

「DPCデータの利活用に向けたサンプリング データセットの整備に関する研究」 (H26-特別-指定-017) 結果概要

産業医科大学
公衆衛生学教室
松田晋哉

研究の背景

- 「新たな情報通信技術戦略 工程表（高度情報通信ネットワーク社会推進戦略本部決定）」（平成22年6月22日）
 - － レセプト情報等データベースの第三者提供に加えて、DPCデータの第三者提供についても提供形態の決定、ガイドライン策定に関する検討を指示
- 「レセプト情報等の提供に関する有識者会議」
 - － DPCデータの第三者への提供のあり方について議論

資料及び方法

- 一般社団法人 診断群分類研究支援機構（以下DPC研究機構）が収集した平成24年度データを用いた（1,057施設、6,852,048件）
 - データ使用に関しては個々の施設とDPC研究機構との間で個別に契約を結び、かつ産業医科大学倫理委員会の審査を受けた。
- 以下の3種類のデータセットを作成
 - 汎用サンプリングデータ
 - 特定の研究目的に対応したデータセット
 - 階層的ランダムサンプリングデータ

汎用サンプリングデータ(コストデータ)

- RecID: 1入院ごとに割り付けられる患者ID(データ識別番号)をハッシュ化
- dpccd: DPC14桁(研究班で再コード化)
- sex: 性別 1=男性、2=女性
- los: 在院日数
- age_c10: 年齢階級 0=0-9歳、1=10-19歳、2=20-29歳、・・・、9=90歳以上
- 診療区分ごとの処理: 1桁目を四捨五入した出来高換算の診療区分10の点数

作成したDPCモデルデータセット (コストデータ:平成24年度研究班データ)

RecID,dpood,sex,los,age_c10,診療区分 10,診療区分 20,診療区分 30,診療区分 40,診療区分 50,診療区分 60,診療区分 70,診療区分 80,診療区分 90,診療区分 97↵

0x00000361C4A65E1EACE479D9ABE96E1BAC7094D0,060020xx01x0xx,2,14,05,4000,2630,1860,200,80260,3080,420,NULL,27290,1630↵

0x0000011A1DC8074432DB42E1B4ED970C9F61F238,040311xxxxxxxx,2,3,06,450,830,150,NULL,NULL,15890,3560,NULL,9400,600↵

0x00000AE286243B9BD1C77E3B0ACC23A08D001AF5,040200xx99x00x,1,2,02,720,100,230,820,NULL,1490,2960,NULL,8300,130↵

0x000009EC750068B73DDFC08562F3B65E7BF11A6,060050xx99x00x,1,6,06,NULL,80,NULL,NULL,NULL,3430,6700,NULL,15470,810↵

0x000013F16A6B17611D25ED27776E76EBB51314FB,060160x003xx0x,1,6,06,630,120,140,100,35430,30,NULL,NULL,14170,860↵

0x00000B0FBBA456778FA0D05FAEDC442A19F26828A,060050xx99x00x,1,12,07,710,90,400,70,NULL,5750,9270,NULL,29830,1790↵

0x0000154D1E4DBDE88E1BC668953F72BEA934414D,060100xx02xx0x,1,2,07,420,NULL,140,NULL,5040,1410,NULL,NULL,6290,70↵

0x0000029379851266BEE123624FF948E788DB44A1,040070xxxxx0xx,1,7,00,0,290,1600,0,NULL,0,0,NULL,24250,1650↵

0x00000AEC5115495D9643EFC832E99F59E627C4CF,060100xx02xx0x,2,3,06,330,NULL,200,80,5030,2070,290,NULL,7640,200↵

0x00000FFF3AE7FE1A6557C9D9F3A1AB2961F8B1DF,110200xx02xxxx,1,10,07,630,1650,170,NULL,20580,1010,NULL,NULL,22310,1580↵

0x00000DBE3B3AE6C20AF8DB60143D62E862F74FB8,060020xx99x00x,1,5,05,NULL,150,60,NULL,NULL,NULL,NULL,12070,280↵

0x00000ED588BEB78C2B7FD230D02BDB18BA750845,130030xx97x1xx,1,22,08,360,980,52150,2280,8420,6420,1050,380,35050,3350↵

0x0000179E7DDC03ACE2E2A6DFCD7022287E882CE0,040080x1xxx0xx,1,4,00,420,800,740,190,NULL,240,210,NULL,15910,660↵

0x0000108118E657F6823AB8F1BD598B490FC7C534,040040xx9904xx,1,13,06,1160,9100,9830,NULL,NULL,1460,210,NULL,20620,2300↵

0x0000208881412A66DE95CCAB7F7A8722DA88C11A,110200xx02xxxx,1,9,07,310,5110,1290,220,21410,3740,410,NULL,20720,1380↵

0x00001D6C23D491B45661E7A7571409886E8065D5,010060x099000x,2,94,08,870,0,NULL,50,NULL,0,280,169310,180310,53780↵

0x00002788BCA4F93D6247346C2CD886D5FDC4260F,160990xx99x00x,2,22,06,270,2600,NULL,NULL,NULL,1020,2590,NULL,44710,4210↵

0x00002C22FD67B36D9D714F3AC4152161E45CDA7,060350xx99x0xx,1,2,06,NULL,2260,240,50,NULL,1840,NULL,NULL,5500,140↵

0x00002D905B8E6CF275ABB9BE3D0C6CBB286C304D,030250xx991xxx,1,2,05,NULL,NULL,NULL,NULL,NULL,3480,NULL,NULL,5490,NULL↵

件数別・分析単位別患者数と分類数の変化 (平成24年度研究班データ)

粒度	全国		都道府県		二次医療圏	
	症例数	分類数	症例数	分類数	症例数	分類数
全件	6,852,048	2,886	6,852,048	2,886	6,852,048	2,886
30件以上	6,750,369	2,233	4,812,590	1,114	3,004,162	700
	98.5%	77.4%	70.2%	38.6%	43.8%	24.3%
50件以上	6,668,617	1,987	4,187,379	806	2,295,911	467
	97.3%	68.8%	61.1%	27.9%	33.5%	16.2%
100件以上	6,482,306	1,595	3,264,631	481	1,449,703	226
	94.6%	55.3%	47.6%	16.7%	21.2%	7.8%

データに含める最低症例数が少なくなると都道府県レベル、二次医療圏レベルでのデータの代表性が著しく低下する

DPC 020110xx97xxx0(白内障 手術有 片眼)における症例基準別カバー率
(平成24年度研究班データ: 都道府県レベル)

年齢階級	ALL	30症例以上		50症例以上		100症例以上	
0-9歳	212	33	15.6%		0.0%		0.0%
10-19歳	260	35	13.5%		0.0%		0.0%
20-29歳	506	190	37.5%	81	16.0%		0.0%
30-39歳	1,123	727	64.7%	549	48.9%	198	17.6%
40-49歳	2,848	2,436	85.5%	2,083	73.1%	1,652	58.0%
50-59歳	9,235	9,168	99.3%	9,007	97.5%	8,339	90.3%
60-69歳	40,664	40,664	100.0%	40,664	100.0%	40,452	99.5%
70-79歳	85,266	85,266	100.0%	85,266	100.0%	85,266	100.0%
80-89歳	54,718	54,718	100.0%	54,718	100.0%	54,621	99.8%
90-99歳	4,779	4,595	96.1%	4,260	89.1%	3,313	69.3%
100歳以上	22		0.0%		0.0%		0.0%
合計	199,633	197,832		196,628		193,841	
%	100%	99%		98%		97%	

DPC 040080x099x0xx(肺炎・急性気管支炎・急性細気管支炎
15歳以上 手術無 手術処置等2無)における症例基準別カバー率
(平成24年度研究班データ: 都道府県レベル)

年齢階級	ALL	30症例以上	50症例以上	100症例以上
10-19歳	2,273	1,922 84.6%	1,476 64.9%	612 26.9%
20-29歳	3,098	2,796 90.3%	2,334 75.3%	1,325 42.8%
30-39歳	4,244	4,031 95.0%	3,799 89.5%	2,654 62.5%
40-49歳	3,855	3,657 94.9%	3,158 81.9%	2,239 58.1%
50-59歳	5,825	5,758 98.8%	5,428 93.2%	4,324 74.2%
60-69歳	16,718	16,718 100.0%	16,684 99.8%	16,336 97.7%
70-79歳	32,432	32,432 100.0%	32,432 100.0%	32,368 99.8%
80-89歳	44,241	44,241 100.0%	44,241 100.0%	44,241 100.0%
90-99歳	16,022	16,022 100.0%	15,979 99.7%	15,529 96.9%
100歳以上	758	335 44.2%	193 25.5%	0.0%
合計	129,466	127,912	125,724	119,628
%	100%	99%	97%	92%

特定の研究目的に対応したデータセットの作成

【Research question】

- くも膜下出血の予後に施設の症例数は関係しているか？
- Stroke care unit (SCU)はくも膜下出血の予後を改善するか？

上記のResearch questionを検証する目的で
6,852,048症例(1057施設)から16,117症例(383施設)を下記の手順で抽出

- 様式1データセットからくも膜下出血症例を抽出
- データ識別番号を用いて各様式1データに対応するE/Fファイルから次頁の情報の有無及び回数

抽出したデータ

ID、性、年齢階級、入院経路、入院形態、緊急入院、救急車による入院、退院先、転帰、主傷病、入院契機病名、医療資源病名、医療資源病名2、入院時併存症1、入院時併存症2、入院時併存症3、入院時併存症4、入院後続発症1、入院後続発症2、入院後続発症3、入院後続発症4、手術1、手術2、手術3、手術4、手術5、在院日数、DPCコード、手術、集中治療利用、ICU利用、SCU利用、HCU利用、JCS30以上、退院時死亡、出来高換算点数、一日当たりコスト、施設患者数合計、施設手術件数

くも膜下出血の退院時死亡に関連する要因の検討を目的として作成した DPCデータセットの例

ID, 性, 年齢階級, 入院経路, 入院形態, 緊急入院, 救急車による入院, 退院先, 転帰, 主傷病, 入院背景病名, 医療資源病名, 医療資源病名 2, 入院時併存症 1, 入院時併存症 2, 入院時併存症 3, 入院時併存症 4, 入院後続発症 1, 入院後続発症 2, 入院後続発症 3, 入院後続発症 4, 手術 1, 手術 2, 手術 3, 手術 4, 手術 5, 在院日数, DPCコード, 手術, 集中治療利用, ICU利用, SCU利用, HCU利用, JCS30以上, 退院時死亡, 出来高換算点数, 一日当たりコスト, 施設患者数合計, 施設手術件数

```
1, 1, 80, 2, 1, 309, 1, 1, 1, 1600, 1600, 1600, NULL, N40, I693, I10, C259, D62, G919, NULL, NULL, K1771, K145, NULL, NULL, NULL, 47, 010020x001x1xx, 2, 1, 1, 0, 0, 0, 0, "402, 489", "8, 564", 115, 93, .
2, 2, 60, 2, 0, 200, 1, 2, 2, R51, R51, 1609, NULL, E119, F329, NULL, NULL, NULL, NULL, NULL, NULL, NULL, NULL, NULL, 23, 010020x099x00x, 99, 0, 0, 0, 0, 0, 0, "113, 333", "4, 928", 115, 93, .
3, 1, 40, 2, 0, 309, 1, 1, 2, 1602, 1609, 1602, NULL, E780, NULL, NULL, NULL, I10, NULL, NULL, NULL, K1771, K1742, K145, NULL, NULL, 56, 010020x001x1xx, 2, 1, 1, 0, 0, 0, 0, "728, 163", "13, 003", 115, 93, .
4, 2, 50, 2, 0, 309, 0, 4, 4, 1600, 1609, 1600, NULL, S0650, J969, G935, NULL, R402, NULL, NULL, NULL, K1771, K168, K1492, K145, K386, 201, 010020x101x1xx, 2, 1, 1, 0, 0, 1, 0, "1, 123, 428", "5, 589", 115, 93, .
5, 2, 70, 2, 1, 302, 0, 1, 1, 1604, 1609, 1604, NULL, I10, I671, B169, NULL, G919, K112, NULL, NULL, K1781, K1781, K1781, NULL, NULL, 171, 010020x003x1xx, 1, 1, 1, 0, 0, 0, 0, "1, 308, 660", "7, 653", 115, 93, .
6, 1, 60, 2, 1, 200, 0, 4, 2, 1608, 1609, 1608, NULL, I259, NULL, NULL, NULL, J189, J42, K210, NULL, K1761, K145, K664, K396, K386, "2, 782", 010020x197x1xx, 97, 0, 0, 0, 0, 1, 0, "3, 073, 527", "1, 105", 115, 93, .
7, 2, 40, 2, 0, 310, 1, 1, 1, 1602, 1609, 1602, NULL, I10, NULL, NULL, NULL, I639, K210, NULL, NULL, K1781, NULL, NULL, NULL, NULL, 23, 010020x003x0xx, 1, 1, 1, 0, 0, 0, 0, "434, 133", "18, 875", 115, 93, .
8, 2, 70, 2, 0, 310, 0, 1, 2, 1601, 1609, 1601, NULL, I10, E785, M4782, NULL, J459, K210, NULL, NULL, K1771, NULL, NULL, NULL, NULL, 30, 010020x001x0xx, 2, 1, 1, 0, 0, 0, 0, "401, 649", "13, 388", 115, 93, .
9, 2, 40, 2, 1, 309, 1, 1, 1, 1604, 1609, 1604, NULL, G910, I10, E780, G629, I638, K210, J42, N141, K1781, K145, K189, NULL, NULL, 54, 010020x002x1xx, 1, 1, 1, 0, 0, 0, 0, "532, 681", "9, 864", 115, 93, .
10, 2, 40, 2, 1, 302, 0, 1, 2, 1608, 1609, 1608, NULL, I10, E854, NULL, NULL, D649, I209, G918, NULL, K1761, K145, NULL, NULL, NULL, 40, 010020x001x1xx, 2, 1, 1, 0, 0, 0, 0, "606, 750", "15, 169", 115, 93, .
11, 2, 50, 2, 0, 302, 1, 5, 6, 1609, 1609, 1601, NULL, K709, NULL, NULL, NULL, NULL, NULL, NULL, NULL, NULL, NULL, NULL, 3, 010020x199x10x, 99, 1, 1, 0, 0, 1, 1, "36, 907", "12, 302", 115, 93, .
12, 2, 70, 2, 1, 302, 1, 4, 4, 1602, 1619, 1602, NULL, J209, I509, R402, NULL, NULL, NULL, NULL, NULL, NULL, NULL, NULL, NULL, 222, 010020x197x0xx, 97, 1, 1, 0, 0, 1, 0, "699, 208", "3, 150", 115, 93, .
13, 1, 70, 2, 1, 309, 0, 4, 2, 1608, 1608, 1608, NULL, K701, NULL, NULL, NULL, G408, NULL, NULL, NULL, K172, NULL, NULL, NULL, NULL, 40, 010020x097x0xx, 97, 1, 1, 0, 0, 0, 0, "418, 701", "10, 468", 115, 93, .
14, 1, 60, 2, 0, 310, 1, 4, 2, 1601, 1601, 1601, NULL, I10, G910, E785, R13, I638, G935, D649, T818, K1771, K168, K1492, K1801, K145, 57, 010020x001x1xx, 2, 1, 1, 0, 0, 0, 0, "675, 292", "11, 847", 115, 93, .
15, 2, 50, 2, 0, 309, 1, 4, 4, 1601, 1609, 1602, NULL, I671, G935, NULL, NULL, G919, J961, NULL, NULL, K1772, K1801, NULL, NULL, NULL, 733, 010020x101x1xx, 2, 1, 1, 0, 0, 1, 0, "2, 476, 319", "3, 378", 115, 93, .
16, 1, 60, 2, 0, 309, 1, 4, 3, 1601, 1609, 1601, NULL, I10, NULL, NULL, NULL, G039, J152, G919, T857, K1771, K1643, K1801, K150, K1742, 923, 010020x101x1xx, 2, 1, 1, 0, 0, 1, 0, "4, 671, 143", "5, 061", 115, 93, .
17, 2, 50, 2, 1, 309, 0, 1, 1, 1601, 1601, 1601, NULL, I10, E780, I802, M179, I639, R571, G919, NULL, K1771, NULL, NULL, NULL, NULL, 51, 010020x001x1xx, 2, 1, 1, 0, 0, 0, 0, "601, 496", "11, 794", 115, 93, .
18, 1, 60, 2, 0, 309, 1, 4, 2, 1608, 1609, 1608, NULL, I10, NULL, NULL, NULL, I638, G910, F068, NULL, K1761, K6101, K1742, NULL, NULL, 84, 010020x001x1xx, 2, 1, 1, 0, 0, 0, 0, "856, 491", "10, 196", 115, 93, .
19, 2, 20, 2, 1, 100, 0, 1, 1, 1608, G282, 1608, NULL, NULL, NULL, NULL, NULL, NULL, NULL, K1643, K172, NULL, NULL, NULL, 19, 010020x097x0xx, 97, 1, 1, 0, 0, 0, 0, "267, 961", "14, 103", 115, 93, .
20, 1, 50, 2, 0, 309, 1, 4, 2, 1608, 1609, 1608, NULL, I611, I10, NULL, NULL, NULL, NULL, NULL, NULL, K1781, K145, NULL, NULL, NULL, 346, 010020x103x1xx, 1, 1, 1, 0, 0, 1, 0, "1, 628, 211", "4, 706", 115, 93, .
```

くも膜下出血の外科重症例における死亡退院に関連する要因の分析 (平成24年度研究班データ)

1,693 の重症外科症例(JCS >= 30)でICU/SCU/HCUの利用がある1693症例						
	B	SE	Odds ratio	95% of CI		
性	-0.333	0.116	0.717	0.572	-	0.899
年齢階級	0.023	0.004	1.024	1.016	-	1.031
救急車による入院	0.261	0.180	1.298	0.912	-	1.847
ICU	0.128	0.135	1.136	0.871	-	1.481
SCU	-0.655	0.161	0.519	0.379	-	0.713
施設の手術患者数	-0.010	0.004	0.990	0.983	-	0.997
定数	-1.314	0.338	0.269	0.138	-	0.522

性: 男性=0, 女性 =1; 年齢階級: 表3と同様

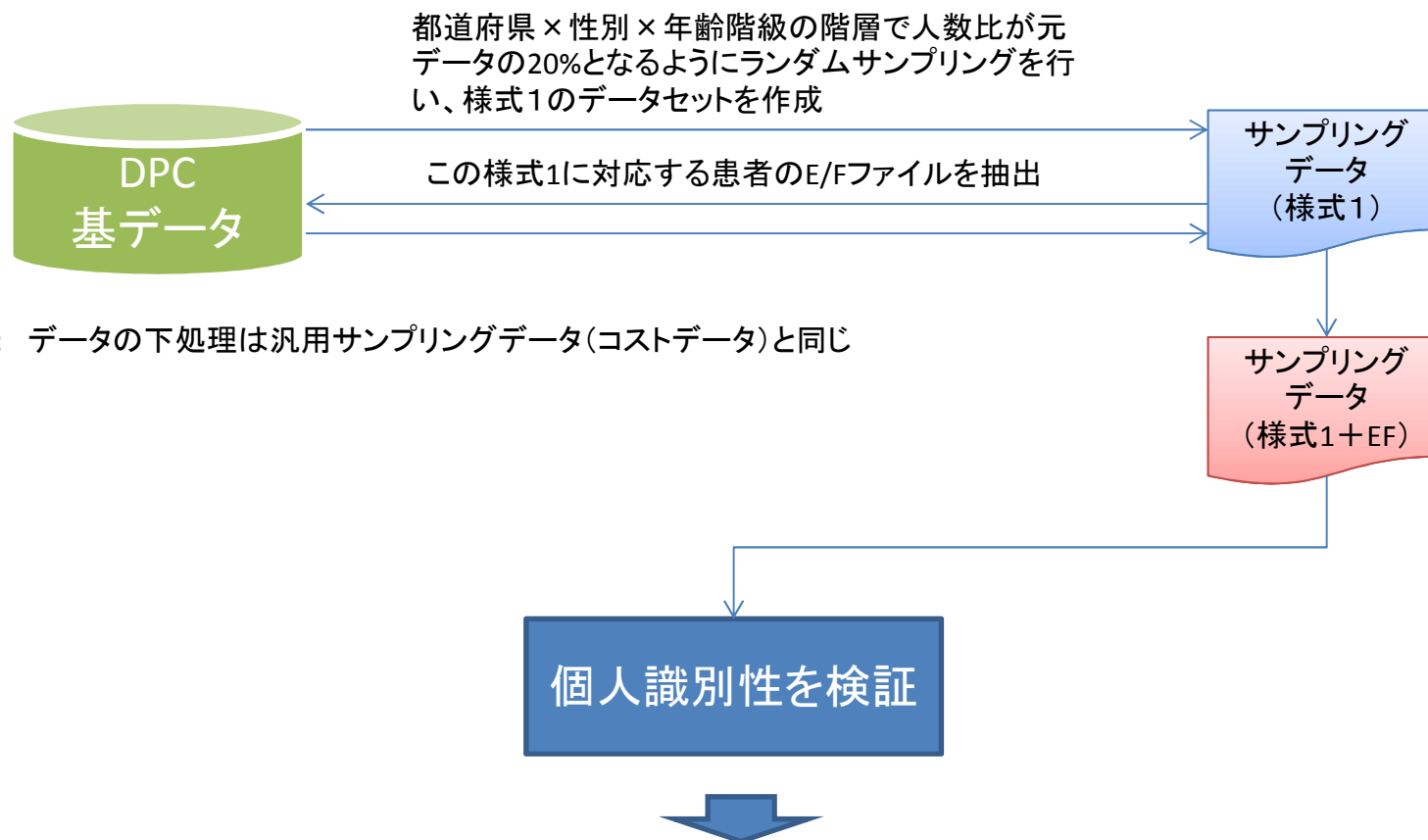
救急車による入院: 0=非該当, 1=該当

ICUへの入院: 0=非該当, 1=該当 (HCU症例が参照)

SCUへの入院: 0=非該当, 1=該当(HCU症例が参照)

施設の手術患者数: くも膜下出血の外科症例

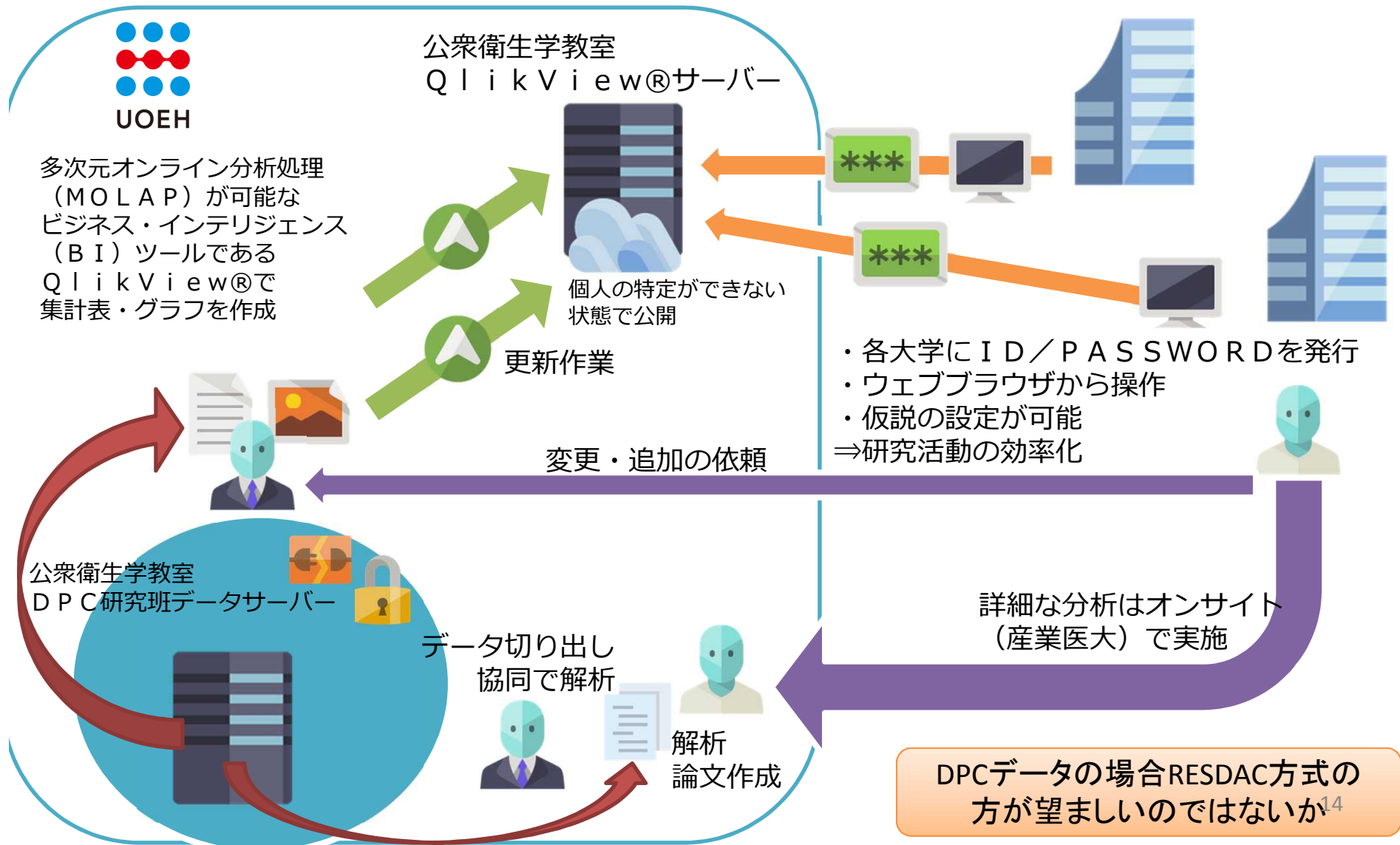
階層的ランダムサンプリングデータ



注：データの下処理は汎用サンプリングデータ(コストデータ)と同じ

- EFフルマッチの場合→個人を特定
- 定義表の項目のみの場合→個人を特定できない

産業医科大学公衆衛生学教室・ 中四国脳神経外科学講座間の連携



結論

- DPCデータの利活用のための サンプリングデータセットの仕様について検討した。
- 利用目的に応じたデータ切り出し処理を行うことで個人情報保護に配慮したデータセットを作成することが可能であることが示された。
- しかしながら、このようなデータから得られる知見は厚生労働省がすでに公開している集計データを用いることで把握できるものが少なくなく、またDPC情報が本来持っている長所を生かすようなものでもない
- したがって、DPCデータの利活用を進めるためにはアメリカのRESDACを倣ってオンサイトでの分析を支援する組織の創設が望ましいのではないか。