

参考資料 1

DPC データ分析システムの概要

内容

1. DPC データ分析システムの要件の設定とシステム設計	27
2. サンプルデータを用いた分析テストラン	33
参考資料1-1. DPC 分析システム概要設計書 「システム構成と機能一覧」	35
参考資料1-2. DPC 分析システムインターフェース説明書 「管理ツールおよび分析ツールの画面遷移図」	57
参考資料1-3. 診断群分類見直しのためのデータ収集 調査項目一覧	85
参考資料1-4. レセプト電算コードマスター	129

1. DPC データ分析システムの要件の設定とシステム設計

1-1. 分析システムの概要

本研究に用いられる大量のデータの効率的な解析を実施するためには、収集されたデータを元に、年齢、性別、診断群分類、在院日数、医療費等の視点からインタラクティブな解析を提供できるデータベースと分析システムが必要である。これらの点に関して、分析に適するデータフォーマット、データ統合の手法、分析システムのあり方等を検討した。また、インタラクティブな分析のインターフェース、分析手順等を検討した。

全体のシステム構成は、参考資料1-1のDPC分析システム概要設計書「システム構成と機能一覧」に示すように、データ保護の観点からの安全性と、各大学病院の分析担当者がネットワークを介して効率的にデータ分析作業に携われるよう利便性を考慮した。データのリモートアクセスはファイアーウォールとVPNを用いることで安全性と利便性の両立を図った。また、強固な認証機能を装備することにより、外部からの侵入を防止しデータを保護する手段を取るとともに、各大学病院毎の分析担当者にそれぞれデータアクセス範囲を設定することで、集計データは他大学と比較可能でかつ自大学の個別データのみにはアクセス出来る仕様とした。

1-2. 分析システムの設計の背景

高齢者医療を含む診断群分類の妥当性の検証には、散布図、箱ひげ図などの集約的分析とアウトライヤー、インライヤー解析などの個別的な分析を、効率的に平行して進めていく必要がある。そこで、診断群分類の各階層での年齢、在院日数、医療費等のばらつきとその要因の分析、医療費区分毎のばらつきへの影響度の検討、ばらつきの医療機関要因の分析等に散布図、箱ひげ図、変動係数等を活用する方法論を検討した。

年齢、診断群分類、在院日数、医療費等の視点からの集約的分析には、入院 ID をキーとして「様式1」データ、「様式2」データ、DPC コードを格納するリレーショナルDBを構築し、それを基礎に分析用アプリケーション作成と集計分析を実施した。EF ファイルデータはレコードが細分化され、そのままでは集計分析に適していないため、「様式2」相当ファイルに集計して利用した。

データ量、分析手法の複雑性から診断群分類毎の集計データをリアルタイムで閲覧するのは実用的ではなかったため、あらかじめバッチ処理によって全ての分析パネルを作成し、分析者のリクエストに応じて必要な集計結果パネルを閲覧する方法が有効であった。分析結果は html 化されていたため、web 参照など汎用性、携帯性に優れていた。

一方、アウトライヤー分析等で必要な個別的解析には、集約的分析の個別データからリンクされた EF ファイルデータ参照機能が効果的であった。箱ひげグラフ上のはずれ値から直接 EF ファイルを参照し、アウトライヤー症例の具体的診療内容を精査することにより、アウトライヤー要因の解析が効率的に実施された。

1-3. 分析システムの具体的な機能

以上の検討を元に、以下のような機能を持つように分析システムを設計した。装備された基本的な機能としては、①利用者管理、②マスタ管理、③データ収集、④データ抽出、⑤データ分析、⑥データエクスポートの6つを用意した。

①利用者管理は、利用者登録と認証により、それぞれの利用者が設定されたアクセス権の範囲内で分析とデータアクセスが可能となる。アクセス権限は、集計データ、統計データ、様式1レベル個別データ、EF明細データレベル個別データ等の別に設定可能とした。

②マスタ管理は、基本マスタ、定義表および点数表に係わるDPCマスタ、施設コード及び名称のマスタ、診療行為コード等に関するマスタを整備し、必要により修正、更新が可能なものとした。

③データ収集は、様式1およびEFファイルをそれぞれ個別に取り込める仕様とし、データの不一致等が検出出来る機能を付加した。データ取り込み時にDPCコードをふり直すことも可能である。

④データ抽出機能は、本分析システムの要となる部分で、基本的には元データにある全ての情報によって必要なデータが抽出可能な使用となっている。様式1の患者プロフィール、診断名、DPC分類、診療情報等、および診療報酬明細レベルの条件による抽出も可能である。これらの条件を条件ファイルとして保存することで、他分析者との共有も可能である。データ抽出ロジックは、GUIを用いた取り扱いやすいものを用意し、データ解析の経験が少ない研究者もDPCデータ分析に参加出来るよう考慮した。

⑤データ分析機能は、参考資料1-1のDPC分析システム概要設計書「システム構成と機能一覧」に示すように、抽出されたデータの集約結果を多角的に様々な視点から視覚的に把握出来るような機能を設定した。集約の軸としては、年齢階級別、性別、施設別、1日あたり診療報酬請求額別、診療区分別診療報酬請求額別集計等とともに、日計表プロット、箱ひげプロットを用意し、さらにはずれ値から直接、個別データにアクセスするドリルスルー機能を設定した。この機能により、はずれ値症例の個

別データ、EF ファイル明細データ等も瞬時にアクセスし分析することが可能となっている。この機能は、アウトライヤーの解析などに強かにサポートすることが期待される。

⑥データエクスポートは設定された抽出条件により抽出された全データを二次利用可能なように CSV フォーマットなどの形で電子媒体に転送する機能として設定した。この機能により、抽出されたデータを用いた複雑な統計解析などの利便性が向上すると考えられる。

1-4. 分析システムを用いた DPC データ分析方法の概要

DPC 分析システムの基本的な使用方法と画面インターフェースは参考資料1-2 DPC 分析システムインターフェース説明書に示す。基本的な運用フローは、①データの取り込み、②抽出条件の設定、③データ集計、④データ分析であり、後半の②、③、④を適宜反復する探索的データ分析により、DPC のあり方や改善方法を検討することが可能であると考えられる。

システムの基本ツールは管理ツールと分析ツールに分けられる。管理ツール部分では、対話的な機能により、ログイン、メインメニュー、マスタ設定、データインポート、データエクスポート等の作業をサポートするようになっている。

分析ツール部分は、同様に対話的な機能により、ログイン、条件シートの設定、集計指示、グラフ分析等の作業をサポートする。条件設定シートは複数の条件をタブを用いて組み合わせることにより、理論的にはあらゆる条件による抽出をサポートしている。グラフ分析部分に置いては、マウスによるポインティングと明細データへのドリルスルーがリンクされているため、インタラクティブかつリアルタイムにアウトライヤーデータを調査、解析することが可能である。

1-5. 分析に用いる原データの詳細

DPC データ分析システムに投入するデータの詳細を参考資料1-3「診断群分類見直しのためのデータ収集 調査項目一覧」に示した。基本的な診療情報に相当する様式1および様式1データを正確に収集するための診療情報入力票には、収集されたデータの定義、コード、意義が示されている。

様式2データは、基本的には診療報酬の集計データであり、診療区分毎の診療行為の状況、一部の特殊な診療行為の状況、加算等の設定情報等に基づくデータの解析に必要であると考えられる。

様式3は施設データに係わる調査票、様式4は保険外診療の状況の調査票、様式5は包括点数情報の調査票であり、これらの一部がデータ分析システムに取り込まれ、分析に使用可能となっている。

様式E、Fはレセプトデータダウンロード・データ仕様として定義されている。これは、診療報酬の明細レベルのデータを実施日の情報とともに保持しているため、診療行為の状況を把握する上で重要なデータを提供している。特に、F データは、個別診療行為データを示しているため、診療報酬上は包括されてしまう、検査、画像診断等の個別明細データを把握することも可能としているため、診療内容の解析には非常に深い情報を提供すると考えられる。

これらのE、Fファイルにおける診療行為を示すデータは、参考資料1-4レセプト電算コードマスターに定義されるコードを用いることが必要とされている。したがって、本 DPC 分析システムに於いては、これらのマスタにおいて定義されたコードに基づき、特定の診療行為等を抽出することが可能である。マスタとしては、診療行為コードマスター、高額医薬品マスタ、材料コードマスターがある。例えば、これらのマスタに基づき、乳幼児加算を算定している症例の抽出、特定の手術を実施している症例の抽出、特定の薬剤を使用している症例の抽出が可能である。

2. サンプルデータを用いた分析テストラン

分析結果パネルは、棒グラフ、散布図、箱ひげ図などの集計グラフと平均、最大、最小、分散、変動係数等の基本統計一覧表を組として作成することにより、正確な検討が可能であった。

具体的な項目をいくつか以下に例示する。

- 性別・年齢別の症例数の分布の集計は、疾患の性・年齢特異性を把握する上で必須であった。
- 病院別在院日数の評価は、医療機関別特異性の把握に有益であった。はずれ値の一覧表を付記することにより、個別症例に関する EF ファイルの詳細データを参照することが可能であり、医療機関毎のアウトライヤー要因の分析と患者年齢の関与等の解析に重要な役割を果たした。
- 在院日数と1日あたり診療報酬請求額のプロット、在院日数と1入院あたり診療報酬請求額のプロット、および在院日数と診療報酬区分別プロット等は、患者の年齢区分毎に比較が可能であり、高齢者に関する包括評価の妥当性、診断群分類の妥当性、短期入院評価の分析に有用であった。
- 診療報酬区分別の1日あたり診療報酬請求額と1入院あたり診療報酬請求額は、変動要因への患者年齢の影響の分析に必要であった。そのほか病院特異性および地域特異性の把握は、患者年齢分布の相違との関連性を比較する上で有用であった。

参考資料 1-1

DPC 分析システム概要設計書

システム構成と機能一覧

システム名

DPC分析システム

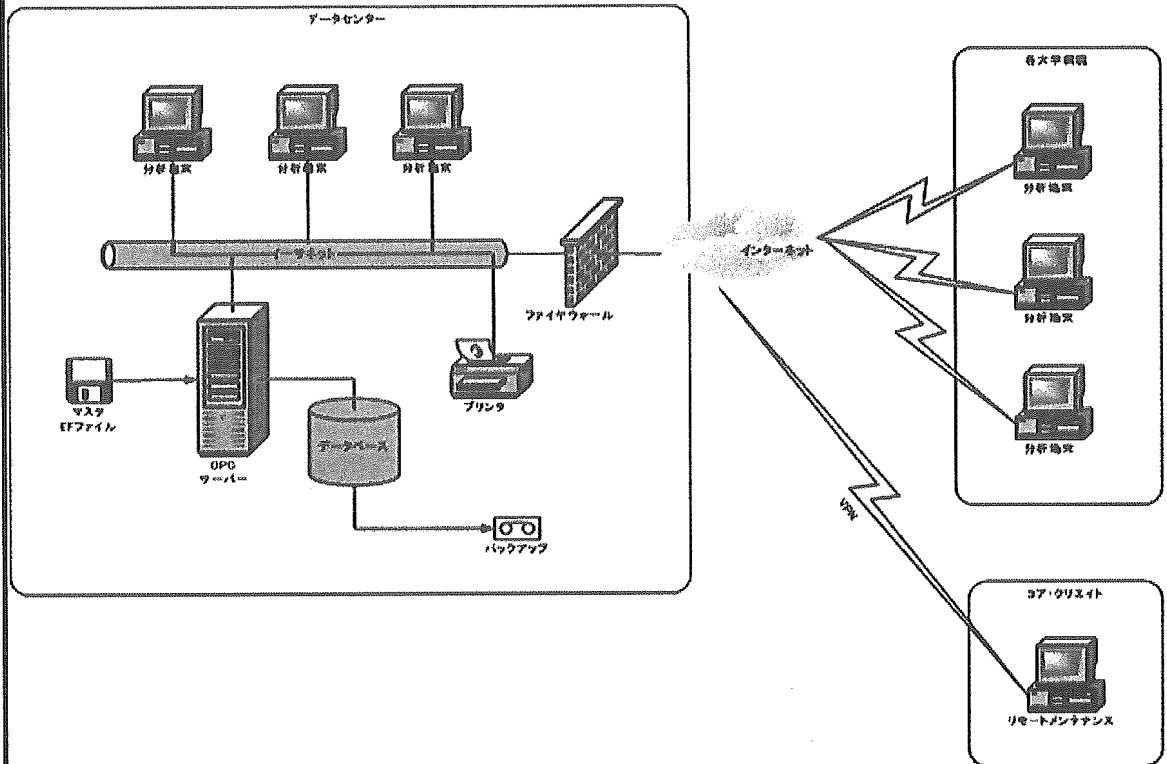
概要設計書

DPC分析システム
概要設計書

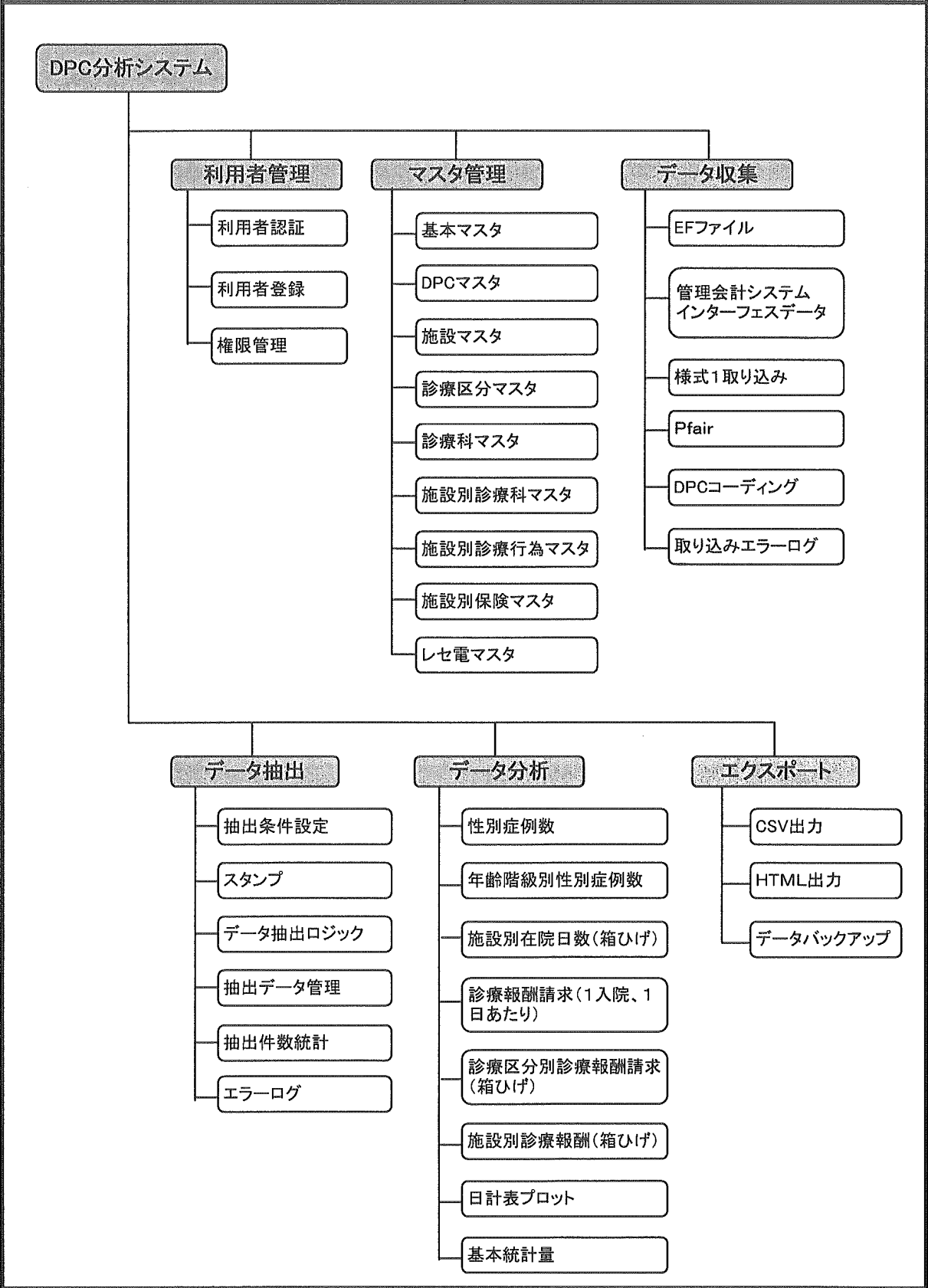
システム名	
DPC 分析システム	目次

1. システム構成図	41
2. 機能構成図	42
3. 機能一覧	43

システム名 DPC分析システム サブシステム名	文書名 システム構成図
-------------------------------	----------------



システム名	文書名
DPC分析システム	機能構成図



システム名 DPC分析システム サブシステム名 利用者管理		機能一覧
No	機能名称	機能概要
1	利用者認証	DPC分析システムを利用するために認証を行う。 この機能はWEBインターフェースでクライアント機能を実装するため、利用者に応じてアクセスできるコンテンツの制限を設ける。 昨年度に開発したDPC分析では、CD-ROMメディアを利用したコンテンツ提供をおこなってきたが、リアルタイムに分析ができない等欠点があるため、WEBインターフェースでの分析システムを実装する。
2	利用者登録	このシステムを利用するユーザ登録、修正、削除をおこなう。
3	権限管理	2. 利用者登録において登録したユーザの権限を管理する。 全てのプログラムに対し、実行、参照、データ更新などの権限を実装する。

システム名 DPC分析システム サブシステム名 マスタ管理		機能一覧
No.	機能名称	機能概要
1	基本マスタ	このシステムで不変的なマスタ。全国の施設に関係なく共通で利用するマスタやシステムマスタ等のメンテナンスをする。
2	DPCマスタ	厚生労働省から提供されるDPCマスタのメンテナンスをおこなう。
3	施設マスタ	特定機能病院のマスタメンテナンスをおこなう。
4	診療区分マスタ	EFファイルで利用する診療区分マスタ、7区分マスタのメンテナンスをおこなう。
5	診療科マスタ	厚生労働省から提供されている診療科マスタのメンテナンスをおこなう。
6	施設別診療科マスタ	施設固有の診療科コードと5. 診療科マスタの対応テーブルをメンテナンスをおこなう。
7	施設別診療行為マスタ	施設固有の診療行為マスタのメンテナンスをおこなう。レセ電コードへの対応テーブルを作成する。
8	施設別保険マスタ	施設固有の保険マスタのメンテナンスをおこなう。
9	レセ電マスタ	レセ電マスタのメンテナンスをおこなう。
10	病名マスタ	Kコードの病名マスタメンテナンスをおこなう、。

システム名 DPC分析システム サブシステム名 データ収集		機能一覧
No	機能名称	機能概要
1	EFファイル取り込み	EFファイル及び様式5の取り込みをおこなう。
2	管理会計システム インターフェース データ取り込み	医事会計基本データ、医事会計明細データ、 DPC明細データ、患者数データ、保険データ、 病名データ、診療科マスタ、診療区分マスタ、 実施部門マスタ、病棟マスタ、病室マスタ、。 点数マスタ、伝票マスタ、病名マスタの 取り込みをおこなう
3	様式1取り込み	様式1のデータ取り込みをおこなう。
4	DPCコーディング	DPCコーディング機能の実装（詳細コード及び 支払用コードのコーディング）
5	取り込みエラーログ	取り込みで発生したエラーのリストを表示及び PDFファイル出力、印刷をおこなう。

システム名 DPC分析システム サブシステム名 データ抽出		機能一覧
No	機能名称	機能概要
1	抽出条件設定	DPC分析をおこなうための前処理としてデータをあらゆる角度から抽出をおこない、抽出されたデータを元に、分析をおこなう。 抽出条件 性別、年齢別、入外別、高額機材、Kコード、施設別、診療科別、診療区分(7区分)、退院後の再来の有無、一般医療、老人医療、小児医療区分別、様式1の副傷病名別 DPCコードでの桁数分類別 入院期間別
2	スタンプ	1. 抽出条件設定で登録した条件をスタンプとして保存をおこなう。以降はこのスタンプを利用して、簡単に抽出をおこなう。
3	データ抽出ロジック	1. 抽出条件設定で条件した内容をもとにデータベースからデータの抽出をおこない、二時利用データベースに格納をおこなう。
4	抽出データ管理	3. データ抽出ロジックにて抽出されたデータの保存、削除、エクスポートの管理をおこなう。
5	抽出件数統計	3. データ抽出ロジックにて抽出されたデータのサマリーを表示。該当患者数や診療区分別件数等
6	エラーログ	3. データ抽出ロジックにて抽出中に発生したエラーの管理をおこなう。

システム名
DPC分析システム
サブシステム名
データ分析

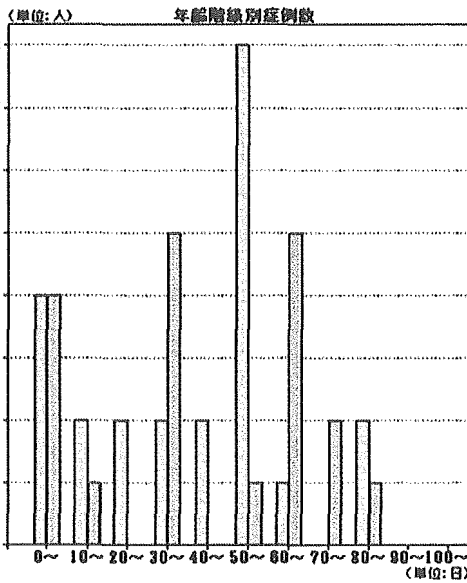
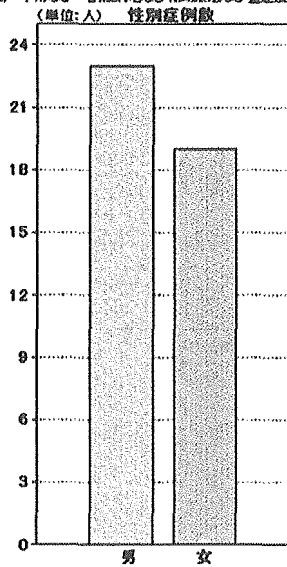
機能一覧

No 機能名称 機能概要

1 性別症例数

2 年齢階級別症例数

①全包括 0100100000
総患数→平均なし→別個病名なし 補助療法なし 重症度なし



性別症例数	
男	女
22	19

年齢階級別症例数	
年齢	症例
0~	4
10~	2
20~	2
30~	2
40~	5
50~	8
60~	1
70~	5
80~	2
90~	2
100~	1

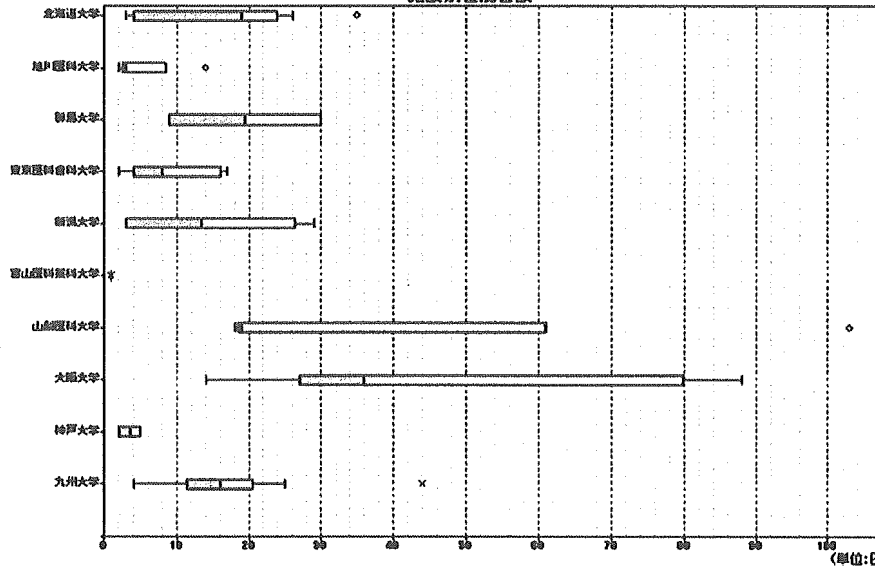
システム名 DPC分析システム サブシステム名 データ分析	機能一覧
--	------

No 機能名称 機能概要

3 施設別在院日数(箱ひげ)

U 高年齢 U100100000
U 高齢者→平均なし→ 別病名なし、補助療法なし、重症度なし

施設別在院日数



	箱ひげ下 下値	箱ひげ上 箱下値	ひげ下値	BOX下値	メジアン	BOX上値	ひげ上値	箱ひげ上 箱上値	箱ひげ上 上値	最大値	#
九州大学	-15.5	-4	4	11.5	16	20.5	25	25.5	30.5	44	
群馬大学	-9	-2.5	2	2	3.5	6	8	6.5	11	5	
大塚大学	-13.5	-2.5	14	27	36	50	55	105.5	149	55	
山形医科大学	-10.5	-0.5	10	10.5	15	21	21	22.5	145	103	
岡山医科大学	1	1	1	1	1	1	1	1	1	1	
新潟大学	-9.5	-2.5	3	3	13.5	25.5	29	25.5	72.5	29	
東京医科大学	-3.5	-1.5	2	4	6	10	17	22	40	17	
群馬大学	-6.5	-2.5	5	5	19.5	30	30	40.5	72	30	
旭川医科大学	-16.5	-4.5	2	2.5	3	6.5	6.5	11.5	20.5	14	
北海道大学	-3.5	-2	3	4	15	24	28	34	64	38	

施設別データ				
	施設コード	患者番号	区分	箱
九州大学	409810997	0005416985	他科(上)	44
山形医科大学	192110952	0000033389	他科(上)	103
岡山医科大学	168010036	0002263652	他科(上)	1