

常業務に十分利用できる品質であることが確かめられた。

4.4 情報セキュリティに関する評価

情報セキュリティに関しては、現在、情報漏えいや外部からの攻撃は観測されていない。また、接続設定に関して、病院情報システム側に特に変更等の作業を行ってもいないために、接続設定作業により病院情報システム端末がウイルス等に感染したリスクも極小化されている。もっとも、情報漏えいや攻撃が精緻な場合、漏えいや障害そのものに気づかないまま推移している可能性もあるため、最終的な評価のためには、さらなる長期にわたる運用実績の蓄積とその評価が必要であろう。

5. おわりに

今回の受託研究により、わずかな投資のみで病院情報システム端末よりインターネットに安全に接続することが可能であることを示すことが出来た。この仕組みは、動画像等の動きのあるウェブページの閲覧には向かないが、静的なページやデータベースの利用等には十分に実用的であり、この仕組みを用いることで、臨床医が日常の診療を行っている外来や病棟から安全に未分類疾患の情報集約に関する研究班の情報サイトを利用できることが明らかとなった。

一方、情報システムは、長期に運用することによって故障の傾向や他のシステムからの影響等の問題が明らかになっていくことがある。そのために、今回得られた知見をより確実なものとしていくためには、実証実験の長期的な継続とさらなる情報収集が欠かせない。たとえば、今回設置したシステムにより、“富良野病院のような250床規模の病院であれば、2人同時接続の機材で十分実用に耐える”といった知見が得られる可能性があるが、結論に至るためにはさらなる運用経験の蓄積が必要である。そこで、今後の検証では、たとえば1ヶ月程度の期間に限定して、生じた問題や外部からの攻撃についての分析を行ったうえで、大きな問題がなさそうであれば期間を2ヶ月に増やし、また、その検証の上で期間を3ヶ月に延長し、といった漸進的な情報収集が必要となるだろう。

また、期間の延長だけでなく、情報システムの接続技術としての質的な向上も検討の余地がある。まず、病院情報システムの広域ネットワークへの間接接続に際して、高性能なリモートデスクトップ製品やシンクライアント製品を利用することにより、接続者数を増やしたり動画像等にも対応させるといった発展が考えられる。また、広域ネットワークへの安全な直接接続を実現することにより、感染症アウトブレイクなどの健康危機時における疫学情報収集の効率化、電子カルテ情報の交換や疫学的な二次利用等への対応も望ましい方向である。

現在、過剰とも言える個人情報保護と情報セキュリティ意識の高まりにより、病院情報システムのネットワーク化が避けられ、医療従事者の負担を増していると共に、医療の質に対しても悪影響が生じている。しかし、個人情報の漏えいは深刻な問題ではあるが、金銭的被害につ

いては損害保険により補填することが可能であり、病院の道義的責任についても、病院情報システムを過剰に守ることによって医療従事者に生じる負担と比してどれだけの優位を有しているか再検討の余地がある。本研究により、わずかな投資により病院情報システムから安全にインターネット接続が行えることが明らかとなりつつある。こうした試みの継続を通じて、日本中の医療従事者における業務負担の軽減と地域医療のさらなる充実が期待される。

謝辞

本研究の遂行に際しご協力を頂いた北海道日本電気ソフトウェアに深く感謝します。

レセプトデータによる特定疾患患者の診断確定までの受療パスの把握

分担研究者 岡本悦司 (国立保健医療科学院経営科学部経営管理室長)

研究要旨

特定疾患の診断が確定し、公費医療の対象になるまでにどのような医療機関を受診し、どのような診断や検査を受けてきたのか、の受療パスを 2008 年度より稼働しているレセプト情報・特定健診等情報データベース(レセプトナショナルデータベース,NDB)を用いて把握する。NDBはこうした研究利用が認められる見通しであるが、その申請と許可の手続きを定める大臣告示は今年度に間にあわなかったため、研究方法と NDB の仕様より申請に必要な SQL プログラムを作成を試みた。次年度においては研究利用ルールが告示された後すみやかに申請し、結果を得る。

A. 研究目的

特定疾患等の、希少・難治疾患患者は診断確定までに、様々な医療機関や診療科を受診し、時間的にも精神的にも負担が大きいと考えられる。それら患者について、診断確定にいたるまでの受療状況や中途の医師が疑った病名、および行なわれた検査等を分析し、それぞれの疾患について辿った受療パスの典型例をパターン化できれば、迅速な診断やスムーズな医療連携に資することができ、患者の時間的・精神的負担の軽減に役立つ。

わが国では、特定疾患や小児慢性疾患(以下、難治疾患)と診断確定すれば、公費負担の対象となり、医療機関から提出されるレセプトには、通常の保険者番号の他に公費負担番号が記載される。それゆえ、膨大なレセプトの中から難治疾患患者のレセプトを抽出することは容易である。

2008 年度より、高齢者医療確保法に基づき電子レセプトデータを毎月国が収集するデータベース(レセプト情報・特定健診等情報データベース、以下 NDB)が構築された。これは医療費適正化計画の「作成、実施、評価」を目的とするが、公益性の高い目的であれば研究利用を認める方針であり、研究利用ルールの大臣告示案が既にパブリックコメントも終了し告示を待つ状態である。また NDB では、個人は特定しないが、ハッシュ関数等の暗号手法を用いて同一人を長期的に追跡できるしくみとなっている。

そこで、NDB に研究利用申請を行い、新規に難治疾患の公費医療を受けだした患者について前の期間にどのような受療をたどったか、レセプトデータより把握し、疾患ごとの医療機関、傷病名そして診療行為(検査等)のパスを作成する。

B. 研究方法

NDB に以下のプログラムを実行し、必要なデータを抽出する。

- 1) 公費負担番号が、特定疾患、小児慢性疾患等に該当するレセプトを抽出する(下表参照)。
- 2) そのうち 2009 年 4 月以降にレセプトが出現しだした新規受給者を抽出する(対象者)
- 3) それら対象者について公費医療開始前の全医科レセプトより、医療機関、傷病そして診療行為レコードを抽出する。

分析手法としては、

- 1) 時系列的に以下のような内容を分析し、疾患ごとの典型的な受療パスをパターン化する。

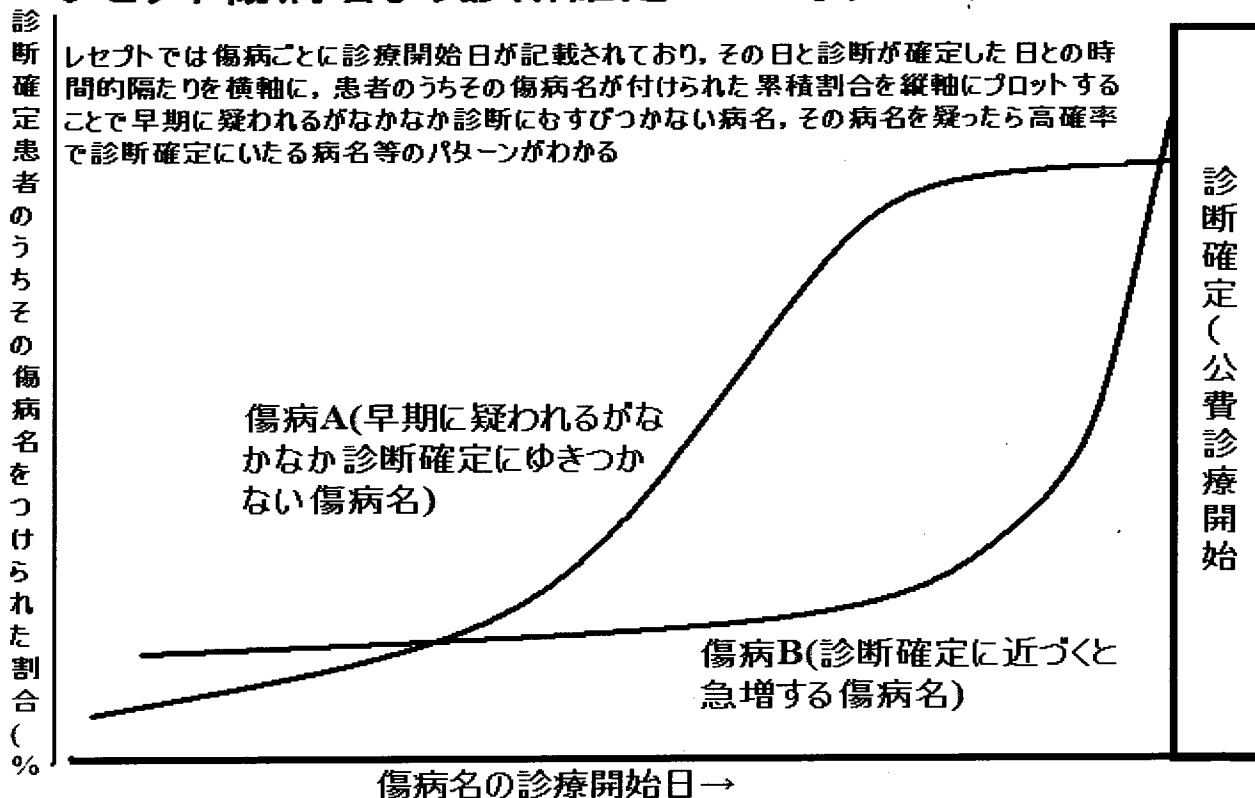
医療機関・診療科のパス(例: A診療所→B病院神経内科→国立Cセンター)

傷病名の変遷(例: A診療所胃腸炎→B病院クローン病疑い→国立Cセンター地中海熱)

主な検査の履歴(例: A診療所腹部単純→B病院胃透視→Cセンター大腸内視鏡)

2) 傷病名と検査については以下のような定量的分析も試みる。すなわちレセプト傷病名にはひとつひとつ診療開始日と転帰(治癒, 死亡, 中止(転医), その他)も記載されるので, たとえばサルコイドーシスの診断確定と公費診療開始までに複数の医療機関においていつどのような傷病名を疑って検査等が行なわれたか, を可視化する。いくつかの傷病について以下のようなプロットを作成することにより, どのような疾患の診断確定までにはどのような傷病名が疑われていったかのプロセスが示される。

レセプト傷病名より診断確定にいたるプロセスの分析法



C. 研究結果

本研究のためには国が高齢者医療確保法に基づき 2008 年 4 月より稼働している「レセプト情報・特定健診等情報データベース」の研究利用が前提となる。研究利用ルールは高齢者医療確保法に基づく大臣告示として出される予定で既にその告示案が示され 2009 年 10 月 14 日パブリックコメントが締め切られた。しかし 2010 年 2 月末日現在未告示であり, 残念ながら本年度の報告書にはまにあわなかった。そこで本年度は研究利用が可能になった時に申請する分析法を SQL で示し, 分析結果は次年度にのぼすこととした。

【参考】パブリックコメントならびに研究利用ルール告示案

<http://search.e-gov.go.jp/servlet/Public?CLASSNAME=Pcm1010&BID=495090161&OBJCD=100495&GROUP=>

本年度においては, レセプトデータの内容と活用法, 公開されたサンプルデータ(DPC レセプト)を用いて SQL プログラムの試作, そしてレセプトナショナルデータベースの仕様書より実際に活用できる SQL の案を示す。

C1. 電子レセプトデータ構造の検討

レセプトナショナルデータベースの活用の前提として、レセプトデータからどのような内容が得られるか検討した。下図はレセプトの見本である。右上の保険者番号に加えて、難病等の公費診療患者には左上の公費①公費②に公費負担者番号が入る。重要な点は公費の対象となる難病も、通常の傷病も同一レセプトに記載される、という点である。左側の点数欄の右側には公費欄があり、提供された診療行為のどれのどの額が公費負担となるか明記される。

傷病名

点数欄

点数

日数

摘要欄

電子で提供されるレセプトは紙レセプトと一対一対応してコード化される。

紙レセプトと電子レセプトは一対一対応

紙レセプトと電子レセプトは一対一対応
電子レセプト (レセコードデータ)

電子レセプト (レセコードデータ)

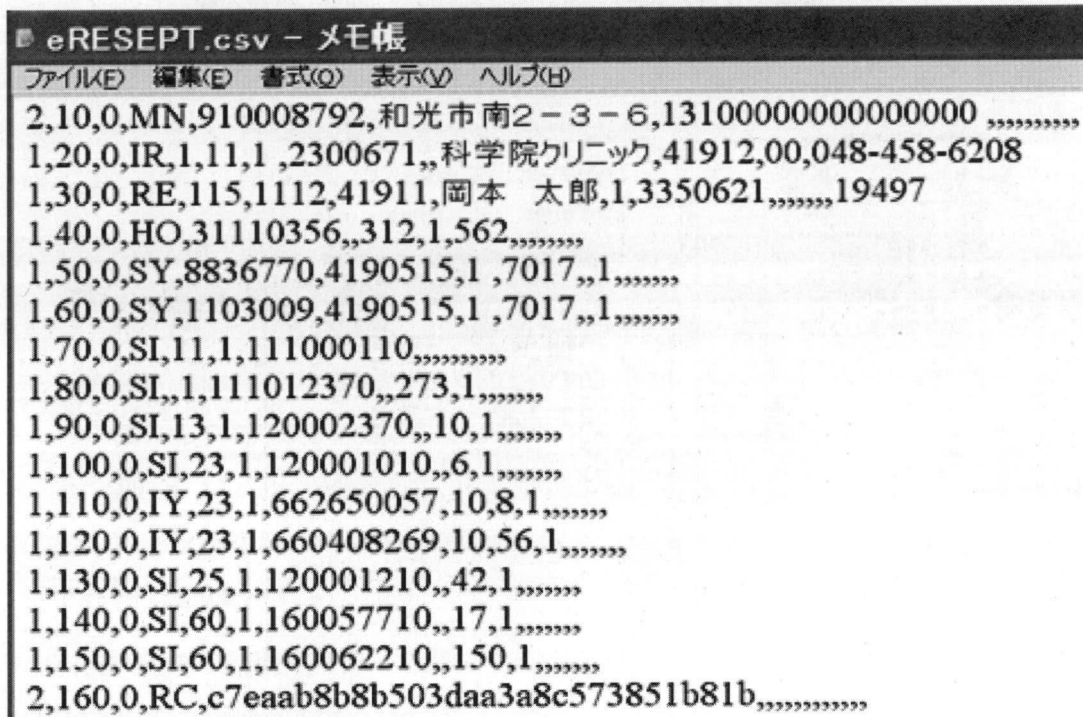
IR.1,13,1,9999996,基金病院41802,00...
RE.55,1116,41711,基金 花子,2,3220725...199,1116 55...01
HO.06139992,55,55,1,1441...
SY.8838506,4120824,1... SYは傷病名、179の傷病名コード診療
SY.8840829,4120824,1... 開始日等記載する。末尾001は主傷病、
6000は公費診療/001, 700007右/左/手/足
SY.2724007,4121116,1_01
SY.5319009,4131117,1...
SY.8835582,4131212,1,2056...
SY.2500013,4140131,1_01
SY.2500014,4171005,1,8002...
SI.12,1,112007410,58,1
SI.12,1,112011010,52,1
SI.12,1,112007170,5,1
SI.13,1,113002010,87,1
SI.13,1,120002370,10,1
SI.21,1,120000710,9,1
IY.21,1,610443013,2,000,16,56
SI.25,1,120001210,42,1
SI.25,1,120003170,45,1
SI.27,1,120001810,8,1
SI.60,1,160095710,12,1
SI.60,1,160061910,155,1
SI.60,1,160022410...
SI.1,160020910...
SI.1,160023410...
SI.1,160020610,62,1
GO.1,1441,99... GOは合計の点

画像データのイメージ

○厚生労働省の定める
規格及び方式
○レセプト電算処理マ
ターコードで記載
・傷病名マスター
・診療行為マスター
・医薬品マスター
・特定器材マスター
・コメントマスター

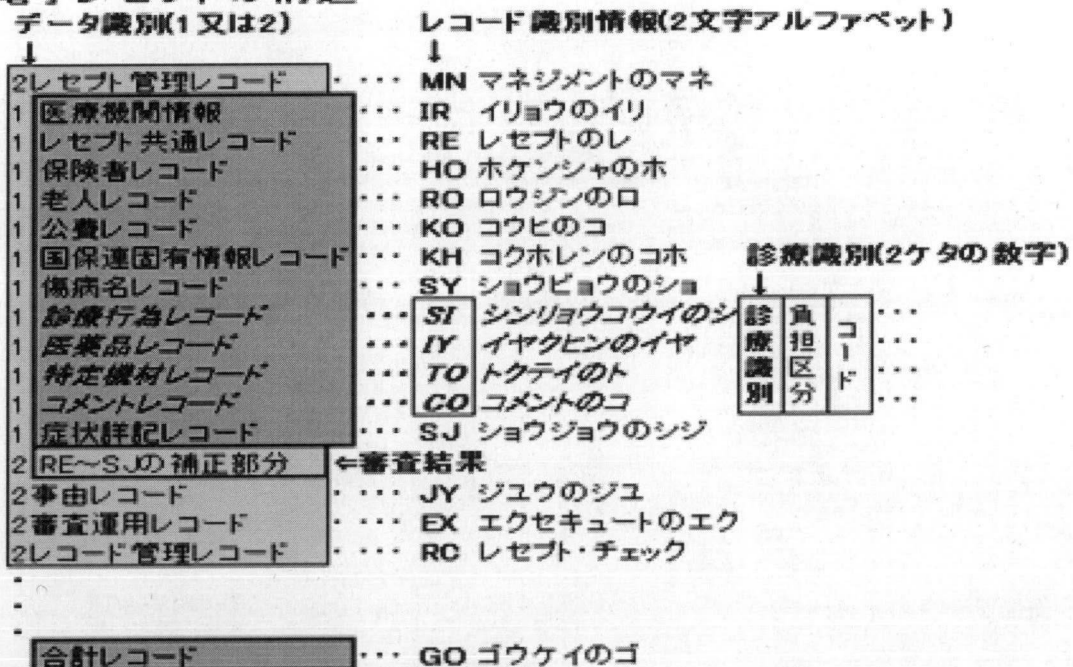
電子レセプトは csv 形式であり、以下はメモ帳で開いたところである。

電子レセプトをメモ帳で開くと



以下は電子レセプトの構造である。アルファベット 2 文字のレコード識別の読み方を理解するだけでも構造がグッと分かりやすくなる。

電子レセプトの構造

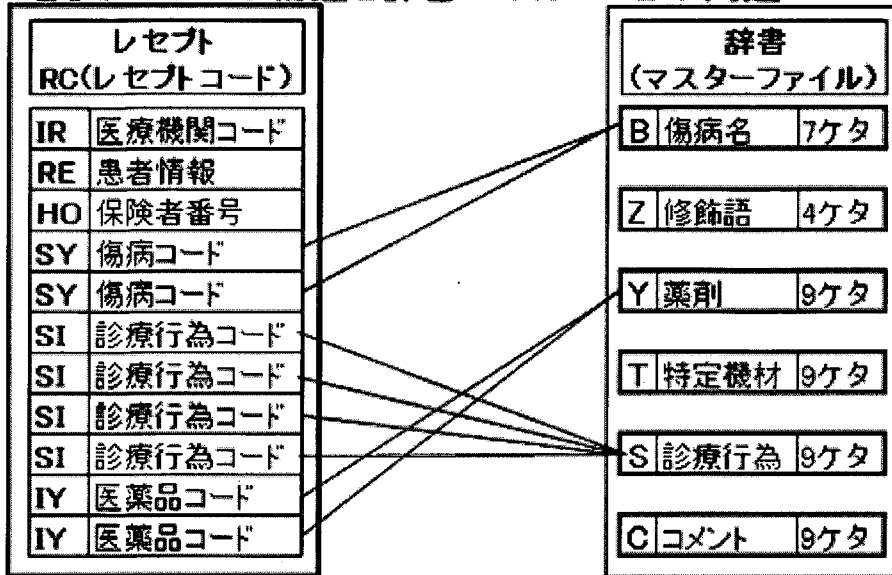


- ※ 太線(橙色)内は医療機関からの請求情報
- ※ 黄色部分は支払基金による追加情報
- ※ 斜体部分は摘要欄情報

数字の羅列にすぎないコードはマスターと呼ばれる辞書でリンクすることで解読される。

コードを辞書(マスター)で解読する

電子レセプトの構造と辞書(マスター)との関連



コードは無秩序に振られているのではなく規則がある。以下の解読法を理解するだけで無意味な数字の羅列も意味ある情報として理解できる。公費(KO)レコード中の保険者番号が次頁の特定疾患治療研究事業のコードである者が難病患者であり、NDB から抽出可能である。

レセプトコード解読表

保険者コード 9ケタ [][][][][][][][][] 法別番号 都道府県 コード 後証番号 【政管健保】 都道府県 社保事務 都道府県 社保事務	法別番号 00国保一般 06健保組合 27老人保健 67国保退職者
医療機関コード 10ケタ [][][][][][][][][][][] 都道府県 点数表 都市区 コード 後証番号	点数表コード 医1 歯3 薬4
医薬品コード 9ケタ [0][][][][][][][][] 医薬品 剤型 薬効分類 コード	剤型コード 内服1 注射4 外用5 衛生材料7
診療行為コード 9ケタ [1][][][][][][][][] 医科 診療種別 コード 取扱区分	取扱区分 基本項目1 合成項目3 準用項目5 加算項目7 通則加算項目9

特定疾患治療研究事業

(法別51)

(一部負担がない患者 601)
(一部負担がある患者 602)

名称	番号			
	法別	府県	機関	検
北海道	51	01	601 602	2 0
青森県	51	02	601 602	1 9
岩手県	51	03	601 602	0 8
宮城県	51	04	601 602	9 7
秋田県	51	05	601 602	8 6
山形県	51	06	601 602	7 5
福島県	51	07	601 602	8 4
茨城県	51	08	601 602	5 3
栃木県	51	09	601 602	4 2
群馬県	51	10	601 602	1 9
埼玉県	51	11	601 602	0 8
千葉県	51	12	601 602	9 7
東京都	51	13	601 602	8 6
神奈川県	51	14	601 602	7 5
新潟県	51	15	601 602	6 4
富山県	51	16	601 602	5 3

名称	番号			
	法別	府県	機関	検
石川県	51	17	601 602	4 2
福井県	51	18	601 602	3 1
山梨県	51	19	601 602	2 0
長野県	51	20	601 602	9 7
岐阜県	51	21	601 602	8 6
静岡県	51	22	601 602	7 5
愛知県	51	23	601 602	6 4
三重県	51	24	601 602	5 3
滋賀県	51	25	601 602	4 2
京都府	51	26	601 602	3 1
大阪府	51	27	601 602	2 0
兵庫県	51	28	601 602	1 9
奈良県	51	29	601 602	0 8
和歌山県	51	30	601 602	7 5
鳥取県	51	31	601 602	6 4
島根県	51	32	601 602	5 3

名称	番号			
	法別	府県	機関	検
岡山県	51	33	601 602	4 2
広島県	51	34	601 602	3 1
山口県	51	35	601 602	2 0
徳島県	51	36	601 602	1 9
香川県	51	37	601 602	0 8
愛媛県	51	38	601 602	9 7
高知県	51	39	601 602	8 6
福岡県	51	40	601 602	5 3
佐賀県	51	41	601 602	4 2
長崎県	51	42	601 602	3 1
熊本県	51	43	601 602	2 0
大分県	51	44	601 602	1 9
宮崎県	51	45	601 602	0 8
鹿児島県	51	46	601 602	9 7
沖縄県	51	47	601 602	8 6

C2. サンプル DPC データによる SQL プログラム

レセプトナショナルデータベースは厚生労働省保険局の端末において SQL によるデータ抽出が可能とされており(調達仕様書による)申請にあたっては SQL プログラムの提出が求められる可能性がある。そこで支払基金がネット上で公開している DPC レセプトを例にアクセス SQL を用いてデータ抽出法を説明する。

医療機関から提出された電子レセプトは通番すらふられておらず、また各行に患者氏名やレセプトを特定するデータが含まれていない。すなわち落としてバラバラになったら元に戻せない書類の束と同じである。それゆえ、データ分析のためには、たとえ一行(レコードという)だけ取り出しても誰のどのレセプトかわかるようにデータを挿入する必要がある。

ID	データ識別	レセプト内通番	枝	レコード	患者名	レセプト固有番号	F2	F3	F4	F5
1	2	10	0	MN			929000010		東京都港区	131421029
2	1	20	0	IR	患者名はレコード識別REのレコード(行)のF5に 含まれている		1	13	1	9999913
3	1	30	0	RE			3	1117	42101	患者3
4	1	40	0	HO			06139992	3	3	19
5	1	50	0	CO			01	1	810000001	特定入院
6	1	60	0	RE			3	1117	42101	患者3
7	1	70	0	HO			06139992	3	3	17
8	1	80	0	BU			060020XXC	4210113	4210129	9
9	1	90	0	SB			1514002	4036		C162
					レセプトを分析可能とするためには全レコード(行)に患者 名やレセプト固有番号を追加しなければならない。		1514002	4036		C162
12	1	120	0	SB			1514002	4036		C162
13	1	130	0	SB			8831494			C240
							2113008			K635
14	1	140	0	SB			7856003			R599

このような作業は Excel でなら VLOOKUP 関数等を用いて容易にできるが、アクセスのような SQL プログラムでこれを行うためには、サブクエリという手法を用いて、以下のように各レセプトの基本情報をヨコ並びに変換する必要がある。

そのためには全レコードに通番をふり,同一人,同一レセプトについてレコードと通番の対応表を作成しなければならない。

レコード識別

IR...医療機関レコード

RE...レセプト基本情報 これをこうしたい...→ IR RE HO KO ...

HO...保険者レコード

KO...公費レコード

⋮
⋮
⋮

このままでは
分析不能

ID レコード識別

1 IR...医療機関レコード

2 RE...レセプト基本情報

3 HO...保険者レコード

4 KO...公費レコード

⋮
⋮
⋮

まず通番を
ふる

レコード識別とIDの対応表を作成する

IR RE HO KO ...

1 2 3 4

1 16 17

35 36 37 38

レコード識別が RE, HO のレコードを抽出する SQL はそれぞれ以下の通りであり,まとめて *RESQL*, *HOSQL*, *KOSQL*, *IRSQL* と略記する。

SELECT * FROM DPC WHERE レコード識別='RE'... *RESQL*

SELECT * FROM DPC WHERE レコード識別='HO'... *HOSQL*

SELECT * FROM DPC WHERE レコード識別='KO'... *KOSQL*

SELECT * FROM DPC WHERE レコード識別='IR'... *IRSQL*

であるが, 第一段階として RE レコードと HO レコードを結合して REHO レコードを作成する。

そのためには二つのクエリを入れ子にする。入れ子にする場合, テーブル名にはその都度 AS に続けて異なる名称を記載しなければならない。HO レコードは必ず RE レコードの一つ次なので RE.ID=HO.ID-1 という条件でジョインできる。また変数名は RE.ID を REID というように表記する。

```
SELECT RE.ID as REID, HO.ID as HOID FROM (RESQL) as RE LEFT JOIN (HOSQL) as HO ON RE.ID=HO.ID-1... REHOSQL
```

上記の SQL を *REHOSQL* と略記して次に KO と結合する。

```
SELECT REHO.REID as REHOREID, REHO.HOID as REHOHOID, KO.ID as KOID FROM (REHOSQL) as REHO LEFT JOIN (KOSQL) as KO ON REHO.HOID=KO.ID-1... REHOKOSQL
```

さらに医療機関レコード(IR)と結合する。IR レコードは続きではなく離れた ID である場合もあるので, LEFT JOIN ではなく直積(CROSS JOIN)を用いる。直積は膨大な組合せが生成されるため REID で GROUP BY 大きく

る。REID 以外の数値は集計関数として MAX を用いる。

```
SELECT REHOKO.REHOREID as REHOKOREID, MAX(REHOKO.REHOHOID) as REHOKOHOID,
MAX(REHOKO.KOID) as REHOKOKOID, MAX(IR.ID) as IRID FROM (REHOKOSQL as REHOKO,
(IRSQL as IR WHERE REHOKO.REHOREID>IR.ID GROUP BY REHOKO.REHOREID . . .
IRREHOKOSQL
```

このようにレコード識別と ID の対応表ができる。

IR	RE	HO	KO
198	199	200	201
328	332	333	
668	669	670	
1117	1122	1123	
1117	1321	1322	
1626	1627	1628	
1835	1836	1837	
1928	1929	1930	
2002	2003	2004	
2117	2118	2119	
2381	2382	2383	
2500	2505	2506	2507
3006	3010	3011	
3185	3189	3190	
3419	3420	3421	
3588	3589	3590	3591
4004	4007	4008	
4124	4127	4128	
4124	4189	4190	
4276	4280	4281	
4276	4416	4417	
4687	4688	4689	
4772	4773	4774	
4834	4835	4836	
4928	4933	4934	4935
5014	5015	5016	5017
5544	5548	5549	
6348	6349	6350	
6490	6491	6492	
6512	6513	6514	
6753	6754	6755	
6920	6921	6922	

上の結果からわかるように、保険者レコード(HO)は全てのレコードにあるが公費負担レコード(KO)は一部であることがわかる。このようにレコード識別と ID の対応表ができると ID を使って患者名や病院名も表示できる。SQL 文中の LEFT JOIN を INNER JOIN に変えることで公費負担患者 5 人のみを抽出した結果である。公費負担番号の最初の 2 ケタは法別番号であり、12 は生活保護、16 は障害者自立支援法に基づく育成医療であることを示す。次の 2 ケタは都道府県であり東京都であることがわかる

IR	RE	HO	KO	公費負担	患者名	病院名
198	199	200	201	16139990	患者9	B病院
2500	2505	2506	2507	12139994	患者4	L病院
3588	3589	3590	3591	16139990	患者23	P病院
4928	4933	4934	4935	12139994	患者16	W病院
5014	5015	5016	5017	12139994	患者28	X病院

ちなみに最終的な SQL 文は以下の通り。

```
SELECT MAX(IR.ID) as IR, REHOKO.REHOREID as RE, MAX(REHOKO.REHOHOID) as HO,
```

MAX(REHOKO.KOID) as KO, MAX(REHOKO.KOF2) as 公費負担者番号, MAX(REHOKO.REHOREF5) as 患者名, MAX(IR.F7) as 病院名 FROM (SELECT REHO.REID as REHOREID, REHO.HOID as REHOHOID, REHO.REF5 as REHOREF5, KO.ID as KOID, KO.F2 as KOF2 FROM

(SELECT RE.ID AS REID, RE.F5 as REF5, HO.ID AS HOID FROM

(SELECT * FROM DPC WHERE レコード識別='RE' and F19 IN ('0','2')) as RE

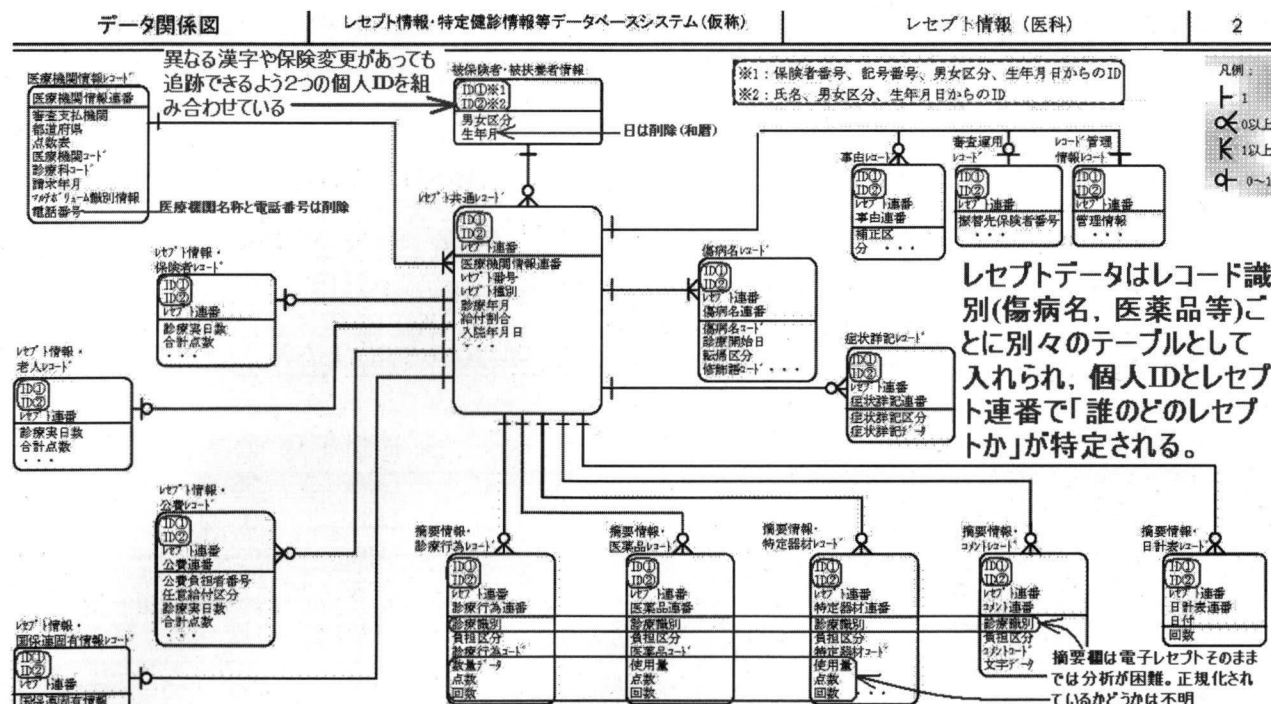
LEFT JOIN

(SELECT * FROM DPC WHERE レコード識別='HO') as HO

ON RE.ID=HO.ID-1) as REHO INNER JOIN (SELECT * FROM DPC WHERE レコード識別='KO') as KO ON REHO.HOID=KO.ID-1) as REHOKO, (SELECT * FROM DPC WHERE レコード識別='IR') as IR WHERE REHOKO.REHOREID>IR.ID GROUP BY REHOKO.REHOREID

C3 レセプトナショナルデータベースの仕様と SQL 案

ナショナルデータベースの仕様は既に公開されているので、実際の SQL 案を作成した。



氏名や生年月日(しかし年月は収集される)等の個人情報(暗号化されて収集される。たとえば、渡辺と渡邊というふうに同一人でも漢字が異なることもあれば、保険の記・番号も変更になることがある。それでも同一人を追跡できるよう、暗号は「保険者番号、記・番号、性、生年月日から生成される ID1」と「氏名、性、生年月日から生成される ID2」が併用される。

また複雑なレセプトデータはレコード識別(医薬品は IY, 傷病名は SY 等)ごとに別テーブルとして個人 ID とレセプト連番が振られるようである。

特定疾患の公費医療を受けている患者 ID とレセプトは以下のように抽出される。

```
SELECT ID1, ID2, レセプト連番 FROM 公費レコード WHERE 公費負担者番号 IN (特定疾患治療研究事業のコード)
```

次に2009年4月前には公費負担がなかった患者を絞って「対象者リスト」という別テーブルにする。そのためにはレセプトの診療年月が必要。診療年月はレセプト共通レコードに含まれているのでレセプト連番でJOINする。2009年4月以降に公費診療を開始した患者とは、公費レコードのあるレセプトの診療年月の最小が2009年4月(レセプトは和暦なので42104)という患者であるから

```
SELECT ID1 as 個人ID, MIN(診療年月) AS 公費開始年月 INTO 対象者リスト FROM レセプト共通レコード INNER JOIN 公費レコード ON レセプト共通レコード.レセプト連番=公費レコード.レセプト連番 WHERE 公費負担者番号 IN (特定疾患治療研究事業コード) GROUP BY ID1 HAVING MIN(診療年月)>=42104
```

得られた対象者リストの患者について、公費開始年月より前の全てのレセプトの情報を抽出する。受療パスに必要なのは、医療機関レコード、傷病レコードそして診療行為レコードであり、いずれもレセプト共通レコードとレセプト連番でJOINし診療年月も取得する。

```
SELECT * FROM 医療機関レコード INNER JOIN レセプト共通レコード ON 医療機関レコード.レセプト連番=レセプト共通レコード.レセプト連番 WHERE 医療機関レコード.ID1 IN (SELECT 個人ID FROM 対象者リスト)
```

上のSQLを医療機関レコードを傷病レコード、診療行為レコードにそれぞれ置き換えて実施する。

特定疾患コードでは特定疾患の公費医療の対象となっていることはわかるが、どの疾患で対象になっているかわからない。特定傷病名のみにしぼるのなら上で作成された傷病レコードで対象者をしぼる。たとえばサルコイドーシスおよび関連疾患の患者に絞るには別に下のようなSARCOIDOSISというテーブルを作成しておき制約条件を追加する。

```
SELECT * FROM 傷病レコード INNER JOIN レセプト共通レコード ON 傷病レコード.レセプト連番=レセプト共通レコード.レセプト連番 WHERE 傷病レコード.ID1 IN (SELECT 個人ID FROM 対象者リスト) AND 傷病コード IN (SELECT SHOBYOCODE FROM SARCOIDOSIS)
```

SHOBYOCODE	ICD10	ICD9
1359007 肝サルコイドーシス	D868	0302
1359008 眼サルコイドーシス	D868	0302
1359002 サルコイドーシス	D869	0302
8834027 サルコイドーシス 性肝肉芽腫	D868	0302
8834028 サルコイドーシス 性虹彩毛様体炎	D868	0302
1359003 サルコイドーシス 性ぶどう膜炎	D868	0302
8834029 サルコイドーシスにおける多発性脳神経麻痺	D868	0302
1359009 心サルコイドーシス	D868	0302
1359012 頭蓋内サルコイドーシス	D868	0302
8838829 肺サルコイドーシス	D860	0302
8839374 皮膚サルコイドーシス	D863	0302
8841116 リンパ節サルコイドーシス	D861	0302

また、対象者リストの対象者ごとに全ての傷病コードとその最初(最小)の診療開始日を抽出するには、

```
SELECT ID1, 傷病コード, MIN(診療開始日) FROM 傷病レコード WHERE ID1 IN (SELECT 個人ID FROM 対象者リスト) GROUP BY ID1, 傷病コード
```

E. 研究発表

本年度においては、実際の結果が得られなかったため、まだ発表されたものはない。

F. 知的財産権の出願・登録

該当なし。

病態解明に向けた原因不明先天異常バイオリソースの構築と潜在的 ゲノム構造異常解析

研究分担者 羽田 明（千葉大学大学院医学研究院教授）
稲澤 譲治（東京医科歯科大学難治疾患研究所教授）

研究要旨

外表奇形を伴う発達遅滞（MCA/MR）の多くは原因不明であり、臨床診断にしばしば困難を伴う。我々はこの原因探索を目的として、全国の医療機関 23 施設とコンソーシアムを形成して症例収集を行い、ゲノム解析を行ってきた。主たる解析手法はアレイ CGH である。既知染色体異常症の原因領域と全染色体のサブテロメア領域に特異的な BAC クローンを収載した診断型アレイを用いて迅速な診断を行うとともに、陰性症例に対しては全ゲノムを 4523 個の BAC でカバーする MCG Whole Genome Array-4500 を用いた二次スクリーニングを行い、結果として全体の約 1/4 に疾患との関連が疑われるゲノム異常を検出した。また、検出されたゲノムコピー数異常を手掛かりとし、疾患に関与する遺伝子 CASK, BMP4, YWHAЕ を指摘するなどの成果を挙げてきた。われわれの研究は、診断が比較的困難であり病態の機序が不明であった MCA/MR 領域の臨床や研究に寄与することが期待できる。

A. 研究目的

本研究の目的は、アレイ CGH その他の手技を用いて MCA/MR の疾患成立に関与すると考えられるゲノム異常の存在を明らかにし、これを端緒として MCA/MR の病態を明らかにしてゆくことである。

B. 研究方法

全国 23 施設と形成したアレイコンソーシアムにおいて MCA/MR 症例を収集し、臨床遺伝専門医によってアレイ解析への適否を検討する。症例はアレイ CGH を主に用いて、ゲノムコピー数変化（copy number variants; CNV）を下記の要領でスクリーニングする。解析対象となるのは、臨床的に診断がつかない、あるいは既知の症候群としては非定型的な症状を示し、染色体核型正常である MCA/MR である。

最初に、既知染色体異常症の原因領域と全染色体のサブテロメア領域に特異的な BAC を

収載した既知先天異常症診断型アレイである MCG Genome Disorder (GD) アレイによる一次スクリーニングを行い、陰性症例に対しては、全ゲノムを 4523 個の BAC でカバーする全ゲノム探索型アレイである Whole Genome Array-4500 (WGA4500) による二次スクリーニングを施行する。このことにより、既に疾患のエンティティが確立された症候群、特に従来法では診断困難であった微細ゲノムコピー数異常症の迅速な診断と、病因性が不明である CNV に対しての詳細な検討の両立が期待できる。

病因性が不明である CNV に対しては、以下のような方法で検討を行ってきた。① FISH 法や定量 PCR などによるゲノム構造異常の詳細な解析 ② 両親・家族解析による当該 CNV が一般健常者集団にも出現しうるものであるか否かの検討 ③ 文献やデータベースへの照応によるゲノム異常と疾患との関連の判断

④RT-PCR や Northern blotting などを用いた発現解析 これらの検討を経て疾患との関連が強く疑われる遺伝子に対しては、マウスやゼブラフィッシュなどのモデル動物の作製、類似の phenotype を示す患者集団における変異解析などを行い、疾患形成の原因となりうる遺伝子の同定を目指す。

また、これらの成果を基盤として、ゲノム異常と表現型の関連 (genotype-phenotype) を明らかにしてゆくためにバイオリソースならびに疾患データベースの拡充を図ってゆく。(倫理面への配慮)

本計画研究は、「ヒトゲノム・遺伝子解析研究に関する倫理指針 (平成 13 年 3 月 29 日 文部科学省・厚生労働省・経済産業省共同告示第 1 号) (平成 16 年 12 月 28 日全部改正) (平成 17 年 6 月 29 日一部改正)」が定める A 群ならびに B 群試料等に相当するヒト細胞を用いるため、「ヒトゲノム研究に関する基本原則」(科学技術会議生命倫理委員会)を十分理解し、「ヒトゲノム・遺伝子解析研究に関する倫理指針」を遵守して実施する。本研究で用いられるヒト試料は、東京医科歯科大学難治疾患研究所ならびに臨床施設において倫理審査委員会にて承認された手続きに基づき収集され、具体的には、臨床施設でインフォームドコンセントを取得され、匿名化処理が施された後、東京医科歯科大学・難治疾患研究所の細胞ならびに DNA バンクに供与され、ヒトゲノム・遺伝子解析研究に用いられるため、個人情報漏洩することはないシステムが担保されている。

C. 研究結果

我々は 2005 年 10 月～2009 年 9 月の 4 年間にわたり、科学技術振興機構 (JST) CREST および新エネルギー・産業総合開発機構 (NEDO) の支援のもとにコンソーシアムから 534 例の MCA/MR 症例を収集し、1 次スクリーニングにおいて 55 例 (10.3%) に CNV を検出した。2009 年 10 月以降は当該厚生労働科学研

究費補助金 (難治性疾患克服研究事業) の支援を受けこれまでに 46 症例を解析し、7 例 (15.2%) に CNV を検出した。これらは診断型アレイである GD アレイにより検出された異常であり、いずれも疾患に関連する CNV (pathogenic CNV) であると考えられた。

また、これらの陰性症例の中から現在までに 2 次スクリーニングにおいて 353 症例を解析し、66 例に 69 箇所 CNV を検出した。これらの CNV が疾患に関連するか、あるいは一般健常者集団の中にも見出されるものであるのかを判別するために種々の評価を行い、pathogenic CNV を有するのは 49 例 (13.9%) であると判断した (図 1)。

これらの検索の中で、疾患関連遺伝子として CASK, BMP4, YWHAЕ, GLI3 のハプロ不全に起因する症例を見出し、前者 2 例を報告した。同一箇所反復する CNV を 3 組 6 例に見出し、これらが新規症候群と定義される可能性を考えて検索を続けている。また、本スクリーニングより胎児期の神経発達に影響を及ぼすことが疑われる候補遺伝子を見出し、モデル動物作製を試行している。

D. 考察

従来先天異常疾患の診断は、主に臨床症状を手がかりとして行われてきたが、非定型的な症候を有するため診断に苦慮する例も決して稀ではない。日本産婦人科医会による先天異常モニタリング調査によれば、全出生の 1.8% に先天異常を有する児が報告されているが (2003 年)、このうち 60～70% は原因が不明である。このような症例に対して従来から行われてきた G 分染による染色体核型分析の解像度はいわゆるバンドレベルの最小でも 5～10Mb 程度にとどまるため、これよりも微細なゲノムコピー数異常を検出することは不可能である。先天異常疾患における染色体異常の検出頻度は対象をどのように選ぶかにもよるが、8%～27% と報告により大きく差があった。

我が国の末梢血リンパ球の染色体検査の異常検出率は10~15%程度である。

アレイ CGH はこのような染色体微細構造異常を高精度・高密度に検出することのできるツールであり、先天異常疾患の領域でも近年広く用いられつつある。一方でアレイ CGH による解析の進展に伴い、一般健常者集団にも相当数の CNV が存在し、さらに de novo CNV も個人あたり少なくとも2,3個は存在することも明らかになってきている (Lee et al., 2007)。このためにアレイ CGH で検出された CNV の臨床的意義の検討は必須となる。

本研究において我々は、原因不明であった MCA/MR 症例に対して診断型アレイと全ゲノムアレイを用いて二段階スクリーニングを行うことにより、1次スクリーニングでは既知の疾患とサブテロメア異常を迅速に診断して臨床に寄与することと、2次スクリーニングでは検出された CNV の臨床的意義を詳細に検討することを両立させることができた。結果として1次スクリーニングでは全体の10.3%に、2次スクリーニングでは13.9%に pathogenic CNV を検出した。このことは、従来臨床的に診断がつかず、核型分析でも異常を見いださなかった MCA/MR 症例において、約1/4に疾患原因を見出す可能性を示唆している。のみならず、今回検出した CNV を端緒として、疾患関連遺伝子の検出や、新規症候群を定義する可能性のある CNV などを見出すことができた。これらの成果は当教室にて推進している genotype-phenotype データベースの拡充に寄与し、先天異常疾患の臨床・基礎研究の両面における重要なリソースになることが期待される。具体的には、類似疾患の集積による新たな疾患概念の構築、療育や遺伝相談への情報提供、病態予測、治療方法の開発などである。MCA/MR はゲノム異常に端を発し胎生期より phenotype が形成されてくる疾患であるため、これらの疾患原因探索はヒトの発生発達の機序解明に貢献し、生命科学の発展に寄与することが期待できる。

E. 結論

アレイ CGH を適切に運用することで、先天異常疾患の臨床ならびに原因探索の両面において有意義な成果を挙げることができた。

【研究協力者】 (敬称略)

旭川医大：蒔田芳男

東京医科歯科大学難治疾患研究所：

林 深、本田尚三、岡本奈那、井本逸勢、

F. 研究発表

1. 論文発表

Honda S, Orii K, Kobayashi J, Hayashi S, Imamura A, Imoto I, Nakagawa E, Goto Y, Inazawa J: Novel deletion at Xq24 including the UBE2A gene in a patient with X-linked mental retardation. *J Hum Genet.* 2010 (in press)

Hayashi S, Kurosawa K, Imoto I, Mizutani S, Inazawa J: Detection of cryptic chromosome aberrations in a patient with a balanced t(1;9)(p34.2;p24) by array-based comparative genomic hybridization. *Am J Med Genet* 2005; 139: 32-36

Honda S, Hayashi S, Kato M, Niida Y, Hayasaka K, Okuyama T, Imoto I, Mizutani S, Inazawa J. Clinical and molecular cytogenetic characterization of two patients with non-mutational aberrations of the FMR2 gene. *Am J Med Genet.* 2007. 143A:687-93.

Hayashi S, Ono M, Makita Y, Imoto I, Mizutani S, Inazawa J. Fortuitous detection of a submicroscopic deletion at 1q25 in a girl with Cornelia-de Lange syndrome carrying t(5;13)(p13.1;q12.1) by array-based comparative genomic hybridization. *Am J Med Genet A.* 2007. 143A:1191-7.

Hayashi S, Honda S, Minaguchi M, Makita Y, Okamoto N, Kosaki R, Okuyama T, Imoto I,

Mizutani S, Inazawa J. Construction of a High-density and High-resolution Human Chromosome X Array for Comparative Genomic Hybridization Analysis. *J Hum Genet.* 2007. 52:397-405.

Tokutomi T, Hayashi S, Imai K, Chida A, Ishiwata T, Asano Y, Inazawa J, Nonoyama S. dup(8p)/del(8q) recombinant chromosome in a girl with hepatic focal nodular hyperplasia. *Am J Med Genet A.* 2007. 143A:1334-7.

Okamoto N, Kubota T, Nakamura Y, Murakami R, Nishikubo T, Tanaka I, Takahashi Y, Hayashi S, Imoto I, Inazawa J, Hosokai N, Kohsaka S, Uchino S. 22q13 Microduplication in two patients with common clinical manifestations: a recognizable syndrome? *Am J Med Genet A.* 2007. 143A:2804-9.

Hayashi S, Mizuno S, Migita O, Okuyama T, Makita Y, Hata A, Imoto I, Inazawa J. The CASK gene harbored in a deletion detected by array-CGH as a potential candidate for a gene causative of X-linked dominant mental retardation. *Am J Med Genet A.* 2008. 146A:2145-51.

Hayashi S, Okamoto N, Makita Y, Hata A, Imoto I, Inazawa J. Heterozygous deletion at 14q22.1-q22.3 including the BMP4 gene in a patient with psychomotor retardation, congenital corneal opacity and feet polysyndactyly. *Am J Med Genet A.* 2008. 146A:2905-10.

邦文総説

林深、井本逸勢、稲澤譲治：ゲノム微細構造異常と先天異常疾患．（株）診断と治療社．小児科診療．第 72 巻第 1 号：31-37，2009（2009/1/1）(7P)

林深、稲澤譲治：アレイ CGH による先天異常の解析．金原出版株式会社．小児科．第 50 巻第 7 号：847-853，2009（2009/9/30）

邦文著書

稲澤譲治、蒔田芳男、羽田 明：アレイ CGH 診断活用ガイドブック-知っておきたい染色体微細構造異常症-．医薬ジャーナル社（大阪・東京）2008，2，15

2. 学会発表

・国際学会 [ポスター]

1. Hayashi S, Imoto I, Inazawa J. Result of exploring pathogenesis of congenital disorders using array-CGH. The American Society of Human Genetics 55th annual meeting, Salt Lake City, October 28, 2005.

2. Hayashi S, Honda S, Imoto I, Inazawa J. Exploring causative genomic aberration in congenital disorders using CGH-arrays. The American Society of Human Genetics 56th annual meeting, New Orleans, October 10, 2006.

3. Hayashi S, Honda S, Imoto I, Inazawa J. Exploring cryptic genomic aberrations related to multiple congenital anomaly with mental retardation using in-house CGH-arrays. The American Society of Human Genetics 57th annual meeting, San Diego, October 25, 2007.

4. Hayashi S, Honda S, Imoto I, Inazawa J. Application of in-house array-CGH for investigation and diagnosis of congenital genomic disorders. The American Society of Human Genetics 58th annual meeting, Philadelphia, Nov 12, 2008.

5. Hayashi S, Honda S, Mizuno S, Okamoto N, Makita Y, Hata A, Imoto I, Inazawa J. Analyses of 499 Cases of Multiple Congenital Anomalies with Mental Retardation using array-CGH for Investigation and Diagnosis. The American Society of Human Genetics 59th annual meeting, Honolulu, October 22, 2009.

・国内学会 [口演]

1. 林深、水口真希、井本逸勢、稲澤譲治 高精度ゲノムマイクロアレイの開発とそれによる遺伝疾患の潜在的微小ゲノム構造異常の探索 第49回日本人類遺伝学会 2004年10月14日、東京
2. 林深、井本逸勢、稲澤譲治 高密度ゲノムアレイを用いた先天異常疾患の解析 第50回日本人類遺伝学会 2005年9月22日(ワークショップ)、倉敷
3. 林深、本田尚三、井本逸勢、水谷修紀、稲澤譲治 高密度ゲノムアレイを用いた先天異常疾患の解析とゲノム多型の検出 第51回日本人類遺伝学会 2006年10月19日、米子
4. 林深、本田尚三、井本逸勢、稲澤譲治:アレイ CGHによる先天異常症の潜在的染色体異常診断とその解析. 日本人類遺伝学会第52回大会. 2007年9月14日(シンポジウム)、東京.
5. 林深、本田尚三、井本逸勢、稲澤譲治:アレイ CGHを用いた多発奇形を伴う精神遅滞症例の解析と診断. 日本人類遺伝学会第53回大会. 2008年9月28日、横浜.
6. 林深、岡本奈那、本田尚三、井本逸勢、蒔田芳男、羽田明、稲澤譲治:アレイ CGHを用いた多発奇形を伴う精神遅滞症例解析の4年間の実績. 日本人類遺伝学会第54回大会. 2009年9月24日、東京.
7. 林深、岡本伸彦、水野誠司、小野正恵、小崎里華、奥山虎之、知念安紹、蒔田芳男、羽田明、井本逸勢、稲澤譲治:小頭症と小脳脳幹部低形成を伴う発達遅滞12例におけるCASK遺伝子の解析. 日本人類遺伝学会第54回大会. 2009年9月26日、東京.

G. 知的財産権の出願・登録状況

なし

未分類疾患の情報集約に関する研究

分担報告書

分担研究名： 境界領域の膠原病の疾患概念確立へのアプローチ

分担研究者 国際医療福祉大学クリニック 鈴木 元

分担金額： 1000 千円

研究要旨

ウィルスや化学物質等に対する免疫応答を契機に自己免疫応答が惹起されたり、化学物質による直接的なシグナル作用により、膠原病類似の生体応答が誘導されたりする事が知られている。後者の機序による疾病は、公衆衛生的な対処が可能な疾病であるが、既知の膠原病あるいはその疑い症例として診断されている可能性がある。本年度は、過去に欧米でアウトブレイクした例（toxic oil syndrome, eosinophili-myalgia syndrome、EMS）を題材にして、EMS と既存の疾患（好酸球性筋膜炎）とのオーバーラップに関して検討した。

A.研究目的

膠原病は、基本的に自己免疫反応をベースに慢性的な血管炎や結合織炎により病態が形成される。しかし、自己免疫応答を介さずに、化学物質が直接細胞のシグナル伝達系に働き、炎症応答類似の生体応答を惹起し、もって膠原病類似の疾病を発症させる場合がある。特異的な自己抗体などのバイオマーカーを欠く膠原病は、元々症状の組み合わせのみで診断される場合が多いため、異なる病因の疾患を纏めただけの「独立疾患」である場合がある。このような「独立疾患」の中には、化学物質により惹起された疾患が紛れ込む可能性がある。化学物質中毒として認知される場合を例外とすると、このような疾病を正しくスクリーニングすることは、容易ではない。さらに、既知の膠原病自体、初発から診断基準を満たす症状や徴候や検査データが揃うまでに1年以上の時間がかかることも希ではない。本研究は、公衆衛生的な対処が必要な化学物質による膠原病類似の疾病が出現した場合に備え、過去の事例を材料にそのような疾患をスクリーニングする手法に関して検討する。

B.研究方法

H21 年度は、文献調査により、アウトブレイクした新疾患を検討する中から、スクリーニング手法を学ぶ事とした。

（倫理面への配慮）

本研究は、文献調査であるので、倫理的問題は該当しない。

C.研究成果

ケース 1. toxic oil syndrome

1981年にスペインのマドリッドでアウトブレイクした toxic oil syndrome は、びまん性間質性ないし肺胞-間質性肺炎像の X 線像を伴う急性呼吸窮迫症候群 (ARDS)、筋痛、好酸球増多、皮疹・浮腫等の皮膚症状、神経症状などを呈する疾患で、既存の疾患としては硬皮症や皮膚筋炎などに類似した¹。約 19,900 名が発症し、1,663 名が死亡した。本症候群では、短期間に家族発症例が続発し、同一地区での患者が多発するといった特徴から、当初は感染症を疑い疫学調査がなされた。複数の感染性微生物が病原菌として疑われ、抗生物質による症例対照研究も行われたが、結果は否定的であった。その後、小児科医が乳幼児に疾患発症が少ないことに着目し、6ヶ月未満で同症候群を疑われた症例と、外傷や手術で入院していた6ヶ月未満の乳幼児を対照とした症例対照研究を実施し、きめ細かい調査(離乳食にオリーブオイルを使っている場合はそのブランド名を含む)を実施した。この結果、無ブランドの5Lボトルの「オリーブオイル」がリスク要因として浮かび上がった。この発見を契機に、成人を含めた疫学調査が実施され、無ブランドの5Lボトルの「オリーブオイル」がリスク因子として確定した。第1症例の入院から約40日後にこの「オリーブオイル」の回収が開始され、その後4~5ヶ月後に新規症例発症がなくなった。

背景としてスペインは、オリーブオイル以外の植物性食用油の輸入を禁止していた。フランス産の菜種油などを工業用目的でスペインに輸入する場合、アニリンやひまし油やメチレンブルーを添加し、食用には適さないようにしていた。この2%アニリン添加菜種油を再精製して「オリーブオイル」として行商やオープンマーケットで売られていた物の中に、toxic oil syndrome を起こした化学物質(アニリン化合物, 3-(N-phenylamino)-1,2-propanediol がもっとも相関)が精製過程で生じたと思われる。しかし、動物実験で病状を再現できなかったこと、保存されていた「オリーブオイル」のなかで化合物の分解が進んだことなどの理由で、2009年段階でもどの化合物が真の病因かは同定されていない²。

本症候群は、急性期こそ呼吸困難と好酸球増加症といった Loeffler 症候群を思わせる症状が主体であるが、中間期や慢性期は筋痛や強皮症様の皮膚症状、関節拘縮、運動障害、シェーグレン症候群など膠原病様の症状が全面に出てくる。血清中に可溶性 IL-2R α や IgE や G-CSF が増加し、肺組織で IL4 mRNA や IL5 mRNA の発現増加と好酸球浸潤や T 細胞 (Th2) 浸潤が認められることなど、免疫学的機序を伺わせる所見もある。ゲノム多型(第6染色体 D6S1050 microsatellite 多型、および arylamine N-acetyltransferase 多型)が症状の重さに相関することも知られている。他方、特異的な自己抗体は発見されていない。

¹ De la Paz MP, Philen RM, Borda IA: Toxic oil syndrome. The perspective after 20 years. *Epidemiologic Reviews* 23 (2): 231-47, 2001.

² Rangan C, Barceloux DB. Food contamination. In Barceloux DG, *Medical Toxicology of Natural Substances: Foods, Fungi, Medical Herbs, Toxic Plants, and Venomous Animals*. Hoboken, NJ: John Wiley & Sons, 2008. p. 5-21.