

# Maximum tolerated dose estimation by Biased coin design and stopping rule in Phase I clinical trial

Soyoung Jeon<sup>a</sup> · Dongjae Kim<sup>a,1</sup>

<sup>a</sup>Department of Biomedicine · Health Science, The Catholic University of Korea

(Received November 25, 2019; Revised December 13, 2019; Accepted December 17, 2019)

---

## Abstract

Phase I clinical trials (Dose Finding Studies) are the first step in administering new drugs developed through animal experiments or in vitro experiments to humans. An important area of interest in designing Phase I clinical trials is determining the dose that provides the greatest efficacy and acceptable safe dose to the patient. In this paper, we propose a method to determine the maximum tolerated dose considering efficacy and safety using Biased coin design and stopping rule. The proposed method is compared with existing methods through simulation.

Keywords: Phase I clinical trials, maximum tolerated dose, Biased coin design, stopping rule

---

## 1. 서론

제1상 임상시험은 ‘투약용량 발견 시험(dose finding study)’라고도 불리는 단계로서, 동물 실험 혹은 시험관 실험을 통하여 개발된 신약물질을 사람에게 실시하는 첫 단계이다. 제1상 임상시험을 디자인하는 데에 있어서 주요 관심 분야는 환자에게 허용할 수 있는 최대의 효능을 가진 용량, 즉 최대허용용량(maximum tolerated dose; MTD)을 결정하는 것이다. 여기서 최대허용용량은 독성이 나타나지 않는 용량 범위 내에서의 최대 복용량을 뜻하거나, 독성 반응이 나타나는 용량 범위 중에서의 최소복용량을 뜻한다 (Lim, 2018). 최대허용용량의 추정은 목표로 정한 독성반응을 초과하지 않는 복용량의 추정이고 (Chevret, 1993), 사람에게 허용할 수 있는 독성 수준에서 최대 효능을 가진 용량 범위의 추정이다 (O’Quigley와 Chevret, 1991). 여기서 목표로 하는 독성반응율은 일반적으로 25–33%를 사용하고 있다 (Storer, 1989).

MTD를 추정하는 기존의 방법에는 Dixon과 Mood (1948)가 제안한 DM 방법이 있다. DM 방법은 한 명의 피험자에게 약을 투여했을 때 독성 반응이 나타나면 용량을 감소시키고 그렇지 않으면 용량을 증가시키는 방법이다. 이 방법은 피험자 한 명의 독성 반응만으로 실험을 진행하므로 계획이 단순하다는 장점이 있지만 MTD를 정확하게 제시하지 못하거나 치료 효과가 없는 낮은 용량 수준에 피험자가 할당될 수 있다는 단점이 있다.

Storer (1989)와 Korn 등 (1994)은 DM 방법을 보완하여 세 명의 피험자를 같은 용량에 배정하는 표준 방법(standard method; SM)을 제안하였다. 표준방법은 멈춤 규칙의 조건을 적용하여, 조건을 만족하

---

<sup>1</sup>Corresponding author: Department of Biomedicine · Health Science, The Catholic University of Korea, 222 Banpo-daero, Seocho-gu, Seoul 06591, Korea. E-mail: [djkim@catholic.ac.kr](mailto:djkim@catholic.ac.kr)

면 실험을 종료하고 한 단계 낮은 수준의 용량을 MTD로 결정하게 된다. 표준방법 중에 SM3 방법은 미국에서 주로 사용하는 방법이다 (Goodman 등, 1995). SM3 방법은 단계적으로 용량을 올려가면서 실험을 하기 때문에 MTD를 조심스럽게 찾을 수 있고 디자인의 사용이 편리하다는 장점이 있다. 반면에 반응이 나타날 때까지 실험을 해야 하므로 실험 기간이 길어질 수 있고, 최대허용용량에서의 독성 발생 확률이 33%에 근접할 것이라는 것과 실험으로 구한 최대허용용량의 독성비율이 실험 전에 정한 목표 독성확률에 가까워짐을 확신할 수 없다는 것이 단점이다.

O'Quigley 등 (1990)은 SM 방법의 단점을 보완하기 위해 continual reassessment method (CRM)를 제안했다. CRM은 시험 전에 독성 반응이 나타날 것으로 생각되는 목표독성확률을 정하고 이전 피험자에게서 얻어진 독성 결과를 이용하여 용량독성곡선(dose-toxicity curve)을 수정한 후에 다음 피험자에게 투여할 용량을 결정하는 축차적인 방법이다. CRM 방법은 목표용량(target dose)을 찾는데 효과적이고 비편향 추정방법(unbiased estimation method)이기 때문에 비교적 객관적이라는 장점이 있다. 하지만 피험자들이 높은 용량수준에 배정될 수 있다는 단점이 있다. Kang (2002)은 이런 단점을 보완하여 수정된 CRM (modified-CRM; MCRM)을 제안했다. 수정된 CRM은 용량 상승 시에 한 단계만 올라가는 보수적인 방법이고 세 명의 피험자를 하나의 군으로 묶어서 실험기간을 단축시킬 수 있다는 장점이 있다. CRM 방법은 통계학적 모형을 이용하기 때문에 계산이 복잡하다는 문제점도 있는데, 이런 문제점을 해결하기 위해 O'Quigley와 Shen (1996)은 우도함수를 이용한 연속재평가 우도방법(CRM likelihood approach; CRML)을 제안했다. CRML은 많은 사전정보를 가지지 않은 경우에 유용하고, 대부분 마지막 추천용량이 CRM의 결과와 비슷하다.

Storer (2001)가 제안한 Best-of-Five 설계는 각 용량 수준에 5명까지의 피험자를 할당할 수 있는 방법으로 '3 + 1 + 1' 설계라고도 불린다. 이 방법은 표준 SM3 보다 용량 증가속도가 빠른 급진적인 방법으로 목표한 MTD에 빠르게 도달할 수 있지만 독성반응율이 높다는 단점이 있다 (Lim, 2018).

무작위 배정 방법 중의 하나인 Biased coin 방법은 두 군간의 대상자수의 차이가 특정값 이상이면 배정비를 변경하여 균형을 맞추는데 목적을 두고 있으며, Efron's Biased Coin Design을 제안한 Efron (1971)은 배정비를 2/3로 하였다. 임상시험에서 Biased coin 방법으로 MTD를 구하는 기존의 방법은 1명의 피험자가 독성반응이 나타났다면 다음 피험자에게 한 단계 낮은 용량을 적용하고, 독성반응이 나타나지 않으면 다음 피험자에게 한 단계 높은 용량을 투여할 것인지 또는 동일한 용량을 투여할 것인지 Biased coin을 이용해 결정하는 방법이다.

본 논문에서는 목표독성확률을 33%로 정하여 1명의 피험자로 시작해 독성반응이 나타나지 않을 경우 다음 피험자에게 사용할 용량을 Biased coin 방법으로 결정하고, 독성반응이 나타날 경우 적절한 멈춤 규칙을 이용해 MTD를 정하는 방법을 제안하였다. 여기서 멈춤규칙은 Best-of-Five 설계의 단점을 보완하기 위해 한 명의 피험자를 추가한 '3 + 2 + 1' 방법이다. Biased coin 방법은 Efron이 제안한 배정비를 사용하여 2/3의 확률로 독성이 나오지 않으면 다음 피험자에게 용량을 증가시키고, 1/3의 확률로 독성이 나오면 동일한 용량을 투여하는 형태로 사용하였다. 제안하는 방법은 독성 반응이 없을 경우 무조건 다음 단계의 용량으로 높이지 않고 같은 용량으로 다음 피험자에게 다시 적용함으로써 기존의 MTD 추정 방법보다 안전성이 높을 것으로 예상된다. 또한 독성 반응이 있을 때는 용량 증가속도가 빠른 '3 + 2 + 1' 방법을 사용하기 때문에 실험에 필요한 피험자수가 기존의 방법보다 적을 것으로 생각된다. 본 논문에서는 기존의 방법 중 SM3, CRM 방법과 제안하는 방법을 모의실험으로 비교하였다.

## 2. 제안하는 방법

본 논문에서는 Best-of-Five 설계의 단점을 보완한 '3 + 2 + 1' 설계와 Biased coin 방법을 이용한 MTD

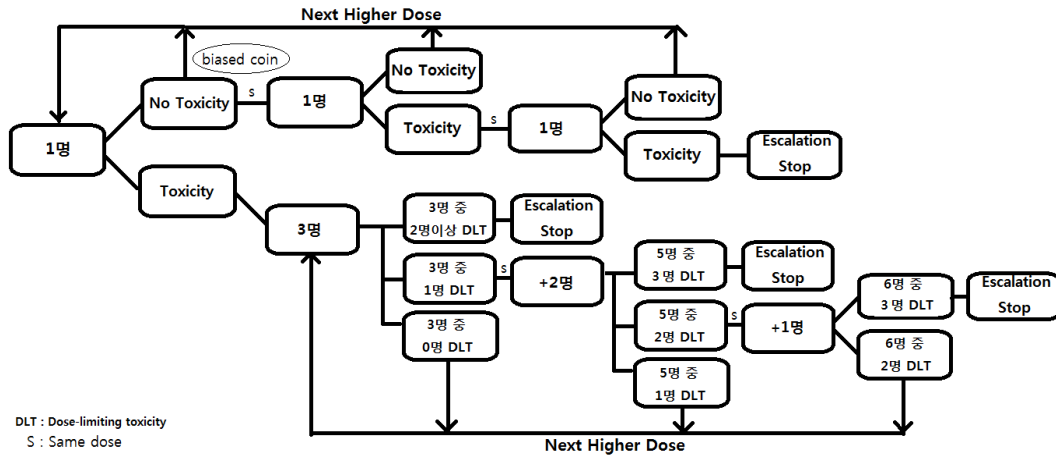


Figure 2.1. Process of BSM.

추정법을 제안하였다. 첫 피험자가 독성반응이 없으면 다음 피험자에게 동일한 용량을 투여할 것인지 아니면 한 단계 높은 용량을 투여할 것인지 Biased coin을 이용해 정한다. 만약 다음 피험자에게 동일한 용량을 투여했는데 독성반응이 없으면 같은 용량 수준에서 두 명의 피험자가 독성반응이 없는 것이므로 다음 피험자에게는 한 단계 높은 용량을 투여한다. 반대로 독성 반응이 있을 때는 같은 용량 수준에 다른 한명의 피험자를 적용한 후 독성반응이 있으면 3명의 피험자 중 2명이 독성반응이 있는 것이므로 실험을 종료하고 한 단계 낮은 용량으로 MTD를 정하고 독성반응이 없으면 3명의 피험자 중 1명만 독성반응이 있는 것이므로 다음 피험자에게 한 단계 높은 용량 수준을 적용한다.

만약 첫 피험자가 독성반응이 있을 경우에는 3명의 피험자에게 동일한 용량을 투여하여 2명 이상 독성반응이 있으면 실험을 종료하고 한 단계 낮은 용량으로 MTD를 정하고, 모두 독성반응이 나타나지 않으면 다음 피험자군에 한 단계 높은 용량을 투여한다. 1명만 독성반응이 있을 경우 다른 2명의 피험자에게 동일한 용량을 투여하여 모두 독성반응이 있으면 총 5명 중 3명이 독성반응이 있는 것이므로 실험을 종료하고 한 단계 낮은 용량으로 MTD를 정하고 모두 독성반응이 없으면 다음 피험자군에 한 단계 높은 용량을 투여한다. 2명 중 1명만 독성 반응이 있을 경우 총 5명 중 2명이 독성반응이 있는 것이므로 다른 1명의 피험자에게 동일한 용량을 투여한다. 여기서 독성반응이 있으면 총 6명 중 3명이 독성반응이 있으므로 실험을 종료하고 한 단계 낮은 용량으로 MTD를 정하고 독성반응이 없으면 다음 피험자군에 한 단계 높은 용량을 투여한다. Figure 2.1은 이 방법을 그림으로 표현한 것이다.

### 3. 모의실험 계획 및 결과

본 논문에서 제안하는 ‘3 + 2 + 1’ 설계와 Biased coin 방법을 이용한 MTD 추정법(BSM 방법)과 SM3, CRM 을 비교하기 위해 모의실험을 실시하였다. 제 1상 임상시험의 모의실험에서 복용량 수준은 6단계가 많이 사용되지만 본 논문에서는 더 많은 비교를 위해 8단계를 사용하였다. 모의실험에서 사용한 dose toxicity curves는 Park (1999)과 Ahn (1998)의 논문에서 사용한 곡선을 8단계 복용량 수준으로 수정하여 Lee와 Kim (2012)이 사용한 곡선을 적용하였다. 모의실험에서 사용한 각 복용량 수준의 독성반응확률을 Table 3.1과 Figure 3.1로 표현하였다.

곡선을 살펴보면, Curve 1은 독성 반응확률이 균일하게 증가하고 Curve 2, 3, 4, 5, 6은 확률이 낮았다가 특정 용량 수준에서 급격하게 증가한다. Curve 7, 8은 첫 용량 수준부터 독성 반응확률이 비교적 높

Table 3.1. Dose toxicity curves

Dose level	Curve 1	Curve 2	Curve 3	Curve 4	Curve 5	Curve 6	Curve 7	Curve 8
1	5	2	1	1	1	5	22	15
2	10	4	1	3	2	25	32	25
3	25	33	5	5	3	50	41	35
4	35	67	10	32	4	60	48	45
5	50	80	25	55	15	70	54	55
6	70	85	80	75	25	80	69	65
7	80	90	90	82	50	90	80	75
8	90	93	95	95	65	95	89	85

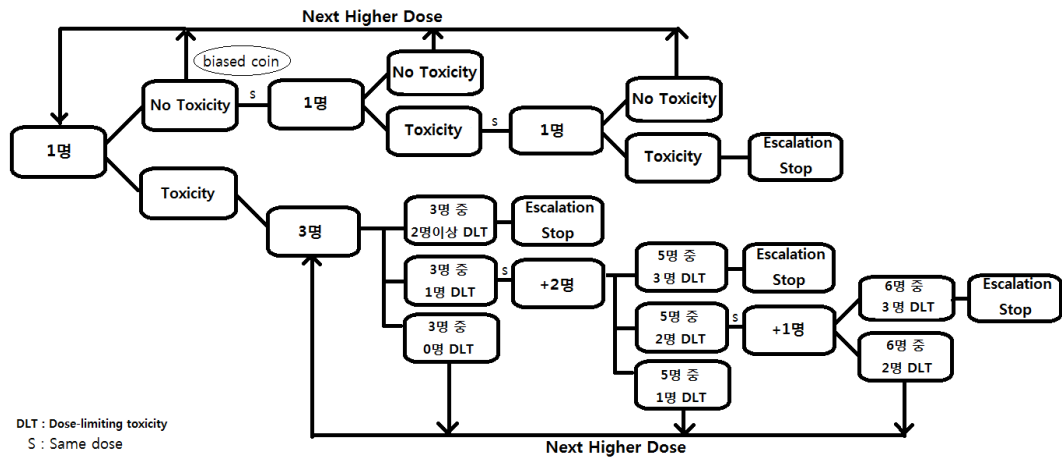


Figure 3.1. Dose toxicity curves.

은 경우이다.

모의실험에서 독성 분포는 SAS의 RANBIN함수를 이용하여 각 복용량 수준에서의 독성 반응확률을 이용한 이항분포난수를 생성했다. 각 복용량 수준에서 멈춤규칙에 의하여 1명, 2명, 또는 3명이 배정되는 BSM과 SM3의 실험을 위해  $Bin(1, p_{ij})$ ,  $Bin(2, p_{ij})$ , 또는  $Bin(3, p_{ij})$ 를 사용하여 난수를 생성하였다. 여기서  $p_{ij}$ 는 Table 3.1에서 제시하는 곡선  $j$  ( $j = 1, 2, \dots, 8$ )의  $i$  ( $i = 1, 2, \dots, 8$ )번째 복용량의 독성 확률이다. BSM에서 Biased coin이 적용된 부분은 SAS에서 2/3의 확률로 1이 나오면 다음 피험자에게 한 단계 높은 용량을 주고, 0이 나오면 다음 피험자에게 동일한 용량을 한번 더 적용하도록 하였다. CRM은 R-package “dfcrm”을 사용하여 모의실험을 실시했다. 이때 독성모델은 제 1상 임상시험에서 가장 많이 사용되는 logistic 함수를 이용했고 초기에 추측되는 독성률(prior estimate)은 Curve 1로 정하였다. CRM은 실험하기 전에 피험자수를 정해야 하는데 1상 임상시험에서는 보통 20명 이하의 작은 수의 피험자를 대상으로 실험이 진행된다. 따라서 본 논문에서는 BSM과 SM3에 이용한 피험자수의 평균값인 13명으로 실험했다. 모의실험의 목표독성확률은 33%이고 10,000번 반복 시행하였다.

용량 수준별 MTD 발생 비율과 피험자수는 Table 3.2에 제시하였다. 각 곡선마다 용량별로 MTD가 선택된 비율을 살펴보면 다음과 같다. 용량간의 독성 수준의 차이가 균일하게 분포하는 Curve 1의 경우 BSM과 CRM은 목표독성확률과 근접하지만 더 높은 확률인 35%에 해당하는 Dose 4에서 MTD 발생 비율이 가장 높았다. SM3는 독성확률이 10%인 Dose 2에서 발생 비율이 가장 높았다. 목표독성확률과 근접하지만 더 낮은 확률인 25%에 해당하는 Dose 3에서 MTD 발생비율이 가장 높은 방법이 없으며

**Table 3.2.** The ratio of MTD at each dose level

		Target toxicity probability = 0.33									Average toxicity	Average subjects
Dose		1	2	3	4	5	6	7	8	None		
Curve 1		5	10	25	35	50	70	80	90			
BSM	% MTD	1.0	10.9	24.5	33.4	24.7	4.8	0.7	0.0	17	4.08	12.62
	%subjects	11.5	13.6	20.1	23.4	19.9	9.7	1.7	0.2			
SM3	% MTD	9.4	36.1	32.8	18.1	3.5	0.2	0.0	0.0	267	2.82	14.34
	%subjects	23.4	26.0	26.9	16.5	6.3	0.9	0.0	0.0			
CRM	% MTD	0.6	10.2	31.3	32.8	22.1	2.9	0.1	0.0		3.74	13.00
	%subjects	12.1	18.2	26.4	21.3	16.8	4.5	0.7	0.1			
Curve 2		2	4	33	67	80	85	90	93			
BSM	% MTD	0.1	18.6	60.4	17.9	2.5	0.4	0.1	0.0	0	3.77	9.68
	%subjects	14.2	15.1	28.2	32.9	8.2	1.1	0.2	0.0			
SM3	% MTD	1.8	55.5	40.8	1.8	0.0	0.0	0.0	0.0	41	2.69	12.38
	%subjects	25.6	26.9	34.4	12.6	0.5	0.0	0.0	0.0			
CRM	% MTD	0.1	17.6	65.4	15.9	0.9	0.1	0.0	0.0		4.23	13.00
	%subjects	9.3	22.4	44.7	18.2	4.7	0.5	0.1	0.0			
Curve 3		1	1	5	10	25	80	90	95			
BSM	% MTD	0.0	0.1	1.1	11.1	77.2	9.7	0.7	0.0	2	3.85	12.29
	%subjects	11.0	11.1	12.0	14.1	20.5	27.8	3.2	0.2			
SM3	% MTD	0.2	2.7	9.1	35.2	52.3	0.4	0.0	0.0	13	2.93	18.69
	%subjects	16.5	16.5	18.2	19.4	20.1	9.3	0.1	0.0			
CRM	% MTD	0.0	0.0	0.9	12.2	73.2	13.7	0.0	0.0		3.41	13.00
	%subjects	8.1	8.5	10.4	15.7	40.3	15.8	1.1	0.1			
Curve 4		1	3	5	32	55	75	82	95			
BSM	% MTD	0.1	0.1	17.9	49.2	27.3	4.8	0.6	0.0	2	3.83	11.70
	%subjects	11.5	12.1	13.0	23.4	27.0	10.9	1.8	0.2			
SM3	% MTD	1.0	2.7	52.6	38.6	5.0	0.1	0.0	0.0	13	2.73	15.81
	%subjects	19.5	20.6	21.3	26.4	11.1	1.1	0.0	0.0			
CRM	% MTD	0.0	0.2	17.6	49.9	29.9	2.4	0.1	0.0		3.86	13.00
	%subjects	8.4	9.1	20.0	31.8	24.3	5.5	0.7	0.1			
Curve 5		1	2	3	4	15	25	50	65			
BSM	% MTD	0.1	0.0	0.2	3.3	13.4	50.0	32.9	0.0	1004	3.99	15.54
	%subjects	8.7	9.0	9.4	9.9	13.2	17.7	22.2	9.9			
SM3	% MTD	0.5	1.0	1.7	18.1	31.6	39.3	7.8	0.0	60	2.92	22.22
	%subjects	13.9	14.3	14.6	14.8	17.3	15.1	8.8	1.3			
CRM	% MTD	0.0	0.0	0.2	2.4	21.0	44.4	28.4	3.6		2.67	13.00
	%subjects	8.2	8.5	9.0	10.5	19.7	25.1	14.6	4.5			
Curve 6		5	25	50	60	70	80	90	95			
BSM	% MTD	9.3	45.7	30.1	11.9	2.6	0.4	0.1	0.0	14	3.86	9.67
	%subjects	15.0	24.5	34.9	18.2	6.1	1.2	0.2	0.0			
SM3	% MTD	40.0	49.8	9.3	0.9	0.0	0.0	0.0	0.0	267	2.69	10.53
	%subjects	31.9	40.5	23.5	3.8	0.3	0.0	0.0	0.0			
CRM	% MTD	6.8	52.3	32.3	7.0	1.5	0.1	0.0	0.0		4.29	13.00
	%subjects	21.3	37.5	27.7	8.9	3.8	0.7	0.1	0.0			

(Continued)

		Target toxicity probability = 0.33									Average toxicity	Average subjects
Dose		1	2	3	4	5	6	7	8	None		
Curve 7		22	32	41	48	54	69	80	89			
BSM	% MTD	21.0	30.4	25.5	14.4	6.9	1.5	0.3	0.0	589	4.10	11.02
	%subjects	16.9	27.1	26.1	17.5	8.4	3.2	0.6	0.1			
SM3	% MTD	54.5	32.3	10.6	2.3	0.4	0.0	0.0	0.0	3347	2.80	10.81
	%subjects	35.7	40.1	18.0	5.1	1.0	0.1	0.0	0.0			
CRM	% MTD	29.5	32.9	20.8	10.2	5.6	0.9	0.1	0.0		4.31	13.00
	%subjects	39.5	26.4	17.9	8.7	5.7	1.6	0.3	0.0			
Curve 8		15	25	35	45	55	65	75	85			
BSM	% MTD	10.9	25.1	30.6	21.7	8.7	2.6	0.5	0.0	230	4.12	11.76
	%subjects	14.5	21.8	25.6	21.3	11.6	4.0	1.0	0.2			
SM3	% MTD	39.9	36.3	18.5	4.5	0.7	0.0	0.0	0.0	1863	2.81	11.85
	%subjects	31.5	36.0	22.0	8.4	1.8	0.2	0.0	0.0			
CRM	% MTD	13.1	30.2	29.6	17.0	8.1	1.7	0.3	0.0		4.01	13.00
	%subjects	27.0	26.3	23.1	12.6	8.0	2.3	0.5	0.1			

(Continued)

로 세 방법 모두 MTD를 잘 추정하지 못했다. Dose 2부터 Dose 4까지 급격하게 독성확률이 증가하는 Curve 2를 보면 BSM과 CRM은 독성확률이 33%인 Dose 3에서 MTD 발생비율이 가장 높았다. 두 방법의 MTD 발생비율은 각각 60.4%, 65.4%로 CRM이 더 높았다. SM3은 독성확률이 4%인 Dose 2에서 MTD 발생비율이 가장 높았다. Dose 6에서 급격하게 독성확률이 증가하는 Curve 3를 보면, 목표독성확률과 근접한 25%에 해당하는 Dose 5에서 세가지 방법 모두 MTD 발생비율이 가장 높았다. 발생비율을 보면 BSM, CRM, SM3가 각각 77.2%, 73.2%, 52.3%로 BSM이 가장 잘 추정하였다. Dose 4, 5에서 급격하게 독성확률이 증가하는 Curve 4의 경우, 32%의 독성확률을 갖는 Dose 4에서 MTD 발생비율이 가장 높은 방법은 BSM과 CRM 이었다. 두 방법의 MTD 발생비율은 각각 49.2%, 49.9%로 거의 동일했다. Dose 5부터 독성확률이 급격하게 증가하지만 전체적으로 확률이 비교적 낮은 Curve 5의 결과를 보면 세가지 방법 모두 25%의 독성확률을 갖는 Dose 6에서 MTD 발생비율이 가장 높았다. 발생비율은 BSM이 50.0%로 다른 방법에 비해 잘 추정하는 것을 알 수 있었다. Dose 2, 3에서 급격하게 독성확률이 증가하는 Curve 6의 경우 세 방법 모두 25%의 확률을 갖는 Dose 2에서의 MTD 발생이 가장 많았다. MTD 발생비율을 보면 BSM, CRM, SM3가 각각 45.7%, 52.3%, 49.8%로 CRM이 가장 높았고 BSM이 가장 낮았다. 첫 용량 수준의 독성확률이 높고 비교적 균일하게 증가하는 Curve 7은 독성확률이 32%인 dose 2에서 BSM과 CRM의 MTD 발생비율이 가장 높았다. 발생비율은 BSM, CRM, SM3가 각각 30.4%, 32.9%, 32.3%로 CRM이 가장 높았다. 첫 용량 수준의 독성확률이 비교적 높고 균일하게 증가하는 Curve 8의 경우 25%의 독성확률을 갖는 Dose 2에서 CRM의 MTD 발생비율이 가장 높았다. 발생비율은 CRM이 30.2%로 다른 방법보다 잘 추정하였다.

BSM과 SM3에서 실험에 필요한 평균 피험자수를 보면, 대부분의 곡선에서 BSM의 피험자수가 더 적었다. 평균 피험자수의 차이가 가장 큰 곡선은 Curve 5였는데 BSM이 15.54명, SM3가 22.22명으로 BSM이 6.68명 적은것으로 나타났다. 그 다음으로 Curve 3에서는 BSM과 SM3가 각각 12.29명, 18.69명으로 BSM이 SM3에 비해 6.4명 적었다. 세 번째는 Curve 4였는데 BSM이 11.7명, SM3가 15.81명으로 BSM이 4.11명 적었다. 다음으로 Curve 2는 BSM과 SM3가 각각 9.68명, 12.38명으로 BSM이 2.7명 적었고 Curve 1은 각각 12.62명, 14.34명으로 1.72명의 차이가 있었다. Curve 6, Curve 7, Curve 8에서는 BSM과 SM3의 평균 피험자수 차이가 1명 미만으로 비슷했다.

BSM과 SM3에서 나타난 결측값의 횟수를 살펴보면, Curve 5를 제외한 모든 곡선에서 BSM이 더 적었다. 결측값은 Dose 1에서 독성반응이 나타나 더 이상 실험을 진행할 수 없는 경우와 Dose 8에 도달하기까지 MTD가 발생하지 않아서 더 이상 용량을 증가시킬 수 없는 경우에 발생한다. 결측값 횟수의 차이가 가장 큰 곡선은 Curve 7로 나타났다. Curve 7에서 BSM과 SM3의 결측값 횟수가 각각 589번, 3,347번으로 BSM이 SM3에 비해 2,758번 적었다. 두 번째로 차이가 큰 곡선은 Curve 8였는데 BSM과 SM3가 각각 230번, 1,863번으로 BSM이 1,633번 적었다. 다음으로 Curve 5는 BSM이 1,004번, SM3가 60번으로 곡선들 중에서 유일하게 SM3가 더 적었다. 그리고 Curve 1에서는 BSM과 SM3가 각각 17번과 267번, Curve 6에서는 각각 14번과 267번으로 두 곡선 모두 BSM이 약 250번 적었다. Curve 2, Curve 3, Curve 4는 결측값 횟수의 차이가 50번 미만으로 두 방법이 비슷하게 나타났다.

용량 수준별로 선택된 피험자 비율을 살펴보면 다음과 같다. Curve 1에서 BSM은 전체 용량 수준중에 Dose 4에서의 피험자 비율이 23.4%로 가장 높았다. SM3와 CRM은 Dose 3에서의 비율이 각각 26.9%, 26.4%로 다른 용량 수준에 비해 높았다. Curve 2는 BSM에서 Dose 4가 32.9%로 가장 높았고 SM3와 CRM은 Dose 3의 비율이 각각 34.4%, 44.7%로 가장 높았다. Curve 3에서 BSM은 Dose 6에서의 비율이 27.8%로 가장 높았고 SM3와 CRM은 Dose 5에서 각각 20.1%, 40.3%로 가장 높았다. Curve 4에서 BSM은 Dose 5의 비율이 27.0%로 가장 컸고 SM3와 CRM은 Dose 4에서 각각 26.4%, 31.8%로 가장 컸다. Curve 5에서 BSM은 Dose 7이 22.2%로 가장 높게 나타났고 SM3은 Dose 5에서 17.3%, CRM은 Dose 6이 25.1%로 가장 높았다. Curve 6에서는 BSM이 Dose 3에서 34.9%로 가장 높았고 SM3와 CRM은 Dose 2에서 각각 40.5%와 37.5%로 가장 높았다. Curve 7에서 BSM과 SM3는 Dose 2에서의 비율이 27.1%와 40.1%로 가장 컸고 CRM은 Dose 1의 비율이 39.5%로 가장 높았다. Curve 8에서 BSM은 Dose 3의 비율이 25.6%로 가장 높았고 SM3는 Dose 2에서 36.0%로 가장 높았다. CRM은 Dose 1에서의 비율이 27.0%로 가장 컸다.

#### 4. 결론 및 고찰

제 1상 임상시험의 주목표 중 하나는 적절한 MTD를 추정하는 것이다. 이때, 사람을 대상으로 실험하기 때문에 높은 안전성과 적은 피험자로 적절한 MTD를 추정하는 것이 매우 중요하다. 본 논문에서는 1명의 피험자를 대상으로 독성반응이 없을 경우 Biased coin 방법을 적용하여 안전성을 높이고, 독성반응이 있을 경우 용량 증가속도가 빠른 '3 + 2 + 1' 방법으로 피험자수를 줄일 수 있는 MTD 추정법을 제안하였다. 제안하는 방법은 모의실험을 통해 SM3, CRM과 8가지 독성 확률 곡선을 이용하여 비교하였다.

모의실험으로 기존의 방법들과 비교했을 때 대부분의 곡선에서 SM3보다 목표독성확률에 가깝게 MTD를 추정했고, 2가지 곡선에서 CRM보다 MTD를 더 잘 추정했다. BSM이 MTD를 가장 잘 추정했던 곡선은 용량이 곡선 후반에 급격하게 증가하는 경우, 곡선 중반부터 용량이 증가하지만 전체적으로 확률이 낮은 경우였다. 곡선 중반에 용량이 급격하게 증가하는 경우에는 BSM과 CRM의 MTD 추정 비율이 거의 동일했는데, 이 경우 CRM은 계산 방법이 복잡하다는 단점이 있기 때문에 BSM이 더 효율적이다. 또한 BSM 방법은 대부분의 곡선에서 기존의 방법보다 평균 피험자수가 적었다. 특히 첫 용량수준의 독성확률이 작고 중간 용량수준에서 급격하게 확률이 증가하는 곡선들에서 피험자수의 차이가 비교적 크게 나타났다. 또한 BSM은 기존의 방법보다 MTD를 추정할 수 없는 경우의 횟수가 더 적었다. 특히 첫 용량 수준의 독성확률이 높고 균일하게 증가하는 곡선들에서 결측값의 횟수에 큰 차이가 있었다.

따라서 본 논문에서 제안하는 MTD 추정법은 대부분의 경우에서 SM3보다 정확성이 더 높았다. 곡선

의 용량이 전체적으로 높을 때보다 첫 용량은 낮지만 중후반에 용량이 증가하는 경우에 CRM보다 정확성이 좋고 효율적이었다. 또한 실험에 필요한 피험자수가 기존의 방법보다 적고 MTD를 추정하지 못하는 경우도 적기 때문에 더 효율적이라는 것을 알 수 있다.

반면에 각 용량 수준별로 선택된 피험자 비율을 보면, BSM 방법이 높은 수준의 용량에 비교적 많은 피험자가 할당된 것을 알 수 있다. 따라서 높은 용량 수준에 더 적은 피험자가 할당될 수 있도록 추가적인 연구가 필요할 것으로 보인다.

## References

- Ahn, C. (1998). An evaluation of Phase I cancer clinical trial designs, *Statistics in Medicine*, **17**, 1537–1549.
- Chevret, S. (1993). The continual reassessment method in cancer phase I clinical trials: a simulation study, *Statistics in Medicine*, **12**, 1093–1108.
- Dixon, W. J. and Mode, A. M. (1948). A method for obtaining and analyzing sensitivity data, *Journal of the American Statistical Association*, **43**, 109–128.
- Efron, B. (1971). Forcing a sequential experiment to be balanced, *Biometrika*, **58**, 403–417.
- Goodman, S. N., Zahurak, M. L., and Piantadosi, S. (1995). Some practical improvements in the continual reassessment method for phase I studies, *Statistics in Medicine*, **14**, 1149–1161.
- Kang, S. (2002). Investigation on the modified continual reassessment method in phase I clinical trial, *The Korean Journal of Applied Statistics*, **15**, 323–336.
- Korn, E. L., Midthune, D., Chen, T. T., Rubinstein, L. V., Christian, M. C., and Simon, R. M. (1994). A comparison of two phase I trial designs, *Statistics in Medicine*, **13**, 1799–1806.
- Lee, N. and Kim, D. (2012). Two-stage maximum tolerated dose estimation by stopping rule in Phase I clinical trial, *The Korean Statistical Society*, **19**, 57–64.
- Lim, H. (2018). *New Drug Development and Clinical Trials*, *Bullsbook*.
- O’Quigley, J. and Chevret, S. (1991). Method for dose finding studies in cancer clinical trials: a review and results of a Monte Carlo study, *Statistics in Medicine*, **10**, 1647–1664.
- O’Quigley, J., Pepe, M. and Fisher, M. (1990). Continual reassessment method: a practical design for phase I clinical trials in cancer, *Biometrics*, **46**, 33–48.
- O’Quigley, J. Shen, L.Z. (1996). Continual reassessment method: a likelihood approach, *Biometrics*, **52**, 673–684.
- Park, I. (1999). *The estimation of maximal tolerated dose in sequential phase I clinical trials* (Master’s Thesis of Science), The Catholic University of Korea.
- Storer, B. E. (1989). Design and analysis of phase I clinical trials, *Biometrics*, **45**, 925–937.
- Storer, B. E. (2001). An evaluation of phase I clinical trial designs in the continuous dose response setting, *Statistics in Medicine*, **20**, 2399–2408.



# 제 1상 임상시험에서 Biased Coin Design과 멈춤규칙을 이용한 MTD 추정법

전소영<sup>a</sup> · 김동재<sup>a,1</sup>

<sup>a</sup>가톨릭대학교 의생명 · 건강과학과

(2019년 11월 25일 접수, 2019년 12월 13일 수정, 2019년 12월 17일 채택)

---

## 요약

‘투약용량 발견 시험(Dose Finding Study)’라고도 불리는 제 1상 임상시험은 동물 실험 혹은 시험관 실험을 통하여 개발된 신약물질을 사람에게 실시하는 첫 단계이다. 제1상 임상시험의 가장 주요한 목적은 환자에게 허용할 수 있고 최대의 효능을 가진 복용량을 결정하는 것이다. 본 논문에서는 이를 고려하여 최대허용용량(MTD)를 결정할 수 있는 적절한 추정방법을 제안하였다. 이 방법은 Biased coin design과 멈춤규칙을 이용하여 MTD를 추정한다. 제안하는 방법은 모의실험을 통해 기존의 방법들과 비교하였다.

주요어: 제 1상 임상시험, 최대허용용량, Biased coin design, 멈춤규칙

---

<sup>1</sup>교신저자: (06591) 서울특별시 서초구 반포대로 222, 가톨릭대학교 의생명 · 건강과학과.  
E-mail: djkim@catholic.ac.kr