

组学大数据平台在肿瘤精准医疗中的应用

中国医学科学院 系统医学研究中心
Center of Systems Medicine, Chinese Academy of Medical Sciences

苏州系统医学研究所
Suzhou Institute of Systems Medicine

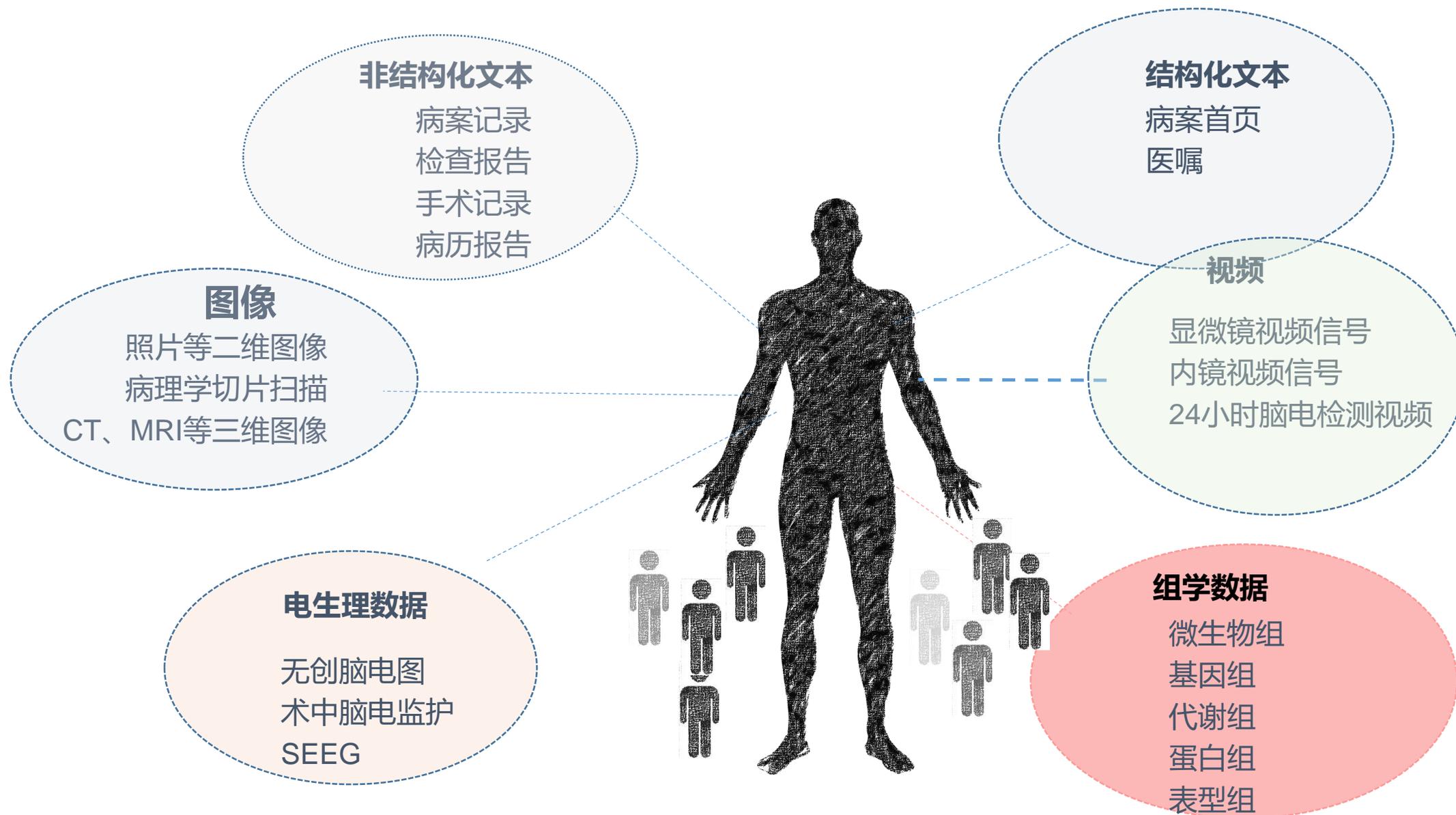


秦晓峰

中国医学科学院医学生物信息大数据中心苏州系统医学研究所



医疗行业产生大量数据



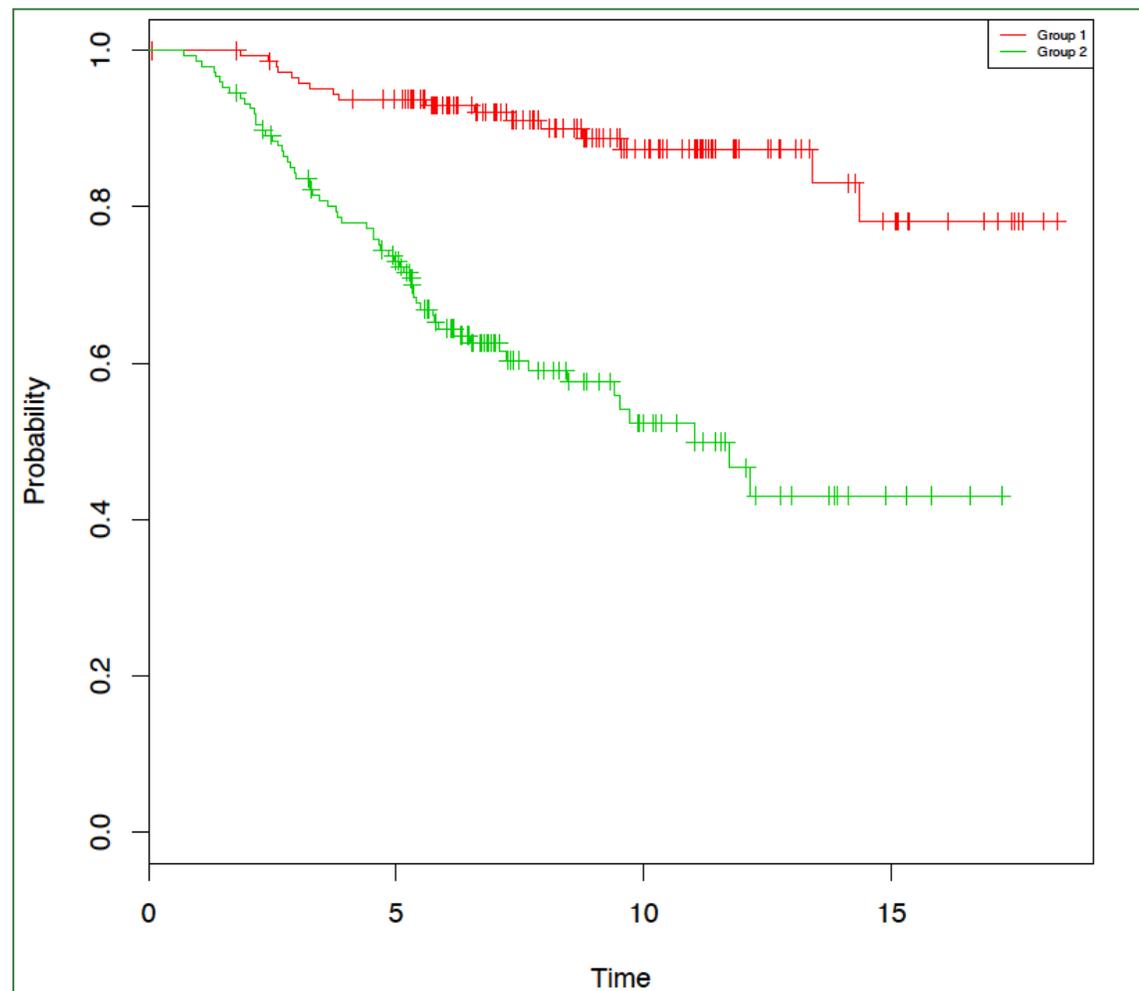
临床数据来源和分析

- 临床数据来源:

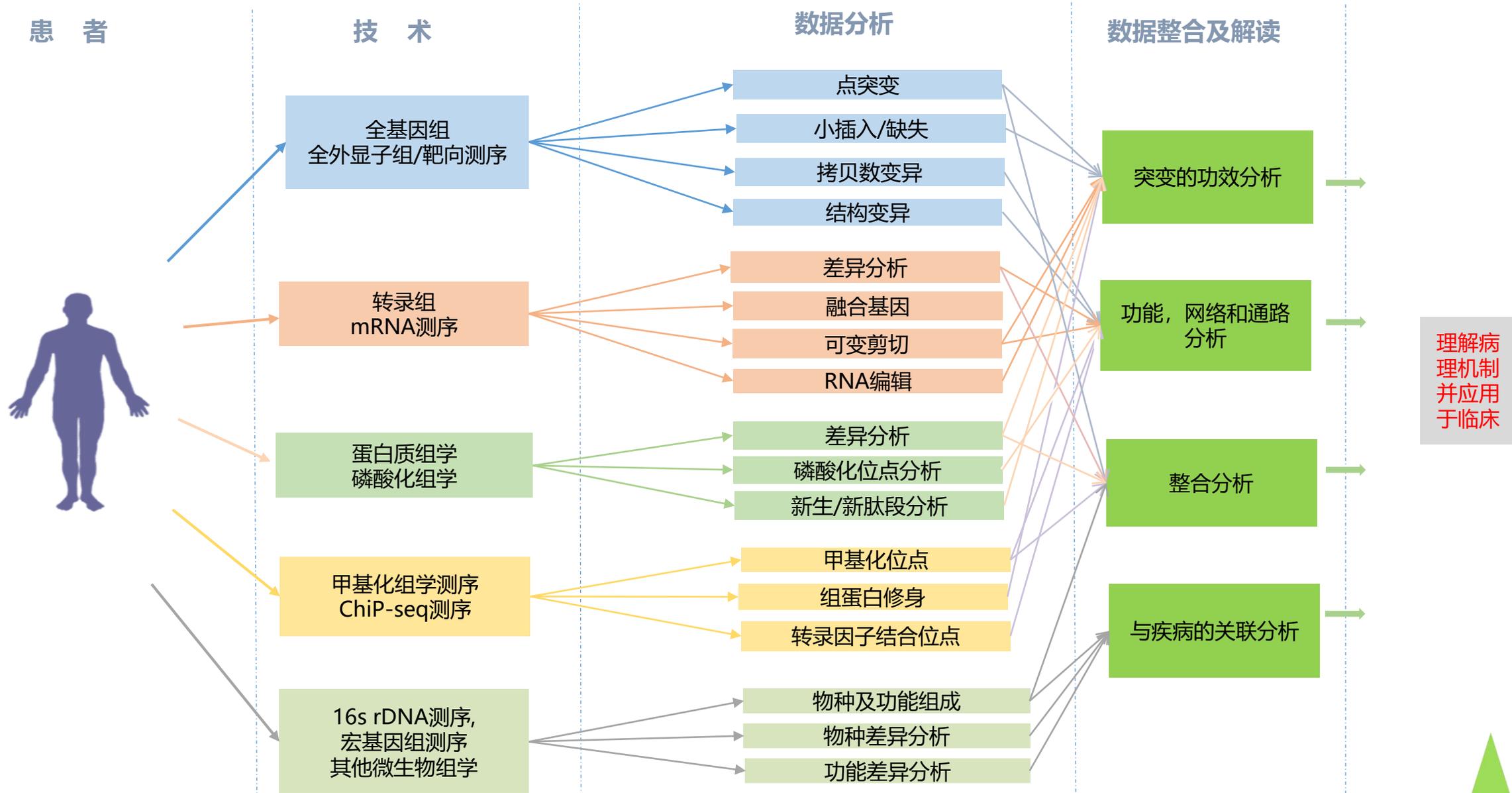
年龄
性别
过敏情况
药物测试
疾病详情
家族史
药物接受和排斥
曾使用剂量水平生存率诊断测试
手术

- 临床数据分析:

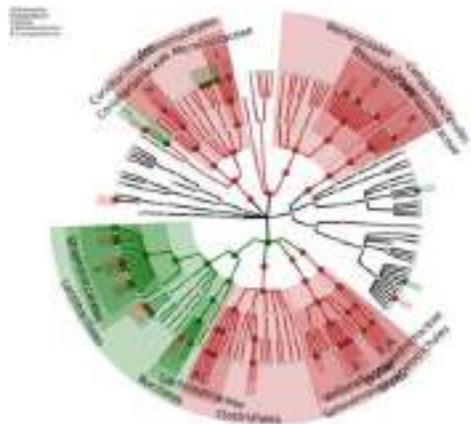
生存分析
预测



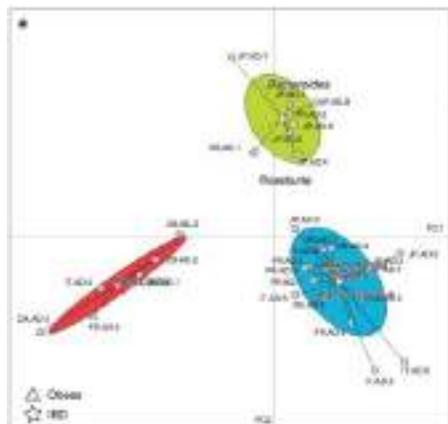
组学数据来源和分析



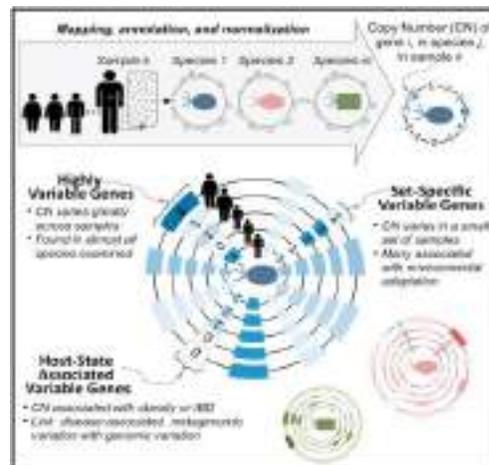
宏基因组深度挖掘



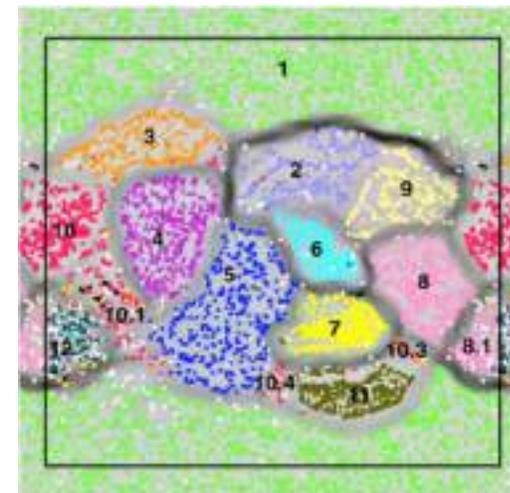
挖掘组间物种、功能差异



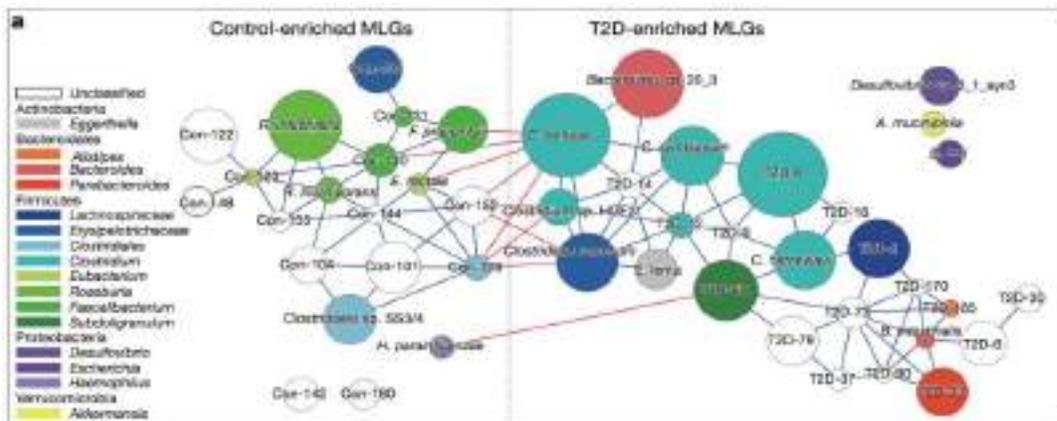
样品聚类分析 (肠型)



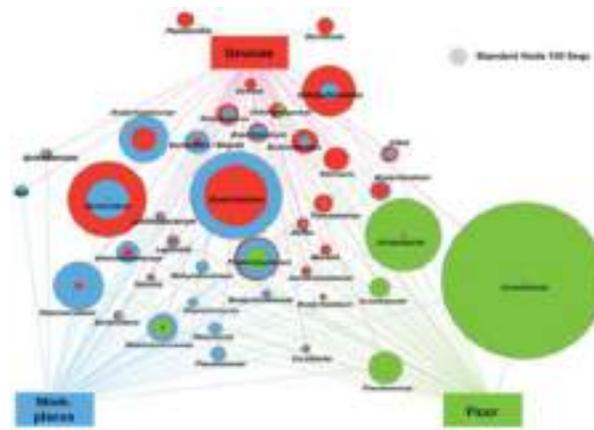
拷贝数变异：挖掘功能变化



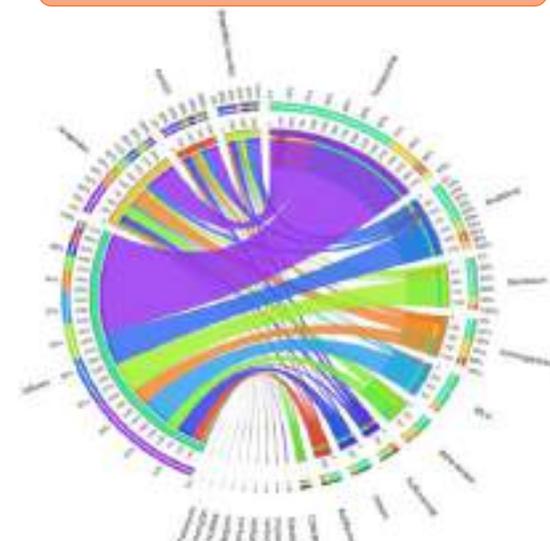
从宏基因组数据中组装单菌



CAG/MGS/MLG分析：从种或菌株层级挖掘物种变化

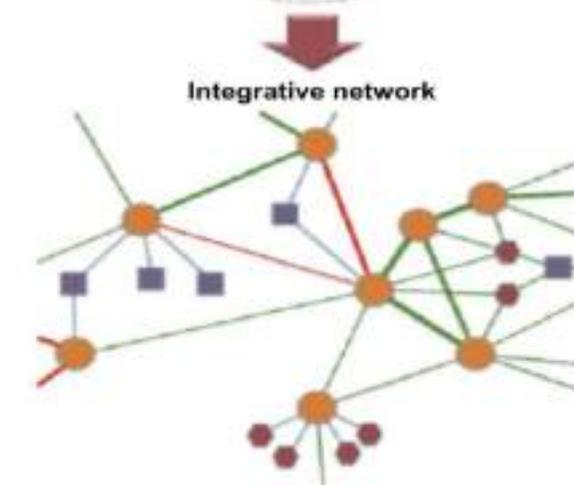
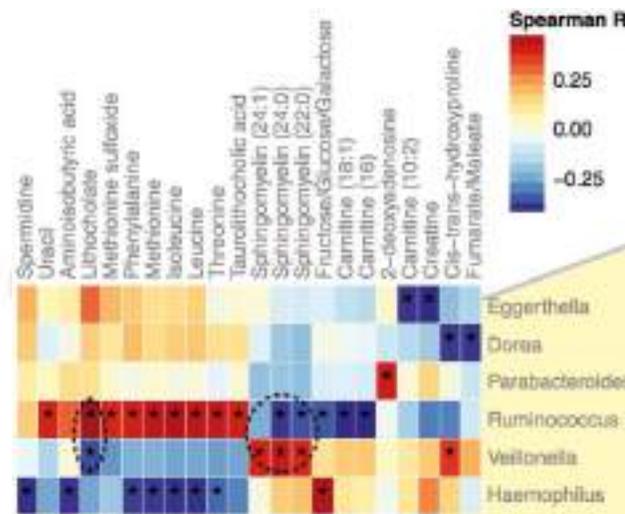
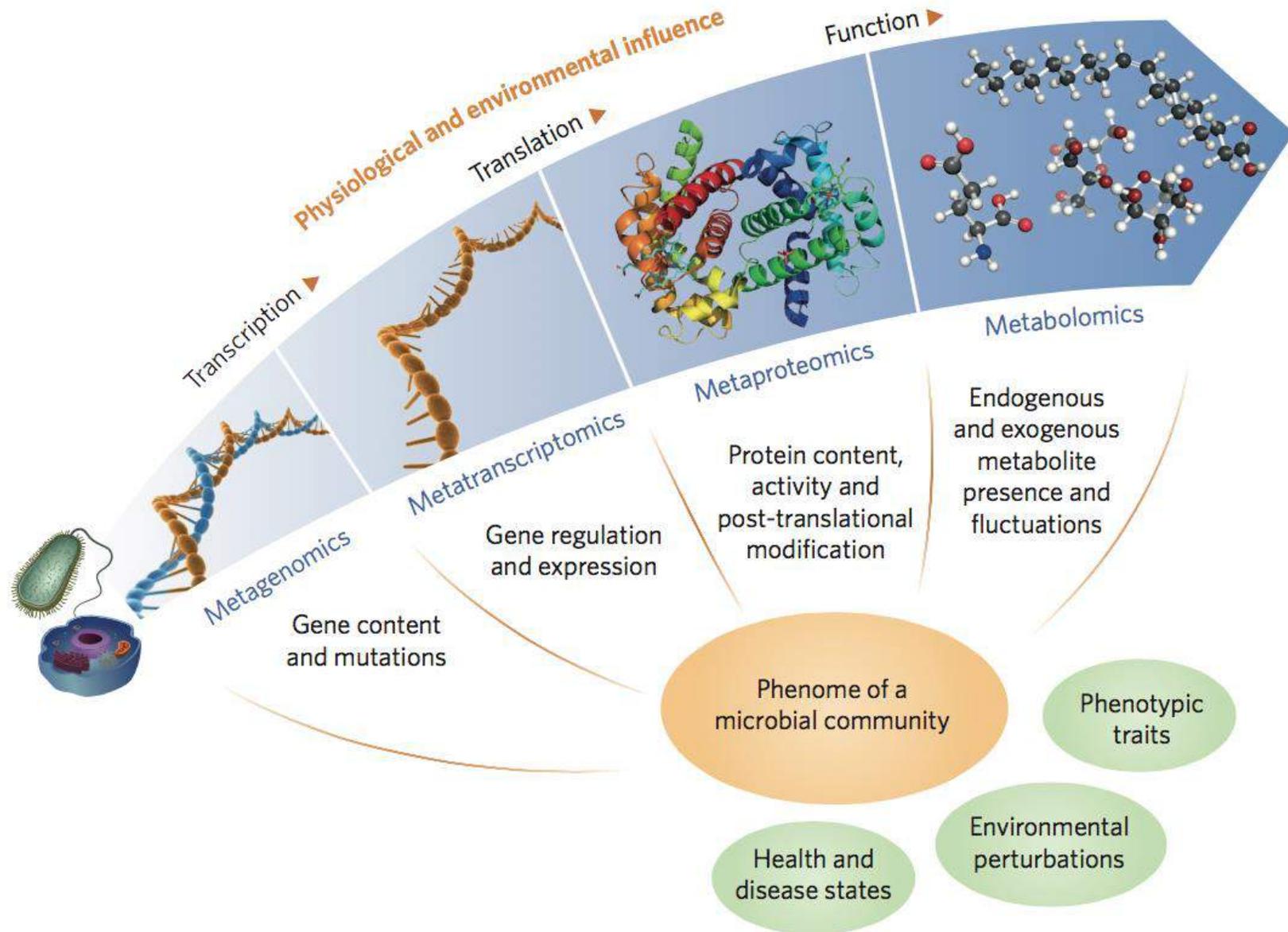


菌群、表型、临床数据关联分析



耐药基因挖掘

多组学联合分析



Integrated networks for multiomic data.

微生物组在肿瘤免疫治疗中的应用

- **期刊:** Science
- **发表时间:** 2017.11
- **实验设计:** 249名接受过抗PD-1免疫疗法的肺癌、肾癌等多种不同的癌症；免疫治疗前69名患者接受了抗生素的治疗；
- **研究技术:** 粪便宏基因组测序
- **验证:** 无菌小鼠FMT（粪菌移植）验证；
- **结果:** 1) 抗生素治疗的患者，癌症很快出现复发，生存时间也更短；
2) 恢复较好的患者体内，*Akkermansia muciniphila* 的益生菌丰度更高，对癌症免疫疗法还有促进作用；
3) 接受了“起效者”粪便的小鼠对于PD-1抑制剂的反应要明显优于接受了“无效者”粪便的小鼠，后者在口服 *Akkermansia muciniphila* 后，能恢复对免疫疗法的反应。

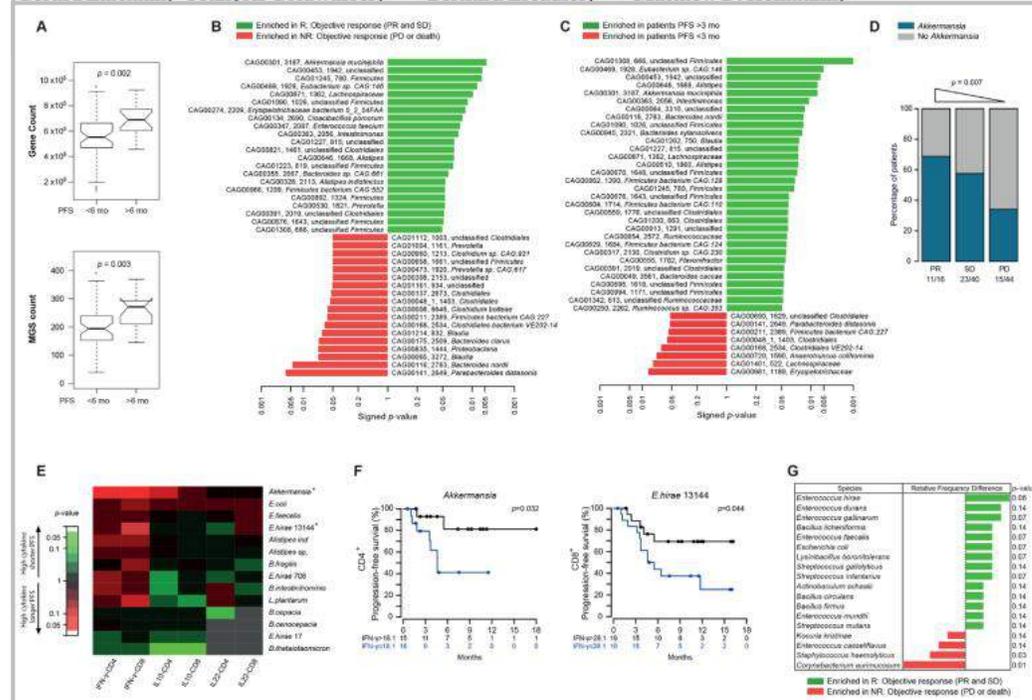
Science

REPORTS

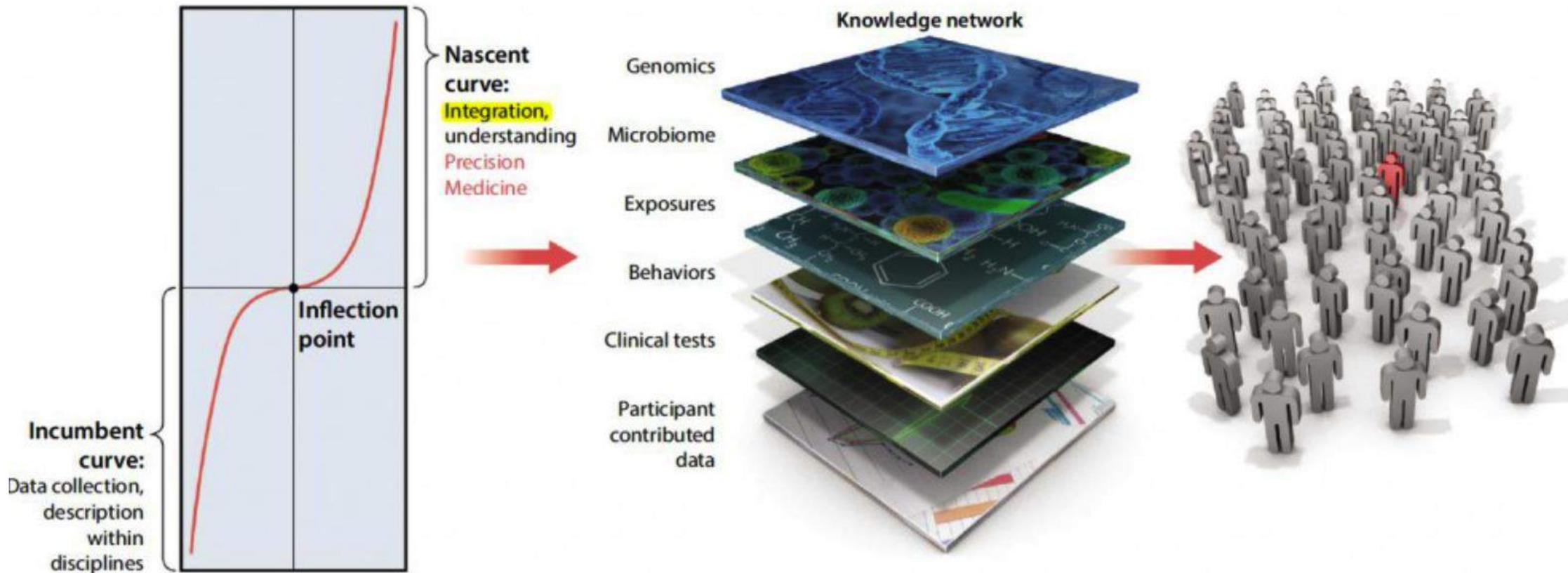
Cite as: B. Routy *et al.*, *Science* 10.1126/science.aan3706 (2017).

Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors

Bertrand Routy^{1,2,3}, Emmanuelle Le Chatelier,⁴ Lisa Derosa,^{1,2,3} Connie P. M. Duong,^{1,2,5} Maryam Tidjani Alou,^{1,2,3} Romain Daillère,^{1,2,3} Aurélie Fluckiger,^{1,2,5} Meriem Messaoudene,^{1,2} Conrad Rauber,^{1,2,3} Maria P. Roberti,^{1,2,5} Marine Fidelle,^{1,3,5} Caroline Flament,^{1,2,5} Vichnou Poirier-Colame,^{1,2,5} Paule Opolon,⁶ Christophe Klein,⁷ Kristina Iribarren,^{8,9,10,11,12} Laura Mondragón,^{8,9,10,11,12} Nicolas Jaquetol,^{1,2,3} Bo Qu,^{1,2,3} Gladys Ferrere,^{1,2,3} Céline Clémenson,^{1,13} Laura Mezquita,^{1,14} Jordi Remon Masip,^{1,14} Charles Naltet,¹⁵ Solenn Brosseau,¹⁵ Coureche Kaderbhai,¹⁶ Corentin Richard,¹⁶ Hira Rizvi,¹⁷ Florence Levenez,⁴ Nathalie Galleron,⁴ Benoit Quinquis,⁴ Nicolas Pons,⁴ Bernhard Ryffel,¹⁸ Véronique Minard-Colin,^{1,10} Patrick Gonin,^{1,20} Jean-Charles Soria,^{1,14} Eric Deutsch,^{1,13} Yohann Loriot,^{1,3,14} François Ghiringhelli,¹⁶ Gérard Zalcman,¹⁵ Francois Goldwasser,^{9,21,22} Bernard Escudier,^{1,14,23} Matthew D. Hellmann,^{24,25}



医生目前面临决策的信息维度大大增加



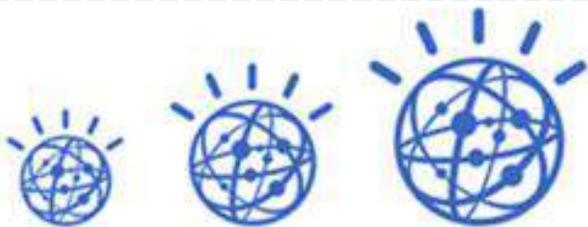
Hawgood S, Hook-Barnard IG, O' Brien TC, Yamamoto KR. Precision medicine: Beyond the inflection point. Science translational medicine 2015;7:300ps17.



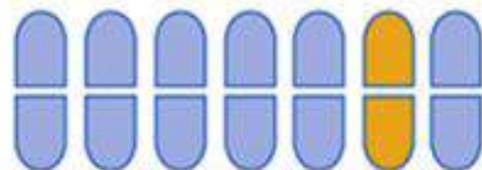
阅读和记忆学习医
疗文献、临床指导和
医学指南



**将病人和临床试验
方案进行匹配**



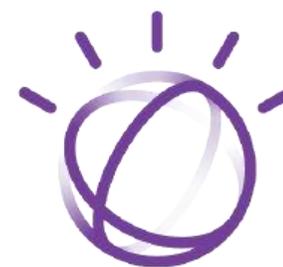
持续不断的学习
从不断增加的病人的
组学数据和临床数据
中不断学习



依据最新用药指
导推荐**潜在的治
疗选择方案**

Watson医生

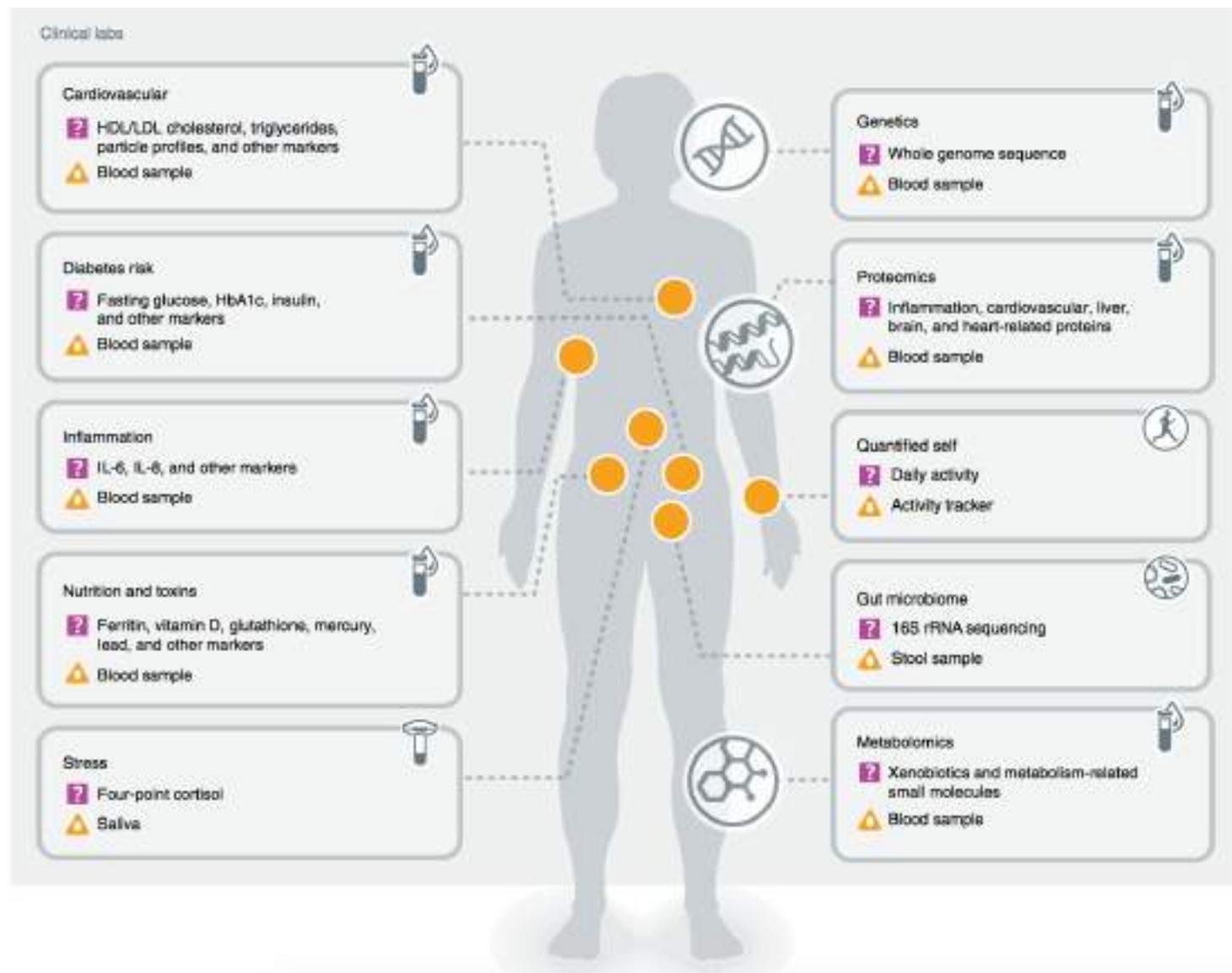
- 由IBM公司开发
- 人工智能系统
- 询问病人的病征、病史
- 人工智能技术
- 自然语言的处理和分析技术
- 从各个渠道搜集到的信息和数据
- 迅速给出诊断提示和治疗意见



针对个人进行纵向密集数据收集可以揭示分子疾病标志物

- 前瞻性
- 108个人
- 全基因组测序分析
- 临床检测分析
- 蛋白质组学分析
- 代谢组学分析
- 微生物群落分析 (对16S rRNA进行测序)
- 参与者配戴活动跟踪器监测日常活动
- 创立相关性网络
- 关联分析
- 鉴定已知和候选标志物
- Meta分析

Price N D, Magis A T, Earls J C, et al. A wellness study of 108 individuals using personal, dense, dynamic data clouds[J]. Nature Biotechnology, 2017, 35(8):747.



在癌症治疗中的联合用药

- 不同癌症分期的医学研究
- 基因和分子诊断
- 肿瘤信息学
- 传统中药
- 数学分析
- 治疗毒性评价
- 个性化用药

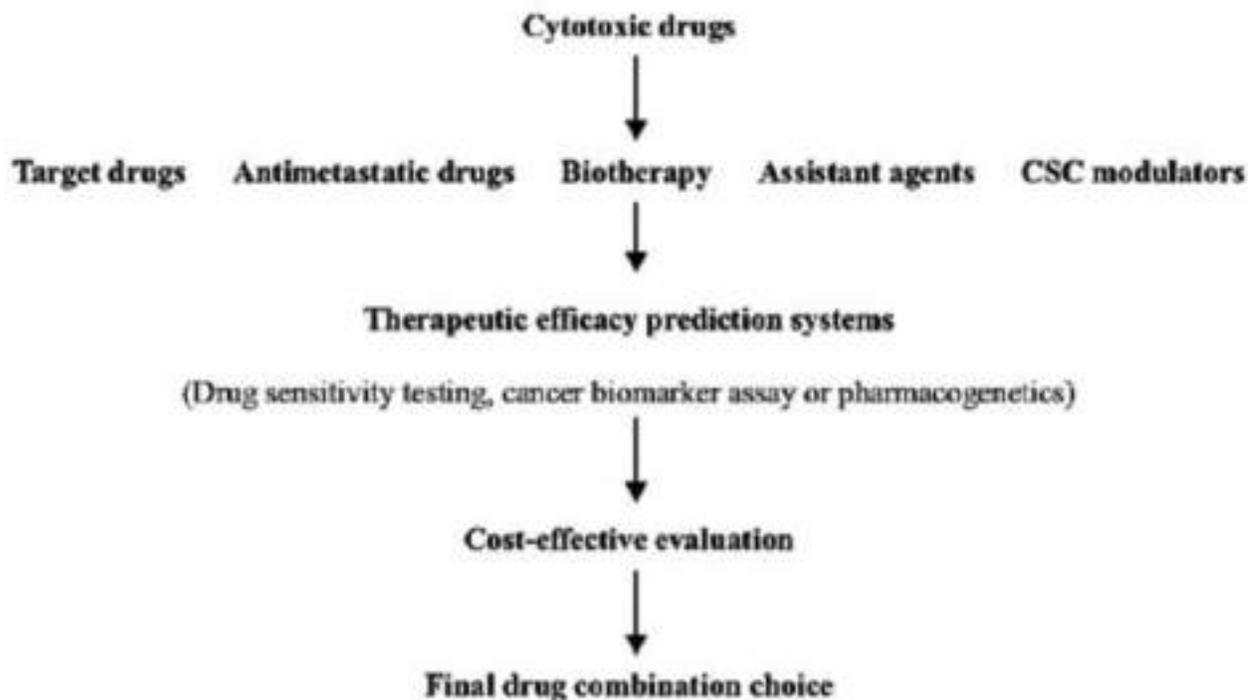
Drug Combination in Clinical Cancer Treatments

Author(s): Da-Yong Lu*, Ting-Ren Lu, Nagendra Sastry Yarla, Hong-Ying Wu, Bin Xu, Jian Ding, Hong Zhu.

Journal Name: Reviews on Recent Clinical Trials

Volume 12 , Issue 3 , 2017

DOI : 10.2174/1574887112666170803145955



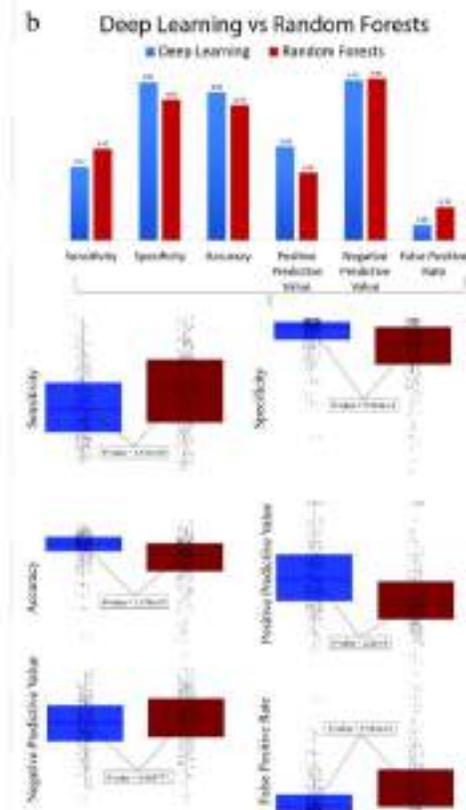
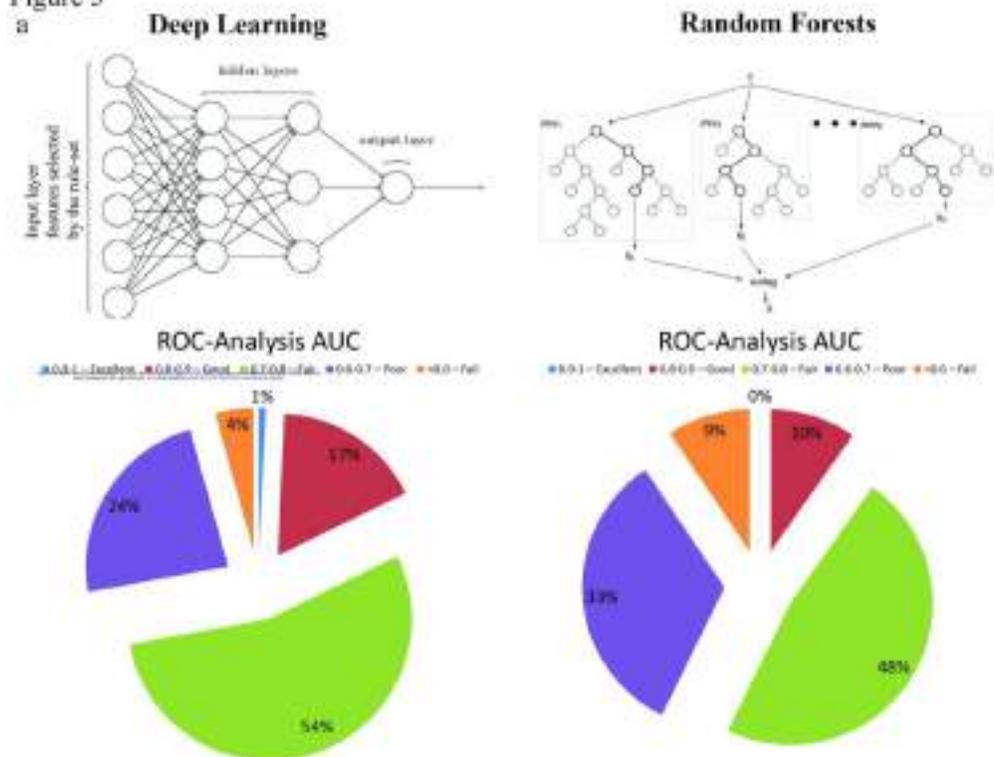
利用深度学习和关联规则挖掘预抗癌药物反应

- 数据来源：药物基因组689个癌症细胞系和139种抗癌药物。来自CCLP和GDSC.
- 规则关联挖掘
- 深度学习
- 预测药物反应

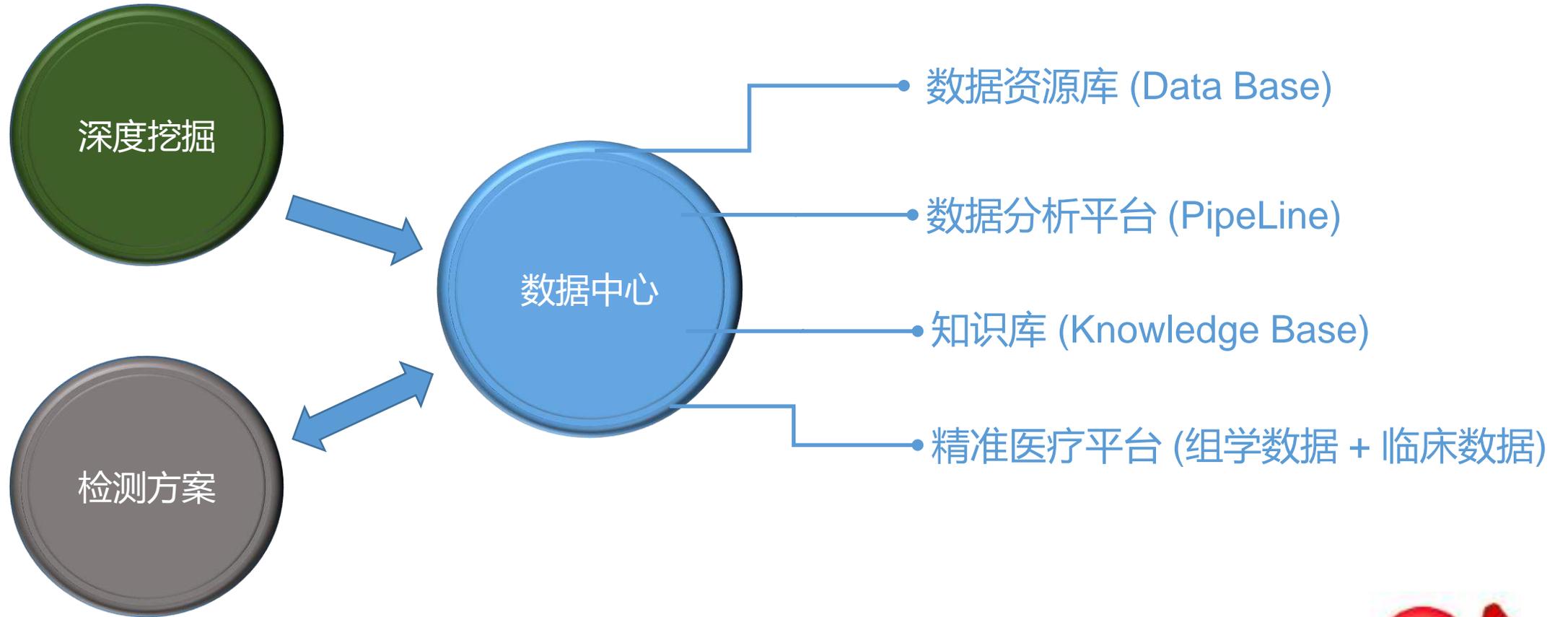
Deep Learning and Association Rule Mining for Predicting Drug Response in Cancer

Konstantinos N Vougas, Thomas Jackson, Alexander Polyzos, Michael Lontos, Elizabeth O Johnson, Vassilis Georgoulas, Paul Townsend, Jiri Bartek, Vassilis G Gorgoulis
doi: <https://doi.org/10.1101/070490>

Figure 5



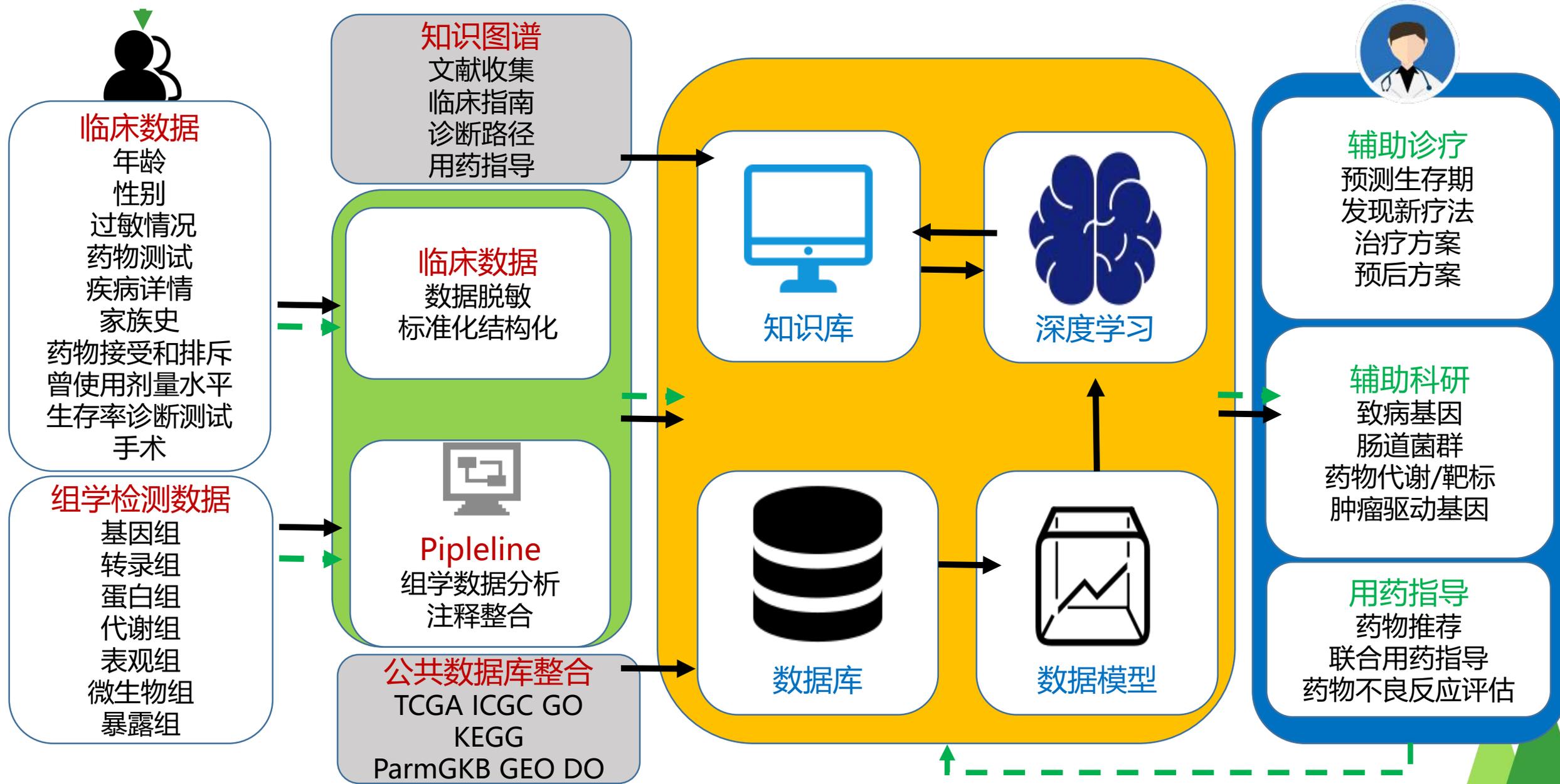
组学大数据平台与精准医疗



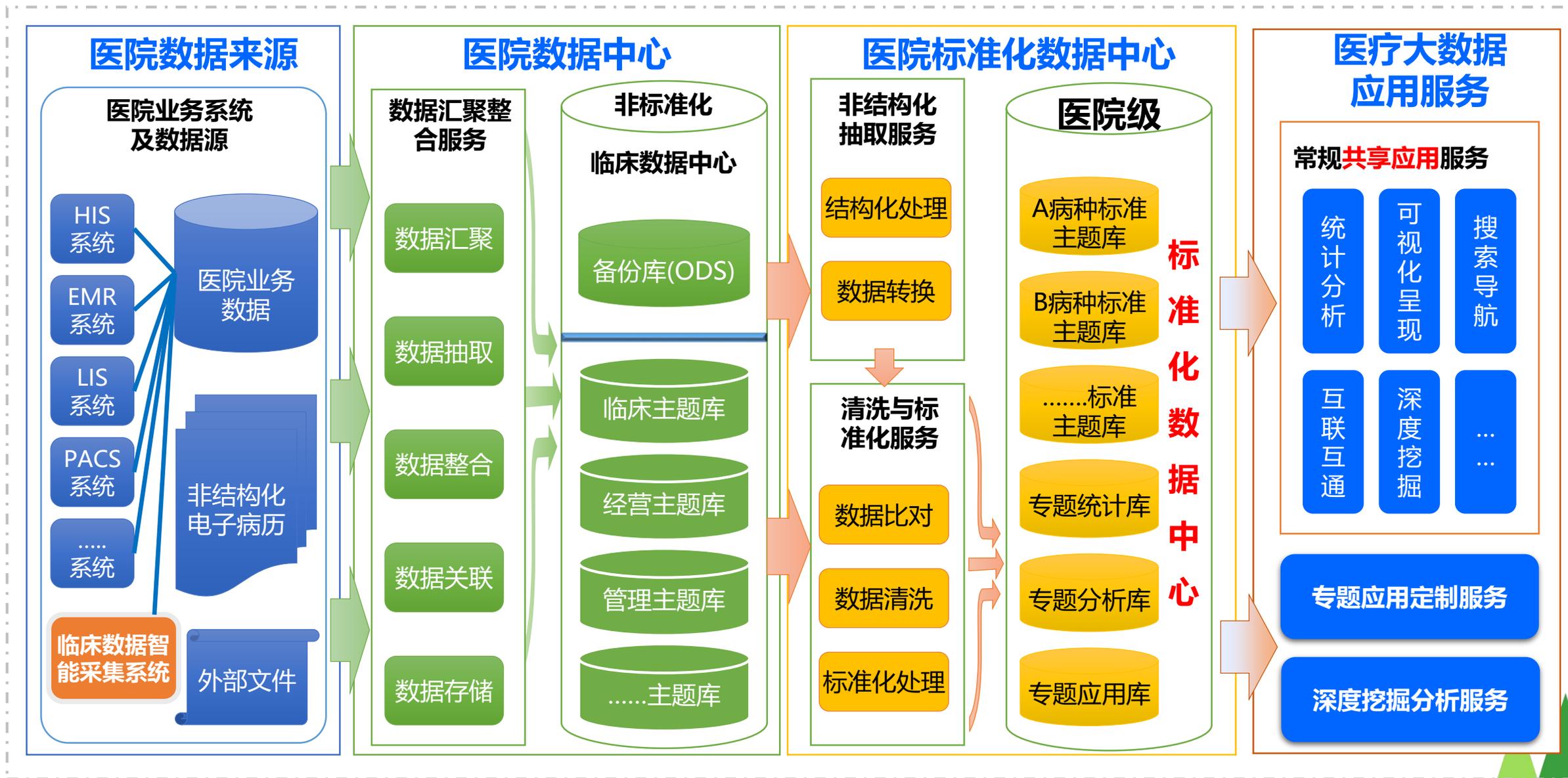
- 荧光定量PCR、基因芯片、SNP分型、二代测序



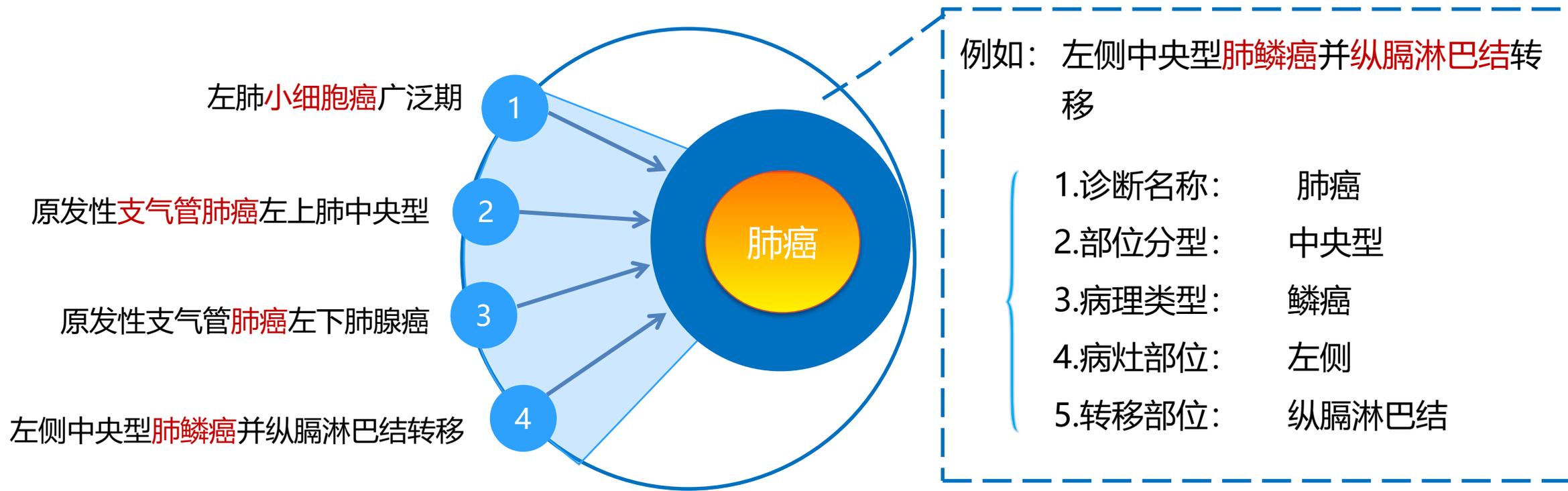
组学大数据平台在肿瘤精准医疗中的应用



医疗大数据来源



医学术语标准化

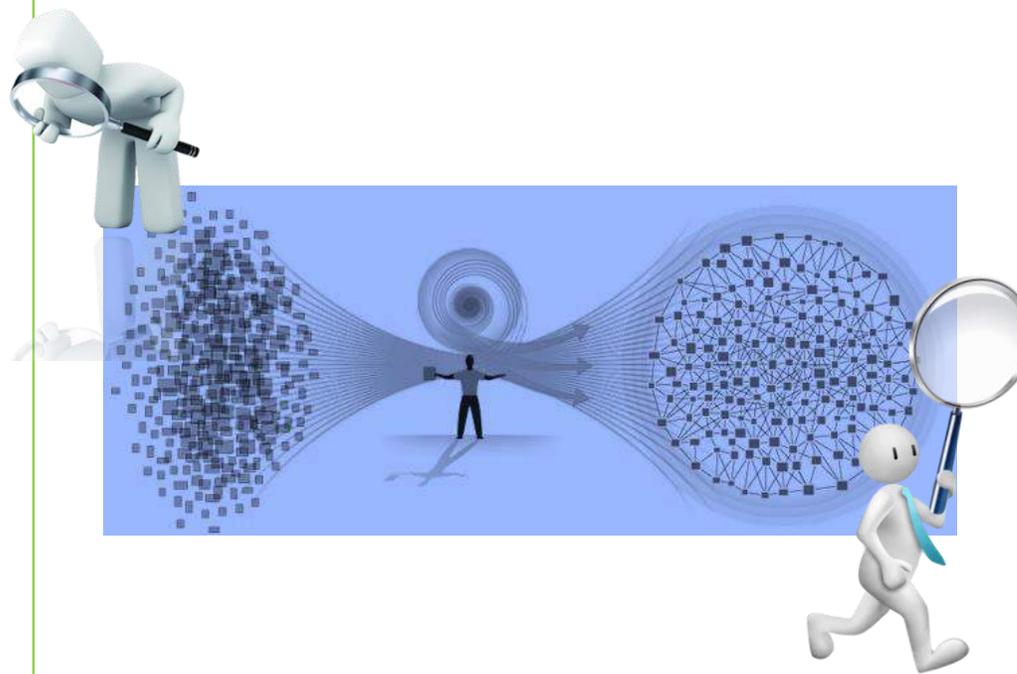


将不同描述的诊断转化成统一的标准化诊断，并且保留诊断中的重要信息。

临床数据结构化

现有电子病历检查描述

患者不慎摔伤，伤后神志不清约30余秒，后可唤醒，体温正常（36.5℃），感头痛明显，左侧鼻腔内有少量鼻血。于我院行腹部B超及胸片，未见明显异常。入院体格检查：肠鸣音正常，约4次/分。



处理后的结构化检查描述

指标名称	结果
摔 伤	有
神志不清	30余秒
体 温	36.5℃
头 痛	明显
左侧鼻腔	少量流血
腹部B超	未见明显异常
胸 片	未见明显异常

将自然语言描述的电子病历转化成结构化、标准化的电子病历

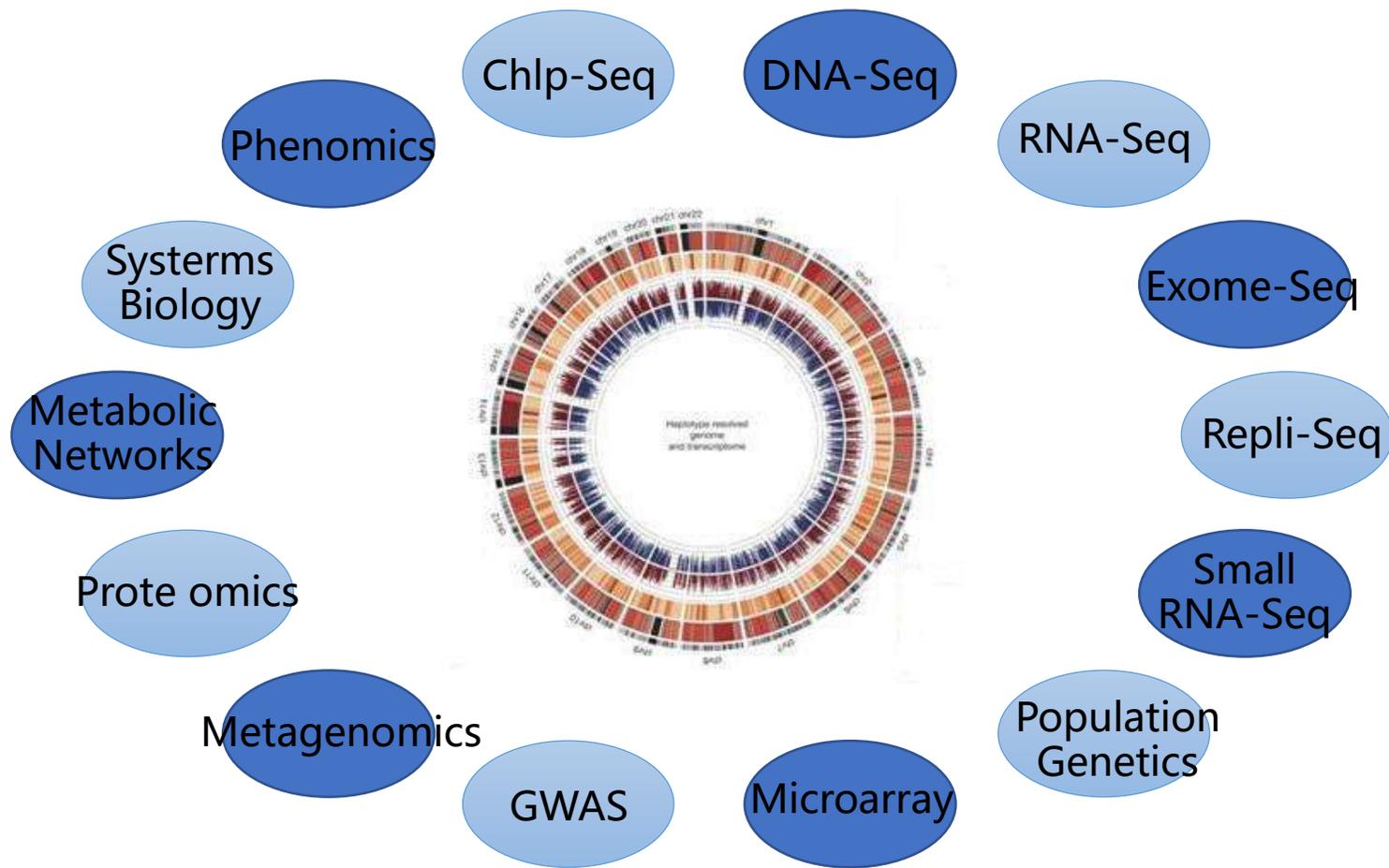
覆盖各种可能的组学

组学数据

- 基因组学数据
- 转录组学数据
- 蛋白组学数据
- 表型组学数据
- 代谢组学数据
- 微生物组学数据
-

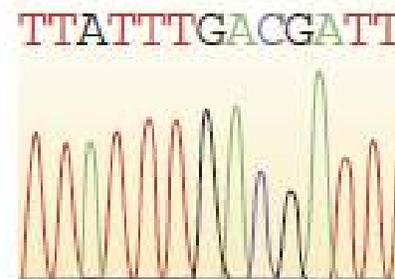
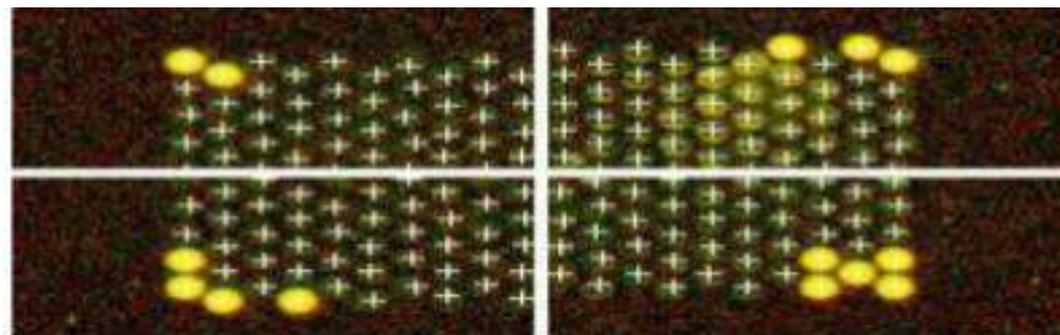
组学数据格式

fasta/fastq/vcf/sam/bam/gff3/gff2/gtf/bed/ ...

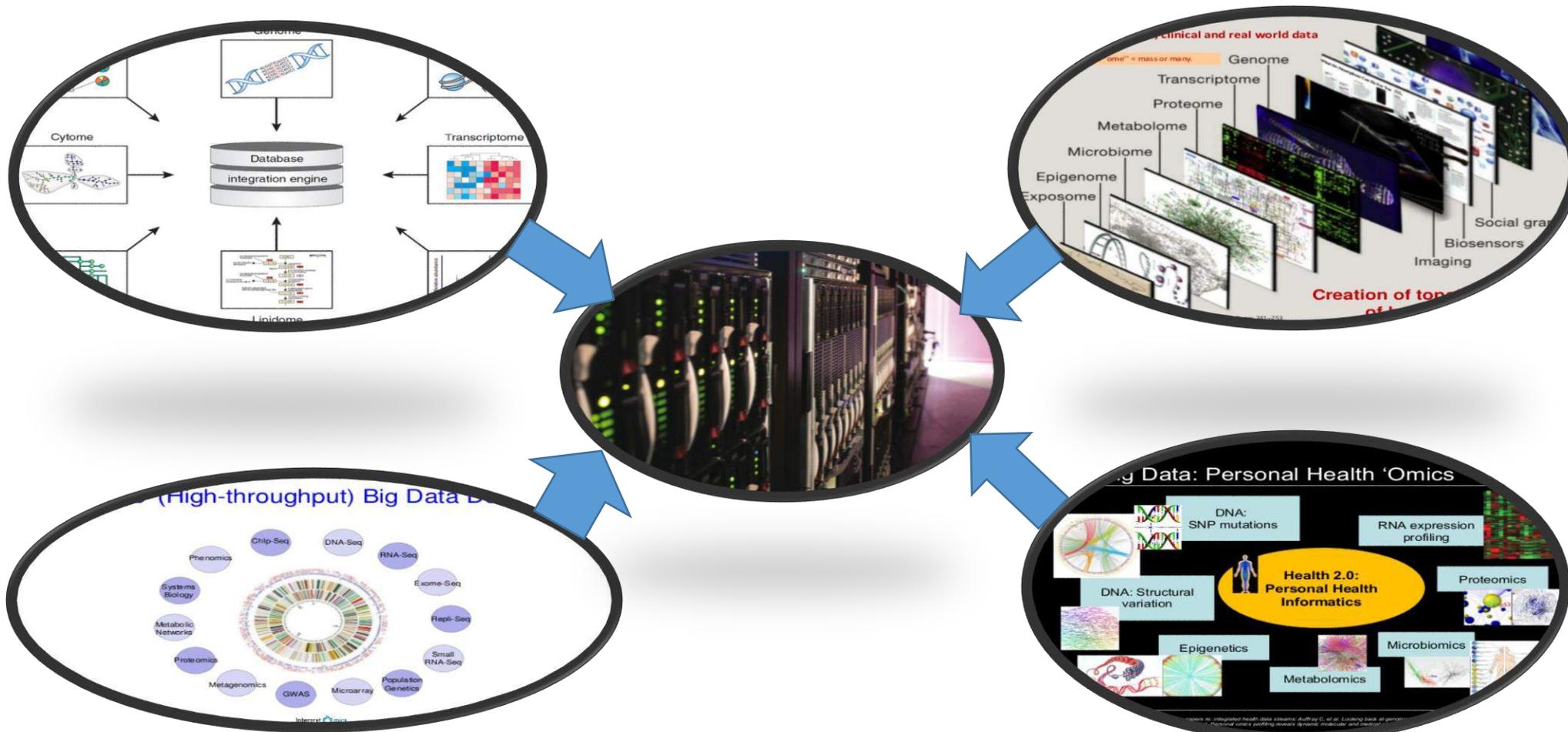


检测

- 荧光定量PCR
- 基因芯片
- SNP分型
- 二代测序
-



组学数据汇集



组学数据来源

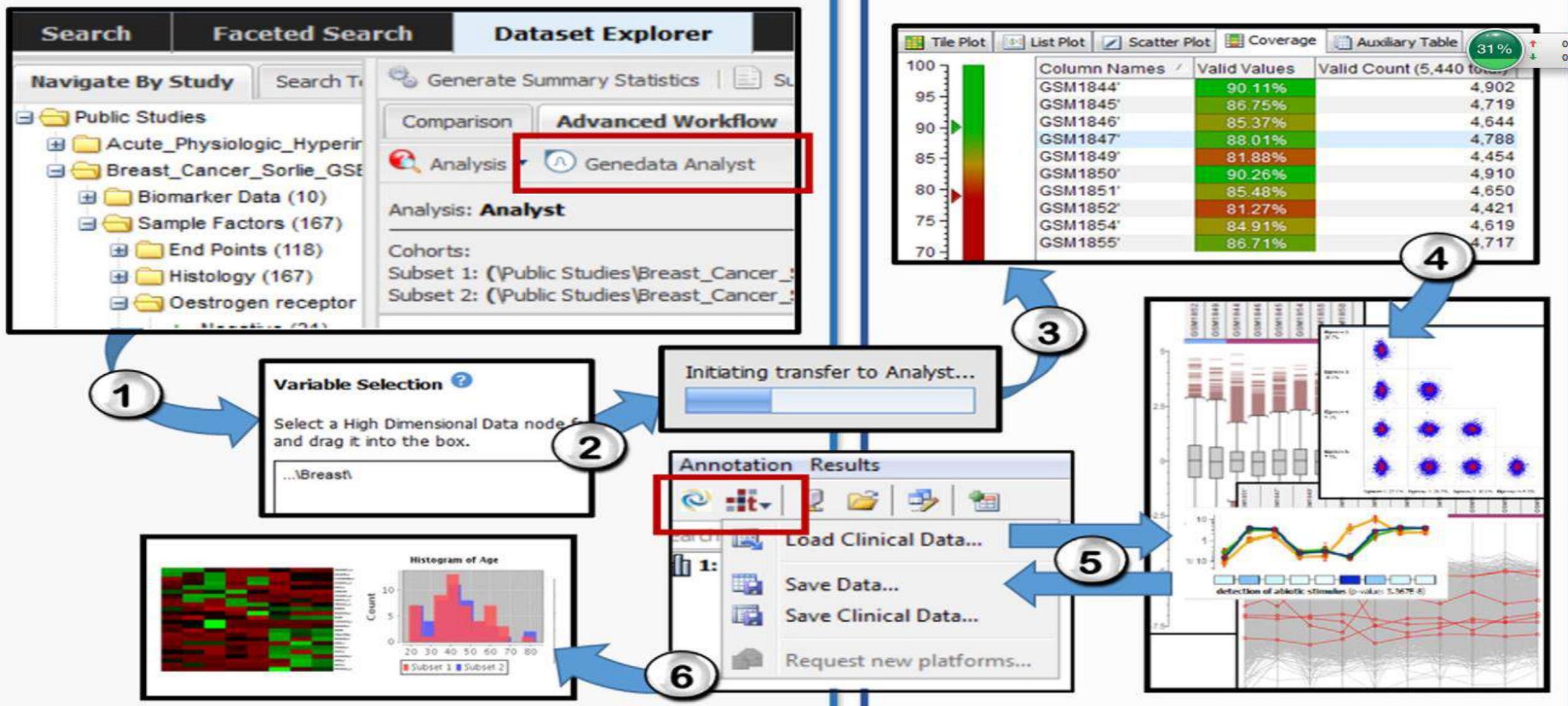
综合数据库

- NCBI美国国家生物技术信息中心
美国国家健康研究所、国家医学图书馆
- EBI欧洲生物信息研究所
欧洲分子生物学实验室
- DDBJ 日本DNA数据库
日本研究机构
- BIGD 生命与健康大数据中心
北京科学院北京基因研究所
-

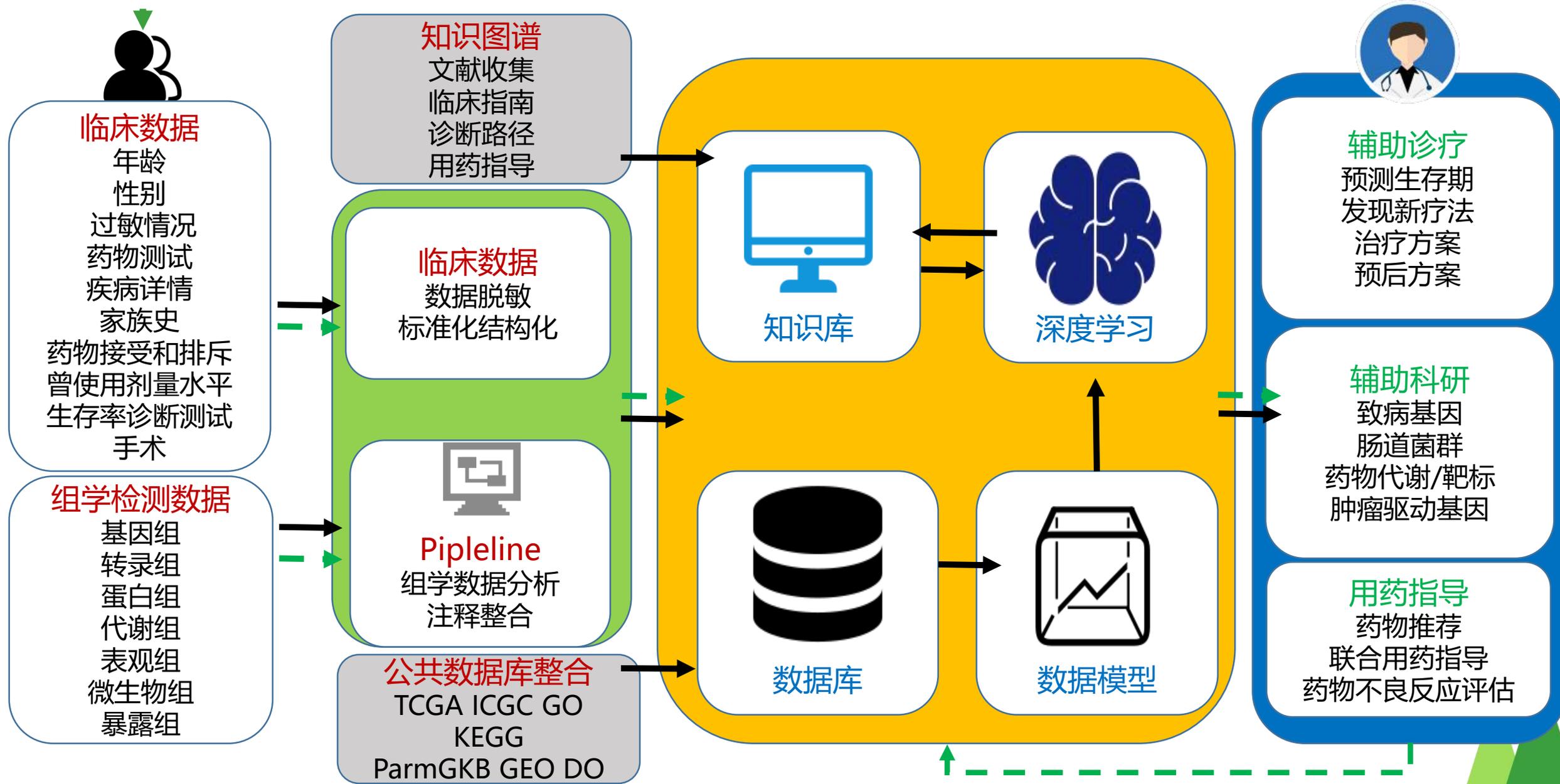
肿瘤相关数据库

- TCGA 癌症基因信息的数据库
- TGDB 肿瘤基因数据库
- ICGC 国际癌症基因组联盟
- Oncomine 肿瘤基因芯片数据库
- CGAP癌症基因数据库
- MethylCancer DNA甲基化与癌症数据库
-

组学数据分析处理流程



组学大数据平台在肿瘤精准医疗中的应用



辅助诊疗

根据类似患者信息，精确匹配**最佳诊疗方案**，快速诊断、准确用药、**提高诊疗水平**。



医生初步检查



辅助诊疗

最佳诊疗
方案推荐



辅助诊疗

- 预测生存期
- 发现新疗法
- 治疗方案
- 预后方案

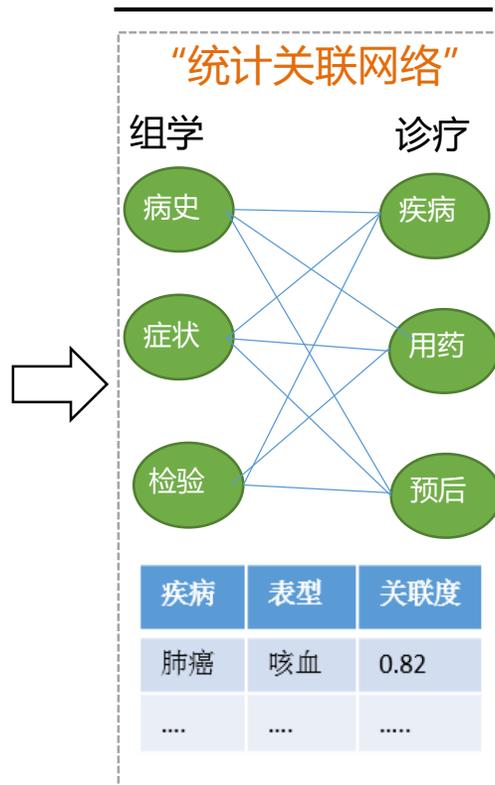
病历系统



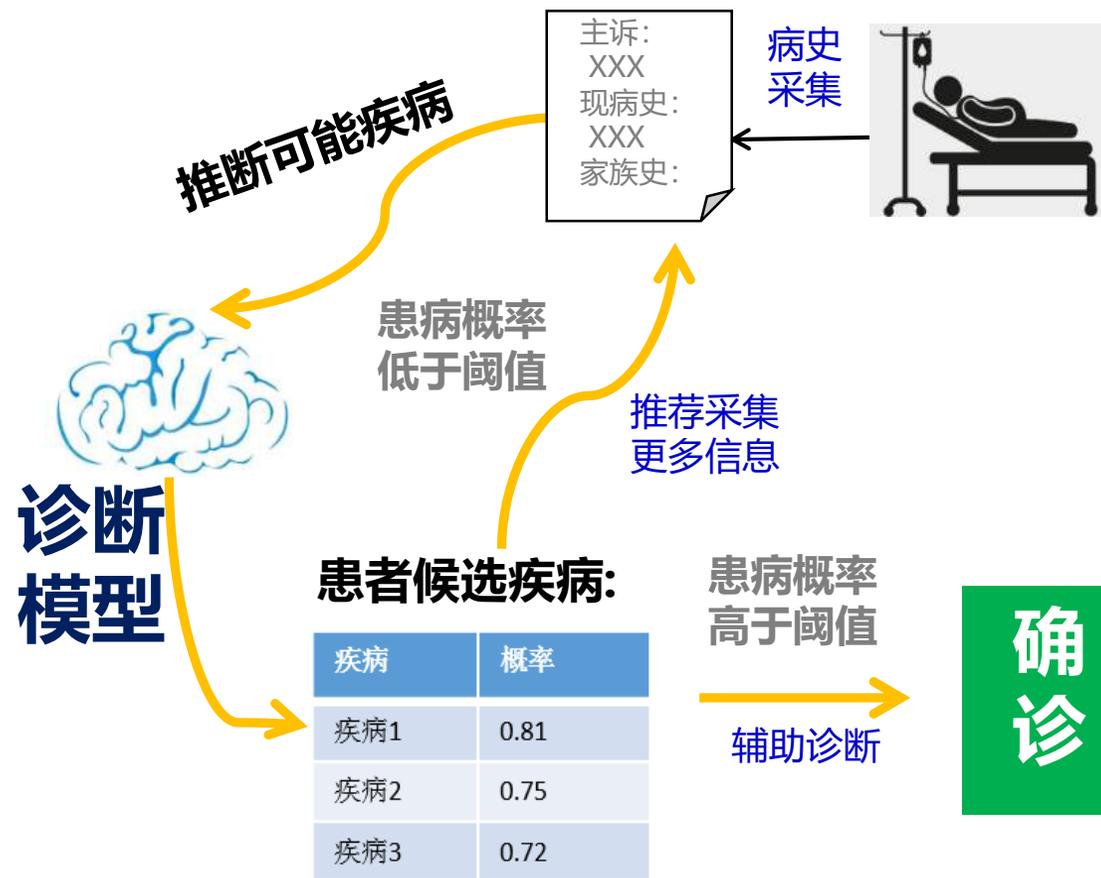
检测系统



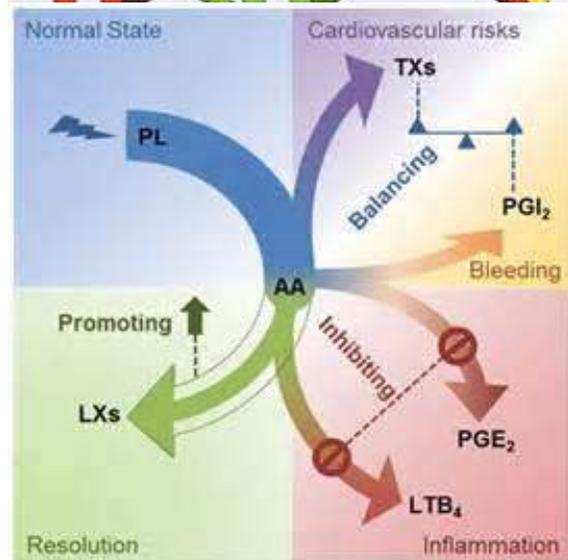
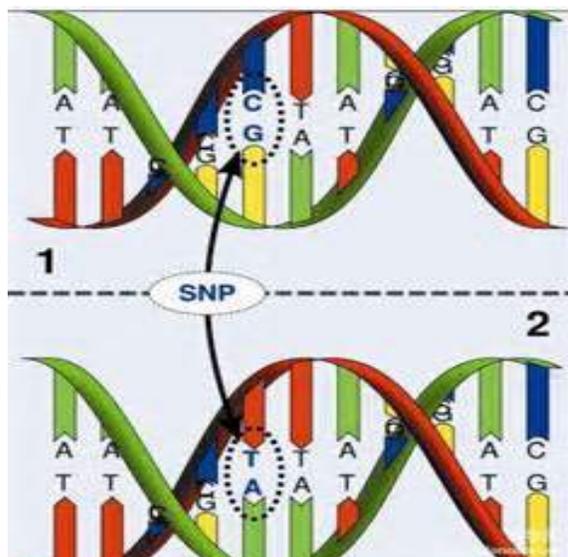
疾病知识图谱



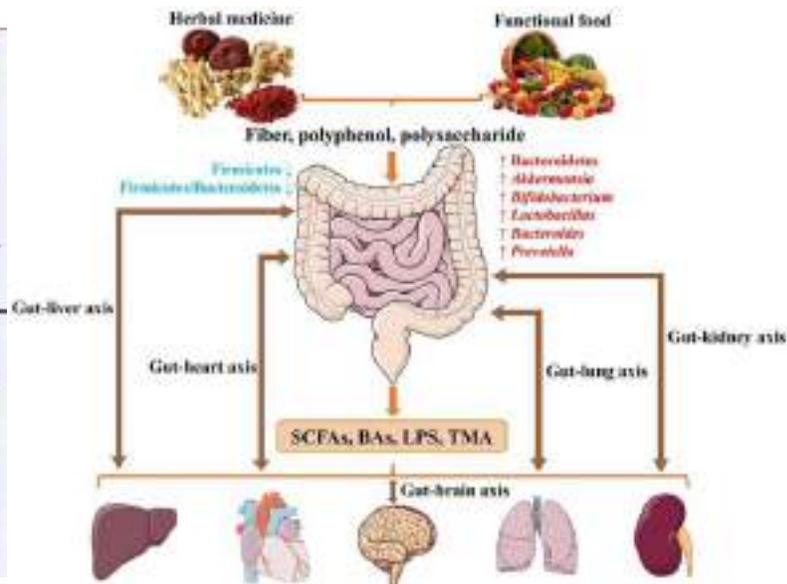
动态辅助诊断



- 致病基因
- 肠道菌群
- 药物代谢/靶标
- 肿瘤驱动基因



Accounts of Chemical Research 2015, 48, 2242



用药指导

- 药物推荐
- 联合用药指导
- 药物不良反应评估

