



2023

CHINA

INTERCHANGE

BEIJING | 25-26 AUGUST



Efficient Data Standardization Through AI

Presented by 朱博贡, Chief Business Architect, 耀乘健康科技



Meet the Speaker

朱博贡

Title: Chief Business Architect

Organization: 耀乘健康科技 (yaocheng.cn)

朱博贡先生作为前睿理健康科技中国的重要技术骨干，于2020年共同创办了上海耀乘健康科技有限公司，任技术副总裁、首席业务架构师，负责耀乘健康的产品及技术研发工作，以及开发运营管理等方面工作。

在加入耀乘健康科技之前，朱博贡先生就职于谷歌（Google）集团超过13年，负责过Ads、Cloud AI等领域的各种项目。离开谷歌后在睿理健康科技中国（谷歌母公司旗下专注于生命科学领域的创业公司）任技术主管。



Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*



Efficient Data Standardization Through AI

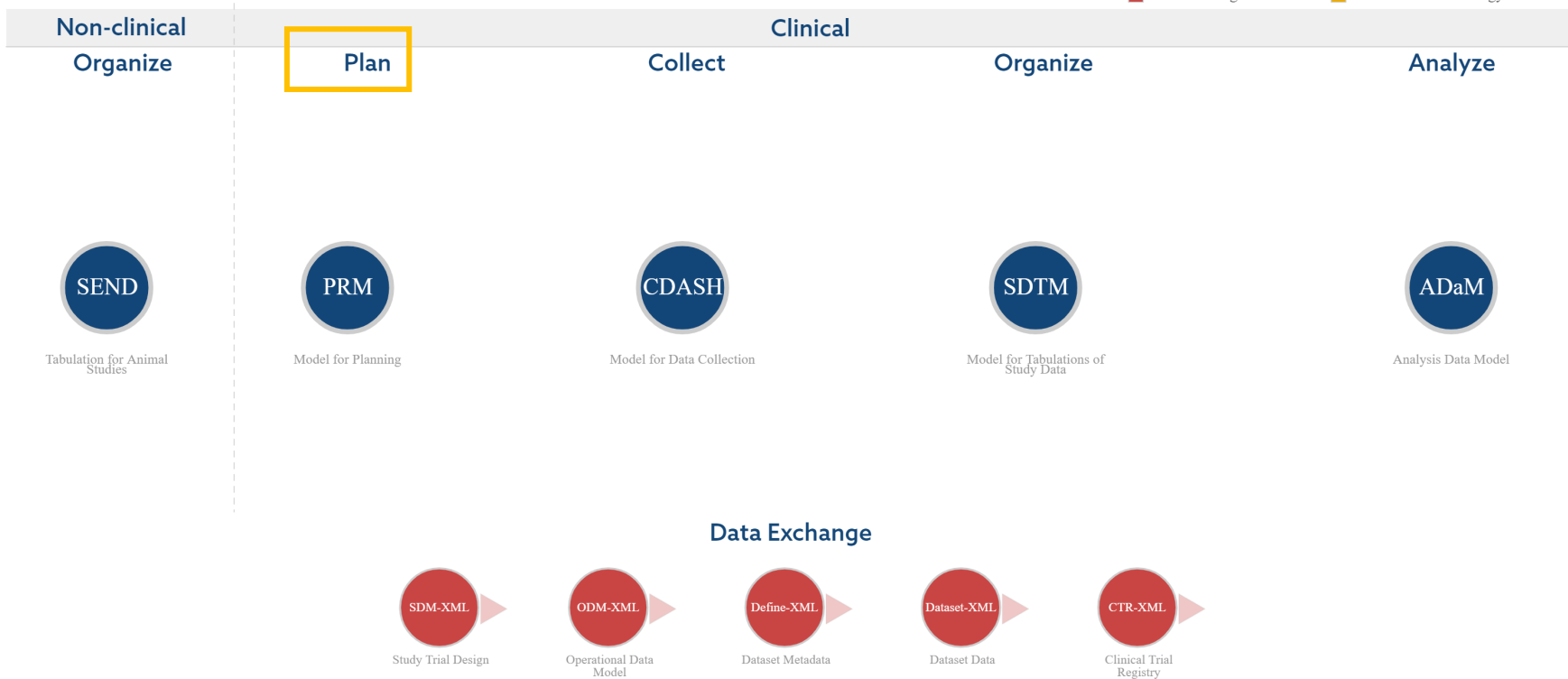
Agenda

1. CDISC 关于 Digitized Protocol 的构想
2. Digitize protocol 过程中的挑战
3. AI 可以辅助解决的问题 及 潜在的应用

CDISC 的标准体系

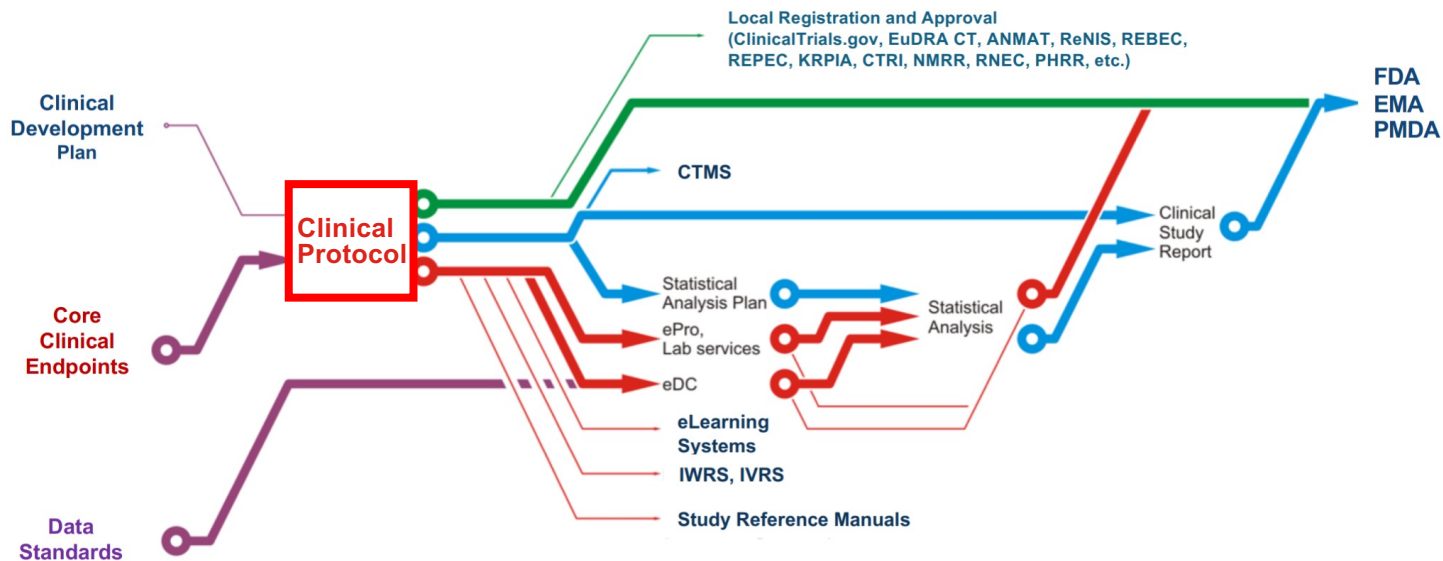
CDISC Standards in the Clinical Research Process

- Foundational Standard
- Therapeutic Area
- Data Exchange
- Controlled Terminology



以方案为中心，基于各类文档的试验体系





Source: <https://www.cdisc.org/sites/default/files/pdf/DDFPublicReview-FINAL.pdf>

Digitized Protocol构想

The Clinical Trial Information Flow

从 Human Readable 到 Machine Readable 的鸿沟

PROTOCOL SYNOPSIS

TITLE: A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, MULTICENTER STUDY TO EVALUATE THE SAFETY AND EFFICACY OF TOCILIZUMAB IN PATIENTS WITH SEVERE COVID-19 PNEUMONIA

PROTOCOL NUMBER: WA42380

VERSION NUMBER: 3

EUDRACT NUMBER: 2020-001154-22

IND NUMBER: 148225

IND NUMBER: NCT04320615

TEST PRODUCT: Tocilizumab (RO4877533)

PHASE: Phase III

INDICATION: Severe COVID-19 pneumonia

SPONSOR: F. Hoffmann-La Roche Ltd

Objectives and Endpoints
 This study will evaluate the efficacy, safety, pharmacodynamics, and pharmacokinetics of tocilizumab (TCZ) compared with a matching placebo in combination with standard of care (SOC) in hospitalized patients with severe COVID-19 pneumonia. Specific objectives and corresponding endpoints for the study are outlined below.

Efficacy Objectives

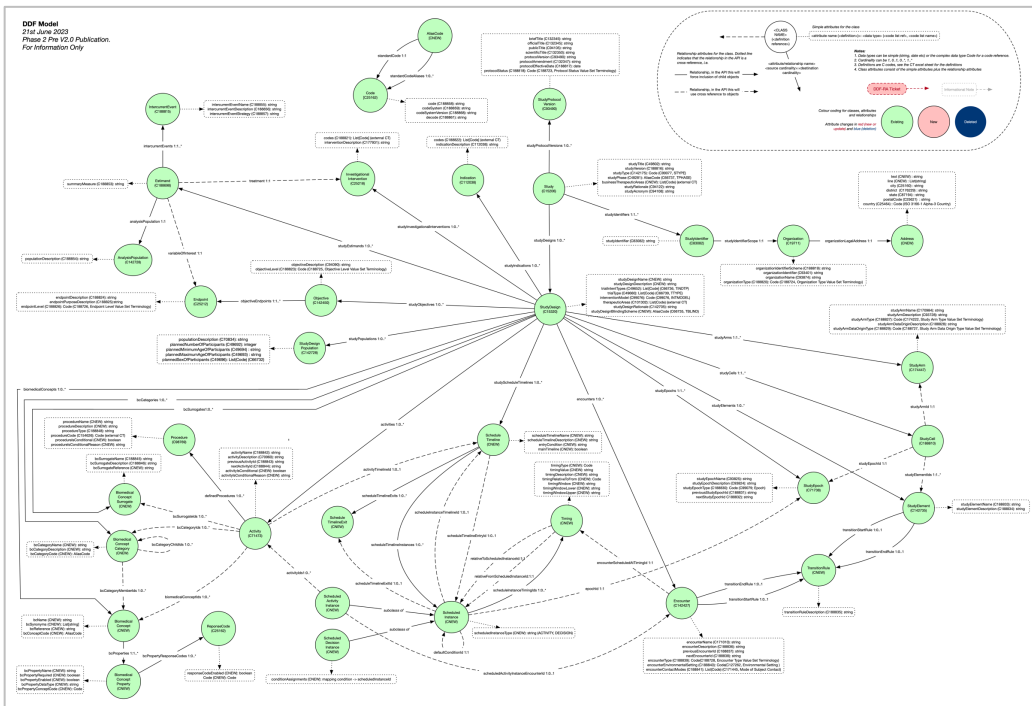
Primary Efficacy Objective
 The primary efficacy objective for this study is to evaluate the efficacy of TCZ compared with placebo in combination with SOC for the treatment of severe COVID-19 pneumonia on the basis of the following endpoint:

- Clinical status assessed using a 7-category ordinal scale at Day 28

Secondary Efficacy Objective
 The secondary efficacy objective for this study is to evaluate the efficacy of TCZ compared with placebo in combination with SOC for the treatment of severe COVID-19 pneumonia on the basis of the following endpoints:

- Time to clinical improvement (TTCI) defined as a National Early Warning Score 2 (NEWS2) of ≤ 2 maintained for 24 hours
- Time to improvement of at least 2 categories relative to baseline on a 7-category ordinal scale of clinical status
- Incidence of mechanical ventilation
- Ventilator-free days to Day 28
- Incidence of intensive care unit (ICU) stay
- Duration of ICU stay

Source: <https://classic.clinicaltrials.gov/ct2/show/NCT04320615>



Source: DDF USDM Reference Architecture v2.0

从 Human Readable 到 Machine Readable 的鸿沟

PROTOCOL SYNOPSIS	
TITLE:	A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, MULTICENTER STUDY TO EVALUATE THE SAFETY AND EFFICACY OF TOCILIZUMAB IN PATIENTS WITH SEVERE COVID-19 PNEUMONIA
PROTOCOL NUMBER:	WA42380
VERSION NUMBER:	3
EUDRACT NUMBER:	2020-001154-22
IND NUMBER:	148225
IND NUMBER:	NCT04320615
TEST PRODUCT:	Tocilizumab (RO4877533)
PHASE:	Phase III
INDICATION:	Severe COVID-19 pneumonia
SPONSOR:	F. Hoffmann-La Roche Ltd
Objectives and Endpoints	
This study will evaluate the efficacy, safety, pharmacodynamics, and pharmacokinetics of tocilizumab (TCZ) compared with a matching placebo in combination with standard of care (SOC) in hospitalized patients with severe COVID-19 pneumonia. Specific objectives and corresponding endpoints for the study are outlined below.	
Efficacy Objectives	
Primary Efficacy Objective	
The primary efficacy objective for this study is to evaluate the efficacy of TCZ compared with placebo in combination with SOC for the treatment of severe COVID-19 pneumonia on the basis of the following endpoint:	
<ul style="list-style-type: none">Clinical status assessed using a 7-category ordinal scale at Day 28	
Secondary Efficacy Objective	
The secondary efficacy objective for this study is to evaluate the efficacy of TCZ compared with placebo in combination with SOC for the treatment of severe COVID-19 pneumonia on the basis of the following endpoints:	
<ul style="list-style-type: none">Time to clinical improvement (TTCI) defined as a National Early Warning Score 2 (NEWS2) of ≤ 2 maintained for 24 hoursTime to improvement of at least 2 categories relative to baseline on a 7-category ordinal scale of clinical statusIncidence of mechanical ventilationVentilator-free days to Day 28Incidence of intensive care unit (ICU) stayDuration of ICU stay	

Source: <https://classic.clinicaltrials.gov/ct2/show/NCT04320615>

```
340 "studyObjectives": [  
341 {  
342   "objectiveId": "Objective_1",  
343   "objectiveDescription": "The primary efficacy objective for this study is to evaluate the efficacy of TCZ compared with placebo in combination with SOC for the treatment of severe COVID-19 pneumonia on the basis of the following endpoint:",  
344   "objectiveLevel": {  
345     "codeId": "Code_352",  
346     "code": "C85826",  
347     "codeSystem": "http://www.cdisc.org",  
348     "codeSystemVersion": "2022-12-16",  
349     "decode": "Trial Primary Objective"  
350   },  
351   "objectiveEndpoints": [  
352     {  
353       "endpointId": "Endpoint_1",  
354       "endpointDescription": "Clinical status assessed using a 7-category ordinal scale at Day 28",  
355       "endpointPurposeDescription": "",  
356       "endpointLevel": {  
357         "codeId": "Code_351",  
358         "code": "C94496",  
359         "codeSystem": "http://www.cdisc.org",  
360         "codeSystemVersion": "2022-12-16",  
361         "decode": "Primary Endpoint"  
362       }  
363     }  
364   ]  
365 },  
366 {  
367   "objectiveId": "Objective_2",  
368   "objectiveDescription": "The secondary efficacy objective for this study is to evaluate the efficacy of TCZ compared with placebo in combination with SOC for the treatment of severe COVID-19 pneumonia on the basis of the following endpoints:",  
369   "objectiveLevel": {  
370     "codeId": "Code_354",  
371     "code": "C85827",  
372     "codeSystem": "http://www.cdisc.org",  
373     "codeSystemVersion": "2022-12-16",  
374     "decode": "Trial Secondary Objective"  
375   },  
376   "objectiveEndpoints": [  
377     {  
378       "endpointId": "Endpoint_2",  
379       "endpointDescription": "Time to clinical improvement (TTCI) defined as a National Early Warning Score 2 (NEWS2) of  $\leq 2$  maintained for 24 hours",  
380       "endpointPurposeDescription": "",  
381       "endpointLevel": {  
382         "codeId": "Code_353",  
383         "code": "C139173",  
384         "codeSystem": "http://www.cdisc.org",  
385         "codeSystemVersion": "2022-12-16",  
386         "decode": "Secondary Endpoint"  
387       }  
388     }  
389   ]  
390 }  
391 ]
```

Source: DDF USDM Reference Architecture v2.0

Human Readable + AI → Machine Readable

ICH M11 Template

114	Table of Contents	
115		
116	1 PROTOCOL SUMMARY.....	19
117	1.1 Protocol Synopsis.....	19
118	1.2 Trial Schema.....	22
119	1.3 Schedule of Activities.....	22
120	2 INTRODUCTION.....	23
121	2.1 Purpose of Trial.....	23
122	2.2 Summary of Benefits and Risks.....	23
123	3 TRIAL OBJECTIVES, ENDPOINTS AND ESTIMANDS.....	25
124	3.1 {Primary/Secondary/Exploratory} Objective + Associated Endpoint (and Estimand).....	25
125	4 TRIAL DESIGN.....	26
126	4.1 Description of Trial Design.....	26
127	4.1.1 Participant Input into Design.....	27
128	4.2 Rationale for Trial Design.....	27
129	4.2.1 Rationale for Comparator.....	27
130	4.2.2 Rationale for Adaptive or Novel Trial Design.....	28
131	4.2.3 Other Trial Design Considerations.....	28
132	4.3 Access to Trial Intervention After End of Trial.....	28
133	4.4 Start of Trial and End of Trial.....	28
134	5 TRIAL POPULATION.....	29
135	5.1 Selection of Trial Population.....	29
136	5.2 Rationale for Trial Population.....	29
137	5.3 Inclusion Criteria.....	29
138	5.4 Exclusion Criteria.....	30
139	5.5 Lifestyle Considerations.....	30
140	5.5.1 Meals and Dietary Restrictions.....	30
141	5.5.2 Caffeine, Alcohol, Tobacco, and Other Habits.....	30
142	5.5.3 Physical Activity.....	30
143	5.5.4 Other Activity.....	30
144	5.6 Screen Failures.....	30
145	6 TRIAL INTERVENTION AND CONCOMITANT THERAPY.....	31
146	6.1 Description of Trial Intervention.....	31

14

ICH M11 Template

147	6.2 Rationale for Trial Intervention.....	31
148	6.3 Dosing and Administration.....	32
149	6.3.1 Trial Intervention Dose Modification.....	32
150	6.4 Treatment of Overdose.....	32
151	6.5 Preparation, Handling, Storage and Accountability.....	32
152	6.5.1 Preparation of Trial Intervention.....	32
153	6.5.2 Handling and Storage of Trial Intervention.....	33
154	6.5.3 Accountability of Trial Intervention.....	33
155	6.6 Participant Assignment, Randomisation and Blinding.....	33
156	6.6.1 Participant Assignment.....	33
157	6.6.2 Randomisation.....	34
158	6.6.3 Blinding and Unblinding.....	34
159	6.7 Trial Intervention Compliance.....	34
160	6.8 Concomitant Therapy.....	34
161	6.8.1 Prohibited Concomitant Therapy.....	35
162	6.8.2 Permitted Concomitant Therapy.....	35
163	6.8.3 Rescue Therapy.....	35
164	6.8.4 Other Therapy.....	35
165	7 DISCONTINUATION OF TRIAL INTERVENTION AND PARTICIPANT	
166	WITHDRAWAL FROM TRIAL.....	36
167	7.1 Discontinuation of Trial Intervention.....	36
168	7.1.1 Criteria for Permanent Discontinuation of Trial Intervention.....	36
169	7.1.2 Temporary Discontinuation or Interruption of Trial Intervention.....	36
170	7.1.3 Rechallenge.....	36
171	7.2 Participant Withdrawal from the Trial.....	37
172	7.3 Lost to Follow-Up.....	37
173	7.4 Trial Stopping Rules.....	37
174	8 TRIAL ASSESSMENTS AND PROCEDURES.....	38
175	8.1 Screening/Baseline Assessments and Procedures.....	38
176	8.2 Efficacy Assessments and Procedures.....	38
177	8.3 Safety Assessments and Procedures.....	38
178	8.3.1 Physical Examination.....	39
179	8.3.2 Vital Signs.....	39

15

ICH M11 Template

180	8.3.3 Electrocardiograms.....	39
181	8.3.4 Clinical Laboratory Assessments.....	39
182	8.3.5 Suicidal Ideation and Behaviour Risk Monitoring.....	39
183	8.4 Adverse Events and Serious Adverse Events.....	39
184	8.4.1 Definitions of AE and SAE.....	39
185	8.4.2 Time Period and Frequency for Collecting AE and SAE Information.....	40
186	8.4.3 Identifying AEs and SAEs.....	40
187	8.4.4 Recording of AEs and SAEs.....	40
188	8.4.5 Follow-up of AEs and SAEs.....	40
189	8.4.6 Reporting of SAEs.....	40
190	8.4.7 Regulatory Reporting Requirements for SAEs.....	40
191	8.4.8 Serious and Unexpected Adverse Reaction Reporting.....	40
192	8.4.9 Adverse Events of Special Interest.....	41
193	8.4.10 Disease-related Events or Outcomes Not Qualifying as AEs or SAEs.....	41
194	8.5 Pregnancy and Postpartum Information.....	41
195	8.5.1 Participants Who Become Pregnant During the Trial.....	41
196	8.5.2 Participants Whose Partners Become Pregnant.....	42
197	8.6 Medical Device Product Complaints for Drug/Device Combination Products.....	42
198	8.6.1 Definition of Medical Device Product Complaints.....	42
199	8.6.2 Recording of Medical Device Product Complaints.....	42
200	8.6.3 Time Period and Frequency for Collecting Medical Device Product Complaints.....	42
201	8.6.4 Follow-Up of Medical Device Product Complaints.....	42
202	8.6.5 Regulatory Reporting Requirements for Medical Device Product Complaints.....	42
203	8.7 Pharmacokinetics.....	43
204	8.8 Genetics.....	43
205	8.9 Biomarkers.....	43
206	8.10 Immunogenicity Assessments.....	43
207	8.11 Medical Resource Utilisation and Health Economics.....	44
208	9 STATISTICAL CONSIDERATIONS.....	45
209	9.1 Analysis Sets.....	45
210	9.2 Analyses Supporting Primary Objective(s).....	45
211	9.2.1 Statistical Model, Hypothesis, and Method of Analysis.....	45

16

Source: 2023 CLINICAL ELECTRONIC STRUCTURED HARMONISED PROTOCOL (DRAFT) M11 Template, CDE

Schedule of Activities (SoA) 的转换示例

The screenshot displays the AURORA Prime interface for a clinical trial protocol. The main content area shows the following tables:

Weight/height

字段组1		wh
Weight (kg)	Double(4,1)	WEIGHT
Height (cm)	Double(4,1)	HEIGHT
BMI	Double(4,1)	BMI

[From Protocol]Height only recorded at screening visit.
 BMI为Weight (kg) / Height (cm) / Height (cm) * 10000

Vital signs

字段组1		vs
Were Vital Signs Collected?	<input type="radio"/> Yes <input type="radio"/> No	VSPERF
Reason Not Collected:	String(\$50)	VSREASND
Date of Assessment:	dd-MMM-yyyy	VSDAT
Systolic Blood Pressure: (mmHg)	Integer(3)	SYSBP
Diastolic Blood Pressure: (mmHg)	Integer(3)	DIABP
Heart Rate: (beats/min)	Integer(3)	PULSE
Respiration Rate: (breaths/min)	Integer(2)	RESP
Temperature (C):	Double(4,1)	TEMP

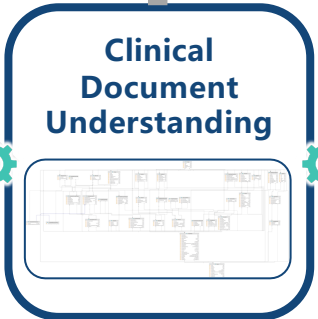
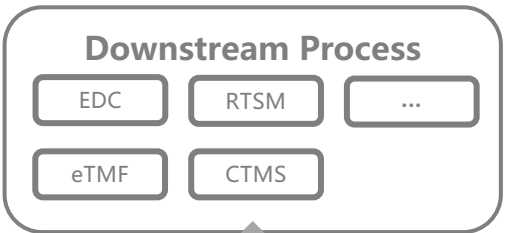
[From Protocol]Vital signs: Blood Vital Signs: blood pressure (BP), heart rate (HR), respiratory rate (RR), and temperature (temp)

ECOG performance status

字段组1		ep
Was the ECOG Performance Status Assessed?	<input type="radio"/> Yes	EPYN

表单数量: 21

4. Objective disease assessments: Chest and abdominal (including pelvic cavity) CT or MRI should be performed during screening; brain CT/MRI should be

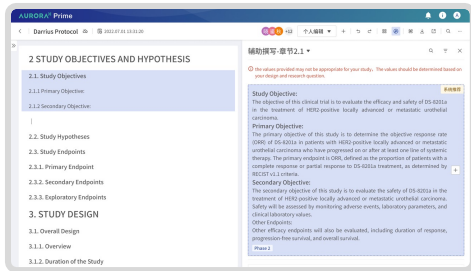


- Academic Paper
- Protocol
- Guiding Principle
- ...

- Entity
- Relation
- Attribute
- ...

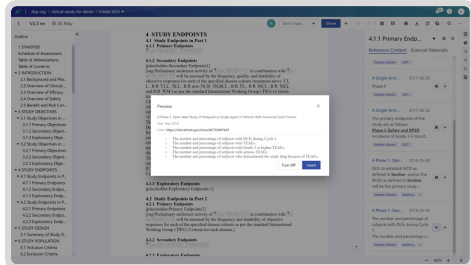


- Objectives
- Endpoints
- I/E Criteria
- SoA
- ...



Content Reuse & Generation

- ✓ From Synopsis to the Main section
- ✓ Related documents, e.g. CSR
- ✓ ...

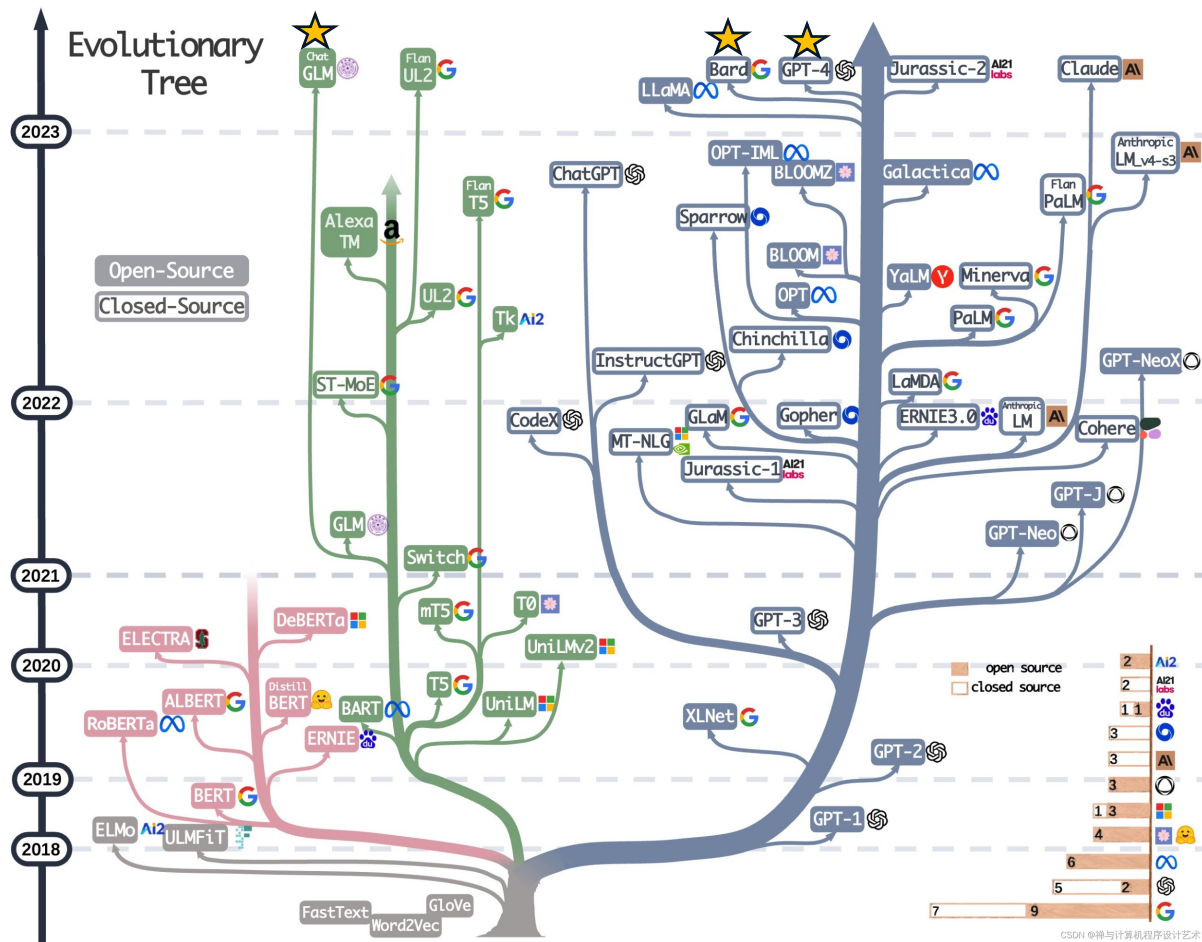


Reference

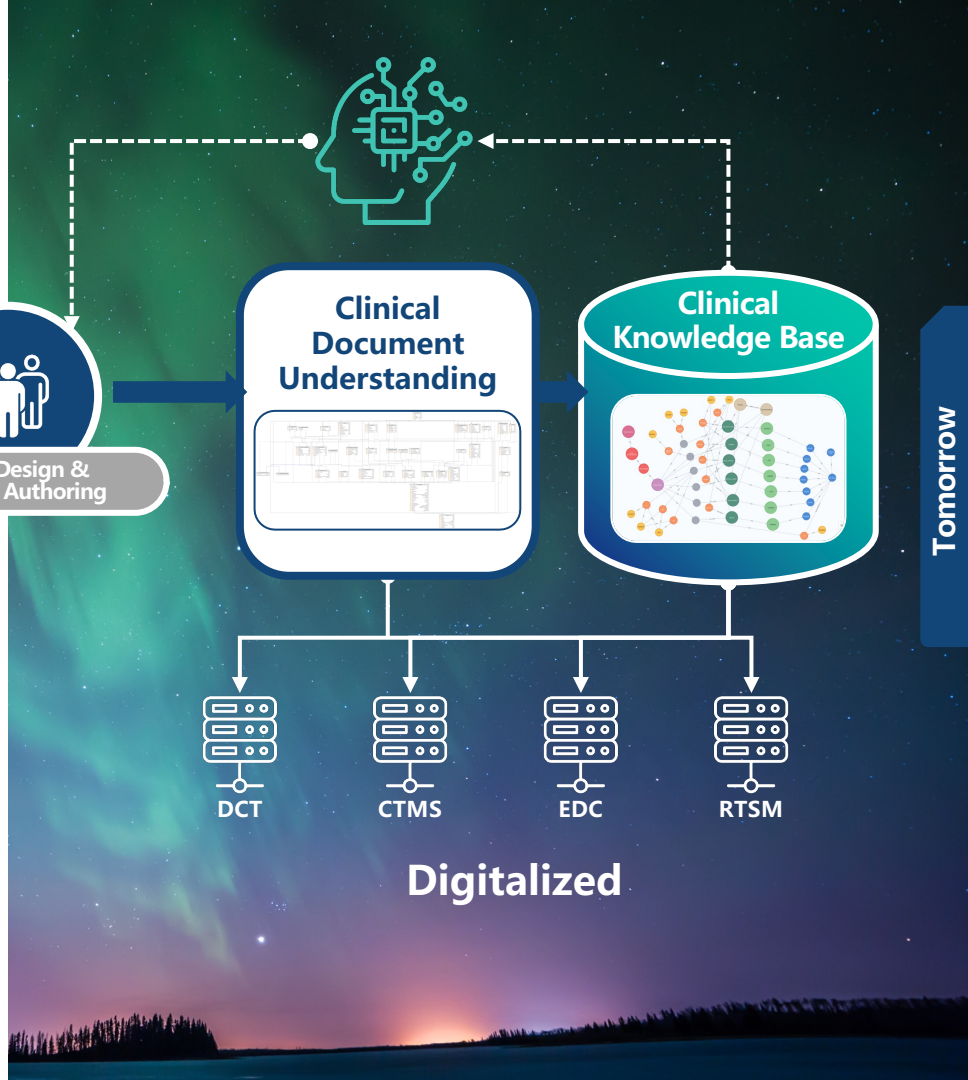
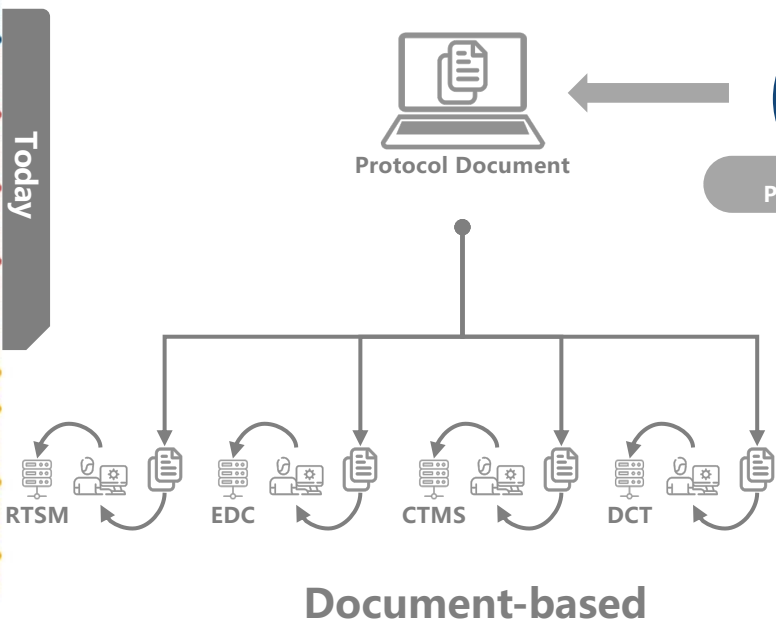
- ✓ What's the popular endpoints of a particular biological target ?
- ✓ What's the most used I/E criteria of a particular indication ?
- ✓ ...

如何选择模型

- 哪个模型的效果好？
- 哪个模型的价格比好？
- 哪个模型适合我的使用场景？



从 Document & Manual Flow 迈向 Digital & Automated Flow



效率 + 合规 → 价值

yaocheng.cn

展台: 11



FDA 正持续与业界展开探讨

FDA Releases Two Discussion Papers to Spur Conversation about Artificial Intelligence and Machine Learning in Drug Development and Manufacturing

Source: [FDA 2023-05-10](#)

Solicit feedback on 3 key areas

1. Human-led governance, accountability, and transparency
2. Quality, reliability, and representativeness of data
3. Model development, performance, monitoring, and validation

器械已有相关指导原则: GMLP

Good Machine Learning Practice for Medical Device Development: Guiding Principles

Share Tweet LinkedIn Email Print

The U.S. Food and Drug Administration (FDA), Health Canada, and the United Kingdom's Medicines and Healthcare products Regulatory Agency (MHRA) have jointly identified 10 guiding principles that can inform the development of Good Machine Learning Practice (GMLP). These guiding principles will help promote safe, effective, and high-quality medical devices that use artificial intelligence and machine learning (AI/ML).

Artificial intelligence and machine learning technologies have the potential to transform health care by deriving new and important insights from the vast amount of data generated during the delivery of health care every day. They use software algorithms to learn from real-world



Download PDF

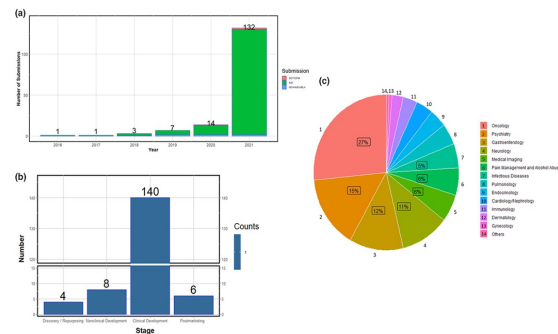
Source: [FDA 2021-10-27](#)

10 guiding principles

1. Multi-Disciplinary Expertise Is Leveraged Throughout the Total Product Life Cycle
2. Good Software Engineering and Security Practices Are Implemented
3. ...

大势所趋

The regulatory uses are real: In 2021, more than 100 [drug and biologic applications](#) submitted to the FDA included AI/ML components. These submissions spanned a range of therapeutic areas, and sponsors incorporated the technologies in different developmental stages.



Source: [Landscape Analysis of the Application of Artificial Intelligence and Machine Learning in Regulatory Submissions for Drug Development From 2016 to 2021](#)