

PROFESSIONAL TRAINING MANUAL

# 미국 FDA 신약 허가 신청을 위한 비임상 SEND 표준화 구축 및 실무 운영 가이드

베트남 등 글로벌 위탁 연구기관(CRO) 기초 데이터 수집부터 최종 SAS Transport File(.xpt) 패키징 및 FDA 보완 대응 전략까지

## 교육 대상 및 핵심 역량

- ✓ 바이오벤처 및 제약사 비임상/규제과학(RA) 담당 실무자
- ✓ 글로벌 CRO 파트너십 구축 및 비임상 데이터 검수 책임자
- ✓ FDA Refusal to File (RTF) 리스크 해소를 위한 예방 아키텍처 수립



## 교육 프로그램 목차

<b>MODULE 01</b>	비임상 SEND 표준과 선제적 도입의 당위성	Page 02
<b>MODULE 02</b>	핵심 비임상 도메인 아키텍처 및 연계 설계	Page 03
<b>MODULE 03</b>	FDA 바이어스 예방을 위한 데이터 표준화 4대 규칙	Page 04
<b>MODULE 04</b>	LIMS 연동형 SEND Ready 데이터 수집 스키마	Page 05
<b>MODULE 05</b>	Pinnacle 21 검증 대응 및 nSDRG 작성 전략	Page 06
<b>MODULE 06</b>	글로벌 계약 및 데이터 전송 계약(DTA) 실무 가이드	Page 07
<b>APPENDIX</b>	실무 검증 워크북 및 자가 평가 진단지	Page 08

### 서문 : 신약 개발자가 마주하는 비임상의 현실

미국 FDA에 IND(임상시험계획승인) 또는 NDA(신약허가신청)를 접수할 때, 제출하는 비임상 원천 데이터는 반드시 CDISC SEND 규격에 맞춰야 합니다. 많은 기업들이 단순히 최종 연구보고서(Study Report)의 변환만을 생각하지만, 이로 인해 수동 매핑 에러, 시간 지연, 막대한 예산 낭비를 겪고 있습니다.

본 교재는 데이터 수집의 첫 단추인 **건강 검진(Health Check) 단계**에서부터 어떻게 해외 파트너 및 위탁 기관을 제어하여 규격에 완벽히 호환되는 데이터를 만들어낼지 규제과학적이고 계약적 차원의 실무 해법을 담고 있습니다.



MODULE 01

## 사후 변환(Retrospective) vs 선제적 수집(Prospective)

비임상 데이터를 완료한 후에 일괄 변환하는 기존의 '사후 변환' 프로세스는 치명적인 병목 현상을 유발합니다. 반면, 데이터 발생 시점부터 표준을 주입하는 '선제적 수집' 프로세스는 FDA 심사 기간 단축과 무결성 유지를 보장합니다.

### 데이터 수집 패러다임의 전환

#### ■ 사후 일괄 변환 (Retrospective)

시험 완료 후 엑셀/PDF 취합 → 수동 변환 스크립트 가동 → Pinnacle 21 검증 수행 → 다량의 데이터 매핑 에러 발생 → 소스 데이터 재확인 요청 (수개월 지연)

#### ■ 선제적 표준 수집 (Prospective)

SEND 표준 수집 템플릿 배포 → 입력 시 정합성 실시간 검증 → 데이터 락(Data Lock)과 동시에 자동 매핑 → .xpt 즉각 변환 및 제출 (RTF 제로)

### 규제 및 비용 관점 비교 매트릭스

비교 영역	사후 일괄 변환	선제적 표준 수집
추가 변환 비용	시험당 \$15,000 ~ \$40,000	초기 세팅 후 추가 비용 0원
데이터 무결성	수기 변환 과정에서 오기 및 누락 다발	감사 추적(Audit Trail) 보존
소요 기간 (Time-to-File)	연구 완료 후 최소 2~3개월 소요	최종 데이터 확인 후 1주일 내 완료
FDA RTF 리스크	매우 높음 (검증 에러 누적)	완전 소멸 (사전 검증 필)



## MODULE 02

## SEND 핵심 도메인 아키텍처 수립

해외 파트너사나 현지 위탁시험기관(CRO)에서 보내오는 동물 건강 검진(Health Check) 및 기초 생리 데이터는 임의 파일이 아닌 **SENDIG(SEND Implementation Guide)** 사양에 부합하는 정밀 도메인(Domain) 연계가 필요합니다.

## TECHNICAL ARCHITECTURE

## ■ 관계형 마스터 데이터 (Relational Data Model)

비임상 데이터를 FDA에 올바르게 연계하기 위해서는 최상위 마스터인 **DM (인구통계 정보)**을 중심으로 기초 설계가 수립되어야 다른 분석 결과 도메인들과 바인딩됩니다.

## DM Domain

고유 식별자, 성별, 종, 도입 연령 관리

## TX/SE Domain

시험군 정의 및 시기별 마스킹 조건

## BW/LB/VS Domain

체중, 혈액생화학, 생체 징후 실데이터

## 실무 도메인 바인딩 가이드라인

## 1. DM (Demographics) - 모든 분석 데이터의 중심축

도입일자(`BRTHDC`), 종(`SPECIES`), 계통(`STRAIN`), 성별(`SEX`) 등 동물의 고유 기초 메타데이터가 완벽하게 규정되고 통일되어야 후속 도메인 분석이 가능합니다.

## 2. TX (Trial Design) - 시험 환경 및 배정 조건 제어

동물이 투약 전에 있다면 투약 단계 전 상태를 표시하는 타임라인 정의 필드(`SE - Subject Elements`)와 군 매핑 상태가 완벽히 얼라인되어 있어야 합니다.

## 3. LB (Laboratory Test Results) - 임상병리 수집 구조

LIMS(임상실험실 관리 시스템)에서 추출된 기기 수치가 CDISC Controlled Terminology(CT) 코드집에 명시된 단위 및 명칭으로 1:1 상호 대응하도록 데이터 딕셔너리를 연동합니다.



## MODULE 03

## 데이터 표준화 4대 대원칙

해외 연구소에서 기록하는 원시 텍스트 형태의 비정형 로우 데이터를 FDA 심사가 통과 가능한 정형 데이터로 승급 시키기 위한 극단적 4대 데이터 처리 규칙입니다.

**Rule 1** 계층 구조형 글로벌 동물 고유 식별자 (USUBJID) 강제 적용

단순한 `1`, `2`, `3` 또는 수컷 `M1`, `F1` 과 같은 숫자는 타 연구 결과와 중복됩니다. FDA 데이터 이식성을 보장하기 위해 다기관 비임상에도 호환되는 고유 번호 형식을 적용합니다.

**형식 예시:** [Study ID]-[Site ID]-[Animal No] → PROJ2026-VNCRO-M015

**Rule 2** ISO 8601 날짜/시간 및 추적 시간 표기 통일

수기로 임의 기입된 현지 타임존 표기를 방지하고, 기계 분석 친화적인 ISO 표준을 강제합니다.

**날짜 예시:** 2026-05-29 | **일시 예시 (UTC):** 2026-05-29T09:30:00Z

**Rule 3** 정량 분석 한계 미만(LLOQ) 데이터의 수학적 분리 처리

LIMS 데이터 출력 시  $<0.01$  이나 N.D. 를 그대로 수치 변수에 채워 넣으면 P21 오류가 납니다. 원본 텍스트는 LBORRES 에 두고, 계산용 수치는 통계 규칙에 의거 가공하여 LBSTRESN 에 기록하고, LBLMT 필드에 "LOQ" 마커를 명시합니다.

**Rule 4** CDISC Controlled Terminology (통제 용어) 원격 적용

수집 도구(Excel 등)의 각 셀 입력부에 CDISC CT 라이브러리를 바인딩하여, 타당하지 않은 영문 오기나 로컬 용어 기입을 원천 예방해야 합니다.



MODULE 04

## 수집 스키마 템플릿 명세

해외 파트너사에 제공할 'SEND 호환 템플릿'의 설계 양식입니다. 이 형식에 준하여 로우 데이터를 수집해야 데이터 이송 속도를 보장받을 수 있습니다.

### 1) 체중 측정 스키마 (BW Domain Schema)

열 이름 (Column)	CDISC 코드	데이터 타입	유효성 제약	실제 데이터 예시
Study ID	`STUDYID`	String	FDA 보고 공식 코드	STUDY-2026-01
Animal ID	`USUBJID`	String	계층형 고유 식별자	ST26-VN-M001
Measurement	`BWORRES`	Float	소수점 2자리 수치	382.50
Original Unit	`BWORRESU`	String	CDISC CT 단위 (g)	g
DateTime	`BWDTC`	String (ISO)	ISO 8601 포맷	2026-05-29T08:15

### 2) 임상화학 검사 스키마 (LB Domain Schema)

STUDYID	USUBJID	LBSEQ	LBTESTCD	LBTEST	LBORRES	LBSTRESC	LBSTRESN	LBSPEC
ST-2601	VN-M01	1	ALT	Alanine Amino...	42	42	42	SERUM
ST-2601	VN-M01	2	AST	Aspartate Amino...	115	115	115	SERUM
ST-2601	VN-M01	3	GLUC	Glucose	<20	LOQ	10	SERUM

**실무 예외 처리 조언:** 위 노란색으로 강조된 GLUC(포도당) 분석 사례처럼 정량 하한 미만(<20)의 결과는 기계적 수치 변환 시 에러를 유발하므로 표준 규정인 **LOQ 분류(LBSTRESC)** 및 보정값 입력 규칙(10)을 준수해 에러를 완전 차단합니다.



## MODULE 05

## Pinnacle 21 검증 트러블슈팅

FDA에 최종 제출된 SEND 데이터셋은 자가 유효성 검사 도구인 **Pinnacle 21**에 의해 엄밀히 평가됩니다. 검출되는 위반(Violations) 수준에 따라 서류 심사 거절 혹은 보완 요구서가 발송됩니다.

### ⚠ 대표 오류 유형 및 즉각적 해결법

#### 🚫 [Error SD0029] USUBJID not found in DM Domain

##### 원인 분석:

혈액 검사(LB)나 체중(BW) 기록에는 존재하는 특정 동물의 ID가 정작 DM(기본 인구통계) 마스터 리스트에 누락된 상태입니다.

##### 해결 가이드:

변환 가동 이전에 DM의 유니크한 ID 어레이와 교차 대조(Cross-Validation)하는 파이프라인 무결성 체크용 정제 단계를 선제 운영하십시오.

#### ⚠ [Warning SD0058] Invalid ISO 8601 date/time format

##### 원인 분석:

해외 위탁 연구원이 수기로 작성한 날짜 텍스트(`2026/05/29` 등)의 포맷 규격 위반입니다.

##### 해결 가이드:

데이터 전처리 정제(ETL) 스크립트 상에 정규표현식 변환 규칙을 내장하여 `YYYY-MM-DD`로 일괄 강제 포매팅 처리합니다.

#### 📖 nSDRG(비임상 데이터 표준화 요약서) 작성 수칙


실제 환경에서 해결할 수 없는 일부 잔여 Warning(예: 장비 기능상의 한계로 발생하는 미매핑 등)은 임의 변조하지 않고, 그 기술적 및 과학적 타당성을 nSDRG(nonclinical Study Data Standardization Reviewer's Guide)에 소명하여 심사 보완 요구를 차단해야 합니다.



## MODULE 06

## 계약서(SOW/DTA) 바인딩 템플릿

해외 CRO나 파트너 기관과의 협업 계약을 맺을 때, 사후에 데이터 무결성 책임 문제로 법적 분쟁을 겪거나 재작업 비용을 전가당하지 않도록 계약서 내에 SEND 준수 요건을 법적 효력 있게 못 박아야 합니다.

 표준 계약 조항 예시 (Standard Legal Clause)

## 제X조 (비임상 데이터 표준 및 인도 규격)

1. "위탁 연구기관(이하 "CRO")"은 본 계약에 따라 수행되는 모든 시험(건강검진, 기초 스크리닝 시험 및 효능 시험 포함)의 데이터 수집, 가공 및 산출물을 CDISC SEND 표준 규격(최신 고시된 SENDIG 버전을 기준으로 함)에 상호 호환될 수 있는 형태로 제공하여야 한다.
2. "CRO"는 데이터 인도 전, 상용 유효성 검사 도구(Pinnacle 21 등)를 활용하여 정합성 검증을 필히 수행하여야 하며, "Reject" 및 "Error" 등급의 결함이 식별되지 않은 깨끗한 상태의 데이터 파일(`.xpt` 또는 이에 매핑 가능한 가독성 높은 데이터 스키마) 및 검증 리포트를 "의뢰사"에게 제출하여야 한다.
3. 만약 인도된 데이터의 포맷 오류나 통제 용어 미준수로 인해 FDA 승인 프로세스 상에서 보완 요구(Deficiency Letter) 또는 RTF가 발생할 경우, "CRO"는 무상으로 즉각적인 데이터 보정 및 재매핑 작업을 전적으로 책임지고 수행하여야 한다.

 실무 이행 관리(Compliance) 조언

단순한 조항 삽입을 넘어, 매 회 검사 차수(예: 검진 완료 후)마다 엑셀 로우데이터의 SEND 이식 유효성 간이 검증을 중간 정산 조건(Milestone Approval)으로 배치하면 파트너사의 성실 이행을 보다 효과적으로 통제할 수 있습니다.



## APPENDIX

## 실무 자가 평가 및 워크북

본 장에서는 교육을 통해 학습한 SEND 데이터 표준화의 핵심 사항을 실제 업무에 적용하기 전 스스로 점검하고 테스트해보는 교육용 자가 평가 문항입니다.

## Q1. 다음 중 FDA 제출용 동물 고유 식별자(USUBJID) 표기법으로 가장 올바른 양식은 무엇인가?

- ① Rat-No.5 (수컷)
- ② M005-VN
- ③ PROJ2026-VNCRO-M005
- ④ 2026-05-29-M5

💡 해설: 고유 식별자는 타 연구 데이터와 절대로 중복되지 않도록 [연구프로젝트코드]-[기관식별코드]-[성별동물번호] 식의 계층식 유니크 설계를 지녀야 합니다.

## Q2. 임상분석 중 LLOQ 미만의 검출값을 데이터 세트에 기입할 때, Pinnacle 21 검증 에러를 피하기 위해 어떠한 표준 절차를 밟아야 하는가?

- 1. 원본 문자열("<0.01")은 LBORRES 필드에 고스란히 둔다.
- 2. 분석 계산용 LBSTRESN 필드에는 하한치의 1/2 등을 수치화하여 보정 기록한다.
- 3. LBLMT 판단 플래그에 "LOQ"를 기입하여 타당성을 소명한다.

💡 해설: 이 세 가지 단계를 모두 일치되게 수행해야 유효성 오류 및 RTF를 원천 차단할 수 있습니다.

## ☑ 내부 실무 이행 최종 마스터 체크

- ✓ 해외 파트너 계약에 DTA 조항 적용 여부
- ✓ LIMS 시스템 내 CDISC CT 바인딩 상태
- ✓ ISO 8601 일시 표준 변환 파이프라인 수립
- ✓ 잔여 Warning에 대한 nSDRG 작성 숙지



