

2019/09/13

INGEMメディカルAIセミナー

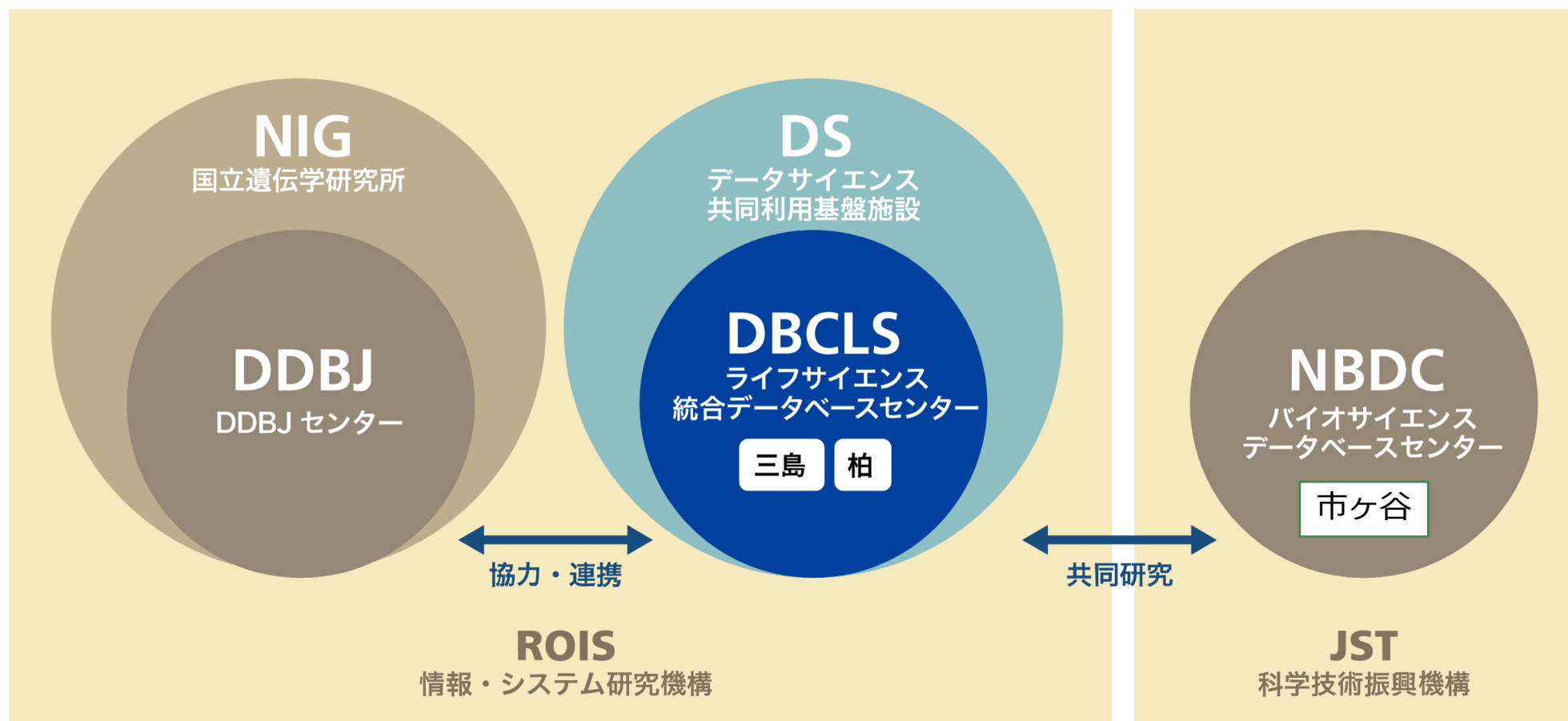
稀少・遺伝性疾患検索システム PubCaseFinderの 研究開発と使い方について

情報システム・研究機構

ライフサイエンス統合データベースセンター

藤原豊史 @fujitoyo

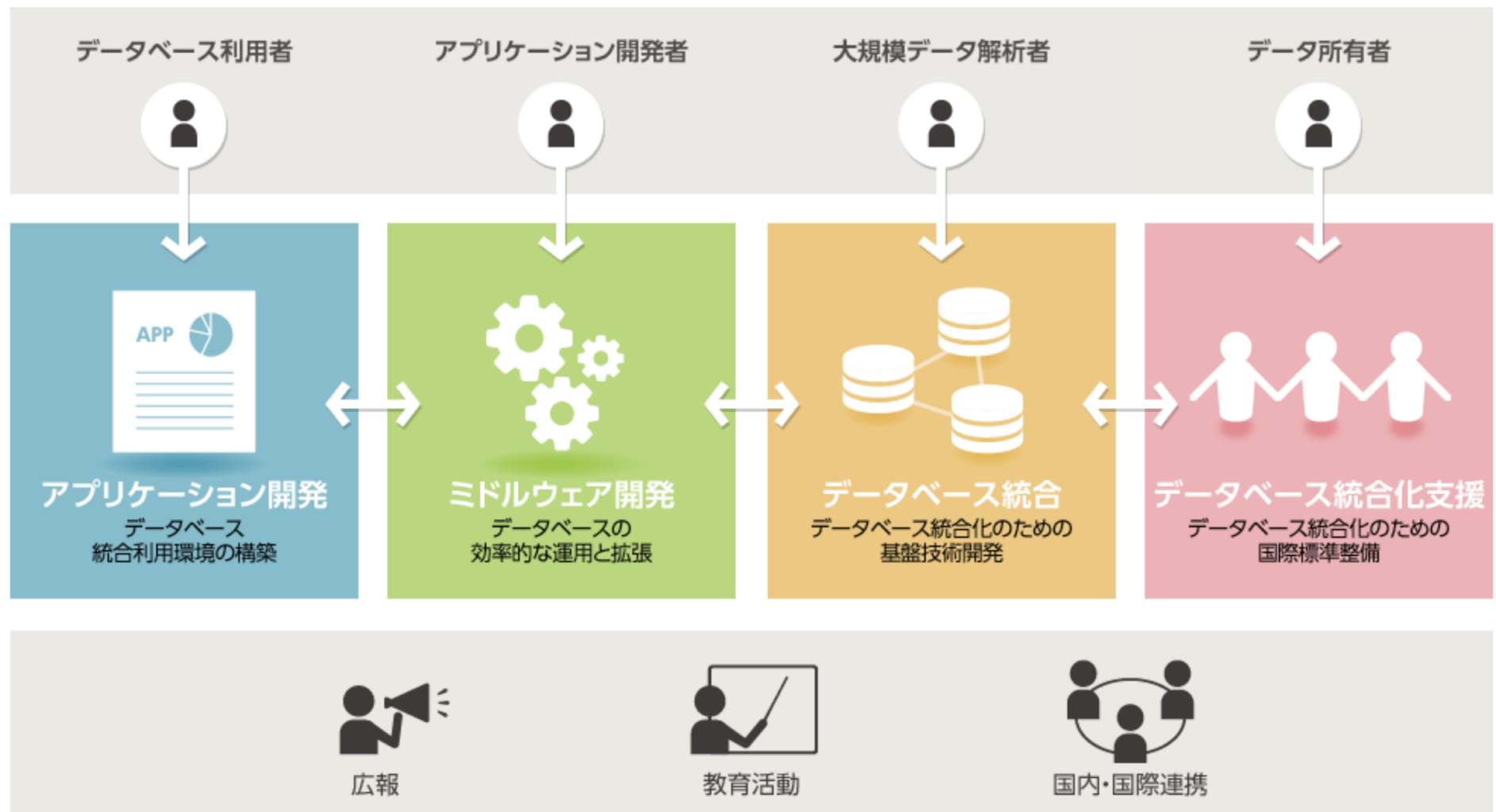
ライフサイエンス統合データベースセンターと関連機関



- 2007年4月 情報・システム研究機構（ROIS）に DBCLS 設立
- 2011年4月 JST にバイオサイエンスデータベースセンター（NBDC）設立
「基盤技術開発プログラム」を ROIS が受託
- 2014年4月 NBDC 「統合データベースにおける基盤技術開発とデータベース運用に係る共同研究」開始
- 2016年4月 ROIS にデータサイエンス共同利用基盤施設設置

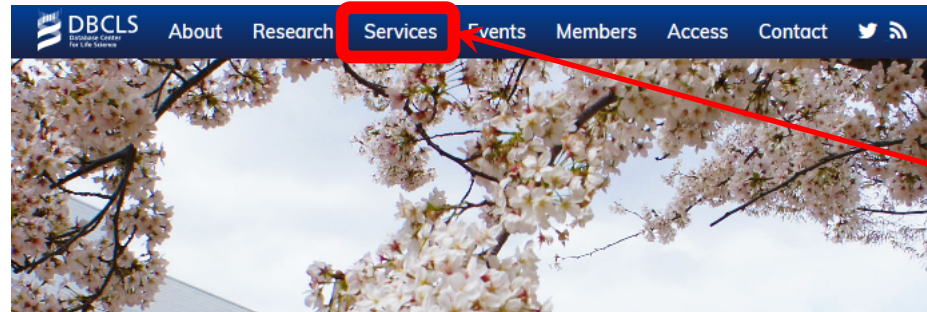
DBCCLSのミッション

□ ライフサイエンスのデータ(ベース)を統合的に利用できる環境の構築



DBCLSのサービス

サービスへのアクセス



Services をクリック

カテゴリ

- すべて
- データ解析ツール
- 文献知識抽出
- コンテンツ
- セマンティックウェブ

ユーザ

- すべて
- データベース利用者
- アプリケーション開発者
- 大規模データ解析者
- データ所有者

TOGO GENOME TogoGenome 生物種とゲノムに関連する多種多様な情報を集約し、ゲノム情報から体系的に新しい発見を実現したシステムです。 ● セマンティックウェブ ● データベース利用者 [詳細] [アクセス]	DBCLS SRA 遺伝子発現解析の基盤となる遺伝子の遺伝子発現量と発現パターンに関する、閲覧できるウェブツールです。 ● データ解析ツール ● データベース利用者 [詳細] [アクセス]	RefEx 遺伝子発現解析の基盤となる遺伝子の遺伝子発現量と発現パターンに関する、閲覧できるウェブツールです。 ● データ解析ツール ● データベース利用者 [詳細] [アクセス]
AOE AOE 公共データベースに登録された遺伝子発現データについて、各遺伝子発現の発現量・発現・比較することができるウェブツールです。 ● データ解析ツール ● データベース利用者 [詳細] [アクセス]	GGRNA 統合遺伝子検索GGRNA 様々なキーワードや塩基配列、アミノ酸配列などから迅速に遺伝子を検索することができ、遺伝子検索エンジンです。 ● データ解析ツール ● データベース利用者 [詳細] [アクセス]	GGGenome 高速塩基配列検索GGGenome ゲノムと転写産物から塩基配列を高速に検索するツールです。 ● データ解析ツール ● データベース利用者 [詳細] [アクセス]
Genndoo Genndoo 遺伝子、疾患について、関連する疾患、薬物、遺伝、生命現象などの情報をキーワードでリスト表示するツールです。 ● データ解析ツール ● データベース利用者 [詳細] [アクセス]	LODQA LODQA 自然言語でSPARQLエンドポイントを検索するためのインタフェースです。 ● 文献知識抽出 ● セマンティックウェブ ● データベース利用者 [詳細] [アクセス]	PubCaseFinder PubCaseFinder 患者の症状を入力するだけで、関連する論文・症例を自動的にリストアップしてくれる症例向け検索システムです。 ● 文献知識抽出 ● セマンティックウェブ ● データベース利用者 [詳細] [アクセス]
CRISPRdirect CRISPRdirect 入力した塩基配列に対してCRISPR-Cas9システムのガイドRNAを設計することができるツールです。 ● データ解析ツール ● データベース利用者 [詳細] [アクセス]	新着論文レビュー 日本人をターゲットとする生命科学分野の論文について、最新の著者発表の論文による日本語のレビューを公開するサービスです。 ● コンテンツ ● データベース利用者 ● データ所有者 [詳細] [アクセス]	領域融合レビュー 生命科学分野における最新の研究成果について、第一線の研究者の発表による日本語のレビューを公開するサービスです。 ● コンテンツ ● データベース利用者 ● データ所有者 [詳細] [アクセス]
TOGO TV 統合TV データベースやウェブサイトの使い方を動画で解説するサービスです。提供している動画は110本を超えています。 ● コンテンツ ● データベース利用者 ● データ所有者 [詳細] [アクセス]	BodyParts3D/Anatomography BodyParts3D/Anatomography 誰でもウェブ上で3Dモデルを制作・共有できるサービスです。作成した3Dモデルは、論文に用いるなど自由に公開することが可能です。 ● コンテンツ ● データベース利用者 ● データ所有者 [詳細] [アクセス]	ライフサイエンスQA ライフサイエンス研究にまつわる疑問や質問、それらへの回答を共有するウェブサイトです。 ● コンテンツ ● データベース利用者 [詳細] [アクセス]
YummyData うまかデータ Umako-Yummy SPARQLエンドポイントの機械的な評価を定期的に行います。 ● セマンティックウェブ ● データベース利用者 ● アプリケーション開発者 [詳細] [アクセス]	umakaviewer UmakoViewer RDFデータの構造を可視化するサービスです。 ● セマンティックウェブ [詳細] [アクセス]	SPARQL Builder SPARQL Builder 直感的なUIとマウス操作でSPARQLクエリを構築するためのSPARQLクエリ構築ツールです。 ● セマンティックウェブ ● データベース利用者 ● アプリケーション開発者 [詳細] [アクセス]
SPARQL support SPARQL support SPARQLクエリをMarkdownによる簡単な記述によりREST API化するためのアプリケーションです。 ● セマンティックウェブ ● データベース利用者 [詳細] [アクセス]	SPARQLList SPARQLList SPARQLクエリをMarkdownによる簡単な記述によりREST API化するためのアプリケーションです。 ● セマンティックウェブ ● データベース利用者 ● アプリケーション開発者 [詳細] [アクセス]	SPANG SPANG コマンドライン環境で利用可能なSPARQLクライアントです。 ● セマンティックウェブ ● データベース利用者 ● アプリケーション開発者 [詳細] [アクセス]
sparql-proxy SPARQL-proxy SPARQLエンドポイントにジョブ管理やキャッシュ機能を持たせて資金に公開するためのアプリケーションです。 ● セマンティックウェブ ● アプリケーション開発者 [詳細] [アクセス]	TOGO STANZA TogoStanza SPARQLクエリを可視化したウェブページに埋め込むスタンプを提供するアプリケーションです。 ● セマンティックウェブ ● アプリケーション開発者 [詳細] [アクセス]	d3sparql.js d3sparql.js SPARQLの検索結果をD3.jsを用いて簡単に可視化するためのJavaScriptライブラリです。 ● セマンティックウェブ ● アプリケーション開発者 [詳細] [アクセス]
TOGO TABLE TogoTable 生物種とゲノムに関連する多種多様な情報を集約し、ゲノム情報から体系的に新しい発見を実現したシステムです。 ● セマンティックウェブ ● データベース利用者 [詳細] [アクセス]	TOGO WS TogoWS 国内外の主要データベースに対して、統一されたウェブサービスのインターフェイスを提供するサービスです。 ● セマンティックウェブ ● アプリケーション開発者 [詳細] [アクセス]	D2RQ Mapper D2RQ Mapper 関係データベース(RDB)を簡単にセマンティックウェブに対応させるウェブアプリケーションです。 ● セマンティックウェブ [詳細] [アクセス]
TogoDB TogoDB 表形式のデータをCSV/TSV形式でインポートすることで、データベースの公開・管理を誰でも手軽に行えるシステムです。 ● セマンティックウェブ ● データベース利用者 [詳細] [アクセス]	PubAnnotation PubAnnotation 文脈アノテーションのレポートリ。特にPubMedに対してのアノテーションを統合している。 ● 文献知識抽出 ● データベース利用者 [詳細] [アクセス]	PUBDICTIONARIES PubDictionaries テキストマイニング用辞書のレポートリ。誰でも辞書のアップロードができ、辞書を持ってテキストアナリシスができます。 ● 文献知識抽出 ● アプリケーション開発者 [詳細] [アクセス]
Allie Allie 文脈中に登場する単語とその正式名称の統合およびその関係性を検索します。 ● 文献知識抽出 ● セマンティックウェブ ● データベース利用者 ● 大規模データ解析者 [詳細] [アクセス]	Colil Colil ある論文が他の論文から引用されているとき、本文ではどのような文脈で引用されているのかについて、簡単に検索します。 ● 文献知識抽出 ● データベース利用者 [詳細] [アクセス]	inMeXes inMeXes 生命科学の文脈に特化した検索エンジン。1万人の力によって迅速に再構築します。 ● 文献知識抽出 ● データベース利用者 [詳細] [アクセス]
NBDC RDF Portal NBDC RDF Portal 日本の研究グループによって開発された生命科学データのRDFを統合しています。 ● セマンティックウェブ ● アプリケーション開発者 ● 大規模データ解析者 [詳細] [アクセス]	LSO RDF Data Portal LSO RDF Data Portal ライフサイエンス分野に特化したSPARQLでアクセスできるサービスです。 ● 文献知識抽出 [詳細] [アクセス]	TOGO VAR TogoVar TogoVarは、日本人ゲノム配列の個人による違い(バリエーション)とそれに隣接する疾患発症率などを収集・整理したデータベースです。 ● セマンティックウェブ ● データベース利用者 [詳細] [アクセス]

日本人ゲノム多様性統合データベース

TOGOVAR A comprehensive Japanese genetic variation database

Home Datasets Downloads Terms Contact About History Help Configuration

Search for disease or gene symbol or rs...

Disease: Breast-ovarian cancer, familial 2 Gene: ALDH2 refSNP: rs114202595 TogoVar: tgv421843 Position(GRCh37/hg19): 16:48258198 Region(GRCh37/hg19): 10:73270743-73376976

Results The number of available data is 10,000 out of 64,094,472.

TogoVar ID	RefSNP ID	Position	Ref / Alt	Type	Gene	Alt frequency	Consequence	SIFT	PolyPhen	Clinical significance
tgv67073503		1: 66507	ATAA... 44bp	Insertion			Intergenic variant			
tgv67073504		1: 66507	ATAT... 46bp	Insertion			Intergenic variant			
tgv67073557	rs374644745	1: 68057	G > A	SNV			Intergenic variant			
tgv67073583		1: 69098	C > T	SNV	OR4F5		Missense variant	0	0.207	
tgv457	rs201219564	1: 69270	A > G	SNV	OR4F5		Synonymous variant			
tgv465	rs140739101	1: 69428	T > G	SNV	OR4F5		Missense variant	0.01	0.984	
tgv478	rs2691305	1: 69511	A > G	SNV	OR4F5		Missense variant	0.82	0	
tgv485	rs190717287	1: 69534	T > C	SNV	OR4F5		Synonymous variant			
tgv491	rs556374459	1: 69555	T > G	SNV	OR4F5		Missense variant	0.04	0.973	
tgv492	rs754025211	1: 69559	G > A	SNV	OR4F5		Missense variant	0.02	0.675	

Filters

Dataset

- ☒ All 64,094,472
- ☒ JGA NGS 4,537,092
- ☒ JGA SNP 1,249,724
- ☒ 3.5KJPN 55,195,066
- ☒ HGV 510,169
- ☒ ExAC 9,087,393
- ☒ Clinvar 436,189

Alternative allele frequency

0.0 ~ 1.0 ☐ Invert range

☐ for all datasets ☒ for any datasets

Variant calling quality

☒ Exclude filtered out variants in selected datasets

Variant type

TogoVar収録データセット一覧

ヒトゲノムリファレンス配列

- GRCh37.p13 (2013/06/28)

Japanese Genotype-phenotype Archive (JGA)の個人ゲノムデータ由来のバリエーションデータセット

データセット名	バージョン/最終更新日	サンプルサイズ	検出バリエーション数	除外後バリエーション数	作成者
JGA-NGS	2018/06/01	125	13,338,968	4,679,025	NBDC
JGA-SNP	2018/06/01	183,884	1,966,919	1,249,724	NBDC

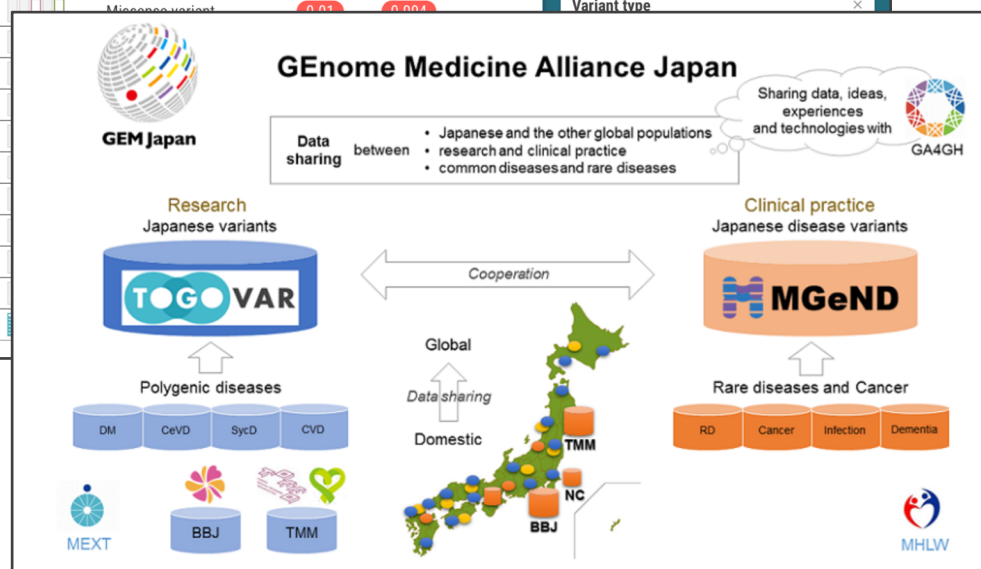
注：アレル数5以下のバリエーションはJGA-NGS/JGA-SNPデータセットから削除しています。

第三者が作成したバリエーションデータセット

データセット名	バージョン/最終更新日	サンプルサイズ	バリエーション数	バリエーション部位数	作成者
ClinVar	2019/06/24	-	443,213	-	NCBI
Exome Aggregation Consortium (ExAC)	Release 1 (2017/02/27)	60,706	10,195,872	9,362,319	Broad Institute
Human Genetic Variation Database (HGVD)	Version 2.30 (2017/08/02)	1,208	554,461	501,556	京都大学
ToMMo 3.5KJPNv2 Allele Frequency Panel (3.5KJPN)	v20181105open; Unfiltered	3,552	64,675,495	60,816,012	東北メディカル・メガバンク機構

注1：ClinVarはVCFファイルに含まれるGRCh37上の位置が決定しているバリエーションのみを掲載しています。

注2：3.5KJPNv2はAutosome、chrX (PAR1+PAR2+XTR)、chrMTのSNVとINDELから構成されます。詳細はSummary of ToMMo 3.5KJPNv2 (v20181105; Unfiltered)をご覧ください。



BioHackathon による国際・国内連携

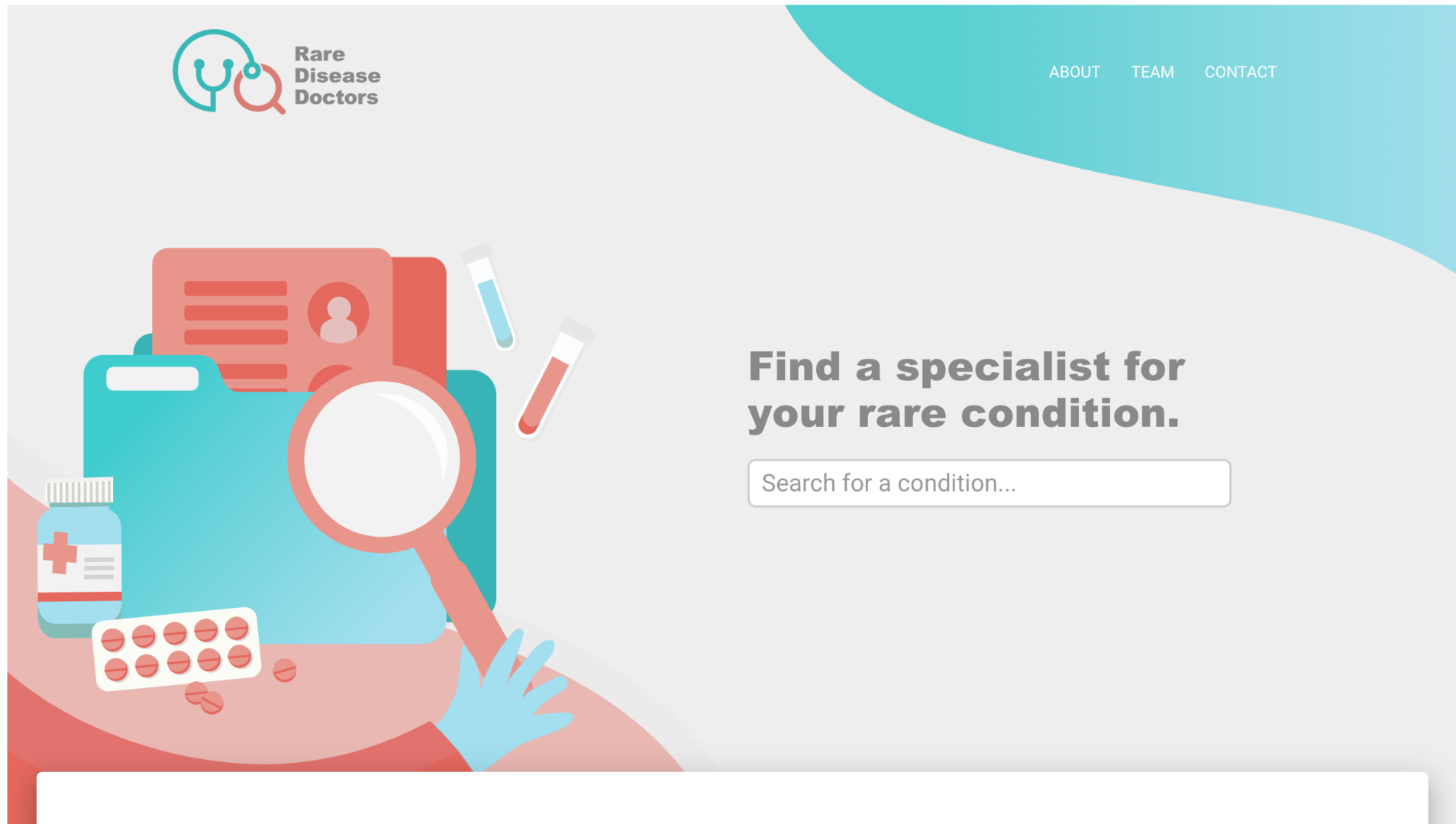
- 1週間程度泊まり込みで開発と議論を徹底的に行う開発者会議
 - 国際版バイオハッカソン（毎年1回、～140名）
 - 国内版バイオハッカソン（毎年1回、～70名）
 - BLAH: Biomedical Linked Annotation Hackathon（毎年1回、～50名）



2018年12月の国際版バイオハッカソン

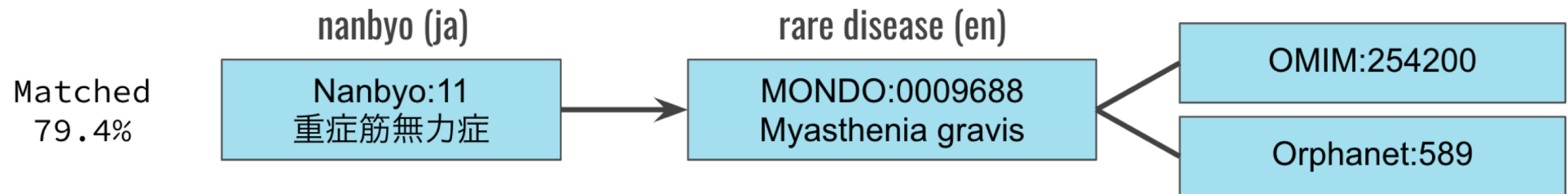
BioHackathon2019 in Fukuoka

<http://rarediseasedoctors.org>



BioHackathon2019 in Fukuoka

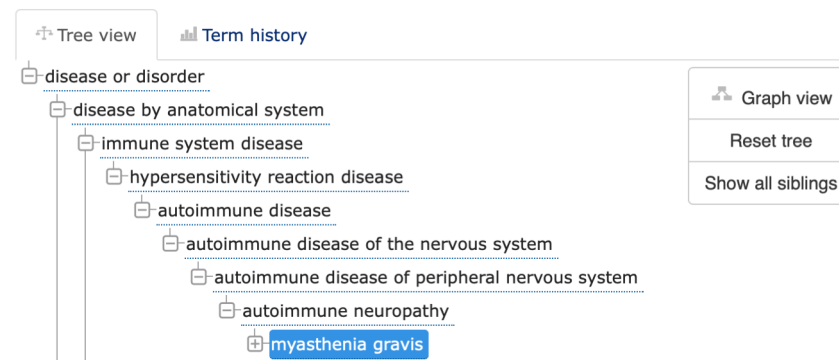
Mapping Nanbyo diseases to MONDO diseases



MONDO

myasthenia gravis

http://purl.obolibrary.org/obo/MONDO_0009688



Term info

def

Myasthenia gravis (MG) is a rare, clinically heterogeneous, autoimmune disorder of the neuromuscular junction characterized by fatigable weakness of voluntary muscles.

hasExactSynonym

acquired myasthenia, myasthenia gravis, autoimmune myasthenia gravis

hasRelatedSynonym

myasthenia gravis; MG, MG

xref

http://purl.obolibrary.org/obo/MESH#_D009157,
http://purl.obolibrary.org/obo/NCIT#_C60989,
http://purl.obolibrary.org/obo/ICD9#_358.00,
http://purl.obolibrary.org/obo/EFO_0004991,
http://purl.obolibrary.org/obo/UMLS#_C0026896,
http://purl.obolibrary.org/obo/COHD_76685,
http://purl.obolibrary.org/obo/MedDRA_10028417,
http://purl.obolibrary.org/obo/SCTID_91637004,
http://purl.obolibrary.org/obo/Orphanet_589,
http://purl.obolibrary.org/obo/ICD10#_G70.00,
http://purl.obolibrary.org/obo/ICD10#_G70.0,
http://purl.obolibrary.org/obo/OMIM_254200,
http://purl.obolibrary.org/obo/ICD9#_358.0,
http://purl.obolibrary.org/obo/GARD_0007122,
http://purl.obolibrary.org/obo/DOID_437

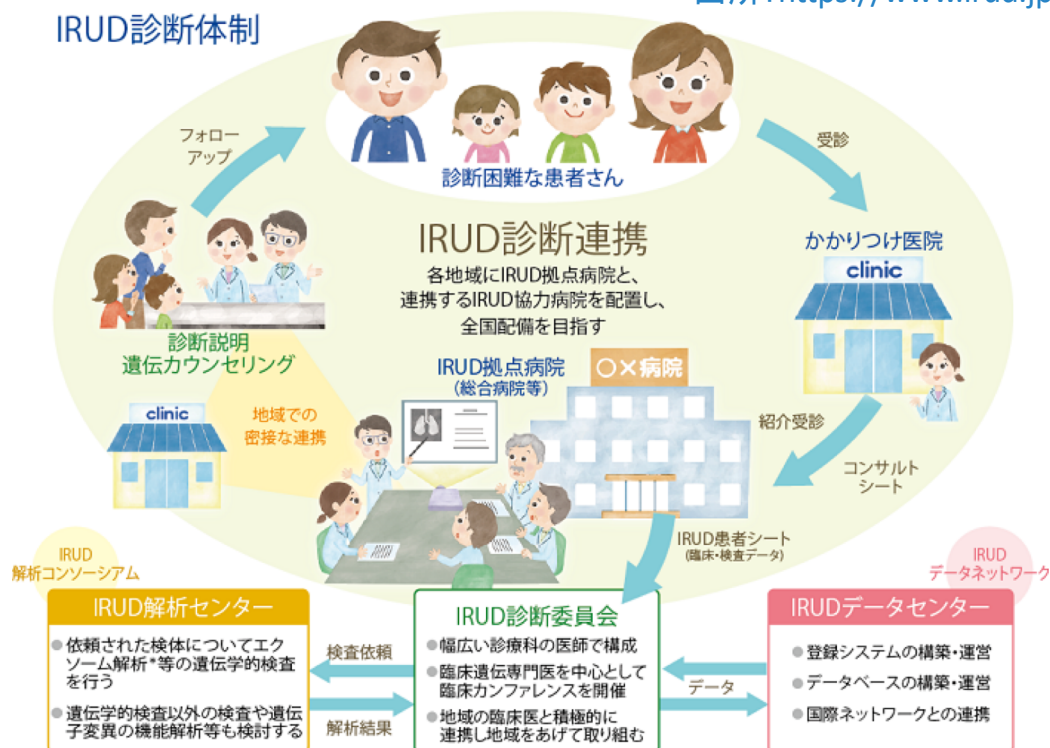
ゲノム解析を用いた診断

□ 日本医療研究開発機構(AMED)が主導する未診断疾患イニシアチブ(IRUD)は、日本全国の未診断患者に対してゲノム解析をすることで、診断しようとする試み

○ 既に6,000件以上のゲノム解析が実施された

○ 診断率: 約33%

出所: <https://www.irud.jp/system.html>

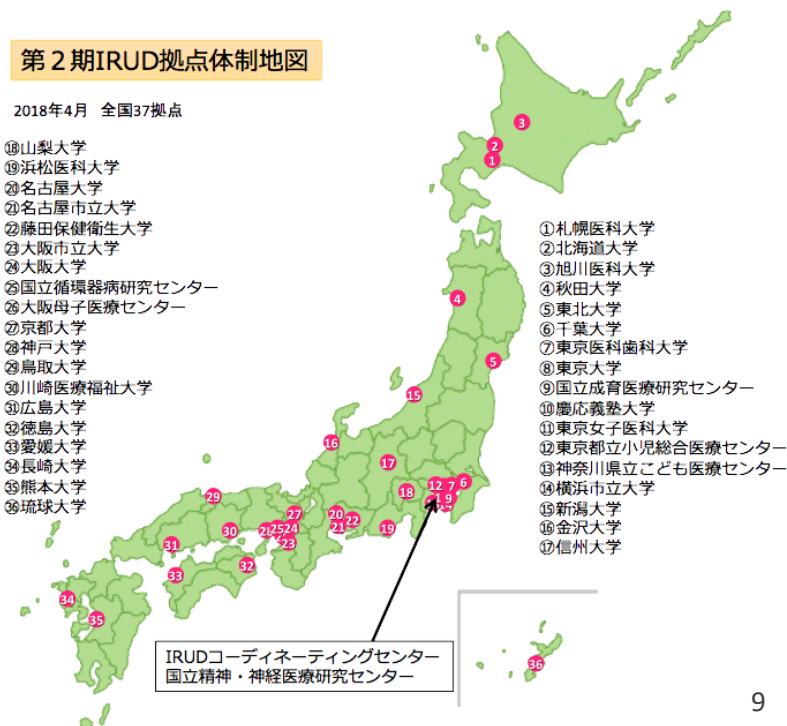


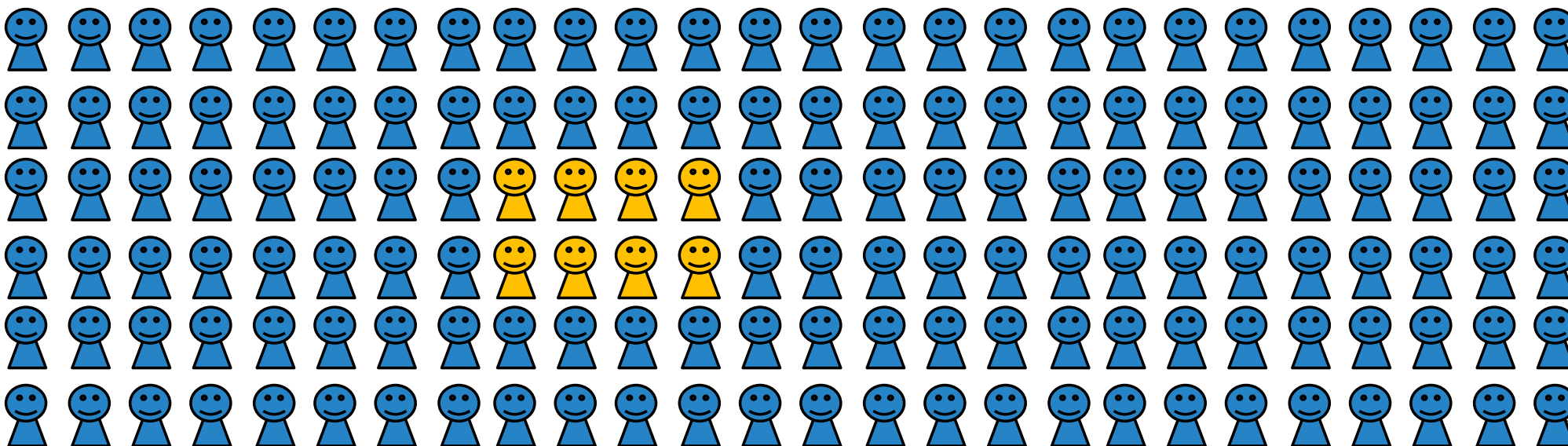
第2期IRUD拠点体制地図

2018年4月 全国37拠点

⑩山梨大学
⑨浜松医科大学
⑧名古屋大学
⑦名古屋市立大学
⑥藤田保健衛生大学
⑤大阪市立大学
④大阪大学
③国立循環器病研究センター
②大阪母子医療センター
①京都大学
③神戸大学
②鳥取大学
①川崎医療福祉大学
④広島大学
③徳島大学
②愛媛大学
④長崎大学
③熊本大学
②琉球大学

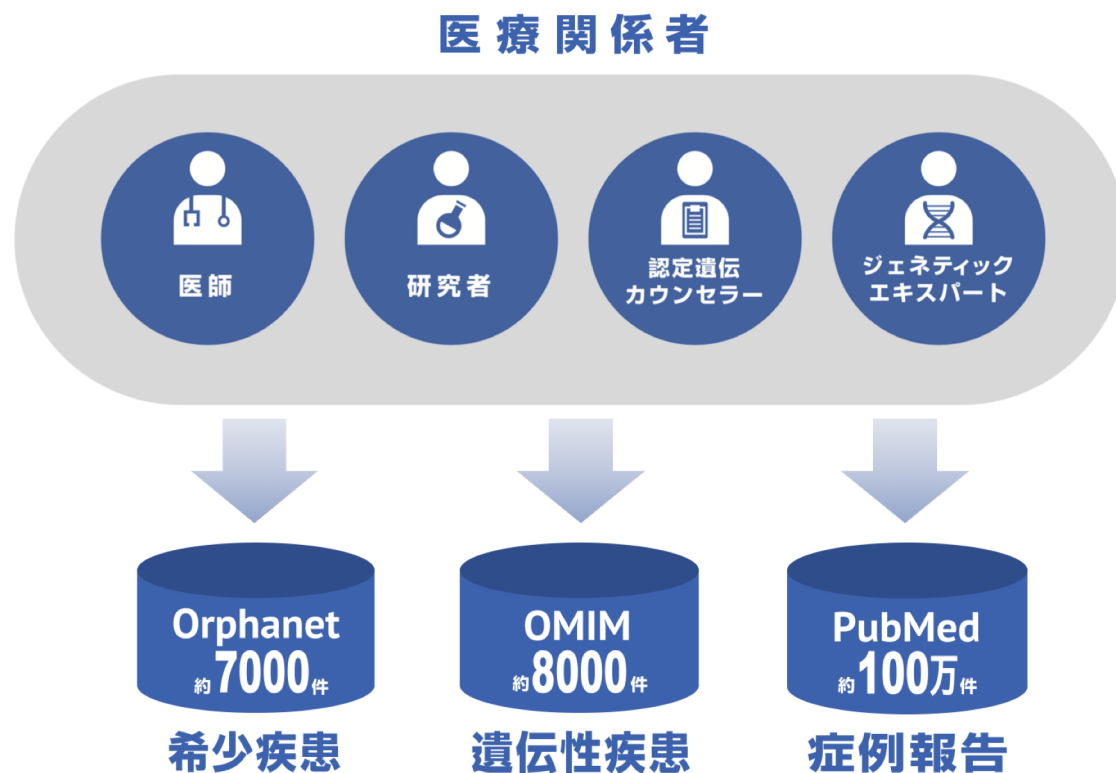
①札幌医科大学
②北海道大学
③旭川医科大学
④秋田大学
⑤東北大学
⑥千葉大学
⑦東京医科歯科大学
⑧東京大学
⑨国立成育医療研究センター
⑩慶応義塾大学
⑪東京女子医科大学
⑫東京都立小児総合医療センター
⑬神奈川県立こども医療センター
⑭横浜市立大学
⑮新潟大学
⑯金沢大学
⑰信州大学





背景・目的

- ❑ 疾患数の多さに比較して各疾患の頻度が低い中で、一人の医療者が十分な症例数の診療経験を積むことは困難であり、症状の組み合わせから診断に至るのは容易ではない（Kosaki, 2017）
- ❑ 早期の診断には、医療関係者が**患者の症状**を元に、**疾患情報**や**過去の症例**を容易に検索できる環境整備が重要（Sebastian, 2016）



PubCaseFinder

患者の 徴候・症状 を入力

HP:0001878 溶血性貧血

HP:0000000

疾患を絞り込む

結果の要約をダウンロード

希少疾患 (Orphanet) 4,066 件

合計: 4,066 件

順位 (類似度)	疾患名 (疾患ID)
1 (100.0%)	Hereditary spherocytosis (ORDO:822)
1 (100.0%)	Sickle cell anemia (ORDO:232)
3 (97.9%)	Alpha-thalassemia (ORDO:846)

PubCaseFinder

患者の 徴候・症状 を入力

HP:0001878 溶血性貧血

HP:0000000

疾患を絞り込む

結果の要約をダウンロード

希少疾患 (Orphanet) 4,066 件

合計: 6969 件

順位 (類似度)	疾患名 (疾患ID)
1 (100.0%)	hereditary spherocytosis typ (OMIM:182900)
2 (91.9%)	hereditary spherocytosis typ (OMIM:616649)
2 (91.9%)	hereditary spherocytosis typ (OMIM:612653)
4 (90.1%)	hemolytic anemia to glucophospha

PubCaseFinderヘルプ利用規約APIお問い合わせ

症例報告徴候・症状疾患原因遺伝子

患者の 徴候・症状 を入力

HP:0001878 溶血性貧血

HP:0000952 Jaundice

HP:0001744 脾腫

HP:0004444 球状赤血球症

HP:0001081 胆石症

症例報告を絞り込む

症例報告を検索クリア

合計: 205 (症例報告)123...21»10 (表示件数)

対応する徴候・症状遺伝子変異キーワード (MeSH)

順位 (類似度)	PMID (PMCID)	
1 (91.9%)	19763011	Hereditary spherocytosis in a 27-year-old woman: case report. Hassan A, Babadoko AA, Isa AH, Abunimye P. Ann Afr Med. 2009;8(1):61-3. [抄録を表示] 黄疸脾腫溶血性貧血球状赤血球症 ヒト女成人栄養補助食品脾腫
2 (90.1%)	27423290	Hereditary Spherocytosis with Splenomegaly and Cholelithiasis in a Young Male of Western Region of Nepal - A Case Report. Ghimire P, Gurung NV, Shrestha S, Poudel SR, Chapagain A. Kathmandu Univ Med J (KUMJ). 2015;13(52):366-8. [抄録を表示] 黄疸胆石症脾腫溶血性貧血
3 (89.3%)	8717295	[Intrathoracic extramedullary hematopoiesis in a case of hereditary spherocytosis]. Takahashi R, Igarashi T, Nakagawa A, Ohuchi H, Nishino M, Murakami S, Yoshida Y, Abe S. Nihon Kyobu Shikkan Gakkai Zassh. 1996;34(1):71-5. [抄録を表示] 胆嚢炎脾腫貧血球状赤血球症 ヒト中年男縦隔腫瘍胸部
4 (86.9%)	26073240	Hereditary Spherocytosis in a Middle-aged Man Complicated with Common Bile Duct Stones. Saitoh T, Hatanaka M, Hasegawa S, Yoshida M, Akizawa T, Minami T, Yamashita K

詳細な疾患情報を取得


Phenotype-Driven Differential-Diagnosis System

[ヘルプ](#)
[利用規約](#)
[API](#)
[お問い合わせ](#)

5 (69.2%)	anemia, congenital dyserythropoietic, type 1a ---- 貧血, 先天性赤血球産生異常性, Ia型	大球性異常造血性貧血 異型赤血球増加症 胎児水腫 脾腫 遷延性新生児黄疸 CDAN1 常染色体劣性遺伝 OMIM:224120 画像検索 (Google) KEGG:H00917 Gene Reviews
7 (67.0%)	cystic fibrosis-gastritis-megaloblastic anemia syndrome ---- 脾嚢胞性線維症-helicobacter pylori 胃炎-巨赤芽球性貧血-知能低下	外分泌性膵不全 巨赤芽球性貧血 胆汁性肝硬変 葉酸欠乏症 常染色体劣性遺伝 , folate deficiency megaloblastic anemia, and intellectual disability. There have been no further descriptions in the literature since 1991. >> 翻訳 (Google) OMIM:219721 画像検索 (Google)
8 (66.7%)	vitamin B12-responsive methylmalonic acidemia type cb1B ---- メチルマロン酸尿症, cb1B型	メチルマロン酸血症 肝腫 貧血 MMAB 常染色体劣性遺伝 An autosomal recessive form of methylmalonic aciduria, caused encoding cob(II)yrinic acid a,c-diamide adenosyltransferase, mit OMIM:251110 画像検索 (Google) Gene Reviews
8 (66.7%)	vitamin B12-responsive methylmalonic acidemia type cb1A ---- メチルマロン酸尿症 cb1A型	メチルマロン酸血症 肝腫 貧血 MMAA 常染色体劣性遺伝 An autosomal recessive form of methylmalonic aciduria, caused encoding MMAA protein. >> 翻訳 (Google) OMIM:251100 画像検索 (Google) Gene Reviews

疾患データベースへのリンク

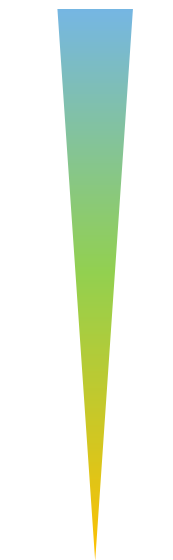
- ☐ OMIM
- ☐ 琉球大学遺伝性疾患DB
- ☐ Orphanet
- ☐ GeneReviews
- ☐ KEGG DISEASE
- ☐ Google画像検索

PubCaseFinder を支える重要なリソース

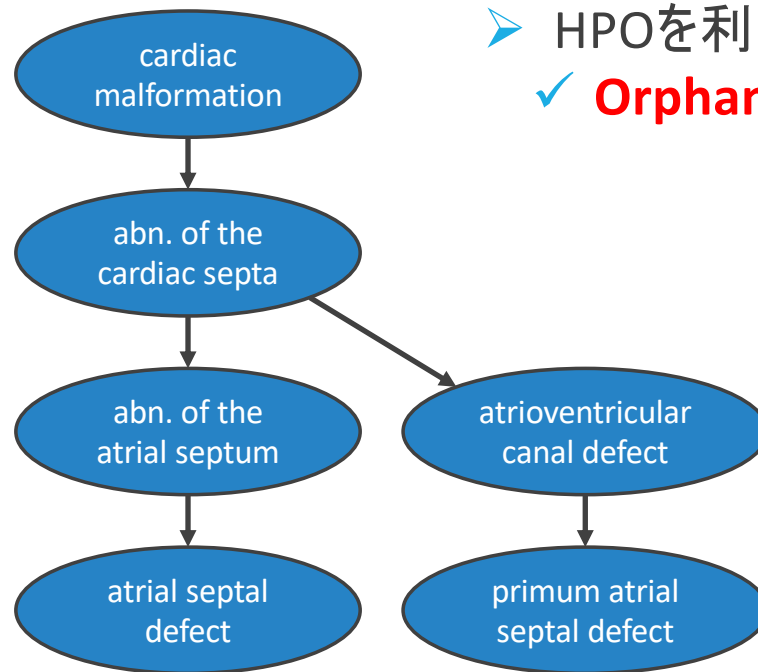
□ Human Phenotype Ontology (HPO)

- 希少・遺伝性疾患に関する症状を**約13,000件**収録 (Sebastian, 2016)

general



specific



- HPOを利用する多数のデータベース (Köhler, 2017)
- ✓ **Orphanet**, **OMIM**, ClinVar, MedGen, GARD... etc.

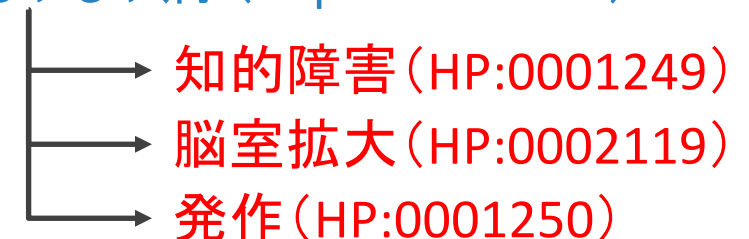


希少疾患



遺伝性疾患

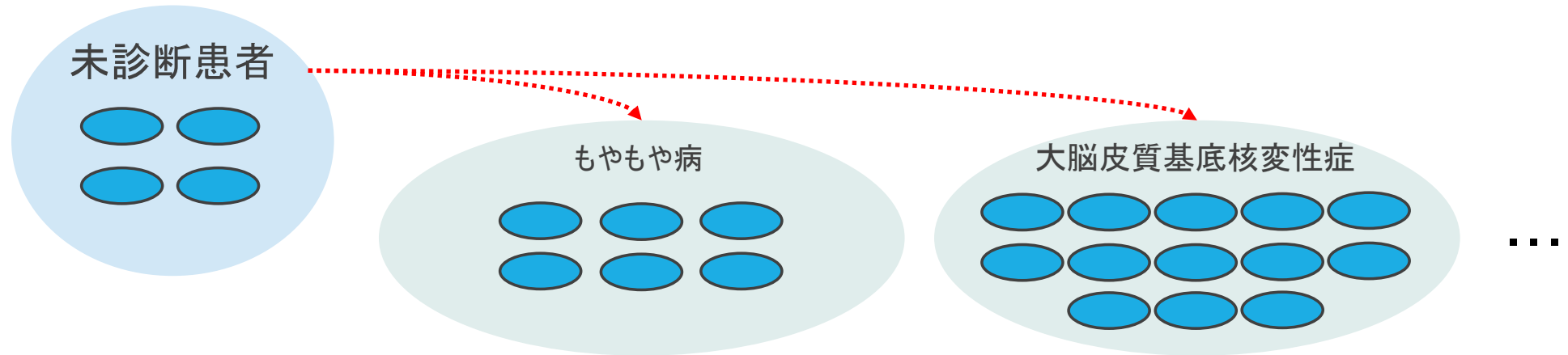
例) もやもや病 (Orphanet:2573)



代表表現: Atria septal defect (心房中隔欠損)
類義語: Atrial septum defect
類義語: ASD

症状セットの類似度計算手法

- 未診断患者と各疾患との類似度を症状セットを元に求める



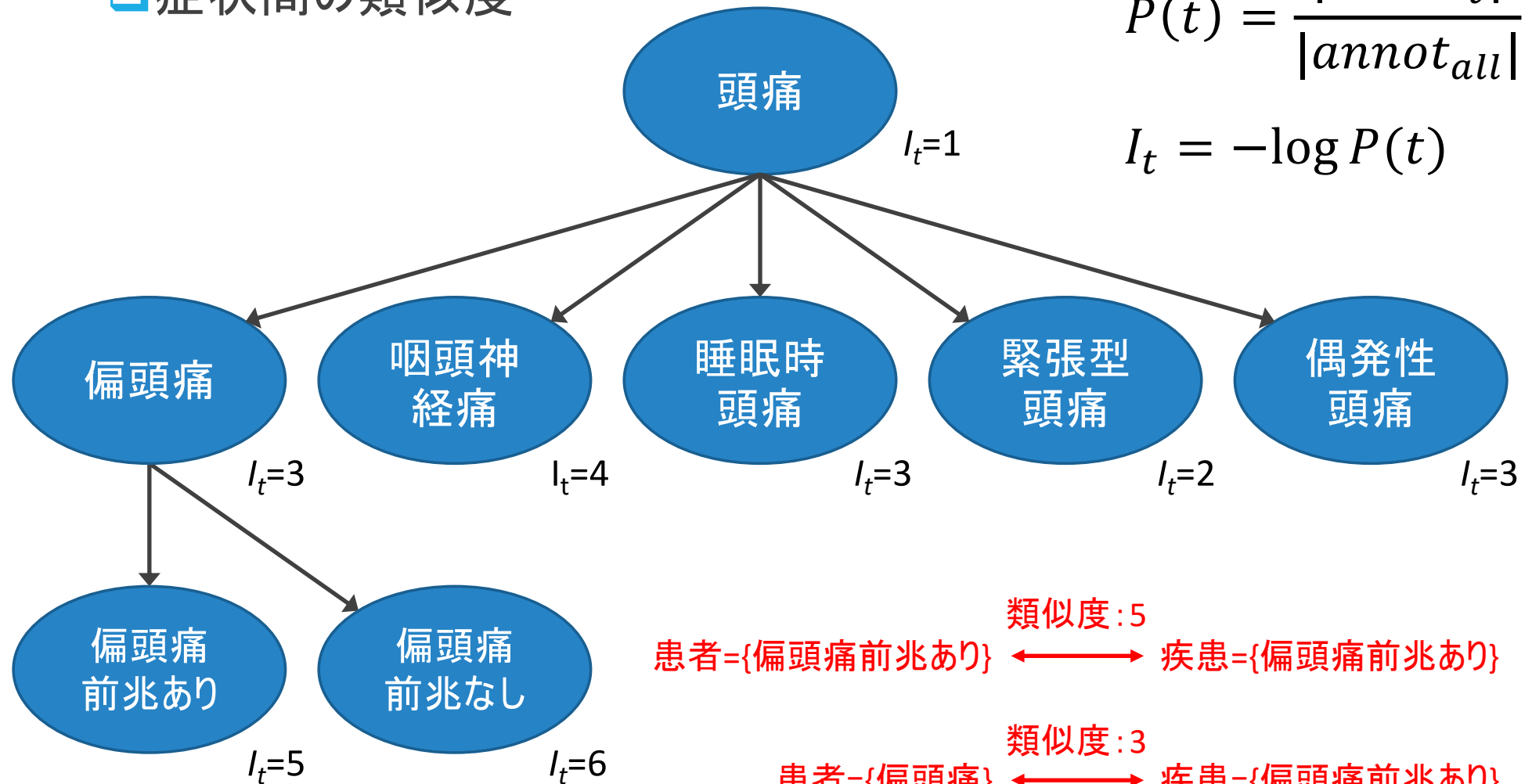
Measure	Equation	Variations	Reference
Resnik(a,b)	$\max_{t \in g^a \cap g^b} IC(t)$	Avg, Max	Rensik, 1995
Lin(a,b)	$\frac{2 * Resnik(a, b)}{IC(a) + IC(b)}$	Avg, Max	Lin, 1998
Jiang-Conrath(a,b)	$\frac{1}{IC(a) + IC(b) - 2 * Resnik(a, b) + 1}$	Avg, Max	Jiang, 1997
simGIC(P,Q)	$\frac{\sum_{t \in g^P \cap g^Q} IC(t)}{\sum_{t \in g^P \cup g^Q} IC(t)}$		Pesquita, 2007
GeneYenta(P,Q)	$\frac{\sum_{t \in T_c} R_t \times \max_{t' \in T_d} sim_{terms}(t, t')}{\sum_{t \in T_c} R_t \times I_t} \times 100$		Gottlieb, 2015

症状セットの類似度計算手法

□ 症状間の類似度

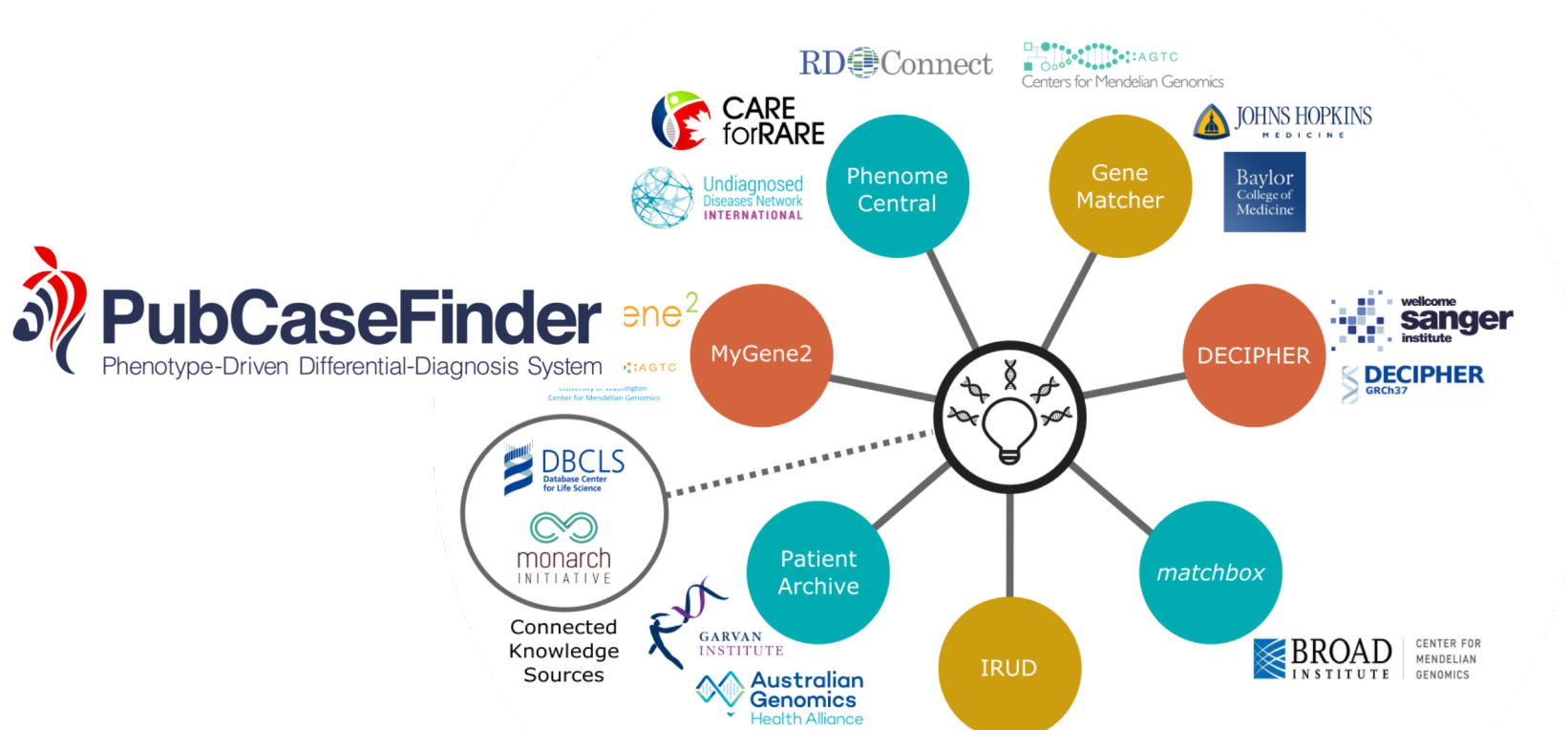
$$P(t) = \frac{|annot_t|}{|annot_{all}|}$$

$$I_t = -\log P(t)$$



HPOと類似度計算手法の活用事例

- 希少疾患は症例数が少ないため、世界規模で症例データを共有する必要がある
- 多くのプロジェクトが症例データ共有プロジェクト Matchmaker Exchangeに参加 (Orion, 2016)



100以上の疾患原因遺伝子が同定された

HPOと類似度計算手法の活用事例

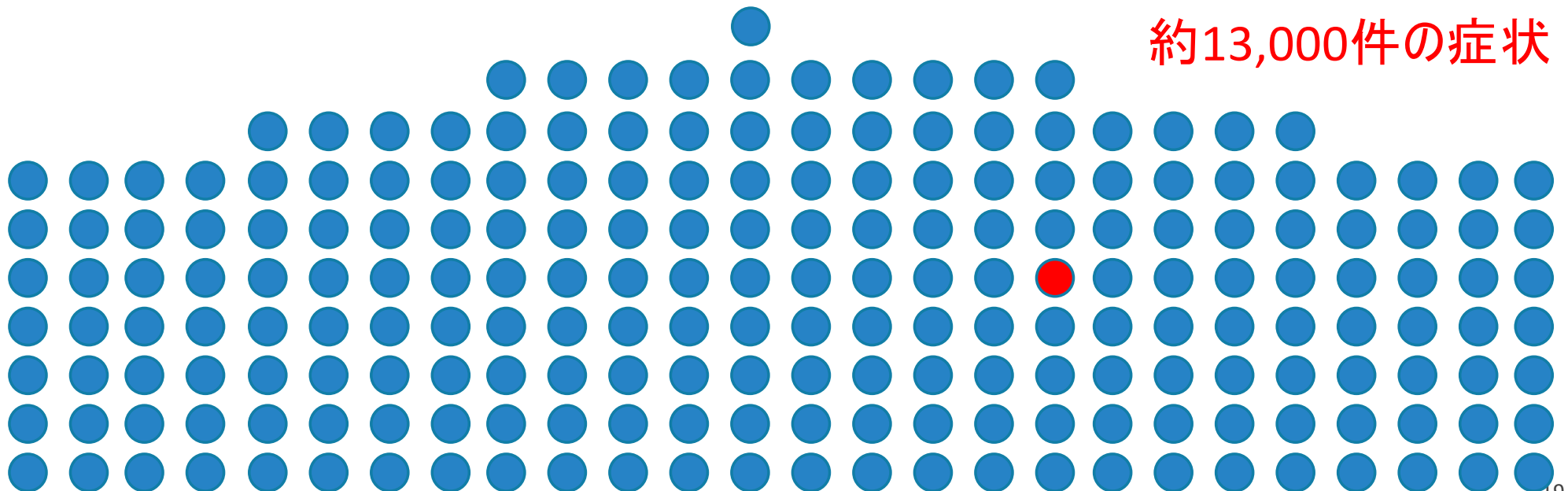
□ 現在、多くの症例データベースが症状をHPOで管理

Name	URL
PhenomeCentral	phenomecentral.org
DDD (Deciphering Developmental Disorders)	www.ddduk.org
DECIPHER (Database of genomic variation and Phenotype in Humans using Ensembl Resources)	decipher.sanger.ac.uk
ECARUCA (European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations)	http://umcecaruca01.extern.umcn.nl:8080/ecaruca/ecaruca.jsp
The 100 000 Genomes Project	https://www.genomicsengland.co.uk/
Geno2MP (Exome sequencing data linked to phenotypic information from a wide variety of Mendelian gene discovery projects)	http://geno2mp.gs.washington.edu
NIH UDP (Undiagnosed Diseases Program)	available via phenomecentral.org
NIH UDN (Undiagnosed Diseases Network)	available via phenomecentral.org
HDG (Human Disease Gene Website series)	www.humandiseasesgenes.com
Phenopolis (An open platform for harmonization and analysis of sequencing and phenotype data)	https://phenopolis.github.io
GenomeConnect (Patient portal developed by ClinGen (67))	www.genomeconnect.org
FORGE Canada & Care4Rare Consortium	available via phenomecentral.org
RD-Connect	platform.rd-connect.eu
Genesis	thegenesisprojectfoundation.org

HPO利用の問題点



顔	?
HP:0010628 顔面麻痺	
HP:0007209 顔面麻痺	
HP:0010828 片側顔面スパスム	
HP:0000324 顔面非対称	
HP:0000282 顔面浮腫	
HP:0011331 片側顔面萎縮	



HPO利用の問題点

□ HPO用語探索支援ツール「PhenoTouch」を開発

PubCaseFinder
Phenotype-Driven Differential-Diagnosis System

(最終更新日: 2018年9月25日) ヘルプ 利用規約 API お問い合わせ

患者の徴候および症状 JPN ▼

HP:0001009 毛細血管拡張 × HP:0001249 知的障害 × HP:0001250 発作 × HP:0002072 舞蹈病 × HP:0002315 頭痛 ×

Clear OK Cancel

上位概念

神経系生理の異常 26 追加 置換

HP:0001009 毛細血管拡張 ×

HP:0002315 頭痛 ×

HP:0002315 頭痛

追加 置換

HPO Id : HP:0002315

症状(日) : 頭痛

症状(英) : Headache

症状定義 : Cephalgia, or pain sensed in various parts of the head, not confined to the area of distribution of any nerve.

Synonym : Headaches

下位概念

追加 置換 0 Thunderclap headache

追加 置換 2 偏頭痛

追加 置換 0 喉頭神経痛

追加 置換 0 睡眠時頭痛

追加 置換 0 緊張型頭痛

追加 置換 0 群発(性)頭痛

追加 置換 0 頭痛 (褐色細胞腫を伴う)

HPO利用の問題点

- ヒトの3Dモデルをタッチすると、該当箇所に関連する症状の一覧を得ることができる

tokyogw.bits.cc

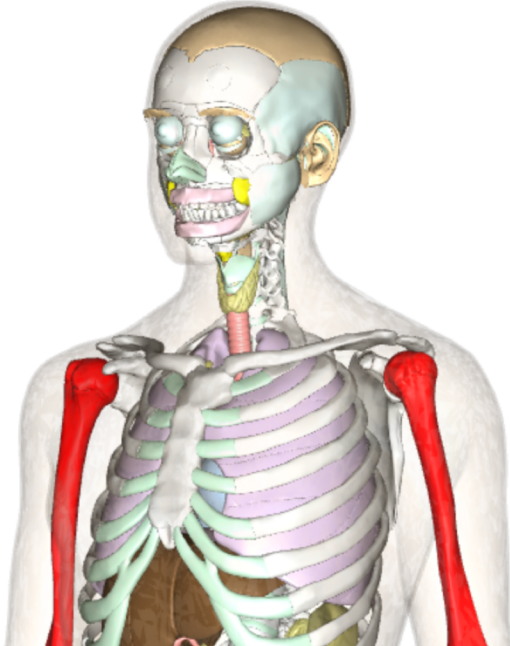
患者の徴候および症状 JPN ▼

HP:0001090 大きな眼 × HP:0009938 落ちくぼんだ頬部 × HP:0002057 目立つ眉間 × HP:0000885 幅広い肋骨 ×

Clear OK Cancel

身体各部位を選択 Close

☒ Internal ☒ Bone ☐ Muscle



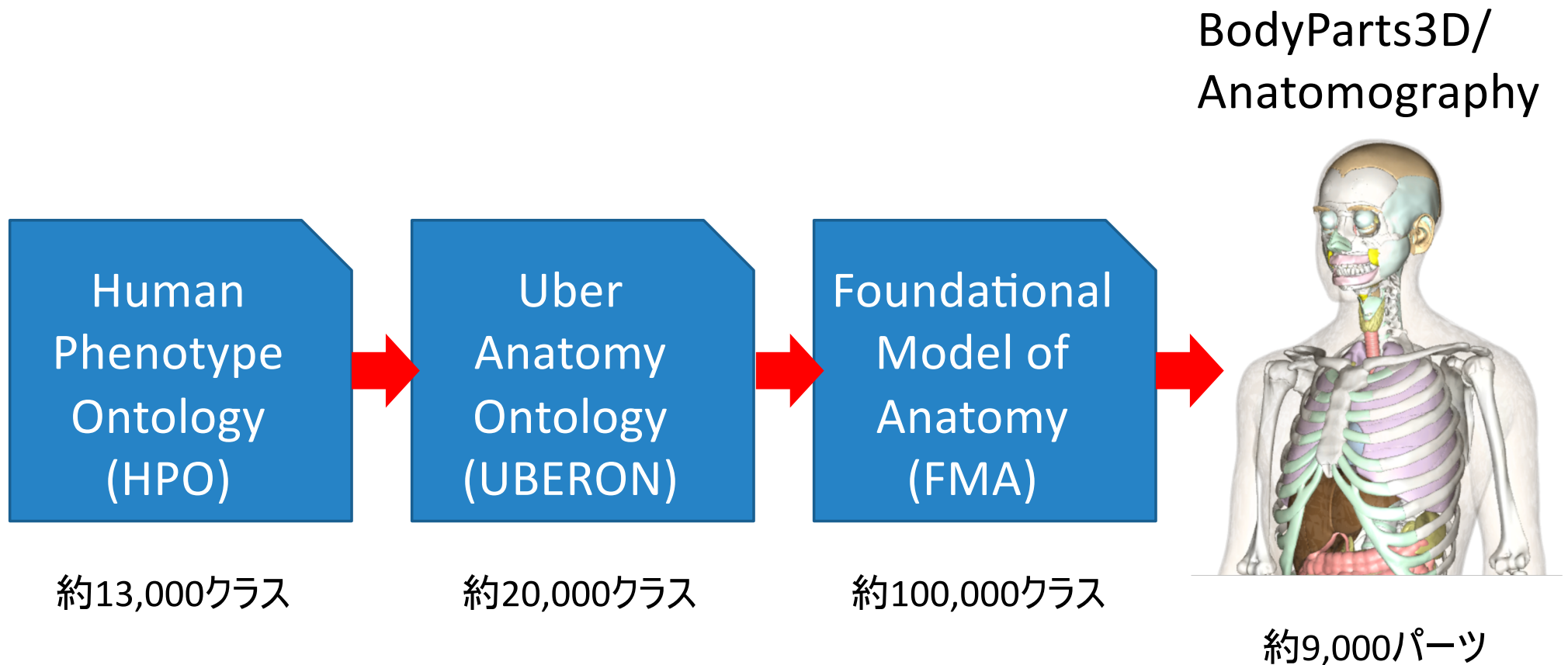
11 長骨
0 上腕骨

FMA:13303 上腕骨

追加	置換	0	三角形の上腕骨
追加	置換	11	上腕骨の異常
追加	置換	3	上腕骨無形成/低形成
追加	置換	3	上腕骨骨化異常
追加	置換	0	上腕骨骨粗鬆症
追加	置換	0	二分した上腕骨
追加	置換	0	先細りの上腕骨
追加	置換	1	変形した上腕骨
追加	置換	0	幅広い上腕骨

HPO利用の問題点

- オントロジーのクロスリファレンスを利用して、HPO termをヒトの Body Partsに対応付けた



HPO利用の問題点

pubcasefinder.dbcls.jp

患者の徴候および症状 JPN ▼

HP:0001903 貧血 × HP:0000952 Jaundice × HP:0000952 黄疸 ×

Clear OK Cancel

上位概念

赤血球の異常 15 追加 置換

HP:0001903 貧血

追加 置換

HPO Id : HP:0001903

症状(日) : 貧血

症状(英) : Anemia

症状定義 : A reduction in erythrocytes volume or hemoglobin concentration.

Synonym: Anaemia, Decreased hemoglobin

下位概念

追加 置換 0 一過性erythroblastopenia

追加 置換 7 不適切な産生貧血

追加 置換 2 赤血球寿命減少による貧血

追加 置換 0 難治性貧血

HP:0001903 貧血 × HP:0000952 黄疸 ×

ex) 毛細血管拡張、知的障害、発作、舞踏病、頭痛

問題点2: 疾患－症状関連情報の不足

- Orphanetの約7,000の希少疾患のうち、**症状が割り当てられているのは約2,500疾患のみ**

【順位】	【類似度】	【疾患】	【症状セット】			
1	0.9	ファブリー病	下痢	発疹	嘔吐	...
2	0.8	アジソン病	脱力感	筋力低下	易疲労感	...
3	0.6	遺伝性膵炎	腹痛	脂肪便	体重減少	...
4	0.5	ウィルソン病	黄疸	浮腫	食欲不振	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮
2478	0.1	カナバン病	痙攣	症状が少ない		

Orphanet
の約60%

Orphanet

出所: Orphadata: Free access data
from Orphanet. © INSERM 1997.

疾患－症状
関連情報

← 不足

ランキング
対象外

ペーテース異常

症状がない

例：もやもや病の症状

□ Orphanet: 5症状

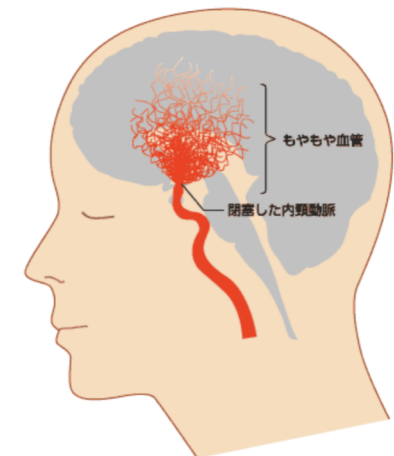
- 脳血管の異常、毛細管拡張症、脳室拡大、てんかん発作、知的障害

□ 難病情報センター：もやもや病（指定難病22）

出所：難病情報センター
<http://www.nanbyou.or.jp/entry/209>

○ 診断・治療方針（医療従事者向け）：症状

- 意識障害、脱力発作、感覚異常、不随意運動、頭痛、嘔気、運動障害、言語障害、視野障害など



もやもや病

出所：東海大学医学部脳神経外科
<http://neurosurgery.med.u-tokai.ac.jp/edemiru/moyamoya/>

文献から疾患に関連する症状を抽出

- 日々大量に出版される文献から手動で情報を抽出するのは限界がある
 - 自動抽出を試みる



3000報以上/日

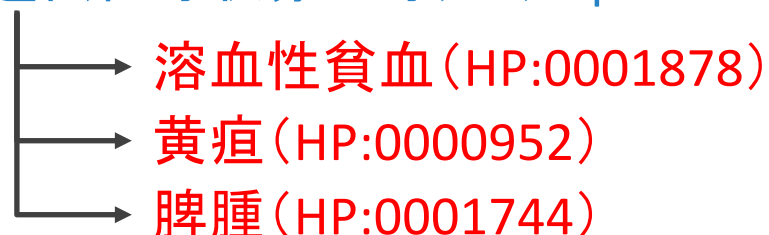
文献から疾患に関連する症状を抽出

□ 課題:「疾患－症状」関連データの不足が大きな課題

➤ 約100万件の症例報告を対象に、テキストマイニング技術で自動取得



例) 遺伝性球状赤血球症 (Orphanet:822)



症例報告 PMID: 12355853

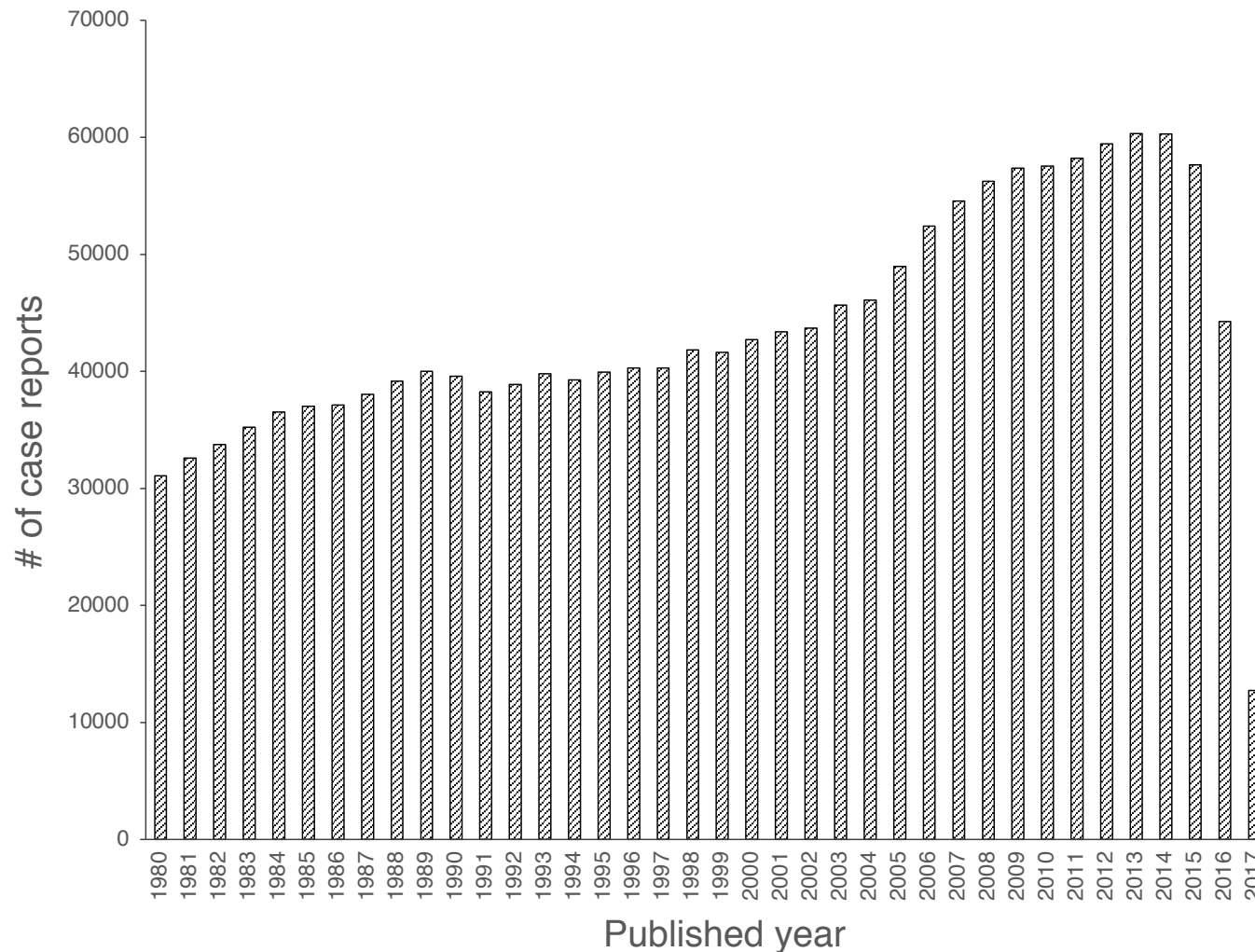
青: 疾患名 赤: 症状

Hereditary spherocytosis is a genetic, frequently familial hemolytic blood disease characterized by varying degrees of hemolytic anemia, splenomegaly, and jaundice. . . .

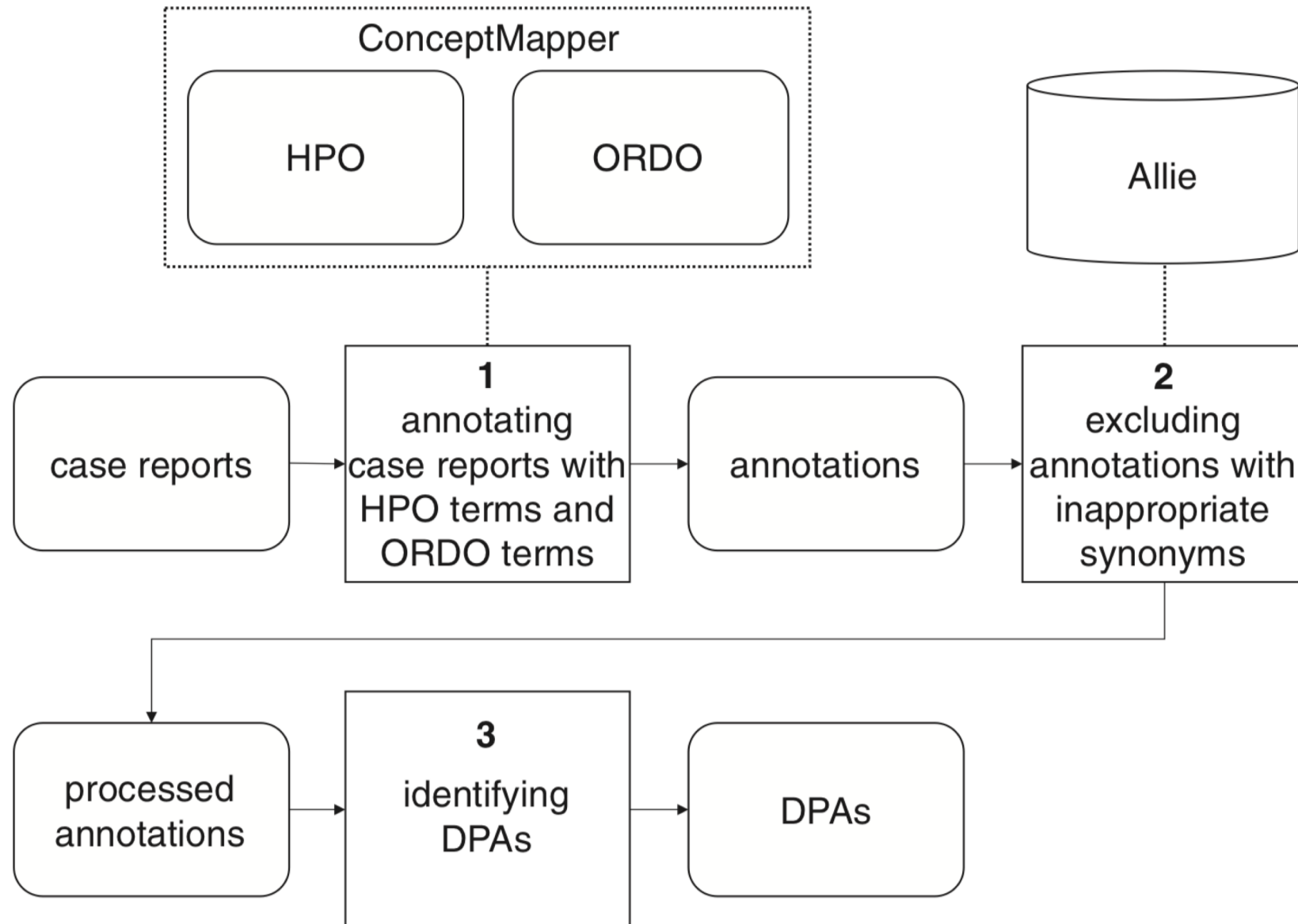


症例報告出版数

- Distribution of the number of case reports published per year in PubMed from 1980 to 2017



文献から疾患に関連する症状を抽出



文献から疾患に関連する症状を抽出

□ オントロジーを用いたアノテーションツール

- ConceptMapper (Tanenblatt, 2010)、MetaMap (Aronson, 2010)、NCBO Annotator (Jonquet, 2009)

□ CRAFT Corpus を利用した、パフォーマンス比較 (Christopher, 2014)

- 8つのオントロジーにおいてF-measureを比較結果、7つのオントロジーでConceptMapperのF-measureが最も高かった

□ HPO gold standard (Tudor, 2015)を用いたツール評価 (藤原, JSAI2017)

System	F-measure	Precision	Recall
NCBO Annotator	0.51	0.54	0.47
MetaMap	0.56	0.51	0.61
ConceptMapper	0.52	0.52	0.51

System	Processing time (sec)
NCBO Annotator	206.0
MetaMap	351.0
ConceptMapper	4.3

100万件の症例報告
の処理に要する時間



17.7 (day)

5.2 (hour)

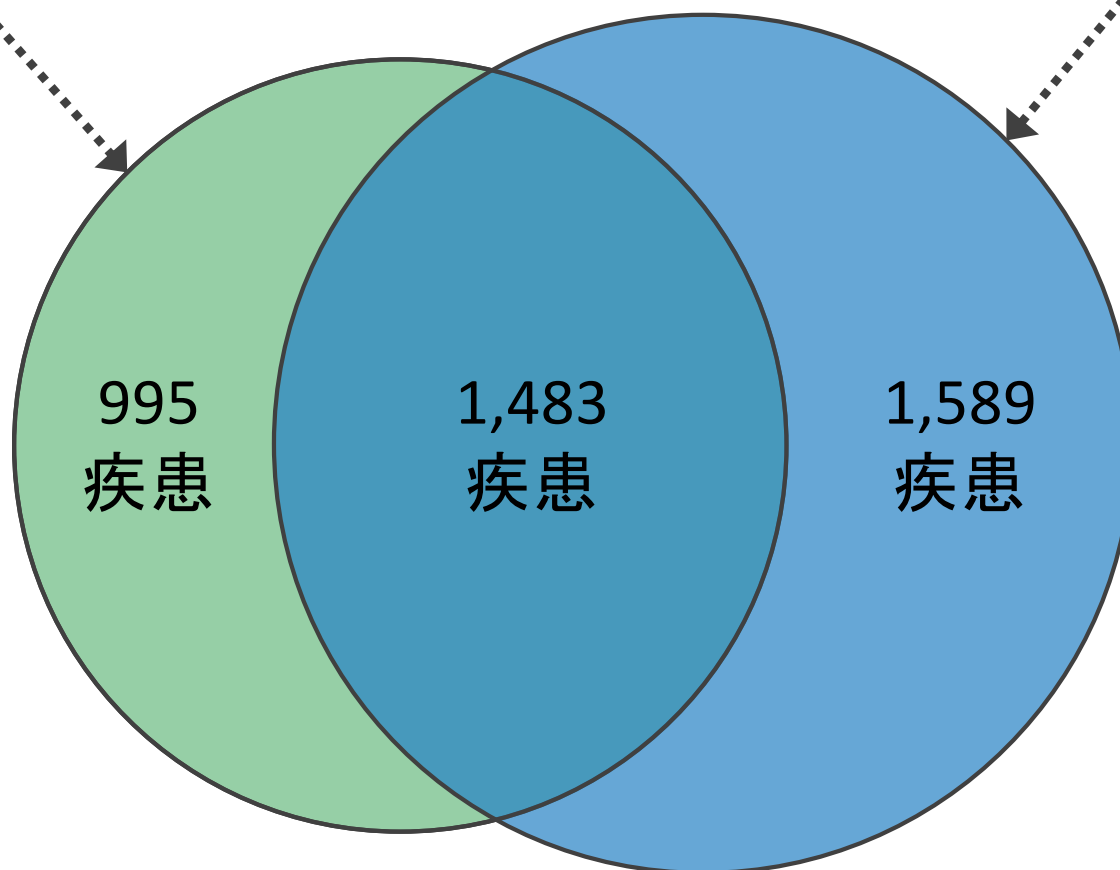
疾患－症状関連情報抽出 結果

【Orphanet】

2,478疾患に症状を付与

【症例報告】

3,072疾患に症状を付与



PubCaseFinderで疾患－症状関連情報を確認

PubCaseFinder
Phenotype-Driven Differential-Diagnosis System

ヘルプ 利用規約 API お問い合わせ

遺伝性疾患 (OMIM) 7,021 件 希少疾患 (Orphanet) 4,057 件

合計: 4,057 件

1 2 3 ... 406 »

10 (表示件数)

順位 (類似度)	疾患名
1 (86.6%)	Hereditary spherocytosis ----- 遺伝性球状赤血球症

溶血性貧血 球状赤血球症 脾腫 薬物

ANK1 EPB42 SLC4A1 PTPN11 S

Hereditary spherocytosis is a congenital symptom-free carriers to severe hemolytic anemia and cholelithiasis. >> 翻訳 (Google)

Orphanet:822 画像検索 (Google)

症状をクリック

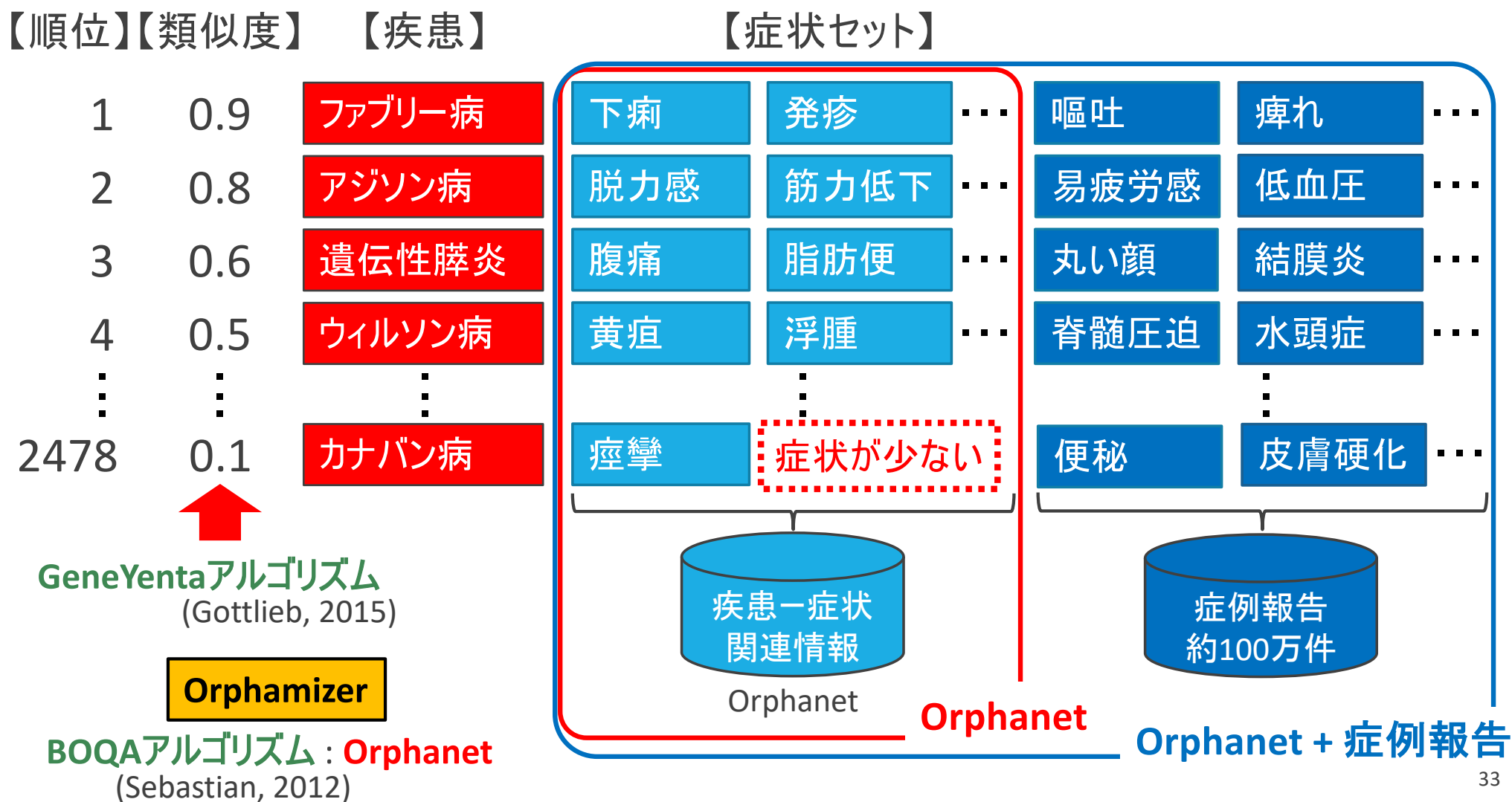
PubCaseFinder
Phenotype-Driven Differential-Diagnosis System

ヘルプ 利用規約 API お問い合わせ

27108201	MIXED_SAMPLE	Adult
Disease-modifying influences of coexistent G6PD-deficiency, Gilbert syndrome and deletional alpha thalassemia in hereditary spherocytosis: A report of three cases.		
Jamwal M, Aggarwal A, Kumar V, Sharma P, Sachdeva MU, Bansal D, Malhotra P, Das R. Clin Chim Acta. 2016;458:51-4.		
Hereditary spherocytosis (HS) is a common inherited hemolytic anemia characterized by heterogeneous clinical presentations with variable degrees of anemia, jaundice, splenomegaly and gallstones.		
27566068 (5002123)	MALE	Middle Aged
Open-heart surgery using a centrifugal pump: a case of hereditary spherocytosis.		
Matsuzaki Y, Tomioka H, Saso M, Azuma T, Saito S, Aomi S, Yamazaki K. J Cardiothorac Surg. 2016;11(1):138.		
Hereditary spherocytosis is a genetic, frequently familial hemolytic blood disease characterized by varying degrees of hemolytic anemia, splenomegaly , and jaundice.		
27906107 (5134285)	FEMALE	Adult
Previously undiagnosed hereditary spherocytosis in a patient with jaundice and pyelonephritis: a case report.		
Tateno Y, Suzuki R, Kitamura Y. J Med Case Rep. 2016;10(1):337.		
The classical clinical features of hereditary spherocytosis are anemia, jaundice, and splenomegaly .		
27423290	MALE	Child
Hereditary Spherocytosis with Splenomegaly and Cholelithiasis in a Young Male of Western Region of Nepal - A Case Report.		
Ghimire P, Gurung NV, Shrestha S, Poudel SR, Chapagain A. Kathmandu Univ Med J (KUMJ). 2015;13(52):366-8.		

取得した疾患－症状関連情報を疾患ランキングシステムに活用

□ 症例報告から取得した疾患－症状関連情報は、ランキング精度の向上に寄与するか？



疾患ランキング精度の評価方法

□ カナダCare4Rareから症状から診断がつかなかった135症例を評価用に頂いた

○ 診断結果 (ORDO ID)、症状 (HPO IDs)

□ Recallで疾患ランキング精度を評価

	症例 1		症例 2		症例 3	
	順位	疾患名	順位	疾患名	順位	疾患名
	1	疾患A	1	疾患A	1	疾患A
診断結果 と一致	2	疾患B	2	疾患B	2	疾患B
	3	疾患C	3	疾患C	3	疾患C
	4	疾患D	4	疾患D	4	疾患D
Top5	5	疾患E	5	疾患E	5	疾患E
	6	疾患F	6	疾患F	6	疾患F
	⋮	⋮	⋮	⋮	⋮	⋮
	2478	疾患Z	2478	疾患Z	2478	疾患Z

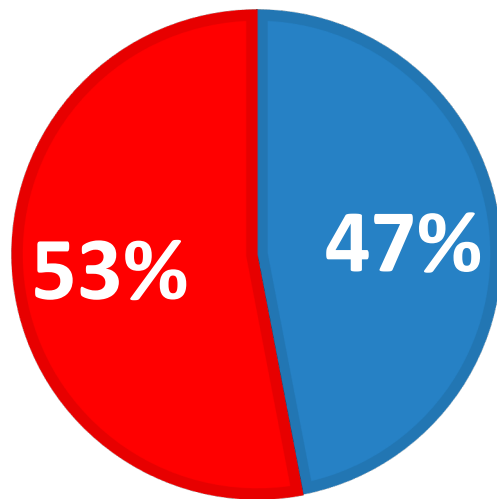
$$RecallX = \frac{TopX \text{ にランキングした症例数}}{\text{全症例数}}$$

疾患ランキング精度の評価

□ 評価用の135症例を利用して、トップ5のRecallを比較

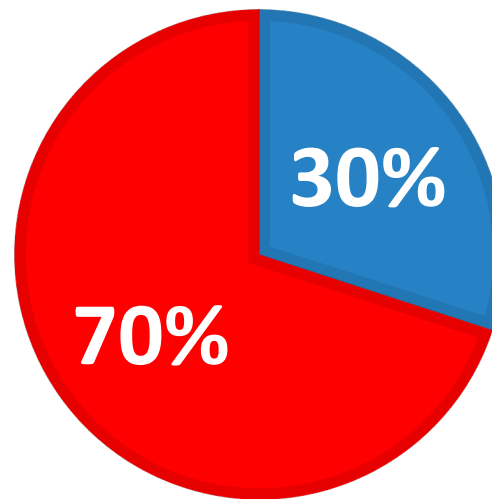
PubCaseFidner
(Orphanet + 症例報告)

■ 正解症例 ■ 不正解症例



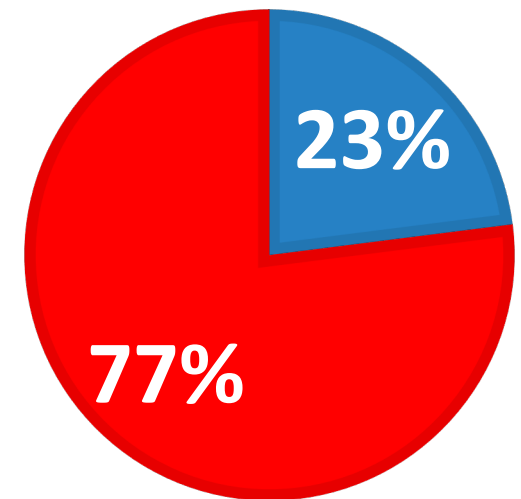
PubCaseFidner
(Orphanet)

■ 正解症例 ■ 不正解症例



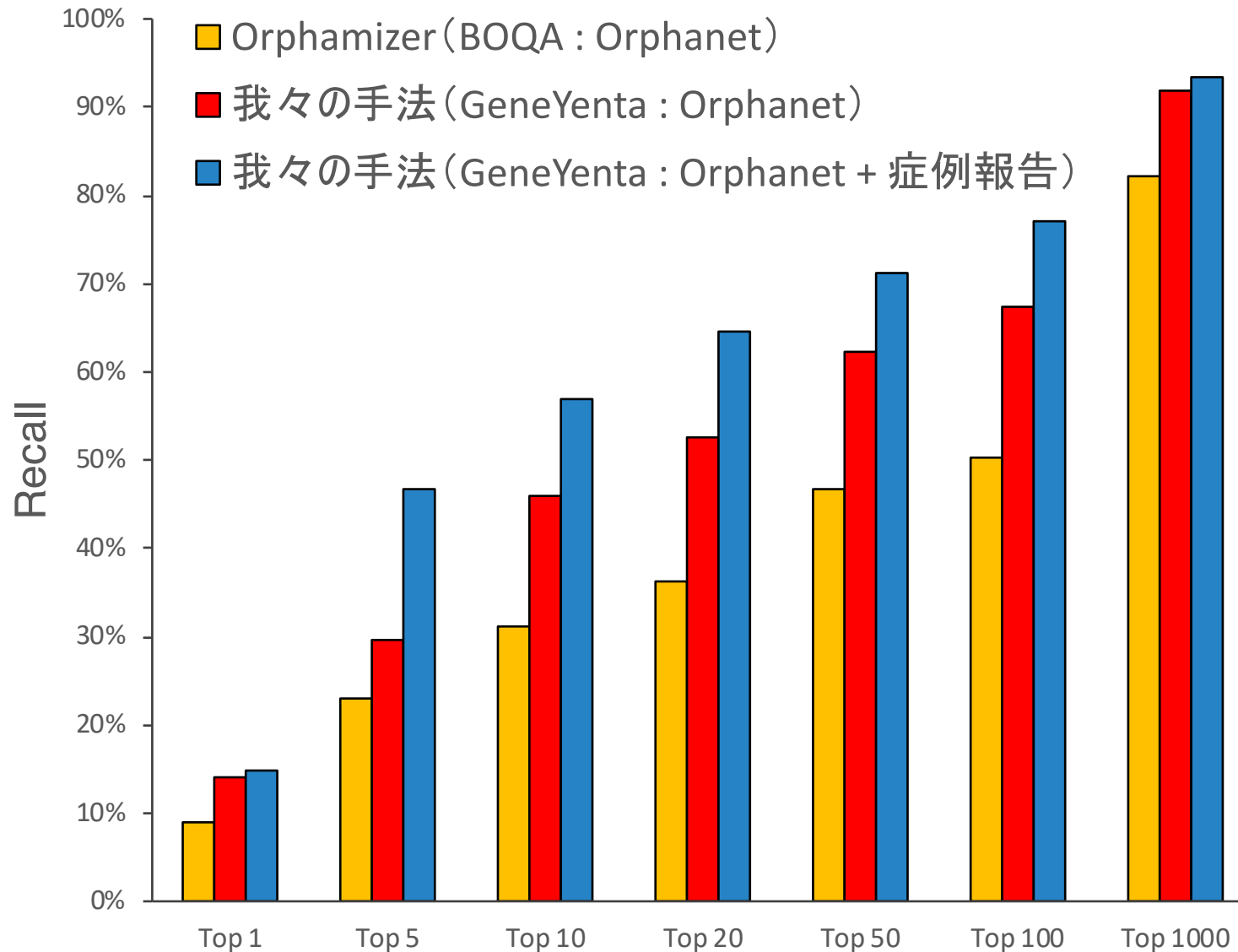
Orphamizer
(Orphanet)

■ 正解症例 ■ 不正解症例



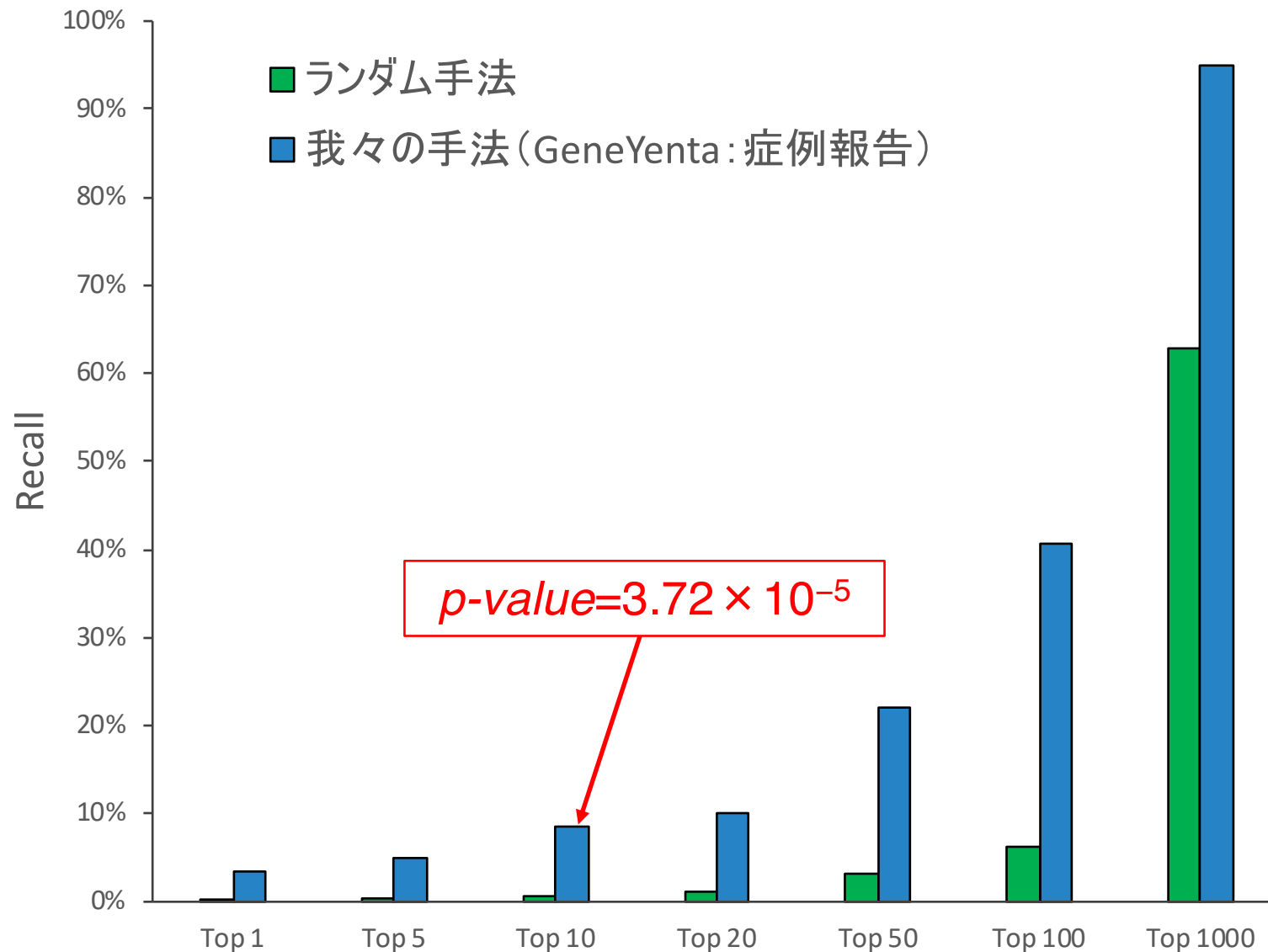
疾患ランキング精度の評価

□ 評価用の135症例を利用して、Recall1,5,10,20,50,100,1000を比較

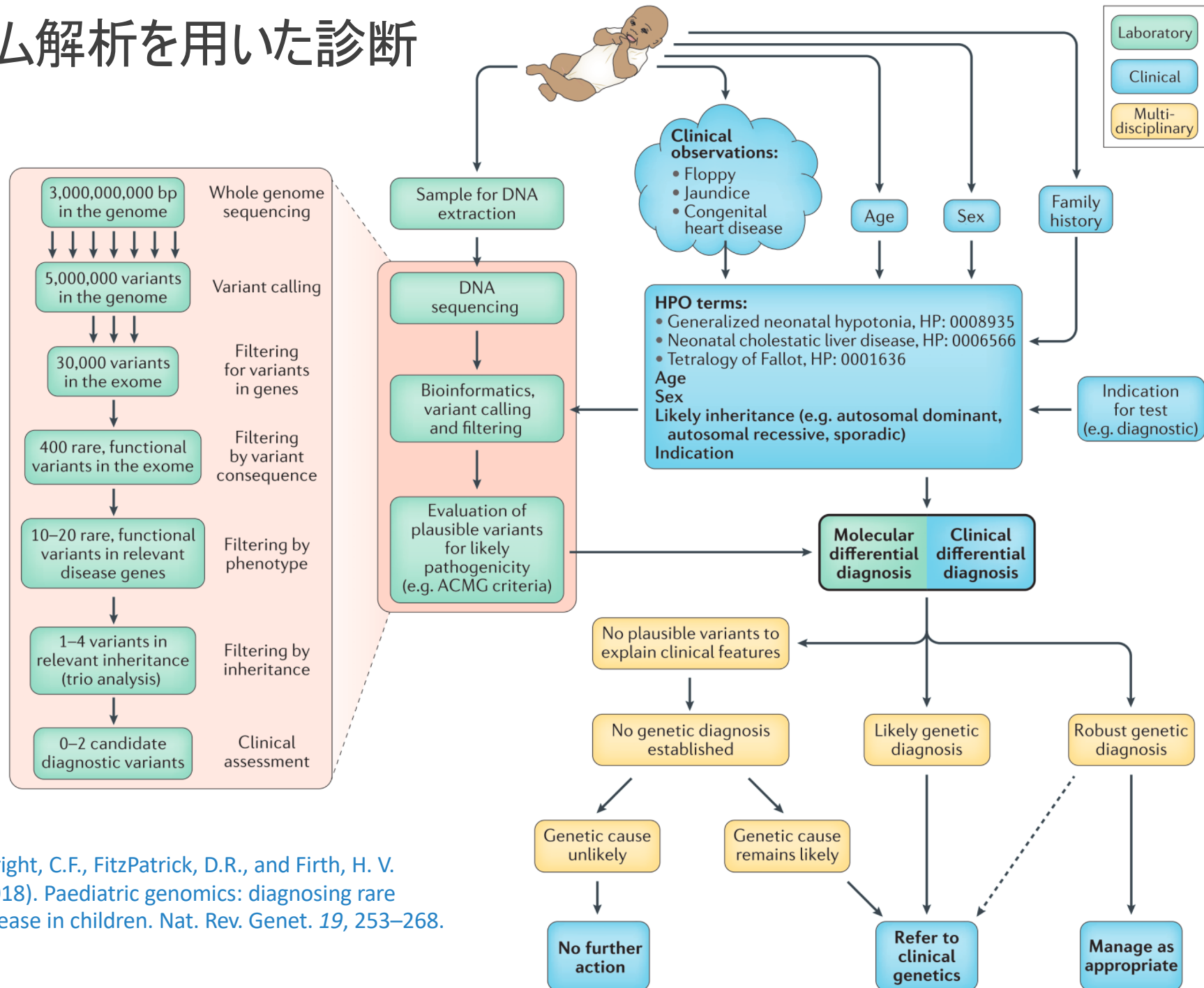


疾患ランキング精度の評価

□ 評価用の59症例を利用して、Recall1,5,10,20,50,100,1000を比較



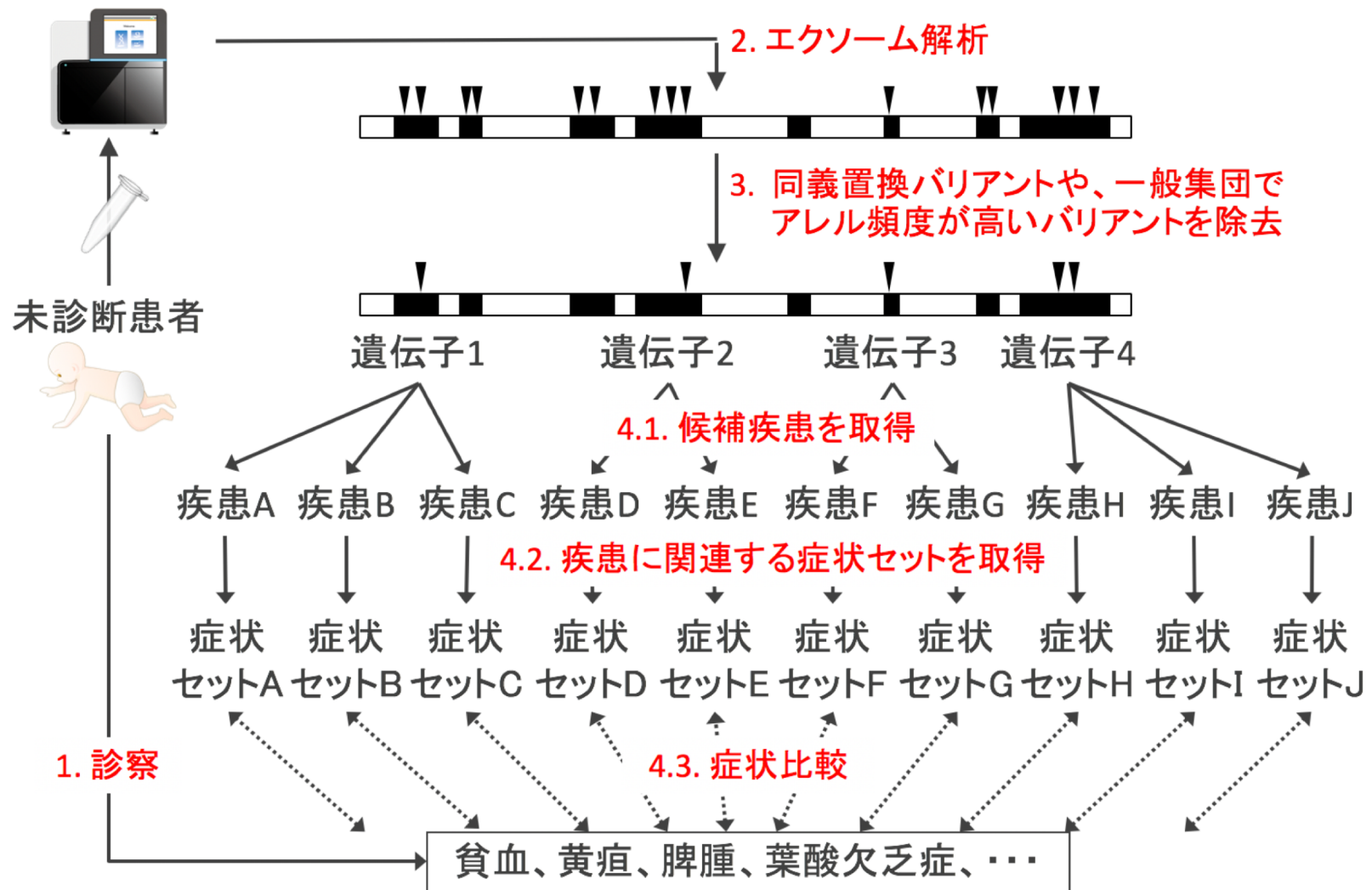
ゲノム解析を用いた診断



出所: Wright, C.F., FitzPatrick, D.R., and Firth, H. V. (2018). Paediatric genomics: diagnosing rare disease in children. Nat. Rev. Genet. 19, 253–268.

ゲノム解析を用いた診断にPubCaseFinderを活用

□ 患者の症状と候補バリエントを入力する



ゲノム解析を用いた診断にPubCaseFinderを活用

患者の 徴候・症状 を入力 + Upload File (HPO ID):

HP:0001878 溶血性貧血 × HP:0000952 Jaundice × HP:0001744 脾腫 × HP:0004444 球状赤血球症 × HP:0001081 胆石症 ×

疾患を絞り込む + Upload File (Entrez Gene ID):

ENT:286 ANK1 (SPH1) × ENT:2038 EPB42 (MGC116735 | MGC116737 | PA) ×
ENT:6521 SLC4A1 (CD233 | FR | RTA1A | SW | WR) × ENT:6708 SPTA1 (EL2) × ENT:6710 SPTB × ENT:59 ACTA2 (ACTSA) ×
ENT:57674 RNF213 (KIAA1554 | NET57) × ENT:7403 KDM6A × ENT:8085 KMT2D (ALR | CAGL114 | MLL4) ×
ENT:5906 RAP1A (KREV-1 | SMGP21) × ENT:5908 RAP1B (K-REV | RAL1B) ×

結果の要約をダウンロード

疾患を検索

クリア

希少疾患 (Orphanet) 9 件

遺伝性疾患 (OMIM) 16 件

合計: 9 件

1

10 (表示件数)

結果をダウンロード

順位 (類似度)	疾患名 (疾患ID)	対応する徴候・症状	疾患原因遺伝子
1 (100.0%)	Hereditary spherocytosis (ORDO:822)	溶血性貧血 球状赤血球症 胆石症 脾腫 黄疸	ANK1 EPB42 SLC4A1 SPTA1 SPTB
Hereditary spherocytosis is a congenital hemolytic anemia with a wide clinical spectrum (from symptom-free carriers to severe hemolysis) characterized by anemia, variable jaundice, splenomegaly and cholelithiasis. >> 翻訳 (Google) 画像検索 (Google) 症例報告検索			
2 (81.3%)	8p11.2 deletion syndrome (ORDO:251066)	仙骨部陥凹 溶血性貧血 球状赤血球症 脾腫	ANK1
8p11.2 deletion syndrome is a contiguous gene syndrome characterized by the association of congenital spherocytosis, dysmorphic features, growth delay and hypogonadotropic hypogonadism. >> 翻訳 (Google) 画像検索 (Google) 症例報告検索			
3 (80.1%)	Hereditary elliptocytosis (ORDO:288)	楕円赤血球症 溶血性貧血 黄疸	

□ 遺伝子-疾患関連情報

○ OMIM: 6298ペア

➤ MedGenから取得

○ Orphanet: 7502ペア

➤ Orphadataから取得

ゲノム解析を用いた診断にPubCaseFinderを活用

- 東京医科歯科大学の宮冬樹先生のところで、疾患遺伝子変異解析の専門家と臨床医が候補遺伝子を絞り込むことでpathogenicなバリエーションを同定した17検体(論文で報告済み)を用いて、PubCaseFinderとPhenIXの精度を比較した

患者		NGS解析			
臨床所見	臨床所見 (HPO)	解析方法	解析パイプライン後の候補バリエーション数	論文で報告した疾患原因遺伝子	報告論文
小脳低形成	HP:0001321,小脳低形成	エクソーム解析 (Trio)	34	<i>HNRNPH2</i>	Bain JM. et al. Am. J. Hum. Genet. 99.3, 728-734 (2016).
てんかん 視神経萎縮	HP:0001250,発作 HP:0000648,視神経萎縮	エクソーム解析 (Trio)	26	<i>RTN4IP1</i>	Okamoto et al. J. Hum. Genet. 62, 927-929 (2017).
皮膚弛緩症	HP:0000973,弛緩性皮膚	エクソーム解析 (Trio)	27	<i>ALDH18A1</i>	Nozaki et al. Brain Dev. 38, 678-684 (2016).
白質異常 皮質盲	HP:0002500,大脳白質の異常 HP:0100704,Cortical visual impairment	エクソーム解析 (Trio)	36	<i>KIF1A</i>	Okamoto N. et al. J. Hum. Genet. 59, 639-641 (2014).
低身長 精神運動発達障害 骨異常	HP:0004322,低身長 HP:0002361,精神運動発達悪化 HP:0000924,骨格の異常	エクソーム解析 (Trio)	16	<i>FAM20C</i>	Tamai et al. Am. J. Med. Genet. A, 176, 682-686 (2018).
macrocephaly cortical dysplasia	HP:0000256,Macrocephaly HP:0002539,Cortical dysplasia	エクソーム解析 (Trio)	36	<i>AKT3</i>	Negishi Y. et al. BMC Med. Genet. 18, 4 (2017).
水無脳症 小脳低形成 脳幹低形成	HP:0002324,水無脳症 HP:0001321,小脳低形成 HP:0002365,脳幹低形成	エクソーム解析 (Trio)	18	<i>TUBA1A</i>	Yokoi S. et al. Sci. Rep. 5, 15165 (2015).
小頭症 脳室周囲異所性灰白質 小眼球症	HP:0000252,小頭 HP:0007165,脳室周囲異所性灰白質 HP:0000568,小眼球	エクソーム解析 (Trio)	21	<i>PLK4</i>	Tsutsumi M. et al. Eur. J. Hum. Genet. 24, 1702-1706 (2016).
半球間裂嚢胞 脳梁低形成	HP:0002079,脳梁低形成	エクソーム解析 (Trio)	34	<i>NFIA</i>	Negishi Y. et al. Hum. Genome Var. 2, 15007 (2015).

PubCaseFinderの解析例

□ 症状を入力

- 発作 (HP:0001250)
- 視神経萎縮 (HP:0000648)

□ 候補遺伝子を入力

- 候補バリエーション: 26
 - 候補遺伝子: 15

```
ACADVL
AMER1
AMMECR1
APC2
CEP170B
FAM200B
LAMA5
NKX2-5
PKD1L3
RTN4IP1
TFF1
THOC2
TTN
ZNF501
ZNF502|
```

患者の 徴候・症状 を入力 + Upload File (HPO ID):

HP:0001250 発作 × HP:0000648 視神経萎縮 ×

疾患を絞り込む + Upload File (Gene Symbol or Entrez Gene ID or Ensembl Gene ID):

ENT:37 ACADVL (ACAD6 | LCACD | VLCAD) × ENT:139285 AMER1 (FLJ39827 | WTX) × ENT:9949 AMMECR1 ×
ENT:10297 APC2 (APCL) × ENT:1482 NKX2-5 (CSX1 | NKX2.5 | NKX4-1) × ENT:84816 RTN4IP1 (NIMP) ×
ENT:57187 THOC2 (THO2 | dJ506G2.1) × ENT:7273 TTN (CMH9 | CMPD4 | FLJ32040 | LGMD2J | MYLK5 | TMD) ×

共有・ダウンロード 症状入力支援 クリア 疾患を検索

遺伝性疾患 (OMIM) 18 件 希少疾患 (Orphanet) 14 件

合計: 18 件 1 2 » (表示件数)

順位 (類似度)	疾患名	対応する徴候・症状	疾患原因遺伝子	遺伝形式
1 (77.9%)	optic atrophy 10 with or without ataxia, mental retardation, and seizures ----- 視神経萎縮 1 +/- 精神遅滞およびけいれん	蓋明 視神経杯蒼白 RTN4IP1 常染色体劣性遺伝 OMIM:616732 画像検索 (Google)		
2 (68.5%)	osteopathia striata with cranial sclerosis ----- 線状骨症-頭蓋骨硬化症	両眼隔離 発作 AMER1 X連鎖優性遺伝 Osteopathia striata with cranial sclerosis (OS-CS) is a bone dysplasia characterized by longitudinal striations of the metaphyses of the long bones, sclerosis of the craniofacial bones, macrocephaly, cleft palate and hearing loss. >> 翻訳 (Google) OMIM:300373 画像検索 (Google) KEGG:H00444		
3 (48.8%)	early-onset myopathy with fatal cardiomyopathy ----- Salih ミオパシー	眼瞼下垂 運動発達遅滞 TTN 常染色体劣性遺伝 OMIM:611705 画像検索 (Google) Gene Reviews		

PhenIXの解析例

□ 症状を入力

- 発作 (HP:0001250)
- 視神経萎縮 (HP:0000648)

□ 候補遺伝子を入力

- 候補バリエーション: 26
 - vcfファイル



Top 11 ranked candidate genes

Variant	Variant analysis	Phenotypic analysis
1) AMER1 Gene relevance score: 1.000 variant score 1.000 total score: 1.000		
AMER1:uc004dvo.3:exon2:c.194G>T:p.G65V (chrX:g.63412973C>A)	Pathogenicity: MISSENSE Mutation Taster: 1.0 (P) Polyphen2: 0.0 (B) SIFT: 0.1 (T) Frequency No frequency data found Variant read depth: 1 Analyse variant with MutationTaster2	- Phenomizer: Semantic similarity score: 1.16 Osteopathia striata with cranial sclerosis [MIM:300373; gene: MIM:300647] X chromosomal
2) TTN Gene relevance score: 0.260 variant score 1.000 total score: 0.630		
TTN:uc002umx.1:exon13:c.659C>T:p.T220I (chr2:g.179505273G>A)	Pathogenicity: MISSENSE Polyphen2: 0.0 (B) SIFT: 0.0 (D) Frequency No frequency data found Variant read depth: 1	- Phenomizer: Semantic similarity score: 0.60 - Observed variants not compatible with mode of inheritance: Cardiomyopathy, dilated, 1G [MIM:604145; gene: MIM:188840], autosomal dominant Cardiomyopathy, familial hypertrophic, 9 [MIM:613765; gene: MIM:188840], autosomal dominant

精度比較

□ 結果

トップ1の正解率

PubCaseFinder: **94%** PhenIX: **35%**

患者				NGS解析				PubCaseFinder解析結果		PhenIX解析結果	
性別	年齢	臨床所見	臨床所見 (HPO)	解析方法	解析パイプライ後の候補バリエーション数	論文で報告した疾患原因遺伝子	報告論文	ランキング疾患数	論文報告遺伝子の順位	ランキング遺伝子数	論文報告遺伝子の順位
男	3歳	小脳低形成	HP:0001321,小脳低形成	エクソーム解析 (Trio)	34	<i>HNRNP2</i>	Bain JM. et al. Am. J. Hum. Genet. 99.3, 728-734 (2016).	24	1	17	10
男	13歳	てんかん 視神経萎縮	HP:0001250,発作 HP:0000648,視神経萎縮	エクソーム解析 (Trio)	26	<i>RTN4IP1</i>	Okamoto et al. J. Hum. Genet. 62, 927-929 (2017).	18	1	11	10
男	10歳	皮膚弛緩症	HP:0000973,弛緩性皮膚	エクソーム解析 (Trio)	27	<i>ALDH18A1</i>	Nozaki et al. Brain Dev. 38, 678-684 (2016).	23	1	8	1
男	5歳	白質異常 皮質盲	HP:0002500,大脳白質の異常 HP:0100704,Cortical visual impairment	エクソーム解析 (Trio)	36	<i>KIF1A</i>	Okamoto N. et al. J. Hum. Genet. 59, 639-641 (2014).	25	1	20	4
男	1歳	低身長 精神運動発達障害 骨異常	HP:0004322,低身長 HP:0002361,精神運動発達悪化 HP:0000924,骨格の異常	エクソーム解析 (Trio)	16	<i>FAM20C</i>	Tamai et al. Am. J. Med. Genet. A, 176, 682-686 (2018).	7	1	7	2
女	1歳11か月	macrocephaly cortical dysplasia	HP:0000256,Macrocephaly HP:0002539,Cortical dysplasia	エクソーム解析 (Trio)	36	<i>AKT3</i>	Negishi Y. et al. BMC Med. Genet. 18, 4 (2017).	33	1	20	2
女	2歳	水無脳症 小脳低形成 脳幹低形成	HP:0002324,水無脳症 HP:0001321,小脳低形成 HP:0002365,脳幹低形成	エクソーム解析 (Trio)	18	<i>TUBA1A</i>	Yokoi S. et al. Sci. Rep. 5, 15165 (2015).	15	1	8	1
女	2歳	小頭症 脳室周囲異所性灰白質 小眼球症	HP:0000252,小頭 HP:0007165,脳室周囲異所性灰白質 HP:0000568,小眼球	エクソーム解析 (Trio)	21	<i>PLK4</i>	Tsutsumi M. et al. Eur. J. Hum. Genet. 24, 1702-1706 (2016).	7	1	9	3
男	4歳	半球間裂嚢胞 脳梁低形成	HP:0002079,脳梁低形成	エクソーム解析 (Trio)	34	<i>NFIA</i>	Negishi Y. et al. Hum. Genome Var. 2, 15007 (2015).	13	1	20	18
女	NA	脳梁欠損 てんかん 知的障害	HP:0001274,脳梁無発生 HP:0001250,発作 HP:0001249,知的障害	エクソーム解析 (Trio)	3	<i>EPG5</i>	Hori et al. Sci. Rep. 14, 3552 (2017).	1	1	1	1
女	NA	多小脳回 異所性灰白質	HP:0002126,多小脳回 HP:0002281,異所性灰白質	エクソーム解析 (Trio)	23	<i>NEDD4L</i>	Kato et al. J. Hum. Genet. 62, 861-863 (2017).	12	1	13	8
女	NA	小頭症 低身長 自閉症	HP:0000252,小頭 HP:0004322,低身長 HP:0000717,自閉症	エクソーム解析 (Trio)	16	<i>ASXL3</i>	Hori I. et al. Am. J. Med. Genet. A, 170, 1863-1867 (2016).	6	2	10	8
女	NA	小頭症 てんかん性脳症	HP:0000252,小頭 HP:0200134,てんかん性脳症	エクソーム解析 (Trio)	46	<i>STAMBP</i>	Hori et al. J. Hum. Genet. 63, 957-963 (2018).	15	1	20	2
男	NA	巨脳症	HP:0001355,巨大脳症	エクソーム解析 (Trio)	34	<i>MYCN</i>	Kato et al. J. Med. Genet. 56, 388-395 (2019).	40	1	15	3
女	NA	小頭症	HP:0000252,小頭	エクソーム解析 (Trio)	5	<i>ASPM</i>	F. Miya et al. Sci. Rep. 5, 9331 (2015).	3	1	2	1
女	NA	脳梁欠損	HP:0001274,脳梁無発生	エクソーム解析 (Trio)	38	<i>EPG5</i>	Hori et al. Sci. Rep. 14, 3552 (2017).	15	1	15	1
男	NA	水頭症	HP:0000238,水頭症	エクソーム解析 (Trio)	25	<i>TUBA1A</i>	Yokoi S. et al. Sci. Rep. 5, 15165 (2015).	8	1	12	1

未知の疾患バリエントを探索

- インターネット上で公開されている未診断症例約25,000件を収集
 - PubCaseFinderからこれら症例を検索可能 (coming soon)

※ 2019/3/22時点

データベース名	症例数
DECIPHER	22,912
MyGene2	2,081
UDN	114

サンプル (UDN)

The screenshot displays the UDN website interface. At the top, the logo 'UDN Undiagnosed Diseases Network' is visible alongside social media links for Facebook and Twitter. A navigation bar includes links for 'About Us', 'How to Apply', 'Sites', 'Participants', 'Research', 'News', 'Contact Us', and 'En Español'. A search bar is located on the right. The main content area is divided into two sections: 'Symptoms / Signs' and 'Genetic Variants of Interest'. The 'Symptoms / Signs' section lists several clinical features: Cleft lip and palate (unilateral), Heart defect (atrial septal defect), Absent thymus, No functional T-cells, and Short neck (Klippel Feil anomaly; specific skeletal anomalies include fusion of the posterior elements of C2-C4 and fusion). The 'Genetic Variants of Interest' section states that clinicians and researchers have identified a genetic change causing the participant's symptoms, citing Liu et al. (2018). Below this, a table provides details about the genetic variant.

Gene	Inheritance Pattern	Position	Transcript	DNA Change	Protein Change
TBX2	autosomal dominant	chr17:59477596	NM_005994.3	c.59G>A	p.R20Q

まとめ



鑑別診断に活用



N of 1 Problem
coming soon



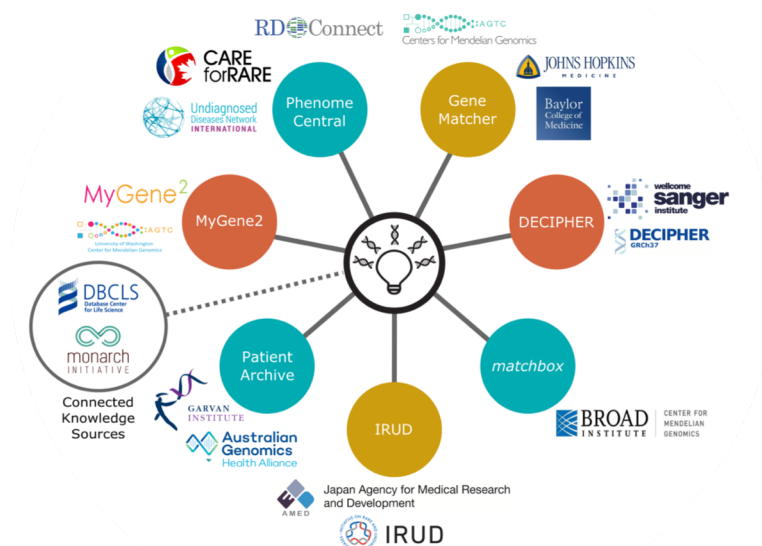
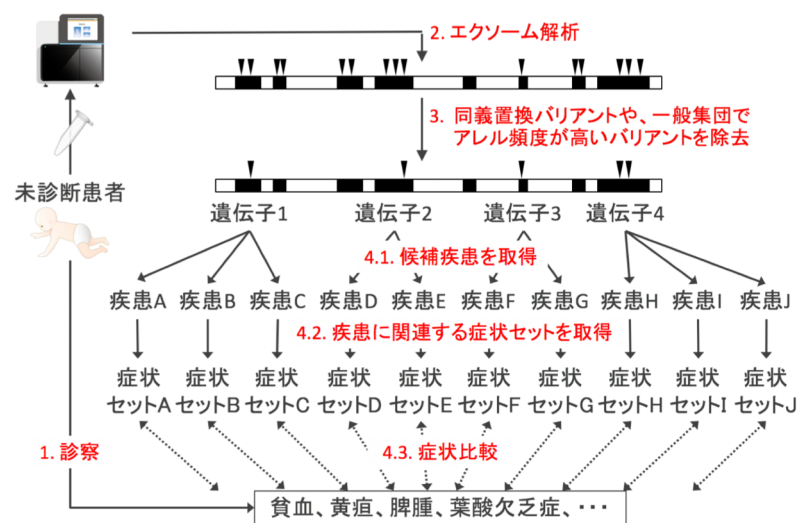
症例報告を探す



簡単な症状入力

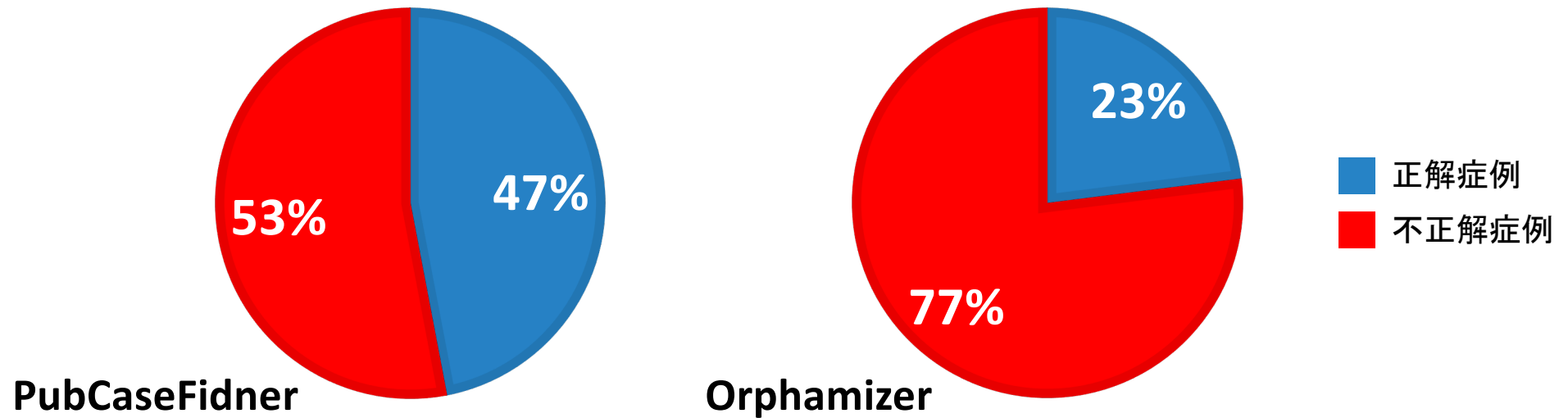


多言語対応

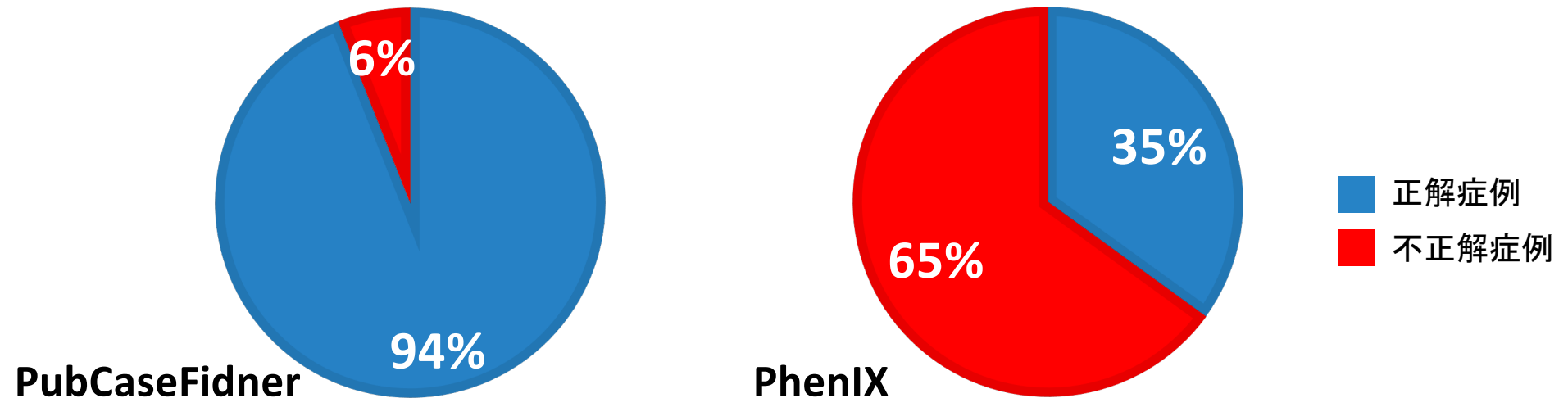


まとめ

■ 症状のみを用いたTop5の正解率比較



■ 症状と遺伝情報を用いたTop1の正解率比較



今後

- 約25,000件の未診断症例を検索対象に追加(9月末リリース予定)
- 日本語症例報告の追加(本年度リリース予定)
 - J-Stageに収納されている約15万件の症例報告
- 疾患ではなく疾患原因遺伝子をランキング表示する機能を追加(来年度リリース予定)
- 希少疾患インフォマティクスハッカソン(2020年2月開催予定)
- 国際版希少疾患インフォマティクスハッカソン(2020年3月開催予定)
 - 予算を確保できれば