

生存分析在药物经济学评价 Markov 模型转移概率计算中的应用



周挺, 马爱霞

中国药科大学国际医药商学院(南京 211198)

【摘要】 在药物经济学评价的研究中, Markov 模型作为一种决策分析模型被广泛应用。Markov 模型中需考虑疾病风险随时间变化的特性时, 各 Markov 健康状态间转移概率参数的计算既是重点也是难点, 而生存分析可以提供一种可行的解决方案。本研究通过探讨生存分析在计算动态 Markov 模型转移概率时, 基于累积概率密度函数的新方法, 并结合胃癌 Markov 模型进行实例分析, 为开展相关研究的学者提供参考。

【关键词】 生存分析; 药物经济学; Markov 模型; 转移概率; 胃癌

The survival analysis applied in calculation of markov model transition probability in pharmaceutical evaluation

ZHOU Ting, MA Aixia

School of International Pharmaceutical Business, China Pharmaceutical University, Nanjing, 211198, P.R.China

Corresponding author: Ma Aixia, Email: aixiama73@sina.com

【Abstract】 Markov model is one of the decision analysis models, which is widely used in pharmacoeconomic evaluation studies. In terms of dealing with changes of disease risks during different times, the transition probabilities among different Markov health states becomes hard to calculate. Nevertheless, survival analysis is an available resolution. In this paper, we introduced how to apply survival analysis in calculation of transition probability in time-dependent model based on cumulative probability with a case analysis on advanced gastric cancer Markov model, and provide more information for researchers to build models.

【Key words】 Survival analysis; Pharmacoeconomics; Markov model; Transition probability; Gastric cancer

药物经济学作为评估医疗卫生产品、服务及项目的一门科学, 强调了不同方案在预防、诊断、治疗和疾病管理时的临床、经济学和人文产出, 为合理配置稀缺的医疗卫生资源提供了重要的信息与证据^[1,2]。在药物经济学评价研究中, 根据数据来源和方法的不同, 通常分为基于临床试验 (trial-based)^[3-5]和基于模型 (model-based)^[6-9]的两类研究。由于疾病的复杂多变及财力投入的有限, 往往难以在临床研究开展期间观察到所有事件终点, 常以中间指标作为替代效果指标^[10,11], 导致不能反映患者的长期疾病风险和干预方案的长期影响。基于上述原因, 收集二手文献及各种数据库的资料构建药物经济

学模型, 模拟疾病发生、发展及治疗长期结果的方案更为多见^[12,13], 其可行性往往更高。

常见的基于模型的药物经济学评价方法包括决策分析模型和计量经济模型^[14]。在决策分析模型中, 针对短期且疾病复杂程度较低的研究多采用决策树模型, 而对复杂的慢性疾病多采用建立 Markov 模型来模拟。在构建 Markov 模型时, 转移概率的计算是模型分析的重难点之一, 其参数通常无法直接获取, 需要根据临床文献或流行病学资料等进行计算和调整。根据转移概率是否存在时间依赖 (time-to-event) 可分为动态模型和静态模型。静态模型往往是对动态模型的简化, 未考虑时间因素对疾病风险产生的影响, 动态模型则充分考虑时间因素带来的影响, 其结果的可靠性更优。通过生存分析来得到 Markov 模型中随时间变化的转移概率参数是可行方案之一, 本文对相关理论和方法进

DOI: 10.7507/1672-2531.201801088

基金项目: 江苏省研究生科研与实践创新计划项目 (编号: KYCX17_0735)

通信作者: 马爱霞, Email: aixiama73@sina.com

行分析探讨,并通过胃癌模型的实例分析为开展相关研究的学者提供参考。

1 Markov 模型与转移概率

1.1 Markov 模型

Markov 过程是一种随机过程,对该过程的模型化被称为 Markov 模型。在 20 世纪 80 年代开始被应用于医疗卫生领域^[15]。一般而言,针对某一疾病构建的 Markov 模型包括: Markov 健康状态、循环周期、转移概率、健康产出及成本等基本要素。

Markov 模型实际上是根据疾病对患者健康影响的差异,将整个疾病过程划分为多个健康状态(被称作 Markov 健康状态,互不包含交叉,在同一时刻的患者仅能处于其中一种状态),通过设定相同的时间间隔作为循环周期,观察多个循环周期下,不同 Markov 健康状态彼此间按照一定的发生几率进行转换,最终得到消耗的成本和效果/效用产出^[14]。需注意,模型中存在吸收态(absorbing state),即只可转入不可从该状态中转出,如死亡。

1.2 转移概率

Markov 健康状态间的转移概率,反映了患者从一个状态过渡到另一个状态的可能性。如图 1 简化的 Markov 气泡图所示,不同椭圆表示不同的健康状态,箭头表示彼此之间的转换路径,而箭头上的数字表示转移概率的大小。

当假设不同循环周期下各健康状态间的转移概率不随时间变化,此时则认为疾病风险恒定,以固定不变的发生率在模型中进行循环,直到满足循环终止的条件。然而在真实世界中,不同健康状态之间的转归变化受到在该状态下的时间长短影响,静态模型中固定转移概率的假设往往简化了疾病发生发展的过程,尽管减轻了模拟难度,但结果的可信程度却大打折扣^[16]。因此,需要从实际出发,根据疾病风险随时间的变化来确定每个周期下不同的转移概率。在缺乏一手数据时,动态模型的转移概率无法直接获得,需结合现有文献资料及相关数据库资料,采用生物统计分析及计量手段等方法来进行计算和调整,如生存分析。

2 生存分析与转移概率

2.1 生存分析

生存分析是将事件结果和观察到该结局所经历时间结合起来分析的一种统计分析方法。因此,生存分析可以有效处理疾病发生风险随时间变化的问题,是获得动态 Markov 模型转移概率的一

种可行方案,生存分析特别适用于癌症的 Markov 模型。

在使用生存分析计算时间依赖转移概率前,需熟悉一系列概念。包括:累积生存函数、累积死亡密度函数、死亡概率函数、风险函数和累积风险函数。累积生存函数 $S(t) = P(T > t)$ 表示患者生存时间大于 t 的概率,其值介于 0 到 1 之间。而累积死亡密度函数 $F(t) = p(T \leq t)$ 与 $S(t)$ 互补,反映了患者存活时间小于或等于 t 的概率,两者相加的概率和为 1。对累积死亡密度函数求导,可得到死亡概率函数 $f(t)$,反映了患者在 t 时刻发生瞬时死亡的可能性。风险函数 $h(t)$ 则是一个条件概率函数,反映了已存活到 t 时刻的患者在该时刻的瞬时死亡率,可用 t 到 $t + \delta t$ 这一非常小的区间内发生死亡的概率极限来表示。累积风险函数 $H(t)$ 指患者从一开始到 t 时刻的死亡风险累积概率,可通过积分方法,使用风险函数 $h(t)$ 从 0 时刻到 t 时刻的曲线下面积来表示,而 $S(t)$ 与 $H(t)$ 的关系可用 $S(t) = [-H(t)]$ 来表示。

2.2 生存分析计算转移概率

根据患者所处健康状态,生存分析描绘的曲线通常分为两类:总体生存(overall survival, OS)曲线和无进展生存(progression-free survival, PFS)曲线。以 OS 曲线为例,如果以时间间隔 u 为一个周期长度,则从 $t-u$ 到 t 这段时间区间,转移概率可分为两大类:患者仍处于存活状态的概率 P_{sts} ,以及患者发生死亡的概率 P_{std} 。其中 $P_{sts} = S(t) / S(t-u)$,而 $P_{std} = 1 - P_{sts}$ 。PFS 曲线可得到类似的结果。

就生存分析而言,通常有 3 类生存分析方法:非参数法、参数法和半参数法。所谓非参数法,是不限制生存时间的分布形式及参数,但假设 2 组或多组的总体生存率曲线相同。参数法则是假定生存时间服从特定的参数分布,然后根据已知分布的特点对影响生存时间的因素进行分析,常用的分布包括 Weibull 分布、指数分布等。而半参数法也无

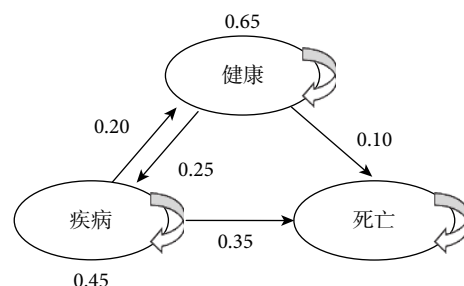


图 1 Markov 模型气泡图

需提前设置分布,主要用于分析影响生存时间和生存率的因素。

为了计算时间依赖转移概率,需对生存时间做出参数分布假设,多采用 Weibull 分布,属于连续性的概率分布。基于 Weibull 分布,其包含两个重要参数:形状参数(γ)和尺度参数(λ),二者均大于 0^[17]。如图 2 所示, γ 参数反映了随着时间增加,风险率(hazard rate)的变化情况,当 $\gamma > 1$ 时风险率随时间增加; $0 < \gamma < 1$ 时则相反,风险率随时间减小;而 $\gamma = 1$ 时等同于指数分布,此时风险率不随时间发生变化^[18]。Weibull 分布下的死亡概率函数 $f(t) = \lambda \gamma t^{(\gamma-1)} \exp(-\lambda t^\gamma)$, 所以其累计概率函数 $H(t) = \lambda t^\gamma$, 可推出 $S(t) = \exp(-\lambda t^\gamma)$, $F(t) = 1 - \exp(-\lambda t^\gamma)$ 。因此,当循环周期为 u 时,从一个状态到另一个状态的转移概率通用公式可表示为 $tp(t_u) = 1 - \exp\{\lambda(t-u)^\gamma - \lambda t^\gamma\}$ 。

3 Markov 动态模型转移概率计算实例

3.1 Markov 模型健康状态及假设

以晚期胃癌为例,按患者的 3 种状态(无进展生存、进展生存和死亡)进行 Markov 模型时间依赖转移概率的计算。笔者曾使用对 OS 和 PFS 曲线结果校正的方法^[19],为避免多重校正的影响,在本研究中将尝试新的基于累积概率计算的方法。

模型基本假设:①以 1 个月作为循环周期间隔;②晚期胃癌患者发生进展后不能退回到无进展状态;③初始队列患者为 1 000 例,循环开始时均处于无进展生存状态。④入组基线年龄为 59 岁,当循环周期达到 60 个周期(5 年)或 99% 的患者进入吸收态(死亡)时循环终止。临床数据源自曲妥珠单抗联合化疗与单独化疗比较治疗晚期胃癌的 III 期临床试验^[20]。

3.2 计算过程及方法

①使用 Get Data Graph Digitizer 软件读取已发表文献报告的曲妥珠单抗联合化疗组与单独化疗组的 PFS 和 OS 曲线^[20],进而获得患者无进展生存

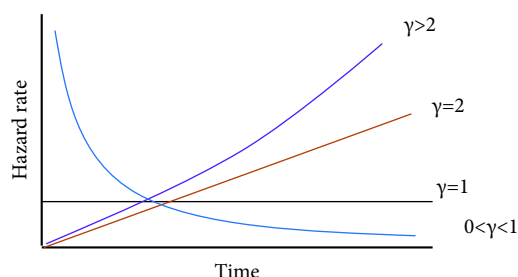


图 2 Weibull 分布形状参数与风险率

时间和总生存时间的样本基础数据(主要涉及不同时间点下处于 PFS 和 OS 的患者比例),为下一步分析做准备。②对生存率和生存时间做数据清理后,使其转变为适合进行生存分析的数据格式,使用 Stata 12 软件进行 Kaplan-Meier 生存分析。③用 Weibull 分布函数模拟新的 PFS 及 OS 曲线,分别得到 PFS 和 OS 的 Weibull 函数的尺度参数 λ_{pfs} 、 λ_{os} 和形状参数 γ_{pfs} 、 γ_{os} 的值,同时比较新拟合曲线与原始曲线的相似度。④将 λ_{pfs} 、 γ_{pfs} 、 λ_{os} 及 γ_{os} 带入转移概率计算公式,得到累积的总生存率 $S(t)$,累积的总死亡率 $F(t)$,累积 PFS 的概率及累积的处于 PD 的概率,进而算得各周期下 3 个状态的时间依赖转移概率。

由于入组人群均为处于 PFS 状态,因此,累积 PFS 的概率与累积 PD 的概率之和等于累积的总生存率 $S(t)$ 。根据累积的总死亡率 $F(t)$ 、累积 PFS 的概率及累积 PD 的概率,可以算得该 1 000 人组成的队列在不同周期下分布的患者数量。基于每个周期处于 PFS 状态、PD 状态和死亡状态的患者数,首先可直接获得每个周期无进展生存到无进展生存的转移概率 P_{pff} 。其次,从无进展生存到进展的概率 P_{ftp} ,无进展生存到死亡的概率 P_{ftd} ,以及 P_{pff} 三者之和为 1。为了得到 P_{ftp} ,通常假设 P_{ftd} 为自然死亡率^[21],进而采用 $(1 - P_{pff} - P_{ftd})$ 算得 P_{ftp} 。最后,为了计算从进展到进展的转移概率 P_{ptp} ,需明确某一周期处于 PD 的患者数(N_{pd2})来源于两部分:上一周期处于无进展生存转移到进展的人数($N_{pfs1} \times P_{ftd}$)和上一周期处于进展的患者仍处于进展的数量($N_{pd1} \times P_{ptp}$)。因此,可得到等式: $N_{pd2} = N_{pfs1} \times P_{ftd} + N_{pd1} \times P_{ptp}$,进而得到每一周期的 P_{ptp} 。

3.3 结果

图 3 和图 4 展示了原始文献和重新拟合的结果。OS 原始曲线中,曲妥珠单抗联合化疗组和单独化疗组的中位生存期分别为 13.8 个月和 11.1 个月,重新拟合的中位生存期分别为 13.4 和 10.8 个月,仅相差 0.4 个月和 0.3 个月;而 PFS 原始曲线中曲妥珠单抗联合化疗组和化疗组的中位生存期分别为 6.7 个月和 5.5 个月,重新拟合后分别为 6.7 和 5.2 个月,仅相差 0 个月和 0.3 个月。结合中位生存期和生存曲线的分布,重新拟合的结果和原始结果吻合度高,再现性较好。

基于 Weibull 分布,可得到曲妥珠单抗联合化疗组的参数: $\lambda_{pfs} = 0.064$, $\gamma_{pfs} = 1.198$, $\lambda_{os} = 0.016$, $\gamma_{os} = 1.421$; 化疗组的参数: $\lambda_{pfs} = 0.091$, $\gamma_{pfs} = 1.215$,

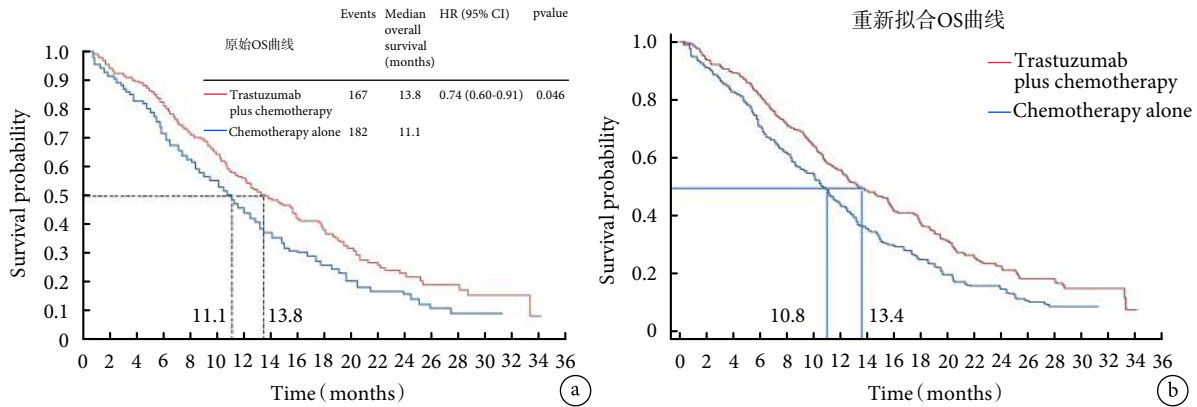


图3 原始 OS 曲线 (左) 与重新拟合后 OS 曲线 (右) 对比

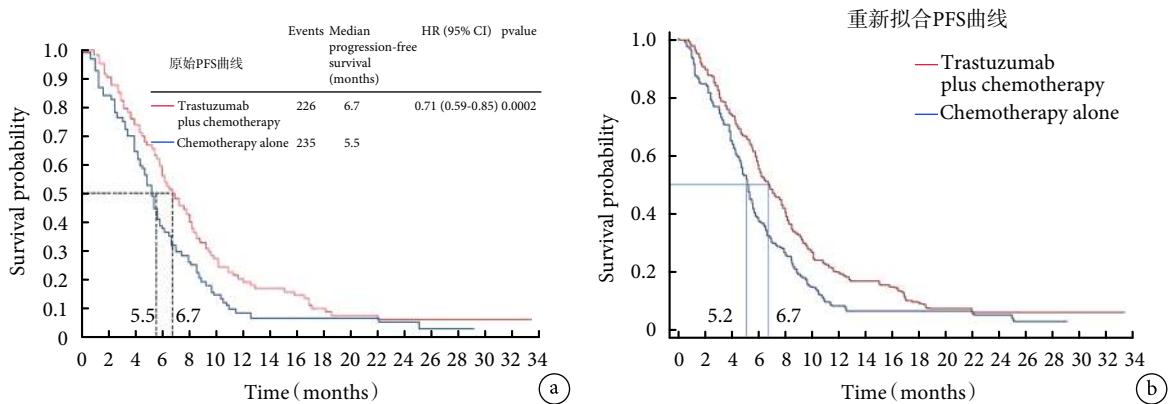


图4 原始 PFS 曲线 (左) 与重新拟合后 PFS 曲线 (右) 对比

$\lambda_{os}=0.035$, $\gamma_{os}=1.255$ 。各周期随时间变化的转移概率和累积概率见表1(因篇幅原因,仅展示第一年前十个周期的结果)。需要注意的是,由于入组人群均处于癌症无进展状态,第一个周期没有一直保留在进展状态的患者,也没有从进展状态转化到死亡的患者,因此第一周期的 Pptd 和 Pptp 可以不考虑。此外, Pftd 为自然死亡率的假设从第二周期开始,第一周期等于累积死亡率。

4 讨论

进行药物经济学评价时,在人力财力投入有限的情况下,采用模型法能够更有效地结合多种资源对疾病的发生、发展进行模拟。Markov 模型作为决策分析模型之一,可得到患者在不同干预手段下的长期成本及健康产出,为医疗卫生决策者提供更充分的决策依据。然而在计算 Markov 模型中各周期下不同健康状态间的转移概率时,如果缺乏大型流行病学调查数据或一手数据,往往难以直接得到。此外,在多数情况下疾病风险会随时间变化,假如不考虑这点,过度简化模型会降低模拟结果可靠性,因此在构建 Markov 模型时,计算具有时间

依赖性的转移概率既是重点也是难点。

生存分析作为一种同时考虑事件结局和发生时间的分析方法,可以捕捉随时间变化的疾病风险变化信息。通过关注患者的生存曲线,对时间变量赋予适宜的分布可得到目标参数,进而计算不同时间区间内疾病转归的发生概率。即使在缺乏临床一手数据的情况下也可通过一系列方法,基于已发表文献的资料重新拟合得到参数。本文案例显示,采用生存分析进行参数计算后重新拟合的 OS 和 PFS 曲线与原始研究的重合度较好,可行性较高。此外,使用累积概率计算转移概率的方法可以避免多重校正带来的影响。

尽管通过生存分析来计算动态模型中随时间变化的转移概率是一种可行的方法,但也存在一些局限。第一,重新拟合获得参数是建立在有公开发表的文献报告了 Markov 模型各干预组生存曲线的基础上,若无法获得或者曲线不完整(如仅汇报了 OS 曲线,无 PFS 曲线)均不能计算完整的累积概率和转移概率。通常仅癌症的研究才会报告生存曲线,而一些进展较慢的疾病,如糖尿病、高血压等则缺乏相关生存曲线报告,无法通过该方法实现时

表 1 动态 Markov 模型各循环周期累积概率及转移概率

组别	循环周期	累积概率 (%)				转移概率 (%)				
		累积 F (t)	累积 S (t)	累积 PD	累积 PFS	Pftp	Pftd	Pftf	Pptd	Pptp
曲妥珠单抗+化疗组										
	1	1.62	98.38	93.76	4.62	4.62	1.62	93.76	-	-
	2	4.27	95.73	86.26	9.47	7.95	0.05	92.00	56.34	43.66
	3	7.48	92.52	78.64	13.89	8.78	0.05	91.17	33.34	66.66
	4	11.04	88.96	71.23	17.73	9.36	0.05	90.58	25.35	74.65
	5	14.84	85.16	64.20	20.96	9.82	0.05	90.13	21.22	78.78
	6	18.80	81.20	57.62	23.59	10.20	0.05	89.75	18.69	81.31
	7	22.83	77.17	51.52	25.65	10.53	0.05	89.42	16.98	83.02
	8	26.90	73.10	45.92	27.18	10.81	0.05	89.13	15.76	84.24
	9	30.96	69.04	40.81	28.23	11.07	0.05	88.87	14.84	85.16
	10	34.97	65.03	36.17	28.85	11.30	0.05	88.64	14.13	85.87
化疗组										
	1	3.48	96.52	91.30	5.22	5.22	3.48	91.30	-	-
	2	8.10	91.90	80.96	10.94	11.27	0.05	88.67	87.67	12.33
	3	13.11	86.89	70.78	16.11	12.53	0.05	87.42	45.42	54.58
	4	18.27	81.73	61.25	20.49	13.41	0.05	86.53	31.74	68.26
	5	23.43	76.57	52.57	24.00	14.11	0.05	85.84	25.02	74.98
	6	28.51	71.49	44.83	26.67	14.68	0.05	85.26	21.05	78.95
	7	33.45	66.55	38.00	28.56	15.18	0.05	84.77	18.44	81.56
	8	38.21	61.79	32.04	29.74	15.62	0.05	84.33	16.62	83.38
	9	42.78	57.22	26.90	30.33	16.01	0.05	83.94	15.28	84.72
	10	47.12	52.88	22.48	30.40	16.36	0.05	83.58	14.27	85.73

PFS: 无进展生存; PD: 疾病进展; Pftp: 无进展生存到进展转移概率; Pftd: 无进展生存到死亡转移概率; Pftf: 无进展生存仍处于无进展生存转移概率; Pptd: 疾病进展到死亡转移概率; Pptp: 疾病进展仍处于进展状态转移概率。

间依赖转移概率的计算。第二, 基于累积概率计算转移概率还需要进行相关的假设, 因为从 OS 曲线和 PFS 曲线得到的发生概率可能融合了多种情况, 要进行剥离和调整, 健康状态越多则剥离难度越大。第三, 重新拟合新的曲线获得相应参数会丢失部分信息, 导致 Markov 模型各循环周期下的生存及死亡人数无法与 OS、PFS 曲线完全一致。但调整转移概率目的就是为了最大限度地接近真实结果, 因此该方法较适合健康状态划明确且状态数量不多的癌症 Markov 模型研究。

综上所述, 采用模型法进行药物经济学评价在简化复杂的疾病发生发展过程的同时, 能够更高效地为医疗卫生决策者面临的资源配置问题提供证据。将生存分析应用于药物经济学评价的 Markov 模型参数计算, 尽管存在一些局限, 对合理解动态 Markov 模型中转移概率时间依赖性的问题, 特别是癌症的药物经济学评价, 仍是有效可行的方法之一。

参考文献

- Berger ML, Bingefors K, Hedblom EC, et al. Health care cost, quality, and outcomes. ISPOR book of terms. ISPOR, 2003.
- 胡善联. 药物经济学. 北京: 高等教育出版社, 2009: 1-200
- Sculpher MJ, Claxton K, Drummond M, et al. Whither trial-based economic evaluation for health care decision making? *Health Econ*, 2010, 15(7): 677-687.
- Petrou S, Trinder J, Brocklehurst P, et al. Economic evaluation of alternative management methods of first-trimester miscarriage based on results from the MIST trial. *BJOG*, 2010, 113(8): 879-889.
- van Dongen JM, van Wier MF, Tompa E, et al. Trial-based economic evaluations in occupational health: principles, methods, and recommendations. *J Occup Environ Med*, 2014, 56(6): 563-572.
- Oostenbrink JB, Rutten-van Mülken MP, Monz BU, et al. Probabilistic Markov model to assess the cost-effectiveness of bronchodilator therapy in COPD patients in different countries. *Value Health*, 2010, 8(1): 32-46.
- Zachariah JP, Samnaliev M. Echo-based screening of rheumatic heart disease in children: a cost-effectiveness Markov model. *J Med Econ*, 2015, 18(6): 410-419.
- Samuelson EM, Odum SM, Fleischli JE. The cost-effectiveness of using platelet-rich plasma during rotator cuff repair: a markov model analysis. *Arthroscopy*, 2016, 32(7): 1237-1244.
- Guerin P, Bourguignon S, Jamet N, et al. MitraClip therapy in mitral regurgitation: a Markov model for the cost-effectiveness of a new therapeutic option. *J Med Econ*, 2016, 19(7): 696-701.
- Birkeland KI, Home PD, Wendisch U, et al. Insulin degludec in type 1 diabetes: a randomized controlled trial of a new-generation ultra-long-acting insulin compared with insulin glargine. *Diabetes Care*, 2011, 34(3): 661-665.
- Heller SR, Colagiuri S, Vaaler S, et al. Hypoglycaemia with insulin aspart: a double-blind, randomised, crossover trial in subjects with Type1 diabetes. *Diabet Med*, 2010, 21(7): 769-775.

- 12 陈永法, 韩洪娜. 达比加群和利伐沙班及华法林预防心房颤动并发脑卒中的经济学评价. *中国新药杂志*, 2016, (11): 1216-1224.
- 13 郭婉婕, 官海静, 刘国恩, 等. 萘环类药物联用右雷佐生治疗乳腺癌的药物经济学评价. *中国卫生经济*, 2017, 36(7): 75-79.
- 14 刘国恩. *中国药物经济学评价指南及导读*: 2015 版. 科学出版社, 2015: 1-50.
- 15 Beck JR, Pauker SG. The Markov process in medical prognosis. *Med Decis Making*, 1983, 3(4): 419-458.
- 16 Briggs AH, Claxton K, Sculpher MJ. *Decision modelling for health economic evaluation*. New York: Oxford University Press, 2006.
- 17 Carroll KJ. On the use and utility of the Weibull model in the analysis of survival data. *Controlled Clin Trials*, 2003, 24(6): 682-701.
- 18 Collett D. *Modelling survival data in medical research*. CRC press, 2015.
- 19 周挺, 马爱霞, 付露阳. 药物经济学评价 Markov 模型中转移概率计算的探讨. *中国卫生经济*, 2017, 36(12): 40-42.
- 20 Bang YJ, Van CE, Feyereislova A, *et al.* Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet*, 2010, 376(9742): 687-697.
- 21 World Health Organization. Country statistics. Available at: <http://www.who.int/countries/chn/en/>.

收稿日期: 2018-01-22 修回日期: 2018-06-06
本文编辑: 张永刚