

등록번호

안내서-0235-03



# 항암제 임상시험 가이드라인 [민원인 안내서]

2023. 11.



식품의약품안전처

식품의약품안전평가원

의약품심사부 종양항생약품과

## 지침·안내서 제·개정 점검표

명칭

항암제 임상시험 가이드라인

아래에 해당하는 사항에 체크하여 주시기 바랍니다.

등록대상 여부	<input type="checkbox"/> 이미 등록된 지침·안내서 중 동일·유사한 내용의 지침·안내서가 있습니까?	<input type="checkbox"/> 예 <input checked="" type="checkbox"/> 아니오
	☞ 상기 질문에 '예'라고 답하신 경우 기존의 지침·안내서의 개정을 우선적으로 고려하시기 바랍니다. 그럼에도 불구하고 동 지침·안내서의 제정이 필요한 경우 그 사유를 아래에 기재해 주시기 바랍니다. (사유 : )	
	<input type="checkbox"/> 법령(법·시행령·시행규칙) 또는 행정규칙(고시·훈령·예규)의 내용을 단순 편집 또는 나열한 것입니까?	<input type="checkbox"/> 예 <input checked="" type="checkbox"/> 아니오
	<input type="checkbox"/> 단순한 사실을 대외적으로 알리는 공고의 내용입니까?	<input type="checkbox"/> 예 <input checked="" type="checkbox"/> 아니오
	<input type="checkbox"/> 1년 이내 한시적 적용 또는 일회성 지시·명령에 해당하는 내용입니까?	<input type="checkbox"/> 예 <input checked="" type="checkbox"/> 아니오
	<input type="checkbox"/> 외국 규정을 번역하거나 설명하는 내용입니까?	<input type="checkbox"/> 예 <input checked="" type="checkbox"/> 아니오
	<input type="checkbox"/> 정보제공 등 직원 교육용 자료입니까?	<input type="checkbox"/> 예 <input checked="" type="checkbox"/> 아니오
☞ 상기 사항 중 어느 하나라도 '예'에 해당되는 경우에 지침·안내서 등록 대상이 아닙니다. 지침·안내서 제·개정 절차를 적용하실 필요는 없습니다.		
지침·안내서 구분	<input type="checkbox"/> 내부적으로 행정사무의 통일을 기하기 위하여 반복적으로 행정사무의 세부기준이나 절차를 제시하는 것입니까? (공무원용)	<input type="checkbox"/> 예(☞지침서) <input checked="" type="checkbox"/> 아니오
	<input type="checkbox"/> 법령, 고시·훈령·예규와 같은 규정 또는 식약처장이 정한 특정한 사안에 대하여 그 절차 등의 내용을 알기 쉽게 풀어 설명하거나 식약처의 입장을 기술하는 것입니까? (민원인용)	<input checked="" type="checkbox"/> 예(☞안내서) <input type="checkbox"/> 아니오
기타 확인 사항	<input type="checkbox"/> 상위 법령을 일탈하여 새로운 규제를 신설·강화하거나 민원인을 구속하는 내용이 있습니까?	<input type="checkbox"/> 예 <input checked="" type="checkbox"/> 아니오
	☞ 상기 질문에 '예'라고 답하신 경우 상위법령 일탈 내용을 삭제하시고 지침·안내서 제·개정 절차를 진행하시기 바랍니다.	

상기 사항에 대하여 확인하였음.

2023 년 11 월 30 일

담당자  
확 인(부서장)

홍 정 희

이 안내서는 '항암제 임상시험'에 대하여 알기 쉽게 설명하거나 식품의약품안전처의 입장을 기술한 것입니다.

본 안내서는 대외적으로 법적 효력을 가지는 것이 아니므로 본문의 기술방식('~하여야 한다' 등)에도 불구하고 민원인 여러분께서 반드시 준수하셔야 하는 사항이 아님을 알려드립니다. 또한, 본 안내서는 2023년 11월 현재의 과학적·기술적 사실 및 유효한 법규를 토대로 작성되었으므로 이후 최신 개정 법규 내용 및 구체적인 사실관계 등에 따라 달리 적용될 수 있음을 알려드립니다.

※ "민원인 안내서"란 대내외적으로 법령 또는 고시·훈령·예규 등을 알기 쉽게 풀어서 설명하거나 특정한 사안에 대하여 식품의약품안전처의 입장을 기술하는 것(식품의약품안전처 지침서등의 관리에 관한 규정 제2조)

※ 본 안내서에 대한 의견이나 문의사항이 있을 경우 식품의약품안전평가원 의약품심사부 중앙항생약품과로 문의하시기 바랍니다.

전화번호 : 043-719-3067, 3078

팩스번호 : 043-719-3050

# 목 차

I. 서론 .....	1
II. 약동학 .....	1
III. 바이오마커 .....	2
IV. 탐색임상시험 .....	3
1. 세포독성 물질 .....	3
1.1. 1상 단독요법 용량 및 용법 결정 시험 .....	4
1.1.1. 주요 목적 .....	4
1.1.2. 시험대상자 선정 .....	4
1.1.3. 투여 경로 및 용법 .....	5
1.1.4. 용량 증가 .....	5
1.1.5. 안전성 평가 .....	5
1.2. 2상 단독요법 치료적 탐색임상시험 .....	5
1.2.1. 목적과 설계 .....	6
1.2.2. 시험대상자 선정 .....	6
1.2.3. 용량 및 용법 .....	6
1.2.4. 항암 활성 평가 .....	7
2. 비세포독성 물질 .....	7
2.1. 1상 단독요법 용량 및 용법 결정 시험 .....	8
2.1.1. 주요 목적 .....	9
2.1.2. 시험대상자 선정 .....	9
2.1.3. 용량 증가 .....	9
2.1.4. 안전성 평가 .....	10
2.2. 2상 단독요법 치료적 탐색임상시험 .....	10
2.2.1. 시험 설계 및 항암 활성 측정 .....	10
2.2.2. 시간 관련 평가변수가 포함된 탐색시험 .....	11
3. 면역조절 화합물 및 단클론항체 .....	12
3.1. 단클론항체 .....	12
3.2. 암백신을 포함한 면역조절 화합물 .....	13

<b>4. 병용요법</b>	<b>15</b>
4.1. 전통적인 세포독성 물질의 병용	15
4.2. 비세포독성 물질을 포함한 병용	16
4.2.1. 병용요법에서 투여량 결정을 위한 독성 및 내약성 평가	17
<b>V. 3상, 치료적 확증임상시험</b>	<b>18</b>
<b>1. 설계</b>	<b>19</b>
1.1. 환자집단	19
1.2. 대조요법	20
1.2.1. 단독요법 및 병용요법	21
1.3. 교차	21
1.4. 무작위배정 및 맹검	22
1.5. 평가변수	22
1.5.1. 1차 평가변수	22
1.5.2. 이차평가변수 및 탐색적 분석	24
1.6. 완치목적으로 시행되는 치료	24
1.6.1. 독성 감소 또는 유사한 정도가 예상되는 경우	25
1.6.2. 독성 증가가 예상되는 경우	26
1.6.3. 독성 증가가 매우 클 것으로 예상되는 경우	26
1.7. 장기 질병조절 목적의 치료	26
1.7.1. 감소 또는 유사한 정도의 독성이 예상되는 경우	26
1.7.2. 독성 증가가 예상되는 경우	26
1.7.3. 독성 증가가 매우 클 것으로 예상되는 경우	27
1.8. 고식적 치료	27
<b>2. 특별 고려사항</b>	<b>28</b>
2.1. 수술 후 보조요법 및 선행요법	28
2.2. 약물 내성 조절제, 방사선/화학요법 감작제 및 화학요법보호제	28
2.3. 종양 예방	29

<b>3. 방법론적 고려사항</b>	<b>29</b>
3.1. 적응적 설계(adaptive design)	29
3.2. 중간분석	30
3.3. 사건까지의 시간 분석 및 반응과 진행 평가	30
3.4. 비열등시험	31
3.5. 치료 결과에 따른 환자군 간의 분석	31
3.6. 작은 인구집단 및 매우 드문 종양에 대한 시험	32
3.7. 외부 대조군의 사용	32
<b>4. 특수 인구군</b>	<b>32</b>
4.1. 고령자 및 허약 환자군	32
4.2. 성별	33
4.3. 장기 기능이 저하된 환자	33
<b>5. 안전성</b>	<b>33</b>
5.1. 종양학에서 안전성, 기본개념 및 평가원칙	33
5.1.1. 기본개념	34
5.1.2. 종양학에서 안전성	34
5.1.3. 기본 평가 원칙	35
5.2. 안전성 자료 수집, 분석 및 보고	35
<b>VI. 용어 정의 및 약어</b>	<b>37</b>
 부록 1. 무진행 생존(PFS) 또는 무질병 생존(DFS) 평가변수 고려사항	 40
부록 2. 항암제 마스터 프로토콜 임상시험	49

## 제·개정 이력

연번	제·개정번호	승인일자	주요내용
1	B1-2006-2-001	2006.01	제정
2	B1-2015-2-017	2015.12.	가이드라인 명칭변경, 법적효력 문구 통일, 양식표준화, 연락처 현행화
3	안내서-0235-02	2019.09.30.	항암제 임상시험 설계 고려사항 제공, 무진행 생존 또는 무질병 생존 평가변수 고려사항 부록 추가
4	안내서-0235-03	2023.11.30.	항암제 마스터 프로토콜 임상시험 부록 추가

## I. 서론

전 세계적으로 암환자 발생률과 사망률이 더욱 빠르게 증가하고 있다. 이에 항암제를 개발하고자 하는 제약회사, 임상시험자에게 항암제 임상시험 설계 시 고려해야 할 사항을 제공하는 것을 목적으로, 2006년 1월 ‘항암제 임상시험 평가지침’을 제정하였다. 이후 다양한 기전을 가진 최신 항암제들이 빠르게 개발되는 국제적 흐름에 맞춘 항암제 임상시험 가이드라인 개정이 요구되었으며, 2015년 12월 비세포독성 물질, 치료 목적 및 독성에 근거하여 3상 치료적 확증 임상시험 구성, 안전성 평가, 약동학, 바이오마커 등 추가, 세포독성 물질의 탐색 임상시험에 대한 내용 간략화 등을 포함하여 전반적인 사항을 개정하였다. 이후 추가된 부록에서는 임상시험 평가변수로 무진행 생존(PFS) 또는 무질병 생존(DFS)을 사용하는 것에 대한 고려사항과 최신 유형의 임상시험 설계로서 항암제 마스터 프로토콜에 대한 안내 및 고려사항 등을 제공한다.

## II. 약동학

일반적으로 다른 의약품과 마찬가지로 항암제에도 동일한 권장사항이 적용되며, 이용 가능한 임상 약리 가이드라인을 참고한다.

과거에는 항암제의 인간 human mass-balance 시험(혈장 및 배설물에서 방사성 표지된 용량을 조사하는 *in vivo* 시험)이 다른 의약품과 동일한 정도로 수행되지 않았다. 약물 간 상호작용 평가를 포함하여 시험약의 임상 약리학의 이해를 위해 이러한 시험에서 얻어지는 정보가 매우 중요하기 때문에 mass-balance 시험을 권장한다.

3상 임상시험 이전에 음식-약물 간 상호작용 시험이 수행되어야 한다. 식이 또는 공복 상태의 투여에 대한 결과 등 식이 및/또는 공복 상태의 투여에 대한 합리적 이유가 제시되어야 한다.

약물 간 상호작용 가능성이 평가되어야 한다. *In vitro* 자료에서 항암제가 중요한 약물 간 상호작용을 일으키거나 또는 상호작용을 받는 경우, 가능한 한 *in vivo*에서도 상호작용을 조사해야 한다.

약물 소실 기전 및 약리 활성을 가진 잠재적인 대사물질의 생성/소실 기전 등에 대한 사전 정보에 근거하여 장기 기능 장애 환자에 대한 시험을 선정해야 한다. 간 장애 환자에서의 시험이 필요하고, 대상 환자군에서 간 전이가 일반적인 경우, 첫 번째 단계로 간에 암세포가 전이된 환자에 대한 시험이 필요하다. 좀 더 진행된 간 질환에 대한 시험이 필요한지 여부는 사례별로 결정한다. 치료 중 악성 복수 또는 흉막액과 같은 다른 내강(third space)에서의 상태가 관찰된다면 이러한 환자들에 대한 약동학을 포함한 탐색적 시험이 권장된다.



집단 약동학(population PK) 분석을 통한 내적 요인에 대한 영향 평가를 권고한다. 혈장 농도 자료는 가능한 한 많은 임상시험으로부터 얻어야 한다. 희박한(환자 당 샘플이 거의 없음) 및 풍부한 자료(전체 혈장 농도-시간 프로파일)를 모두 사용할 수 있다. 시험에서 공변량 요소로 나이, 체중, 성별, 신장 기능, 빌리루빈, 간 효소, 유전자형, 가용성 수용체/리간드, 종양 부담(tumor burden), 염증 지표 등이 포함될 수 있다.

유효성 및 안전성과 관련된 약동학/약력학(pharmacokinetics/pharmacodynamics, PK/PD) 분석을 위한 PK 및 PD(바이오마커 및 임상적 지표) 샘플링이 권장된다. 이 정보는 약물의 노출-반응(exposure-response) 관계를 이해하는데 도움이 되며, 과도한 독성이 나타나거나 치료 효과가 없을 수 있는 환자들에 대한 치료 전략을 합리적으로 선택하는데 도움이 될 수 있다. PK/PD 정보를 제공하고, 3상 용량 선택을 뒷받침하기 위하여 2상 무작위배정 임상시험에서 노출-유효성 및 노출-안전성 분석/모델링이 권장된다(Ⅳ. 탐색임상시험 2. 비세포독성 물질, 3. 면역조절 화합물 및 단클론항체 참조). 궁극적으로 약물의 PK/PD를 완전히 규명하고 요약하기 위해서 모든 개발 단계에서 얻어진 PK 및 PD 자료에 대한 종합적인 분석이 권장된다. 수집된 모든 자료를 효율적으로 활용하기 위해서는 PD 자료(예, 연속 평가변수인 종양 축소)에 대한 종적인 PK/PD 분석이 권장된다. PK/PD 관계를 확인하기 위한 검정력과 공변량 효과에 대하여 시험 설계 평가에 기반한 시뮬레이션이 권장된다. 정보적 중도절단(informative censoring)으로 이어지는 높은 탈락률로 인해 누락된 자료의 처리는 종적인 분석 및 민감도 분석에서 매우 중요하다. 즉, 종양 축소에 대하여 초기 시점의 자료 사용을 고려해야 한다.

### Ⅲ. 바이오마커

유익성-위험성을 최적화하기 위해서는 적절한 대상 환자군(target population)을 확인하는 것이 필수적이다. 임상 약물 개발의 모든 단계에서 바이오마커를 효과적으로 사용하여 이를 확인할 수 있다. 바이오마커는 목적에 따라 통상적인 생물학적 절차, 병리학적 절차 또는 특정 치료 개입에 따른 약리학적 반응 등을 객관적으로 측정하고 평가할 수 있어야 한다. 다양한 서로 다른 진단법으로 적절한 바이오마커를 확인하고 측정할 수 있다(예, 전사발현 프로파일링[expression profiling of transcripts], 차등 항원 발현[differential antigen expression], 유전적 진단법[genetic diagnostics], 차세대 염기서열[next generation sequencing] 등).

약리학적 분류와 관계없이 오늘날 새로운 분자약물의 임상 개발은 중개연구(translational research)로부터 시작하는 것으로 여겨진다. 이는 대부분의 경우 시험할 가설 및 바이오마커 후보가 있음을 의미한다. 바이오마커의 활용은 바이오마커 상태에 따른 임상시험대상자의 전향적 층화, 생물학적 유효 용량 결정, 기전 또는 개념의 초기 증명, 독성 평가, 질병의 자연적 과정의 지표 등과 같이 광범위하다. 대상 환자군을 확인하고, PK 및 PD 변동성을 설명하기 위한 노력이 필수적일지라도 약물 개발 프로그램(기술적 및 임상적 검증) 계획 시 그 결과를 확인해야

하는 필요성을 간과해서는 안 된다. 시험대상자 층화의 경우, 비임상 및 임상 개발 단계 초기에 바이오마커 선정의 근거가 확립되었다면 음성 환자군(negative population)에 대한 확증이 요구되지 않을 수 있으며, 관심 있는 바이오마커가 있는 환자군에서 이러한 시험을 수행할 수 있다.

초기 임상시험에서는 바이오마커를 탐색적 목적으로 시험하지만 기술적/정량적인 신뢰성은 필수적으로 보장되어야 한다. 혈청 바이오마커 또는 기타 생물학적 검체는 유용할 수 있지만 따로 정당화되지 않는 한 종양 검체가 바이오마커 평가의 필수적인 부분을 차지할 것으로 예상된다. 그러나 단일 생검은 종양의 이질성으로 인해 대표성을 띄지 않을 수 있다. 비임상시험에서 정상 조직에서 관찰된 변화와 종양 특징 간의 상관관계가 있는 것으로 나타나는 경우, 초기 임상시험에서 정상 조직 샘플이 사용될 수 있다. 초기 약물 개발 단계에서는 기능적 영상(functional imaging)의 역할이 잘 정립된 것으로 간주되진 않지만 이를 사용하는 것이 권장된다.

바이오마커 진단법의 개발은 초기 임상 개발 단계부터 고려해야 하며, 기술의 임상 적용을 극대화해야 한다. 진단 분석법은 적절하게 품목 허가 시 제공되어야 한다.

확증임상시험에 사용하기 위한(예, 유효성 평가로 활용) 바이오마커는 원칙적으로 잘 설계된 전향적 임상시험에서 주의 깊고 엄격하게 검증되어야 하며, 체계적인 평가가 이루어지는 것이 바람직하다.

## IV. 탐색임상시험

탐색임상시험은 약물 개발에 필수적이다. 1/2상 탐색임상시험과 3상 확증임상시험 간의 구분은 본 가이드라인에 따른 것이다. 그러나 탐색적 목적이 3상 임상시험에서의 중요한 부분이 될 수 없다는 것을 의미하는 것은 아니다. 마찬가지로 가설 수립, 검증 및 확증 또한 2상 임상시험에서 수행할 수 있다.

0상 임상시험으로 불리는 마이크로 용량의 탐색임상시험은 조직 분포 및 수용체 결합 등과 같은 특정 상황에서 유용할 수 있다. 예를 들어, 약물이 중추신경계 등과 같은 보호부위를 침투할 가능성이 있는지 여부를 초기에 확인하는 것이 중요하다고 판단되는 경우 또는 낮은 약물 농도에서 약리 활성에 대한 초기 자료를 얻을 수 있는 가능성이 있는 경우가 있다.

### 1. 세포독성 물질

여기에서는 전통적인 세포독성 제제(즉, 단기간 노출 이후 DNA 복제, 유사분열 등에 대한 간섭을 통해 비가역적인 치명적 세포 손상을 초래하는 물질)를 다룬다. 이러한 물질에서 독성 및 종양 반응은 적절한 활성 지표로 간주된다.

개념적으로 본 항은 독소 함유 단클론항체(monoclonal antibody coupled toxin product)와 같은 좀 더 표적화된 세포독성 물질도 관련이 있다. 그러나 이러한 경우에는 종양 항원 발현 및 전구 물질 활성화 경로 또한 고려해야 한다.

비세포독성 물질에 대해서는 활성/저항성의 전제조건을 규명하고, 저항성에 대한 표지자를 확인하기 위한 비임상 및 임상시험이 권장된다.

### 1.1. 1상 단독요법 용량 및 용법 결정 시험

1상 임상시험 설계 시 기본 개념은 용량 결정을 위하여 독성이 수용 가능한 평가변수이어야 한다는 것이다. 1상 임상시험의 주요 목적은 용량제한독성 및 다음 단계의 임상시험을 위한 적정 용량을 결정하는 것이다. 비록 환자들 간의 PK 변동성 차이가 크에도 불구하고 일반적으로 이러한 목적을 달성하는 것이 어렵지 않지만 이후 임상시험을 위한 합리적인 투여 용량 결정이 복잡한 경우도 종종 있다.

초기 용량은 고정 용량(flat doses) 또는 체표면적(body surface area, BSA) 당 용량을 사용할 수 있다. BSA 당 용량이 환자 간 약물 노출 변동성을 줄일 수 있다는 것에 대한 과학적 뒷받침은 약하며, BSA가 높은 환자에서는 약물의 과다 노출, BSA가 낮은 환자에서는 과소 노출을 초래할 수 있다. 실제 PK 자료를 사용한 모델링과 시뮬레이션을 통해 약물 노출의 변동성에 BSA 또는 체중이 중요한지 탐색될 수 있을 것으로 예상된다.

가능한 경우, PD 평가변수 사용이 용량 선택에 도움이 될 수 있다.

#### 1.1.1. 주요 목적

- 정해진 투여 일정과 투여 방법에 따라 최대내약용량(maximal tolerated dose, MTD), 용량제한 독성(dose limiting toxicity, DLT) 및 2상 권장용량(recommended phase 2 dose, RP2D)을 확인한다.
- 용량 및 용법에 따른 빈번한 부작용과 표적 장기 독성을 규명한다. 또한 중증도, 지속기간, 가역성도 결정한다.
- 용량 및 시간 의존성을 포함한 PK의 초기 규명. 적절하게 목표 효과 및 이상반응과 관련한 PK/PD와 투여 경로 별 약물 노출

#### 1.1.2. 시험대상자 선정

일반적으로 1상 임상시험은 확립된 치료 대안이 없는 암 환자에게 수행된다.

### 1.1.3. 투여 경로 및 용법

인간에게 최초 투여 시 타당한 비임상 자료를 근거하여 투여 경로, 투여 용법을 선정한다. 대부분의 경우 정맥투여는 생체이용률에 따른 변화를 배제할 수 있기 때문에 인간에게 최초 투여 시 정맥투여를 권장한다.

투여 용법 결정 시 같은 계열 약물에서 얻어진 경험이 도움이 될 수 있다. 주기 의존성(cycle dependency) 및 *ex vivo* 종양/정상 조직의 세포독성 비율(ratio tumor/normal tissue cytotoxicity)과 관련된 비임상 자료 등을 활용할 수 있다.

### 1.1.4. 용량 증가

독성이 매우 약하거나 또는 유의한 독성이 나타나지 않는 경우, 활성이 없는 용량에 노출되는 환자 수를 줄이기 위해 동일 환자에서의 용량 증가가 적절할 수 있다. 비임상 자료에서 축적 독성이 없다고 증명된 경우에 한하여 DLT 평가기간 종료 후에 동일 환자 내 용량 증가가 허용될 수 있다.

수용할 만한 독성이 발생하였다면 환자가 회복될 때까지 기다린 후에 재투여가 가능하며, 이때는 동일 용량으로 최소 2주기 이상 투여하는 것이 바람직하다.

### 1.1.5. 안전성 평가

이상반응을 평가하기 위한 최소한의 요건에는 증상평가, 신체검사, 심전도, 혈액 및 뇨의 검사실 검사, 적절한 영상학적 평가 등이 포함된다. 비임상 자료를 근거로 추가적으로 필요한 검사를 수행해야 한다. 비임상시험 또는 유사 계열 약물에서 QTc에 관한 특별한 징후가 없다면 정교한 QTc 시험이 기대되지는 않지만 정기적인 모니터링의 부분으로서 심전도를 포함하는 것이 권고된다. 투여 부위의 국소독성에 대해 별도로 기록한다. 독성은 일반적으로 인정되는 시스템(예, 미국 국립암연구소[National Cancer Institute]의 이상반응에 대한 일반용어기준[Common Terminology Criteria for Adverse Events, CTCAE])에 따라 등급화 한다.

독성에 영향을 주는 인자들(장기 기능 장애, 병용요법)도 적절하게 평가되어야 하고, 이러한 인자들은 이후 2상/3상 임상시험에서 더 상세히 설명되어야 한다.

## 1.2. 2상 단독요법 치료적 탐색임상시험

2상 임상시험은 다양한 암종 또는 선택된 암종에 대한 단독요법의 항암 활성을 평가하거나,

다제 병용요법 또는 다학적(multimodality) 치료의 적용 가능성이나 항암 활성을 평가할 수 있다.

본 항에서는 이후 치료적 확증임상시험으로 나아갈 물질인지 파악하기 위해 특정 암종 환자에 대한 단독요법의 항종양 효과를 평가하는 것이 주요 목적인 임상시험에 대해 초점을 맞춘다.

### 1.2.1. 목적과 설계

2상 임상시험은 다양한 시험 설계를 사용할 수 있다. 초기 임상시험은 치료 활성 및 내약성에 대한 최초 근거를 제시하도록 설계되어야 한다. 특히 1개 확증적 중추시험(confirmatory pivotal trial)만을 수행할 것으로 예상되는 경우에는 무작위 대조군을 포함시킬 것을 권장한다 (V. 3상, 치료적 확증임상시험 1.2. 대조요법 참조).

임상시험의 목적은 다음과 같다:

- 표적 종양 유형에서 반응 가능성 평가(및 다른 관련 유효성 평가) 및 향후 임상시험(암의 초기 단계에서의 시험, 병용요법, 표준 치료와의 비교)에 대한 필요성 여부 결정
- 적절한 경우, 약물유전학(pharmacogenomics) 및 바이오마커 특성 평가
- 안전성 및 항암 활성과 관련하여 용량 및 용법 의존성을 보다 명확히 규명
- 약물 부작용을 보다 명확히 규명
- PK 및 PK/PD를 보다 명확히 규명(II. 약동학 참조)
- 최적의 투여 경로 결정

### 1.2.2. 시험대상자 선정

국제적으로 합의된 진단 기준에 따라 표적 질환, 병기 및 이전 치료 병력에 대하여 명확하게 규정되어야 한다.

안전성과 항암 활성에 대한 타당하고 과학적 근거가 있는 경우, 대안적 치료요법이 가능한 환자에서도 시험을 수행할 수 있다. 여기에는 수술이 예정된 치료받지 않은 환자에서의 선행요법이 포함될 수 있으며, 이 경우 수술 지연이 환자에게 불리하지 않아야 한다. 환자의 안전성과 치료 이익은 항상 보장되어야 하며, 상세하고 타당한 설명이 시험계획서에 명시되어야 한다. 이러한 경우, 기능적 영상과 같은 항종양 활성을 예민하게 측정할 수 있는 측정법을 사용한다.

### 1.2.3. 용량 및 용법

용량 및 용법은 명확하게 정의되어야 한다. 특별한 주의(환자에게 수액 공급, 차광 및 온도 주의 등)가 필요한 의약품 투여와 시험기간 중 병용 금기 약물에 대한 상세한 내용이 기술되어야 한다.

- 관찰되는 독성의 중증도에 따른 용량 감량에 대한 지침이 제공되어야 한다.
- 적절한 경우, 독성이 낮은 경우 용량 증량에 대한 지침을 고려할 수 있다.
- 고위험 환자(예, 표적 장기 독성의 위험이 높거나 시험약 성분의 대사 또는 배설 기능에 문제가 있는 환자)에 대한 별도의 고려가 있어야 한다.
- 누적 독성에 대한 모든 증거는 기록해야 하며, 총 약물 투여 용량의 함수로 평가되어야 한다. 이는 표적 장기 또는 기능에 따라 특별히 시험되어야 한다.

#### 1.2.4. 항암 활성 평가

객관적 반응률(objective response rate, ORR)은 국제 기준(예, RECIST, Volumetric RECIST 또는 WHO 기준)에 따라 기록되어야 한다. 특정 상황에서는 이러한 기준을 변경할 수 있으나 타당성이 설명되어야 한다.

ORR을 평가할 때, 배정된 대로 분석(intention to treat, ITT) 원칙을 준수해야 한다. 단일군 시험에서는 1차 유효성 평가변수로 계획서 순응(per-protocol) 분석 집단의 ORR이 보고될 수 있다. 시험 목적에 따라 종양 반응의 독립적 외부 검토가 권장된다.

일반적으로 반응 지속기간(duration of response), 종양 진행까지 시간(time to tumor progression, TTP)/무진행 생존(progression-free survival, PFS), 확증된 객관적 반응률(confirmed ORR), 전체생존(overall survival, OS) 관련 자료는 보고되어야 한다. 종양 바이오마커 및 활성에 대한 다른 약력학적 측정치의 사용이 권장된다.

악성 혈액종양에서는 많은 경우 질병 특이적 반응 기준이 불가피하지만 일부 질환에서는 그 기준에 대한 완전한 합의가 아직 이뤄지지 않았다. 따라서 이러한 문제에 대해 국제실무단(international working group)이 정한 기준을 따르는 것이 중요하다. 특히 덜 보수적인(less conservative) 질병 특이적 반응 기준이 사용된다면, 의뢰자는 약물 개발 측면에 중점을 둔 타당성을 제시해야 할 것으로 예상된다.

2상 무작위배정 임상시험 시 베이스라인에서 증상이 있는 환자의 경우, 증상 조절에 대한 평가가 권장된다.

## 2. 비세포독성 물질

비세포독성 물질이란 항호르몬 제제, antisense 제제, 신호전달 억제제, 혈관생성 억제제, 세포주기 억제제, 면역 조절제 등에 이르는 매우 다양한 기전을 가지는 약물군을 일컫는다. 이러한 비세포독성 물질을 사용한 임상시험 설계에서 고려할 점은 독성이 용량 및 용법 결정 시험의

적절한 평가변수가 아닐 수 있으며, ORR이 항종양 활성의 적절한 측정방법이 아닐 수 있다.

세포독성 항암화학요법과 비교하여 비세포독성 물질은 일반적으로 지속적으로 투여하며, 독성 프로파일이 다르기 때문에 DLT는 여러 번의 투여 주기 이후에 처음 발생할 수 있다. 내약성과 독성으로 RP2D를 결정하는 경우 이러한 특성이 매우 중요하며, DLT 및 MTD의 정의와 관련된 대체 전략이 필요할 수 있다.

이러한 이유로 비세포독성 물질의 초기 약물 개발 단계는 더 복잡하며 비임상시험에서 확인된 개별 물질의 약리학적 특성에 따라 맞춤화(tailored) 되어야 한다. 예를 들어, 용량 및 용법을 결정하기 위하여 바이오마커 평가에 근거한 항종양 활성 측정이 초기에 필요할 수 있기 때문에 전통적인 세포독성 물질과 같은 1상 및 2상 임상시험 간의 엄격한 구분이 덜 적절할 수 있다.

그러나 세포독성 물질과 관련하여 논의된 대부분의 사항들(시험대상자 선정에 대한 제한사항, 투여 경로에 대한 권고사항, 독성 및 항종양 활성 평가 등)은 비세포독성 물질에도 적용되므로 여기서는 이에 대해 더 기술하지 않는다.

## 2.1. 1상 단독요법 용량 및 용법 결정 임상시험

비임상 자료와 이용 가능한 경우의 건강한 자원자에서의 자료를 사용하여 환자 대상 임상시험을 설계해야 한다(예, 선정기준 및 개시용량 관련). 세포독성 물질에서의 가이드라인과 마찬가지로 이용 가능한 확립된 치료요법이 있는 환자는 일반적으로 제외되어야 한다. 전통적인 세포독성 물질에 불응성인 경우 명백하게 관련이 없는 물질에 대해 일부 교차 내성이 나타날 수 있다. 이는 용량/농도-효과 관계를 정의하는데 영향을 미칠 수 있다. 기회시험(window of opportunity studies) 등을 포함하여 이들 임상시험의 분석 민감도를 높이기 위해 민감성이 높고 윤리적으로 수용 가능한 모든 방법을 수행할 것을 권장한다. 필요한 경우, 약물 활성과 관련이 있다고 추정되는 표적의 발현을 측정하는 것이 이에 해당된다.

PD 측정에는 생화학적 측정(비임상시험에서 검증된 수용체 결합, 효소 억제, 신호전달 하위 단계 등), 기능적 영상검사, 단백질체학(proteomics), 면역학적 측정(항체 또는 T세포 반응) 등이 해당될 수 있다. 집단 PK/PD 시험도 권장된다. 비임상 모델에서 세포증식억제를 보이는 물질의 경우, 임상시험에서 종양 감소를 보이기 위해 장기간의 투여가 필요할 수 있다. 만약 임상시험에서 예상치 않은 조기 종양 감소가 발견되는 경우, 이는 조기 반응과 관련한 기저 작용기전을 탐색하는 추가 시험이 필요하다는 신호이다.

돌연변이된 BRAF 등과 같은 약물 활성의 단일 주요 표적이 있는 물질은 약물 개발이 보다 수월할 것으로 인정되는 반면, 다중 표적 물질의 경우 약리학적 근거가 탐색임상시험 프로그램

(적절한 치료 대상 환자군을 확인하기 위해 선정된 바이오마커 등)에 반영되는 것이 예상된다.

### 2.1.1. 주요 목적

- 내약성, 안전성, PK 그리고 가능한 경우 약물 활성의 PD 측정이 적절한 목적이다.
- 전통적인 세포독성 약물과 마찬가지로 가능한 항종양 활성을 규명하기 위해 전통적인 방법과 병행하여 종양 표지자 및 민감한 영상 기술의 사용을 권고한다. 기관 간의 변동성을 줄이기 위하여 기술(예, 기능적 영상 기술, 바이오마커 분석 등)의 표준화를 시행하는 것이 권고된다.

### 2.1.2. 시험대상자 선정

초기 임상시험은 비임상에서의 내약성 및 독성학 결과와 추정된 약리학적 특성에 근거하여 때때로 건강한 자원자를 대상으로 수행할 수도 있다.

선정기준 및 시험대상자 수는 시험 목적에 따라 결정되어야 하며, 향후 시험에 사용될 용량 및 용법에서의 PK 및 PD 변동성 또한 고려되어야 한다.

약리적으로 입증되지 않는 경우, 접근 가능한 종양(주 병변 및/또는 전이 병변)의 생검 분석은 확증임상시험의 적절한 대상 환자군을 확인하기 위한 시험에서 핵심적인 역할을 할 것으로 예상된다. 이는 중요한 사항으로 실시기관, 시험자 및 시험대상자 모집 시에 고려되어야 한다.

### 2.1.3. 용량 증가

지금까지의 사용 경험에 따르면 대부분의 물질은 종양 선택성이 없는 것으로 나타났다. 따라서 내약성과 독성은 용량 및 용법 결정 시험에서 여전히 중요한 척도이다. 그러나 권장용량을 결정하기 위해 MTD까지 증량하는 것이 적절하지 않은 경우가 있다. 이 경우, 관련 동물 모델의 PD 및 안전성 자료, 초기 및 후속 용량 코호트에서의 인간 PK/PD 자료에 근거하여 용량 증가를 할 수 있다. 작용기전 기반의 PK/PD 모델링 또한 의사결정 시 유용할 수 있다.

관련 독성을 수집하고 유용한 RP2D에 도달하기 위해 MTD 및 DLT의 개념을 사전에 어떻게 정의할 것인지에 대해 신중하게 고려해야 한다.

다수의 분자표적물질(molecularly targeted agent, MTA) 및 면역조절요법은 지속적 및/또는 장기간 투여가 필요하다. 또한 면역조절제(immunomodulator)로 인한 면역반응 등과 같은 특정 유형의 제제-특이적 독성은 종종 첫 치료 주기 이후에 나타난다. 따라서 첫 번째 주기의 급성 독성에



전형적으로 중점을 둔 세포독성 제제에 대한 표준 정의가 적용되지 않을 수 있다. 장기간의 낮은 등급의 독성은 내약성 및 의도한 용량 강도를 유지할 가능성에 영향을 미치므로 이러한 독성은 DLT와 MTD의 정의에서 언급될 필요가 있다.

MTA의 1상 임상시험에서 환자의 과반수 이상이 첫 번째 주기 이후 3-4 등급의 독성을 보이는 것으로 관찰되었다. 따라서 DLT를 길게 관찰하는 좀 더 폭넓은 DLT 정의를 고려할 필요가 있다. 첫 번째 주기의 급성 독성, 내약성에 영향을 미치는 지연 독성, 및 후기 중증 독성 간의 구분이 유용할 수 있다. 첫 주기의 이상반응(AE)에 근거한 용량 증가는 여전히 합리적일 수 있으며, 이로 인해 유효 용량 강도에 빠르게 도달할 필요성과 이후 용량 감소의 필요성 간의 균형을 이룰 수 있다. 따라서 AE는 치료 주기별로 항상 보고되어야 하며, RP2D는 발생 가능한 AE의 종합적 평가를 기반으로 결정되어야 한다.

PK 및 독성/내약성의 개인별 변동성으로 인해 환자의 약 75%가 용량 감소 없이 RP2D에 내약성을 가지는 것이 수용 가능한 것으로 고려된다.

#### 2.1.4. 안전성 평가

‘Ⅳ. 탐색임상시험 1. 세포독성 물질 1.1. 1상 단독요법 용량 및 용법 결정시험’에 논의된 일반 원칙이 적용되지만 약리학적으로 예측 가능한 AE가 더 다양하므로 이를 시험 계획 시 고려해야 한다. 예를 들어, 면역관문 억제제(check point inhibitor)는 자가 면역반응이 예측 가능하며, 항혈관 신생물질은 혈관관련 이상반응, 고혈압 및 단백뇨가 예상된다.

### 2.2. 2상 단독요법 치료적 탐색임상시험

이러한 시험을 단순하게 설명하기 위해 용량/노출 범위는 용량제한독성의 유무에 관계없이 약리학 적 활성/표적 점유율에 따라 결정된다고 가정한다. 달리 입증되지 않는 한, 적절한 대상 환자군을 찾아내기 위한 활동을 지속한다.

#### 2.2.1. 시험 설계 및 항암 활성 측정

대부분의 종양에서 부분 반응을 만족하는 자연적 종양감소가 드물기 때문에, ORR은 항종양 활성을 측정하는 더 확실한 방법이다. 따라서 탐색적 목적의 무작위배정된 대조군이 없는 시험은 해석 가능한 것으로 간주되며, 세포독성 물질에서 언급된 가이드라인을 참고한다. 이와 무관하게 무작위배정된 대조군을 시험에 포함하는 것이 권장되며, 선택된 바이오마커가 예후 예측적(prognostic) 및/또는 반응예측적(predictive) 인지 여부를 탐색하기 위해 대조군에 특별히 관심을 기울일 수 있다(V. 3상, 치료적 확증임상시험 1.2. 대조요법 참조).

하지만 TTP와 PFS는 원칙적으로 기저 종양의 성장 속도 및 항암제의 활성화에 영향을 받는다. 또한, 확인된 질병진행이 선정기준인 경우, 대부분의 환자에서 기저 종양의 성장 속도를 확인하기 어렵고, 과거 자료(historical data)는 해석이 더 어려울 것이다. 따라서 무작위배정된 대조군이 없는 시험에서의 TTP/PFS 자료의 해석은 문제가 된다. 특히 유방암에서는 TTP/PFS와 동일한 문제가 있음에도 불구하고, 임상적 유효 반응률(clinical benefit response rate, CBR; 완전 반응, 부분 반응 및 6개월 시점에 질병 진행 없음)은 잘 확립된 항종양 활성 측정법이며, 시험 간 비교를 위해 사용되기도 한다.

## 2.2.2. 시간 관련 평가변수가 포함된 탐색임상시험

주로 종양 성장 억제를 유도하는 물질에 대한 탐색임상시험에 적용 가능한 이상적인 설계는 아마도 아직 없을 것이다. 다음의 몇 가지 설계 대안이 찬반양론 속에 논의되고 있지만, 규제 관점에서는 아래와 같은 사항은 원칙적으로 허용된다. 설계와 관계없이 종양 진행이 문서화된 환자만 등록하는 것을 권장한다.

- 무작위배정, 용량 비교시험(예, 약리적 활성이 있을 것으로 여겨지는 최저 용량과 이보다 높은 용량을 비교하는 시험)에서 TTP/PFS의 차이를 보이는 경우, 절대적이지는 않지만 명백한 활성의 증거가 될 것이다.
- 정의된 기간 동안 시험약물을 투여한 후 질병이 진행되지 않은 환자를 대상으로 무작위로 치료를 중단하는 단일군 시험. 그러나 이 설계는 환자나 시험자가 수용하기 쉽지 않으며, 일부 약물의 경우 잔류효과(carry-over effect)가 있을 수 있다.
- 이전에 치료 경험이 있는 환자를 대상으로, 동일 환자 내에서의 TTP/PFS 비교는 활성 증거를 제공할 수 있다. 가장 최근 이전 치료의 TTP를 시험약 치료 시의 TTP/PFS와 비교한다. 그러나 시간에 따른 종양 성장 속도가 적어도 비슷하다는 기본 가정이 항상 성립하지 않는다는 것을 유의해야 한다. 탐색적 목적에서 이것은 중요한 문제가 되지 않는다. 이전 치료에 대한 1차 내성 및 2차 내성을 가진 환자를 모집하는 것을 권장하며, 이로 인해 적어도 어느 정도는 시험 집단이 관련성이 있음이 보장된다. 또한 기존 치료에 조기 실패한(1차 내성) 환자는 종양 성장 속도의 유동성 및 영상 기술에 따른 변동성 때문에 TTP에서 일부 역전이 보일 수 있다는 것을 유의해야 한다.

특정 적응증에 대해서는 이전에 치료받지 않은 환자도 동일 환자 내 비교(즉, 치료 없이 질병이 진행된 환자에게 진행까지 시험약 치료를 실시)가 타당할 수 있다.

- 특정 집단에서 활성을 나타내는 것으로 알려진 약물(타당한 경우 위약/최적의 지지요법(best supportive care, BSC))과 비교하는 무작위배정 2상 임상시험이 또 다른 대안이 된다. 종양 진행은 최대 종양 반응(best tumor response)과 비교하여 판정되기 때문에 종양 감소를 유도하는 약물에 비해 순수한 성장 억제 약물이 TTP/PFS에 더 유리함을 유의해야 한다. 따라서 종양 감소가 유도된 환자보다 성장 억제 약물의 치료에 실패한 환자에서 종양 진행

시의 종양 부담이 더 클 것이다.

- 더 이상 적용할 수 있는 명확한 방법이 없는 경우, 내부 대조군이 없어도 TTP/PFS 및 CBR은 2상 임상시험에서 항종양 활성 측정방법으로 허용되어야 한다. 이러한 경우, 사용된 방법론 등에 대한 체계적인 문헌 검토가 권장된다.

이론적으로 결과 변수로 ORR을 사용하는 2상 임상시험과 유사한 통계적 접근법이 적용 가능하다. 그러나 조기 종료(early escape)의 기준을 세우는 것이 더 어렵다. 사전 정의된 시점에서 무진행 환자의 비율을 합리적으로 정확히 추정하기 위하여 시험대상자 수는 충분해야 한다. 치료하지 않았을 때의 질병 진행률을 추정하는 것은 더 문제가 되며 “기대되는 활성(promising activity)”을 정의하기가 더 어렵다.

이러한 시험에서 ORR 및 종양 진행에 대한 전통적인 기준을 사용하는 것이 권고되며 독립적 검토가 권고된다. 다만, 염증성 부종으로 인한 외형적인 종양 크기 증가와 같은 “가성진행(pseudoprogression)”이 특정 약물에서는 활성의 첫 징후일 수 있다. 이전 임상시험에서 이러한 경우가 나타나면 시험계획서에 이를 반영하는 것이 허용된다. 주요 활성 측정방법으로 ORR 및 TTP를 사용하는 것이 동시에 종양/PD 표지자를 사용하는 것과 모순되는 것으로 간주되어서는 안 된다.

무작위배정 설계가 적절하다고 고려되는 경우, 건강관련 삶의 질(health-related quality of life, HRQoL) 또는 증상 조절 평가가 중요한 정보를 제공할 수 있다.

기회 시험(window of opportunity studies)의 경우 및 기능 종양 영상 및/또는 바이오마커 등과 같이 약리학적 활성에 대한 민감한 측정법이 있고, 그 측정법에 민감하다고 알려진 종양을 가진 대상 환자군이 있는 경우, 하나 또는 되도록 그 이상의 시험약 용량에 대한 위약-대조 시험이 가능할 수 있다. 민감한 측정법이 ORR과의 관련성에 대하여 충분히 검증되지 않았더라도, 규제 관점에서 탐색적 목적으로는 허용될 수 있으며, 엄밀한 용량 비교뿐만 아니라 활성이 없는 경우의 조기 종료(early escape)도 가능하게 한다. 복합변수(예, 바이오마커 또는 영상 또는 증상)의 사용 여부를 포함하여 질병 진행에 대한 기준을 계획서에 명확히 정의하는 것이 권장된다.

### 3. 면역조절 화합물 및 단클론항체(immune modulating agents and monoclonal antibodies)

본 항에서는 탐색임상시험에 대한 가이드라인을 제공하고자 하는 것이 주목적이지만, 확증임상시험에 대한 일부 관련된 측면의 가이드라인을 제공한다.

#### 3.1. 단클론항체

단클론항체는 항체의존 세포매개 세포독성(antibody-dependent cell-mediated cytotoxicity, ADCC) 및/또는 성장인자/항세포사멸 수용체 신호체계 차단 등을 통해 직접적으로 종양세포에 영향을 주거나, 종양이나 종양 지지 구조의 성장인자 표적 또는 T세포 억제 신호 차단(예, 항CTLA4) 등을 통해 간접적인 영향을 줄 수 있다.

단클론항체의 주요 활성을 규명하기 위해 *in vitro* 비임상시험을 수행해야 한다. 이러한 시험에는 다음의 관련 분석이 포함될 수 있다:

- 1) 표적 항원과의 결합: 표적의 (과)발현 여부를 확인하기 위해 종양 세포 또는 혈장을 검사해야 하며, 표적 발현과 활성 간의 관계를 조사해야 한다.
- 2) 원치 않은 표적: 종양 특이성 파악이 어려울 수 있으나, *in vitro* 에서 “원치 않은” 표적을 걸러내어 안전성 평가를 용이하게 할 수 있다.
- 3) Fab 관련 기능(예, 수용성 리간드의 중화, 수용체 활성화 또는 차단)
- 4) Fc 관련 기능(예, 항체의존 세포매개 세포독성[ADCC], 보체의존 세포독성[complement-dependent cytotoxicity, CDC], 보체 활성화[complement activation])

표적-매개성 특징(target-mediated disposition)은 단클론항체에서도 볼 수 있다. 비용량 비례 약동학적 동태(non-dosage proportional PK behavior)의 양상에 대한 적절한 특성 규명은 광범위한 종류의 표적을 가진 암환자들이 시험되는 후기 단계 시험 때까지 가능하지 않을 수 있다. 따라서 다른 유형 및 다른 병기의 종양을 대상으로 하는 다양한 임상 개발 프로그램에서 단클론항체 PK에 대한 지속적인 평가를 권장한다.

단클론항체의 제거(clearance)는 전형적으로 FcRn IgG 순환과 면역원성(항약물항체(anti-drug-antibody, ADA))등에 의해 영향을 받으며, 환자 건강상태 인자(예, 알부민, 수용성 수용체/리간드, 질병 유형 및 중증도, 종양부담 등)에 의해서도 영향을 받을 수 있다. 이러한 요인들에 대한 정보는 단클론항체의 노출 및 반응의 성격을 이해하는데 도움이 될 수 있다. 다른 임상 분야에서 단클론항체의 면역원성에 대한 경험은 분석법 선택, 활성 상실에 대한 표지 및 예측 가능한 안전성 문제 등과 관련해 고려되어야 한다.

### 3.2. 종양 백신을 포함한 면역조절 화합물

암치료 백신을 포함한 면역치료요법은 기존 악성 질환에 대한 특정 항종양 면역반응을 유도하는 것이 목적이다. 이러한 면역요법은 통상적으로 암환자의 선천적인 면역반응 뿐만 아니라 후천적인 T와 B세포 면역반응을 유도하는 것이 목적이다. 합성 펩타이드, 재조합의약품, 바이러스 유사 물질, 면역조절항체, 유전자 치료 및 세포기반 제품과 같이 사용된 약물의 성격은 매우 다양하다. 자가 항원으로부터 통상적으로 파생된 종양 항원에 대한 내성을 피하기 어렵기 때문에 암백신은 종종 사이토카인이나 toll형 수용체 작용제(toll-like receptor agonist)와 같은

약리학적 활성을 가진 보조제와 병용된다. 면역관용현상(immune tolerance)을 피하기 위한 또 다른 방법은 T세포 저해 신호를 단클론항체와 같은 것으로 차단하는 것이다. 이에 따른 T세포 활성화 및 증식은 바람직하거나 바람직하지 않은 면역 자극 효과를 유발한다(바람직한 항종양 효과 이외에 대장염, 내분비계 장애 등과 같은 면역 관련 독성 발생).

1상 임상시험에서 계획된 개시 용량 및 용법을 정당화하기 위한 비임상시험으로서 개념입증(proof-of-concept)을 위한 *in vitro* 및 *in vivo* 시험이 제시되어야 한다. 또한 사례별로 “최소 기대 생물학적 효과 수준(minimal anticipated biological effect level, MABEL)” 접근법 및 관련 물질의 비임상 및 임상 자료를 사용하여 개시 용량의 근거를 뒷받침할 수 있다.

인간 주 조직적합 복합체(major histocompatibility complex, MHC) 분자에 제시되어야 하는 인간 특이적 항원에 근거한 제품의 경우, 예측 동물 모델을 사용할 수 없는 경우가 종종 있다. 그럼에도 불구하고 가능한 경우, 동종 항원을 사용한 동물 모델이나 인간 MHC를 이식한 동물을 비임상 약리학 및 독성학 시험을 위해 고려할 수 있다. 인간 종양과 건강한 조직에서 표적 항원 발현 정도 차이에 대한 정보가 제공되어야 한다. 적절한 예측 가능 동물 모델이 없는 경우, 인간 세포를 이용한 *in vitro* 시험(예, *in vitro* T세포 점화분석(priming assay))이 개념입증을 나타내는데 적절할 수 있다.

초기 단계 임상시험의 목적은 안전성 확인 및 원하는 면역반응을 유도하는 용량과 용법을 결정하는 것이다. 일반적으로 용량결정 시험은 RP2D를 확립하는데 필요하다. 적절한 용량 및 용법을 결정하기 위해서는 항원 특이적 T세포의 유도 또는 체액성 반응의 존재 등과 같은 면역반응을 모니터링 하는 것이 중요하다. 이를 위해서는 다중 모니터링 분석법이 필요할 수 있으며, 주의 깊게 조사해야 한다. 분석 방법은 임상시험 계획서에 상세하게 기술되어야 한다.

치료 전후에 시행된 종양 생검은 표적 조직에서 면역 활성화의 범위 및 유형을 평가하는데 결정적 역할을 할 것으로 예상되며, 가능한 항종양활성에 대한 조기 표지자 역할을 할 수 있다.

높은 종양 부담을 가진 환자를 등록하여 종양 반응을 유도하는 것은 현실적으로 어려울 수 있기 때문에, 최소 또는 적은 종양 부담을 가진 환자를 포함시키는 것이 좋을 수 있다. 예로는 비소세포폐암 환자에서 종양 완전 절제 후 보조요법으로 암 면역요법을 적용한 경우 또는 항암화학요법에 반응이 있는, 수술 불가능한 비소세포폐암 환자에 적용한 경우다. 이전에 세포독성 치료를 여러 번 받지 않은, 국소적이고 측정 가능한 질병을 가진 환자에 대하여 명확하게 시험약 치료를 사용하는 임상시험 설계는 신중하게 정당화되어야 한다. 다른 제제에 대해서는, 확증임상시험을 개시하기 전에 항종양 활성화에 대한 근거가 필수적이다.

암환자들은 대개 질병이 진행된 경우 치료를 중단한다. 효과적인 면역반응 및 임상반응을 유도

하기까지 전형적인 세포독성 약물보다 더 오랜 시간이 필요할 수 있다(지연효과). 따라서 생물학적 활성이나 임상 효과가 시작되기 전에 질병이 진행될 수 있다. 진행이 느린 질병의 경우, 암 면역치료법을 중단하는 것은 적절하지 않을 수 있다. 이 경우, “느리게 진행되는 질병 (slowly progressive disease)” 및/또는 시험탈락 기준에 대한 세부적인 정의가 임상시험 계획서에 포함되어야 하며, 환자에 대한 면밀한 모니터링이 요구된다. “느리게 진행되는 질병”은 시험 중인 질병의 경과에 따라 정의되어야 한다. 확증임상시험에서 적절히 입증한 경우에는 수정된 질병 진행에 대한 정의 기준을 인정할 수 있지만 OS가 권장되는 결과 측정법이다.

자가 면역반응(세포 및 체액성) 유도 및 내성 유도와 같은 예측 가능한 독성은 임상 개발 동안 주의 깊게 모니터링 해야 한다.

## 4. 병용요법

전통적인 세포독성 약물은 감내할만한 독성 수준에서 항종양 효과를 증가시키기 위해 병용요법으로 사용되어 왔다. 독성이 부분적으로 중복되지 않고 활성/내성의 필수요소가 부분적으로 중복되지 않는 약물과 병용하는 병용 요법이 사용되어 왔다. 학계 및 규제기관이 이러한 접근법을 허용했지만, 병용요법이 단일 약물들의 순차적 사용보다 더 나은 장기적 결과를 낼 수 있을 지 여부는 알려져 있지 않다.

### 4.1. 전통적인 세포독성 물질의 병용

이용 가능한 대체 요법이 있는 환자들을 대상으로 할 때에는 병용요법을 구성하는 각 개별 약물의 활성이 입증되어야 한다.

탐색적 단계는 병용요법의 MTD와 RP2D 결정과 ORR 및 PFS/TTP를 이용한 항암 활성의 예비 평가를 포함한다. 새로운 병용요법의 항종양 효과는 가정에 근거하지만, 병용요법의 독성은 종종 개별 약물의 독성에 근거하여 예상할 수 있다. 만약 PK 상호작용을 배제할 수 있고, 용량-반응/독성 양상이 알려지지 않았다면, 용량 결정 시험은 각 약물의 단독요법 권장용량의 1/2 용량에서 시작할 수 있다. 또한 한 약물의 단독요법 최대 권장용량과 다른 약물의 단독요법의 권장용량을 줄여서(50% 미만) 용량결정시험을 시작할 수 있다. 약물의 투여 순서가 잠재적인 PK 상호작용과 항암 활성의 관점에서 중요할 수 있으므로, 시험 설계에서 투여 순서를 고려해야 한다.

유익성-위험성을 최적화하는 병용요법의 약물 간 용량 강도(dose intensity)의 균형을 이루는 획일적 방법은 존재하지 않는다. 따라서 단독요법 항암 활성이 가장 높은 약물에 용량 강도(dose intensity)의 우선권을 부여하는 것이 허용된다.

만약 병용요법에서 약물 중 하나가 단독요법 치료로 인정된 경우, 단독요법과 병용요법을 비교하는 무작위배정 2상 임상시험이 효과적일 수 있다. 확증임상시험을 위해서는 최상의 이용 가능한 근거-기반 대조요법과의 비교가 예상된다.

## 4.2. 비세포독성 물질을 포함한 병용

만약 상반되는 강력한 생물학적/약리학적 논란이 없는 경우, 비세포독성 약물과 병용된 항암화학요법은 통상적으로 “최선의 이용 가능한 요법” 이어야 한다. 항암화학요법의 용량 강도/전신노출이 변경되지 않는다면, 모든 환자가 적절한 치료를 받을 것이라고 추정할 수 있다. 따라서 이러한 관점에서 시험대상자의 선정기준을 제한할 필요는 없다.

이전의 비임상 및 임상 경험에서 PD 지표 등이 항암 활성화에 정보를 제공하는 경우, 이는 탐색 임상시험 계획의 일부로 포함되어야 한다. 비세포독성 약물의 표적 발현이 세포독성 약물에 의해 영향을 받는지에 대한 시험이 포함될 수 있으며, 그 반대의 경우에 대한 조사도 포함될 수 있다.

비임상 모델에서 부가 활성화(add-on activity)의 예측 가능성에 대한 현황을 고려할 때, 시험요법과 단독 화학요법을 비교하는 무작위배정 2상 임상시험이 권장된다. 이러한 임상시험에서, 전통적인 항종양효과 지표(ORR 및 TTP)가 종양 표지자 및 종양 대사활성 등의 민감한 측정 등과 함께 적절히 추가되는 것이 권고된다.

항암화학요법에 대한 비세포독성 약물의 부가 활성화가 임상시험에서 입증되었을 때, 일반적으로 새로운 적응증에 대하여 임상시험을 수행할 경우 추가적으로 무작위배정 2상 임상시험을 실시할 필요가 없어진다. 하지만 표적 발현의 중요성과 표적 억제가 악성 종양들 간에 다를 수 있다는 것을 감안해야 한다.

비세포독성 약물에 대한 표적의 발현이 항암화학요법에 따라 서로 다르게 나타날 수 있다면, 부가 시험(add-on study)을 실시하기 전에 새로운 항암화학요법에 대한 표적 발현을 시험하는 것이 권장된다.

부가 효과(add-on effects)에 대한 기전과 전제조건을 이해하기 위한 시험은 권장되며, 이는 향후 시험에서 대상 환자군의 특성을 구체화할 수 있다.

일부 비세포독성 약물의 경우, 병용요법은 항암 활성을 최적화하기 위해 필요할 뿐만 아니라, 실제로 활성을 얻기 위해서 필요할 수 있다. 이러한 약물의 경우, 예를 들면, 단독요법에서의 표적 포화 및 병용요법에 대한 비임상적 독성이 적절한 시작 용량/용법을 결정하는 데 사용될 수 있다.

그렇지 않을 경우, 단독요법과 마찬가지로 용량/용법의 탐색 및 치료적 탐색임상시험이 필수적으로 진행될 수 있다.

만약 강력한 생물학적 및/또는 약리학적 비임상 및 초기 개념-입증(early proof-of-principle) 임상자료에 의해 뒷받침 되는 경우, 두 개의 새로운 약물이 공동 개발 프로그램에서 병용될 수 있다.

다음의 세 개의 시나리오를 예상할 수 있다:

단일증강(uni-enhancement)이란 한 병용요법에서 항암 활성 자체가 없거나 최소한인 약물 B가 상대 약물 A의 항암 활성을 향상시키는 경우(예, 내성 발생의 억제 등을 통해)를 말하며, 약물 B의 이러한 기여는 적절한 비임상 모델로부터 얻은 자료에 의해 확립되어야 한다. 대조요법과 비교를 위해서는 2상 임상시험이 권장되는 반면, 약물 B의 단독요법에 대한 2상 임상시험 자료는 불필요할 수 있다. 적절한 2상 임상시험 설계는 AB 대 A 대 대조치료의 세 개 군으로 하는 무작위배정시험이 될 수 있다.

공동증강(co-enhancement)이란 병용요법의 각 약물들이 그 자체로는 약간의 항종양 활성을 갖지만 병용요법 시에 항암 활성이 상당히 향상되는 경우 고려된다. 2상 임상시험에서 새로운 병용요법은 효과적인 용량의 단독요법과 적합한 대조요법과 비교되어야 한다: AB 대 A 대 B 대 대조치료. 2상 임상시험의 결과를 기반으로 한 개 또는 양 쪽의 단독요법은 3상 임상시험에서는 불필요할 수 있다. 한쪽 약물(B)의 단독요법이 3상 임상시험의 일부인 경우(A+B 대 B 대 대조), 동일한 단독요법이 2상 임상시험에 포함될 필요는 없다(A+B 대 A 대 대조치료).

합성치사(synthetic lethality)란 병용요법의 각 약물들이 그 자체로는 종양 활성이 없거나 최소이지만, 병용요법으로서 잠재적 항암 활성을 보여주는 경우를 말한다. 만약 비임상 및 임상시험에서 해당 병용요법의 용량/노출 정도를 상당히 상회하는 수준에서 “비활성”을 나타내며 병용요법에 명백한 활성을 보이는 경우, 2상 및 3상 임상시험에서 각 약물의 기여도 확인은 불필요할 수 있다.

동일한 표적이 암종마다 다른 영향력을 가질 수 있으므로, 같은 병용요법도 새로운 적응증에 대해서는 새로운 시험을 통해 입증될 필요가 있다.

#### 4.2.1. 병용요법에서 투여량 결정을 위한 독성 및 내약성 평가

약물의 종류와 상관없이 용량 결정을 위한 적절하고 유용한 약동학적 1차 평가변수가 없는 경우, 용량은 주로 독성 및 내약성에 의해 결정되며, 세포독성 물질을 포함한 병용요법의 경우, IV.



탐색임상시험 1.1. 세포독성 물질 1상 단독요법 용량 및 용법 결정 시험의 일부 내용과 비세포 독성 물질을 포함한 병용요법의 경우, IV. 탐색임상시험 2.1. 비세포독성 물질 단독요법 용량 및 용법 결정 시험의 가이드라인이 적용된다. 예를 들어 장기간의 치료는 용량 제한 및 후기 이상반응을 파악하는데 필요할 수 있다.

상기에서 논의한 바와 같이, 치료요법의 일부를 구성하는 개별 약물의 최적 용량 강도는 효과 또는 안전성의 관점에서 경험적으로 확인하는 것은 거의 불가능하다.

내약 가능한 요법을 파악하는 것과는 별도로 목적에는 관찰된 독성과 관련해 용량 감소를 고려하기 위해 관찰된 이상반응을 초래한 약물에 대한 확인도 포함시켜야 한다. 단독요법으로 사용된 약물의 독성 프로파일이 일부 가이드라인을 줄 수는 있지만, 계열 경험(class experience), 작용기전 등도 또한 고려해야 한다.

## V. 3상 치료적 확증임상시험

3상 임상시험은 실제 임상에서 적용될 수 있도록 정의된 대상 환자군에서(보조적 치료를 포함한) 시험약물의 유익성과 위험성을 규명하기 위한 목적으로 계획되어야 한다. 이러한 임상시험은 기본적으로 무작위배정 대조군 비교시험이어야 한다.

본 항 1.6.-1.7.에서는 치료의 목적에 따라 완치, 장기 질병조절, 고식적 치료에 대한 논의로 구성되었다.

완치 목적이 아닌 의도로 치료가 시행된 일부 종양의 경우, 여전히 실제 임상 현장에서 잘 확립된 약물요법이 존재하며 이 요법들은 항종양 활성에 있어 주요 차이점을 보여주고 있다. 이는 임상에서 치료요법 선택은 유효성과 안전성에 근거해야 한다는 점을 반영하는 것이다. 따라서 선정된 대조요법과 비교하려는 시험요법에서 예상되는 내성/독성을 계획 단계에서부터 고려하는 것이 타당하다. 안전성 자료는 최초의 확증임상시험 수행 전에는 다소 제한적일 수 있으나, 대체적으로 주요 독성은 파악이 되어야 하며, 이는 대안적 표준요법과 비교해 시험요법의 상대적 독성을 개략적으로 추정하는데 충분해야 한다.

본 문서에서는 각 시험군의 예상되는 독성에 따라 다음의 3개의 분류를 사용하였다: 감소되거나 비슷한 독성, 증가된 독성 및 매우 증가된 독성. 각 상태의 다양성 등으로 인해 여기서는 정확한 정의는 제공하지 않는다. 하지만, 대부분의 경우, “매우 증가된 독성(major increase in toxicity)”은 시험요법으로 인해 치료관련 사망의 증가, 삶의 질에 장기적 영향을 주는 비가역적 부작용, 또는 환자의 상태에 심각한 손상 발생 등으로 인식된다. 2차 종양 발생의 위험성을 포함하여 다른 문제들도 고려해야 한다. 이러한 분류는 확증임상시험 계획에 대한 가이드라인을

제공하고, 적절한 유익성-위험성 평가를 위한 고려사항을 제시하는 것이다.

## 1. 설계

### 1.1. 환자집단

진단, 치료개시 기준, 선정기준, 반응기준 및 대조요법 선택 등은 과학적 증거 및/또는 일반적으로 인정되고 있는 최신의 진료가이드라인 등에 근거하여 정당화되어야 한다. 또한, 바이오마커를 이용한 탐색임상시험은 3상 임상시험의 계획 단계에서 유익성-위험성을 최적화하기 위한 환자 선정기준에 대한 자료를 제공할 수 있을 것이다.

치료군 간의 차이를 발견할 수 있는 시험의 검정력을 증가시키기 위해서는 시험대상 집단의 이질성(수행능력, 동반질환, 장기의 기능장애 등)을 최소화 하는 것이 바람직하다. 이는 임상적으로 대표성을 가지는 환자들의 등록 의도와 실제 등록 가능한 환자가 있는지가 균형을 이루어야 한다. 따라서 시험자는 실제 임상에서 시험약 치료를 받을 가능성이 있는 환자군을 임상 시험에 등록할 것을 권장한다. 수행능력 등의 제한사항은 추후 허가사항에 반영되어야 한다.

환자들은 관련 종양 인자들(예, 수행능력, 동반질환, 장기 부전, 병기, 분화도, 표적의 발현, 예후에 중요한 바이오마커 및/또는 종양 민감성, 기존 치료 효과[반응/내성/난치 등])에 따라 특징을 분류해야 한다. 중요하고 이미 알려진 예후 변수들에 근거한 층화가 고려되어야 하며, 층화에 사용된 것이 아닌 다른 공변량을 분석에 포함시켜야 하는 경우 이러한 요소들이 계획서나 통계분석 계획에 기술되어야 한다.

만약 탐색임상시험이 종양 표현형/유전자형에 근거해 특정 환자들을 포함/제외시키는 기준을 제공하는 경우, 나중에 의약품 허가사항(labelling)에 반영될 것이다. 활성 표적이 없는 환자들이 포함된 임상시험의 경우, 예를 들어 하위 그룹분석을 통해 표적 발현이 항종양 활성에 무관하다는 점이 입증되지 않았다면, 제한된 의약품 허가사항(labelling)이 여전히 적절할 수 있다.

만약 임상시험의 등록을 정하는 바이오마커가 현장 또는 지역에서 시행 가능하다면 민감도 분석을 가능하도록 하기 위해 바이오마커에 대한 중앙 평가(central assessment)를 보완하여 임상 시험 중에서도 평가가 이루어지는 것이 권장된다.

일부 희귀한 상태의 경우, 임상시험의뢰자는 대안적인 기준을 사용해 표적 집단을 정의하고자 할 수 있다. 예를 들어 특정한 분자학적으로 잘 정의된 구조를 표적으로 하는 시험약의 활성을 조사하는 시험의 경우, 해당되는 표적을 가진 상이한 암종의 환자들을 등록하는 것도 가능할 수 있다

하지만 상이한 암종에서 표적의 결정적 역할이 입증되어야 한다. 적절한 대조치료 선택의 가능한 결과에 대해서도 고려해야 하며, 상이한 암종을 가진 환자 하위 집단 내의 유익성-위험성에 대한 결론을 내릴 수 있도록 비임상 및 약리학적 자료를 포함한 모든 이용 가능한 입증 자료에 근거해 시험을 설계할 것을 권장한다. 이와 같은 대상 환자의 비전통적인 선정기준을 사용한 확증임상시험을 시작하기 전에 규제기관과 논의되어야 한다.

일부의 표적 적응증(target indication)이 매우 작은 환자 그룹에만 적용될 수 있으며, 그룹이 너무 작아 “예외적인 상황”이 적용될 수 있다. 활성 표적이 이러한 드문 상태에서만 발현되는 경우가 아니라면, 의뢰자는 일반적으로 이러한 소그룹에 대한 시험은 병행하여 진행하도록 하거나, 특히 안전성과 관련 더 종합적인 평가를 허용하는 적응증에서 유익성-위험성이 확립된 경우 진행하도록 권장될 것이다.

## 1.2. 대조요법

대조요법은 통상적으로 증거에 기반한 최선의 이용 가능한 치료법을 선택해야 한다. “증거에 기반한 최선의 이용 가능한(best available, evidence-based)”이란 무작위배정 임상시험을 통해 유리한 유익성-위험성이 확고하게 입증된 허가 받은 요법이어야 하는 것은 아니지만, 널리 사용되고 있으며 적어도 유익성-위험성의 관점에서 대안적인 치료법만큼 좋아야 한다.

지역에 따라 다르게 선호되는 표준요법이 존재한다는 사실은 알려져 있다. 우월성 시험(시험요법 대 대조요법)에서는 대조요법이 상기에서 정의되었듯이 증거에 기반한 것이면 문제가 되지 않는다. Add-on studies(대조요법+시험요법 대 대조요법)는 지역에 따라 선호하는 몇몇 개의 대조요법을 설정하는 것도 가능하다. 이 경우 타당한 확증적 임상적/약리학적 근거가 필요하며 규제기관과 논의를 권장한다. 1개 이상의 대조요법이 사용되는 경우, 층화가 권고되며 전체적인 우월성 결과는 한 개의 대조요법의 열등한 결과에 의해 좌우되어서는 안 된다.

만약 목적이 비열등한 효과를 입증하는 것이라면, 선정된 대조요법은 반드시 적절한 비열등 경계가 정의될 수 있는 치료법이어야 한다. 대부분의 경우, 이 대조요법은 이전의 무작위배정 임상시험들을 통하여 그 당시 대조요법과 비교하여 그 우월성을 이미 입증했던 치료방법이어야 한다. 이에 대해서는 V. 3상, 치료적 확증임상시험 3. 방법론적 고려사항 3.3. 사건까지의 시간 분석 및 반응과 진행 평가를 참조한다.

최선의 이용 가능한 대조요법들 중에서 유사한 주기 기간을 가진 요법이 우선적으로 고려되어야 하는데 그 이유는 동일한 종양평가 일정으로 수월하기 때문이다. 만약 임상시험의 목적이 내성 및 독성을 향상시키는 것이 아닌 경우, 시험요법과 유사한 독성이 예상되는 요법이 선호된다. 이 경우 이중맹검 조건 하에서도 시험 진행이 가능하다. 활성 대조요법이나 BSC에 시험약을

추가하여 진행하는 시험(add-on studies)인 경우는 의미가 있다고 판단되면 위약을 사용하도록 권장된다.

실제 임상에서 환자들이 특정 요법에 따라 치료를 받고 있음에도 근거가 확립된 대조요법이 없는 경우도 있다. 이러한 경우 BSC를 대조요법으로 사용할 수는 있으나, 종종 반응률이 입증된 현재 사용하는 치료 요법이 더 바람직하다. 하나의 대조요법을 선택할 수 없는 경우, 시험자의 최선의 선택이 옵션이며, 이러한 경우, 낮은 독성을 가진 대조요법이 선호되며, 환자관련 평가 변수의 우월성이 입증되어야 한다.

증거에 기반한 치료법이 없다는 것은 환자들이 여러 치료 요법에서 실패했다는 의미이다. 이러한 상황에서는 다수요법에 실패한 환자(last line)에서 시행된 BSC/시험자의 최선의 선택 등과 비교한 무작위배정 임상시험 보다는 “구제요법” 단일군 연구에 의해 뒷받침되는, 질환이 덜 진행된 환자들에서 적절하게 수행된 무작위배정 시험으로부터 품목 허가에 필요한 자료를 얻는 것이 더 유익하고 쉬울 것이다.

### 1.2.1. 단독요법 및 병용요법

시험요법이 단독요법이든 혹은 병용요법이든, 그 시험요법은 효과뿐만 아니라 유익성-위험성과 관련해 “최선의 이용 가능한(best available)” 대조요법과 비교되어야 한다.

만약 시험약(A)을 확립된 요법(B)에 추가할 경우, AB 대 B 비교 시 AB의 우월성과 유익성-위험성에 대하여 유리함이 입증되어야 한다. B의 용량 강도(dose intensity)와 유익성-위험성에 관련된 이용 가능한 자료에 근거하여 논의되어야 한다. 전통적으로 이런 종류의 시험에는 A만으로 구성된 3의 치료군(A alone third arm)을 포함시키지 않는다. 그러나 이 또한 이용 가능한 탐색임상시험 자료에 근거하여 정당화 되어야 한다. 대체 시험(substitution study)인 경우, 즉 확립된 요법(BC)의 구성요소(C)를 시험약(A)로 대체하고, 비열등성(BC 대 BA)이 목적인 경우, BC의 활성화에 대한 C의 기여도가 명확히 정의되어 있어야 한다.

드물게는 완전히 새로운 조합인 AB를 대조요법과 비교하는 경우, 확실한 비임상 및 1/2상 임상 자료로 시험요법에 두 가지 약물을 병용하는 필요성을 설명해야 한다.

### 1.3. 교차(cross-over)

적합한 유익성-위험성 평가를 위해서는 OS에 불리함이 없는 경우에 한해, 병이 진행되었을 때 교차를 시행해야 한다.

## 1.4. 무작위배정 및 맹검

무작위배정 및 층화는 널리 인정되는 일반적 규정에 따라야 한다. 많은 경우 비교되는 치료 간 명백한 독성의 차이, 혹은 안전성에 관한 우려 때문에 이중맹검 설계는 불가능할 수 있다. 만약 임상시험이 공개시험으로 수행될 경우, 시험의 평가변수 선정, 독립적 검토, 민감도 분석의 실시 및 기타 공개시험 자체가 갖는 가능한 비뚤함을 줄이기 위한 여러 방법들이 고려되어야 한다.

## 1.5. 평가변수

### 1.5.1. 1차 평가변수(primary endpoints)

확증임상시험에서는 시험약의 임상적 유익성을 증명하여야 한다. 따라서 임상시험의 1차 평가변수가 선정기준에 타당한 대상 환자군에 대하여 임상적 유익성을 확실하고 유의하게 나타낼 수 있다는 충분한 증거가 입증되어야 한다.

타당한 1차 평가변수에는 완치율(cure rate), 전체 생존(OS) 및 무진행 생존(PFS)/무질병 생존(disease-free survival, DFS)이 포함된다. 확실하게 증명된 OS의 긍정적인 효과는 임상적 및 방법론적 관점에서 임상시험의 가장 설득력 있는 결과이다. PFS/DFS의 연장도 환자에게 이익으로 간주된다. 1차 평가변수는 시험요법의 상대적 독성에 따라 선택되어야 하나, 예를 들어, 질병 진행 후의 예상 생존율, 가능한 차선의 요법 및 현재의 상황도 고려해야 한다. 선택된 1차 평가변수와 상관없이 모든 관련된 결과에 있어서 치료 효과의 정도가 유익성-위험성 평가의 근거가 된다.

만약 PFS/DFS가 1차 평가변수라면 OS는 2차 평가변수로 설정해야 하며, 그 반대도 마찬가지이다.

OS를 2차 평가변수로 사용할 경우, 대개는 우월한 경향을 보여줌으로써, OS 관련 치료 효과에 부정적 영향이 없다는 것을 보여주는 것이 권장된다. PFS에 상당한 효과가 있는 경우, 또는 질병진행 후 장기 생존이 예상되는 경우 및/또는 안전성이 명백하게 우월한 경우, OS에 대한 정확한 추정이 반드시 품목 허가에 필요하지 않을 수 있다. 추후 OS에 대한 확인이 권장된다.

OS가 1차 평가변수로 보고된 경우, PFS에 대한 효과도 일관성이 있을 것으로 기대된다. 만약 그렇지 않을 경우가 예상되면, 예를 들어 어느 특정 면역조절요법의 경우, 임상시험 계획서에 분명하게 기술되어야 한다.

어떤 상황에서는 질병의 진행이 아주 천천히 관찰되기 때문에 자주 질병 진행을 평가해야 하며 이는 환자들에게 부담이 될 수 있다. 이러한 경우, 미리 구체적으로 충분히 근거가 있는 정해진 시기에서의 사건 발생률(event rate)을 1차 평가변수로 사용할 수 있다. 만약 어느 특정 시점에서의 사건 발생률이 1차 평가변수로 선택되었다면 모든 환자들이 해당 기간 동안에 시험에 참여하고 있을 것을 권장한다. 고정 시점에서의 평가가 1차 결과 측정방법으로 사용된 경우, 사건 발생까지의 시간 분석에서 시험자가 평가한 PFS는 2차 평가변수로 보고되어야 한다.

종양의 약물내성 프로파일은 치료요법의 영향을 받으며, 이는 후속 치료요법의 활성화와 관련될 수 있다. 예를 들어 난소암의 1차 치료, 비소세포폐암 및 일부 혈액질환과 같이 제한된 주기수의 치료가 일반적인 질환에서 유지/장기간 요법을 위약/치료를 하지 않는 군과 비교하는 경우에서 분명해진다. 유지요법에서 병이 진행된 경우는 적어도 이 유지요법에 내성이 생겼다는 의미이며, 질병의 진행을 막는 치료와는 다를 수 있다. 후속 치료에 대한 교차 내성의 정도는 시험요법과 대조요법이 서로 다를 수 있다.

규제 관점에서 이러한 사항은 유지요법 또는 “유도(induction)” 주기의 증가와 같은 새로운 개념을 도입한 설계에서 중요하다. 따라서 가능하다면 이러한 시험은 생존의 관점에서 환자의 유의성을 입증하도록 설계되어야 한다. 만약 이것이 실행 가능하지 않는 경우, 후속 치료요법에서의 PFS(PFS2)와 같은 평가변수가 정해져야 한다. 이를 위해서는 질병이 진행된 후 합의된 후속 치료가 대조 및 유지치료군에서 사용되는 것이 이상적이다. 후속 치료에 미치는 부정적 효과와 유지요법에 대한 내약성 및 독성을 극복하기 위해서는 시험군에서 무작위배정 시점에서부터 PFS2까지의 시간이 대조군 보다 월등히 우월해야 한다. 규제적인 경험의 제한적이고 방법론적인 이슈들이 예상되기 때문에 규제기관과 논의를 권장한다.

만약 유지요법에 사용된 시험약물이 재발 시에 단일제제로서 사용될 수 있는 경우, 유지치료와 같은 초기 치료와 진행 시점에 사용한 지연 치료(deferred therapy)를 비교하기를 권장한다.

시험 계획서 안에 후속 치료를 정하고 PFS2까지 예정대로 환자들을 평가하는 것이 실행 가능하지 않을 수 있다. 이 경우 후속 치료까지의 시간을 PFS2의 대리변수로 사용할 수 있다. 시험군과 대조군에서 PFS2를 비교할 때 PFS2의 커지기 쉬운 변동성을 고려해야 한다.

후속 치료의 선택은 환자의 질병 진행 시점에서의 수행능력(performance status)을 반영할 수 있다. 이는 실제 임상 현장과도 관련되기 때문에 유지치료요법과 같은 새로운 개념을 소개하는 시험뿐만 아니라 대부분의 시험에서 후속 치료까지의 시간을 수집하도록 권장한다. 이 경우, 증례기록서에 특정 후속 치료를 선택한 이유를 포함한다면 도움이 될 것이다.

이러한 모든 경우에 자료의 완성도(maturity)에 대한 논의가 필요하다. 일반적으로 공격적인 질

환은 치료 유무와 상관없이 조기 질병 진행을 보이기 때문에, PFS1에서 열등한 결과를 보여준 군이 초기 PFS2에서는 우월하다는 비뚤린 결과를 보여 줄 수 있다.

TTP와 TTF와 같은 변수들이 드물게 1차 평가변수로 대체될 수 있다. 하지만 이 경우는 충분히 그 타당성이 입증되어야 한다.

임상시험 시작 시점에서 종양 관련 증상을 지니고 있는 환자를 대상으로 할 때, 증상 조절이 항암 활성화와 연관되어 있다면 증상 조절 여부가 유효한 치료 활성화의 척도가 될 수 있다. 또한, 비뚤림의 가능성을 최소화한다는 조건하에서 기존에 많은 치료를 받은 환자를 대상으로 하는 시험에서는 증상 조절이 시험의 1차 평가변수가 될 수 있다. 몇몇 경우 증상이 동반된 종양 진행까지 시간(time to symptomatic tumor progression) 역시 환자 유익성에 대한 적절한 평가 지표가 될 수 있다.

종양 반응과 연관된 상황들, 예를 들어 사지 보존 수술(limb-saving surgery)이 환자 유익성 평가의 적절한 1차 지표가 될 수도 있다. 그러나 이러한 분석은 임상시험의 주요 목적인 치료 전략의 전체적 성공에 대한 초점을 흐릴 수 있기 때문에 일반적으로 특정 부위 혹은 특정 원인과 관련된 사건들의 분석은 피해야 한다.

종양 부담과의 연관성이 확실히 입증된 바이오마커(예, 다발성 골수종과 M-단백)는 다른 종양 부담의 지표들과 조합하여 총체적 종양 반응/진행 평가에 사용할 수 있다. 그러나 새로운 계열의 물질에 대해서는 그 표지자가 종양 부담 측정의 적절한 지표이며, 평가에 이용 시 비뚤림이 발생하지 않음(예를 들어 종양 표지자의 선택적 억제에 의한 현상이 아님)을 입증하여야 한다.

### 1.5.2. 2차 평가변수 및 탐색적 분석

OS나 PFS 중 어느 것이 1차 평가변수이든 간에 ORR과 특정 기간(예, 3개월, 6개월)의 종양반응 안정화 비율은 반드시 보고되어야 한다. 특히 고식적 요법(palliative setting)인 경우에는 일반적으로 인정되는 측정방법을 이용한 HRQoL/PRO(patient reported outcome) 평가가 유용할 수 있다.

## 1.6. 완치목적으로 시행되는 치료

고등급(high grade)의 림프종이나 생식세포종양이 있는 환자 또는 수술 후 보조요법 치료를 하는 경우 새로운 치료를 개발하는 궁극적 목적은, 완치율 및 생존을 향상시키거나, 치료 효과를 유지 하면서 관련된 독성을 감소시키는 것이다. 그럼에도 불구하고 급성 골수성 백혈병과 같이 치료가 복잡한 경우에는 비교적 유효한 시험약물이 이러한 유효성 평가변수들에 미치는 영향을 보여주는 것이 어려울 수 있다.

시험약물을 단독요법으로 사용하는 경우는 거의 없을 것이고, 이미 확립된 요법에 추가하는 요법이나 수정된 요법이나 확립된 요법의 일부를 시험약물로 대체할 것으로 예상된다. 이러한 맥락에서 유지요법이 확립되지 않았다고 판단되는 경우에는 이러한 유지요법을 추가요법으로 간주할 수 있다.

급성 백혈병 치료에서 완전 반응(complete response, CR) 획득 실패, 재발, 재발과 관련 없는 사망은 무사건 생존(event-free survival, EFS) 분석에서 사건(event)으로 간주된다. 사전에 정해진 관해 유도치료 단계에서 CR을 얻지 못한 환자들은 원점(time 0)에서 사건이 발생한 것으로 간주될 것이다.

고등급 림프종이나 고형 종양을 가진 환자에서는 PFS가 결과 지표로 사용될 수 있다. 예정된 기간/주기 안에 최소한 부분 반응(partial response, PR)을 얻지 못한 경우는 일부 계획서에서는 치료 실패로 간주되어 최소한 PR을 얻은 환자들만 지속적으로 치료를 받게 된다. 1차 분석에서는 PR에 도달하지 못한 환자들은 추적을 종료하거나 질병진행 또는 사망 사건이 발생할 때까지 추가 후속 치료(next line)를 하는 것을 권고한다.

완치율 향상이 치료 목적인 경우에는 미리 정해 놓은 시점에서의 DFS가 결과 지표로 사용된다.

#### 1.6.1. 독성 감소 또는 유사한 정도가 예상되는 경우

많은 경우 확립된 요법(AB)에서 A를 시험물질인 X로 대체하는(XB) 대체 시험 설계가 예상된다. 이 경우 규제 관점에서는 비열등성 시험 설계가 수용 가능하며, 대부분의 경우 PFS를 1차 평가 변수로 받아들일 수 있다.

관해 유도치료 후 공고치료 및/또는 유지치료를 하는 경우에는 시험약물 종료 이후에 투여되는 치료에 의한 교란 효과가 발생할 수 있기 때문에 PFS가 아닌, 보다 더 적절한 변수를 선택할 수 있다. 즉, 만일 추가 치료가 예정된 경우라면, 시험약물 종료 후 CR(그리고 특별히 정한 경우 CR+PR)이 적합한 1차 평가변수가 될 수 있다. 이러한 경우에는 시험물질이 공고치료를 미칠 수 있는 영향에 대해 항상 설명되어야 하며, CR에 관한 결과는 PFS에 의해 뒷받침되어야 한다.

CR의 정의는 확립된 임상 기준을 따를 것을 권고한다. 그러나 적용 가능한 경우, 예를 들어 분자적 기준에 의해 정의되는 미세잔존질환(minimal residual disease, MRD) 같은 보조적인(supportive) 증거를 찾아야 한다. 다른 바이오마커에 대해서는 표준화를 통해서 실험실 내 및 실험실 간의 변동성을 최소화해야 한다.

#### 1.6.2. 독성 증가가 예상되는 경우



대체 또는 부가요법 설계를 적용할 수 있다. 대부분의 경우는 PFS 또는 OS의 우월성이 입증되어야 하며, 사건 발생까지의 기간이 연장되는 유익성이 독성 증가를 상쇄할 만큼 충분히 크다는 것이 입증되어야 한다.

만일 시험약물 종료 후에 시행한 치료가 전체 결과에 교란을 줄 가능성이 있다면, 또는 OS의 추이와 관련이 있는 관해 유도치료 후 CR의 주요 증가도 결과 지표로 충분할 수 있다. 이 방법은 만약 대상 환자군이 작다면 특별히 고려할 수 있다.

### 1.6.3. 독성 증가가 매우 클 것으로 예상되는 경우

완치율 증가나 OS 향상 입증이 목표가 되어야 한다. 대상 환자군이 작은 시험의 경우에는 PFS의 주요 증가가 생존에 유리한 경향과 관련 있는 적절한 보조 자료로서 충분히 사용될 수 있다.

## 1.7. 장기 질병 조절 목적의 치료

확립된 표준치료가 가능하고 후속 치료 선택이 충분히 효과적일 수 있는 진행된 유방암, 대장암, 저등급(low grade) 림프종 및 만성백혈병 같은 질환에서의 조기 치료 같은 상황이 이에 해당된다.

### 1.7.1. 독성 감소 또는 유사한 정도가 예상되는 경우

대체요법이나 단독요법 시험이 예상된다. 규제 관점에서는 비열등성 설계가 받아들여질 수 있으며, PFS가 적합한 1차 평가변수로 고려된다. 독성이 상당히 감소된 경우, 자료가 타당하다고 판단된다면 품목 허가 이후 축적된 생존 자료를 제출할 수 있다.

### 1.7.2. 독성 증가가 예상되는 경우

목적은 최소한 PFS의 우월성을 입증하는 것이다.

생존 자료는 품목허가 시점에 이용 가능해야 한다. 모든 경우에서 완전한 생존 자료를 기대할 수는 없으나, 왜 그런 상황인지에 대한 타당한 설명이 제시되어야 한다. 이러한 경우에는 품목허가 후 생존 관련 추적 관찰이 예상된다. 만일 치료와 관련한 사망의 증가가 없다는 것을 합리적으로 확실하게 입증하지 못한다면, 허가 전에 유익성-위험성 평가를 위한 이용 가능한 완전한 생존 자료가 있어야 한다.

유지치료가 확립되지 않은 분야에서 시험을 수행하는 특수한 경우에는 대체 유효성 평가변수가 더 적절할 수 있다(V. 3상, 치료적 확증임상시험 1.5. 평가변수 참조). 강력한 유도요법 이후에

치료 휴약 기간을 장기간 연장하는 것이 목적이 될 수 있다.

### 1.7.3. 독성 증가가 매우 클 것으로 예상되는 경우

주목적은 생존 향상을 입증하는 것이다.

경우에 따라서는 대상 환자군이 작고, 질병 진행 시점에서 시험약물을 포함한 여러 가지 이용 가능한 효과적인 후속 치료제가 존재하는 경우는 예후가 좋을 것으로 예상되기 때문에 이 목적을 달성하지 못할 수 있다. 만일 PFS를 1차 평가변수로 선택했다면 이에 대한 타당성이 필요하다. 또한 치료 관련 사망 가능성을 평가하기 위해서 계획 단계부터 주의 깊은 논의가 필요하다. 비록 PFS 연장의 이익이 매우 큰 경우에만 받아들여기로 했을지라도, 양호한 생존 경향이 예상되는 경우 가능한 OS를 추정할 수 있을 만큼 환자수가 충분한 것이 좋다.

### 1.8. 고식적 치료(Palliative therapy)

생존에 대한 예후가 불량하고, 충분히 문서화된 대조요법을 확인하기 어려운 마지막 단계에 있는 환자들의 설정이 주로 해당된다. 또 다른 경우는 명확하거나 가능한 모호하지 않은 정의에 의한 강력한 완치요법을 받기에 적합하지 않은 환자들도 포함된다.

확립된 대조요법이 없는 경우, 시험자가 선택한 최선의 방법 또는 위약이 포함되거나 포함되지 않은 BSC가 허용될 수 있다.

BSC를 대조요법으로 시험이 수행되는 경우, 시험의 목적은 OS의 연장 및/또는 전반적 증상 조절 또는 HRQoL 개선하는 것이다. 후자는 가능한 비뚤림을 줄이기 위해 모든 노력을 기울여야 한다. 시험의 목적에 관계없이 이 환자군에서의 시험은 치료가 견딜만해야 한다.

대조요법이 확립된 것은 아니지만 효과적으로 알려진 경우에는 PFS의 우월성이 받아들여질 수 있다. 이 경우, 다음 사항이 유의성-위해성 평가에서 고려되어야 한다: 대조요법의 유용성을 보여주는 증거, 대조요법보다 우월한 PFS 유의성의 정도, 내약성/독성 프로파일, 질병 진행 후의 생존 및 상태의 빈도.

시험이 시작될 때 수행능력이 나쁜 환자들은 고식적 치료가 적합할 것으로 여겨질 수도 있지만, 치료반응이 상당히 좋을 수 있고, 추후에 완치 목적의 치료가 가능해질 수도 있다. 이러한 환자들을 어떻게 다루어야 할지에 대해서 분석 계획에서 정의되어야 한다.

## 2. 특별 고려사항

## 2.1. 수술 후 보조요법 및 선행요법

보조요법의 궁극적인 목적은 완치율(cure rate)을 증가시키는 것이다. DFS의 연장이 대상 환자에게 의미가 있다고 판단되면 시험 계획 시 DFS 연장이 완치율에 긍정적인 효과를 입증하는 것이 가능한지 여부를 고려하는 것이 중요하다. 즉, 재발률이 현저하게 안정기(plateau)에 도달한 시점에 분석 시행을 고려한다.

보조요법이 재발 시점에서 선택할 수 있는 치료 요법을 제한할 수 있으므로, OS 자료가 보고되어야 한다. 보조화학요법이 표준치료로 확립된 분야에서는(예, 유방암 및 대장암) DFS 및 이용 가능한 안전성과 재발 이후의 후속 치료에 따른 PFS를 포함한 생존 자료에 근거하여 시험군에서의 유익성-위험성이 긍정적으로 간주되는 경우, 품목 허가 이후에 추가로 완전한 OS 자료를 보고할 수 있다. 주요 독성에 대한 우려가 있는 일부 경우에는 OS에 대한 긍정적인 효과가 입증되어야 한다.

선행요법의 목적에는 전반적인 치료 결과 개선(PFS/DFS 및 OS)과 장기 보존(더 보존적인 수술) 등이 있을 수 있다. 만약 장기 보존이 주요 목적이라면, 최소 비열등한 무질병 생존/무진행 생존(DFS/PFS)이 입증되어야 한다. 수술 후 보조요법의 경우처럼 흔히 정해진 주기 수가 부여된다. 시험의 목적에 따라서는 지정된 치료 기간 이후 종양 축소가 관찰되지 않을 경우 치료중단을 허용할 수 있다.

수술 시의 병리학적 완전반응(pCR)을 2차 평가변수로 보고하는 경우, 치료가 중단된 환자는 비반응자로 간주되어야 한다.

## 2.2. 약물 내성 조절제, 방사선/화학요법 감작제 및 화학요법보호제

원칙적으로 약물 내성 조절제 및 방사선/화학요법 감작제에 해당하는 시험약(A)에 대한 확증적 임상시험 설계는 기존에 확립된 요법인 B보다 AB가 항암 활성 및 유익성-위험성 평가에서 유익성이 입증되어야 한다. 약동학적 상호작용 또는 항암 활성과 관련 없는 약력학적 상호작용이 존재하는 경우, 유사한 독성 수준에서 AB와 B를 비교하기 위해 AB의 병용요법 중 B의 용량을 조절할 필요가 있을 수 있다. 그러나 약동학적 상호작용과 관련된 전체 효과들이 B의 혈중 농도 변화에 따라 측정된다면(즉, 분포에 변화 없는 경우), B의 유사한 노출 수준에서 AB 대 B를 비교하기 위하여 B의 용량 조절이 선호된다.

화학요법보호제의 경우, 정상 조직이 종양 조직에 비해 독성으로부터 더 보호된다는 것을 입증해야 한다. 세포독성이 심한 물질의 경우, 용량과 관련된 효과의 차이보다는 용량과 관련된 독성의 차이를 관찰하는 것이 더 쉽다. 즉, 항암 활성의 감소 없이 정상 세포에 대한 보호가 이루어짐을 입

증하기 위하여, 항암 활성과 관련된 측정 항목에 대해 좁은 신뢰성 구간을 설정한 대규모의 임상 시험이 필요하다. 따라서 안전성이 개선되고 동시에 항암 활성에 대하여 비열등하다는 가설을 입증할 공동 1차 평가변수들이 필요하다. 한편, 몇몇 경우 시험군에 투여하는 세포독성 물질의 용량을 증가시켰을 때 독성의 증가 없이 항암 활성이 증강됨을 입증함으로써 조직 보호의 차이를 설득력 있게 보여주는 것이 더 쉬운 방법일 수 있다.

그러나 약동학적 상호작용이 없고, 화학요법보호제가 종양과 상호작용이 없다는 것이 입증되면 (예, 종양세포 내 표적이 없는 경우) 종양 보호에 대한 공식적인 비열등성의 입증 대신 독성 감소만을 입증하는 것이 인정될 수도 있다.

## 2.3. 종양 예방

이에 대한 규제 경험은 제한적이나, 개념적으로는 보조요법 상태와 유사하다. 그러므로 관찰된 종양 발생 위험의 감소가 예방 요법의 부작용을 상회하는지 판단하기 위해 종양 발생 위험이 있는 대상군이 명확히 정의되어야 한다. 종양 예방이 변경된 생물학적 양상을 나타내는 종양을 대상으로 할 수 있으므로, 종양의 표현/유전형에 대한 비교 자료가 요구되고 생존에 대한 자료가 필요할 수 있다. 이러한 시험 계획 시에는 규제기관과의 협의가 권장된다.

## 3. 방법론적 고려사항

흔히 하나의 적응증을 얻기 위해서는 한 개의 임상시험이면 가능한 것으로 예상된다. 그러나 하나의 핵심 임상시험을 근거로 품목 허가를 얻기 위해서는 표준적으로 적용되는 통계학적 유의성에 대한 일반적 기준을 초과하는 수준의 유효성 입증이 필요하다. 이는 비열등시험이나 BSC/시험자의 최적의 선택과 비교할 때 특히 중요하다. 다른 적응증에 대해 시행된 확증임상 시험으로부터 얻어진 보조적 증거들이 평가에 고려될 수 있다는 것은 잘 알려져 있다. 이러한 시험들의 보조 자료로서의 가치는 상황에 따라 다를 수 있으며 그 결과들이 새로운 적응증에도 적용될 수 있는지에 대해서는 논의가 있어야 한다.

### 3.1. 적응적 설계(adaptive design)

만약 2상/3상 임상시험이, 확증단계의 적절한 투여량과 같은, 2상 임상시험의 단일하고 복잡하지 않은 문제를 다루기 위한 목적으로만 설계된 경우, 적응적 설계는 약물 개발 효율을 증가시킬 수 있다.

더 복잡한 문제를 다루어야 하는 경우(예, 적절한 대상 환자군에 대한 정의를 포함하는 경우), 또는 다수의 문제가 있는 경우(예, 바이오마커 양성 종양 샘플에 대한 참고치 결정 및 환자수

재조정)는 적응적 설계 접근법이 이익이 되는지 검토해야 하며, 과학적 조언을 구하는 것을 고려해야 한다. 품목 허가를 위한 자료로써 독립적인 유효성/안전성 추가 시험의 필요성 또한 고려해야 한다.

최근 새로운 유형의 적응적 설계으로서 마스터 프로토콜을 활용한 임상시험이 도입되고 있으며, 해당 세부 내용은 부록 2를 참고한다.

### 3.2. 중간분석

중간분석은 3상 임상시험에서 흔히 실시되지만, 무용성(futility) 또는 차이의 발견으로 인한 조기 종료는 매우 민감한 사안이다.

만약 장기간 시험에서 기대되는 사건 수의 대부분이 관찰되었고, 차이가 나타났다면, 이러한 상황은 이 임상시험이 충분히 완성도(mature)가 있고 훨씬 더 긴 관찰기간을 갖더라도 시험 결과가 바뀌지 않을 것을 시사한다고 받아들여질 수 있다. 좀 덜 완성된 자료로부터 얻어진 중간 분석을 해석하는 데는 문제가 있다.

예를 들어, 시험 계획 단계에서 치료 효과를 낮게 예상했는데, OS에 있어 통계적으로 유의한 결과가 대표성을 가진 성숙된 자료들이 완성되기 전에 너무 일찍 입증되면 난처한 상황이 발생할 수 있다.

일반적으로 무용성(futility) 이외의 이유로 PFS 데이터에 기반한 중간분석을 시행하는 것은 권장되지 않는다.

### 3.3. 사건까지의 시간 분석 및 반응과 진행 평가

PFS/DFS를 1차 평가변수로 하는 임상시험에서는 영상 기법이나 시험 방문에 관한 균형 유지와 시험계획서에 정해진 일정을 따르는 것이 필수적이며 위반사항은 보고되어야 한다.

상기(Ⅳ. 탐색임상시험 2. 비세포독성 물질 2.2. 2상 단독요법 치료적 탐색임상시험 2.2.2. 시간 관련 평가변수가 포함된 탐색임상시험)에서 논의되었듯이, 주로 종양을 줄이는 물질과 주로 성장을 억제하는 물질 간의 PFS에 관하여 비교하는 경우, 질병의 진행시점에서 종양 부담에 있어서는 성장을 억제하는 물질이 유리하다. 현재까지 ORR과 PFS 간의 결과가 명백히 불일치할 때 이를 비교하는 규제 경험이 없고 이러한 차이를 보정하는 방법이 확립되어 있지 않다. 만약 탐색임상시험이 이런 결과를 보여주는 경우, OS와 같은 대체 평가변수가 고려되어야 한다.

시험치료와 대조치료 간에 작용 기전의 차이가 있으면 종양 부담 및 항종양 활성의 측정과 관련하여 문제가 발생할 수 있다. 앞에서 다룬 치료 초기의 종양 부종(early tumor swelling)이 한 가지 예이다. 이와 같은 문제가 예상되는 경우는 언제라도 표준적인 계측법(RECIST, WHO)을 변형할 필요가 있는데, 핵심 시험(pivotal study)이 시작되기 전에 규제기관과 합의에 이르러야 한다. 마찬가지로, 종양 측정 기술이 독립적 판단을 보장할 수 없는 경우라면 규제기관과 가능한 다른 방법에 대해 논의하는 것이 권장된다.

고등급의 신경교종 등과 같은 종양 관련 부종이 이슈가 되는 경우, 항상 거짓반응(pseudo-response) 가능성을 고려해야 하며 이를 고려한 업데이트된 반응 및 진행 평가 기준이 적용되어야 한다. 만약 동 기준이 아직 수립되지 않은 경우, 대안적 방법에 대한 논의를 위해 과학적 자문을 받을 것을 권고한다.

### 3.4. 비열등시험

평가변수로서 PFS의 비열등성 경계는 대조요법의 효과 및 여러 유형의 사건들(사망, 새로운 전이, 표적 병변의 진행, 임상적 질병진행)에 대한 논의를 포함시켜야 한다. 질병진행 프로파일에서 차이점이 예상되는 경우, 이는 적절하게 확보된 비열등성 경계와 신뢰할 만한 추정을 위해 요구되는 사건 발생 건수를 확보할 수 있는 적절한 샘플 크기 등과 함께 계획 단계에서 고려해야 한다.

시험군과 대조군이 유사한 활성을 가질 것으로 예상하는 비열등시험의 평가에서는 시험 민감도(즉, 차이를 확인하기 위한 시험의 능력)의 중요성을 고려할 때, 하위그룹(subgroup) 분석에 대한 사전 계획을 세우는 것이 중요하다(예를 들면, 치료요법 간의 활성 차이를 추적하기 힘들 수 있기 때문에 베이스라인에 나쁜 수행능력, 동반 질환 등과 같은 나쁜 예후를 가진 환자를 배제하는 것). 마찬가지로 계획서 순응(per protocol) 분석 계획을 정의하여 계획서 위반, 순응도 문제 등이 차이점을 확인할 가능성을 감소시키지 않도록 해야 한다. 이러한 분석은 일관성을 입증하기 위한 목적으로 진행해야 한다.

### 3.5. 치료 결과에 따른 환자군 간의 분석

치료 결과에 따라 환자를 분류하여 사건발생까지의 시간 변수(OS 혹은 PFS 등)를 비교하는 것은 문제점이 있다. 종양 반응, 용량 강도, 독성 혹은 순응도 등의 결과들은 치료, 환자, 종양 간의 상호작용을 나타내므로 치료 자체의 효과는 따로 떼어낼 수 없다. 그럼에도 불구하고, 독성으로 인해 용량을 감량했음에도 명백하게 생존이 향상되거나, 반응이 있는 환자에서 생존이 향상되지 않는 것과 같은 어떤 예상치 못한 결과들은 유용한 자료가 될 수 있다. 예상치 못했던 결과들을 찾아내는 것이 이와 같은 탐색적 분석을 시행하는 이론적 근거가 된다.

서로 다른 치료를 받는 환자군 간에 반응지속기간을 비교하는 것은 유용한 자료로 간주된다. 반응지속기간은 각 시험군에 대한 신뢰구간과 함께 보고되어야 하나, 이 비교는 무작위배정군에 대한 것이 아니므로 각 군 간의 통계학적 유의성에 대한 분석은 피해야 한다. 반응을 보이지 않는 환자에서는 반응유지기간을 0으로 하여 시험 그룹 간의 통계적 비교를 한다.

### 3.6. 작은 인구집단 및 매우 드문 종양에 대한 시험

종양의 표현형이나 표적 발현과 관련한 제한점으로 인하여 종양 자체가 드물거나 적응증이 매우 협소한 경우가 있다. 이때는 항종양 활성의 차이를 검출하기 위한 적절한 규모의 무작위 배정시험을 수행하기에 충분한 환자 수를 모집하는 것이 불가능하다. 어떤 경우에는 소규모, 무작위 배정, 대조군 비교 임상시험이 가장 최선책이지만, 다른 경우에는 환자 내 TTP 분석(또는 무작위 대조군 비교와의 동시 진행)이 더 좋은 대안이 될 수도 있다. 후자의 경우, 직전 요법의 TTP는 시험요법군의 질병진행이나 사망 시간과 비교될 수 있다. 시험요법 직전에 실시된 요법 및 직전 요법에서 질병진행 등의 임상적 적절성은 독립적으로 판단되어야 하며, 분석을 위해 시험계획서에 적절한 조건을 명확히 정의해야 한다. 우월성이 입증되어야 한다.

유익성-위험성 관련 기록 작성 방법에 대해서는 일반적인 해결책이 없으므로, 과학적 자문을 받을 것을 권고한다.

### 3.7. 외부 대조군의 사용

외부 대조군(문헌 대조군 포함)은 비뚤림을 제어하는 것이 불가능하므로 치료 효과가 현저하고 질병의 일반적인 경과를 충분히 예측 할 수 있는 경우에만 사용하도록 제한한다.

현저한 효과는 암치료 분야에서는 드물게 보고되고 있으나, 그러한 효과가 종종 발견된다는 것은 알려져 있는 사실이며 특히 고도로 숙련된 관찰자에게 잘 나타난다. 이러한 경우 무작위 배정, 대조군 비교시험을 통한 전향적 확증임상시험은 시험자, 환자 및 윤리위원회가 수용할 수 없을 뿐 아니라 불필요하다.

## 4. 특수 인구군

### 4.1. 고령자 및 허약 환자군

고령 환자가 임상에서 신약으로 치료를 받을 것으로 예상되는 경우, 동반 질환이 있는 고령자들을 포함해 충분히 많은 수의 고령자를 등록시켜 유익성-위험성 평가가 가능하도록 하는 것이 권장된다.

품목 허가 신청에 유용한 자료를 제공하기 위해 모든 합당한 노력을 기울여야 하겠으나, 만약 유익성-위험성이 고령자 인구나 동반 질환이 있는 대상 집단에서 명확히 평가될 수 없는 경우에는, 이러한 점이 의약품 허가사항에 반영되어야 하며 허가 후 임상시험이 필요할 수 있다.

고령 환자로부터 얻은 데이터는 PK 분석(예, 집단 PK 분석의 일부로서)에 이용 가능해야 하며, 안전성 프로파일에 대한 설명에도 부작용 프로파일 및 결과(예, 용량 감소, 약물 투여 지연 또는 동반 치료의 개시 지연)의 심각성을 포함시켜야 한다. 젊은 환자 집단과 함께, 고령자 집단의 치료 효과의 일관성 및 안전성 프로파일에 대하여 연령별로 분석되어야 한다.

몇몇 화합물은 저하된 장기의 기능에 대해 감수성이 낮은 등 약동학적 특성 때문에 특히 고령자의 치료에 적합할 수 있다. 이러한 경우 고령자만을 대상으로 한 임상시험이 권장된다. 이러한 경우 중 일부에서는 적절한 대조요법을 선정하는 것이 어려울 수 있고, PFS/OS 이외의 다른 측정 지표가 더 타당할 수 있다는 사실이 알려져 있다. 이러한 경우 개발 계획에 대해 규제기관과의 협의가 권장된다.

고령 여부와 별개로 명백히 수행능력이 나쁜 허약한 환자군은 전형적인 확증 시험에는 거의 포함되지 않는 취약군이다. 허가 관점에서는 이러한 환자군에서의 임상시험이 권장된다.

## 4.2. 성별

몇몇 종양 및/또는 치료요법의 경우 성별에 따른 항암 활성 간 차이가 보고된 바 있다. 성별에 따른 치료의 차이가 있을 것으로 예상되는 경우는 그 점이 시험 설계에 고려되어야 한다. 그렇지 않다면 여성과 남성의 비율이 실제 질병의 유병률을 반영하도록 하고, 의뢰자는 성별에 따른 하위그룹(subgroup) 분석(효과 및 안전성)을 실시하도록 권장된다.

## 4.3. 장기 기능이 저하된 환자

II. 약동학을 참조한다.

## 5. 안전성

### 5.1. 종양학에서 안전성, 기본개념 및 평가원칙

안전성 자료의 질과 정보의 유용성은 확증임상시험에서 규제적 유익성-위험성 평가를 위해 사용되는 경우뿐 만 아니라 약물 개발의 초기 단계에서도 매우 중요하다.



### 5.1.1. 기본개념

약물이상반응(adverse drug reaction, ADR)의 개념에는 인과관계의 의미가 포함되며, 임상시험에서 약물에 대한 인과관계가 있거나 없거나 이상반응(adverse event, AE)에 대한 정보는 항상 수집되어야 하고 심각성 정도가 기록되어야 한다. 인과관계에 대한 평가 이후, 일부 AE는 ADR로 결정된다. 또한, 치료유발 이상반응(treatment-emergent adverse event, TEAE)의 개념은 임상시험 시작 시에는 없었거나 시작 후 심각도가 증가된 AE를 말한다.

종양학에서 AE에 대한 현재의 표준 등급 부여 시스템은 NCI-CTCAE 독성기준이다. 내약성과 관련해 환자보고결과(patient reported outcome, PRO)를 사용해 추가로 다루어질 수 있다.

내약성의 개념은 환자의 삶의 질이나 일상 활동에 영향을 주는(예, 설사, 점막염, 신경장애 등) ADR을 말한다. 하지만 환자가 반드시 느낄 필요는 없는 혈구감소로 정의될 수도 있고, 의도된 용량 및 일정으로 약물을 주입하는 데 장애가 되는 경우를 의미할 수도 있다. 내약성에는 용량 조정 및 투여 중단을 등과 같은 기타 결과들도 반영되어야 한다.

### 5.1.2. 종양학에서 안전성

종양학에서 AE와 시험약물과의 인과관계에 대한 판단은 기저 종양질환 자체에 의한 증상 및 시험약물 이외 병용치료로부터 발생하는 독성 등으로 인해 어려울 수 있다. 이러한 어려움은 비무작위배정 임상시험에서 더 심하며, 이는, 항암제의 안전성 프로파일을 이해하는데 장애물이 될 수 있다. 특정 약물부작용이 첫 번째에서 두 번째 치료 주기 동안 가장 두드러지게 나타나고 이후 치료에서 내약성이 생기는 것은 드물지 않은 현상이다. 한편, 약물의 장기적인 치료 효과를 갖는 사람들에게 있어서는 누적독성이 발생된다. 이런 측면에서 ADR의 누적 발생률만으로는 제품의 안전성 프로파일을 충분히 설명할 수 없다.

현재 약리학적 치료의 주요군에는 세포독성제, 표적치료제, 면역조절제 등이 포함되며, 또한 재조합형 바이러스 치료제 및 세포치료제 등과 같은 발전된 치료법들이 있다. 상이한 용량 요법 및 약리학적 작용 방식이 상이한 방식으로 독성과 내성 프로파일 등에 영향을 주며, 이는 안전성 자료의 수집과 보고 계획에 포함되어야 한다. 전통적인 세포독성제는 통상적으로 일정 간격으로 투여되며, 주로 급격하지만 일시적인 독성으로, 다음 치료 주기 이전에 회복이 된다. 따라서 세포독성제의 안전성 프로파일은 질병진행까지 또는 제한된 치료 기간 동안 지속적으로 투여되는 표적치료제 또는 면역조절제 등과는 다른 양상을 보인다. 일부 약물에서는 내약성이 주요 쟁점이 되지만, 다른 약물의 경우, 잠재적으로 생명을 위협하는 부작용이 주요 쟁점이 된다. 이러한 두 가지 종류의 독성은 종합적으로 조사되어야 한다. 주요 약리학적 치료군의 빈번한 병용투여는 안전성 수집과 분석에 복잡성과 요구를 가중시킨다.

### 5.1.3. 기본 평가 원칙

유익성-위험성 균형 평가에 있어 비록 낮은 독성 등급일지라도 내약성에 영향을 주는 흔한 ADR 대 드물게 심각한 또는 생명을 위협하는 ADR에 대한 가중치는 질환의 시기에 따라 달라진다. 즉 완화치료에 있어 좋은 내약성은 우선권을 갖는 반면, 완치 목적의 치료에서는 치료의 완료에 대한 위험을 증가시키지 않는다면 내약성은 덜 강조될 수 있다. 상응하여 완화치료에서 드물게 심각한 또는 심지어 치명적인 ADR이 어떤 경우 수용 가능한 위험으로 고려될 수 있다. 반면에 보조요법에서는 많은 환자들이 수술 단독만으로 완치되고 표준 보조요법을 통해 더 많은 수가 완치되므로 생명을 위협하는 ADR이 높은 경우 바람직하지 않다. 선행요법에서 이익/위험 평가는 종양의 수술 가능성에 따라 영향을 받기 때문에 더욱 복잡해진다. 국소적으로 진행된 또는 염증성 유방암과 같은 일차적으로 수술이 불가능한 종양을 가진 환자들에게는 높은 위험도가 어느 정도 받아들여질 수 있다.

## 5.2. 안전성 자료 수집, 분석 및 보고

누적 독성을 포함한 모든 독성이 기술되어야 한다. 수집된 자료에서 질환관련 사건으로 추정된 것을 배제하는 것은, 비록 그것이 합리적 가정에 근거한 것 일지라도, 약물과의 관련성을 감지하는 것을 방해할 수 있으므로 바람직하지 않다. 완치가 목적인 경우 독성에 대해 장기간에 걸친 추적 확인이 중요하다. 늦게 나타나는 독성은 일반적으로 치료 후 수 년 후에 발생 할 수 있으며, 2차 악성종양과 특정 장기 독성(예, 중추신경계, 심혈관계)을 포함한다. 장기 독성을 경험하는 환자의 수는 시간이 경과함에 따라 증가할 수 있으며, 따라서 허가 후 약물감시 활동(post licensure pharmacovigilance activities)의 목표 가운데 하나가 된다.

독성 등급별 누적 빈도에 기초한 부작용 보고에 더하여, 특정 항암제의 안전성 프로파일에 대한 완벽한 이해를 위해서는 보완적인 척도가 필요하다. 무엇보다도 치료의 진행에 따라 특정한 AE의 발병률, 유병률 및 중증도가 어떻게 변화하는지, 그리고 용량 감소가 어느 정도 AE를 완화시키는지 이해하는 것이 중요하다. 독성과 약물 투여와의 상관관계를 이해하는 것이 핵심이다.

핵심 사건들 즉, 흔하면서 치료 주기별 내성 및 안전성에 영향을 미치는 사건들이 의미가 있다. 예를 들어, 제한된 기간 동안만 발생하는 피로 또는 3등급 설사는 내성에 큰 영향을 미치지 않을 수 있으나, 장기간의 피로 또는 2등급 설사는 유익성-위험성 균형에 주요한 쟁점이 될 수 있으며, 따라서 구체적인 분석이 필요하다. 치료 기간 당 또는 치료 주기별 독성의 발병률 및 유병률, 독성 발생까지의 소요 시간, 그리고 독성 지속 기간(등급별을 포함)과 같은 척도들을 통상적으로 고려해야 한다. PRO 또한 평가에 유용할 수 있다.

만약 AE 발현 특성상 합리적으로 정당화될 수 있다면, AE에 대한 시간 조정 분석, 예를 들어

서로 다른 종료 일자별(cut-off dates) 발병률 또는 100 환자 년당 사건 비율(event rates per 100 patient-years) 과 같은 것들 역시 적용될 수 있다. 하지만 모든 AE에 대해서 그렇게 상세하게 보고할 필요는 없다. 선별 기준에는 투여 철회, 감량, 또는 중단으로 귀결되는 사건들, 중대한 이상반응(serious adverse event, SAE), 그리고 내성 또는 유익성-위험성 균형에 영향을 미칠 가능성이 있는 사건들이 포함될 수 있다.

축발된 약물 부작용으로 인한 용량 감량의 영향을 평가하는 것이 중요하다. 또한, 용량 감량을 고려하는 경우, 종적인 PK/PD 자료는 추가적인 통찰을 제공할 수 있다. 또한 항구토제 또는 성장인자의 사용과 같은 예방적인 조치의 기대 효과를 보고하는 것이 요구될 수 있다.

때로는 호중구감소증과 관련된 감염의 중증도, 검사 자료, 입원률 및 입원 기간, 자원 활용(예, 수혈) 그리고 회복 및 치명률을 포함하는 결과 등과 같은 핵심적인 부작용을 추가적으로 기술해야 할 필요가 있을 수 있다.

보다 집중적인 세포독성/면역억제치료를 받는 환자에 대해서는, 발생 가능한, 개연성이 있는 또는 입증된 감염의 빈도와 유형(바이러스, 세균, 진균)에 대한 감시를 실시해야 한다. 장기적인 면역 결핍을 유발하는 것으로 알려져 있거나 의심되는 약물에 대해서는, 치료를 종료한 후 최대 1년까지 기회 감염에 대한 감시를 고려해야 한다. 면역관문억제제와 같은 면역조절제의 경우, 잠재적으로 발생할 수 있는 면역관련 설사/대장염, 발진, 점막염, 간 독성, 뇌하수체염 그리고 기타 내분비이상에 대한 인식과 감시가 중요하다.

안전성 평가를 촉진하기 위해서 특정 시점 즉, 3개월, 6개월, 그리고 1년의 규정된 시점에서 중추시험(pivotal studies)으로 부터 얻어진 누적 AE의 발생률을 포함해야 한다. 치료법을 시행하는 시간이 상당히 길거나 짧은 경우에는, 추가적, 또는 대체적인 시점(예, 1개월, 5년)을 고려해야 한다.

## Ⅵ. 용어 정의 및 약어

객관적 반응률(objective response rate, ORR)

CR과 PR이 관찰된 환자들의 비율

건강관련 삶의 질(health-related quality of life, HRQoL)

1차 내성(primary resistance)

객관적 반응 혹은 증식억제가 없는 진행(progression)

내성(resistance)

치료 종료 후 정의된 기간내 진행(progression)

2차 내성(secondary resistance)

증명된 객관적 반응 혹은 증식억제 후 진행

무진행 생존(progression-free survival, PFS)

무작위배정부터 객관적인 종양진행 혹은 사유를 불문한 사망까지의 시간

2차 무진행 생존(PFS2)

무작위배정부터 후속 치료 중 객관적 종양진행 또는 사유를 불문한 사망까지의 시간. 일부의 경우, 후속 치료 기간이 PFS를 대체할 수 있음

무질병 생존(disease-free survival, DFS)

무작위배정부터 재발 혹은 사유를 불문한 사망까지의 시간

부분 반응(partial response, PR)

분자표적물질(molecularly targeted agent, MTA)

비세포독성 물질(non-cytotoxic)

세포독성 물질에 속하지 않는 항암물질

세포증식억제제(cytostatic)

비임상시험에서 종양세포생존에 직접적인 영향을 미치지 않으나 세포분열을 저해하는 항암물질

세포독성 물질(cytotoxic)

비임상시험에서 단기간의 노출로 DNA 복제, 세포분열 등 저해를 통해 비가역적이고 치명적인 장애를 일으키는 항암물질

안정병변(stable disease, SD)

약동학(pharmacokinetics, PK)

약력학(pharmacodynamics, PD)

약물내성 조절제(drug resistance modifier; chemosensitizer)

항종양화학물에 약력학적 작용을 미쳐 활성을 증가시키나 자기 자신은 항종양활성을 가지고 있지 않은 물질

용량제한독성(dose limiting toxicities, DLT)

완전 반응(complete response, CR)

2상 권장용량(recommended phase 2 dose, RP2D)

이상반응(adverse event, AE)

임상적 유효 반응률(clinical benefit response rate, CBR)

CR 또는 PR 또는 장기간의 SD. “장기간의 SD” 는 조건 특정적으로 정의되며, 일례로 유방암의 경우는 통상적으로 24주 이상임

전체 생존(overall survival, OS)

무작위배정부터 사망까지의 시간

종양 진행까지 시간(time to tumor progression, TTP)

무작위배정부터 종양진행이 관찰될 때까지의 시간으로서 다른 원인(기저 악성종양과 관련 없는)으로 인한 사건(사망)은 제외한다.

체표면적(body surface area, BSA)

최대내약용량(maximum tolerated dose, MTD)

일반적으로 6명의 시험대상자 중 적어도 2명에서 용량제한독성(DLT)이 나타나는 용량으로 정의되며, 그 이상의 용량 증가는 수행되지 않는다.

최적 지지요법(best supportive care, BSC)

항생제, 영양학적 유지, 대사장애 환자의 보정, 최적증상조절 및 통증관리가 포함되나, 종양특이적인 치료는 포함되지 않는다.

치료 실패까지 시간(time to treatment failure, TTF)

무작위배정으로부터 사망, 진행, 독성으로 인하여 치료의 중단 혹은 부가적인 새로운 항암치료를 할 때까지의 시간

항체 의존 세포독성(antibody dependent cellular cytotoxicity, ADCC)

화학요법보호제(chemoprotectant)

항종양활성에 영향을 미치지 않으면서 정상조직에서 항종양화학물의 활성을 방해하는 물질

환자보고결과(patient-reported outcome, PRO)

## 부록 1. 무진행 생존(PFS) 또는 무질병 생존(DFS) 평가변수 고려사항

### 1. 서론

임상시험 평가변수로 무진행 생존(progression-free survival, PFS) 또는 무질병 생존(disease-free survival, DFS)을 사용하는 것에 대한 용어정의, 평가 빈도 및 방법, 발견비뒀림(ascertainment bias), 편차 데이터 및 누락 데이터의 처리 및 영상 검토(radiology review) 등과 관련한 고려사항을 제시하는 것이다. 1차 평가변수의 선택에 대한 가이드라인과 1차 평가변수로 PFS/DFS를 사용하는 것의 타당성이 다루어진다.

### 2. 평가변수 정의

PFS란 전통적으로 무작위배정(또는 비무작위 시험의 등록) 시점으로부터 객관적인 질병진행 또는 모든 사유로 인한 사망(이 중 빨리 도래한 시점)까지의 기간으로 정의된다. DFS란 무작위배정(또는 비무작위 시험의 등록)시점으로부터 객관적인 질병의 재발 또는 모든 사유로 인한 사망(이 중 빨리 도래한 시점)까지의 기간으로 정의된다. 진행과 재발의 시점은 위 기준이 충족되는 문서화된 증거가 있는 첫째 날을 사용하며, 이는 1회 또는 그 이상의 방문 누락, 치료의 중단 또는 새로운 항암치료 이후 진행이 관찰된 상태에도 마찬가지이다.

가능한 한 진행에 대한 용어 정의는 확립된 반응평가 기준(예, RECIST, WHO 기준, EBMT 기준, RANO)을 준수해야 하며, 비영상학적 기준에 대한 명확한 정의도 제공되어야 한다. 약물의 종류, 병변의 위치와 종류, 임상시험의 목적에 따라 적절히 변형된 기준이 특수 상황에 맞도록 있어야 하며 좀 더 적합한 기준이 고려되어야 한다. 예를 들어, 객관적인 임상적, 생화학적 혹은 영상학적 기준이 추가적으로 진행의 평가에 사용될 수 있으며, 모든 진행에 대한 기준이 객관적이어야 하며, 용어 정의는 임상시험계획서 내에 명확하게 미리 정의하는 것이 중요하다.

PFS가 상이한 사건의 복합으로 정의되므로(예, 새로운 병변, 기존 병변 또는 질환의 진행, 사망) 요약표를 사용하여 개별 유형의 사건에 대해 정의하는 것을 권고한다.

추적 방법에 지장을 줄 수 있는 특정 유형의 약물[예를 들어, 상이한 작용기전을 통해 영상에 나타나는 병변의 조영증강에 영향을 주는 약물, 또는 면역조절제 등과 같이 조영증강에 지장을 줄 뿐만 아니라 반응의 특정 패턴(예를 들어, 병변의 뚜렷한 진행 이후에 반응을 보이는 경우)을 보이는 항암제]의 경우, 다른 추적방법 또는 평가변수 등이 고려되어야 할 필요가 있다.

사건까지의 기간(time to event, TTE)은 통계적 분석을 위한 평가변수로 정의하기에 적절하다. 특정 시점에 특정 사건을 경험한 환자들의 비율에 근거한 다른 접근법들은 일부 경우에는 장점이

있을 수 있지만 한계가 있으며, 확정된 시점 접근법의 사용을 고려하고 있는 의뢰자는 규제기관의 과학 자문을 받을 것을 권고한다.

## 2.1. 데이터 확보

증례기록서(CRF)에 수집된 정보는 임상시험계획서에 전적으로 따라야 하며, 인체용 약품등록을 위한 기술적 요건의 조화를 위한 국제회의(International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, ICH) E9 가이드라인에 근거하여 모든 계획된 분석을 실행하기 위해 필요한 데이터에 초점이 맞춰져야 한다.

## 2.2. 간격확인 진행(Interval-detected progressions)

일반적으로 진행의 정확한 시점은 알려지지 않을 것이며, 대신 진행은 특정 시간 간격 중(예, 실제 질환이 진행한다면, 추적검사를 위한 2회의 방문 사이에 발생할 것이다)에 발생한 것으로 알려질 것이다. 일반적으로 1차 평가분석이 목적인 경우, 간격 중도절단(interval censoring)은 문제가 되지 않으며, 사건 발생 시점에 대한 분석은 재발이 확인된 시점으로 한다.

실질적인 진행의 관찰이 정해진 추적 방문시점에서(screen-detected progression)가 아니라 방문과 방문사이에서 이루어지는 경우(interval-detected progression), 예정된 평가 시점에서 벗어나서 평가를 하는 경우가 실제 임상에서 발생할 수 있다. 예정된 평가 시점에서 벗어나서 평가를 하는 경우가 실제 임상에서 발생할 수 있고, 실질적인 진행의 관찰이 정해진 추적 방문시점에서가 아니라 방문과 방문사이에서 이루어질 수 있다(간격확인 진행, interval-detected progression). 이러한 두 가지 종류의 진행이 혼재되어 있는 것은 치료효과 차이에 대한 잘못된 결론을 초래하는 발견비뒤틀림(detection bias)이 발생할 수 있기 때문에 우려스러운 부분이다. 예를 들어, 치료군 간의 독성 또는 증상 완화에 중요한 차이가 있는 경우, 진행은 더 높은 독성 또는 증상을 가진 치료군에서 더 조기에 관찰된다. 시험자들은 또한 시험군(experimental arm)에 유리한 비뒤틀림을 가지고 대조군(control arm) 환자들을 더 빈번하게 검사(또는 시험군 환자에 대한 평가 지연) 할 수 있다. 환자의 임상상태가 변화되어 질병상태의 예정되지 않은 평가를 초래하게 될 때, 적절하게 맹검이 되지 않은 임상시험은 발견비뒤틀림(ascertainment bias)의 위험이 있다.

예정되지 않은 평가로 인한 비뒤틀림의 문제는 적절한 시험 설계와 시험 수행으로 최소화되어야 한다. 임상시험은 가능한 경우 적절하게 맹검되어야 하고, 평가일정은 신중히 고려되어야 한다. 예정된 평가 간의 간격이 종양 진행까지 기간의 중앙값에 비해 상대적으로 짧은 경우, 간격사이에서 발견되는 진행이 감소할 것이므로, 예정에 없는 재발이 중대한 문제가 되지 않을 것이다.

예정된 평가시기 사이의 계획되지 않았던 평가에서 진행이 확인된다면, 1차 분석 목적에서 진



행일은 평가가 예정된 시점이 아니라 진행이 확인된 시점을 근거로 해야 한다. 이러한 접근법은 예정된 대로 분석(intention-to-treat, ITT) 원칙에 더 가깝고, 정보적 중도절단 가능성이 더 적다. 계획된 평가시간을 기반으로 한 대안적 분석과 간격 중도절단의 사용은 추가 분석(supportive analyses)에 포함되며, 치료군의 계획된 평가에 대한 순응도의 기술적 분석이 제공되어야 한다.

추측된 진행시간의 정확성을 유지하며 비뚤림을 최소화하는 것을 목적으로 추가 분석 내에서 추적관찰 패턴의 예측을 벗어난 차이를 어떻게 다룰지에 대해 다양한 접근법이 제안되었고, 이와 같은 분석에 대한 사전 계획을 정의하는 것을 고려해야 한다. 하지만 데이터 사후분석(*post hoc data analysis*)은 발견비뚤림을 보상하는데 있어 그 가치가 제한적이다.

### 2.3. 정보적 중도절단(informative censoring)

실제 시험에서 모든 무작위 대상자에 대한 PFS 사건의 관찰은 거의 불가능하므로 중도절단된 생존 데이터를 초래하게 된다. 집단 간의 생존기간 비교를 위해 일반적으로 사용되는 방법들은 중도절단이 실제 생존기간과 연관된 어떤 요인과도 관련이 없는 경우(즉, 중도절단이 “충분한 정보를 주지 않는” 것으로 알려진 경우)에만 유효하다. 반대로 충분한 정보를 주는 중도절단(informative censoring)은 치료 차이의 범위에 대한 부정확한 결론을 초래할 수 있다. 정보적 중도절단을 교정하는 만족할 만한 방식은 없으며, 이는 시험의 적절한 설계 및 수행을 통해 최소화해야 한다. 이와 유사한 고려사항이 OS 데이터 분석에 적용된다.

충분한 정보를 주지 않는(uninformative censoring) 중도절단을 가정하는 것은 일반적으로(분석을 위한 데이터종료 시점에 어떠한 사건도 관찰되지 않은 완전한 추적관찰을 받은 환자에 대한) “관리 중도절단(administrative censoring)”에 적용된다. 하지만, 이것은 분석을 위한 데이터종료 시점에 불완전한 추적관찰이 된 환자에서는 적용되지 않는다(조기 중도절단, premature censoring). 이러한 환자들에서 충분한 정보를 주지 않는 중도절단 가정은 표준생존 분석 접근법을 사용해 체계적으로 조사되어야 한다(예를 들어, Kaplan-Meier 곡선을 사용한 시간 대 중도절단 조사, 공변량 전체에 걸쳐 중도절단 패턴을 검사함, 종양부담과 기타 치료군 전체의 차이점 등과 같은 PFS 사건과 관련된 공변량과 중도절단 간의 탐색). 유용한 중도절단의 범위는 임상시험보고서에서 논의되어야 한다.

임상시험계획서의 계획된 치료를 잘 따르지 못하는 경우(non-compliance)가 일어날 수 있다. 그 예로는, 시험대상자가 잘못된 시험약을 받거나(또는 투여 되지 않거나), 예정된 종료 이전에 치료에서 중도탈락 하거나, 진행이 입증되기 전에 치료를 변경하는 것 등을 들 수 있다.

진행이 판정되기 전에 시험요법을 취소하는 사건은 정보를 제공하며, 통계적 분석에서는 이러한 사건을 중도절단 하는 것이 적절한가에 대해 항상 의문을 제기해야 한다. 모든 상황에서 최적

으로 이러한 문제를 처리할 수 있는 방법은 없지만, PFS/DFS의 1차분석을 위해 분석정의가 설정될 때에 가능한 한 배정된 대로 분석(ITT) 원칙이 준수되어야 한다. 특히, 모든 무작위 환자에서 결과 데이터(outcome data)는 의도된 평가일정에 따라 수집되어야 하며, 진행 또는 재발 일은 시험 위반 및 시험약의 중단 또는 치료법의 변경과 관계없이 객관적인 진행 또는 재발이 처음 발생한 시간에 근거해 배정되어야(assigned) 한다. 특정 시험에서 더욱 정확성을 기하기 위해 상이한 접근법이 고려되는 경우, 계획 단계에서 규제기관의 과학적 자문을 과학적 자문을 받을 것을 권고한다.

시험계획서에서 예상되더라도 시험요법을 취소한 환자에서 진행에 대한 신뢰할 만한 데이터를 얻는 것은 어려울 수 있다. 이러한 이유 및 기타의 이유로 중도절단 규칙을 포함한 누락된 데이터를 처리하기 위한 방법을 사전에 정하고 정당화할 필요가 있다. 이러한 방법들은 시험의 목적에 적합하고, 정보의 비뚤림 및 유실(loss)을 최소화하기 위해 선택되어야 한다. 이에는 PFS/DFS 분석에 있어서 사건으로서 정의된 진행/재발 전에 치료법의 취소 또는 변경을 고려한 방법들이 포함된다. 잠재적 비뚤림은 항상 언급되어야 하며, 민감도 분석은 다른 방법을 사용하여 수행되어야 한다. 보조 분석은 다음과 같은 것을 포함해 수행한다. 예를 들어, 질병 진행일을 예정된 방문일로 정의하거나, 간격 중도절단 분석, 질병 진행을 무작위배정 후 사전에 정한 특정 시간에 할당하는 단일시점(single time point) 분석, 중도절단된 환자를 사건에 대해 고위험 또는 저위험군으로 가정하고 민감도 분석을 수행하는 것들이 포함 될 수 있다.

## 2.4. 1차 분석과 민감도 분석

1차 평가분석의 목적을 위해서, 진행일은 진행이 문서로 확인된 시간을 기준으로 정해야 하며, 평가 예정 시간을 기준으로 정해서는 안 된다.

1차 평가분석을 위한 전략은 임상시험이 시작되기 전에 명확하게 기술되어야 한다. 통계적 분석계획을 적절히 고려하는 것이 중요하며, 이에는 시험 계획 단계에서 편차 및 누락 데이터의 처리 문제를 다루는 민감도 분석이 포함된다.

맹검시험에서 시험종료 시점에 맹검 검토(blind review)를 수행하는 것은 예측하지 못한 이슈들이 다루어 질 수 있도록 분석에 선택된 방법들로 다루어진 데이터와 분석 범위를 검토하게 하는 유용한 기회를 제공할 수 있다. 반면, 공개 시험(open-label trial) 종료 시점에서(데이터 세트에 무작위화 코드를 적용하기 전에) 검토가 수행될 수 있는데, 이에 따른 통계적 분석에 대한 수정(amendments) 또는 업데이트는 신중하게 검토되어야 한다. 그 이유는 그러한 것들이 데이터-주도(data-driven)라는 가능성을 배제하기 어렵기 때문이다. 이러한 시험들에서는 시험 계획서를 작성할 때 최선의 성실함이 요구되며, 축적된 데이터의 지식에 근거해서 분석의 중요 측면을 수정하는 경우 통계 분석 계획에 문제가 증가될 수 있음을 고려해야 한다. 이러한

데이터분석 이슈를 처리하고 문서화하는 방식은 임상시험을 위한 통계적 원칙에 대한 지침서에서 제공된 일반지침을 준수해야 한다(1998년 ICH). 명목상 맹검방식으로 수행되지만, 부작용 프로파일이 기능적 맹검해제(unblinding)를 초래하는 시험의 경우는 주의를 요한다.

현재, 규제 관점에서 볼 때 민감도 분석을 위해 권고되는 여러 가지 가능한 접근법이 있으며 (참조, 예를 들면 2007년 미국 FDA) 민감도 분석의 범위는 시험결과가 견고함(robust)을 입증할 수 있을 정도로 충분해야 하며 임상상황에 따라야 하고 관찰되는 임상데이터의 예상된 성격 등에 부합해야 한다(예, 환자의 중도탈락 패턴). 제시된 분석 범위에서 얻은 결론들 간에 차이가 있는 경우는 그에 대해 설명이 필요하며, 상이한 분석 및 분석 세트의 중요성은 또한 시험설계에 따라 정해져야 한다(우월성 또는 비열등성).

민감도 분석은 사용된 방법 내의 중요한 가정들을 다룰 수 있도록 설계되어야 하며, 이에는 편차 및 누락 데이터, 정보를 주지 않는 중도절단, 비례위험(proportional hazard) 등에 대한 취급 및 해당하는 경우, 일정에 없는 평가를 다루는 것 등이 포함된다.

민감도 분석은 임상시험계획서 또는 통계적 분석 내에 규정되어야 하며, 모든 변경사항은 시험 보고서 내에 정당화해야 한다.

## 2.5. 중간분석

중간분석은 안전성을 관찰하고 ‘무용성(futility)’을 평가하며 시험을 조기에 중단할 정도의 충분한 효능에 대한 증거가 있는지 여부를 고려하기 위해 종양 임상시험에서 통상적으로 사용된다. 중간분석의 시기, 목표 및 수행은 규제 지침에 부합하게 항상 정당화되어야 하지만, 일반적인 안전성에 대한 모니터링 측면과 무용성 평가측면에서 이견은 없다. 증명된 유효성에 대해 임상시험을 조기 종료하도록 설계된 중간분석은 더욱 논란(challenging)이 된다. 이러한 목적의 중간분석이 원칙적으로(in principle) 용인되는 경우, 1차 평가변수로서 PFS는 특별한 고려를 해야 하기 때문에 일반적으로 PFS에 대한 중간분석은 추천되지 않는다. 그럼에도 불구하고 중간분석이 필요하고 정당화 된 경우, 다음의 특별한 이슈들이 다루어져야 한다:

- 1) 데이터세트는 배정된 대로 분석(ITT) 시험집단과 특정 하위그룹에 대한 일관된(robust) 결론을 확보하는 데 충분한 완성도가 있어야 한다(내적 일관성(internal consistency)).
- 2) 종종 증거의 수준이 사용 가능하도록 하는 결과 요건을 갖춘 오직 하나의 확증(중추)시험만이 존재할 것이다.
- 3) OS, 안전성 및 기타 2차 평가변수 등에 대한 자료는 유익성-위해성에 대한 규제적 결정을 위해서는 완전하지 않거나 충분하지 못할 수 있다.

따라서 중간분석의 시기 및 목표가 PFS에서 통계적 중요성 확인만을 고려하여 계획되어서는 안 된다. 적절한 정당화에는 PFS 및 OS 데이터의 성숙도 및 하위그룹에서의 이용 가능한 안전성과 2차 평가변수의 증거가 포함되어야 한다(‘진행 후 추적관찰 및 치료’ 참조). 이러한 고려로 인해 특정 중간분석은 비실용적일(impractical) 수 있다.

### 3. 평가의 빈도 및 방법

종양평가의 방법 및 빈도는 치료 주기가 서로 다른 경우라도 전체 시험 군에서 동일해야 한다.

PFS 평가는 특정 종양에 가능한 모든 병소가 치료 전에 평가되어야 하며, 침범 부위는 다른 부위와 함께 임상적 및 영상학적으로 적용된 동일한 방법을 사용하여 추적관찰 기간에도 체계적으로 평가되어야 한다. 마찬가지로, DFS의 평가는 유사 질병 부위도 추적관찰 동안 체계적으로 평가하는 것이 요구된다. 특정 약물이 전이 패턴에 영향을 주는 것으로 의심할 만한 사유가 있는 경우, 체계적 평가를 위해 추가적인 관련 부위를 사전에 명시해야 한다.

따라서 평가빈도는 예측된 치료효과를 탐지할 정도로 적절해야 한다. 평가시점과 진행평가를 위한 최적의 빈도는 시험 별로 결정해야 하며, 시험의 목적 및 치료일정, 질병진행의 특정 패턴 등을 고려해야 한다. 예를 들어, PFS의 중앙값이 짧은 경우, 최초의 방문은 적시에 이루어져야 하며, 진행을 정확히 평가할 필요성과 환자를 침습적이며 자원이 많이 소모되는 진단 절차에 노출시키는 것을 최소화 할 필요성 사이에 균형을 유지해야 한다. 방문 빈도가 증가되면 예정된 방문들 사이의(계획되지 않은) 평가의 가능성을 감소시키는 반면, 방문 주기가 일상적인 진료행위와 일치되지 않는 경우, 계획된 방문 전체를 놓칠 가능성이 높아질 것이다. 통계적인 관점에서 보면 평가 빈도의 증가는 질병의 진행률이 매우 빠른 경우가 아닌 한 통계적 검정력을 증가시키지 않는다.

시험계획서에 정의된 일정을 준수하는 것은 필수적이며 일정으로부터 벗어나는 경우는 보고해야 한다. 방문일정에 대한 순응도는 분석 시에 서술적으로 조사되어야 하며 시험결과에 대한 영향도 탐색되어야 한다.

### 4. 맹검된 독립적 중앙평가(blinded independent central review, BICR)

다수의 병소가 추적되어야 하는 진행성 질환에서 특히 진행에 대한 평가는 측정오차(measurement error)가 발생할 수 있다. 일반적으로 이러한 측정오차를 최소화하기 위한 노력을 기울여야 한다. 오차를 범하지 않기 위한 모든 합리적인 노력을 했음에도 불구하고 중대한 측정 오차가 발생할 수 있다. 하지만 규제적 측면에서는 측정오차가 전체 치료군에 균등하게 발생하고 치료효과가 충분히 큰 경우라면 이는 주요 문제가 되지 않을 수 있다. 하지만 측정오차

가 전체 치료군 간에 상이한 경우, 이는 분석결과를 해석하는데 어려움을 초래할 수 있다. 마찬가지로 만약 치료효과가 적거나 미미한 경우, 심한 측정오차는 유익성-위해성 평가에 지장을 줄 수 있으며, 따라서 적절한 방법표준화 및 현지 평가(local evaluation)를 수행하는 시험자들에 대한 교육을 통해 측정오차를 최소화하기 위한 모든 합리적인 노력을 기울여야 한다.

시험자들의 질병진행에 대한 평가는 특정 치료군에 유리하도록 하는 시스템적 비뚤림을 유발하여 부정확한 치료군 간 비교결과를 초래할 수 있다. 이러한 이유로 가능한 한 언제라도 적절한 맹검(masking techniques)이 사용되어야 한다. 시험자 비뚤림은 적절한 이중맹검 무작위시험에서는 일반적으로 이슈가 되지 않는다. 하지만 상이한 약물들의 특정한 효과로 인해 맹검을 할 때 항암제 시험들은 문제가 많다. 사실 치료들의 상이한 독성 프로파일로 인해 종종 치료군 배정을 완벽하게 은폐하는 것은 불가능 할 수 있으며, 최선의 치료와 비교하는 시험은 특히 이런 종류의 비뚤림의 위험이 있다.

이러한 비뚤림을 파악하고 감소시키기 위한 하나의 전략은 모든 환자에 대해 모든 관련 데이터의 완전한 BICR을 수행하는 것이다. 이러한 전략은 중요한 시험자 비뚤림이 예상되거나 보통의 치료효과가 기대되는 경우에 권고된다. 사건의 대다수가 임상적 진행에 근거하는 것이 아니라 영상에 근거해서 발견되는 경우, BICR은 더욱 그 의미가 있을 것이다.

하지만 중대한 시험자 비뚤림이 있는 경우, 진행에 대한 완전한 BICR도 정보적 중도절단을 막을 수 없는 경우가 발생한다. 왜냐하면 현지(local)에서 평가된 진행 시점에 환자가 시험계획서에서 탈락되면, 이 시점 이후로 더 이상의 가용한 검사실, 영상 또는 임상 평가 데이터가 없는 경우가 있기 때문이다. 이러한 점을 제거하는 방법 중의 하나는 실시간 BICR을 수행하는 것이고, 또 다른 방법은 현지에서 진행이 확인된 이후 추가 영상을 수집하는 것이다. 하지만 환자들이 진행 이후의 추가 추적에서 누락될 수 있기 때문에 이는 비현실적일 수 있다.

현지 시험자 평가(local investigator assessment)의 비뚤림은 시험자 평가와 BICR간의 불일치 측면에서 조사될 수 있다. 이러한 목적을 위한 통계에는 초기 불일치율(early discrepancy rate, EDR)과 후기 불일치율(late discrepancy rate, LDR)이 포함되며, 이들은 각각 현지 평가에서 BICR에 비해 조기 또는 후기에 진행을 선언하는 빈도에 근거한다.

특정 상황에서 완전한 BICR을 하지 않는 가능한 전략 중 하나는 샘플에만 근거하여 BICR을 수행하는 것이다(점검(audit)). 만약 불일치통계에 대한 검토를 통해 현지 시험자 평가에 비뚤림이 없다는 것이 입증되는 경우, 완전한 BICR을 안할 수 있다. 비뚤림이 점검(audit)에 근거해 배제될 수 없는 경우, 완전한 BICR은 다른 분석을 위한 근거를 제공하기 위해 추후에 실행될 수 있다. 다른 전략에는 맹검된 현지 또는 국가-특정(country-specific)의 영상 검사가 포함된다. 현재 이러한 접근법의 현실적인 실행에 대한 광범위한 경험이 없으므로 이러한 접근법의 적절

성에 대한 규제적 지침은 특히 시험의 통일성 확보, 샘플이 생성되는 방식, 중요한 방향적 불일치가 배제될 수 있는지 여부를 판단하기 위해 사용되어야 하는 통계 및 지표(metrics) 등에 대한 논의 시행 전에 사례별(case-by-case)로 마련되어야 한다.

일반적으로 완벽한 BICR이 가능한 경우, 1차 분석은 독립적인 평가를 통해 얻어진 결과에 근거해 계획될 수 있다. 중대한 시험자 비몰림이 합리적으로 배제될 수 있는 경우, 시험자 평가는 1차 분석을 위해 사용될 수 있도록 설계될 수 있다. 전략과 상관없이 BICR을 통해 정해진 결과의 역할 및 BICR의 범위와 관련된 결정 규칙은 임상시험계획서에 사전에 명시해야 한다.

일반적으로 시험의 질에 대한 확신은 BICR로부터 얻은 시험결과가 어떤 중요 정도에서도 시험자 평가와 다르지 않은 경우 증가한다.

독립적 심사를 위한 절차는 전향적으로 정의되어야 하며, 임상시험 기본문서에 기술되어야 한다.

## 5. 효과의 크기

효과의 크기는 준-모수절차(semi-parametric procedure)에 근거하여 PFS의 생존함수(survivor function)에 대한 추정치를 도면에 그래프로 표현하고 위험률의 추정, 사건까지의 중앙시간(median time-to-event)(예, 상위 사분위 및 하위 사분위) 및 특정 시간대에 무사건 환자 비율의 추정(예, 1년째 무사건 환자 %) 등에 의해 정량화되어야 한다. 임상적 관점에서 PFS 중앙값은 PFS 기간의 분포의 위치에 대한 요약지표(summary measure)로 고려되어야 하지만 중앙값들(medians)의 차이에 지나치게 의존하는 것은 피해야 한다. 그 이유는 전체적으로 생존 곡선을 고려하는 것에 비해 일반적으로 유용하지 않기 때문이다. 어떠한 경우이든 요약지표의 선택은 정당화되어야 하며 사전에 명시되어야 한다.

생존함수에 대한 비모수추정은 계단함수이므로, 모수적 비례위험모델링(parametric proportional hazards modelling) 및 기타 “매끄럽게 하는(smoothing)” 기술들은 탐색분석에 있어 유용할 수 있다.

PFS에서 확인하고자 하는 임상적으로 유의한 차이(clinically relevant difference)는 종종 사소한 이슈가 아니다. 그 이유는 이러한 평가변수가 대개 환자에게 임상적으로 직접적 관련이 없는 검사실적, 영상학적 또는 임상적 평가에 의해 이루어질 수 있기 때문이다. 이러한 차이는 임상적 및 역학적 근거에 기반하여 전향적으로 정당화되어야 할 필요가 있으며, 이는 OS에 대한 대리성 입증, 증상관련 기대효과, 치료변경, 감정적 영향, 또는 건강관련 삶의 질 등이 포함된다.

분석시점에 종합적인 PFS 평가변수의 다른 구성요소(예, 새로운 병소, 크기 증가, 사망 또는

원인관련 사망) 및 OS, 증상조절, 건강관련 삶의 질(적절한 경우) 등과 같은 추가적인 평가변수들에 대해 보조적으로 기술된 요약표(supportive descriptive summary tables)와 대조적 위험분석(competing risk analyses)을 이용할 수 있어야 한다.

## 6. 진행 후 추적관찰 및 치료

PFS에서의 큰 효과는 일반적으로 OS에 대한 효과와도 관련이 있는 것으로 예측되므로 만약 이에 해당되지 않는다면 합리적인 설명을 제공해야 한다.

PFS의 관점에서 치료를 비교할 때, 시험약으로 치료하는 것이 PFS에 이득이 있더라도 OS에 대해서는 더 나쁜 결과를 보일 수 있음을 고려해야 한다. 이는 장기적 독성, 진행 이후 사용된 치료에 대한 상이한 내성 프로파일 또는 증가된 전이 가능성을 초래하는 생물학적 변화 등에 기인할 수 있을 것이다.

따라서 PFS가 1차 평가변수인 경우 OS에 부정적 영향이 없음을 충분히 확인할 수 있도록 가능한 모든 환자들을 사망 시까지 추적관찰 해야 한다(가이드라인 본문 참조). 후속 치료를 하는 경우, 혹은 특히 후속 치료의 효과에 대한 확신이 없는 경우, 객관적 반응을 측면에서 후속 치료의 결과, 후속 치료 후 무진행생존(PFS2) 등도 실제적 측면에서 이용가능 해야 한다. PFS2는 무작위배정(또는 비무작위 시험에서 등록) 시점으로부터 두 번째 객관적 질병진행 또는 사유를 불문한 사망까지의 시점 중 더 빨리 도래하는 시점까지의 기간으로 정의된다. 생존해 있는 환자 및 두 번째 객관적 질병진행이 관찰되지 않은 환자는 생존해 있는 것으로 알려진 시점과 두 번째 객관적 질병진행이 관찰되지 않은 시점에서 중도절단 된다. OS와 PFS2가 신뢰할 만큼 결정될 수 없는 경우, 후속 치료 종료 시의 결과를 살펴봄으로써 추가 치료의 유의한 효과가 없다는 점을 배제할 수 있다. 이러한 분석을 위해 사건은 후속 치료의 종료 또는 중단, 두 번째 질병진행 또는 모든 사유에 의한 사망 중 더 빨리 도래하는 사건으로 정의된다.

진행 후 시험군으로의 편도교차(one-way cross-over)는 OS와 다른 장기간(long-term) 2차 평가변수에 대한 기타 후속 비교에 방해를 줄 것이다. 따라서 이러한 종류의 교차는 시험의 객관성을 충족시키기 위해서 일반적으로 회피되어야 한다. 그럼에도 불구하고 교차가 필요하다고 판단되는 경우 PFS, OS 및 기타 중요한 2차 평가변수 관련 이용 가능한 데이터가 과학적 및 규제적 측면에서 시험목적에 충족시키고, 적절한 결론이 도출 될 수 있을 정도로 충분한 신뢰가 있어야 한다. 이러한 상황에서 OS 분석은 계획된 2차 분석 또는 계획된 공동-1차 분석(co-primary analysis)에 근거할 수 있다.

## 부록 2. 항암제 마스터 프로토콜 임상시험

### 1. 서론

환자 개개인에게 맞춤형 치료를 제공하는 정밀의료(Precision medicine)의 개념이 항암제 임상 시험에 도입되면서, 시험대상자의 선정 단계에서 생체표지자(Biomarker) 또는 특정 유전자 변이 여부를 확인하고 시험약물에 좋은 반응을 보일 수 있는 환자군을 선발하는 경우가 증가하고 있다. 그러나 적절한 생체표지자 또는 유전자 변이를 갖는 시험대상자 수는 매우 제한적이므로 과거 전통적인 임상시험과는 다른 적응형 설계(Adaptive design)가 포함된 ‘마스터 프로토콜’이 도입되었다. 이와 같은 새로운 형태의 임상시험은 시험 결과의 질적인 향상과 함께 시험기간의 단축을 통해 새로운 의약품의 개발을 가속화한다.

본 가이드라인은 최신 유형의 임상시험 설계에 대한 일반적인 안내와 함께 마스터 프로토콜 임상시험의 설계 및 수행 시 고려사항을 포함하고 있다.

의약품의 제품화를 가속화하기 위해 확장 코호트를 사용하여 사람에게 처음 투여하는 (First-In-Human, FIH) 또는 초기 단계(Early stage)의 임상시험은 본 가이드라인에 포함되지 않는다.

### 2. 정의

‘마스터 프로토콜 임상시험’은 전체 시험 구조 안에서 하나 이상의 시험약물 및/또는 하나 이상의 질병 환자군에의 안전성 및 유효성을 동시에 평가하기 위해 설계된 임상시험을 의미한다.

과거 개별 요법에 대해 독립적인 시험을 수행하는 전통적인 임상 프로토콜의 패러다임 하에서는 시험대상자가 하나의 프로토콜에 모집 및 선별된 후 일부 선정/제외기준을 충족하지 못하면 다른 시험에의 참여 기회도 놓칠 수 있다. 또한 각각의 임상시험에 대해 별도로 데이터가 수집 및 시험되는 과정이 반복되기 때문에 한명의 시험대상자에 대해 중복된 정보가 수집되지만 서로 공유되지는 않는다는 문제가 있었다.

그러나 ‘마스터 프로토콜’은 공통 스크리닝 플랫폼을 사용하여 시험대상자가 자격이 있는 모든 시험을 식별하는 것이 핵심이자 주요 장점이다. 아울러 데이터의 중앙 집중화를 통해 더 적은 자원을 활용하되 다양한 측면에서 균일한 결정을 함으로서 데이터의 품질이 향상된다.



### 3. 마스터 프로토콜 임상시험의 유형

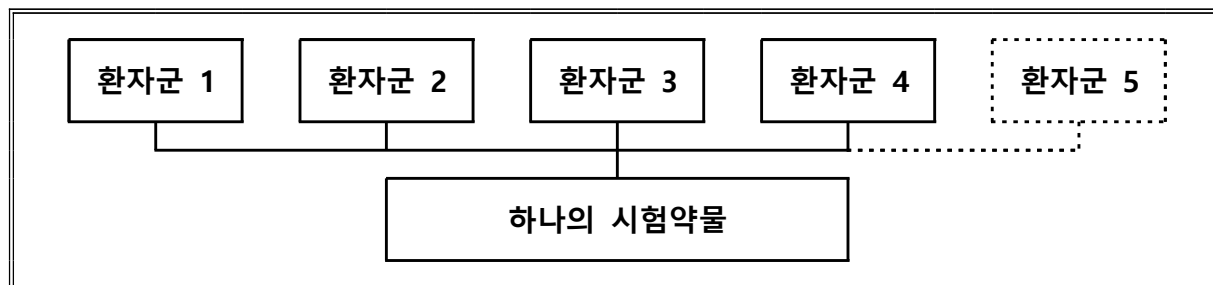
이 가이드라인에서는 대표적인 마스터 프로토콜 임상시험 유형으로 아래의 3가지를 소개한다.

#### 3.1. 바구니형 임상시험(Basket trial)

여러 질병 또는 단일 질병의 세부유형(subtype)에서 하나의 시험약물을 연구하기 위해 설계된 임상시험을 말한다. 동 시험에는 질환의 종류, 병기, 조직학적 유형, 이전 치료 횟수, 유전자 또는 생체표지자 특징이 서로 다른 다양한 환자군이 모두 포함된다.

일반적으로 1차 평가변수는 전체 반응률(Overall Response Rate, ORR)로 설정하며, 개별 환자군에서의 효과를 별도로 관찰한다. 만약 하위 연구에서 강력한 실마리 정보가 관찰되면 해당 환자군의 연구를 확장하는 것을 고려할 수 있다. 다만 각각의 환자군은 구체적인 목표와 함께 해당 환자군에서의 과학적인 근거, 표본 크기에 대한 타당성을 확보해야 하며, 무용성(Futility) 및 유효성(Efficacy) 평가를 위한 중단 규칙을 포함하는 상세한 통계 분석 계획(Statistical Analysis Plan, SAP)을 사전에 설정해야 한다.

그림 1. 바구니형 설계가 적용된 마스터 프로토콜 임상시험 예시



\* 환자군은 다양한 질환의 아형에서 프로토콜의 정의에 따른 하위군으로 나뉘며, 점선으로 표시된 환자군 5는 시험의 진행과 동시에 추가적인 하위군을 추가·변경할 수 있음을 의미함

#### 3.2. 우산형 임상시험(Umbrella trial)

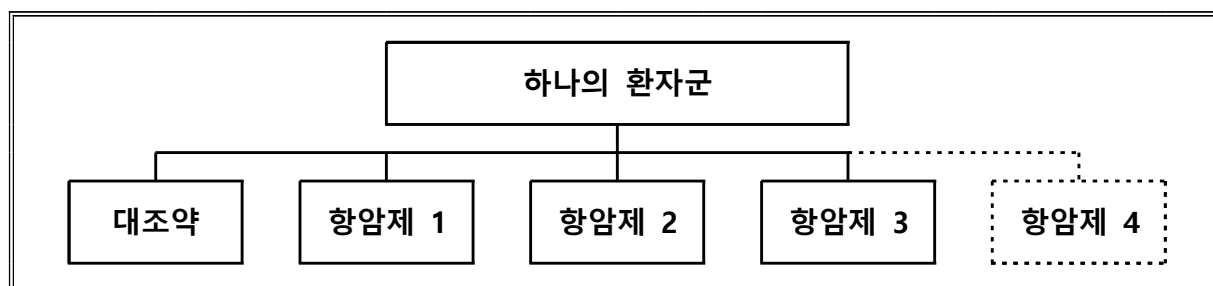
단일 질병의 맥락에서 다양한 시험약물을 연구하기 위해 설계된 임상시험을 말한다. 동 시험에는 전통적으로 정의된 하나의 환자군(단일 암종 또는 하나의 조직학적 아형)이 포함된다.

일반적으로 공통의 대조군으로서 표준 치료요법을 설정한 후 시험에 진입하며, 따라서 무작위 배정으로 표준 치료요법과의 비교가 가능한 특징이 있다. 대조군은 임상시험의 결과를 현재의

의료 환경에서 해석이 가능하도록 최신의 표준 치료요법을 반영해야 하며, 별도의 군을 추가로 설정할 수도 있다. 윤리적인 측면에서도 이전 표준 치료요법에 환자를 무작위 배정 하는 것은 바람직하지 않다.

유효성의 비교 분석은 시험약물과 공통 대조군 간에 수행되며, 시험약물 간의 비교는 권장되지 않는다.

그림 2. 우산형 설계가 적용된 마스터 프로토콜 임상시험 예시



\* 프로토콜에서 정의된 단일 질환의 환자군에 대해 다양한 항암제가 투여되며, 점선으로 표시된 항암제 4는 시험의 진행과 동시에 추가적인 시험약물을 추가·변경할 수 있음을 의미함

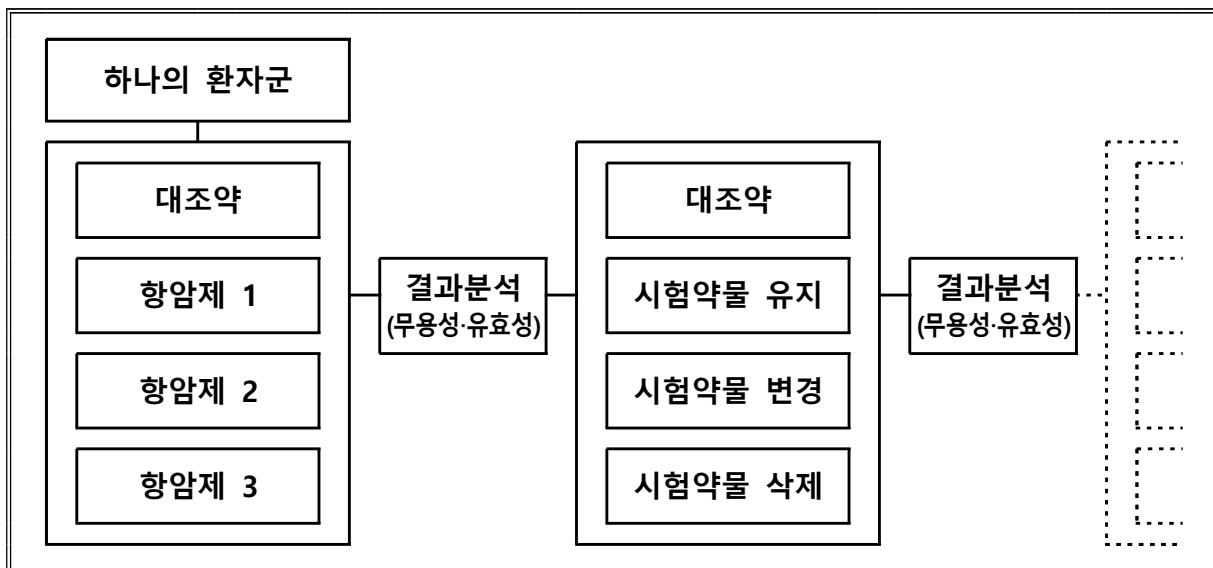
### 3.3. 플랫폼 임상시험(Platform trial)

단일 질병의 맥락에서 의사결정 알고리즘을 기반으로 다양한 약물을 도입 또는 삭제가 허용되는 임상시험을 말한다.

동 설계는 전통적으로 정의된 하나의 환자군이 포함된다는 면에서 우산형 임상시험과 유사하나, 임상시험의 초기에 생성되는 정보에 따라 향후 임상시험의 진행이 조정된다는 차이점이 있다.

임상시험의 고정된 종료일이 없이 시간이 지남에 따라 시험이 진행되고 그 결과에 따라 계층이 추가 또는 삭제될 수 있다. 예를 들어 특정 약물이 유효성 기준을 충족하면 대조약(표준 치료요법)에 포함된 환자들이 해당 약물로 대체될 수 있으며, 무용성 분석 결과에서 치료적 이점이 낮다고 판단되는 경우 해당 치료에 배정된 환자들은 시험이 중단되거나 새로운 시험약물 또는 대조약 군으로 전환될 수 있다.

그림 3. 플랫폼 설계가 적용된 마스터 프로토콜 임상시험 예시



\* 임상시험에 등록된 환자에서의 결과를 통해 환자군을 확장·변경하거나 시험약물을 유지·삭제 및/또는 새로운 약물로 교체 등이 가능함

#### 4. 마스터 프로토콜의 설계 시 고려사항

일반적으로 마스터 프로토콜의 설계 이전에 시험약물 및/또는 대상 환자군에 대한 2상 권장용량 (Recommended Phase 2 Dose, RP2D)이 결정되어야 한다. 설계의 특성상 다양한 환자군 또는 시험약물이 동시에 시험에 포함되므로 해당 범위를 모두 포괄하는 안전성, 약물학, 비임상시험 자료 등이 제출되어야 한다.

##### 4.1. 새로운 병용요법

두 개 이상 시험약물의 새로운(novel) 조합의 경우 병용요법에 대한 강력한 과학적 증거와 함께 개별 약물이 항종양 활성이 있는 경우 개별약물에 대한 RP2D가 설정되었는지 확인해야 한다. 만약 RP2D가 설정되지 않았을 경우에는 해당 내용을 포함하는 시험의 구성(Dose-finding)도 가능하며, 이와 같이 마스터 프로토콜 내에 용량 결정 요소가 포함되는 경우에는 효능 평가를 진행하기 전에 병용요법에 대해 안전성 평가를 위한 도입군으로서 최소 6명의 안전성 정보를 활용하는 것이 바람직하다.

##### 4.2. 생체표지자를 표적으로 하는 시험약물

생체표지자의 분석 방법은 타당한 밸리데이션으로 사전에 검증되어야 한다. 설정된 분석 방법이

부적절한 경우(예: 낮은 정확성 및 정밀성) 시험약물의 안전성 및 유효성 결과에 대한 신뢰성에 영향을 준다. 만약 분석 방법이 국내에 허가되지 않았거나 허가된 사용목적 외의 범위인 경우 마스터 프로토콜의 설계 전에 관련 부서와 사전에 협의해야 한다.

표적하는 생체표지자와 시험약물의 작용기전을 바탕으로 치료 반응을 예측할 수 있는 환자군에 대한 과학적 근거자료를 제출해야 한다. 특히 하나 이상의 생체표지자를 보유한 시험대상자를 하위군에 어떻게 배정할지에 대해 사전에 명확하고 자세하게 기술해야 하며 해당 근거자료를 반드시 구비해야 한다. 이때는 동시에 발현되는 동반 유전자 돌연변이의 임상적 연관성을 고려할 수도 있고, 낮은 빈도의 생체표지자를 가진 환자를 우선으로 배정할 수도 있다.

### 4.3. 통계학적 고려사항

1차 평가변수가 객관적 반응률인 비무작위 프로토콜의 경우, 계획된 표본 크기는 전체 반응률의 95% 신뢰구간을 하한으로 임상적으로 중요하지 않은 반응을 배제할 수 있도록 충분해야 한다. 분석 계획에는 무용성 분석을 포함하는 것이 권장되며, 효과가 없는 약물에 대한 노출은 최대한 제한해야 한다.

마스터 프로토콜은 다양한 시험약물을 동시에 평가하므로 임상시험을 시작하기 전에 적절한 중간 분석을 설정하고, 중간 분석의 결과 또는 새로운 외부 데이터의 근거 수준에 따라 어떻게 시험약물을 추가·삭제하거나 환자군을 확장 또는 중단할 것인지를 사전에 계획해야 한다.

## 5. 마스터 프로토콜의 국내 임상시험 승인 사례

최근 다양한 사례의 마스터 프로토콜 임상시험이 국내 승인되었다. 이 가이드라인에서 소개하고 있는 마스터 프로토콜 임상시험의 각 유형 별로 국내 승인 사례를 소개한다.

## 1. 바구니형 임상시험 국내 승인 사례(1)

(시험제목) X 유전자 변이가 있고 이전에 치료 경험이 있는 국소 진행성 절제 불가능 또는 전이성 고형암 시험대상자를 대상으로 한 신물질 A 및 기허가 의약품 B의 병용요법에 대한 제2상 바구니형 임상시험

- (배경) 동 임상시험의 시험대상자는 종양 조직학 및 X 유전자 변이 상태에 기반하여 각각 별도의 코호트에 등록되며, 1단계로 최대 약 12명의 반응 평가 가능 시험대상자가 등록됨. 1단계 중 특정 코호트에서 유의미한 항종양 반응이 관찰되는 경우, 해당 암종 및 X 유전자 변이 유형에서 시험 요법의 활성 및 안전성을 추가로 규명하기 위해 해당 코호트에 반응 평가 가능 시험대상자를 최대 총 30명까지 추가 등록함(2단계 확장)
- X 유전자는 이미 다수의 고형암에서 검증된 바 있는 항종양 표적으로, X 유전자를 타겟으로 하는 다수의 생물학적제제 및 소분자 약물 등이 이미 품목허가 된 바 있음. 기허가 의약품 B의 경우 항-X 생물학적제제로서 X 유전자 변이 고형암 환자의 표준 치료요법 중 하나임

### ■ (설계)

종양 유형	X 유전자		시험대상자 수	
	과다발현(증폭)	돌연변이	1단계	2단계 확장
자궁암	코호트 1		12	30
위암	코호트 2		12	30
비소세포폐암	코호트 3	코호트 4	12	30
직결장암	코호트 5		12	30
담도암	코호트 6		12	30
기타 고형암	코호트 7	코호트 8	30	-
선택적 추가 종양*	코호트 9	코호트 10	12	30

\* 기타 고형암 코호트(7 또는 8)에 특정 종양 유형을 가진 충분한 수의 시험대상자가 등록되는 경우, 해당 종양 유형을 선택적 추가 종양 코호트(9 또는 10)로 변경하여 별도로 평가 가능

- (시험방법) 신물질 A 및 기허가 의약품 B의 병용투여 방법은 고정되어 있으며, 종양 치료 반응은 RECIST v1.1에 따라 평가함
- 베이지안 접근법을 이용하여 무익성 기준이 설정되어있으며, 중간분석 결과 반응률(ORR)이 < 20%인 경우 현행 표준치료의 반응률 보다 더욱 양호할 가능성이 없다고 판단함(1단계 12명 중 최소 3명의 확인된 반응 필요)

## 1. 바구니형 임상시험 국내 승인 사례(2)

(시험제목) 희귀 난치성 고형암 환자를 대상으로 질병 예측 생체 표지자를 평가하고 약물 A의 치료 반응을 확인하기 위한 제2상 바구니형 임상시험

- (배경) 현재 이용 가능한 모든 표준 치료요법에도 불구하고 질병이 진행되었거나 불내성을 보이거나, 또는 표준 치료요법이 없는 희귀 난치성 고형암에 대하여 생체 표지자의 종류와 무관하게 암종 별로 최대 50명의 시험대상자가 등록됨(1단계). 중간분석에서 생체 표지자 종류에 따라 반응률의 경향이 관찰되는 경우, 해당 생체 표지자 코호트는 최대 50명까지 추가 등록함(2단계 확장)
- 약물 A는 대표적인 항암화학요법제로서 특정 생체 표지자를 타겟으로 하지 않는 기허가 의약품임
- 모든 시험대상자는 등록 시점에 반드시 생체 표지자 확인을 위한 종양 검체를 제출해야 함
- (설계)

종양 유형	생체 표지자의 종류와 무관	생체 표지자의 보강 확장*	시험대상자 수	
			1단계	2단계 확장
희귀암 1	코호트 1		50	
희귀암 2	코호트 2		50	
희귀암 3	코호트 3		50	
희귀암 4	코호트 4	코호트 8	50	50
희귀암 5	코호트 5		50	
희귀암 6	코호트 6		50	
희귀암 7	코호트 7	코호트 9	50	50

\* 코호트 4 또는 7에 특정 생체 표지자 종류가 있는 시험대상자가 다수 등록되고 해당 코호트 환자에서 통계적으로 유의미한 유효성 경향이 관찰되는 경우, 해당 생체 표지자가 있는 환자를 추가 코호트 8 또는 9로 확장하여 등록 및 평가 가능

- (시험방법) 약물 A의 투여 방법은 고정되어 있으며, 종양 치료 반응은 RECIST v1.1에 따라 평가함
- 일차 유효성 평가변수는 객관적 반응률(ORR)이며, 중간분석과 별도로 무용성을 관찰하고 약물 A의 투여에서 관찰되는 종양 반응 패턴을 확인하여 개별 환자의 치료 결정을 내릴 목적으로 면역 관련 RECIST(irRECIST)를 사용할 수 있음

## 2. 우산형 임상시험 국내 승인 사례

(시험제목) 전이성 Y 암 환자에서 신물질 A와 다양한 기허가 의약품의 병용요법에 대한 안전성 및 유효성을 평가하기 위한 1b/2상, 라벨 공개, 우산형 임상시험

- (배경) 신물질 A는 단독요법으로서 동 임상시험의 시험대상자인 전이성 Y 암에 대한 2차 요법이 품목허가 된 바 있음. 동 임상시험에서 신물질 A와의 병용요법을 탐색하고자 하는 기허가 의약품은 총 5가지로, 단독요법으로서 전이성 Y 암에 대한 2차 요법이 품목허가 되었으나 각각의 의약품은 모두 다른 항암 작용기전을 가지고 있음. 따라서 동 임상시험 중 1b상 파트를 통해 병용요법(신물질 A의 부가요법)의 2상 권장용량(RP2D)을 결정하고, 2상 파트를 통해 안전성 및 유효성을 평가하고자 함
- 신물질 A는 이전 임상시험에서 단독요법으로서 RP2D가 결정된 바 있으며, 전이성 Y 암 환자에서 2차 치료제로서 표준 치료요법 대비 우월성이 확인됨

### ■ (설계)

1b상 (N=6)	코호트 1	→	코호트 2	→	코호트 3
용량 증량	신물질 A 80mg + 병용약물 저용량	DLT ≤33%	신물질 A 100mg + 병용약물 저용량	DLT ≤33%	신물질 A 100mg + 병용약물 상용량
* 신물질 A는 단독요법으로 상용량이 100mg이며, 이상반응 발생 시 80mg로 감량함					
** 병용약물의 저용량은 국내 허가사항에 따른 상용량 대비 1단계 감량된 용량을 의미함					

- 환자의 선정/제외기준 및 용량 제한 독성(DLT)의 기준 등은 신물질 A와 병용하는 약물의 작용기전 및 해당 약물의 단독요법에서 관찰된 이상반응의 프로파일 등을 고려하여 별도로 설정함
- 2상 파트의 경우 개별 병용약물에 따라 결정된 RP2D 용량으로 수행되며 총 5개의 코호트 별로 최소 30명, 최대 50명의 환자를 평가할 계획임
- (시험방법) 신물질 A 및 병용약물의 투여 방법은 가변적이며 공통 대조군(표준 치료요법)은 설정되지 않았음
- 1차 유효성 평가변수의 경우 1b상 파트에서는 DLT가 관찰된 시험대상자의 수로 설정하고, 2상 파트에서는 6개월 시점에서의 PFS 비율로 설정함
- 1b상 파트에서 각 병용요법에 대한 개별 코호트의 모든 시험대상자가 DLT 관찰 기간(제1주기, 28일)을 완료한 후 모든 가용 안전성 및 PK 자료를 검토하고, 다음 코호트 또는 2상 파트의 진행 여부를 결정함

### 3. 플랫폼형 임상시험 국내 승인 사례

(시험제목) 진행성 Z 암 환자에서 면역요법 기반의 병용요법을 평가하는 제2상, 탐색적, 라벨 공개, 플랫폼형 임상시험

- (배경) 동 임상시험의 시험대상자는 진행성 Z 암 환자로서, 1차 표준 치료요법으로 면역 관문 억제제가 확립되어 있음. 그러나 장기간의 관해 유지 또는 생존을 보이는 환자가 매우 드물고 대부분의 환자가 질병의 진행을 경험함에 따라 새로운 병용요법을 탐색하고자 함
- 현재 동 임상시험에서 탐색하고자 하는 병용약물은 총 3가지로, 별도의 임상시험을 통해 병용요법의 2상 권장용량(RP2D)이 확인된 바 있음

#### ■ (설계)

	대상 환자군	예비단계 (N=약 20)	확장단계 (N=약 50)
하위시험 A	이전 치료경험이 없고 치료 가능한 유전체 변이가 없는 환자(1L)	대조군	예비단계 유효성 후
		시험군 1	
		시험군 2	
		:	
하위시험 B	이전 치료경험을 받은 후 질병이 재발된 환자로 치료 가능한 유전체가 없거나 또는 있는 경우 해당 치료를 1가지 이상 완료한 환자(2L+)	대조군	예비단계 유효성 후
		시험군 1	
		시험군 2	
		:	
* 하위시험의 경우 각각의 하위연구별 계획서를 별도로 수립함. 여러가지 병용요법에 대한 효율적인 시험이 가능하도록 향후 하위연구 또는 투여군이 추가 가능함			

- 1개의 하위연구에 등록되어 시험에서 투여를 받는 동안 질병의 진행, 임상적 이익의 상실 또는 허용 불가능한 독성을 경험하는 환자의 경우, 하위연구별 시험계획서의 적합성 기준에 부합하는지 여부를 재검토하여 플랫폼 시험 내의 다른 하위연구에 등록 가능함

#### ■ (시험방법) 일차 유효성 평가변수는 RECIST v1.1에 따라 평가한 객관적 반응률(ORR)임

- 개별 하위연구에 따른 표본 크기 및 목표 반응률이 별도로 설정되어 있으며, 중간분석 결과 또는 표준 치료요법의 변화 등 새로운 유효성 정보에 따라 변경될 수 있음



## 항암제 임상시험 가이드라인[민원인 안내서]

발 행 일 2023 월 11 월

발 행 인 박 윤 주

편집위원장 김 영 림

편 집 위 원 (의약품심사부 종양항생약품과)

홍정희, 백주현, 박소라, 김지명, 서정원, 구민지, 이소라, 김경아, 이주협

도움주신 분 식품의약품안전평가원 임상연구과

발 행 처 식품의약품안전평가원 의약품심사부 종양항생약품과



"청렴한 식약처  
국민 안심의 시작"

### 공익신고자 보호제도란?

– 공익신고자등(친족 또는 동거인 포함)이 공익신고등으로 인하여 피해를 받지 않도록 **비밀보장**,  
**불이익보호조치**, **신변보호조치** 등을 통하여 보호하는 제도

#### ◆ 보호조치 요구 방법

우편(30102) 세종특별자치시 도움5로 20 정부세종청사 7동, 국민권익위원회 공익보호지원과 / 전화 044-200-7773 / 팩스 044-200-7949

【공직자 부조리 및 공익신고안내】 ★★ 신고자 및 신고내용은 보호됩니다.

▶ 부조리 신고 : 식약처 홈페이지 "국민신문고" > 공직자 부조리 신고" 코너

▶ 공익 신고 : 식약처 홈페이지 "국민소통" > 신고센터 > 부패·공익신고 상담" 코너