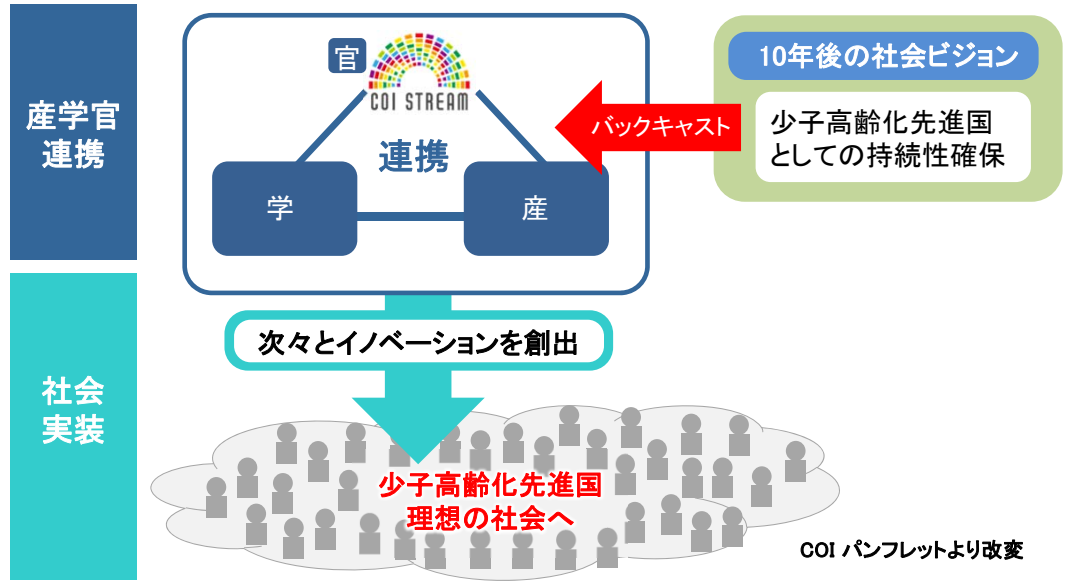
COI STREAM

ヘルスビッグデータを用いた健康長寿イノベーション

東京大学医科学研究所ヒトゲノム解析センター
DNA情報解析分野 教授
宮野　悟

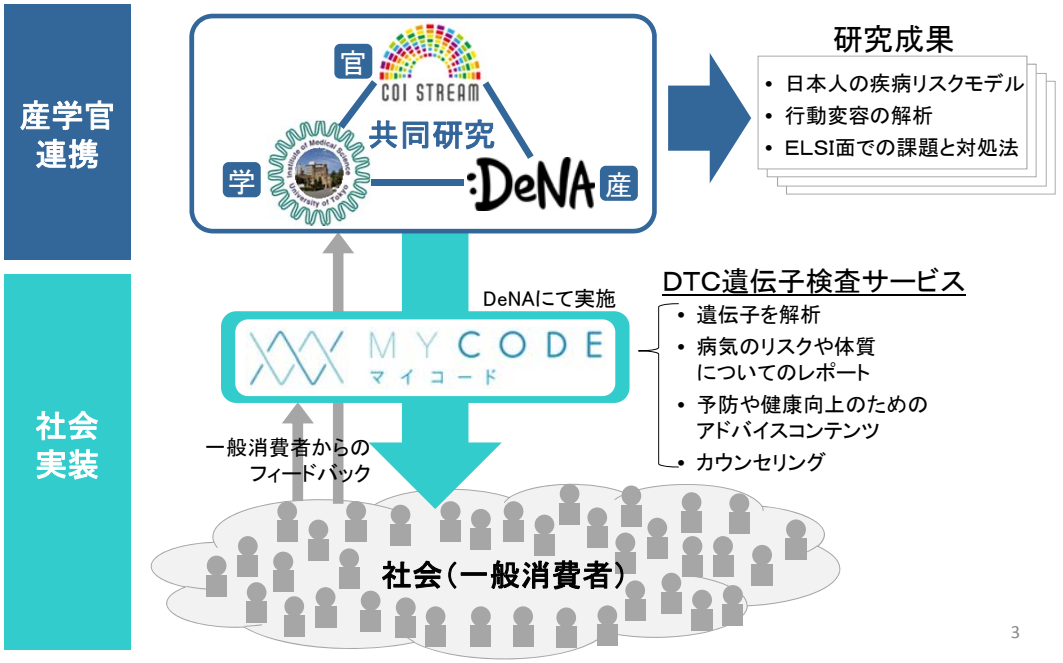
COI STREAM
革新的イノベーション創出プログラム



10年後の日本が目指すべき社会の姿を見据え、企業だけでは実現できない革新的なイノベーションを産学連携で実現していくこと

2

COI-Tにおける共同研究と研究成果の社会実装



3

共同研究の内容

1. 日本人DNAデータを用いた日本人のための疾病リスク予測モデルの研究
2. DNAの一部を調べる消費者直販型(DTC)のサービスによる消費者の健康に関する行動変容の解析
3. ヘルスビッグデータ社会実装にあたっての ELSI 面での課題抽出と対処法の研究

4

1. 日本人DNAデータを用いた日本人のための疾病リスク予測モデルの研究

PubMed*に登録されており、ゲノムワイド関連解析 (GWAS) を用いた論文のある

- 282種の疾患
- 437種の体質

これらの疾患・体質の研究論文を調査

・単一遺伝子疾患
・家族性腫瘍
・小児疾患
・先天性疾患　　は除外

*PubMed: 米国国立医学図書館（National Library of Medicine）が作成している医学生物学の文献データベース。

5

エビデンスのレベル分類

I	システマティック・レビュー／RCT のメタアナリシス
II	1つ以上のランダム化比較試験
III	非ランダム化比較試験による分析疫学的研究（コホート研究）
IVa	分析疫学的研究（症例対照研究、横断研究）
IVb	記述研究（症例報告やケース・シリーズ）
V	患者データに基づかない、専門委員会や専門家個人の意見
VI	

介入研究

観察研究

GWASによるゲノムと疾患・体質の関連解析

Minds 診療ガイドライン作成の手引き 2007 より
http://minds4.jcqhc.or.jp/minds/glgl/glgl.pdf

6

米国 NIH GWASカタログ

genome.gov
National Human Genome Research Institute
National Institutes of Health

http://www.genome.gov/gwastudies/

約1,900のゲノムと疾患、ゲノムと体質をGWASにより解析した論文に基づくSNP のデータベース

前立腺がんの例

Date Added to Catalog (since 11/25/08)	First Author/Date/ Journal/Study	Disease/Trait	Initial Sample Size	Replication Sample Size	Region	Reported Gene(s)	Mapped Gene(s)	Strongest SNP-Risk Allele	Context	SNP type	Risk Allele Frequency in Controls	P-value	OR Type	OR or beta-coefficient and [95% CI]	Platform [SNPs passing QC]	CNV
11/25/08	Eeles RA February 10, 2008 *Nat Genet* Multiple newly identified loci associated with prostate cancer susceptibility.	Prostate cancer	1,854 cases, 1,894 controls	3,268 cases, 3,366 controls	10q11.23	*MSMB*	-	rs10993994-T	nearGene-5	Novel	0.40	9×10^{-29}	Y	1.25 [1.17-1.34]	Illumina [541,129]	N
					19q13.33	*KLK3*		rs2735839-G		Novel	0.85	2×10^{-18}	Y	1.2 [1.10-1.33]		
					8q24.21	*Intergenic*		rs4242384-C		Known	0.09	3×10^{-16}	Y	1.88 [NR]		
					8q24.21	*Intergenic*	-	rs6983267-G	ncRNA;intron	Known	0.49	9×10^{-13}	Y	1.42 [NR]		
					11q13.3	*Intergenic*		rs7931342-G		Novel	0.51	2×10^{-12}	Y	1.19 [1.11-1.27]		
					17q12	*HNF1B*	-	rs7501939-C	intron	Known	0.57	9×10^{-12}	Y	1.41 [NR]		
					6q25.3	*SLC22A3*	-	rs9364554-T	intron	Novel	0.29	6×10^{-10}	Y	1.17 [1.08-1.26]		
					7q21.3	*LMTK2*	-	rs6465657-C	intron	Novel	0.46	1×10^{-9}	Y	1.12 [1.05-1.20]		
					Xp11.22	*NUDT10, NUDT11*		rs5945619-C		Novel	0.36	2×10^{-9}	Y	1.19 [1.07-1.31]		
					3p12.1	*Intergenic*		rs2660753-T		Novel	0.11	3×10^{-8}	Y	1.18 [1.06-1.31]		
					8q24.21	*Intergenic*	-	rs1016343-T	ncRNA	Known	0.18	1×10^{-7}	Y	1.37 [NR]		
					17q24.3	*Intergenic*	-	rs1859962-G	intron	Known	0.46	1×10^{-6}	Y	1.26 [NR]		

7

論文調査体制

共同研究

科学的根拠調査の専任チーム

東大医科研

東京大学 医科学研究所
THE INSTITUTE OF MEDICAL SCIENCE THE UNIVERSITY OF TOKYO

東大医科研教員

face to face MTG
電話、Mail等

DeNA

DeNA生命系博士号保有者5名

東京大学 医科学研究所

東大医科研教員
東大医科研病院臨床遺伝指導医

（専任チーム外のサポート）

- **全体の face to face MTG 15回以上**
- **専任チームの face to face MTG 50回以上**
- **電子mail 数千通以上**

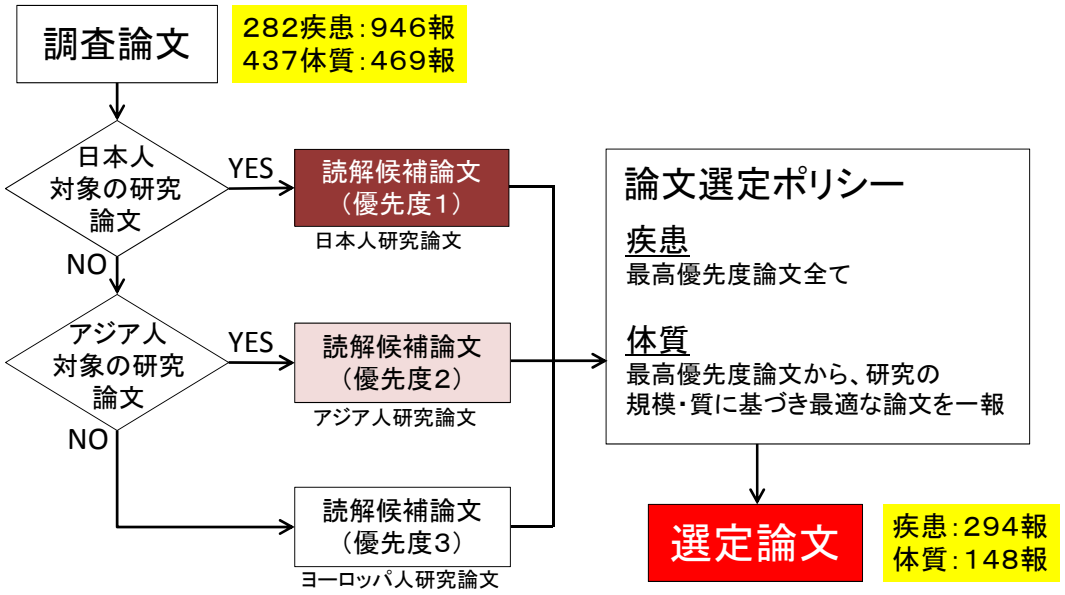
データ整理など、一部業務委託

外部委託先

生物・生命系博士号保有者3名

8

論文選定フロー



9

論文内容の分析

- 各論文から以下の情報を抽出することで GWAS カタログにはない情報を得る
  - 対象疾患（または、体質）と統計学的に有意に関連する SNP の情報

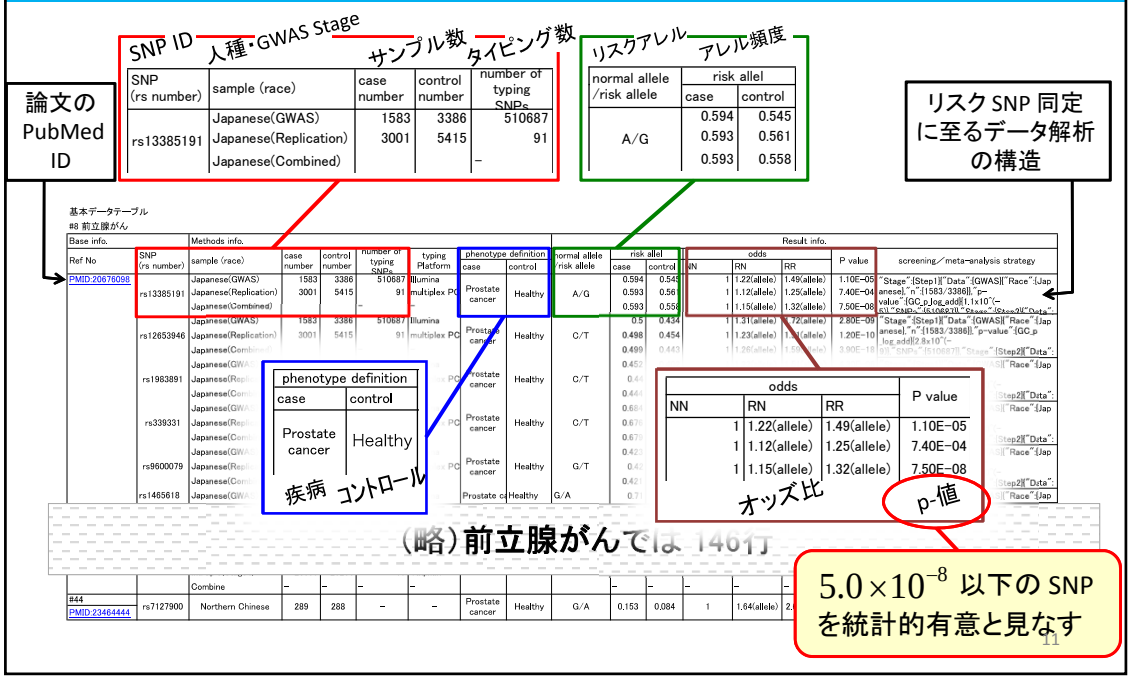
SNP ID　リスクアレル　アレル頻度　オッズ比

  - 関連 SNP 同定に至ったデータ解析のプロセス

サンプルサイズ　人種　解析 SNP 数
SNPアレイのプラットフォーム
論文の研究デザイン

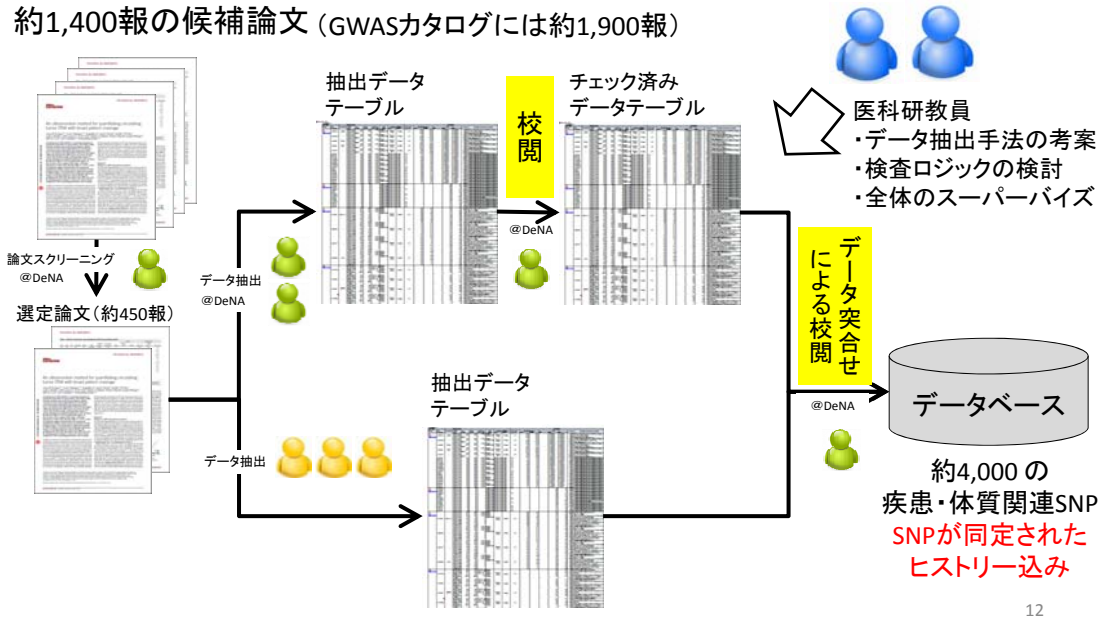
10

抽出データテーブル



11

品質管理手順と情報の更新



12

7割以上の論文は4回分析
8割以上の論文は2回以上分析

	疾患	体質	計
候補論文数	946	469	1415
選定論文数	294	148	442
GWAS カタログに登録ない論文数	6	4	8
1回精読した論文数	53	17	70
2回精読した論文数	57	-	57
4回精読した論文数	184	131	315
著者に確認した件数	23 件	5 件	28 件
総抽出SNP数	3,141	555	3,696

13

米国 NHI GWAS カタログ

卵巣がん（漿液性浸潤性上

rs757210-G

G がリスクアレルと書かれているが・・・

Date Added to Catalog (since 11/25/08)	First Author/Date/ Journal/Study	Disease/Trait	Initial Sample Size								in Controls			Platform [SNPs passing QC]	CNV
07/12/13	Pharoah PD April 01, 2013 *Nat Genet* GWAS meta-	Ovarian cancer	3,769 European ancestry cases,	211 East Asian cases, 972 East Asian	17q12	*HNF1B, ACACA, C17orf78, TADA2A,*	HNF1B	rs757210-G	intron	0.37	8 x 10-10 (Serious invasive)	1.12 [1.08-1.17]	Illumina [2,508,744] (Imputed)	N

【論文のデータ】

Table 2 ORs and tests of association by histological subtype and population for the most strongly associated SNPs at 8q21, 10p12 and 17q12

SNP	Locus	Ref/alt allele	RAF	Subtype	Ancestry	OR	95% CI	P
rs757210	17q12	G/A	0.37	All invasive	European	1.05	1.02–1.09	0.00090
			0.37	Serous	European	1.12	1.08–1.17	8.1×10^{-10}
			0.37	Endometrioid	European	0.98	0.91–1.04	0.47
			0.37	Clear cell	European	0.80	0.72–0.88	3.9×10^{-6}
			0.37	Mucinous	European	0.89	0.81–0.99	0.027
			0.29	Serous	Asian	1.10	0.87–1.37	0.43
			0.53	Serous	African	1.08	0.75–1.54	0.69
			0.37	Serous	Mixed	0.86	0.73–1.02	0.093

reference and alternative allele.

Ref/alt allele
G/A

このSNPはマイナーアレル(A)がリスクと判断されるが GWAS カタログはそうはなっていない。
最終的に著者に問合わせ、マイナーアレル(A)がリスクであることを確認

【dbSNPの情報】

	Sample Ascertainment				Genotype Detail				Alleles	
ss#	Population	Individual Group	Chrom. Sample Cnt	Source	A/A	A/G	G/G	HWP	A	G
ss14298599	HapMap-CEU	European	120	IG	0.150	0.567	0.283	0.2810	0.433	0.567

論文で解析された CEU サンプルで A がマイナーアレルであると確認

14

論文の分析により GWAS カタログにはない情報を抽出

【ベーチェット病　Behçet's disease】

pubmed_id	snp	race	stage	case_num	ctrl_num	phenotype_case	phenotype_ctrl	Normal_allele	Risk_allele	Odds_nn	Odds_rn	Odds_rr	p_value
20622879	rs1495965	Japanese	GWAS	612	740	Behcets's disease	Healthy controls	A	G	1	1.56	2.43	1.20E-08
		Turkish	Replication	1215	1279	Behcets's disease	Healthy controls	A	G	1	1.25	1.56	8.40E-05
		Korean	Replication	119	140	Behcets's disease	Healthy controls	A	G	1	1.35	1.82	9.40E-02
		Japanese, Turkish, Korean	Combined	1946	2159	Behcets's disease	Healthy controls	A	G	1	1.35	1.82	1.90E-11
	rs1800871	Japanese	GWAS	612	740	Behcets's disease	Healthy controls	C	T	1	1.64	2.69	9.50E-09
		Turkish	Replication	1215	1279	Behcets's disease	Healthy controls	C	T	1	1.36	1.85	3.10E-07
		Korean	Replication	119	140	Behcets's disease	Healthy controls	C	T	1	1.51	2.28	4.40E-02
		Japanese, Turkish, Korean	Combined	1946	2159	Behcets's disease	Healthy controls	C	T	1	1.45	2.10	1.00E-14
	rs4959053	Japanese	GWAS	612	740	Behcets's disease	Healthy controls	G	A	1	3.18	10.11	1.80E-26

Step 1: GWAS ステージ　日本人データ ←

Step 2: Replication ステージ　トルコ人データ

Replication ステージ　韓国人データ

Step 3: Combined ステージ　日本人＋トルコ人＋韓国人データ

日本人のみのデータからの結果を抽出可能

【ブルガダ症候群　Brugada syndrome】

pubmed_id	snp	race	stage	Case_num	Ctrl_num	phenotype_case	Phenotype_ctrl	Normal_allele	Risk_allele	Odds_nn	Odds_rn	Odds_rr	p_value
23872634	rs10428132	European	Discovery GWAS	312	1115	Brugada syndrome	Control	G	T	1	3.00	9.00	6.79E-26
		European	Replication 1	594	806	Brugada syndrome	Control	G	T	1	2.35	5.52	1.66E-30
		Japanese	Replication 2	208	1016	Brugada syndrome	Control	G	T	1	2.56	6.55	1.56E-16
		European,Japanese	Combined meta-analysis	416	2937	Brugada syndrome with symptoms	Control	G	T	1	2.84	8.07	1.15E-39

Step 1: GWAS ステージ　ヨーロッパ人データ

Step 2: Replication ステージ　ヨーロッパ人データ

Replication ステージ　日本人データ ←

Step 3: Combined ステージ　ヨーロッパ人＋日本人データ

15

構築した疾病リスク予測モデル 体質解析モデル 289

158疾患

日本人データに基づくモデル	43
東アジア人データに基づくモデル	30
ヨーロッパ他データに基づくモデル	85

131体質

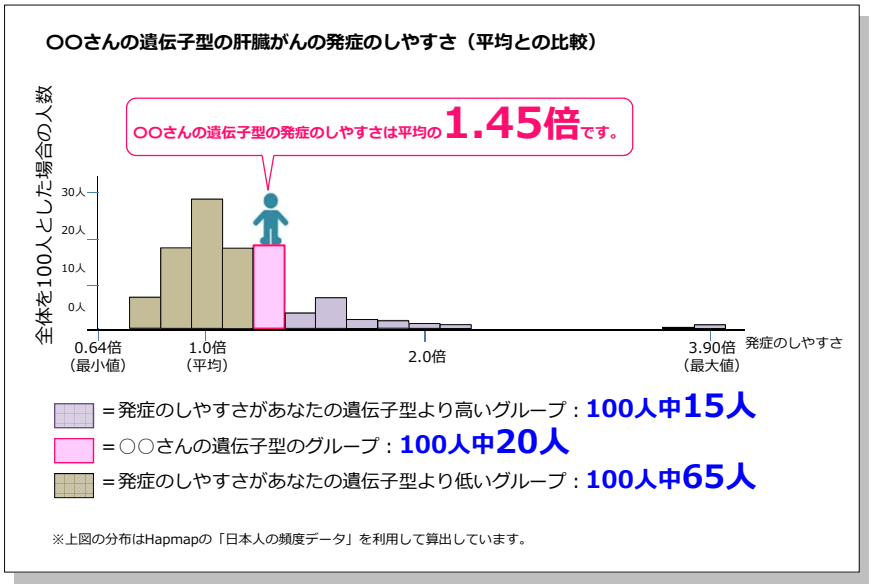
日本人データに基づくモデル	27
東アジア人データに基づくモデル	16
ヨーロッパ他データに基づくモデル	88

- 1/4が日本人データに基づくモデル
- 東アジア人対応116モデル
- モデルに用いられている SNP の約20%は、論文の分析によって GWAS カタログにはない情報を抽出したもの
- 9件の GWAS カタログの誤記を修正したものを含む

16

予測モデルからの結果の表示
～言い過ぎず、今分かることを～

伝え方の例：日本人の中では発症しやすい方か？



17

3. ヘルスビッグデータ社会実装にあたっての ELSI 面での課題抽出と対処法の研究

1. 現行法・指針等と調和したDTC検査起業化にとって必要な要件に関する助言。
2. 諸外国におけるDTC検査の規制状況に関する調査研究
3. 診療、研究での対応と調和した、DTC検査における偶発的・二次的所見への対応方策の検討

18

3. ヘルスビッグデータ社会実装にあたっての ELSI 面での課題抽出と対処法の研究

1. 現行法・指針等と調和したDTC検査起業化にとって必要な要件に関する助言
2. 諸外国におけるDTC検査の規制状況に関する調査研究
3. 診療、研究での対応と調和した、DTC検査における偶発的・二次的所見への対応方策の検討

19

諸外国でのDTC検査規制状況

米国（2007年6月の調査）	
DTC検査への規制なし	25州とコロンビア特別区
条件／制限付きで容認 （検査内容、検査を発注できる人、目的、CLIA認証など）	13州
DTC検査を禁止 （医業該当、消費者は検査を発注できないなど）	12州

イギリス	保健省	DTC企業を交えた協議を実施し、2010年に旧・人類遺伝学協議会が「DTC検査サービスに関する原則の共通枠組み」を公表。
フランス	生命倫理関連法(2004改正)	DTC検査を禁止（医業該当）
ドイツ	遺伝子診断法(2009)	DTC検査を禁止（医業該当）
韓国	保健福祉部 遺伝子検査評価院(2005-)	生命倫理法に基づき、実施機関の登録と遺伝学的検査の質評価を実施。DTC販売されていた14種類の検査に関して実施を禁止。

米国については、Genetics and Public Policy Center による2007年調査を参照。イギリス、フランス、ドイツ、韓国については、Borry et al, 20:715–721, 2012；内閣官房・旧医療イノベーション室勉強会用資料（2012, 武藤香織作成）

20

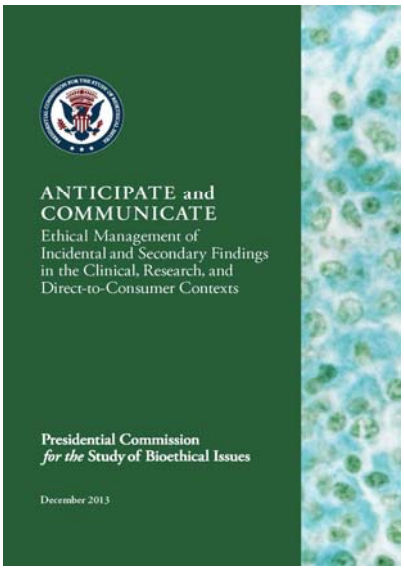
諸外国でのDTC検査規制状況

- 米国
  - 2013年11月22日　FDAは、23andMe に対してサービスの中止を求めた。その理由は、
    - 検査デバイスが医療機器に該当するため要承認
    - BRCAの遺伝リスクや薬剤応答の結果の誤りが重大な過誤をもたらすおそれ
    - FDAが求めた追加情報の提供に未対応
- 中国
  - 2014年2月9日　CFDAは、人のゲノム解析について、シークエンサー、遺伝子検査デバイス等の医療機器登録と、承認を受けるまで中止するよう求める通知を出した。同時に、製品試用の場として臨床応用特区を公募予定であると表明した。
  - 7月2日、BGI社のシークエンサーが医療機器として承認を受けた。
  - 複数の企業が臨床応用特区に申請中。

21

「診療、研究、DTC検査における偶発的・二次的所見のマネジメント」

米国「生命倫理問題の研究に関する大統領諮問委員会（PCSBI）」（2013.12）



- DTC検査における偶発的・二次的所見を論じた随一の報告書。この報告書では、「所見」を5種類に分類し、対応方策を検討。
- 本事業に関連して重要なのは、以下の2つ。
- 「予期されぬ偶発的所見」
  - 試料採取時には知られていなかった、新しく発見された遺伝的な健康リスクが同定された場合など、現時点では予想できない局面。
- 「予期される二次的所見」
  - 網羅的なゲノム解析を実施して、既知の遺伝子変異が見つかる場合、親子・きょうだいで申込み、結果を照合してみると血縁関係がないことが発覚する場合など、事前にある程度予想できる局面。

22

DTC検査における偶発的所見・二次的所見に関する勧告

米国「生命倫理問題の研究に関する大統領諮問委員会（PCSBI）」（2013.12）

Direct-to-Consumer Recommendations

Consent in the Direct-to-Consumer Context

Government Regulation in the Direct-to-Consumer Context

Industry-Wide Best Practices in the Direct-to-Consumer Context

- 事業者は、開示しない予定の所見の範囲について、顧客に前もって伝達するべきである（勧告15）
- 連邦政府は、DTC検査サービスの安全性を監視する機構を、州政府は消費者保護の体制を確立すべきである（勧告16）
- 事業者は、偶発的・二次的所見のマネジメントに関連する業界のベストプラクティスづくりに取り組むべきである。これには、具体的な手順、診療サービスへの紹介基準を含むべきである。（勧告17）

これらの勧告も参考にしつつ、診療や研究での取り組みと調和しながら、DTC検査事業における「所見」への対応を検討する予定。

23

COI STREAM

- 利益相反などについて
- 23andMe との相違

24

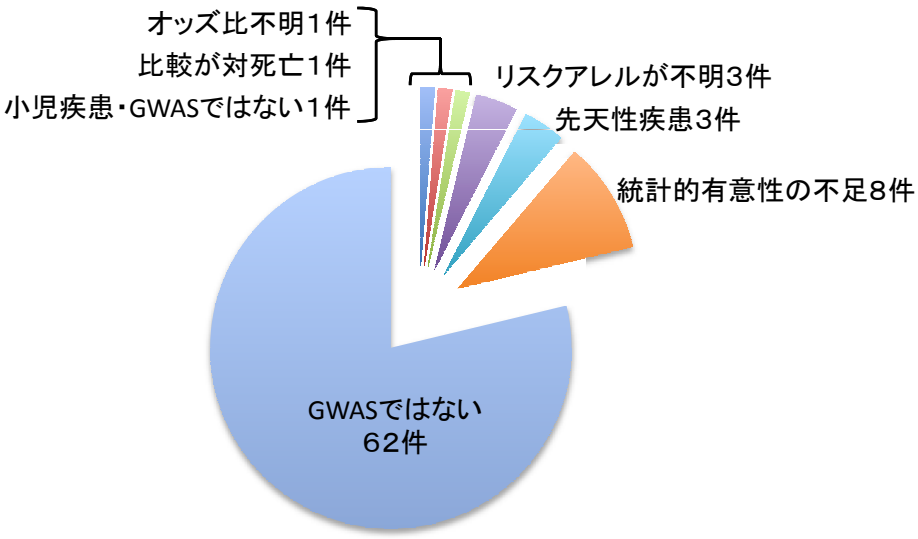
利益相反などについて

- 東大医科研と DeNA は、共同研究契約を結んでいます。
課題名「予知と予防による健康長寿社会実現」
- 東大医科研 COI STREAM 研究者は、DeNA、DeNA LS の株式、ストックオプションなどを保有せず、また、コンサルテーションなどによる金銭等を受領する行為はしないというポリシーを取っています。
- 東大医科研 COI STREAM の委託費は、DeNA、DeNA LS の行う事業に使われていません。

DeNA：株式会社ディー・エヌ・エー
DeNA LS：DeNA ライフサイエンス

25

23andMe では提供しているが開発した疾病リスク予測モデルには含まれない項目



26

23andMe では提供しているが開発したモデルには含まれない疾患・体質

疾患

Attention-Deficit Hyperactivity Disorder	オッズ比不明
Sudden Cardiac Arrest	比較が対死亡
Sjögren-Larsson Syndrome	小児疾患・GWASではない
Gallstones	
Progressive Supranuclear Palsy	リスクアレルが不明
Psoriasis	
Cleft Lip and Cleft Palate	
Hypospadias	先天性疾患
Narcolepsy	
Gaucher Disease	
Hypertrophic Cardiomyopathy (MYBPC3 25bp-deletion)	
Male Infertility	
Obsessive-Compulsive Disorder	
Oral and Throat Cancer	
Peripheral Arterial Disease	統計的有意性の不足
Preeclampsia	
Tardive Dyskinesia	
Agenesis of the Corpus Callosum with Peripheral Neuropathy	
Alpha-1 Antitrypsin Deficiency	
Autosomal Recessive Polycystic Kidney Disease	
Autosomal recessive spastic ataxia of Charlevoix-Saguenay	
Back Pain	
Beta Thalassemia	
Bloom's Syndrome	
BRCA Cancer Mutations (Selected)	
Breast Cancer Risk Modifiers	
Canavan Disease	
Congenital Disorder of Glycosylation Type 1a (PMM2-CDG)	
Connexin 26-Related Sensorineural Hearing Loss	
Cystic Fibrosis	
D-Bifunctional Protein Deficiency	GWASではない
Developmental Dyslexia	
Dihydrolipoamide Dehydrogenase Deficiency	
DPD Deficiency	
Factor XI Deficiency	
Familial Dysautonomia	
Familial Hypercholesterolemia Type B	
Familial Hyperinsulinism (ABCC8-related)	
Familial Mediterranean Fever	

Fanconi Anemia (FANCC-related)	
G6PD Deficiency	
Glycogen Storage Disease Type 1a	
Glycogen Storage Disease Type 1b	
GRACILE Syndrome	GWASではない
Hashimoto's Thyroiditis	
Hemochromatosis (HFE-related)	
Hereditary Fructose Intolerance	
Intrahepatic Cholestasis of Pregnancy	
LAMB3-related Junctional Epidermolysis Bullosa	
Larynx Cancer	
Leigh Syndrome, French Canadian Type (LSFC)	
Limb-girdle Muscular Dystrophy	
Malaria Complications	
Malaria Resistance (Duffy Antigen)	
Maple Syrup Urine Disease Type 1B	
Medium-Chain Acyl-CoA Dehydrogenase (MCAD) Deficiency	
Mucolipidosis IV	
Neural Tube Defects	
Neuronal Ceroid Lipofuscinosis (CLN5-related)	
Neuronal Ceroid Lipofuscinosis (PPT1-related)	
Niemann-Pick Disease Type A	
Nijmegen Breakage Syndrome	
Pendred Syndrome	
Phenylketonuria	
Placental Abruption	
Primary Hyperoxaluria Type 2 (PH2)	
Rhizomelic Chondrodysplasia Punctata Type 1 (RCDP1)	
Salla Disease	
Sarcoma	
Sickle Cell Anemia & Malaria Resistance	
Squamous Cell Carcinoma	
Tay-Sachs Disease	
Torsion Dystonia	
TTR-Related Cardiac Amyloidosis	
TTR-Related Familial Amyloid Polyneuropathy	
Tyrosinemia Type I	
Usher Syndrome Type I (PCDH15-related)	
Usher Syndrome Type III	
Zellweger Syndrome Spectrum	

体質

Memory	統計的有意性の不足
Nicotine Dependence	日本人で再現性なし
Smoking Behavior	SNPアレイにプローブなし
Avoidance of Errors	
Breastfeeding and IQ	
Cluster Headaches	
Eating Behavior	
Food Preference	
Lactose Intolerance	
Measures of Intelligence	
Muscle Performance	GWASではない
Non-ABO Blood Groups	
Norovirus Resistance	
Odor Detection	
Pain Sensitivity	
Reading Ability	
Resistance to HIV/AIDS	
Response to Diet	
Response to Exercise	

27

まとめ

- 289 の疾病・体質についてSNPデータから疾病リスク予測、体質解析のモデルを構築した
  - 論文の分析、品質管理、情報の更新法の確立
  - 結果の表示法の検討
- DTC遺伝子検査で取り扱う項目の検討・絞り込み
  - 単一遺伝子疾患、家族性腫瘍、小児疾患、先天性疾患、精神疾患、薬剤応答、才能などを除外
- 医業との関係、DTC遺伝子検査に対する倫理的・法的・社会的問題については継続して議論

28