

前向きコホート、医療情報、個人の疾患感受性多型情報を統合するオントロジーの開発

An Ontology for Prospective Cohort, Health Information, and Personal Susceptibility to Diseases

高井 貴子^{*1}

Takako Takai-Igarashi

^{*1} 東北大学東北メディカル・メガバンク機構

Tohoku Medical Megabank Organization, Tohoku University

IBO (Integrated Biobank Ontology), Tohoku Medical Megabank Project, ontology, genome cohort, biobank.

1. 背景

東北メディカル・メガバンク事業は、15万人の健康人を対象とした大規模ゲノムコホート研究を行い、提供された試料と情報を管理するバイオバンクを構築することにより、創薬研究や個別化医療・個別化予防等の次世代医療体制の構築を目指すことを主たる目的としている。コホート研究は、健康状態や環境要因およびゲノム情報などを、病気の発症前に収集し、その人たちがその後どのような病気を発症したかを時間的にすべて前向きに検討するもので、病気の早期発見と早期治療を実現することができる。

一般にバイオバンクは複数の学問領域に跨る情報を格納するので、その検索においてはオントロジーに基づく機械支援が必要となるが、その研究は端緒にすぎたばかりである。欧州ではバイオバンク連携における情報共有を目的として MIABIS オントロジー (OMIABIS) が開発されたが、その対象は生体試料の収集・保管・分配に限定されている。本研究では MIABIS を参照しつつ、健康情報、疾患情報、ゲノム情報を更に統合したオントロジー IBO (Integrated Biobank Ontology) の開発を目的としている。

生命医学においては、オントロジーの直交性、交換性の標準が BFO (Basic Formal Ontology)、IAO (Information Artifact Ontology) および PATO (Phenotype And Trait Ontology) により定義されている。BFO は、生命医学領域のトップオントロジーである。IAO は、生命医学の研究成果の中核を成す表現物 (記述や解釈) を、実態 (もの) と峻別するオントロジーである。PATO は表現型 (質) の形式化の標準を与えている。

IBO はこれら先行研究の成果を踏まえて開発するが、更に以下 A~D の問題解決が必要と考えている。

- コホート研究で収集する調査情報は「表現」であるのか。
- 対象者の表現型 (質) と遺伝子 (もの) は、どう関係するのか。
- 未病から発症に至るプロセスは、Smith および Mizoguchi の疾患定義で表現できるか。
- 疾患リスクは「性質」であるのか。

本報告では IBO の開発と、課題 C の解析結果について報告する。対象疾患は、被災が発症原因となる PTSD (Post Traumatic Stress Disorder)、および (原発開放隅角) 緑内障、歯周炎 (骨吸収を中心とする病態) とした。

2. 方法

2.1 参照したオントロジー

IBO は、既存のオントロジーで定義済みの概念については、その再利用を基本とする。上位オントロジーは BFO とし、BFO 準拠の以下のオントロジーより概念を取得し再構成した。

オントロジー名	略称	ドメイン
Basic Formal Ontology	BFO	生命医学の上位概念
Relation Ontology	RO	生命医学の関係概念
Information Artifact Ontology	IAO	生命医学の表現概念
Phenotype And Trait Ontology	PATO	生命医学の表現型
Ontology General Medical Science	OGMS	疾患・医療関連概念
Ontology for Genetic Susceptibility Factor	OGSF	疾患感受性遺伝因子
Experimental Factor Ontology	EFO	生命医学の実験結果
Mammalian Phenotype Ontology	MP	ほ乳類の (疾患関連) 表現型
Human Phenotype Ontology	HP	ヒトの (疾患関連) 表現型
Gene Ontology, Biological Process	GO-BP	生体機能の分子プロセス
Uber Anatomy Ontology	UBERON	動物の解剖学的構造
Ontologized MIABIS	OMIABIS	バイオバンク

2.2 オントロジー構築ツール

IBO の構築には Protégé 5.0 を用いた。既存オントロジーの取り込みには MIREOT 1.1.0 プラグインを用いた。

2.3 ドメイン知識の獲得

PTSD、緑内障、歯周炎のドメイン知識は、総説、教科書および専門家へのインタビューより獲得した。以下に参照した総説のうち主要なものを示す。

疾患	参照した主な総説 [第一著者姓 発行年 雑誌名]
PTSD	Parsons 2013 Nat Neurosci, Almlil 2014 Int J Neuropsychopharmacol, Malan-Muller 2014 Genes

連絡先: 高井貴子, 東北大学, 宮城県仙台市青葉区星陵町 2-1, 022-273-6281, takai@megabank.tohoku.ac.jp

	Brain Behav, Logue 2015 Neuropsychopharmacology, Reznikov 2016 Neuropsychopharmacology.
緑内障	Abu-Amero 2015 Int J Mol Sci, Choi 2015 Biomed Res Int, Chung 2015 Biomed Res Int, Iglesias 2015 Eye (Lond), Kumar 2016 Gene.
歯周炎	Kajiya 2010 J Oral Microbiol, Graves 2011 Journal of Dental Research, Kayal 2013 Biomed Research International, Barbato 2015 Clin Cases Miner Bone Metab.

3. 結果と考察

3.1 オントロジーの構造

IBO は OWL で記述し、各概念クラスの属性は owl:Restriction を使って記述した。ゲノムコホート、バイオバンク、疾患、に関する概念クラスは、以下の階層に定義されている。

Continuant の下位概念	
Generically Dependent Continuant の下位概念	
Measurement	: コホート研究で収集する質問票の回答、検体検査の結果、生理学検査の結果
Specifically Dependent Continuant の下位概念	
Abnormal State	: 疾患の因果関係を構成する病態
Disease	: 疾患
Predisposition to Disease	: 疾患になりやすい体質
Risk Factor	: 疾患リスク因子
Symptom	: 症状
Habit	: 生活習慣
Drug Role	: 医薬品ロール
Independent Continuant の下位概念	
Gene	: 遺伝子
Molecular Entity	: 分子
Cell	: 細胞
Organism Part	: 解剖学的構造を持つ生体の部分
Organism	: 生物体
Population	: ヒトの集団
Specimen	: 生体試料
Occurrent の下位概念	
Biological Process	: 分子・細胞レベルのプロセス
Bodily Process	: ヒトの生体レベルのプロセス
Disease Course	: 疾患の内部プロセス
Behavior	: ヒトの習性

owl:Restriction で用いる属性の種類は以下である。

ObjectProperty の下位概念	
has part	: 全体の部分である
participate in	: プロセスに参加するものである
realized by	: 具現されたものである
measurement of	: 測定/観測された結果である
preceded by	: 因果関係において先行する

3.2 疾患の定義

前向きコホートにおいては、未病から疾患に至るプロセスを制御する原因分子やマーカー分子の発見が研究目標となる。IBO における疾患の表現モデルは、Smith [Ceusters 2010] および Mizoguchi [Mizoguchi 2011] の定義 (疾患はプロセスを enact する) を基本とし、コホート研究で収集する質問票や検査データ、

および生体分子の機能との関係を追加した。図 1 に示すように疾患は Continuant 概念と Occurrent 概念の組み合わせで表現される。Disease は Disease Course により具現される。Disease は複数の Abnormal State から構成され、Disease Course は複数の Biological Process から構成される。Abnormal State である概念クラスの間、および Biological Process である概念クラスの間には因果関係が存在する。Abnormal State は Measurement の下位概念により観測される。Biological Process は Molecular Entity の下位概念により特徴付けられる。

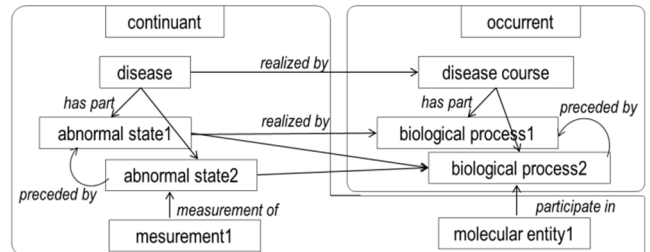


図 1. IBO における疾患表現モデル。

続いて、PTSD、緑内障、歯周炎を対象として、Abnormal State である概念クラス、および Biological Process である概念クラスについて、既存のオントロジーである MP および GO-BP に対してマッピングを行い、ほぼ全ての概念クラスがマップされることが分かった(図 2)。MP および GO-BP は遺伝子のアノテーションが充実しており、遺伝子の実験データ解析に頻用される。本結果は、遺伝子の解析データから MP および BO-BP の推定を介して、疾患の推定が可能となることを示す。

abnormal state (category)	abnormal state (MP class)	MP ID	biological process (GO-BP class)	GO ID
fibrotic phenotype observed at the optic nerve head	abnormal optic nerve morphology	MP:0001330	transforming growth factor-beta secretion	GO:0038044
	retinal fibrosis	MP:0006186	fibroblast proliferation	GO:0048144
	abnormal extracellular matrix morphology	MP:0013258	gene expression involved in extracellular matrix organization	GO:1901148
			regulation of extracellular matrix organization	GO:1903053
	abnormal trabecular meshwork morphology	MP:0005203	transforming growth factor beta receptor signaling pathway	GO:0007179
		regulation of extracellular matrix disassembly	GO:0010715	

図 2. 緑内障の Abnormal State のひとつである「視神経乳頭の線維化」が MP および GO-BP の概念クラスの再構成で表現できることを示す。

一方、MP あるいは GO-BP にマップされない概念クラスを以下に示す。ヒトの病態に特有の概念であると考察される。

PTSD	hyper-vigilance, heightened arousal, trauma exposure, memory consolidation
緑内障	loss of vision
歯周炎	invasion of bacteria, pocket formation, loss of soft tissue, loss of tooth

本報告では「C. 未病から発症に至るプロセスは、Smith および Mizoguchi の疾患定義で表現できるか」について、ひとつの表現した結果を示すことができたが、更に解析を深める必要がある。

参考文献

[Ceusters 2010] Ceusters et al: Foundations for a realist ontology of mental disease, J Biomed Semantics, 1:10, 2010.
[Mizoguchi 2011] Mizoguchi R, et al: River Flow Model of Diseases, Proc. of 2nd International Conference on Biomedical Ontology (ICBO2011), Buffalo, NY, USA, 2011.