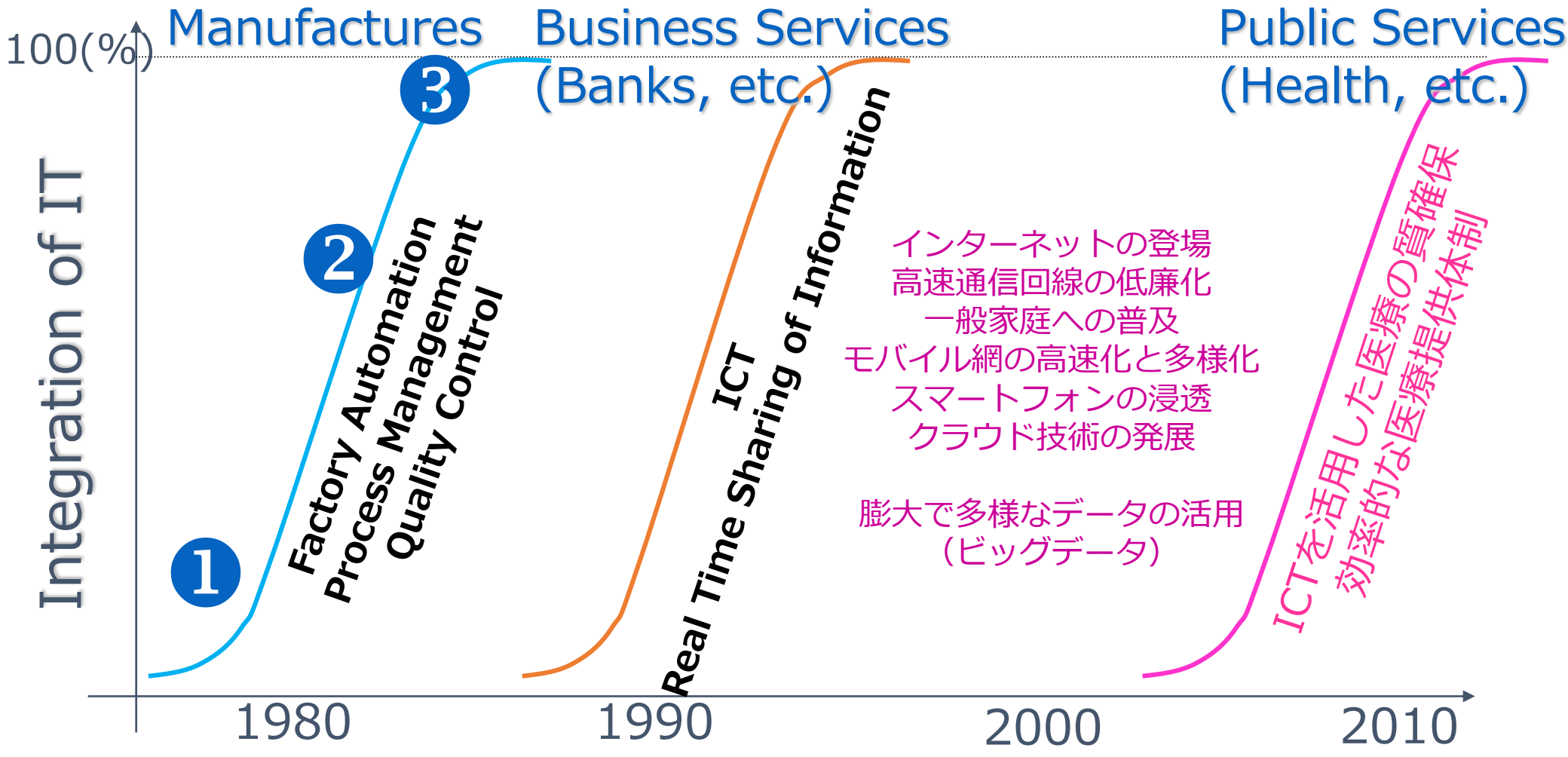
SDMコンソーシアムの役割とその社会的な背景

岐阜大学大学院医学系研究科
医療情報学分野
紀ノ定保臣



❶ IT as gadget

❷ Trojan horse networks

❸ Final full integration of ICT in Business (Re-Engineering, Law, ...)

医療の質向上に向けたビッグデータの活用

A)　「医療の質向上に向けたビッグデータの活用」を因数分解する
　　→　「医療の質」、「質の向上」、「ビッグデータ」、「活用」

①　「医療の質」の定義：

「構造（ストラクチャー）」

「過程（プロセス）」

「成果（アウトカム）」　A. Donabedianが提唱（1966年）

②　質を向上させるための手法　→　「構造・過程・成果」の改善

③　特に「過程（プロセス）」が重要　←　マネジメントが必要

④　「過程（プロセス）」の改善　←　ビッグデータの活用

ビッグデータとは

1．既存の技術では管理することが困難なデータ

2．管理することが困難になる要因　：　３V　　＋　Value：価値

- Volume：データの量が大きいこと
- Variety　：データの種類が多様である
  診療記録のテキストデータ，検査の数値データ，画像など
  (non-structured data)　　(structured data)
- Velocity：　データの発生頻度（速度）や更新の頻度が速いこと

3．データの特性と技術について

- ３Vの面で管理が困難なデータ（データの特性）
- それらのデータを蓄積・処理・分析するための技術
- それらのデータを分析し，有用な意味や洞察を引き出せる人材や組織

4．ビッグデータを蓄積・処理・分析するための技術

- 大規模データを効率よく，高速に処理する基盤技術（例：Hadoop）
- 柔軟で拡張性に優れたデータベース技術（例：Key-Value，NoSQL）
- 機械学習
- 統計解析

5．データサイエンティスト

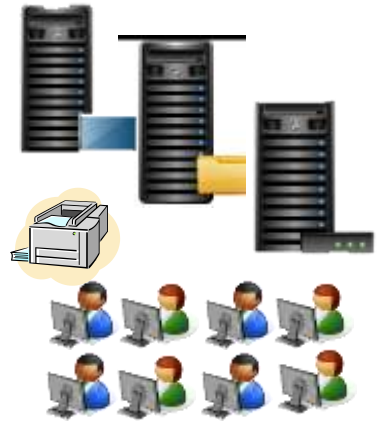
- データを分析し，有用な意味や洞察を引き出せる人材や組織

バランス・スコアカードの視点から見た
医療ビックデータ活用型リアルタイム・ダッシュボード・システム

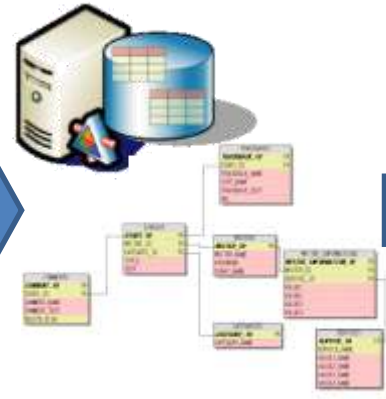
Balanced Scorecard

Finance （財務）
Outcome（成果）
Process （過程）
Structure（構造）

EMR



Clinical Data Repository



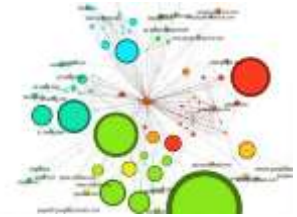
Real Time Dashboard

View

病院全体として・・・
地域全体を見ながら

オーダー・実施内容は？
外来・病棟機能は？
患者プロファイルは？
患者ADL？，医療コスト？

患者
スタッフ
診療情報
クリニカルパス



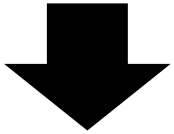
分析力が競争力となる時代の到来

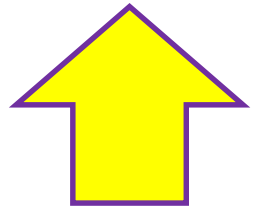
- データ分析に基づく意思決定
- サービスの開発力
- 経営判断
- サービスの設計

など

ビッグデータの課題

- 統計・データ分析の専門家不足
- プログラミングの高度化・多様化
- セキュリティ・マネジメントに対する懸念





【ビッグデータ時代の要素技術：Hadoop】

- HDFS（Hadoop Distributed File System）
  ：分散ファイルシステム
- MapReduce ： 分散処理フレームワーク
- 大量に，深く，素早くを支える

【次世代のBIツール】

- 大規模リアルタイムデータ分析基盤（リアルタイムDWH）

医療の質向上に向けたビッグデータの活用

B)　「医療の質向上に向けたビッグデータの活用」を因数分解する
　　→　「医療の質」、「質の向上」、「ビッグデータ」、「活用」

①　「医療の質」の定義：

「構造（ストラクチャー）」

「過程（プロセス）」

「成果（アウトカム）」

②　質を向上させるための手法　→　「構造・過程・成果」の改善

③　特に「過程（プロセス）」が重要　🡐　マネジメントが重要

④　「過程（プロセス）」の改善　　🡐　ビッグデータの活用

⑤　活用するための手法　→　「価値(Value)ある情報」を取り出すこと

診療現場にマネジメントを導入する

Management

・管理

・経営

Process（過程）

・Administrative Process（管理過程）

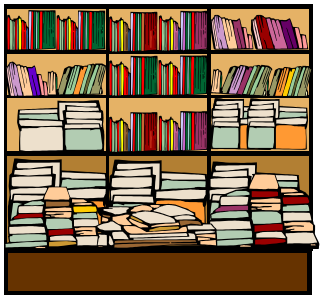
・Clinical Process（診療過程）

医療の質向上

診療過程の可視化・分析・課題の発見・改善

収益の向上に貢献

- 保険診療のルールを理解
- 診療録の記載
- 診療録の保存



診療録管理

診療情報管理

※看護記録・その他諸記録を含む

診療過程の可視化・分析・課題の発見・改善
（プロセス改善型アプローチ）

診療過程の可視化	医事会計システム／オーダエントリ-システム／電子カルテシステム
分析	DWH（データウェアハウス）／Dash Board（ダッシュボード）
課題の発見	データ・サイエンティスト／診療情報管理士
改善	PDCA／病院長のリーダシップ

医療ビッグデータ
- 診療データの活用
- 医療の質向上
- 患者満足度の向上

データサイエンス
- 統計科学
- データマイニング
- プロセス改善

データマイニング
- 分類分析
- クラスタ分析
- アソシエーション分析

診療過程の可視化・分析・課題の発見・改善（仮説アプローチ）

課題の明確化	仮説のアプローチ（課題解決の目的・仮説・検証・分析／手段）
データの収集	医事会計システム／オーダエントリ-システム／電子カルテシステム どのようなデータが必要か
分析	DWH（データウェアハウス）／Dash Board（ダッシュボード） 何を明らかにすれば良いのか
改善	PDCA／病院長のリーダシップ

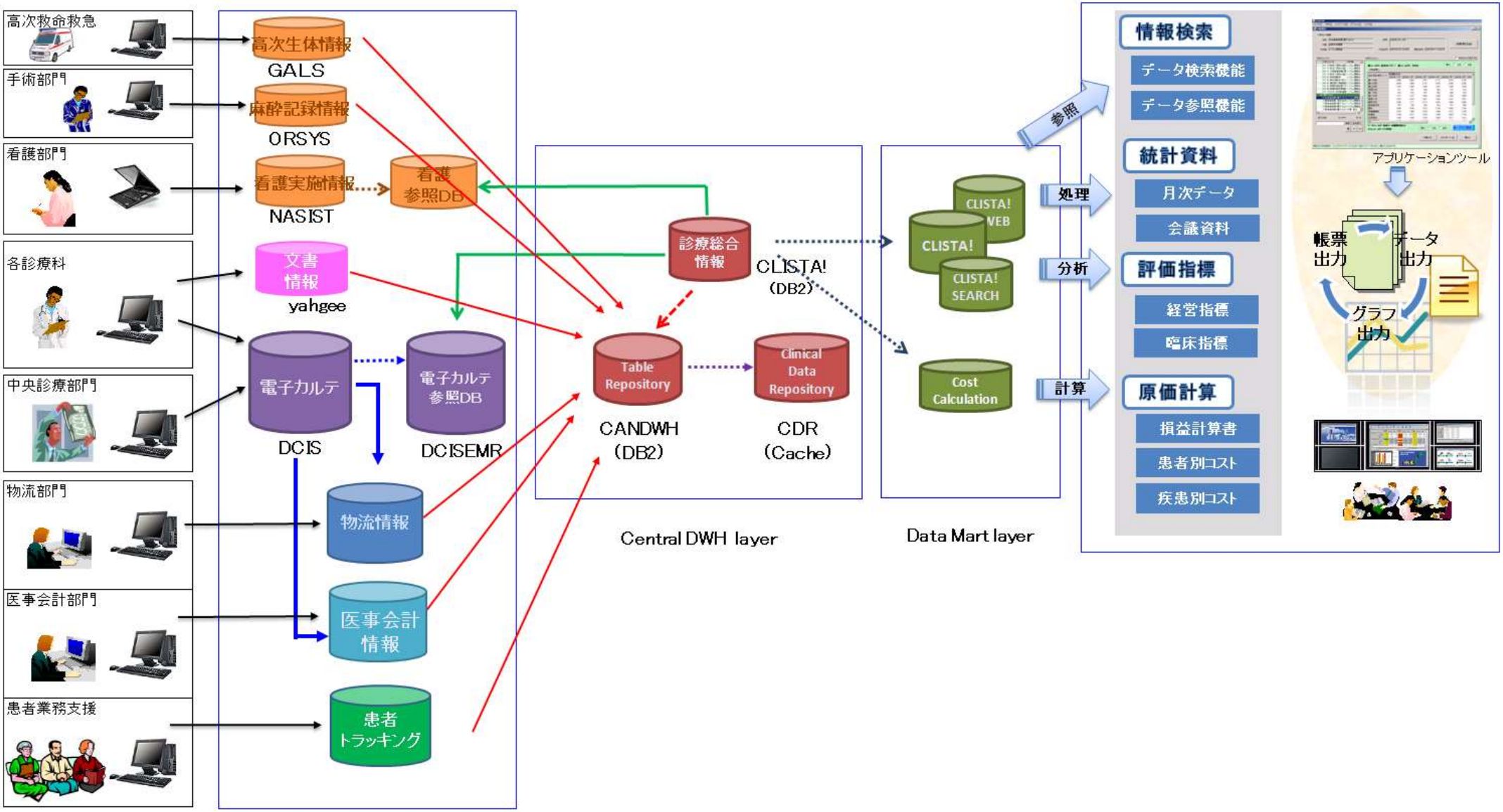
ロジックツリーの活用
- 目的：課題の解決
- 仮説：仮説の検証
- 手段：分析

データサイエンス
- 統計科学
- データマイニング
- プロセス改善

データマイニング
- 分類分析
- クラスタ分析
- アソシエーション分析

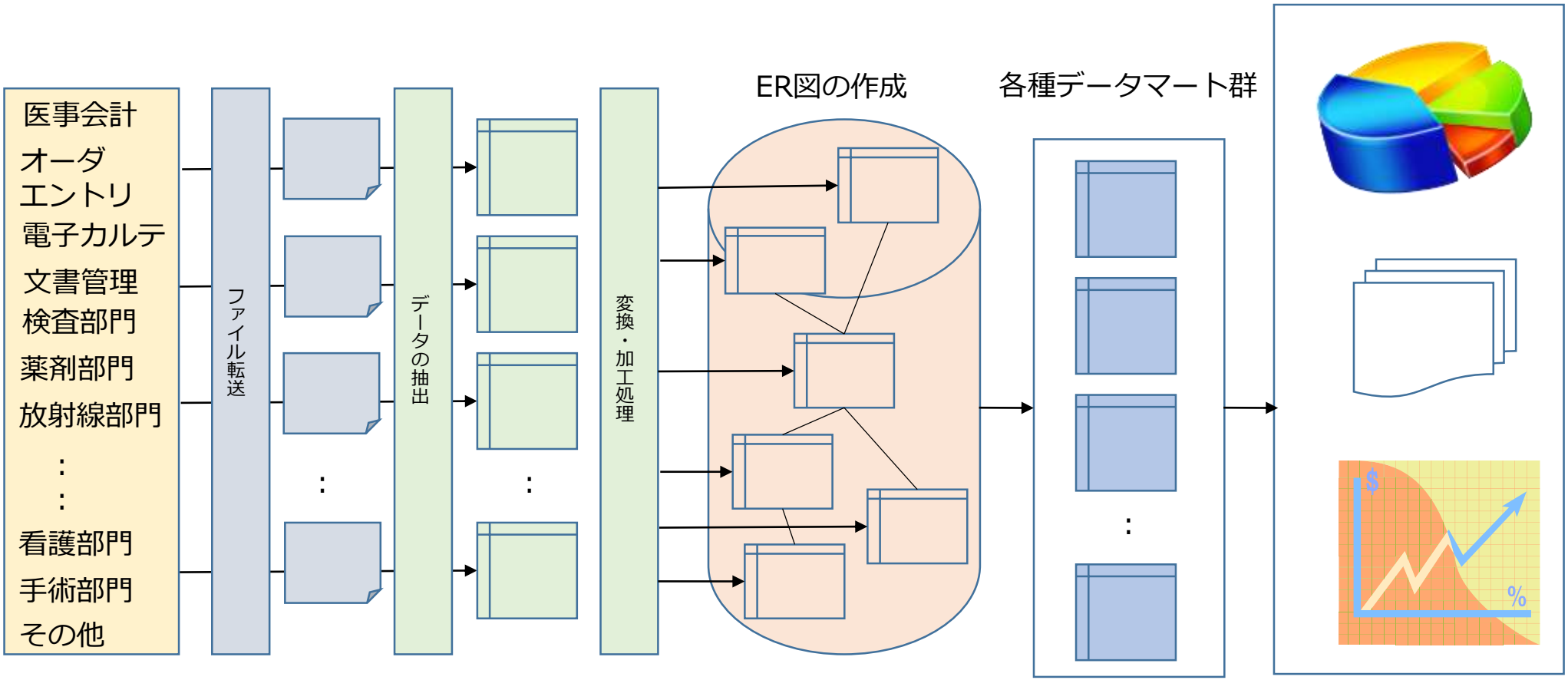
岐阜大学病院におけるDWH概要図

新DWH概念図



データを分析するシステムの概要

関連システム	ETL	DWH	データマート	BI・分析ツール
	E: データの抽出 T: 変換 L: 移行	データウェアハウス	使用頻度の多いデータセット群	OLAP, ピボットテーブル



電子カルテデータに基づく質指標から見た外来看護の変化

-外来看護必要度のデータと看護ケアデータの経年変化から見た報告-

【看護の変化】外来での手術オリエンテーション、入院の準備の説明など、外来看護師の役割の増加

【在院日数の短縮】在宅療養に必要な自己管理のための指導の増加

【患者の高齢化】↑

【入院中にセルフケア能力の確立が出来ない】看護計画立案件数の増加、看護計画説明件数の増加

【継続看護が必要な患者】↑

【医療依存度の高い患者の増加】ドレーン留置、在宅酸素、人工呼吸器装置、輸液ポンプ、心電図モニターの使用等、その他の医療機器、医療用品、循環コントロール、呼吸コントロール、呼吸ケアの必要性が増加

【転倒・転落防止の予防、移動時の介助、見守りの必要な外来患者】↑

安全ケア、移動ケアの増加

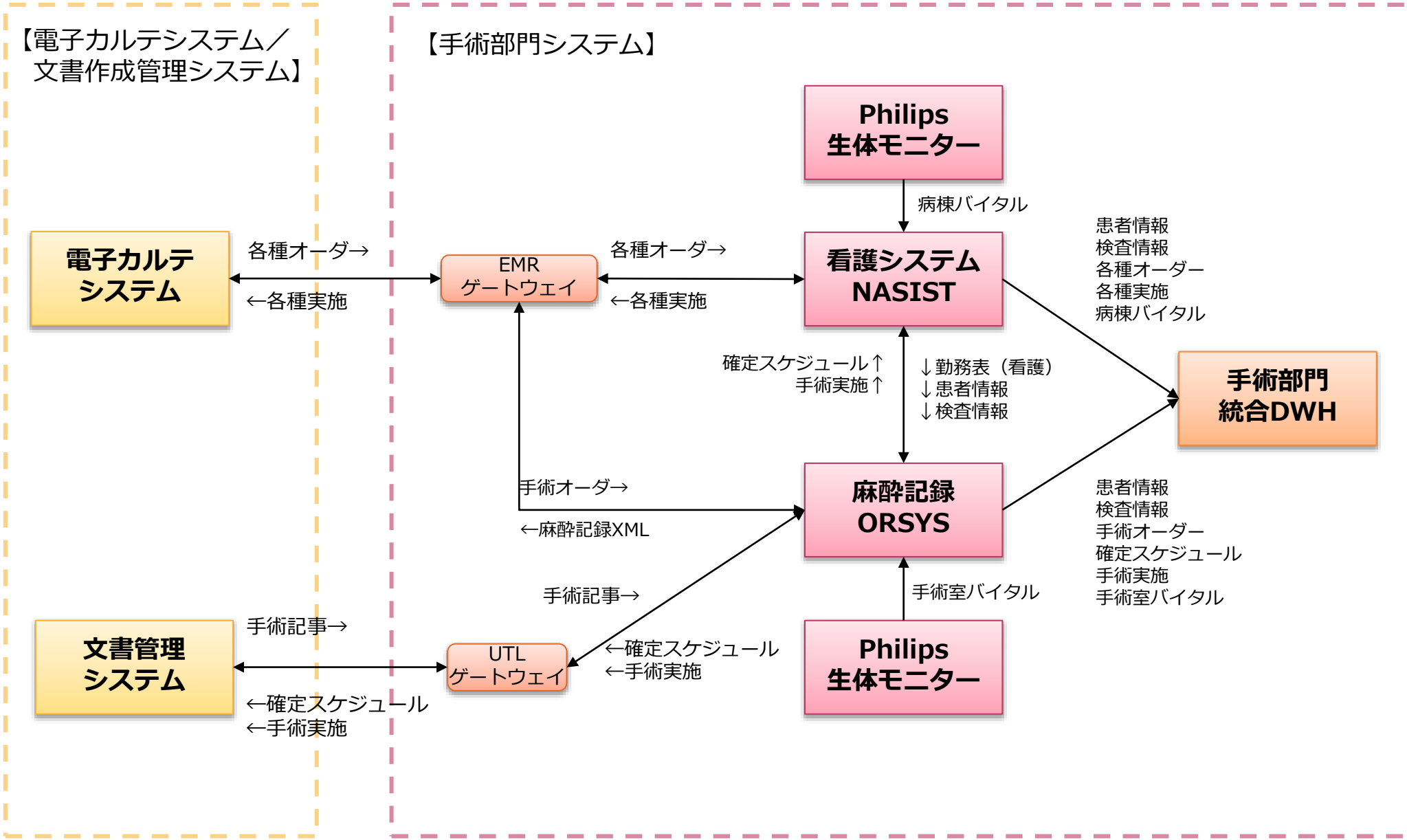
【看護師の予防意識】↑

意思決定支援 → 用語の定義の再確認

カンファレンスの実施 → 看護の質向上

看護計画の立案・説明 → 質評価の提示

手術・看護部門システムDWH概要図



基幹システム

電子カルテサーバー

- 患者基本情報
- 患者プロフィール
- 入退院情報
- 外来受診情報
- 病名
- 術式
- 処方
- 注射
- 検体検査結果
- 検査レポート
- 過去手術記録
- カルテ記事

- 文書通知
- 実施情報
- DPC情報

検査部門システム

放射線

生理

内視鏡

循環器

病理

- DICOM画像
- 心電図波形
- 脳波PDF

各診療科向けシステム

泌尿器科

眼科

耳鼻科

歯科

- 耳鼻科動画像
- DICOM画像

病棟システム

ベットサイドモニタシステム

- リアルタイム波形

手術部門

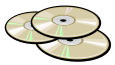
麻酔記録

術中画像キャプチャと映像配信

- 麻酔記録情報
- 手術記録情報
- DICOM画像
- ストリーミング動画

外部持ち込みデータの取込み

フイルム



CD/DVD

- DICOM画像

スキャンセンター（仮称）

- 部門システムサーバの仮想化
- 各種記録の統合一元管理
- チーム医療を支援するワークフローの導入
- ワークフローに沿った文書・画像の管理と参照機能の実現

画像・診断結果報告書・文書等統合管理システム

・電子カルテシステムで管理されている診療録や検査結果，ドキュメントビューで作成される大量の文書データに加え，院内の様々な部署で発生する画像や診断結果報告書等を統合管理。

・手術記録や看護記録，服薬指導記録や栄養管理記録，医療連携センターによる入退院支援記録などの蓄積と活用

・多職種間での情報共有とチーム医療

電子カルテ端末（外来・病棟 全端末）

診療文書データベース
ドキュメントビュー

- 診療情報提供書
- 診断書/証明書
- 同意書
- 入院診療計画書
- 術前サマリ
- 手術記録
- 退院サマリ
- DPC様式１

など、1500種類以上

統合検査ビューワ
クリニカルコックピット

- 全検査結果参照
- 全端末での生体モニタ参照
- 患者説明
- カンファレンス

生体支援センター

褥瘡対策支援

NST業務支援

SSI業務支援

栄養管理室

NCM業務支援

NST業務支援

緩和ケアチーム

緩和ケアチーム支援

がんセンター

がん登録

医事課

診断書/公費文書管理

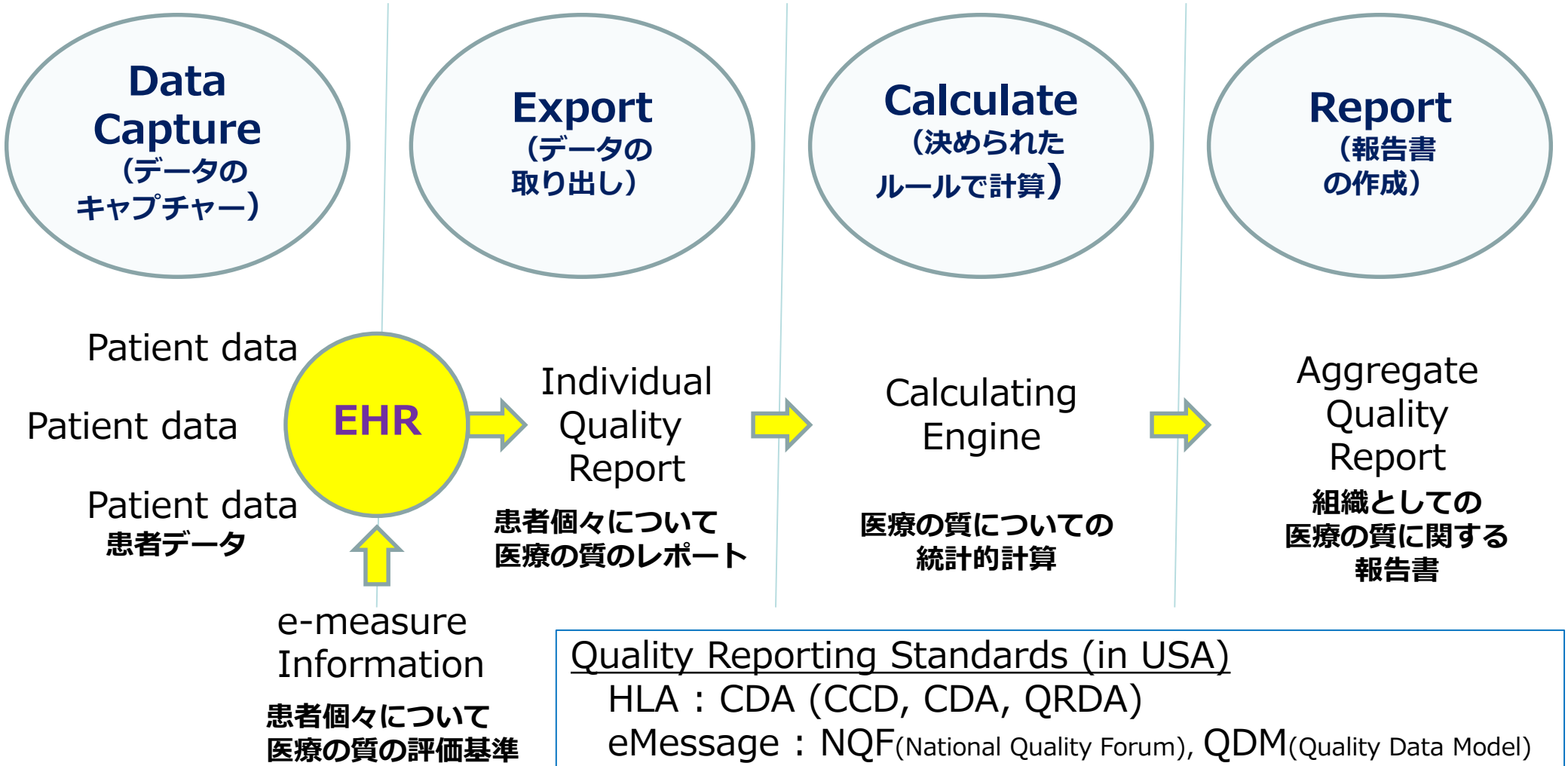
DPC確認

医療連携センター/手術部，等

データが活用できる病院になるためには

1）　データが取り出せるシステムが必要
　　①　取り出せるデータを普段から作っておく
　　②　DWH　と　ダッシュボード
2）　データを分析することができるシステムと人材が必要
　　①　データ分析ツール
　　②　データ分析ツールを使いこなせる人材（データサイエンティスト）
3）　データを分析し，分析結果を知識にする組織文化が必要
　　①　組織戦略としての取り組みが不可欠
　　②　継続性が必要であり，組織の文化にすることが重要
4）　分析結果に基づいた病院運営のマネジメントが必要
　　①　データを活用することにより，人材が育ち，ツールが充実する
　　②　「マネジメント」を正しく理解し，実践する組織力は強くなる
5）　そのような取り組みが，成功しなければならない
　　①　マネジメント運営が成功しなければ，職員はついてこない
　　②　「マネジメント」を理解している職員を増やすことが重要

診療データの効果的な活用と医療の質報告書へ



Quality Reporting Standards (in USA)

HLA：CDA (CCD, CDA, QRDA)

eMessage：NQF(National Quality Forum), QDM(Quality Data Model)

http://www.qualityforum.org/QualityDataModel.aspx

Meaningful Stage 2：

HL7 consolidated CDA implementation guide (www.hl7.org)

手術部門におけるValue Chain

１．手術室の効率的な運用

２．周術期の効果的な管理

３．病棟における適切な退院時期の決定

数値データが精度良く取得できる環境の整備

手術室の問題点

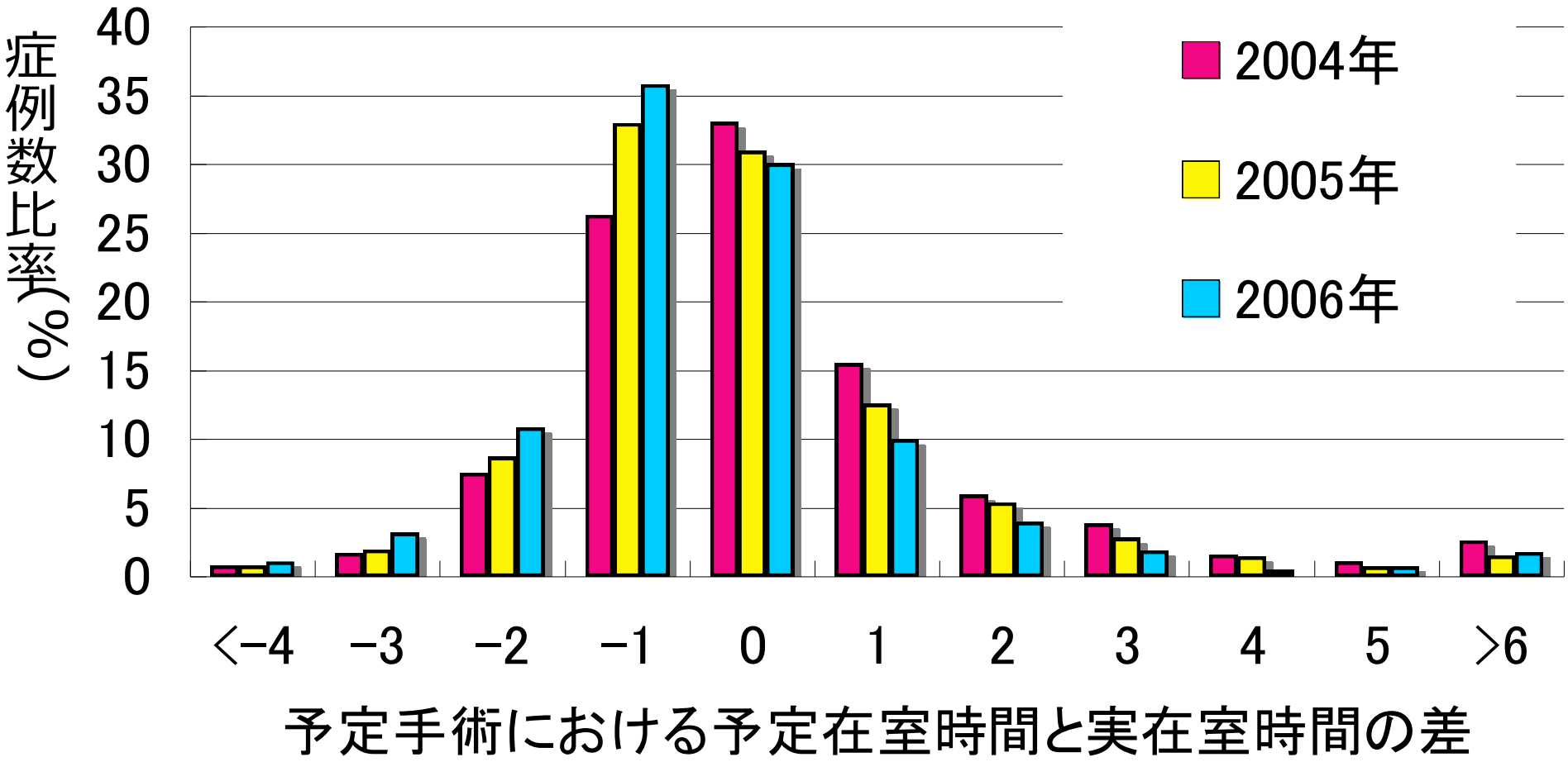
1．患者が手術室に在室中，責任区分が明確でない時間帯が存在する

2．予定手術時間を大きく超過する手術例が存在する

↓

手術室における時間区分の定義の厳格化

実在室時間と予定在室時間



年を追って予定在室時間は，実在室時間に収斂している。
しかし，なお大幅な時間超過手術も認められる。

時間区分の定義

麻酔準備完了時刻　：　麻酔導入およびライン確保等が終了し，体位確保、麻酔器の移動などを開始する時刻

手術予定在室時間　：　外科系診療科医師が手術申込時に申告する手術室在室予定時間

硬膜外麻酔時間　：　患者入室から硬膜外麻酔終了まで

麻酔準備時間　：　患者入室から麻酔準備終了時刻まで

手術準備時間　：　麻酔準備終了時刻から手術開始まで

診療科時間　：　麻酔準備終了から手術終了まで

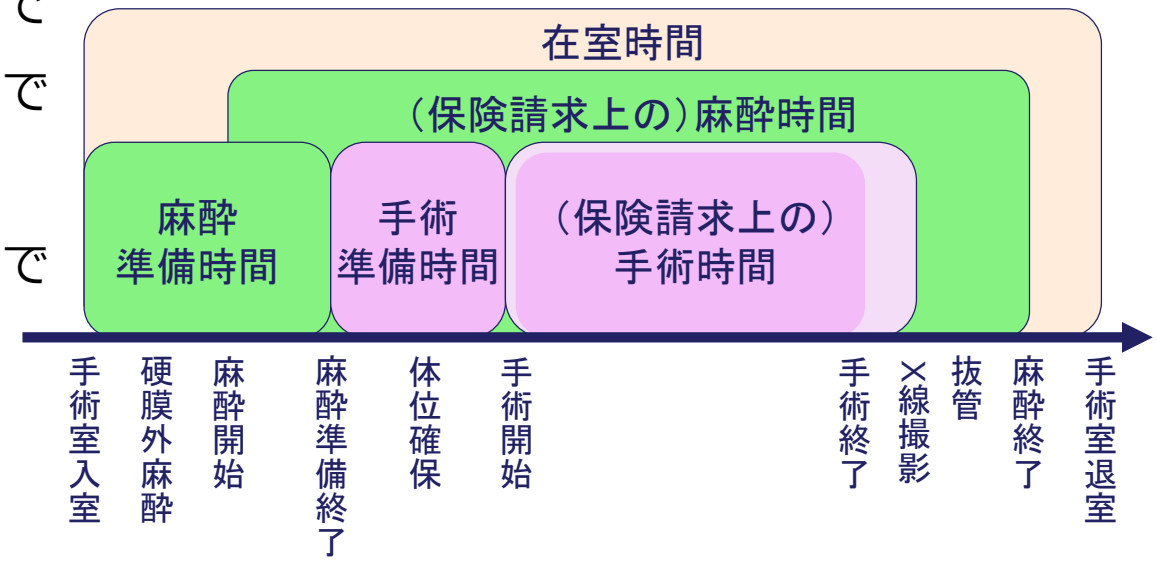
手術時間　：　手術開始から手術終了まで

麻酔時間　：　麻酔開始から麻酔終了まで

抜管時間　：　手術終了から抜管まで

覚醒時間　：　手術終了から麻酔終了まで

在室時間　：　手術室入室から退室まで



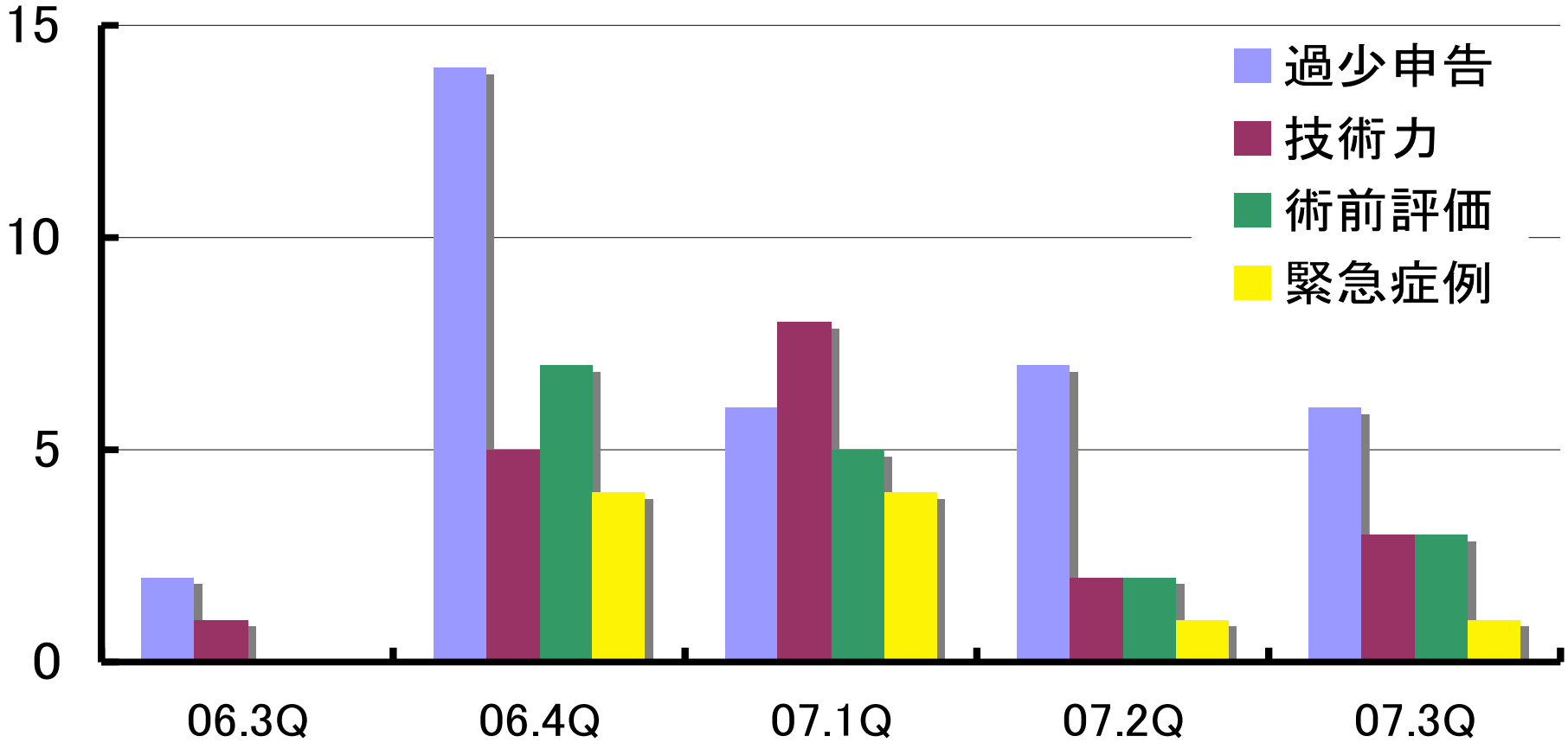
3時間ルール

- 目的
  - 正確な術前評価や綿密な準備により，予定手術時間を正確に申告し，手術の質を向上させる。また医療資源の効率利用も促す。
- 手術
  - 予定申告時間よりも3時間以上延長（短縮）した場合に，その理由を手術室に提出する
- 麻酔
  - 90分以上麻酔導入に要した場合に，理由を提出する

- 平成18年9月より施行した

３時間以上延長した原因の推移

当初は、予定在室時間の過少申告が最多であった。
しかし，いずれの理由でも件数は徐々に減少している。



再手術との相関

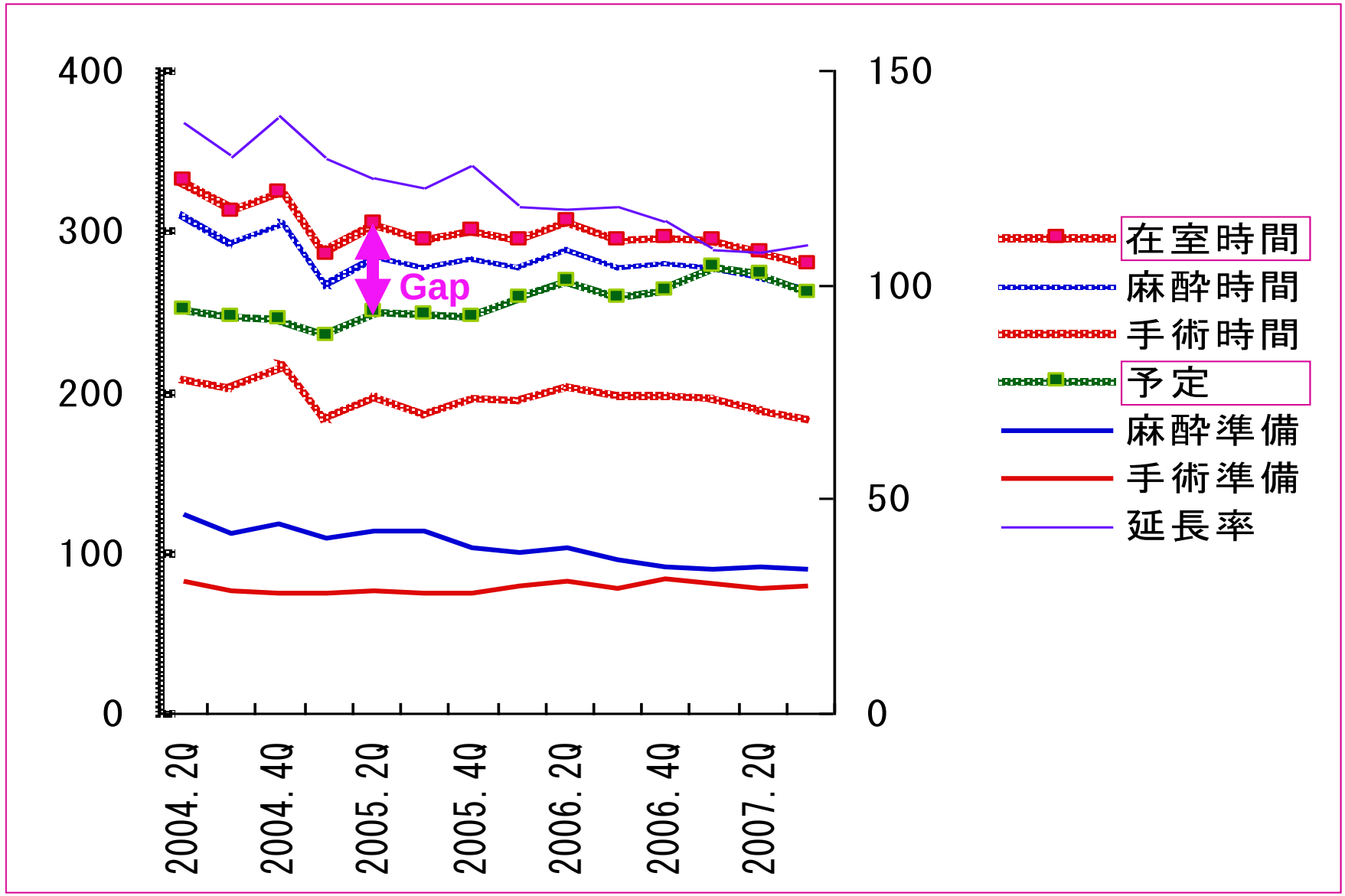
1週間以内に緊急手術を実施した麻酔科管理症例数

超過時間	＜3時間	＞3時間	計
再手術なし	7669	767	8436
1週間以内の再手術 (%)	172 (2.2%)	42* (5.5%)	214 (2.5%)
計	7841	809	8650

＊ $p<0.001$

3時間超過手術は，1週間以内に再手術となる率が有意に高い

３時間ルールの導入効果



Reproducible Research

データ収集

DWH

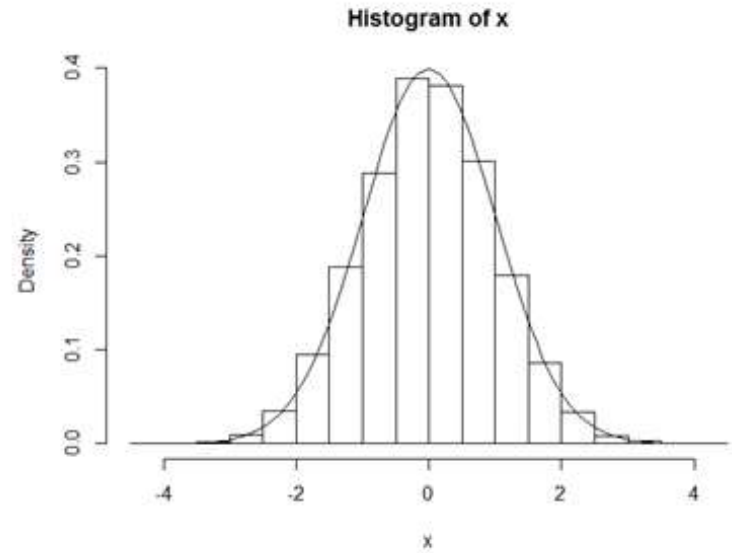
データ整形・
データ解析

画像・表・数値
データの作成

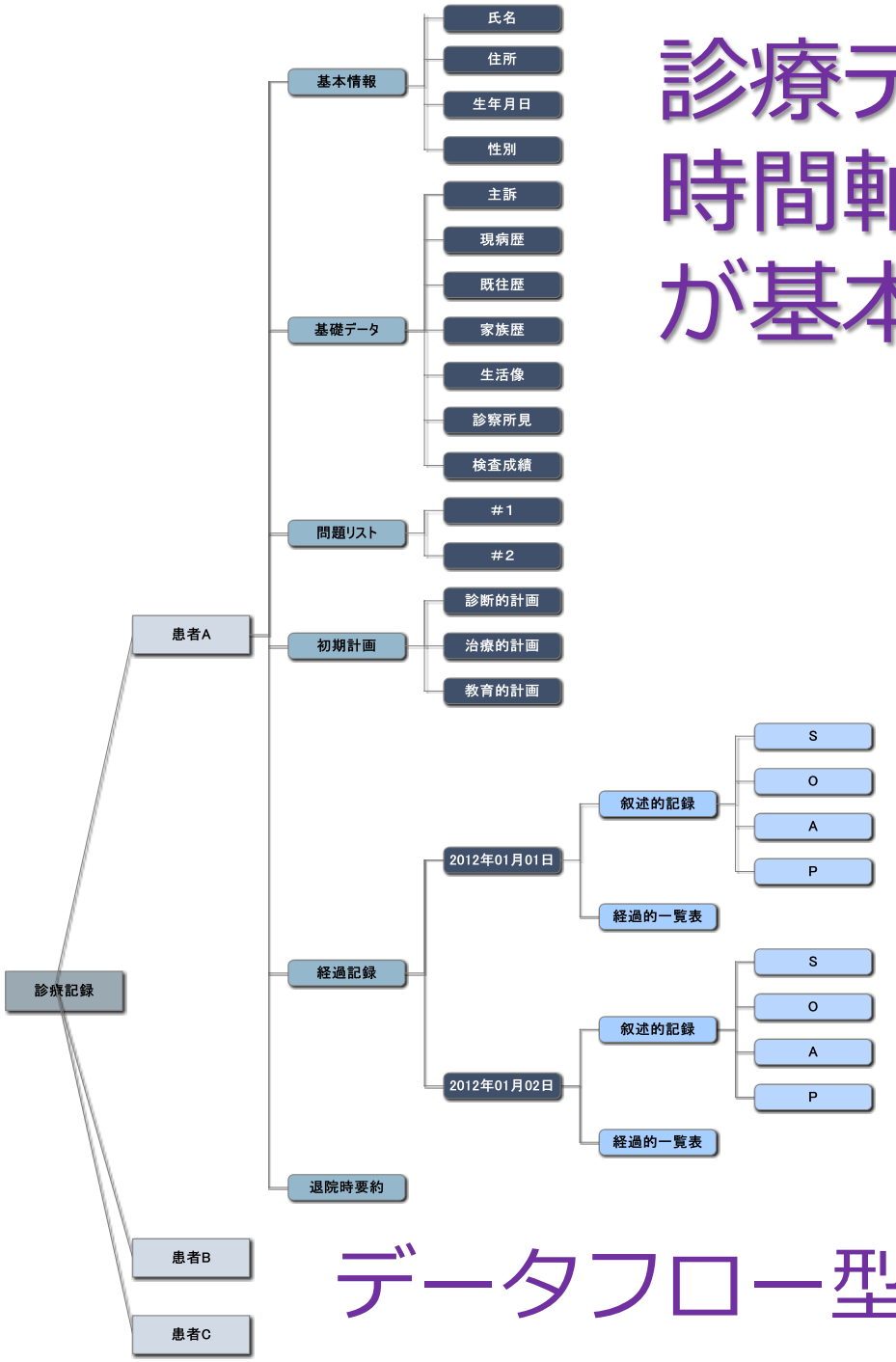
質レポート完成

- Example of Mark Down FIle
- Table:LOS
- Number of Data : 10,000
- X <- {Data}
- hist(x, probability=TRUE)
- curve(dnorm(x), add=TRUE)

ドキュメント作成



診療データの発生は時間軸に沿った木構造が基本



正規化

対応可能

Relational Database

データフロー型で処理ができないだろうか

Data Table

Nodes | Edges | Configuration | Add node | Add edge | Clear graph

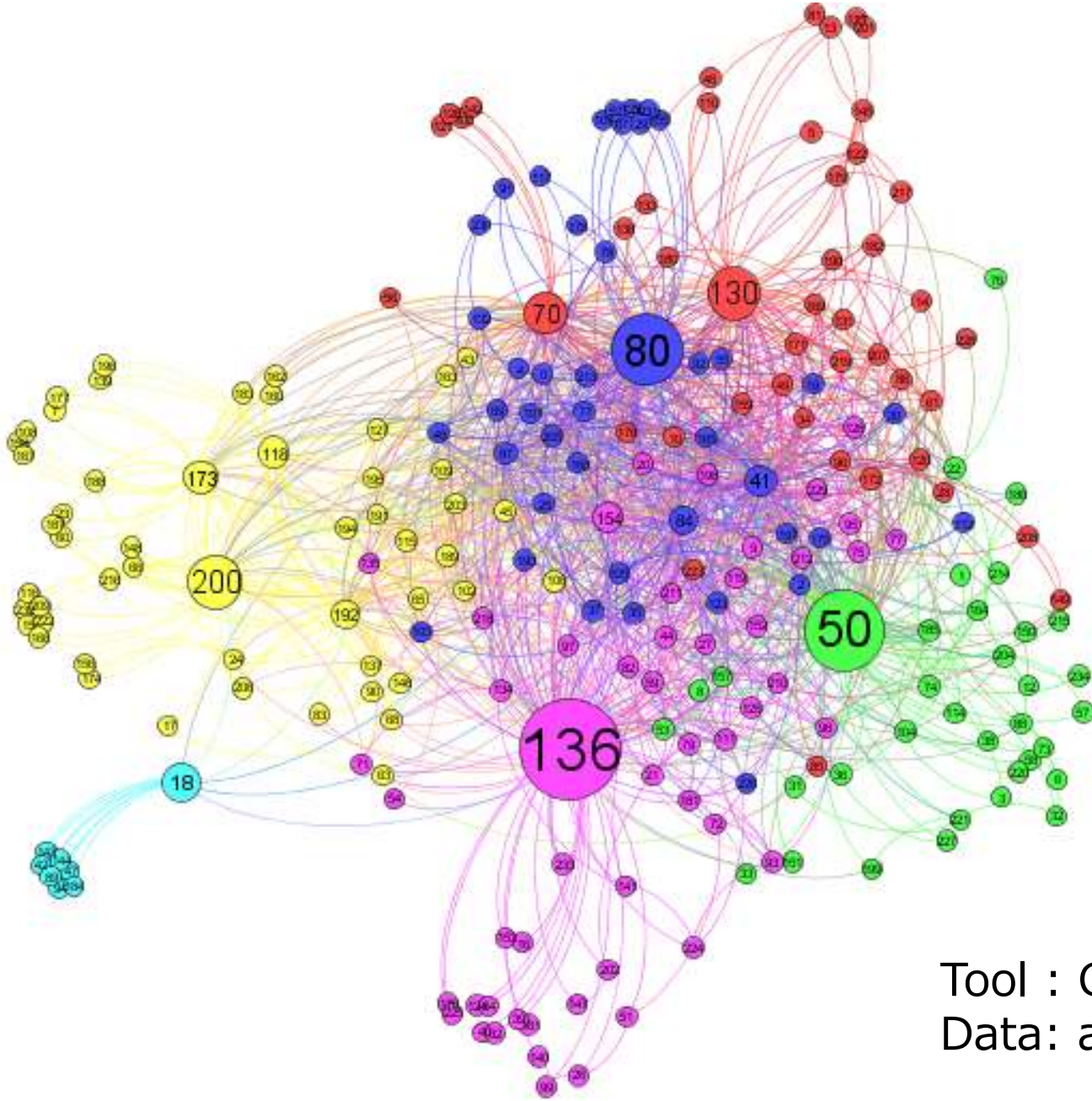
Nodes	Id	Label	tooltip
0	0	0	LIT(lngx=-92.224444,laty=34.729444)
1	1	1	CAK(lngx=-81.442222,laty=40.916111)
2	2	2	ALB(lngx=-73.8,laty=42.733333)
3	3	3	CHO(lngx=-78.45,laty=38.133333)
4	4	4	ABQ(lngx=-106.609167,laty=35.040278)
5	5	5	AEX(lngx=-92.548611,laty=31.3275)
6	6	6	APN(lngx=-83.55,laty=45.083333)
7	7	7	ACV(lngx=-124.1,laty=40.983333)
8	8	8	AVL(lngx=-82.541667,laty=35.436111)
9	9	9	AUS(lngx=-97.67,laty=30.194444)
10	10	10	GRB(lngx=-88.133333,laty=44.483333)
11	11	11	BWI(lngx=-76.668333,laty=39.175278)
12	12	12	BGR(lngx=-68.816667,laty=44.8)
13	13	13	PAH(lngx=-88.773056,laty=37.060278)
14	14	14	PFN(lngx=-85.683333,laty=30.216667)
15	15	15	BLI(lngx=-122.533333,laty=48.8)
16	16	16	BJI(lngx=-94.933333,laty=47.516667)
17	17	17	BTM(lngx=-112.4975,laty=45.954722)
18	18	18	BIL(lngx=-108.533333,laty=45.8)
19	19	19	BHM(lngx=-86.75,laty=33.566667)
20	20	20	FNT(lngx=-83.743611,laty=42.965556)
21	21	21	BIS(lngx=-100.75,laty=46.783333)

Data Table

Nodes | Edges | Configuration | Add node | Add edge | Clear graph

Source	Target	Type	Id
106	79	Undirected	1001
106	41	Undirected	1002
106	70	Undirected	1004
106	35	Undirected	1006
106	80	Undirected	1007
107	80	Undirected	1008
19	11	Undirected	101
109	105	Undirected	1010
109	49	Undirected	1011
109	87	Undirected	1012
109	84	Undirected	1013
109	106	Undirected	1014
109	48	Undirected	1015
109	50	Undirected	1016
109	20	Undirected	1017
109	41	Undirected	1019
109	45	Undirected	1021
109	70	Undirected	1022
109	200	Undirected	1024
109	80	Undirected	1025
109	24	Undirected	1027
110	13	Undirected	1028
110	70	Undirected	1029
110	80	Undirected	1030
111	84	Undirected	1034
111	50	Undirected	1035
111	44	Undirected	1036

Meaningful Use of Big data



Tool : Gephi 0.8
Data: airlines.graphml

SDMとは

Semantic Data Modeling

意味を考慮したデータモデリング

セマンティックス ➔ 意味
データモデル ➔ データの構造

現状のDWHの問題点

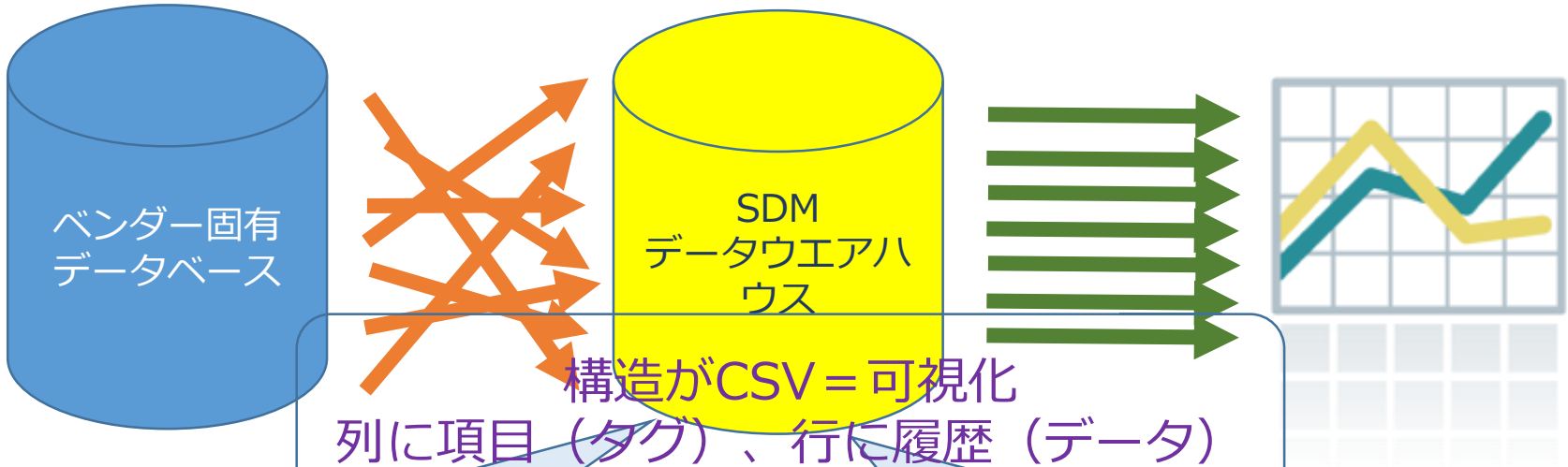
- 複数のシステムのデータが統合されていないため、それぞれのデータ抽出の負荷が大きい
- 項目名と定義、構造が統一されていないため、抽出したデータが関係づけられない
- テーブルごとに粒度と頻度がまちまちであるため、指標を算出することができない
- 検索ツール側で加工されており、特定の検索ツールを用いないとデータを抽出することができないため、非定型の検索ができない
- 独立したデータベースとサーバーを持っていないため、抽出元のシステムが停止すると検索ができない
- モデリングされたデータ構造を持っていないため、検索に時間がかかるうえ全文検索ができない
- セマンティックスを意識していないため、意味が違う項目が並んでいて、ユーザーが選択と並べ替えをする必要がある

SDMの特徴：可視化データベース

日付など、定義が必要なものは定義という項目を作り、データに定義の内容を記入する。
（例）検査日定義：検体採取日

項目名（タグ）と内容を一致させる
（例）項目１→血糖値

アプリケーション層　ETL　データウエアハウス層　ビジネスインテリジェンス層

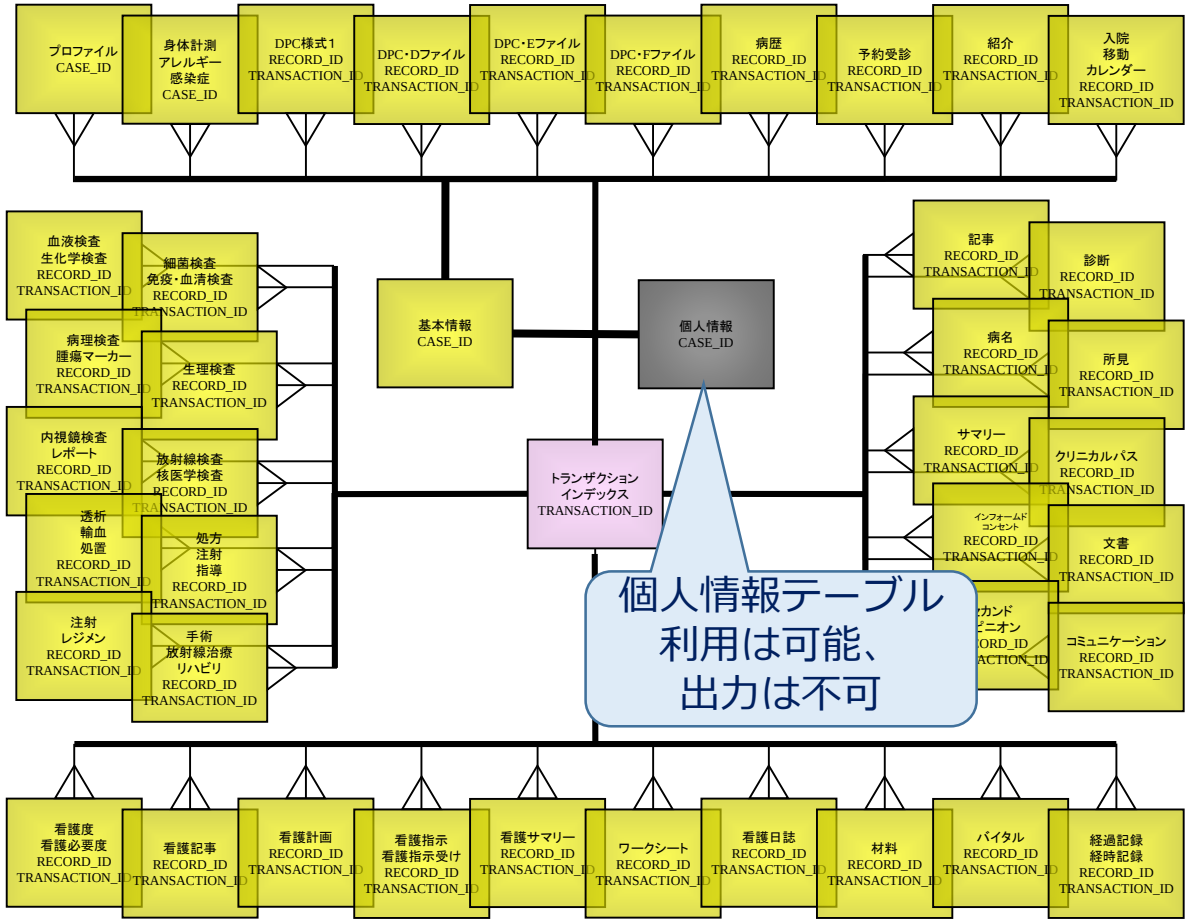


構造がCSV＝可視化
列に項目（タグ）、行に履歴（データ）

コードデータは文字データへマスター変換
（例）病名コード：C349　→　肺癌

2値、１,０を、＋,－へ変換
（例）そばアレルギー：＋

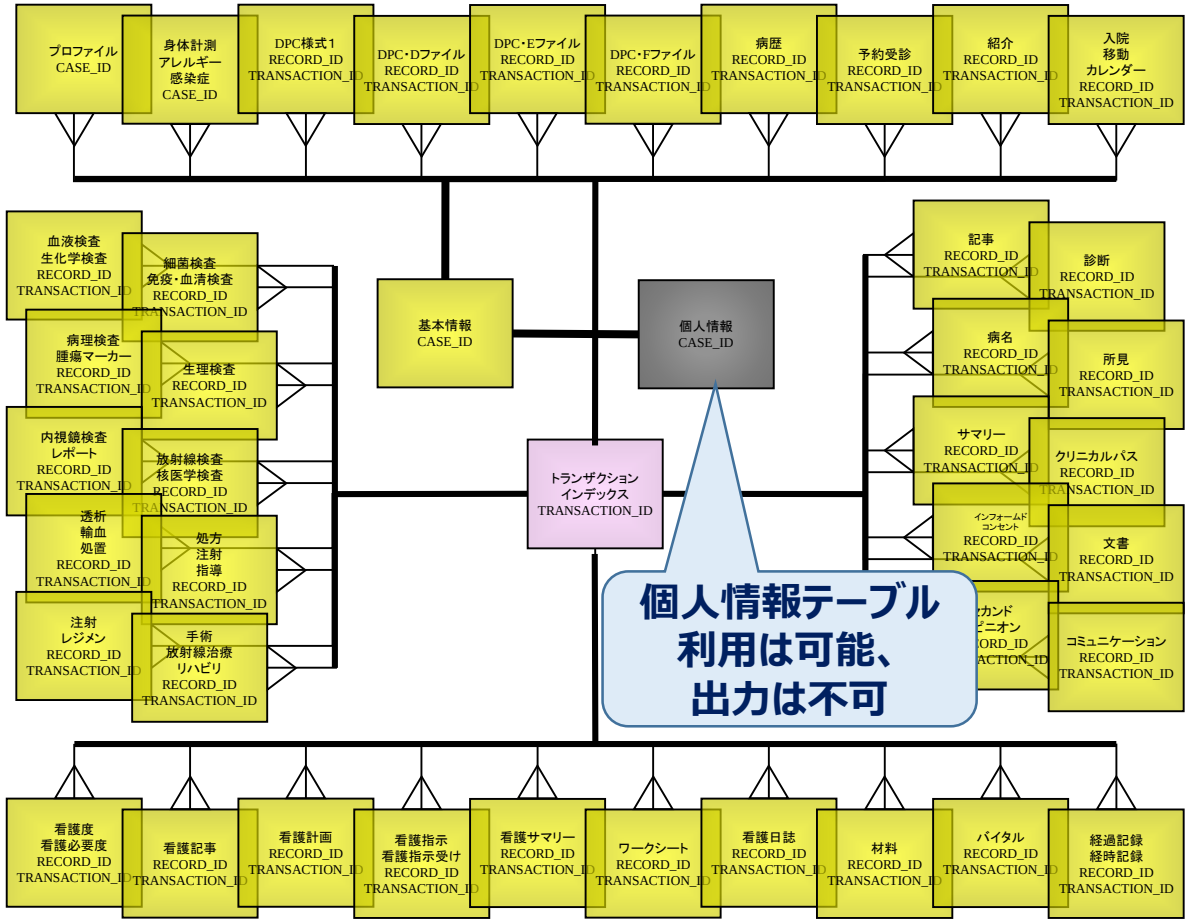
データモデリングにより高速化



検索速度向上のための設計ルール

- 全体を１つのER図で表現する
  - トランザクションインデックスは各テーブルのサブセット
- 階層を減らす
  - 処方のレシピ（RP)は薬品のテーブルへ記載してRPテーブルを削減
  - マスターは変換した結果を記載して削減
- 列数の多いテーブルは垂直分割をする
  - プロフィールなどはグループで分割（アレルギー、感染症など）
- 行数の多いテーブルは水平分割する
  - 記事、文書系など、意味が違うものを別テーブルにする
- カレンダーテーブル（日次）を作成する
  - 周術期における相対日、入院期間における相対日などの計算を簡素化する
- キー日時(KEY_DATE)を各テーブルに設置し、高速インデックスに設定する
  - 検索は期間を絞り込むのが一般的であるため
- 集計の軸になる項目はあらかじめインデックスを付与する
  - 症例ID、病棟、診療科など集計の単位になる項目のインデックス化

データモデリングにより高速化



テーブル定義のルール

- 項目（タグ）名はデータの内容を表す表現とする
  - 例：○　GENERAL_ANESTHESIA 、
  - 　　×　MASUI_1
- コードの場合マスター変換の結果とともに記録する
  - 例：DISEASE_CODE C349　DISEASE 肺癌
- フラグの場合テキストに変換する
  - 例：○　CLINCAL_OUTCOME 軽快
  - 　　×　5
- 2値の場合＋－で表現する
  - 例：○　FOOD_ALLERGY　＋
  - 　　×　食物アレルギー 1
- あいまいな項目名の場合、その定義をデータとして記録する
  - 例：　KEY_DATE_DEF　検体採取日
- 各レコードには必ずオーナーを記録する
  - 2次利用における稟議規定には必須
- キーは抽出元を特定するキー、テーブルの行を特定するキー、上位階層のテーブルのキーを記録する
  - 全テーブルのリレーションが表現できるモデルの実現のため

SDMによる課題解決

- 項目名と定義、構造が統一されているため、複数のシステムからのデータを統合でき、ユーザー負担を劇的に軽減できる
- すべてのテーブルが一つのER図で関連付けられているので、検索ツールに依存することがなく、非定型検索が可能となる
- 独立したデータベースとサーバーを持つため、カルテ停止時においても診療情報を検索することができる
- データモデリングの手法を用いているため、データベースの能力に頼らずに、高速な全文検索をすることができる
- 項目名と定義、構造が統一されていて検索ツールに依存しないため、ユーザー間で分析手法の交換が可能となる
- セマンティックスモデルを用いているため、必要な項目だけが簡単に抽出できる

SDMコンソーシアムの役割

- SDM-DWH利用者の情報交換を推進
  - セミナーの主催
  - イベントへの参加
  - ユーザー会の主催
  - 会員への情報発信
  - 知識データベースの作成と運営
- SDM-DWHを用いた情報利用技術者の育成
  - 勉強会の主催
  - 研修の設置と運営
  - 資格認定
  - プロジェクト斡旋