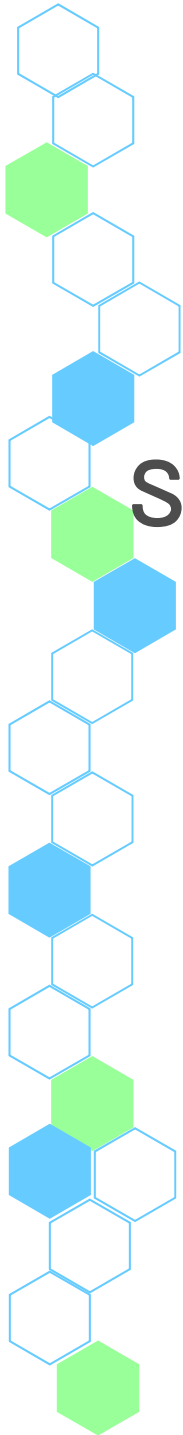




もしも、
Spotfire とVisual AnalyticsでOpenData を
眺めてみたら
(公開用/簡略版)

塩野義製薬
北西 由武

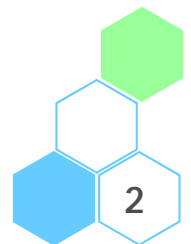
SASユーザー総会2016



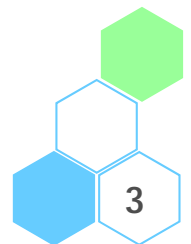
Contents



1. BIツールのポジショニング
2. オープンデータ × 可視化と仮説
3. 医療ビッグデータ(Real World Data) × 可視化と予測
4. まとめ

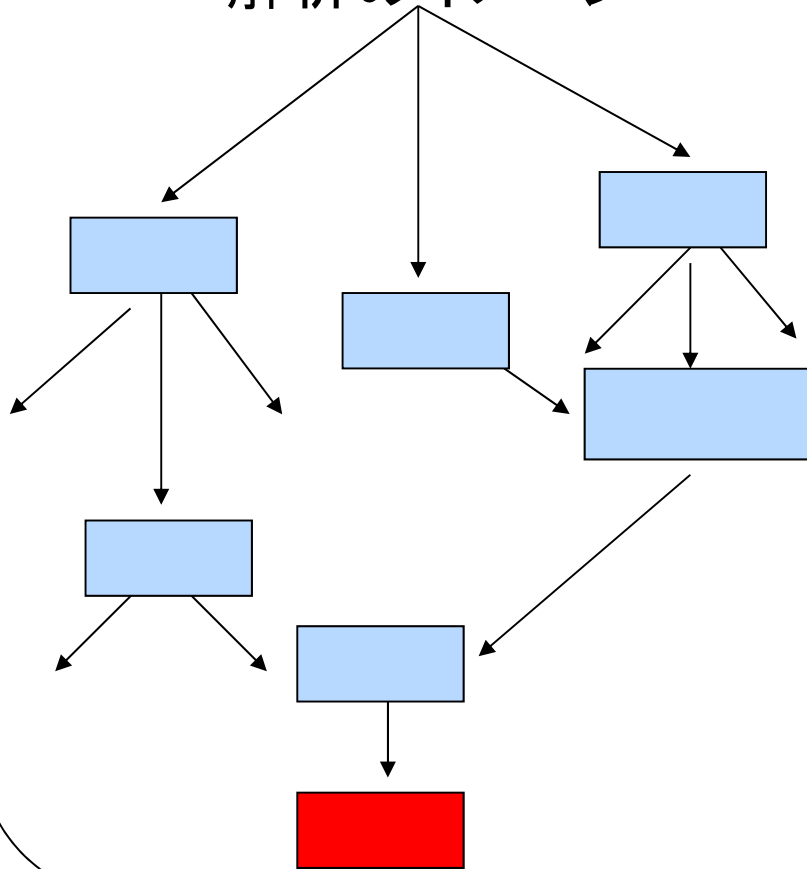


BIツールのポジショニング

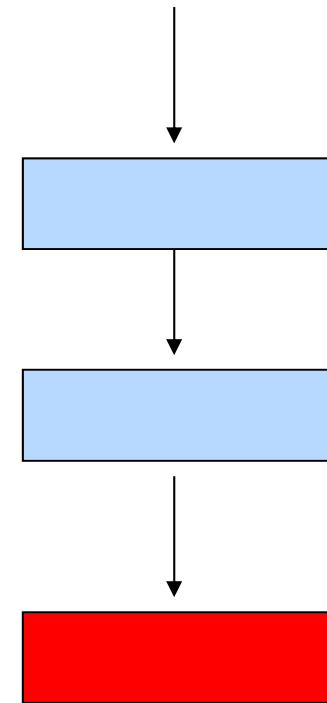


BIツールを使う利点

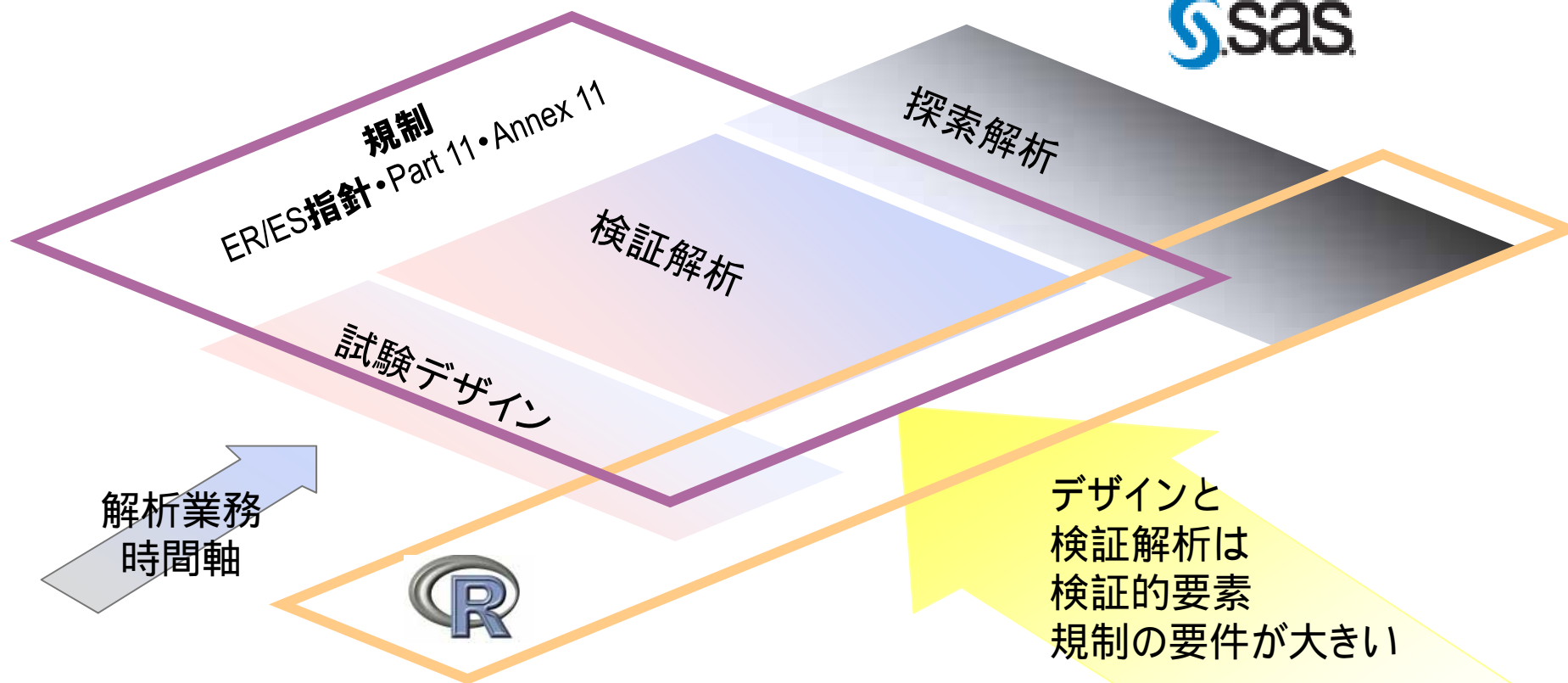
BIツールを使っての
解析のイメージ



オフラインプログラミングでの
解析のイメージ

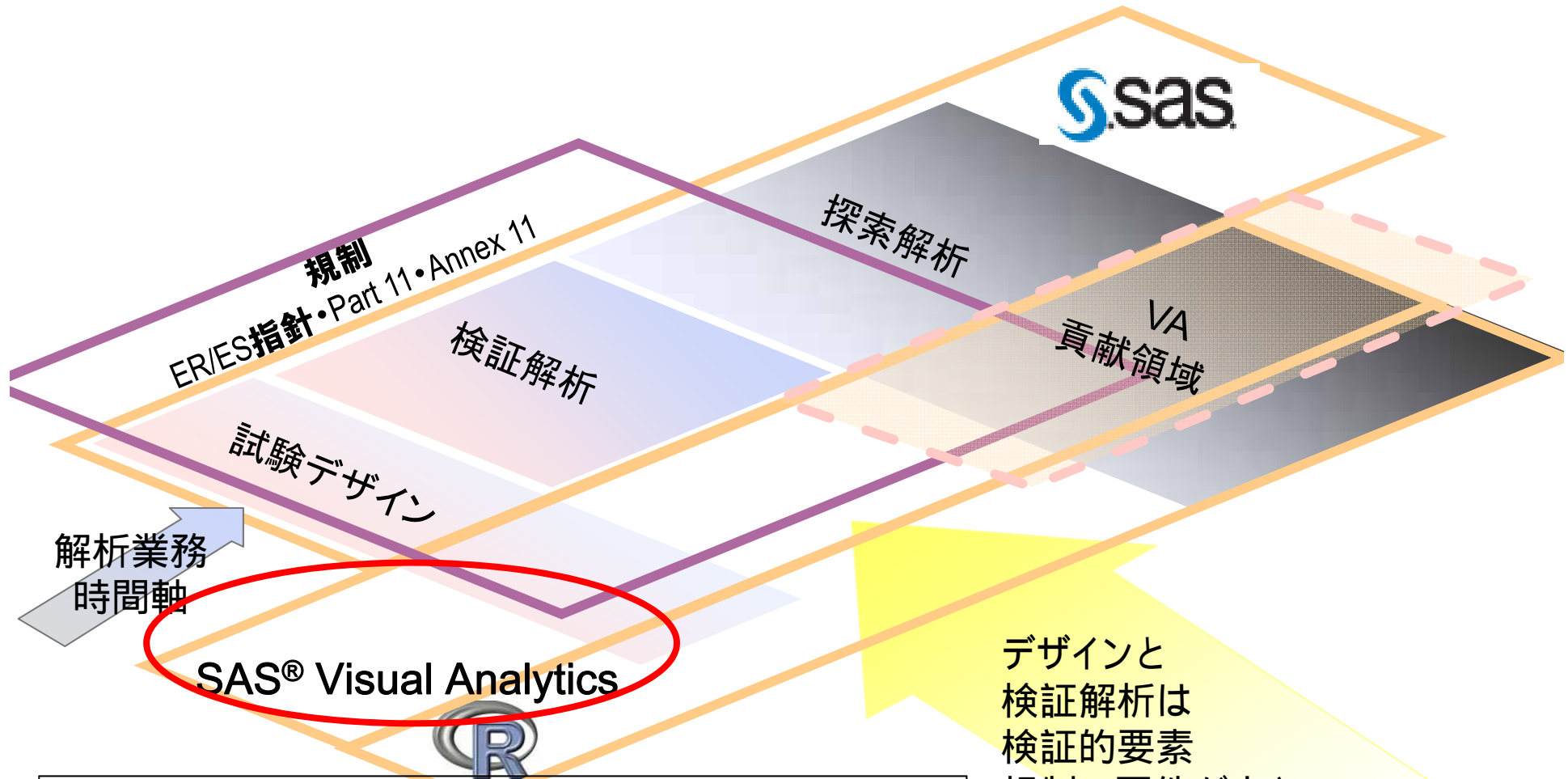


BIツール導入以前(解析コア業務)



- 規制に絡んだ解析業務が多い。
- グラフ化はSASかRを用いる。プログラミングに結構手間と時間がかかる。
- Rの利用者はごく一部。
- Excelでの解析は個人のスキルに依存する。

BIツール導入以降(解析コア業務)



デザインと
検証解析は
検証的要素
規制の要件が大きい

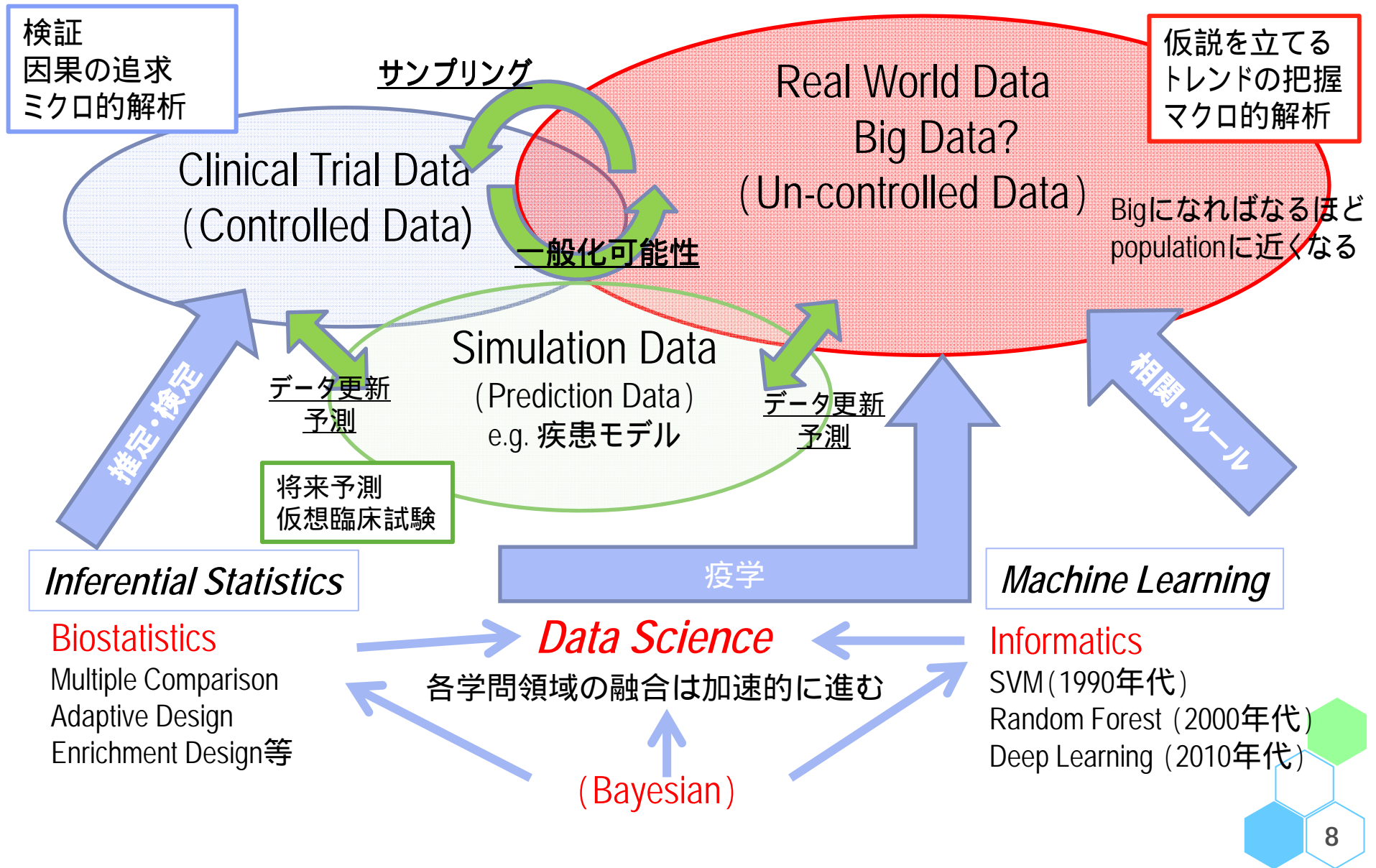
- 発想を伴った規制外の探索解析が増えた。
- 解析人口の拡大(解析部門以外もSpotfireで解析)
- 探索解析の効率化,リソース投入が可能になった。
- 情報のシェアが容易になった。

BIツールの開発業務での利用場面



利用場面	こんなときに	こんな活躍
モニタリング	<ul style="list-style-type: none"> ・安全性の監視(外れ値が一目瞭然) ・進捗確認(EDCの補助) 	<ul style="list-style-type: none"> ・臨床検査値を推移図やeDISHなどで視覚化, 有害事象の発現頻度の確認 ・来院の推移図など
症例検討会	医師とともに問題症例の検討	気になった症例の情報を即時に表示できる
解析作業時	解析資料の作成時にバリデーションの補助ツールとして	プログラムなしでクロス集計ができるので, 周辺度数の確認に使える
探索解析	より効果の出やすい集団の特定など	集団を絞って関連情報を確認(効果の高い症例をマークして背景情報から共通項を探すなど)
会議	<ul style="list-style-type: none"> ・ゴールが見えないときに ・視覚で訴えたいときに 	<ul style="list-style-type: none"> ・あーでもない, こーでもないを即実践 ・キレイな図が簡単に
タスクリソース管理	<ul style="list-style-type: none"> ・FTE換算と仕事の波の視覚化 ・予実確認 	・タスクの波が可視化され, 一目瞭然なので, リソース不足に陥る前に手を打てる.
予算管理	・損益分岐点を可視化したいときに	・図で示したほうが明確化され, 切迫感がでる.
解析資料	・SAS作成グラフの代用	・凝った詳細なグラフはSpotfireで置き換えて, 必要と判断すれば, SASで作り直す.
申請資料 (計画解析)	...	今のところは使っていない

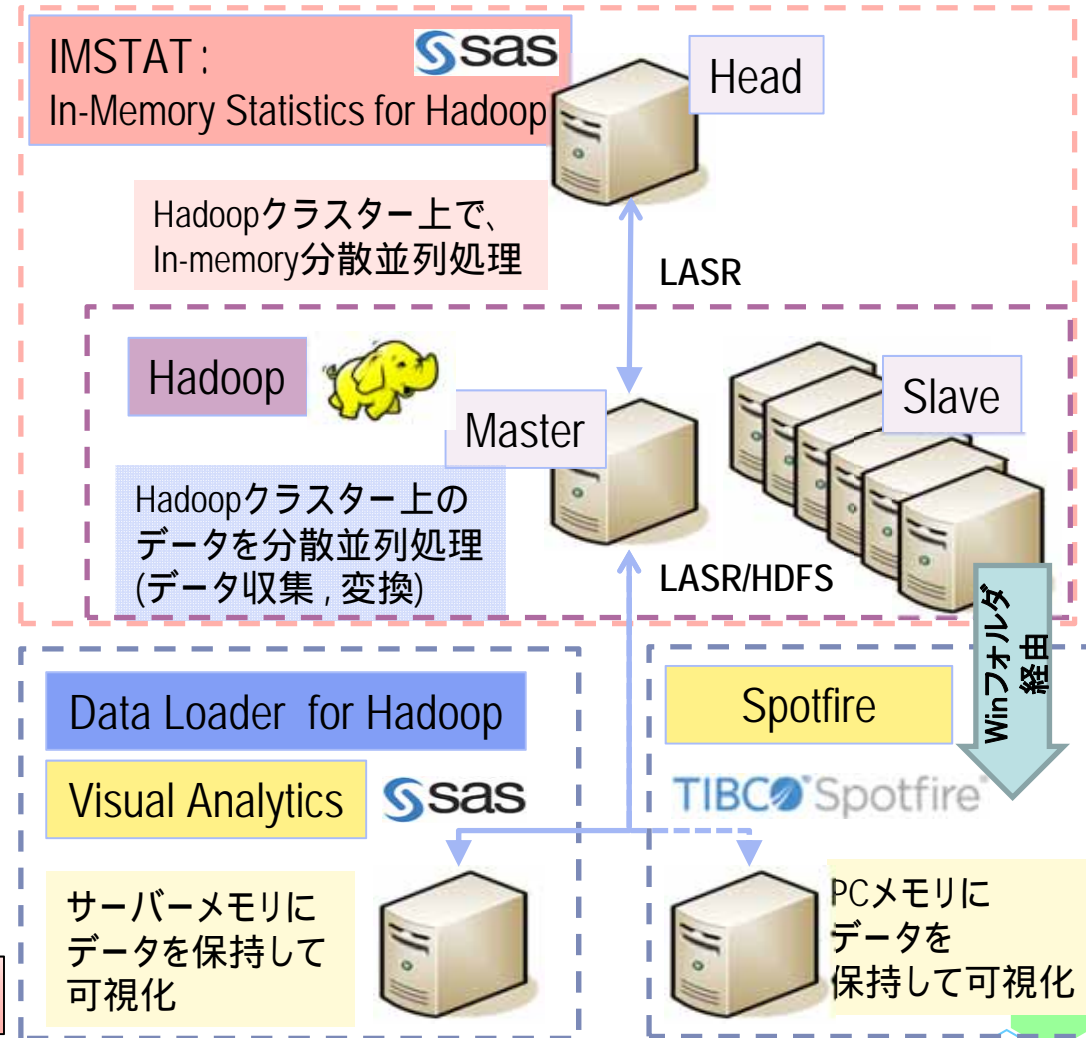
ビッグデータの活用: データ駆動型医薬品開発



製薬企業 × BigDataを見据えたシステム@Shionogi

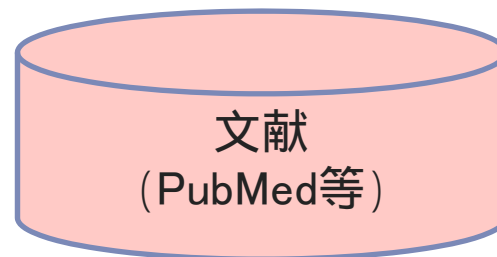
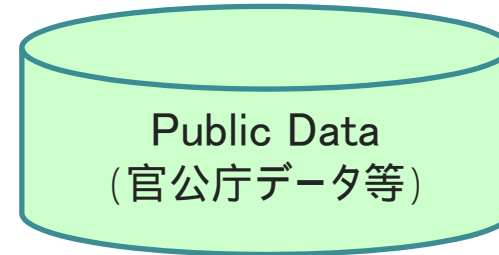
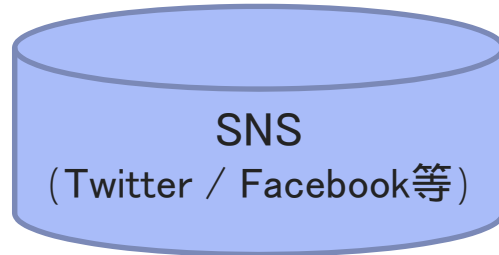


- 初期のHadoopユーザ: 2 ~ 3名
- Hadoopユーザ育成に時間を要する.
- ソフトウェア・プログラム言語の選択と集中が必要
- IMSTAT / Data Loader の導入により Hadoopを有効活用し, ビッグデータ解析環境の利便性を飛躍的に向上
- Hadoop環境に対しては順次拡張可能
- 可視化をVAとSpotfireで使い分ける.

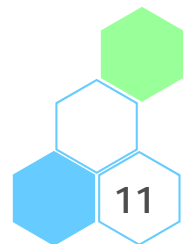


オープンデータ × 可視化と仮説

文献データとSNS(テキストマイニング&空間統計)



そのような知見があるかどうか？



医療ビッグデータ(*Real World Data*) × 可視化と予測

医療ビッグデータとは？

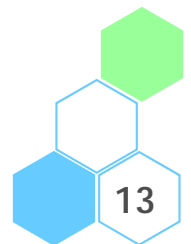


医師の診療行為から生まれるデータであり、カルテ、レセプト(診療報酬明細)などをソースとして得られるもの

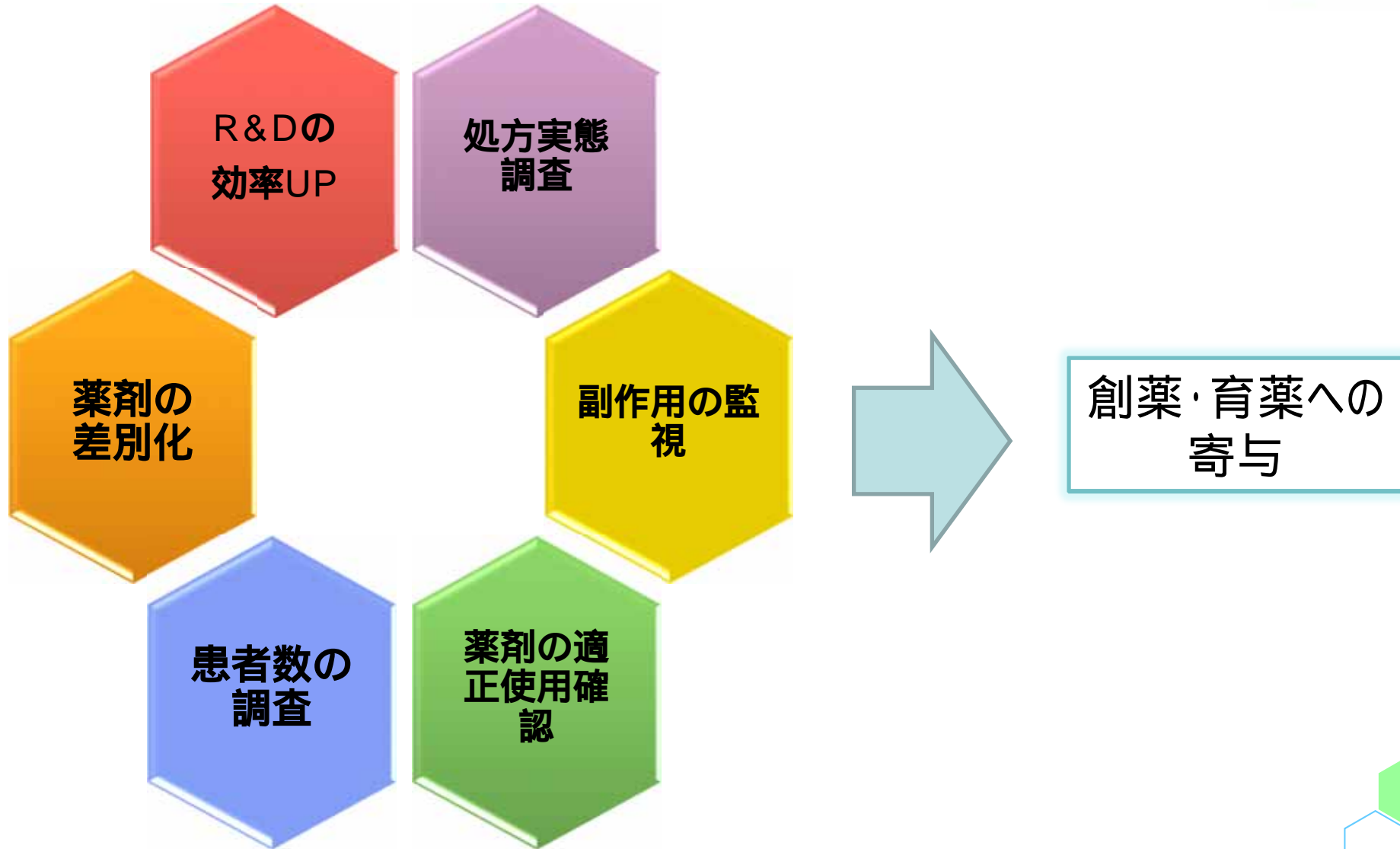
患者さん一人ひとりの薬剤処方歴, 手術歴, 診断歴, 入院歴といった情報が蓄積されたデータ

2014年に閣議決定された健康・医療戦略の中にも利用促進の旨が盛り込まれている

Real World Data (RWD) と呼ばれることも多い



RWDの活用



医療ビッグデータの解析事例



OSIM2

OMOP*が公開している, MarketScan® Research Databases等の
商用の医療データベースをもとにシミュレーションから作成された
データ

オープンデータとして公開されている

100万人分の処方, 診断, 患者背景のデータが含まれる

処方データ: 118,541,933件

診断データ: 175,029,957件



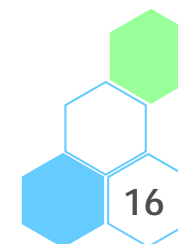
医療ビッグデータの解析事例



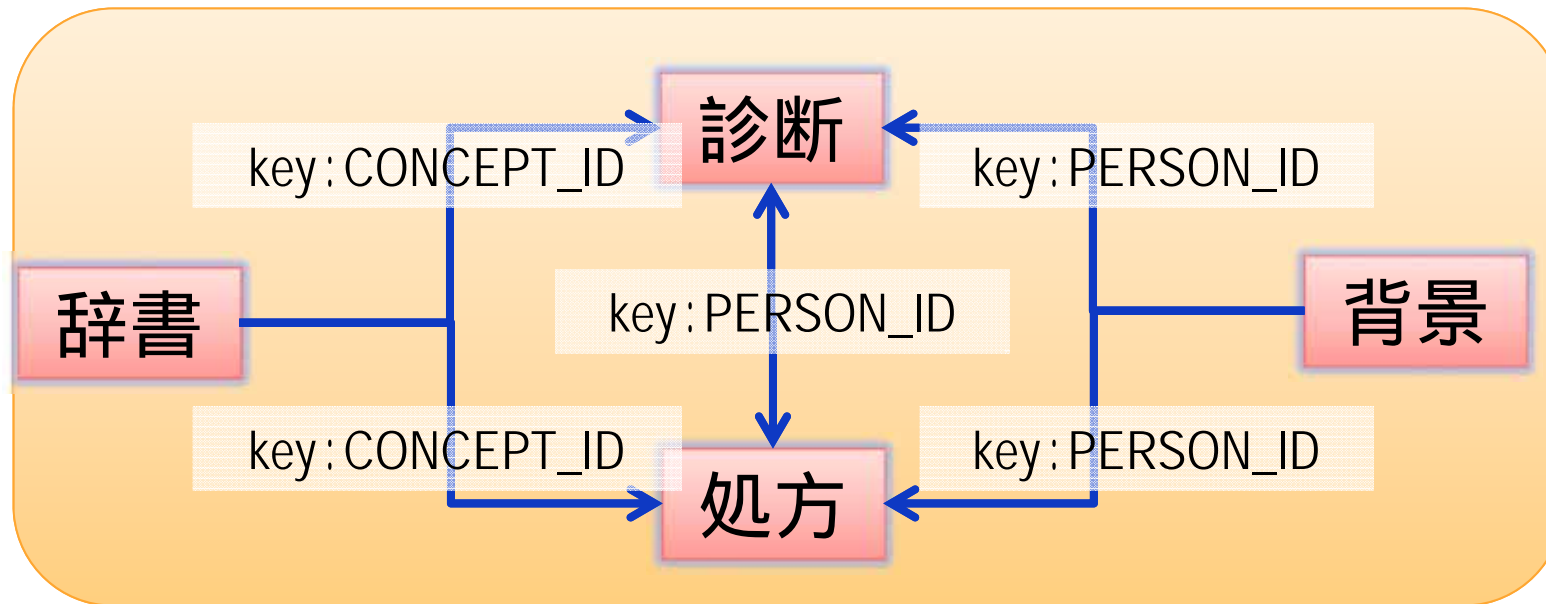
実際のRWD解析ではデータの特徴，解析目的に対するFeasibilityの確認が必須ですが，本発表では割愛している部分もございます

“RWDで出来ること”に焦点をあててご紹介いたします

説明時の解釈はあくまで一例です



データの説明～結合～



解析対象データ

同じ疾患に対して色々な種類の薬剤が発売されている時

1. どの薬剤が最も使われているか？
 - 薬剤ごとに使っている人の差異があるかもしれない
2. 服用薬をある薬剤から他の薬剤へ変更した人はどれくらいいるか？
 - 変更しているのはどういう人か？



SAS VAのサンキーダイアグラムを利用

疾患遷移

症例単位の疾患をパスウェイで表現する.



サンキーダイアグラムが表現に適している .

注目している疾患の患者数がどれくらいいるか

1. 市場性の予測

より細かい患者像に絞って検討することができる

2. アンメットニーズの把握

人口予測モデルとの組み合わせ

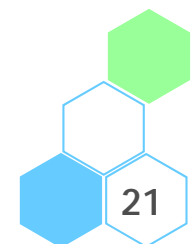
患者数推定 ~ 考え方 ~



男性($i=1$)		基準集団 (国勢調査など) (人)	観察集団 (OSIMのデータ) (人)	患者数
年齢(j)		p_{ij}	N_{ij}	X_{ij}
~5歳	1	p_{11}	N_{11}	X_{11}
6~9歳	2	p_{12}	N_{12}	X_{12}
10~14歳	3	p_{13}	N_{13}	X_{13}
:	:	:	:	:

年齢と性のズレを調整する

$$\text{推定患者数} = \sum_j \sum_i \frac{p_{ij}}{N_{ij}} * X_{ij}$$



【RWD事例】患者数推定（若年性認知症）



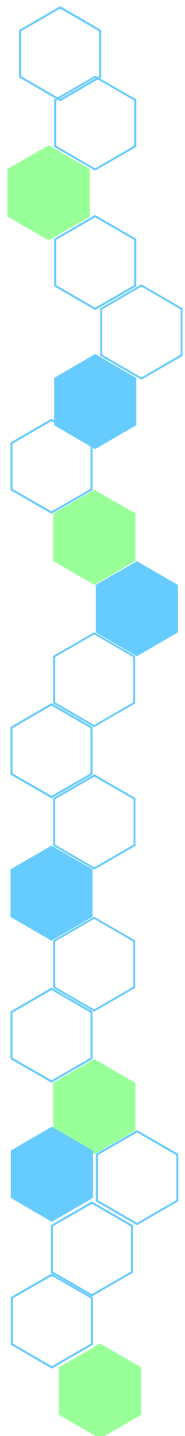
人口ピラミッドの予測 × RWDの罹患率

→ 将来患者数の予測からアンメットニーズを把握できる



まとめ

- 身近なデータを可視化するだけでも“新たな気づき”を得られる場合がある.
- オープンデータも多種多様であり, 組み合わせて適切に可視化することで, 仮説立案の助けになる.
データ×データ×データ...
- 医療ビッグデータは産官学共に注目しており, 今後, 臨床試験の在り方が変わり得るかもしれない. ただし, 現状, 項目がシンプルなため, どう可視化するかが重要.
- Spotfire vs Visual Analyticsは, 現在のところ, 各々特色があり, 甲乙付け難い **引き分け.**



ナニワ

データサイエンス研究会

- End of File -

