

HTAにおける Real World Dataの応用可能性

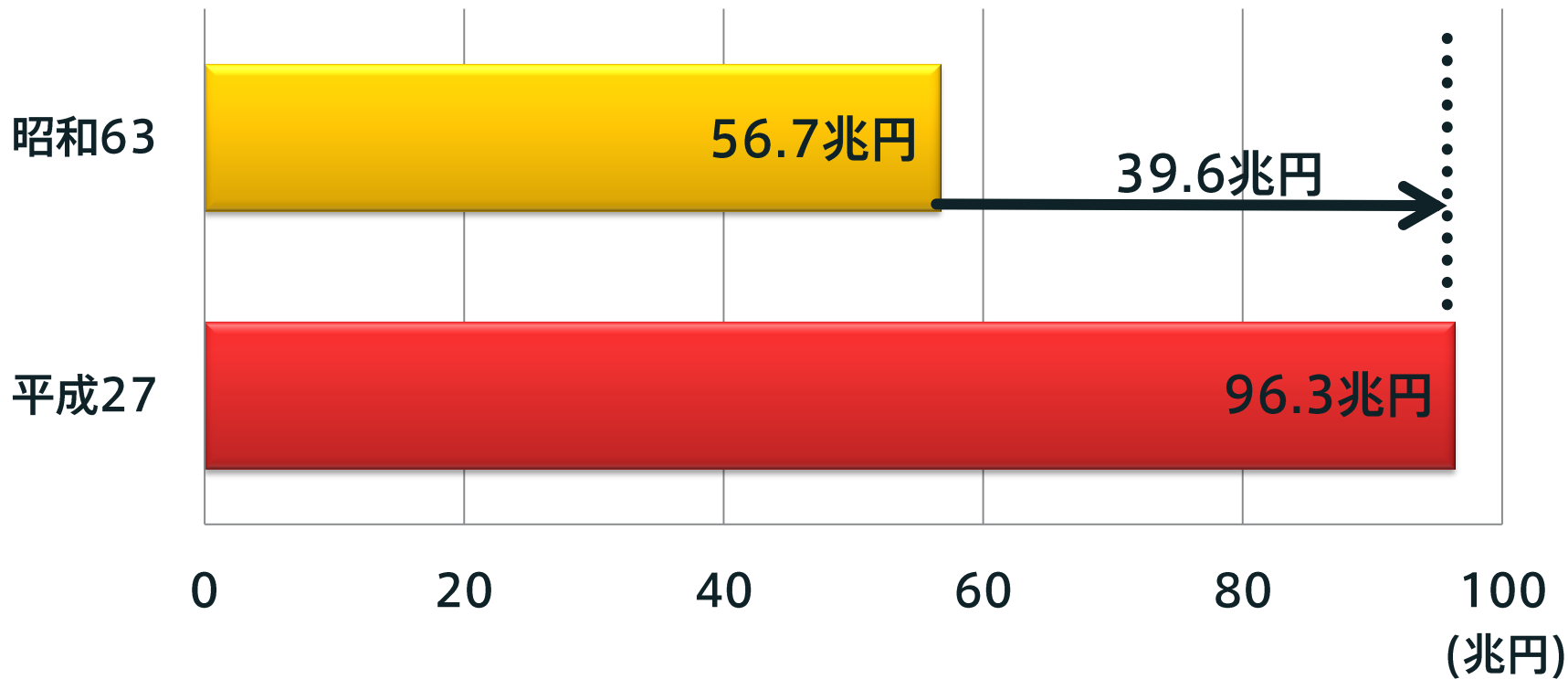
株式会社データビークル
西内 啓

自己紹介

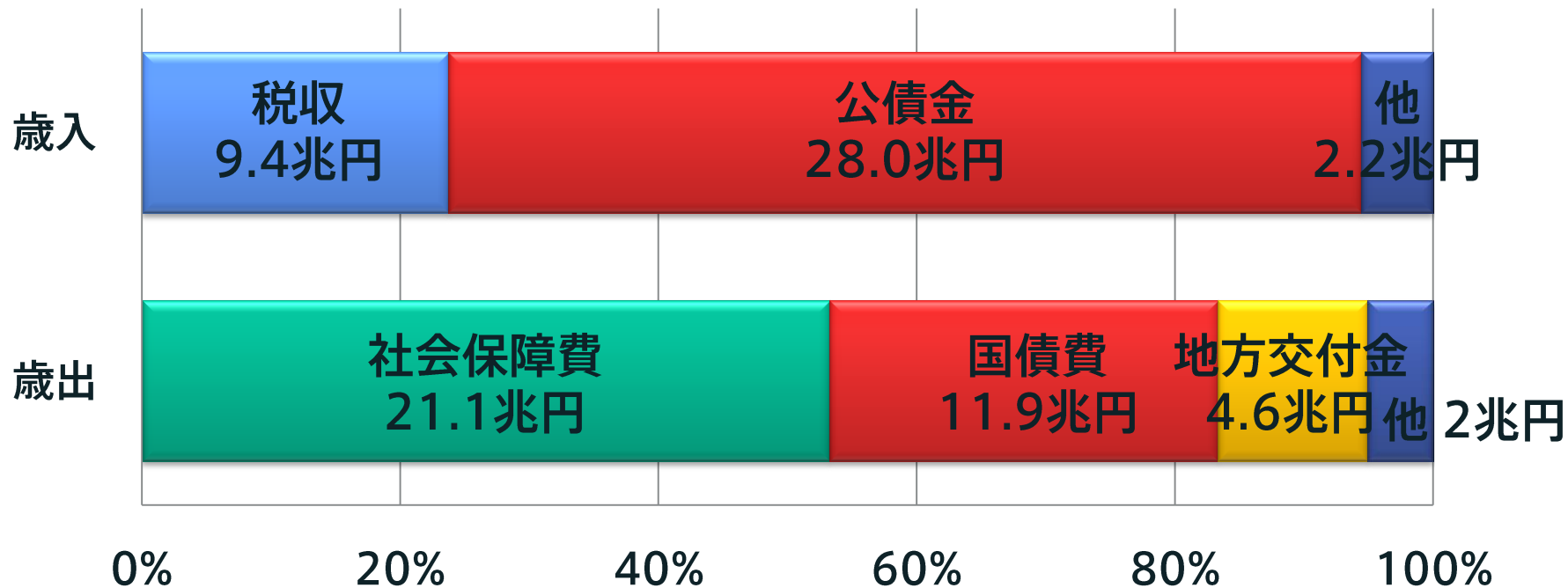
- 平成16 : 東京大学医学部卒(生物統計学)
- 平成20 : 東京大学医学部助教
- 平成22 : ハーバード大学客員研究員
- 平成22～ : 在野へ
- 平成26 : データビークル創業

なぜ今HTAか

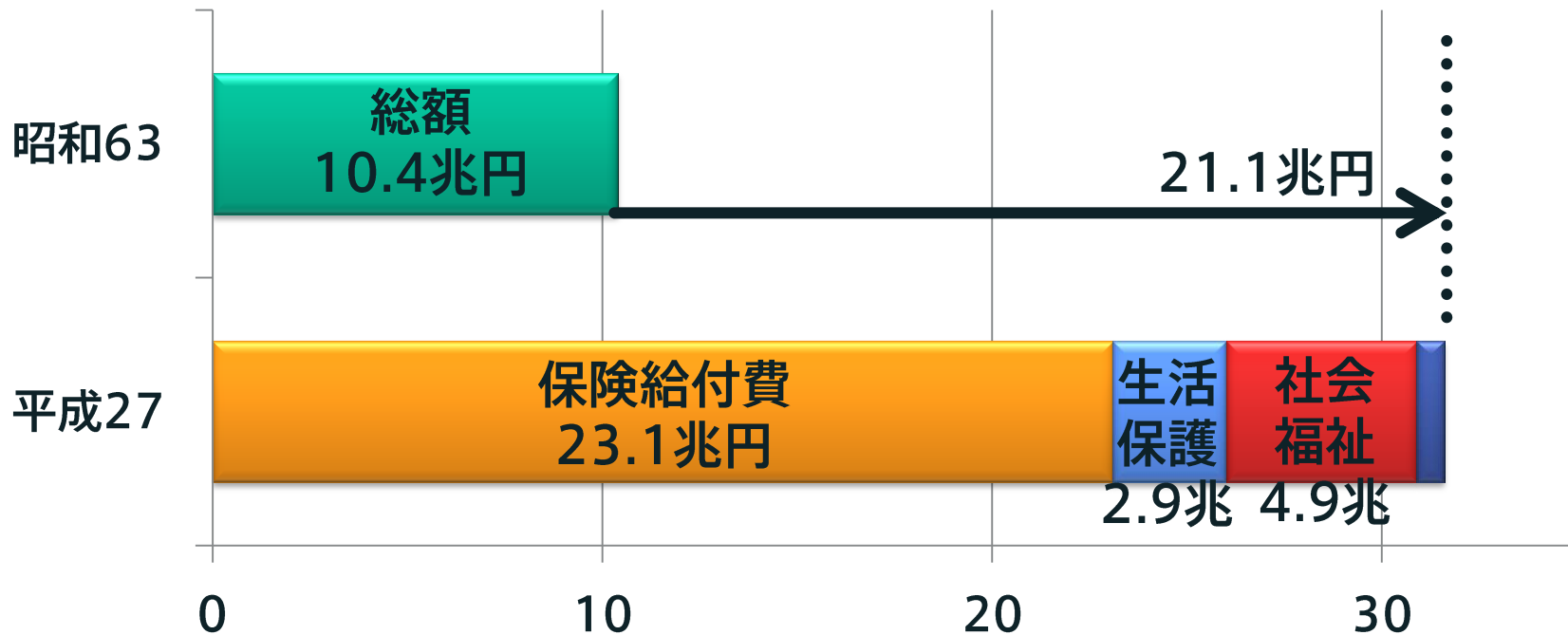
一般会計歳入歳出規模の比較



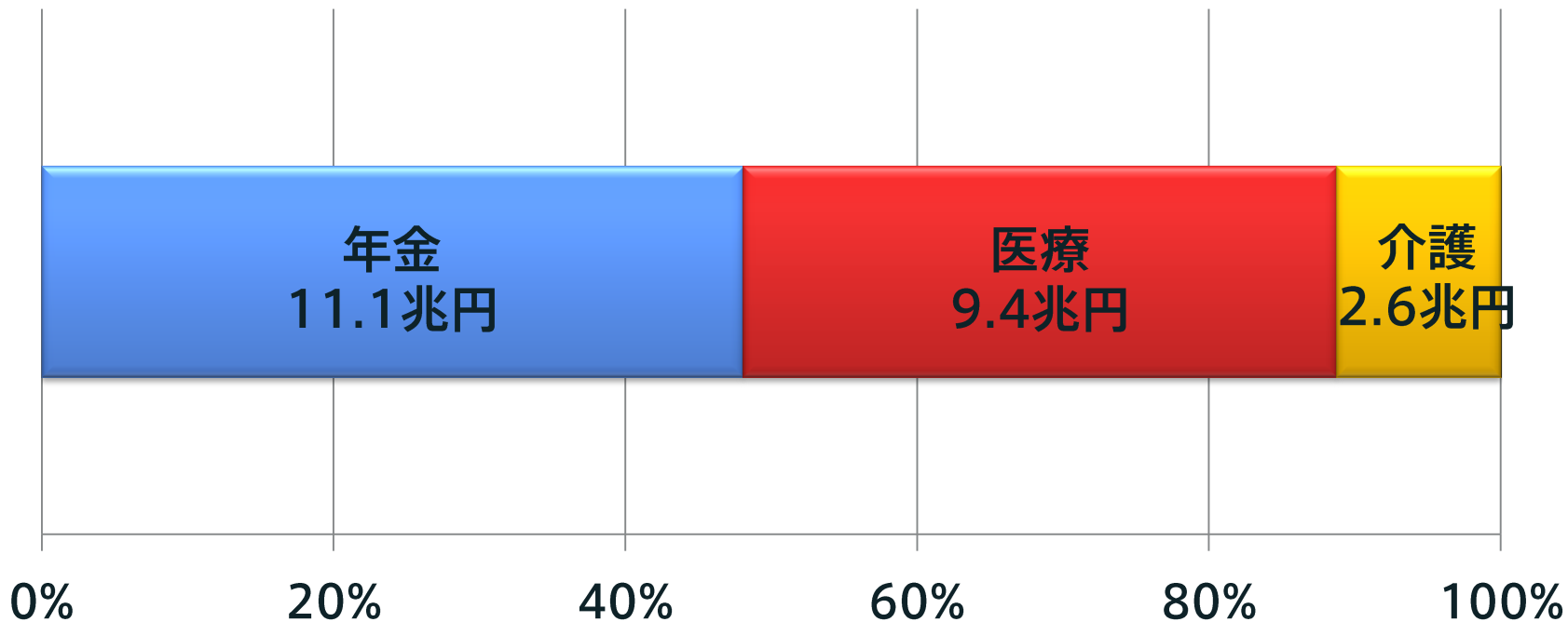
増分39.6兆円の内訳



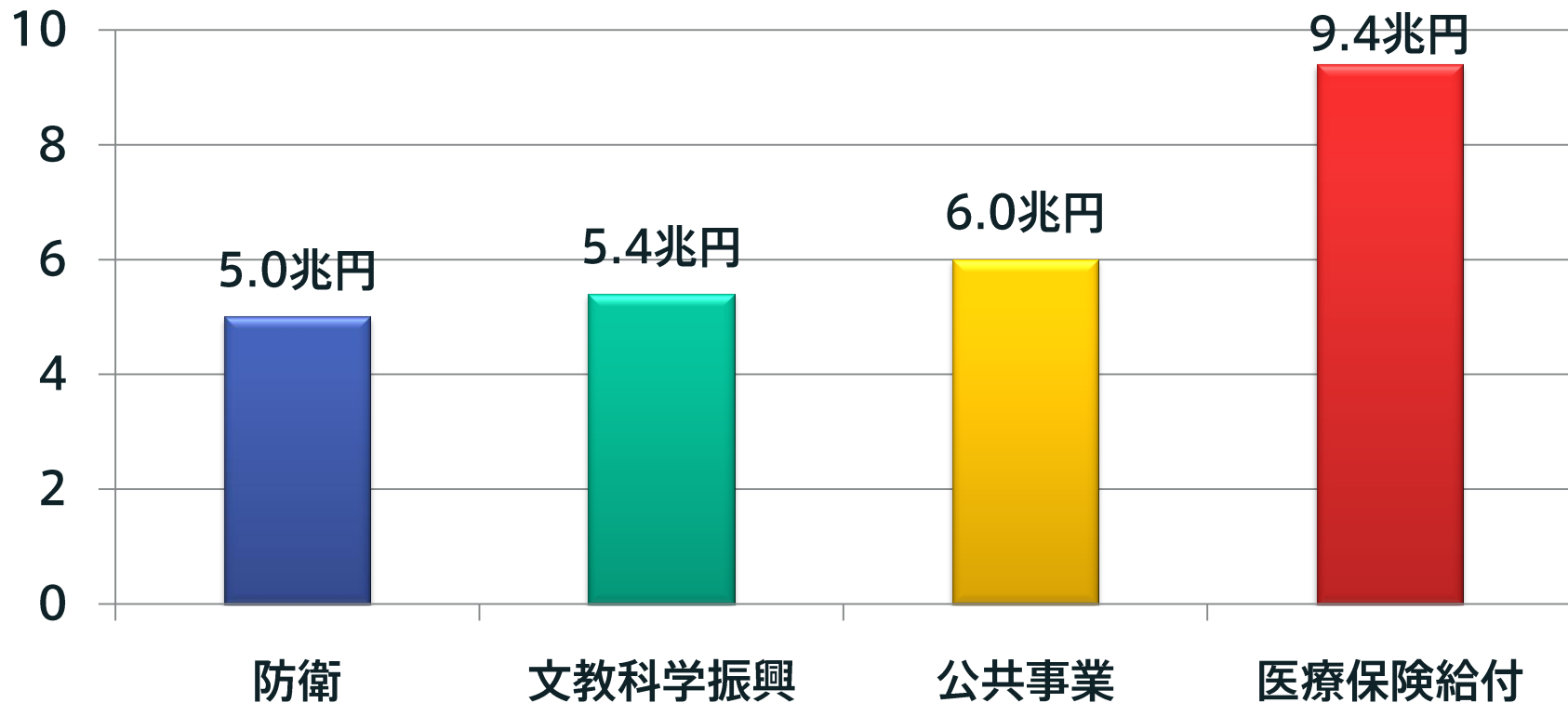
社会保障費の内訳



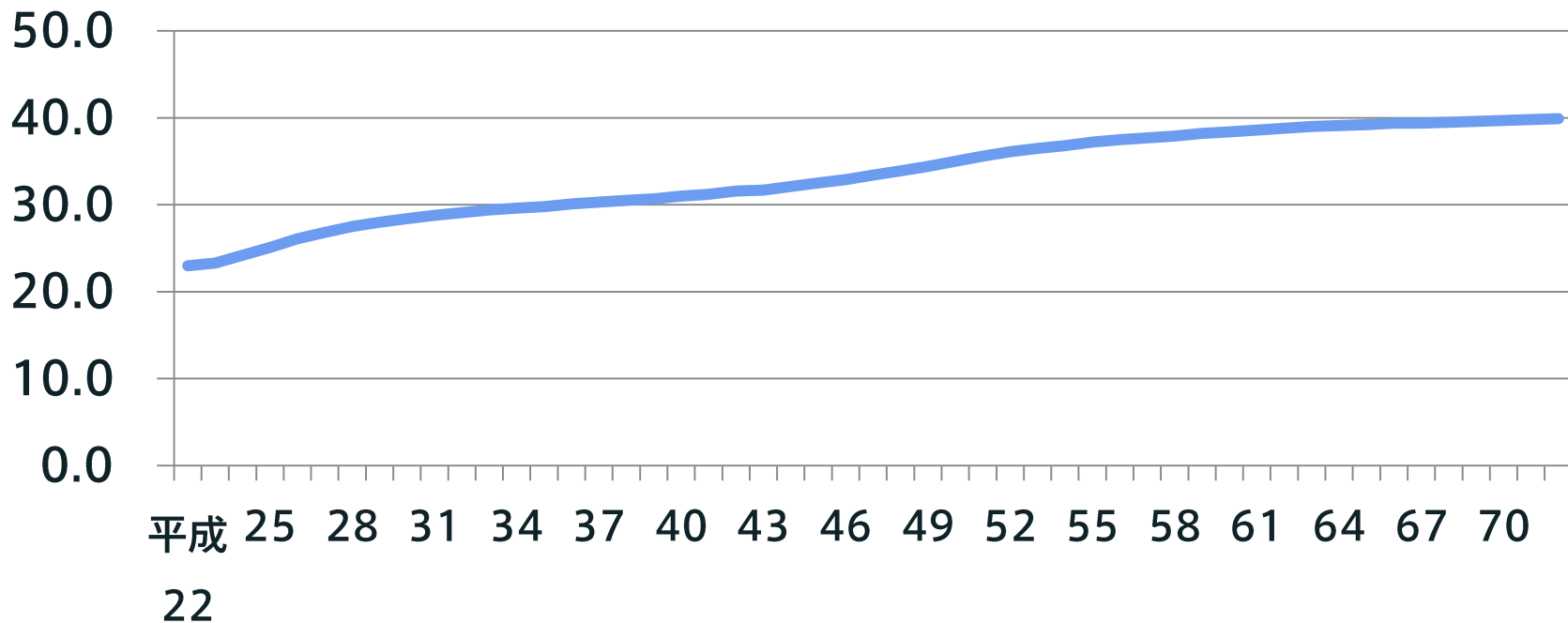
保険給付費の内訳



他の予算との比較(H27)

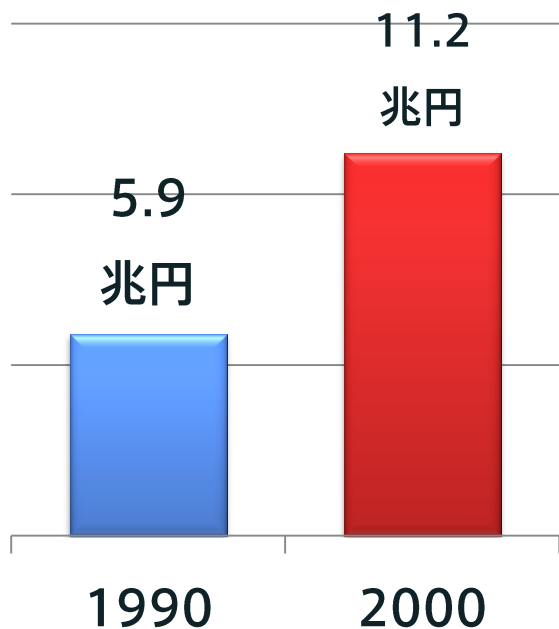


今後の高齢化率の予測値

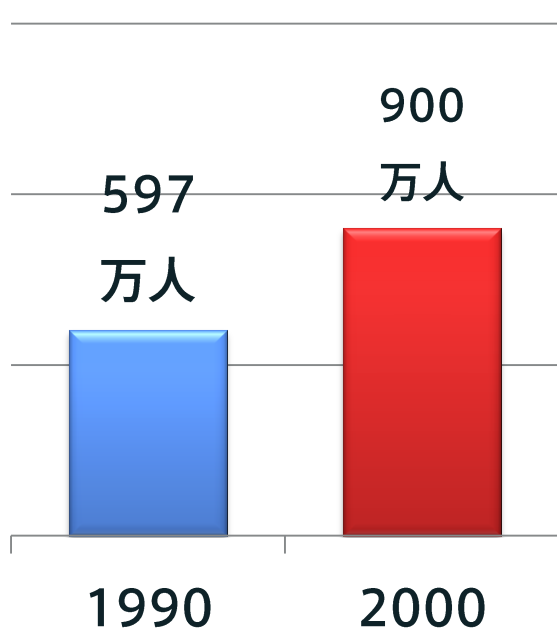


医療費の増大は高齢化のせい？

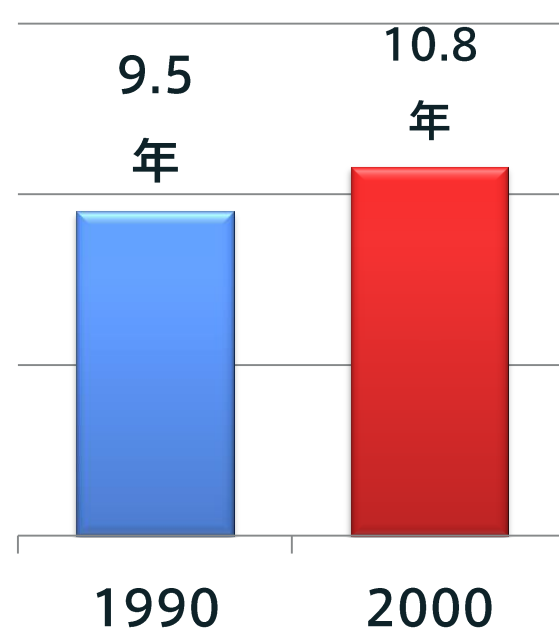
後期高齢者医療費



後期高齢者人口



75歳時余命(男)



もし財政が破綻したら・・・



今のうちに
ソフト
ランディング

Stuckler D & Basu S(著), 橘&臼井(訳) 2014.

お金が有限なら . . .

1人の命



同じコストで
5人の命



HTAとは

医療技術評価機関国際ネットワーク (INAHTA)の定義

医療技術の開発,普及,使用による
医学,経済,社会,倫理的意義を
分析する学際的な政策研究分野

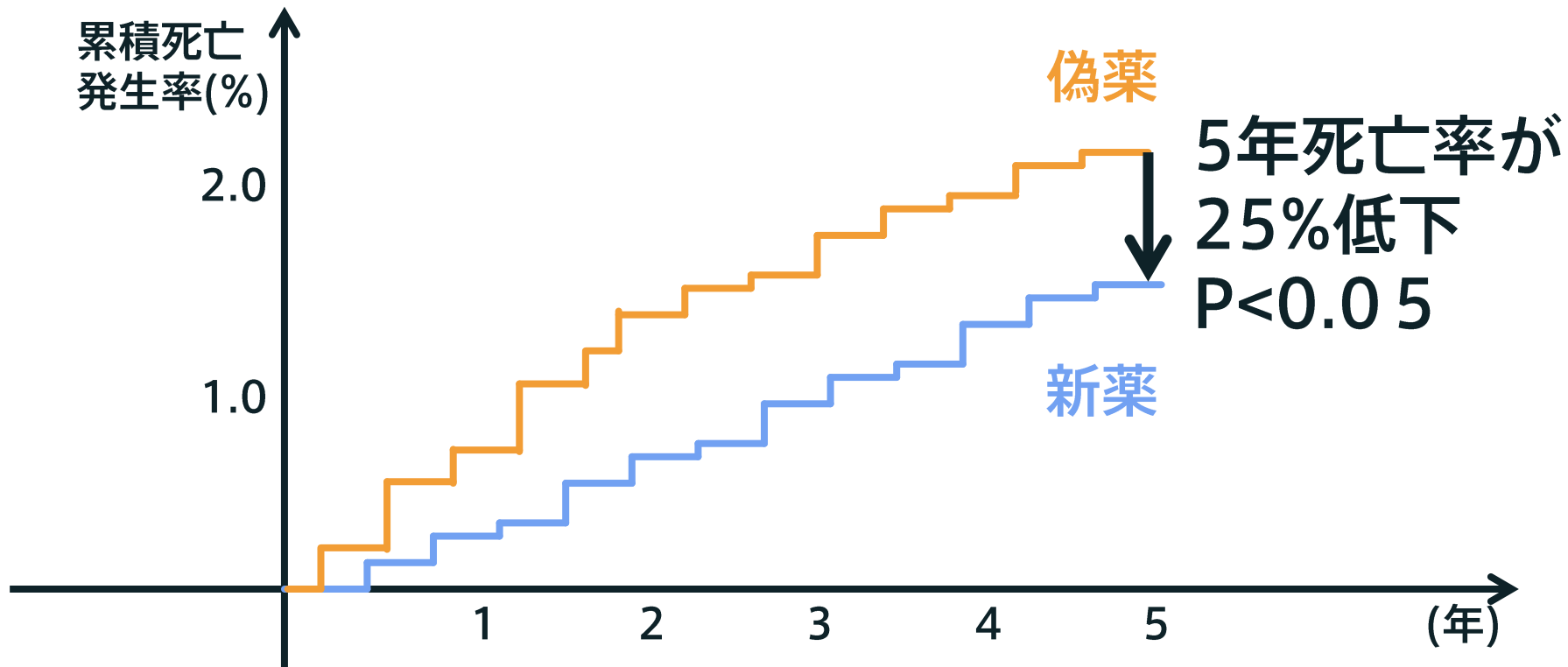
ハーバードで聞いた表現

医療技術の

P-valueだけじゃなく

Value(価値)それ自体を議論

p-valueで評価する医療技術



医療技術のvalueは？

プラセボ群
5年死亡率
2.0%

1万人いたら5年で200人が死亡

新薬投与群
5年死亡率
1.5%

1万人いたら5年で150人が死亡
つまり**50人分の命**が救われる

そのためのコストは？

新薬1錠50円



一日2錠を5年分



1万人に投与



公的負担は9割



16.4億円

一人あたり
3280万円の
社会的コスト

50人分の命

日本人に聞いてみました

Q. 他人の1QALYに対していくらまでの社会的コストなら許容できますか？

平均値
約600万円

中央値
約100万円

HTA普及の歴史

- 1990初頭 カナダ,豪州で経済評価必須化
- 1999 英国国立医療技術評価機構
(NICE) 創設
- 2008 韓国でもHTA導入
台湾,タイなども追随
- 2012 中医協で . . .

Real World Dataって？

RealじゃないWorldの例

注意深い管理下



内的妥当性と外的妥当性

内的妥当性

比較する群間での
偏りのなさ

ランダム化すれば
大丈夫

外的妥当性

研究参加者自体の
偏りのなさ

ランダム抽出？
全数調査？

全数っぽいデータたち

レセプト

アプリ
利用ログ

センサー
データ

健診記録

電子カルテ

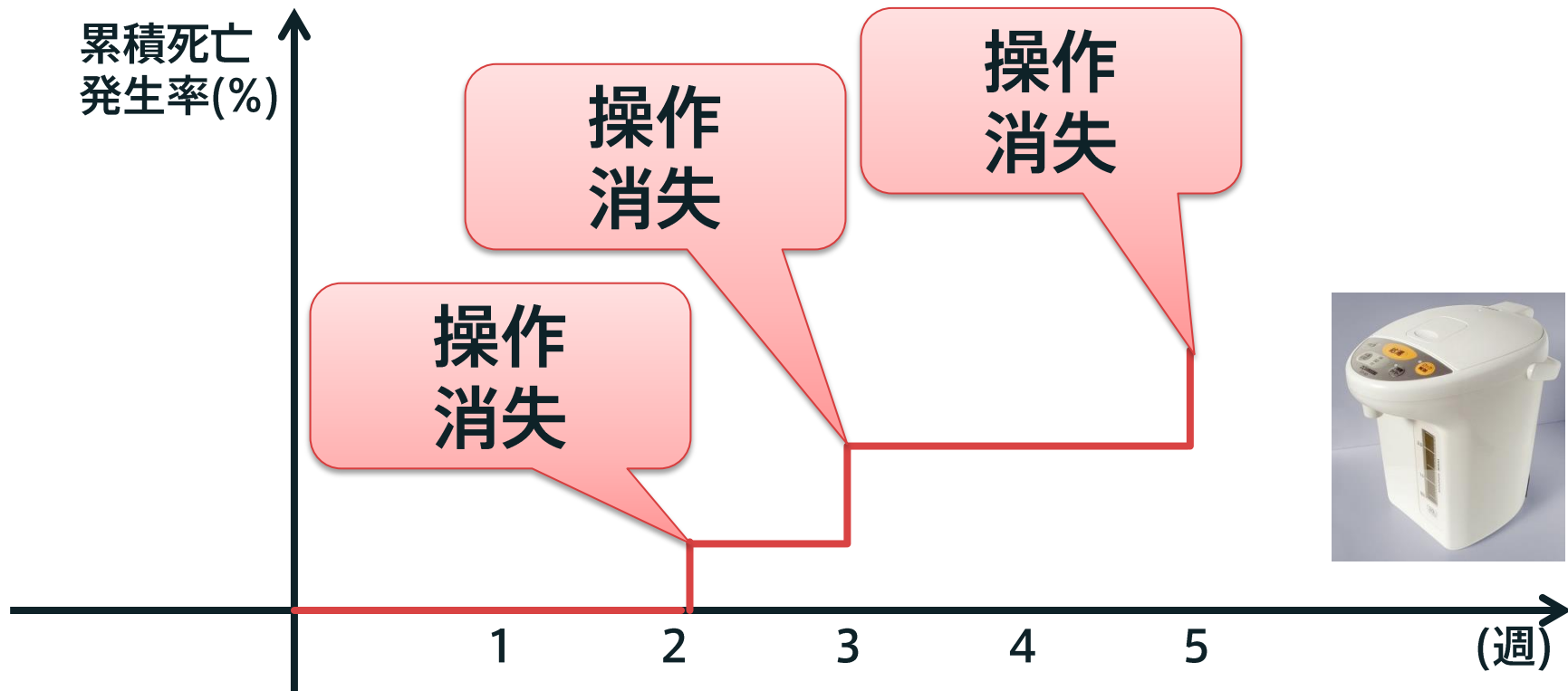
検索履歴

センサーデータの例

操作がWeb上に
発信/記録され
お年寄りの生存確認に！



電気ポットで生存時間分析



アプリの利用ログ

アドヒアランスが
どんどん蓄積！



検索履歴の利用例

google.org インフルトレンド

[Google.org ホーム](#) (英語)

[デング熱トレンド](#)

インフルトレンド

ホーム

国/地域を選択 ▼

[インフルトレンドの仕組み](#)

[よくある質問](#)

インフルエンザの流行状況

きわめて大規模

大規模

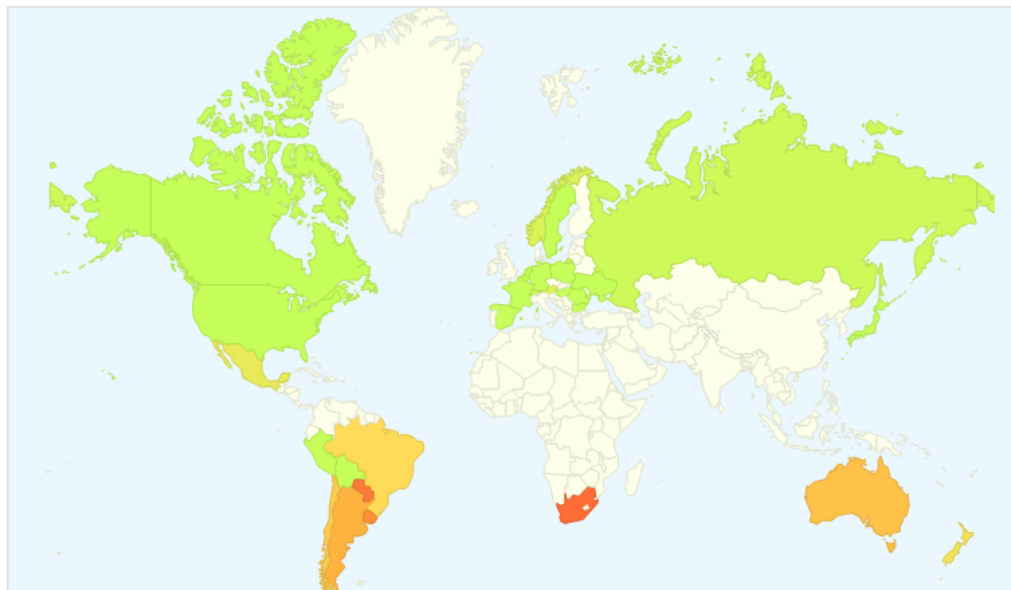
中規模

小規模

ごく小規模

世界のインフルエンザの流行を探る

Google では、特定の検索キーワードでの検索数がインフルエンザの流行の指標となることを発見しました。Google インフルトレンドでは Google 検索の集計データを使用してインフルエンザの流行を予測します。[詳細](#)



ただし . . .

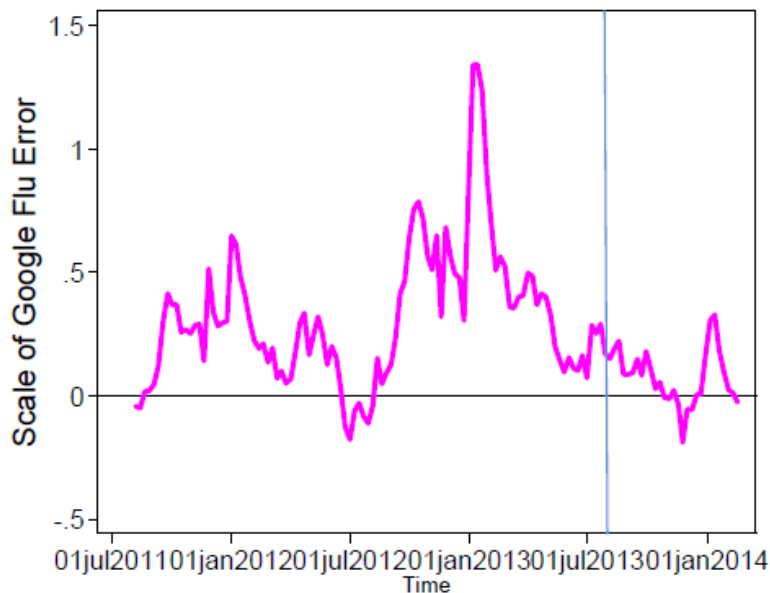
Google Flu Trends Still Appears Sick: An Evaluation of the 2013-2014 Flu Season

David Lazer,^{1,2*} Ryan Kennedy,^{1,2,3} Gary King,² Alessandro Vespignani¹

Electronic copy available at: <http://ssrn.com/abstract=2408560>

内容の抜粋

- 過大評価になりがち
(5割増ぐらい)
- 単純な自己回帰モデルの方が良い精度



カイザー・ファングによる “ビッグデータ”の限界

- O : あくまで観察データ
- C : 管理されていない
- C : 全数なように「思える」
- A : 元々別な用途のためのもの
- M : 複数のデータソースを結合

レセプトや電子カルテも . . .

O : あくまで観察データ

C : 管理されていない

C : 全数なように「思える」

A : 元々別な用途のためのもの

M : 複数のデータソースを結合

主な電子カルテベンダ

富士通

パナソニック ヘルスケア

ソフトウェア・サービス

日本電気

シーエスアイ

アピウス

日本アイ・ビー・エム

東芝医療情報システムズ

ワイズマン

日立製作所

亀田医療情報

ビー・エム・エル

三洋電機

システムロード

日本の健康保険

組合健保

共済

国保

船員保険

後期高齢者

協会健保

データの理想(縦に結合)

ID	Birthdate	Gender	...	Created_at
TKY-00001	1960/11/2	男性		2001/4/1
TKY-00002	1989/3/1	女性		2001/4/2
TKY-00003	1974/6/12	男性		2001/4/2

重複なし

同じ列名で同じ形式のデータ

ID	Birthdate	Gender	...	Created_at
OSK-00001	1991/2/28	女性		2010/1/4
OSK-00002	1943/9/3	女性		2010/1/5
OSK-00003	1956/7/29	男性		2010/1/9

データの現実(縦に結合)

ID	Birthdate	Gender	...	Created_at
0000001	1960/11/2	男性		2001/4/1
0000002	1989/3/1	女性		2001/4/2
0000003	1974/6/12	男性		2001/4/2

重複あり

異なる列名,異なる形式,異なるコード体系

ID	Age	Sex	...	Created_at
0000001	39	2		4 Jan 2010
0000002	42	2		5 Jan 2010
0000003	53	1		9 Jan 2010

データの理想(横に結合)



異なるデータソース,異なるIDの 同一人物を特定する苦勞

ID	氏名	住所	...
00809342	齊藤 和子	東京都調布市つつじヶ丘2丁目3-	...

ID	氏名	住所	...
05169160	齋藤 和子	調布市つつじヶ丘 二丁目三番地ノ	...

ID	氏名	住所	...
069534	齋藤和子	東京都調布市つつじが丘2-3-	...

ID	姓	名	住所	...
A1523189	鈴木	和子	東京都文京区弥生1丁目1-	...

Real World Dataの理想

一元的に個人等を特定する一意なID

項目名,コード体系,データ型等の標準化

研究者や実務者のアクセシビリティ

そこに至るハードル

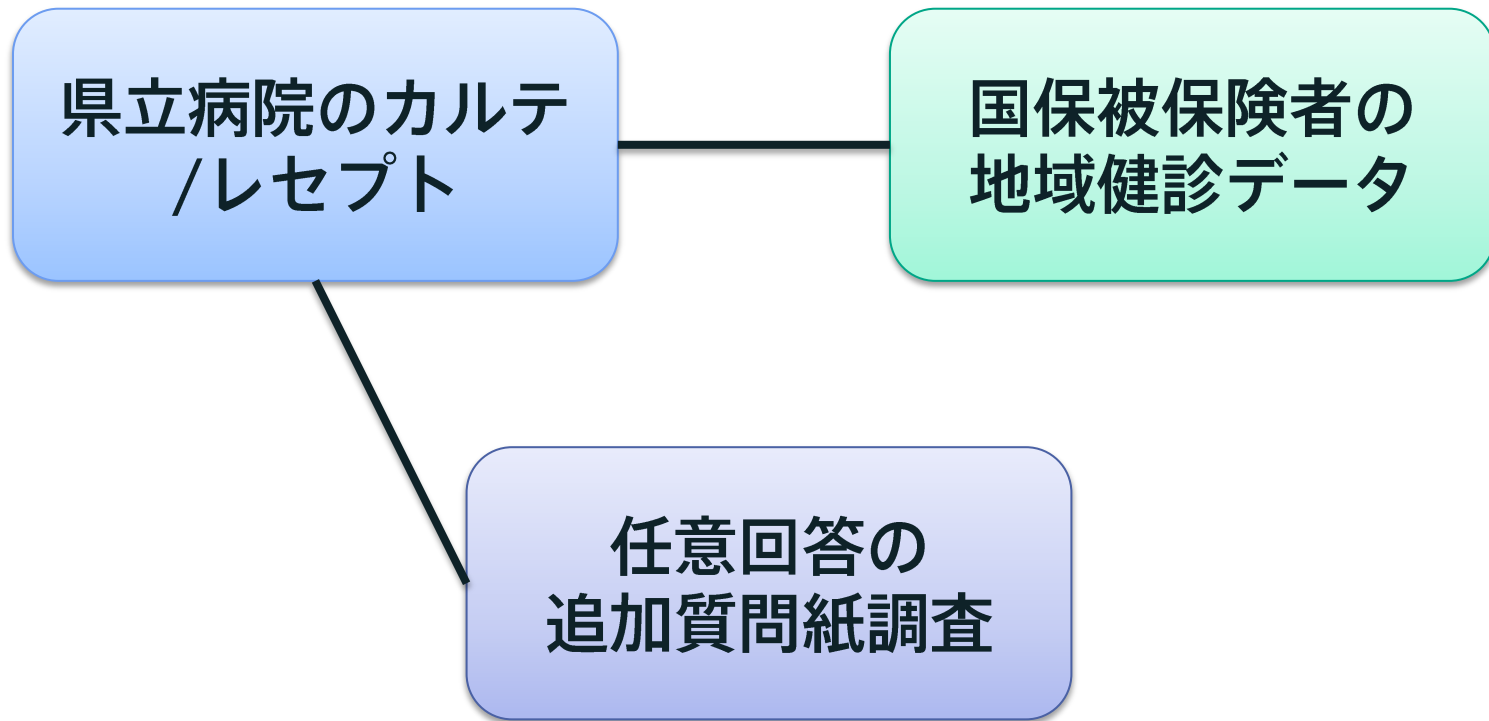


うまくいかない考え方



データが完全になるまで
分析できません！

現実的なミニマムスタートの例



統計手法も味方につけよう

欠測

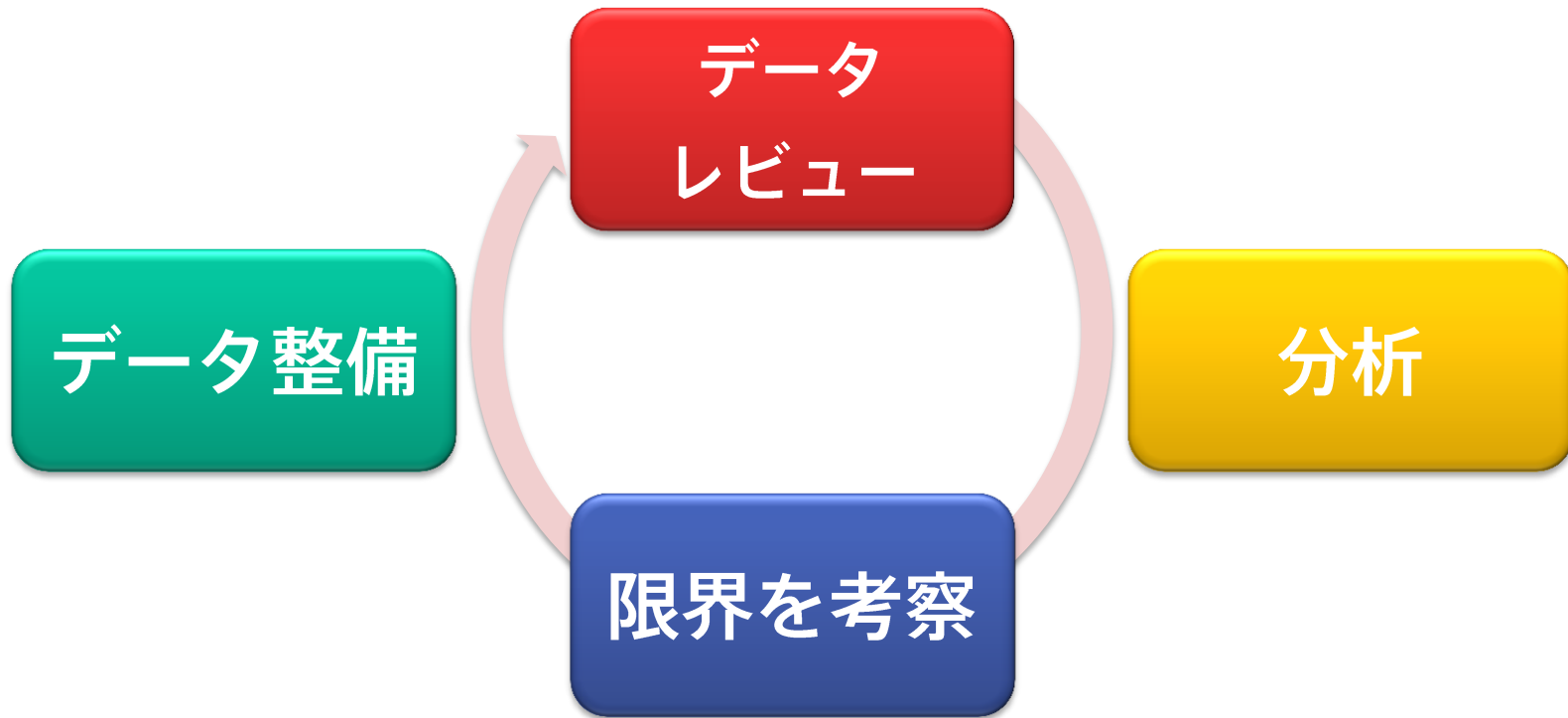
打ち切り

選択バイアス



手法で
対処可能な
ものも

継続的なサイクルを



Real World Dataについてのまとめ

キレイでも万能でもない

対処すべき方法と価値はある

分析と整備に不断の努力を