

# 백혈병환자 대상의 제 1상 임상시험 연속재평가 방법

이주형<sup>a</sup>, 송혜향<sup>1,a</sup>

<sup>a</sup>가톨릭대학교 의생명과학교실 의학통계학과

## 요약

제 1상 임상시험 계획에서 신약제의 최대허용용량을 추정하기 위해 전통적 방법인 표준 3+3계획과 모형 중심의 베이지안 방법을 적용시킨 연속재평가방법이 주로 사용되고 있다. 본 논문에서는 치료불응 또는 재발된 백혈병이나 골수형성 이상증후군 환자를 대상으로 시행한 제 1상 임상시험의 연구계획을 개관하였다. 단순한 표준 3+3계획으로 환자확보 기간이 길어져 연구를 완성하지 못하고 때 이르게 중단할 수밖에 없는 실정에 반하여, 최근에 제시된 표준 3+1+1계획과 Rolling-6 계획은 연구기간을 크게 단축시킬 수 있다. 제 1상 임상시험의 용량수준 선택과정에서 너무 보수적인 연구자의 태도는 최대허용용량의 정확한 추정을 불가능하게 한다. 새로운 환자에게 투여할 용량수준을 결정할 시점에서 용량제한 독성반응이 아직 나타나지 않은 환자의 관측시간 및 늦게 나타나는 환자의 독성반응 시간을 모두 감안한 연속재평가방법인 TITE-CRM 계획이 유용하며, 이러한 CRM 계획으로 진행되는 임상시험 시행 중 모의실험으로 각 용량수준에서 용량제한 독성반응율이 과대 또는 과소 추정되는지를 파악할 수 있음이 장점이다. 백혈병환자 대상의 임상연구에서 채택되는 제 1상 임상시험의 여러 연구계획의 장, 단점을 제시한다.

주요용어: 제 1상 임상시험, 최대허용용량, 표준계획, 연속재평가방법, 백혈병.

## 1. 서론

제 1상 임상시험은 신약제(drug or drug combination)의 동물실험 후 인간에게 처음 적용하는 단계로써 신약제의 독성(toxicity), 즉 안전성(safety)을 알아보고자 하며 제 1상 임상시험의 주된 목적은 최대허용용량(maximal tolerated dose; MTD)의 추정이다. 따라서 신약제의 성공적 개발은 제 1상 임상시험 결과에 의해 좌우된다고 할 수 있다. 그러므로 연구 대상인 질병에 가장 적합한 임상시험 계획을 선택하여 소수 환자로 최대허용용량 MTD를 추정함과 더불어 독성반응율을 파악하며, 이러한 추천용량(recommended dose)으로 제 2상 임상시험에서 신약제의 효능을 알아보게 된다. 최대허용용량 MTD의 추정에는 용량제한 독성반응(dose-limiting toxicity; DLT)을 우선 정의해야 하며, 이 DLT는 신약제의 복용으로써 특정한 부작용이 나타나는, 따라서 더 이상의 용량증가를 불가능하게 하는, 용량을 제한토록 하는 독성반응을 뜻한다. 정의된 DLT를 이용하여 구체적으로 미국에서 시행되는 제 1상 임상시험에서는 최대허용용량 MTD는 많아야 33%의 환자가 DLT를 경험하는 수준으로 결정하고 있으나 상황과 질병에 따라 33%보다 낮은 대략 20~33% 사이의 값을 목표로 정할 수 있으며, 이 수준이 독성반응 목표수준(target toxicity level)이다. 이러한 배경에는 용량증가에 따라 치료 효능(efficacy)과 독성반응이 동시에 증가함이 전제된다.

제 1상 임상시험의 목적인 ‘환자가 참을 수 있는 독성반응에서 가장 효능이 높은 용량수준’, 즉 MTD 추정은 일반적으로 질병의 특성에 따라 채택되는 여러 가지 요소, 예를 들어서 출발 용량수준,

<sup>1</sup> 교신저자: (137-701) 서울 서초구 반포동 505, 가톨릭대학교 대학원 의생명과학교실, 의학통계학과, 인간유전체 다형성 연구소, 교수. E-mail: hhsong@catholic.ac.kr

각 용량수준에 할당되는 환자수, 용량증가 방법에 의해 결정된다. 본 논문에서는 저자가 함께 참여하는 백혈병(leukemia) 말초혈액 조혈모세포이식 또는 골수이식(bone marrow transplantation; BMT) 환자를 대상으로 시행하는 제 1상 임상시험의 여러 계획을 알아보며, 더불어 백혈병 골수이식 환자 대상의 임상시험에서의 애로점 및 해결방안을 논의하고자 한다. 단순 악성종양(solid malignancies)의 경우와는 다르게 백혈병의 치료 방법인 골수이식(BMT) 중 공여자의 골수를 이식받는 경우에 이식 후 발병하는 이식대숙주병(graft-versus-host disease; GVHD)의 관리가 어려우며, 따라서 공여자 및 환자의 여러 위험요인이 질병의 치유에 영향을 미친다. 그러므로 제 1상 임상시험은 이식 전의 치료 또는 이식 후의 치료 과정에서 진행될 수 있다. 예를 들어서 백혈병 중에서 주로 성인에게 흔한 급성골수성 백혈병(acute myeloid leukemia; AML)은 사전 질병 경험없이 또는 과거의 골수형성이상증(myelodysplasia)이나 항암치료 후에 발병하며, 예후가 좋지 않고 치료가 매우 어렵기 때문에 신약제의 개발이 급선무이다. 이와 같이 항암치료가 실패하거나 재발 급성골수성 백혈병환자의 치료로써 세디라니브(cediranib) 신약제의 MTD를 결정하고자 다기관 참여의 제 1상 임상시험 연구가 Fiedler 등 (2010)에 의해 최근에 시행되었다.

## 2. 표준계획

가장 기본적인 제 1상 임상시험 계획은 전통적 계획(traditional design) 또는 표준계획(standard design)이며 MTD를 추정하기 위해 각 용량수준에 기본적으로 3명 또는 6명의 환자가 배정되고, 연구 대상의 질병에 따라서 조금씩 변형된 계획을 채택한다 (Goodman 등, 1995). 이 표준계획은 Park과 Song (1999)에 설명하였듯이 Dixon과 Mood (1948), Wetherill (1963), Storer (1989)에 걸쳐 연구되었으며, 용량-독성반응 곡선(dose-toxicity curve) 형태에 대한 사전 정보 없이 출발할 수 있는 단순한 계획으로써 바로 전에 할당된 환자군에서 관측되는 독성반응에 따라 용량증가 또는 감소하며 진행한다. 최근까지도 임상에서 주로 사용되는 계획은 표준 3+3계획으로 Park과 Song (1999)에서 Standard Method 6(SM6)와 동일한 계획이다. 이 밖에도 표준 3+1+1 계획, 가속적정(accelerated titration) 표준 계획, Rolling-6 계획 등이 최근까지 제시되었다.

### 2.1. 여러 표준계획

표준 3+3계획 (또는 표준 3+3 up-and-down계획)은 동물실험 자료에 근거하여 안전하다고 추측되는 낮은 용량수준에서 3명의 환자로 출발한다. 3명 환자 중 아무도 DLT가 아닌 경우에 다음 환자군은 한 단계 높은 용량수준을 복용하며, 3명 중 1명이 DLT인 경우에는 추가로 3명이 동일 용량수준에 할당된다. 총 3~6명 중 2명 이상이 독성반응을 나타내어 33% 이상이 DLT이면 연구를 종료하고 한 단계 낮은 용량수준을 MTD로 결정하며 이 MTD를 제 2상 임상시험의 추천용량으로 제시한다. 이 계획에서 처음 복용의 3명 환자를 2명으로 변경시키고 또한 추가 3명을 4명으로 변경시키면 표준 2+4계획이 된다. 이와 같은 방법으로 표준 3+2계획 또는 표준 2+3계획이 가능하다.

각 용량수준에 5명 환자가 할당될 수 있는 표준 3+1+1계획은 주로 많이 사용되는 표준 3+3계획보다도 용량증가 속도가 빨라지는 공격적인 계획으로써, 출발 용량수준에서의 3명의 환자 중 1명 또는 2명 환자가 DLT인 경우에 동일 용량수준에 추가로 1명 환자를 더 할당하며 이러한 4명의 환자 중 2명 환자가 DLT인 경우에 다시 추가로 1명 환자를 더 할당한다 (Storer, 2001). 처음 3명 환자 중 아무도 DLT가 아닌 경우에, 4명 환자 중 1명, 또는 5명 환자 중 2명이 DLT인 경우에는 용량을 증가하게 되고 반면에 3명 이상이 DLT인 경우에는 연구를 종료한다.

일반적으로 제 1상 임상시험의 출발 용량수준이 매우 낮기 때문에 일부 환자는 치료효과가 없는 저용량수준에 할당되는 문제점을 Mick과 Ratain (1993)이 지적하였고, 따라서 임상시험의 초기에 빠른

용량증가가 가능한 가속적정 표준계획이 제시되었다 (Simon 등, 1997). 이 가속적정 표준계획은 임상 시험의 초기 낮은 용량수준에 단지 1명 환자를 할당하여 용량증가의 시간을 단축시키고 또한 할당된 환자가 DLT를 보이지 않는 경우에는 동일 환자의 용량수준을 증가시킴으로써 할당되는 총 환자수를 줄이는 방법(앞에서 설명한 환자내 용량증가방법)이며, 이러한 연구 초기의 가속 단계가 끝나면 표준 3+3계획을 시행한다. 그러나 동일 환자가 처음 용량수준과 나중의 용량수준의 결과에 복합적으로 작용하기 때문에 결과해석에 어려움이 따른다.

소아환자의 경우 특히 임상시험 기간을 단축하고자 하는 시도는 자연스러운 것이며 최근 Skolnik 등 (2008)이 Rolling-6 계획을 제안하였다. 현시점까지 할당된 환자의 DLT 결과를 모두 관측하지 않았어도 임상시험 대상 환자가 확보되는 순간에 환자할당을 지연시키지 않고 적절한 용량수준에 6명까지의 환자를 할당하는 매우 적극적인 진행방법이다. 일반적으로 성인 환자에게서 습득된 정보, 특히 안전성에 대한 정보가 충분히 확보된 상태에서 소아환자의 임상시험을 진행한다는 Lee 등 (2005)에 근거하여 이 Rolling-6 계획은 제안되었다. Skolnik 등 (2008)에 제시된 이 계획의 몇 가지 예로서는 현재 용량수준에 3명 환자가 할당되고 아무도 DLT가 아닌 경우에 투입되는 4번째 환자를 증가된 용량에 할당하고, 현재 용량수준에서 만약 1명이 DLT이거나 또는 3명 환자 중 한 명이라도 DLT이면 새로운 환자는 현재 용량수준에 할당한다. 그러나 만약 2명 또는 그 이상에서 DLT이면 감소된 용량에 투입한다. 표준 3+3계획은 3명씩 환자의 DLT 결과를 모두 관측한 후에 환자할당이 이루어지므로 그 기간 사이에 확보된 환자는 임상시험에 투입되지 못하고 소실될 수 있으나 Rolling-6 계획에서는 6명 환자가 모두 채워진 경우를 제외하고는 계속 환자할당을 진행할 수가 있어 환자소실이 줄어든다.

제 1상 임상시험 연구에서 구체적인 용량수준의 결정은 휘보나치 수열(Fibonacci sequence)에 근거하며 바로 전 투여용량의 100% 증가 후, 67%, 50%, 40%, 30~35% 등의 증가로 서서히 증가하지 만 이와 다른 증가수열도 제시될 수 있다. 구체적으로 Siu 등 (2002)의 임상연구에서는 최소 용량수준(3.6mg)에서 휘보나치 수열로 용량증가하였으나 7번째 용량수준(51.2mg)까지 한 명의 환자도 독성 반응이 없었으므로 제 1상 임상시험 진행 중 규정을 변경하여 8번째, 9번째 용량수준은 전 투여용량의 100% 증가였다.

표준계획의 장점은 미리 정해진 규정에 따라 용량증가 또는 감소하므로 매우 단순하며 프로그램이 요구되지 않는다는 점이다. 한편, 표준계획의 단점은 용량증가 또는 감소가 예측할 수 없는 방향으로 움직이므로 결정된 최대허용용량 MTD 추정량에 대한 불확실성 때문에 독성반응율이 제 1상 임상시험에서 추구하는 목표수준, 예를 들어서 33%를 겨냥할 수 없다는 점이다 (Goodman 등, 1995; Ahn, 1998; Storer, 2001). 또한 Rolling-6 계획은 예외이지만 다른 여러 표준계획에서는 환자의 DLT 결과를 관측한 후에 용량증가 또는 감소하기 때문에 임상시험 조건을 만족하는 환자가 확보되어도 놓칠 수 있고 임상시험 완결까지의 기간이 길어진다는 단점이 있다.

## 2.2. 표준계획으로 시행된 임상시험의 예

표준계획으로 시행된 임상시험의 몇 가지 예로서 임상시험의 예외점을 짐작할 수 있다. Attia 등 (2009)은 악성종양 환자에게 병합약제로 제 1상 임상시험을 실시하여 MTD를 결정하고자 표준 3+3계획을 채택하였다. 신약제와 관련된 독성반응 DLT를 정의하고 적어도 28일간의 일주기(cycle)를 완료하지 못한 환자는 다른 환자로 대치하였다. MTD는 6명 환자 중 1명 이하가 DLT인 용량수준에서 최대 용량수준으로 정한다고 서술하였고, 이는 앞에서의 설명과 일치하는 MTD 정의이다.

Yamamoto 등 (2010)은 림프종백혈병(leukemia lymphoma) 환자 중 재발환자의 치료제로써 KW-0761 항체의 제 1상 임상시험을 표준 3+3계획으로 진행하였다. 용량수준은 시작용량 0.01mg로부터 0.1, 0.5, 1.0mg의 네 단계이며, 최고용량수준인 1.0mg에서 1명만이 독성반응(발진, 열성 백혈구감소증, 또는 단계 4의 백혈구감소증)을 나타내어 MTD에 도달되지 못했음을 논문에서 보고하였다. 더욱

높은 용량수준이 연구에서 감안되어 여러 용량수준 가운데 하나가 MTD이어야 하는데, 용량수준을 잘못 선택하여 MTD를 결정할 수 없었던 연구이다.

Brochstein 등 (2010)은 백혈병 소아의 동종(allogeneic) 골수이식 후에 발생한 부작용인 급성이식대 숙주병(acute graft-versus-host disease; acute GVHD)의 직접적 치료로써 시플리주마브(siplizumab) 신약제에 대한 MTD를 결정하고자 제 1상 임상시험 연구를 시행하였다. 이 연구에서 채택한 용량수준은 0.012, 0.04, 0.12mg의 세 수준으로 각 용량수준에 5명의 환자를 할당하며 마지막 환자가 21일 이상 관측되어 독성반응을 파악할 수 있을 때 용량증가 또는 감소한다. 각 용량수준에서 2명 이상이 신약제와 관련된 부작용이 발생하면 용량증가할 수 없으며 또한 전반적 안전성에서 염려되는 패턴이 보이면 용량증가할 수 없도록 계획하였다. 대부분의 임상연구가, Brochstein 등 (2010)이 언급하였듯이, 급성 GVHD가 발생하기 전 단계의 예방조치에 노력을 기울이고 있기 때문에, 실제로 급성 GVHD의 직접적 치료에 대한 연구가 매우 드문 실정이어서 직접적 치료에 대한 연구가 시급한데도 불구하고 4년 반 정도의 연구기간동안 10명 환자의 환자확보에 그쳤다. 또한 수집된 자료에서 신약제의 안전성에 대해 염려되는 부분이 있어서 연구는 중단되고 MTD는 결정할 수 없었다. 이 연구에서 실제적인 용량증가는 연구후원기관이 전반적인 안전성에 대한 자료에 근거하여 결정하였다고 보고하였고, 구체적인 용량증가 과정이 제시되지 않았다. 따라서 표준 3+2계획, 표준 2+3계획, 또는 표준 3+1+1계획 중 어떤 계획으로 시행되었는지 알 수 없다. 매우 더딘 환자확보는 연구를 중단하는 결과를 초래하였다.

Kantarjian 등 (2010)의 연구는 치료불응 또는 재발(refractory-relapse) 급성백혈병(acute leukemia) 또는 골수형성 이상증후군(myelodysplastic syndrome; MDS) 환자에게 항암치료 대신 사파씨타빈(sapacitabine) 경구약을 투여하는 제 1상 임상시험 연구를 표준 3+3계획으로 진행하였다. 한 주기(cycle) 치료는 3~4주이며, 환자에게 도움이 되는 한에는 주기를 반복하여 지속하게 된다. 어떠한 용량 스케줄, 즉 한 번에 7일간 지속, 또는 두 주에 걸친 3일간 지속과 4일간의 휴식기를 결들인 용량 스케줄이 더욱 적합한 것인지에 대한 확실한 정보가 없는 상태이므로 두 가지 다른 용량 스케줄로 임상 시험을 시행하였으며, 매일 2회 복용으로써 첫 번째 스케줄인 7일간 복용의 1회 용량수준은 (75, 100, 125, 175, 225, 275, 325, 375)mg이며, 이 용량수준에 할당된 환자수는 각각 (3, 1, 3, 3, 3, 6, 9, 7)명으로 총 35명이 참여하였다 (100mg에 단 1명만이 할당된 이유는 악성종양 연구로부터 알려진 안전성의 정보에 따른 것이라고 한다). 두 번째 스케줄의 1회 용량수준은 (375, 425, 475)mg이며, 이 용량수준에 할당된 환자수는 각각 (3, 6, 3)명으로 총 12명이 참여하였다. 첫 번째 용량 스케줄에서는 375mg에서 DLT, 특히 사망이 관측되어 325mg으로 용량감소하여 3명을 다시 할당하고서 연구를 멈추었고, 제 2상 연구의 추천용량은 325mg이다. 한편 두 번째 용량 스케줄에서는 475mg에서 1명의 DLT가 관측되고 다른 환자들에서도 안전성에 대한 위험이 있어 425mg으로 용량감소하여 3명을 다시 할당한 후 연구를 멈추었고, 제 2상 연구의 추천용량은 425mg이다. 두 번째 용량 스케줄은 전형적인 표준 3+3계획의 연구이지만, 첫 번째 용량 스케줄은 약간 벗어나긴 해도 표준 3+3계획에 비슷하게 진행되었다. Kantarjian 등 (2010)은 논문에서 멈춤용량을 MTD로 정의하고 이와 달리 한 단계 아래 용량수준을 추천용량으로 정의하였으나, 일반적으로 한 단계 아래 용량수준이 추천용량이면서 동시에 MTD로 정의하며 이는 본 논문에서 채택한 정의이다.

### 3. 모형중심의 CRM계획

모형중심의 베이지안 방법을 적용시킨 임상시험 계획은 임상시험에 참여하는 모든 환자의 독성반응 자료를 사용하여 다음 환자의 용량수준을 결정한다. 환자가 확보되는대로 진행할 수도 있어 앞에서의 표준계획보다 빠르게 진행된다. 우선 용량-독성반응 곡선에 대한 모형을 가정하며, 용량-독성반응 곡선이란 신약제의 용량  $d_k$  ( $k = 1, \dots, K$ )에 따른 독성반응을  $p_k$ 를 나타내는 곡선이며 일반적으로

S자 형태를 지닌 로지스틱 함수로 표현한다. 특히 하나의 모수( $\beta$ )만으로 표현한 단순 로지스틱 곡선  $p_k = \phi(d_k; \beta) = e^{3+\beta d_k} / (1 + e^{3+\beta d_k})$ 을 사용한다. 여기서 모수  $\beta$ 는 용량-독성반응 곡선의 기울기를 나타내며, 작은  $\beta$ 값은 용량  $d_k$ 에 따른 독성반응의 완만한 증가를 뜻하고, 큰  $\beta$ 값은 용량에 따른 독성반응의 급격한 증가를 뜻한다.

### 3.1. 여러 CRM계획

베이저안 방법인 CRM계획은  $\beta$ 의 사전분포(prior distribution)를 가정하고서 각 환자의 DLT 독성반응 자료가 수집되는 매순간에 현시점까지의 모든 환자 자료를 사용하여  $\beta$ 의 사후분포(posterior distribution)를 베이지 정리를 이용하여 구한 후, 다음 환자의 용량증가 또는 감소를 결정하는 방법이다. 이러한 베이저안 방법을 적용시킨 제 1상 임상시험 계획을 연속재평가방법(continual reassessment method; CRM)이라 한다 (O'Quigley 등, 1990; O'Quigley와 Chevret, 1991). O'Quigley 등 (1990)의 논문에서는 하나의 모수를 가지는 용량-독성반응 곡선으로써  $p_k = \phi(d_k; \beta) = [(\tanh d_k + 1) / 2]^\beta$  곡선과 로지스틱 곡선을 고려하였다. 미리 정해진 환자수로 임상시험을 실시한 후 조건에 가장 적합한 용량수준을 MTD로 정한다.

CRM계획에서 과격한 증가나 감소를 피하고 현재의 용량수준에서 단지 한 단계 증가 또는 감소에 국한시키는 방법도 제시되었다. 환자수를 미리 정하지 않는 방법에서 환자 투입을 멈추고 MTD를 결정하는 시점 결정에 대해 여러 가지가 제안되었으며, 가장 단순한 방법은 동일 용량수준에 6명 환자가 할당되었을 때 멈춘다 (Korn 등, 1994). 또 다른 방법은 MTD로 추정된 용량수준에서의 DLT 확률의 정도(precision)가 원하는 수준에 도달된 시점에서 또는 용량수준이 더 이상 변하지 않을 때 멈추는 것이다 (Piantadosi 등, 1998). 이 방법에서 MTD로 결정된 용량수준의 신뢰구간을 제시할 수 있으며 이는 모형 중심의 제 1상 임상시험 계획의 장점이라 하겠다. 그러나 만약 용량-독성반응 곡선에 대한 모형이 잘못 가정된 경우에, 특히 고용량수준에 환자가 할당되기 쉽기 때문에 이를 보완하고자 수정된 형태의 CRM계획이 제안되었다. 최근의 문헌에 보고되는 CRM계획을 채택한 제 1상 임상시험 연구는 대부분 수정 CRM(modified CRM; mCRM)에 의한 계획이다. 본 논문에서 mCRM계획이라 일컫는 이 수정 CRM계획은 과도용량 투여를 억제하도록 고려된 용량증가(escalation with overdose control; EWOC) 방법이다 (Babb 등, 1998; Babb과 Rogatko, 2001, Rogatko 등, 2005; Tighiouart 등, 2005). 따라서 단순히 EWOC 방법이라고 일컫기도 한다.

환자에게 발생하는 DLT가 관측된 기간의 자료를 모두 감안한 연속재평가방법이 TITE-CRM(time-to-event CRM)계획이며, 아직 DLT가 관측되지 않은 환자의 경우에는 임상시험에서 미리 정해진 관측기간까지 도달하지 않은 부분 관측기간을 감안하므로 모든 환자의 정보를 활용하게 되어 임상시험 기간이 단축되고, 적은 환자수에 근거하여 MTD를 결정할 수 있음이 밝혀졌다 (Cheung과 Chappell, 2000). 대부분의 악성종양 환자는 항암요법과 방사선치료가 병합된 치료를 받는다. 이러한 병합치료에서 독성반응이 특히 더디게 관측되는 질병의 상황에서 독성반응을 관측하기까지 상당한 시간이 요구되므로 확보된 환자를 쉽게 소실하게 되는 표준계획보다도 CRM계획이 더욱 바람직하다. Normolle과 Lawrence (2006)는 이러한 질병 연구에 적합한 TITE-CRM계획을 자세히 설명하고 있다.

여러 CRM계획인 CRM계획, mCRM계획, TITE-CRM계획의 장점은 어떤 규정을 정하여 단순히 용량증가를 시도하는 표준 3+3계획보다도 각 환자의 독성반응 자료가 수집된 매순간마다 현시점까지의 모든 환자의 독성반응 자료 또는 환자에게 발생하는 DLT가 관측된 기간의 자료를 이용하여 용량증가 또는 감소를 결정하는 방법이므로 특히 mCRM계획과 TITE-CRM계획은 효능이 없는 저용량수준과 독성이 강한 고용량수준에 되도록 적은 환자수로 MTD를 결정할 수 있는 효율적인 계획이며, 표준계획보다도 추정된 MTD의 편의와 평균제곱오차가 작다 (Cheung과 Chappell, 2000). 한편, 여러 CRM계획의 단점은 규정에 따라 진행되는 단순한 표준계획에 비해 통계적 이론이 복잡하여 임상연구를 진

행하는 과정에서, 다시 말하면 매 환자 확보 매순간마다 용량증가와 감소의 결정에 프로그램의 활용이 반드시 요구된다. 특히 환자에게 발생하는 DLT가 관측된 기간의 자료를 모두 감안하는 TITE-CRM 계획의 성공적인 적용에는 임상시험의 과정 중 모의실험의 시행이 필수적이며 통계인의 개입이 요구된다.

### 3.2. CRM계획으로 시행된 임상시험의 예

EWOC 방법을 적용한 mCRM계획으로서 시행된 제 1상 임상시험 계획은 Siu 등 (2002)의 연구가 처음이며, 임파중을 포함한 진행성 악성종양 환자가 기존치료에 반응이 없거나 효과적인 치료를 없는 경우에 SAM486A라는 새로운 억제제로 치료하는 제 1상 임상시험이다. 또한 이러한 mCRM계획으로 시행된 백혈병 환자 대상의 제 1상 임상시험 연구가 최근에 와서 증가하고 있다 (O'Donnell 등, 2002; Song 등, 2003; Lonial 등, 2010).

O'Donnell 등 (2002)은 비골수절제(nonmyeloablative) 골수이식(BMT) 후의 이식거부(graft rejection)와 이식대숙주병(GVHD)가 고용량 싸이클로포스파미드(cyclophosphamide; Cy)에 의해 억제되는 동물실험 연구결과에 근거하여, 제 1상 임상시험으로 골수이식전 Cy치료에 대한 연구를 시행하였다. 연구에서 이식전 Cy치료로 인해 공여자의 이식이 성공하고 이식거부가 없게 되는 Cy치료 MTD를 결정하는 것이 목적이다. 처음 2명 환자 중 2명이 이식거부를 보이거나 또는 3명 환자 중 2명이 이식거부를 보일 때를 효능이 없는 경우로 정의하였다. 여기서 최저용량수준은 Cy치료를 받지 않는 것이고 이 최저용량수준에 3명 환자가 할당되었다. 이 최저용량수준에서 2명이 효능이 없어 Cy치료의 용량수준을 14.5mg로 높였다. O'Donnell 등 (2002)의 연구에서 높은 용량수준에 10명의 환자가 할당되었고 처음 6명의 환자에서 효능이 없었으며 연구는 종료되었다. 실제로 두 용량수준으로 연구가 시행되었다고 믿기 어려우며, mCRM계획을 적용하였다고 설명하였으나 논문에 명확한 용량증가의 과정이 제시되지 않았으며 어떤 용량수준이 MTD인지 확실하지 않다.

Song 등 (2003)은 림프종백혈병 환자의 국부방사선치료에 대한 임상연구를 실시하였다. 림프종백혈병 환자는 일반적으로 전신투사 방사선치료를 받게 되며, 골수이식(BMT) 치료가 실패하여 재발된 환자를 대상으로 이식전 질병이 나타났던 부위에 국한하여 5일 동안 방사선량을 1000cGy에서 2000cGy까지를 투사하며 구체적인 방사선량을 제 1상 임상시험으로 알아내고자 한다. DLT는 골수이식(BMT) 후 30일 이내에 나타나는 단계 3 또는 그 이상의 독성반응으로 정하였고, 이 임상시험에서 독성반응의 목표수준은 대략 20%로 정하였다. 총 6명 환자가 MTD 용량수준에서 할당되어야만 이 용량을 추천용량으로 제시할 수 있다. mCRM계획으로 진행된 이 연구는 총 21명 환자 중 1명은 질병이 급속히 진전되어 골수이식(BMT)을 시행할 수 없어 연구에서 제외되었다. 결과적으로 총 20명 환자 자료 중 2000cGy수준에서 6명 환자 중 3명이 단계 3 또는 그 이상의 독성반응을 보였으며, 1500cGy수준에서 6명 환자 중 1명이 독성반응을 보여 MTD 용량수준은 1500cGy수준이다. 논문에 mCRM계획을 적용하였다고 설명하고 있지만 다음의 4장에 제시된 여러 가지 사항의 언급이 없고, Song 등 (2003) 논문의 표 1에 제시된 용량증가 및 감소과정이 제시된 결과를 보면 표준계획으로 진행된 연구가 아닌가 의문시된다.

Lonial 등 (2010)의 연구는 기존의 약제로는 치료반응이 없는 다발성골수종(multiple myeloma) 환자 대상이며, 치료효과가 높은 병합약제를 찾는 과정에서 기존의 고용량 메파란(mephalan) 단독약제에 추가로 보테조미브(bortezomib)의 투여용량에서 효능이 있는지와 투여시점을 파악하고자 실시한 제 1상 임상시험이다. 보테조미브 투여시점에 대해서는 메파란 투여 전과 후의 두 군에 환자를 랜덤 할당하며, 투여용량에 대해서는 과도용량 투여를 억제하는 EWOC 용량증가 방법으로 제 1상 임상시험을 전과 후의 각 군에서 독립되게 실시한다. 전투여군에 19명과 후투여군에 20명이 할당되었다. DLT는 고용량 메파란 투여후 30일 이내에 나타나는 단계 4의 점막염(mucositis), 또는 치료와 관련된

사망(treatment-related mortality; TRM)으로 정하였고, 이러한 독성반응은 고용량 메파란 투여후 100일에 평가하였다. 추가약제인 보테조미브의 용량수준은 1.0, 1.3, 1.6mg로 정하고 각 환자는 한 가지 용량수준에만 할당된다. EWOC 방법에 의한 mCRM계획에서 비혈액학적 독성 단계 3 또는 혈액학적 독성 단계 4인 확률의 목표수준은 0.33으로 정하였고, 사전분포는 구간 (0.4, 1.6)mg에 걸친 일양분포(uniform distribution)로, 초기 DLT 독성반응율은 1.0, 1.3, 1.6mg에 걸친 일양분포로 정하였다. 병합 약제의 가장 낮은 용량수준에서 전투여군에서는 1명이 단계 5의 대사성(metabolic) DLT 독성반응이, 후투여군에서도 1명이 DLT 독성반응이 최소용량수준인 1.6mg에서 관측되었고, 다른 용량수준에서는 DLT 독성반응이 없었다. 투여시점이 서로 다른 전후 두 군의 독성반응과 치료효과가 차이가 없었으며 연구자들은 임상에서의 관심인 여러 의학적 수치로 병합치료제의 효과뿐만 아니라 안전함을 밝혔다. 전투여군에서는 (1.0, 1.3, 1.6)mg 용량수준에 (5, 3, 11)명이, 후투여군에서는 (1.0, 1.3, 1.6)mg 용량수준에 (3, 3, 14)명이 할당되었음을 보고하였고, MTD의 명확한 제시는 없으나 두 군 모두에서 최대용량이 MTD인 것으로 짐작된다.

Tevaarwerk 등 (2011)의 연구는 악성종양의 연구로써 치료 상황이 골수이식과 크게 다른 점이 없는 데, 이 연구에서 채택한 제 1상 임상시험 계획은 환자에게 발생하는 DLT 사건이 관측되기까지의 시간을 고려한 연속재평가방법인 TITE-CRM계획으로 진행되었으며, 연속적 용량수준의 계산을 허용하였다. 특히 이 TITE-CRM계획은 여러 주기에 걸쳐 지속되는 치료로 인해 독성반응이 늦게 나타나는 경우에도 모든 환자의 관측기간을 감안하여 MTD를 추정할 수 있는 계획이다. 이 연구의 용량증가 또는 감소에 대해서는 4장에서 설명한다.

#### 4. TITE-CRM계획의 진행과정과 모의실험

환자의 관측기간을 비롯한 모든 정보를 활용하는 TITE-CRM계획의 진행에서 가정하게 되는 사전 분포 형태 및 여러 가지 사항이 MTD를 크게 잘못 예측하는 경우를 방지하기 위해 모의실험은 필수적인 요소이다. 이러한 모의실험을 설명하고 몇 가지 경우에서 모의실험을 제시한다. TITE-CRM계획의 진행과정에서 우선 정해져야 하는 사항은 (1) 용량제한 독성반응인 DLT의 임상적 정의, (2) 독성반응 목표수준(target toxicity level), (3) 환자의 독성반응을 알아보는 일정 관측기간, (4) 용량수준수, (5) 각 용량수준별 초기 DLT 독성반응율, (6) 첫 환자의 시작 용량수준, (7) 용량증가, 용량감소 및 멈춤의 규정과 제한점이다. 이 밖에도 전체 환자수, 사전 분포, 분포의 모수 추정방법, 환자의 관측기간을 감안한 가중방법이 정해져야 한다. 임상의와 통계인이 연구를 시작하기 전에 과거 임상연구에 근거하여 이러한 여러 가지 사항을 함께 결정한다. 새로운 환자의 다음 용량수준은 통계인이 계산한 결과로써 알게 되며 임상의에게 이 용량수준을 전달하여 임상시험이 진행된다.

이제 계산과정을 간략히 설명한다. 기호를 정의하여 위의 (2) 독성반응 목표수준을  $p^* \in (0, 1)$ 로, (3) 환자의 독성반응을 관측하는 일정한 관측기간을  $T$ 로, (4) 용량수준의 수를  $K$ 로, (5) 각 용량수준을  $d_1, \dots, d_K$ 로 나타낸다. 각 용량에서의 초기 DLT 독성반응율을  $\hat{p}_{01}, \dots, \hat{p}_{0K}$ 로 나타내며,  $J$ 번째 환자할당으로 추정된 사후(posterior) DLT 독성반응율을  $\hat{p}_{J1}, \dots, \hat{p}_{JK}$ 로 나타낸다.

CRM계획은 용량-독성반응 모형으로 모수가 하나인 단순 로지스틱 모형을 선택할 때 용량에 따라 독성반응 곡선을 단조함수로 증가한다. 즉

$$p_k = \phi(d_k; \beta) = \frac{e^{3+\beta d_k}}{1 + e^{3+\beta d_k}} \quad (4.1)$$

이로부터 0번째 단계의 초기 독성반응율  $\hat{p}_{0k}$ 의 함수로 표현한 용량수준  $d_k$ 는 다음과 같다.

$$d_k = \log \left( \frac{\hat{p}_{0k}}{1 - \hat{p}_{0k}} \right) - 3. \quad (4.2)$$

베이지안 방법에 의한  $\beta$ 의 추정에는 사전분포(prior distribution)  $\pi(\beta)$ 를 지정해야 하며 임상시험의 시작 단계에서는 단순하게 정규분포  $N(1, \sigma^2)$ 로 지정하고 임상시험이 진행되면서 수정할 수 있다. 이제  $J$ 번째 환자까지의 자료가 확보되고  $J+1$ 번째 환자가 할당되어야 할 때  $J$ 번째 환자까지의 용량수준  $\{x_1, \dots, x_J\}$ 와 독성반응의 결과  $\{y_1, \dots, y_J\}$ 를 알고 있으며, 여기서  $y_j = 1$ 은 독성반응이 있음을 나타내고  $y_j = 0$ 은 독성반응이 없음을 나타낸다. 또한  $\{u_1, \dots, u_J\}$ 은 이 시점에서의 각 환자의 관측기간을 나타내며, 여기서  $0 < u_j \leq T$ 이다. 각 환자의 관측기간에 따라 다른 가중을 고려하여 독성반응이 관측된 환자에게는 가중 1을, 아직 독성반응을 관측하지 못한 환자에게는 부분 관측기간을 감안한 관측기간에 따라 직선적으로 증가하는 가중을 부여할 수 있겠다. 즉 이러한 직선적 가중함수는 다음과 같다.

$$w_j = \begin{cases} \frac{u_j}{T}, & \text{if } y_j = 0, \\ 1, & \text{if } y_j = 1. \end{cases} \quad (4.3)$$

연구의 특성에 따라 이와 다른 증가형태의 가중을 가정할 수도 있다. 이제 환자의 모든 정보를 이용한 가중우도함수(weighted likelihood function)는 분포함수(distribution function)에 가중을 부여하여 다음과 같다.

$$L_J(\beta) = \prod_{j=1}^J (w_j \phi(x_j; \beta))^{y_j} (1 - w_j \phi(x_j; \beta))^{1-y_j}. \quad (4.4)$$

$J$ 번째 환자까지의 자료를 이용한  $\beta$ 의 추정값은 베이지정의를 이용하여 다음과 같은 사후 평균(posterior mean)이 적절하며, 이 때 앞에서 지정한 사전분포  $\pi(\beta)$ 를 사용하게 된다.

$$\hat{\beta}_J = \frac{\int_0^\infty \beta L_J(\beta) \pi(\beta) d\beta}{\int_0^\infty L_J(\beta) \pi(\beta) d\beta}. \quad (4.5)$$

일반적으로 적분이 수리적으로 쉽게 해결될 수 없기 때문에 프로그램을 사용하여  $\hat{\beta}_J$ 를 구한다. 이제  $\hat{\beta}_J$ 를 이용하여 각 용량수준 ( $k = 1, \dots, K$ )에 해당하는 DLT 확률  $p_{jk} = \phi(d_k; \hat{\beta}_J)$ 이 추정되며, 이로부터 새로이  $J+1$ 번째 환자가 할당받게 되는 용량수준은 목표수준인  $p^*$ 에 가장 가까우면서도  $p^*$ 를 결코 초과하지 않는  $\hat{p}_{jk}$ 를 구하고 이에 해당하는 용량수준  $d_k$ 가 된다.

모의실험이 가능한 TITE-CRM계획의 프로그램으로, 비용없이 사용가능한 것은 현재 세 가지를 들 수 있다. 첫 번째는 Johns Hopkins의 Oncology Biostatistics, Sidney Kimmel Comprehensive Cancer Center 출처의 CRM(<http://www.cancerbiostats.onc.jhmi.edu/software.cfm>)이고, 두 번째는 MD Anderson Cancer Center의 Department of Mathematics 출처의 CRM(<http://biostatistics.mdanderson.org/SoftwareDownload>)이며, 세 번째는 University of Michigan Comprehensive Cancer Center 출처의 SAS 프로그램(<http://roadrunner.cancer.med.umich.edu/wiki/titecrm.v7.tar.gz>)이 있다.

이제 mCRM계획 또는 TITE-CRM계획에서의 모의실험을 세 번째 SAS 프로그램을 이용하여 제시한다. CRM계획의 적용으로 임상시험의 초기에 구해진 독성반응율의 추정값이 참값에 가까울 때 환자가 확보되는대로 용량증가와 감소로 인해 빠른 시일에 MTD에 도달할 수 있고 또한 많은 환자가 MTD 용량수준 또는 근방의 용량수준에 할당될 수 있음이 가장 바람직하다. CRM계획에서 가정하는 여러 사항에 달려 있으며, 몇 가지 사항이 달라지면서 임상시험에서 어떠한 결과에 도달하는지를 모의실험으로 알아본다.

Tevaarwerk 등 (2011)의 연구에서와 같이 용량수준수를 8단계로 정하고, 초기독성반응율을 I, II, III의 세 가지로 정하였으며, 초기독성반응율 I은 참독성반응율에 비해 낮은 용량수준의 독성반응율은



표 1: TITE-CRM 계획 모의 실험에서의 모수

모수	용량수준								
	1	2	3	4	5	6	7	8	
참독성반응율	0.005	0.032	0.100	0.210	0.299	0.345	0.470	0.640	
초기독성반응율	I	0.001	0.010	0.090	0.230	0.350	0.400	0.560	0.750
	II	0.001	0.050	0.100	0.200	0.300	0.350	0.450	0.650
	III	0.010	0.020	0.070	0.150	0.240	0.300	0.420	0.500

표 2: TITE-CRM 계획 모의 실험 결과

모수	용량수준								MTD 미정	
	1	2	3	4	5	6	7	8		
일양분포, 목표수준 0.3										
초기독성반응율	I	0.001	0.010	0.090	0.230	0.350	0.400	0.560	0.750	
	II	0.001	0.050	0.100	0.200	0.300	0.350	0.450	0.650	
	III	0.010	0.020	0.070	0.150	0.240	0.300	0.420	0.500	
사전 일양분포, 초기독성반응율 II										
목표수준	0.15	0.007	0.064	0.288	0.480	0.123	0.025	0.008		0.005
	0.20		0.030	0.125	0.479	0.234	0.109	0.002		0.003
	0.25		0.002	0.058	0.383	0.306	0.171	0.080		
	0.30		0.001	0.019	0.183	0.346	0.277	0.171	0.003	
사전 정규분포, 목표수준 0.3										
표준편차 $\sigma$	0.3		0.001	0.019	0.182	0.313	0.289	0.191	0.004	0.001
	0.5		0.001	0.010	0.187	0.307	0.261	0.206	0.013	0.015
	1.0			0.022	0.189	0.260	0.253	0.203	0.057	0.016
	1.5			0.020	0.190	0.259	0.249	0.203	0.061	0.018

더욱 낮고 높은 용량수준의 독성반응율은 더욱 높은 경우이며, II는 참독성반응율과 비슷한 경우이고, III은 모든 용량수준에서 참독성반응율보다 낮은 경우이다. 각 경우에 대한 1,000번의 모의실험에서 지정한 모수와 모의실험 결과가 각각 표 1과 2에 제시되었다.

독성반응 목표수준을 0.3으로 정하고 초기독성반응율 I, II, III에 따라 결정된 MTD를 모의실험으로 알아본 결과, MTD는 5가 옳은 용량수준인데 I의 경우에는 용량수준 4가 MTD로 채택되는 확률이 초기독성반응율 II, III의 경우보다 높아 35.6%이다. 한편 초기독성반응율 III의 경우에는 참독성반응율과 비슷한 경우인 II에 비해서 고용량수준 6 또는 8이 MTD로 채택되는 확률이 높았다. 이러한 경향은 독성반응 목표수준을 0.25로 지정하여도 표에 제시하지는 않았으나 동일하다. 이제 초기독성반응율 III의 경우에 독성반응 목표수준을 여러 가지로 변경하여 MTD 결과를 살펴보았을 때, 참 독성반응율에 해당하는 용량수준과 또한 한 단계 높은 용량수준이 MTD로 결정되는 경향을 볼 수 있었다. 마지막으로 사전분포를 정규분포로 정한 경우를 살펴보면 일양분포를 지정한 경우보다도 고용량수준이 MTD로 지정되는 경향을 볼 수 있고, 특히 표준편차를 크게 지정하는 경우에는 고용량수준을 잘못되게 MTD로 지정하는 경향이 더욱 심화되어 MTD가 결정되지 못하는 경우의 확률도 2% 정도가 되었다.

## 5. 여러 계획의 비교

임상에서 다루는 여러 질병의 특성이 매우 다양하고, 해당 질병에 대한 임상적 치료 및 치료 대상자가 다를 수 있는 상황에서 제 1상 임상시험의 여러 계획들을 함께 비교한다는 것은 어려운 과제이다. 이는 Storer (2001)의 모든 상황 또는 여러 가지 용량-독성반응 곡선에 대해 최적인 단 하나의 계획은 존재하지 않는다는 언급과 같은 뜻이다. 여러 계획을 모의실험으로 비교한 연구자들은 대부분 현실 상황

을 염두에 두고 연구하기에 모의실험의 모수가 여러모로 다양할 수 밖에 없으며, 이를 염두에 두고서 여러 논문의 모의실험 결과를 대략 요약하는 것에 그친다. 여러 계획을 비교한 논문을 언급한 후에 모의실험 결과를 간략히 설명한다.

CRM계획이 처음 제시된 후 표준 3+3계획과 CRM계획의 MTD 추정 결과가 여러 연구자들에 의해 모의실험으로 비교되었다 (Storer, 2001; Ahn, 1998; Goodman 등, 1995). 최근의 Gerke와 Siedentop (2008) 연구도 표준 3+3계획과 CRM계획을 모의실험으로 비교한 연구이다. 한편, MTD가 추정되기 까지 긴 연구기간이 요구되는 표준 3+3계획의 단점을 보완하고자 Skolnik 등 (2008)은 Rolling-6 계획을 제안하였으며, 이러한 임상시험 기간의 단축은 특히 소아암 환자의 연구에 바람직하다. 따라서 Skolnik 등 (2008)은 모의실험에서 표준 3+3계획과 Rolling-6 계획을 비교하였다.

최근에 Onar-Thomas와 Xiong (2010)는 표준 3+3계획, Rolling-6 계획 및 CRM계획을 모의실험으로 비교하였다. 한편 독성반응이 늦게 나타나는(late-onset toxicity) 질병에 대한 제 1상 임상시험에서 모든 환자의 관측기간을 감안한 TITE-CRM계획을 제안한 Cheung과 Chappell (2000)은 기존의 CRM계획과 TITE-CRM계획을 비교한 모의실험 결과를 제시하였고, Normolle과 Lawrence (2006)는 TITE-CRM계획을 표준 3+3계획과 비교하였다.

이제 여러 계획을 비교한 모의실험 결과를 설명한다. Normolle과 Lawrence (2006)는 독성반응이 늦게 나타나는, 다시 말하면 추가로 네, 다섯 명의 환자가 연구에 투입된 시점에서야 비로써 첫 번째 환자의 독성반응이 관측되는 경우에 TITE-CRM계획은 표준 3+3계획보다도 임상시험의 기간이 단축되고 더욱 정확히 MTD를 추정하며, MTD 용량수준 또는 그 이상의 용량수준, 그러나 위험수준을 넘지 않는 용량수준에 할당되는 환자가 많다고 보고하였다. 그러나 TITE-CRM계획에서 기대되는 이러한 바람직한 결과는 연구 시작 전에 결정하는 용량수준이 독성반응 목표수준을 포함한다는 조건하에서 성립한다.

Onar-Thomas와 Xiong (2010)가 표준 3+3계획, Rolling-6 계획 및 CRM계획을 모의실험으로 비교한 결과를 살펴보면, 앞의 두 계획은 CRM계획과 비교해 볼 때 MTD 보다도 낮은 용량수준에 많은 환자수가 할당된다고 보고하였다. 또한 임상시험의 기간은 Rolling-6 계획이 표준 3+3계획보다도 짧지만 그러나 Rolling-6 계획과 CRM계획을 비교할 때 환자확보의 속도에 따라 결과가 달라진다. 환자확보가 매우 빠른 경우에는 Rolling-6 계획의 임상시험 기간은 CRM계획보다도 짧지만, 더디거나 중간 속도의 환자확보의 경우에는 그 반대가 될 수 있음을 보고하였다. 최근에 제시된 표준 3+3계획과 CRM계획을 비교한 Gerke와 Siedentop (2008) 논문에서는 표준 3+3계획은 MTD 추정에 있어 부정확한 방법이지만 MTD의 참값보다 낮은 용량수준을 MTD 추정값으로 택하는 확률이 가장 높아 이러한 측면에서 안전한 방법이라고 보고하였다. 또한 Gerke와 Siedentop (2008)는 환자수에 대해 언급하였으며, 모든 용량수준에서 전반적으로 낮은 독성반응의 경우에는 CRM 계획보다도 표준 3+3계획이 더욱 적은 환자수가 요구되지만, 그렇지 않은 경우에는 CRM 계획이 더욱 적은 환자수가 요구되며 MTD까지의 용량수준 단계가 클수록 요구되는 이러한 두 계획에서의 환자수의 차이는 더욱 벌어진다고 보고하였다.

모의실험이 의존하지 않고도 알 수 있는 자명한 사실이기도 하고 동시에 여러 논문에 걸쳐 보고된 일반적인 결과는, 표준 3+3계획과 표준계획의 변형 계획들은 독성반응 목표수준을 정하지 않고 시작하는 계획이기 때문에 사전정보에 근거한 CRM계획보다는 MTD 추정량이 부정확하다. 그러나 용량수준이 2~4 단계라면 모형에 근거한 CRM계획의 장점이 크지 않으며 이는 축적된 자료가 충분하지 않은 것에 따른 것이다. 또한 독성반응의 확률이 낮아서 독성반응 사건 발생이 매우 드문 경우에 CRM계획은 장점이 될 수 없다. 한편 용량수준의 수가 5, 6 또는 그 이상인 경우, 또는 독성반응의 확률이 전반적으로 낮지 않은 경우에는 모형에 근거한 CRM계획은 장점을 지닌다.

## 6. 결론

제 1상 임상시험에서 MTD 추정값은 일반적으로 질병의 특성에 따라 채택되는 여러 가지 요소, 예를 들어서 용량수준수, 출발 용량수준, 각 용량수준에 할당되는 환자수, 용량증가 방법에 의해 결정된다. 현실적으로 매우 더딘 환자확보는 연구를 중단하는 결과를 초래함을 임상연구에서 볼 수 있었으며 이러한 상황에서는 표준 3+3계획보다도 용량증가 속도가 빠를 수 있는 표준 3+1+1 계획이나 최근에 와서 제시된 Rolling-6 계획이 더욱 적합하다. Rolling-6 계획은 한 용량수준에 6명 환자가 모두 채워진 경우를 제외하고는 계속 환자할당을 진행할 수가 있어 환자소실을 방지할 수 있고 임상시험의 기간이 단축될 수 있는 계획이다.

제 1상 임상시험에서 적절한 용량수준의 선택에는 많은 사려가 요구된다. 특히 말기암이나 백혈병 환자의 제 1상 임상시험은 주로 치료가 불응인 환자를 대상으로 저용량수준이나 고용량수준에 되도록 적은 환자가 할당되고 MTD 용량수준 근방에 많은 환자가 할당되어야 한다. 그렇기 위해서는 용량수준은 MTD 용량수준을 품는 선택이어야 한다. 그러나 임상연구에서 때때로 보수적인 용량수준을 선택하여 용량수준의 최대값이 MTD로 결정되어 MTD의 참값은 이 용량수준의 최대값 이상일 수 있음을 볼 수 있다. 임상연구자는 논문에 MTD를 결정할 수 없었다고 보고하거나 또는 용량수준의 최대값을 MTD로 보고하는데 이는 실제로 MTD 용량수준이라 판단되지 않는다. 이러한 입장에서 단지 두 용량수준으로 시행된 임상연구는 결코 바람직한 일이 아니다.

신약제의 경우에 논문에 보고된 바도 없기 때문에 임상연구자는 구체적인 정보가 부족하여 최적의 용량 스케줄을 모르고 있다. 예를 들어서 한 주의 나날에 걸친 복용, 또는 며칠간의 휴식을 사이에 둔 두 주에 걸친 복용이 바람직한지 알 수가 없다. 여러 용량 스케줄에서의 용량수준의 선택은 어려운 과제 중 하나이다. 표준 3+3계획으로 진행된 Kantarjian 등 (2010)의 연구에서 두 가지 다른 용량 스케줄이 있고 두 번째 용량 스케줄의 시작 용량수준이 첫 번째 용량 스케줄에서 밝혀진 MTD로 결정되어 두 번째 용량 스케줄의 시행이 첫 번째 스케줄이 거의 끝나는 시점인 것으로 짐작되며 결국 환자의 안전성에 문제가 있어 12명의 환자에서 연구를 멈추었다. 이러한 경우에는 첫 번째 용량 스케줄부터 모든 정보를 이용할 수 있는 CRM계획의 이용이 바람직하며 이로부터 과악한 용량-독성반응을 곡선의 형태에 근거하여 두 번째 용량 스케줄도 CRM계획으로 정해가는 것이 안전하다. 즉 두 다른 용량 스케줄의 용량-독성반응 곡선의 차이를 모의실험의 활용으로 임상시험의 시행 중에 빨리 과악한 후, 용량수준을 결정하고 변경하는 것이 도움이 된다.

다수 통계인에 의해 지속적으로 연구되어진 CRM계획의 임상연구에의 활발한 적용은 2000년도 이후부터 찾아 볼 수 있다. 본 논문에서 설명한 백혈병환자의 제 1상 임상시험에서 Lonial 등 (2010)의 다발성골수종 환자의 연구와 Tevaarwerk 등 (2011)의 악성종양의 최근 논문에서 비로소 CRM계획의 적용과정의 자세한 서술을 볼 수 있다. 더욱 활발한 CRM계획과 TITE-CRM계획의 이용이 장차 임상연구에서 촉구된다 하겠다.

## 참고 문헌

- Ahn, C. (1998). An evaluation of phase I cancer clinical trial designs, *Statistics in Medicine*, **17**, 1537–1549.
- Attia, S., Morgan-Meadows, S., Holen, K. D., Bailey, H. H., Eickhoff, J. C., Schelman, W. R., Traynor, A. M., Mulkerin, D. L., Campbell, T. C., McFarland, T. A., Huie, M. S., Cleary, J. F., Tevaarwerk, A. J., Alberti, D. B., Wilding, G. and Liu, G. (2009). Dose-escalation study of fixed-dose rate gemcitabine combined with capecitabine in advanced solid malignancies, *Cancer Chemotherapy and Pharmacology*, **64**, 45–51.

- Babb, J., Rogatko, A. and Zacks, S. (1998). Cancer phase I clinical trials: Efficient dose escalation with overdose control, *Statistics in Medicine*, **17**, 1103–1120.
- Babb, J. S. and Rogatko, A. (2001). Patient specific dosing in a cancer phase I clinical trial, *Statistics in Medicine*, **20**, 2079–2090.
- Brochstein, J. A., Grupp, S., Yang, H., Pillemer, S. R. and Geba, G. P. (2010). Phase-I study of siplizumab in the treatment of pediatric patients with at least grade II newly diagnosed acute graft-versus-host disease, *Pediatric Transplantation*, **14**, 233–241.
- Cheung, Y. K. and Chappell, R. (2000). Sequential designs for phase I clinical trials with late-onset toxicities, *Biometrics*, **56**, 1177–1182.
- Dixon, W. J. and Mood, A. M. (1948). A method for obtaining and analyzing sensitivity data, *Journal of the American Statistical Association*, **43**, 109–126.
- Fiedler, W., Mesters, R., Heuser, M., Ehninger, G., Berdelb, W. E., Zirrgiebele, U., Robertson, J. D., Puchalskig, T. A., Collinsf, B., Jurgensmeierf, J. M. and Serve, H. (2010). An open-label, phase I study of cediranib (RECENTIN tm) in patients with acute myeloid leukemia, *Leukemia Research*, **34**, 196–202.
- Gerke, O. and Siedentop, H. (2008). Optimal phase I dose-escalation trial designs in oncology—a simulation study, *Statistics in Medicine*, **27**, 5329–5344.
- 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.
- Kantarjian, H., Garcia-Manero, G., O'Brien, S., Faderl, S., Ravandi, F., Westwood, R., Green, S. R., Chiao, J. H., Boone, P. A., Cortes, J. and Plunkett, W. (2010). Phase I clinical and pharmacokinetic study of oral sapacitabine in patients with acute leukemia and myelodysplastic syndrome, *Journal of Clinical Oncology*, **28**, 285–291.
- 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, D. P., Skolnik, J. M. and Adamson, P. C. (2005). Pediatric phase I trials in oncology: An analysis of study conduct efficiency, *Journal of Clinical Oncology*, **23**, 8431–8441.
- Lonial, S., Kaufman, J., Tighiouart, M., Nooka, A., Langston, A. A., Heffner, L. T., Torre, C., McMillan, S., Renfro, H., Harvey, R. D., Lechowicz, M. J., Khoury, H. J., Flowers, C. R. and Waller, E. K. (2010). A phase I/II trial combining high-dose melphalan and autologous transplant with bortezomib for multiple myeloma: A dose- and schedule-finding study, *Clinical Cancer Research*, **16**, 5079–5086.
- Mick, R. and Ratain, M. J. (1993). Model-Guided determination of maximum tolerated dose in phase I clinical trials: Evidence for increased precision, *Journal of the National Cancer Institute*, **85**, 217–223.
- Normolle, D. and Lawrence, T. (2006). Designing dose-escalation trials with late-onset toxicities using the time-to-event continual reassessment method, *Journal of Clinical Oncology*, **24**, 4426–4433.
- O'Donnell, P. V., Luznik, L., Jones, R. J., Vogelsang, G. B., Leffell, M. S., Phelps, M., Rhubarb, P., Cowan, K., Piantados, S. and Fuchs, E. J. (2002). Nonmyeloablative bone marrow transplantation from partially HLA-mismatched related donors using posttransplantation cyclophosphamide, *Biology of Blood and Marrow Transplantation*, **8**, 377–386.
- Onar-Thomas, A. and Xiong, Z. (2010). A simulation-based comparison of the traditional method, rolling-6 design and a frequentist version of the continual reassessment method with special attention to trial duration in pediatric phase I oncology trials, *Contemporary Clinical Trials*, **31**, 259–270.
- O'Quigley, J. and Chevret, S. (1991). Methods 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, L. (1990). Continual reassessment method: A practical design for phase I clinical trials in cancer, *Biometrics*, **46**, 33–48.

- Park, I. H. and Song, H. H. (1999). Estimation of maximal tolerated dose in sequential phase I clinical trials, *The Korean Communications*, **6**, 543–564.
- Piantadosi, S., Fisher, J. D. and Grossman, S. (1998). Practical implementation of a modified continual reassessment method for dose-finding trials, *Cancer Chemotherapy and Pharmacology*, **41**, 429–436.
- Rogatko, A., Babb, J. S., Tighiouart, M., Khuri, F. R. and Hudes, G. (2005). New paradigm in dose-finding trials: Patient-specific dosing and beyond phase I, *Clinical Cancer Research*, **11**, 5342–5346.
- Simon, R., Freidlin, B., Rubinstein, L., Arbut, S. G. and Christian, M. C. (1997). Accelerated titration designs for phase I clinical trials in oncology, *Journal of National Cancer Institute*, **89**, 1138–1147.
- Siu, L. L., Rowinsky, E. K., Hammond, L. A., Weiss, G. R., Hidalgo, M., Clark, G. M., Moczygemba, J., Choi, L., Linnartz, R., Barbet, N. C., Sklenar, I. T., Capdeville, R., Gan, G., Porter, C. W., Von Hoff, D. D. and Eckhardt, S. G. (2002). A Phase I and pharmacokinetic study of SAM486A, a novel polyamine biosynthesis inhibitor, administered on a daily-times-five every-three-week schedule in patients with advanced solid malignancies, *Clinical Cancer Research*, **8**, 2157–2166.
- Skolnik, J. M., Barrett, J. S., Jayaraman, B., Patel, D. and Adamson, P. C. (2008). Shortening the timeline of pediatric phase I trials: The rolling six design, *Journal of Clinical Oncology*, **26**, 190–195.
- Song, D. Y., Jones, R. J., Welsh, J. S., Haulk, T. L., Korman, L. T., Noga, S., Goodman, S., Herman, M., Mann, R., Marcellus, D., Vogelsang, G., Ambinder, R. F. and Abrams, R. A. (2003). Phase I study of escalating doses of low-dose-rate, locoregional irradiation preceding cytoxan-TBI for patients with chemotherapy-resistant non-Hodgkin's or Hodgkin's lymphoma, *International Journal of Radiation Oncology, Biology, Physics*, **57**, 166–171.
- 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.
- Tevaarwerk, A., Wilding, G., Eickhoff, J., Chappell, R., Sidor, C., Arnott, J., Bailey, H., Schelman, W. and Liu, G. (2011). Phase I study of continuous MKC-1 in patients with advanced or metastatic solid malignancies using the modified time-to-event continual reassessment method (TITE-CRM) dose escalation design, *Investigative New Drugs*, Published online: 12 January.
- Tighiouart, M., Rogatko, A. and Babb, J. S. (2005). Flexible Bayesian methods for cancer phase I clinical trials: dose escalation with overdose control, *Statistics in Medicine*, **24**, 2183–2196.
- Wetherill, G. B. (1963). Sequential estimation of quantal response curves (with discussion), *Journal of the Royal Statistical Society, Series B*, **25**, 1–48.
- Yamamoto, K., Utsunomiya, A., Tobinai, K., Tsukasaki, K., Uike, N., Uozumi, K., Yamaguchi, K., Yamada, Y., Hanada, S., Tamura, K., Nakamura, S., Inagaki, H., Ohshima, K., Kiyoi, H., Ishida, T., Matsushima, K., Akinaga, S., Ogura, M., Tomonaga, M. and Ueda, R. (2010). Phase I study of KW-0761, a defucosylated humanized anti-CCR4 antibody, in relapsed patients with adult T-cell leukemia-lymphoma and peripheral T-cell lymphoma, *Journal of Clinical Oncology*, **28**, 1591–1598.

## Continual Reassessment Method in Phase I Clinical Trials for Leukemia Patients

Joo-Hyoung Lee<sup>a</sup>, Hae-Hiang Song<sup>1,a</sup>

<sup>a</sup>Department of Medical Life Sciences, Division of Biostatistics, Catholic University

---

### Abstract

The traditional method of 3+3 standard design and model-based Bayesian continual reassessment method (CRM) are commonly used in Phase I clinical trials to identify the maximal tolerated dose (MTD) of a new drug. In this paper we review clinical examples of Phase I trials that were carried out in patients with refractory or relapsed leukemia and myelodysplastic syndrome. The recently proposed 3+1+1 design and rolling-6 design can shorten the trial duration, when a very slow accrual of patients with a simple 3+3 standard design may result in the untimely termination of trials. Too conservative approaches in determining the dose levels in Phase I clinical trials can leave clinical investigators unable to accurately determine the MTD. When determining future patient doses, the designs that use a time-to-event CRM can cooperate late toxicities by accounting for the proportion of the observation period of each enrolled patient. With the CRM design, simulations under different scenarios during the trial are important in detecting the under- or over-estimation of the initial estimate of the dose-limiting toxicity rate for each dose level. We present the advantages and drawbacks of the designs used in Phase I clinical trials for leukemia patients.

**Keywords:** Phase I clinical trial, maximal tolerated dose, standard design, continual reassessment method, leukemia.

---

<sup>1</sup> Corresponding author: Professor, Department of Medical Life Sciences, Division of Biostatistics, Catholic University, Medical College, Seoul, Integrated Research Center for Genome Polymorphism, The Catholic University of Korea, 505 Banpo-Dong, Seocho-Gu, Seoul 137-701, Korea. E-mail: hhsong@catholic.ac.kr