 项目;其中 12 篇(92.31%)有基础资料可比性,对组间情况进行了统计学分析比较,另 1 篇(7.69%)文献组间均衡性的比较仅简单描述为:“部位、狭窄程度、病理类型、转移”等方面“相近”。

2.2.2.7 治疗方法:8 篇(61.54%)讲述未使用其他外放疗或化疗;1 篇(7.69%)术前未使用外放疗或化

疗,余 4 篇(30.77%)未提及是否合并其他放疗。

2.2.2.8 随访和失访: 交代随访和失访的分别为 13 篇(100%)和 1 篇(7.69%)。随访时间:6 ~ 39 个月。1 篇(7.69%)给出了失访的人数,但对具体原因未作说明。

2.2.2.9 统计学方法的应用: 交代统计学方法的 12 篇(92.31%),1 篇(7.69%)未使用统计学方法。其中采用 *t* 检验者 5 篇 (38.46%), 卡方检验者 7 篇 (53.85%), 秩和检验 3 篇 (23.08%), 另 1 篇 (7.69%) 未提及具体的统计学方法就直接给出了 *P* 值。

2.2.2.10 结论推导: 有 12 篇(92.31%)文献主要结论的推导以统计推断的结果为依据,1 篇 (7.69%)未使用统计学方法,仅为叙述后推导而出。12 篇中有 1 篇(7.69%)文献未进行结果的外推性讨论,11 篇(84.62%)进行了外推性讨论且在推导方面同时考虑了统计学意义和临床意义。

3 讨论

上述结果表明我国 ^{125}I 粒子支架治疗食管癌的临床治疗研究仍以叙述性论文为主,真正的随机对照研究数量较少,仅有的 2 篇 RCT 文献按照 Jadad 量表评分和 CONSORT 评价,结果都属于较低质量文献。目前相关文献存在如下不足:①试验设计较差。大多数是回顾性的统计分析,故试验结果的说服力不够。②随机和盲法的实施存在严重不足。仅有的 2 篇 RCT 文献中虽提到随机分组,但对具体情况未作任何交代,其余 2 篇文献中所谓“随机”只是随意性抽取部分病例作对照组,其没有 1 篇是真正完全随机化的。所有文献均未提到盲法控制和随机分配的隐匿方法,所以无法避免人为因素对结果产生的影响。③纳入及排除标准不统一甚至缺乏。难以保证研究的可重复性和研究结果的可靠性。④疗效标准及疗效判定指标不统一。而相对统一的吞咽困难程度评价参考标准,就有 3 个版本(Stooler 评分标准、Neuhans 评分标准、Cwikiel 分级法^[15])。未见具体针对肿瘤大小改变、淋巴结以及远距离转移等改变情况设定的疗效判定标准。⑤基础资料可比性缺乏统一性的标准。虽然大多数文献都介绍了 2 组在一些方面没有统计学差异,但对具体项目分类、项目数量等都无统一的标准,研究中的偏倚等很难避免。⑥对随访尤其是失访情况不重视。⑦统计学方法不统一。数据和结论的可信度都直接受影响,其结果的可推广性亦受到较大影响。⑧实验的

样本量有限。大多数试验例数在 60 例以下,且都是单中心的研究。不同医疗中心的医疗观点、地区环境、人口自然特征等因素都可能会影响结果,制定统一标准、多中心联合、增加样本量都是不错的选择。故目前急需开展高质量的严格设计和执行的大规模、多中心、前瞻性、设盲的 RCT 研究。

[参考文献]

- [1] Parkin DM, Pisani P, Ferlay J. Global cancer statistics[J/OL]. CA Cancer J Clin, 49: 33 - 64, 1.
- [2] Ferlay J, Shin HR, Bray F, et al. Globocan 2008, cancer incidence and mortality worldwide:IARC cancerbase[DB/OL], [2011-02-20].http://globocan.iarc.fr/.
- [3] He J, Gu D, Wu X, et al. Major causes of death among men and women in china[J]. N Engl J Med, 2005, 353: 1124 - 1134.
- [4] 郭金和, 滕皋军, 何仕诚, 等. 食管内照射支架的研制及临床应用的初步结果[J]. 中华放射学杂志, 2004, 38: 916 - 920.
- [5] 郭金和, 滕皋军, 朱光宇, 等. 食管内照射支架的研制及动物实验研究[J]. 中华放射学杂志, 2006, 40: 550 - 555.
- [6] Guo JH, Teng GJ, Zhu GY, et al. Self-expandable stent loaded with ^{125}I seeds: feasibility and safety in a rabbit model[J]. Eur J Radiol, 2007, 61: 356 - 361.
- [7] Guo JH, Teng GJ, Zhu GY, et al. Self-expandable esophageal stent loaded with ^{125}I seeds: initial experience in patients with advanced esophageal cancer[J]. Radiology, 2008, 247: 574 - 581.
- [8] Julian PT Higgins PT, Green S. Cochrane handbook for systematic reviews of interventions [EB/OL][2011-05-16]. http://www.cochrane-handbook.org.
- [9] Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary [J]. Control Clin Trials, 1996, 17: 1 - 12.
- [10] Altman DG, Schulz KF, Moher D, et al. The revised consort statement for reporting randomized trial: explanation and elaboration[J]. Ann Intern Med, 2001, 134: 663 - 694.
- [11] 宋永彬, 柳立军, 王志康, 等. 自膨式携带放射性粒子带膜食管支架治疗晚期食管癌的研究[J]. 中国医科大学学报, 2007, 36: 82 - 83.
- [12] 李 滨, 庄丽维, 王彦君, 等. ^{125}I 同位素粒子介入治疗老年晚期食管癌的临床研究[J]. 中国老年学杂志, 2009, 29: 1881 - 1882.
- [13] Hoffmann W, Dittler HJ, Neuhaus H, et al. Implantation of self-expanding esophageal metal stents for palliation of malignant dysphagia[J]. Endoscopy, 1992, 24: 405 - 410.
- [14] Spitzer WO, Dobson AJ, Hall J, et al. Measuring the quality of life of cancer patients: a concise QL-index for use by physicians [J]. J Chronic Dis, 1981, 34: 585 - 597.
- [15] Cwikiel W, Stridbeck H, Tranberg KG, et al. Malignant esophageal strictures: treatment with a self-expanding nitinol stent [J]. Radiology, 1993, 187: 661 - 665.

(收稿日期:2011-03-03)