结构化数据医院——融合材料领域知识的结构化数据质量治理平台

V1.0

用户操作手册

文档编号：2025-R0374
最后修订日期：2025 年 7 月 14 日

目录

关于本文档

主题

欢迎使用结构化数据医院——融合材料领域知识的结构化数据质量治理平台的用户操作手册，该手册包含了需要了解和使用结构化数据医院——融合材料领域知识的结构化数据质量治理平台的相关知识。

本文档包含了如下章节：

- **第 1 章** 软件概述
- **第 2 章** 数据挂号
- **第 3 章** 数据体检
- **第 4 章** 数据上传

序言中介绍了使用该文档的一些帮助信息。

读者

使用结构化数据医院——融合材料领域知识的结构化数据质量治理平台的最终用户和系统管理员。

第 1 章　软件概述

本章概述　本章介绍了结构化数据医院——融合材料领域知识的结构化数据质量治理平台的功能和操作界面特点。

内容

1.1 模块介绍

在材料科学研究中，数据不仅支撑着新材料的发现与性能预测，也逐步成为推动材料智能设计的核心要素。然而，目前材料数据普遍存在评测标准不一、质量参差不齐等问题，严重制约了基于机器学习的在材料领域中的运用。因此，如何在不依赖大量实验或高成本计算的前提下，提升材料数据的质量，已成为材料智能研究面临的关键挑战。因此，我们研发了“结构化数据医院——融合材料领域知识的结构化数据质量治理平台”。该平台包含数据挂号、数据体检、数据诊治三大核心功能模块：首先用户在数据挂号模块中填写并提交相关数据与信息；随后，在数据体检模块中，系统将会从可溯源性、时间敏感性、完整性、一致性、准确性、均衡性、规范性、冗余性、洞察力九大维度对数据进行质量评估，生成面向机器学习的数据画像报告和数据体检报告；最后，在数据诊治模块中用户可以分别从九大维度开展定向数据质量治理，系统性提升数据质量。平台基于 B/S 架构，采用 Vue.js、Element UI、Flask 与 Spring Boot 等技术构建，具备良好的扩展性与可维护性。同时，平台内置多种基于 Python 开发的数据质量治理算法，为用户提供了全面、高效的数据质量管理服务。平台旨在为材料领域的机器学习研究提供可靠、高质量的数据支撑。

具体来说，结构化数据医院——融合材料领域知识的结构化数据质量治理平台设计与实现了如图 1-1 所示的数据挂号、数据体检和数据诊治这三大模块。

1. 数据挂号

此模块是一个用于实现上传结构化数据相关信息的工具。用户在此依次填写数据基本信息、用户信息和来源信息，完成结构化数据的上传。系统支持上传.xlsx 和.csv 格式的结构化数据。同时，包含数据模板预览和下载以及上传数据的预览。

2. 数据体检

此模块是一个用于对数据进行全方位数据质量体检的工具。平台首先计算 27 个元特征维度全面系统地刻画数据的特点，然后从可溯源性、时间敏感性、完整性、一致性、准确性、均衡性、规范性、冗余性和洞察力九大质量维度定性定量评估数据的质量，依次展示评估结果，形成数据画像报告和体检报告，并提供针对性诊治建议，引导用户前往相应科室进行质量提升。

3. 数据诊治

此模块是一个用于对数据进行全方位数据质量治理的工具。在数据诊治模块，平台将分

别从可溯源性、时间敏感性、完整性、一致性、准确性、均衡性、规范性、冗余性、洞察力九个治理维度对数据进行全面治理。

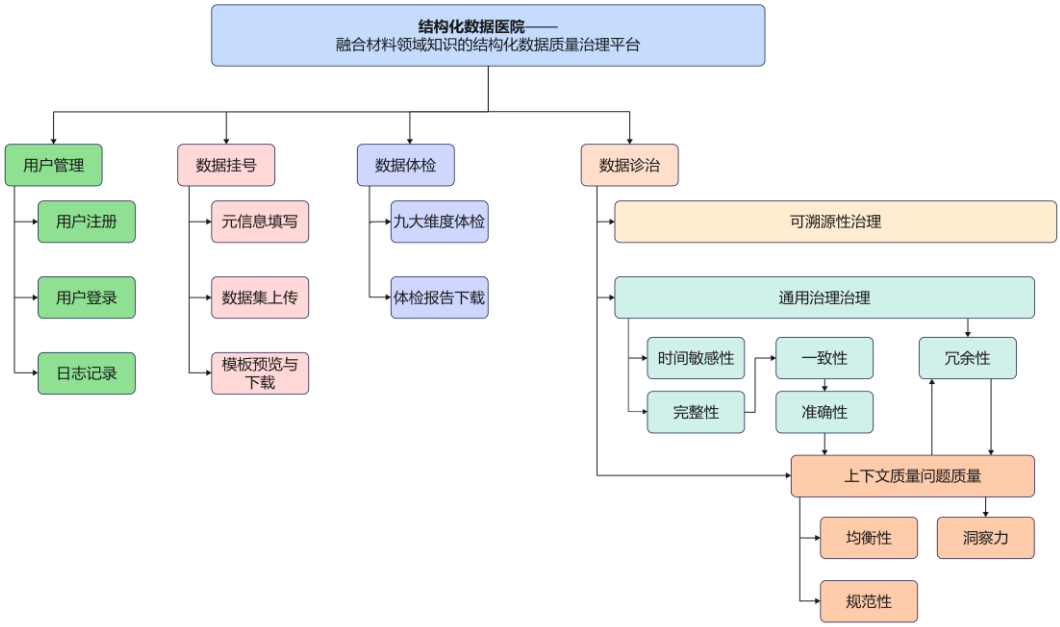

图 1-1 系统软件架构

1.2 用户层级、角色与相应权限

使用本系统的用户主要分为两个不同层级，即：管理员，普通用户，如表 1-1 所示。

表 1-1 用户层级

用户层级	用户名	密码
管理员	admin	admin123
普通用户	123@qq.com	111

不同层级角色的用户各自享有不同权限。例如对于同一功能，管理员可对其他用户信息进行编辑，而普通用户只可修改自身用户信息。

如在本手册阅读过程中，您发现所描述的内容与您所实际看到的内容不一致，可能是由于权限差异。

1.3 登录界面

如图 1-2 所示是结构化数据医院——融合材料领域知识的结构化数据质量治理平台的初始登陆页面。

对于用户，需选择对应账号及密码输入（表 1-1），以登录进入主界面。

图 1-2 登陆界面

1.4 主界面和导航栏

平台主界面由顶部区域、侧边导航栏模块构成，如图 1-3 所示。顶部区域包括平台 logo 和退出按钮。侧边导航栏包括数据挂号、数据体检和数据诊治三大模块，用户可根据研究需求选择对应模块开展具体工作。首先用户在数据挂号模块中填写并提交相关数据与信息；随后，在数据体检模块中，系统将会从可溯源性、时间敏感性、完整性、一致性、准确性、均衡性、规范性、冗余性、洞察力九大维度对数据进行质量评估，生成面向机器学习的数据画像报告和数据体检报告；最后，在数据诊治模块中用户可以分别从九大维度开展定向数据质量治理，系统性提升数据质量。

图 1-3 主界面

第 2 章　数据挂号模块

本章概述

本章主要介绍了结构化数据医院——融合材料领域知识的结构化数据质量治理平台的数据挂号部分的内容。我们提供了不同信息填写的表单以及模板预览与下载，让用户更好地上传数据。

本章内容

2.1 数据上传

（1）上传数据基本信息。在对应页面中输入数据的基本信息，包括数据集名称、样本数量、特征数量、目标属性、材料类别、获取方式、机器学习任务、时序特征名称（如果有时序特征）、数据集描述等，并上传测试数据集，如图 2-1 所示。

（2）上传用户信息。点击“下一步”，进入用户信息上传页面，在对应页面中输入用户的信息，包括提交者、所在单位、邮箱、电话等，如图 2-2 所示。

（3）上传数据来源信息。点击“下一步”，进入数据来源信息上传页面，在对应页面中输出数据来源信息，包括数据分类、数据来源等点击“提交”，存储所填写的信息，完成数据挂号，如图 2-3 所示。

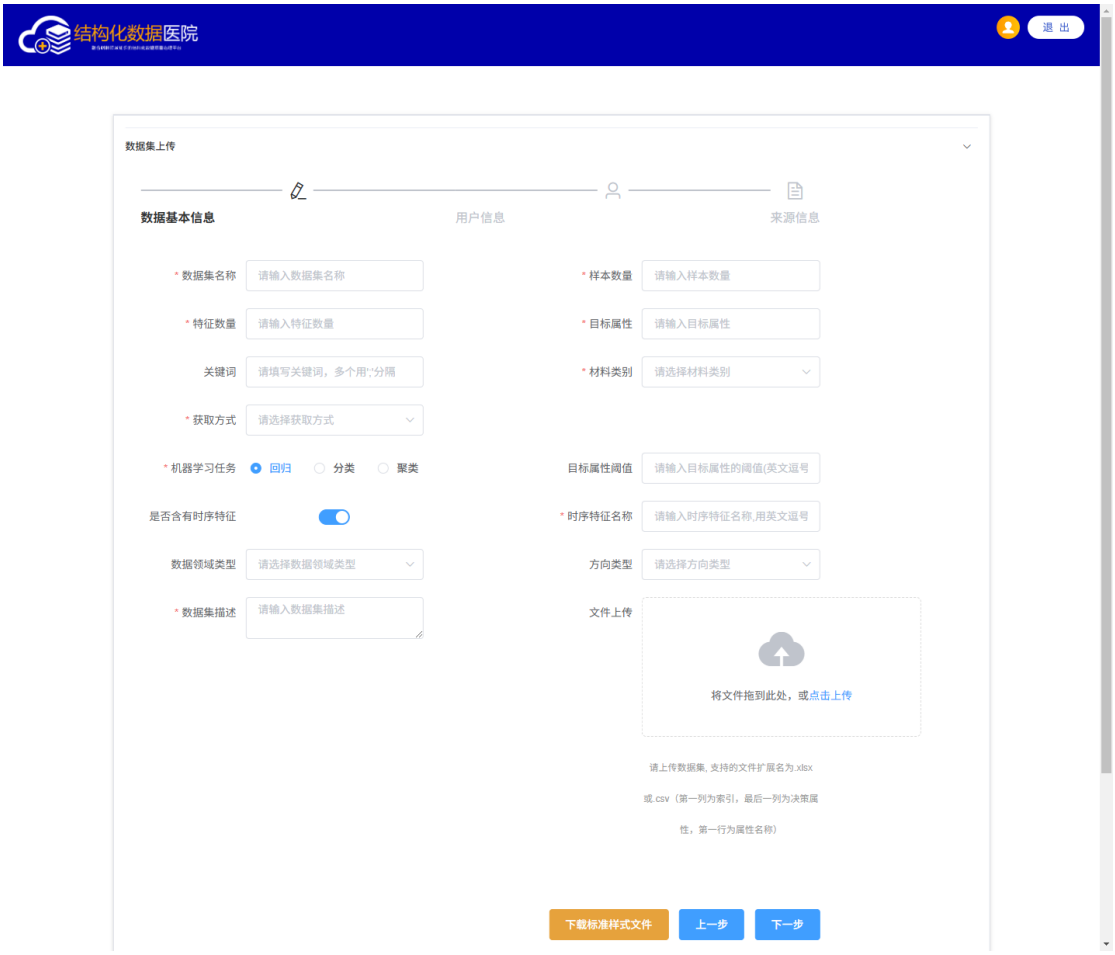

图 2-1 数据基本信息上传页面

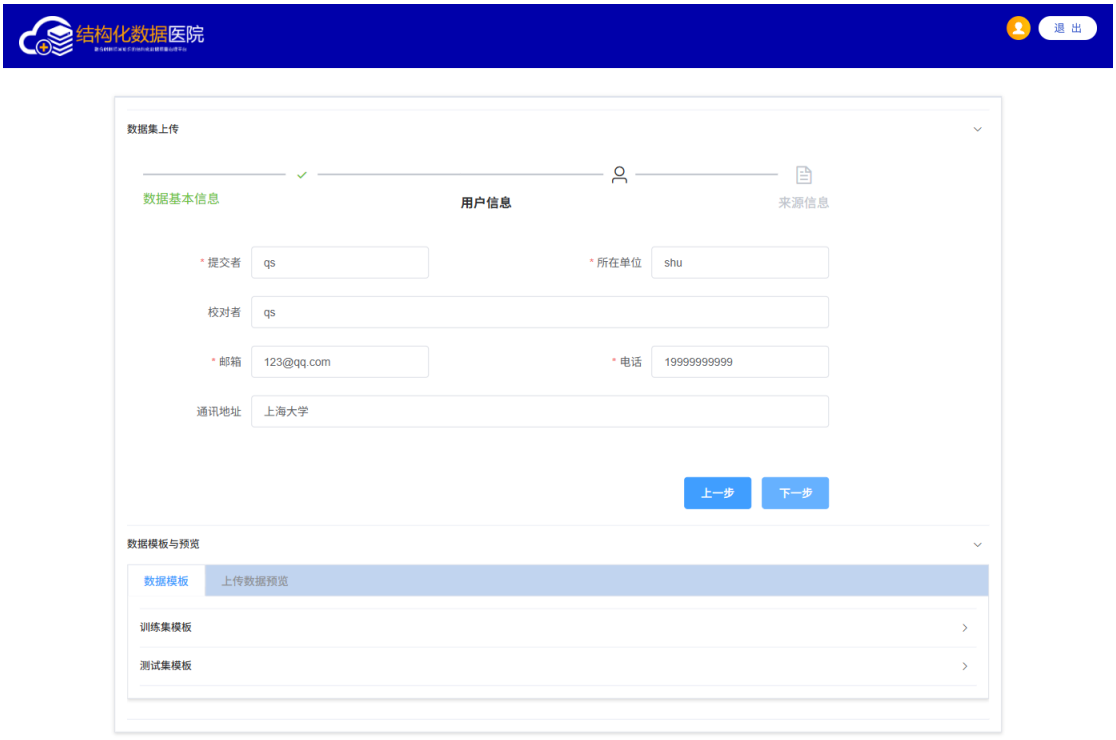

图 2-2 用户信息上传页面

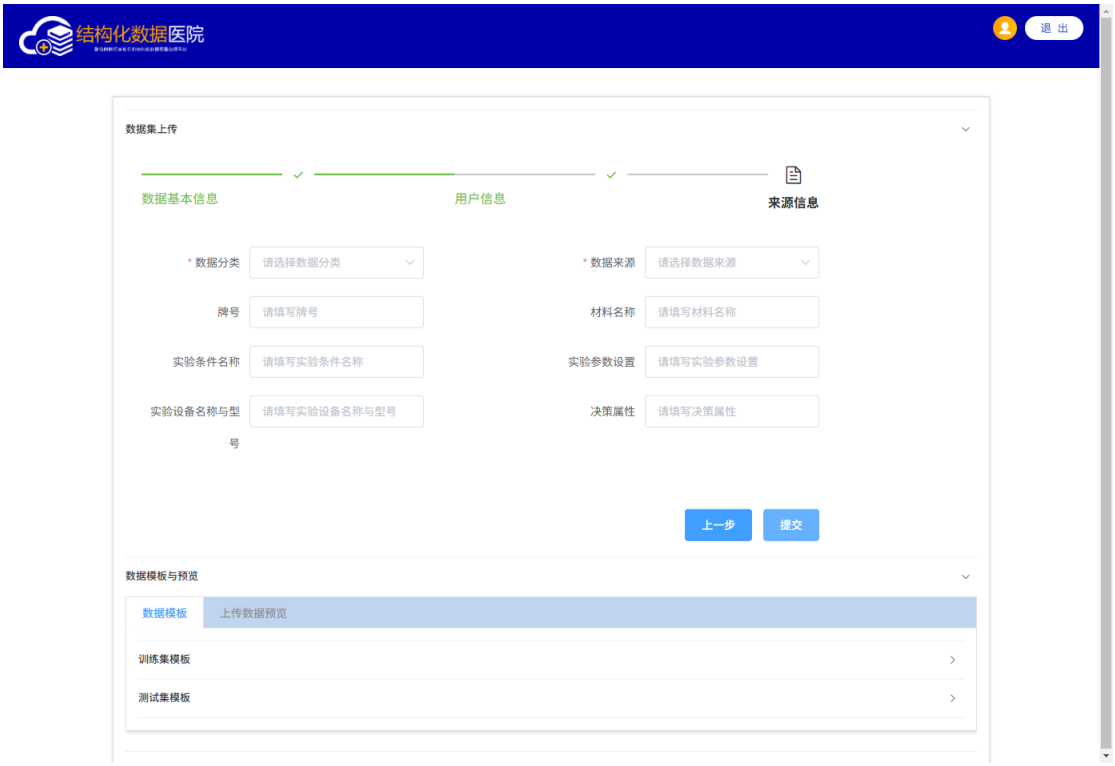

图 2-3 数据来源信息上传页面

2.2 数据模板预览和下载

用户首先在数据模板与预览区域点击左边“下拉”标签即可看到训练集和测试集的数据模板，如图 2-4 所示。点击“下载模板”便可进行数据模板的下载，如图 2-5 所示。

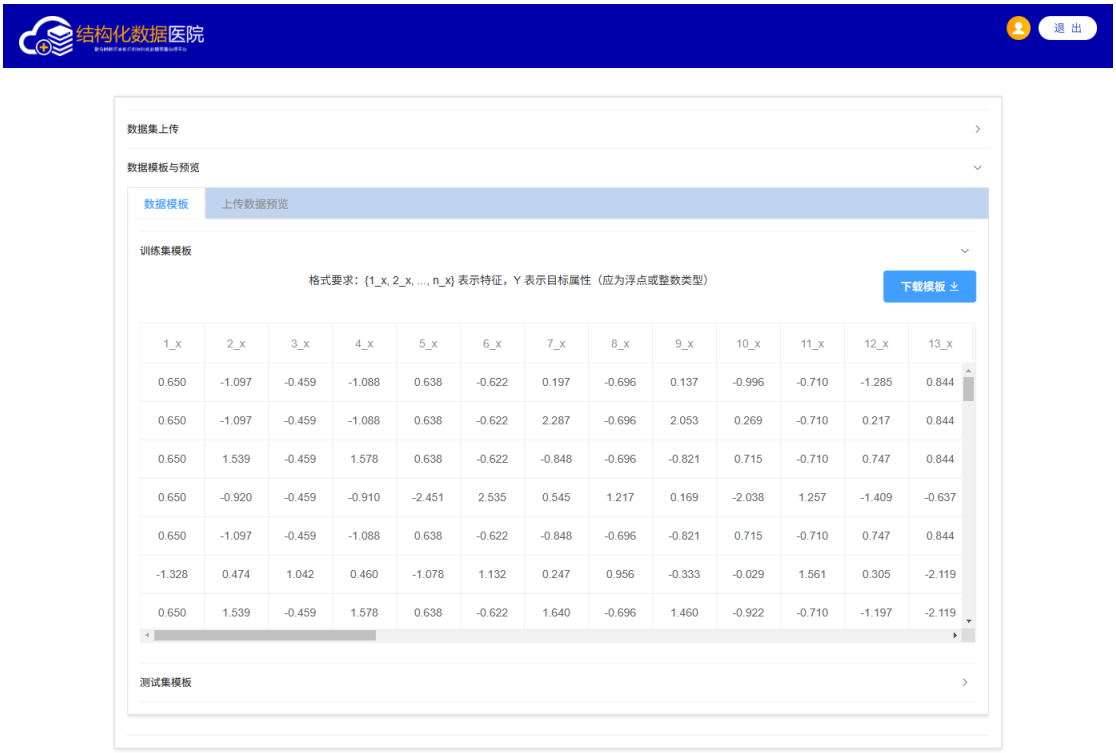

图 2-4 数据模板预览

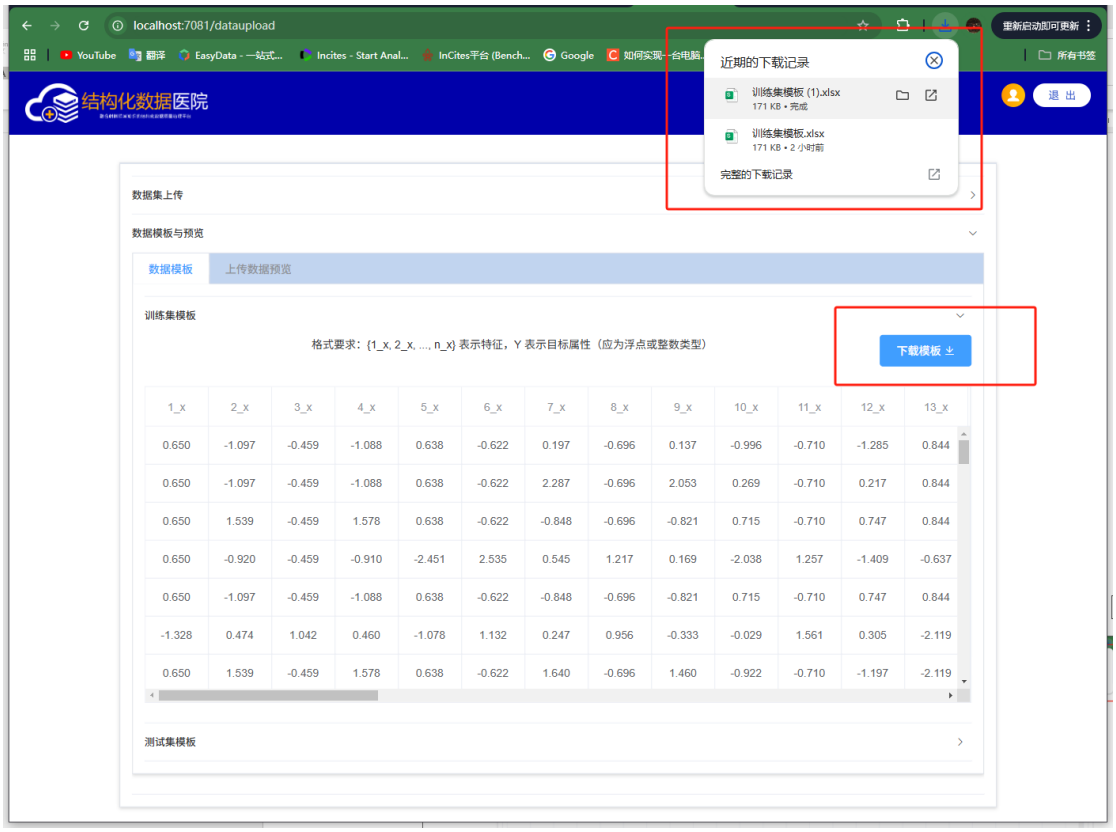

图 2-5 数据模板下载

第 3 章　数据体检模块

本章概述

本章主要介绍了结构化数据医院——融合材料领域知识的结构化数据质量治理平台的数据体检部分的内容。我们首先计算 27 个元特征维度全面系统地刻画数据的特点，然后从可溯源性、时间敏感性、完整性、一致性、准确性、均衡性、规范性、冗余性和洞察力九大质量维度定性定量评估数据的质量，依次展示评估结果，形成数据画像报告和体检报告。

本章内容

3.1 数据体检

点击系统主页面中“数据体检”进入数据体检页面。平台将依次从可溯源性、时间敏感性、完整性、一致性、准确性、均衡性、规范性、冗余性和洞察力九大质量维度定性定量评估数据的质量。此处不需要人工操作，系统将根据上传的数据自动计算并展现结果。

3.1.1 可溯源性

平台在可溯源性维度展示了元信息完整率，包括基础信息完整率、材料知识完整率和机器学习任务信息完整率。此处以 nasicon 数据集为例进行测试，运行结果如下图 3-1 所示，在数据挂号部分我们上传的信息通过数据体检评估可知其基础信息完整率、材料知识完整率和机器学习任务信息完整率都是百分百。

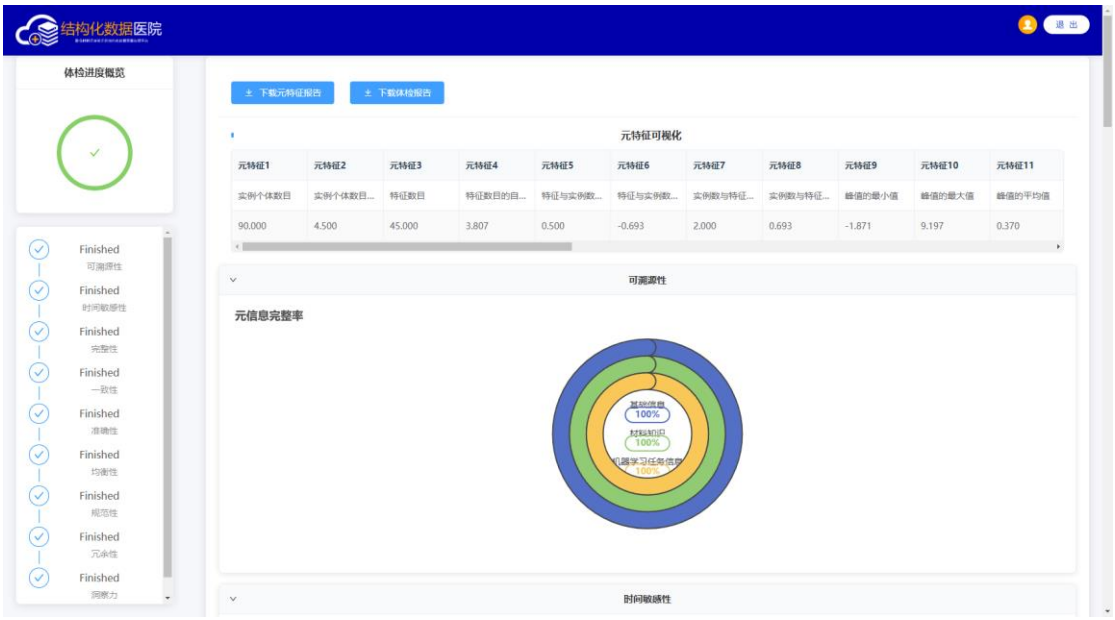

图 3-1 数据体检可溯源性

3.1.2 时间敏感性

平台在时间敏感性维度展示了时间序列的最小嵌入维。此处假设 nasicon 数据集的“a”特征列为其时间序列，并以此为例进行测试，运行结果如下图 3-2 所示，可知“a”的最小嵌入维为 4。

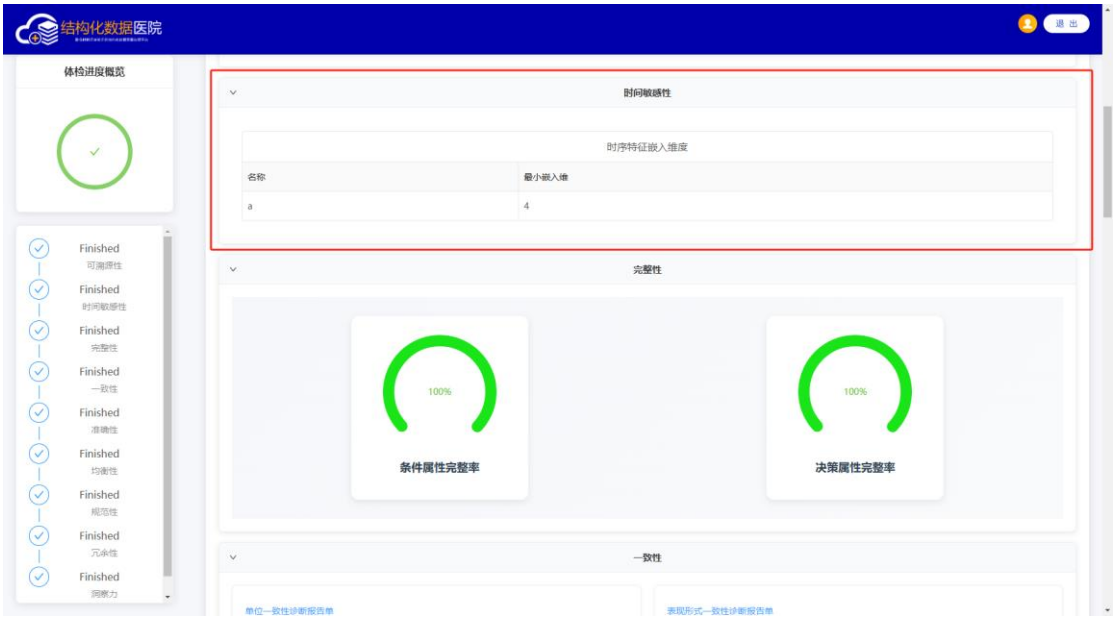

图 3-2 数据体检时间敏感性

3.1.3 完整性

平台在完整性维度展示了条件属性完整率和决策属性完整率。此处以 nasicon 数据集为例进行测试，运行结果如下图 3-3 所示，可知我们的条件属性完整率和决策属性完整率都为百分百。

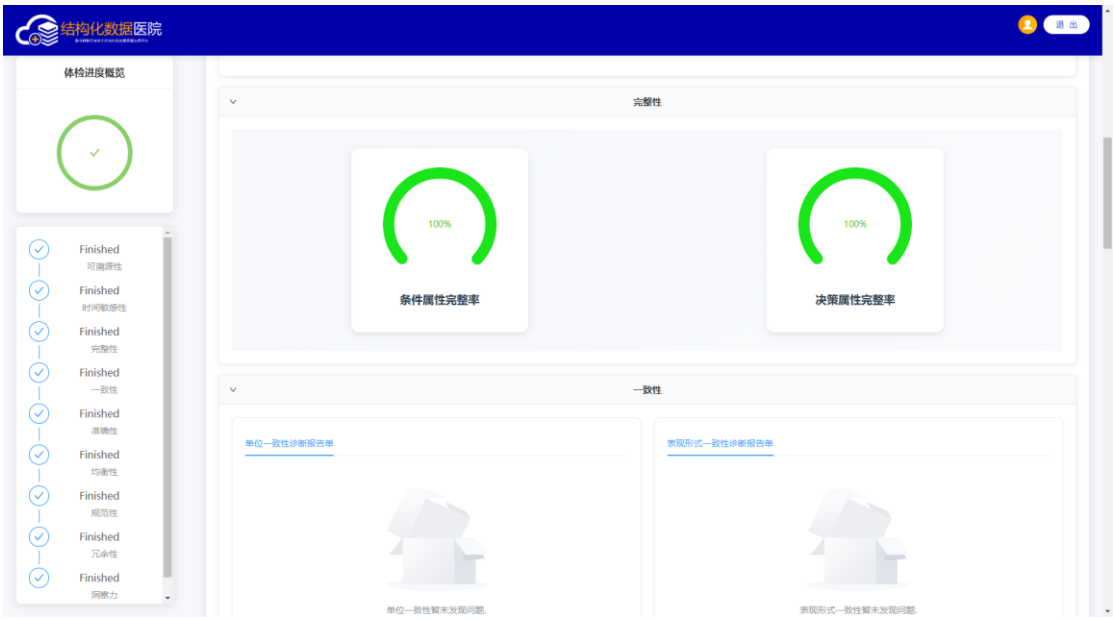

图 3-3 数据体检完整性

3.1.4 一致性

平台在一致性维度展示了单位一致性诊断报告单和表现形式一致性诊断报告单。此处以 nasicon 数据集为例进行测试，运行结果如下图 3-4 所示，可知单位一致性和表现形式一致性都不存在问题。

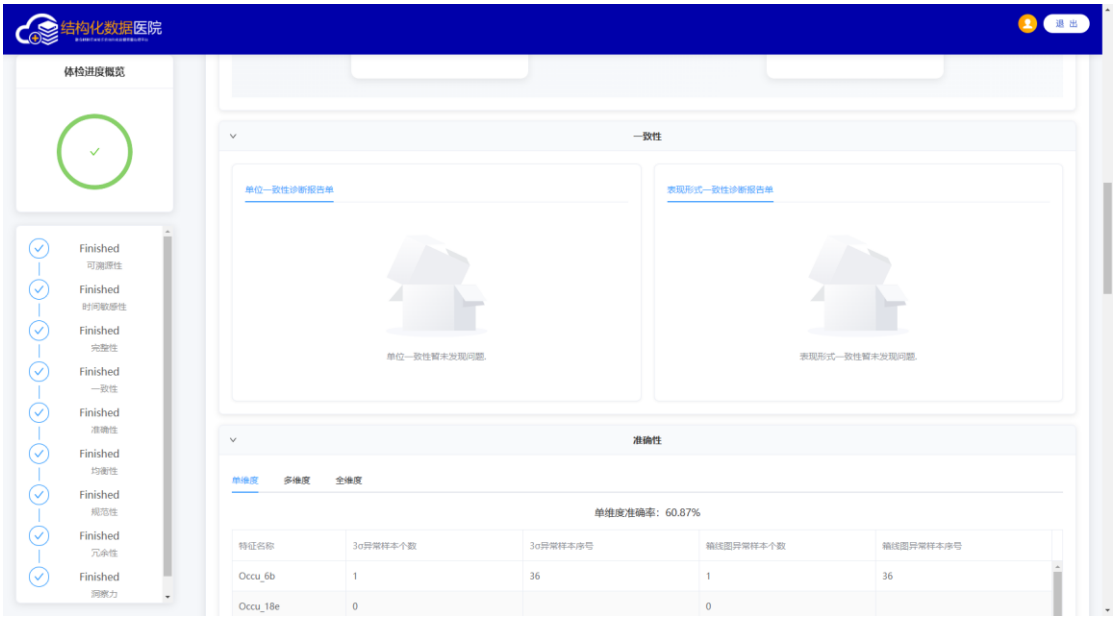

图 3-4 数据体检一致性

3.1.5 准确性

平台在准确性维度展示了单维度准确率、多维度准确率和全维度准确率。此处以 nasicon 数据集为例进行测试，单维度运行结果如下图 3-5 所示，可知单维度准确率为 60.87%；多维度运行结果如下图 3-6 所示，可知多维度准确率为 89.08；全维度运行结果如下图 3-7 所示，

可知全维度准确率为 86.67%。

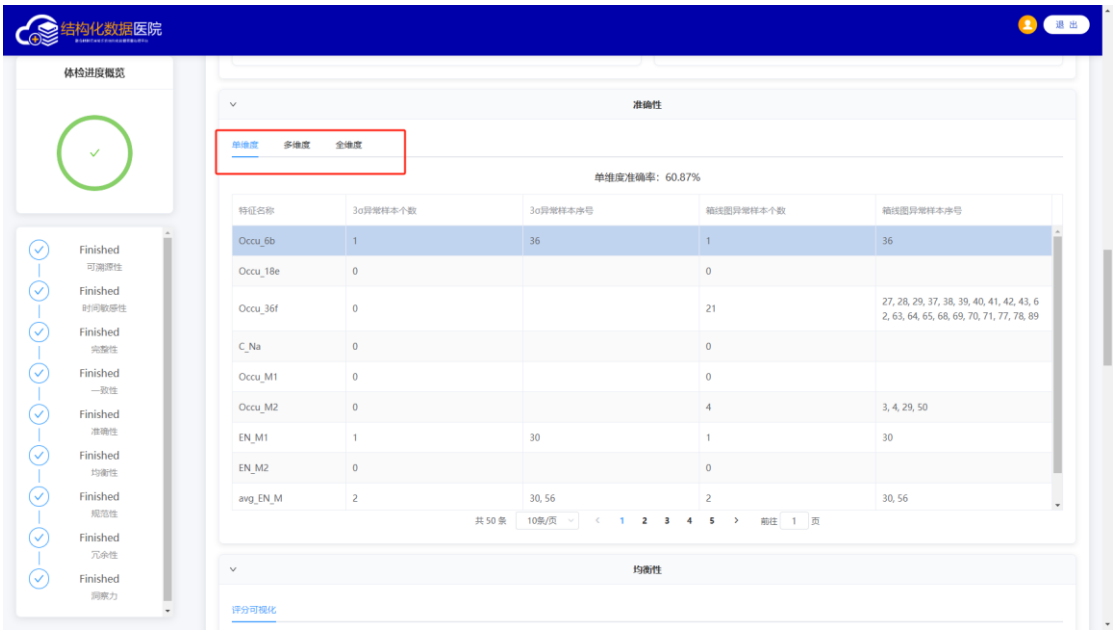

图 3-5 数据体检准确性单维度

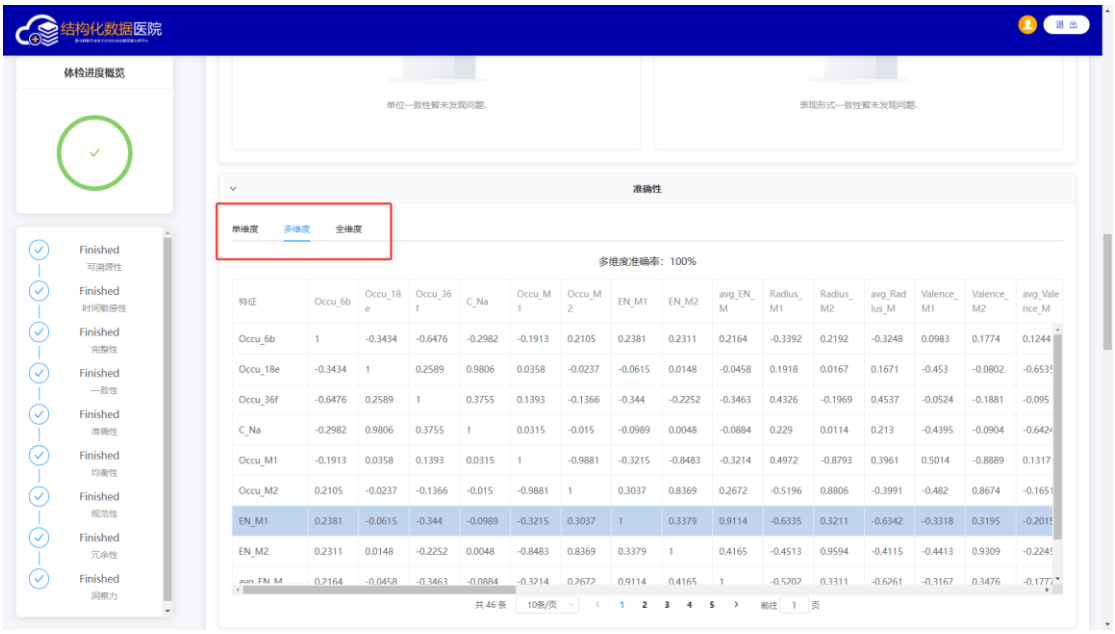

图 3-6 数据体检准确性多维度

图 3-7 数据体检准确性全维度

3.1.6 均衡性

平台在均衡性维度展示了数据均衡性指标。此处以 nasicon 数据集为例进行测试，运行结果如下图 3-8 所示，可知数据均衡性指标为 7.41%。

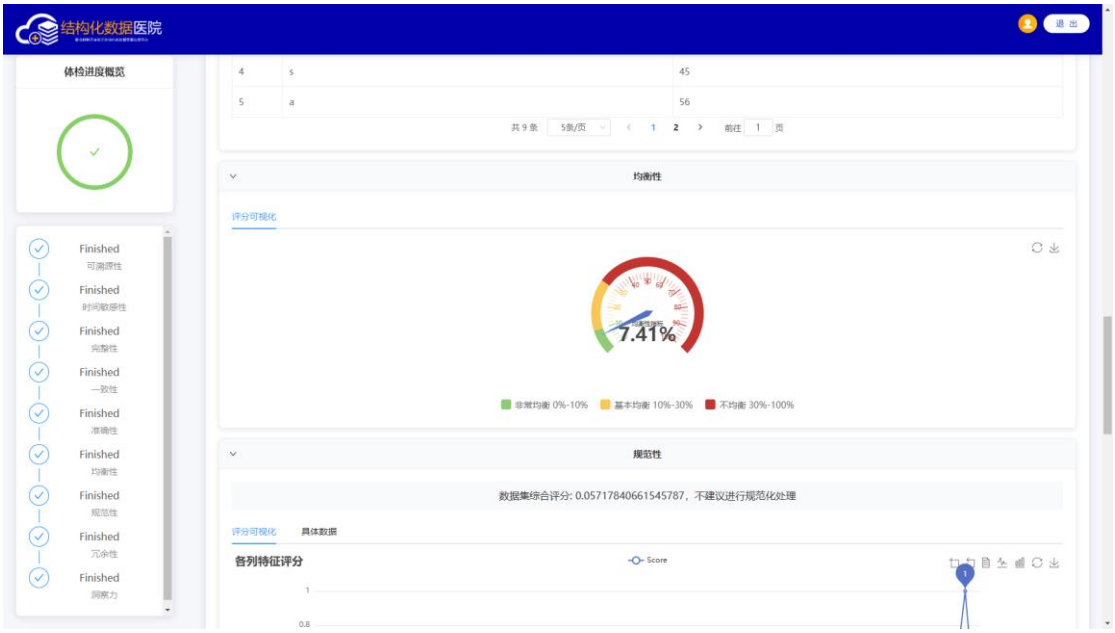

图 3-8 数据体检均衡性

3.1.7 规范性

平台在规范性维度展示了数据集综合评分以及各列特征评分可视化和具体数据。此处以

nasicon 数据集为例进行测试，运行结果如下图 3-9 所示，可知数据集综合评分为 0.057。

图 3-9 数据体检规范性

3.1.8 冗余性

平台在冗余性维度展示了冗余样本绝对数、冗余样本相对数、冗余特征绝对数、冗余特征相对数及其评分可视化和具体数据。此处以 nasicon 数据集为例进行测试，运行结果如下图 3-10 所示，可知数据集冗余样本绝对数为 29，冗余特征绝对数为 22。

图 3-10 数据体检冗余性

3.1.9 洞察力

平台在洞察力维度展示了该数据集在回归模型上 RMSE 最小的前三个模型及其对应的 RMSE 和参数。此处以 nasicon 数据集为例进行测试，运行结果如下图 3-11 所示。

图 3-11 数据体检洞察力

3.2 元特征报告

平台在数据体检界面提供了下载元特征报告功能，点击“下载元特征报告”即可生成元特征报告。此处以 nasicon 数据集为例进行测试，生成元特征报告如下图 3-12 所示。

图 3-12 数据体检的元特征报告

3.3 体检报告

平台在数据体检界面提供了下载体检报告功能，点击“下载体检报告”即可生成体检报告。此处以 nasicon 数据集为例进行测试，生成体检报告如下图 3-13 所示。

图 3-13 数据体检的体检报告

第 4 章　数据诊治模块

本章概述

本章主要介绍了结构化数据医院——融合材料领域知识的结构化数据质量治理平台的数据诊治部分的内容。用户可通过此模块对数据体检中数据存在的质量问题从可溯源性、时间敏感性、完整性、一致性、准确性、均衡性、规范性、冗余性、洞察力九个治理维度分别进行质量治理。

本章内容

4.1 可溯源性诊治

点击平台页面左侧导航栏中的“可溯源性科室”-“检查报告”，进入到可溯源性页面。平台在可溯源性维度展示了元信息完整率和数据溯源表格展示。此处以 nasicon 数据集为例进行测试，系统可通过在数据挂号上传的信息来进行元信息完整率的计算以及数据溯源中各个字段的信息显示，运行结果如下图 4-1 所示。

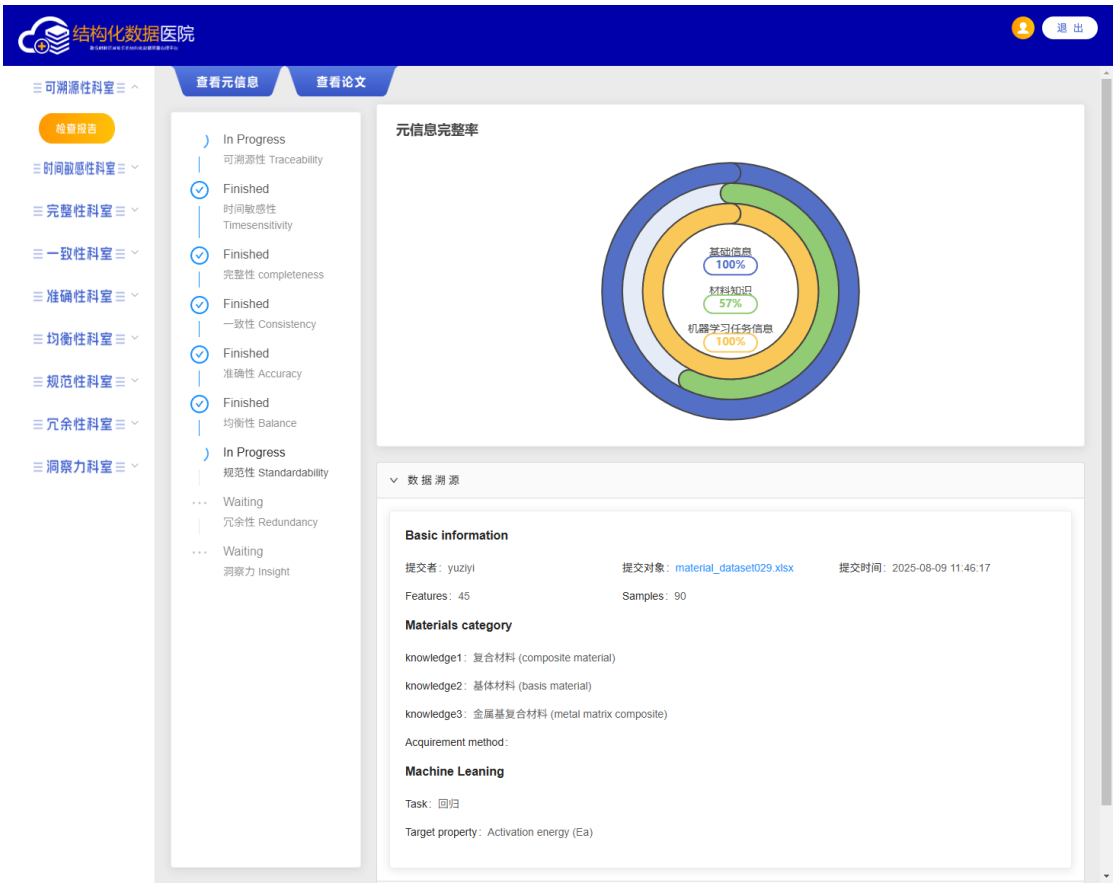

图 4-1 数据诊治可溯源性

4.2 时间敏感性诊治

（1）点击平台页面左侧导航栏中的“时间敏感性科室”-“导医台”，进入到时间敏感性文件选择页面，运行结果如下图 4-2 所示。

图 4-2 数据诊治时间敏感性导医台

（2）点击“选择文件”，选择数据挂号上传的文件，点击“前往 go”，进入到时间敏

感性最小嵌入维计算页面。平台会根据在数据上传时选择的时序特征，通过 CAO 算法计算其最小嵌入维数，并在页面上进行显示，运行结果如下图 4-3 所示。

图 4-3 数据诊治时间敏感性体检报告

4.3 完整性诊治

（1）点击平台页面左侧导航栏中的“完整性科室”-“导医台”，进入到完整性文件选择页面，运行结果如下图 4-4 所示。

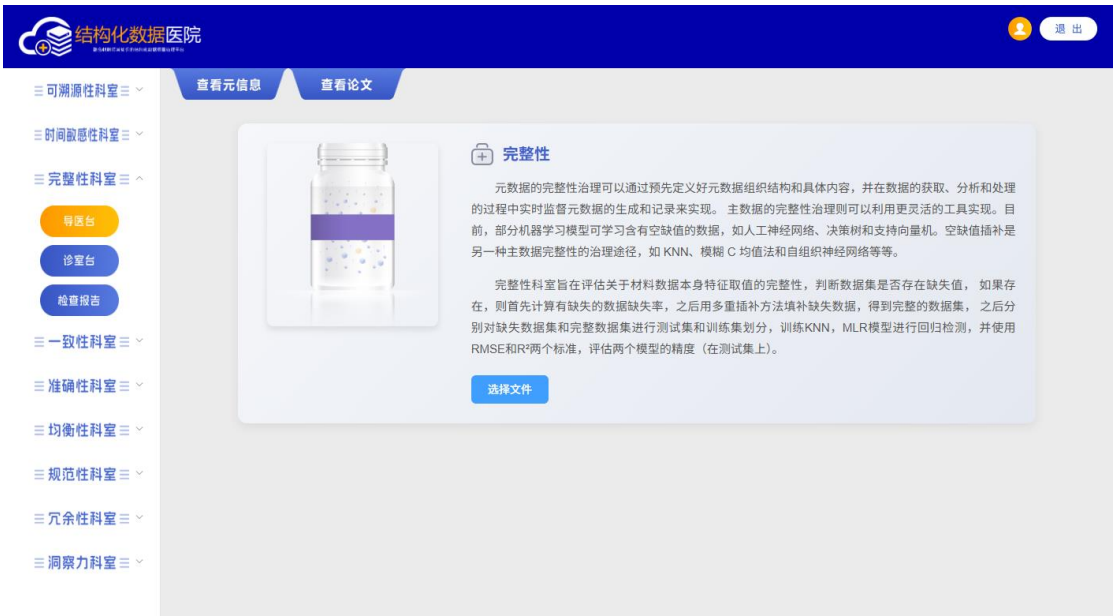

图 4-4 数据诊治完整性导医台

（2）点击“选择文件”，进入到选择文件页面，运行结果如下图 4-5 所示。在待分析的数据集中选择数据挂号上传的数据集，点击“分析”，等待平台分析完成之后会将此数据集

加入到已分析的数据集中，运行结果如下图 4-6 所示。在已分析的数据集中点击“报告”，会跳转到数据详情页面，展示说明文档、数据完整性概览、特征相关性矩阵、特征缺失率、Top5 重要特征、筛选条件/决策属性，运行结果如下图 4-7 所示。

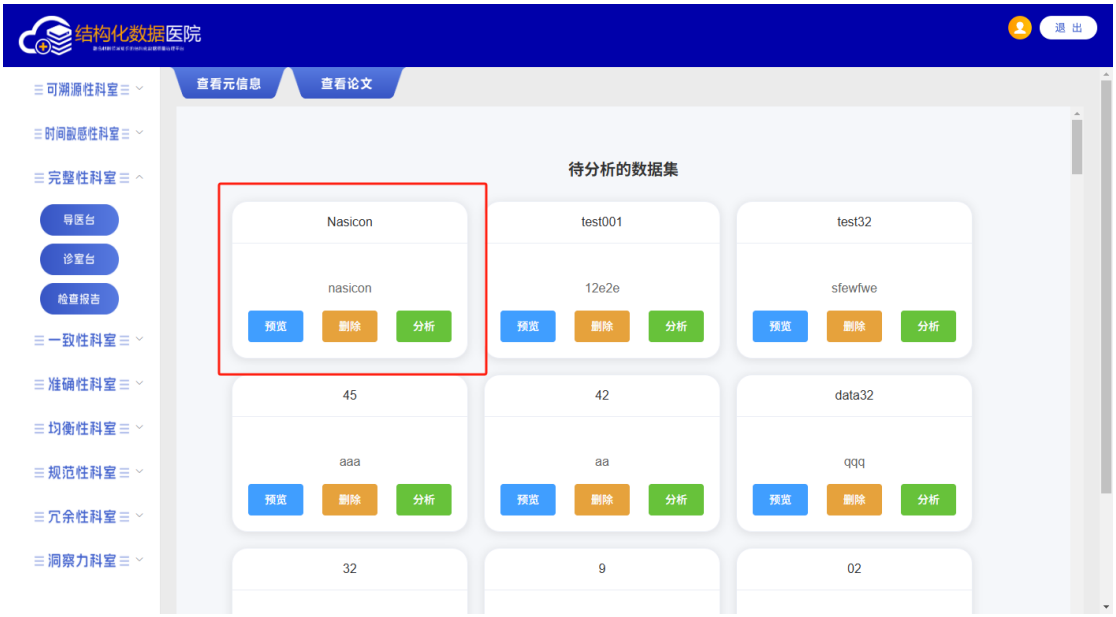

图 4-5 数据诊治完整性分析数据集

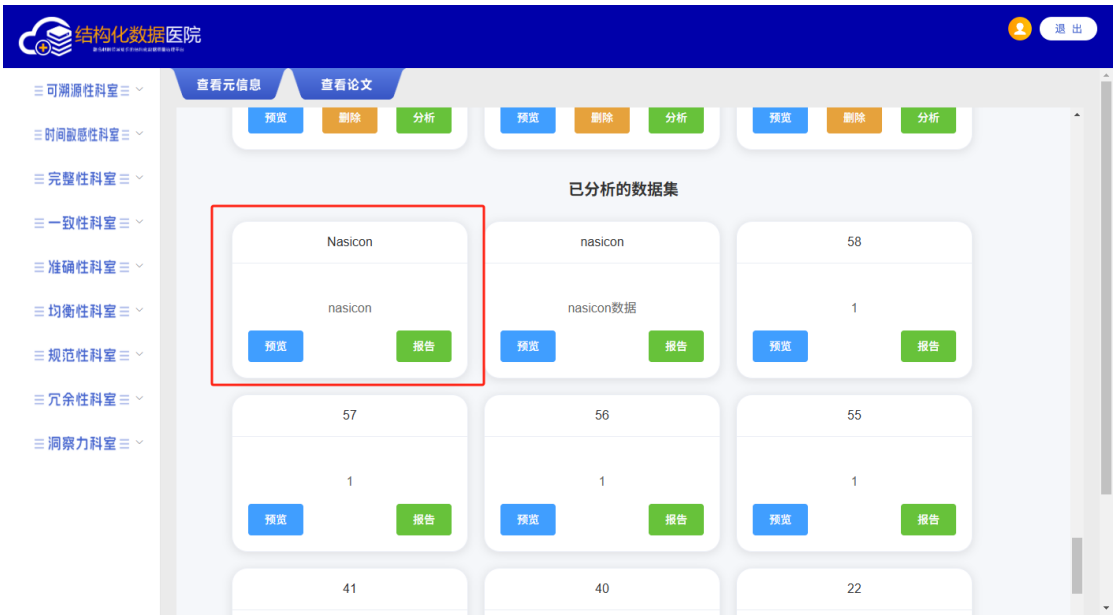

图 4-6 数据诊治完整性查看已分析数据集

图 4-7 数据诊治完整性分析结果

（3）点击平台页面左侧导航栏中的“完整性科室”-“诊室台”，进入到算法填补页面，运行结果如下图 4-8 所示。平台提供双模式填补策略，可以选择推荐算法和手动填补两种方式对数据进行填补，此处以推荐算法为例。勾选待填补数据集前面的方框，点击“生成推荐算法”，平台会根据领域知识自动生成适合该数据集的最佳填补算法。再次勾选待填补数据集前面的方框，点击“一键填补”，平台对待填补数据集进行填补，运行结果如下图 4-9 所示。

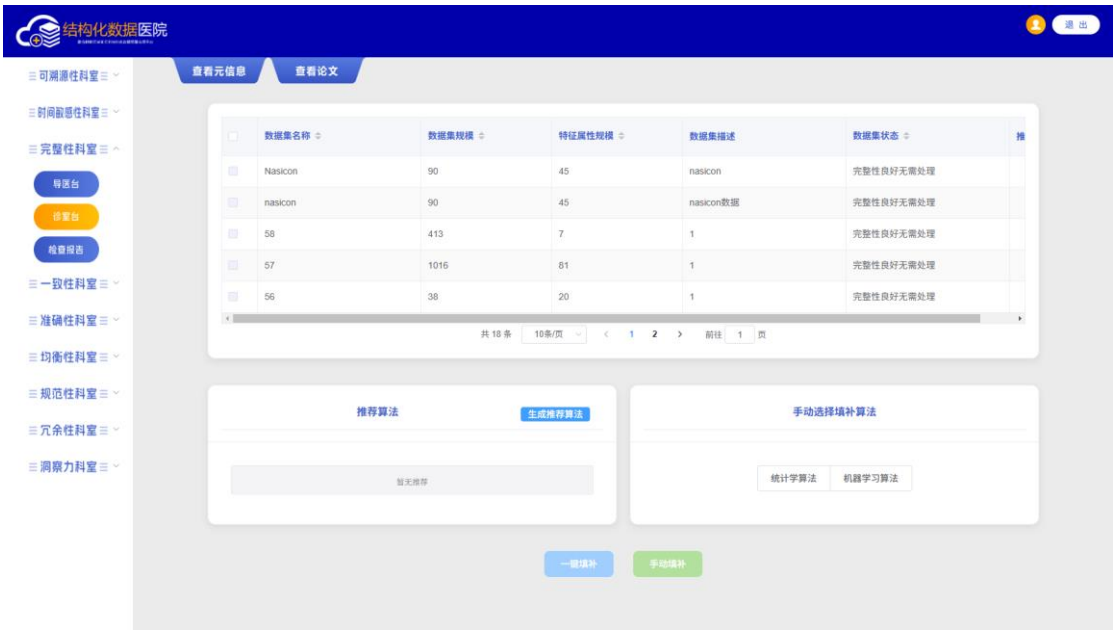

图 4-8 数据诊治完整性诊室台

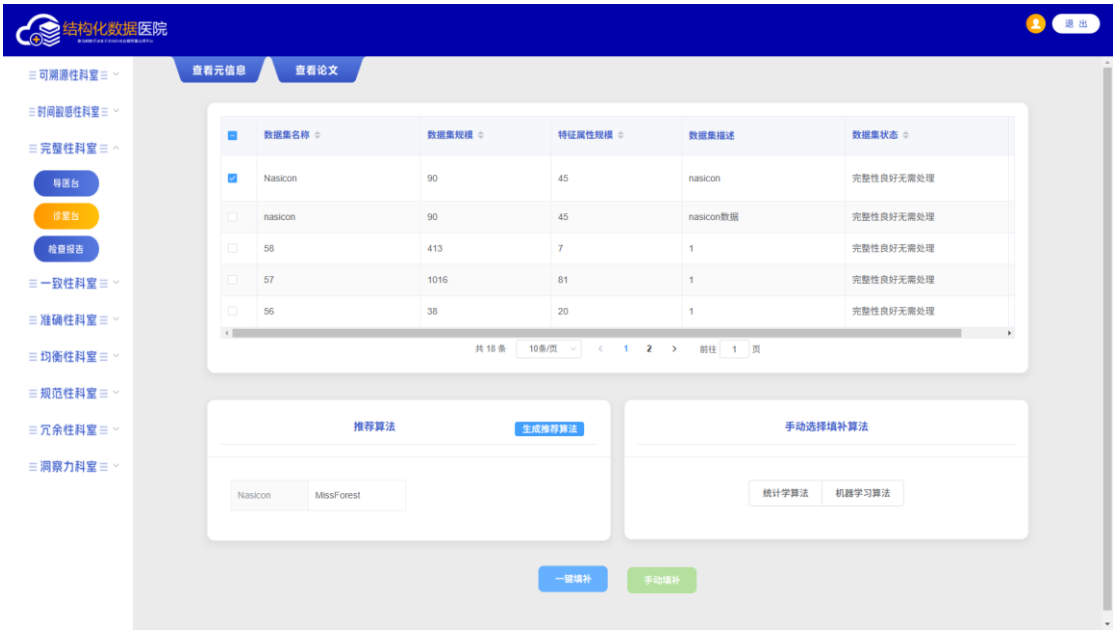

图 4-9 数据诊治完整性诊治结果

（4）点击平台页面左侧导航栏中的“完整性科室”-“检查报告”，进入到填补效果评估页面，运行结果如下图 4-10 所示。在“选择数据集”中选择刚刚填补后的数据集，在“选择回归模型”中根据需求选择相应的模型，在“选择评估指标”中根据需求选择相应的指标，点击“开始评估”，平台会根据用户选择进行填补效果评估，并在“评估结果”区域可视化呈现填补效果，运行结果如下图 4-11 所示。

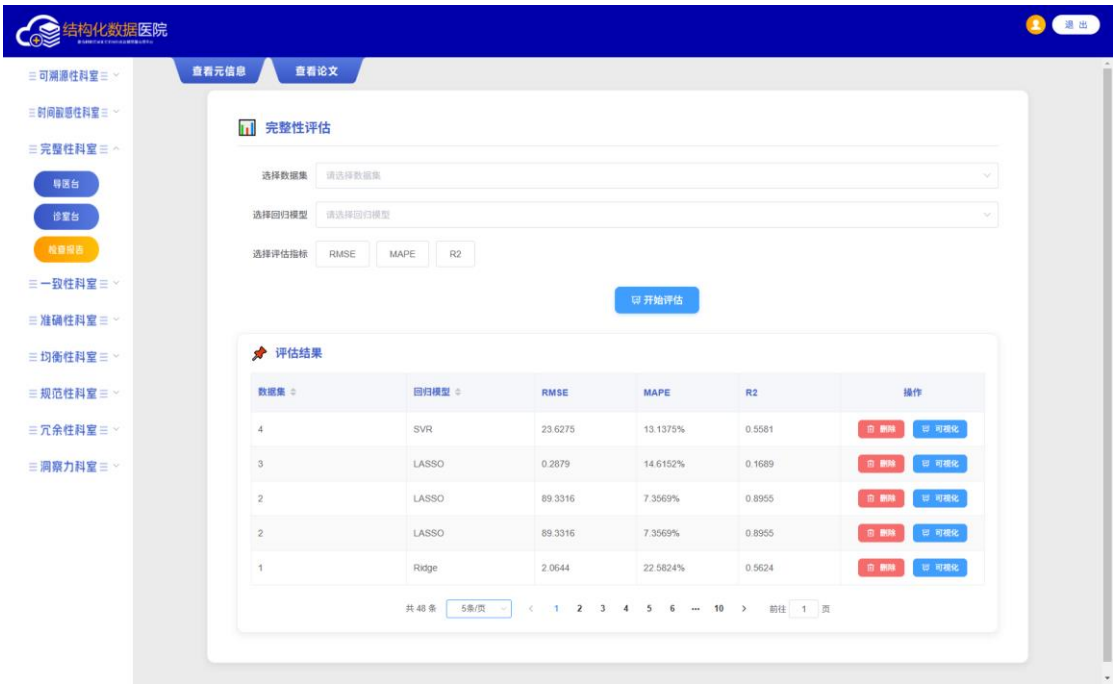

图 4-10 数据诊治完整性检测报告页面

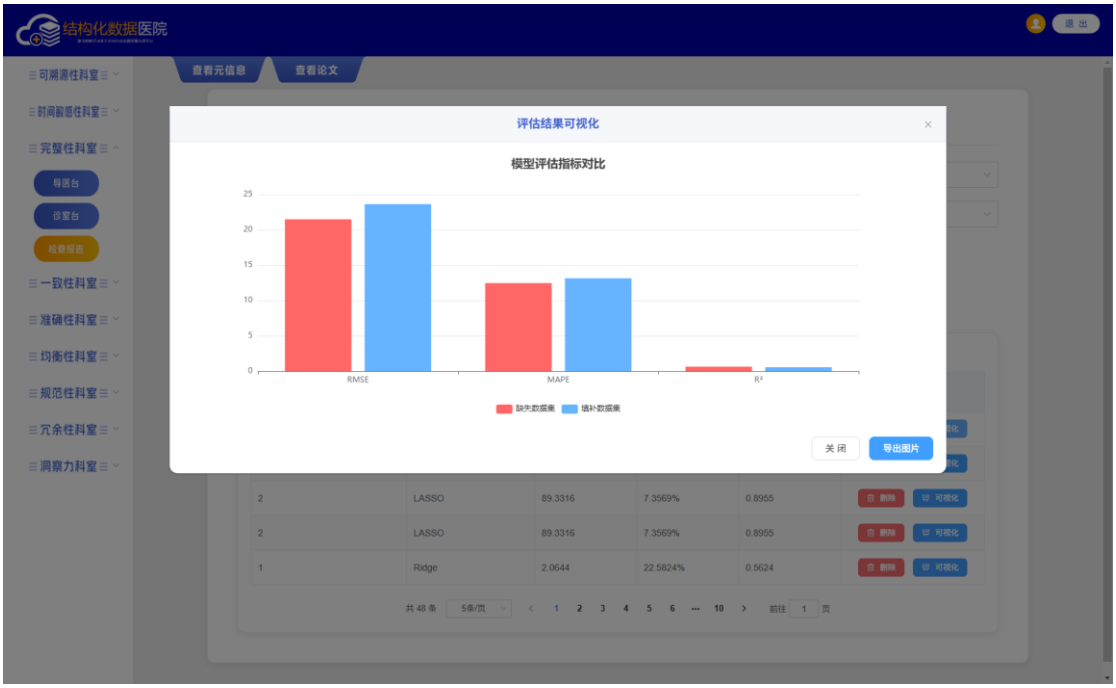

图 4-11 数据诊治完整性评估结果

4.4 一致性诊治

（1）点击平台页面左侧导航栏中的“一致性科室”-“导医台”，进入到一致性文件选择页面，在一致性治理中平台实现了对于数据量纲不一致和表现形式不一致的分别治理，运行结果如下图 4-12 所示。

图 4-12 数据诊治一致性导医台

（2）点击“量纲不一致”区域的“选择文件”，选择数据挂号上传的文件，点击“前往”，进入到量纲不一致治理页面，运行结果如下图 4-13 所示。在“单位不一致判断”中根据自身是否需要治理的需求选择“是”或者“否”，此处以需要治理为例，选择“是”，平

台会在“单位一致性诊断报告单”区域展示存在单位问题的数据，并在“单位一致化结果”区域展示治理后的数据集，运行结果如下图 4-14 所示。

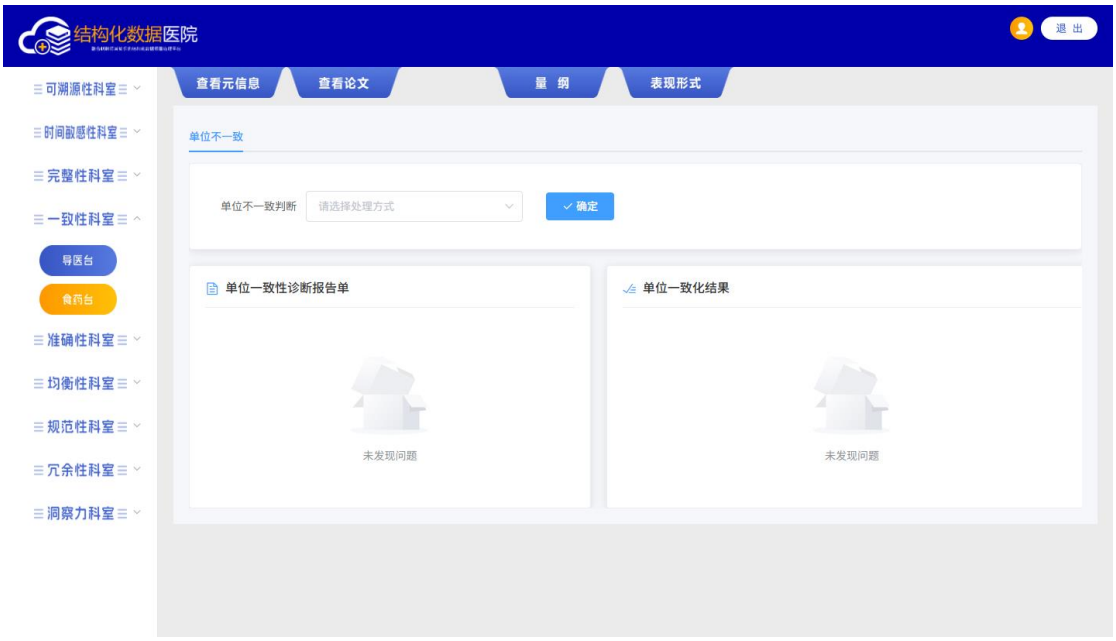

图 4-13 数据诊治量纲一致性食药台

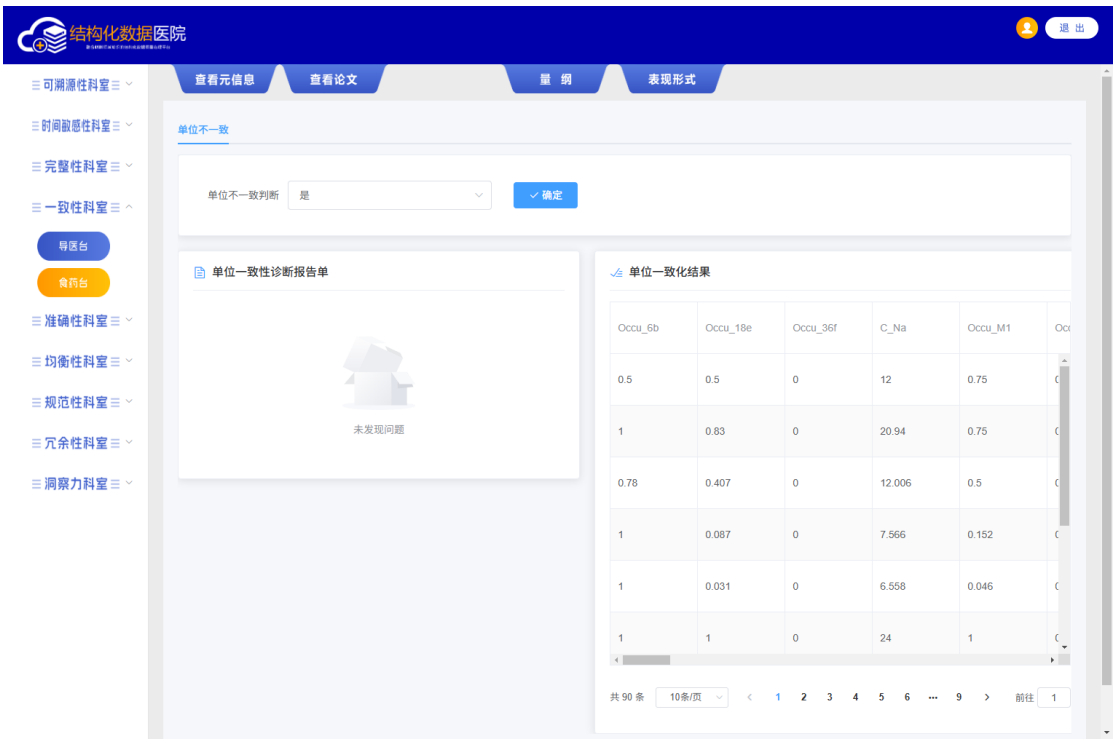

图 4-14 数据诊治量纲一致性诊治结果

（3）点击平台页面上方的“表现形式”，进入到表现形式不一致治理页面，运行结果如下图 4-15 所示。在“表现形式不一致判断”中根据自身是否需要治理的需求选择“是”或者“否”，此处以需要治理为例，选择“是”，平台会在“表现形式一致性诊断报告单”区域展示存在表现形式问题的数据，并在“表现形式一致化结果”区域展示治理后的数据集，运行结果如下图 4-16 所示。

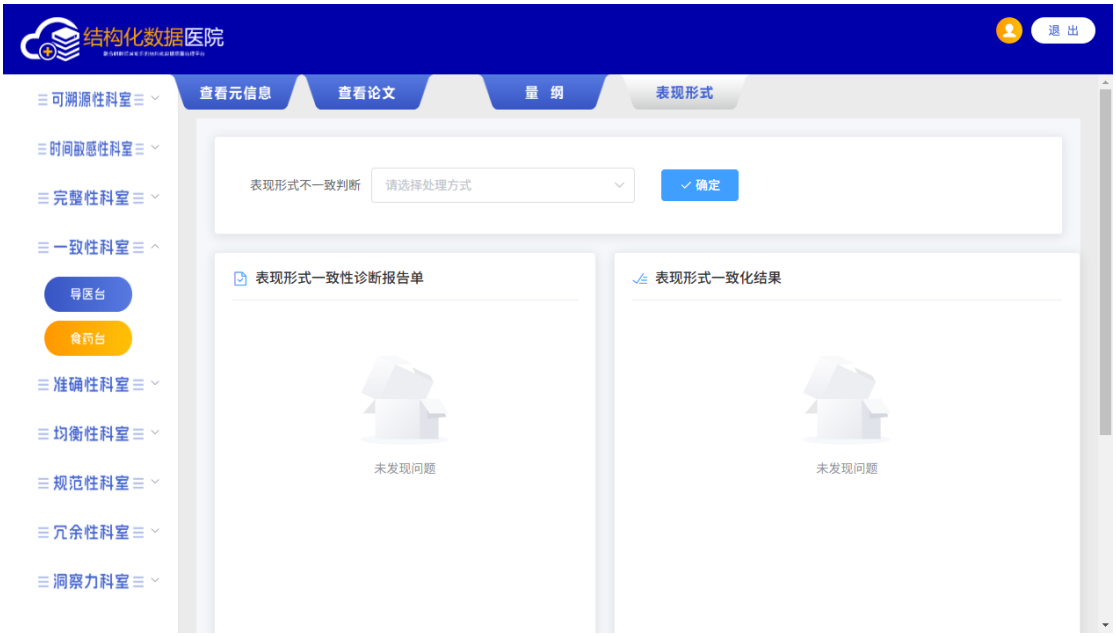

图 4-15 数据诊治表现形式一致性食药台

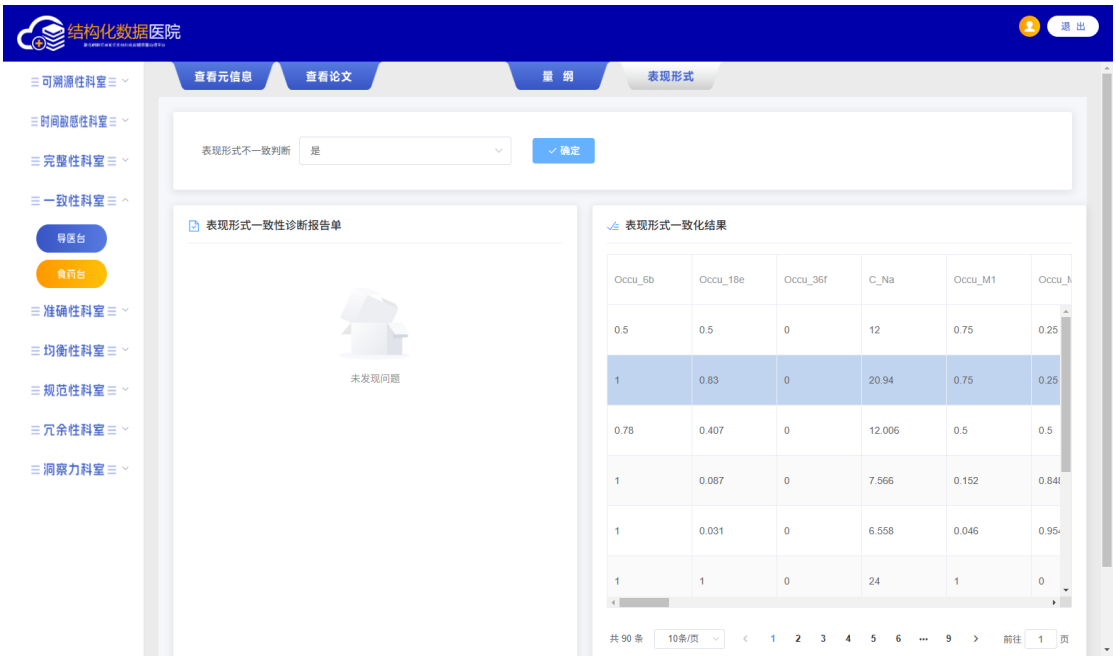

图 4-16 数据诊治量纲一致性诊治结果

4.5 准确性诊治

（1）点击平台页面左侧导航栏中的“准确性科室”-“导医台”，进入到准确性文件选择页面，运行结果如下图 4-17 所示。

图 4-17 数据诊治准确性导医台

（2）点击“选择文件”，选择数据挂号上传的文件，页面会展示当前数据的单描述符规则和相似性识别规则（如果没有立即显示，请等待几秒钟后刷新页面），用户可根据领域知识对其进行修改，运行结果如下图 4-18 所示。点击“前往 Go”，进入到数据准确性分维度检测页面，运行结果如下图 4-19 所示。

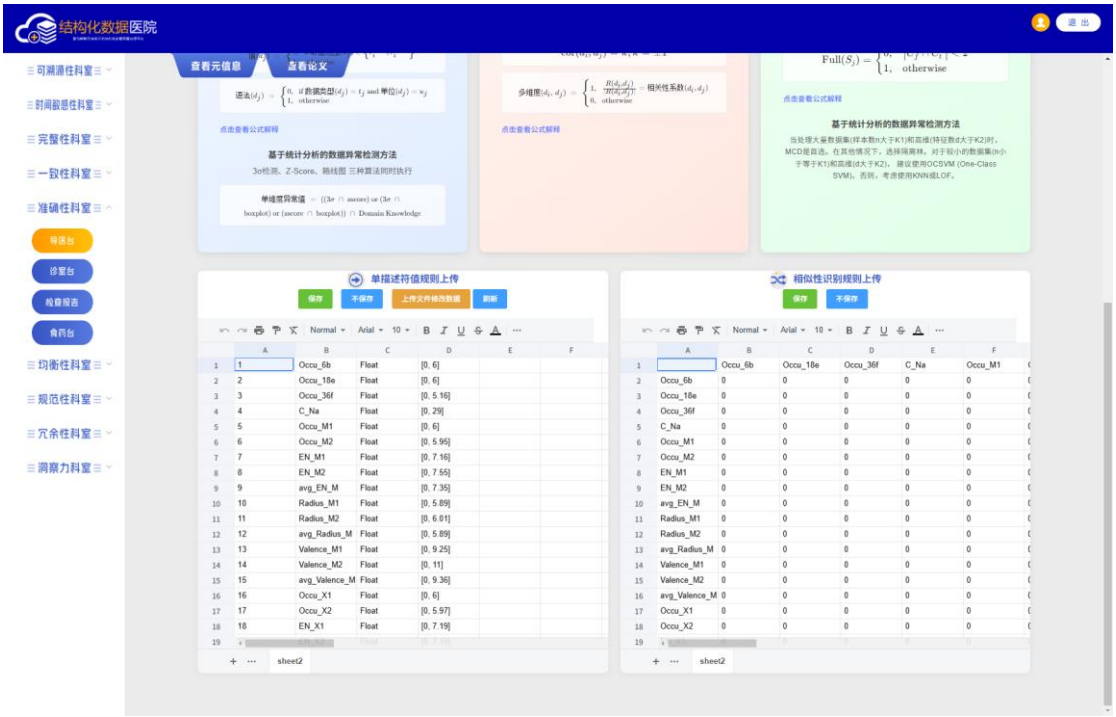

图 4-18 数据诊治准确性导医台规则

图 4-19 数据诊治准确性诊室台

（3）点击“单维度”区域的“检测”，平台会根据异常检测算法和所修改的领域知识对异常值进行检测并展示在页面上，运行结果如下图 4-20 所示。点击“多维度”区域的“检测”，平台会根据异常检测算法和所修改的领域知识对异常特征对进行检测并展示在页面上，运行结果如下图 4-21 所示。点击“全维度”区域的“检测”，平台会根据异常检测算法和所修改的领域知识对异常样本进行检测并展示在页面上，运行结果如下图 4-22 所示。

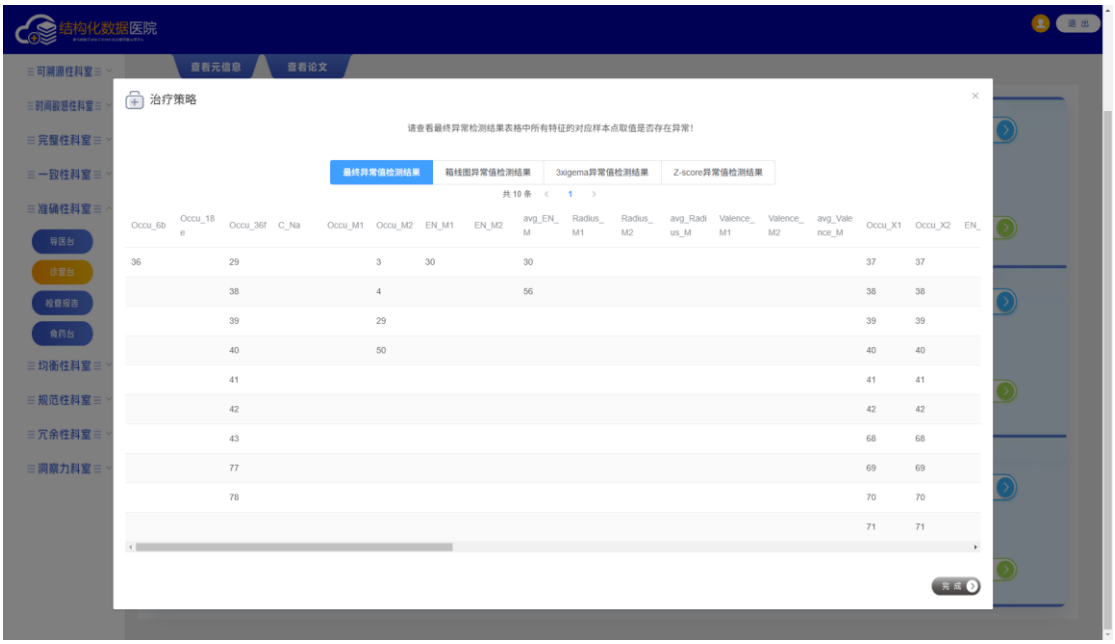

图 4-20 数据诊治准确性诊室台单维度

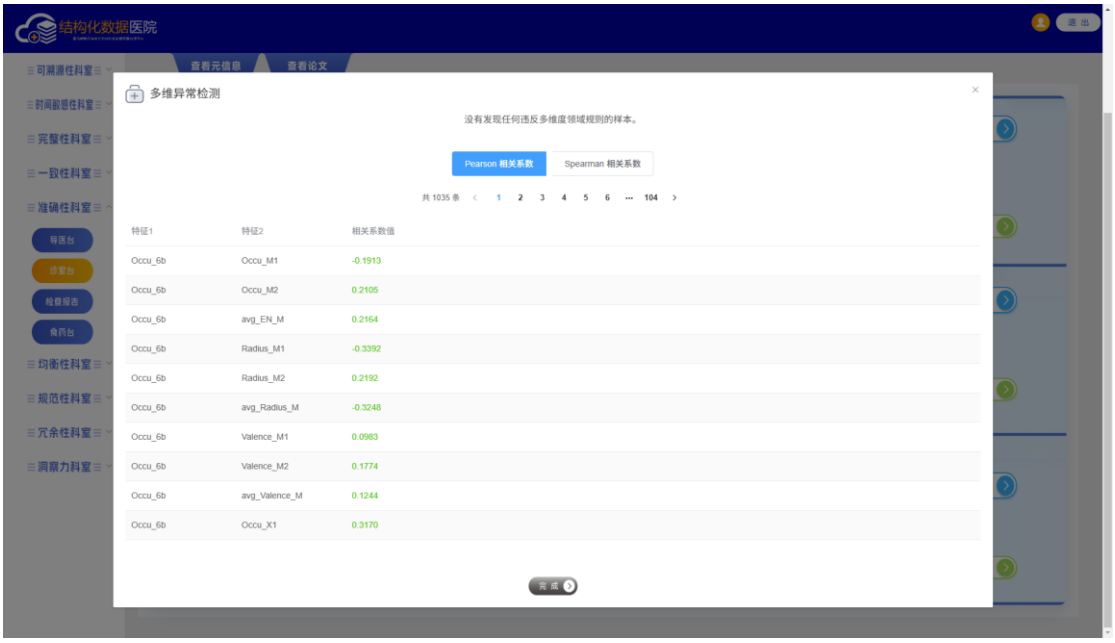

图 4-21 数据诊治准确性诊室台多维度

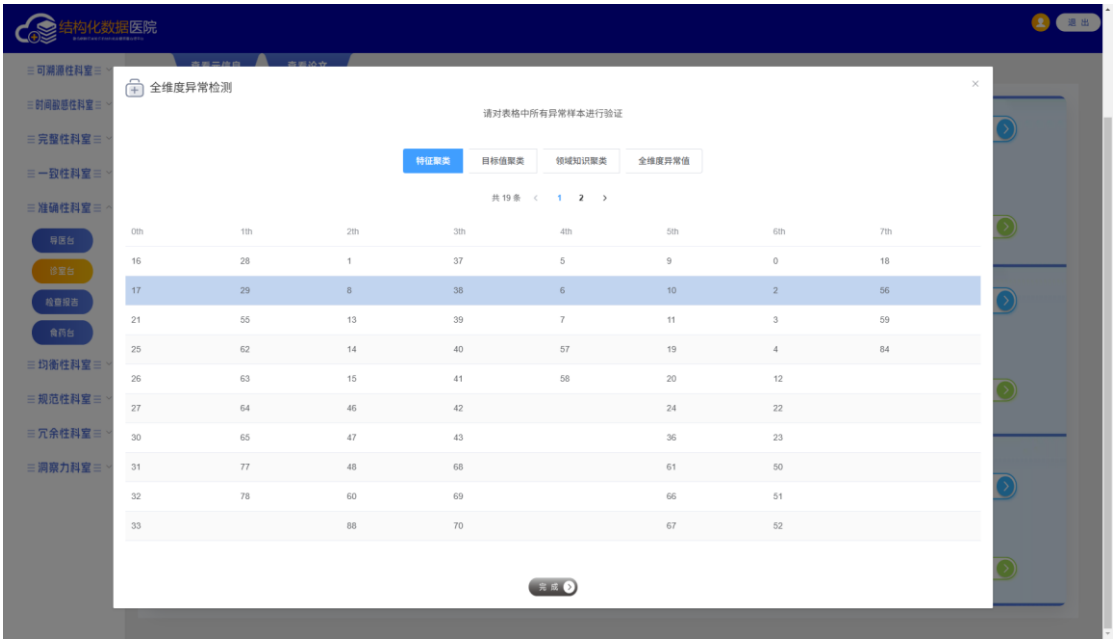

图 4-22 数据诊治准确性诊室台全维度

（4）点击“单维度”区域的“查看详情”，进入到单维度准确性治理页面，页面会呈现字段准确性以及各特征的异常率等相关单维度准确性指标，并根据检测结果提供治疗策略，运行结果如下图 4-23 所示。点击页面上方的“多维度”，进入到多维度准确性治理页面，页面会呈现特征准确性等相关多维度准确性指标，并根据检测结果提供治疗策略，运行结果如下图 4-24 所示。点击页面上方的“全维度”，进入到全维度准确性治理页面，页面会呈现样本准确性等相关全维度准确性指标，并根据检测结果提供治疗策略，运行结果如下图 4-25 所示。

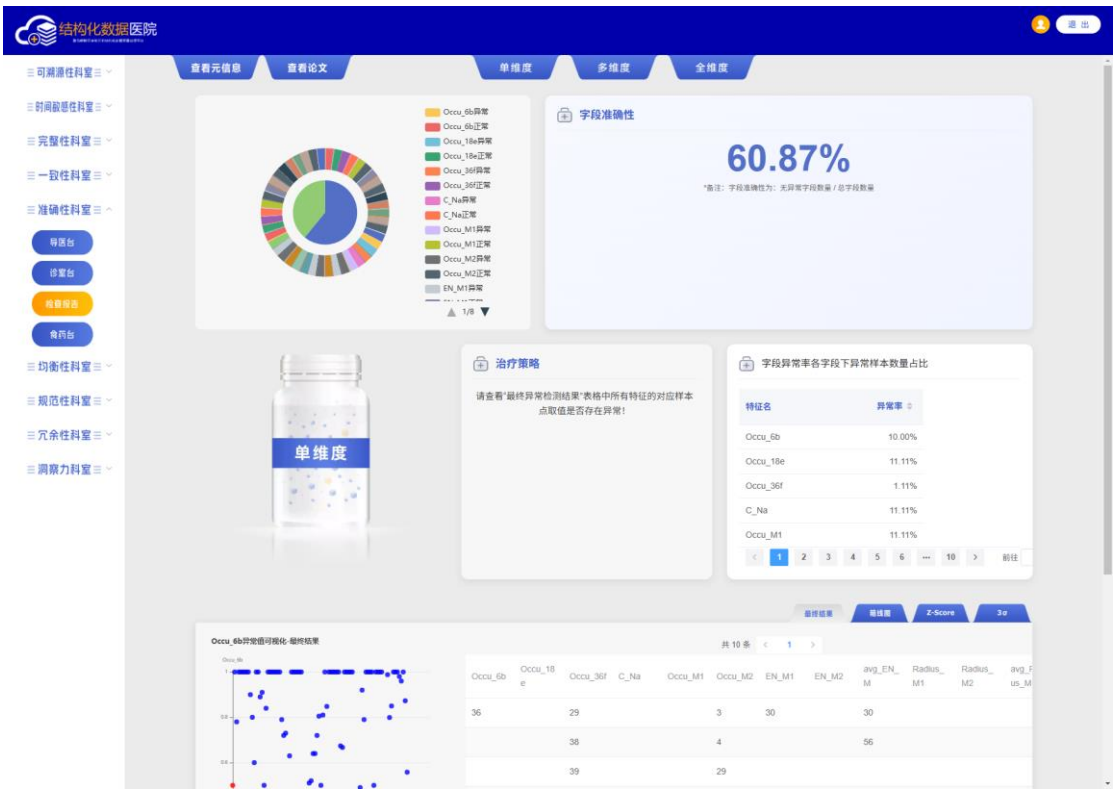

图 4-23 数据诊治准确性检查报告单维度

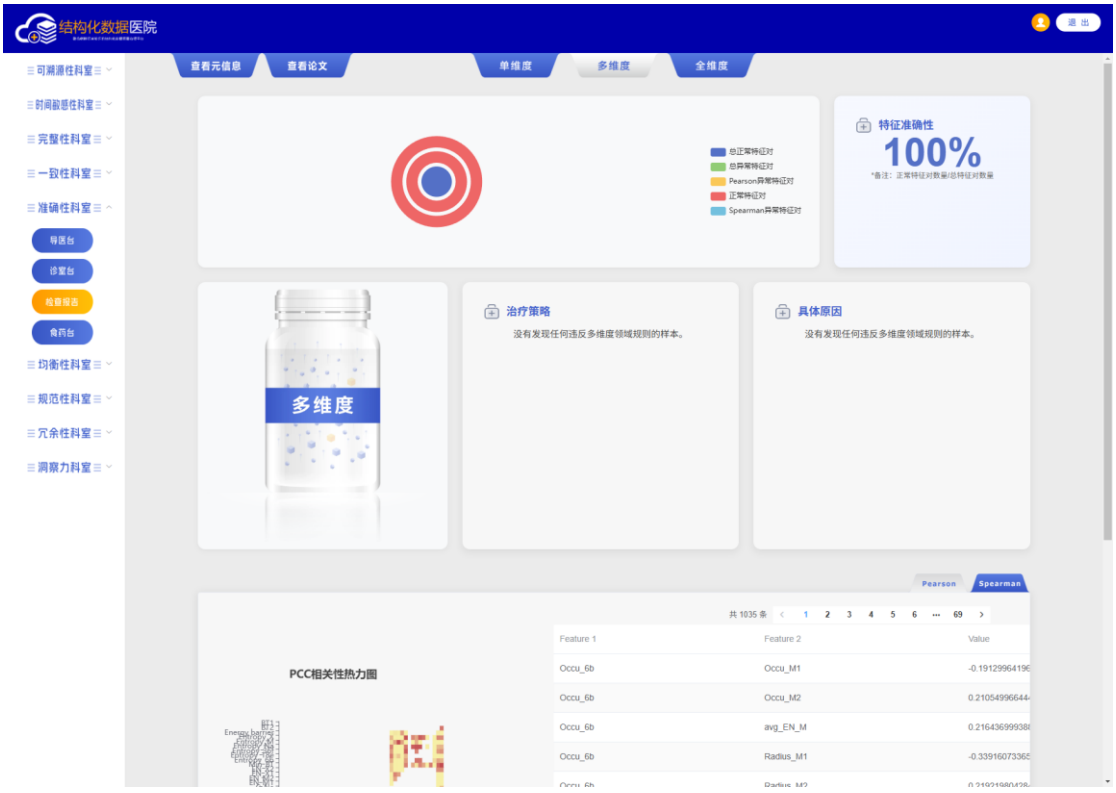

图 4-24 数据诊治准确性检查报告多维度

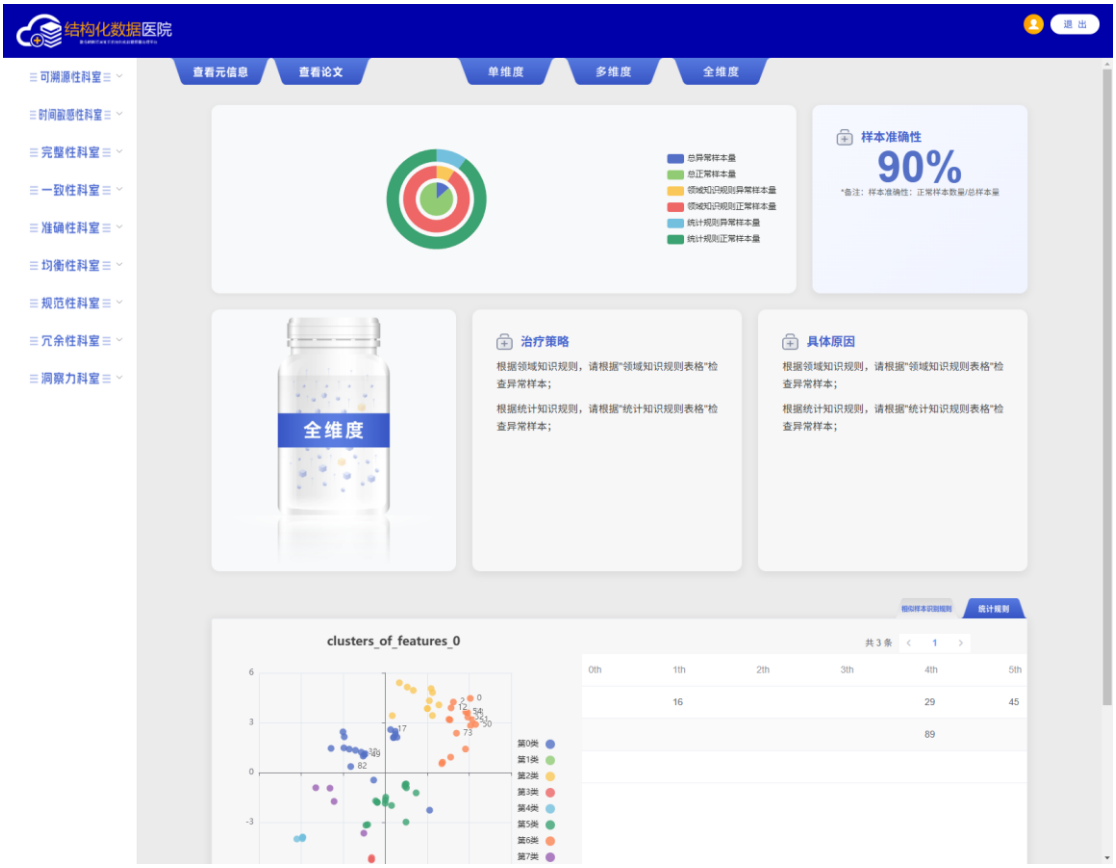

图 4-25 数据诊治准确性检查报告全维度

（5）点击平台页面左侧导航栏中的“准确性科室”-“食药台”，进入到数据准确性治理页面，页面鼠标各个维度的食药瓶上会自动呈现需要治理的样本。点击“一键修正”平台会对有问题的数据进行治理，运行结果如下图 4-26 所示。

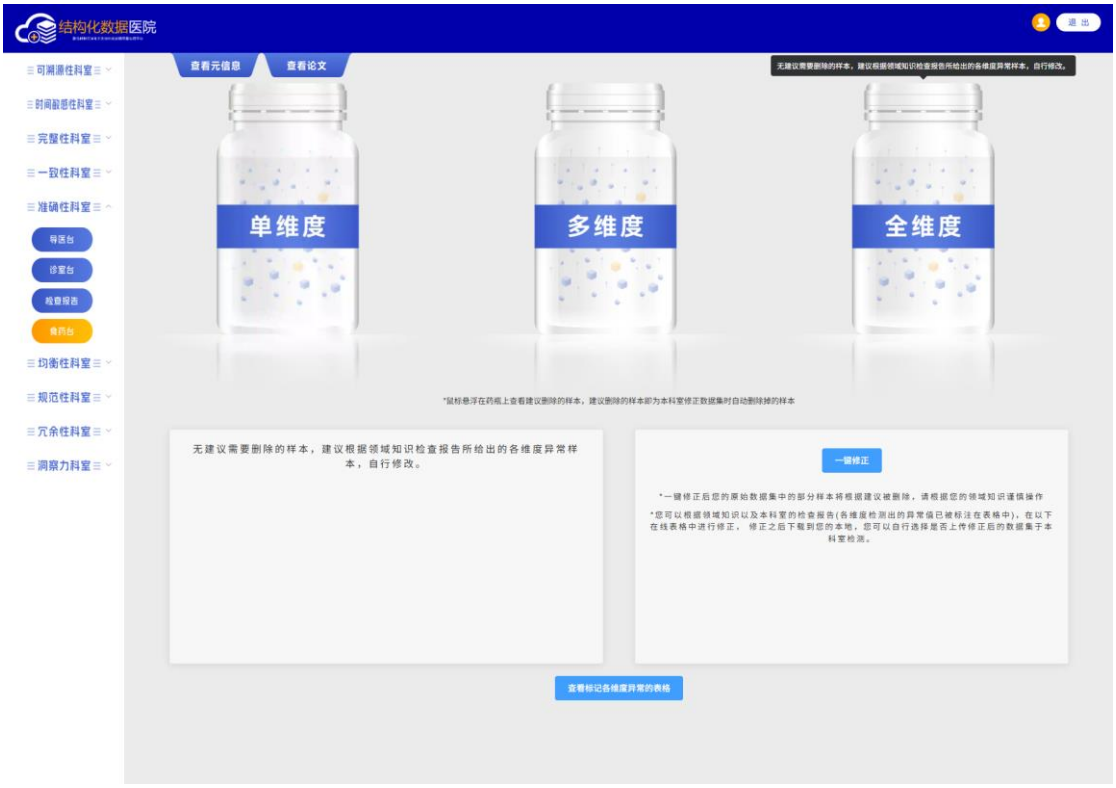

图 4-26 数据诊治准确性食药台

4.6 均衡性诊治

（1）点击平台页面左侧导航栏中的“均衡性科室”-“导医台”，进入到均衡性文件选择页面。点击“选择文件”，选择数据挂号上传的文件，点击“均衡性检测”，平台会计算选择数据的均衡性指标并进行显示，运行结果如下图 4-26 所示。

图 4-26 数据诊治均衡性导医台

（3）点击“前往治理”，平台进入均衡性治理页面，可根据是否需要对样本进行均衡性治理的需求选择“不对样本做修改”或者“采样”，此处以“采样”为例进行测试。点击“采样”-“确定”，平台根据均衡性指标判断如何进行治理并进行治理，随后点击“选择文件”，选择刚刚治理后生成的文件，点击“再次均衡性检测”，平台会展示治理后的均衡性指标，运行结果如下图 4-27 所示。

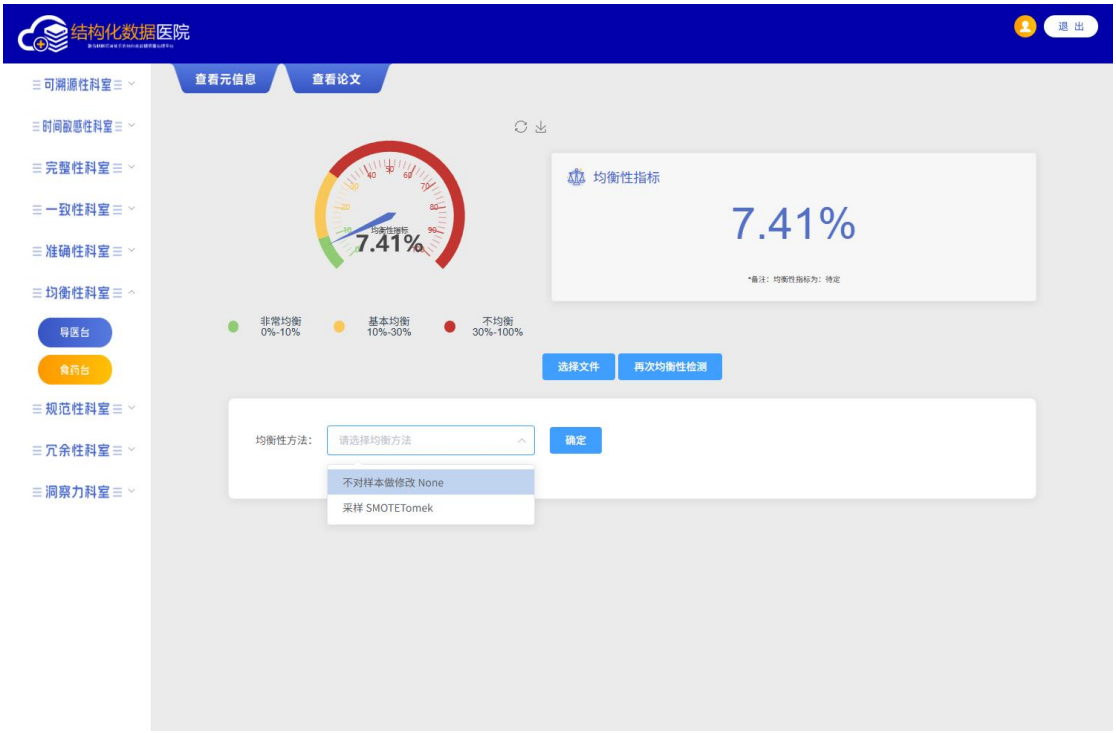

图 4-26 数据诊治均衡性食药台

4.7 规范性诊治

（1）点击平台页面左侧导航栏中的“规范性科室”-“导医台”，进入到规范性文件选择页面，在规范性治理中平台实现了对于取值规范化和划分规范化的分别治理，运行结果如下图 4-28 所示。

图 4-26 数据诊治规范性导医台

（2）点击“取值规范化”区域的“选择文件”，选择数据挂号上传的文件，点击“前往 Go”，进入到取值规范化治理页面，运行结果如下图 4-27 所示。在“规范化方法”中根据自身需求选择不同的规范化方法，此处以最小-最大规范化为例。选择“最小-最大规范化”，点击“确定”，平台会展示治理后的数据集，运行结果如下图 4-28 所示。

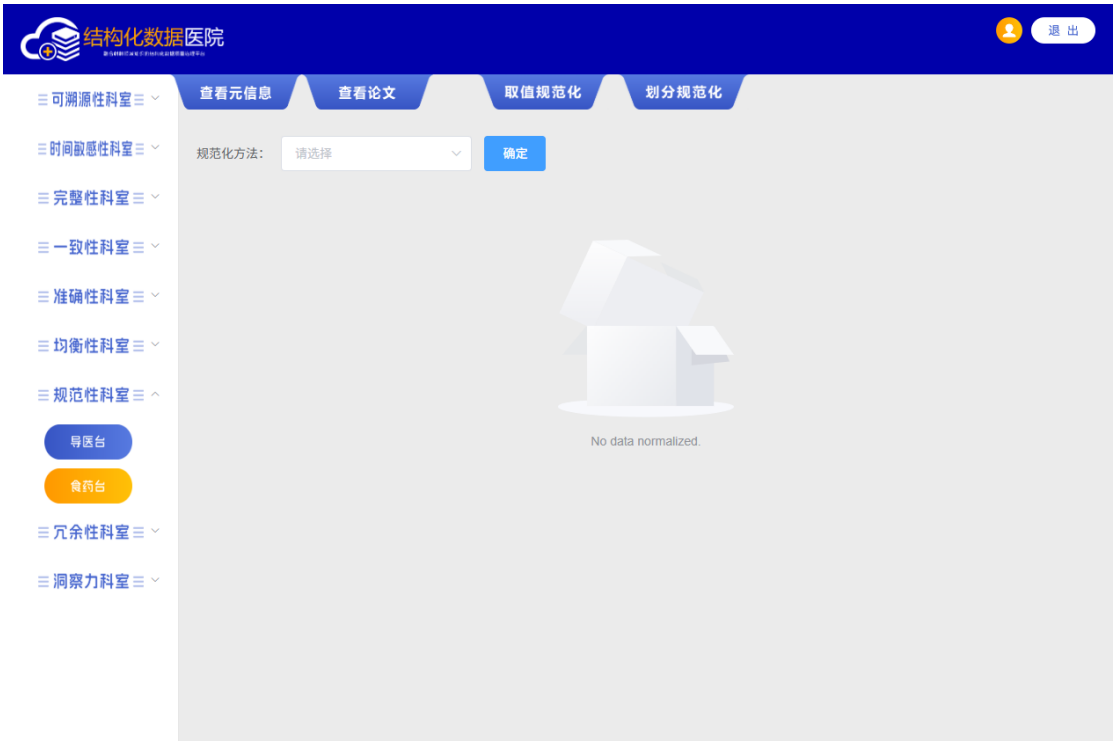

图 4-27 数据诊治规范性取值规范化食药台

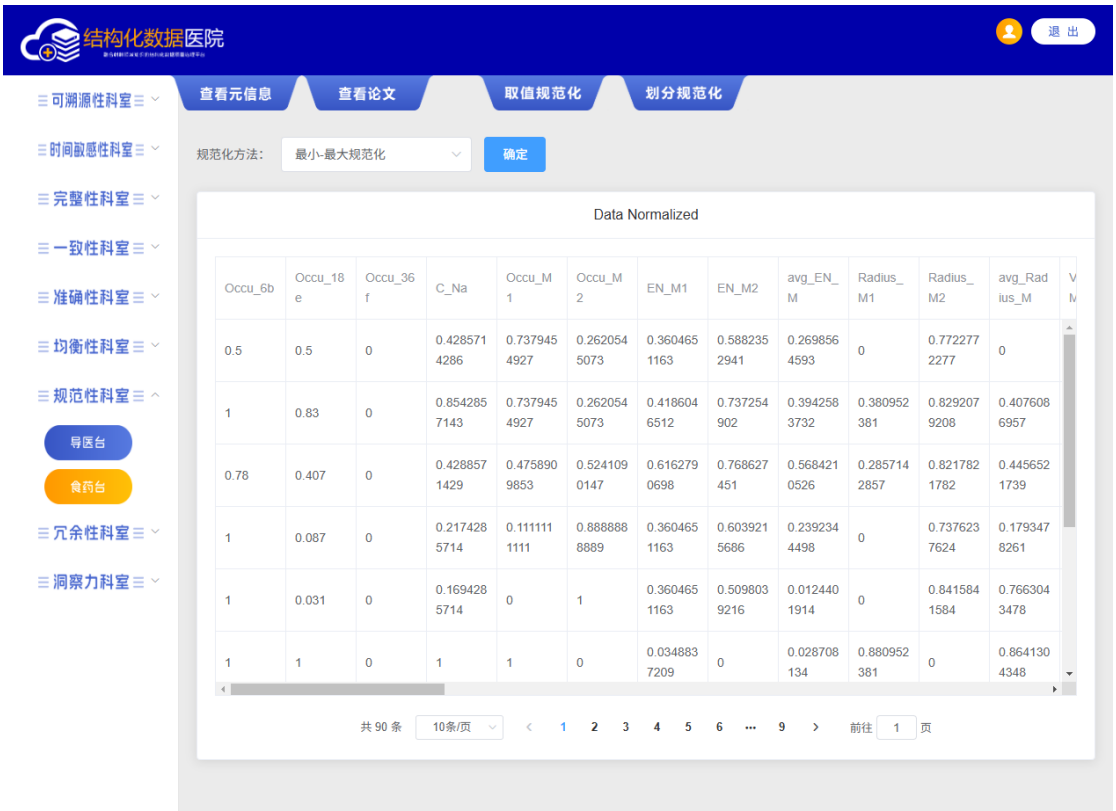

图 4-28 数据诊治规范性取值规范化治理结果

（3）点击平台页面上方的“划分规范化”，进入到划分规范化治理页面，运行结果如下图 4-29 所示。在“规范化方法”中根据自身需求选择对应的划分方法，此处以留一法为例进行测试。选择“留一法”，点击“确定”，平台会在“训练集”、“测试集”和“验证集”区域展示划分后的数据，运行结果如下图 4-30 所示。

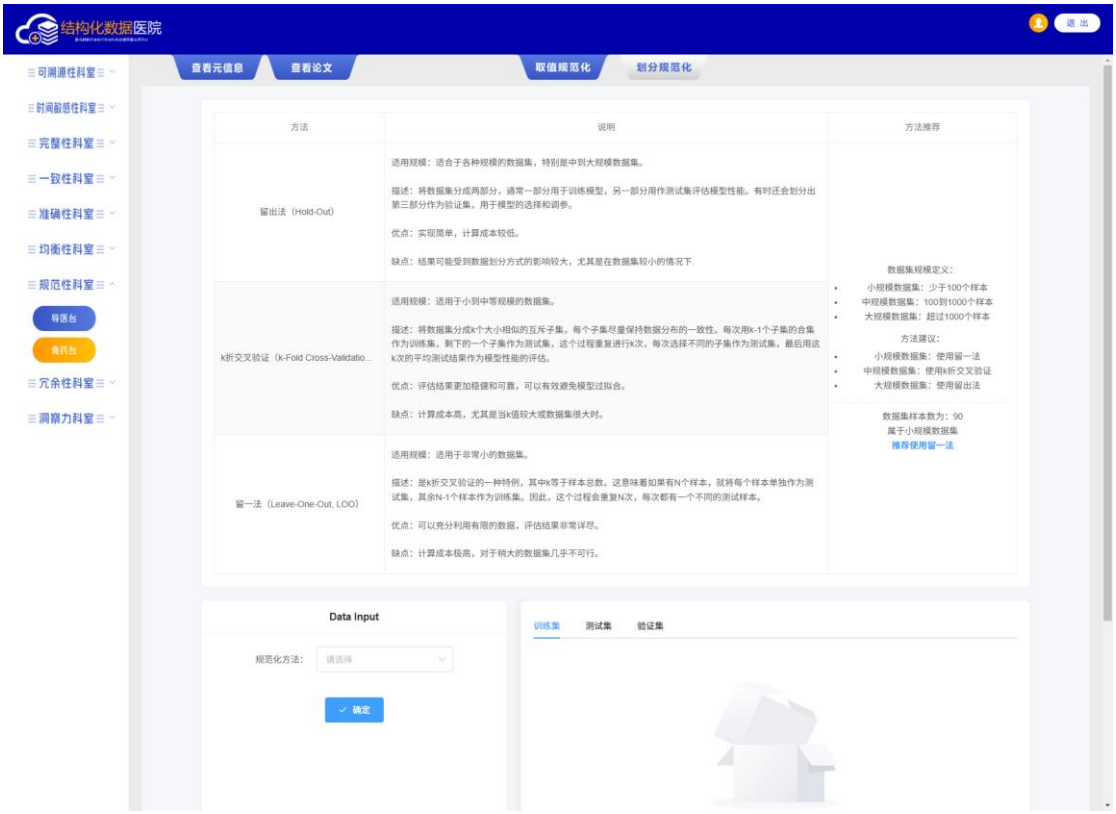

图 4-29 数据诊治规范性划分规范化食药台

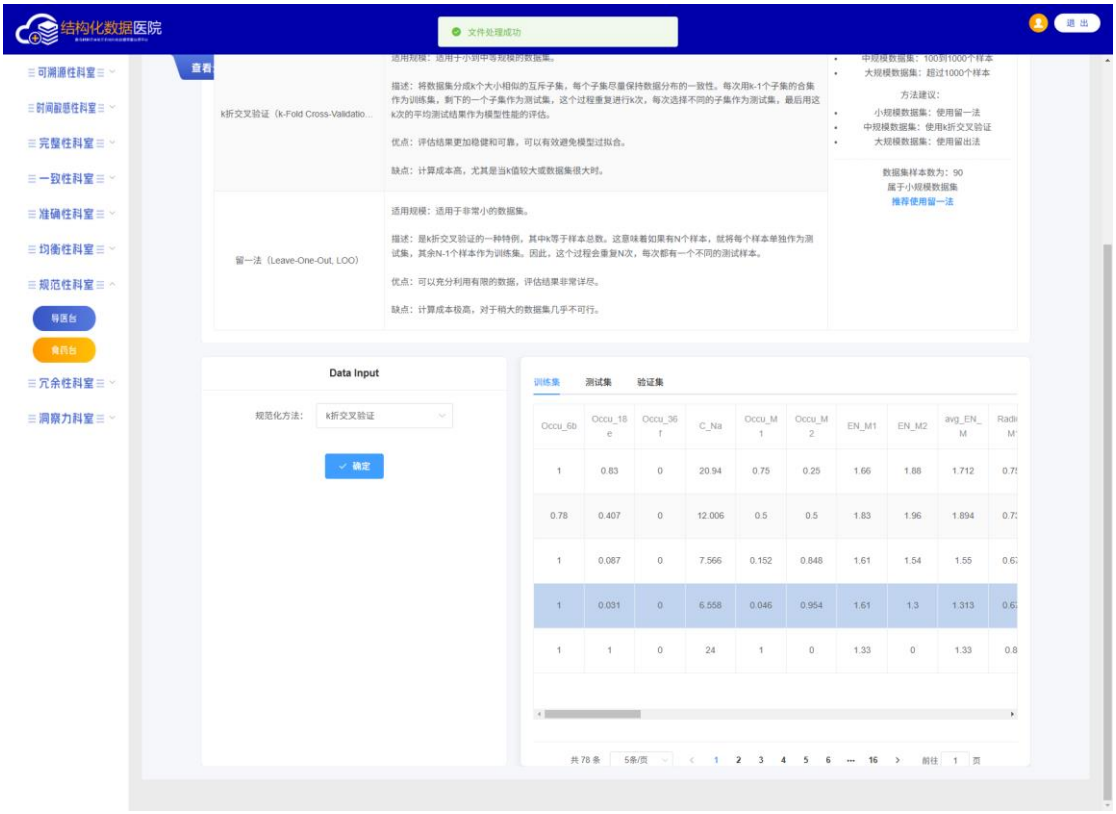

图 4-30 数据诊治规范性划分规范化治理结果

4.8 冗余性诊治

（1）点击平台页面左侧导航栏中的“冗余性科室”-“检查报告”，进入到冗余性数据管理界面，平台实现了对于所有上传数据集的查询、查看、下载等功能，运行结果如下图 4-31 所示。

图 4-30 数据诊治冗余性检查报告页面

（2）点击平台页面左侧导航栏中的“冗余性科室”-“导医台”，进入到冗余性文件选择页面，运行结果如下图 4-31 所示。

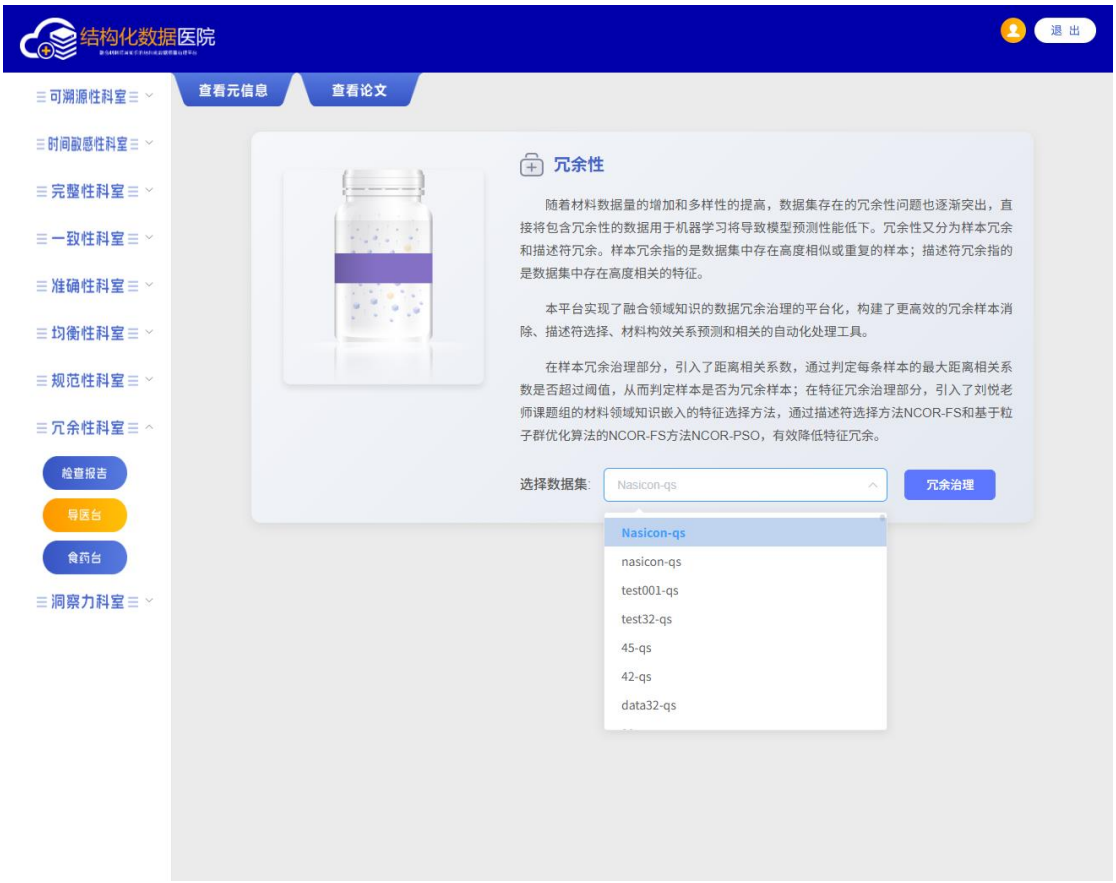

图 4-31 数据诊治冗余性导医台

（3）在“选择数据集”区域选择需要进行冗余性治理的数据，点击“冗余治理”，进入到数据集评估页面，运行结果如下图 57 所示。点击“下一步”，可在“距离相关系数阈值”区域根据需求输入对应的值，点击“下一步”平台开始计算冗余样本并进行呈现，运行结果如下图 58 所示。可在冗余样本概况区域根据自身需求对冗余样本进行勾选，并点击“删除选中项”对其进行删除，删除之后点击“前往可视化分析”，进入到冗余样本治理后的详情展示页面，运行结果如下图 4-32 所示。

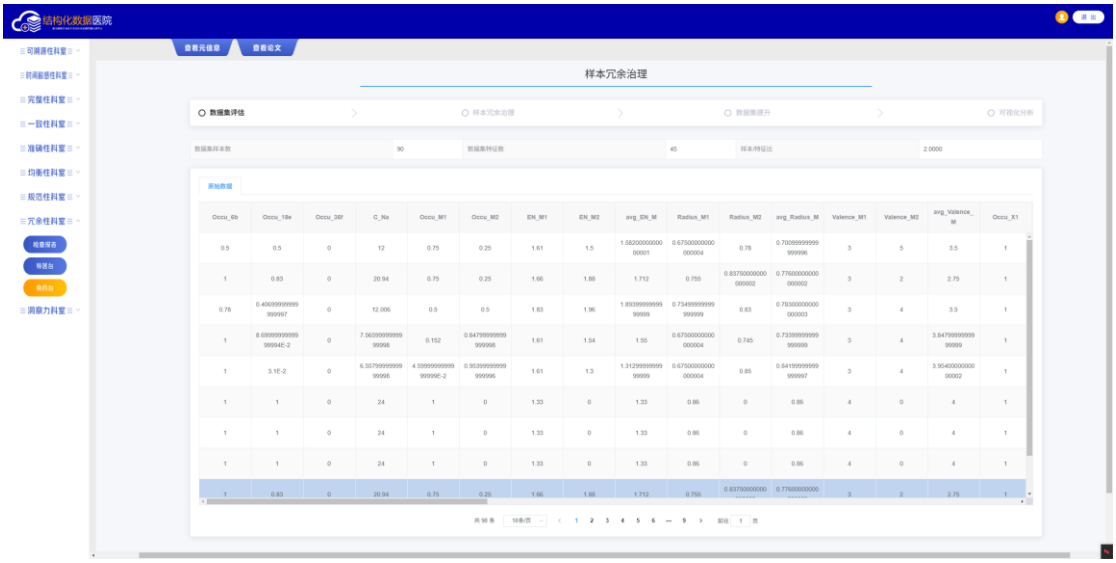

图 4-32 数据诊治冗余性食药台样本冗余

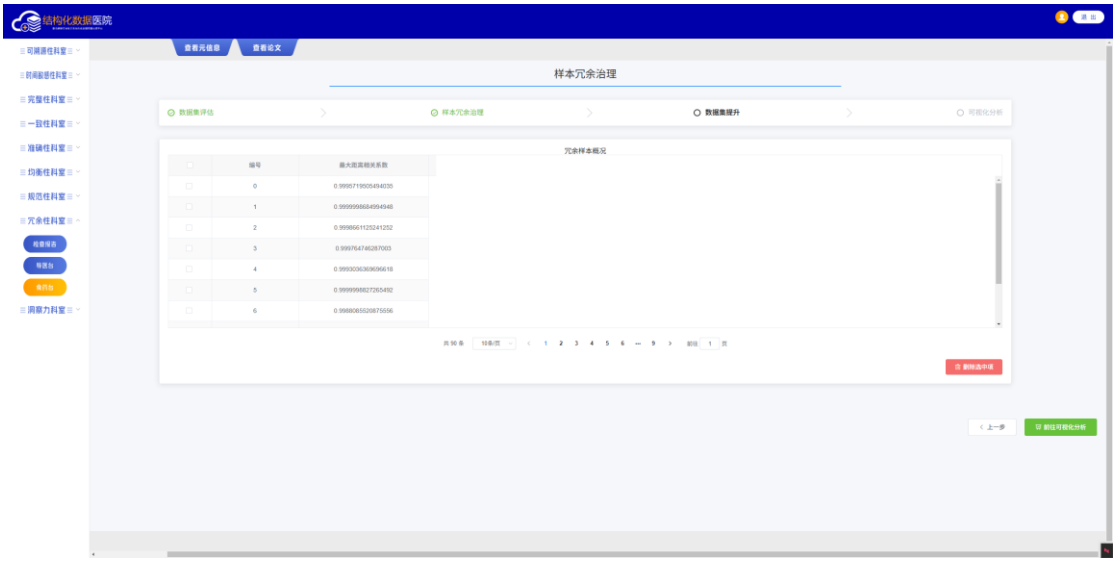

图 4-33 数据诊治冗余性食药台样本冗余

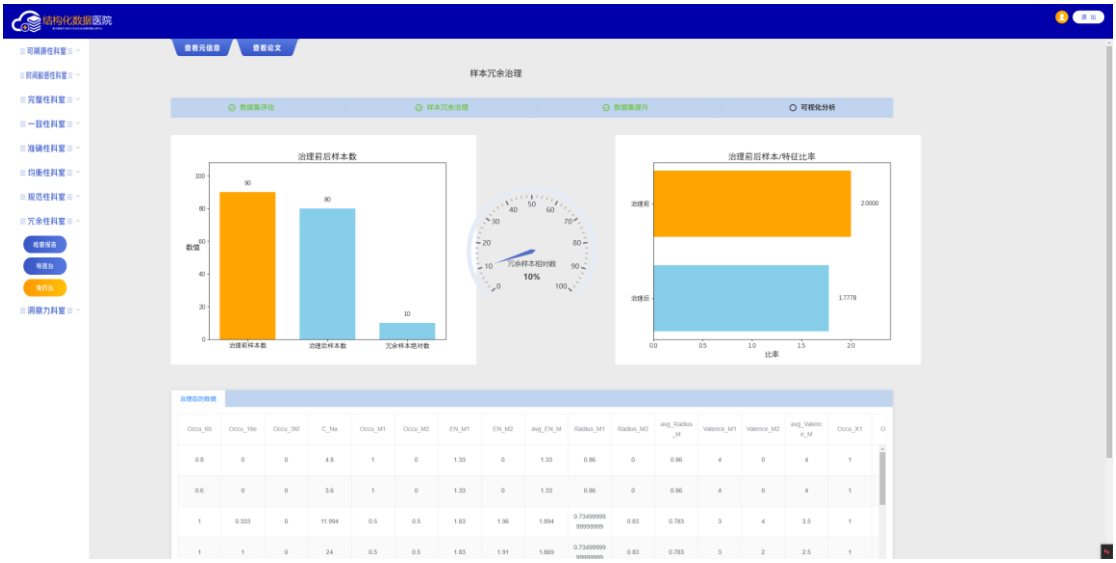

图 4-34 数据诊治冗余性食药台样本冗余

（4）点击“进入特征冗余治理”，进入到数据集评估页面，运行结果如下图 60 所示。点击“下一步”进入到 KNCOR 上传页面，可根据自身需求上传 KNCOR 文件，点击“选择文件”后选取对应文件即可，运行结果如下图 4-35 所示。点击“下一步”，进入到特征子集查看和添加页面，运行结果如下图 4-36 所示。点击“下一步”，进入到计算最优特征子集页面，点击“开始计算”，平台自动计算该数据集最有特征子集并进行呈现，运行结果如下图 4-37 所示。点击“下一步”，进入到冗余特征治理后的详情展示页面，运行结果如下图 4-38 所示。

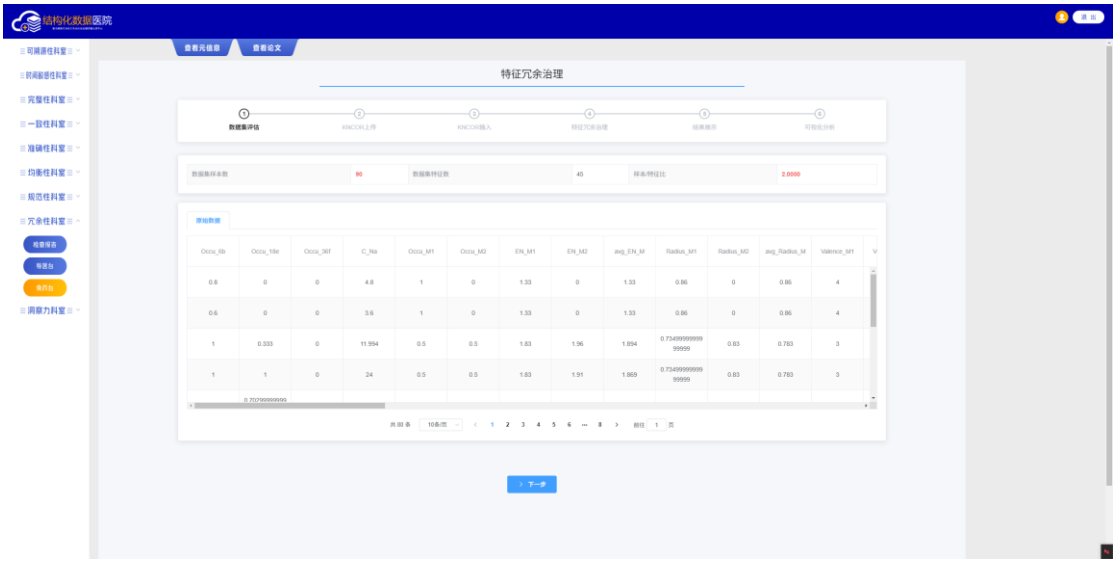

图 4-35 数据诊治冗余性食药台样本冗余

图 4-36 数据诊治冗余性食药台特征冗余

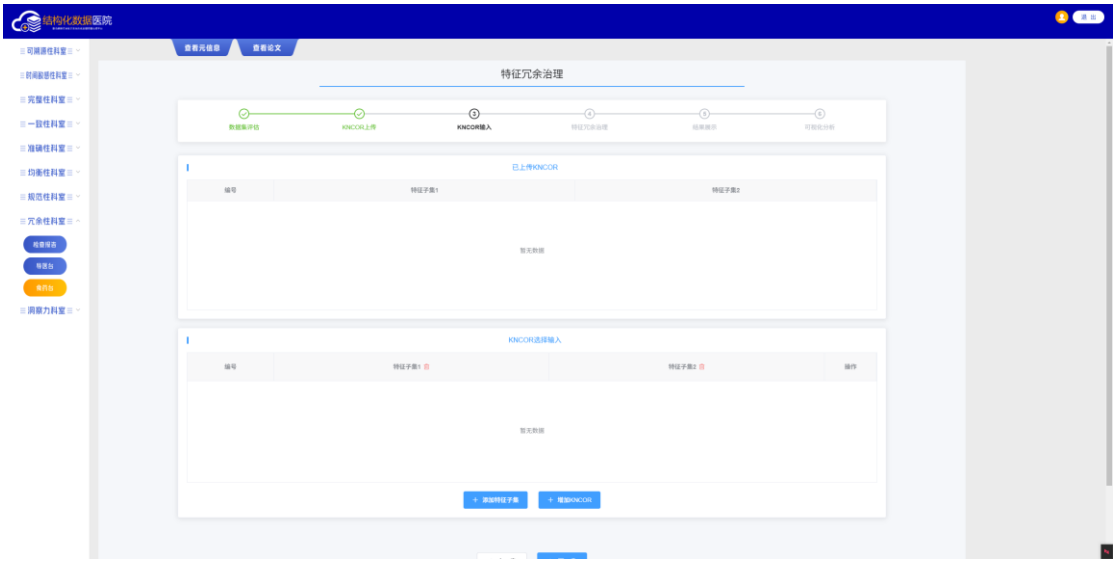

图 4-37 数据诊治冗余性食药台特征冗余

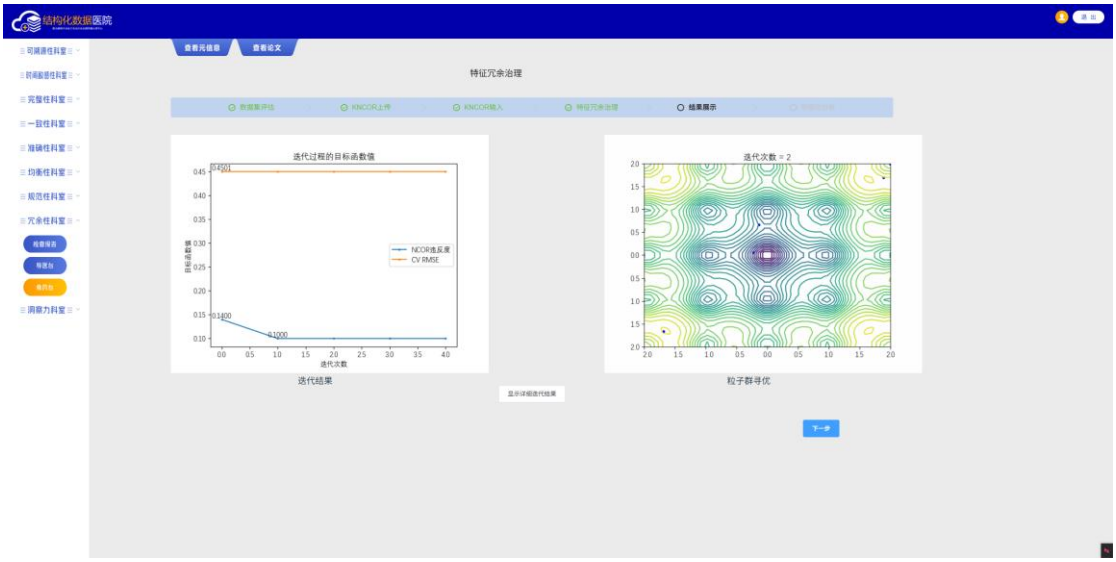

图 4-38 数据诊治冗余性食药台特征冗余

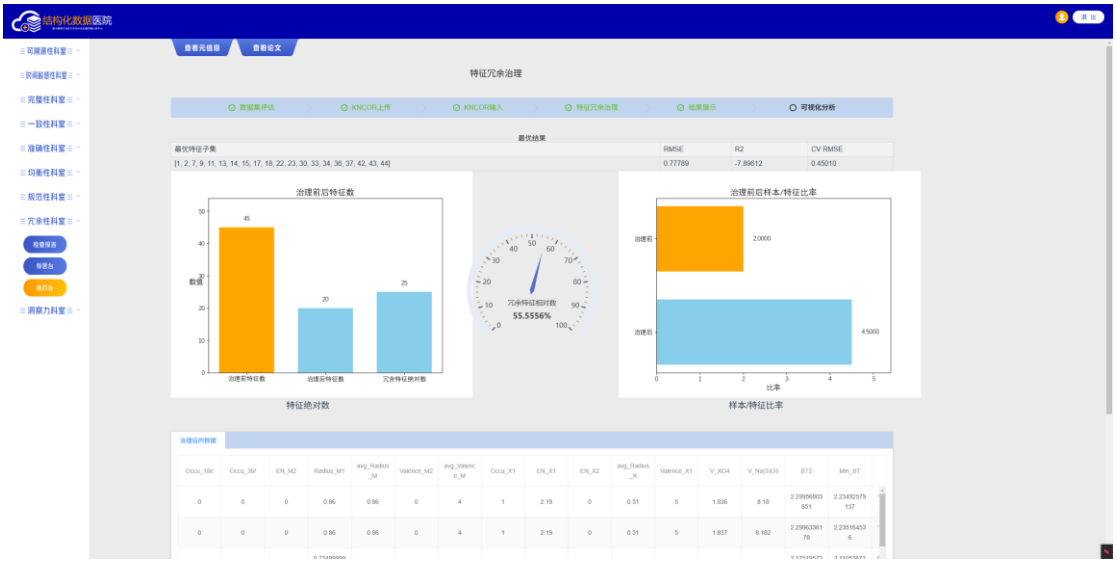

图 4-39 数据诊治冗余性食药台特征冗余

4.8 洞察力诊治

（1）点击平台页面左侧导航栏中的“洞察力科室”-“导医台”，进入到文件和功能选择页面，运行结果如下图 4-40 所示。分别选择划分规范性所生成的训练集和验证集，并根据下游任务选择“自动回归”或“自动聚类”，进入到洞察力检测页面，运行结果如下图 4-41 所示。

图 4-40 数据诊治洞察力导医台

图 4-41 数据诊治诊室台

（2）点击“自动回归”区域的“检测”，进入到元特征计算页面，运行结果如下图 4-42 所示。点击“元特征计算”，平台开始计算数据的元特征并进行呈现，运行结果如下图 4-43 所示。

图 4-42 数据诊治诊室台元特征计算前

图 4-43 数据诊治诊室台元特征计算后

（3）点击平台左侧导航栏中的“洞察力科室”-“检查报告”，进入到算法推荐页面，运行结果如下图 4-44 所示。点击“确认推荐”，平台开始计算每个算法的评分进行推荐并展示在页面上，运行结果如下图 4-45 所示。

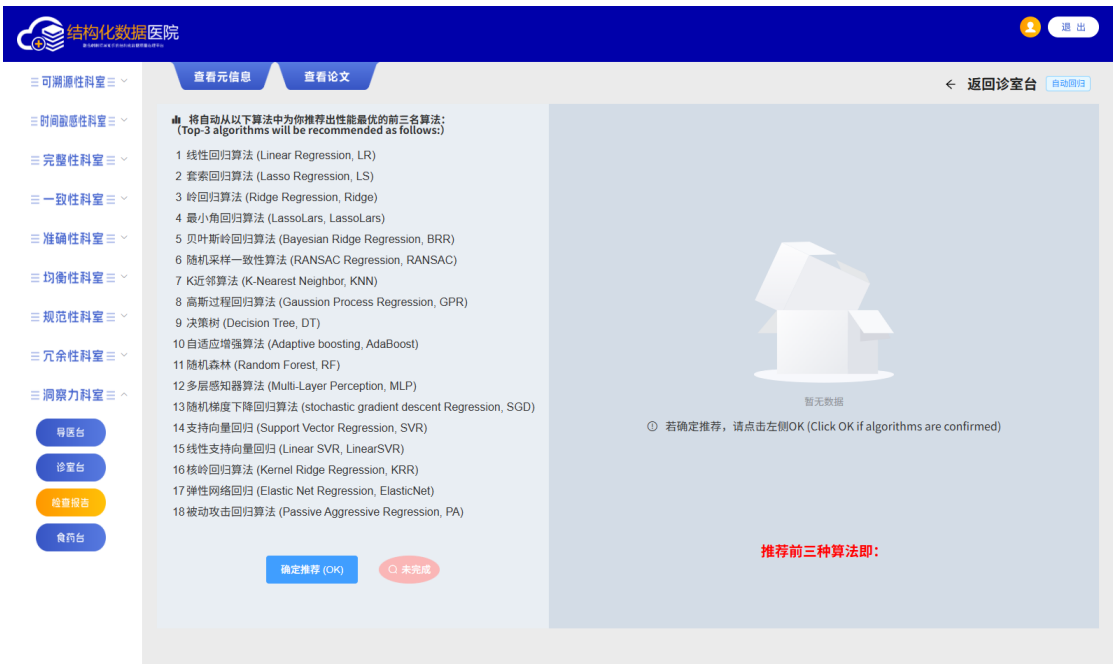

图 4-44 数据诊治检查报告前

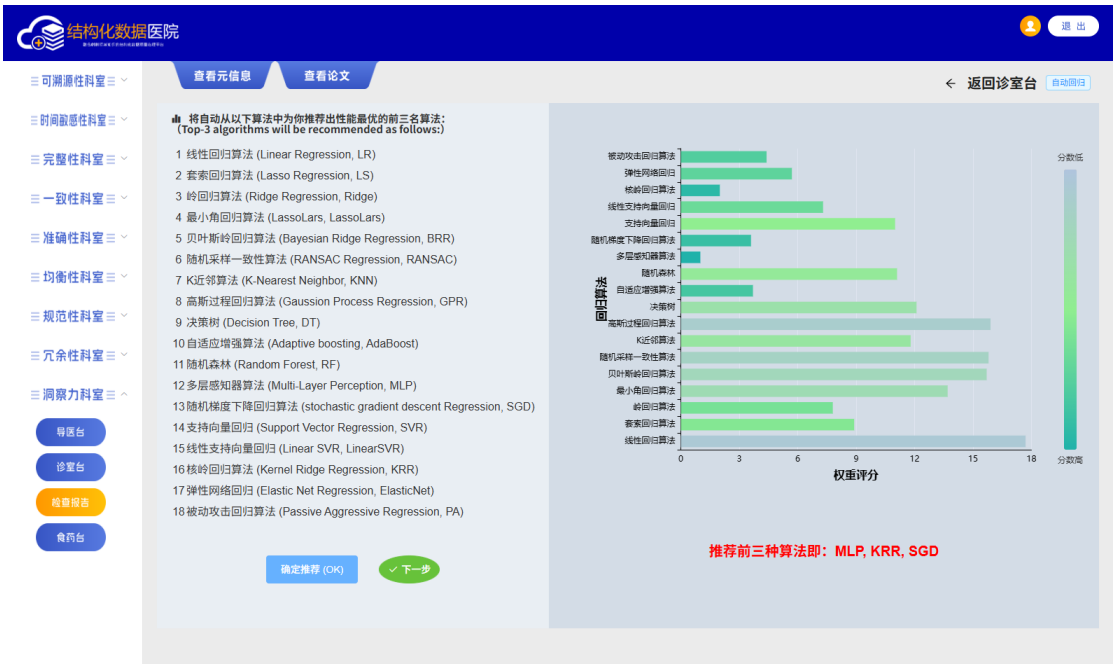

图 4-45 数据诊治检查报告后

（4）点击“下一步”，进入到模型对比效果页面，运行结果如下图 4-46 所示。点击“开始预测”，平台计算在前三种模型上该数据集对应的均方根误差等结果并呈现在页面上，运行结果如下图 4-47 所示。

图 4-46 数据诊治食药台

图 4-47 数据诊治食药台预测后