

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2014-83194
(P2014-83194A)

(43) 公開日 平成26年5月12日(2014.5.12)

(51) Int.Cl. F I テーマコード (参考)
A 6 1 B 5/00 (2006.01) A 6 1 B 5/00 G 4 C 1 1 7
G 0 6 Q 50/24 (2012.01) G 0 6 Q 50/24 1 0 0

審査請求 未請求 請求項の数 10 O L (全 25 頁)

(21) 出願番号	特願2012-233886 (P2012-233886)	(71) 出願人	503360115 独立行政法人科学技術振興機構 埼玉県川口市本町四丁目1番8号
(22) 出願日	平成24年10月23日(2012.10.23)	(74) 代理人	110000947 特許業務法人あーく特許事務所
		(72) 発明者	合原 一幸 千葉県習志野市谷津4-8-8-208
		(72) 発明者	陳 洛南 兵庫県神戸市東灘区向洋町中7丁目2-1 5番館808号室
		(72) 発明者	劉 銳 東京都世田谷区北沢1-13-17-10 2
		Fターム(参考)	4C117 XA05 XB09 XB12 XE02 XE44 XF30 XJ12 XJ27 XJ37

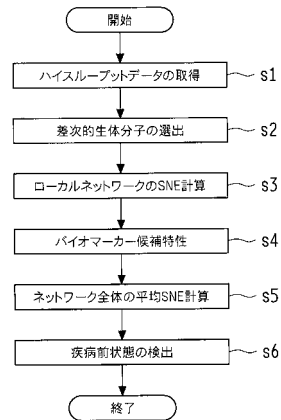
(54) 【発明の名称】 ネットワークエントロピーに基づく生体の状態遷移の予兆の検出を支援する検出装置、検出方法及び検出プログラム

(57) 【要約】

【課題】健康状態から疾病状態への状態遷移の予兆を示す疾病前状態を高精度に検出することが可能な検出装置、検出方法及び検出プログラムを提供する。

【解決手段】生体に関する遺伝子、タンパク質等の測定データをハイスループットデータとして取得する取得処理(s1)と、差次的生体分子の選出処理(s2)と、ローカルネットワークのSNE計算(s3)と、バイオマーカー候補特性(s4)と、ネットワーク全体の平均SNE計算(s5)と、疾病前状態であるか否かを判定し、検出する検出処理(s6)とを実行する。

【選択図】図4



【特許請求の範囲】**【請求項 1】**

生体に関する測定により得られた複数の因子項目の測定データに基づいて、測定対象である生体の状態遷移の予兆の検出を支援する検出装置であって、

前記各因子項目の測定データの時系列変化が所定の基準以上である因子項目を選出する選出手段と、

前記選出手段にて選出した因子項目の相関関係に基づき求まる各因子項目間の動的な結合関係を示すネットワークにおいて、因子項目毎に隣接する他の因子項目との間の統計力学的な微視的エントロピーを計算する微視的計算手段と、

前記微視的計算手段にて計算した微視的エントロピーの減少の程度が所定の検出基準を超える場合に、状態遷移の予兆として検出する検出手段と

を備えることを特徴とする検出装置。

10

【請求項 2】

請求項 1 に記載の検出装置であって、

前記微視的計算手段にて計算した微視的エントロピーの減少の程度が所定の選択基準を超える場合に、当該微視的エントロピーに係る因子項目を生体の症状の指標となるバイオマーカーの候補として選択する選択手段を更に備え、

前記検出手段は、前記選択手段が選択した因子項目に係る微視的エントロピーの減少の程度が所定の検出基準を超える場合に、状態遷移の予兆として検出する

ことを特徴とする検出装置。

20

【請求項 3】

請求項 1 又は請求項 2 に記載の検出装置であって、

前記微視的計算手段にて計算した因子項目毎の微視的エントロピーに基づいて、選出した因子項目全体の代表値となる巨視的エントロピーを統計的に計算する巨視的計算手段を更に備え、

前記検出手段は、前記巨視的計算手段により計算した巨視的エントロピーの減少の程度が第 1 の検出基準を超える場合であって、前記微視的計算手段にて計算した微視的エントロピーの減少の程度が第 2 の検出基準を超えると、状態遷移の予兆として検出する

ことを特徴とする検出装置。

30

【請求項 4】

請求項 1 乃至請求項 3 のいずれかに記載の検出装置であって、

前記因子項目間の相互作用を記録したデータベースにアクセスする手段を更に備え、

前記微視的計算手段は、前記データベースに記録されている因子項目間の相互作用に基づいて、因子項目間の動的な結合関係を示すネットワークを導出する手段を含む

ことを特徴とする検出装置。

【請求項 5】

請求項 1 乃至請求項 4 のいずれかに記載の検出装置であって、

前記微視的計算手段は、

各因子項目について、隣接する他の全ての因子項目に係る測定データの状態変化の分布を示す確率密度関数に基づく測定データの確率と該確率の対数との積の総和に基づいて、微視的エントロピーを計算する

ことを特徴とする検出装置。

40

【請求項 6】

請求項 5 に記載の検出装置であって、

前記微視的計算手段は、

因子項目毎に、以前の摂動に基づいて決定される閾値に対する変化の大小により測定データを 2 値化し、

2 値化した測定データが多変量正規分布に従うとして確率密度関数を計算し、

計算した確率密度関数を重積分した遷移確率に基づいて定常分布となる前記測定データの確率を計算する

50

ことを特徴とする検出装置。

【請求項 7】

請求項 1 乃至請求項 6 のいずれかに記載の検出装置であって、
前記各因子項目のそれぞれの測定データが、有意性をもって経時的に変化しているか否かを検定する差次検定手段を更に備え、

前記選出手段は、経時的变化に有意性があると検定された因子項目を選出することを特徴とする検出装置。

【請求項 8】

請求項 1 乃至請求項 7 のいずれかに記載の検出装置であって、
前記因子項目は、遺伝子に関する測定項目、タンパク質に関する測定項目又は代謝物に関する測定項目である

ことを特徴とする検出装置。

【請求項 9】

生体に関する測定により得られた複数の因子項目の測定データに基づいて、測定対象である生体の状態遷移の予兆の検出を支援する検出装置を用いた検出方法であって、

前記検出装置は、

前記各因子項目の測定データの時系列変化が所定の基準以上である因子項目を選出する選出ステップと、

前記選出手段にて選出した因子項目の相関関係に基づき求まる各因子項目間の動的な結合関係を示すネットワークにおいて、因子項目毎に隣接する他の因子項目との間の統計力学的な微視的エントロピーを計算する微視的計算ステップと、

前記微視的計算手段にて計算した微視的エントロピーの減少の程度が所定の検出基準を超える場合に、状態遷移の予兆として検出する検出ステップと

を実行することを特徴とする検出方法。

【請求項 10】

コンピュータに、生体に関する測定により得られた複数の因子項目の測定データに基づいて、測定対象である生体の状態遷移の予兆の検出を支援させる検出プログラムであって、

コンピュータに、

前記各因子項目の測定データの時系列変化が所定の基準以上である因子項目を選出する選出ステップと、

前記選出手段にて選出した因子項目の相関関係に基づき求まる各因子項目間の動的な結合関係を示すネットワークにおいて、因子項目毎に隣接する他の因子項目との間の統計力学的な微視的エントロピーを計算する微視的計算ステップと、

前記微視的計算手段にて計算した微視的エントロピーの減少の程度が所定の検出基準を超える場合に、状態遷移の予兆として検出する検出ステップと

を実行させることを特徴とする検出プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、生体に関する測定により得られた複数の因子項目の測定データに基づいて、測定対象である生体の状態遷移の予兆の検出を支援する検出装置、検出方法及び検出プログラムに関する。

【背景技術】

【0002】

様々な研究結果によると、多くの疾病、特に複雑な疾病の悪化進行のプロセスは、気候システム、生態システム、経済システム等のシステムと同様に、ある臨界閾値を超えた時点、いわゆる分岐点に達すると、状態遷移が突然発生し、健康な安定状態から急激に疾病状態に変化する（例えば、非特許文献 1～5 参照）。このような複雑な疾病の動的メカニズムに関する研究において、疾病悪化（例えば、喘息発作、癌の発病）の進行プロセスを、

10

20

30

40

50

時間に依存した非線形動力学システムとしてモデル化し、モデル化したシステムの動態を解析することによって、分岐点での状態遷移で病気が急激に悪化することが既に判明している（非特許文献1、6参照）。

【0003】

図1は、疾病の進行プロセスを概念的に示す説明図である。図1のaは、疾病の進行プロセスを模式的に示している。図1のb、cおよびdは、進行プロセスの過程において、前述のシステムの安定性をポテンシャル関数として示し、横軸にシステムの状態変数を取り、縦軸にポテンシャル関数の値をとって概念的に示した模式図である。図1のaに示すように、疾病悪化の進行プロセスは、正常状態（健康状態）、疾病前状態、疾病状態として表すことができる。正常状態において、システムは安定して、図1のbに示すように、ポテンシャル関数の値が最小値になる。疾病前状態において、システムは、図1のcに示すように、ポテンシャル関数の値が高くなる。したがって、外乱の影響を受けやすい状態であり、小さな外乱を受けるだけで状態遷移してしまう分岐点の付近、即ち、正常状態の限界に位置している。但し、当該疾病前状態は、適切な処置によって、しばしば正常状態に回復することができる。一方、疾病状態において、システムは再び安定化して、図1のdに示すように、ポテンシャル関数の値が大局的最小値になる。そのため、正常状態からの分岐による状態遷移で生じたこの疾病状態が、正常状態へ回復することは困難である。

10

【0004】

したがって、もし疾病前状態を検知して、疾病状態に遷移する前に、疾病状態に遷移しつつあることを患者に告知することができれば、適切な措置をとることによって、患者を疾病前状態から正常状態へ回復させることができる可能性が高い。

20

【0005】

即ち、分岐点（臨界閾値）を検出することができれば、状態遷移の予測が可能であり、病気の早期診断を実現することができる。しかしながら、複雑疾病の場合、状態遷移の予測は極めて困難である。その理由は以下の通りである。

【0006】

・第一に、疾病前状態は、正常状態の限界であり、分岐点に達する前に、著しい変化は検出し難い。そのため、従来のバイオマーカー、スナップショット測定等の手法による診断では、正常状態と疾病前状態とを区別することが難しい。

・第二に、様々な研究がなされている中で、分岐点を予測するための早期診断用の警告信号を精度高く検出することができる信頼性の高い疾病モデルがまだ開発されていない。特に、同じ疾病でも、個人によって、疾病悪化の進行プロセスが異なるため、モデルベースの診断方法は、成功する確率が低い。

30

・第三に、疾病前状態の検出対象は患者であり、通常、一人の患者から得られるサンプルの数が限られているため、長期間にわたって、予測に必要な十分なサンプルを採取することが困難である。

【0007】

それに対して、本願発明者らは、正常状態から疾病状態へ遷移する前の疾病前状態を示す警告信号となるバイオマーカーの候補を検出する方法を提案した（非特許文献7）。当該手法によれば、疾病状態に遷移する直前に現れる動的ネットワークバイオマーカー（DNB）を検出することによって、病気の早期診断を実現することができる。

40

【先行技術文献】

【非特許文献】

【0008】

【非特許文献1】ヴェネガス・ジェイ・ジー（Venegas, J.G.）等著，「破壊変化に対する前奏曲のような喘息の自己組織（Self-organized patchiness in asthma as a prelude to catastrophic shifts.）」，（英国），ネイチャー（Nature），ネイチャーパブリッシンググループ（Nature Publishing Group），2005年，434巻，p. 777 - 782

【非特許文献2】マックシャーリー・ピー・イー（McSharry, P.E.）、スミス・エル・エー（Smith, L.A.）、タラッセンコ・エル（Tarassenko, L.）著，「痙攣発作の予測：非線形

50

方法が適切であるか。(Prediction of epileptic seizures: are nonlinear methods relevant.)」, (英国), ネイチャー・メディシン (Nature Medicine), ネイチャーパブリッシンググループ (Nature Publishing Group), 2003年, 9巻, p. 241 - 242

【非特許文献3】ロベルト・ピー・ビー (Roberto, P.B.), エリセオ・ジー (Eliseo, G.), ジョセフ・エフ・シー (Josef, C.), 「物流回帰の変化点推定のための転移モデル (Transition models for change-point estimation in logistic regression.)」, (米国), 医学における統計 (Statistics in Medicine), ワiley-Blackwell, 2003年, 22巻, p. 1141 - 1162

【非特許文献4】ペック・エス (Paek, S.) 等著, 「音響神経腫に対するガンマナイフ手術後、聴覚保存 (Hearing preservation after gamma knife stereotactic radiosurgery of vestibular schwannoma.)」, (米国), キャンサー (Cancer), ワiley-Blackwell, 2005年, 1040巻, p. 580 - 590

【非特許文献5】リウ・ジェイ・ケイ (Liu, J.K.), ロビット・アール・エル (Rovit, R.L.), クッドウェル・ダブリュー・ティー (Couldwell, W.T.), 「脳下垂体卒中 (Pituitary Apoplexy)」, (米国), セミナーズ・イン・ニューロサージェリー (Seminars in neurosurgery), ティーメ (Thieme) 2001年, 12巻, p. 315 - 320

【非特許文献6】タナカ・ジー (Tanaka, G.), ツモト・ケイ (Tsumoto, K.), ツジ・エス (Tsuji, S.), アイハラ・ケイ (Aihara, K.) 著, 「前立腺癌に対する間歇的なホルモン療法のハイブリッドシステムモデルの分岐点分析。(Bifurcation analysis on a hybrid systems model of intermittent hormonal therapy for prostate cancer.)」, (米国), フィジカル・レビュー (Physical Review), アメリカ物理学会, 2008年, 237巻, p. 2616 - 2627

【非特許文献7】陳洛南 (Luonan Chen), 劉銳 (Rui Liu), 劉治平 (Zhi-Ping Liu), 李美儀 (Meiyi Li), 合原一幸 (Kazuyuki Aihara) 著, 「動的ネットワークバイオマーカーによる複雑な病気の突然の悪化に対する早期予兆信号の検出 (Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers.)」, サイエントフィックリポート (SCIENTIFIC REPORTS), 2012年3月29日, インターネット <URL: <http://www.natureasia.com/ja-jp/srep/abstracts/35129>>

【発明の概要】

【発明が解決しようとする課題】

【0009】

このように疾病に関する早期発見および早期治療に有効な疾病前状態に対しては、様々な検出方法の提案が社会的に求められている。なお、非特許文献7に示す動的ネットワークバイオマーカー (DNB: Dynamical Network Biomarker) の検出方法では、生体サンプルから測定した遺伝子等のデータに含まれるノイズの影響によって、検出精度が低下することがある。また、大量なハイスループットデータから、DNBとなる条件を満たすDNBの候補を検出するため、膨大な計算量が必要となる。

【0010】

本発明は、このような状況に鑑みてなされたものであり、遷移状態において、着目する因子および該因子に動的に直接接続する接続因子の状態変化をローカルネットワークエントロピーとして捉え、ローカルネットワークエントロピーに基づいて疾病状態へ遷移する前の疾病前状態を検出する。これにより、新たな方法で疾病前状態を検出することが可能な検出装置、検出方法及び検出プログラムを提供することを目的とする。

【課題を解決するための手段】

【0011】

上記目的を達成するために、本発明に係る検出装置は、生体に関する測定により得られた複数の因子項目の測定データに基づいて、測定対象である生体の状態遷移の予兆の検出を支援する検出装置であって、前記各因子項目の測定データの時系列変化が所定の基準以

10

20

30

40

50

上である因子項目を選出する選出手段と、前記選出手段にて選出した因子項目の相関関係に基づき求まる各因子項目間の動的な結合関係を示すネットワークにおいて、因子項目毎に隣接する他の因子項目との間の統計力学的な微視的エントロピーを計算する微視的計算手段と、前記微視的計算手段にて計算した微視的エントロピーの減少の程度が所定の検出基準を超える場合に、状態遷移の予兆として検出する検出手段とを備えることを特徴とする。

【0012】

上記特徴を有する検出装置では、生体の状態を複数の因子項目間の動的な結合関係を示すネットワークにおいて、統計力学的に計算した微視的エントロピーに基づいて、生体の状態遷移の予兆の検出を支援することができる。

10

【0013】

また、前記微視的計算手段にて計算した微視的エントロピーの減少の程度が所定の選択基準を超える場合に、当該微視的エントロピーに係る因子項目を生体の症状の指標となるバイオマーカーの候補として選択する選択手段を更に備え、前記検出手段は、前記選択手段が選択した因子項目に係る微視的エントロピーの減少の程度が所定の検出基準を超える場合に、状態遷移の予兆として検出することを特徴とする。

【0014】

したがって、計算の対象となる因子項目を絞り込むことができるので、ノイズの混入を低減して検出精度を向上させることが可能であり、しかも、計算量を低減することができるので、処理負荷の軽減および計算の高速化を実現することができる。

20

【0015】

また、前記微視的計算手段にて計算した因子項目毎の微視的エントロピーに基づいて、選出した因子項目全体の代表値となる巨視的エントロピーを統計的に計算する巨視的計算手段を更に備え、前記検出手段は、前記巨視的計算手段により計算した巨視的エントロピーの減少の程度が第1の検出基準を超える場合であって、前記微視的計算手段にて計算した微視的エントロピーの減少の程度が第2の検出基準を超えるとときに、状態遷移の予兆として検出することを特徴とする。

【0016】

したがって、巨視的エントロピーに基づきシステム全体が不安定となる状態を検出することができる。

30

【0017】

また、前記因子項目間の相互作用を記録したデータベースにアクセスする手段を更に備え、前記微視的計算手段は、前記データベースに記録されている因子項目間の相互作用に基づいて、因子項目間の動的な結合関係を示すネットワークを導出する手段を含むことを特徴とする。

【0018】

したがって、因子項目の関係に基づくネットワークを構築することができる。

【0019】

また、前記微視的計算手段は、各因子項目について、隣接する他の全ての因子項目に係る測定データの状態変化の分布を示す確率密度関数に基づく測定データの確率と該確率の対数との積の総和に基づいて、微視的エントロピーを計算することを特徴とする。

40

【0020】

したがって、統計力学および情報理論上のエントロピーをネットワークのエントロピーに拡張して用いることができる。

【0021】

また、前記微視的計算手段は、因子項目毎に、以前の摂動に基づいて決定される閾値に対する変化の大小により測定データを2値化し、2値化した測定データが多変量正規分布に従うとして確率密度関数を計算し、計算した確率密度関数を重積分した遷移確率に基づいて定常分布となる前記測定データの確率を計算することを特徴とする。

【0022】

50

したがって、因子項目の大きな変化に基づいてネットワークのエントロピーを計算することができる。

【0023】

また、前記各因子項目のそれぞれの測定データが、有意性をもって経時的に変化しているか否かを検定する差次検定手段を更に備え、前記選出手段は、経時的变化に有意性があると検定された因子項目を選出することを特徴とする。

【0024】

したがって、著しい変化を示した因子項目を選出することができる。

【0025】

また、前記因子項目は、遺伝子に関する測定項目、タンパク質に関する測定項目又は代謝物に関する測定項目であることを特徴とする。

【0026】

したがって、因子項目を遺伝子、タンパク質、または代謝物とすることによって、生体内の生物学的変化を定量的に把握することができる。

【0027】

本発明に係る検出方法は、生体に関する測定により得られた複数の因子項目の測定データに基づいて、測定対象である生体の状態遷移の予兆の検出を支援する検出装置を用いた検出方法であって、前記検出装置は、前記各因子項目の測定データの時系列変化が所定の基準以上である因子項目を選出する選出ステップと、前記選出手段にて選出した因子項目の相関関係に基づき求まる各因子項目間の動的な結合関係を示すネットワークにおいて、因子項目毎に隣接する他の因子項目との間の統計力学的な微視的エントロピーを計算する微視的計算ステップと、前記微視的計算手段にて計算した微視的エントロピーの減少の程度が所定の検出基準を超える場合に、状態遷移の予兆として検出する検出ステップとを実行することを特徴とする。

【0028】

上記特徴を有する検出方法では、生体の状態を複数の因子項目間の動的な結合関係を示すネットワークにおいて、統計力学的に計算した微視的エントロピーに基づいて、生体の状態遷移の予兆の検出を支援することができる。

【0029】

本発明に係る検出プログラムは、コンピュータに、生体に関する測定により得られた複数の因子項目の測定データに基づいて、測定対象である生体の状態遷移の予兆の検出を支援させる検出プログラムであって、コンピュータに、前記各因子項目の測定データの時系列変化が所定の基準以上である因子項目を選出する選出ステップと、前記選出手段にて選出した因子項目の相関関係に基づき求まる各因子項目間の動的な結合関係を示すネットワークにおいて、因子項目毎に隣接する他の因子項目との間の統計力学的な微視的エントロピーを計算する微視的計算ステップと、前記微視的計算手段にて計算した微視的エントロピーの減少の程度が所定の検出基準を超える場合に、状態遷移の予兆として検出する検出ステップとを実行させることを特徴とする。

【0030】

上記特徴を有する検出プログラムをコンピュータにて実行することにより、コンピュータが検出装置として動作するので、生体の状態を複数の因子項目間の動的な結合関係を示すネットワークにおいて、統計力学的に計算した微視的エントロピーに基づいて、生体の状態遷移の予兆の検出を支援することができる。

【発明の効果】

【0031】

本発明によれば、生体サンプルを診断対象から採取し、採取した生体サンプルに対する測定により得られる複数の因子項目の測定データに基づいて、生体の状態を複数の因子項目間の動的な結合関係を示すネットワークについて、統計力学的に微視的なエントロピーを計算する。そして、計算したエントロピーの時系列変化に基づいて、生体の状態遷移の予兆の検出を支援する。これにより、疾病前状態の検出する方法を新たに提案して、疾病

10

20

30

40

50

に関する早期発見および早期治療を実現することが可能である等、優れた効果を奏する。

【図面の簡単な説明】

【0032】

【図1】図1は、疾病の進行プロセスを説明する模式図である。

【図2】図2は、SNEとDNBとの関係を示す図表である。

【図3】図3は、疾病の進行プロセスにおけるDNBとSNEの特性を例示する模式図である。

【図4】図4は、実施の形態におけるDNBの検出方法の一例を示すフローチャートである。

【図5】図5は、実施の形態における差次的生体分子の選出処理の一例を示すフローチャートである。

【図6】図6は、実施の形態におけるローカルネットワークのSNE計算処理の一例を示すフローチャートである。

【図7】図7は、実施の形態におけるバイオマーカー候補の選択処理の一例を示すフローチャートである。

【図8】図8は、本発明に係る検出装置の構成例を示すブロック図である。

【図9】図9は、本発明に係る検出装置による生体の状態遷移の検出処理の一例を示すフローチャートである。

【発明を実施するための形態】

【0033】

<基本理論>

本発明の発明者らは、1つのサンプルから多数（例えば、数千個）のデータ、即ち高次元データを得ることが可能なゲノムハイスループットテクノロジーを利用して、分岐理論に基づいて、複雑疾病の時間発展の数理モデルを構築し、分子ネットワークレベルにおける疾病悪化の進行メカニズムの研究を行った。その結果、疾病前状態において、状態遷移発生前の直前分岐（突然悪化）状態を検出することができる動的ネットワークバイオマーカー（DNB：Dynamical Network Biomarker）の存在を解明した。当該動的ネットワークバイオマーカーを疾病前状態の警告信号として利用すれば、疾病モデルが不要で、数少ないサンプルだけで複雑疾病の早期診断を実現することができる。

【0034】

通常、疾病の進行プロセスに係るシステム（以下、システム（1）という）は、下記の式（1）で表すことができる。

【0035】

$$Z(k+1) = f(Z(k); P) \quad \dots \text{式(1)}$$

【0036】

式（1）において、 $Z(k) = (z_1(k), \dots, z_n(k))$ は、時刻 k ($k = 0, 1, \dots$) で観測されたシステム（1）の動的状態を表す変数であり、タンパク質発現量、遺伝子発現量、代謝物発現量等の情報とすることができる。詳細には、タンパク質、遺伝子等の分子に係る濃度、個数等の情報である。Pは、システム（1）の状態遷移を駆動する緩やかに変化するパラメータであり、例えば、SBP, CNV等の遺伝的要因、メチル化、アセチル化等の非遺伝的要因等の情報とすることができる。f = (f₁, ..., f_n) は Z(k) の非線形関数である。

【0037】

正常状態、疾病状態はそれぞれ、状態方程式 $Z(k+1) = f(Z(k); P)$ の不動点アトラクタで表すことができる。複雑疾病の進行プロセスは、非常に複雑な動的特性を有するため、関数 f は、数千個の変数を有する非線形関数である。しかも、システム（1）を駆動するパラメータ P の要素は特定し難いものである。そのため、正常状態および疾病状態のシステムモデルを構築し解析することは大変困難である。

【0038】

ところで、システム（1）には、次の（A1）～（A3）に示す特性を有する不動点が

10

20

30

40

50

存在する。

【0039】

(A1) Z^* は、システム(1)の不動点とすれば、 $Z^* = f(Z^*; P)$ 、

(A2) P_c をシステムが分岐する閾値とすれば、 $P = P_c$ の時、ヤコビ行列 $f(Z; P_c) / Z | Z = Z^*$ の1つの実数固有値又は1対の複素共役の固有値の絶対値が1になる、

(A3) $P \neq P_c$ の時、一般的に、システム(1)の固有値の絶対値は1ではない。

【0040】

本願発明者らは、上記特性から、システム(1)が状態遷移点近くになると、次のような特異な特性が現れることを理論的に解明した。即ち、システム(1)が状態遷移点近くになると、システム(1)の各変数 z_1, \dots, z_n のそれぞれをノードとして構成したネットワーク(1)において、一部のノードからなる支配グループ(サブネットワーク)が現れる。状態遷移点近くで現れる支配グループは、理想的には、次の(B1)~(B3)に示す特異な特性を有する。

10

【0041】

(B1) z_i と z_j は支配グループに属するノードである場合、

$$PCC(z_i, z_j) = \pm 1;$$

$$SD(z_i) = \dots;$$

$$SD(z_j) = \dots.$$

(B2) z_i は支配グループに属するノードであるが、 z_j は支配グループに属するノードではない場合、

20

$$PCC(z_i, z_j) = 0;$$

$$SD(z_i) = \dots;$$

$$SD(z_j) = \text{有界値}.$$

(B3) z_i と z_j は支配グループに属するノードではない場合

$$PCC(z_i, z_j) = \dots, (-1, 1);$$

$$SD(z_i) = \text{有界値};$$

$$SD(z_j) = \text{有界値}.$$

【0042】

ここで、 $PCC(z_i, z_j)$ は z_i と z_j との間のピアソン相関係数であり、 $SD(z_i)$ 、 $SD(z_j)$ は z_i と z_j との標準偏差である。

30

【0043】

即ち、ネットワーク(1)において、上記特異な特性(B1)~(B3)を有する支配グループの出現は、システム(1)が臨界遷移状態(疾病前状態)に在ることの兆候として捉えることができる。したがって、当該支配グループを検出することにより、システム(1)の状態遷移の予兆を検出することができる。即ち、当該支配グループを状態遷移、即ち、疾病悪化直前の疾病前状態を示す警告信号とすることができる。即ち、システム(1)がいくら複雑であっても、また駆動パラメータ要素が不明であっても、警告信号となる支配グループだけを検出すれば、システム(1)の数理モデルを直接扱うことなく、疾病前状態であることを特定することができる。疾病前状態であることを特定することにより、疾病に対する事前対策および早期治療を実現することが可能となる。非特許文献7として示したように、本願発明者らは、当該疾病前状態を示す警告信号となる当該支配グループを、動的ネットワークバイオマーカー(以下、DNBと略称する)と称している。非特許文献7にて開示したDNBとは、特定のタイミングでのみ実効的な結合を生じる論理的な動的結合関係を示すネットワークを、バイオマーカーとして用いるものである。

40

【0044】

以上のように、DNBは、特異な特性(B1)~(B3)を有する支配グループであって、複数のノードからなるサブネットワークとして、システム(1)が疾病前状態に在るときに、ネットワーク(1)に現れるものである。ネットワーク(1)において、各ノード(z_1, \dots, z_n)を、遺伝子、タンパク質、代謝物等の生体分子について測定の対

50

象となる因子項目とすれば、DNBは、上記特異な特性(B1)~(B3)を満たした一部の生体分子に係る因子項目からなるグループ(サブネットワーク)である。

【0045】

上記の特異な特性(B1)~(B3)を直接利用して、DNBとなる候補を検出する手法は非特許文献7に既に開示されている。当該手法によれば、生体サンプルから疾病状態に遷移することを警告するDNBを検出することができる。しかしながら、測定データにノイズが含まれている場合、検出精度が悪化する。また、大量な測定データから上記条件(B1)~(B3)を満たすDNBを検出する必要があるため、計算量が膨大であり、検出効率が高くない。

【0046】

それに対して、本願発明者らは、DNBを精度高くかつ効率よく検出することができる、遷移状態に基づくローカルネットワークエントロピーを用いた疾病前状態の検出方法を提案した。次に、当該方法について具体的に説明する。以降に説明にするローカルネットワークエントロピーとは、特定のタイミングでのみ実効的な結合を生じる論理的な動的結合関係を示すネットワークにおいて、その一つのノードに着目して統計力学的に計算する微視的エントロピーである。なお、本願では、更にローカルネットワークエントロピーからネットワーク全体の巨視的エントロピーの計算も行っている。

【0047】

<遷移状態に基づくローカルネットワークエントロピー>

上述したシステム(1)の動的振る舞いは、分岐点近傍では下記の式(2)として近似的に表すことができる。

【0048】

$$Z(t+1) = A(P)Z(t) + \epsilon(t) \quad \dots \text{式(2)}$$

【0049】

式(2)において、 $\epsilon(t)$ はガウスノイズであり、Pはシステム(1)の非線形関数fのヤコビアン行列Aを制御するパラメータベクトルである。そして、Zの変化量を $z_i(t) = z_i(t) - z_i(t-1)$ と記すと($i = 1, 2, \dots, n$)、分岐理論および中心多様体理論に基づいて、次の(C1)および(C2)の結論を証明することができる。

【0050】

(C1) Pが状態遷移点または分岐点の近辺に存在しない場合、

任意のノードi, j ($i = j$ の場合を含む)に関し、 $z_i(t+T)$ は $z_i(t)$ から統計的に独立である($i, j = 1, 2, \dots, n$)。

(C2) Pが状態遷移点または分岐点の近辺に存在する場合、

・ノードiとjとがいずれも支配グループまたはDNBに存在するとき、 $z_i(t+T)$ と $z_j(t)$ との間の相関は高くなる。

・ノードiとjとがいずれも支配グループまたはDNBに存在しないとき、 $z_i(t+T)$ は $z_j(t)$ から統計的に独立である。

【0051】

上記結論(C1)および(C2)に基づき、本願発明者らは遷移状態に着目して、ローカルネットワークエントロピー(以下、必要に応じてSNE(State-transition-based local network entropy)と称する。)を用いてDNBをより高精度に、かつ効率良く検出する方法を見いだした。以下に、遷移状態に基づくSNEの概念およびSNEとDNBとの関係について説明する。

【0052】

・遷移状態(transition states)について

時刻tにおける任意の変数 z_i に対して、式(3)および式(4)として示す次の条件を満たす $x_i(t)$ にて遷移状態を示すものとする。

【0053】

$$|z_i(t) - z_i(t-1)| > d_i \quad \text{であれば、} \quad x_i(t) = 1 \quad \dots \text{式(3)}$$

10

20

30

40

50

$|z_i(t) - z_i(t-1)| \geq d_i$ であれば、 $x_i(t) = 0$...式(4)

【0054】

式(3)および式(4)において、 d_i は時刻 t でノード i の変化が大きいかなんかを定める閾値である。本発明において、 $X(t) = (x_1(t), \dots, x_n(t))$ を、時刻 t におけるシステム(1)の「遷移状態」と定義する。上記したDNBの特異な特性および結論(C1)および(C2)から、次の(D1)および(D2)の遷移状態の特性を導くことができる。

【0055】

(D1) ノード i とノード j とが支配グループまたはDNBに属するノードである場合、遷移状態 $x_i(t+T)$ と $x_j(t)$ との間の相関が急激に増加し、しかも

$$p(x_i(t+T) = 1 | x_j(t) = 1) > p(x_i(t+T) = 0 | x_j(t) = 1)$$

$$p(x_i(t+T) = 0 | x_j(t) = 0) > p(x_i(t+T) = 1 | x_j(t) = 0)$$

ここで、 $\{0, 1\}$ 、 p は遷移確率である。

(D2) ノード i とノード j とがいずれも支配グループまたはDNBに属するノードではない場合、遷移状態 $x_i(t+T)$ が統計的に $x_j(t)$ から独立であり、しかも

$$p(x_i(t+T) = i | x_j(t) = j) = p(x_i(t+T) = i)$$

ここで、 $i, j \in \{0, 1\}$ 、 $i, j \in (0, 1)$ である。

【0056】

なお、システムが正常状態であれば外乱から速やかに回復できるが、疾病前状態では小さな外乱にも敏感になっているため、上記閾値 d_i は正常状態の「小さな変化」と疾病前状態の「大きな変化」とを区別できるように設定しなければならない。ここでは、システムが正常状態($t = t_0$)において、各ノード k で、 $p(|z_k(t_0)| > d_k) = \alpha$ とし、各閾値 d を下記の式(5)のように設定する。式(5)により設定される各閾値 d によって、以前の状態 $z_i(t-1)$ から判別の対象となる時点の状態 $z_i(t)$ までの間に大きな変化又は状態遷移が生じているかなんかを判別することになる。なお、式(5)において、 i_1, i_2, \dots, i_m は、ノード i とリンクする m 個の隣接ノードを示している。

【0057】

【数1】

$$p(|z_i(t_0)| > d_i, |z_{i_1}(t_0)| > d_{i_1}, \dots, |z_{i_m}(t_0)| > d_{i_m}) \leq \alpha \dots \text{式(5)}$$

【0058】

例えば、正常状態の時期に採取したサンプルに基づいて、正常状態における摂動に対しては、 $\alpha = 0.5$ となるように、各閾値 d_i を設定する。

【0059】

・ローカルネットワークについて

ノード i が m 個のノードとリンクを有する、即ち、ノード i が m 個の隣接ノード (i_1, i_2, \dots, i_m) を有する場合、ノード i を中心とするネットワークをローカルネットワークと定義する。この場合、ノード i を中心とするローカルネットワークの時刻 t における遷移状態は、 $X^i(t) = (x_i(t), x_{i_1}(t), \dots, x_{i_m}(t))$ となる。以下、数式を簡潔化するために、「 i 」を省いて、 $X^i(t)$ を $X(t)$ と記述する。

【0060】

各ノード i の接続関係は、ノード間の相互作用に基づいて設定される。例えば、ノードとしてタンパク質を用いる場合、タンパク質間の相互作用を示した PPI (Protein-Protein Interaction) 等のデータベースに記録されている情報を用いることができる。このようなデータベースは、例えば、BioGrid (www.thebiogrid.org)、TRED (www.rulai.cshl.edu/cgi-bin/TRED/)、KEGG (www.gnome.jp/kegg)、HPRD (www.hprd.org) 等のウェブページから入手することも可能である。タンパク質をノードとして用いる場合、タンパク質間の相互作用を示した PPI 等のデータベースに基づいて、隣接

10

20

30

40

50

ノード並びに隣接ノードに基づくローカルネットワークおよび全体のネットワークが設定されるが、他の因子をノードとする場合、該当する因子に関するデータベースを用いることになる。

【0061】

当該時刻 t における遷移状態 $X(t)$ に基づく次の時刻 $t+1$ における遷移状態は、 $2^m + 1$ 通りの可能性がある。これらの可能な遷移状態はそれぞれ確率事象であり、 $\{A_u\}_{u=1,2,\dots,2^{m+1}}$ を用いて下記の式(6)として示すことができる。

【0062】

$$A_u = \{x_{i1} = 0, x_{i2} = 1, \dots, x_{im} = m\} \quad \dots \text{式(6)}$$

但し、 $x_{ij} \in \{0, 1, 2, \dots, m\}$

10

【0063】

したがって、ローカルネットワークにおいて、離散確率過程は、下記の式(7)となる。

【0064】

$$\{X(t+i)\}_{i=0,1,\dots} = \{X(t), X(t+1), \dots, X(t+i), \dots\}$$

...式(7)

但し、 $X(t+i) = A_u, u \in \{1, 2, \dots, 2^{m+1}\}$ 。

【0065】

即ち、システム(1)が正常状態または疾病前状態である場合、当該離散確率過程は、マルコフ過程(stochastic Markov process)であり、マルコフマトリックス $P = (p_{uv})$ で定義することができる。それによって、状態 u から状態 v への遷移レートを下記の式(8)として示すことができる。

20

【0066】

【数2】

$$p_{u,v}(t) = \Pr(X(t+1) = A_v | X(t) = A_u) \quad \dots \text{式(8)}$$

但し、 $u, v \in \{1, 2, \dots, 2^{m+1}\}$,

$$\sum_v p_{u,v}(t) = 1,$$

Pr: 離散確率.

30

【0067】

ローカルネットワークエントロピーについて

一定期間において、上述したローカルネットワークの状態遷移マトリックスが定常であり、変化しないと仮定する。 $p_{u,v}(t)$ は状態遷移マトリックスにおける u 行目 v 列目の要素であり、任意の2つの可能な状態 A_u と A_v との間の遷移確率である。したがって、特定の期間(正常状態または疾病前状態)において、下記の式(9)で示す確率過程は定常マルコフ過程である。

40

【0068】

【数3】

$$\text{確率過程 } \{X(t)\}_{t \in [t_1, t_2]} \quad \dots \text{式(9)}$$

【0069】

そして、下記の式(10)を満たす定常分布 $\pi = (\pi_1, \dots, \pi_{2^{m+1}})$ が存在する。

【0070】

【数4】

$$\sum_v \pi_v P_{u,v} = \pi_u \quad \cdots \text{式(10)}$$

【0071】

それによって、下記の式(11)で示すローカルネットワークエントロピーを定義することができる。

【0072】

【数5】

$$H_i(t) = H(x) = - \sum_{u,v} \pi_v P_{u,v} \log P_{u,v} \quad \cdots \text{式(11)}$$

10

【0073】

式(11)において、インデックス「i」はローカルネットワークの中心ノードiを示し、Xはローカルネットワークの状態遷移過程 $X(t)$, $X(t+1)$, ..., $X(t+T)$ を示す。式(11)として示したローカルネットワークエントロピーは、統計力学的な微視的エントロピーを拡張した概念である。

【0074】

20

以下、上記ローカルネットワークエントロピーをSNEと称する。上述したように、一定な期間において、確率過程 $X(t)$, $X(t+1)$, ... はマルコフ過程であるため、マルコフ連鎖の性質などから、下記の式(12)が得られる。

【0075】

【数6】

$$H_i(t) = H(x) =: \lim_{T \rightarrow \infty} \frac{1}{T} H(X(t), X(t+1), \dots, X(t+T)) \quad \cdots \text{式(12)}$$

【0076】

30

そのため、SNEは条件付きエントロピーであり、状態遷移に依存する平均遷移エントロピーとすることができる。従って、下記の式(13)で表すことができる。

【0077】

$$H_i(t) = H(X(t) | X(t-1)) = H(X(t), X(t-1)) - H(X(t-1))$$

...式(13)

【0078】

また、SNEは以下の(E1)~(E3)に示す性質を有する。

【0079】

40

(E1) システム(1)が正常状態または疾病状態である場合、システム(1)は小さな外乱に対して安定であるため、 $X(t)$ と $X(t-1)$ とはほぼ独立である。

【0080】

従って、

$$H(X(t), X(t-1)) = H(X(t)) + H(X(t-1)) > 0 \text{ となるため}$$

、 $H_i(t) = H(X(t))$ となり、SNEの値が大きく減少することがない。

【0081】

(E2) システム(1)が疾病前状態である場合、システム(1)は小さな外乱に対して不安定であることから、 $X(t)$ と $X(t-1)$ とは強く相関する。

50

【 0 0 8 2 】

従って、

$H(X(t), X(t-1)) = H(X(t-1))$ となるため、

$H_i(t) = 0$ となり、 $H_i(t)$ が急激に減少する。

【 0 0 8 3 】

(E3) 各ローカルネットワークの SNE の平均値を全体のネットワークの SNE とすることができる。即ち、下記の式 (14) に示すように、 n 個のノードを有するネットワーク全体の SNE である $H(t)$ は、各ノードを中心とするローカルネットワークの SNE である $H_i(t)$ の平均値とする。

【 0 0 8 4 】

【 数 7 】

$$H(t) = \frac{1}{n} \sum_{i=1}^n H_i(t) \quad \cdots \text{式(14)}$$

【 0 0 8 5 】

式 (14) として示したネットワーク全体の SNE は、統計力学的な巨視的エントロピーを拡張した概念である。

【 0 0 8 6 】

以上のように定義される SNE と、疾病前状態を検出することができる他の方法としての DNB との関係について説明する。本願において、ネットワークにおけるノードは、DNB に係る自ノードおよび他のノードとの関係から以下の 4 つのタイプに分類することができる。

【 0 0 8 7 】

・タイプ 1 (DNB コアノード) : それ自体が DNB ノードであり、それとリンクしているノードもすべて DNB ノードであるノードを「DNB コアノード」と定義する。

・タイプ 2 (DNB 境界ノード) : それ自体が DNB ノードであるが、それとリンクしているノードの中に、少なくとも 1 つは DNB ノードではないノードを含むノードを「DNB 境界ノード」と定義する。

・タイプ 3 (非 DNB コアノード) : それ自体は DNB ノードではないが、少なくとも 1 つの DNB とリンクしているノードを「非 DNB コアノード」と定義する。

・タイプ 4 (非 DNB 境界ノード) : それ自体は DNB ノードではなく、DNB ノードともリンクしていないノードを「非 DNB 境界ノード」と定義する。

【 0 0 8 8 】

一方、システムが正常状態における SNE を H^{no} (X) と記し、疾病前状態における SNE を H^{pre} (X) と記すと、それぞれを下記の式 (15) および式 (16) で表すことができる。

【 0 0 8 9 】

10

20

30

【数 8】

$$H^{nor}(x) = -\sum_{u,v} \pi_u^{nor} p_{u,v}^{nor} \log p_{u,v}^{nor} \dots \text{式(15)}$$

但し、

$$p^{nor}(x_i(t)=1) = p^{nor}(|\Delta z_i| \geq d_i) = \alpha,$$

$$p^{nor}(x_i(t)=0) = p^{nor}(|\Delta z_i| < d_i) = 1 - \alpha.$$

10

$$H^{pre}(x) = -\sum_{u,v} \pi_u^{pre} p_{u,v}^{pre} \log p_{u,v}^{pre} \dots \text{式(16)}$$

但し、

$$p^{pre}(x_j(t)=1) = p^{pre}(|\Delta z_j(t)| \geq d_j) \rightarrow 1,$$

$$p^{pre}(x_j(t)=0) = p^{pre}(|\Delta z_j(t)| < d_j) \rightarrow 0.$$

20

【0090】

図2は、SNEとDNBとの関係を示す図表である。図2は、上記関係式およびDNBの特性に基づいて証明されるSNEとDNBとの間の関係を、タイプ毎に、ノードの種類、中心ノードの状態遷移およびローカルSNEの状態を対応付けて示している。

【0091】

図2において、 α は、 $(0, 1)$ となる定数である。即ち、中心ノードがタイプ1のDNBコアノードであるローカルネットワークにおいて、中心ノードの状態遷移は1に近い値であり、SNEは急激に0まで減少する。中心ノードがタイプ2のDNB境界ノードであるローカルネットワークにおいて、中心ノードの状態遷移は1に近い値であり、SNEは減少する。中心ノードがタイプ3の非DNBコアノードであるローカルネットワークにおいて、中心ノードの状態遷移は所定の定数 α に近い値であり、SNEは減少する。中心ノードがタイプ4の非DNB境界ノードであるローカルネットワークにおいて、中心ノードの状態遷移は所定の定数 α に近い値であり、SNEはあまり変化しない。

30

【0092】

図3は、疾病の進行プロセスにおけるDNBとSNEとの特性を例示する模式図である。図3は、疾病の進行プロセスにおけるDNBとSNEとの特性を概念的に示している。図3には、遺伝子などを示すノード $z_1 \sim z_6$ からなるネットワークが示されている。

【0093】

図3の(a)は正常状態、疾病前状態および疾病状態を示し、疾病前状態から正常状態へ回復することが可能であるが、疾病状態から疾病前状態へ回復することは困難であることを示している。

40

【0094】

図3の(b)は正常状態におけるノード $z_1 \sim z_6$ を丸印にて示しており、各ノードの標準偏差(丸印内の斜線の密度で表す)および各ノード間の相関係数を示している。

【0095】

図3の(c)は疾病前状態におけるノード $z_1 \sim z_6$ を示し、 $z_1 \sim z_3$ の標準偏差が大きくなっており(丸印内の斜線の密度が密になっている)、互いの相関係数が高くなっている(リンクの線が太くなっている)が、他のノードとの間の相関係数が低減している

50

(他のノードとつなぐ線が細くなっている)ことを示している。したがって、疾病前状態においてDNBのメンバーである $z_1 \sim z_3$ が顕在化している。

【0096】

図3の(d)は、疾病状態におけるノード $z_1 \sim z_6$ を示しており、ノード $z_1 \sim z_3$ の標準偏差は正常状態よりわずかに大きくなっているが、各ノード間の相関係数が均等になっていることを示している。

【0097】

図3の(e)は、従来のバイオマーカーによる診断方法を示している。図3の(e)において、横軸は、特定のタンパク質の濃度等の従来のバイオマーカーの指標を示しており、左から右へ行くほど、濃度等の指標が高くなることを示している。図3の(e)に示すように、従来のバイオマーカーによる診断方法では、疾病前状態において、丸印で示す正常サンプルと、星印で示す疾病前サンプルとを明確に区別することはできない。

10

【0098】

図3の(f)は、SNEを指標として用い、図3の(e)に示したサンプルを再配置した例を示しており、右から左へ行くほど、SNEが低くなり、指標としての警告の程度が高まっていることを示す。図3の(f)に示すように、SNEを用いて、丸印で示す正常サンプルと、星印で示す疾病前サンプルとは明確に分離されている。従って、SNEを指標として用いることにより、疾病前状態を検出することが可能となる。

【0099】

図3の(g)は、疾病進行プロセスにおけるネットワークの平均SNEの値の変化を、横軸に時間を取り、縦軸にSNEの値をとったグラフとして示している。正常状態および疾病状態において、SNEの値が高く、システムが高いロバスト性を有することを示しており、それに対して、疾病前状態では、SNEの値が急激に減少し、システムのロバスト性が低くなっていることを示している。

20

【0100】

< SNEによる疾病前状態の検出 >

上述したように、システムが状態遷移点近傍、即ち疾病前状態になると、DNBノードからなる支配グループが現れ、システムを正常状態から疾病状態へ引き込む。上記図2の図表に記載したSNEとDNBとの関係を利用して、各サンプリング時刻 t において、ネットワーク全体における各ノード i を中心ノードとするローカルネットワークのSNEをそれぞれ計算すれば、SNEの値が急激に減少するノードをDNBノードとして検出することができる。さらに、算出した各ローカルネットワークのSNEから、ネットワーク全体の平均SNEを計算し、その値が急激に減少した場合、多数のDNBノードが存在し、DNBノードからなる支配グループが現れていることを意味するため、システムが疾病前状態にあると判断することができる。また、図2の図表に示したように、タイプ1~4のノードにおいて、SNEの値が増加することがない。そのため、ネットワーク全体の平均SNEの計算において、値が減少したローカルネットワークのSNEのみを使用すれば、ノイズの混入を防止することができるため、検出の精度の向上を見込むことができ、しかも、計算量を低減することができるため、計算の効率を向上することができる。

30

【0101】

< DNBの検出方法 >

次に、前述の理論に基づいて、SNEにより疾病前状態を検出する具体的な方法について説明する。図4は、実施の形態における疾病前状態の検出方法の一例を示すフローチャートである。本発明の検出方法においては、まず、生体に関する測定により測定データを得ることが必要である。DNAチップなどのハイスループット技術を利用すれば、1つの生体サンプルから2万個以上の遺伝子の発現量を測定することができる。統計的に分析するために、本発明において、測定対象から複数の生体サンプルを異なる時刻で採取し、採取した生体サンプルを測定して得られた測定データを集計して統計データを取得する。本発明におけるDNBの検出方法は、主に、図4に示すように、ハイスループットデータの取得処理(s1)と、差次的生体分子の選出処理(s2)と、ローカルネットワークのS

40

50

NEの計算処理(s 3)と、バイオマーカーの候補の選択処理(s 4)と、ネットワーク全体の平均SNEの計算処理(s 5)と、疾病前状態であるか否かを判定し、検出する検出処理(s 6)とを含む。次に、これらの各処理について詳細に説明する。

【0102】

ステップs 1のハイスループットデータの取得処理は、検出対象となるサンプルをケースサンプルとし、参照用サンプルをコントロールサンプルとし、各サンプルから、それぞれのハイスループット生理学的データ、即ち、生体分子の発現量の測定データ(例えば、マイクロアレイデータ)を取得する処理である。参照用サンプルとは、検査の対象となる患者から事前に採取したサンプル、採取時に最初に採取したサンプル等のサンプルであり、測定装置の校正等の目的を有するコントロールサンプルとして用いられる。コントロールサンプルは、必ずしも必要ではないが、誤差要因を排除し、測定の信頼性を向上させるために有用である。

10

【0103】

ステップs 2の差次的生体分子の選出処理は、発現量に著しい変化を示した生体分子を選択する処理である。図5は、実施の形態における差次的生体分子の選出処理の一例を示すフローチャートである。図5は、図4に示したステップs 2の差次的生体分子の生体処理を詳細に示したものである。

【0104】

図5に示すように、まず、n個のケースサンプルからそれぞれ測定したハイスループットデータ(生体分子の発現量)に基づく統計データをD1cとし、コントロールサンプルから測定したデータをDrとする(s 21)。次に、各ケースサンプル生体分子D1cに対して、t検定を行い、コントロールサンプルのハイスループットデータDrに比べて発現量に著しい変化を示した生体分子D2cを選出する(s 22)。ステップs 22では、発現量に著しい変化を示した生体分子D2cを選出する方法として、t検定を例示しているが、特に手法を限定するものではない。例えば、U検定等の他の検定方法を適用することも可能であり、このようなノンパラメトリック手法による検定は、母集団となるD1cが正規分布に従わない場合に、特に有効である。また、t検定をする場合においても、有意水準の値は、0.05、0.01等の値を適宜設定することができる。

20

【0105】

次に、誤発現率FDRを用いて、各ケースサンプル生体分子D2cに対して、複数の比較(multiple comparisons)又は複数のt検定を補正し、補正後の各ケースサンプル遺伝子又はタンパクデータD3cを選出する(s 23)。次に、2倍変更方法(The two-fold change method)を用いて、補正後の各ケースサンプル遺伝子又はタンパクデータD3cから標準偏差SDが比較的大きく変化するDcを差次的生体分子として選出する(s 24)。ここで選出された差次的生体分子Dcは、コントロールサンプルの生体分子Drに比べて、著しい差異を示すだけでなく、自らの平均値からも大きく逸脱している。ステップs 23においても、検定方法はt検定に限るものではない。

30

【0106】

次に、ローカルネットワークSNEの計算(図4のステップs 3)を行う。図6は、実施の形態におけるローカルネットワークのSNE計算処理の一例を示すフローチャートである。まず、ステップs 24にて選択した差次的生体分子Dcに係る測定データを下記の式(17)にて正規化する(ステップs 31)。なお、以降の計算においては、式(17)にて正規化したデータを用いる。

40

【0107】

$$A = (D_{\text{case}} - \text{mean}(N_{\text{control}})) / SD(N_{\text{control}}) \quad \dots \text{式}(17)$$

但し、Dcase: 遺伝子、タンパク質の濃度等の測定データ

mean(Ncontrol): コントロールサンプルの平均値

SD(Ncontrol): コントロールサンプルの標準偏差

【0108】

そして、ステップs 24で選択した差次的生体分子Dcの中のそれぞれの生体分子を中

50

心ノードとしたローカルネットワークを P P I 等のデータベースを用いて導出する（ステップ s 3 2）

【0109】

また、選択された N 個の中心ノードのそれぞれについて閾値 $d = \{ d_1, \dots, d_N \}$ を決定する（ステップ s 3 3）。ステップ s 3 3 で決定される各ノードの閾値 d は、前述の式（3）および式（4）にて遷移状態を決定するための閾値であり、正常状態である時刻 t_1 にて採取したサンプルに基づいて、正常状態における摂動に対し、各ノード k で、 $p(|z_k(t_1)| > d_k) =$ として規定される が例えば 0.5 となるように決定される。

【0110】

そして、導出した各ローカルネットワークについて、下記の式（18）に示す確率密度関数 f（式（18）の左辺を省略して表記）をそれぞれ計算する（ステップ s 3 4）。なお、ステップ s 3 4 において、確率密度関数 f は、正規化した測定データを、閾値 d にて 2 値化したデータを用いて計算される。

【0111】

【数 9】

$$f_{t_k}(Z) = \frac{1}{(2\pi)^{k/2} |\Sigma(t_k)|^{1/2}} \exp\left(-\frac{1}{2}(Z - u(t_k))^T \Sigma(t_k)^{-1} (Z - u(t_k))\right) \quad \dots \text{式(18)}$$

10

20

【0112】

式（18）は、k 個のノードを有するサンプルが多変量正規分布に従うと仮定して計算した確率密度関数 f である。式（18）において、 $Z = (z_1, \dots, z_N)$ であり、各ローカルネットワークの各時刻 t における平均値 $\mu(t_k) = (\mu_1(t_k), \dots, \mu_N(t_k))$ である。また、 $\Sigma(t_k)$ は、ローカルネットワークにおける共分散行列である。そして、式（18）に示す確率密度関数 f を、それぞれの積分領域で重積分することにより、下記の式（19）に示すように、時刻 t_k における各ローカルネットワークの遷移確率 $p_{u,v}(t_k)$ を計算する（ステップ s 3 5）。

【0113】

【数 10】

$$\begin{aligned} p_{u,v}(t_k) &= \Pr(X(t_k) = A_v | X(t_{k-1}) = A_u) \\ &= \frac{\Pr(X(t_k) = A_v, X(t_{k-1}) = A_u)}{\Pr(X(t_{k-1}) = A_u)} \\ &= \frac{\int \dots \int_{Z \in \Omega_u, \tilde{Z} \in \Omega_v} f_{t_k}(Z, \tilde{Z}) dz_1 \dots dz_N d\tilde{z}_1 \dots d\tilde{z}_N}{\int \dots \int_{Z \in \Omega_u} f_{t_k}(Z) dz_1 \dots dz_N}, \quad \dots \text{式(19)} \end{aligned}$$

30

【0114】

式（19）において、 $Z = (z_1, \dots, z_N)$ 、 $Z \sim = (z \sim_1, \dots, z \sim_N)$ 、u および v は、それぞれ状態 A_u および A_v における積分領域を示す（なお $Z \sim$ は Z の上方に \sim を付けるものである。）。式（19）では、条件付き多変量正規分布またはガウシアンカーネル推定を用いて遷移確率 $p_{u,v}(t_k)$ を求めている。さらに、遷移確率 $p_{u,v}(t_k)$ から下記の式（20）により、時刻 t_k における定常分布 $v(t_k)$ を計算する（ステップ s 3 6）。

【0115】

40

【数 1 1】

$$\sum_v \pi_v(t_k) p_{u,v}(t_k) = \pi_u(t_k) \quad \dots \text{式(20)}$$

【0 1 1 6】

そして、計算により得られた遷移確率 $p_{u,v}(t_k)$ および定常分布 $\pi_v(t_k)$ を用いて下記の式(21)により、時刻 t_k における各ノード i ($i = 1, \dots, N$) を中心とする各ローカルネットワークのエントロピー $H(t_k)$ を計算する(ステップ s 3 7)。

【0 1 1 7】

【数 1 2】

$$H_i(t_k) = - \sum_{u,v} \pi_v(t_k) p_{u,v}(t_k) \log P_{u,v}(t_k) \quad \dots \text{式(21)}$$

【0 1 1 8】

次に、バイオマーカーの候補の選択(図4のステップ s 4)を行う。図7は、実施の形態におけるバイオマーカー候補の選択処理の一例を示すフローチャートである。バイオマーカーの候補の選択処理においては、ステップ s 3 6 の計算に基づき、正常状態である時刻 $t-1$ から疾病前状態か否か判定の対象となる時刻 t までの間の時系列変化から SNE の値が急激に 0 まで減少したローカルネットワークの中心ノードを求め、求めた中心ノードを DNB のメンバーとして記憶する(ステップ s 4 1)。さらに、SNE の値が減少したローカルネットワークの SNE を「SNE グループ」のメンバーとして記憶する(ステップ s 4 2)。ステップ s 4 1 ~ s 4 2 の処理は、ローカルネットワークの SNE の減少の程度が所定の選択基準を超える場合に、当該 SNE に係る因子項目を生体の症状の指標となるバイオマーカーの候補として選択する処理である。なお、ステップ s 4 1 および s 4 2 の記憶は、後述する検出装置の記録部又は記憶部に対して行われる。

【0 1 1 9】

次に、ネットワーク全体の平均 SNE の計算(図4のステップ s 5)を行う。ステップ s 5 では、各時刻において、上記ステップ s 4 2 にて記憶した SNE グループのメンバーの SNE のみを使って、ネットワーク全体の平均 SNE を上記の式(14)により計算する。SNE グループのメンバーの SNE のみを使用することにより、ノイズの混入を防止して精度を向上させることが可能であり、しかも計算量を低減することが可能となる。

【0 1 2 0】

そして、疾病前状態か否かの検出処理(図4のステップ s 6)を行う。具体的には、上記ステップ s 5 で計算した各時刻でのネットワーク全体の SNE の中に、所定の検出基準を超える程度に値が急激に減少した SNE が存在するか否かを判断し、急激に減少した SNE が存在する場合、SNE が減少した時刻で、システムが疾病前状態であると判定する。即ち、疾病前状態であることを検出する。他方、値が急激に減少した SNE が存在しない場合、即ち、減少の程度が所定の検出基準を超えない程度である場合、システムが疾病前状態ではないと判定する。このようにして急激に減少した SNE が存在する場合、医療診断において、疾病前状態である可能性が高い等の診断を下すことを支援し、更なる検査等の診断を行う上での契機とすることができる。

【0 1 2 1】

< 検出装置 >

以上詳述した SNE による疾病前状態の検出方法は、コンピュータを用いた検出装置として本発明を具現化することができる。図8は、本発明に係る検出装置の構成例を示すブロック図である。図8に示す検出装置1は、パーソナルコンピュータ、サーバコンピュータに接続されるクライアントコンピュータ、その他各種コンピュータを用いて実現される

10

20

30

40

50

。検出装置 1 は、制御部 1 0、記録部 1 1、記憶部 1 2、入力部 1 3、出力部 1 4、取得部 1 5、通信部 1 6 等の各種機構を備えている。

【 0 1 2 2 】

制御部 1 0 は、C P U (Central Processing Unit) 等の回路を用いて構成され、検出装置 1 全体を制御する機構である。

【 0 1 2 3 】

記録部 1 1 は、H D D (Hard Disk Drive) 等の磁気記録機構、S S D (Solid State Disk) 等の不揮発性半導体記録機構等の不揮発性の補助記録機構である。記録部 1 1 には、本発明に係る検出プログラム 1 1 a 等の各種プログラムおよびデータが記録されている。さらに、記録部 1 1 には、疾病前状態の検出に用いられる因子項目間の関係を示した関係データベース 1 1 b が記録されている。関係データベース 1 1 b は、例えば、因子項目間の関係としてタンパク質間の相互作用を示した P P I 等のデータベースが用いられる。制御部 1 0 は、関係データベース 1 1 b にアクセスし、記録されている各因子項目間の関係を読み取ることができる。

10

【 0 1 2 4 】

記憶部 1 2 は、S D R A M (Synchronous Dynamic Random Access Memory)、S R A M (Static Random Access Memory) 等の揮発性の主記憶機構である。

【 0 1 2 5 】

入力部 1 3 は、キーボード、マウス等のハードウェア、およびドライバ等のソフトウェアを備える入力機構である。

20

【 0 1 2 6 】

出力部 1 4 は、モニタ、プリンタ等のハードウェア、およびドライバ等のソフトウェアを備える出力機構である。

【 0 1 2 7 】

取得部 1 5 は、外部から各種データを取得する機構である。具体的には、L A N (Local Area Network) 等の内部通信網を介してデータを取り込む L A N ポート、測定機器に接続可能なパラレルケーブル等の専用線に接続するポート等の各種ハードウェア、およびドライバ等のソフトウェアである。

【 0 1 2 8 】

通信部 1 6 は、インターネット等の外部通信網を介してデータを取り込む L A N ポート等の各種ハードウェア、およびドライバ等のソフトウェアである。なお、L A N ポートを用いて取得部 1 5 を構成する場合、取得部 1 5 および通信部 1 6 を共用することも可能である。また、通信部 1 6 は、外部通信網を介して接続されるウェブサーバ等の外部記録装置に記録されている関係データベース 1 6 a から情報を取り込むことができる。即ち、制御部 1 0 は、関係データベース 1 6 a にアクセスし、記録されている各因子項目間の関係を読み取ることができる。

30

【 0 1 2 9 】

そして、記録部 1 1 に記録されている検出プログラム 1 1 a を、記憶部 1 2 に記憶し、制御部 1 0 の制御に基づき実行することにより、コンピュータは、検出プログラム 1 1 a に係る各種手順を実行し、本発明の検出装置 1 として機能する。なお、便宜上、記録部 1 1 および記憶部 1 2 として区別しているが、両者とも各種情報の記録という同様の機能を有するものであり、装置の仕様、運用形態等に応じていずれの機構に記録させるかは、適宜決定することが可能である。

40

【 0 1 3 0 】

図 9 は、本発明に係る検出装置 1 による生体の状態遷移の検出処理の一例を示すフローチャートである。本発明に係る検出装置 1 は、前述の S N E による疾病前状態の検出処理を実行するものである。検出装置 1 の制御部 1 0 は、取得部 1 5 により、生体に関する測定により得られた複数の因子項目についての測定データを取得する (S c 1)。ステップ S c 1 は、図 4 のステップ s 1 として示したハイスループットデータの取得処理に対応する。なお、ここではコンピュータ処理の対象として表現するため、因子項目として表現し

50

ているが、ここでいう因子項目とは、前述の D N B のノードとなり得る遺伝子に関する測定項目、タンパク質に関する測定項目、代謝物に関する測定項目等の測定項目を示している。

【 0 1 3 1 】

制御部 1 0 は、取得した因子項目のそれぞれの測定データが、有意性をもって経時的に変化しているか否かを検定し、検定した結果に基づいて差次的生体分子を選出する (S c 2)。ステップ S c 2 は、図 4 のステップ s 2 として示した差次的生体分子の選出処理に対応する。

【 0 1 3 2 】

したがって、ステップ S c 2 の処理は、制御部 1 0 が、各因子項目の測定データ、並びに因子項目および時系列毎に予め設定されている参照データとの比較結果に基づいて、有意性に関する検定を行い (S c 2 1)、経時変化に有意性があると検定された因子項目を選出する処理 (S c 2 2) を含む。即ち、図 5 に示した各種処理を実行する。なお、検出装置 1 が参照データとして処理するデータとは、コントロールサンプルであり、例えば、最初に取得したサンプルをコントロールサンプルとする等の設定に基づき、検出装置 1 は、当該サンプルに対して参照データとしての取り扱いを行う。

【 0 1 3 3 】

制御部 1 0 は、選出した因子項目の時系列変化の相関関係に基づき求まる各因子項目間の動的な結合関係を示すネットワークにおいて、因子項目毎に隣接する他の因子項目との間の統計力学的な微視的エントロピーとして、ローカルネットワークの S N E を計算する (S c 3)。ステップ S c 3 は、図 4 のステップ s 3 として示したローカルネットワークの S N E の計算処理に対応する。

【 0 1 3 4 】

したがって、ステップ S c 3 の処理は、制御部 1 0 が、関係データベース 1 1 b 又は 1 6 a にアクセスし、記録されている因子項目間の相互作用に基づいて、因子項目間の動的な結合関係を示すネットワークを導出する (S c 3 1)。さらに、制御部 1 0 は、因子項目毎に、以前の摂動に基づいて決定される閾値に対する変化の大小により測定データを 2 値化し (S c 3 2)、2 値化した測定データが多変量正規分布に従うとして確率密度関数を計算し (S c 3 3)、計算した確率密度関数を重積分した遷移確率に基づいて定常分布となる測定データの確率を計算する (S c 3 4)。そして、制御部 1 0 は、各因子項目について、隣接する他の全ての因子項目に係る測定データの状態変化の分布を示す確率密度関数に基づく測定データの確率と該確率の対数との積の総和に基づいて、ローカルネットワークの S N E を計算する (S c 3 5)。即ち、図 6 に示した各種処理を実行する。

【 0 1 3 5 】

制御部 1 0 は、計算したローカルネットワークの S N E の減少の程度が所定の判定基準を超える場合に、当該ローカルネットワークの中心となる因子項目を生体の症状の指標となるバイオマーカーの候補として特定する (S c 4)。ステップ S c 4 は、図 4 のステップ s 3 として示したローカルネットワークの S N E 計算処理のうち、ステップ s 3 3 ~ ステップ s 3 4 の処理に対応する。

【 0 1 3 6 】

したがって、ステップ S c 4 の処理は、制御部 1 0 が、ローカルネットワークの S N E の値が、急激に 0 まで減少した場合に、当該ローカルネットワークの中心となる因子項目を D N B のメンバーとして記録部 1 1 又は記憶部 1 2 に記憶し (S c 4 1)、更に、S N E の値が減少したローカルネットワークの S N E を「 S N E グループ」のメンバーとして記録部 1 1 又は記憶部 1 2 に記憶する (S c 4 2) 処理を含む。

【 0 1 3 7 】

制御部 1 0 は、因子項目毎の微視的エントロピーとして計算した各 S N E に基づいて、選出した因子項目全体の代表値となる巨視的エントロピーを全体の S N E として統計的に計算する (S c 5)。ステップ S c 5 は、図 4 のステップ s 4 として示したネットワーク全体の平均 S N E を、前述の式 (1 4) を用いて計算する処理に対応する。

10

20

30

40

50

【 0 1 3 8 】

そして、制御部 1 0 は、ステップ S c 4 1 にて記憶した因子項目およびステップ S c 4 2 にて記憶した S N E に基づいて、症状変化の予兆として、疾病前状態を検出する (S c 6)。ステップ S c 6 は、図 4 のステップ s 6 の疾病前状態であるか否かの判定処理に対応する。

【 0 1 3 9 】

したがって、ステップ S c 6 の処理は、制御部 1 0 が、ステップ S c 4 で計算した各時刻でのネットワーク全体の S N E の中に、値が急激に減少した S N E が存在するか否かを判断し、急激に減少した S N E が存在する場合、S N E が減少した時刻で、システムが疾病前状態であると判定し、他方、値が急激に減少した S N E が存在しない場合、システムが疾病前状態ではないと判定する処理を含む。

10

【 0 1 4 0 】

そして、制御部 1 0 は、判定結果を含む検出結果を出力部 1 4 から出力し、処理を終了する。また、医師は、出力された検出結果に基づき更なる検査、診察、診断、治療等の行為の要否を判断することができ、患者は、出力された検出結果に基づき自身の身体の状態を把握することが可能となる。

【 0 1 4 1 】

前記実施の形態は、本発明の無数に存在する実施例の一部を開示したに過ぎず、疾病の種類、検出すべき目的等の様々な要因を加味して適宜設計変更することが可能である。特に、因子項目としては、生体に関する測定により得られた情報であれば、様々な測定データを用いることができる。例えば、前述の遺伝子、タンパク質、代謝物に関する測定データに限らず、C T スキャン等の測定装置にて出力される体内画像に基づいて、各部位の様々な状況を数値化することにより、測定データとして用いることが可能である。

20

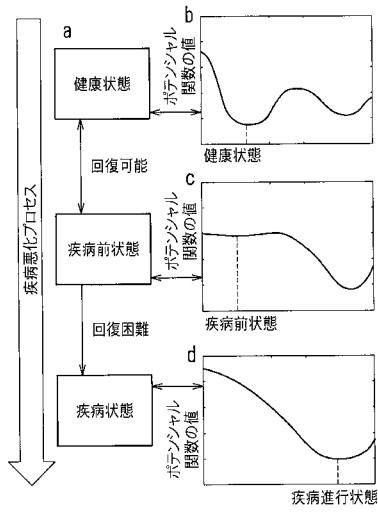
【 符号の説明 】

【 0 1 4 2 】

- 1 検出装置
- 1 0 制御部
- 1 1 記録部
- 1 1 a 検出プログラム
- 1 1 b 関係データベース
- 1 2 記憶部
- 1 3 入力部
- 1 4 出力部
- 1 5 取得部
- 1 6 通信部
- 1 6 a 関係データベース

30

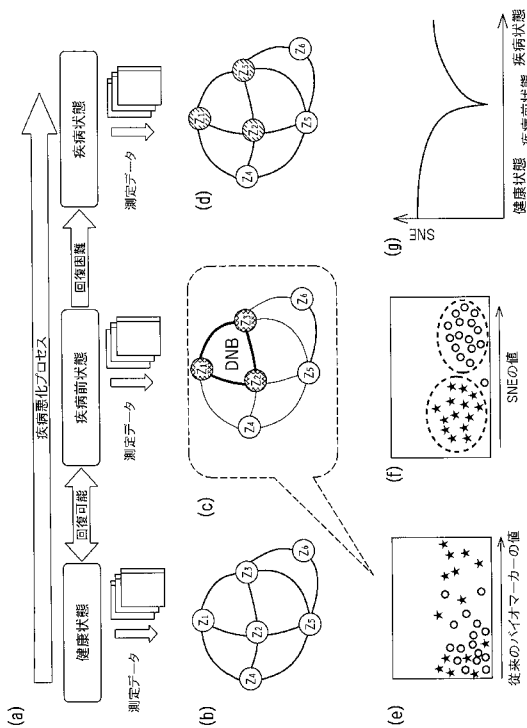
【 図 1 】



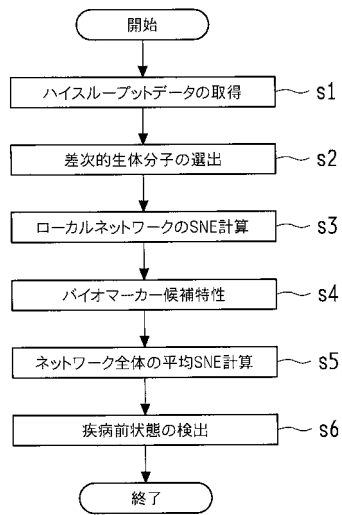
【 図 2 】

タイプ	ノード種類	中心ノードの状態遷移	ローカルSNE
1	DNB	$P(z(t) - z(t-1) > d) \rightarrow 1$	急激に0まで減少
2	DNB	$P(z(t) - z(t-1) > d) \rightarrow 1$	減少
3	非DNB	$P(z(t) - z(t-1) > d) \rightarrow \beta$	減少
4	非DNB	$P(z(t) - z(t-1) > d) \rightarrow \beta$	あまり変化しない

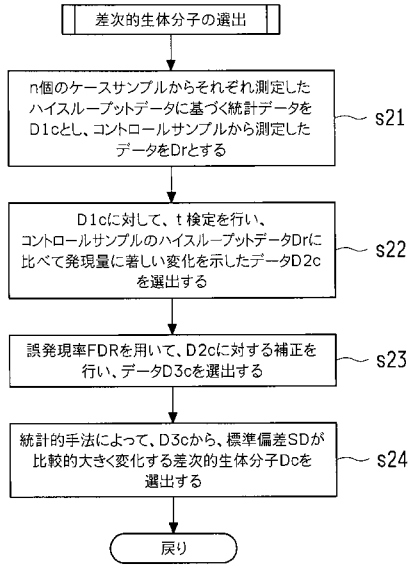
【 図 3 】



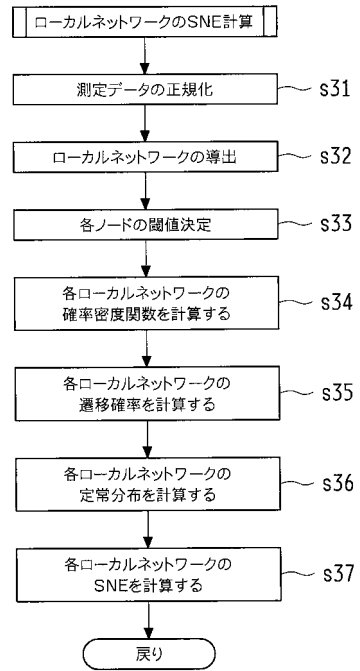
【 図 4 】



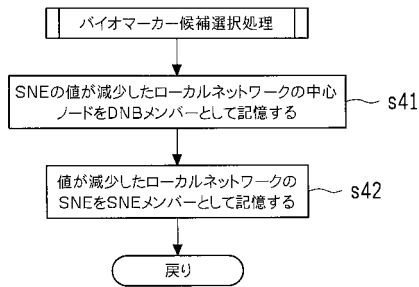
【 図 5 】



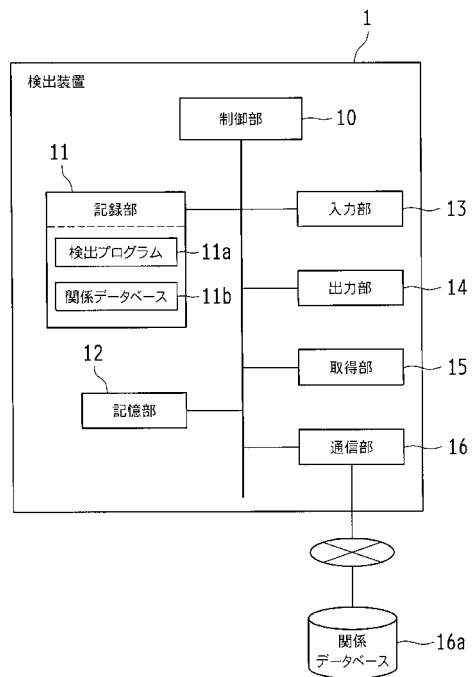
【 図 6 】



【 図 7 】



【 図 8 】



【 図 9 】

