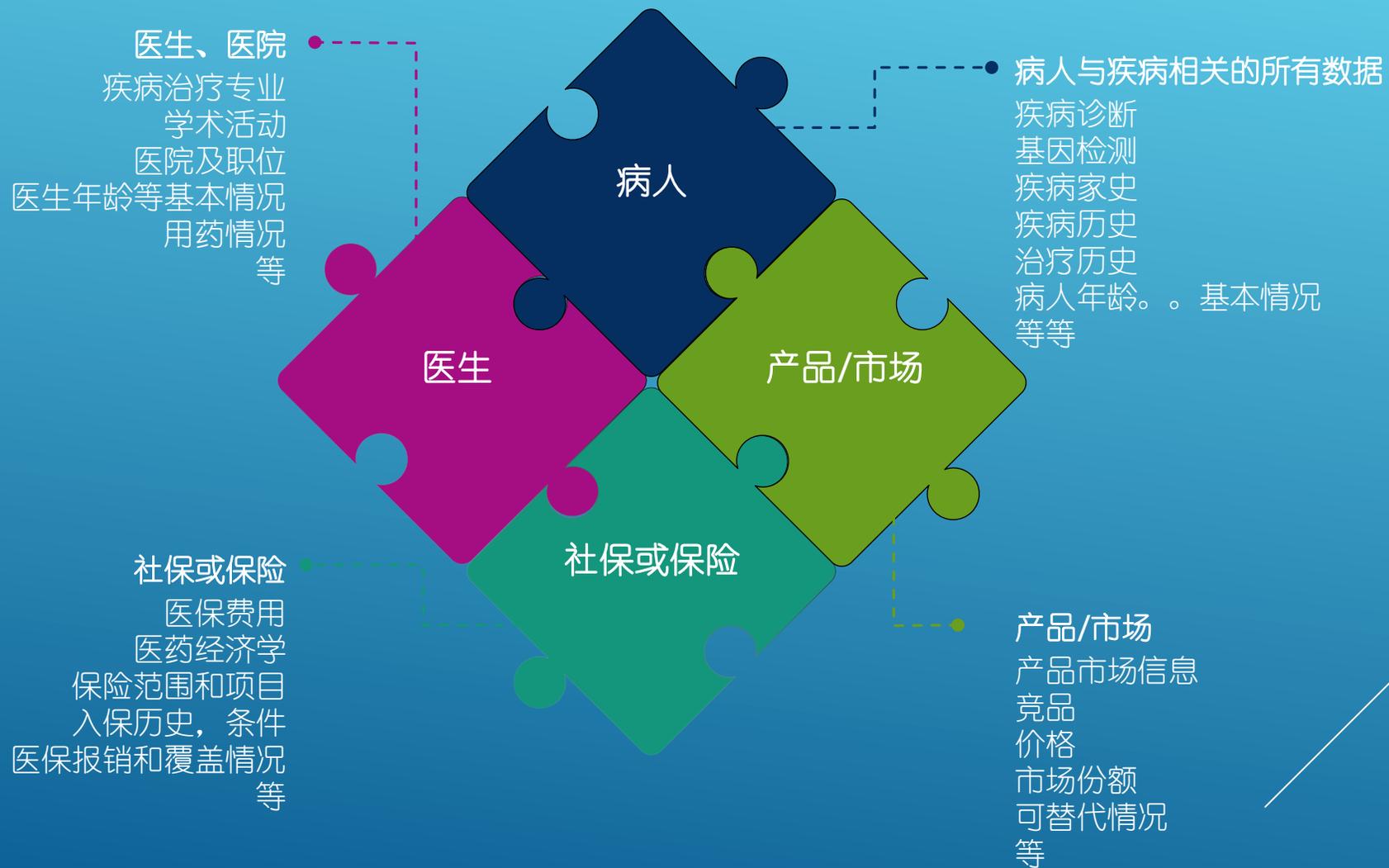


医药大数据的前生今世和未来

严浩瑄

百时美施贵宝, BI&A Oncology Lead

医药大数据需要什么样的数据内容



医药大数据的前生 - 有什么数据都行

分散

- 单个产品销售数据
- 少数医院
- 缺乏病人诊断

样本

- 覆盖面窄
- 小数据样本
- 需要慢慢积累，但旧数据容易过时

无结构

- 数据之间没有连接键
- 内部数据和外部数据缺乏衔接

医药大数据的今世 - 数据慢慢多起来了但依旧价值有限

内部数据

- 可靠性，历史数据堆积
- 投入多，数据输入真实可靠性不够
- 与外部数据缺乏连接

外部数据

- 数据覆盖非常有限
- 各自来源，数据结构，ID编码，等都不一样
- 相互之间有不一致地方，很难判断取舍

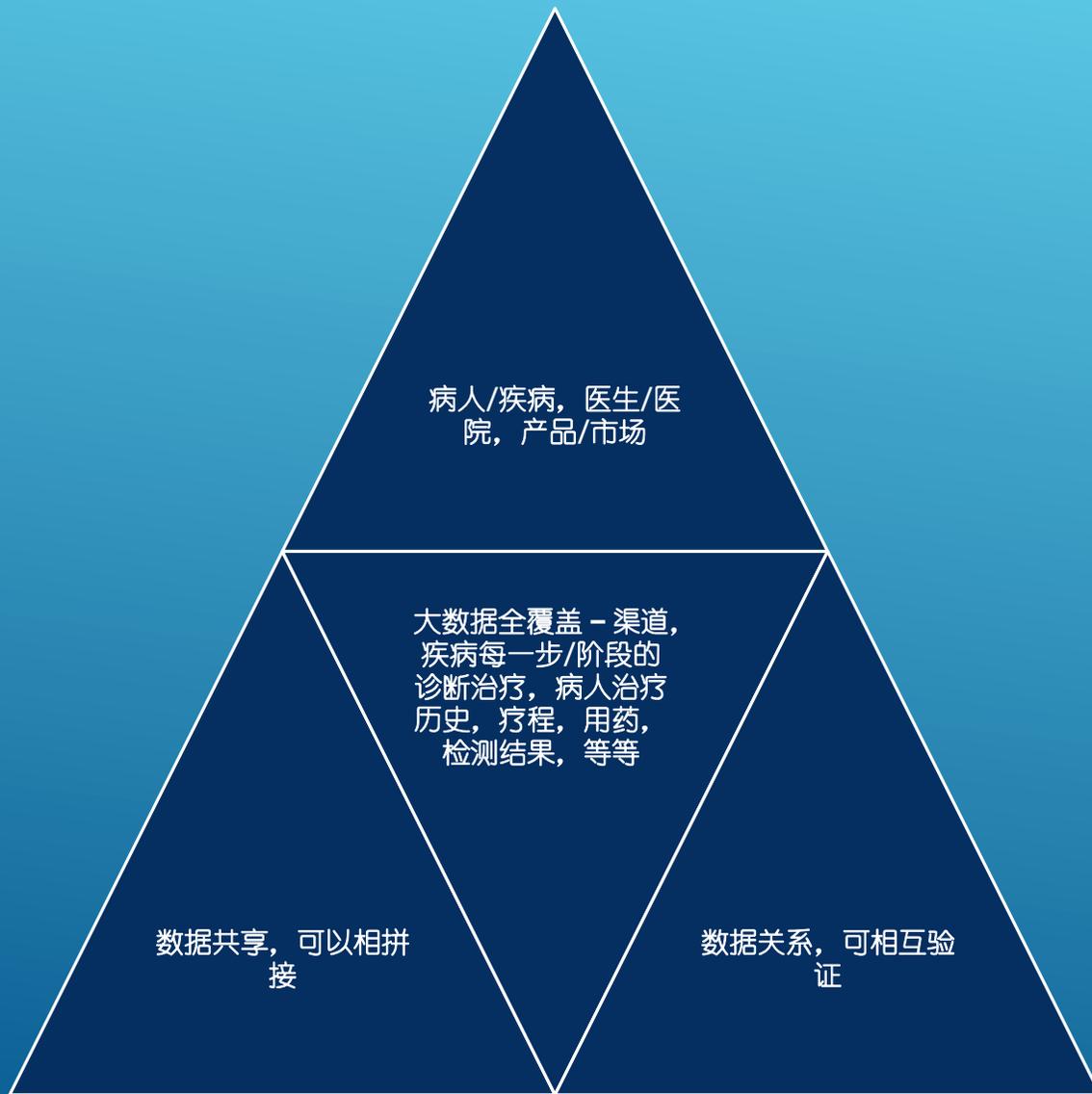
及时性和通用性

- 数据来源往往有些不稳定
- 及时性更新比较有挑战，有时候缺乏一致性
- 很难通用，比如和其他数据联合

医药大数据的未来 - 大数据首先需要大大大和全全全

- **大数据的全数据模式：样本=全部；**
- 大数据不用随机分析法这样的捷径，而采用所有数据的方法；
- 这里的“大”是相对的——相扑比赛所有数据存储还不需
要一个TB，但是**是所有的数据！**
- 在大数据时代采用随机采样法，就像在汽车时代骑马一样，虽然特定情况下仍可采样随机采样法，但是慢慢地我们会放弃它

医药大数据的未来 - 大数据作为资产，需要有价值数据内容，全覆盖，并可以共享和相通



医药大数据肿瘤领域的需求

病人	时间日期	诊断	治疗用药	生物基因检测	治疗进展/几线	门诊或住院
医院/医生	病历	治疗专业	费用	医保	治疗理念	产品评价
产品	上市日期	适应症	剂量	价格	慈善	治疗效果/临床数据
市场/竞品	用药量	病人量	市场份额	趋势	分瘤种分治疗线	上市时间
社保/保险	报销情况	病人支付和报销比例	采购价	适应症	竞品准入	其他规定

肿瘤医药大数据要能够帮助回答制药公司市场和商业决策核心问题，比如

1. 产品份额，病人份额，趋势变化，竞品表现，分瘤种分治疗线
2. 分瘤种和治疗线的产品使用/病人来源
3. 治疗流程与治疗标准，用药量，疗程，慈善，经济情况与用药决定
4. 治疗效果进程，用药选择决定，联用 或 单用及其比例和选择标准

大数据的未来需要做到什么？

大数据带来的机遇和挑战

政策支持和医药行业对大数据认识

隐私条例的制定，管理和严格遵守

企业大数据环境，人才和建设

数据的价值，真实性和整合力

增加企业竞争力和创利润能力

Agricultural

一万年

Industry

十八世纪中叶

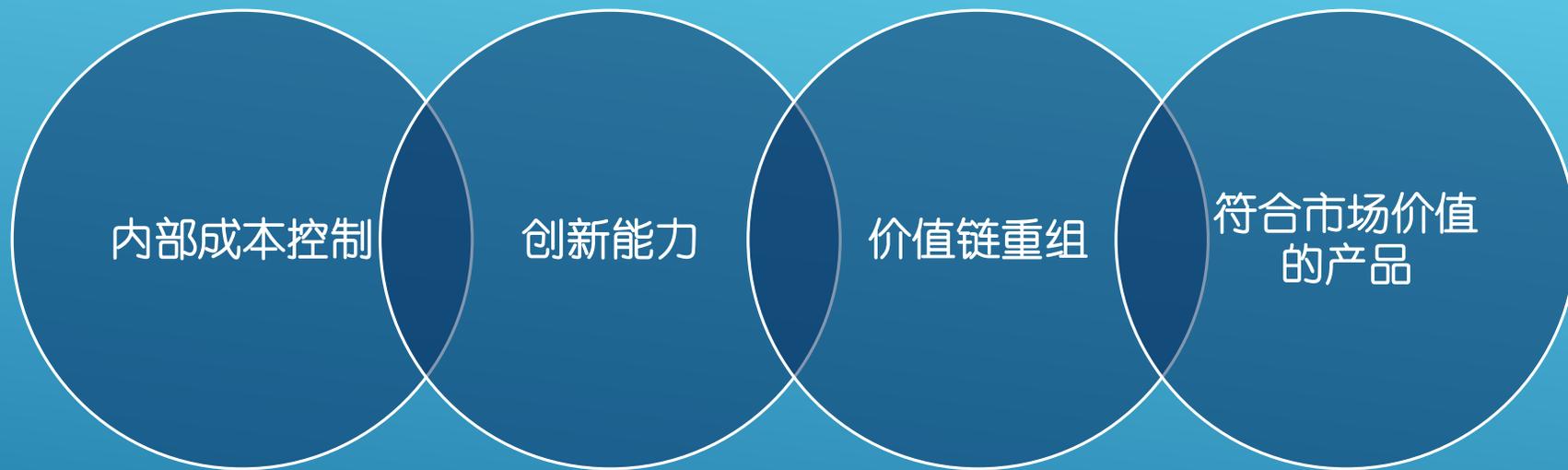
Information

二十世纪中后期

Data

2013年

制药公司的核心竞争力由什么决定？



公司如何应用资金投入，达到有效的结果和目标，实现投资最大化

大数据资产是公司保证实现上述核心价值实现的必须的竞争力之一