

## ライフサイエンス分野の統合データベース整備事業 『統合医科学データベース構築方式の開発』



東京医科歯科大学 田中 博

平成21年1月21日

# 平成20年度の成果

## ①疾患・臨床医科学データベースに特化した統合技術開発に関する研究

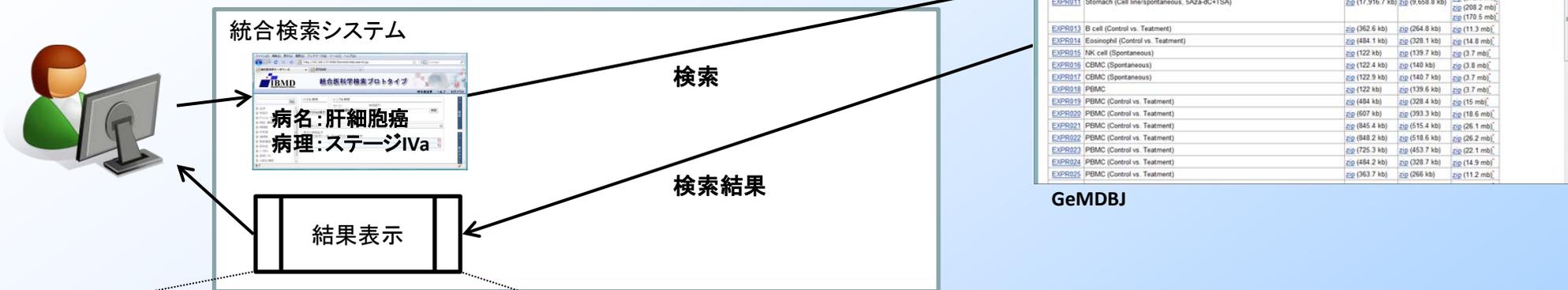
1. 肝細胞癌、大腸癌に特化し、ターミノロジー、シソーラスなどを収集整理。自動作業化のためのハードルを明確化。
2. 国際的で広範囲な倫理規定の基本調査を行い、草案に向けたステップ項目を策定し、各項目ごとの必要案を検討した。現在は、具体的な基準内容の検討。
3. 検索エンジン設計の要素化、各要素技術のISO, WHOの場での国際標準化の推進。

## ②疾患データベースの高度化と実証的統合に関する研究

プロトタイプシステムの機能向上に向けて、作業中である。

1. 症例を増加し、癌200症例、神経疾患400症例を統合化。
2. GEMDBJとの統合に向けた連携方式の策定。
3. 検索GUIについては、より直感的な操作を可能とする新たなパネル方式を開発。

## 国立がんセンター研究所 GeMDBJとの連携



Genome Medicine Database of Japan (GeMDBJ) - Mozilla Firefox

https://gemdbj.nbio.go.jp/dgdb/DownloadSite.do

発現解析

実験ID	実験名	GeMDBJ形式	MAGE-ML形式	celファイル
EXP001	Normal tissue series.	zip (6,397.2 kb)	zip (3,572.7 kb)	zip (204.5 mb)
EXP002	Kidney (Non-cancerous, Adenocarcinoma)	zip (3,696.8 kb)	zip (2,037 kb)	zip (158.8 mb)
EXP003	Kidney (Cell line)	zip (406.4 kb)	zip (264.3 kb)	zip (17.6 mb)
EXP004	Stomach (Cell line)	zip (1,206 kb)	zip (687.3 kb)	zip (49.3 mb)
EXP005	Blood, Leukocyte (Normal, Acute myeloid leukemia)	zip (3,945.5 kb)	zip (2,159.8 kb)	zip (159.2 mb)
EXP006	Blood, Cells Panel (Allthia)	zip (1,208.8 kb)	zip (703.3 kb)	zip (37.2 mb)
EXP007	Blood, PBMC (LPS/CyC stimulation time course)	zip (1,277.5 kb)	zip (796 kb)	zip (58 mb)
EXP008	Blood vessel, Cell line, Umbilical vein endothelial cell, HUVEC	zip (88.4 kb)	zip (63.7 kb)	zip (3.1 mb)
EXP009	Brain, Brain cortex (Amyotrophic lateral sclerosis as control, Alzheimer's disease)	zip (4,378.6 kb)	zip (2,537.6 kb)	zip (169.1 mb)
EXP010	Blood, PBMC (LPS stimulation time course)	zip (1,230.5 kb)	zip (712.5 kb)	zip (40.8 mb)
EXP011	Stomach (Cell line/spontaneous, S4za-dC+TSA)	zip (11,916.7 kb)	zip (9,658.8 kb)	zip (143.9 mb)
EXP012	B cell (Control vs. Treatment)	zip (362.6 kb)	zip (264.8 kb)	zip (11.3 mb)
EXP014	Eosinophil (Control vs. Treatment)	zip (484.1 kb)	zip (328.1 kb)	zip (14.8 mb)
EXP015	NK cell (Spontaneous)	zip (122 kb)	zip (139.7 kb)	zip (3.7 mb)
EXP016	GMAC (Spontaneous)	zip (122.4 kb)	zip (140 kb)	zip (3.8 mb)
EXP017	GMAC (Spontaneous)	zip (122.9 kb)	zip (140.7 kb)	zip (3.7 mb)
EXP018	PBMC	zip (122 kb)	zip (139.6 kb)	zip (3.7 mb)
EXP019	PBMC (Control vs. Treatment)	zip (484 kb)	zip (328.4 kb)	zip (15 mb)
EXP020	PBMC (Control vs. Treatment)	zip (607 kb)	zip (393.3 kb)	zip (18.6 mb)
EXP021	PBMC (Control vs. Treatment)	zip (484.2 kb)	zip (315.4 kb)	zip (26.1 mb)
EXP022	PBMC (Control vs. Treatment)	zip (482.2 kb)	zip (318.6 kb)	zip (26.2 mb)
EXP023	PBMC (Control vs. Treatment)	zip (726.3 kb)	zip (453.7 kb)	zip (22.1 mb)
EXP024	PBMC (Control vs. Treatment)	zip (484.2 kb)	zip (328.7 kb)	zip (14.9 mb)
EXP025	PBMC (Control vs. Treatment)	zip (263.7 kb)	zip (266 kb)	zip (11.2 mb)

GeMDBJ

### 結果画面(サマリー)

データベース名	病名	症例ID	0/1%	分子情報	詳細
GenDBJ	肝細胞癌	1D			
GenDBJ	肝細胞癌	2D			
GenDBJ	肝細胞癌	3D			
GenDBJ	肝細胞癌	4D			
GenDBJ	肝細胞癌	5D			
GenDBJ	肝細胞癌	6D			

### 結果画面(分子情報)

遺伝子名 Gene1 解析

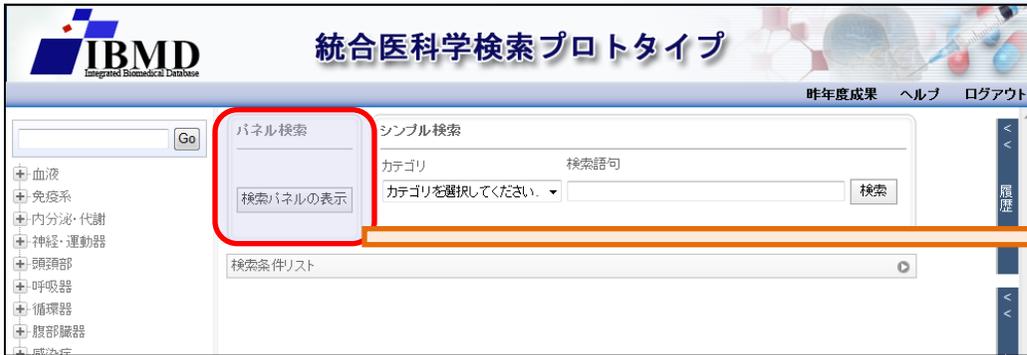
症例ID	Gene1
1	2
2	2.1
3	-
4	4.4
5	5.8
6	7.2

遺伝子発現量

クラスタリング結果

クリックで分子情報解析開始

分子情報クリックで画面遷移

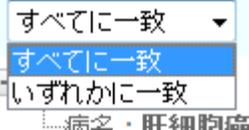


## パネル検索GUI

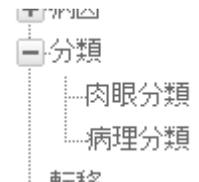
検索語句の入力は  
カテゴリ名をクリック



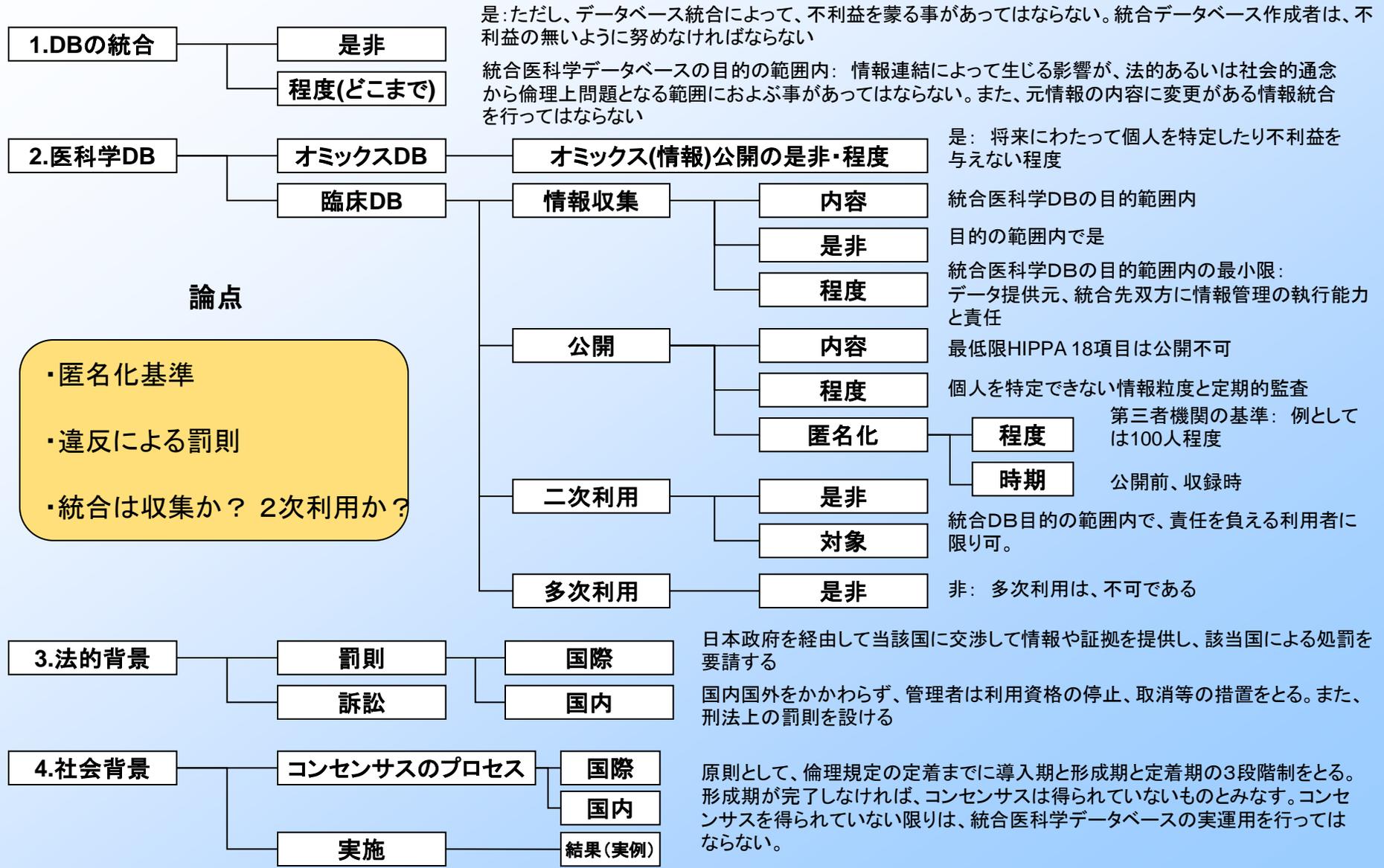
結合条件の選択



カテゴリの階層を表すツリー構造を見ながら  
検索条件が入力可能



# 統合医科学DB 倫理規定草案検討の論点概要



# 平成21年度の目標

## ①疾患・臨床医科学データベースに特化した統合技術開発に関する研究

1. 統合化要素技術として疾患オントロジーの整備
2. プロトタイプシステムの評価に基づいた、要素技術の問題点の把握
3. 中核機関を含め、登録研究者や一般公開に向けた統合DB倫理規定の具体的確立の検討

## ②疾患データベースの高度化と実証的統合に関する研究

1. セマンティック検索エンジンの高度化を行い、プロトタイプシステムを改良し公開
2. 要件定義に基づいたGeMDBJ、パーキンソン病データベースとの高度な連携方式の開発
3. 分子から臨床生活情報までの網羅的疾患情報の組織化方法の確立
4. 新規データベースの統合
  - － 癌研究会癌研究所のデータベースとの具体的な統合を目指す

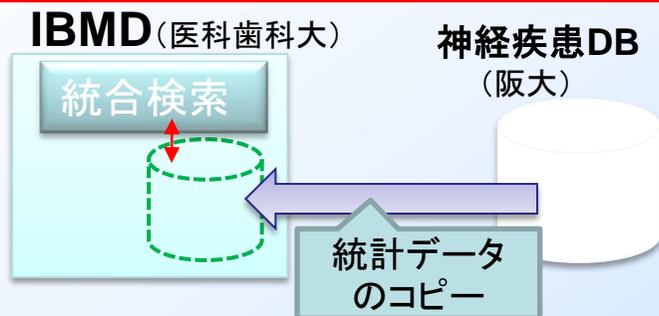
# 大阪大学

	H20年度成果	H21年計画
①疾患・臨床医科学データベースに特化した統合技術開発に関する研究	多階層の総合医科学データベースを統合構築するために必要な統合技術(疾患DB情報モデル、オントロジー、セマンティックなど)の要件定義	プロトタイプシステムの試作と問題点の検討 統合化要素技術として疾患オントロジーの整備とプロトタイプとして開発したセマンティック検索エンジンの高度化
②疾患データベースの高度化と実証的統合に関する研究	パーキンソン病データベースについて、症例に基づいた検索などの具体的要求と課題の検討	統合検索システムとの連携を行うとともに、倫理的問題の解決し、登録研究者へ公開する

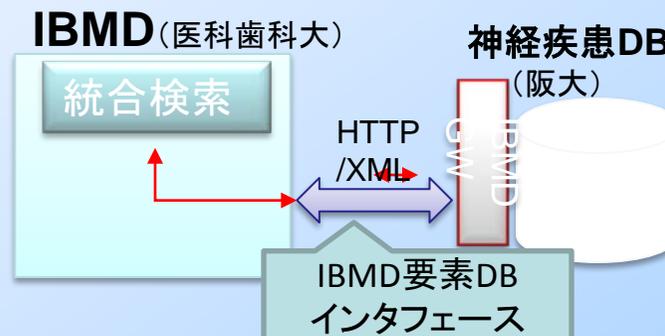
# H20年度成果概要(大阪大学)

## ①疾患・臨床医科学データベースに特化した統合技術開発に関する研究

- ・医科歯科大で作成した統合の要件定義(IBMD要素DBインタフェース)に沿って、パーキンソン病DB統合のための要件定義(IBMD GW)を作成



H19年度に試作した統合方式  
(暫定的なプロトタイプ)



H20年度に作成した統合の要件定義  
(H21年度に実装予定)

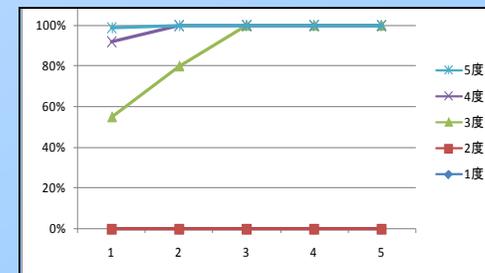
## ②疾患データベースの高度化と実証的統合に関する研究

- ・DBの公開症例数を100症例(H19年度時点)から400症例まで増加させる。
  - ・DBの経年情報表示に「グラフ表示機能」を追加する。
- ※共にH20年度内に公開予定。

重症度の割合

重症度	1年目	2年目	3年目	4年目	5年目
1度	0% (0/27人)	0% (0/5人)	0% (0/2人)	0% (0/2人)	0% (0/1人)
2度	0% (0/27人)	0% (0/5人)	0% (0/2人)	0% (0/2人)	0% (0/1人)
3度	55% (15/27人)	80% (4/5人)	100% (2/2人)	100% (2/2人)	100% (1/1人)
4度	37% (10/27人)	20% (1/5人)	0% (0/2人)	0% (0/2人)	0% (0/1人)
5度	7% (2/27人)	0% (0/5人)	0% (0/2人)	0% (0/2人)	0% (0/1人)

現状の経年情報表示例

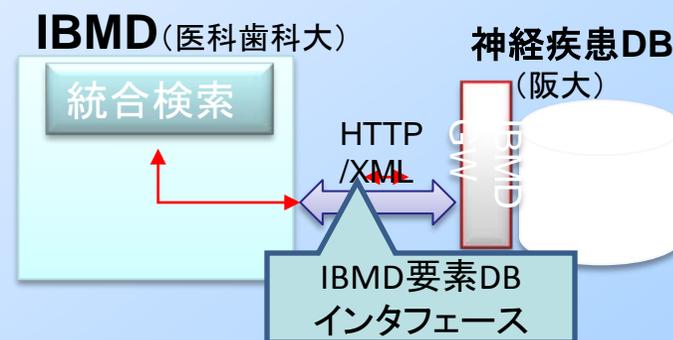


グラフ表示(イメージ図)

## H21年度計画概要(大阪大学)

### ①疾患・臨床医科学データベースに特化した統合技術開発に関する研究

- ・H20年度に作成したパーキンソン病DB統合のための要件定義(IBMD GW)に基づいてプロトタイプシステムを試作し、問題点を検討する。
- ・統合化要素技術として、疾患オントロジーの整備とプロトタイプとして開発したセマンティック検索エンジンの高度化を行う。



H20年度に作成した要件定義に基づいた統合化

### ②疾患データベースの高度化と実証的統合に関する研究

- ・実施項目①の要件定義に基づき、これまで開発した神経疾患DBを統合検索システムと連携させる。
- ・H20年度までに公開した症例(非連結匿名化・統計情報のみ)について、これまでに実施してきた「統合DB向け公開の為の患者への再同意」、「公開倫理案」の結果に基づいて、倫理的問題を解決し、登録研究者へ公開する。