

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5月20日

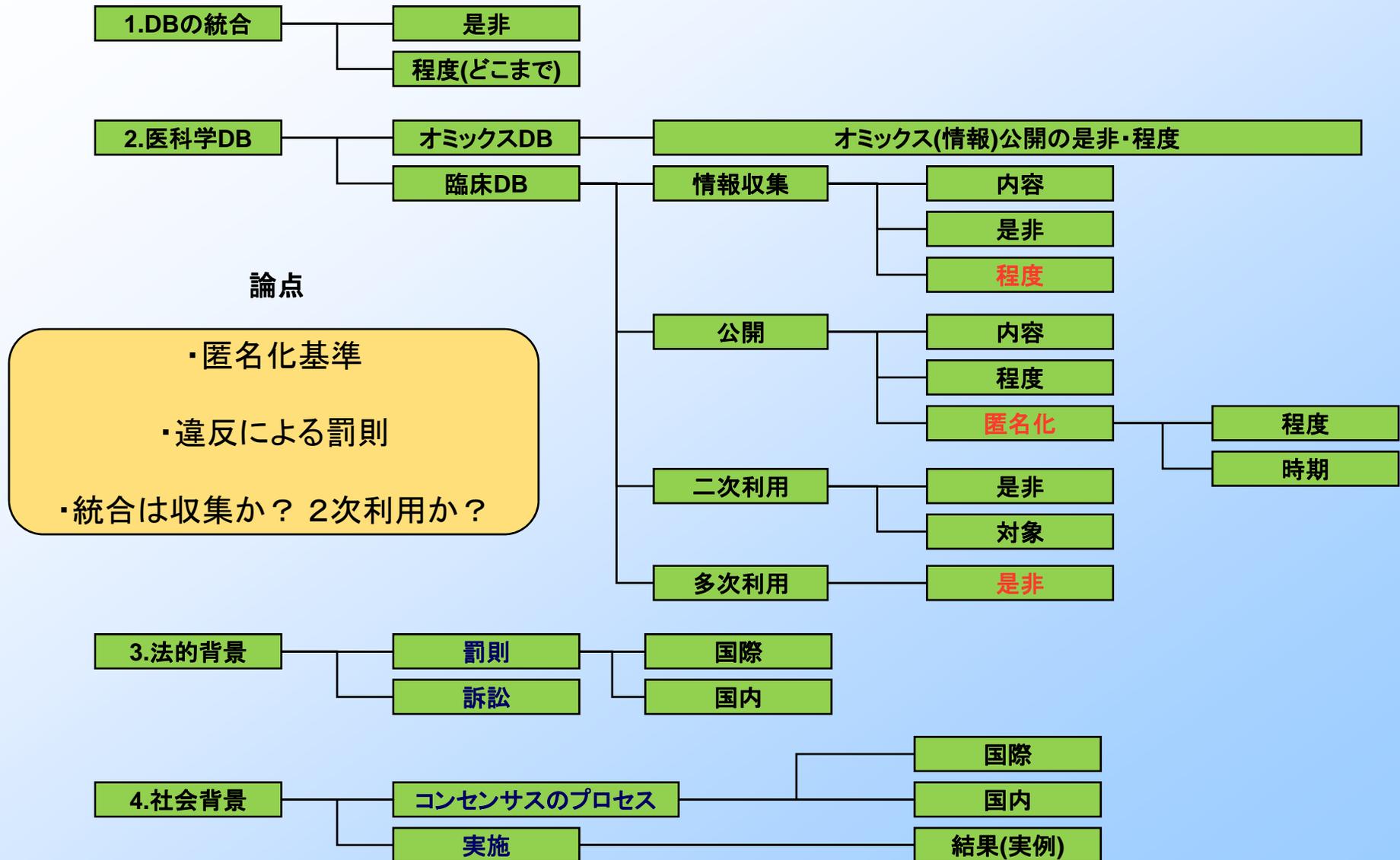
# 平成20年度の成果

## 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

昨年度の調査、検討結果に基づき、統合技術の要件定義を行った。

1. 国際的で広範囲な倫理規定の基本調査を行い、倫理規定草案を策定した。
2. 倫理規程案を元に、システム外部設計書を作成した。
3. 外部設計書を元に、内部設計書を作成した。
4. 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を行った。

# 統合医科学DB 倫理規定草案検討の論点概要



# 倫理規定として予定される項目

## ・確保すべき3要素

正確性 ⇒ 正確かつ最新の内容に保つ義務

安全性 ⇒ 患者等個人の人権の安全管理のための措置  
取扱い者に対して監督する義務

透明性 ⇒ 取得時の利用目的の通知 本人からの開示・訂正請求に応じる義務

## ・情報を連結する事で匿名性が損なわれる可能性を構造的にゼロにする

⇒ 連結不可能匿名化

## ・利用目的による利用者の制限

⇒ 医療従事者でも利用者登録を必要とする、監査証跡(アクセス記録等)のとれる公開体制

## ・個人を特定して悪用した場合に関する罰則

⇒ 安全性の基準としての匿名性の定量的評価

## ・情報取扱いについての対象者への説明と、これに基づく同意の取得について

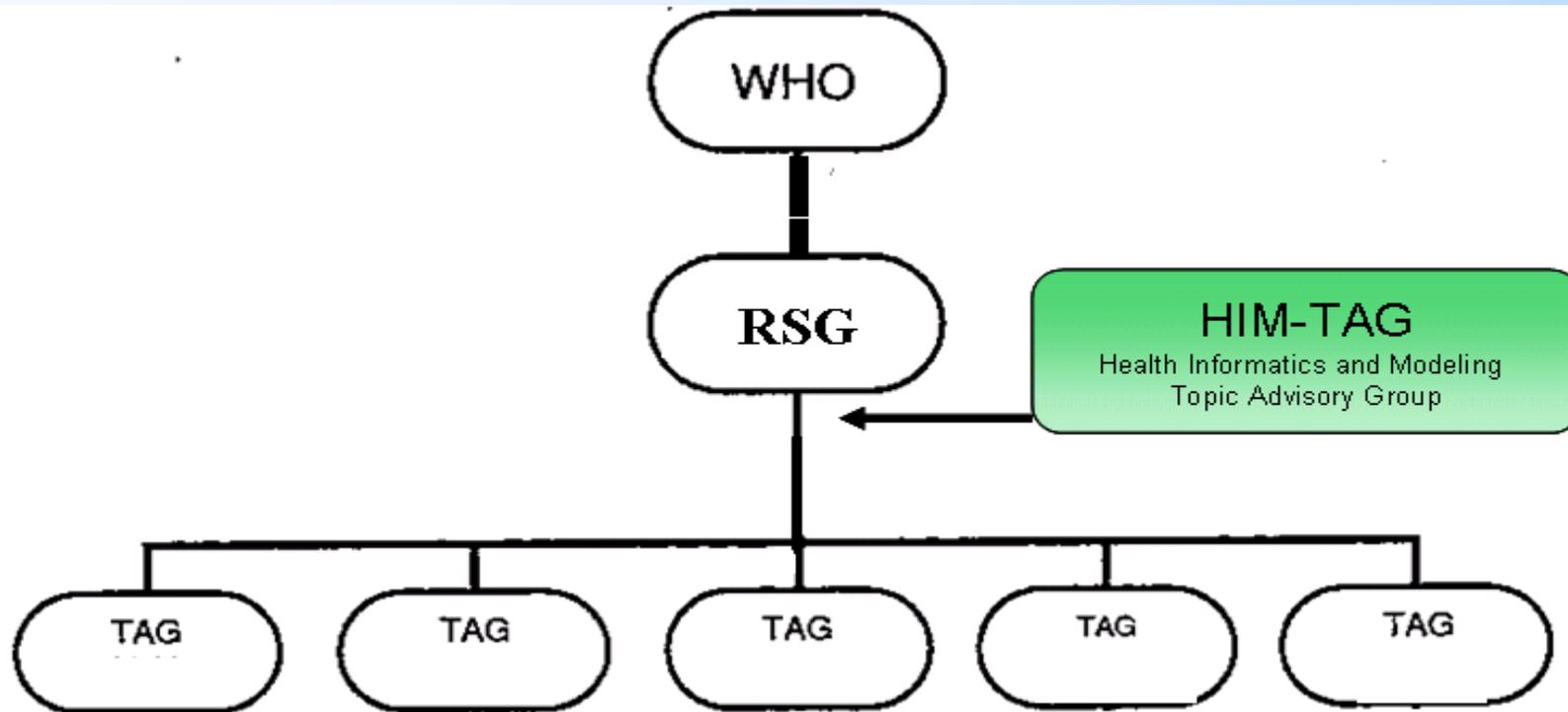
⇒ 現実問題として全ての場で同意を得る事ができるかの検討

## ・倫理審査委員会との連携

⇒ DB公開前の段階から連携し、倫理上の問題やプライバシー権の侵害の可能性について審査

## 国際標準化の推進(WHO)

### ICD11 and HIM-TAG



## TAG-HIM present status

- 1. Content Model Group
  - Treat the original use case of ICD11 (Stefany)
- 2. Information group
  - Alan Rector (Univ. of Manchester, UK), John, Chris, SCT
- 3. Content Model front end for each TAG,
  - Categorical Structure for Rare disease
    - Jean Marie with Rare Disease TAG
  - **IBMDB model is modified for internal medicine**
    - **Jun Nakaya and Hiroshi Tanaka with Internal Medicine TAG**
  - Robert will take care of these things
- 4. SCT(SNOMED) Coordination
  - IHTSDO harmonization panel will take this issue.
    - (Kent, Alan Rector (Univ. of Manchester, UK), Chris, Olivier)
  - GO and other ontologies will cover the remained area.

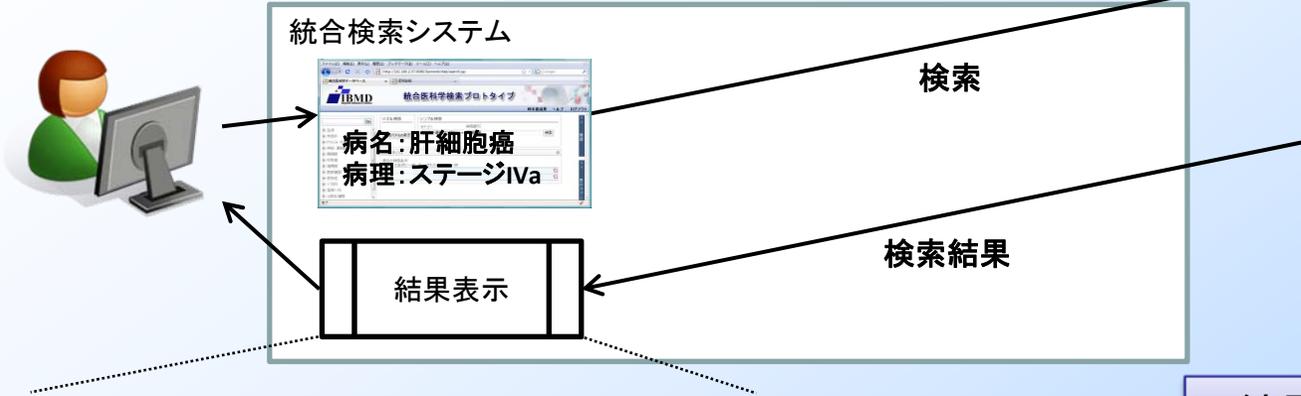
# 平成20年度の成果

## 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試験公開によって得られた要求、課題より要件定義を行ない、以下の検討、開発を行った。

1. 網羅的疾患分子病態DB(447症例)、パーキンソン病DB(400症例)への症例数追加に加え、新たにGeMDBJ(440症例)を統合した。
2. 対象疾患を絞り、肝細胞癌、大腸癌、口腔癌、パーキンソン病、頭頸部癌において、語数247語、関連数190のオントロジー情報を整備した。
3. 複数の要素データベースを統一して検索可能なインターフェースを検討した。
4. 検索GUIとして、新たにパネル検索方式を実装した。
5. 新たに分子情報クラスタリングの機能を追加した。

## 国立がんセンター研究所 GeMDBJとの連携



Genome Medicine Database of Japan (GeMDBJ) - Mozilla Firefox

ファイル(E) 編集(E) 表示(V) 検索(S) ブックマーク(B) ツール(T) ヘルプ(H)

https://gemdbj.nbio.jp.jp/dp/DownloadSite.do

発現解析

実験ID	実験名	GeMDBJ形式	MAGE-ML形式	cellファイル
EXPR001	Normal tissue series	zip (6,397.2 kb)	zip (3,572.7 kb)	zip (204.5 mb)
EXPR002	Kidney (Non-cancerous, Adenocarcinoma)	zip (3,696.8 kb)	zip (2,037 kb)	zip (158.8 mb)
EXPR003	Kidney (Cell line)	zip (406.4 kb)	zip (264.3 kb)	zip (17.6 mb)
EXPR004	Stomach (Cell line)	zip (1,206 kb)	zip (687.3 kb)	zip (49.3 mb)
EXPR005	Blood, Leukocyte (Normal, Acute myeloid leukemia)	zip (3,345.5 kb)	zip (2,159.8 kb)	zip (159.2 mb)
EXPR006	Blood, Cells Panel (Allthia)	zip (1,208.8 kb)	zip (703.3 kb)	zip (37.2 mb)
EXPR007	Blood, PBMC (LPS/Cyc stimulation time course)	zip (88.4 kb)	zip (726 kb)	zip (58 mb)
EXPR008	Blood vessel, Cell line, Umbilical vein endothelial cell, HUVEC	zip (1,237.5 kb)	zip (83.7 kb)	zip (3.1 mb)
EXPR009	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)
EXPR010	Blood, PBMC (LPS stimulation time course)	zip (1,230.5 kb)	zip (712.5 kb)	zip (40.8 mb)
EXPR011	Stomach (Cell line/spontaneous, S-Aza-dC+TSA)	zip (17,916.7 kb)	zip (9,658.8 kb)	zip (212.9 mb)
EXPR012	B cell (Control vs. Treatment)	zip (362.6 kb)	zip (264.8 kb)	zip (11.3 mb)
EXPR013	Eosinophil (Control vs. Treatment)	zip (484.1 kb)	zip (328.1 kb)	zip (14.8 mb)
EXPR014	NK cell (Spontaneous)	zip (122 kb)	zip (139.7 kb)	zip (3.7 mb)
EXPR015	CGMC (Spontaneous)	zip (122.4 kb)	zip (140 kb)	zip (3.8 mb)
EXPR016	CGMC (Spontaneous)	zip (122.9 kb)	zip (140.7 kb)	zip (3.7 mb)
EXPR017	CGMC (Spontaneous)	zip (122.9 kb)	zip (139.6 kb)	zip (3.7 mb)
EXPR018	PBMC	zip (484 kb)	zip (328.4 kb)	zip (15 mb)
EXPR019	PBMC (Control vs. Treatment)	zip (607 kb)	zip (393.3 kb)	zip (18.6 mb)
EXPR020	PBMC (Control vs. Treatment)	zip (484 kb)	zip (328.4 kb)	zip (15 mb)
EXPR021	PBMC (Control vs. Treatment)	zip (484.2 kb)	zip (328.7 kb)	zip (14.9 mb)
EXPR022	PBMC (Control vs. Treatment)	zip (484.2 kb)	zip (328.7 kb)	zip (14.9 mb)
EXPR023	PBMC (Control vs. Treatment)	zip (484.2 kb)	zip (328.7 kb)	zip (14.9 mb)
EXPR024	PBMC (Control vs. Treatment)	zip (484.2 kb)	zip (328.7 kb)	zip (14.9 mb)
EXPR025	PBMC (Control vs. Treatment)	zip (484.2 kb)	zip (328.7 kb)	zip (14.9 mb)

GeMDBJ

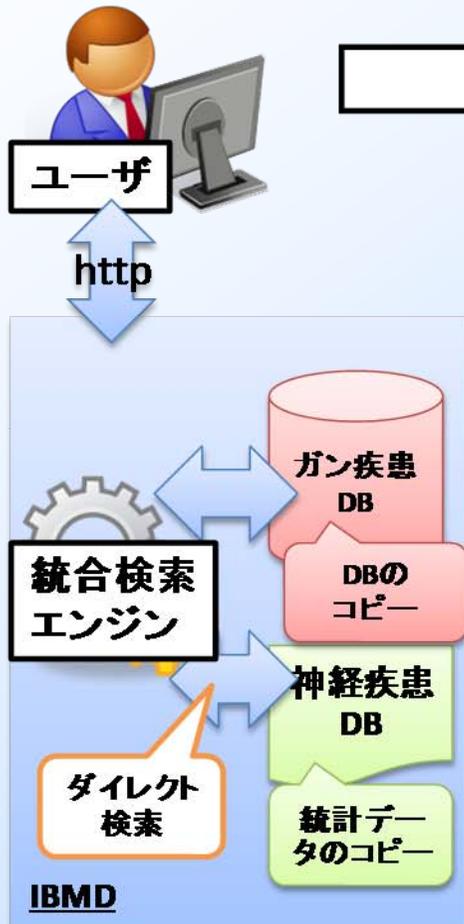
結果画面(サマリー)

結果画面(分子情報クラスタリング)

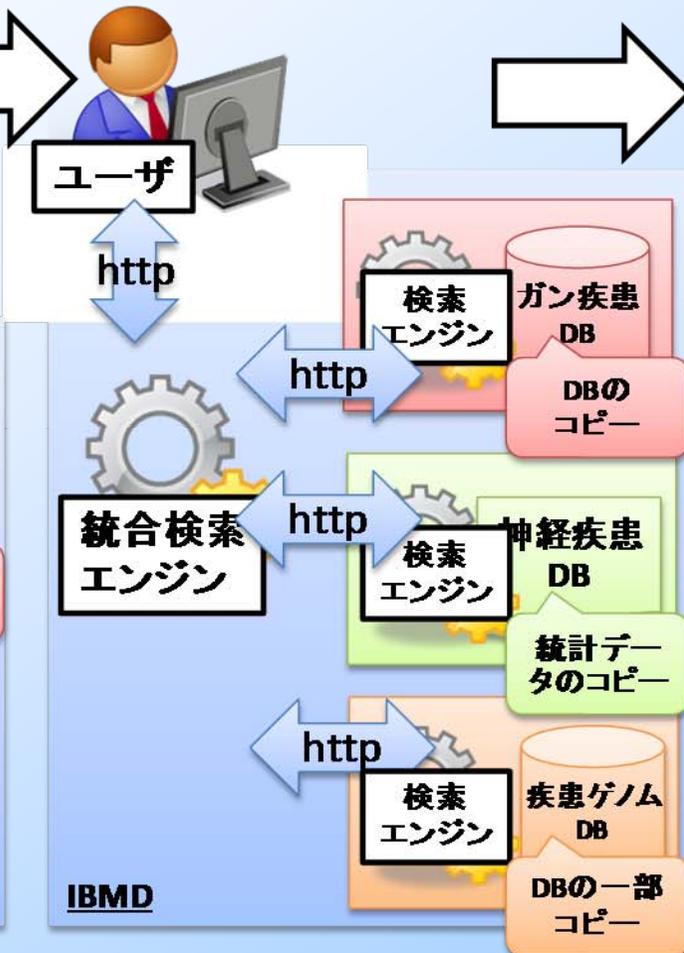
分子情報クリック  
で画面遷移

# 統合検索インタフェースプログラム

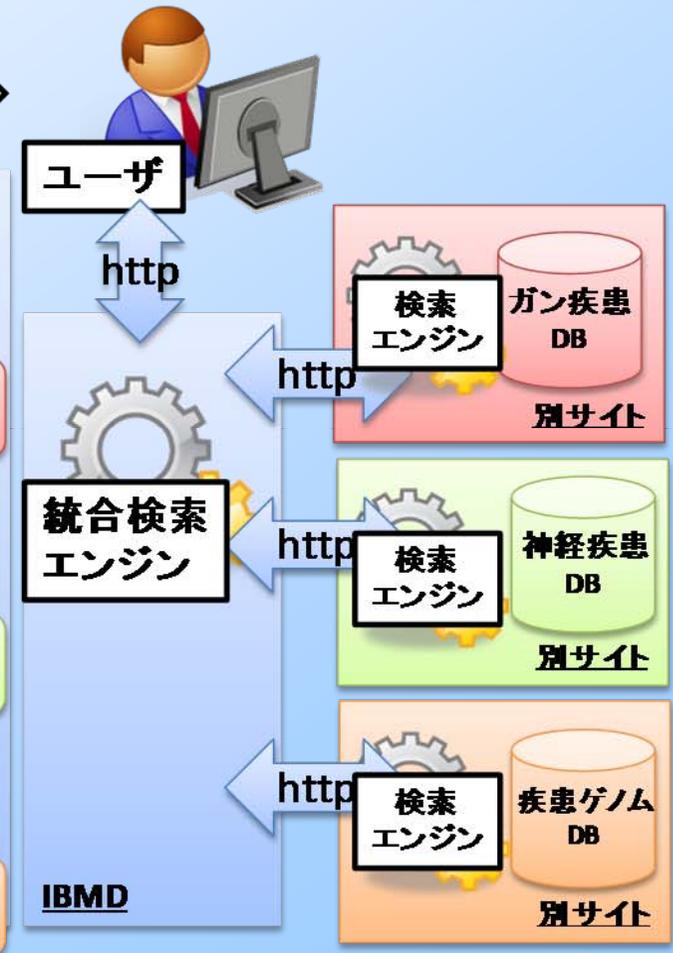
H19年度



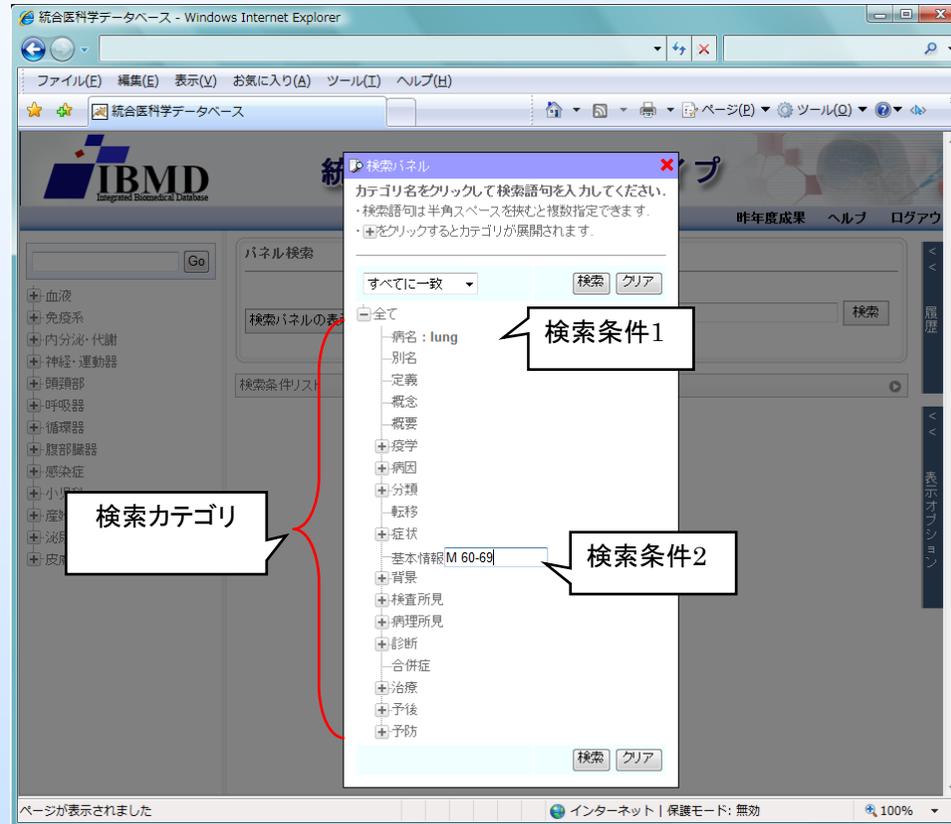
H20年度



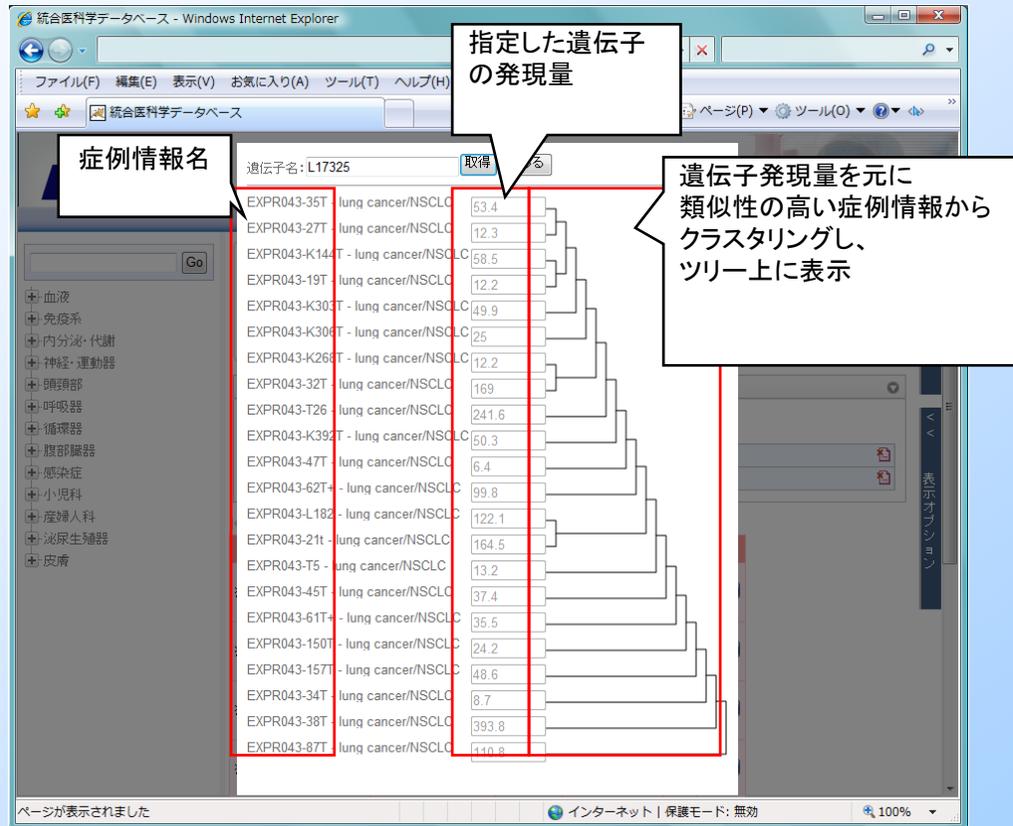
H21年度(予定)



# パネル検索GUI



# 分子情報クラスタリング



# 検索実施例1

The screenshot shows the IBMD search interface with the following search criteria:

- 検索条件: 病名: Hepatocellular, 組織所見(顕微鏡): V=0

The search results table is as follows:

データベース名	病名	組織所見(顕微鏡)	検索キーワード	D/S	詳細
網羅的疾患分子病態DB	Hepatocellular carcinoma(肝細胞癌)/Hepatocellular carcinoma(肝細胞癌)	...=1/w=0/+/...	病名:Hepatocellular 組織所見(顕微鏡):V=0	D	[Icons]
疾患ゲノムデータベース	liver cancer/Hepatocellular carcinoma	...=0/w=0	病名:Hepatocellular 組織所見(顕微鏡):V=0	D	[Icons]
疾患ゲノムデータベース	liver cancer/Hepatocellular carcinoma	...=0/w=0	病名:Hepatocellular 組織所見(顕微鏡):V=0	D	[Icons]
疾患ゲノムデータベース	liver cancer/Hepatocellular carcinoma	...=1/w=0	病名:Hepatocellular 組織所見(顕微鏡):V=0	D	[Icons]
疾患ゲノムデータベース	liver cancer/Hepatocellular carcinoma	...=0/w=0	病名:Hepatocellular 組織所見(顕微鏡):V=0	D	[Icons]
疾患ゲノムデータベース	liver cancer/Hepatocellular carcinoma	...=0/w=0	病名:Hepatocellular 組織所見(顕微鏡):V=0	D	[Icons]

Annotations on the screenshot:

- 検索条件**: Points to the search criteria input fields.
- 検索結果**: Points to the search results table.
- 網羅的疾患分子病態DB 81症例**: Points to the first result row.
- 疾患ゲノムデータベース 54症例**: Points to the subsequent result rows.

# 検索実施例2

The screenshot shows the IBMD search interface. On the left, there is a navigation menu with categories like '血液' (Blood), '免疫系' (Immune system), '内分泌・代謝' (Endocrine/Metabolism), '神経・運動器' (Neurology/Musculoskeletal), '臓器' (Organ), and '呼吸器' (Respiratory). The main search area has two tabs: 'パネル検索' (Panel Search) and 'シンプル検索' (Simple Search). Under 'パネル検索', there is a '検索条件リスト' (Search Condition List) with the following criteria: '病名: Hepatocellular', '基本情報: 40代 男', and '組織所見(顕微鏡): w=0 vp=1'. Below this is a table of search results.

データベース名	病名	基本情報	組織所見(顕微鏡)	検索キーワード	D/S	詳細
網羅的疾患分子病態DB	Hepatocellular carcinoma(肝細胞癌)/Hepatocellular carcinoma(肝細胞癌)	...3, 40代男	.../+vp=1/w=0/+/...	病名:Hepatocellular 基本情報:40代男 組織所見(顕微鏡):w=0 組織所見(顕微鏡):vp=1	D	
疾患ゲノムデータベース	liver cancer/Hepatocellular carcinoma	M/40-49	...te/vp=1/w=0	病名:Hepatocellular 基本情報:40-49 基本情報:男 組織所見(顕微鏡):w=0 組織所見(顕微鏡):vp=1	S	

検索条件  
検索結果

網羅的疾患分子病態DB  
1症例

疾患ゲノムデータベース  
1症例

# 検索実施例3

検索条件  
検索結果

検索条件

現在の検索条件  
 全ての条件に一致  
 いずれかの条件に一致  
 病名 : Hepatocellular  
 組織所見(顕微鏡) : vp=0 well

データベース名	病名	組織所見(顕微鏡)	検索キーワード	D/S	詳細
網羅的疾患分子病態DB	Hepatocellular carcinoma(肝細胞癌)/Hepatocellular carcinoma(肝細胞癌)	..ma/well/tr.../+vp=0/w...	病名:Hepatocellular 組織所見(顕微鏡):vp=0 組織所見(顕微鏡):well	D	
網羅的疾患分子病態DB	Hepatocellular carcinoma(肝細胞癌)/Hepatocellular carcinoma(肝細胞癌)	..od>well/tr...-/vp=0/w...	病名:Hepatocellular 組織所見(顕微鏡):vp=0 組織所見(顕微鏡):well	D	
疾患ゲノムデータベース	liver cancer/Hepatocellular carcinoma	well/vp=0/w...	病名:Hepatocellular 組織所見(顕微鏡):vp=0 組織所見(顕微鏡):well	D	
疾患ゲノムデータベース	liver cancer/Hepatocellular carcinoma	well/vp=0/w...	病名:Hepatocellular 組織所見(顕微鏡):vp=0 組織所見(顕微鏡):well	D	
疾患ゲノムデータベース	liver cancer/Hepatocellular carcinoma	well/vp=0/w...	病名:Hepatocellular 組織所見(顕微鏡):vp=0 組織所見(顕微鏡):well	D	

網羅的疾患分子病態DB  
22症例

疾患ゲノムデータベース  
10症例

クリックで  
詳細結果表示

詳細結果

「組織所見(顕微鏡)」の詳細情報

検査時間	検査材料	組織診断	組織所見	病理診断名	組織学的分化度	組織構造	発育様式	皮膚形成	門脈浸潤(vp)	静脈浸潤	vp/vv	肝動脈浸潤	胆管浸潤	切除断端(SM)	pT	pM	pステージ
60代 00:00:00				Hepatocellular carcinoma	mod=well	trabecular	eg	+	vp=0	=0	-	0	0	-			I

# 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

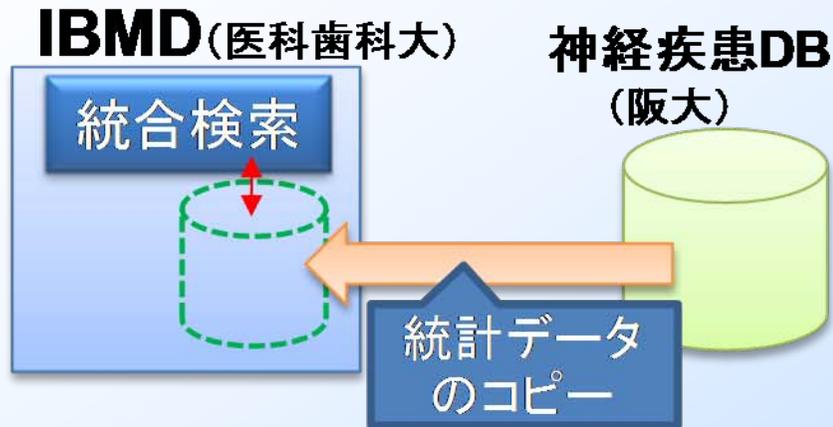
- プロトタイプの評価を通して、疾患オントロジー及びプロトタイプとして開発したセマンティック検索エンジンの高度化、特に分子情報に関するデータを取り扱うためのインタフェースの高度化を行う。

## 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

- 症例に基づいた統合検索システムを構築し、東京医科歯科大学の統合検索システムの登録研究者に向けて公開を行う。
- がん疾患とパーキンソン病を中心として、国内の主要な疾患データベースを参照できるような仕組みを検討する。

# 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

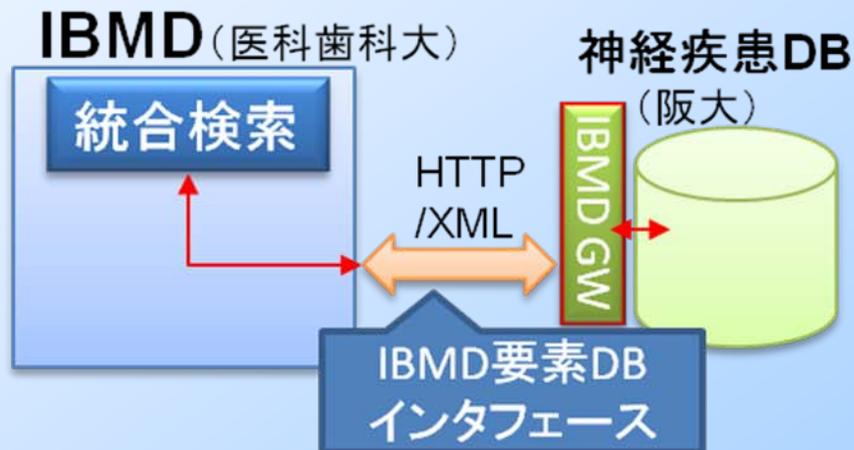


### データ移動型統合方式

H19年度に暫定的なプロトタイプとして試作した統合方式

H19年度と同様のデータ移動型統合方式でのデータ更新作業を実施

→神経疾患DB(阪大)からIBMDに移動させるデータの更新は問題なく行えることを確認



### インタフェース型統合方式

H20年度に作成した要件定義に基づく統合方式 (H21年度に実装予定)

統合データベースの継続的な運用・統合検索の高度化に対応するためのシステム改良の柔軟性などを考慮し、H21年度にはインタフェース型統合方式に移行

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

## 神経疾患DBの要素DBインタフェース定義(一部)

各機能毎に統合DBと要素DBの対応関係を定義

要素DB外接機能一覧								
No	機能名	インタフェース名	区分	データソース	システムフロー	Request 項目数	Result 項目数	概要
1	件数問い合わせ	CountGet()	照会	阪大 DB	External System	N	1	一覧取得resの参照元に記載する項目が検索条件に一致する臨床情報の件数取得
2	一覧取得	SearchGet()	照会	阪大 DB	External System	N	M	検索条件に一致する臨床ID及び項目内容取得
3	詳細情報取得	DetailGet()	照会	阪大 DB	External System	N	M	リクエストに記載された臨床IDの詳細項目内容取得

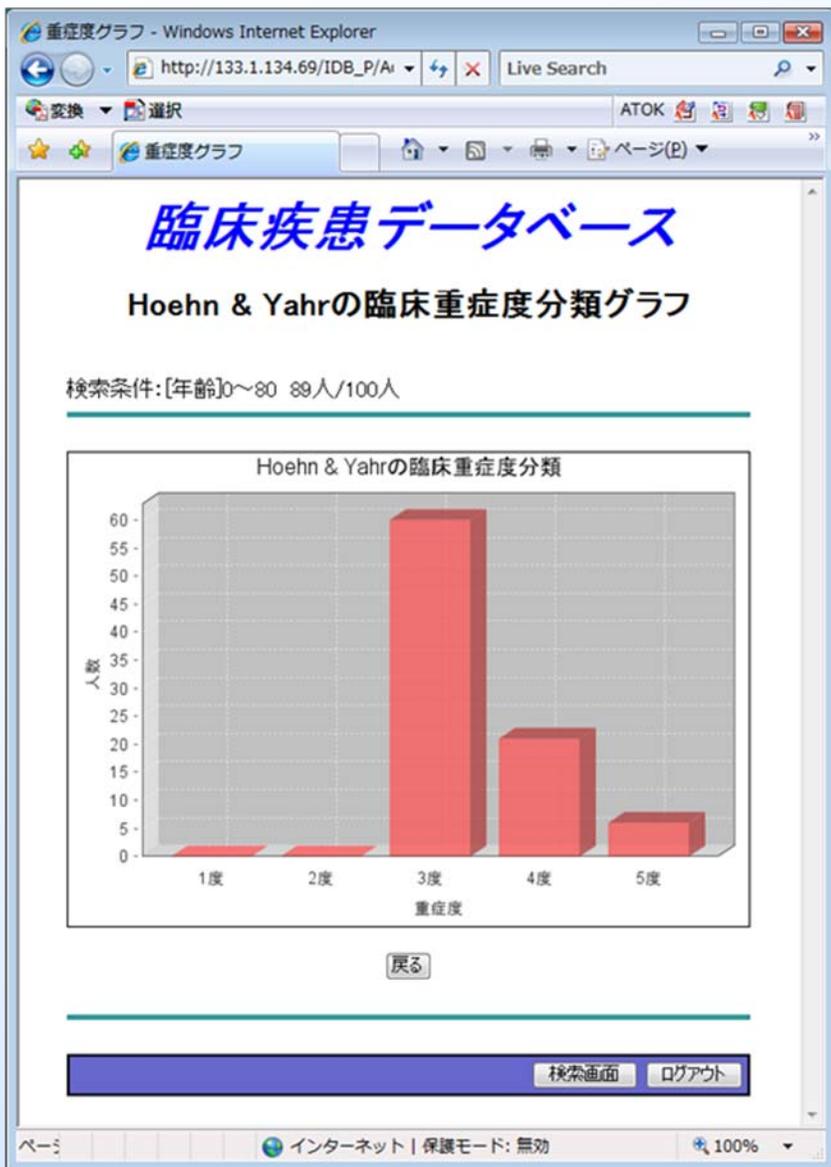
### 外部システム間 I/F項目定義

No	項目名	長さ	必須	参照元		補足説明	
				テーブル名	項目名		
1	臨床ID	文字列	8	Yes	MML臨床情報	uid	文書ユニークID
2	病名	文字列	-	Yes	詳細情報取得(病名)参照		
3	別名	文字列	-	No	なし		
4	定義	文字列	-	No	なし		
5	概念	文字列	-	No	なし		
6	概要	文字列	-	No	なし		
7	疫学	文字列	-	No	なし		
8	発症経路	文字列	-	No	なし		
9	原因論	文字列	-	No	なし		
24	感染経路	文字列	-	No	なし		
25	分類	文字列	-	No	MML臨床情報	mmlDh:diseaseClass	診断分類
26	肉眼分類	文字列	-	No	なし		
27	病理分類	文字列	-	No	なし		
28	転移	文字列	-	No	なし		
29	症状	文字列	-	No	MML臨床情報	mmlDh:clinicalConditio	臨床症状
30	時期別	文字列	-	No	MML臨床情報	mmlDh:developmentA	発症年齢と経過
31	重症度別	文字列	-	No	MML臨床情報	mmlDh:parkinsonDegr	パーキンソン病重症度
32	臓器別	文字列	-	No	MML臨床情報	mmlDh:autonomicinfo	自立神経情報
33	発病様式	文字列	-	No	なし		
34	基本情報	文字列	-	No	詳細情報取得(基本情報)参照		
35	背景	文字列	-	No	なし		
36	病歴	文字列	-	No	MML臨床情報	mmlDh:DiagnosticHist	診断歴情報
37	生活歴	文字列	-	No	MML臨床情報	mmlDh:life	生活状況
38	家族歴	文字列	-	No	MML臨床情報	mmlDh:family	家族歴
39	検査所見	文字列	-	No	詳細情報取得(検査所見)参照		
40	血液	文字列	-	No	なし		

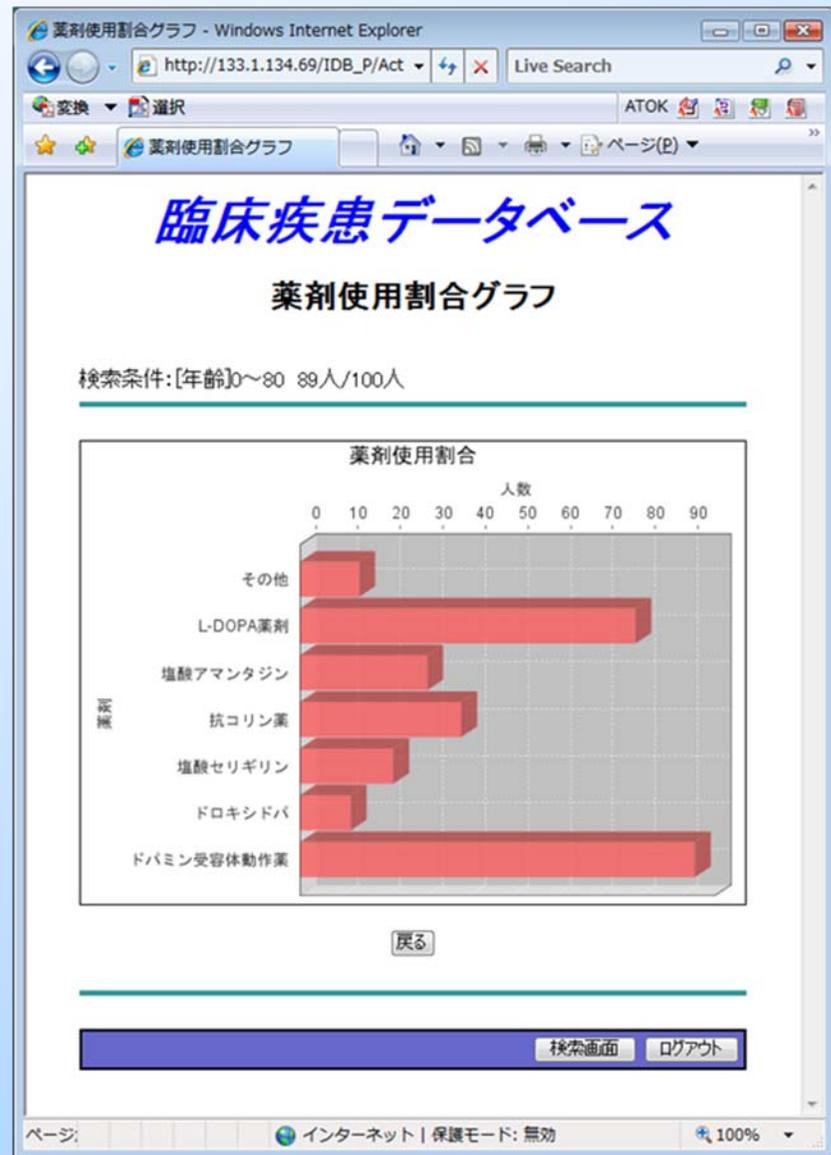
オントロジーで定義された疾患テンプレート情報

要素DB(神経疾患DB)のデータとの対応





臨床重症度分類のグラフ表示例

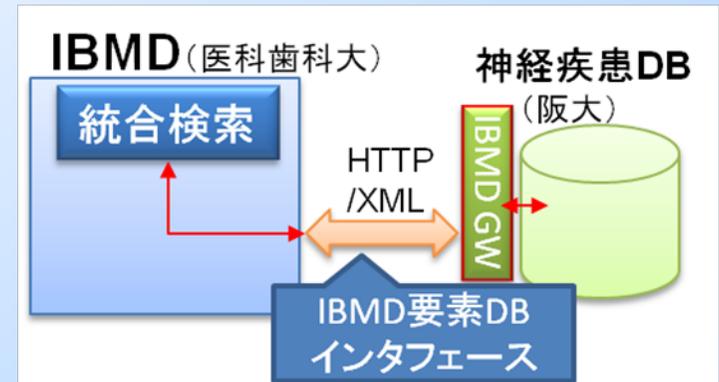


薬剤使用割合のグラフ表示例

## 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

平成20年度に実施した多階層の統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、疾患DBオントロジー、セマンティクスなど)の要件定義に基づいたプロトタイプシステムの試作を通し、問題点の検討を行う。



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。
- 平成20年度に公開した疾患データベースについて、検索などを含む具体的要求と課題の検討を進め、倫理的な問題の解決(患者同意を含め)を行う。
- 神経疾患(パーキンソン病)DBの公開患者数を増加させる(連結匿名化400症例)。

## H21年度計画：特記事項（大阪大学）

### 分担機関・東大グループとのデータ連携

大阪大学／東京大学のそれぞれで開発しているパーキンソン病DBについて、相互のDBが持つデータ(スキーマおよび内容)の相互運用性比較を行った上で、DBの相互連携の為のAPI提供の可能性、補正データの収集などを検討する。

### 利用者から見て使い勝手の良いDBの在り方の検討

パーキンソン病に着目し、

- パーキンソン病の治療ガイドライン(神経学会6月公開予定)、パーキンソン病患者の会、治療薬提供製薬会社等のホームページで公開されている関連情報などへのリンク付け

- 中核機関Webサイト内DB横断検索サービス内へのAPIを介した連携など、利用者が必要とする情報に効率的にアクセスできる仕組みを検討する。

### DB公開方法(アクセス管理ポリシー)の再考

神経疾患DBの統計値の公開については、パスワード認証なしで公開するようアクセス管理ポリシーの変更を検討する(公開時期は平成21年度の上期を予定)。