

(19) 日本国特許庁 (JP)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2016-517556

(P2016-517556A)

(43) 公表日 平成28年6月16日 (2016.6.16)

(51) Int. Cl.	F I	テーマコード (参考)
<b>G06Q 50/22 (2012.01)</b>	G06Q 50/22 106	4C341
<b>G06F 17/30 (2006.01)</b>	G06F 17/30 220Z	5L099
<b>A61G 12/00 (2006.01)</b>	A61G 12/00 Z	

審査請求 有 予備審査請求 未請求 (全 26 頁)

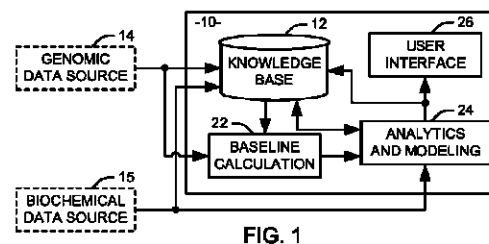
(21) 出願番号	特願2015-560310 (P2015-560310)	(71) 出願人	503178185
(86) (22) 出願日	平成26年2月27日 (2014.2.27)		ノースロップ グラマン システムズ コーポレーション
(85) 翻訳文提出日	平成27年8月28日 (2015.8.28)		NORTHROP GRUMMAN SYSTEMS CORPORATION
(86) 国際出願番号	PCT/US2014/019022		アメリカ合衆国 22042-4511
(87) 国際公開番号	W02014/149497		バージニア州 フォールズ チャーチ フェアビュー パーク ドライブ 2980
(87) 国際公開日	平成26年9月25日 (2014.9.25)	(74) 代理人	100105957
(31) 優先権主張番号	13/837,370		弁理士 恩田 誠
(32) 優先日	平成25年3月15日 (2013.3.15)	(74) 代理人	100068755
(33) 優先権主張国	米国 (US)		弁理士 恩田 博宣
		(74) 代理人	100142907
			弁理士 本田 淳

最終頁に続く

(54) 【発明の名称】 健康学習システム及び方法

## (57) 【要約】

健康学習システム及びそれに関連する方法が提供される。個人から採取された血液試料に対する生化学アッセイを計画された間隔で実行し、複数の生化学パラメータのそれぞれについて、個人を表す値の時系列を提供する。個人に関連する臨床パラメータが知識ベースから抽出される。個人のゲノムパラメータが特定される。複数の生化学パラメータの複数のサブセットのそれぞれの予期時系列が、少なくとも臨床パラメータ及びゲノムパラメータから計算される。生化学パラメータの複数のサブセットのそれぞれについて、個人を表す値の時系列が、計算された予期時系列と比較されて、個人についての複数の病気のそれぞれの可能性を特定する。複数の病気のうちの少なくとも1つの病気の可能性がユーザに伝達される。



**【特許請求の範囲】****【請求項 1】**

個人化されたヘルスケアサポートを提供する方法であって、

個人から採取された血液試料に対する生化学アッセイを計画された間隔で実行して、複数の生化学パラメータのそれぞれについて、前記個人を表す値の時系列を生成し、

前記個人に関連する複数の臨床パラメータを知識ベースから抽出し、

前記個人の複数のゲノムパラメータを特定し、

少なくとも前記複数の臨床パラメータ及び前記複数のゲノムパラメータから、前記複数の生化学パラメータの複数のサブセットのそれぞれの予期時系列を計算し、

前記生化学パラメータの複数のサブセットのそれぞれについて、前記個人を表す値の時系列を、計算された前記予期時系列と比較して、前記個人についての複数の病気のそれぞれの可能性を特定し、

ヘルスケア行動指針を作成できるように前記複数の病気のうちの少なくとも 1 つの病気の可能性を出力すること

を備える方法。

**【請求項 2】**

前記複数の病気のうちの少なくとも 1 つの病気の可能性を出力することは、前記複数の病気のうちの少なくとも 1 つの病気の可能性をユーザに伝達することを備える、請求項 1 に記載の方法。

**【請求項 3】**

前記ユーザは前記個人であり、前記少なくとも 1 つの病気の可能性をユーザに伝達することは、伝達された前記少なくとも 1 つの病気の可能性に基づいて、ヘルスケア治療行動指針を前記個人に提供することを備える、請求項 2 に記載の方法。

**【請求項 4】**

前記ヘルスケア治療行動指針を前記個人に提供することは、伝達された前記少なくとも 1 つの病気の可能性に基づいて、次の生化学アッセイの実行タイミングを前記個人に指示することを備える、請求項 3 に記載の方法。

**【請求項 5】**

前記ヘルスケア治療行動指針を前記個人に提供することは、特定のタイプの医療関係者からの治療を求めるように前記ユーザに指示することを備える、請求項 3 に記載の方法。

**【請求項 6】**

前記ユーザは臨床医であり、前記少なくとも 1 つの病気の可能性をユーザに伝達することは、伝達された前記少なくとも 1 つの病気の可能性に基づいてケアプロトコルを前記臨床医に提供することを備える、請求項 2 に記載の方法。

**【請求項 7】**

前記複数の病気のうちの少なくとも 1 つの病気の可能性を出力することは、前記複数の病気のうちの少なくとも 1 つの病気の可能性を、知識ベース、データベース、又はディスプレイのうちの少なくとも 1 つに供給することを備え、出力された前記可能性は、予測モデルを更新するための少なくとも 1 つの教師なし学習プロセスに利用可能である、請求項 1 に記載の方法。

**【請求項 8】**

前記個人に関連する複数の個人をそれぞれ表す複数の生化学アッセイ系列から、複数のコホートパラメータを抽出することを更に備え、前記複数の生化学パラメータの前記複数のサブセットのそれぞれの予期時系列は、少なくとも前記臨床パラメータ、前記ゲノムパラメータ、及び前記コホートパラメータから計算される、請求項 1 に記載の方法。

**【請求項 9】**

前記生化学パラメータの複数のサブセットのうちの選択された 1 つのサブセットについて、選択された前記生化学パラメータの前記個人を表す時系列及び選択された前記生化学パラメータの計算された前記予期時系列のそれぞれのグラフィックス表現を前記ユーザに表示して、計算された前記予期時系列を計画された前記生化学アッセイからの測定値と比較

10

20

30

40

50

できるようにすることを更に備える、請求項 1 に記載の方法。

【請求項 1 0】

前記ユーザは臨床医であり、前記方法は更に、前記予期時系列の計算に使用された前記臨床パラメータ及び前記ゲノムパラメータのうちの選択された一つのパラメータから、新しい値を前記臨床医が選択できるようにし、前記臨床パラメータ、前記ゲノムパラメータ、及び前記コホートパラメータのうちの前記選択された一つのパラメータの前記新しい値を反映するように、前記予期時系列の前記グラフィック表現を変更することを備える、請求項 1 に記載の方法。

【請求項 1 1】

前記知識ベースが複数の患者からなる母集団の生化学アッセイ、ゲノムパラメータ、臨床パラメータ、及び測定された臨床結果を有するように、前記値の時系列、前記複数のゲノムパラメータ、及び測定された臨床結果のそれぞれを前記知識ベースに格納することを更に備える、請求項 1 に記載の方法。

【請求項 1 2】

前記個人を表す値の時系列を計算された前記予期時系列と比較して前記個人の複数の病気のそれぞれの可能性を特定することは、

計算された前記予期時系列からの前記値の時系列の偏差を特定し、

前記複数の病気のうちの 1 つに関連する予測モデルであって、前記偏差から導出される少なくとも 1 つのパラメータから、前記複数の病気のうちの前記関連する 1 つの病気の可能性を特定するように構成されている前記予測モデルに前記偏差を入力として適用すること

を備える、請求項 1 1 に記載の方法。

【請求項 1 3】

前記知識ベースにデータマイニングアルゴリズムを適用して、前記臨床パラメータ、前記ゲノムパラメータ、及び前記コホートパラメータのうちの 1 つを病気に関連付ける少なくとも 1 つの因果事例を識別し、

識別された前記因果事例に従って予測モデルを改良すること  
を更に備える、請求項 1 2 に記載の方法。

【請求項 1 4】

プロセッサに動作可能に接続された少なくとも 1 つの非一時的コンピュータ可読媒体のセットに記憶された機械実行可能命令として実現される個人化ヘルスケアシステムであって、

知識ベースであって、対象となる複数の病気のそれぞれについて、当該病気を有する患者の組及び当該病気を有さない患者の組を含む複数の患者からなる母集団の記録と、一患者に関連するとともに計画された間隔で実行された生化学アッセイから取得した複数の生化学パラメータの値の時系列、複数の遺伝子マーカー、及び複数の臨床パラメータを含む所与の記録とを備える知識ベースと、

所与の患者について、少なくとも、当該患者に関連する前記臨床パラメータ及び前記ゲノムパラメータから、前記生化学パラメータの複数のサブセットのそれぞれの予期時系列を計算するように構成されたベースライン計算構成要素と、

計算された前記予期時系列からの前記値の時系列の偏差を特定し、前記複数の病気のうちの 1 つの病気に関連する予測モデルであって、当該病気を有する患者の組及び当該病気を有さない患者の組のそれぞれに関連するデータから導出されたものであり、かつ、前記偏差から導出された少なくとも 1 つのパラメータから前記患者が前記病気を有する可能性を特定するように構成された前記予測モデルに前記偏差を入力として適用するように構成された分析及びモデリング構成要素と、

特定された前記患者が前記病気を有する可能性をユーザに提供するように構成されたユーザインタフェースと  
を備えるシステム。

【請求項 1 5】

前記知識ベースには、所定の時間期間後に前記患者が前記病気を有するか否かを少なくとも反映した測定臨床結果が提供され、前記分析及びモデリング構成要素は、複数の教師なし学習アルゴリズムを前記知識ベースに対して実行して、前記臨床パラメータ及び前記ゲノムパラメータのうち的一方を前記病気に関連付ける少なくとも1つの因果事例を特定するように構成されたデータマイニング構成要素を備える、請求項14に記載のシステム。

【請求項16】

前記分析及びモデリング構成要素は、前記ユーザが前記ユーザインタフェースを通して利用可能な分析構成要素であって、前記知識ベース及び関連データベースからデータを検索して、前記データマイニング構成要素によって特定される前記少なくとも1つの因果事例をサポートするエビデンスを提供するように構成された前記分析構成要素を更に備える、請求項15に記載のシステム。

10

【請求項17】

前記分析及びモデリング構成要素は規則エンジンを更に備え、前記規則エンジンは、関連する規則セットに従って、前記データマイニング構成要素によって特定される因果事例を評価して、前記分析構成要素を用いて評価される場合、前記因果事例に関連するどの変数が、実施可能な結果を提供する最も高い可能性を提示するかを特定する、請求項16に記載のシステム。

【請求項18】

前記患者に関連する前記複数の臨床パラメータは、前記患者の年齢、体重、血圧、及び体温のうちの少なくとも1つを含む、請求項16に記載のシステム。

20

【請求項19】

前記ユーザインタフェースは、前記病気の前記可能性、ヘルスケア治療行動指針、及び計画された次の生化学アッセイのそれぞれを伝達するように構成される患者ダッシュボードを備える、請求項14に記載のシステム。

【請求項20】

情報抽出構成要素を組み込み、非構造化リサーチ源を、前記知識ベースと互換性を有するテンプレートに変換するリサーチインタフェースを更に備える、請求項14に記載のシステム。

【請求項21】

前記ユーザインタフェースは、特定された前記患者が前記病気を有する可能性に基づいて、推奨されるケアプロトコルを臨床医に伝達するように構成される臨床判断サポート構成要素を備える、請求項14に記載のシステム。

30

【請求項22】

患者に個人化ヘルスケアを提供する方法であって、

患者から採取された血液試料に対する生化学アッセイを計画された間隔で実行して、複数の生化学パラメータのそれぞれについて、前記患者を表す値の時系列を生成し、

前記患者に関連する複数の臨床パラメータを知識ベースから抽出し、

前記患者の複数のゲノムパラメータを特定し、

前記患者に関連する複数の個人をそれぞれ表す複数の生化学アッセイ系列から、複数のコホートパラメータを抽出し、

40

少なくとも前記臨床パラメータ、前記コホートパラメータ、及び前記ゲノムパラメータから、前記複数の生化学パラメータの複数のサブセットのそれぞれの予期時系列を計算し、

計算された前記予期時系列からの前記値の時系列の偏差を特定し、

病気に関連する予測モデルであって、前記偏差から導出された少なくとも1つのパラメータから前記病気の可能性を特定するように構成されている予測モデルに前記偏差を入力として適用し、

前記病気の可能性に基づいてヘルスケア治療行動指針をユーザに提供することを備える方法。

【発明の詳細な説明】

50

## 【技術分野】

## 【0001】

本発明は、ヘルスケア情報管理に関し、特に健康学習システム及び方法に関する。

## 【背景技術】

## 【0002】

ヘルスケア業界は、治療的、予防的、リハビリ的、及び苦痛緩和的なケアによって患者を治療するための商品及びサービスを提供する。近代のヘルスケアセクタは、多くのサブセクタに分割され、個人及び人々の健康ニーズを満たすように、熟練した専門家及び専門職の助手の合同チームに依存している。ヘルスケア業界は、世界中で最大の且つ最速で成長中の業界である。ヘルスケアは、大半の先進国の国内総生産（GDP）の10%超を消費し、国の経済の膨大な部分を形成し得る。現在、米国は、GDPの17%超をヘルスケアに費やしており、この量は毎年、略6%で成長すると見込まれている。ヘルスケアのこの増大を鈍化させ、最終的には逆転させるために、多くの試みがなされてきたが、大半の試みは失敗したか又は今までのところはまだ影響を及ぼしていない。費用の影響に起因して、ヘルスケアは多くの場合、ケアの集約度がより低く、費用がはるかに低い疾病の早期段階ではなく、疾病の後期段階で、現行の技術及び応用に基づいて、提供される。

10

## 【発明の概要】

## 【課題を解決するための手段】

## 【0003】

本発明の一局面によれば、個人化されたヘルスケアサポートを提供する方法が提供される。個人から採取された血液試料に対する生化学アッセイを計画された間隔で実行して、複数の生化学パラメータのそれぞれについて、前記個人を表す値の時系列を生成する。前記個人に関連する複数の臨床パラメータを知識ベースから抽出し、前記個人の複数のゲノムパラメータを特定し、少なくとも臨床パラメータ及びゲノムパラメータから、複数の生化学パラメータの複数のサブセットのそれぞれの予期時系列を計算する。生化学パラメータの複数のサブセットのそれぞれについて、前記個人を表す値の時系列と、計算された予期時系列とを比較して、前記個人についての複数の病気のそれぞれの可能性を特定する。複数の病気のうちの少なくとも1つの病気の可能性がユーザに伝達される。

20

## 【0004】

本発明の別の局面に従う健康学習システムは、複数の患者からなる母集団の記録を格納した知識ベースを含む。所与の記録は、一患者に関連するとともに計画された間隔で実行された生化学アッセイから取得した複数の生化学パラメータの値の時系列、複数の遺伝子マーカー、及び複数の臨床パラメータを含む。複数の患者からなる母集団は、対象となる複数の病気のそれぞれについて、当該病気を有する患者の組及び当該病気を有さない患者の組を含む。ベースライン計算構成要素は、所与の患者について、少なくとも、当該患者に関連する臨床パラメータ及びゲノムパラメータから、生化学パラメータの複数のサブセットのそれぞれの予期時系列を計算するように構成されている。分析及びモデリング構成要素は、計算された予期時系列からの前記値の時系列の偏差を特定し、複数の病気のうちの1つの病気に関連する予測モデルにこの偏差を入力として適用するように構成されている。予測モデルは、当該病気を有する患者の組及び当該病気を有さない患者の組のそれぞれに関連するデータから導出されたものであり、また、特定された偏差から導出された少なくとも1つのパラメータから、前記患者が病気を有する可能性を特定するように構成されている。ユーザインタフェースが、特定された患者が病気を有する可能性をユーザに提供するように構成されている。

30

40

## 【0005】

本発明の更に別の局面によれば、個人化ヘルスケアの方法が提供される。患者から採取された血液試料に対する生化学アッセイを計画された間隔で実行して、複数の生化学パラメータのそれぞれについて、前記患者を表す値の時系列を生成する。前記患者に関連する複数の臨床パラメータが知識ベースから抽出される。前記患者の複数のゲノムパラメータが特定される。前記患者に関連する複数の個人をそれぞれ表す複数の生化学アッセイ系列

50

から、複数のコホートパラメータが抽出される。複数の生化学パラメータの複数のサブセットのそれぞれの予期時系列が、少なくとも臨床パラメータ、コホートパラメータ、及びゲノムパラメータから計算される。計算された予期時系列からの前記値の時系列の偏差が特定される。病気に関連する予測モデルにこの偏差が入力として適用される。この予測モデルは、特定された偏差から導出された少なくとも1つのパラメータから、前記病気の可能性を特定するように構成されている。ユーザには、前記病気の可能性に基づくヘルスケア治療行動指針が提供される。

【図面の簡単な説明】

【0006】

【図1】本発明の一局面に従うヘルスケア学習システムを示す。

10

【図2】本発明の一局面に従うヘルスケア学習情報処理システムの一実装形態を示す。

【図3】本発明の一局面に従う個人化ヘルスケアサポートを提供する方法を示す。

【図4】本発明の一局面に従うヘルスケア学習システムにおいて新しい因果事例を発見し適用する方法を示す。

【図5】図1～図4に開示されるシステム及び方法の例を実施可能なハードウェア構成要素の例示的なシステムを示す概略ブロック図である。

【発明を実施するための形態】

【0007】

多くの国々でのヘルスケアは、個人に対して診断活動及び臨床活動及び臨床活動をいつどのように実行するかについての指針である医療プロトコールによって決定されている。しかし、これらのプロトコールはせいぜい、その個人の任意のかかなりの知識を表面的に参照して作成される。本発明者らは、ゲノミクスがケアのカスタマイズ化に有用であり得るが、測定時での健康状態へのより多くの洞察を提供する他のデータと共に、補足ゲノミクスデータが必要であることを見出した。ゲノミクスは主に、病気のリスクについてであるのに対し、他の測定値は現在の健康状態についてである。

20

【0008】

したがって、本発明者らは、プロテオミクス及び他のソースから導出された個人についてのデータにより、新しいタイプの医療プロトコールが可能になり得ると判断した。このプロトコールは、柔軟性に欠け、個人の医療状態ではなく母集団に基づく一般化に頼る今日の医療プロトコールに取って代わるものとして、個人の深い医療知識：個人の現在の医療状態及びプロテオーム状態並びに経時的な各自の傾向及び病歴の両方に適応する。そのような新しい個人化された医療プロトコールによる医療行為は、平均的な個人健康を改善しながら、それと同時に大きな費用節減を提供するものと期待される。

30

【0009】

一実装形態では、低コストで、侵襲性が最小のプロテオームに基づくテストを定期的に、例えば、個人のゲノミクスによって与えられる各スケジュールで、患者の母集団のそれぞれに対して実施することができ、個人化されたヘルスケアへの新しい手法に使用することができる。この定期的なテストからの結果を使用して、個人化された長期医療データを作成し、任意の後の診断の効率を大いに改善する。結果は、個人の深い医療知識並びに経時的な各自の傾向及び病歴を使用して個人化医療プロトコールを提供するのに使用可能な個人化医療データを提供するとともに、治療又は生活習慣の変更を必要とし得る特定の病気の進行の早期指標を提供するのに使用することができる。基本的に、プロテオームに基づくテストは、個人化ケアへの「門番」として機能する。この手法は、現在の医療モデルを反応的な症状に基づく手法から、個人化情報に基づく予測／予防的手法にシフトさせる。

40

【0010】

一実装形態では、システムは、1滴の血液からの数千ものタンパク質の同時アッセイからデータを受信することができる。これらのデータを電子ヘルスケア記録及び他のソース内のデータと組み合わせることで、各個人化患者についての現在の情報及び長期情報の両方を提供することができる。現在の医療知識のデータベース及びベストプラクティスは、

50

検知された医療信号を現在及び予測の病気及び診断に関連付ける「因果事例」の組と組み合わせ使用されることで、全て1回の血液テストから、数千もの病気へのヘルスケア治療行動指針についての高速で正確且つ個人化された診断及び推奨を同時に提供することができる。この手法は、任意の手順な診断手順から今日利用可能なデータよりもはるかに多くのデータに基づいて診断及び推奨を行うことにより、結果を改善するとともに、一連の高価なテストを安価な1回のテストで置換することにより、費用を低減し、早期検出及び介入を可能にし、専門家及び機関にわたるテスト結果の共有を機械化し、臨床判断実行でのばらつきを低減し、現在、不必要にスクリーニングされ、テストされ、且つ治療される広範囲の個人を大幅に低減する。さらに、このデータを全て中央知識ベースに保持することにより、新しい因果事例への研究に大いに役立つことができる。

10

#### 【0011】

ヘルスケアも実施に膨大なバリエーションがある。患者の病気を示す個人データを有するだけでは十分ではない。推奨される行動指針は、適切な実施を一貫して反映するとともに、エビデンスに基づく基準を組み込まなければならない。しかし、データは、医療行為での正当性がないバリエーションが、害、さらには死亡を招くことに加えて、全体の米国医療支出の30%~50%を占めることを示す。品質の改善は、単に母集団の平均を改善するのではなく、バリエーションの低減によって始まり、これは、本発明の一局面に従う健康学習システムによって対処される。国営医療システムの規模でバリエーションを低減するにはまず、新しい発見に基づくベストプラクティスのエビデンスを作成し、そのエビデンスを患者及び医療提供者への推奨に組み込む必要がある。

20

#### 【0012】

本提案のシステムは、従来のヘルスケア手法よりも優れた幾つかの利点を提供する。例えば、本システムでは、消費者が、客観的データに基づいて、いつヘルスケアシステムに入って詳細な診断及び治療を探る必要があるかを知ることができる。現在、人々はこの判断を大半が主観的又は信頼性が低い最小の情報（例えば、人々がどう「感じる」か、「何が痛むか」、体温等）のみを用いて下している。本システムでは、患者の入口地点である医療関係者が、侵襲が最小の技術及び高度の確実性を有するデータソースに基づいて、広範囲の病気及び要因にわたって患者の変化及び非正常状態を素早く評価し、患者を適切なテスト、スクリーニング、専門医師、及び処置に送り、それにより、時間、金銭、及び不満を低減できるようにもする。医療行為でのバリエーションは、詳細な診断情報を特定の患者層に適用される臨床有効性のエビデンスと結合することによって大幅に低減することができ、ヘルスケアシステムがケア提供を改善させるとともに、ケア提供の的を絞れるようにする。ヘルスケアシステムの効率も、病気の人々を疾患の過程でより早期に、より先見的に識別しモニタリングすることによって上げることができ、それにより、病気の人々はより効率的に治療に入り、より強い治療の使用が低減される。最後に、医師の実務は、個人のヘルスケア選択肢に自動的にリンクすることができる、新しい「因果事例」、すなわち、健康因子、疾患、及び傾向を示す測定された健康指標についての最新の情報をコンピュータデータベースに定期的に組み込むことを通して、より効率的且つ有効にすることができる。

30

#### 【0013】

図1は、本発明の一局面に従う健康学習システム10を示す。このシステムは、少なくとも1つの非一時的コンピュータ可読媒体のセットに記憶された機械実行可能命令であって、プロセッサ、専用ハードウェア、又は専用ハードウェア及びソフトウェア構成要素の組合せによって実行される機械実行可能命令として実現可能であると認識されよう。システム10は、患者の母集団のそれぞれの記録を格納した知識ベース12を含む。知識ベース12は、遺伝子マーカーを見つけるための個人の遺伝子マッピングを表すゲノムデータソース14及び生化学アッセイから導出される個人の様々な生化学パラメータのレベルを表す生化学データソース15のそれぞれから受信されたデータを含むことができる。本発明の一局面によれば、生化学アッセイは、健康な患者であっても使用可能な時系列生化学パラメータを提供することが促されるように、定期的な間隔で計画することができる。

40

50

## 【0014】

したがって、各記録は、患者に関連する、定期的な間隔で実行される生化学アッセイからとられる複数の生化学パラメータのそれぞれの値の時系列、複数の遺伝子マーカー、及び複数の臨床パラメータを含むことができる。複数の臨床パラメータは、例えば、電子健康記録データベースから抽出し、前の診断及び処置、臨床観測、長期バイオメトリックパラメータ（例えば、年齢、体重、血圧、体温、血糖値等）、及び家族の医療履歴を含むことができる。患者の母集団が、対象となる複数の病気のそれぞれについて、当該病気を有する患者の組及び当該病気を有さない患者の組を含むことができることが認識されよう。患者記録に加えて、知識ベース12は、様々な疾患の罹患率及び測定された結果と、利用可能なパラメータと医療研究から引き出された病気との因果リンクについてのデータとを表す統計を含むこともできる。一実装形態では、研究インタフェース（図示せず）を提供して、データを利用可能な医療研究から抽出することができ、研究インタフェースは、情報抽出構成要素を含み、ジャーナル記事等の非構造化研究ソースを知識ベースと互換性のあるテンプレートに変換する。

10

## 【0015】

ベースライン計算構成要素22は、所与の患者について、少なくとも、患者に関連する臨床パラメータ及びゲノムパラメータから、患者の生化学パラメータの予期時系列を計算するように構成される。システム10は、生化学アッセイに基づいて多数の病気について患者を並列して評価するが、あらゆる生化学パラメータがあらゆる状況及び患者に関連するわけではないことが認識されよう。したがって、ベースライン計算構成要素22は、利用可能な生化学パラメータの複数のサブセットのそれぞれの予期時系列を選択的に計算して、処理リソースを保存し得る。

20

## 【0016】

分析及びモデリング構成要素24は、計算された予期時系列からの値の時系列の偏差を特定し、偏差を入力として、複数の病気のうちの各病気に関連する1つ又は複数の予測モデルに適用するように構成される。各予測モデルは、当該病気を有する患者の組及び当該病気を有さない患者の組のそれぞれに関連する知識ベース12内のデータから導出することができる。例えば、予測モデルは、知識ベースからのデータでトレーニングされる、回帰モデル、人工神経ネットワーク、サポートベクターマシン、及び統計分類器等の適切な教師あり学習アルゴリズムを含むことができる。各予測モデルは、測定されたバイオメトリックパラメータと、偏差からのベースラインとの偏差に従って、複数の疾患のうちの1つの疾患の可能性を予測する。例えば、予測モデルは、距離尺度（例えば、ユークリッド、マハラノビス、マンハッタン）のうちの1つ又は複数に対して動作することができ、測定された時系列と予期時系列との差を予測特徴として使用することができる。代替的には、幾つかの最新データ点にわたる時系列の差を特徴として使用することができる。一般に、2つの時系列間の差を表す幾つかの記述統計学を計算可能であり、これらの任意の測定値が予測特徴として有用であり得ることが認識されよう。所与のモデルが、計算された偏差以外のパラメータも同様に含むことができ、これらの追加のパラメータを知識ベースから引き出すことができることが認識されよう。一実装形態では、予測モデリングの結果は、治療の行動指針及び測定された臨床結果で補足して、追加の因果事例を生成するに当たり使用するために、知識ベース12にフィードバックすることができる。

30

40

## 【0017】

一実装形態では、分析及びモデリング構成要素24はデータマイニング構成要素（図示せず）を含むことができ、データマイニング構成要素は、複数の教師なし学習アルゴリズムを知識ベース12に対して実行して、臨床パラメータ及びゲノムパラメータのうちの一方を病気に関連付ける少なくとも1つの因果事例を特定するように構成される。特定された因果事例は、主題専門家によって確認されると、既存の予測モデルの改良又は新しい予測モデルの生成に使用することができる。新たに生成された因果事例の機関によるレビューに向けて、分析及びモデリング構成要素は分析構成要素（図示せず）を含むこともでき、分析構成要素は、ユーザインタフェース26を通してユーザに提供され、知識ベース1

50



2 及び関連するデータベース（図示せず）からデータを検索するように構成される。主題専門家のガイド下で、分析構成要素は、知識ベース及びデータベースに対して様々な問い合わせを行い、所与の因果事例をサポート又は否定するエビデンスを提供することができる。一実装形態では、分析及びモデリング構成要素は規則エンジン（図示せず）も含み、規則エンジンは、関連する規則セットに従ってデータマイニング構成要素によって特定される因果事例を評価して、因果事例に関連するどの変数が、分析構成要素を用いて評価される場合に実行可能な結果を提供する最高の可能性を提示するかを特定する。関連すると信じられるパラメータに分析を制限することにより、この規則エンジンを使用して、処理リソースを保存するとともに、知識ベース 1 2 に格納されたデータの相互関係を特定するに当たり、誤検出の可能性を低減することができる。

10

#### 【0018】

ユーザインタフェース 2 6 は、患者が病気を有する、特定された可能性をユーザに提供するように構成される。ユーザインタフェース 2 6 は視覚化ツールを含むことができ、それにより、ユーザは、生化学パラメータ値の予期時系と、生化学パラメータ値の実際の時系列とのグラフィカル比較を見ることができる。一実装形態では、ユーザインタフェースは患者ダッシュボード（図示せず）を含み、患者ダッシュボードは、病気の特定された可能性、ヘルスケア治療行動指針、及び計画された次の生化学アッセイのそれぞれを伝達するように構成される。したがって、患者に、生化学分析に基づいて適切な時間にヘルスケアシステムに入るように指示することができる。患者ダッシュボードは、任意の診断された疾患及び推奨される治療選択肢についての情報へのリンクを含むこともできる。

20

#### 【0019】

ユーザインタフェース 2 6 は臨床判断サポート構成要素（図示せず）を含むこともでき、この構成要素は、患者が病気を有する、特定された可能性に基づいて、推奨されるケアプロトコルを臨床医に伝達するように構成される。ヘルスケアシステムでの全ての利害関係者に提供される知識ベース 1 2 及び予測モデルからデータが作られることにより、ユーザインタフェース 2 6 は、臨床医及び患者への推奨される行動指針の透明性を保証するとともに、研究者が知識ベースに格納されているデータに容易にアクセスして、新しい因果事例及び予測モデルを生成できるようにすることを保証することができる。

#### 【0020】

図 2 は、本発明の一局面に従うヘルスケア学習情報処理システム 5 0 の一実装形態を示す。図示の実装形態では、システム 5 0 は、各データインタフェース 6 2 ~ 6 5 を通して、破線の輪郭で示されるシステム外部の複数のデータソース 5 2 ~ 5 6 からデータを受信し、そのデータを処理して、これらのリソースからの蓄積データに基づいて推奨を患者、臨床医、及び研究者に提供する。第 1 のデータソース 5 2 は電子医療記録データベースを含み、各電子医療記録データベースは、例えば、前の診断及び処置、臨床観測、長期バイオメトリックパラメータ、及び家族の医療履歴を含む複数の患者の医療データを含む。情報処理システムと互換性を有することができる電子医療記録データベースの例としては、国軍保健長期的技術アプリケーション（A H L T A）、復員軍人健康情報システム及び技術アーキテクチャ（V I S T A）、及び大きな患者ベースを有する大きなヘルスケア組織によって保持される同様のそのようなデータベースを挙げることができる。これらのデータベースからの記録は、電子医療記録データベース（E M R D）インタフェース 6 2 を通して提供して、検索された記録を、ヘルスケア情報処理システム 5 0 に関連付けられた知識ベース 6 8 に適切なフォーマットに変換することができる。一実装形態では、電子医療記録データベースに格納された完全な記録は、インタフェースによって臨床的に関連する観測の組に切り詰められる。

30

40

#### 【0021】

データソースは、患者の大きな母集団からとられるバイオメトリックアッセイ 5 3 を含むこともできる。図示の実装形態では、プロテオームアッセイが利用されるが、薬理ゲノムアッセイ、メタボロミクスアッセイ、エピゲノムアッセイ、インタラクトームデータ、転写学データ、及びマイクロバイオームデータを含め、他のバイオメトリックアッセイを

50

利用することも可能なことが認識されよう。一実装形態では、プロテオームアッセイ 53 は、約 10 万のタンパク質を検出し、計画された間隔で施されて、10 万のタンパク質のそれぞれの血液レベルの時系列を提供することができる。アッセイインタフェース 63 は、知識ベース 68 用にアッセイデータをフォーマットし、アッセイの識別情報に知識ベース内の対応する患者記録を関連付けることができる。アッセイインタフェース 63 はまた、プロテオームデータを知識ベース 68 によって利用されるスケールに正規化することができる。一実装形態では、プロテオームアッセイは、知識ベース 68 に提供される臨床的に重要な特徴のベクトルに低減することができ、完全なアッセイは、圧縮され、患者ファイルから完全なアッセイへのタイムスタンプ付きラインと共に別個の大容量記憶装置に記憶される。

10

#### 【0022】

システム 50 は、患者の母集団からのゲノムデータ 54 を利用することもできる。例えば、ゲノミクスデータは、各患者について、適切なアッセイを介して捕捉され、ゲノミクスインタフェース 64 を通してシステムに提供することができる。ゲノミクスインタフェース 64 は、既知の遺伝子マーカーをゲノムから抽出し、抽出されたデータを知識ベース 68 用にフォーマットし、例えば、患者記録から抽出されたマーカーへのリンクを介して、遺伝子情報の識別情報に知識ベース内の対応する患者記録を関連付ける。

#### 【0023】

母集団健康データソース 55 からの情報及び統計は、健康データインタフェース 65 を通して提供することができる。母集団健康データソース 55 は、例えば、様々な疾患の罹患率及び測定結果を表す構造化又は半構造化データを含む。母集団健康データソース 55 の例としては、米国立がん研究所によって保持される監視疫学遠隔成績 (SEER) プログラム、米疾病対策センターによって保持される行動危険因子サーベイランスシステム (BRFSS)、米医療研究・品質調査機構によって保持される医療費及び医療資源利用プロジェクト (HCUP)、及び食品医薬局有害事象報告システム (FAERS) を挙げることができる。健康データインタフェース 65 は、これらのリソースで保持されている構造化及び半構造化データを、システム 50 に関連付けられた知識ベース 68 に適切なフォーマットに変換することができる。

20

#### 【0024】

最後に、様々な疾患の因果ファクタに関するデータを医療研究文献 56 から捕捉し、研究インタフェース 66 を通して知識ベース 68 に提供することができる。医療研究文献の例示的なソースは、米国立医学図書館からのメドライン (Medline) コレクション、パブメド (PubMed) コレクション、ジンバンク (GenBank) 配列データベース、及び全米バイオテクノロジー情報センターによって保持される遺伝子発現オムニバスリポジトリ、欧州バイオインフォマティクス研究所によって保持されるアレイエクспレス及びインタープロ (InterPro) データベース、米国立アレルギー感染症研究所によって保持されるインポート (Import) 免疫学データベース及び DAVID (Database for Annotation, Visualization, and Integrated Discovery)、及びユニプロット (UniProt) 知識ベース、並びにウィキペディア、WebMD、医療組織ウェブサイト、及び同様の情報ソース等のインターネット公開物を含むことができる。医療研究データ 56 は非構造化データを含むことができるため、研究インタフェース 66 は情報抽出構成要素を含むことができ、この構成要素は、ジャーナル記事等の非構造化研究ソースを知識ベース 68 と互換性のあるフォーマットに変換する。情報抽出構成要素は、非構造化ソースを個々の言葉又は語句に分解し様々な言葉又は語句の文脈及び意味を解釈し、抽出された情報を使用して、非構造化ソースを表すテンプレートを生成する。一実装形態では、生成されるテンプレートは、非構造化ソースに関連する分野の人間である専門家によってレビューして、知識ベース 68 に提供される情報が正確であることを保証することができる。

30

40

#### 【0025】

知識ベース 68 は、大規模並列システムとして実施して、低応答時間と、データ量の増

50

大に対して大きなスケーラビリティとを提供する。一実装形態では、知識ベース68は複数の地理的に遠い地域キャッシュを含むことができ、それにより、地方の臨床医は、所与の患者母集団に関連するデータに容易に素早くアクセス可能である。各キャッシュは、マスタ知識ベースに動作可能に接続されて、研究者は全体としてのデータを分析することができ、各キャッシュは、計画された予約に従ってマスタ知識ベースによって供給されるデータであることができる。緊急治療室及び他の非計画ケアソースからの要求を優先して、患者情報へのリアルタイム又は準リアルタイムアクセスを可能にする。キャッシュ内の情報は、最近最も使用されていないデータが置換されるように置換することができる。知識ベース68は、健康な母集団についてのデータ及び疾病症候群、アレルギー反応、又は何らかの他の望ましくない臨床結果を有する個人の母集団についてのデータの両方を含む様々な患者からの、少なくとも臨床観測、プロテオミクス、及びゲノミクスを格納する。知識ベース68は、知識ベース内のアクティブデータの混合、例えば、通知サブシステムによってサポートされるトリガー及びスケーラブル規則エンジンを使用する規則エンジンを含み得る。

10

20

30

40

50

#### 【0026】

本発明の一局面によれば、分析及びモデリング構成要素70は知識ベース68と対話して、データ間の関係を特定することができる。分析及びデータモデリング構成要素70の機能は、所与の患者についての様々な任意の病気の可能性が、患者に関連するデータをより大きな母集団からのデータと比較することによって予測することができる、本明細書では「フォワード分析」と呼ばれるものと、患者の大きな母集団からのデータがマイニングされて、臨床パラメータと識別された病気との関係を特定する「バックワード分析」とに大まかに分割することができる。

#### 【0027】

フォワード分析プロセスの一例では、ベースライン計算器72は、所与の患者について、プロテオームアッセイ53からの臨床的に関連するタンパク質のレベル等のバイオメトリックパラメータの予期される長期進行を計算するように構成することができる。一般に、ベースラインは、同様の状況の患者、すなわち、患者と同じ場所で生活又は仕事する患者、同様の遺伝子マーカを有する患者、同様の病歴を有する患者、又は他の様式で患者と共通する臨床的に関連するパラメータを有する患者のコホートに記録されたバイオメトリックパラメータの融合に従って特定される。ベースラインは1つ又は複数の統計モデルを介して計算することができ、統計モデルは、このデータを利用して、診断及び病気のみならず、体重、血圧、及び血糖値等の長期記録パラメータ、患者の遺伝子、並びに年齢及び居住地等の患者の経歴パラメータも含め、病歴を所与として、複数の臨床的に関連するバイオメトリックパラメータのそれぞれの患者に適切なレベル又はレベル範囲を特定する。

#### 【0028】

知識ベース68が多数の患者記録を含むことが予期されることが認識されよう。したがって、一実装形態では、タンパク質毎に、知識ベース68に単に問い合わせ、バイオメトリックパラメータの患者の値前後の定義された範囲内のそのタンパク質のベースラインを確立することに関連する全て又は閾値数のバイオメトリックパラメータを有する全て又は所定数の記録を返すことができる。タンパク質の時系列は、検索された全ての記録の平均をとり、ベースラインを提供することができる。

#### 【0029】

バイオメトリックパラメータのベースラインが計算されると、計算されたベースライン及びバイオメトリックパラメータの測定された複数の系列を一連の予測モデル73に提供することができる。予測モデル73は、測定されたバイオメトリックパラメータとベースラインとの偏差に従って、複数の疾患のうちの1つの病気の可能性を予測する、回帰モデル、人工神経ネットワーク、サポートベクターマシン、及び統計分類器等の任意の適切な教師あり学習アルゴリズムを含むことができる。一実装形態では、予測モデル73は、類推アルゴリズムを含むことができ、このアルゴリズムは、患者の測定されたバイオメトリックパラメータ、遺伝子マーカ、及び医師による臨床観測を、病気の有無が分かっている

他の患者からのバイオメトリックパラメータ、遺伝子データ、及び観測と比較して、その患者が病気を経験し得る可能性を特定する。予測モデル 73 によって評価される病気は、1 つ又は複数の疾患オントロジー 74 から引き出すことができる。疾患オントロジーは、国際疾病分類 (ICD)、精神疾患の分類と診断の手引 (DSM)、国際医薬用語集 (MeSH)、バイオオントロジー (BioOntology)、及びオープン生物学オントロジー等の既存のリソースから編纂することができる。

#### 【0030】

システムが厳密な疾患オントロジーに限定されないことが認識されよう。多くの病理学的状態が症状によって定義され、不正解な分類に繋がる。例えば、慢性疲労症候群が、異なる、恐らくは関連しない病理学のホストの包括的クラスであることが起こり得る。自閉症及び統合失調症等の他の疾患が、ある範囲の症状強度に沿って存在し、これもまた、異なる根本原因を有する状態をグループ化し得る。このため、システムは、これらの不正確な症状提示ではなく、基礎となる生物学的データによって病理学を定義する補完的方法を提供することができる。特に、生物学的データ (例えば、ゲノム、プロテオーム、メタボロミクス) の独自の組合せが統計学的に処理され、結果及び症状に関連付けられて、より精密な病理学的分類を提供する。生物学的状態を直接、病理学的分類にリンクすることにより、症状の根本となる生物学的原因に直接対処する治療を割り当てることができる。

10

#### 【0031】

システムによって実行されるバックワード分析は、前は結び付けられていなかった予測子間の結び付きについて、知識ベース 68 に格納されているデータを分析する 1 つ又は複数のデータマイニングアルゴリズム 76 を含むことができる。データマイニングアルゴリズム 76 によって特定される結び付きを利用して、システムによって実行されるフォワード分析で使用される新しい因果事例を定義することができる。このプロセスは完全に自動化することができ、新しい因果事例は予測モデル 73 に自動的に、又は半教師あり様式で統合され、半教師あり様式では、新たに発見された各因果事例は、主題専門家によってレビューされてから、予測モデルに組み込まれる。データマイニングアルゴリズム 76 は、例えば、異常検出アルゴリズム、アソシエーション規則学習、クラスタリングアルゴリズム、及びシーケンシャルパターンマイニングを含むことができる。

20

#### 【0032】

一実装形態では、新しい因果事例は、治療、タンパク質発現が変わる場合に生成され、次に、結果、相関、スコア付け、推奨、及び因果の重み付けのうちの何れかの調整として、知識ベースに繰り返し入力される。この情報は、研究者が仮説を評価できるようにし、新しいバイオマーカの識別等の後続研究を示唆する。システムが新しいデータを取り込み、処理するにつれて、分析及びデータマイニングアルゴリズムが自動的に実行されるため、興味を引く関係が出現する。研究者はログインして、更新された傾向リストと、出現した統計学的に有意な関係とを持ち出す。これらのリストは、研究者が関係の背後にある意味を探求し、将来の研究プロジェクトについての仮説を作成する機会として機能し、それにより、研究の生産性を加速化する。

30

#### 【0033】

システム 50 は分析構成要素 77 も含み、この構成要素は、データを知識ベース 68 から検索して、データマイニング構成要素 76 及び研究者によって識別された因果事例を確認するように構成される。このため、分析構成要素 77 は、所与の遺伝子配列とライブラリ配列との共通性を見つける基本ローカルアラインメントサーチツールと、ベースラインタンパク質アッセイと後の人生で診断された疾患との相関を自動的に発見し、ベースラインタンパク質アッセイと遺伝子配列との相関を自動的に発見し、ゲノムを疾患又はアレルギー反応に相関付けることにより、新しい遺伝子マーカを発見する様々なカスタム分析アルゴリズムとの統合を含むことができる。さらに、分析構成要素 77 は、臨床治療結果に関連するタンパク質レベル変化を追跡するアルゴリズムを含むことができ、それにより、タンパク質及び疾患との生物学的関係を探求し、遺伝子突然変異に関連付け因果生物学的相互作用の知識を用いてより効果的な薬剤を開発する。最後に、分析構成要素 77 は、統

40

50

計学的分析及び分析ツールを含み、データマイニング構成要素 7 6 及び他の分析ツールによって生成された仮説を確認するに当たり、研究者を支援することができる。一実装形態では、分析ツールは、高度信号処理アルゴリズムを含み、雑音の多いデータ及び神経スパイク尺度から相関を抽出することができる。

#### 【0034】

既知の副作用にも拘わらず、薬剤が処方されることが多い。本発明者らは、副作用を呈する可能性が最も高い人についての知識が、本発明によるヘルスケア学習情報処理システム 5 0 の能力内にあるとともに、特に代替の薬剤が存在する場合、かなりの価値があると判断した。同様に、これは、誰がよく、且つ／又は副作用なしで反応し得るかを予測することとなろう。このために、知識ベース 6 8 は、システム 5 0 からフィードバックされた結果データを収集するように設計される。肯定的で副作用がない結果は、特定の遺伝子突然変異又はベースラインタンパク質レベルに固有であり得、したがって、医師の治療推奨をサポートする追加の情報として機能するとともに、研究及び発見の分野を示唆することができる。したがって、結果は、個人の患者の特定の遺伝子突然変異及びタンパク質レベルにリンクすることができ、それにより、プロテオミクス及びゲノミクスからの患者の反応を予測することができる。

10

#### 【0035】

システムが、有意な相関について数百人～数千人の個人を繰り返しテストし得ることが認識されよう。含まれる変数が多いほど、洞察があり使用可能な関係を発見する確率が高くなるが、偽陽性の確率も増大する。「複数比較」というこの問題を補正する標準的な方法は、有意テスト値を何らかの補正係数で乗算することである。例えば、ボンフェローニ補正では、p 値は、実行される独立テストの数で乗算される。不都合なことに、これは偽陰性の確率を増大させる。したがって、実行される独立有意性テストが多くなるほど、非有意性の背景雑音に埋もれることになる関心を引く関係が多くなる。

20

#### 【0036】

本発明の一局面によれば、規則エンジン 7 8 は、専門家生成規則と機械生成規則との混合を含み、分析前の洞察力に富むか、又は実施可能な結果について、どのタイプの変数が最良の確率を提示するかを学習する重みは、連続して展開され調整される。自動規則エンジン 7 8 は、1 つの研究実験前にどのテストを実行すべきかを判断するに当たり、専門研究者の尽力を補足することが予期される。テストの全体数を低減することにより、処理性能も最適化される。最終的に、規則エンジン 7 8 は、ヘルスケアに基づく統計規則を作成するに当たり、統計設計と機械インテリジェンスとを仲介する。

30

#### 【0037】

様々な分析及びモデリングプロセス 7 0 の結果は、知識ベース 6 8 に提供して、患者の記録に追加されるとともに、任意の関連する医療データベース 5 2 に提供することができる。これらの記録は一般に、これらのファクタが分かるようになると、治療記録及び患者結果で補足される。結果は、各視覚化構成要素 8 2 ～ 8 4 にも提供される。一実装形態では、研究者視覚化構成要素 8 2 は、このシステムで収集された遺伝データ及びプロテオームデータに適用される分析検索エンジン 7 7 によって発見された知識を、容易に理解可能な視覚的様式で提示する。研究者視覚化構成要素 8 2 は、分析検索アルゴリズムのユーザインタフェースを提供して、タンパク質アッセイと、遺伝子配列と、診断との間の相関を発見することができる。研究者視覚化構成要素 8 2 は、特定の治療結果に関連するタンパク質レベル変化を表示する様々な表示及びグラフィカル操作ツールを含むこともでき、それにより、研究者は、タンパク質及び疾患への生物学的関係を探求し、結果及びタンパク質を遺伝子突然変異に関連付け、因果の生物学的相互作用の知識を用いてより有効な薬剤を開発することができる。研究者視覚化構成要素 8 2 は、結果データがシステムにフィードバックされるため、データベースにわたる変数間に出現した統計学的関連の定期的なレポートを提供することもできるとともに、単純に、価値のある科学的データベースからの関連するデータ及び発見にアクセスすることができる。

40

#### 【0038】

50

臨床医判断サポート構成要素 8 3 により、臨床医は、所与の患者及び関連するサポート情報についてのフォワード分析プロセスの結果にアクセスすることができる。例えば、臨床医判断サポート構成要素 8 3 は、臨床医に、患者の臨床観測、最新のタンパク質アッセイ、地理的位置、及び関連する環境要因と一致する疾病のリストを可能性順に表示することができる。臨床医は、現在のタンパク質アッセイと測定されたか、又は帰属するベースラインアッセイとの比較及び／又はタンパク質アッセイの患者の履歴と予期されるタンパク質アッセイの通常の時系列との比較を表示するように、臨床医判断サポート構成要素 8 3 に指示することもできる。臨床医判断サポート構成要素 8 3 は、患者のゲノム及び性別、体重、及び年齢等の外因性変数からのマーカー等の、計算されたベースラインアッセイでの有意な値を表示することもできる。判断サポート構成要素 8 3 は、患者が所定の時間期間にわたりオフィスとコンタクトしていないとき、又は計画された生化学アッセイを提供していないときを臨床医に通知することもできる。一実装形態では、この通知は、臨床医に過度の負担をかけないように、ある時間期間にわたって編纂して、リスト形態で提供することができる。

10

#### 【0039】

患者ダッシュボード 8 4 は、フォワード分析プロセスの結果及びサポートデータを患者に提示することができる。このために、患者には、リスク上昇の任意の発見、発見をサポートしたゲノムパラメータ、生化学的パラメータ、及び臨床的パラメータ、並びにリスク上昇に関連する疾患又は結果、潜在的な治療、及び発見をサポートするパラメータに関連する情報へのリンクを提示することができる。例えば、患者には、処方された薬剤に関連する副作用についての情報へのリンクを提供することができる。健康スクリーニング結果についての任意の推奨及び患者に提供される潜在的な行動指針は、確実性加重且つリスクベースの加重を含み、患者による情報を受けた上での判断に役立つことができる。ダッシュボード 8 4 は、S/MIME 等の暗号化電子メールサービスを介して、患者が臨床医に質問して、前に訪問したときに受け取った情報をはっきりさせるためのインタフェースを提供することもできる。ダッシュボード 8 4 は、計画された生化学アッセイ、臨床医との予約、又は薬剤の摂取若しくは補充についてのリマインダーを患者に提供することもできる。一実装形態では、患者は、ダッシュボード 8 4 を通して症状の観察を記録することができ、同様に、患者の電子医療記録でのデータのレビュー、修正、及び補足を行うこともできる。

20

30

#### 【0040】

所与の患者の医療結果が分かった後、新しい結果を反映するように、知識ベース 6 8 を更新することができる。このために、測定臨床結果の組 8 6 を知識ベース 6 8 に提供して、既存の患者データを補強することができる。測定臨床結果は、例えば、予測後、設定された時間期間後に、対象となる病気を患者が有するか否かを反映することができる。システムに入力される新しい医療研究及び新しい患者記録と共に、これらの患者結果 8 6 は、知識ベース 6 8 に、分析及びモデリング構成要素 7 0 によって発見される新しい因果事例のベースを提供することができる。

#### 【0041】

一使用事例では、研究所が患者の血液を採取し、ゲノムアッセイ 5 4 及びプロテオームアッセイ 5 3 を提供する。一実装形態では、プロテオームアッセイ 5 4 は、各テスト場所でのオーバーヘッドが比較的低く、少量の血液試料から数千ものタンパク質のレベルを同時に特定することができる低コストの容易に繰り返し可能なアッセイを使用して実行することができる。テストを広くアクセス可能なものにすることができる。テストは、低コストでアクセス可能なように設計されるため、複数の個人の大きな集団の長期データを効率的に編纂することができる。データが正規化され処理されると、臨床的観測に鑑みて解釈される患者のタンパク質レベル及び遺伝子マーカーが、予測モデル 7 3 を通して所与の病気の可能性の増大を示すか否かを判断することができる。この使用事例では、患者が、高リスクの特定のタイプの癌に関連する遺伝子マーカーと、そのタイプの癌に関連するタンパク質の上昇とを有すると判断される。知識ベース 6 8 は、このタイプの癌の生存率が、3 ヶ月

40

50

以内に診断される場合に有意に高いことを示す情報を含むことができる。

#### 【0042】

癌の高リスクが識別されると、レポートが生成され、患者に通知される。患者は、患者ダッシュボード84にログインして、レポートを見ることができ、レポートは、診断と、疾病、高リスクの識別に使用されたプロテオーム及び遺伝子データ、及び潜在的な治療についての情報へのリンクとを含むことができる。レポートは、患者が癌専門医への訪問を計画すべきことの推奨を含むこともできる。同様に、かかりつけ医及び／又は患者を治療する癌専門医等の患者に関連する臨床医は、臨床医判断サポート構成要素83を通してアラートを受信することができる。アラートは、診断に関連する全体リスクスコアと、診断の依拠となった特定の遺伝子マーカ及びタンパク質を、関連する研究へのリンクと共に含むとともに、このデータを表示する視覚化ツールを含むサマリレポートにリンクすることができる。臨床医の治療判断及び臨床結果は、経過観察での訪問からの情報と、患者及び臨床医からのコメントと共に、知識ベース68にフィードバックすることができる。次に、これらの発見は、更なる分析のために、研究者視覚化構成要素82を通して利用可能な様々なツールを通して研究者に提供することができる。

10

#### 【0043】

第2の使用事例では、研究者は、最近出現したデータ傾向を示すサマリレポートを見て、食物アレルギーに関連するタンパク質が上昇した患者について、高い率の非有害プラミペキソール反応を発見し得る。次に、研究者は、プラミペキソールによく反応する患者によって共有される遺伝子突然変異とタンパク質上昇との既知の関係について、研究者視覚化構成要素82のテキストマイナーを介して、利用可能なジャーナル記事内のテキストを検索するとともに、知識ベース内のデータ及び関連のあるデータソース内のデータを検索する。既知の関係が見つからないと仮定すると、研究者は、任意の識別されたタンパク質が線維筋痛を有する幾人かの患者で上昇し、遺伝子突然変異を有する患者でアレルギーへの感受性を増大させるという仮説を用いて、これもまた上昇し得る未確認タンパク質を探すためのテストを開発し実行することができる。

20

#### 【0044】

研究者は、研究の結果及び特定された仮説を知識ベースに提供し、プロテオミクス研究所に識別されたタンパク質のアプタマーを開発するように要求することができる。アプタマーが生成されると、計画されたプロテオームアッセイを受けている複数の患者からの結果を集計して、研究者の仮説を確認又は否定することができる。仮性をサポートするか、又は否定するエビデンスを提供するために、知識ベース68からの他の情報をマイニング又はクエリしてもよいことが認識されよう。仮説が確認されると仮定すると、例えば、研究者視覚化構成要素82を通しての知識ベース68のクエリを通して、更なる研究を行い、このタンパク質のレベルを低減するために利用することができる薬剤を見つけることができる。次に、この発見は、薬剤と線維筋痛との既知の関係として知識ベース68にフィードバックすることができる。

30

#### 【0045】

これが全て行われた後、線維筋痛と診断された患者は、一般的な薬剤への反応がよくなると臨床医によって判断され得る。臨床医は、ドーパミン作動薬を処方したいことがあるが、有効性及び副作用について懸念する。臨床医は、ゲノムアッセイ若しくはプロテオームアッセイ用に血液を採取するように患者に指示するか、又は患者の計画されたアッセイからの既存のゲノムデータ及びプロテオームデータを利用し得る。この情報から、患者が、ドーパミン作動薬プラミペキソールに反応する患者に関連する遺伝子突然変異を共有するが、ドーパミン作動薬ロピニロールによく反応する患者に関連するマーカを有さないことを特定し得る。アレルギーへの感受性の増大に関連するタンパク質も、患者で上昇していることが見つけられることもある。グルテンを食事からなくした事例の60%で、タンパク質発現レベルが低減したことを示す知識ベース内の情報を自動的に検索して、臨床医及び患者に提供することができる。

40

#### 【0046】

50

この情報は全て、複数の治療選択肢と共に、臨床医判断サポート構成要素 8 3 において臨床医に提供することができ、各治療選択肢には、予測モデル 7 3 から生成される、治療が好ましい臨床結果に繋がる可能性を表すスコアが関連付けられる。2 つの高スコア治療は、患者にグルテンフリーの食事をさせるとともに、プラミペキソールを処方することを含み得る。したがって、臨床医は、グルテンフリーの食事の任意の恩益効果を考慮して、プラミペキソールの投与量を低減し、何れかの選択肢を選択するか、又は選択肢を組み合わせ得る。プラミペキソールが処方される限り、副作用に関連するタンパク質のレベルを追跡することができ、例えば、薬剤の効果が明確になるまで、患者のプロテオームアッセイの頻度を増大する。

#### 【0047】

患者には、診断と、臨床医によって行われた治療判断と、予約スケジュールとを有するサマリリポートを提供することができる。このリポートは、線維筋痛、薬剤に関連する副作用及び相互作用、並びにグルテンフリーダイエットを推進するためのアドバイスを説明するオンラインリソース等の診断及び治療に関連する情報へのリンクを含むことができる。診断に使用される遺伝子マーカー及びタンパク質レベルについての情報を提供することもできる。患者は、患者ダッシュボード 8 4 を使用して、痛み及び疲労感等の症状レベルを経時記録することができる。さらに、関連するタンパク質のレベルを経時追跡して、進歩への患者の気付きを維持し、恐らくはコンプライアンスを促進することができる。患者の記録した症状及び臨床医の観測は、知識ベース 6 8 にフィードバックして、選択された治療の有効性及び任意の副作用の発生率を評価するに当たり使用することができる。

#### 【0048】

図示のシステム 5 0 は、幾つかの利点を提供する。例えば、システムは、単一の血液試料から多くの因果事例をテストすることにより、規模の経済を可能にする。システムは、生体センサのばらつき及び任意の他の潜在的な誤差源を含め、測定値及び推奨確実度の定量化、集計、及び開示を行い、推奨に関連する信頼度が患者及び臨床医にとって有意味であることを保証することができ、システムは、時間の経過に伴って推奨の正確性を改善することができる。その結果、システムは、ミッションクリティカルな使用をサポートし、利用可能な因果事例及び新しいターゲット母集団の包含の両方において、時間の経過と共に予期されるデータの増大に伴ってスケーリングするのに十分に高い信頼性、能力、及び可用性を有することができる。

#### 【0049】

図 1 及び図 2 において上述した上記構造的及び機能的特徴に鑑みて、図 3 及び図 4 を参照して方法の一例がよりよく理解される。説明を簡単にするために、図 3 及び図 4 の方法は、順次実行されるものとして示され説明されるが、他の例では、幾つかの動作を本明細書に示され説明される順序とは異なる順序で、且つ／又は同時に行うことが可能であるため、本発明が図示の順序によって限定されないことを理解され認識されたい。

#### 【0050】

図 3 は、本発明の一局面に従う個人化されたヘルスケアサポートを提供する方法 1 0 0 を示す。1 0 2 では、個人から採取された血液試料に対する生化学アッセイが、計画された間隔で行われ、複数の生化学パラメータのそれぞれの値の時系列を生成する。一実装形態では、生化学アッセイは、アッセイが低コストであり、臨床環境外で容易に実行可能なように、1 滴の血液から多数のタンパク質レベルを測定するベースラインタンパク質アッセイである。したがって、生化学アッセイへの患者アクセスを好都合にして、完全な値の時系列を生成するに当たりコンプライアンスを促進することができる。

#### 【0051】

1 0 4 では、知識ベースから、個人に関連する複数の臨床パラメータが抽出される。パラメータは、診断された疾患又は症状の臨床的観測等のカテゴリー並びに年齢、体温、体重、血圧、コレステロールレベル、及び他のそのようなデータ等の間隔又は比率データであることができる。一実装形態では、複数のコホートパラメータは、その個人に関連する複数の個人を表す記録からの知識ベース内の対応する複数の生化学アッセイ系列から抽出

10

20

30

40

50



することができる。例えば、コホートパラメータは、患者に関連する人々の集合、患者の近くで生活するか、又は仕事する人々の集合、及び患者と共通する病気又は遺伝子マーカーを共有する人々の集合のうちの1つ又は複数にわたる所与の生化学パラメータの平均時系列を含むことができる。

#### 【0052】

106では、個人の複数のゲノムパラメータが特定される。一実装形態では、これは、生化学パラメータの導出に使用されるものと同じ血液試料から行うことができる。知識ベースが患者の母集団の生化学アッセイ、ゲノムパラメータ、及び臨床パラメータを含むように、値の各時系列及び複数のゲノムパラメータを知識ベースに格納することができることが理解されよう。

10

#### 【0053】

通常、化学的分析及び生物学的分析を使用して、生物学的試料の特有の特徴を特定する。次に、特徴を代表的な定量値に変換し、情報処理システムに提供して、データマイニング、機械学習、及び他の計算機能を含む計算及び統計分析を行うことができる。生化学試料から導出される特有の特徴を特定し、特徴を他の試料と突き合わせて、又は参照データセットにわたって比較する多くの方法が、生化学の分野の当業者には既知である。例えば、異なる生物学的試料からの複数の質量スペクトルを比較し、複数の試料にわたる共通特徴を識別することは、参照状態として使用することができ、一方、特有の特徴を識別することは、異常な状態を検出するための潜在的なバイオマーカーとして機能することができる。特徴は、複数の個人にわたり、及び／又は時間的に特定の個人に関して比較することができる。本明細書に記載のように、様々なタイプの生化学パラメータが既知であり、分析に使用可能である。本発明は、異なるタイプの複数の生化学アッセイを組み合わせ、特有の特徴に時間的成分を包含することにより、大幅に改善された生化学特有の特徴を生成する。

20

#### 【0054】

108では、複数の生化学パラメータの複数のサブセットのそれぞれについて、予期時系列が、少なくとも臨床パラメータ及びゲノムパラメータから計算される。例えば、予期時系列は、患者の臨床パラメータ及びゲノムパラメータに関連する様々な特徴を有する患者からの値の時系列の加重結合として特定することができ、重みは類似性に応じて選択され、類似性は、例えば、患者と知識ベース内の様々な他の患者との多変量距離尺度として特定される。代替的には、関連する生化学パラメータに関して、患者の値から予め定義される範囲内の値を有する患者について、知識ベースに問い合わせることができる。予期時系列は、検索された記録の非加重平均（例えば、平均又はメジアン）であってもよい。

30

#### 【0055】

一実施形態では、本発明は、まず、生化学アッセイを特徴ベクトルとして表現することにより、予期時系列を計算し、各特徴ベクトルは、生化学パラメータの組に対応する複数の係数を有する。次に、本発明は、特定の病気を有する患者の大きな母集団から導出される病理学的特徴ベクトルを含む複数の組のクラスタを生成する。特定の各クラスタの特徴ベクトルメンバは、特徴ベクトルとクラスタ重心との間のユークリッド距離計算によって測定される特有類似性を有する。同様に、K平均クラスターリングアルゴリズム等の周知の教師なしクラスターリング法を使用することもできる。更に別の代替は、マハラノビス距離を類似性（相関）の測定に使用することであり、これには、一般的に規模不変であるという利点がある。さらに、生化学アッセイに関連するデータセット及び特徴ベクトルの組み合わせは、多変量ベクトル又は行列として多次元で表現することができ、クラスターリング及び距離の計算は、生化学アッセイ特徴ベクトル集合にわたる多変量ベクトル又は行列を融合し相関付けることによって実行することができる。統計分析の分野の当業者には既知のより多くの距離測定及び特徴ベクトルタイプがある。本明細書に記載される実施形態は、例としてのみ示され、一般性を失わずに、様々な代替を使用可能なことが理解される。

40

#### 【0056】

一連のクラスタ重心が経時追跡され、クラスタ遷移パスによって特徴付けられる時間的

50

測定がここで導入される。個々の患者の生化学アッセイの値の時系列は、新しい血液試料が採取される都度（例えば、1年毎）、関連する特徴ベクトルの最近傍クラスタへの距離を計算することにより、予期時系列と比較することができる。計算への強化として、大きなバイオインフォマティクスデータセット内に沢山あり、システムの結果／価値に実質的に寄与しない不必要な特徴は、除去することができ、それにより、結果を改善する。データ分析の分野の当業者に周知の教師あり機械学習及びデータマイニングを実行するために、多くの他の方法が利用可能である。

#### 【0057】

110では、生化学パラメータの複数のサブセットのそれぞれについて、個人を表す値の時系列は、計算された予期時系列と比較されて、個人の複数の病気のそれぞれの可能性が特定される。例えば、計算された予期時系列からの値の時系列の有意な偏差を特定し、複数の病気のうちの1つに関連する予測モデルに入力として適用することができ、予測モデルは、有意な偏差から導出される少なくとも1つのパラメータから、複数の病気のうちの1つの病気の可能性を特定するように構成される。一実装形態では、予測モデルは、主題専門家によって仲介される教師なし学習プロセスによって生成し、改良することができる。例えば、データマイニングアルゴリズムを知識ベースに適用して、臨床パラメータ、ゲノムパラメータ、及びコホートパラメータのうちの1つを病気に関連付ける少なくとも1つの因果事例を識別することができる。因果事例が、例えば、1つ又は複数の分析ツールを適用して、知識ベースからエビデンスを検索することにより、主題専門家によってレビューされ、検証されると、識別された因果事例に従って予測モデルを改良するか、又は生成することができる。

10

20

#### 【0058】

112では、複数の病気のうちの少なくとも1つの病気の可能性がユーザに伝達される。一実装形態では、ユーザは個人であり、伝達は、伝達された少なくとも1つの病気の可能性に基づくヘルスケア治療行動指針、伝達された少なくとも1つの病気の可能性に基づいて次の生化学アッセイを計画すべきときの個人への指示、及び個人が治療を探るべき医療関係者のタイプについての推奨のうちの任意又は全てを含むことができる。別の実施形態では、ユーザは臨床医であり、伝達は、伝達された少なくとも1つの病気の可能性に基づく臨床医に推奨されるケアプロトコールを含む。

30

#### 【0059】

一実装形態では、伝達は、生化学パラメータの複数のサブセットのうちの選択される1つについて、選択された生化学パラメータについて個人を表す時系列と、選択された生化学パラメータの計算された予期時系列とのそれぞれのグラフィカル表現をユーザに表示するように構成されるユーザインタフェースを通して行われ、それにより、計算された予期時系列を、計画された生化学アッセイからの測定値と容易に比較することができる。ユーザインタフェースにより、臨床医は、予期時系列の計算に使用されたパラメータのうちの選択される1つから新しい値を選択し、選択されたパラメータの新しい値を反映するように、予期時系列のグラフィカル表現を変更することが可能であり得る。これにより、臨床医は、可能な治療及び生活習慣の変更の、患者の健康への効果を特定可能であり得る。これらのツールが、新しい因果事例を探すことにおける支援のために、研究者に提供することもできることが更に認識されよう。

40

#### 【0060】

図4は、本発明の一局面に従うヘルスケア学習システムにおいて新しい因果事例を発見し適用する方法150を示す。152では、ヘルスケア学習システムに関連する知識ベースが、知識ベース内の患者の測定された臨床結果で更新することができる。例えば、測定された結果は、ユーザインタフェースを介してシステムに直接入力するか、又は医療記録データベースから検索することができる。154では、教師なし学習プロセスを知識ベースに対して実行して、潜在的な因果事例を発見する。教師なし学習プロセスは、例えば、異常検出アルゴリズム、アソシエーション規則学習、クラスタリングアルゴリズム、及びシーケンシャルパターンマイニングを含むことができる。

50

## 【0061】

156では、分析者は、1つ又は複数の分析を知識ベースに対して実行して、潜在的な因果事例を確認するように促される。例えば、研究者に、テキスト又はグラフィカル表現としてレビューに利用可能な適切なサポートデータと共に、最近出現したデータ傾向を示すサマリレポートを提供し得る。次に、研究者は、テキストマイナーを介して利用可能なジャーナル記事内のテキストを検索するか、又は関心があると分かった任意の出現傾向についての仮説を立てるために、知識ベース内の関連データの1つ又は複数のクエリを考案することができる。次に、研究者は、テストを開発して実行し、仮説を確認することができ、研究の結果及び特定された仮説は知識ベースに供給される。因果事例を表す仮説が確認される場合、1つ又は複数の予測モデルは、158では、新しい発見を反映するように更新される。

10

## 【0062】

図5は、図1及び図2に示される健康学習システム等の、図1～図4に開示されるシステム及び方法の例を実施可能なハードウェア構成要素の例示的なシステム200を示す概略ブロック図である。システム200は、様々なシステム及びサブシステムを含むことができる。システム200は、パーソナルコンピュータ、ラップトップコンピュータ、モバイル装置、タブレットコンピュータ、ワークステーション、コンピュータシステム、アプリケーション、特定用途向け集積回路（ASIC）、サーバ、サーバブレードセンタ、サーバファーム等であることができる。

20

## 【0063】

システム200は、システムバス202、処理ユニット204、システムメモリ206、メモリ装置208及び210、通信インタフェース212（例えば、ネットワークインタフェース）、通信リンク214、ディスプレイ216（例えば、ビデオ画面）、並びに入力装置218（例えば、キーボード及び／又はマウス）を含むことができる。システムバス202は、処理ユニット204及びシステムメモリ206と通信することができる。ハードディスクドライブ、サーバ、スタンドアロンデータベース、又は他の不揮発性メモリ等の追加のメモリ装置208及び210も、システムバス202と通信することができる。システムバス202は、処理ユニット204、メモリ装置206～210、通信インタフェース212、ディスプレイ216、及び入力装置218を相互接続する。幾つかの例では、システムバス202は、ユニバーサルシリアルバス（USB）ポート等の追加のポート（図示せず）も相互接続する。

30

## 【0064】

処理ユニット204は、計算装置であることができ、特定用途向け集積回路（ASIC）を含むことができる。処理ユニット204は、命令セットを実行して、本明細書に開示される例の動作を実施する。処理ユニットは処理コアを含むことができる。

## 【0065】

追加のメモリ装置206、208、及び210は、データ、プログラム、命令、テキスト又はコンパイルされた形態でのデータベースクエリ、及びコンピュータの動作に必要であり得る任意の他の情報を記憶することができる。メモリ206、208、及び210は、メモリカード、ディスクドライブ、コンパクトディスク（CD）、又はネットワークを介してアクセス可能なサーバ等のコンピュータ可読媒体（一体化されるか、又はリムーバブル）として実施することができる。特定の例では、メモリ206、208、及び210は、テキスト、イメージ、ビデオ、及び／又はオーディオを含むことができ、それらの部分は、人間が理解可能なフォーマットで利用可能であり得る。

40

## 【0066】

追加又は代替として、システム200は、通信インタフェース212を通して外部データソース又はクエリソースにアクセスすることができ、外部データソース又はクエリソースは、システムバス202及び通信リンク214と通信することができる。

## 【0067】

動作に当たり、システム200を使用して、本発明による健康学習システムの1つ又は

50

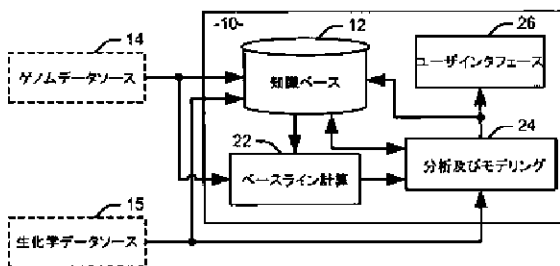
複数の部分を実施することができる。複合アプリケーションテストシステムを実施するコンピュータ実行可能論理は、特定の例により、システムメモリ 206 及びメモリ装置 208、210 のうちの 1 つ又は複数に存在する。処理ユニット 204 は、システムメモリ 206 並びにメモリ装置 208 及び 210 に端を発する 1 つ又は複数のコンピュータ実行可能命令を実行する。本明細書で使用される場合、「コンピュータ可読媒体」という用語は、命令を処理ユニット 204 に提供して実行させることに関与する媒体を指す。

#### 【0068】

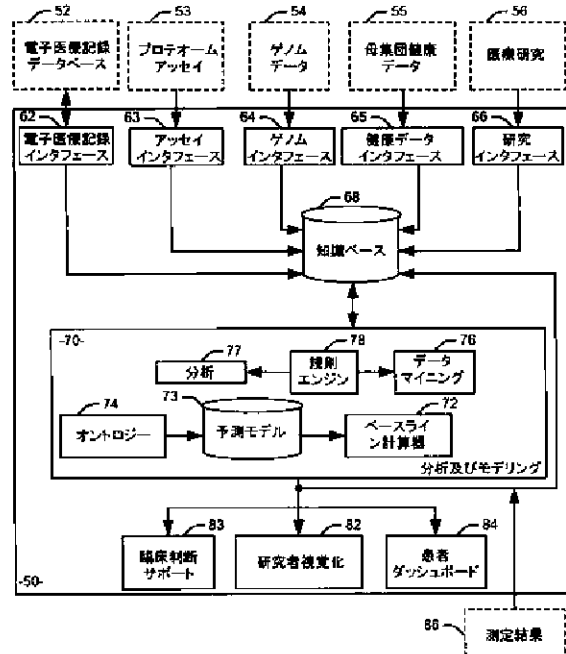
上述したものは、本発明の例である。当然、本発明を説明するために、構成要素及び方法論の考えられるあらゆる組合せを説明することは不可能であり、本発明の多くの更なる組合せ及び順序が可能なことを当業者は認識しよう。したがって、本発明は、添付の特許請求の範囲内にある全てのそのような代替形態、変更形態、及び変形形態を包含することが意図される。

10

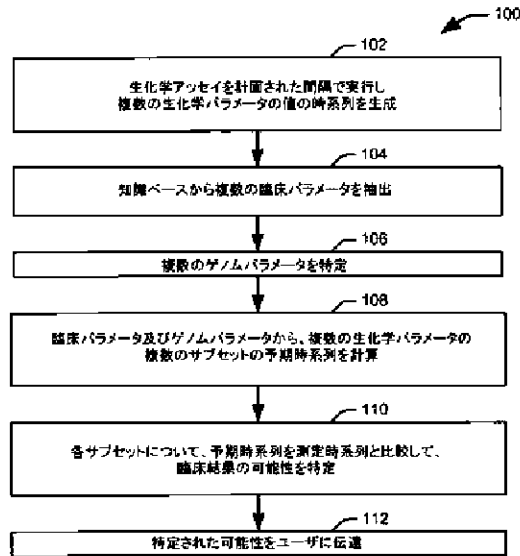
【図 1】



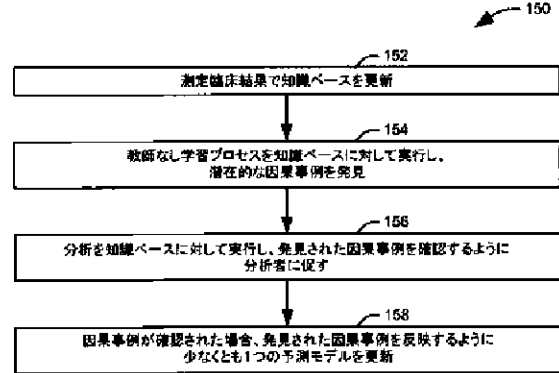
【図 2】



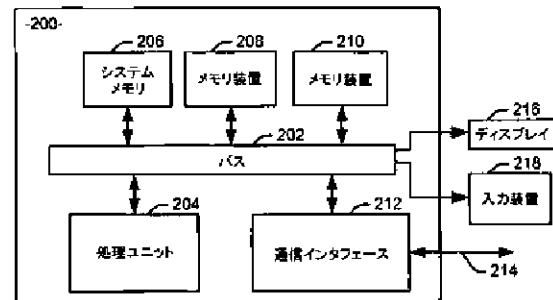
【図 3】



【図 4】



【図 5】



## 【手続補正書】

【提出日】平成27年8月28日(2015.8.28)

## 【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

個人化されたヘルスケアサポートを提供する方法であって、

個人から採取された血液試料に対する生化学アッセイを計画された間隔で実行して、複数の生化学パラメータのそれぞれについて、前記個人を表す値の時系列を生成し、

前記個人に関連する複数の臨床パラメータを知識ベースから抽出し、

前記個人の複数のゲノムパラメータを特定し、

少なくとも前記複数の臨床パラメータ及び前記複数のゲノムパラメータから、前記複数の生化学パラメータの複数のサブセットのそれぞれの予期時系列を計算し、

前記生化学パラメータの複数のサブセットのそれぞれについて、前記個人を表す値の時系列を、計算された前記予期時系列と比較して、前記個人についての複数の病気のそれぞれの可能性を特定し、

ヘルスケア行動指針を作成できるように前記複数の病気のうちの少なくとも1つの病気の可能性を出力すること

を備える方法。

【請求項2】

前記複数の病気のうちの少なくとも1つの病気の可能性を出力することは、前記複数の病気のうちの少なくとも1つの病気の可能性を、知識ベース、データベース、又はディスプ

レイのうちの少なくとも1つに供給することを備え、出力された前記可能性は、予測モデルを更新するための少なくとも1つの教師なし学習プロセスに利用可能である、請求項1に記載の方法。

【請求項3】

前記個人に関連する複数の個人をそれぞれ表す複数の生化学アッセイ系列から、複数のコホートパラメータを抽出することを更に備え、前記複数の生化学パラメータの前記複数のサブセットのそれぞれの予期時系列は、少なくとも前記臨床パラメータ、前記ゲノムパラメータ、及び前記コホートパラメータから計算される、請求項1に記載の方法。

【請求項4】

前記ユーザは臨床医であり、前記方法は更に、前記予期時系列の計算に使用された前記臨床パラメータ及び前記ゲノムパラメータのうちの選択された一つのパラメータから、新しい値を前記臨床医が選択できるようにし、前記臨床パラメータ、前記ゲノムパラメータ、及び前記コホートパラメータのうちの前記選択された一つのパラメータの前記新しい値を反映するように、前記予期時系列の前記グラフィック表現を変更することを備える、請求項1に記載の方法。

【請求項5】

前記知識ベースが複数の患者からなる母集団の生化学アッセイ、ゲノムパラメータ、臨床パラメータ、及び測定された臨床結果を有するように、前記値の時系列、前記複数のゲノムパラメータ、及び測定された臨床結果のそれぞれを前記知識ベースに格納することを更に備える、請求項1に記載の方法。

【請求項6】

前記個人を表す値の時系列を計算された前記予期時系列と比較して前記個人の複数の病気のそれぞれの可能性を特定することは、

計算された前記予期時系列からの前記値の時系列の偏差を特定し、

前記複数の病気のうちの1つに関連する予測モデルであって、前記偏差から導出される少なくとも1つのパラメータから、前記複数の病気のうちの前記関連する1つの病気の可能性を特定するように構成されている前記予測モデルに前記偏差を入力として適用すること

を備える、請求項5に記載の方法。

【請求項7】

前記知識ベースにデータマイニングアルゴリズムを適用して、前記臨床パラメータ、前記ゲノムパラメータ、及び前記コホートパラメータのうちの1つを病気に関連付ける少なくとも1つの因果事例を識別し、

識別された前記因果事例に従って予測モデルを改良することを更に備える、請求項6に記載の方法。

【請求項8】

プロセッサに動作可能に接続された少なくとも1つの非一時的コンピュータ可読媒体のセットに記憶された機械実行可能命令として実現される個人化ヘルスケアシステムであって、

知識ベースであって、対象となる複数の病気のそれぞれについて、当該病気を有する患者の組及び当該病気を有さない患者の組を含む複数の患者からなる母集団の記録と、一患者に関連するとともに計画された間隔で実行された生化学アッセイから取得した複数の生化学パラメータの値の時系列、複数の遺伝子マーカー、及び複数の臨床パラメータを含む所与の記録とを備える知識ベースと、

所与の患者について、少なくとも、当該患者に関連する前記臨床パラメータ及び前記ゲノムパラメータから、前記生化学パラメータの複数のサブセットのそれぞれの予期時系列を計算するように構成されたベースライン計算構成要素と、

計算された前記予期時系列からの前記値の時系列の偏差を特定し、前記複数の病気のうちの1つの病気に関連する予測モデルであって、当該病気を有する患者の組及び当該病気を有さない患者の組のそれぞれに関連するデータから導出されたものであり、かつ、前記

偏差から導出された少なくとも1つのパラメータから前記患者が前記病気を有する可能性を特定するように構成された前記予測モデルに前記偏差を入力として適用するように構成された分析及びモデリング構成要素と、

特定された前記患者が前記病気を有する可能性をユーザに提供するように構成されたユーザインタフェースと  
を備えるシステム。

【請求項9】

前記知識ベースには、所定の時間期間後に前記患者が前記病気を有するか否かを少なくとも反映した測定臨床結果が提供され、前記分析及びモデリング構成要素は、複数の教師なし学習アルゴリズムを前記知識ベースに対して実行して、前記臨床パラメータ及び前記ゲノムパラメータのうち的一方を前記病気に関連付ける少なくとも1つの因果事例を特定するように構成されたデータマイニング構成要素を備える、請求項8に記載のシステム。

【請求項10】

前記分析及びモデリング構成要素は、前記ユーザが前記ユーザインタフェースを通して利用可能な分析構成要素であって、前記知識ベース及び関連データベースからデータを検索して、前記データマイニング構成要素によって特定される前記少なくとも1つの因果事例をサポートするエビデンスを提供するように構成された前記分析構成要素を更に備える、請求項8に記載のシステム。

【請求項11】

前記分析及びモデリング構成要素はルールエンジンを更に備え、前記ルールエンジンは、関連するルールの組に従って、前記データマイニング構成要素によって特定される因果事例を評価して、前記分析構成要素を用いて評価される場合、前記因果事例に関連するどの変数が、実施可能な結果を提供する最も高い可能性を提示するかを特定する、請求項10に記載のシステム。

【請求項12】

前記患者に関連する前記複数の臨床パラメータは、前記患者の年齢、体重、血圧、及び体温のうちの少なくとも1つを含む、請求項10に記載のシステム。

【請求項13】

前記ユーザインタフェースは、前記病気の前記可能性、ヘルスケア治療行動指針、及び計画された次の生化学アッセイのそれぞれを伝達するように構成される患者ダッシュボードを備える、請求項8に記載のシステム。

【請求項14】

情報抽出構成要素を組み込み、非構造化リサーチ源を、前記知識ベースと互換性を有するテンプレートに変換するリサーチインタフェースを更に備える、請求項8に記載のシステム。

【請求項15】

前記ユーザインタフェースは、特定された前記患者が前記病気を有する可能性に基づいて、推奨されるケアプロトコルを臨床医に伝達するように構成される臨床判断サポート構成要素を備える、請求項8に記載のシステム。

## 【国際調査報告】

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2014/019022

## A. CLASSIFICATION OF SUBJECT MATTER

INV. G06F19/28  
ADD. G06F19/24 G06F19/26

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	J. HERRERO: "GEPAS: a web-based resource for microarray gene expression data analysis", NUCLEIC ACIDS RESEARCH, vol. 31, no. 13, 1 July 2003 (2003-07-01), pages 3461-3467, XP055123687, DOI: 10.1093/nar/gkg591 abstract page 3462, column 2, paragraph 4 - page 3463, column 2, paragraph 2; figures 1,3 ----- -/--	1-22

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

## \* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

18 June 2014

Date of mailing of the international search report

01/07/2014

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040,  
Fax: (+31-70) 340-3016

Authorized officer

Schmitt, Constanze



## INTERNATIONAL SEARCH REPORT

International application No

PCT/US2014/019022

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ZIV BAR-JOSEPH ET AL: "Studying and modelling dynamic biological processes using time-series gene expression data", NATURE REVIEWS GENETICS, vol. 13, no. 8, 18 July 2012 (2012-07-18), pages 552-564, XP055123855, ISSN: 1471-0056, DOI: 10.1038/nrg3244 abstract page 557, column 1, paragraph 1 - paragraph 3; figures 1, 3; table 1 -----	1-22
A	Edwin P Romijn ET AL: "Expression clustering reveals detailed co-expression patterns of functionally related proteins during B cell differentiation: a proteomic study using a combination of one-dimensional gel electrophoresis, LC-MS/MS, and stable isotope labeling by amino acids in cell culture (SILAC)", Molecular & cellular proteomics : MCP, 1 September 2005 (2005-09-01), pages 1297-1310, XP055123812, United States DOI: 10.1074/mcp.M500123-MCP200 Retrieved from the Internet: URL:http://www.ncbi.nlm.nih.gov/pubmed/15961381 [retrieved on 2014-06-17] abstract -----	1-22
A	I. G. COSTA ET AL: "Constrained mixture estimation for analysis and robust classification of clinical time series", BIOINFORMATICS, vol. 25, no. 12, 28 May 2009 (2009-05-28), pages i6-i14, XP055123849, ISSN: 1367-4803, DOI: 10.1093/bioinformatics/btp222 abstract page i6, column 1, paragraph 3 - page i8, column 1, paragraph 4; figures 1,6 -----	1-22

## フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US

(72)発明者 シーゲル、ニール ジー.

アメリカ合衆国 90274 カリフォルニア州 ローリング ヒルズ エステーツ ゴールデン  
スパー プレイス 19

(72)発明者 シェカール、サム エス.

アメリカ合衆国 20854 メリーランド州 ポトマック ビーチノル レーン 10606

(72)発明者 ユー、ジェフリー シー.

アメリカ合衆国 94552 カリフォルニア州 カストロ バレー クレストフィールド サー  
クル 25720

Fターム(参考) 4C341 LL30

5L099 AA04