

様に行なった。

- ・ 分析 ICU-5: 施設限定+変数①②
- ・ 分析 ICU-6: 施設限定+変数①②③
- ・ 分析 ICU-7: 施設限定+変数①② ④
- ・ 分析 ICU-8: 施設限定+変数①②③④

(2) ER (救急医療部門) データ

施設調査データに関しては、設置・運営形態や ICU の併設状況等の分析を、10 月実績調査データに関しては、1 次・2 次・3 次患者の利用状況 (24 時間・退院死亡状況、入院時心停止状況) 等の分析を、患者調査データに関しては、性別、年齢、主たる傷病名、受診後 24 時間以内の APACHE スコア、AIS スコア等の分析を実施した。

次に、施設の特性 (機能) を分類するために、階層的クラスター分析を用いて、3 分類を想定し、3 つになるように集約 (グループ) 化をおこなった。その後分類されたクラスター間で施設特性について比較をおこなった。分類に用いた変数は、救急センターに該当しているか非該当かで分け、ER の設置形態、ICU 併設、ER の運営形態であり、グループ内平均関連法を用いて、階層的クラスター分析を行なった。

C. 結果

(1) ICU (集中治療室) データ

① ICU の施設に関する調査 (詳細は II 章参照)

今回の調査における解析対象施設は 129 施設であった。ベッド数において、欠損値が 4 施設でみられ、残りの 125 施設では、2 床~67 床の範囲で、6 床が 30 施設 (24.0%) と最も多く、次いで 10 床 (22 施設、17.6%) であった。10 床未満の病院は、全体の約 55% であった。

運営形態では、治療方針では 10 施設が、カンファレンスでは 19 施設が欠損で、両変数が揃っていたのは 116 施設であった。そのうち、治療方針の 7 項目全て専任/

専従医が決定する施設が 58 施設あり、カンファレンスの項目数としては、5 項目全て当てはまる施設が 33 施設と最も多く、ついで 4 項目 (28 施設)、3 項目 (20 施設) であった。

● 分析 ICU-1

(医師数・運営形態)

各クラスター (CL) の施設数は、CL 1 が 32 施設、CL 2 が 19 施設、CL 3 が 4 施設であった。また、施設の特性は以下の通りであった。

ICU の運営形態では、日勤中常に専任/専従医がいる施設数は、CL 1 は 29 施設、CL 2 は 15 施設、CL 3 は 3 施設であった。治療方針に関して、全ての項目について当てはまるのは CL 1 の 32 施設で、CL 2 は 0 施設、CL 3 は 4 施設であった。特に治療方針の個々の項目で見ても差がないようであった。また、1 床あたりの常勤医数は、CL 1 では平均値 0.35 人で、最小値 0.07、最大値 1.00 であった。CL 2 は平均値 0.26 人で 0.06~1.00 で、CL 3 は平均値 2.04 人で 1.90~2.17 であった。

● 分析 ICU-2

(医師数・運営形態・機器数)

各クラスターの施設数は、CL 1 が 42 施設、CL 2 が 3 施設、CL 3 が 2 施設であった。また、施設の特性は以下の通りであった。

ICU の運営形態では、日勤中常に専任/専従医がいる施設数は、CL 1 は 37 施設、CL 2、CL 3 は全施設であった。治療方針に関して、全ての項目について当てはまるのは CL 3 の全 3 施設で、CL 1 の 35 施設、CL 2 の 17 施設は 7 項目全てを専任/専従医が決定するわけではなかった。また、1 床あたりの常勤医数は、CL 1 では平均値 0.28 人で、最小値 0.06、最大値 1.00 であった。CL 2 は平均値 2.00 人で 1.90~2.17 で、CL 3 は平均値 0.21 人で 0.17~0.25 であった。また、1 床あたりの機器数は、CL 1 の人工呼吸器は 1.23 (0.50~1.23) 台、血液浄化装置は 0.40 (0.08~1.00) 台であった。CL 2 の人工呼吸器は

0.83 (0.50~1.00) 台、血液浄化装置は 0.28 (0.07~0.60) 台であった。CL3 の人工呼吸器は 2.67 (2.33~3.00) 台、血液浄化装置は 5.46 (5.25~5.67) 台であった。

● 分析 ICU-3

(医師数・運営形態・看護師数)

各クラスターの施設数は、CL1 が 17 施設、CL2 が 2 施設、CL3 が 32 施設であった。また、施設の特性は以下の通りであった。

ICU の運営形態では、日勤中常に専任/専従医がいる施設数は、CL1 は 13 施設、CL2 は全 2 施設、CL3 は 29 施設であった。治療方針に関して、全ての項目について当てはまるのは CL2 の全 2 施設、CL3 の全 32 施設で、CL1 の全 17 施設は 7 項目全てを専任/専従医が決定するわけではなかった。また、1 床あたりの常勤医数は、CL1 では平均値 0.27 人で、最小値 0.06、最大値 1.00 であった。CL2 は平均値 2.05 人で 1.93~2.17 で、CL3 は平均値 0.35 人で 0.07~1.00 であった。1 勤務帯の 1 床あたり看護師数は、CL1 が 0.65 (0.45~0.77) 人、CL2 が 0.64 (0.56~0.72) 人、CL3 が 0.64 (0.44~0.99) 人であった。

● 分析 ICU-4

(医師数・運営形態・機器数・看護師数)

各クラスターの施設数は、CL1 が 16 施設、CL2 が 27 施設、CL3 が 1 施設であった。また、施設の特性は以下の通りであった。

ICU の運営形態では、日勤中常に専任/専従医がいる施設数は、CL1 は 12 施設、CL2 は 27 施設、CL3 は全 1 施設であった。治療方針に関して、全ての項目について当てはまるのは CL2 の全 27 施設、CL3 の全 1 施設で、CL1 の全 16 施設は 7 項目全てを専任/専従医が決定するわけではなかった。また、1 床あたりの常勤医数は、CL1 では平均値 0.25 人で、最小値 0.06、最大値 1.00 であった。CL2 は平均値 0.44 人で 0.07~2.17 で、CL3 は平均値 0.25

人 (1 施設) であった。また、1 床あたりの機器数は、CL1 の人工呼吸器は 1.00 (0.50~2.00) 台、血液浄化装置は 0.39 (0.10~1.00) 台であった。CL2 の人工呼吸器は 1.31 (0.50~2.75) 台、血液浄化装置は 0.39 (0.07~1.00) 台であった。CL3 の人工呼吸器は 3.00 (1 施設) 台、血液浄化装置は 5.25 (1 施設) 台であった。1 勤務帯の 1 床あたり看護師数は、CL1 が 0.66 (0.45~0.77) 人、CL2 が 0.63 (0.44~0.99) 人、CL3 が 0.67 (1 施設) 人であった。

● 分析 ICU-5：施設限定

(医師数・運営形態)

各クラスターの施設数は、CL1 が 31 施設、CL2 が 2 施設、CL3 が 19 施設であった。また、施設の特性は以下の通りであった。

ICU の運営形態では、日勤中常に専任/専従医がいる施設数は、CL1 は 28 施設、CL2 は全 2 施設、CL3 は 15 施設であった。治療方針に関して、全ての項目について当てはまるのは CL1 の全 31 施設、CL2 の全 2 施設で、CL3 の全 19 施設は 7 項目全てを専任/専従医が決定するわけではなかった。また、1 床あたりの常勤医数は、CL1 では平均値 0.35 人で、最小値 0.07、最大値 1.00 であった。CL2 は平均値 1.91 人で 1.91~1.93 で、CL3 は平均値 0.26 人で 0.06~1.00 であった。

● 分析 ICU-6：施設限定

(医師数・運営形態・機器数)

各クラスターの施設数は、CL1 が 41 施設、CL2 が 3 施設、CL3 が 2 施設であった。また、施設の特性は以下の通りであった。

ICU の運営形態では、日勤中常に専任/専従医がいる施設数は、CL1 は 36 施設、CL2 は全 3 施設、CL3 は全 2 施設であった。治療方針に関して、全ての項目について当てはまるのは CL2 の全 3 施設で、CL1 の 17 施設、CL3 の 1 施設は 7 項目全てを専任/専従医が決定するわけではなかった。また、1 床あたりの常勤医数は、CL1

では平均値 0.29 人で、最小値 0.06、最大値 1.00 であった。CL2 は平均値 2.00 人で 1.90~2.17 で、CL3 は平均値 0.21 人で 0.17~0.25 であった。また、1床あたりの機器数は、CL の人工呼吸器は 1.23 (0.50~2.75) 台、血液浄化装置は 0.40 (0.08~1.00) 台であった。CL2 の人工呼吸器は 0.83 (0.50~1.00) 台、血液浄化装置は 0.28 (0.07~0.60) 台であった。CL3 の人工呼吸器は 2.67 (2.33~3.00) 台、血液浄化装置は 5.46 (5.25~5.67) 台であった。

● 分析 ICU-7: 施設限定

(医師数・運営形態・看護師数)

各クラスターの施設数は、CL1 が 17 施設、CL2 が 2 施設、CL3 が 31 施設であった。また、施設の特性は以下の通りであった。

ICU の運営形態では、日勤中常に専任/専従医がいる施設数は、CL1 は 13 施設、CL2 は全 2 施設、CL3 は 13 施設であった。治療方針に関して、全ての項目について当てはまるのは CL2 の全 2 施設、CL3 の全 31 施設で、CL1 の全 17 施設は 7 項目全てを専任/専従医が決定するわけではなかった。また、1床あたりの常勤医数は、CL1 では平均値 0.27 人で、最小値 0.06、最大値 1.00 であった。CL2 は平均値 2.05 人で 1.93~2.17 で、CL3 は平均値 0.35 人で 0.07~1.00 であった。1勤務帯の 1床あたり看護師数は、CL1 が 0.65 (0.45~0.77) 人、CL2 が 0.64 (0.56~0.72) 人、CL3 が 0.64 (0.44~0.99) 人であった。

● 分析 ICU-8: 施設限定

(医師数・運営形態・機器数・看護師数)

各クラスターの施設数は、CL1 が 16 施設、CL2 が 26 施設、CL3 が 1 施設であった。また、施設の特性は以下の通りであった。

ICU の運営形態では、日勤中常に専任/専従医がいる施設数は、CL1 は 12 施設、CL2 は 25 施設、CL3 は全 1 施設であった。治療方針に関して、全ての項目について

当てはまるのは CL2 の全 26 施設、CL3 の全 1 施設で、CL1 の全 16 施設は 7 項目全てを専任/専従医が決定するわけではなかった。また、1床あたりの常勤医数は、CL1 では平均値 0.25 人で、最小値 0.06、最大値 1.00 であった。CL2 は平均値 0.45 人で 0.07~2.17 で、CL3 は平均値 0.25 人(全 1 施設)であった。また、1床あたりの機器数は、CL1 の人工呼吸器は 1.00 (0.50~2.00) 台、血液浄化装置は 0.39 (0.10~1.00) 台であった。CL2 の人工呼吸器は 1.33 (0.50~2.75) 台、血液浄化装置は 0.39 (0.07~1.00) 台であった。CL3 の人工呼吸器は 3.00 台、血液浄化装置は 5.25 台(全 1 施設)であった。1勤務帯の 1床あたり看護師数は、CL1 が 0.66 (0.45~0.77) 人、CL2 が 0.63 (0.44~0.99) 人、CL3 が 0.67 人(全 1 施設)であった。

② ICU の患者に関する調査

今回の調査における解析対象件数は 8,026 件であった。施設別では、最大 304 件を扱っている施設があり、最も少ない施設では 10 件であった。10~50 件を扱った施設で全体の約 30%を、51~100 件を扱った施設で約 36%を、101~200 件を扱った施設では約 14%を、201 件以上の施設で約 16%を占めていた。

対象患者の年齢は、8 歳~103 歳で平均 66.4 歳であった。4 分位は 59 歳、69 歳、77 歳であった。

また、年齢カテゴリーでの分布は、表 1 のようになっていた。これで見ると、50 代以降の患者で全体の約 70%を占めていることが分かる。

入室経路では、欠損値 183 件が見られた。回答の有った 7,843 件のうち、手術室の 3,231 件 (41.2%) が最も多く、次いで、救急外来 2,556 件 (32.6%)、病棟 1,320 件 (16.8%)、検査室など院内の病棟以外 396 件 (5.0%)、転院搬送 340 件 (4.3%) となっていた。

APACHE II スコアにおいては、180 件で

欠損値となった。残りの 7846 件における平均値は 19.1、最小値 0、最大値 59 であった。スコアを過去文献[1]において院内死亡率に差があるとされているカテゴリーに、分割した時の分布は、表 2 の通りであった。最も割合が多かったのが、15~19 のスコアの患者で、23.9%であった。0~24 のスコアの患者で、全体の 73% 強を占めていた。150 件以上扱った病院における APACHE II スコアの状況は、図 1 の通りである。施設内においてばらつきがみられ、ほぼ 30 点以下に収まっていた。

退室時に人工呼吸器をつけていた割合は 2.8% であり、退室先は、表 3 の通りであった。退室先は院内が 93.2% であった。

また、特定機能病院とその他の病院において施設ごとの APACHE II スコアのカテゴリーの割合の分布から、施設により ICU を利用する患者の重症度の違いが見られた。

(2) ER (救急医療部門) データ

① ER の施設に関する調査

今回の調査における解析対象施設は 148 施設であった。そのうち、救命救急センターに該当するのは欠損値の 4 施設を除いた 144 施設中 48 施設 (33.3%) であった。ER の設置形態では、欠損値 5 施設を除いた 143 施設において、外来のみの ER であるのが 74 施設 (51.7%) で、病棟と外来を備えている ER であるのが、69 施設 (48.3%) であった。また、ICU が、ER と独立である施設が 101 施設 (74.3%)、併設であるのが 35 施設 (25.7%) であった (欠損値は 12 施設で見られた)。

医師による運営形態の違いでは、専従医と院内医師の協力型をとっている施設が最も多く 72 施設 (50.0%) で、次いで院内医師協力型が 69 施設 (46.6%)、専従医のみ 3 施設 (2.0%) であった (4 施設で欠損値がみられた)。

クラスター分析の結果、救命救急セン

ターに該当する各クラスター (CL) の施設数は、CL1 は 21 施設、CL2 は 2 施設、CL3 は 25 施設であった。また、施設の特性は以下の通りであった。

専従医がいる施設は、CL1 は 19 施設 (90.5%)、CL2 が 0 施設、CL3 が 24 施設 (96.0%) であった。また、ER に病棟を設置している施設は、CL1 は 20 施設 (95.2%)、CL2 は 0 施設、CL3 は 21 施設 (84.0%) であった。また、CL1 は全 21 施設が ICU を併設しており、CL2、CL3 は ICU が独立している施設であった。

救命救急センターに該当しない各クラスター (CL) の施設数は、CL1 は 62 施設、CL2 は 17 施設、CL3 は 8 施設であった。また、施設の特性は以下の通りであった。

専従医がいる施設は、CL1 は 16 施設 (25.8%)、CL2 が 6 施設 (35.3%)、CL3 が 8 施設 (100%) であった。また、ER に病棟を設置している施設は、CL1 は 2 施設 (3.2%)、CL2 は全 17 施設、CL3 は全 8 施設であった。また、ICU を併設している施設は、CL1 が 6 施設 (9.7%) で、CL2 が 0 施設、CL3 が全 8 施設であった。

② ER 10 月実績調査

今回の調査の解析対象は、148 件の施設における 117,556 件であった。(うち、13 施設は 0 症例)。また、1 次患者が 78.5%、2 次患者が 15.8%、3 次患者が 5.8% であった。症例数が 100 件以上の施設においては、1 次患者が 29% という施設から 97% という施設までばらつきがみられ、3 次患者では 100% という施設が見られた。

1 次患者が 100 件以上の施設では、1 次患者の 24 時間以内死亡割合が 0~5%、2 次患者が 100 件以上の施設では 2 次患者の 24 時間以内死亡割合は 0~5%、3 次患者が 100 件以上の施設では 3 次患者の 24 時間以内死亡割合は 0~16% であった。同様に各患者層が 100 件以上の施設での退院時死亡割合は、1 次患者ではほとんどの施設 (94 施設) において 0% で最

高でも3%であった。また、2次患者で0~21%、3次患者で0~18%であった。

③ ERの患者に関する調査

今回の調査における解析対象症例は、8,410件であった。施設別では、最大488症例を扱っており、最も少ない施設では1件(3施設)であった。1~50件を扱った施設で全体の約20%を占め、51~100件を扱った施設で約30%を占め、101~200件を扱った施設で約27%を占め、201件以上を扱った施設で約23%を占めていた。

対象患者の年齢は、0歳~104歳で、平均62.5歳であった。4分位は51歳、68歳、79歳であった。また、年齢カテゴリーでの分布は、表4のようになった。50歳以上の症例で全体の約75%を占めていた。

入院の原因となった主病名の分布は、表5の通りであった。最も症例が多かった脳内出血は510件あり、循環器系の疾患が上位10疾病中8疾病含まれていた。

APACHE IIスコアにおいては、欠損値が366件見られ、残りの8,044件における平均値は17.3、最小値0、最大値57であった。スコアのカテゴリーでの分布は、表6の通りであった。APACHE IIスコアが0~4の患者の割合が最も少なく306件であり、0~24のスコアの患者で全体の80%を占めていた。症例数が150例以上の施設におけるAPACHE IIスコアの状況は、図2の通りである。施設内においてばらつきが見られ、施設ごとでもばらつきが見られた。

AISスコアの状況は各部位で、表7~表12のようになっていた。全ての部位においてAISスコア0が最も多かった。

また、特定機能病院とその他の病院において施設ごとのAPACHE IIスコアのカテゴリーの割合の分布から、ERに来院する患者の重症度に違いがあることが伺えた。特定機能病院では最も軽いカテゴリー

の患者は見られなかった。これは、特定機能病院のERではより重症の患者を受け入れていることを示唆していると考えられる。

D. 考察

ICU、ERデータともに、施設間で利用する患者の状態像に大きなばらつきがみられており、施設によって果たしている機能の違いがあることが示唆される。APACHE IIスコアの分布状況を見ると、ICUの利用状況では、特定病院とそれ以外の病院で重症度の違いはあまりみられなかったが、ERでは特定病院には、最も軽いカテゴリーの患者が見られず、特定病院が重症の患者を受け入れていることが示唆された。一方、施設の運営形態や常勤医数、1床あたりの医療機器数、看護師数などの構造的な要因を用いて施設を分類することも可能であった。

E. 結論

ICU、ERデータともに、今回のような大規模なデータでの分析はあまりなく、分析結果は急性期病院として具備すべき機能を評価・検討する際の資料として大きな役割を担えると考えられる。また、欠損値が多く、データ入力時に入力忘れ、ご入力などを防ぐための入力画面のインターフェース上の対策などを含めて、データ収集時点での対策も同時に考慮する必要があると考える。

F. 参考文献

1. Knaus WA et al. APACHE II: a severity of disease classification system. Crit Care Med 1985; 13:818-829.

表1. 年齢カテゴリーごとの分布 (ICU)

	度数	有効パーセント	累積パーセント
0～9歳	1	0.00	0.00
10代	67	0.8	0.8
20代	186	2.3	3.2
30代	353	4.4	7.6
40代	525	6.5	14.1
50代	1227	15.3	29.4
60代	1992	24.8	54.2
70代	2418	30.1	84.3
80代	1085	13.5	97.9
90代	166	2.1	99.9
100歳以上	6	.1	100.0
合計	8026		

表2. APACHE IIスコアの分布 (ICU)

	度数	有効パーセント	累積パーセント
0～4	94	0.1	0.1
5～9	806	10.2	10.3
10～14	1830	23.3	33.6
15～19	1877	23.9	57.5
20～24	1237	15.8	73.3
25～29	924	11.8	85.1
30～34	587	7.5	92.6
35以上	491	6.4	100.0
欠損値	180		
合計	8026		

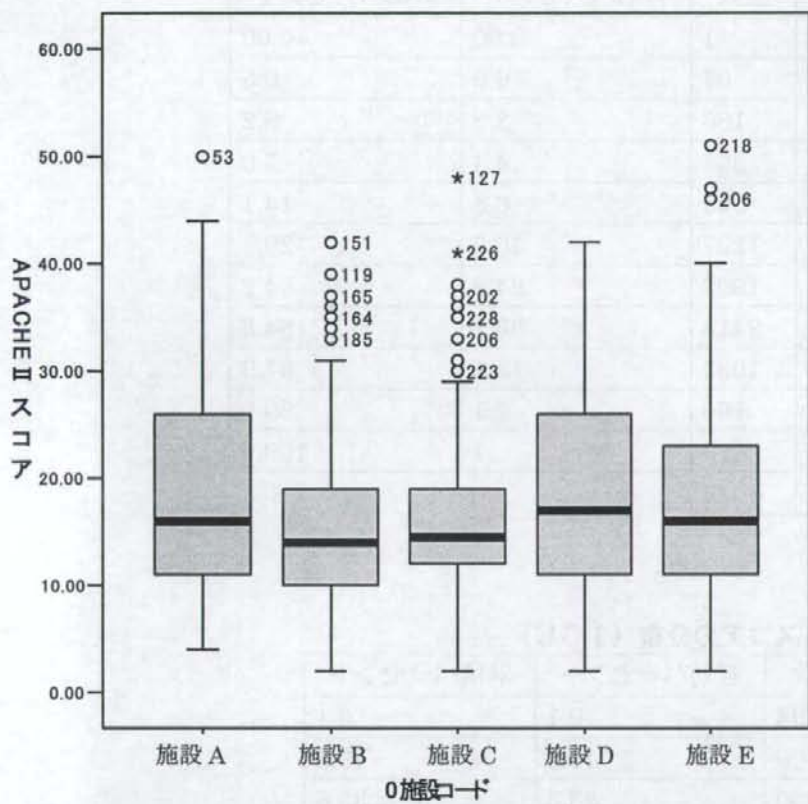


図 1. 症例数 150 例以上の施設における APACHE II スコアの状況 (ICU)

表 3. 退室時転帰の状況 (ICU)

	度数	有効パーセント	累積パーセント
一般病棟	5823	82.8	82.8
HCU (Step Down)	733	10.4	93.2
転院	94	1.3	94.5
死亡	385	5.5	100
欠損値	991		
合計	8026		

表4. 年齢カテゴリーごとの分布 (ER)

	度数	有効パーセント	累積パーセント
0～3歳	115	1.4	1.4
4～10歳	122	1.5	2.8
10代	247	2.9	5.8
20代	409	4.9	10.6
30代	561	6.7	17.3
40代	615	7.3	24.6
50代	1128	13.4	38.0
60代	1481	17.6	55.6
70代	1980	23.5	79.2
80代	1415	16.8	96.0
90代	325	3.9	99.9
100歳以上	11	0.1	100.0
欠損値	1		
合計	8410		

表5. 入院の原因となった主たる病名 ICD-10 (上位抜粋) (ER)

ICD-10	病名	度数	有効パーセント	累積パーセント
I61	脳内出血	510	6.1	6.1
I63	脳梗塞	475	5.6	11.7
I21	急性心筋梗塞	441	5.2	16.9
S06	頭蓋内損傷	357	4.2	21.1
I50	心不全	339	4.0	25.1
I46	心停止	268	3.2	28.3
I20	狭心症	232	2.8	31.1
I60	くも膜下出血	214	2.5	33.6
T50	利尿薬, その他及び詳細不明の薬物, 薬剤及び生物学的製剤による中毒	209	2.5	36.1
I71	大動脈瘤及び解離	155	1.8	37.9

表 6. APACHE II スコアの分布 (ER)

	度数	有効パーセント	累積パーセント
0~4	306	3.8	3.8
5~9	1310	16.3	20.1
10~14	2080	25.9	45.9
15~19	1660	20.6	66.6
20~24	1015	12.6	79.2
25~29	699	8.7	87.9
30~34	457	5.7	93.6
35 以上	517	6.4	100.0
欠損値	366		
合計	8410		

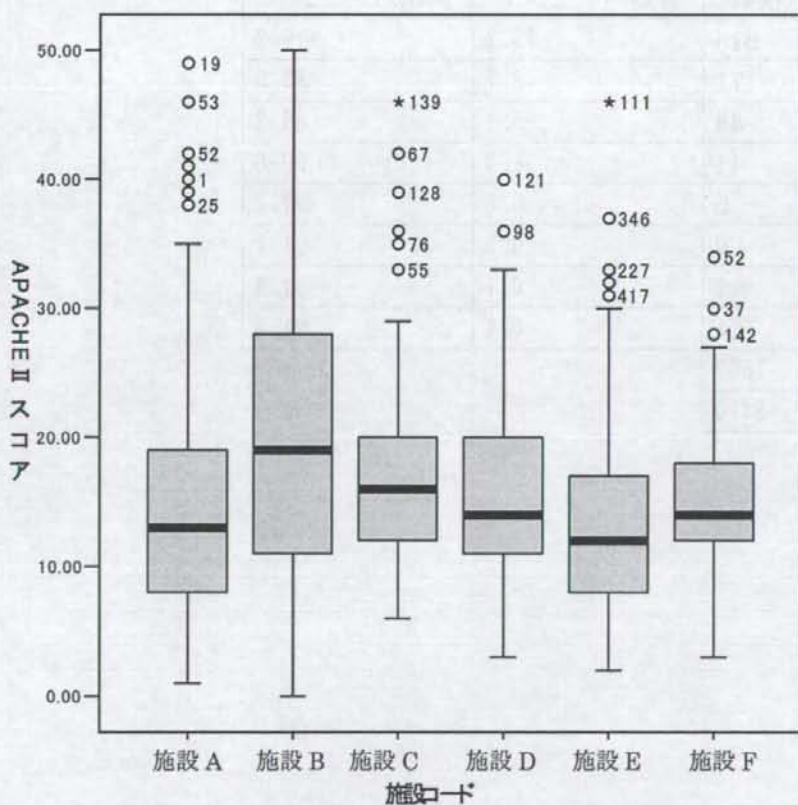


図 2. 症例数 150 例以上の施設における APACHE II スコア (ER)

表 7. 頭頸部最大 AIS の分布 (ER)

頭頸部最大 AIS	度数	有効パーセント	累積パーセント
0	436	53.3	53.3
1	66	8.1	61.4
2	63	7.7	69.1
3	90	11.0	80.1
4	96	11.7	91.8
5	42	5.1	96.9
6	18	2.2	99.1
9	7	0.9	100.0
欠損値	7592		
合計	8410		

表 8. 顔面最大 AIS の分布 (ER)

顔面最大 AIS	度数	有効パーセント	累積パーセント
0	615	75.2	75.2
1	71	8.7	83.9
2	44	5.4	89.2
3	11	1.3	90.6
4	5	0.6	91.2
5	0	0.0	91.2
6	1	0.1	91.3
9	1	0.1	91.4
欠損値	7662		
合計	8410		

表 9. 胸部最大 AIS の分布 (ER)

胸部最大 AIS	度数	有効パーセント	累積パーセント
0	550	67.2	67.2
1	40	4.9	72.1
2	27	3.3	75.4
3	82	10.0	85.5
4	42	5.1	90.6
5	11	1.3	91.9
6	7	0.9	92.8
9	3	0.4	93.2
欠損値	7648		
合計	8410		

表 10. 腹部最大 AIS の分布 (ER)

腹部最大 AIS	度数	有効パーセント	累積パーセント
0	598	73.1	73.1
1	25	3.1	76.2
2	54	6.6	82.8
3	48	5.9	88.6
4	14	1.7	90.3
5	4	0.5	90.8
6	2	0.2	91.1
9	2	0.2	91.3
欠損値	7663		
合計	8410		

表 11. 四肢最大 AIS の分布 (ER)

四肢最大 AIS	度数	有効パーセント	累積パーセント
0	381	46.6	46.6
1	66	8.1	54.6
2	194	23.7	78.4
3	160	19.6	97.9
4	26	3.2	101.1
5	8	1.0	102.1
6	1	0.1	102.2
9	0	0.0	102.2
欠損値	7574		
合計	8410		

表 12. 体表最大 AIS の分布 (ER)

体表最大 AIS	度数	有効パーセント	累積パーセント
0	507	62.0	62.0
1	219	26.8	88.8
2	31	3.8	92.5
3	11	1.3	93.9
4	2	0.2	94.1
5	3	0.4	94.5
6	1	0.1	94.6
9	4	0.5	95.1
欠損値	7632		
合計	8410		

DPC 分析用データセットの作成・開発について

堀口 裕正 東京大学大学院医学系研究科 医療経営政策学講座(研究協力者)
橋本 英樹 東京大学大学院医学系研究科 臨床疫学・経済学分野(分担研究者)
康永 秀生 東京大学大学院医学系研究科 医療経営政策学講座(研究協力者)
石川 ベンジャミン 光一 国立がんセンター(分担研究者)
藤森 研司 北海道大学 医療マネジメント寄附研究部門(分担研究者)

研究要旨

本研究班において、収集した DPC データは、データセットの量が大きく、一般的な研究者が保有する分析環境（コンピュータの能力やデータを保管するストレージの量等）では処理が行えない状況となっている。また、その膨大なデータのうち、矛盾するレコードや、研究で使用するには留意が必要なデータも混じっている。

そこで、いくつかのデータ処理を行うことによって、データを分析可能なものに絞り込み、さらに分析に必要な様々な処理を加えてデータセットを作成し、さまざまな研究が実施しやすい環境を構築することを行った。

昨年度平成 16 年と 18 年のデータについては分析用データセットの完成をみた。本年度は平成 17 年・19 年・20 年のデータについて分析用のデータセットの作成を行い、分析に供することができた。その際、キー情報の重複や必要なデータの欠損のある症例などは、分析に影響度が高いので、それらの症例情報については確実に除去を行った。

このデータセットの完成で、DPC データの精度を向上させ、より高度な分析を実施することが可能となると考えられる。

A. 目的

本研究班において、収集した DPC データは、まず基本調査として入院情報の D/E/F ファイル・様式 1・3・4 ファイル（以下それぞれ FF1/FF3/FF4 ファイルと表記）と多岐にわたっている。また各病院 4~10 月のデータと収集時期の範囲も広く、多数の医療機関が参加している。さらに、付随する調査として ICU/ER・リハビリ・CP・外来機能といった各種調査が付随し、それぞれが、基本調査の集計データを分析にあたって必要としている。

平成 21 年 3 月現在、本研究班では平成 15 年度調査から平成 20 年度調査までの 6 年度の調査データを医療機関との契約の元で保有し、分析に活用しているが、平成 20 年調査データの分量はすべてテキストのデータで、800GB を超えた量となっており、一般的な研究者が保有する分析環境（コンピュータの能力やデータを保管するストレージの量等）では処理が行えない状況となっている。

そこで、本研究では、昨年度より本研究班の保有する DPC 調査データから、

- 1、利用できるデータを選択し、
- 2、その 1 つ 1 つのデータが研究に利用する際に留意すべきポイントがあるかをチェックして情報を付加し
- 3、分析に必要な情報を容易にとりだせるよう用途別に分割し、

さまざまな統計処理を行うための数値を計算しやすいよう整理して共通分析用データセットとして作成し、本研究班においてより精度の高い、高度な分析を行うための環境を整備することを試みている。本年度は平成 20 年データを中心として、そのデータベース構造が変化した部分に対応し、他の年度との経年比較が行えるような処理を行うことを目的とした。

B.方法

以下の処理を、研究班保有の DPC データに対して行い、分析用データセットを作成することとした。

1、データの取り込み

参加医療機関から提供された DPC データ (FF1/3/4/D/E/F ファイル) を DB に取り込む

2、エラーチェック

提出されたデータのエラーチェックを行い、エラーデータをデータセット内から除外する。また研究に使用する際留意する項目に対してチェックしフラグを付与する。

どのような条件についてエラー・留意とするかについては本年の研究として検討を行う。

3、DPC コード情報の一体的保有

平成 20 年度「DPC 松田研究班版 DPC コーダーの開発について」で作成した DPC コーディングツールを利用した DPC コードを分析用データセット内に取り込み、他のデータと一体的に保有する

以上の処理を行い、分析用データセットを作成する。

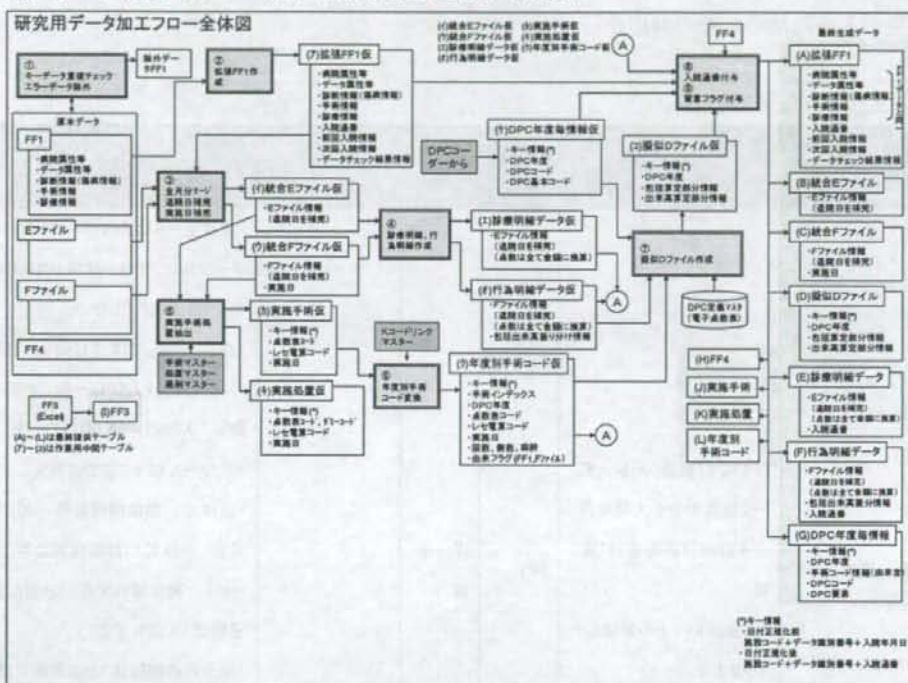
分析用データセットには次のファイルが含まれる

- | | |
|---------------|--------------------------------|
| (A) 拡張 FF1 | 様式 1 (FF1) 情報に留意フラグ等の情報を付加したもの |
| (B) 統合 E ファイル | E ファイルに退院日及び期間内入院回数を付加したもの |
| (C) 統合 F ファイル | F ファイルに退院日・実施日等を付加したもの |
| (D) 疑似 D ファイル | D ファイル作成ルールに基づいて擬似的に D ファイルを生成 |
| (E) 診療明細データ | 統合 E ファイルの点数部分を金額に置き換えたもの |
| (F) 行為明細データ | 統合 F ファイルの点数部分を金額に置き換えたもの |
| (G) DPC 年度毎情報 | 運用時期別に生成された DPC コード情報 |
| (H) FF4 | 様式 4 |
| (J) 実施手術 | F ファイルから手術関係のレコードのみを抽出したもの |

(K) 実施処置 F ファイルから処置・薬剤関係のレコードのみを抽出したもの

(L) 年度別手術コード 手術コードを診療報酬の運用年度別に変換したもの

尚、データ処理のフローは下に示す図のとおりである。



開発及び運用環境

研究班が収集している DPC データは数百病院分にもおよび、膨大なサイズとなるため、以下の環境での開発及び運用を行っている。

OS: Windows Server 2003 R2

DB: SQL 2005 Server Enterprise Edition

開発環境: Visual Studio 2005

設置場所: 東京大学医学系研究科医療経営政策学講座サーバー室内

なお、研究データの守秘性の観点から、データは入室に指紋認証が必要で、窓のない部屋内にサーバー・クライアントともに設置し、大学のネットワークとも物理的に隔絶された管理区域内の研究専用 LAN 内で作業を行っている。

C. 結果

本年度の研究期間においては平成 19/20 年度調査及び平成 15・17 年調査の松田班保有のデータに対して B. 方法で示した処理を行った。

以下特記すべき事項に関して記載する。

1、エラーチェックについて

本研究データセット作成については、以下のルールの下で、データエラーチェックおよび留意フラグ付与を行った。特に平成20年度のデータについては調査における様式1の様式変更が入ったため、その対応を行った。

表1 エラーチェック仕様

番号	チェック名	チェック内容	対象データ	種別	フラグ内容	プロセス
ERR010	FF1 キー重複	FF1にて「施設コード+データ識別番号+入院年月日+診療情報番号」が重複 (過去のFF1との重複もチェックする。)	FF1	キー重複		<ul style="list-style-type: none"> ・重複データはFF1データを除外データFF1テーブルに、エラー情報と共に格納し、以降の処理には使用しない。 ・パターンとしては、「日帰り入院+同一日再入院」、「入院日一致、退院日不一致」、「入院日一致、退院日一致」があるが、ツールはそこまでは判定しない。 ・親様式1(診療情報番号=0)がエラーの場合、子様式1(診療情報番号≥1)のレコードも一緒に除外する。(過去に仮確定した子様式1も除外する。) ・過去月の親様式1との重複の場合は、今回の親様式1および子様式1のみ除外とし、過去月の親様式1および子様式1はそのままとする。 ・子様式1の重複は該当の子様式1のみ除外とし、親様式1および他の子様式1には影響を及ぼさない。
ERR020	FF4 キー重複	FF4にて「施設コード+データ識別番号+入院年月日」が重複	FF4	キー重複		<ul style="list-style-type: none"> ・重複データはFF1データを除外データFF1テーブルに、エラー情報と共に格納し、以降の処理には使用しない。 ・親様式1と子様式1の両方を除外する。(過去に仮確定した子様式1も除外する。)
ERR030	Eファイルキー重複	Eファイルの「施設コード+データ識別番号+入院年月日+データ区分+順序番号+該当月」が重複	Eファイル	キー重複		<ul style="list-style-type: none"> ・重複データはFF1データを除外データFF1テーブルに、エラー情報と共に格納し、以降の処理には使用しない。 ・親様式1と子様式1の両方を除外する。(過去に仮確定した子様式1も除外する。)

番号	チェック名	チェック内容	対象データ	種別	フラグ内容	プロセス
ERR040	Fファイルキー重複	Fファイルの「施設コード+データ識別番号+入院年月日+データ区分+順序番号+行為明細番号+該当月」が重複	Fファイル	キー重複		<ul style="list-style-type: none"> ・重複データはFF1 データを除外データFF1 テーブルに、エラー情報と共に格納し、以降の処理には使用しない。 ・親様式1と子様式1の両方を除外する。(過去に仮確定した子様式1も除外する。)
ERR110	在院日数1日未満	入院日>退院日	FF1	除外		<ul style="list-style-type: none"> ・エラーデータは FF1 データを除外データ FF1 テーブルに、エラー情報と共に格納し、以降の処理には使用しない。 ・親様式1(診療情報番号=0)がエラーの場合、子様式1(診療情報番号≥1)のレコードも一緒に除外する。(過去に仮確定した子様式1も除外する。) ・子様式1のエラーの場合は、該当子様式1は除外するが、親様式1および他の子様式1は除外しない。 ・子様式1で退院日が'00000000'である場合は、許容する。親様式1はこのエラーとなる。
ERR120	入院時年齢0歳未満	生年月日>入院日	FF1	除外		<ul style="list-style-type: none"> ・エラーデータは FF1 データを除外データ FF1 テーブルに、エラー情報と共に格納し、以降の処理には使用しない。 ・親様式1(診療情報番号=0)がエラーの場合、子様式1(診療情報番号≥1)のレコードも一緒に除外する。(過去に仮確定した子様式1も除外する。) ・子様式1のエラーの場合は、該当子様式1は除外するが、親様式1および他の子様式1は除外しない。

番号	チェック名	チェック内容	対象データ	種別	フラグ内容	プロセス
ERR130	年月日誤り	実在しない年月日(13月1日、7月32日など)、およびSQLserverのdatetime型で扱えない日付(1753年1月1日以前)	ALL	除外		<ul style="list-style-type: none"> ・エラーデータはFF1データを除外データFF1テーブルに、エラー情報と共に格納し、以降の処理には使用しない。 ・対象は全ての日付情報とする ・あくまで実在しない年月日のチェックのみであり、手術日が入院日と退院日の間にあるか、などの関連チェックは行なわない。 ・'00000000'は許容する。 ・親様式1(診療情報番号=0)がエラーの場合、子様式1(診療情報番号≥1)のレコードも一緒に除外する。(過去に仮確定した子様式1も除外する。) ・子様式1のエラーの場合は、該当子様式1は除外するが、親様式1および他の子様式1は除外しない。 ・様式1開始日、様式1終了日をチェック対象に追加。
ERR140	7月から12月退棟でない、または、4月以降入棟でない(調査対象期間退棟でない、または、EF提出期間(4月～12月)以外に入棟している)	(様式1開始日が当年4月以降、かつ、様式1終了日が当年7月以降12月以前)でない	FF1	除外		<ul style="list-style-type: none"> ・エラーデータはFF1データを除外データFF1テーブルに、エラー情報と共に格納し、以降の処理には使用しない。 ・子様式1のエラーの場合は、該当子様式1は除外するが、親様式1および他の子様式1は除外しない。
ERR160	統括診療情報番号異常(3日以内再入院集約レコード)	統括診療情報番号が0以上の数字でない	FF1	除外		<ul style="list-style-type: none"> ・他のエラーデータ除外より最初に判定する。 ・エラーデータはFF1データを除外データFF1テーブルに、エラー情報と共に格納し、以降の処理には使用しない。
ERR170	様式1対象期間1日未満	様式1開始日<様式1終了日	FF1	除外		<ul style="list-style-type: none"> ・エラーデータはFF1データを除外データFF1テーブルに、エラー情報と共に格納し、以降の処理には使用しない。

番号	チェック名	チェック内容	対象データ	種別	フラグ内容	プロセス
						<ul style="list-style-type: none"> ・親様式1(診療情報番号=0)がエラーの場合、子様式1(診療情報番号≧1)のレコードも一緒に除外する。(過去に仮確定した子様式1も除外する。) ・子様式1のエラーの場合は、該当子様式1は除外するが、親様式1および他の子様式1は除外しない。
ERR180	親様式1において様式1開始日、様式1終了日が入院日、退院日に一致していない	統括診療情報番号=0かつ(様式1開始日が入院年月日と一致しない または 様式1終了日が退院年月日と一致しない)	FF1	除外		<ul style="list-style-type: none"> ・エラーデータは FF1 データを除外データ FF1 テーブルに、エラー情報と共に格納し、以降の処理には使用しない。 ・親様式1(診療情報番号=0)がエラーの場合、子様式1(診療情報番号≧1)のレコードも一緒に除外する。(過去に仮確定した子様式1も除外する。) ・子様式1のエラーの場合は、該当子様式1は除外するが、親様式1および他の子様式1は除外しない。
ERR190	様式1開始日が入院と退院の範囲外である	様式1開始日<入院年月日 または 様式1開始日 > 退院年月日	FF1	除外		<ul style="list-style-type: none"> ・エラーデータは FF1 データを除外データ FF1 テーブルに、エラー情報と共に格納し、以降の処理には使用しない。 ・親様式1(診療情報番号=0)がエラーの場合、子様式1(診療情報番号≧1)のレコードも一緒に除外する。(過去に仮確定した子様式1も除外する。) ・子様式1のエラーの場合は、該当子様式1は除外するが、親様式1および他の子様式1は除外しない。
ERR200	様式1終了日が入院と退院の範囲外である	様式1終了日<入院年月日 または 様式1終了日 > 退院年月日	FF1	除外		<ul style="list-style-type: none"> ・エラーデータは FF1 データを除外データ FF1 テーブルに、エラー情報と共に格納し、以降の処理には使用しない。 ・親様式1(診療情報番号=0)がエラーの場合、子様式1(診療情報番号≧1)のレコードも一緒に除外する。(過去に仮確定した子様式1も除外する。)