



医疗大数据与临床决策 在制药企业中的应用案例分享

谷成明 博士

辉瑞大中华区副总裁
医学部负责人

仅限学习交流使用



真实世界数据贯穿我们的一生

Exogenous data

1,100 Terabytes

Generated per lifetime

Influence on outcomes

60%

Genomics data

6 TB

Per lifetime

30%

Clinical data

0.4 TB

Per lifetime

10%



1ZB=1024EB=1024PB=1024TB=1024GB





企业对于数据的关注在发生变化





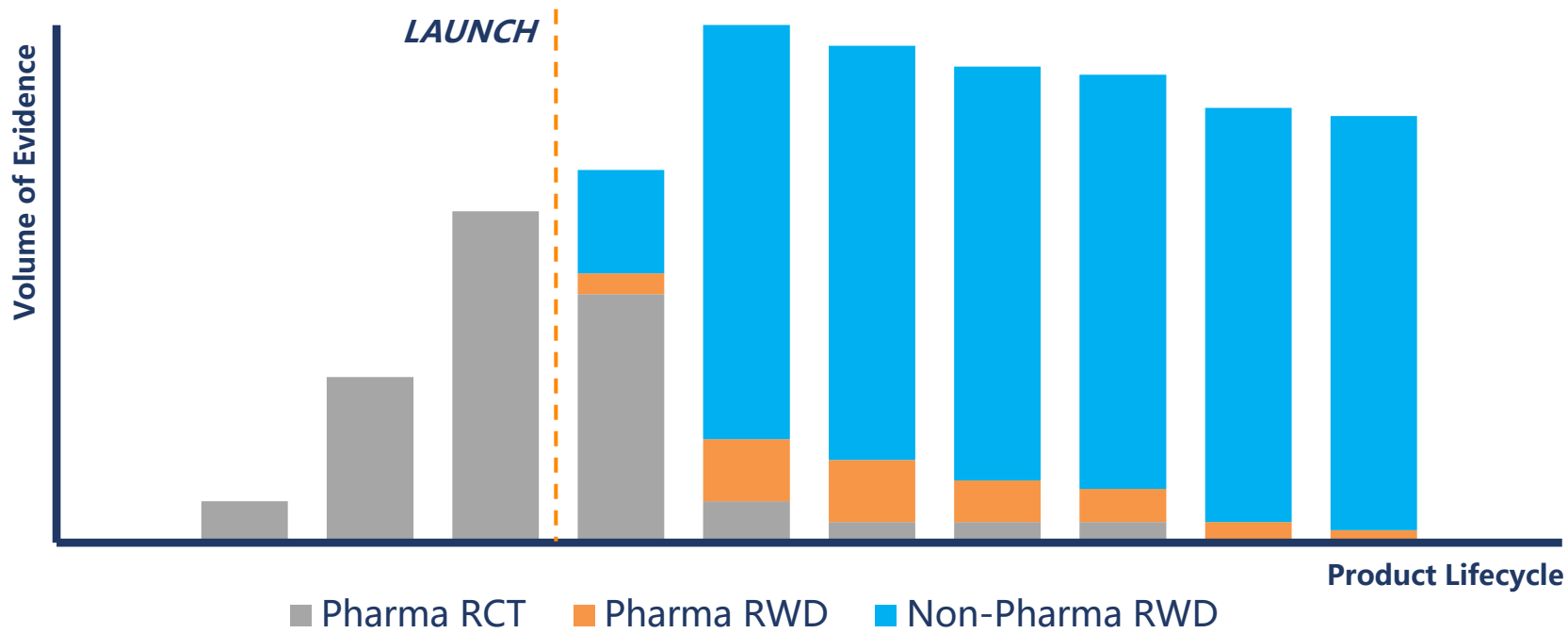
数据的应用趋势

	NEJM	Lancet	JAMA	BMJ
Clinical Trials	57	57	28	18
Observational study				
Prospective Cohort	6	0	6	8
Registry / EMR /Administrative database (including CMS)	10	7	16	20
Other design	9	9	4	8
Meta-analysis / Complex design	3	8	9	20

- 越来越多的基于登记注册、电子病历及管理保险数据的科研成果产生
- 混合型研究设计和多数据源的分析使用



制药企业研究的发展趋势





数据的应用趋势





大数据和人工智能在医疗上都能做哪些事儿？

解决方案

医院

1. 医疗服务质量管理
2. 精准医疗（个性化诊疗）
3. 临床科研平台

政府/保险

1. 政策分析
2. 合理用药
3. 精准医保

企业

1. 真实世界数据分析
2. 效率提升平台
3. 新业务模式合作

健康管理机构

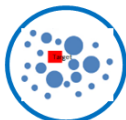
1. 风险筛查
2. 个性化照护
3. 认知决策支持



预测建模



风险因素发现



相似患者分析



风险分层



疗效分析



序列模式挖掘



药物有效性



依从性分析



决策支持



政策分析



精准医保

数据探索

队列构建

数据清洗

数据填补

特征构建

特征选择

数据
库构
建

EMR

登记研究数据

医保数据

个人行为数据

基因数据

知识
库构
建

临床指南

医学文献

针对特定临床结局发现新的风险因素/保护因素

AF患者出现缺血性卒中风险

发现27个新的潜在风险因素

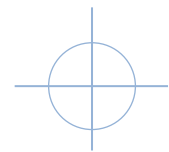
期刊论文拟投 JACC/Circulation

风险因素

- 既往出血性卒中
- 吸烟
- 心肌梗死 & 高脂血症
- 短暂性脑缺血发作 & 高脂血症

保护因素

- 地高辛&他汀
- 心律平&他汀
- 左心室射血分数
- PCI时间, BMI





机器学习下的预测模型优化

	All	P value
Age	1.060	<0.001
Killip > 1	1.681	<0.001
Anterior infarction	1.275	<0.001
HR	1.022	<0.001
SBP	0.982	<0.001
Female	1.622	<0.001
Survival Cardiac Arrest	3.246	<0.001

Grouping Rules Mining

```
Age <= 59.930 [模式: 0] (5.497)
├─ Killip <= 1.500 [模式: 0] (4.636)
│   └─ Female in [0.000] [模式: 0] => 0 (4,124; 0.735)
│   └─ Female in [1.000] [模式: 1] => 1 (512; 0.578)
├─ Killip > 1.500 [模式: 1] (861)
│   └─ cardArrest in [0.000] [模式: 1] => 1 (828; 0.743)
│   └─ cardArrest in [1.000] [模式: 2] => 2 (33; 0.879)
└─ Age > 59.930 [模式: 1] (7.553)
    └─ Killip <= 2.500 [模式: 1] (6.757)
        └─ Age <= 65.016 [模式: 1] (2.057)
            └─ Killip <= 1.500 [模式: 1] (1.691)
                └─ HR <= 71.500 [模式: 0] => 0 (711; 0.592)
                └─ HR > 71.500 [模式: 1] => 1 (980; 0.629)
            └─ Killip > 1.500 [模式: 1] => 1 (366; 0.842)
        └─ Age > 65.016 [模式: 1] (4.700)
            └─ HR <= 98.500 [模式: 1] => 1 (4,245; 0.786)
            └─ HR > 98.500 [模式: 1] (455)
                └─ Killip <= 1.500 [模式: 1] => 1 (259; 0.745)
                └─ Killip > 1.500 [模式: 2] => 2 (196; 0.556)
    └─ Killip > 2.500 [模式: 2] (798)
        └─ Killip <= 3.500 [模式: 1] (415)
            └─ HR <= 87.500 [模式: 1] => 1 (178; 0.702)
            └─ HR > 87.500 [模式: 2] (237)
                └─ SBP <= 133.500 [模式: 2] => 2 (135; 0.748)
                └─ SBP > 133.500 [模式: 1] => 1 (102; 0.529)
        └─ Killip > 3.500 [模式: 2] => 2 (381; 0.782)
```



无统计学意义

缩减3个因素以方便使用，但AUC无明显变化
0.826 → 0.817

- Weight：加入体重并无法增加AUC
- Cr：肌酐指标需要做实验室检查
- Time:从出现症状到医院的时间不便评估

THE LANCET

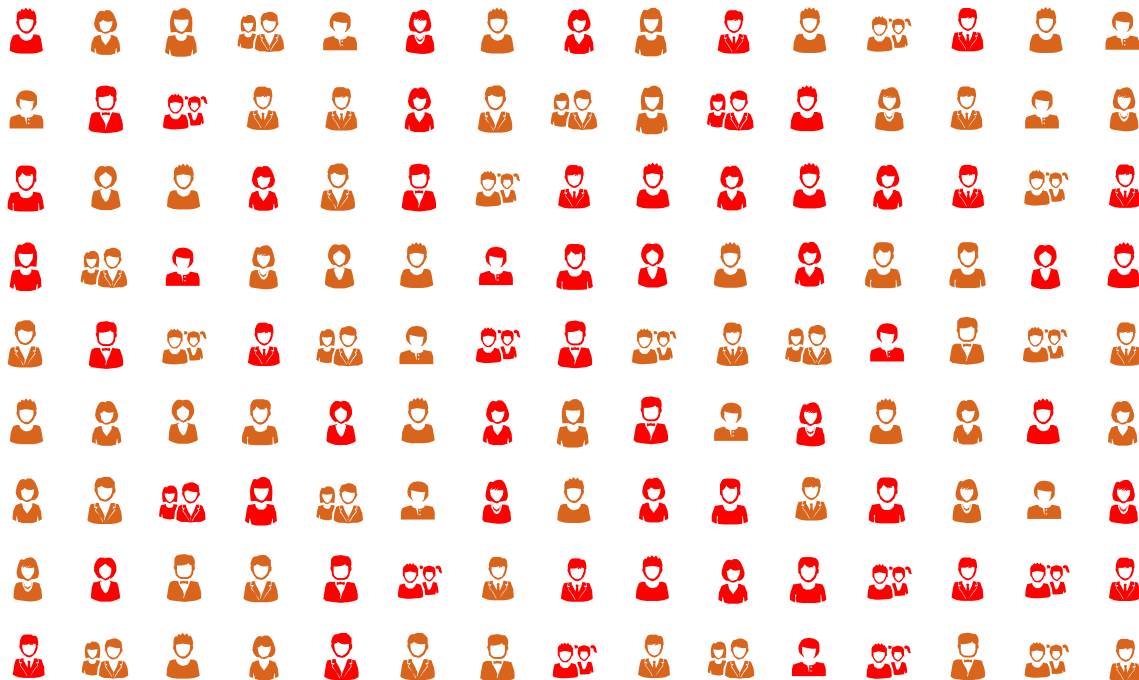
Pass Lancet first round editor committee review.

- New model AUC = 0.783 ,
similar to TIMI (0.717) scoring model.

相似患者分群



= 红色 单身 男性 领带



维度1 色彩 红

维度2 配偶 无

维度3 性别 男

维度4 领带 有

⋮

维度10

⋮

⋮

⋮

维度10,000

⋮

医生

机器

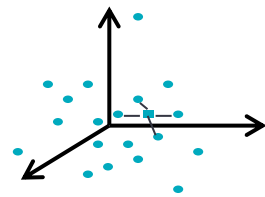
A circle is centered at the origin of a Cartesian coordinate system. The circle is drawn with a blue line and is centered at the intersection of the horizontal and vertical axes. The circle's radius is approximately equal to the distance from the origin to the top of the vertical axis.



10,000+分析维度



$$\begin{bmatrix} \mathbf{x}_1^Q \\ \mathbf{x}_2^Q \\ \vdots \\ \mathbf{x}_N^Q \end{bmatrix}$$

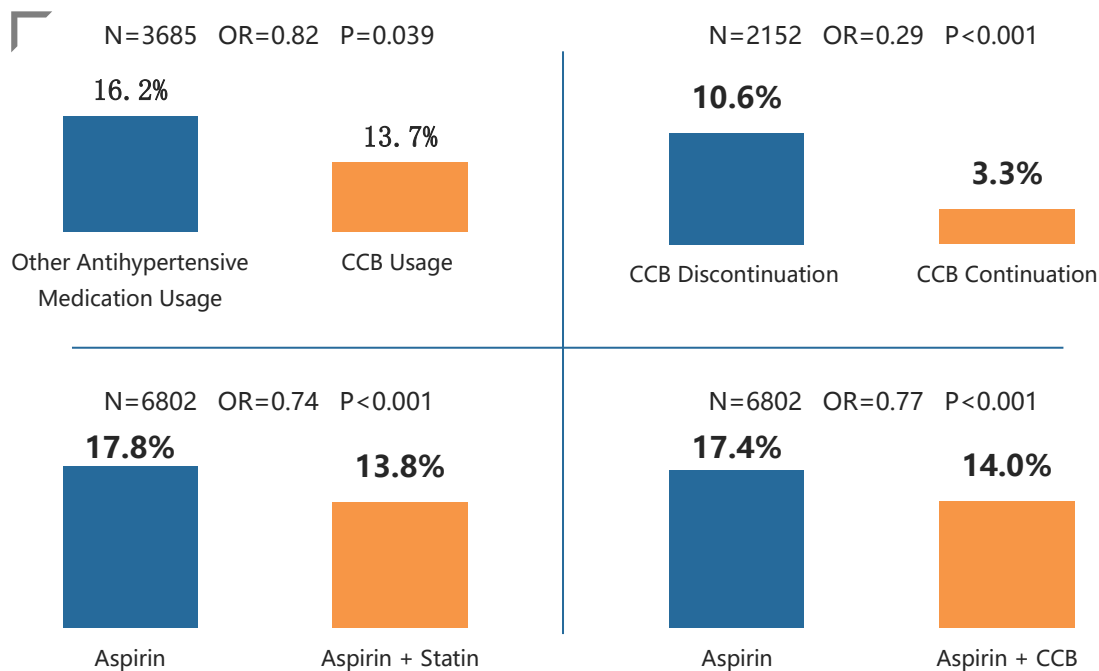


Patient similarity assessment in clinical factor/feature space

药物有效性分析

ACUTE ISCHEMIC STROKE PATIENTS IN Tiantan Hospital

Target to 1-year mortality risk for AIS patients



CCB疗效精准化分析

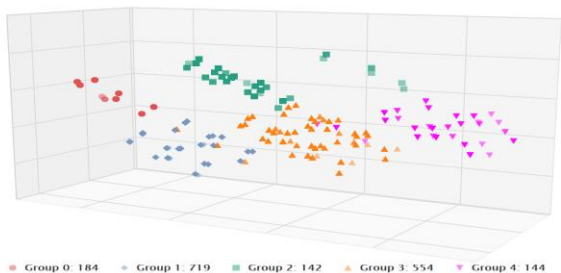
AF患者精准分群：特定人群使用CCB类药物降低患者风险达50%

背景

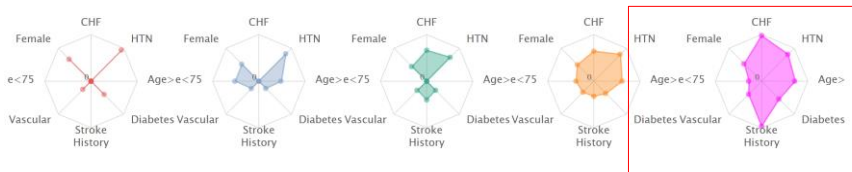
1703名房颤伴高血压的患者中，试图探寻使用CCB类药物疗效，但结果中对于全人群疗效上无差异

1.分群

自动学习精确的患者相似性特征，并根据特征将患者分为子群

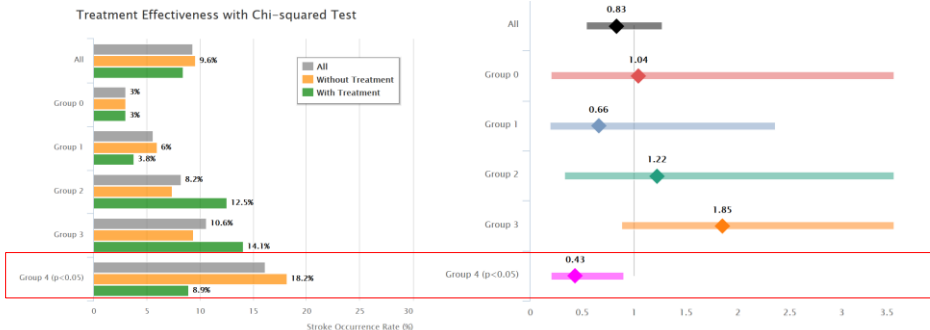


3.特征



2.结果

采取CCB治疗显著将发生卒中的风险降低50%

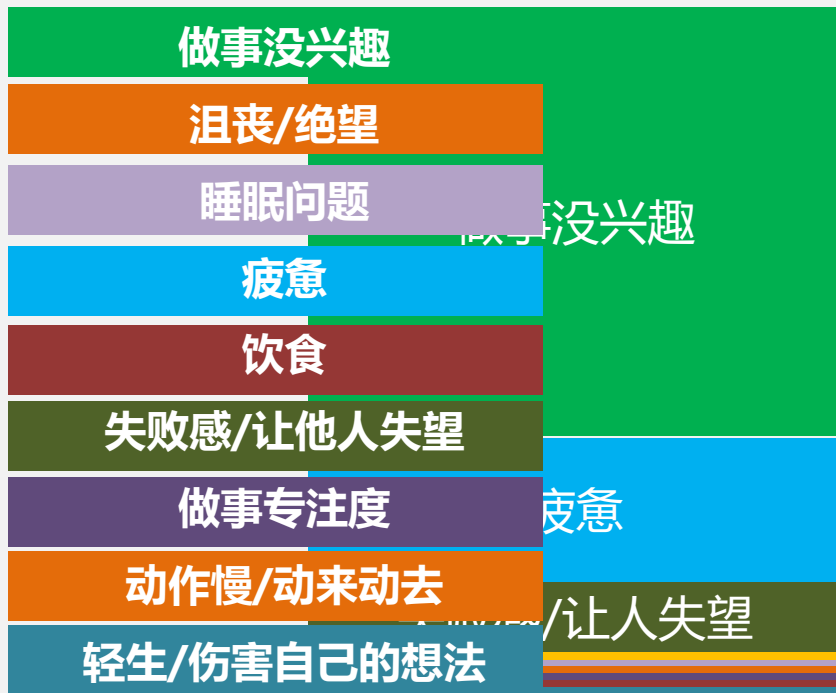


4.应用

临床 科研



诊断抑郁的关键问题



1:1:1:1:1:1:1:1:1

6:2:1:0.1:0.1:.....





什么才是自杀的关键问题

PHQ9

1 做事没兴趣

2 沮丧/绝望

3 睡眠问题

4 疲惫

5 饮食

6 失败感/让他人失望

7 做事专注度

8 动作慢/动来动去

9 轻生/伤害自己的想法

A 疼痛

B 眩晕

C 心跳

D 胸闷

E 肠胃不适

F 体重明显改变

G 疲劳

H 睡眠

I 性功能

J (女性)月经相关

K 其他





诊断自杀的关键问题

PHYSICIAN

轻生/伤害自己的想法

- 总人群中**33.95%**会轻生或伤害自己的想法

Category	%	n
0.000	66.048	82519
1.000	33.952	42419
Total	100.000	124938

- 特征人群：**失败感/让他人失望**
→ **60.41%**会轻生或伤害自己的想法

Category	%	n
0.000	78.305	66864
1.000	21.695	18525
Total	68.345	85389

问题 6
Adj. P-value=0.000, Chi-square=5672.524, df=1

失败感/让他人失望

Category	%	n
0.000	39.584	15655
1.000	60.416	23894
Total	31.655	39549

- 特征人群：**失败感/让他人失望 + 动作慢/动来动去**
→ **70.82%**会轻生或伤害自己的想法

Adj. P-value=0.000, Chi-square=2210.329, df=1

Adj. P-value=0.000, Chi-square=1169.020, df=1

做事没兴趣

动作慢/动来动去

Category	%	n
0.000	82.028	53259
1.000	17.972	11669
Total	51.968	64928

Category	%	n
0.000	66.492	13605
1.000	33.508	6856
Total	1.147	20461

Category	%	n
0.000	46.375	11099
1.000	53.625	12834
Total	1.266	23933

Category	%	n
0.000	29.175	4556
1.000	70.825	11060
Total	12.499	15616

- 特征人群：**无失败感/让他人失望 + 无做事没兴趣**
→ **82.02%**不会轻生或伤害自己的想法

Adj. P-value=0.000, Chi-square=952.553, df=1

Adj. P-value=0.000, Chi-square=223.232, df=1

Adj. P-value=0.000, Chi-square=327.520, df=1

Adj. P-value=0.000, Chi-square=332.182, df=1

疲劳

做事专注度

做事没兴趣

做事没兴趣

Category	%	n
0.000	84.738	41070
1.000	15.262	7397
Total	38.793	48467

Category	%	n
0.000	74.048	12189
1.000	25.952	4272
Total	13.175	16461

Category	%	n
0.000	69.334	10649
1.000	30.666	4710
Total	12.293	15359

Category	%	n
0.000	57.938	2956
1.000	42.062	2146
Total	4.084	5102

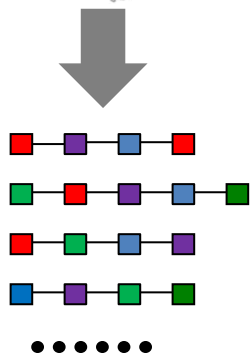
Category	%	n
0.000	54.313	4559
1.000	45.687	3835
Total	6.719	8394

Category	%	n
0.000	42.088	6540
1.000	57.912	8999
Total	12.437	15539

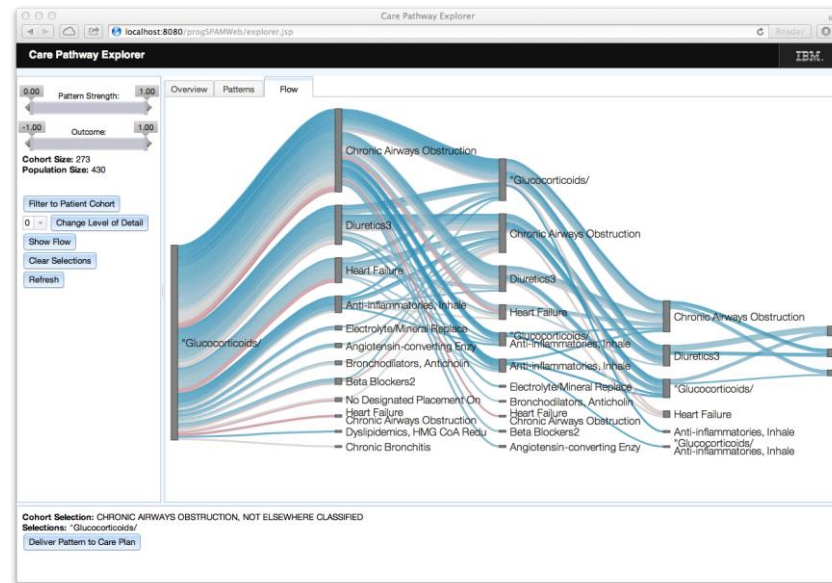
Category	%	n
0.000	44.892	1059
1.000	55.108	1300
Total	1.888	2359

Category	%	n
0.000	26.379	3497
1.000	73.621	9760
Total	10.611	13257

自动寻找不同患者群的最优个性化治疗路径



Temporal Event
Pattern Mining
Engine

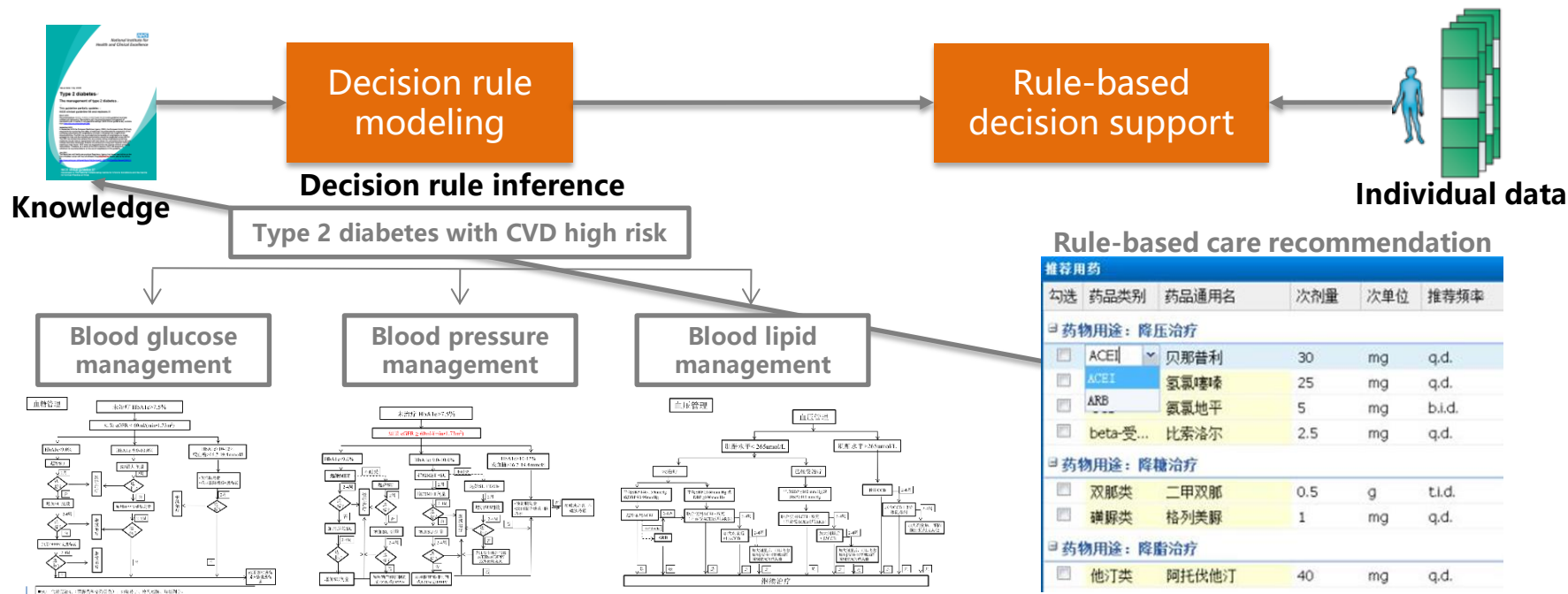


Goal	Care Pathway	Drug used	Drug recommended	Drug information	Data evidence*
Blood sugar management	Blood sugar badly controlled -> oral hypoglycemic drug combination	Metformin	Metformin + Sulfonylureas	Glucophage + Glimepiride	52% : 67%
Blood pressure management	Blood pressure abnormally high -> life style intervention	/	Life style intervention	Life style intervention	86% : 65%
Blood lipid management	Blood lipid abnormal -> Statin	/	Atorvastatin	Lipitor	72% : 87%

* : the first is the percentage for similar patients who used this treatment, and the second is the percentage for similar patients who used this treatment and got good outcome

知识驱动的临床决策支持系统

2型糖尿病知识库



数据驱动的临床决策支持系统

2型糖尿病用药推荐

系统

慢病管理

随访列表

就诊

评估

开处方

转诊

个人档案

挂号

家庭签约

王仕平 医生

今日随访任务：患者32人

姓名	计划日期	风险评估	风险水平	评估
叶尹 三里河三区30楼	已过期 22 天	心脏年龄	85	2015

处方药推荐

#	用途	类别	药品通用名	剂量	单位	频率	药品
1	降糖	双胍类	二甲双胍	0.5	g	t.i.d.	盐
		磺脲类	格列齐特缓释片	30	mg	t.i.d.	
2	降脂	他汀类	阿托伐他汀	40	mg	q.d.	阿
3	降压	ARB	厄贝沙坦	150	mg	q.d.	阿托
		双克	氢氯噻嗪	12.5	mg	q.d.	阿托
4	抗血小板	阿司匹林	阿司匹林肠溶片	75	mg	q.d.	阿托
		氯吡格雷	硫酸氢氯吡格雷	50	mg	q.d.	拜

打印处方

- 基于临床指南的糖尿病综合管理
- 基于大数据分析的疾病风险评估和用药推荐等
- 26万糖尿病患者，10年临床纪录

药物推荐：降脂

他汀类

贝特类

概要

相似患者

临床指南

医学文献

药物信息

对于该处方的建议：**推荐**

来自相似病患分析、临床指南、医学文献的证据

根据《2016欧美高危人群血脂管理指南》和《2013年中国糖尿病指南》，对此类病人需要使用**高强度他汀**治疗。

提到他汀治疗的文献数目：

2796篇

[Pubmed Link](#)

相关数据集中，与当前患者相似的患者数：

26%

[2329/8958人](#)

相似患者中，使用该处方的患者数：

48%

[1113/2329人](#)

使用该处方后，患者血脂达标率：

90%

[1002/1113人](#)

使用该处方后，患者依从性：

85%

[946/1113人](#)

智能化验单

传统看箭头

信息单一，指标固化，不易读懂

首都医科大学附属北京朝阳医院 检验科生化室 地址：门诊楼			
号：C0596130729		样品编号：	
		标本类型：	
测试结果			
单位	参考值		
总胆固醇 (CHOL)	5.08	mmol/L	3.62-5.70
低密度脂蛋白胆固醇 (LDL-C)	1.40	mmol/L	1.03-1.55
高密度脂蛋白胆固醇 (HDL-C)	2.94	mmol/L	1.81-3.36
甘油三酯 (TG)	1.91	mmol/L	0.56-2.26
载脂蛋白A1 (ApoA1)	1.24	g/L	1.00-1.70

智能显价值

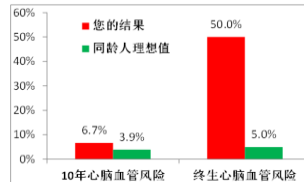
动脉粥样硬化性心脑血管疾病风险评估报告	
患者信息	姓名：张某某 性别：男 年龄：55岁 职业：工程师
检验信息	检验日期：2023-10-27 检验科室：检验科生化室 标本类型：空腹静脉血
动脉粥样硬化性心脑血管疾病风险评估结果	
1. 10年ASCVD风险智能评估结果	
您的10年ASCVD风险为6.7%，属于中危人群。同龄人理想值为3.9%。	
2. 个性化健康处方	
根据您的风险评估结果，建议您采取以下措施：1. 调整饮食结构，减少饱和脂肪和胆固醇的摄入。2. 增加运动量，每周至少进行150分钟中等强度有氧运动。3. 戒烟限酒，保持心理平衡。4. 定期监测血压、血糖、血脂等指标。5. 遵医嘱服用降脂药物，将LDL-C控制在1.8mmol/L以下。	

- 与HIS&LIS的无缝对接自动抓取
- 10年ASCVD风险智能评估，结果可视化
- 个性化健康处方，提供个性化健康建议

智能化风险评估

动脉粥样硬化性心脑血管疾病风险评估结果

- 您10年内患心脑血管疾病（冠心病、脑梗死等）的风险是**6.7%**（同龄人理想值是3.9%），为同龄人理想值的**1.7倍**
- 您的终生患心脑血管疾病风险是**50%**（同龄人理想值是5%），为同龄人理想值的**10倍**
- 您属于冠心病、脑卒中等动脉粥样硬化性疾病的**中危人群**



个性化检测标准值

患者建议

1. 您的心脑血管主要危险因素及理想目标值如下：

主要危险因素	您的结果	您的理想目标值
低密度脂蛋白胆固醇 LDL-C	4.0mmol/L	<3.4mmol/L
血压	160/90mmHg	<140/90mmHg
空腹血糖	5.0mmol/L	<3.9-5.6mmol/L
糖化血红蛋白 HbA1c	6%	<7%
吸烟状况	正在吸烟	戒烟
体重指数	26.1	18.5-23.9

个性化健康处方

2. 健康处方：

- 改善生活方式：合理饮食，控制脂肪摄入，适量运动，体重达标，戒烟限酒，心理平衡
- 血脂管理：根据该风险评估结果的危险程度，决定是否启动药物调脂治疗，将降低低密度脂蛋白胆固醇 LDL-C 水平作为首要干预靶点，首选他汀类调脂药物（例如阿托伐他汀）；血脂异常需要长期的治疗 and 生活方式干预，自行减量或停药会增加发生心肌梗死或脑卒中等风险；定期监测，管理好您的坏胆固醇 LDL-C，使其达到理想目标值
- 血压管理：定期监测血压，长期服用降压药物，防止血压波动，平稳控制清晨6点至10点的血压，减少心脑血管事件的发生
- 定期门诊随访，您的下次随访时间：_____年_____月_____日

检验医师：_____ 临床医师：_____ 日期：_____

根据《中国成人血脂异常防治指南（2016年修订版）》



分级诊疗政策对就医行为改变的分析

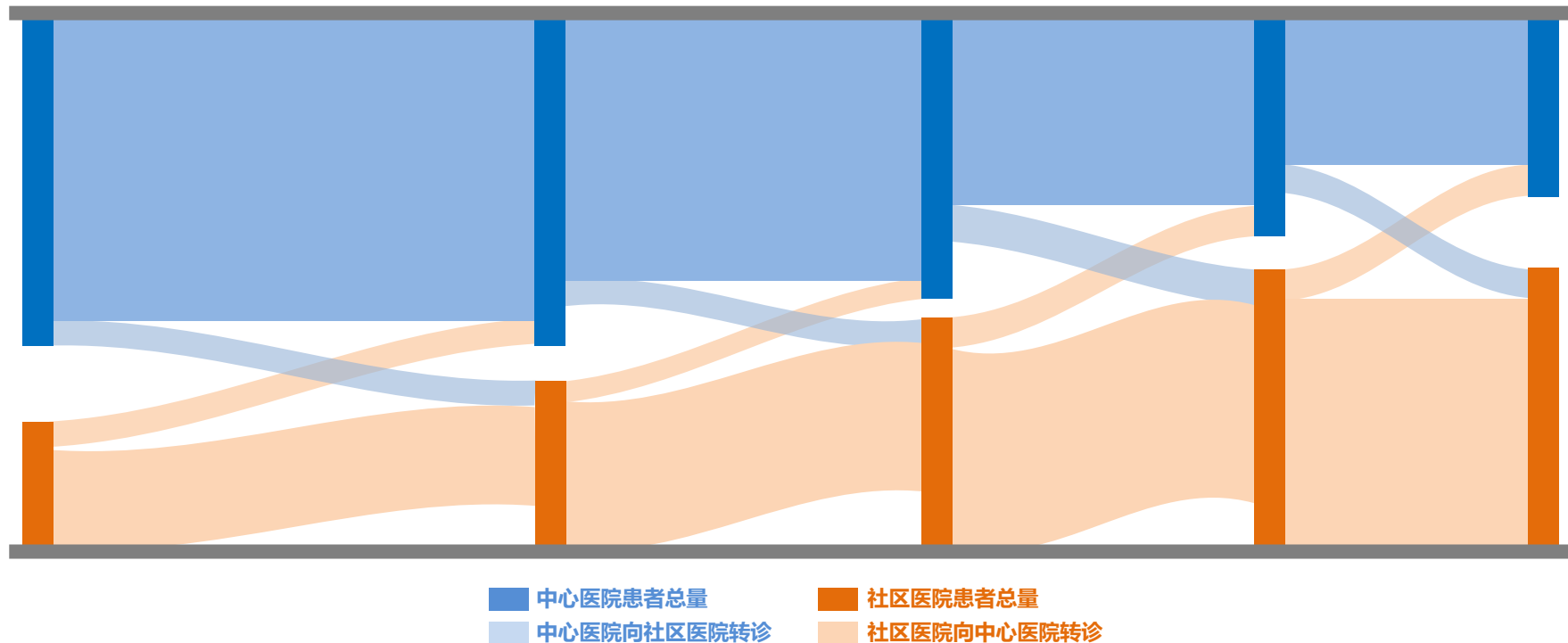
2012/08/01:
正式提出慢病一体化管理

2014/10/01:
提出慢病1+1+x概念

2015/04/01:
高血压、糖尿病普通
门诊下放社区

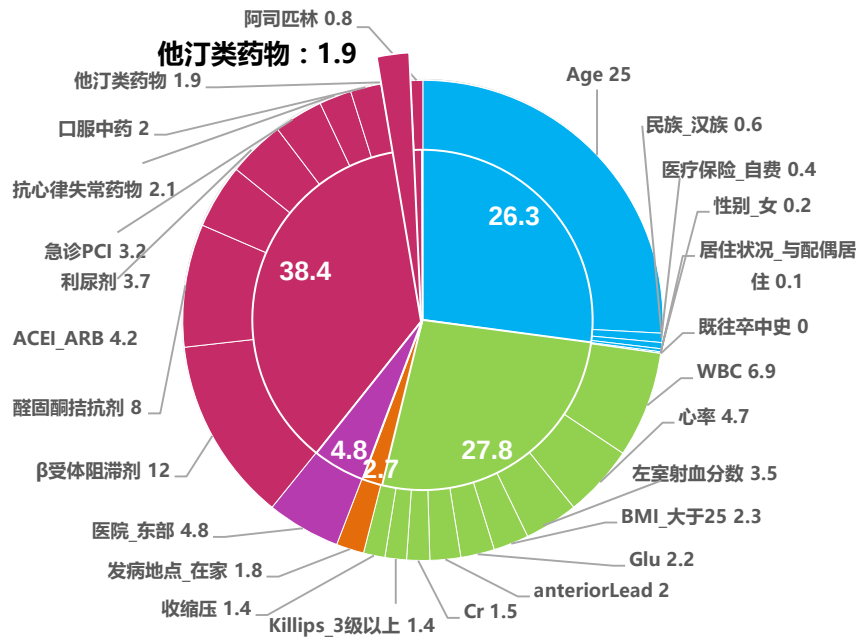
2015/07/01:
基层医疗机构设置
健康管理师服务

2015/11/01:
政策实施至今



医院管理指标贡献度比较

省级医院



Sociodemographic
社会人口

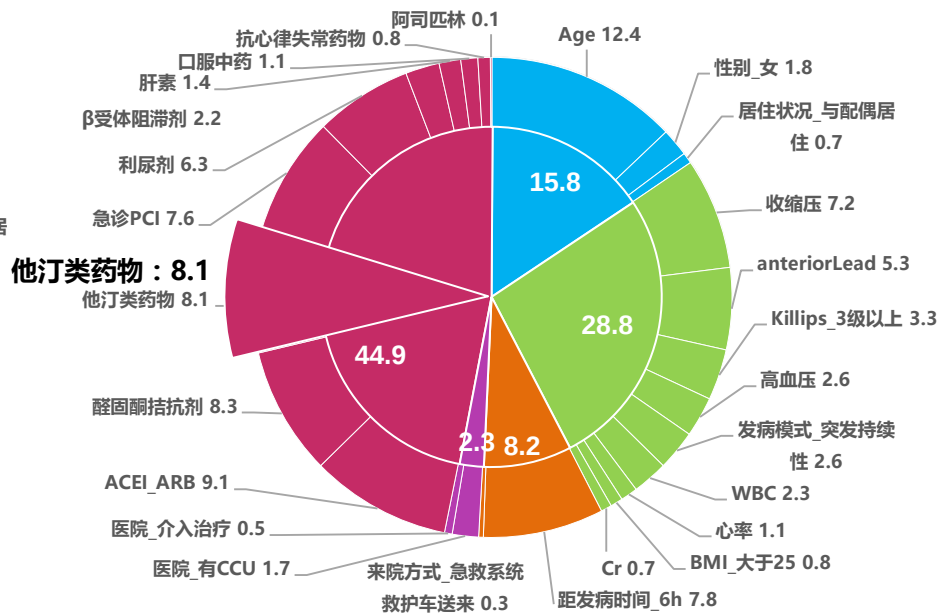
Clinical Condition
临床情况

Admission Pathway
院前情况

Hospital
医院

Treatment
治疗

县级医院



Sociodemographic
社会人口

Clinical Condition
临床情况

Admission Pathway
院前情况

Hospital
医院

Treatment
治疗



心脑血管疾病患者合理用药分析

——基于2005-2014年福建省直单位参保人群的数据分析

陈纯¹ 张勇² 黄绍中³ 谢小平⁴

(¹福建医科大学药学院 福州 350122; ²福建省医疗保险管理中心 福州 350001; ³福建省社会医疗保险协会 福州 350001; ⁴辉瑞投资有限公司医学部 北京 100000)

【摘要】目的：调查分析2005-2014年期间福建省直单位参保人群的心脑血管疾病(cardiovascular disease, CVD)的合理用药情况，探讨可行的控费策略。方法：(1)从福建省医疗保险管理中心就医数据库中筛选2005-2014年期间符合高胆固醇症诊断的人群，提取这些个体的门诊、住院用药记录，分析患者的他汀类用药情况；(2)筛选符合高血压、糖尿病、血脂异常、心肌梗塞或卒中诊断的个体，提取这些个体的所有门诊和住院记录(合计70,052名患者的15,468,958人次门诊记录，127,304人次住院记录)，导入



THANKS