

ICS 11.120
CCS C10

T/CNPHARS

中国药理学学会团体标准

T/CNPHARS 0001—2025

群体药动学-药效学分析标准操作规范

Standard operation procedure for
population pharmacokinetic-pharmacodynamic analysis

2025 - 08 - 14 发布

2025 - 08 - 15 实施

中国药理学学会 发布

目 次

前 言	II
引 言	3
1 范围	4
2 规范性引用文件	4
3 术语和定义	4
4 符号	5
5 标准操作规范	5
5.1 概述	5
5.2 前期工作	6
5.3 基础模型	11
5.4 最终模型	15
5.5 模型评价	18
5.6 模型应用	24
5.7 分析报告	25
5.8 质量控制和风险管理	27
附 录 A （规范性） 质量控制要点	29
附 录 B （资料性） 参数估算的常用方法	34
附 录 C （资料性） 协变量筛选的常用方法	35
参 考 文 献	36

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国药理学学会定量药理学专业委员会提出。

本文件由中国药理学学会归口。

本文件起草单位：上海交通大学医学院附属胸科医院，长春金赛药业有限责任公司，葛兰素史克（上海）医药研发有限公司，强生（中国）投资有限公司，上海中医药大学，北京协和医院，复旦大学附属华山医院，上海强世信息科技有限公司，亘喜生物科技（上海）有限公司，海瑞宁康生物医药，北京大学药学院，北京大学第三医院，皖南医学院弋矶山医院，北京大学人民医院，山东大学，中南大学，复旦大学药学院，苏州大学附属第一医院，牛津大学纳菲尔德医学部热带疾病和全球健康中心，美国加州大学旧金山分校。

本文件主要起草人：焦正，丁俊杰，马广立，周绚，李禄金，王鲲，李良，任宇鹏，郑青山，胡蓓，张菁，谢海棠，周田彦，刘东阳，赵维，阳国平，方翼，相小强，陈锐，赵亮，王亚宁，刘昭前，缪丽燕。

本文件每五年更新一次。若遇重大技术事件或发生变故，则及时更新。

引 言

群体药动学-药效学分析是通过数学建模的方法，将经典药动学-药效学理论与统计学原理和方法相结合的一种分析技术。其核心在于考察和建立研究对象的药动学和药效学的定量关系、并定量分析其内在和外在的影响因素和随机效应。相较于经典药动学-药效学分析，群体药动学-药效学分析能够充分利用稀疏采样的药动学、药效学或药动学-药效学数据，开展深入研究。这一优势使其在针对患者群体，尤其是儿童、老年人、孕产妇、危重症患者等特殊人群进行药动学-药效学研究时，能够发挥更为重要的作用。

群体药动学旨在定量描述药物在体内吸收、分布、代谢和排泄的过程，并定量考察药动学在研究群体中的变异程度及其来源。而群体药效学旨在定量描述药物效应随药物暴露量和时间的变化过程，并考察在研究群体中药物效应的变异程度及其来源。两者相结合，可定量描述研究群体中的药物、机体和疾病之间的关系，为药物研发和临床用药决策提供重要依据，是“模型引导的新药研发（model informed drug development）”和“模型引导的精准用药（model informed precision dosing）”中的重要分析技术。

群体药动学-药效学分析是一个融合药理学、药物治疗学、数学建模与计算机科学等多学科理论与知识的“学习与验证”的迭代过程。其分析流程复杂且耗时，亟需通过规范化和标准化提高分析效率与质量。自1999年美国食品药品监督管理局率先颁布群体药动学技术指导原则以来，欧洲、日本及中国等药品监管机构也相继发布了相关技术指导文件。然而，现有的技术指南多侧重于基于原则和框架性体要求，缺乏具体的标准操作流程与规范，影响了群体药动学-药效学分析的质量及结果解读。因此，本规范基于当前对群体药动学-药效学分析的理论认知与实践经验，旨在为相关研究与实践活动提供科学指导，促进其合理开展与应用。

随着新技术（如人工智能等）的引入以及定量药理学学科的持续发展，群体药动学-药效学分析的重要性与应用范围将日益扩大。为确保持续的科学性、合理性与实用性，需依据科学判断深入开展研究，并根据学科发展需求，适时对本规范进行更新，为相关领域提供有力支持。

群体药动学-药效学分析标准操作规范

1 范围

本标准规定了群体药动学-药效学（PPK-PD）分析的内容架构、操作流程与技术要求，适用于药物研发及临床用药决策全过程中PPK-PD分析工作的开展。

PPK-PD分析通过整合药物研发各阶段、不同试验设计所获得的PK、PD或PK-PD数据，实现对目标PPK-PD特征的系统表征，识别并量化PK-PD变异性的来源，为不同场景下的剂量选择与给药方案优化提供科学依据。将PPK-PD分析与生理药动学模型、疾病进展模型、临床试验模拟等其他相关技术结合，可实现对多源信息的有效整合与管理，从而科学制定新药研发策略。此外，将PPK-PD分析与贝叶斯方法结合，可支持用药方案的制定与调整、用药依从性评估等，为临床用药决策提供定量依据。

本标准旨在规范PPK-PD分析的实施过程，确保其科学性、规范性及结果可靠性，为相关研究与实践活动提供技术指导。

2 规范性引用文件

本文件没有规范性引用文件。

3 术语和定义

下列术语和定义适用于本文件。

3.1 群体药动学-药效学分析

PPK-PD分析是建立在经典药动学-药效学分析基础之上，并与统计学模型相结合，考察目标群体中PK-PD的群体特征。“群体”指根据研究目的所确定的研究对象的集合。“群体特征”包括群体平均值或典型值，也包括由于不同个体在内因因素和外因因素等方面的差异所导致的变异。群体分析方法可定量解析群体中变异的大小及影响因素的作用。

3.2 基础模型

基础模型是指表征数据整体特征的模型，包括结构模型与统计学模型两部分。

3.3 最终模型

最终模型指在基础模型构建后，通过系统地筛选所有潜在协变量，最终仅保留具有临床意义的协变量、并经模型评价和优化之后所形成的模型。

3.4 模型评价

模型评价旨在客观评估模型在特定领域的预测能力和稳健性，并判断模型的潜在偏倚是否会对决策产生实质性影响。

3.5 固定效应

固定效应是一类特定的或可测量的、相对明确的模型参数。群体模型中，固定效应参数定义了结构模型参数的群体典型值，如清除率、分布容积、吸收速率和生物利用度等。固定效应参数还包括了协变量参数。

3.6 随机效应

随机效应是一类未知的、难以测量或无法观测的因素，以量化固定效应参数无法解释的变异或模型预测误差。

3.7 协变量

协变量是可能影响药动学和药效学的因素，包括内在因素（如年龄、性别、体重、种族、基因多态性、疾病类型、并发症、肝肾功能等）和外因因素（如合并用药、吸烟、饮酒、饮食等）。

3.8 个体预测值

基于个体经验贝叶斯法估计的预测数据。

3.9 群体预测值

基于群体参数估计的预测数据。

3.10 收缩

采用经验贝叶斯法估算时，个体参数趋向于群体典型值的度量指标。

3.11 条件权重残差

模型诊断的度量指标，综合考虑了模型拟合算法、观测值和实测值。

4 符号

下列符号适用于本文件。

BLQ: 低于定量下限数据

BSV: 个体间变异

DV: 因变量

EDA: 探索性数据分析

EBE: 经验贝叶斯法

IPRED: 个体预测值

NONMEM: 非线性混合效应模型计算软件

OBS: 观测值

OFV: 目标函数值

PD: 药效学

PK: 药动学

PK-PD: 药动学-药效学

PMAP: 群体建模分析计划

PMAR: 群体建模分析报告

PPK: 群体药动学

PPD: 群体药效学

PPK-PD: 群体药动学-药效学

PRED: 群体预测值

RUV: 残差变异

VPC: 可视化预测检验

5 标准操作规范

5.1 概述

群体药动学-药效学分析主要涵盖以下六大步骤：分析前准备、基础模型、最终模型、模型评价、模型应用和分析报告，见下图。此外，质量控制、风险管理和伦理应贯穿整个分析过程，确保分析全过程的准确、可靠和可追溯。

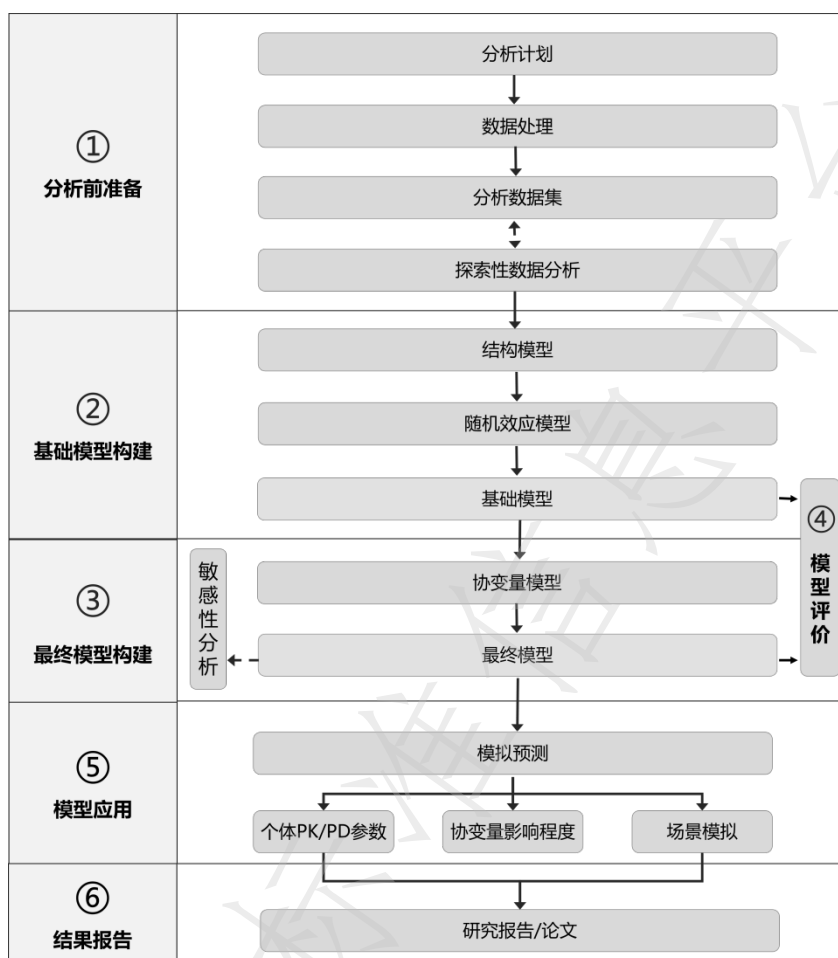


图1 群体药动学-药效学分析流程

注：虚线箭头代表在特定情况下所需的步骤，双向箭头则表示可能会发生的循环往复的步骤。群体药动学-药效学分析是持续的“学习与确认”（learn and confirm）过程，随着数据的不断积累，必要时应对模型进行更新和优化。

鉴于 PPK-PD 分析是“学习与确认”的迭代过程，故可根据分析目标，将 PPK-PD 分析分为探索性分析和确证性分析两大类。其中，探索性分析通常是不断积累数据和探索“药物-人体-疾病”相互作用规律的过程；而确证性分析一般指对药物研发或临床用药等决策产生重大影响的分析。

在新药临床研究与临床用药决策，不同类型的分析可具有不同的流程与评价标准。探索性分析具有较高的灵活性，其流程可根据具体分析目标进行适当调整，以便更好地挖掘数据中的潜在规律和特征。而在进行确证性分析时，应根据已有的分析数据和分析目标，明确分析类型和遵循的规范，严格按照标准规范流程开展工作，以保证分析结果的可靠性和可重复性。除非特别说明，下文均指确证性分析。探索性分析可以借鉴确证性分析的流程和规范，灵活地开展分析。

PPK-PD分析过程中，建议依据研究背景、分析目标、以及已获得的药物和疾病相关信息，科学合理地判断模型参数的重要性。重要参数是指对分析目标具有显著影响的参数，而非重要参数则是对分析目标无显著影响的参数。

5.2 前期工作

开展 PPK-PD 分析之前，应事先制订相应的群体建模分析计划（Population Modelling Analysis Plan, PMAP，以下简称“分析计划”），准备分析数据集，并开展探索性数据分析，为后续的建模奠定基础。其一般流程和参见图2。



图2 前期工作流程

注：双向箭头表示步骤间可能会循环往复

5.2.1 分析计划

通常，分析计划应包括以下内容：

- (1) 研究背景：提供药物研发的背景信息，包括药物的基本信息（如名称、剂型、规格、理化性质等）、药物的作用机制、拟开发的适应症和疾病相关信息、所处的研发阶段、前期非临床和临床研究获得的相关信息等。
- (2) 分析目标：待解决的药物研发或临床实践中的问题。
- (3) 分析内容：包括且不限于建立模型描述药物在研究对象中的 PK 或 PK-PD 关系、分析协变量对 PK 或 PK-PD 的影响、定量考察药物暴露量和疗效终点的关系、探索暴露量和安全性的关系等。
- (4) 研究假设：包括且不限于 PK 特征，PK-PD 之间的关系、协变量的影响等。
- (5) 数据来源：纳入研究项目的名称、研究设计、研究对象及其特征、给药方案、PK-PD 采样方案、生物样本分析方法等相关信息。
- (6) 数据准备：包括源数据集的格式（见 5.2.2.1）及其转换格式（如转换为.csv 文件）和方式、数据集变量的定义、数据集修改更新的标识、数据集质量控制规范和要求、数据集应包括的内容（如剂量、浓度-时间数据、人口学数据、潜在的协变量信息、PK-PD 相关信息等，并提供详细的数据集内容的说明）、以及数据的处理办法（包括且不限于缺失数据、低于定量下限数据和离群数据的处理方法等）。
- (7) 分析软件：包括专业分析软件 NONMEM、Phenoix NLME、Monolix 和统计分析软件 R、R package、SAS、Fortran、PsN、Pirana 等相关软件（含版本号）以及其在分析过程中所起的作用。
- (8) 分析方法：包括数据探索性分析、基础模型构建、协变量模型构建、最终模型构建，模型选择标准、模型评估方法等。此外，应对可能影响最终结论的假设进行详细说明。一般，不应限制分析过程中对模型假设进行的修正或增减。如对假设进行变更或调整，须在研究报告中进行说明和评价。
- (9) 模型应用：应用构建的模型拟解决的问题，如评估协变量的影响、模拟不同给药方案下的 PK/PD 变化、比较不同人群的暴露量等。

探索性分析时，由于对药物 PK 和 PD 的认识相对有限，PMAP 可视为一个动态的“活文件”，分

析方法的描述可适当简化。随着研发过程中对药物 PK 和 PD 理解的逐步深入，可进行相应的补充和优化。探索性分析时，PMAP 应重点明确分析目标和旨在解决的研发或临床用药过程中的关键问题，无需过多涉及技术细节。同时，分析人员在分析过程中应详尽记录 PMAP 的任何修订，以确保分析过程的可追溯性和透明度。

对于确证性分析，例如针对确证性临床试验等前瞻性研究的分析，以及用于指导药物开发和与监管机构沟通的分析，PMAP 应明确分析过程中的关键假设和评价标准，并在数据分析之前定稿，以确保分析的严谨性和规范性。

表 1 准备和提交群体模型分析计划的时间节点

分析目的	准备和提交的时间节点
<ul style="list-style-type: none"> ● 药物研发中的探索性分析 ● 探索性学术研究、用于指导临床用药决策 	开展分析前完成初步的分析计划，在分析过程中可根据分析目的、数据的可及性等进一步调整、更新计划。
<ul style="list-style-type: none"> ● 药物研发和监管中的确证性研究，供监管部门评审，以寻求共识，包括且不限于确证性临床研究，基于PK暴露量的推断等。 ● 确证性学术研究、用于临床用药决策 	研究分析前定稿，明确分析过程中的关键假设和评价标准。

5.2.2 分析数据集

在开展群体分析时，纳入分析的数据取决于建模目的。通常情况下，临床早期试验收集密集采样的数据，可用于建立准确的基础结构模型。而在晚期临床试验中，收集的稀疏采样数据包含了重要的协变量信息，适用于建立协变量模型。PMAP 需明确说明为达到分析目的所需的具体分析数据。

5.2.2.1 数据整理

建议遵循临床数据交换标准协会（Clinical Data Interchange Standards Consortium, CDISC）制定的、适用于群体分析、且符合分析数据模型（Analysis Data Model, ADaM）标准的基本数据结构实施指南（ADaM popPK Implementation Guide v1.0 | CDISC）。若未遵循相关指南，则应制定数据申请表（Data Request Form, DRF），以明确所用的变量及其定义和源数据变量等，以便数据的整理和编辑。在制订 DRF 时，应全面考虑数据分析的需求，纳入后续分析所需的指示变量和时间变量等，从而使后续的数据分析更为高效便捷。例如，采用 NONMEM 软件进行分析时，可纳入多次用药相关的 AD DL、SS 和 II 等变量。

5.2.2.2 完整分析数据集与验证数据集

在群体建模过程中，数据纳入策略需遵循“最大化信息利用原则”（Maximum Information Utilization Principle）。具体而言，应优先采用完整分析数据集（Full Analysis Dataset），系统整合所有可获取的研究信息进行建模。完整数据集可最大限度保留群体的异质性信息，如特殊人群、亚组信息值等，确保模型的预测稳健性。模型评估过程中，可用外部数据集，即建模过程中未纳入的数据，或其他研究产生的数据来考察模型的预测效能。

5.2.2.3 低于定量下限的数据

低于定量下限（Below Limit of Quantification, BLQ）的数据是指测量值低于方法的定量下限。在研究设计及执行过程中，特别是针对验证性研究，所采用的检测方法应满足研究目的的要求，并尽可能减少产生超出预期范围或可接受限度的 BLQ 数据。然而，BLQ 数据的出现往往难以完全避免。在数据分析阶段，应基于具体研究目的选择合理的 BLQ 数据处理策略，并评估该策略对分析结果可能产生的影响。

通常情况下，BLQ 数据的占比 $\leq 10\%$ ，且无系统性趋势分布，则可直接在参数估计过程中排除这些 BLQ 数据（MI 法）。如果 BLQ 数据的占比 $> 10\%$ 时，应首先排查研究设计、生物检测、受试者依从性等方面的影响。当排查后认为数据仍可用于分析，则应评估 BLQ 处理方法对模型参数估算的影响。另外，应注意某些情况下不应排除 BLQ 数据。例如，临床疗效的应答率直接从 PK 或 PD 数据衍生而

来，且应答率定义为 PK 或 PD 指标值高于（或低于）特定临界值时，将任何 BLQ 数据排除在群体分析之外，均可导致错误估算临床疗效应答率。在此情况下，应将所有 BLQ 数据纳入分析之中。

除了上述直接排除 BLQ 数据以外，常见的处理方式还包括如下方法：

- (1) BLQ 数据高于或低于定量限的似然（M3 法）；
- (2) 将所有 BLQ 值替换为定量下限的一半（M5 法）；
- (3) 每个受试者的首个 BLQ 数据设为定量下限的一半，舍去余下的 BLQ 数据（M6 法）。

根据分析目的，也可尝试其他 BLQ 数据的处理方法，如 M2 法（似然函数假设所有数值都在 LLOQ 处被删失）和 M4 法（BLQ 数据低于定量限且大于零的似然）。

相较于药动学 BLQ 数据的处理，药效学 BLQ 数据的处理需更为谨慎。药效学 BLQ 数据常为生物标志物数据（连续性变量）。一般情况下，应依据分析目标，结合药理学、生理学等理论和方法，参考上述药动学 BLQ 的处理原则，灵活处理药效学 BLQ 数据。BLQ 数据的处理方式须在分析计划中事先约定。同时，建议必要时采用敏感性分析等方法，证明 BLQ 数据处理方法的合理性。

此外，在编制数据分析文件时，应确认药动学和药效学定量下限的数值，特别是汇总多项研究数据时，不同研究的定量下限数值可能不同。在分析数据文件中应保留所有 BLQ 数据。

5.2.2.4 缺失数据

如果出现缺失数据，应尽可能查找。但有时缺失数据难以复得，须对缺失数据进行合理和科学的处理。缺失数据可包括缺失采样时间数据、缺失药动学或药效学观测数据、缺失给药数据和缺失协变量数据。

在试验过程中，时间应采用 24 小时制记录、并按照日期和时间的顺序记录。缺失采样时间数据时，由于无法计算给药后时间，故应在试验中及时进行数据核查减少此类错误。给药数据的缺失可能来源于临床试验时的记录错误，应重新核查记录。确认数据的过程中，分析人员应与试验执行者、数据管理者与程序设计师等相关人员共同商讨。

对于缺失用药时间、用药剂量和采样时间的数据，常采用原定的给药方案或者采样时间来替代缺失数据。但该方法可能会引入参数估计的偏差。然而，缺失药动学或药效学观测数据，应被直接舍弃。

对于协变量缺失数据，应重点关注缺失值的出现是否随机，尤其应关注缺失值是否在特定人群或场景下频繁出现。可接受的缺失比例取决于协变量的性质（例如，连续型或非连续型、随时间变化型或稳定型等）和协变量对分析的重要程度等。对于缺失率大于 10% 的协变量，可考虑不对其进行协变量分析。对于缺失率小于 10% 的协变量，常用的处理方法包括删除所有含缺失协变量的记录、或对缺失协变量进行填补等。常见的填补方法包括：

- (1) 对所有缺失值采用群体的中位数或平均数进行填补，当涉及特定协变量（如体重）时，应根据性别选择不同的中位数或平均数；
- (2) 对于不随时间变化的协变量，可用个体历史数据作为填补值；
- (3) 对于随时间变化的协变量（如体重、肝肾功能、年龄等），可采用前后数据的插值或最近一次的观测值进行填补；
- (4) 根据协变量之间的相关性进行填补。
- (5) 此外，还可采用多重插补法（Multiple Imputation）进行数据填补。即通过生成多个合理填补的数据集、分别分析后汇总结果的统计方法，其核心在于保留缺失不确定性并减少填补偏差。

不恰当的缺失值处理可能导致参数估计出现偏差，进而影响研究结论的准确性。在临床试验过程中，应采取多重措施减少缺失值的发生。此外，缺失值不仅会增加结果解释的复杂性，还可能对研究的科学性和可信度造成负面影响。因此，在制定分析计划时，应提前考虑缺失值的处理策略，并根据需要进行敏感性分析，以评估缺失值可能造成的影响。同时，缺失数据的处理方式以及开展的敏感性分析结果，均应详细记录于群体建模分析报告（PMAR）中。

5.2.2.5 对数转化数据

除了采用原始数据进行建模以外，当数据范围跨多个数量级时，可考虑对数据进行对数转换。对数转换可增加模型与参数估计的稳定性。

5.2.2.6 非计划数据

非计划数据指未按照研究计划采集的数据，如偏离采样时间窗等方案偏离数据。非计划数据处理方

式与按照研究计划采集数据相同。临床研究中如果收集了非计划数据，经确认数据无误后，一般也须纳入分析。如果非计划数据不纳入分析，则应在报告中说明原因。非计划数据对模型参数估算的影响可通过敏感性分析予以评估。

5.2.3 探索性数据分析

5.2.3.1 一般原则

理解数据特征是进行 PK-PD 分析的基础，因此开展群体分析之前，首先应进行探索性数据分析（exploratory data analysis, EDA）。EDA 是通过可视化分析和统计学分析，探索和揭示数据的特征、为后续的建模或决策提供基础。开展 EDA 的意义包括：

- (1) 直观地呈现数据的分布特征，初步判断结构模型和辨识潜在的离群值；
- (2) 确认协变量的分布范围、协变量之间相关性，以避免协变量分析中的混杂效应带来的偏倚；
- (3) 比较不同亚组之间的数据差异，有助于初步评估可能影响 PK/PD 特征的因素；
- (4) 初步验证模型假设是否合理，及时调整模型假设。

可视化分析即采用图表化的方法对数据进行检视，考察数据集的特征。基于 PK-PD 分析数据的类型，常用的可视化分析参见下表 2。

表 2 常用的可视化分析

数据类型	制图类型	示例
连续型	散点图、折线图、箱线图、直方图、密度图	血药浓度、血压、肿瘤大小、血糖水平
离散型		
计数	条形图、点图、泊松分布图	哮喘加重次数、癫痫发作次数
有序分类	条形图、堆叠条形图、饼图、累积概率图、马赛克图	疼痛强度分级、不良事件分级
二元	柱状图、堆积柱状图、ROC 曲线、瀑布图	治疗反应（是/否）、不良事件的发生（是/否）
时间-事件	Kaplan-Meier 图、累积风险图、泳道图（Swimmer plot）、森林图	药效持续时间、无进展生存期

常用的统计学分析包括计算变量的中位数、范围、四分位数、算术平均数、算术标准差、几何平均数、几何标准差等，对数据特征进行统计学描述。

5.2.3.2 因变量

群体分析中因变量通常指药动力学指标（如血药浓度）和药效学指标（如血压、血糖水平、不良事件的发生、总生存期）。通过对药物及其主要活性代谢产物的血药浓度、有效性与安全性指标随时间作图，或血药浓度对药效学指标作图，可加深对药物 PK-PD 特征的理解。

5.2.3.2.1 药动力学

PK 数据常为血药浓度，应采用常规尺度和半对数尺度两种形式，对其随时间的变化进行可视化分析。每个受试者的血药浓度数据点采用连线的方式，描述血药浓度随时间的变化趋势。必要时，可按照首次给药后时间或末次给药后时间，对血药浓度随时间的变化进行分析，亦可根据亚群体进行分层分析。此外，用于内部建模和外部验证的数据也需作分层分析。

5.2.3.2.2 药效学

PD 数据可包括体内生物标志物水平、临床终点指标或替代终点指标，需基于研究目的和方法进行科学合理的选择。根据数据类型，PD 数据可为连续型数据和离散型数据两大类，具体示例见表 2。建议对药效的潜在影响因素，如研究对象的基线特征、药物暴露量等进行分析。以暴露量为例，基于数据类型，常用的可视化分析包括如下：

- (1) 连续型数据：绘制药效效应与暴露量的散点图，并添加观测数据的非参数回归曲线；
- (2) 计数数据：根据暴露分类，绘制计数频率直方图，比较不同暴露下的计数频数分布区别；

- (3) 有序分类数据：可按照事件等级（序数）分层，并对其对应药物暴露量四分法或三分位数作图（条形图或堆叠条形图）。此外，还可针对每个等级，考察其概率和暴露量的关系；
- (4) 二元数据：按有效/无效（疗效）或出现不良事件/未出现不良事件（安全）进行分层，对相应的药物暴露量进行作图（如箱线图）。此外，还可对暴露量四分位数或三分位数分组，将观察到的反应概率与暴露量进行对比，并叠加单变量逻辑（logistic）趋势线；
- (5) 时间-事件数据：按暴露量四分位数或三分位数绘制 Kaplan-Meier 生存曲线。

5.2.3.3 协变量

协变量常为连续型协变量或离散型协变量。连续型协变量需提供描述性统计结果，并采用直方图或箱线图描述连续型协变量的总体分布等。离散型协变量亦需提供描述性统计结果，并计算每个协变量各分类的频数和频率。必要时，协变量可按照亚群体进行分层分析，并采用图表形式，探索协变量之间的相关性。如果纳入多个来源的数据进行分析，须比较不同数据来源中协变量分布特征的差异。

5.3 基础模型

5.3.1 定义和流程

基础模型包括结构模型和随机效应模型。其中，随机效应模型又称为统计学模型，包括了个体间变异模型、残差变异模型、场景间变异模型、以及其相应的协方差矩阵等。基础模型构建的流程见下图 3。

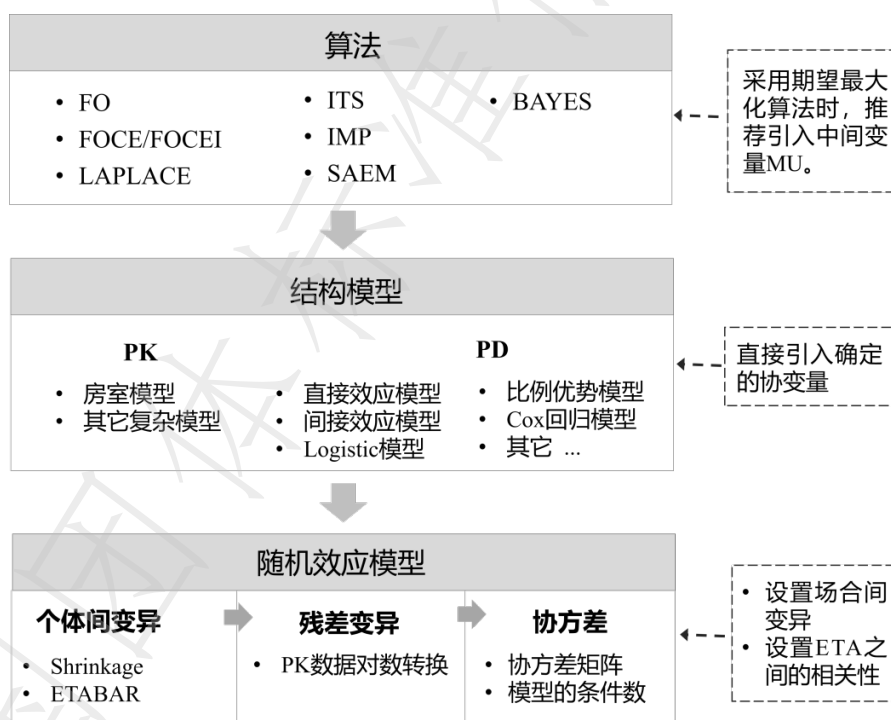


图 3 基础模型构建的流程

注：虚框部分代表在特定情况下的建议步骤。

基础模型的假设应科学合理，并且基础模型能表征体内药物浓度和PD指标随时间变化的过程。根据基础模型，可推断药物浓度和PD指标与时间的关系、以及随机变异（如个体间变异与残差变异）的大小。基础模型建立之后，方可进行协变量模型的构建。

5.3.2 算法

5.3.2.1 分类及特点

本文主要以 NONMEM 软件为例，介绍常用的参数估算方法。NONMEM 7 中的主要算法包括一阶

估算法 (first-order estimation, FO), 一阶条件估算法 (first-order conditional estimation, FOCE), 含个体间变异和残差变异交互作用的一阶条件估算法 (FOCE with η - ε interaction, FOCE-I), 迭代二步法 (iterative two-stage estimation, ITS), 重要性抽样最大期望法 (Importance sampling expectation-maximization, IMP), 随机逼近期望最大化法 (Stochastic approximation expectation-maximization, SAEM), 贝叶斯法 (Bayesian estimation, BAYES) 等。其中, SAEM、IMP 以及处理 BQL 数据的 M3 法, 均采用了基于二阶的拉普拉斯近似法 (LAPLACE)。

其他软件如 Phoenix NLME 中, 也使用 FO 算法。对于 FOCE 算法, NLME 又分为 FOCE-ELS 和 FOCE-LB 算法。其中, 前者与 NONMEM 中的 FOCE 相同, 而后者的计算速度更快。而与 SAEM 和 IMP 类似的算法分别为 QRPEM (quasi-random parametric expectation -maximization) 和 MCPPEM (Monte Carlo parametric expectation-maximization)。Monolix 中主要使用 SAEM 法, S-ADAPT 软件使用 MCPPEM 法。

上述算法均为参数法, 即假设参数呈正态分布或对数正态分布, 模型的解释性强、统计推断成熟、计算效率高、易于协变量建模, 且监管部门接受度高。但参数法强依赖分布假设, 如误设则易致偏差。

非参数法则可避免对数据分布的假设、模型误设的鲁棒性强、可更易辨识亚群、多峰/偏态分布。常用的非参数法包括: 非参数最大似然法 (non-parametric maximum likelihood, NPML) 和非参数期望最大化法 (non-parametric expectation maximization, NPEM)。但是, 相较于参数法, 非参数法计算复杂、解释性较弱、统计推断工具有限。常用的参数估算方法见附件 B。其选择取决于数据的多寡与模型的复杂程度。一般而言, 数据量越少则参数估算方法越重要。而对于数据量丰富的情况, 不同参数估算方法可提供近似的结果。

5.3.2.2 中间变量的引用

NONMEM 软件更新 EM 和 BAYES 算法后, 推荐引入中间变量 MU, 以提高 EM 和 BAYES 算法的计算效率。引入 MU 是一种处理 PPK-PD 模型中参数的标准化方法, 使模型参数在对数尺度上表达, 且也确保了模型结构的一致性。对于呈对数正态分布的参数, 如清除率 (CL) 或分布容积 (V), MU 法尤为适用。使用 MU 时应注意:

- (1) 除了残差变异外, 尽可能在 THETA 上使用 MU。
- (2) THETA 和 MU 之间的关系可以是线性、对数线性或其他形式, 取决于具体的参数化方式。以清除率 CL 的计算为例, CL 呈对数正态分布, 常见的参数化方式是先对 CL 进行对数转换, 即 $MU = \text{LOG}(CL)$, 然后将 MU 转换回原始尺度, 得 CL, 即 $CL = \text{EXP}(MU + \text{ETA})$ 。
- (3) 在 \$PK 或 \$PRED 中定义 MU, 而不在 \$ERROR 中定义。
- (4) 通常应避免条件限制, 即避免在 IF...THEN 结构内定义。如果必须在 IF...THEN 中定义 MU, 需确保在所有条件中明确赋值, 避免未定义行为。

5.3.3 结构模型的建立

建模的第一步是选择结构模型。通常应根据分析目的、数据类型和数据特征选择结构模型。同时, 模型假设包括体内动力学假设、药理学假设、生理学假设、疾病假设、数据处理假设和统计假设等应科学合理。

表 3 常用的 PK-PD 模型

数据类型	常用模型	示例
连续型	直接效应模型 (线性或非线性回归模型)、间接效应模型、效应室模型	血糖、血压、血脂、受体占有率、肿瘤体积
非连续型		
计数	Poisson 回归、负二项分布模型 (negative binomial distribution model)	不良事件发生、哮喘加重次数
有序分类	比例优势模型 (proportional odds model), 马尔科夫链模型 (Markov chain model)	疼痛严重程度、不良反应分级
二分类	Logistic 模型	有效/无效、客观缓解率
时间事件	生存分析: 比例风险回归模型 (proportional hazards model; Cox 回归模型)、参数法	生存/死亡、总生存率、不良事件发生率

注：当有序分类数据的分类数量较多（通常为6~10个或以上）时，其数值范围相对明确且连续性较强，因此在某些情况下可被视为连续数据进行建模。然而，这种简化处理可能忽视数据的离散性和有序性，从而导致模型的不合理性。此时，可考虑引入Beta分布。Beta分布是一种定义在区间[0, 1]上的连续概率分布，可通过线性变换扩展至任意有限区间。Beta分布由两个形状参数 α 和 β 控制，能够灵活地描述各种形状分布。同时避免将离散数据简单地视为连续数据所带来的偏差。

药动学结构模型常首先考虑经典的房室模型，若药物具复杂药动学行为时，也可基于科学合理的假设构建复杂模型。药效学结构模型的选择常基于药物的作用机制和药效学数据的类型。基于数据类型，常用的模型如表3所示。非连续型数据可采用LAPLACE算法进行分析。此外，药效学模型还可包含基线效应、安慰剂效应、疾病进展、耐药等因素。

由于PK-PD模型常较PK模型更复杂，故应根据建模数据和建模目的，选择合理的建模策略。在PK-PD的建模过程中，可采用同步拟合PK和PD参数的方法，也可采用序贯拟合，即先拟合PK参数再拟合PD参数。建模方法的选择须综合考虑参数的拟合精度与计算耗时之间的平衡。当药物的PK行为不受PD影响时，建议采用序贯拟合。由于许多生物大分子药物的PK可受到PD效应的影响，则应考虑采用同步拟合的策略。

另外，在PK数据缺失的情况下，除了建立剂量-效应模型外，还可以考虑建立K-PD模型。此类模型通过将药物浓度（缺失）-时间关系假设为简单的一级动力学过程（虚拟一室模型），在一定程度上保留了药物浓度-效应关系模型的性质。因此，与剂量-效应模型相比，K-PD模型可以更好地描述在潜在PK影响下随时间变化的药物效应。然而，应注意PK-PD模型仍然是定量药物暴露-效应关系的金标准。在PK-PD关系的探索阶段，不建议使用K-PD建模分析的方法。

对于新型或先进治疗药品（如部分生物制剂、细胞治疗药物和基因治疗药物）的PPK-PD建模，亦应遵循与传统药物相同的基本原则。然而，鉴于这些药物独特的体内PK/PD行为，包括生物制剂的靶点介导非线性清除、细胞治疗药物的体内扩增-衰退动力学、以及核酸治疗药物的递送系统依赖性组织分布与基因调控效应延迟等，建模时需根据具体机制和特征，选择相应的结构模型，并充分考虑PK与PD数据变异的来源。

关于模型选择和评估的具体方法，参见5.5.4节。

5.3.3.1 参数化

建模过程中，通过数学形式对模型参数进行结构化定义的过程称为参数化（parameterization）。其核心目的是通过合理的数学表达，增强模型的生物可解释性、数值稳定性及协变量整合能力。建议对常见的PK参数，如CL与V进行参数化，以简化后续的参数拟合与协变量建模的过程。又如，建立二房室模型时，应将中央室体积与外周室体积参数化为 V_c 和 V_p ，而不用稳态表观分布容积，以便更好地发现与解释协变量效应。此外，应考察不同的初值对参数估算值的影响，防止参数估算时的局部最小化。

5.3.3.2 结构模型中的协变量

如果已知某些协变量对PK或PD参数产生较大的影响，则建议在构建结构模型时引入这些协变量。例如，对于主要通过肾脏清除的药物，在结构模型中可考虑将该药的CL分为肾脏清除率与非肾清除率，并在肾脏清除率中引入肌酐清除率的影响，从机制上表征药物的消除方式。又如，药物代谢酶的基因多态性可影响药物的清除，则可将此引入结构模型。

此外，结构模型中还可考虑采用异速生长缩放法（allometric scaling）引入体重、去脂体重或理想体重等表征体型大小（body size）的指标。例如，体重数据跨度较大，或模型中同时包含儿童和成人的数据，分析时可先固定CL的体重幂指数为0.75，V的体重幂指数为1.0。之后，对幂指数进行重新估算。若可显著提高模型拟合度且参数估算值合理，则建议采用幂指数估算值。此外，成熟度模型（maturation model）也可用于儿童PK参数的估算。

5.3.4 随机效应模型的建立

5.3.4.1 一般原则

PPK-PD建模的最大特点之一是将预测值和实测值之间的差异分为两层，分别进行估计。第一层差异是个体参数偏离群体典型值的差异，也称为个体间变异（BSV）。BSV符合均值为零，方差为OMEGA的正态分布。第二层差异是个体内不可预测的随机波动，即残差变异（也称为个体内变异，RUV）。

RUV 符合均值为零，方差为 SIGMA 的正态分布。这两类差异均可用数学函数式表征，称为随机效应模型。评价随机效应时，还应考虑其随机效应的假设是否成立。建模过程中，随机效应模型可逐步完善。

5.3.4.2 个体间变异

建议采用指数模型来表示 PK 参数的 BSV，以保证 PK 参数均为正数。BSV 可加到数据支持的任一 PK 参数上。例如，CL 与 V 是最常引入 BSV 的 PK 参数。如果吸收相有足够的参数，则 BSV 也可以引入至吸收相参数（如一级吸收速率常数 k_a ）。当受试者同时有口服给药和静脉给药的数据时，可尝试估计口服生物利用度 F 及其 BSV。

对于 PD 参数的 BSV，可用加和型、指数型等数学关系式描述。但由于受试对象对药物的敏感性不同或其疾病状态不同，不同类型受试者的 BSV 可能不同，可分别设置不同的 BSV。

有时，BSV 估算值很小，甚至接近零，也并不代表该参数没有 BSV，而可能表明数据不足以准确估算 BSV，或者数据量少导致个体间变异的收缩（shrinkage）。此时，如采用最大似然法进行估算，则可将 BSV 固定为零。但是，BSV 固定为零时，期望最大化（Expectation Maximization）算法在 NONMEM 7 中较为低效。NONMEM 7.3 以前的版本建议将 BSV 固定到极小值维持算法的有效性。NONMEM 7.3 及之后的版本提供了在将 BSV 固定为 0 时仍然保留期望最大化算法有效性的方法。例如，在使用 SEAM 算法时，采用 \$ANNEAL 启动模拟退火算法，为固定为 0 的 OMEGAs 关联的 THETAs 提供全局搜索，获得最能反映真实数据特征的参数取值范围。

5.3.4.3 OMEGA 矩阵

根据研究目的、研究对象和药物特征，建议在引入结构模型协变量之后，再通过全 OMEGA 矩阵（full omega block）考察 BSV 之间的相关性。如果由于数据量不足，难以进行全 OMEGA 矩阵的拟合时，则建议至少考察 CL 与 V 的 BSV 之间的相关性。当参数的 BSV 间存在相关性时，如果采用仅有对角线元素的 OMEGA 结构，将忽略 BSV 间的相关性，导致模拟时产生不合理的参数组合。在构建协变量模型之后，应重新审视 OMEGA 矩阵，BSV 间的相关性可被某些协变量替代。

5.3.4.4 残差变异

PK 数据进行对数转换有利于残差变异模型的稳定。PK 数据对数转化时，常使用加和型或混合型模型。现行 NONMEM 软件并不支持直接计算个体加权残差（IWRES），可用“THETA”化的“SIGMA”计算 IWRES，或者采用 NONMEM 软件计算提供的 CIWRES、CIWRESI、CWRES 或 CWRESI。

PD 数据进行残差变异建模的过程基本与 PK 数据相同。但应确定 PD 指标是否可为负值。例如细胞计数等不可为负值。将此类数据进行对数转换，可确保模拟得到数据无负值。

一般，确定结构模型后再确定残差变异模型。在优化结构模型时，可同时优化残差变异模型。当所有的 PK 数据都进行对数转化时，推荐首选加和型模型表征 PK 的残差变异，若不能达到既定要求，则应考察其他类型的数学模型。

5.3.4.5 场合间变异

场合间变异（Inter-occasion variability, IOV）是一种随机变异，不应随时间推移而呈现趋势性变化。模型中是否纳入 IOV 取决于试验目的和设计，且应将 IOV 对参数的影响和参数的时变性加以区分。IOV 可在引入协变量前纳入模型，并在整个建模过程中均予以考虑。一般，通过研究设计和数据收集的特点来定义“场合”。当整合多个临床试验的数据时，应注意场合的定义，并合理解释 IOV 的影响。

5.3.4.6 协方差

协方差计算中最重要的结果是参数的协方差矩阵。观测值越多，协方差估算越准确。协方差估算能够定量表征参数估算的精确性。

协方差运算成功时，应审视变量间的相关系数矩阵。如相关系数接近 1，则说明模型存在过参数化（over-parameterization）且不稳定。协变量运算成功时也可能是获得了局部最小值，非全局最小值。此时，建议查看模型的条件数（conditional number）。如果条件数过大（例如 > 1000），则表明协方差运算结果可能不够准确，建议考虑其他方法计算参数的精确性。条件数的计算见 5.3.4.7 节。

协方差运算失败表明模型可能过参数化，应简化结构模型或者随机效应模型。有时，改变初值可使运算收敛，但应注意是否仅获得局部最小值。有时，受限于数据本身，协方差运算一直失败，但仍能提

供一些有用的信息，如参数拟合值等。在此情况下，参数的精确性可通过自助法获取，参见 5.5.3 节。

5.3.4.7 相关系数矩阵的特征值

协方差运算的另一个重要结果是相关系数的特征值。模型的最大特征值除以最小特征值等于条件数。条件数标志了模型的稳定性。当条件数 > 1000 时，模型很有可能过参数化。通过相关系数矩阵，可查找其原因。但是，条件数 > 1000 的模型也有价值。如针对不确定性的建模、采用有限的信息数据建立机制性模型等。

5.4 最终模型

5.4.1 定义和流程

最终模型指在基础模型构建之后，通过系统地筛选所有潜在协变量，最终仅保留具有临床意义的协变量、并经模型评价和优化之后所形成的模型。最终模型应能准确反映药物的 PK/PD 特征，量化协变量对模型参数的影响，并明确协变量与临床用药方案间的关系。

有时，无需进行协变量筛选（如分析数据集中不包含协变量），将不包括协变量的模型作为最终模型。此外，模型中的随机变异通常由个体间变异、残差变异和场景间变异描述。模型中引入的协变量可解释部分随机变异的来源。最终模型构建的一般流程见图 4。



图 4 最终模型构建的流程

5.4.2 协变量模型

5.4.2.1 一般原则

构建协变量模型前，须综合考量药物特征、研究对象的特征、临床前数据、早期临床数据、以及数据探索性分析结果，明确待考察的协变量。待考察的协变量常包括人口统计学特征、实验室检验结果等，如年龄、性别、种族、体型（体重、身高、BMI、去脂重量等）、肾功能、肝功能、基因型、药物制剂类型、合用药物、以及服药前是否进食等。

考察协变量前，应探讨每个待考察协变量的科学意义，包括药理学、生物学、病理学与临床意义等。例如，研究药物是通过肾脏清除的药物，研究对象包括了肾功能不全患者，则可将反映肾功能的肌酐清除率以连续变量或分类变量形式纳入至模型。

构建协变量模型时，应避免过参数化。过参数化将导致收敛失败与协方差运算失败。当两个或以上的协变量高度相关时，或者协变量分布不均衡时，应合理地选择协变量，避免共线性。例如，体重、体质指数、去脂体重和身高是高度相关的变量，常用于表征体型的大小。将上述协变量中的二个或二个以

上的指标纳入模型中可产生共线性问题。一般，常采用上述协变量中的临床常用指标进行后续分析，不宜将连续变量不合理地转化为分类变量进行分析。

对于二房室模型，协变量通常在中央室清除率（CL）与中央室分布容积（V_c）中纳入。尽管房室间清除率（Q）与外周室分布容积（V_p）较少受协变量的影响，但如有充足证据时亦可考虑纳入，例如体重等。

药物的 PD 效应可受到多种因素的影响，如疾病进展、生物反馈系统、耐药性、安慰剂等。为了描述药效随时间的变化，在 PD 建模过程中须考虑上述因素的影响。

在确定是否纳入协变量时，需综合考量统计学意义、药理学机理和临床相关性等多方面因素。即使某些协变量因样本量等限制缺乏统计学意义，但具有坚实的药理学机制支持，且具有明确的临床意义，依据研究目的仍可将其纳入模型。反之，当筛选出难以用现有理论知识阐释的显著协变量时，必须审慎解读相关结果，通常不应将其纳入模型。此外，协变量模型构建时，模型评价的方法和注意事项请参见 5.5 节。

5.4.2.2 协变量的筛选

建议采用系统化的协变量筛选方法，以确保协变量模型建立过程的统一性和一致性。选择协变量筛选方法时，应注意不同筛选方法的优势和局限性。此外，协变量筛选前，应确保基础模型的稳定性，并明确协变量的分布特征和协变量间的相关性，避免出现共线性的问题。应用较为广泛的协变量筛选方法包括以下两种方法：

- (1) 全量模型估计法（full model estimation, FME）：将所有具潜在影响的协变量同时整合入模型，然后作进一步协变量分析，构建最终模型。
- (2) 逐步法（stepwise covariate modelling, SCM）：SCM 包括前向纳入（forward inclusion）和逆向剔除（backward elimination）两个过程。前向纳入过程中，逐一考察和纳入协变量，形成全协变量模型（full covariate model, FCM）。逆向剔除过程是在全协变量模型基础上，逐一剔除协变量，形成最终模型。在前向纳入过程中，检验水平通常设定为 $\alpha = 0.05$ 或 $\alpha = 0.01$ ；而逆向剔除过程为降低假阳性风险，其检验水平设定更为严格，一般采用 $\alpha = 0.01$ 、 $\alpha = 0.005$ 或 $\alpha = 0.001$ 。在具体分析中，检验水平的标准设定需依据研究目的、样本量等多方面因素进行综合考量。

SCM 筛选协变量时，前向纳入过程是逐一考察协变量，故难以对特定组合的协变量效应进行评估。如果某些协变量组合对药物 PK 或者 PK/PD 特征有显著影响，可以基于协变量组合生成新的协变量，用于后续的分析。

相较于 SCM 法，FME 法侧重于模型参数估算值的合理性和准确性（包括相对标准误和自助法的置信区间），可直接评估协变量的临床意义，而不依赖于目标函数值（OFV）的变化等以数据为基础的评判标准。FME 法只需运行一次即可完成协变量的评估，结果易于解读。

此外，亦有其他协变量筛选方法，包括瓦尔德似然比检验近似法[Wald's approximation to the Likelihood Ratio Test (LRT), WAM]、广义加和模型法（generalized additive modelling, GAM）、拉索法（least absolute shrinkage and selection operator, LASSO）等。近年来，也有学者提出应用机器学习算法筛选协变量。附录 B 总结了多种筛选协变量方法的特点，供分析人员参考使用。

5.4.2.3 协变量模型的构建

协变量模型应能描述数据的整体趋势，准确表征协变量的影响。建立协变量模型的过程中，可使用加和型、比例型、指数型、幂函数型等数学关系式，确保参数估算值的合理性，如 PK 参数值应为正数。

协变量模型构建时可采用中心化方式。中心化的协变量能在不影响参数估算的情况下，有利于参数估算结果的解读，并提高参数估算的稳定性。一般而言，协变量的中位数（连续型）或者众数（非连续型）的数值可作为中心化的参考值。在协变量的数值范围跨度很大（如儿童与老年人的混合群体）时，应综合考虑数据特征和分析目标，确定中心化的参考值。

5.4.2.4 协变量参数的边界

一般而言，不应协变量参数设定边界值，而可用特定方式进行数据转化，如协变量的对数转化可使协变量参数值的更合理。有时，可设定边界来限定参数估算的范围，减少计算量，避免产生不合理的参数估算值。

5.4.2.5 协变量模型的充分性

使用全协变量模型构建协变量模型时，需确认全协变量模型是否为候选模型中 OFV 最低的模型。由于全协变量模型中纳入了最多的协变量，因此如果全协变量模型的 OFV 比基础模型高，则说明全协变量模型在运算中仅获得了局部最小值。

5.4.2.6 协变量模型的解读

协变量模型成功构建后，应谨慎阐释模型参数的统计学意义与临床相关性。针对协变量对参数的影响、以及模型估计值的置信区间解读，应遵循以下原则：

- (1) 当协变量对参数的影响未达统计学意义，且置信区间范围较窄时，需结合临床背景，判断该结果是否反映真实的无因果关联性，亦或样本量不足导致的 II 类错误。
- (2) 当协变量对参数的影响达统计学意义，且置信区间范围较窄时，应进一步评估效应量大小是否达到预先定义的临床界值，确认其临床意义。
- (3) 当协变量对参数的影响达统计学意义，但置信区间范围较宽时，应警惕估计精度不足带来的决策风险。建议通过敏感性分析验证结果的稳健性，并考虑扩大样本量或优化研究设计以提高参数估计精度。

应注意“窄”和“宽”的界定，需根据目标参数（如 AUC、 C_{max} ）的临床可接受变异范围，而非单纯依赖统计学标准。此外，无论协变量的影响程度有多大，均应说明协变量的临床意义以及相关原因。

结果呈现时，推荐采用森林图（forest plot）等进行可视化展示。森林图可整合效应点估计值、置信区间及预设临床界值参照线，有利于直观判断各协变量对 PK 或 PK-PD 参数的相对影响强度及其不确定性范围。此外，建议在图中明确标注临床决策阈值，以利于统计推断与临床意义的综合解读。

5.4.2.7 混合模型

当总体研究对象由若干个具有显著差异的亚群体（又称为“混合群体”）构成、且已收集的协变量信息无法准确区分时，可应用混合模型（mixture model）描述不同亚群体的 PK-PD 行为。但是，如采用某些协变量能准确分组，则应通过协变量进行区分，而不采用混合模型。

在 NONMEM 中，可通过为每个亚群体指定不同的特征参数来构建混合模型，估算个体属于每个亚群体的概率。例如：药物代谢酶具有基因多态性，导致两个代谢亚群体：快代谢组和慢代谢组，但是收集的数据中缺乏个体的基因多态性信息。此时，可采用混合模型估计两个亚群体的清除率以及每个亚群中的个体比例。构建混合模型时，应注意如下六个方面：

- (1) 模型复杂性：混合模型提升了模型的复杂程度，需额外的参数描述每个亚群体及其隶属的概率。
- (2) 参数可辨识性：区分亚群体时，可能由于参数值差异很小或者数据量不足估，可导致难以区分“亚群体”，而无法估算参数值，即参数不可辨识。
- (3) 计算收敛：随着模型复杂性的提升，混合模型的收敛计算难度增大。在数据量有限或数据噪音较大的情况下，寻找一个稳定且可靠的收敛解变得尤为困难。
- (4) 过拟合：过拟合是指模型在特定数据集上过度拟合，丧失了对新数据的泛化能力。因此，选择合适的模型并进行有效的验证是避免过拟合的关键步骤。
- (5) 模型评价：应注意自助法中亚群的比例限定以及可视化预测检验（Visual Predictive Check, VPC）的合理性。
- (6) 结果解释：解释临床亚群体形成的原因时，混合模型常面临困难，故应用时受到限制。

5.4.3 随机效应模型

加入协变量模型后，随机效应模型可能需要优化，其构建方法与基础模型部分相似，参见 5.3.4 节。

5.4.4 其他

5.4.4.1 离群值

离群值（outlier）指在同一个受试者中明显偏离其他观测值的异常观测值，或是某一受试者明显偏离其他受试者的观测值。建议将 CWRES 绝对值大于 5 或 6 的数据定义为离群值。在建模过程中，包括基础模型、全协变量模型与最终模型的构建过程中均应对离群值予以关注。离群值通常不应被剔除，但

有时离群值可影响模型拟合的成功，应予以剔除。剔除离群值后，应开展敏感性分析，并将结果记录于分析报告中。

5.4.4.2 参数化

在建模过程中，可遇到最小化失败的情况。而过参数化是导致最小化失败的常见原因。根据相关信息，可采取相应的解决方法。具体如下：

- (1) 梯度 (gradient)：在最小化过程中出现以下的梯度值是非正常情况，说明现有数据不足以支持某个参数的估算、或代码错误、或数据有误。
 - 初始迭代或最后一次迭代时，梯度为零
 - 梯度在极大值与极小值之间波动，例如从 $1.0E+07$ 到 $-1.0E+07$
 - 梯度较初值的变化很小，例如从 $1.0E+07$ 到 $1.0E+06$
- (2) 初值：在最小化过程遭遇失败时，可考虑采用重新设定初始条件的方法，探索参数的估计值。若识别出问题参数，通过调整这些参数的初始值，可促使最小化过程的成功。在 NONMEM 软件应用中，相较于 OMEGA 参数，调整 THETA 参数的初始值通常更为有效。
- (3) 协变量：如果协变量与参数之间缺乏明确的相关性，那么单纯地对协变量与参数之间的数学关系式进行调整，无法改善模型的不稳定性。
- (4) 边界值：若由于撞界致最小化终止，应检查梯度为零的参数，然后移除该参数或放宽该参数的边界值。
- (5) 舍入误差：如果最小化终止是由于舍入误差 (rounding errors) 引起，则可改变 NONMEM 输出的最低有效数字的个数 (number of significant digits, NSIG)。例如，如果参数中最小的 NSIG ≥ 2 时，则可考虑在 \$EST 中设置 NSIG=2，或改变对应参数的初始值。如果最小的 NSIG < 2 时，则模型可能过参数化。此时，可以考虑将 NSIG 低的参数固定或简化模型。
- (6) 目标函数值趋向无穷：若目标函数值趋向于无穷大导致错误最小化问题出现，则表明不存在普适性的统一解决方案。针对此类情况，必须依据具体问题的背景和条件，进行个案分析与判断。

5.4.4.3 先验信息

在 NONMEM 7 及更高版本中，引入了一种基于先验信息的贝叶斯建模方法，即 PRIOR 法。该法通过整合先验信息，改善模型的稳定性和预测能力。数据量相对有限时，PRIOR 法能充分利用既往的研究结果进行模型构建。采用 PRIOR 法时，应综合考虑以下几个方面：

- (1) 数据的质量：虽然 PRIOR 法可以在数据量有限的情况下改善模型的稳定性，但高质量的数据仍然是构建准确模型的关键。
- (2) 先验信息的合理性：先验信息应基于可靠的来源，如既往研究结果或专家共识等。不合理的先验信息可能会导致模型的偏倚。
- (3) 计算量：相较于不含先验信息的模型，引入先验信息往往需要更多的计算资源和时间。

此外，先验信息也可通过将参数固定为特定值的方式引入模型。无论使用何种方法引入先验信息，建议在可行的情况下进行敏感性分析，并对方法选择的依据进行解释、记录于分析报告中。

5.5 模型评价

5.5.1 定义

模型评价指标客观评价模型在目标应用领域中的预测能力，或判断模型偏倚是否会对决策产生实质性影响。模型评价是群体分析的重要环节，应贯穿于模型构建与应用的全过程。根据评价方式，模型评价通常可分为诊断图评价和数值统计评价。而具体评价的内容以及标准应根据研究目标，科学合理地进行选择。例如，若研究目标仅为描述数据并评估潜在的协变量效应，则模型评价内容可相对简化；而若模型最终用于支持给药方案设计的模拟，则需要更全面的评估内容和更严格的评估标准。此外，对于确证性研究，必须在分析计划中事先约定评价方式。



图5 模型评价

注：虚线箭头代表在特定情况下需完成的步骤。

5.5.2 一般原则

应依据研究目的客观对待评价结果。若模型评价结果显示构建的模型无法满足研究目的时，需对模型进行优化和调整或考虑补充试验数据。此外，应基于已知的生理学和药理学理论，对模型参数估算值进行初步评估，避免违反常识的参数估计值。例如，二房室模型中的中央室血浆药物分布容积一般应小于3 L等。

通常，PK模型的评价方法与采用连续数据的PD模型的方法一致。但在处理非连续PD数据时，由于PD模型在拟合过程中无法生成预测结果，基于预测的诊断图可能不适用。在这种情况下，通常采用基于模型模拟的方法来进行模型评价。例如，评估肿瘤患者的生存期时，可通过对比实际观测获得的Kaplan-Meier曲线与模型模拟的曲线来进行模型评估。

由于非连续PD数据涵盖多种类型，因此具体的评价方法也各有差异。应采用能反映其数据特性的评价方法，对比观测值与模型模拟值。在评价非连续数据的模型时，通常只能提供基于群体或亚群体层面的评价结果，而难以提供针对个体的评价结果。

在评估PD模型时，药物的药理特性也应作为考量因素。例如，对于使用国际标准化比值(International Normalized Ratio, INR)作为PD指标的抗凝药物模型，须能表征INR的最小值和最大值；对于部分表征抑制作用的模型，应验证其最大效应是否能超100%；而对于降糖药物的模型，需检查其最大效应是否与真实的生理状况相吻合。

由于PD数据通常比PK数据更为复杂且多变，如肿瘤生长数据往往呈现更大的异质性和非线性特征。因此，PD模型的诊断图往往难以呈现出一致的趋势性，基于预测值的诊断图可能表现不佳。PD参数也往往不像PK参数呈现已知的分布趋势。在这种情况下，应当谨慎解读PD模型的评价结果。

5.5.3 诊断图评价

5.5.3.1 一般原则

以评价指标分类，诊断图评价可分为以下四类：

- (1) 基于预测值的模型评价：如群体预测值（population prediction, PRED），个体预测值（individual prediction, IPRED）；
- (2) 基于残差的模型评价：如权重残差（weighted residuals, WRES），条件权重残差（conditional weighted residuals, CWRES）等；
- (3) 基于经验贝叶斯估计值（empirical Bayesian estimates, EBE）的模型评价；
- (4) 基于模拟的模型评价。

诊断图的应用过程中须关注以下问题：

- (1) 绘制诊断图时，应辅以必要的参考线和数据趋势线，在科学合理的条件下，分别采用普通尺度与对数尺度绘图；
- (2) 检视参考线和数据趋势线之间的关系，拟合不佳时，可优化结构模型或随机效应模型；
- (3) 关注是否违反非线性混合模型中随机变异分布的假设，包括：多峰、不对称、均值不为零或方差不齐等；
- (4) 离群值，必要时作进一步的敏感性分析和模型优化；

绘制诊断图时，可根据以下因素开展分层分析，便于比较模型或鉴别拟合不佳的影响因素：

- (1) 实验设计：实验中含不同研究组或不同治疗组时，可按照不同组别进行分层分析。
- (2) 协变量：根据是否纳入协变量进行分层分析、或者根据连续变量数值大小分为若干组进行分层分析，以检视协变量的影响。
- (3) 时间：如果研究周期很长，可根据时间分层，以审视不同时期的影响。

建模过程中，建议对于重要的模型开展以下诊断图评价：

- (1) 观测值-群体预测值（OBS-PRED）散点图；
- (2) 观测值-个体预测值（OBS-IPRED）散点图；
- (3) WRES 或 CWRES 对 PRED 的散点图；
- (4) WRES 或 CWRES 对 TIME 或者给药后时间（time after dose, TAD）的散点图；
- (5) OBS、IPRED 或者 PRED 对时间的散点图（相互叠加或者并排放置）；
- (6) IWRES 绝对值对 IPRED 的散点图；
- (7) IWRES、WRES 和 CWRES 的频数图或者 QQ 图；
- (8) 可视化预测检验（visual predictive checks, VPC）、预测值校正的可视化预测检验（prediction-corrected VPC, pcVPC）等。

对于基础模型和最终模型均需开展上述模型评价，建议并排展示基础模型和最终模型的诊断图，以便于检视协变量纳入之后的改进。最终模型还应包括以下拟合优度图：

- (1) 经典贝叶斯参数估计值（EBE）与对应 BSV 的散点图
- (2) 经典贝叶斯参数估计值（EBE）和对应 BSV 的分布图
- (3) 经典贝叶斯估算的个体 BSV 值对协变量作图
- (4) 个体拟合曲线

在使用经典贝叶斯参数估计值（EBE）进行诊断图评估时，须注意 BSV 和 RUV 的收缩效应。同时，应注意模型诊断图对缺失数据的处理方式。一般情况下，应先将缺失数据剔除后再做模型诊断图。如果在模型诊断分析时，缺失数据未被剔除且默认输出为 0 时，则应谨慎解读诊断图所呈现的结果。

5.5.3.2 基于预测的诊断图

5.5.3.2.1 观测值-群体预测值散点图

观测值-群体预测值（OBS-PRED）散点图可评估 PRED 能否准确描述数据的总体趋势。若 PRED 显示系统偏差，则模型难以准确描述数据的整体趋势。此时应考虑对结构模型或者统计学模型进行修正。

此外，将协变量加入前后的 OBS-PRED 散点图并列展示，可直观地检视协变量的影响。

除了模型误设 (model misspecification) 之外，PRED 在对角线两侧的分布还受到多种因素的影响，包括参数的变异程度和数据的删失情况等。此外，趋势线未涉及预测误差的方差齐性问题，也未考虑数据可能源自不同的个体。

5.5.3.2.2 观测值-个体预测值散点图

观测值-个体预测值 (OBS-IPRED) 散点图旨在评估 IPRED 是否能准确描述个体观测值。IPRED 在对角线两侧的分布仅揭示了残差变异的大小。若个体数据量不足，可导致模型过度拟合。故采用此诊断图作为评判依据前，重要参数 BSV 的收缩一般应 $< 30\%$ 。此外，IPRED-OBS 散点图仅反映 RUV 的大小，未充分考虑协变量对 BSV 来源的解释。因此，无论是基础模型还是最终模型，IPRED-OBS 散点图是相似的，主要用于评估基础模型的拟合情况。

5.5.3.2.3 个体药时曲线

个体药时曲线对于评价个体的拟合程度十分重要。观测值应均匀分布在个体预测曲线周围。PRED 和 IPRED 的药时曲线可交叠或者并排呈现。

5.5.3.3 基于残差的诊断图

基于残差的诊断图能对预测误差进行直观的评价。不同的残差诊断图能显示模型在结构或者统计学的不同特性。

5.5.3.3.1 加权残差-时间散点图

加权残差 (WRES) 对自变量的散点图可用来判断结构模型的合理性。对于 PK/PD 模型，时间是关键自变量，可以用药后时间 (time after dose, TAD) 或者是试验中一段持续的时间。如果时间跨度很大，则可将时间数转化后作图。如果采样密集，那么可以针对不同的给药时间段作图。例如首次给药后或者是末次给药后的时间。

基于 FO 算法估算的 WRES 值有时不准确，可导致错误地提示模型误设。而条件权重残差 (conditional weighted residual, CWRES) 或含交互作用的条件权重残差 (conditional weighted residual with interaction, CWRESI) 是直接基于 FOCE 算法估算而来，故一般应采用 CWRES/CWRESI 来代替 WRES。

5.5.3.3.2 加权残差-群体预测值散点图

CWRES/CWRESI 对 PRED 散点图优于 WRES-PRED 散点图。散点图中应有水平参考线与 CWRES/CWRESI 的平滑线。如果 PRED 的数值范围很广，则应考虑对 PRED 进行了对数转化后作图。

此外，建议采用条件权重残差 (conditional individual weighted residual, CIWRES) 或含交互作用的条件权重残差 (conditional individual weighted residual with interaction, CIWRESI) 对 IPRED 散点图，验证残差模型中的假设，包括均数为 0、以及方差齐性等。

5.5.3.3.3 加权残差的直方图与 Q-Q 图

CWRES/IWRES 的直方图可描述残差的分布。残差分布应为单峰、且均值为零的对称分布。当使用 CWRES/IWRES 直方图评估个体内随机效应的分布时，应注意 CWRES 直方图中不对称的现象可能会更严重。此外，应说明采用核心密度光滑线 (kernel density smooth line) 还是正常密度估计方法 (normal density approximation)。根据观测值数量，可对直方图进行适当分组。另外，IWRES 可通过 QQ 图进行正态分布审视。BSV 收缩值较大时，IWRES 会更偏离正态分布。

5.5.3.4 基于经验贝叶斯估计的诊断图

个体参数的估算值是通过经验贝叶斯方法综合先验群体参数、残差分布和个体观测值计算而来，也依赖于相关参数的个体间变异 (BSV)。

5.5.3.4.1 个体间变异和残差变异

EBE 向群体均值趋近的现象被称为 BSV 收缩。同样，当数据量不充分时，IWRES 的分布向零趋近

的现象被称为 RUV 收缩。当具临床意义的重要参数（如清除率）的 EBE 收缩超过 30% 时，基于 EBE 的模型评价结果（如 IPRED）可能不可靠。EBE 收缩还可影响协变量和暴露-效应关系的分析，因此须结合其他模型评价方法，综合评估收缩的影响。在对不同研究设计的试验数据进行分析时，有时仅有一个数据子集包含了特定参数的丰富信息。此时，应根据研究设计对数据进行分层分析，并计算各分组的收缩。

5.5.3.4.2 个体间变异和参数的散点图

分析过程中，可采用散点图考察 BSV 和参数间的相关性，并在协变量模型的建立过程中，应对 BSV 和参数间的相关性予以重视。

5.5.3.4.3 个体间变异的直方图与 Q-Q 图

BSV 的直方图和 QQ 图可检验参数的 BSV 是否呈正态分布，以及 BSV 模型是否合理。

5.5.3.4.4 个体间变异与协变量散点图

在构建协变量模型的过程中，建议绘制 BSV 与协变量的相关性图，探索参数与协变量之间的相互关系。其展示方式应依据变量的性质来决定。例如，连续型协变量宜采用散点图，而分级型协变量则宜采用箱型图。

5.5.3.4.5 其他

在计算参数时，可采用统计学检验 BSV 的均值是否等于零（ETABAR）。针对重要的 PK 参数（如清除率等），BSV 的均值不等于零且 p 值具有统计学意义时，则应进一步考察其原因，必要时优化或修正模型。

5.5.3.5 基于模拟的诊断图

5.5.3.5.1 定义

基于模拟的模型诊断是通过构建的模型生成模拟数据，评估模型的拟合优度。这种方法有助于确定模型是否能准确描述观测数据，并评估其预测能力。基于模拟的模型诊断具有多种形式，能够解决不同的问题。在执行时，需明确通过模拟解决的问题，以便选择恰当的评估方法。常用方法包括可视化预测检验（Visual Predictive Check, VPC）、数值预测检验（Numerical predictive checks, NPC）、正态预测分布误差（normalized prediction distribution errors, NPDE）、后验预测检验（Posterior Predictive Check, PPC）等。

5.5.3.5.2 可视化预测检验

5.5.3.5.2.1 定义

VPC 是通过对比观测数据与模拟数据的中位数和特定百分位数，来评估模型是否准确表征了建模数据的集中趋势和变异程度。VPC 过程中的模拟采用了最终模型参数估算值，不考虑参数的不确定性。此外，VPC 使用观测数据的相关信息（给药记录、时间、协变量等）来生成个体模拟数据，无法预测新的应用场景。

5.5.3.5.2.2 一般原则

进行 VPC 时，应遵循以下一般原则：

- (1) 模拟次数：模拟次数取决于评价目的。大多数情况下，只需模拟 1000 次。如果分析目标要求达到更高的预测准确度，则需要更多的模拟次数。
- (2) 百分位数：百分位数的选取依赖于每个分组内的观测值数量。例如，每个分组观测值的数量达 10-99 个，可选 80% (10%-90%) 区间，每个分组观测值的数量达 100~200 时，可选 90% (5%-95%) 区间，每个组内观测值的数量大于 200 时，可选 95% (2.5%-97.5%) 区间。
- (3) 分组：计算分位数时，需对自变量进行分组。一般根据时间进行分组，将采样时间相近的观测值合并同一组内。例如，如果将给药后 3.9 与 4.1 小时的采样观测值置于同一组内。但对于

口服药物的吸收相，则一般不宜将时间窗设置得过宽。当稀疏采样且仅有谷浓度数据时，可采用首次给药后时间，非末次给药后时间开展 VPC。

- (4) 分层：应对重要的协变量进行 VPC 分层分析，还可对重要的试验设计特征（例如剂量、试验人群等）进行分层分析，考察上述因素对模型拟合的影响。在此过程中，应注意不同分层中分组内的观测值数量。
- (5) pcVPC：如果根据多个协变量分层，可能导致 VPC 的诊断效力下降。此时，可采用 pcVPC，使用分组内中心时间点对应的数值来标准化观测值与模拟值。在模型中包含协变量时，推荐使用 pcVPC，并注意 pcVPC 中 y 轴的数值不是实际的观测值。
- (6) BLQ 数据：当存在 BLQ 数据的情况下，应对高于定量下限的连续型观测值和 BLQ 数据分别进行 VPC。
- (7) 此外，应注意上述因素是相互关联的，开展 VPC 时应进行综合考虑。选择 VPC 类型的依据须在分析计划和报告中明确阐述。

5.5.3.5.2.3 结果解读

VPC 的重点考察模拟数据与观测数据之间相似性，在结果解读时应注意以下问题：

- (1) $(100-n)\%$ 的观测值在预测区间之外；例如， $n=95$ 时，5% 的观测值在 2.5%-97.5% 的百分位数区间之外，其中分别有 2.5% 的观测值在 2.5% 之下和 97.5% 之上。
- (2) 在 $n\%$ 区间之外的观测值应随机分布，无显著聚集的现象。
- (3) 如果模型结构合理，那么每组内的模拟值与观测值的均值/中位数相近。
- (4) VPC 难以判断所有形式的模型误设。当缺失数据多或者采用分级数据时，可考虑使用其他形式的 VPC 或其他评价方法。
- (5) VPC 难以发现模型的过参数化。相反，参数越多，VPC 呈现的结果越好。

5.5.3.5.3 其他

除了 VPC 以外，还有其他基于模拟的评价方法，例如 NPC、NPDE、PPC 等。

NPC 的原理与 VPC 相似，但侧重于数值统计量的比较。NPC 通过基于待评价模型的模拟数据，构建多个预测区间（如 0%、20%、40%、50%、60%、80%、90%、95%），然后统计观测数据在预测区间以外的数量和百分比，并与理论值作比较。

NPDE 是基于整体预测分布的评价标准和检验方法。该法通过对预测差异（prediction discrepancy）进行标准化转换，考虑了同一个体内观测值的影响。若模型有效，则 NPDE 应服从标准正态分布。

PPC 也是一种基于模拟的模型评价技术。该法首先从模拟数据获取 PK 或 PK/PD 参数的后验预测分布，然后与原始数据进行比较。通常选取多个参数对模型的预测性进行评估。例如 C_{max} 、 C_{min} 、 t_{max} 、AUC、 EC_{50} 或 E_{max} 等。PPC 定义贝叶斯 p 值为后验预测分布（模拟）统计量 \geq 观测统计量的概率，以评估两者的相符程度。贝叶斯 p 值接近 0 或 1 提示模型可能无效。通过绘制直方图，可直观地呈现模拟数据与观测数据的分布特征，也可计算理论值与观测平均值发生偏倚的概率，评估模型的准确性。

5.5.4 数值统计评价

5.5.4.1 定义

数值统计评价也适用于整个建模过程，可评估模型的拟合程度、参数估计的稳定性，以及是否违反模型假设等。常用的数值统计评价包括目标函数值、参数估算值及其标准误、收缩值和自助法等。

5.5.4.2 目标函数值

目标函数值（OFV）代表模型与数据的拟合程度。针对同一数据集，OFV 越小表明拟合度越佳。此外，基于最大似然（Maximum Likelihood, ML）理论，嵌套模型间的 OFV 变化符合卡方分布，故 OFV 可用于嵌套模型间的比较。OFV 变化 > 3.84 （自由度=1，自由度是模型间参数数量的差异），表明 OFV 较低的模型在统计上更优（ $p < 0.05$ ）。但是，该假设依赖于算法（如 NONMEM），可能存在较大的 I 类错误。因此，仅依赖 OFV 可能会产生误导，应避免采用 p 值评价嵌套模型间的差异，须结合多种评价方法进行综合评价。

对于非嵌套模型，可用赤池准则（Akaike Information Criteria, AIC）对模型进行评估：

$$AIC = OFV + 2 \times p$$

式中 p 代表模型的参数数量。AIC 较小的模型被视为是较好的模型。

尽管 OFV 对模型选择的意义有限，但有时仍可提供有效信息。例如，纳入某个参数能使 OFV 降低 10 以上，则此参数可提高模型的拟合度。如果有充分的生理学、药理学等依据支持该参数的纳入，则纳入标准还可适当放宽。

5.5.4.3 参数估算值及其标准误

参数估算值应符合生理学、药理学等相关理论。一般情况下，具临床意义的重要参数估算值的相对标准误 < 30%，非重要参数估算值的相对标准误 < 50%。参数不可辨识或模型过拟合均可导致参数估计值的不可靠。

5.5.4.4 个体间变异的收缩

收缩表征了 BSV 趋近于群体均值的程度。> 30% 的收缩表明数据未含足够信息准确估计 BSV，导致个体预测值可能存在偏差。

5.5.4.5 自助法

自助法（bootstrap）又称为自举法，是一种常用的统计学方法。该法通过从原始数据集中有放回地抽取多个样本（即自助样本），生成与原始数据集大小相同的新数据集，再通过计算参数估算值的置信区间，用于评估参数的精确度和稳健性。自助法的优势在于无需对参数分布进行假设，即可获得参数的经验分布。有时，为了获得合理的参数分布，可根据重要协变量（如性别、合用药物等），进行分层，生成自助法数据集。但是，自助法计算量大，难以辨识复杂模型中不确定性的来源。一般，自助法每个参数的 2.5~97.5% 区间应包含原始参数估算值，且应满足预先设定的稳健率（如 > 80%）。

如由于计算时间过长等原因使自助法应用受限时，可通过抽样重要性重抽样法（Sampling Importance Resampling, SIR）估算参数的精确性。在 SIR 法中，可将 NONMEM 估算的协方差初步计算结果或少数次数自助法的结果作为初始点。如不输入初始点，可大幅增加 SIR 法的计算时间。

5.6 模型应用

模型应用时群体药动学-药效学分析的重要组成部分，常见的模型应用见下图6。



图6 模型应用

5.6.1 估算个体 PK/PD 参数

PPK 模型可用于估算受试者的个体 PK 参数（如清除率、分布容积等）支持后续的 PPK-PD 建模分析，也可用于估算受试者的个体暴露量（如 C_{max} 、 C_{min} 和 AUC 等），支持后续的暴露-效应分析。暴露量估算时应注意：

- (1) 用药方案的终止或改变;
- (2) 由于内在或外在因素的影响, 暴露量可随时间的变化而发生变化。

5.6.2 考察协变量对 PK/PD 的影响

PPK-PD 模型能够阐释内在和外在因素引起的 PK/PD 变异, 并评估协变量对 PK/PD 的影响。例如, 通过固定效应及其参数的估算精度, 可绘制反映协变量对 PK 和 PD 参数影响的森林图, 解释协变量影响和临床意义, 更精准地反映个体差异, 为个体化药物治疗方案的制定提供有力支持。

5.6.3 模拟

5.6.3.1 一般原则

模拟作为群体模型的关键应用之一, 可避免直接进行数学求解, 开展不同条件下的假设性分析, 预测药物的暴露量和效应, 评估治疗方案的有效性和安全性, 为提高药物研发效率、优化治疗方案、保障用药安全以及促进合理监管等方面发挥作用。开展模拟前应制定详细的计划, 确保分析过程的顺利进行、以及对模拟结果的充分认识。模拟分析过程中应注意以下内容:

- (1) 模拟的患者特征与真实患者的典型特征相匹配;
- (2) 模拟场景在临床实践中具有可行性;
- (3) 研究设计、样本量和研究终点须合理;
- (4) PK-PD 指标的设置须科学合理;
- (5) 根据模拟目标, 所采用的模型可以是包含变异的群体模型, 也可以是仅呈现群体平均水平、而不考虑变异的群体模型。

应依据不同假设条件下的模拟结果, 进行综合分析, 制定相应的决策。基于多学科团队对模拟结果的解读, 能显著提升决策的制定和完善。此外, 研究结果应尽可能通过图表形式展现和总结, 以便于不同研究背景专业人员的理解, 并为决策提供更丰富、更直观的量化支持。

5.6.3.2 类型

根据分析目标, 利用多维度信息, 执行模拟分析, 涵盖但不限于固定效应、参数的不确定性、个体间变异、场合间变异、残差变异等方面。

5.6.3.2.1 基于固定效应估计的模拟

若模拟分析旨在呈现典型受试者药物暴露或 PD 指标随时间变化的趋势, 则仅需依据 PPK/PD 模型的固定效应参数估计值进行模拟。然而, 基于这些典型参数的预测值并不代表研究群体的平均预测值。平均预测值是通过整合模型的固定效应、个体间变异以及残差变异, 进而模拟得出的平均药物浓度或 PD 指标随时间变化的曲线。

5.6.3.2.2 基于固定效应估计不确定性的模拟

若模拟分析旨在揭示典型受试者药物暴露量或 PD 指标达到或维持在特定阈值的概率, 或旨在阐释协变量对模型参数的作用, 须将参数估计的不确定性纳入考量。例如, 针对 PPK 模型固定效应及其参数不确定性的模拟分析, 能构建反映协变量对 AUC 或其他暴露量影响的森林图, 进一步理解协变量对药物暴露的影响。

5.6.3.2.3 基于随机变异的模拟

在进行模拟分析以展示个体药物浓度或 PD 指标的预测范围时, 须考虑模型的固定效应和个体间变异。若分析的目的是进一步揭示个体药物浓度或 PD 指标的实际观测值的范围, 则需额外考虑模型的残差变异。若目标是预测未来人群的观测值范围, 还应考虑个体间变异的不确定性。

此外, 在模拟分析中, 应考虑随机效应之间的相关性, 以避免受试者中出现与实际不符的参数组合。对于需要考虑个体间变异和协变量效应的模拟分析, 应在具有真实人口学特征的模拟群体中进行。人口统计学特征值可从相关数据库中获取、或者对原始研究中的受试者个体进行重抽样, 亦或通过目标人群中的协变量分布及协变量之间的相关性来生成。

5.7 分析报告

分析报告通常包括九大部分：摘要、研究背景、目的、数据、分析方法与假设、结果、讨论、结论、附录（图 7）。

分析报告		
• 摘要	• 数据	• 讨论
• 背景	• 分析方法与假设	• 结论
• 目的	• 结果	• 附录

图 7 分析报告内容

5.7.1 摘要

摘要作为报告的核心，必须以简洁明了的语言呈现，涵盖研究的背景、目的、主要假设、研究设计、数据与方法概览，以及关键结果和结论。此外，对于决策具有显著影响的图表也应包括在内。

5.7.2 背景

围绕研究目标，详细描述待分析药物的特性，探讨研究背景以及进行群体分析的意义。

5.7.3 目标

明确研究目标，涵盖主要目标和次要目标。若研究方案与研究目标存在偏差，则必须提供相应的合理解释。常见的研究目标包括：分析协变量对 PK/PD 的影响，通过模拟确定合理的用药方案等。

5.7.4 数据

详细阐明用于群体分析的临床试验信息，包括研究设计、样本量（例如受试者数量和血药浓度样本量）、采样方案、受试者状况、给药方案（如药物名称、剂量、给药途径、给药间隔、用药依从性）、药物浓度检测方法以及定量下限、药效评估方法等。提供人口学特征的统计学描述，并利用图表等方法探索协变量的特征及其相关性。阐述衍生变量的计算方法、数据格式、质量控制流程以及数据整理方式。若存在不同的分析数据集，需明确描述其间的差异，尤其是不被纳入分析的数据及其原因。对于原始观测数据的任何形式转换，必须详细解释，并提供充分的依据。

5.7.5 分析方法与假设

为确保分析结果的可重现性，应提供详尽的细节。建议首先概述分析流程和分析内容，提供总体的分析框架。然后，应详细描述处理各类缺失值、低于定量下限的数据以及异常值的策略和方法。在模型拟合方面，需阐述所使用的算法（例如 FOCE、BAYES 等）、假设（如参数分布假设）、模型选择的标准，以及所用的数据分析软件及其版本。

结构模型应通过示意图和公式详细展示。此外，协变量模型、协变量筛选方法和标准（如 OFV、临床相关性）、个体间变异、残差变异以及时间相关的随机变异模型的假设和选择标准也应详细说明。模型评价应具体说明所用的诊断图法和统计方法，包括敏感性分析等。对于复杂的模型化分析，应详细说明分析流程。模拟方案的描述应详尽，包括虚拟人群的生成方法等。在分析报告中，须重点阐述显著影响研究结论的假设，并对假设的任何改动进行详细说明和评价。

5.7.6 结果

通常涵盖结构模型、协变量筛选过程中的关键步骤，以及最终模型、参数估计值的精确度（如相对标准误差 RSE%、95% 置信区间）等。最终模型的评估和模型应用是重点，包括模拟、关键协变量对 PK 的影响、用药方案调整等。分析结果通常以图表形式展现，图表中的符号或缩写须作明确注解，以便于理解。有时可采用彩色图表以区分不同指标。结果应聚焦于研究目标和模型应用的内容，其他次要内容可置于附录中。

5.7.7 讨论

讨论部分应深入解释研究结果、而非重复描述建模分析结果，并在已有研究的基础上，阐述研究结

果的临床意义，如协变量对药物清除的影响而导致给药方案的调整等。围绕研究目标，评估构建的模型在解决问题方面的效能，包括用于建模的数据是否充分及其潜在的局限性；与既往相同或类似研究的比较等。

5.7.8 结论

应以简洁清晰的语言，描述核心结果，涵盖其所解决的问题，如对药品说明书撰写所提供的支持、用药方案的调整等。

5.7.9 附录

分析数据集的格式；基础模型与最终模型的代码、输出结果；分析过程中重要的图和表。

5.7.10 其他

如果研究报告拟提交监管部门审评时，还应考量评审关注的焦点，并确保涵盖相关技术指南中推荐的关键要素。美国食品药品监督管理局（FDA）的临床药理和定量药理评审人员，常运用基于问题的评审方法，针对下列问题进行评估：

- (1) 目前提供的临床药理学信息在多大程度上构成了有效性和安全性的关键性或辅助性证据？
- (2) 推荐的用药方案是否适用于目标适应症的普通患者群体？
- (3) 基于患者内在因素进行分组的亚群体，是否需要不同的用药方案和治疗管理策略？
- (4) 是否存在具有临床意义的食物与药物或药物与药物间的相互作用？若有，则如何解决？

分析报告应包含以下内容的相关证据或解决方案：

- (1) 支持药物的药理活性；
- (2) 识别不同患者亚组，包括具显著不同治疗效果的亚组、风险与获益平衡的亚组、具显著毒性的亚组，或治疗效果处于边缘状态的亚组；
- (3) 对于仅进行了一项充分设计且控制良好的临床试验的药物，利用剂量-效应或暴露-效应关系分析作为支持性证据，以评估药物的风险获益比；
- (4) 支持或证明用药方案选择的合理性；
- (5) 确定影响药物暴露和药效的内在和外在因素；
- (6) 为特殊人群、特定协变量亚群（如年龄、体重、肾脏/肝脏功能）的剂量选择和调整提供依据。

将研究结果写为研究论文发表时，建议遵循 *Clinical Pharmacokinetics* 杂志刊载的有关 PPK/PD 论文撰写的指导意见、以及国际医学期刊编辑委员会（International Committee of Medical Journal Editors, IJCME）发布的《学术研究实施与报告和医学期刊编辑与发表的推荐规范》等相关要求。

5.8 质量控制和风险管理

5.8.1 质量控制

5.8.1.1 定义

良好的质量控制（Quality Control, QC）确保了分析的准确性和可靠性，同时保障了数据、代码和结果的规范记录和保存，使整个分析过程具有可追溯性和重现性。

通常，分析过程的 QC 可以分阶段进行，并采用基于风险的方法来确定 QC 的范围。例如，根据分析过程中可能出现错误的概率、发现错误的难易程度以及错误对结果的潜在影响，对风险进行评估。在 QC 过程中识别出的任何错误均应在修正后，再次进行部分或全面的 QC。该过程应重复执行，直至所有 QC 均通过为止。此外，QC 人员应与分析人员保持独立。

5.8.1.2 一般流程

首先，对分析数据开展 QC，确认数据集的格式和准确性、并评估数据的完整性。面对大数据集时，如果时间和人力资源有限，无法进行全面的数据 QC，则应进行风险评估。

在模型开发完成后，需要对控制文件和输出文件进行 QC。在此阶段，应对分析过程或内容进行审

查，以确保分析过程遵循了分析计划。在模型开发过程中，由于时间限制，可能难以对所有模型都进行QC。此时，应重点专注报告中所讨论的关键模型，包括且不限于基础模型、协变量模型、最终模型和模拟预测时所用的模型等。

在最终阶段，进行分析报告的QC至关重要。此阶段涉及对分析结果的图和数据列表作详尽核查。若分析内容及报告需提交至监管机构，必须确保数据和报告符合相关法规要求。此外，推荐实施同行评审，以进一步评估报告的质量、内容、研究方法、分析的准确性和合理性。同时，建议由多专业背景的人员包括专业的QC人员、医学专家和定量药理分析人员，对研究报告进行全面核查。

QC过程中，发现的任何错误都应评估其对建模过程和结果的潜在影响。对结果无显著影响的小差错（如药物浓度值的记录错误），建议仅在报告的附录中注明。然而，如果错误对建模结果有重大影响（如重要协变量的记录错误），则须重新运行部分或全部模型。对于结果分析和报告中的错误，须重新绘制图表，并及时更新报告内容。QC的具体要点请参考附录B。

5.8.2 风险管理

风险管理是确保PPK-PD分析在药物研发和临床决策中可靠性和有效性的重要保障。风险管理应贯穿整个分析过程中，包括数据编辑、模型构建、参数估计、模型评价、模型应用、结果报告等各个环节。在风险管理的过程中，应加强多学科专业人员的协作，建立透明化的沟通机制，以确保分析的科学性、严谨性和实践应用的可行性。

风险管理主要包括风险的识别、评估、应对、监控和沟通五大方面。通过系统化分析潜在风险事件（如数据质量缺陷、模型假设偏倚等），量化风险发生的可能性与影响程度，对风险进行评估和分级；再基于风险评估结果，制定并实施应对措施，包括技术性措施（如调整算法、简化模型结构等）与流程型措施（如增加质量控制或模型评价环节等）。此外，须建立持续监测机制，以确保风险处于可控范围。同时，应向监管机构、临床研究者等相关利益方，及时、准确地报告风险管理情况，进行有效的沟通和协调，实现风险共担与决策共识。

5.8.3 伦理

PPK-PD分析过程中的伦理问题不仅影响研究的科学性，还直接影响受试者的权益和研究的公信力。因此，应将伦理考量贯穿于整个数据分析的全过程，严格遵守相关的法律和制度，如《个人信息保护法》，《通用数据保护条例》（欧盟），《健康保险可携性和责任法案》（美国）等，以确保研究的科学性、伦理性和社会价值。

- (1) 数据隐私与保密是核心伦理问题之一，涉及患者个体信息的保护。研究者必须采取加密、匿名化等手段，防止数据泄露，确保患者隐私不被侵犯。
- (2) 知情同意是临床研究的基本伦理原则，受试者应充分了解研究目的、潜在的风险和获益，并在自愿的基础上参与研究。
- (3) 研究诚信至关重要，任何数据篡改、选择性报告或伪造结果的行为都将严重损害研究的科学性和伦理基础。研究者应遵循严格的科学规范，确保数据的真实性和完整性。
- (4) 研究透明度与可重复性也是伦理要求的重要组成部分，研究者应详细记录整个分析过程，包括且不限于分析方法、数据来源及结果解释，以确保研究的可追溯性和可验证性。
- (5) 研究数据的共享可以提高科学研究的效率和质量，但同时也增加了数据泄露和滥用的风险。因此，在共享数据时，必须采取适当的保护措施，确保数据的匿名化和安全性。
- (6) 数据的二次分析可能带来新的发现，但需确保原始数据收集时的伦理考虑仍然适用。在使用患者数据进行新研究时，应尊重患者的隐私和权益。
- (7) 个体化治疗的应用场景下，PPK-PD模型可用于给药方案的设计和调整，但涉及患者的基因、生理和病理特征等敏感信息，需谨慎处理，避免对患者造成不必要的歧视或误解。同时，应在临床决策支持系统中设置人工复核机制，避免模型偏见导致的治疗不平等。
- (8) 研究人员对结果解读应负有伦理责任。不当的模型解读可致错误的临床决策。应建立风险沟通机制，对不确定结果进行清晰阐述，避免过度解读导致的患者伤害。

附录 A (规范性) 质量控制要点

以下的内容主要以NONMEM软件为例，其他分析软件亦可作参考。

A1 数据

常见问题
<ol style="list-style-type: none"> 1. 拟分析的临床研究应与纳入群体分析的数据集保持一致。 2. 所有符合群体分析标准的受试者都应被纳入分析数据集。 3. 每个研究对象的给药事件发生时间（日期）应记录于首个观测值之前。 4. 除备注行外，单元格不应含有字母或非数字字符。 5. 关键变量，如自变量和因变量，不应存在空白单元格，须填写完整。 6. 不应出现不合理的值，例如体重和年龄不应为负数。 7. 每个研究对象中不随时间改变的变量（如编号、性别等）应保持一致。
数据格式和单位
<ol style="list-style-type: none"> 1. 应合理选取变量的有效数字。 2. 每列数据的单位应保持一致。若数据源自不同的研究或多个研究中心，浓度、剂量以及协变量等数据的格式和单位应统一。若存在不一致，应已恰当地转换为研究计划中指定的单位。 3. 剂量和浓度的单位应匹配。如果不匹配，观察室比例因子（Fn）应作相应处理。 4. 当药物以盐的形式应用时，AMT 中的剂量应根据摩尔质量进行了校正。 5. 针对原药及其代谢产物进行同步建模分析时，AMT 值应作摩尔质量校正。 6. 针对大分子药物及相关受体建立靶点介导的药物处置（TMDD）模型分析时，AMT 值应已作摩尔质量校正。 7. 合并多个研究的数据时，药物浓度或实验室指标的检测方法应保持一致。如不一致，则应作标准化处理。 8. 当 RATE 通过 AMT/DUR 求算时，时间单位应确保正确。
自变量
<p>每个受试者中自变量（时间）应按照从先到后的顺序排列。</p>
因变量
<ol style="list-style-type: none"> 1. NONMEM 中因变量（DV）值应为同一类型。若否，则数据集中应有标识 DV 类型的指示变量，并且应用指示变量进行计算。 2. 若 DV 值包含不同类别的数据，则可通过绘制 DV 值与指示变量的散点图来核查指示变量。如果不同指示变量下的 DV 值出现重叠，则说明指示变量未能正确区分不同类别的 DV 值。 3. 应检查 DV 值中是否存在离群值。如果存在离群值，则应检查数据的准确性。在分析计划中应明确离群值的定义。通过基于分类变量或连续变量的分组，绘制栅栏图，有助于识别离群值。
协变量
<ol style="list-style-type: none"> 1. 每个协变量的范围（包括最大值和最小值）都应符合生理学标准。 2. 当协变量应表现出时间依赖性时，所有数据应遵循统一的规则进行标准化，例如是否考虑基线值等。 3. 随着时间的推移，若观察到协变量发生显著且非连续的变化，则这些变化可用于辨识错误或异常值。 4. 按 ID 绘制每个协变量经时变化的栅栏图，观察是否存在随时间变化的协变量。不随时间改变的协变量不应观察到此类变化。 5. 协变量的直方图分布若呈现多峰形态，需核实其准确性，单位是否存在错误标记。例如，部分研究对象身高的单位是厘米，而另一部分研究对象的单位是米。这样可致身高分布的多峰。 6. 分类变量的类别应与数据集保持一致。例如，某变量有四个类别，则该变量应有四个不同的值。
日期
<ol style="list-style-type: none"> 1. 必须核查时间与日期的一致性。研究日期应与试验设计相符。例如，若在 2010 年 1 月 1 日进行给药，而 24 小时后的采样日期标记为 2011 年 1 月 2 日，则存在错误。 2. 所有日期和时间的格式应统一。由于不同国家和地区可能采用不同的日期表示方法，例如 2012 年 4 月 3 日，在欧洲可能被记录为 3-4-2012，而在美国则为 4-3-2012。在合并数据集时需尤为注意。 3. 计算首次给药后时间的方法和单位应保持统一（例如，使用小时或天等）。

缺失数据和衍生数据
<ol style="list-style-type: none">1. 应检查是否存在值为零的协变量，且这些值应准确无误。2. 缺失值应已按照分析计划进行合理处理。3. 当数据被删除时（如异常值），应正确地对其处理和标识，并在报告中记录删除的具体原因。4. 对于 BLQ 数据，应按分析计划进行处理。例如，设定为零或最低定量下限的半值等方法进行填补。5. 衍生变量应按分析计划，采用合理的方法进行计算。
NONMEM 软件的特定用法
<ol style="list-style-type: none">1. 应在 EVID 列中恰当地标记给药记录（例如，EVID = 1）。2. 非给药事件应在 EVID 列中恰当地标记（例如，EVID = 2）。3. 应正确处理 DV 值缺失的情况。例如，DV 值缺失时，应在 MDV 列中用 1 进行标记。4. 给药记录应真实反映研究设计。例如，在使用 ADDL 和 II 时，应记录准确，并正确使用时间单位。例如，II 和 TIME 应使用相同的时间单位。5. CMT 应与事件发生的房室相对应。例如，对于一级吸收模型，吸收室 CMT = 1，中央室 CMT = 2。6. 对于静脉输注给药或零级吸收药物，RATE 应正确赋值。例如，静脉输注 AMT = 1000，RATE = -2，将持续输注时间定义为模型参数，并添加变量 DUR，在控制文件中设置 D1 = DUR。同时，还应注意时间单位的前后一致性。

A2 控制文件、输出文件和数据后处理

Fortran 编译器和第三方软件
<ol style="list-style-type: none"> 1. 记录所使用的 Fortran 编译器。 2. 运行 NONMEM 安装目录下的 CONTROL5, 检查输出结果是否与标准输出结果一致: 3. 正确安装并使用其他第三方处理软件, 例如 PDx-POP、PsN、Pirana 等。
控制文件
<ol style="list-style-type: none"> 1. 确保\$DATA 指定文件中的数据列的数量与\$INPUT 中输入数据集保持一致, 且数据列的顺序与\$INPUT 中的数据项顺序相匹配。 2. ADVAN 和 TRANS 选项与计划拟定的模型和参数设置相符合。例如, ADVAN1 对应静脉给药的一房室模型, 而 ADVAN2 对应口服给药的一房室模型。 3. 比例因子 (Sn) 和采样房室的容积单位保持一致, 如满足 $S_2=V_2$ 的关系。亦或应用了比例因子, 以确保剂量和血药浓度值单位的一致性。例如, 剂量单位为 mg, 体积单位为 L, 浓度单位为 mg/L, 则 ADVAN2 的 S_2 表示为 $S_2=V_2/1000$。 4. 在随机效应模型中, 如果 BSV-RUV 具相关关系, 则应在\$EST 中启用 INTERACTION 选项。 5. 在使用微分方程编写控制文件时, CMT 的值与房室相对应 (例如, 1 代表给药室, 2 代表口服给药模型的观察室)。 6. 在\$SUB 语句中的 TOL 值大于\$EST 语句中的 NSIG 值。 7. 如果残差模型采用双向转换法 (TBS, transform-both-side), 应确保因变量和残差模型均使用了相同的转换方法。例如, 如果 TBS 采用对数转换, 则因变量和残差模型分别为 $Y=f(\text{LOG}(F), \text{EPS}(n))$ 和 $Y=\text{LOG}(F)+\text{EPS}(1)$。 8. 使用多个\$EST 语句时, 估算方法及相关选项已在报告中明确体现。 9. 模拟时应注意以下问题: <ol style="list-style-type: none"> (1) MSF 文件已被用作输入参数。控制文件使用了 TRUE=FINAL, 或将参数固定为最终模型的参数, 并且这些参数数值正确; (2) 定义模拟的次数 (subproblems); (3) 定义模拟的种子数 (seeds); (4) 使用正确的概率密度函数 (probability density function), 如正态分布或均匀分布; (5) 后处理所需的参数已在\$TABLE 语句中完整列出。
输出列表
<ol style="list-style-type: none"> 1. 列表文件中呈现的受试对象总数和观察事件总数与数据集保持一致。 2. 列表文件内未含有任何警告或错误信息。 3. 列表文件所展示的观察事件数量与数据集中未被注释的事件数量相符。 4. \$COV 步骤已正确执行。若执行失败, 模型仍可视为可接受。 5. 最终参数估算的有效数字已确定。 6. THETA、OMEGA 和 SIGMA 的估算值在合理范围内。 7. 已估计标准误。 8. 收缩 (shrinkage) 在可接受的范围内。 9. 所有 ETABAR 的值在统计学上均无显著意义。 10. 最终模型的条件数小于 1000。 11. 列表文件中记录的起始和终止日期时间与输出文件中的日期时间相匹配。
输出文件
<ol style="list-style-type: none"> 1. 若未启用 FIRSTONLY 选项, 输出文件中的数据行数与数据集中的行数应相匹配, 同时应考虑\$INPUT 中由 IGNORE 或 ACCEPT 指令而被排除的数据。 2. 输出文件与控制文件中的时间戳在时间与日期上保持一致。 3. 输出文件与输入文件的数值精度保持一致。
数据后处理

1. 数据由后处理软件正确解读：
 - (1) 读取的数据行数与输出数据集的行数相匹配；
 - (2) 数据变量的数量与输出数据集的列数相符；
 - (3) 原始数据与输出文件的数值精度保持一致；
 - (4) 输出数据集中的变量，如 CWRES，与后处理软件中读取的数据相匹配。应注意变量名称可能不完全相同，但数据必须一致。
2. 核查报告中的附图：使用了正确的控制文件和数据进行绘制：
 - (1) 如果数据后处理软件是基于脚本的，脚本应有适当的注释，包括且不限于运行日期、控制文件的名称，以及用于绘制图表的 NONMEM 输出文件的路径；
 - (2) 如果数据后处理软件是基于图形用户界面的，应有稽查追踪记录，以确保使用正确的 NONMEM 输出文件来绘制图表。
3. 当 MDV=1 时，DV、PRED 和所有残差值被设定为 0，并且这些值在后处理前被正确地重新标记为缺失。
4. 使用正确的残差值进行绘图。例如，报告中的 CWRES 是指 CWRES 而不是 WRES。

A3 分析报告

报告结构
<ol style="list-style-type: none"> 1. 报告应该足够详细，确保能被第三方评估。 2. 建模过程中的每个假设和决定均应记录、讨论和证明。 3. 报告应该具有高质量，从而“可视为最终模型很好的描述了数据的特征，且结果和结论正确有效”。
报告质量
<ol style="list-style-type: none"> 1. 如有可能，需重新运行报告中的 NONMEM 控制文件并比较结果，以验证报告中的结果与 NONMEM 输出结果的一致性。 2. 应提供 NONMEM 和 Fortran 编译器的版本信息。 3. 在数据分析过程中使用第三方软件时，例如 PDx - Pop 和 PsN 等，应标明软件的版本。 4. 提供分析计划之外识别出的异常值。 5. NONMEM 中选择的 ADVAN 应与报告中的模型匹配，如 ADVAN1 对应一室模型，ADVAN2 对应口服吸收的一室模型。 6. 控制文件中的协变量模型应与报告中的模型相同。 7. 控制文件中指定的估算方法应与报告中使用的估算方法一致。 8. 最小化过程应成功完成，报告中应记录任何错误和警告。 9. 提供收缩值，若收缩超过预定义临界值，应在报告中说明。 10. 提供 ETABAR 统计结果，且应记录 $p < 0.05$ 的情况。 11. 提供所有可估计模型参数值及其标准误。 12. 应以统一形式报告方差项，例如方差的百分比系数或方差的对数值。 13. 提供最终模型的条件数。 14. 附录中的所有模型控制文件和输出文件应与报告正文保持一致。 15. 诊断图应与模型相匹配。 16. 如果进行了模拟，报告中应记录种子数和模拟次数。 17. 模型运行概况的表格应与 NONMEM 列表文件中的信息保持一致。
报告内容
<ol style="list-style-type: none"> 1. 早期 <ol style="list-style-type: none"> (1) 充分描述研究内容以支持拟解决的研究问题； (2) 报告提交监管部门时，应包含监管部门所要求的信息； (3) 具说服力和逻辑性； (4) 充分阐述研究背景以明确研究主题； (5) 关键信息应在突出位置呈现； (6) 文档设计和视觉效果应令人满意。 2. 晚期 <ol style="list-style-type: none"> (1) 数据应完整、准确且保持一致性； (2) 所有附图应明确标注并易于理解； (3) 文档格式应保持一致； (4) 语言表达应清晰且正确无误； (5) 所有分歧和矛盾应妥善解决。 3. 其他 <ol style="list-style-type: none"> (1) 如果部分文档已被审核，则在之后的审阅中应避免被重复审阅； (2) 避免总是从文档开头部分开始审阅，可从文档其他部分入手，防止对部分内容的过度关注； (3) 报告的撰写应在分析结束之前，或实时进行。研究目的、方法、数据特征等部分可以预先准备，为已确定的图表在文档中预留位置。分析完成后，只需将相应的结果或解释等插入即可； (4) 当报告的某些部分完成时，可进行阶段性审阅，而非待整个报告完成后才开始审阅。如果无法实时撰写报告，建议与团队成员讨论结果，并针对结果的解读和临床意义等达成共识； (5) 建议向审阅者提供具体的意见，指出审阅者无需专注于方法学，而应重点关注方法流程、结果等。同时，可明确告知哪些部分尚处于草拟阶段，不适宜审阅，或者哪些部分不需审阅等； (6) 培训审阅者如何进行审阅并提出建设性建议，如何从作者的角度提供修改意见，避免模糊的表述，详细阐述其建议，并提供明确的修改细节； (7) 建议制定审阅清单，并对文档整体提出意见，避免将重点放在语法上，提升报告的整体质量； (8) 建议采用文档共享编辑审阅系统，以管理审阅、修改意见和文档版本，从而提高工作效率。

附录 B (资料性) 参数估算的常用方法

方法	描述	优点	缺点	适用场景
参数法				
FO	忽略随机效应和协变量间的相互作用,且根据群体平均值线性化模型。	快速,计算量小	非线性模型准确性略低	低变异的数据,一般线性模型
FOCE/FOCE-I	通过调节个体参数来估算个体间变异。FOCE-I 估算考虑了个体间变异和随机随机变异之间的相互作用。	较 FO 更精确,对于非线性模型也更精确	较 FO 需更大的计算量	较为通用,适用于中等复杂程度的非线性模型
LAPLACE	与 FOCE 相似,使用二阶泰勒展开来提高精度的近似方法	较 FOCE 具有更高的准确性	较 FOCE 需更大的计算量	准确性高的非线性模型,并可用于离散型或分类性数据(概率模型)的拟合
ITS	二阶段迭代计算方法:一阶段为个体拟合,二阶段为群体拟合。	能处理复杂的模型以及高变异性	较前述方法需更大算力	个体和群体变异大的模型
IMP	采用加权模拟近似似然。	灵活,能够处理复杂模型	需较大算力,但计算结果可能不稳定	用于 FOCE 或 LAPLACE 不收敛的复杂模型
SAEM	基于模拟的稳健方法,通过随机抽样来近似似然。	能够处理复杂模型与稀疏数据	需较大算力,且计算结果更稳定	数据稀疏或不完整的复杂模型,广泛用于 PK-PD 建模。使用时常结合 IMP 算法。
非参数法				
NPML	非参数最大似然法:离散点估计方法	亚群识别能力强、计算效率较高	结果离散化、高维计算挑战	代谢分型、多峰分布验证
NPEM	非参数最大期望值法:连续密度估计方法	连续分布描述、揭示复杂形态	计算量大、分箱影响结果	连续非正态分布分析、长尾分布研究
贝叶斯法				
BAYES	结合先验分布与似然估算参数。	能提供完整的参数后验分布	需较大算力,且需要先验信息	具先验信息,或完整概率参数描述

FO: first-order estimation, 一阶估算法; FOCE: first-order conditional estimation, 一阶条件估算法; FOCE-I: FOCE with η - ϵ interaction, 含个体间变异和残差变异交互作用的一阶条件估算法; ITS: iterative two-stage estimation, 迭代二步法; IMP: Importance sampling expectation-maximization, 重要性抽样最大期望法; SAEM: Stochastic approximation expectation-maximization, 随机逼近期望最大化法; LAPLACE: Laplacian approximation, 拉普拉斯近似法; NPML: non-parametric maximum likelihood, 非参数最大似然法; NPEM: non-parametric expectation maximization, 非参数期望最大化法; BAYES: Bayesian estimation, 贝叶斯法。

附录 C
(资料性)
协变量筛选的常用方法

方法	定义	优点	缺点
全模型估算法 (full model estimation, FME)	同时拟合所有协变量, 并作出推论	<ul style="list-style-type: none"> 一次拟合略去协变量建模过程 方便解读输出过程 一步可得最终模型 	<ul style="list-style-type: none"> 必须在事前仔细分辨协变量的效应
逐步法 (Stepwise covariate modelling, SCM)	包括正向纳入与逆向剔除二个步骤	<ul style="list-style-type: none"> 方便实施 (PsN工具包中含自动化处理工具) 	<ul style="list-style-type: none"> 选择性偏差 多次未经校正的比较 最终模型之前要得出一个全协变量模型
瓦尔德似然比检验近似法 (Wald's approximation to the LRT, WAM)	通过全协变量模型获得所有限制模型的 LRT 值。最终模型的确立基于 NONMEM 计算的 SBC 值	<ul style="list-style-type: none"> 协变量较多时, 估算拟合次数较少 	<ul style="list-style-type: none"> 需要全协变量模型的协方差矩阵
广义加和模型 (generalized additive modelling, GAM)	通过广义的线性加法, 将个体参数的贝叶斯预测值与协变量进行回归	<ul style="list-style-type: none"> 原理简单、实施便捷 	<ul style="list-style-type: none"> 无法避免收缩和参数间的相关性 无法处理随时间变化的协变量
拉索法 (least absolute shrinkage and selection operator, LASSO)	所有协变量须标准化, 使其均值为 0, 标准差为 1	<ul style="list-style-type: none"> 在数据集较小时, 比 SCM 能获得预测能力更佳模型 	<ul style="list-style-type: none"> 处理大数据集时无优势
基于相关性检验的条件采样逐步筛选法 (conditional Sampling use for Stepwise Approach based on Correlation tests, COSSAC)	利用基础模型获得的信息, 根据从条件分布中采样的个体参数与协变量之间的相关性, 选择待测试的协变量, 并交替进行正向纳入和逆向剔除。	<ul style="list-style-type: none"> 可获得与 SCM 算法相似的结果, 同时大幅减少运算次数; 基于条件分布中采样的个体参数进行参数-协变量之间的相关性检验, 不易受收缩的影响 	<ul style="list-style-type: none"> 不含随机效应的参数无法采用此法
模型构建的随机近似算法 (Stochastic Approximation for Model Building Algorithm, SAMBA)	基于最大似然估计的逐步模型筛选算法, 在每次迭代中, 利用当前模型估计群体参数, 从条件分布中随机抽样个体参数, 然后结合观测数据评估并选择模型的更新策略, 逐步逼近最优模型结构。	<ul style="list-style-type: none"> 采用线性模型的逐步法选择变量, 运行速度快, 适用于高维协变量 	<ul style="list-style-type: none"> 协变量数量较多或协变量之间具高度相关性时, 有时无法收敛至全局最优解

参 考 文 献

- [1] 国家药品监督管理局.《模型引导的药物研发技术指导原则》，2020年12月.
 - [2] 国家药品监督管理局.《群体药代动力学研究技术指导原则》，2020年12月.
 - [3] CDISC, <https://www.cdisc.org/adam-poppk-implementation-guide-v1-0>
 - [4] European Medicines Agency, EMA. Guideline on reporting the results of population pharmacokinetic analyses. 2007年6月
 - [5] U.S. Food and Drug Administration, FDA: Population Pharmacokinetics Guidance for Industry. 2022年2月
 - [6] Pharmaceuticals and Medical Devices Agency, PMDA: Guideline on Population Pharmacokinetic and Pharmacodynamic Analysis. 2022年2月
 - [7] Guidelines for the quality control of population pharmacokinetic-pharmacodynamic analyses: an industry perspective. AAPS J. 2012;14(4):749-58.
 - [8] Establishing best practices and guidance in population modeling. CPT Pharmacometrics Syst Pharmacol. 2013;2(7):e51.
 - [9] Kris M Jansen, Sarah C McLeay, Michael A Barras, Bruce Green. Reporting a population pharmacokinetic-pharmacodynamic study: a journal's perspective. Clin Pharmacokinet. 2014;53(2): 111-122.
 - [10] Rusten IS, Musuamba FT. Scientific and regulatory evaluation of empirical pharmacometric models: An application of the risk informed credibility assessment framework. CPT Pharmacometrics Syst Pharmacol. 2021 10(11): 1281–1296.
 - [11] 新药研发中群体药动学/药效学研究的一般考虑. 中国临床药理学与治疗学, 2019, 24(11):1201-1220.
 - [12] 焦正主编. 基础群体药动学和药效学分析 (第二版). 科学出版社. 北京. 2024年11月第2版
-