

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号
特表2024-534035
(P2024-534035A)

(43)公表日 令和6年9月18日(2024.9.18)

(51)国際特許分類
G 1 6 H 50/50 (2018.01)

F I
G 1 6 H 50/50

テーマコード (参考)
5 L 0 9 9

審査請求 未請求 予備審査請求 未請求 (全101頁)			
(21)出願番号	特願2024-508966(P2024-508966)	(71)出願人	522454068
(86)(22)出願日	令和4年8月16日(2022.8.16)		インシトロ インコーポレイテッド
(85)翻訳文提出日	令和6年4月8日(2024.4.8)		アメリカ合衆国 カリフォルニア 9 4 0
(86)国際出願番号	PCT/US2022/075006		8 0 , サウス サンフランシスコ, イ
(87)国際公開番号	WO2023/023507		ースト グランド アベニュー 2 7 9 ,
(87)国際公開日	令和5年2月23日(2023.2.23)		スイート 2 0 0
(31)優先権主張番号	63/233,707	(74)代理人	100147485
(32)優先日	令和3年8月16日(2021.8.16)		弁理士 杉村 憲司
(33)優先権主張国・地域又は機関	米国(US)	(74)代理人	230118913
			弁護士 杉村 光嗣
(81)指定国・地域	AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA ,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,A T,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR ,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC, 最終頁に続く	(74)代理人	100226263
			弁理士 中田 未来生
		(72)発明者	フランチェスコ パオロ カサーレ
			アメリカ合衆国 カリフォルニア州 9 4
			0 8 0 サウス サンフランシスコ イー
			最終頁に続く

(54)【発明の名称】 発見プラットフォーム

(57)【要約】

本開示は発見プラットフォームに関するものであり、これには、弱い又は未知の遺伝的駆動因子を伴う複雑な疾患等の着目表現型について研究するために医用画像データを用いるための機械学習手法が含まれる。表現型との関係で着目コバリエントを識別する例示的な方法は次のステップを含む：臨床被験者群から取得された共変量クラスについてのコバリエント情報及び表現型に関する対応する表現型画像データを受信するステップと；表現型画像データを訓練済み教師なし機械学習モデル内へと入力して潜在空間内で複数の埋め込みを取得するステップであって、各埋め込みは表現型画像データにて反映されている表現型状態に対応する、ステップと；臨床被験者群についてのコバリエント情報、複数の埋め込み、及び1つ以上の線形回帰モデルに基づいて、複数の候補コバリエントの各候補コバリエントと表現型状態との間の関連を決定して、着目コバリエントを識別するステップ。

【選択図】 図 1

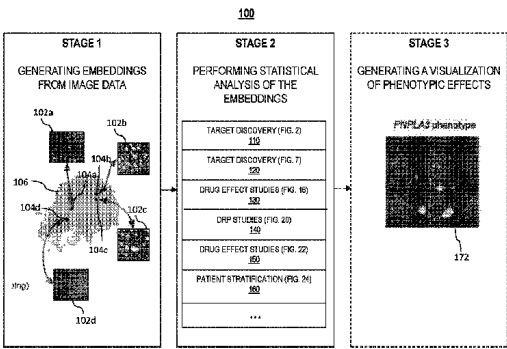


FIG. 1

【特許請求の範囲】**【請求項 1】**

着目疾患の進行に関して治療を評価する方法であって：

被験者プラセボ群にプラセボが施される前に撮像された前記被験者プラセボ群についての複数のベースラインプラセボ医用画像と、前記被験者プラセボ群に前記プラセボが施された後に撮像された前記被験者プラセボ群についての複数のフォローアッププラセボ医用画像とを取得するステップと、

前記複数のベースラインプラセボ医用画像及び前記複数のフォローアッププラセボ医用画像に基づいて複数のプラセボ進行埋め込みを取得するステップと、

被験者治療群に前記治療が施される前に撮像された前記被験者治療群についての複数のベースライン治療医用画像と、前記被験者治療群に前記治療が施された後に撮像された前記被験者治療群についての複数のフォローアップ治療医用画像とを取得するステップと、

前記複数のベースライン治療医用画像及び前記複数のフォローアップ治療医用画像に基づいて複数の治療進行埋め込みを取得するステップと、

前記複数の治療進行埋め込みに基づいて患者が前記プラセボ又は前記治療を受けたのかを決定するための分類モデルを生成するステップとを含む、方法。

【請求項 2】

請求項 1 に記載の方法において、前記分類モデルの出力は薬物反応表現型について示す、方法。

【請求項 3】

請求項 1 ～ 2 のいずれか 1 つに記載の方法において、さらに：

前記分類モデルに基づいて、前記治療と前記着目疾患の前記進行との間の相関メトリックを決定するステップを含む、方法。

【請求項 4】

請求項 3 に記載の方法において、前記相関メトリックは P 値である、方法。

【請求項 5】

請求項 3 ～ 4 のいずれか 1 つに記載の方法において、さらに：

前記相関メトリックを所定の閾値と比較するステップを含む、方法。

【請求項 6】

請求項 5 に記載の方法において、さらに：

前記比較に基づいて、前記治療と前記着目疾患の進行との間の関連を識別するステップを含む、方法。

【請求項 7】

請求項 6 に記載の方法において、さらに：

前記関連に基づいて前記治療を新規被験者について処方するステップを含む、方法。

【請求項 8】

請求項 6 ～ 7 のいずれか 1 つに記載の方法において、さらに：

前記関連に基づいて前記治療を施すステップを含む、方法。

【請求項 9】

請求項 6 ～ 8 のいずれか 1 つに記載の方法において、さらに：

前記関連に基づいて前記治療を調整するステップを含む、方法。

【請求項 10】

請求項 6 ～ 9 のいずれか 1 つに記載の方法において、さらに：

前記関連に基づいて医学的提案を提供するステップを含む、方法。

【請求項 11】

請求項 6 ～ 10 のいずれか 1 つに記載の方法において、さらに：

前記関連に基づいてレポートを生成するステップを含む、方法。

【請求項 12】

請求項 1 ～ 11 のいずれかに記載の方法において、前記着目疾患は非アルコール性脂肪性肝炎（NASH、non-alcoholic steatohepatitis）である、方法。

【請求項 13】

請求項 1～12 のいずれかに記載の方法において、前記複数のプラセボ進行埋め込みを取得するステップは：

前記複数のベースラインプラセボ医用画像を訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数のベースラインプラセボ埋め込みを取得するステップと、

前記複数のフォローアッププラセボ医用画像を前記訓練済み教師なし機械学習モデル内へと入力して前記潜在空間内での複数のフォローアッププラセボ埋め込みを取得するステップと、

前記複数のベースラインプラセボ埋め込みを 1 つ以上の機械学習モデル内へと入力して前記潜在空間内での複数の予測フォローアッププラセボ埋め込みを取得するステップと、

前記複数のフォローアッププラセボ埋め込みと前記複数の予測フォローアッププラセボ埋め込みとの間の差を算出することによって前記複数のプラセボ進行埋め込みを決定するステップを含む、方法。

10

【請求項 14】

請求項 13 に記載の方法において、前記 1 つ以上の機械学習モデルは訓練済み線形モデルを備える、方法。

【請求項 15】

請求項 13～14 のいずれかに記載の方法において、前記 1 つ以上の機械学習モデルは前記訓練済み教師なし機械学習モデルを含む、方法。

【請求項 16】

請求項 13～15 のいずれかに記載の方法において、前記複数の治療進行埋め込みを取得するステップは：

前記複数のベースライン治療医用画像を前記訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数のベースライン治療埋め込みを取得するステップと、

前記複数のフォローアップ治療医用画像を前記訓練済み教師なし機械学習モデル内へと入力して前記潜在空間内での複数のフォローアップ治療埋め込みを取得するステップと、

前記複数のベースライン治療埋め込みを前記訓練済み線形モデル内へと入力して前記潜在空間内での複数の予測フォローアップ治療埋め込みを取得するステップと、

前記複数のフォローアップ治療埋め込みと前記複数の予測フォローアップ治療埋め込みとの間の差を算出することによって前記複数の治療進行埋め込みを決定するステップを含む、方法。

20

30

【請求項 17】

請求項 13～17 のいずれかに記載の方法において、前記訓練済み教師なし機械学習モデルは対照モデルである、方法。

【請求項 18】

請求項 17 に記載の方法において、前記対照モデルは SimCLR モデルである、方法。

【請求項 19】

請求項 13～18 のいずれかに記載の方法において、前記訓練済み線形モデルはベースライン埋め込みを受信し及び予測フォローアップ埋め込みを出力するように構成されている、方法。

40

【請求項 20】

請求項 19 に記載の方法において、前記訓練済み線形モデルは線形混合モデルである、方法。

【請求項 21】

請求項 19 に記載の方法において、前記被験者プラセボ群は第 1 のプラセボ群であり、前記訓練済み線形モデルは前記第 1 のプラセボ群とは異なる第 2 のプラセボ群からの医用画像データを用いて訓練されている、方法。

【請求項 22】

請求項 1～21 のいずれかに記載の方法において、前記分類モデルは、入力進行埋め込みを受信し、また、前記プラセボ又は前記治療を患者が受けたのかを示す分類結果を出力

50

するように構成されている、方法。

【請求項 23】

請求項 1～21 のいずれかに記載の方法において、前記複数のベースラインプラセボ医用画像、前記複数のフォローアッププラセボ医用画像、前記複数のベースライン治療医用画像、及び前記複数のフォローアップ治療画像は、生検医用画像である、方法。

【請求項 24】

治療についての薬物反応表現型（DRP）との関係で着目コバリエントを識別する方法であって：

臨床被験者群から取得された共変量クラスについてのコバリエント情報を受信するステップと、

前記臨床被験者群から複数のベースライン医用画像及び複数のフォローアップ医用画像を受信するステップと、

前記複数のベースライン医用画像及び前記複数のフォローアップ医用画像に基づいて複数の進行埋め込みを取得するステップと、

前記複数の進行埋め込みを訓練済み分類モデル内へと入力して前記臨床被験者群の DRP 値について示す複数の分類結果を取得するステップと、

前記臨床被験者群についての前記コバリエント情報、前記複数の分類結果、及び 1 つ以上の機械学習モデルに基づいて、複数の候補コバリエントの各候補コバリエントと前記 DRP 値との間の関連を決定して、前記着目コバリエントを識別するステップを含む、方法。

【請求項 25】

請求項 24 に記載の方法において、前記 1 つ以上の機械学習モデルは 1 つ以上の線形回帰モデルを備える、方法。

【請求項 26】

請求項 24 に記載の方法において、前記複数の候補コバリエントは複数の候補ミスセンスバリエントを備える、方法。

【請求項 27】

請求項 24 に記載の方法において、前記複数の候補コバリエントは複数の候補遺伝子を備える、方法。

【請求項 28】

請求項 24～27 のいずれかに記載の方法において、前記共変量クラスは人口統計情報、臨床共変量、又はゲノムデータを備える、方法。

【請求項 29】

請求項 24～28 のいずれかに記載の方法において、さらに：

前記識別された着目コバリエントに基づいて新規被験者にて着目疾患を診断するステップを含む、方法。

【請求項 30】

請求項 24～29 のいずれかに記載の方法において、さらに：

前記識別された着目コバリエントに基づいて治療を開発するステップを含む、方法。

【請求項 31】

請求項 24～30 のいずれかに記載の方法において、さらに：

前記識別された着目コバリエントに基づいて前記治療を施す、調整する、又は適用するステップを含む、方法。

【請求項 32】

請求項 24～31 のいずれかに記載の方法において、さらに：

前記識別された着目コバリエントに基づいて医学的提案を提供するステップを含む、方法。

【請求項 33】

請求項 24～32 のいずれかに記載の方法において、さらに：

前記識別された着目コバリエントに基づいて生物学的標的を識別するステップを含む、

10

20

30

40

50

方法。

【請求項 3 4】

請求項 2 4～3 3 のいずれかに記載の方法において、前記複数のベースライン画像及び前記複数のフォローアップ画像は生検画像を備える、方法。

【請求項 3 5】

請求項 2 4～3 4 のいずれかに記載の方法において、前記複数のベースライン医用画像及び前記複数のフォローアップ医用画像に基づいて前記複数の進行埋め込みを取得するステップは：

前記複数のベースライン医用画像を訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数のベースライン埋め込みを取得するステップと、

前記複数のフォローアップ医用画像を前記訓練済み教師なし機械学習モデル内へと入力して前記潜在空間内での複数のフォローアップ埋め込みを取得するステップと、

前記複数のベースライン埋め込みを訓練済み線形モデル内へと入力して前記潜在空間内での複数の予測フォローアップ埋め込みを取得するステップと、

前記複数のフォローアップ埋め込みと前記複数の予測フォローアップ埋め込みとの間の差を算出することによって前記複数の進行埋め込みを決定するステップとを含む、方法。

【請求項 3 6】

請求項 3 5 に記載の方法において、前記訓練済み教師なし機械学習モデルは対照モデルである、方法。

【請求項 3 7】

請求項 3 6 に記載の方法において、前記対照モデルは SimCLR モデルである、方法。

【請求項 3 8】

請求項 3 5 に記載の方法において、前記訓練済み線形モデルはベースライン埋め込みを受信し及び予測フォローアップ埋め込みを出力するように構成されている、方法。

【請求項 3 9】

請求項 3 5 に記載の方法において、前記訓練済み線形モデルは線形混合モデルである、方法。

【請求項 4 0】

請求項 2 4～3 9 のいずれかに記載の方法において、前記訓練済み分類モデルは、入力進行埋め込みを受信し、また、患者がプラセボ又は前記治療を受けたかを決定するように構成されている、方法。

【請求項 4 1】

請求項 2 4～4 0 のいずれかに記載の方法において、前記着目コバリアントを識別するステップは：前記複数の候補コバリアントの候補コバリアントについて：

前記臨床被験者群の DRP 値及び前記コバリアント情報に基づいたモデルを生成するステップと、

前記モデルに基づいて相関メトリックを決定するステップとを含む、方法。

【請求項 4 2】

請求項 4 1 に記載の方法において、前記相関メトリックは P 値である、方法。

【請求項 4 3】

請求項 4 1 に記載の方法において、さらに：

前記相関メトリックを所定の閾値に対して比較して前記候補コバリアントが前記着目コバリアントであるかを決定するステップを含む、方法。

【請求項 4 4】

着目疾患の進行に関して治療を評価する方法であって：

医用画像を取得するステップであって、該医用画像は：（a）被験者プラセボ群にプラセボが施される前に撮像された前記被験者プラセボ群についての複数のベースラインプラセボ医用画像と、（b）前記被験者プラセボ群に前記プラセボが施された後に撮像された前記被験者プラセボ群についての複数のフォローアッププラセボ医用画像と、（c）被験者治療群に前記治療が施される前に撮像された前記被験者治療群についての複数のベース

10

20

30

40

50

ライン治療医用画像と、（d）前記被験者治療群に前記治療が施された後に撮像された前記被験者治療群についての複数のフォローアップ治療医用画像とを備える、ステップと、

前記医用画像を訓練済み教師なし機械学習モデル内へと入力して複数の埋め込みを取得するステップであって、各埋め込みは前記医用画像の1つ以上にて反映されている前記着目疾患との関係での表現型状態に対応する、ステップと、

前記複数の埋め込みを1つ以上の機械学習モデル内へと入力して複数の予測連続医療診断スコアを取得するステップであって、各予測連続医療診断スコアは前記着目疾患の状態について示す、ステップと、

前記複数の予測連続医療診断スコアに基づいて、複数のプラセボ進行スコア及び複数の治療進行スコアを決定するステップと、

前記複数のプラセボ進行スコア及び前記複数の治療進行スコアを前記治療と関連付けるステップと、

前記関連に基づいて、前記複数のプラセボ進行スコアと前記複数の治療進行スコアとの間の相関メトリックを決定するステップとを含む、方法。

【請求項45】

請求項44に記載の方法において、前記医用画像を訓練済み教師なし機械学習モデル内へと入力して前記複数の埋め込みを取得するステップは：

前記（a）を前記訓練済み教師なし機械学習モデル内へと入力して複数のベースラインプラセボ埋め込みを取得することと、

前記（b）を前記訓練済み教師なし機械学習モデル内へと入力して複数のフォローアッププラセボ埋め込みを取得することと、

前記（c）を訓練済み教師なし機械学習モデル内へと入力して複数のベースライン治療埋め込みを取得することと、

前記（d）を前記訓練済み教師なし機械学習モデル内へと入力して複数のフォローアップ治療埋め込みを取得することを含む、方法。

【請求項46】

請求項44～45のいずれかに記載の方法において、前記1つ以上の機械学習モデルは訓練済み線形回帰モデルを備える、方法。

【請求項47】

請求項44～46のいずれか1つに記載の方法において、前記1つ以上の機械学習モデルは前記訓練済み教師なし機械学習モデルを含む、方法。

【請求項48】

請求項46～47のいずれか1つに記載の方法において、前記複数の埋め込みを前記1つ以上の機械学習モデル内へと入力するステップは：

前記複数のベースラインプラセボ埋め込みを前記訓練済み線形回帰モデル内へと入力して複数のベースラインプラセボスコアを取得することと、

前記複数のフォローアッププラセボ埋め込みを前記訓練済み線形回帰モデル内へと入力して複数のフォローアッププラセボスコアを取得することと、

前記複数のベースライン治療埋め込みを前記訓練済み線形回帰モデル内へと入力して複数のベースライン治療スコアを取得することと、

前記複数のフォローアップ治療埋め込みを前記訓練済み線形回帰モデル内へと入力して複数のフォローアップ治療スコアを取得することを含む、方法。

【請求項49】

請求項48に記載の方法において、前記複数のプラセボ進行スコア及び前記複数の治療進行スコアを決定するステップは：

前記複数のベースラインプラセボスコアと前記複数のフォローアッププラセボスコアとの間の差を決定して前記複数のプラセボ進行スコアを決定することと、

前記複数のベースライン治療スコアと前記複数のフォローアップ治療スコアとの間の差を決定して前記複数の治療進行スコアを決定することを含む、方法。

【請求項50】

10

20

30

40

50

請求項 4 8 に記載の方法において、前記複数のプラセボ進行スコア及び前記複数の治療進行スコアを決定するステップは：

前記被験者プラセボ群内の各被験者について、前記被験者プラセボ群内の前記被験者のベースラインプラセボスコア及びフォローアッププラセボスコアに少なくとも基づいて適合化された線形モデルの勾配を決定することと、

前記被験者治療群内の各被験者について、前記被験者治療群内の前記被験者のベースラインプラセボスコア及びフォローアッププラセボスコアに少なくとも基づいて適合化された線形モデルの勾配を決定することを含む、方法。

【請求項 5 1】

請求項 4 4 ～ 5 0 のいずれかに記載の方法において、前記複数の予測医療診断スコアは、予測フォローアップスコアと観測フォローアップスコアとの間の差として対応するベースラインスコアについて調整がなされて算出された疾患進行スコアを含む、方法。

【請求項 5 2】

請求項 4 4 ～ 5 1 のいずれかに記載の方法において、前記複数のプラセボ進行スコア及び前記複数の治療進行スコアを前記治療と関連付けるステップは、患者が前記治療を受けたかについての指示を受信し及び予測疾患進行スコアを出力するように構成されたモデルを生成するステップを含む、方法。

【請求項 5 3】

請求項 5 2 に記載の方法において、前記相関メトリックは前記モデルの P 値である、方法。

【請求項 5 4】

請求項 4 4 ～ 5 2 のいずれかに記載の方法において、さらに：

前記相関メトリックを所定の閾値と比較するステップを含む、方法。

【請求項 5 5】

請求項 5 4 に記載の方法において、さらに：

前記比較に基づいて、前記治療と前記着目疾患との間の関連を識別するステップを含む、方法。

【請求項 5 6】

請求項 5 5 に記載の方法において、さらに：

前記関連に基づいて前記治療を施す、調整する、又は適用するステップを含む、方法。

【請求項 5 7】

請求項 5 5 に記載の方法において、さらに：

前記関連に基づいて医学的提案を提供するステップを含む、方法。

【請求項 5 8】

請求項 4 4 ～ 5 7 のいずれかに記載の方法において、前記着目疾患は非アルコール性脂肪性肝炎（NASH、non-alcoholic steatohepatitis）である、方法。

【請求項 5 9】

請求項 4 4 ～ 5 8 のいずれかに記載の方法において、前記訓練済み教師なし機械学習モデルは対照モデルである、方法。

【請求項 6 0】

請求項 5 9 に記載の方法において、前記対照モデルはSimCLRモデルである、方法。

【請求項 6 1】

請求項 4 6 ～ 6 0 のいずれかに記載の方法において、前記訓練済み線形回帰モデルは線形混合モデルである、方法。

【請求項 6 2】

請求項 4 6 ～ 6 1 のいずれかに記載の方法において、前記訓練済み線形回帰モデルは複数の割り当て済み医療診断スコアに基づいて適合化される、方法。

【請求項 6 3】

請求項 6 2 に記載の方法において、前記複数の割り当て済み医療診断スコアは 1 人以上の医師によって提供される、方法。

10

20

30

40

50

【請求項 6 4】

請求項 6 3 に記載の方法において、前記複数の割り当て済み医療診断スコアの各割り当て済み医療診断スコアは予め定義された値のセットから選択される、方法。

【請求項 6 5】

請求項 4 4 ～ 6 4 のいずれかに記載の方法において、前記複数の予測連続医療診断スコアは、複数の予測線維症スコア、複数の予測小葉内炎症スコア、又は複数の予測脂肪症スコアを含む、方法。

【請求項 6 6】

着目患者サブグループを識別する方法であって：

臨床被験者群から取得された複数の医用画像を訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数の埋め込みを取得するステップと、

前記複数の埋め込みをクラスタリングして 1 つ以上の埋め込みクラスタを生成するステップと、

前記 1 つ以上の埋め込みクラスタに対応する 1 つ以上の患者サブグループを識別するステップと、

前記 1 つ以上の患者サブグループの各患者サブグループをコバリエントと関連付けて前記着目患者サブグループを識別するステップとを含む、方法。

【請求項 6 7】

請求項 6 6 に記載の方法において、前記訓練済み教師なし機械学習モデルは対照モデルである、方法。

【請求項 6 8】

請求項 6 7 に記載の方法において、前記対照モデルは SimCLR モデルである、方法。

【請求項 6 9】

請求項 6 6 ～ 6 8 のいずれかに記載の方法において、前記コバリエントは着目治療であり、また、前記着目患者サブグループは前記着目治療が相当なインパクトを及ぼすサブグループである、方法。

【請求項 7 0】

請求項 6 9 に記載の方法において、前記 1 つ以上の患者サブグループの各患者サブグループを前記コバリエントと関連付けるステップは：

患者サブグループについて、前記患者サブグループ内の患者が前記着目治療を受けたかについての指示を受信し及び予測疾患進行を出力するように構成されたモデルを、生成するステップと、

前記モデルを評価して前記患者サブグループが前記着目患者サブグループであるかを決定するステップとを含む、方法。

【請求項 7 1】

請求項 7 0 に記載の方法において、前記モデルを評価するステップは、前記モデルの相関メトリックを決定すること及び前記相関メトリックを所定の閾値に対して比較することを含む、方法。

【請求項 7 2】

請求項 7 1 に記載の方法において、前記相関メトリックは P 値である、方法。

【請求項 7 3】

請求項 7 0 ～ 7 2 のいずれかに記載の方法において、前記生成されたモデルは前記患者サブグループ内の被験者の疾患進行値によって訓練される、方法。

【請求項 7 4】

請求項 7 3 に記載の方法において、前記疾患進行値は前記患者サブグループ内の前記被験者の医療診断スコアを含む、方法。

【請求項 7 5】

請求項 7 3 に記載の方法において、前記疾患進行値は前記患者サブグループ内の前記被験者の進行スコアを含む、方法。

【請求項 7 6】

10

20

30

40

50

請求項 7 3 に記載の方法において、前記疾患進行値は前記患者サブグループ内の前記被験者の D R P 値を含む、方法。

【請求項 7 7】

請求項 6 6 ～ 6 8 のいずれかに記載の方法において、前記コバリエントは着目疾患の進行であり、また、前記着目患者サブグループは前記着目疾患の前記進行と有意な関連を有するサブグループである、方法。

【請求項 7 8】

請求項 7 7 に記載の方法において、前記 1 つ以上の患者サブグループの各患者サブグループを前記コバリエントと関連付けるステップは：

患者サブグループについて、患者が前記患者サブグループに属するかについての指示を受信し及び予測疾患進行を出力するように構成されたモデルを、生成するステップと、

前記モデルを評価して前記患者サブグループが前記着目患者サブグループであるかを決定するステップとを含む、方法。

【請求項 7 9】

請求項 7 8 に記載の方法において、前記モデルを評価するステップは、前記モデルの相関メトリックを決定すること及び前記相関メトリックを所定の閾値に対して比較することを含む、方法。

【請求項 8 0】

請求項 7 9 に記載の方法において、前記相関メトリックは P 値である、方法。

【請求項 8 1】

請求項 7 8 ～ 8 0 のいずれかに記載の方法において、前記生成されたモデルは前記臨床被験者群の疾患進行値によって訓練される、方法。

【請求項 8 2】

請求項 8 1 に記載の方法において、前記疾患進行値は、前記患者サブグループ内の臨床被験者の医療診断スコア、前記患者サブグループ内の臨床被験者の進行スコア、又は前記患者サブグループ内の臨床被験者の D R P 値を含む、方法。

【請求項 8 3】

請求項 6 6 ～ 6 8 のいずれかに記載の方法において、前記コバリエントは有害副作用であり、また、前記着目患者サブグループは前記有害副作用と有意な関連を有するサブグループである、方法。

【請求項 8 4】

請求項 8 3 に記載の方法において、前記 1 つ以上の患者サブグループの各患者サブグループを前記コバリエントと関連付けるステップは：

患者サブグループについて、前記患者サブグループ内の患者が前記患者サブグループに属するかについての指示を受信し及び前記患者に前記有害副作用が生じるかを予測するように構成されたモデルを、生成するステップと、

前記モデルを評価して前記患者サブグループが前記着目患者サブグループであるかを決定するステップとを含む、方法。

【請求項 8 5】

請求項 8 4 に記載の方法において、前記モデルを評価するステップは、前記モデルの相関メトリックを決定すること及び前記相関メトリックを所定の閾値に対して比較することを含む、方法。

【請求項 8 6】

請求項 8 5 に記載の方法において、前記相関メトリックは P 値である、方法。

【請求項 8 7】

請求項 6 6 ～ 6 8 のいずれかに記載の方法において、前記コバリエントは有害副作用であり、また、前記着目患者サブグループは治療後に前記有害副作用を経験することについて有意な関連を有するサブグループである、方法。

【請求項 8 8】

請求項 8 7 に記載の方法において、前記 1 つ以上の患者サブグループの各患者サブグル

10

20

30

40

50

ープを前記コバリエントと関連付けるステップは：

患者サブグループについて、前記患者サブグループ内の患者が前記治療を受けたかについての指示を受信し及び前記患者に前記有害副作用が生じるかを予測するように構成されたモデルを、生成するステップと、

前記モデルを評価して前記患者サブグループが前記着目患者サブグループであるかを決定するステップとを含む、方法。

【発明の詳細な説明】

【技術分野】

【0001】

概して本開示は発見プラットフォームに関するものであり、より具体的には、弱い又は未知の遺伝的駆動因子を伴う複雑な疾患等の着目表現型について研究するために医用画像データをを用いるための機械学習手法に関する。

【関連出願の相互参照】

【0002】

本願は、2021年8月16日に出願された米国仮特許出願第63／233，707号の優先権の利益を主張するのであり、その内容の全体が参照によって取り込まれる。

【背景技術】

【0003】

人間を悩ます多くの疾患は、部分的には遺伝子によって左右される。具体的には、1つ以上の遺伝的バリエーション又は変異体を有する個体は、疾患について発病し易かったり或いはその発病後に病勢進展がより速まり得る。特に、複数の遺伝的バリエーションによって影響を受けることのある疾患については、特定の遺伝的背景を有する対象者に疾患がもたらされ得る原因となる具体的な因子を識別すること並びに発症後その対象者においてその疾患が進行していくこととなろう様相を予測することが困難であると認識されている。また、他の点を除いて疾患をもち得る対象者での軽微な遺伝的バリエーションによって、効果面及び安全性面の両方に関して、所定の治療的介入に対して欠く対象者がどの程度良好な応答を示すかに影響がもたらされ得る。したがって、特定の遺伝的バリエーションが疾患の発生及び進行にどのように影響するかについての理解、転じては候補治療が疾患の退行又は臨床対象者における有害反応に関してどのように影響するかについての理解が欠けると、効果的な薬剤の開発に関しての相当の障害となり得る。

【0004】

組織病理画像のコンピュータ分析における近時の進展により、特定の疾患を患っている様々な人間の対象にて発現されている疾患表現型についてより理解が深まった。対象者の遺伝的バリエーションについての、また、疾患感受性リスクにどのような影響を複数の遺伝的バリエーションが与え得るかについてのコンピュータ分析における近時の進展により、特定の疾患の精密な遺伝的根拠について理解がより良くなった。されども、これらのタイプの進展をより良く活用していくための余地は残っており、疾患標的を識別したり、疾患の発現や人間の対象における疾患進行の様相を予測したり、異なる遺伝的背景を有する対象者間での治療的候補に対しての起こり得る応答を予測したり、特定の治療を受けるべき適切な患者コホートを識別したり、結果を最適化するために臨床試験を一般的に設計するための活用をなし得る。

【発明の概要】

【0005】

表現型との関係で着目コバリエントを識別する例示的な方法は次のステップを含む：臨床被験者群から取得された共変量クラスについてのコバリエント情報及び表現型に関する対応する表現型データを受信するステップと；表現型データを訓練済み教師なし機械学習モデル内へと入力して潜在空間内で複数の埋め込みを取得するステップであって、各埋め込みは表現型データにて反映されている表現型状態に対応する、ステップと；臨床被験者群についてのコバリエント情報、複数の埋め込み、及び1つ以上の線形回帰モデルに基づいて、複数の候補コバリエントの各候補コバリエントと表現型との間の関連を決定して、

着目コバリエントを識別するステップ。

【0006】

一部の実施形態では、前記表現型は着目疾患、遺伝子発現、メタボロミクス、プロテオミクス、又はリピドミクスを含む。

【0007】

一部の実施形態では、前記表現型データは医用画像データ、組織病理データ、臨床バイオマーカーデータ、又はゲノミックバイオマーカーデータを含む。

【0008】

一部の実施形態では、前記共変量クラスは人口統計情報、臨床共変量、又はゲノムデータを備える。

【0009】

一部の実施形態では、各候補コバリエント (covariant) と前記表現型との間の前記関連 (association) を決定することは次のステップを含む：前記複数の埋め込みの各埋め込みを線形回帰モデル内へと入力して、複数の埋め込みの各埋め込みについて予測連続スコアを受信して、複数の連続スコアを取得するステップと、前記複数の予測連続スコアを、前記臨床被験者群によって発現されている候補コバリエントと関連付ける（例えば、関連について試す）ステップと、前記表現型と前記候補コバリエントとの間の相関メトリックを、前記関連に基づいて、決定するステップであって、前記相関メトリックは前記候補コバリエントが前記表現型に対して及ぼすインパクトについて示す、ステップ。

【0010】

一部の実施形態では、前記候補コバリエントと前記表現型との間の前記関連を決定することは次のステップを含む：前記複数の埋め込みを前記複数の候補コバリエントの各候補コバリエントと関連付けて、前記複数の候補コバリエントのサブセットを識別するステップと、前記サブセット内の各候補コバリエントを前記表現型と関連付けて、前記少なくとも1つの着目コバリエントを識別するステップ。

【0011】

一部の実施形態では、方法はさらに次のステップを含む：前記少なくとも1つの着目遺伝的バリエントに基づいて、前記表現型を表す複数のシミュレーション画像を生成するステップと、ディスプレイ上に複数のシミュレーション画像を表示するステップ。

【0012】

一部の実施形態では、方法はさらに次のステップを含む：前記少なくとも1つの着目遺伝的バリエント (variant) と前記表現型との間の関係 (relationship) を識別するステップ。

【0013】

一部の実施形態では、前記関係は因果関係である。

【0014】

一部の実施形態では、方法はさらに次のステップを含む：前記関係に基づいて新規被験者について診断を提供するステップ。

【0015】

一部の実施形態では、方法はさらに次のステップを含む：前記関係に基づいて治療を開発するステップ。

【0016】

一部の実施形態では、方法はさらに次のステップを含む：前記関係に基づいて治療を施す、調整する、又は適用するステップ。

【0017】

一部の実施形態では、方法はさらに次のステップを含む：前記関係に基づいて医学的提案を提供するステップ。

【0018】

一部の実施形態では、方法はさらに次のステップを含む：前記関係に基づいて前記着目疾患の治療のための生物学的標的を識別するステップ。

10

20

30

40

50

【0019】

一部の実施形態では、前記着目疾患は非アルコール性脂肪性肝炎（NASH、non-alcoholic steatohepatitis）である。

【0020】

着目疾患との関係で少なくとも1つの着目遺伝的バリエントを識別する例示的な方法は次のステップを含む：臨床被験者群から取得された複数の医用画像を訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数の埋め込みを取得するステップであって、各埋め込みは前記複数の医用画像の1つ以上に反映されている前記着目疾患との関係での表現型状態に対応する、ステップと、前記複数の埋め込みの各埋め込みを訓練済み線形回帰モデル内へと入力して、前記複数の埋め込みの各埋め込みについて予測連続医療診断スコアを受信して、複数の予測医療診断スコアを取得するステップであって、各予測連続医療診断スコアは前記着目疾患の状態について示す、ステップと、前記複数の予測連続医療診断スコアを、複数の医用画像の取得元の臨床被験者群によって発現されている複数の候補遺伝的バリエントの各候補遺伝的バリエントと、関連付けるステップと、前記着目疾患と各候補遺伝的バリエントとの間の相関メトリックを前記関連に基づいて決定し、前記少なくとも1つの着目遺伝的バリエントを前記複数の候補遺伝的バリエントから識別するステップであって、前記相関メトリックは前記候補遺伝的バリエントが前記着目疾患に対して及ぼすインパクトについて示す、ステップ。

10

【0021】

一部の実施形態では、方法はさらに次のステップを含む：前記相関メトリックを所定の閾値と比較するステップ。

20

【0022】

一部の実施形態では、方法はさらに次のステップを含む：前記比較に基づいて、前記着目遺伝的バリエントと前記着目疾患との間の関係を識別するステップ。

【0023】

一部の実施形態では、前記関係は因果関係である。

【0024】

一部の実施形態では、方法はさらに次のステップを含む：前記関係に基づいて新規被験者にて前記着目疾患を診断するステップ。

【0025】

一部の実施形態では、方法はさらに次のステップを含む：前記関係に基づいて治療を開発するステップ。

30

【0026】

一部の実施形態では、方法はさらに次のステップを含む：前記関係に基づいて治療を施す、調整する、又は適用するステップ。

【0027】

一部の実施形態では、方法はさらに次のステップを含む：前記関係に基づいて医学的提案を提供するステップ。

【0028】

一部の実施形態では、方法はさらに次のステップを含む：前記関係に基づいて前記着目疾患の治療のための生物学的標的を識別するステップ。

40

【0029】

一部の実施形態では、前記着目疾患は非アルコール性脂肪性肝炎（NASH、non-alcoholic steatohepatitis）である。

【0030】

一部の実施形態では、前記複数の医用画像は生検画像を備える。

【0031】

一部の実施形態では、前記生検画像は1つ以上の臨床試験に対応する。

【0032】

一部の実施形態では、方法はさらに次のステップを含む：前記複数の画像の医用画像を

50

複数の画像タイルに分割するステップと、前記複数の画像タイルの各画像タイルを前記教師なし機械学習モデル内へと入力して、各画像タイルについてタイル埋め込みを受信して、複数のタイル埋め込みを取得するステップと、前記タイル埋め込みを集約して前記複数の埋め込みの埋め込みを取得するステップ。

【0033】

一部の実施形態では、前記タイル埋め込みを集約するステップは前記タイル埋め込みを平均化するステップを含む。

【0034】

一部の実施形態では、前記教師なし機械学習モデルは対照モデルである。

【0035】

一部の実施形態では、前記対照モデルはSimCLRモデルである。

【0036】

一部の実施形態では、前記訓練済み教師なし機械学習モデルは少なくとも部分的には前記複数の医用画像に基づいて訓練されている。

【0037】

一部の実施形態では、前記教師なし機械学習モデルは前記複数の医用画像に基づいて微調整されている。

【0038】

一部の実施形態では、前記線形回帰モデルは線形混合モデルである。

【0039】

一部の実施形態では、前記訓練済み線形回帰モデルは前記複数の埋め込み及び前記複数の埋め込みに対応する複数の割り当て済み医療診断スコアに基づいて適合化される。

【0040】

一部の実施形態では、前記複数の割り当て済み医療診断スコアは1人以上の医師によって提供される。

【0041】

一部の実施形態では、前記複数の割り当て済み医療診断スコアの各割り当て済み医療診断スコアは予め定義された値のセットから選択される。

【0042】

一部の実施形態では、前記複数の予測連続医療診断スコアは、複数の予測線維症スコア、複数の予測小葉内炎症スコア、又は複数の予測脂肪症スコアとされる。

【0043】

一部の実施形態では、前記複数の予測医療診断スコアは、臨床試験中の別個の測定での予測医療診断スコアの差として取得された疾患進行を含む。

【0044】

一部の実施形態では、前記複数の予測医療診断スコアは、前記臨床試験中に取得された別個の測定を反映する予測医療診断スコア間の差として算出された疾患進行スコアを含む。

【0045】

一部の実施形態では、前記複数の予測医療診断スコアは、前記臨床試験中に各個人について取得された別個の測定を反映する予測医療診断スコアについて訓練された線形モデルによって決定された勾配として取得された疾患進行スコアを含む。

【0046】

一部の実施形態では、前記バリエーション特有モデルは線形モデルである。

【0047】

一部の実施形態では、前記バリエーション特有モデルは、前記複数の予測医療診断スコアと前記候補遺伝的バリエーションについて示す複数の値とに基づいて適合化される。

【0048】

一部の実施形態では、前記関連メトリックを決定するステップは、前記バリエーション特有モデルに基づいてP値を決定するステップを含む。

10

20

30

40

50

【0049】

着目疾患との関係で少なくとも1つの着目遺伝的バリエーションを識別する例示的な方法は次のステップを含む：臨床被験者群から取得された複数の医用画像を訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数の埋め込みを取得するステップであって、各埋め込みは前記複数の医用画像の1つ以上にて反映されている前記着目疾患との関係での表現型状態に対応する、ステップと、前記複数の埋め込みを複数の候補遺伝的バリエーションの各候補遺伝的バリエーションと関連付けて、前記複数の候補遺伝的バリエーションについてのサブセットを識別するステップであって、前記複数の遺伝的バリエーションのサブセットは前記複数の医用画像にて反映されている組織学的特徴と関連付けられている、ステップと、前記複数の候補遺伝的バリエーションの前記サブセットの各候補遺伝的バリエーションを前記着目疾患と関連付けて少なくとも1つの着目遺伝的バリエーションを前記サブセットから識別するステップとを含む、方法。

10

【0050】

一部の実施形態では、方法はさらに次のステップを含む：前記少なくとも1つの着目遺伝的バリエーションに基づいて、前記着目疾患を表す複数のシミュレーション画像を生成するステップと、ディスプレイ上に複数のシミュレーション画像を表示するステップ。

【0051】

一部の実施形態では、方法はさらに次のステップを含む：前記少なくとも1つの着目遺伝的バリエーションと前記着目疾患との間の関係を識別するステップ。

【0052】

一部の実施形態では、前記関係は因果関係である。

20

【0053】

一部の実施形態では、方法はさらに次のステップを含む：前記関係に基づいて新規被験者にて前記着目疾患を診断するステップ。

【0054】

一部の実施形態では、方法はさらに次のステップを含む：前記関係に基づいて治療を開発するステップ。

【0055】

一部の実施形態では、方法はさらに次のステップを含む：前記関係に基づいて治療を施す、調整する、又は適用するステップ。

30

【0056】

一部の実施形態では、方法はさらに次のステップを含む：前記関係に基づいて医学的提案を提供するステップ。

【0057】

一部の実施形態では、方法はさらに次のステップを含む：前記関係に基づいて前記着目疾患の治療のための生物学的標的を識別するステップ。

【0058】

一部の実施形態では、前記着目疾患は非アルコール性脂肪性肝炎（NASH、non-alcoholic steatohepatitis）である。

【0059】

一部の実施形態では、前記複数の医用画像は生検画像を備える。

40

【0060】

一部の実施形態では、前記生検画像は1つ以上の臨床試験に対応する。

【0061】

一部の実施形態では、方法はさらに次のステップを含む：前記複数の画像の医用画像を複数の画像タイルに分割するステップと、前記複数の画像タイルの各画像タイルを前記教師なし機械学習モデル内へと入力して、各画像タイルについてタイル埋め込みを受信して、複数のタイル埋め込みを取得するステップと、前記タイル埋め込みを集約して前記複数の埋め込みの埋め込みを取得するステップ。

【0062】

50

一部の実施形態では、前記タイル埋め込みを集約するステップは前記タイル埋め込みを平均化するステップを含む。

【0063】

一部の実施形態では、前記教師なし機械学習モデルは対照モデルである。

【0064】

一部の実施形態では、前記対照モデルはSimCLRモデルである。

【0065】

一部の実施形態では、前記訓練済み教師なし機械学習モデルは少なくとも部分的には前記複数の医用画像に基づいて訓練されている。

【0066】

一部の実施形態では、前記教師なし機械学習モデルは前記複数の医用画像に基づいて微調整されている。

【0067】

一部の実施形態では、前記複数の埋め込みを前記複数の遺伝的バリエーションの各遺伝的バリエーションと関連付けて前記複数の遺伝的バリエーションについての前記サブセットを識別するステップは：埋め込みを受信し及び前記候補遺伝的バリエーションの値を出力するように構成されたバリエーション特有モデルを、前記複数の候補遺伝的バリエーションの候補遺伝的バリエーションについて、生成するステップと、前記バリエーション特有モデルを評価して前記候補遺伝的バリエーションを前記サブセット内に含めるべきかを決定するステップ。

【0068】

一部の実施形態では、バリエーション特有モデルを評価することは次のステップを含む：前記バリエーション特有モデルに基づいて相関メトリックを算出するステップと、前記相関メトリックを所定の閾値と比較するステップ。

【0069】

一部の実施形態では、前記相関メトリックは前記バリエーション特有モデルと関連付けられているP値である。

【0070】

一部の実施形態では、前記複数の遺伝的バリエーションの前記サブセットの各遺伝的バリエーションを前記着目疾患と関連付けて前記少なくとも1つの着目遺伝的バリエーションを識別するステップは：前記遺伝的バリエーションについて示す値を受信し及び前記着目疾患に関する医療診断スコアを出力するように構成されたバリエーション特有モデルを、前記サブセット内の遺伝的バリエーションについて、生成するステップと、前記バリエーション特有モデルを評価して前記候補遺伝的バリエーションが前記少なくとも1つの着目遺伝的バリエーションであるかを決定するステップ。

【0071】

一部の実施形態では、バリエーション特有モデルを評価することは次のステップを含む：前記バリエーション特有モデルに基づいて相関メトリックを算出するステップと、前記相関メトリックを所定の閾値と比較するステップ。

【0072】

一部の実施形態では、前記相関メトリックは前記バリエーション特有スコア予測モデルと関連付けられているP値である。

【0073】

着目疾患の進行に関して治療を評価する例示的な方法は次のステップを含む：プラセボ群にプラセボが施される前に撮像された被験者プラセボ群についての複数のベースラインプラセボ画像と、前記プラセボ群に前記プラセボが施された後に撮像された前記被験者プラセボ群についての複数のフォローアッププラセボ画像とを取得するステップと、前記複数のベースラインプラセボ画像及び前記複数のフォローアッププラセボ画像に基づいて複数のプラセボ進行埋め込みを取得するステップと、前記治療群に前記治療が施される前に撮像された被験者治療群についての複数のベースライン治療画像と、前記治療群に前記治療が施された後に撮像された前記被験者治療群についての複数のフォローアップ治療画像

10

20

30

40

50

とを取得するステップと、前記複数のベースライン治療画像及び前記複数のフォローアップ治療画像に基づいて複数の治療進行埋め込みを取得するステップと、前記複数の治療進行埋め込みに基づいて患者が前記プラセボ又は前記治療を受けたのかを決定するための分類モデルを生成するステップであって、前記分類モデルの出力は薬物反応組織学的表現型（DRP）について示す、ステップと、前記分類モデルに基づいて、前記治療と前記着目疾患の前記進行との間の相関メトリックを決定するステップ。

【0074】

一部の実施形態では、前記相関メトリックはP値である。

【0075】

一部の実施形態では、方法はさらに次のステップを含む：前記相関メトリックを所定の閾値と比較するステップ。

【0076】

一部の実施形態では、方法はさらに次のステップを含む：前記比較に基づいて、前記治療と前記着目疾患の進行との間の関連を識別するステップ。

【0077】

一部の実施形態では、方法はさらに次のステップを含む：前記関連に基づいて前記治療を新規被験者について処方するステップ。

【0078】

一部の実施形態では、方法はさらに次のステップを含む：前記関連に基づいて前記治療を施すステップ。

【0079】

一部の実施形態では、方法はさらに次のステップを含む：前記関連に基づいて前記治療を調整するステップ。

【0080】

一部の実施形態では、方法はさらに次のステップを含む：前記関連に基づいて医学的提案を提供するステップ。

【0081】

一部の実施形態では、方法はさらに次のステップを含む：前記関連に基づいてレポートを生成するステップ。

【0082】

一部の実施形態では、前記着目疾患は非アルコール性脂肪性肝炎（NAFLD、non-alcoholic steatohepatitis）である。

【0083】

一部の実施形態では、前記複数のプラセボ進行埋め込みを取得するステップは：前記複数のベースラインプラセボ画像を訓練済み教師なし機械学習モデル内へと入力して潜在空間内の複数のベースラインプラセボ埋め込みを取得するステップと、前記複数のフォローアッププラセボ画像を前記訓練済み教師なし機械学習モデル内へと入力して前記潜在空間内の複数のフォローアッププラセボ埋め込みを取得するステップと、前記複数のベースラインプラセボ埋め込みを訓練済み線形モデル内へと入力して前記潜在空間内の複数の予測フォローアッププラセボ埋め込みを取得するステップと、前記複数のフォローアッププラセボ埋め込みと前記複数の予測フォローアッププラセボ埋め込みとの間の差を算出することによって前記複数のプラセボ進行埋め込みを決定するステップ。

【0084】

一部の実施形態では、前記複数の治療進行埋め込みを取得するステップは：前記複数のベースライン治療画像を前記訓練済み教師なし機械学習モデル内へと入力して潜在空間内の複数のベースライン治療埋め込みを取得するステップと、前記複数のフォローアップ治療画像を前記訓練済み教師なし機械学習モデル内へと入力して前記潜在空間内の複数のフォローアップ治療埋め込みを取得するステップと、前記複数のベースライン治療埋め込みを前記訓練済み線形モデル内へと入力して前記潜在空間内の複数の予測フォローアップ治療埋め込みを取得するステップと、前記複数のフォローアップ治療埋め込みと前記

10

20

30

40

50

複数の予測フォローアップ治療埋め込みとの間の差を算出することによって前記複数の治療進行埋め込みを決定するステップ。

【0085】

一部の実施形態では、前記教師なし機械学習モデルは対照モデルである。

【0086】

一部の実施形態では、前記対照モデルはSimCLRモデルである。

【0087】

一部の実施形態では、前記訓練済み線形モデルはベースライン埋め込みを受信し及び予測フォローアップ埋め込みを出力するように構成されている。

【0088】

一部の実施形態では、前記訓練済み線形モデルは線形混合モデルである。

【0089】

一部の実施形態では、前記プラセボ群は第1のプラセボ群であり、前記線形モデルは前記第1のプラセボ群とは異なる第2のプラセボ群からの画像データを用いて訓練されている。

【0090】

一部の実施形態では、前記分類モデルは、入力進行埋め込みを受信し、また、前記プラセボ又は前記治療を患者が受けたのかを示す分類結果を出力するように構成されている。

【0091】

一部の実施形態では、前記複数のベースラインプラセボ画像、前記複数のフォローアッププラセボ画像、前記複数のベースライン治療画像、及び前記複数のフォローアップ治療画像は、生検画像である。

【0092】

治療についての薬物反応組織学的表現型（DRP）との関係で着目コバリエントを識別する例示的な方法は次のステップを含む：臨床被験者群から取得された共変量クラスについてのコバリエント情報を受信するステップと、前記臨床被験者群から複数のベースライン画像及び複数のフォローアップ画像を受信するステップと、前記複数のベースライン画像及び前記複数のフォローアップ画像に基づいて複数の進行埋め込みを取得するステップと、前記複数の進行埋め込みを訓練済み分類モデル内へと入力して前記臨床被験者群のDRP値について示す複数の分類結果を取得するステップと、前記臨床被験者群についての前記コバリエント情報、前記複数の分類結果、及び1つ以上の線形回帰モデルに基づいて、複数の候補コバリエントの各候補コバリエントと前記DRP値との間の関連を決定して、前記着目コバリエントを識別するステップ。

【0093】

一部の実施形態では、前記複数の候補コバリエントは複数の候補ミスセンスバリエントを備える。

【0094】

一部の実施形態では、前記複数の候補コバリエントは複数の候補遺伝子を備える。

【0095】

一部の実施形態では、前記共変量クラスは人口統計情報、臨床共変量、又はゲノムデータを備える。

【0096】

一部の実施形態では、方法はさらに次のステップを含む：前記識別された着目コバリエントに基づいて新規被験者にて着目疾患を診断するステップを含む、方法。

【0097】

一部の実施形態では、方法はさらに次のステップを含む：前記識別された着目コバリエントに基づいて治療を開発するステップ。

【0098】

一部の実施形態では、方法はさらに次のステップを含む：前記識別された着目コバリエントに基づいて前記治療を施す、調整する、又は適用するステップ。

10

20

30

40

50

【0099】

一部の実施形態では、方法はさらに次のステップを含む：前記識別された着目コバリエントに基づいて医学的提案を提供するステップ。

【0100】

一部の実施形態では、方法はさらに次のステップを含む：前記識別された着目コバリエントに基づいて生物学的標的を識別するステップ。

【0101】

一部の実施形態では、前記複数の医用画像は生検画像を備える。

【0102】

一部の実施形態では、前記複数のベースライン画像及び前記複数のフォローアップ画像に基づいて前記複数の進行埋め込みを取得するステップは次のステップを含む：前記複数のベースライン医用画像を訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数のベースライン埋め込みを取得するステップと、前記複数のフォローアップ医用画像を前記訓練済み教師なし機械学習モデル内へと入力して前記潜在空間内での複数のフォローアップ埋め込みを取得するステップと、前記複数のベースライン埋め込みを訓練済み線形モデル内へと入力して前記潜在空間内での複数の予測フォローアップ埋め込みを取得するステップと、前記複数のフォローアップ埋め込みと前記複数の予測フォローアップ埋め込みとの間の差を算出することによって前記複数の進行埋め込みを決定するステップ。

10

【0103】

一部の実施形態では、前記教師なし機械学習モデルは対照モデルである。

20

【0104】

一部の実施形態では、前記対照モデルはSimCLRモデルである。

【0105】

一部の実施形態では、前記訓練済み線形モデルはベースライン埋め込みを受信し及び予測フォローアップ埋め込みを出力するように構成されている。

【0106】

一部の実施形態では、前記訓練済み線形モデルは線形混合モデルである。

【0107】

一部の実施形態では、前記訓練済み分類モデルは、入力進行埋め込みを受信し、また、患者がプラセボ又は前記治療を受けたかを決定するように構成されている。

30

【0108】

一部の実施形態では、前記着目コバリエントを識別するステップは：前記複数の候補コバリエントの候補コバリエントについて：前記臨床被験者群のDRP値及び前記コバリエント情報に基づいたモデルを生成するステップと、前記モデルに基づいて相関メトリックを決定するステップ。

【0109】

一部の実施形態では、前記相関メトリックはP値である。

【0110】

一部の実施形態では、方法はさらに次のステップを含む：前記相関メトリックを所定の閾値に対して比較して前記候補コバリエントが前記着目コバリエントであるかを決定するステップ。

40

【0111】

着目疾患の進行に関して治療を評価する例示的な方法は次のステップを含む：医用画像を取得するステップであって、該医用画像は：（a）プラセボ群にプラセボが施される前に撮像された前記被験者プラセボ群についての複数のベースラインプラセボ画像と、（b）前記プラセボ群に前記プラセボが施された後に撮像された前記被験者プラセボ群についての複数のフォローアッププラセボ画像と、（c）治療群に前記治療が施される前に撮像された前記被験者治療群についての複数のベースライン治療画像と、（d）前記治療群に前記治療が施された後に撮像された前記被験者治療群についての複数のフォローアップ治療画像とを備える、ステップと、前記医用画像を訓練済み教師なし機械学習モデル内へと

50

入力して複数の埋め込みを取得するステップであって、各埋め込みは前記医用画像の1つ以上にて反映されている前記着目疾患との関係での表現型状態に対応する、ステップと、前記複数の埋め込みを訓練済み線形回帰モデル内へと入力して複数の予測連続医療診断スコアを取得するステップであって、各予測連続医療診断スコアは前記着目疾患の状態について示す、ステップと、前記予測連続医療診断スコアに基づいて、複数のプラセボ進行スコア及び複数の治療進行スコアを決定するステップと、前記複数のプラセボ進行スコア及び前記複数の治療進行スコアを前記治療と関連付けるステップと、前記複数の疾患進行スコアと前記治療との間の相関メトリックを前記関連に基づいて決定するステップ。

【0112】

一部の実施形態では、前記医用画像を訓練済み教師なし機械学習モデル内へと入力して前記複数の埋め込みを取得するステップは：前記（a）を前記訓練済み教師なし機械学習モデル内へと入力して複数のベースラインプラセボ埋め込みを取得することと、前記（b）を前記訓練済み教師なし機械学習モデル内へと入力して複数のフォローアッププラセボ埋め込みを取得することと、前記（c）を訓練済み教師なし機械学習モデル内へと入力して複数のベースライン治療埋め込みを取得することと、前記（d）を前記訓練済み教師なし機械学習モデル内へと入力して複数のフォローアップ治療埋め込みを取得することを含む。

10

【0113】

一部の実施形態では、前記複数の埋め込みを前記訓練済み線形回帰モデル内へと入力するステップは：前記複数のベースラインプラセボ埋め込みを前記訓練済み線形モデル内へと入力して複数のベースラインプラセボスコアを取得することと、前記複数のフォローアッププラセボ埋め込みを前記訓練済み線形モデル内へと入力して複数のフォローアッププラセボスコアを取得することと、前記複数のベースライン治療埋め込みを前記訓練済み線形モデル内へと入力して複数のベースライン治療スコアを取得することと、前記複数のフォローアップ治療埋め込みを前記訓練済み線形モデル内へと入力して複数のフォローアップ治療スコアを取得することを含む。

20

【0114】

一部の実施形態では、前記複数のプラセボ進行スコア及び前記複数の治療進行スコアを決定するステップは：前記複数のベースラインプラセボスコアと前記複数のフォローアッププラセボスコアとの間の差を決定して前記複数のプラセボ進行スコアを決定することと、前記複数のベースライン治療スコアと前記複数のフォローアップ治療スコアとの間の差を決定して前記複数の治療進行スコアを決定することを含む、方法。

30

【0115】

一部の実施形態では、前記複数のプラセボ進行スコア及び前記複数の治療進行スコアを決定するステップは：前記プラセボ群内の各被験者について、前記プラセボ群内の前記被験者のベースラインプラセボスコア及びフォローアッププラセボスコアに少なくとも基づいて適合化された線形モデルの勾配を決定することと、前記治療群内の各被験者について、前記治療群内の前記被験者のベースラインプラセボスコア及びフォローアッププラセボスコアに少なくとも基づいて適合化された線形モデルの勾配を決定することを含む。

40

【0116】

一部の実施形態では、前記複数のプラセボ進行スコア及び前記複数の治療進行スコアを前記治療と関連付けるステップは、患者が前記治療を受けたかについての指示を受信し及び予測疾患進行スコアを出力するように構成されたモデルを生成するステップを含む。

【0117】

一部の実施形態では、前記相関メトリックはモデルのP値である。

【0118】

一部の実施形態では、方法はさらに次のステップを含む：前記相関メトリックを所定の閾値と比較するステップ。

【0119】

一部の実施形態では、方法はさらに次のステップを含む：前記比較に基づいて、前記治

50

療と前記着目疾患との間の関連を識別するステップ。

【0120】

一部の実施形態では、方法はさらに次のステップを含む：前記関連に基づいて前記治療を施す、調整する、又は適用するステップを含む、方法。

【0121】

一部の実施形態では、方法はさらに次のステップを含む：前記関連に基づいて医学的提案を提供するステップ。

【0122】

一部の実施形態では、前記着目疾患は非アルコール性脂肪性肝炎（NASH、non-alcoholic steatohepatitis）である。

【0123】

一部の実施形態では、前記教師なし機械学習モデルは対照モデルである。

【0124】

一部の実施形態では、前記対照モデルはSimCLRモデルである。

【0125】

一部の実施形態では、前記線形回帰モデルは線形混合モデルである。

【0126】

一部の実施形態では、前記訓練済み線形回帰モデルは複数の割り当て済み医療診断スコアに基づいて適合化される。

【0127】

一部の実施形態では、前記複数の割り当て済み医療診断スコアは1人以上の医師によって提供される。

【0128】

一部の実施形態では、前記複数の割り当て済み医療診断スコアの各割り当て済み医療診断スコアは予め定義された値のセットから選択される。

【0129】

一部の実施形態では、前記複数の予測連続医療診断スコアは、複数の予測線維症スコア、複数の予測小葉内炎症スコア、又は複数の予測脂肪症スコアとされる。

【0130】

着目患者サブグループを識別する例示的な方法は次のステップを含む：臨床被験者群から取得された複数の医用画像を訓練済み教師なし機械学習モデル内へと入力して潜在空間内の複数の埋め込みを取得するステップと、前記複数の埋め込みをクラスタリングして1つ以上の埋め込みクラスタを生成するステップと、前記1つ以上の埋め込みクラスタに対応する1つ以上の患者サブグループを識別するステップと、前記1つ以上の患者サブグループの各患者サブグループをコバリアントと関連付けて前記着目患者サブグループを識別するステップ。

【0131】

一部の実施形態では、前記教師なし機械学習モデルは対照モデルである。

【0132】

一部の実施形態では、前記対照モデルはSimCLRモデルである。

【0133】

一部の実施形態では、前記コバリアントは着目治療であり、また、前記着目患者サブグループは前記着目治療が相当なインパクトを及ぼすサブグループである。

【0134】

一部の実施形態では、前記1つ以上の患者サブグループの各患者サブグループを前記コバリアントと関連付けるステップは：患者サブグループについて、前記患者サブグループ内の患者が前記着目治療を受けたかについての指示を受信し及び予測疾患進行を出力するように構成されたモデルを、生成するステップと、前記モデルを評価して前記患者サブグループが前記着目患者サブグループであるかを決定するステップとを含む。

【0135】

10

20

30

40

50

一部の実施形態では、前記モデルを評価するステップは、前記モデルの相関メトリックを決定すること及び前記相関メトリックを所定の閾値に対して比較することを含む。

【0136】

一部の実施形態では、前記相関メトリックはP値である。

【0137】

一部の実施形態では、前記生成されたモデルは前記患者サブグループ内の被験者の疾患進行値によって訓練される。

【0138】

一部の実施形態では、前記疾患進行値は前記患者サブグループ内の前記被験者の医療診断スコアを含む。

【0139】

一部の実施形態では、前記疾患進行値は前記患者サブグループ内の前記被験者の進行スコアを含む。

【0140】

一部の実施形態では、前記疾患進行値は前記患者サブグループ内の前記被験者のDRP値を含む。

【0141】

一部の実施形態では、前記コバリエントは着目疾患の進行であり、また、前記着目患者サブグループは前記着目疾患の前記進行と有意な関連を有するサブグループである。

【0142】

一部の実施形態では、前記1つ以上の患者サブグループの各患者サブグループを前記コバリエントと関連付けるステップは：患者サブグループについて、患者が前記患者サブグループに属するかについての指示を受信し及び予測疾患進行を出力するように構成されたモデルを、生成するステップと、前記モデルを評価して前記患者サブグループが前記着目患者サブグループであるかを決定するステップとを含む。

【0143】

一部の実施形態では、前記モデルを評価するステップは、前記モデルの相関メトリックを決定すること及び前記相関メトリックを所定の閾値に対して比較することを含む。

【0144】

一部の実施形態では、前記相関メトリックはP値である。

【0145】

一部の実施形態では、前記生成されたモデルは前記臨床被験者群の疾患進行値によって訓練される。

【0146】

一部の実施形態では、前記疾患進行値は、前記患者サブグループ内の被験者の医療診断スコア、前記患者サブグループ内の被験者の進行スコア、又は前記患者サブグループ内の被験者のDRP値を含む。

【0147】

一部の実施形態では、前記コバリエントは有害副作用であり、また、前記着目患者サブグループは前記有害副作用と有意な関連を有するサブグループである。

【0148】

一部の実施形態では、前記1つ以上の患者サブグループの各患者サブグループを前記コバリエントと関連付けるステップは：患者サブグループについて、前記患者サブグループ内の患者が前記患者サブグループに属するかについての指示を受信し及び前記患者に前記有害副作用が生じるかを予測するように構成されたモデルを、生成するステップと、前記モデルを評価して前記患者サブグループが前記着目患者サブグループであるかを決定するステップとを含む。

【0149】

一部の実施形態では、前記モデルを評価するステップは、前記モデルの相関メトリックを決定すること及び前記相関メトリックを所定の閾値に対して比較することを含む。

10

20

30

40

50

【0150】

一部の実施形態では、前記相関メトリックはP値である。

【0151】

着目疾患との関係で少なくとも1つの着目生物学的標的を識別する例示的な方法は次のステップを含む：臨床被験者群から取得された複数の医用画像を訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数の埋め込みを取得するステップであって、各埋め込みは前記複数の医用画像の1つ以上にて反映されている前記着目疾患との関係での表現型状態に対応する、ステップと、前記複数の埋め込みを複数の候補生物学的標的の各候補生物学的標的と関連付けて、前記複数の候補生物学的標的についてのサブセットを識別するステップであって、前記複数の生物学的標的のサブセットは前記複数の医用画像にて反映されている表現型特性と関連付けられている、ステップと、前記複数の候補生物学的標的の前記サブセットの各候補生物学的標的を前記着目疾患と関連付けて前記着目疾患の発病又はその進行に関して機能的インパクトを及ぼす少なくとも1つの着目生物学的標的を前記サブセットから識別するステップと、調節する生物学的標的を識別するステップであって、前記調節は、前記着目疾患に関しての前記少なくとも1つの生物学的標的の機能的インパクトに関して変更、オフセット、緩和、補足や補完するように設計されている、ステップ。

10

【0152】

例示的なシステムは：1つ以上のプロセッサと、メモリと、1つ以上のプログラムとを備えるのであって、前記1つ以上のプログラムは前記メモリ内に記憶されており且つ前記1つ以上のプロセッサによって実行されるように構成されており、前記1つ以上のプログラムは上述された方法のいずれかを行うための命令を含む。

20

【0153】

例示的な非一時的コンピュータ可読記憶媒体は1つ以上のプログラムを記憶するのであって、該1つ以上のプログラムは命令を備えるのであり、該命令は電子装置の1つ以上のプロセッサによって実行されると該電子装置に上述された方法のいずれかを行わせる。

【図面の簡単な説明】

【0154】

特許又は出願のファイルは、少なくとも1つのカラーで表された図面を含む。

【0155】

カラー図面を伴うこの特許又は特許出願公報のコピーは、請求及び必要費用の支払いがなされたらば、当局から提供される。

30

【0156】

【図1】いくつかの実施形態による、例示的な発見プラットフォームのアーキテクチャについて示す図である。

【図2】いくつかの実施形態による、着目疾患との関係で遺伝的バリエーションを識別するための例示的な方法について示す図である。

【図3A】いくつかの実施形態による、着目疾患との関係で遺伝的バリエーションを識別するための例示的なワークフローについて示す図である。

【図3B】いくつかの実施形態による、着目疾患との関係で遺伝的バリエーションを識別するための例示的なワークフローについて示す図である。

40

【図4A】いくつかの実施形態による、例示的な教師なし機械学習モデルについて示す図である。

【図4B】いくつかの実施形態による、例示的な対照学習アルゴリズムの訓練のためのデータアーキテクチャについて示す図である。

【図5】いくつかの実施形態による、医療診断スコアを予測するための例示的な線形回帰モデルの適合化について示す図である。

【図6】いくつかの実施形態による、医療診断スコアを予測するための例示的なバリエーション特有モデルの適合化について示す図である。

【図7】いくつかの実施形態による、着目疾患との関係で少なくとも1つの着目遺伝的バ

50

リアントを識別するための例示的な処理について示す図である。

【図 8 A】いくつかの実施形態による、着目疾患との関係で遺伝的バリエーションを識別するための例示的なワークフローについて示す図である。

【図 8 B】いくつかの実施形態による、着目疾患との関係で遺伝的バリエーションを識別するための例示的なワークフローについて示す図である。

【図 9】いくつかの実施形態による、例示的な線形回帰モデルの適合化について示す図である。

【図 10】いくつかの実施形態による、例示的な G A N モデルについて示す図である。

【図 11】いくつかの実施形態による、予測されたシミュレーション画像の生成について示す図である。

【図 12】図 12 A は、いくつかの実施形態による、増大する線維症スコアを可視化するための予測画像タイルについての例示的なセット及び 3 つのシミュレーション画像を伴う例示的なシリーズについて示す図である。図 12 B は、いくつかの実施形態による、増大する P N P L A 3 タイルスコアを可視化するための予測画像タイルについての例示的なセット及び 3 つのシミュレーション画像を伴う例示的なシリーズについて示す図である。

【図 13】いくつかの実施形態による、様々なスコアと関連付けられている組織学的効果を可視化するための予測シミュレーション画像の例示的なシリーズについて示す図である。

【図 14】いくつかの実施形態による、医療診断スコアを予測するように構成された様々な線形モデルの性能について示す図である。

【図 15】図 15 A は、いくつかの実施形態による、研究、サイト、及び病理スコア効果を伴うバリエーションコンポーネントモデルについて示す図である。図 15 B は、いくつかの実施形態による、サイト及び研究影響について調整がなされる埋め込みについての例示的な全ゲノム的関連研究 (G W A S、genome-wide association study) について示す図である。図 15 C は、いくつかの実施形態による、例示的な全フェノーム的関連研究 (P h e W A S、phenome-wide association study) について示す図である。

【図 16】いくつかの実施形態による、着目疾患との関係で治療を評価するための例示的な方法について示す図である。

【図 17】いくつかの実施形態による、着目疾患との関係で治療を評価するための例示的な処理について示す図である。

【図 18 A】いくつかの実施形態による、例示的な進行埋め込み生成について示す図である。

【図 18 B】いくつかの実施形態による、例示的な進行埋め込み生成について示す図である。

【図 18 C】いくつかの実施形態による、例示的な進行埋め込み生成について示す図である。

【図 18 D】いくつかの実施形態による、例示的な進行埋め込み生成について示す図である。

【図 19】いくつかの実施形態による、線形モデルについての例示的な訓練処理について示す図である。

【図 20】いくつかの実施形態による、治療についての薬物反応組織学的表現型 (D R P、Drug Response histological Phenotype) との関係で着目コバリエーションを識別する例示的な方法について示す図である。

【図 21】いくつかの実施形態による、治療についての薬物反応表現型 (D R P、Drug Response Phenotype) との関係で着目コバリエーションを識別する例示的な方法について示す図である。

【図 22】いくつかの実施形態による、着目疾患の進行との関係で治療を評価するための例示的な方法について示す図である。

【図 23】いくつかの実施形態による、着目疾患の進行との関係で治療を評価するための例示的な方法について示す図である。

10

20

30

40

50

【図 2 4】いくつかの実施形態による、着目患者サブグループを識別するための例示的な方法について示す図である。

【図 2 5】いくつかの実施形態による、3つの患者クラスタについて示す図である。

【図 2 6】図 2 6 A は、いくつかの実施形態による、例示的な縦型の表現分析について示す図である。図 2 6 B は、いくつかの実施形態による、例示的な遺伝子関連研究について示す図である。

【図 2 7】いくつかの実施形態による、DRPと発現との間の関連についての研究について示す図である。

【図 2 8】いくつかの実施形態による、zスコアの比較について示す図である。

【図 2 9】いくつかの実施形態による、例示的な電子装置について示す概略図である。

10

【発明を実施するための形態】

【0157】

以下の詳細な説明は、当業者が様々な実施形態をもたらし並びにこれらを使用することを可能とするために提示されている。具体的な送致、手法、及び用途は、例示的に示されているに過ぎない。開示の例に対しての様々な改変は当業者にとっては明らかであり、本発明の精神及び範囲から逸脱することなく本明細書にて画定した一般原理を他の例や用途に適用できる。したがって、様々な実施形態は、本明細書にて説明及び示された例に限定されることは意図されておらず、むしろ特許請求の範囲と整合する範囲を与えられるべきである。

【0158】

20

本明細書にて、発見プラットフォームを提供することに向けられた方法、システム、電子機器、非一時的記憶媒体、及び装置が開示される。発見プラットフォームは多遺伝子疾患等の複雑な疾患に適用でき、標的識別や臨床試験間分析を可能としたり解釈容易性を向上させたりすることができる。発見プラットフォームは弱い類いの又は未知の遺伝的駆動因子に適用できる。例えば、NASHは未知の遺伝的アーキテクチャを伴う疾患である。一部の実施形態では、発見プラットフォームは、着目遺伝的バリエーションとNASH等の着目疾患との間の関係（例えば、因果関係）を識別できる。識別された関係は、新規被験者において着目疾患が発生する尤度について決定するために用いることができ、特定の症状又は他の疾患関連因子の存在をもとに新規被験者に疾患があるかをより高い信頼度を伴って診断できる。また、新規被験者にて遺伝的バリエーションが識別された場合、着目疾患についての診断又は予後診断を適宜提供でき、これには疾患がどのように進行することが期待され得るかについての予後診断が含まれ得る。例えば、NASHについて着目遺伝的バリエーションが発見された場合、新規被験者についてゲノム検査を行ってバリエーションを検出できる。バリエーションが存在する場合、システムは、疾患の発現を予測でき、診断を提供でき、及び／又は新規被験者においてどのように疾患が進行するかについて予後診断を提供できる。

30

【0159】

一部の実施形態では、発見プラットフォームは複数の段階を備えることができる。第1段階では、例示的なシステム（例えば、1つ以上の電子装置）は、着目疾患等の着目表現型に関する医用画像データに基づいて、埋め込みを生成する。埋め込みとは、変数からベクトル（数値アレイ）への写像である。上述のように、埋め込みが指すのは、医用画像データにて反映されている着目疾患との関係での表現型状態についてのベクトル発現である。埋め込みは、医用画像データのリッチなセマンティック情報（例えば、画像内にて反映されている組織の微細構造の特徴）を捕捉する一方で、下流側の分析に関して関連性を有さない情報（例えば、画像の向き）は排する。例示的な実装例では、着目疾患は非アルコール性脂肪性肝炎（NASH、non-alcoholic steatohepatitis）であり、また、医用画像はいくつかの臨床試験からのヘマトキシリン・エオジン（H&E、hematoxylin & eosin）染色肝生検からのものである。結果として得られる教師なしでの埋め込みは、標的識別や臨床試験間分析を可能にし、また、上述のように解釈容易性を向上させ得る。

40

50

【0160】

一部の実施形態では、システムは、医用画像データを対照学習アルゴリズム等に入力することによって埋め込みを生成する。対照学習モデルは、画像データから埋め込みを抽出できるのであり、また、上述のように、埋め込みはさもなければそのようなデータに割り当てられ得る生物学的エンドポイント又はラベル（例えば、着目疾患の進行）に関して線形的に予測できる。適した対照学習モデルは、同じサンプル画像の異なる増強結果からの埋め込みの間の類似性を最大化し、且つ、異なるサンプル画像の埋め込み間の類似性を最小化するように訓練される。例えば、モデルは、回転、反転、クロッピング、及び色ジッタリングに関して不変となる埋め込みを画像から抽出できる。

【0161】

一部の実施形態では、埋め込みに関しては、下流側の分析に用いられる前に平均アグレーション及び／又は正規化をなすことができる。一部の実施形態では、埋め込みの正規化は分散安定化変換を行うことを伴い、これによってラベルの生物学的エンドポイントを線形的に予測するその能力が向上し得る。本明細書では、正規化によって、埋め込みに基づいて適合化された線形予測モデルの性能を向上させ得る。一部の実施形態では、正規化済みの埋め込みを用いて適合化された線形モデルは、教師あり機械学習モデルと同様の又はそれより優位な予測能力を有するのであり、また、後述のように生成及び適用に関してより演算的に効率的である。

【0162】

第2段階では、システムは、例えば1つ以上の線形回帰モデルを用いて、埋め込みについての統計分析を行う。画像データではなく埋め込みを用いて統計分析をなすことによって、いくつかの技術的利点がもたらされる。第1に、埋め込みは、画像データのリッチなセマンティック情報（例えば、画像内にて反映されている組織の微細構造の特徴）を捕捉する一方で、下流側の分析に関して関連性を有さない情報（例えば、画像の向き）は排する。さらに、埋め込みは、それ自体が表す画像データよりも相当に小さなサイズのものとなっている。例示的な実装例では、埋め込みは2048次元のベクトルとすることができる一方で、対応する医用画像は膨大な画素数に対応するデータを含んでいる（例えば、数万画素、数十万画素、数百万画素等）。

【0163】

さらに、埋め込みによれば、システムは、入力として埋め込みを受信し及び予測を出力するように構成された線形回帰モデルを生成（例えば、適合化）することができる。一部の実施形態では、線形回帰モデルは線形混合モデルであり、これらは治療及び潜在的な共変量効果を含む表現型的バリエーションについての埋め込みについての統計分析のための柔軟な枠組みをもたらす。本明細書にて説明されるように、埋め込みに基づいて生成されている線形モデルは、教師あり機械学習モデル（例えば、画像データを受信するように構成されたニューラルネットワーク）と同様の又はそれより優位な予測力を提供できるのであり、また、教師あり機械学習モデルよりも訓練及び適用に関してより演算的に効率的である。

【0164】

一部の実施形態では、発見プラットフォームの第2段階は、着目疾患について示す連続スコア等の医用画像についての細かい粒度のラベルを取得するために埋め込みを用いることを伴う。例えば、線形モデルは、埋め込みと該埋め込みに関連付けられている病理学者によって割り当てられた離散的医療診断スコアとに基づいて、生成（例えば、適合化）されることができる。そして、モデルは、埋め込みに適用されて、連続医療診断スコアを予測することができる。

【0165】

予測連続スコア（predicted continuous score）は、病理学者によって割り当てられた離散スコアに対して強力な利点を有している。具体的には、予測スコア（predicted score）は、連続的な値をとるのであり、故に病理学者によって割り当てられた離散スコアよりもニュアンスが多く捕捉されることとなる。埋め込み（及び画像データ）に連続的

10

20

30

40

50

スコアを割り当てる能力の結果、より高い精密性及び下流側の分析における向上した統計能力がもたらされるのであり、例えば示される疾患状態の各々と遺伝的バリエーションとの間により緊密な関連を取得し得る。例えば、NASH及び肝線維症の重症度は病理学者によってNASH CRN及びIshakステージ序数スコアによって組織学的に評価されるのであって、例えば次のものがある：Ishak線維症スコア（0～6）、脂肪症スコア（0～3）、小葉内炎症スコア（0～3）、及びバリーニグスコア（0～2）。これらのメトリックの量的分析は、その手法の疾患分類が解像度の低さ故に困難を含む。線形モデルに関しては訓練をなして、画像データ（例えば、H&E肝生検画像データ）から病理スコアについて予測をもたらし得る連続スコアを生成できる。連続スコアは、疾患進行についてより精密な定義付けを可能とし、縦型表現分析及び遺伝関連研究について促進し得る。

10

【0166】

一部の実施形態では、システムは、候補遺伝的バリエーションと着目疾患との間での関連検定を行う。各医用画像（例えば、組織学画像）は、撮像元の人間対象からの関連遺伝子を伴う。関連検定は、候補遺伝的バリエーションと着目疾患について示す連続スコアとに基づいて線形モデルを生成することを伴う。システムは、全ての着目候補遺伝的バリエーション（例えば、100,000種、100万種、1,000万種等のバリエーション）についてバリエーション特有モデル（例えば、100,000種、100万種、1,000万種等のモデル）を生成できる。各モデルを評価して、各候補遺伝的バリエーションと着目疾患との間に有意な関連があるかについて決定して、1つ以上の着目遺伝的バリエーションを識別できる。

20

【0167】

一部の実施形態では、システムは、複数の埋め込みを複数の候補遺伝的バリエーションの各候補遺伝的バリエーションと関連付けて、埋め込みと有意な関連を有する候補遺伝的バリエーションについてのサブセットを識別する。候補遺伝的バリエーション（candidate genetic variant）と埋め込みとの間の関連について評価することによって、システムは、画像内に反映されている組織学的差異と関連付けられている候補遺伝的バリエーションについてのサブセットを識別する。一部の実施形態では、システムは、各候補バリエーション（candidate variant）についてバリエーション特有モデルを生成することによって関連検定を行うのであって、これは埋め込みを受信し及び候補遺伝的バリエーションの値を出力するように構成されている。そして、バリエーション特有モデルは評価されて、各候補遺伝的バリエーションと埋め込みとの間に有意な関連があるかについて決定される（例えば、バリエーション特有モデルと関連付けられているP値に基づいて）。システムは、全ての候補遺伝的バリエーション（例えば、100,000種、100万種、1,000万種等のバリエーション）についてバリエーション特有モデル（例えば、100,000種、100万種、1,000万種等のモデル）を生成して、候補遺伝的バリエーションについてのサブセットを識別できる。さらに、システムは、サブセット内の各候補遺伝的バリエーションを着目疾患と関連付けて、少なくとも1つの着目遺伝的バリエーションを識別できる。一部の実施形態では、システムは、サブセット内の候補遺伝的バリエーションについてバリエーション特有スコア予測モデルを生成できるのであって、これは候補遺伝的バリエーションについて示す値を受信し及び着目疾患に関する医療診断スコアを出力するように構成されている（例えば、上述の連続スコアに基づいて）。そして、モデルについては評価して、候補遺伝的バリエーションと着目疾患との間に有意な関連があるかについて決定する。

30

40

【0168】

第2段階では、他の関連検定手順を実施できる。一部の実施形態では、関連検定は次のものに基づき得る：出力が埋め込みであり入力共変量である単変量線形モデル、出力が埋め込みであり入力共変量である多変量線形モデル、又は出力が共変量であり入力共変量である線形モデル。また、関連検定手順は、線形混合モデルやロジスティック回帰等の線形モデルの拡張や非線形モデル（例えば、ランダムフォレストやSVM等）に基づいていることもできる。関連検定手順の適用によって、各埋め込み次元と検査されたあらゆる共変量との間での、又は、全体としての埋め込み次元の全てと検査されたあらゆる共変量との間での関連についてのP値をもたらし得る。多重仮説検定手順（例えば、Bonf

50

erroni型やBenjamini Hochberg型)を通じて決定される統計的に有意な関連によれば、高コンテンツ表現型データセット(例えば、医用画像データ)におけるバリエーションと関連付けられている因子をもたらし得る。

【0169】

第3段階においてシステムは、識別済み着目ジェネリック等の識別済み着目コバリアントの組織学的効果を説明するために画像に関して可視化をなすか可視化を支援できる。一部の実施形態では、システムは、線形モデルを用いて測定エンドポイントについて予測し得る生検画像タイルを識別する。線形補間によって複数の埋め込みを生成して、着目疾患の進行を表すことができる。埋め込みを、疾患の表現型状態を表す一連の画像に変換することができる。一連の画像はアニメーションとして表示でき、関連する組織学的な変化について可視化表示を提供できるのであり、これによらずにしては病理スコアについての関連研究によって検出ができない場合がある。可視化によって、モデル及び組織学的な変化に基づいての仮説の生成に用いられる特徴の解釈が支援され得る。したがって、システムは、第2段階では疾患ラベルと関連付けられていないバリエーションを発見でき、また、第3段階では新規の可視化ツールによってそれらの効果の特徴付けることができる。

【0170】

一部の実施形態では、シミュレーション画像の形式での予測画像は、訓練済み敵対的生成ネットワーク(GAN)モデルの生成器コンポーネントによって生成される。生成器は、埋め込み(例えば、画像タイル埋め込み)に条件付けられたシミュレーション画像を生成でき、他の特徴を一定に保ちつつ表現型に沿った補間を可能とできる。例えば、生成器は、(条件としての)埋め込み x 及び標準的な正規分布からサンプリングされたノイズベクトル u を受信し、並びに、シミュレーション画像を出力するように構成されていることができる。1つの例示的な実装例では、埋め込み x は2048次元の埋め込みであり、ノイズベクトル u は標準的な正規分布からサンプリングされた512次元のベクトルである。

【0171】

一部の実施形態では、予測画像は、各画像の予測スコア及び／又はそのような画像の埋め込みとの関係で生成された予測特徴に基づいてランキングされた実際の医用画像から選択されることができる。可視化される画像は、ランキングされた画像の一部又は全部とされ得る。例えば、ランキングの上位 N 件の画像を表示できる。別の例を挙げるに、上位 N 件及び下位 M 件の画像を表示できる。ランキングに基づいて代替的な画像サブセットを選択することもできる。

【0172】

本明細書にて説明される実施形態は、単に例示的なものにすぎず、また、発見プラットフォームは任意の着目表現型と共変量との間の関連を発見するために活用できる。本明細書にて説明される一部の例について述べるに：表現型データには医用画像が含まれ；着目表現型は着目疾患(例えば、NASH)であり、これは医療診断スコア(例えば、線維症スコア)によって表されることができ；着目共変量は着目遺伝的バリエーションである。もっとも、本明細書にて説明される手法は別の着目表現型と別の共変量との間の関連を発見することに関して活用することもできるということに留意されたい。例示的な表現型データとしては次のものが含まれるがこれらには限定されないものであり、また、他の列挙が限定的であるとも示唆はされない：生体医用画像(例えば、MRI、X線、CTスキャン)、病理組織学データ(例えば、H&E染色、トリクローム染色)等の生検試料から生成された医用画像、臨床バイオマーカーデータ(例えば、プロテオミクスやcfDNAを含む血液検査測定値、認知／精神評価スコア、マイクロバイオーーム評価等)、及びゲノムバイオマーカーデータ(例えば、バルクRNA-seq、メチル化データ、ゲノム配列データ、エピジェネティック配列データ等)。例示的な表現型としては次のものが含まれるがこれらには限定されないものであり、また、他の列挙が限定的であるとも示唆はされない：着目疾患、遺伝子発現、メタボロミクス、プロテオミクス、トランスクリプトミクス、又はリポミクス等。例示的な共変量としては次のものが含まれるがこれらには限定されないものであ

り、また、他の列挙が限定的であるとも示唆はされない：人口統計情報（例えば、年齢や性別）、臨床共変量（例えば、疾患状態、臨床スコア又は血液バイオマーカー）、ゲノムデータ（例えば、遺伝子データ、発現データ、メチル化データ、等）等。

【0173】

一部の実施形態では、識別済みの関係は、バイオマーカーや疾患介入用の標的の識別に関して用いられ得る。例えば、発病又は進行との関係で因果関係にあると識別された特定の遺伝的バリエーションに関しては、治療的介入に関してそれらが適するかをさらに評価できる。また、遺伝的バリエーションが着目疾患に対して及ぼす機能的インパクトを考慮して、治療的介入のための追加的な生物学的標的を識別することができる。そのような生物学的標的は、例えば、着目遺伝的バリエーションが所在する遺伝子によって転写されるたんぱく質の場合がある。そのような生物学的標的は、着目疾患の少なくとも1つの遺伝的バリエーションの機能的インパクトに関して変更、オフセット、緩和、補足や補完するものと見込まれている他の遺伝子、たんぱく質、又は代謝産物を含み得る。一部の実施形態では、識別済みの関係は、治療を開発するために用いられ得る。例えば、関連付けられている遺伝的バリエーションを有する被験者群に以前施された候補治療のインパクトを、候補治療に関しての変更型又は類似型の開発に際して、考慮できる。以前施された候補治療の効果に関しての縦型の及びクロスセクショナル型のデータの双方は、着目疾患の予測状態又は進行と比較でき、量的観点から、着目疾患の状態又は進行に関して着目遺伝的バリエーションがどの程度まで以前施された候補治療のインパクトに影響を与えるかを決定できる。そのような候補治療に関しての変更型又は類似型については、疾患及びその被験者群との関係での遺伝的バリエーションのインパクトを考慮して、強化された治療的效果又は減じられた有害副作用を有するように選択できる。また、遺伝的バリエーションを反映している疾患モデルは、治療的候補をスクリーニングするために用いられ得る。例えば、ゲノム特徴との関連付けの場合、統計的に有意な関連付けは新規な候補薬剤標的又は人間の疾患に関与する主要経路の発見につながり得る。同様に、関連付けられている遺伝的バリエーションを有する被験者群に以前施された候補治療（candidate therapies）のインパクト（疾患状態に対しての有害なインパクト又は副作用を含む）は、組合せ型治療の開発に用いられ得る。したがって、本明細書にて説明される手法は、異なる遺伝的背景を有する被験者間で治療的候補（therapeutic candidates）に対しての起こるであろう応答を予測するために、特定の治療を受けるべき適切な患者コホートを識別するために、及び結果を最適化するために臨床試験を一般的に設計するために用いることができる。

【0174】

発見プラットフォームは、治療を選択的に施し、調整し、又は適用するために用いられ得る。一部の実施形態では、識別済みの関係は医学的提案を提供するために用いられ得る。医学的提案は、患者に対しての治療及び／若しくは療法の提案、並びに／又は、支援をもらうために医療専門家に連絡せよとの指示を含み得る。一部の実施形態では、識別済みの関係に基づいてレポートが生成され得る。人口統計及び臨床的な特徴と関連付けられている場合、有意な関連は、新たな関連付けの発見につながったり（例えば、性別、年齢、コレステロール水準等との関連）、及び／又は、データセット内のテクニカルバイアス（例えば、特定の臨床センタとの関連）の探知につながり得る。

【0175】

一部の実施形態では、識別済みの関係（relationship）は、着目疾患の治療のための生物学的標的（例えば、薬剤標的）を識別するために用いられ得る。一旦、遺伝的バリエーションが特定の疾患と関連付けられている（associated）ものとして識別されたらば、遺伝的バリエーションは該疾患の治療のための標的（例えば、薬剤標的）として用いられ得る。一部の実施形態では、疾患と相関付けられている（correlated）遺伝的バリエーションについては研究して、遺伝子機能や疾患病理（例えば、遺伝子型／表現型相関）についてさらなる理解を得る、及び／又は患者における疾患の治療のための遺伝的バリエーションの治療的ターゲティングの有望性を査定する。例えば、遺伝的バリエーションは機能喪失バリエーションたり得るのであり、これは遺伝的バリエーションを構成する遺伝子の発現の欠如故に疾患病理を

付与する。一部の実施形態では、遺伝的バリエーション遺伝子型との相関する疾患表現型に対しての様々な薬剤の効果を評価するために薬剤スクリーニングがなされる。一部の実施形態では、疾患治療用薬剤は、該薬剤による治療後の疾患表現型の緩和に基づいて選択される。一部の実施形態では、遺伝的バリエーションはNASH疾患と相関付けられている。一部の実施形態では、NASH疾患と相関付けられている遺伝的バリエーションはNASH疾患治療用の薬剤標的である。生物学的標的は関連付けられている遺伝的バリエーションそれ自体とされ得るも、標的は次のものをも含み得る：（a）バリエーションを含む遺伝子によって転写されるたんぱく質又は代謝産物、（b）バリエーションによって惹起された不足をオフセット／補うことができるたんぱく質又は代謝産物、（c）バリエーション及びその機能的インパクトの逆作動薬である遺伝子、たんぱく質又は代謝産物等。

10

【0176】

以下の説明では、例示的方法、パラメータ、及びこれらに類することについて記載されている。もっとも、そのような説明は、本開示の範囲の限定として意図されるものではなく、代わりに、例示的实施形態の説明として提供されることを認識されたい。

【0177】

以下の説明は、用語「第1」、「第2」等を使用して、種々の要素を説明するが、これらの要素は、用語によって限定されるべきではない。これらの用語は、1つの要素を別の要素から区別するためにのみ用いられているにすぎない。例えば、第1のグラフィカル表現は、種々の説明される実施形態の範囲から逸脱することなく、第2のグラフィカル表現と称され得るのであり、同様に、第2のグラフィカル表現は、第1のグラフィカル表現と称され得る。第1のグラフィカル表現及び第2のグラフィカル表現は両方ともグラフィカル表現であるも、それらは同一グラフィカル表現ではない。

20

【0178】

本明細書の種々の説明される実施形態の説明において使用される用語は、特定の実施形態を説明する目的のためのものにすぎず、限定することは意図されていない。種々の説明される実施形態の説明および添付の請求項において使用されるように、単数形「a」、「an」、および「the」は、文脈によってそうでないことが明確に示されない限り、複数形態も同様に含むように意図される。また、用語「及び／又は」は、本明細書に使用されるように、関連付けられる列挙されたアイテムのうちの1つ以上のあらゆる可能性として考えられる組み合わせを指し、包含することを理解されたい。さらに、「includes（～を含む）」、「including（～を含む）」、「comprises（～を備える）」及び／又は「comprising（～を備える）」との語は、本明細書において使用されるとき、述べられた特徴、整数、ステップ、動作、要素、及び／又は構成要素の存在を規定するが、1つ以上の他の特徴、整数、ステップ、動作、要素、構成要素、及び／又はその群の存在または追加を除外しないことを理解されたい。

30

【0179】

「if（～の場合）」との語は、随意的には、文脈に応じて、「when（～のとき）」または「upon（～に応じて）」若しくは「in response to determining（～を決定することに応じて）」又は「in response to detecting（～を検出することに応じて）」を意味するように解釈される。同様に、「if it is determined（～であることが決定される場合）」又は「if [a stated condition or event] is detected（[述べられた条件または事象]が検出される場合）」は、随意的には、文脈に応じて、「upon determining（～を決定することに応じて）」又は「in response to determining（～を決定することに応じて）」若しくは「upon detecting [the stated condition or event]（[述べられた条件または事象]を検出することに応じて）」又は「in response to detecting [the stated condition or event]（[述べられた条件または事象]を検出することに応じて）」を意味するように解釈される。

40

50

【0180】

図1はいくつかの実施形態による、例示的な発見プラットフォームのアーキテクチャについて示す。ステージ1では、例示的なシステム（例えば、1つ以上の電子装置）は、着目疾患等の着目表現型に関する医用画像データに基づいて、埋め込みを生成する。埋め込みとは、医用画像データにて反映されている着目疾患との関係での表現型状態についてのベクトル表現である。埋め込みは、医用画像データのリッチなセマンティック情報（例えば、画像内にて反映されている組織の微細構造の特徴）を捕捉する一方で、下流側の分析に関して関連性を有さない情報（例えば、画像の向き）は排する。図1に示されるように、4つの医用画像102a, 102b, 102c, 102dは4つの埋め込みに変換できるのであり、これらは埋め込み空間106においてそれぞれ4つのポイント104a, 104b, 104c, 104dとして各々表されるのであり、該空間は本明細書では交換可能な態様で潜在空間106とも称される。示されている例では、着目疾患は非アルコール性脂肪性肝炎（NASH、non-alcoholic steatohepatitis）であり、また、医用画像はいくつかの臨床試験からのH&E染色肝生検からのものである。結果として得られる教師なしでの埋め込みは、標的識別や臨床試験間分析を可能にし、また、上述のように解釈容易性を向上させ得る。

10

【0181】

一部の実施形態では、システムは、医用画像データを対照学習アルゴリズム等に入力することによって埋め込み（例えば、ポイント104a-104dによって表される埋め込み）を生成する。対照学習モデルは、画像データから埋め込みを抽出できるのであって、上述のように埋め込みはさもなければそのようなデータに割り当てられ得る生物学的エンドポイント又はラベル（例えば、着目疾患の進行）に関して線形的に予測できる。適した対照学習モデルは、同じサンプル画像の異なる増強結果からの埋め込みの間の類似性を最大化し、且つ、異なるサンプル画像の埋め込み間の類似性を最小化するように訓練される。例えば、モデルは、回転、反転、クロッピング、色ジッタリング、又は他の画像オーグメンテーション、又はこれらの組み合わせに関して不変となる埋め込みを画像から抽出できる。

20

【0182】

一部の実施形態では、埋め込みに関しては、下流側の分析に用いられる前に平均アグレーション及び／又は正規化をなすことができる（例えば、ステージ2）。一部の実施形態では、埋め込みの正規化は分散安定化変換を行うことを伴うのであり、単純な回帰ベースド分析を下流側にて適用することが可能となる。本明細書では、正規化によって、埋め込みに基づいて適合化された線形予測モデルの性能を向上させ得る。一部の実施形態では、正規化済みの埋め込みを用いて適合化された線形モデルは、教師あり機械学習モデルと同様の又はそれより優位な予測能力を有するのであり、また、後述のように生成及び適用に関してより演算的に効率的である。

30

【0183】

一部の実施形態では、予測画像は着目共変量決定のために識別され得る。一部の実施形態では、予測特徴は組織病理データにてタイルレベルで生成され得る。これは様々な手法を用いて達成され得る。例えば、平均アグレーションタイル埋め込みを取得して生検埋め込みを生成できる。生検埋め込みから共変量を予測できるように、このデータを用いて線形モデルを適合化させることができる。また、タイル埋め込みに線形モデルを適用して、タイルレベルスコアを生成できる。別の例を挙げるに、機械学習モデルの複数インスタンスをタイル埋め込みに直接的に適合化させることができる。例えば、各タイルについてスコア及び重みの両方を予測するモデルに関して適合化をなし得るのであり、そしてスコアに関して加重平均がとられる。スコア及び重みの両方は、2次元の予測特徴と看做され得る。

40

【0184】

一部の実施形態では、予測画像はそれらの画像と関連付けられている予測特徴に基づいて識別され得る。例えば、最高の又は最低の予測スコアを有するタイルを解析して予測画像を識別できる。別の例を挙げるに、着目条件を定義できる（例えば、高い線維症スコア

50

対低い線維症スコア、第1の遺伝シーケンス対第2の遺伝シーケンス、及びこれらに類する事項）。第1の着目条件との関係での所与のタイルの特徴の確率（例えば、 P （タイル特徴 | 条件1））及び第2の着目条件との関係での所与のタイルの特徴の確率（例えば、 P （タイル特徴 | 条件2））についてのモデルに関して適合化をなし得る。そして、タイル画像を可視化できる（例えば、ユーザインタフェース上でのレンダリング）。一部の実施形態では、 P （タイル特徴 | 条件1）／ P （タイル特徴 | 条件2）は、所与の特徴を伴うタイルの観測確率が条件1又は条件2（又は定義されている場合には他の条件）下でよりあり得ることを表し得る。該比は、十分に大きい又は小さい場合には条件1又は条件2の尤度がより大きいことを示すことになる。一部の実施形態では、低い予測確率を有するタイルについてはフィルタリングして外れ値を無視することができる。

10

【0185】

ステージ2では、システムは（例えば、ステージ1で用いられる1つ以上の電子装置と同一又は類似の1つ以上の電子装置）、例えば1つ以上の線形回帰モデルを用いて、埋め込みについての統計分析を行うことができる。画像データではなく埋め込みを用いて統計分析をなすことによって、いくつかの技術的利点をもたらされる。第1に、埋め込みは、画像データのリッチなセマンティック情報（例えば、画像内にて反映されている組織の微細構造の特徴）を捕捉する一方で、下流側の分析に関して関連性を有さない情報（例えば、画像の向き）は排する。さらに、埋め込みは、それ自体が表す画像データよりも相当に小さなサイズのものとなっている。例示的な実装例では、埋め込みは2048次元のベクトルとすることができる一方で、対応する医用画像は膨大な画素数に対応するデータを含んでいる（例えば、数万画素、数十万画素、数百万画素等）。よって、埋め込みを記憶することによって、医用画像データを用いる場合に比して、メモリを節約しつつ分析の実行に要する処理時間をも減じることができる。

20

【0186】

さらに、埋め込みによれば、システムは、入力として埋め込みを受信し及び有用な様々な予測を出力するように構成された線形回帰モデルを生成（例えば、適合化）することができる。一部の実施形態では、線形回帰モデルは線形混合モデルであり、これらは治療及び潜在的な共変量効果を含む表現型的バリエーションについての埋め込みについての統計分析のための柔軟な枠組みをもたらす。本明細書にて説明されるように、埋め込みに基づいて生成されている線形モデルは、教師あり機械学習モデル（例えば、画像データを受信するように構成されたニューラルネットワーク）と同様の又はそれより優位な予測力を提供できるのであり、また、教師あり機械学習モデルよりも訓練及び適用に関してより演算的に効率的である。

30

【0187】

一部の実施形態では、ステージ2は、着目疾患について示す連続スコア等の医用画像102a-102dについての細かい粒度のラベルを取得するために埋め込みを用いることを伴う。例えば、線形モデルは、埋め込みと該埋め込みに関連付けられている病理学者によって割り当てられた離散的医療診断スコアとに基づいて、生成（例えば、適合化）されることができる。そして、モデルは、埋め込みに適用されて、連続医療診断スコアを予測することができる。

40

【0188】

多くの場合、予測連続スコア（predicted continuous score）は、病理学者によって割り当てられた離散スコアに対して強力な利点を有している。具体的には、予測スコア（predicted score）は、連続的な値をとるのであり、故に病理学者によって割り当てられた離散スコアよりもニュアンスが多く捕捉されることとなる。埋め込み（及び画像データ）に連続的スコアを割り当てる能力の結果、より高い精密性及び下流側の分析における向上した統計能力がもたらされるのであり、例えば示される疾患状態の各々と遺伝的バリエーションとの間により緊密な関連を取得し得る。例えば、NASH及び肝線維症の重症度は病理学者によってNASH CRN及びIshakステージ序数スコアによって組織学的に評価されるのであって、例えば次のものがある：Ishak線維症スコア（整数たる0～6）、脂

50

肪症スコア（整数たる0～3）、小葉内炎症スコア（整数たる0～3）、及びバルーニングスコア（整数たる0～2）。これらのメトリックの量的分析は、その手法の疾患分類が解像度の低さ故に困難を含む。他方で、線形モデルに関しては訓練をなして、画像データ（例えば、H & E 肝生検画像データ）から病理スコアについて予測をもたらし得る連続スコアを生成できる。連続スコアは、疾患進行についてより精密な定義付けを可能とし、縦型表現分析及び遺伝関連研究について促進し得る。

【0189】

一部の実施形態では、ステージ2はブロック110～160を含む。各ブロックは1つ以上のコンピューティングシステムによってなされる機能についての機能的表現である。一部の実施形態では、同じ1つ以上のコンピューティングシステムが2つ以上のブロックのオペレーションを実効化することができる。例えば、ステージ2はブロック110を含む。ブロック110では、システム（例えば、1つ以上のコンピューティングシステム）は、候補遺伝的バリエーションと着目疾患との間での関連検定を行うように構成されていることができる。関連検定は、候補遺伝的バリエーションと着目疾患について示す連続スコアとに基づいて線形モデルを生成することを伴う。システムは、全ての着目候補遺伝的バリエーション（例えば、100,000種、100万種、1,000万種のバリエーション）についてバリエーション特有モデル（例えば、100,000種、100万種、1,000万種のモデル）を生成できる。各モデルについては評価して、各候補遺伝的バリエーションと着目疾患との間に有意な関連があるかについて決定できる。ブロック110の詳細は図2を参照して説明される。

【0190】

一部の実施形態では、ステージ2はブロック120を含む。ブロック120では、システムは、複数の埋め込みを複数の候補遺伝的バリエーションの各候補遺伝的バリエーションと関連付けて、埋め込みと有意な関連を有する候補遺伝的バリエーションについてのサブセットを識別するように構成されていることができる。候補遺伝的バリエーション（candidate genetic variant）と埋め込みとの間の関連について評価することによって、システムは、（あるならば）画像内にて反映されている組織学的差異と関連付けられている候補遺伝的バリエーションについてのサブセットを識別できる。本明細書にて説明される手法は、分析を特定の診断スコアに絞ることによっては発見できないであろう組織学に影響を及ぼすバリエーションを識別することができる。

【0191】

一部の実施形態では、システムは、各候補バリエーション（candidate variant）についてバリエーション特有モデルを生成することによって関連検定を行うのであって、これは埋め込みを受信し及び候補遺伝的バリエーションの値を出力するように構成されている。そして、バリエーション特有モデルは評価されて、各候補遺伝的バリエーションと埋め込みとの間に有意な関連があるかについて決定される（例えば、バリエーション特有モデルと関連付けられているP値に基づいて）。システムは、全ての着目遺伝的バリエーション（例えば、100,000種、100万種、1,000万種のバリエーション）についてバリエーション特有モデル（例えば、100,000種、100万種、1,000万種のモデル）を生成できる。

【0192】

ブロック120では、さらに、システムは、サブセット内の各候補遺伝的バリエーションを着目疾患と関連付けて、少なくとも1つの着目遺伝的バリエーションをサブセットから識別できる。一部の実施形態では、システムは、サブセット内の候補遺伝的バリエーションについてバリエーション特有スコア予測モデルを生成できるのであって、これは候補遺伝的バリエーションについて示す値を受信し及び着目疾患に関する医療診断スコアを出力するように構成されている（例えば、上述の連続スコアに基づいて）。そして、モデルは評価されて、候補遺伝的バリエーションと着目疾患との間に有意な関連があるかについて決定する。ブロック120の詳細は図7を参照して説明される。

【0193】

ブロック130では、システムはさらに、着目疾患の進行との関係で治療を評価できる。疾患の進行は、本明細書にて説明される進行埋め込みを用いて定量化することができる

。システムは、入力進行埋め込みを受信しプラセボ又は治療を示す分類結果を出力するモデルからの予測として、薬剤応答表現型（DRP）として帰属させるように構成されていることができる。システムは、DRPと治療との間に有意な関連があるかについて決定できる。有意な関連がある場合、下流側の分析において治療をさらに分析できる（例えば、ブロック140）。ブロック130の詳細は図16を参照して説明される。

【0194】

ブロック140では、システムはさらに、治療と関連付けられているDRPとの関係で着目コバリエントを識別するように構成されていることができる。DRPについての帰属決定は、進行埋め込みが利用可能である限り臨床試験データセットを用いることによって行うことができる。DRPと分子データ（例えば、表現及びジェネティクス）との間の有意な関連は関連検定を介して取り出すことができる。発現ありの関連付けでは、プラセボ対薬剤の差次的発現分析にて検出できなかった遺伝子を識別できる。一部の場合では、DRP分析では、真のプラセボ対薬剤（placebo-vs-drug）の差次的発現分析の症例対照として相関付けられている遺伝子セットを識別できる。一部の場合では、DRP分析では、より大きなコホートの分析故により大きな遺伝子セットを識別できるのでありDRP相関の解釈を支援し得る。

【0195】

一部の実施形態では、小規模臨床試験における治療対プラセボ（treatment vs placebo）の分析のzスコアに対してのより大規模な臨床試験における帰属DRP分析のそのの比較を用いて、相関を識別することができる（例えば、図28との関係で見られること）。ブロック140の詳細は図20を参照して説明される。

【0196】

ブロック150では、システムはさらに、着目疾患の進行との関係で治療を評価するように構成されていることができる。疾患進行は連続医療診断スコアによって定量化できる。高解像度NASHスコア及び様々な治療の間の有意な関連は、関連検定を介して取り出される。この処理では、病理学者によって割り当てられた離散スコアを用いては検出できなかった、医療診断スコアに対しての薬剤の影響を取り出すことができる。連続スコアは、疾患進行についてより精密な定義付けを可能とし、縦型表現分析（例えば、図26A）及び遺伝関連研究（例えば、図26B）について促進し得る。ブロック150の詳細は図22を参照して説明される。

【0197】

ブロック160では、システムはさらに、着目患者サブグループを識別するように構成されていることができる。システムは、患者画像データから埋め込みを取得し及び埋め込みのクラスタを識別して患者サブグループを識別できる。各患者クラスタアイデンティティ及び疾患バイオマーカー、遺伝的バリエーション、及び発現レベル間の有意な関連が関連検定にて取り出される。この手順では、患者セグメント及び関連付けられている臨床ラベル及び分子ドライバが取り出される。ブロック160の詳細は図24を参照して説明される。

【0198】

ステージ2では、他の関連検定手順を実施できる。一部の実施形態では、関連検定は次のものに基き得る：出力が埋め込みであり入力共変量である単変量線形モデル、出力が埋め込みであり入力共変量である多変量線形モデルや、出力が共変量であり入力埋め込みである線形モデル。また、関連検定手順は、線形混合モデルやロジスティック回帰等の線形モデルの拡張や非線形モデル（ランダムフォレストやSVM等）に基づいていることもできる。関連検定手順の適用によって、各埋め込み次元と検査されたあらゆる共変量との間での、又は、全体としての埋め込み次元の全てと検査されたあらゆる共変量との間での関連についてのP値をもたらし得る。多重仮説検定手順（例えば、Bonferroni型やBenjamini Hochberg型）を通じて決定される統計的に有意な関連によれば、高コンテンツ表現型データセット（例えば、医用画像データ）におけるバリエーションと関連付けられている因子をもたらし得る。

【0199】

随意的なステージ3では、システムは、シミュレーション画像172等のシミュレーション画像を生成して、識別済み着目ジェネリック等の識別済み着目コバリアントの組織学的効果を可視化するように構成されていることができる。図1に示されているように、潜在空間内の埋め込みを生成してそしてこれを画像に変換して着目疾患との関係での表現型状態を可視化できる。一部の実施形態では、システムは、線形モデルを用いて測定エンドポイントについて予測し得る生検画像タイルを識別する。線形補間によって複数の埋め込みを生成して、着目疾患の進行を表すことができる。埋め込みは、一連の画像に変換することができる。一連の画像はアニメーションとして表示でき、関連する組織学的な変化について可視化表示を提供できるのであり、これによらずにしては病理スコアについての関連研究によって検出ができない場合がある。例示的な実装例について述べるに、手順では、ステージ2における発見された着目遺伝的バリエーション、生検埋め込み（分析済み生検が行とされ、また、埋め込み次元が列とされる行列）、及びタイル埋め込み（対応するタイルが行とされ、また、埋め込み次元が列とされる行列）が入力として受け付けられて、また、一連の $n=256 \times 256$ タイル画像が出力される。可視化によって、モデル及び組織学的な変化に基づいての仮説の生成に用いられる特徴の解釈が支援され得る。したがって、システムは、ステージ2では疾患ラベルと関連付けられていないバリエーションを発見でき、また、ステージ3では新規の可視化ツールによってそれらの効果の特徴付けることができる。

10

【0200】

一部の実施形態では、シミュレーション画像は、訓練済み敵対的生成ネットワーク（GAN）モデルの生成器コンポーネントによって生成される。生成器は、埋め込み（例えば、画像タイル埋め込み）に条件付けられた画像を生成でき、他の特徴を一定に保ちつつ表現型に沿った補間を可能とできる。例えば、生成器は、（条件としての）埋め込み x 及び標準的な正規分布からサンプリングされたノイズベクトル u を受信し、並びに、シミュレーション画像を出力するように構成されていることができる。1つの例示的な実装例では、埋め込み x は2048次元の埋め込みであり、ノイズベクトル u は標準的な正規分布からサンプリングされた512次元のベクトルである。

20

【0201】

一部の実施形態では、シミュレーション画像はランキングされる。シミュレーション画像のランキングは、シミュレーション画像の提示及び／又はさらなる分析のための記憶に関しての根拠として用いられ得る。一部の実施形態では、シミュレーション画像は医療診断スコアに基づいてランキングされる。

30

【0202】

本明細書にて説明される実施形態は、単に例示的なものにすぎず、また、発見プラットフォームは任意の着目表現型と共変量との間の関連を発見するために活用できる。本明細書にて説明される一部の例について述べるに：表現型データには医用画像が含まれ；着目表現型は着目疾患（例えば、NASH）であり、これは医療診断スコア（例えば、線維症スコア）によって表されることができ；着目共変量は着目遺伝的バリエーションである。もっとも、本明細書にて説明される手法は別の着目表現型と別の共変量との間の関連を発見することに関して活用することもできるということに留意されたい。例示的な表現型データとしては次のものが含まれる：医用画像（例えば、MRI、X線、CTスキャン）、病理組織学データ（例えば、H&E染色、トリクローム染色）、臨床バイオマーカーデータ（例えば、プロテオミクスやcfDNAを含む血液検査測定値、認知／精神評価スコア、マイクロバイオーム評価等）、及びゲノムバイオマーカーデータ（例えば、バルクRNA-seq、メチル化データ、ゲノム配列データ、エピジェネティック配列データ等）。例示的な表現型には次のものが含まれる：着目疾患、遺伝子発現、メタボロミクス、プロテオミクス、トランスクリプトミクス、又はリポミクス等。例示的な共変量クラスには次のものが含まれる：人口統計情報（例えば、年齢や性別）、臨床共変量（例えば、疾患状態、臨床スコア又は血液バイオマーカー）、ゲノムデータ（例えば、遺伝子データ、発現データ、メチル化データ、等）等。

40

50

【0203】

図2は、いくつかの例による、着目疾患との関係で着目遺伝的バリエーションを識別するための例示的な処理200について示す。処理200は、例えば、ソフトウェアプラットフォームを実装する1つ以上の電子装置を使用して、実施される。いくつかの実施例では、処理200は、クライアント-サーバシステムを使用して実施され、処理200のステップは、任意の様式において、サーバと1つ以上のクライアントデバイスとの間で分割される。したがって、処理200の部分がクライアントサーバシステムの特定の装置によって実行されると説明されているも、処理200はそのように限定されることは要さないことに留意されたい。他の例では、処理200はクライアント装置のみ又は複数のクライアント装置のみを用いて実行される。処理200では、いくつかのブロックは、随意に、組み合わせられ、いくつかのブロックの順序は、随意に、変更され、いくつかのブロックは、随意に、省略される。いくつかの実施例では、付加的ステップが、処理200と組み合わせて実施されてもよい。故に、図示（および下記により詳細に説明）されるような動作は、本質的に、例示的であって、したがって、限定として見なされるべきではない。

10

【0204】

図2を参照するに、例示的なシステム（例えば、1つ以上の電子装置）は、臨床被験者群から取得された複数の医用画像を取得できる。医用画像は、着目疾患の状態を表す。一部の実施形態では、複数の医用画像は、臨床被験者群からの生検試料の複数の生検画像を含む。例えば、生検では1つ以上の組織スライドを被験者から取得でき、また、1つ以上のデジタル画像を撮って各組織スライドを撮像できる。

20

【0205】

図3Aに示される例示的なワークフローでは、生検1-nがなされている。生検1-nは複数の被験者（例えば、がん患者）及び／又は複数回の訪問（例えば、スクリーニング訪問回やフォローアップ訪問回）に対応し得る。示されている例では、着目疾患は非アルコール性脂肪性肝炎（NASH、non-alcoholic steatohepatitis）であり、また、生検はいくつかの臨床試験からのH&E染色肝生検である（もっとも、類似のワークフローは他の着目疾患について実装され得る）。各生検（例えば、生検1）は、1つ以上の生検画像（例えば、生検1についての医用画像302）を結果としてもたらす。したがって、生検1-nは、医用画像302、医用画像352等を含む複数の医用画像を結果としてもたらす。

30

【0206】

医用画像は、様々なデータと関連付けられることができる。例えば、図3Aを参照するに、関連データ304は生検1と関連付けられており、データ354は生検nと関連付けられており以下同様。後述のように、データは着目疾患及び被験者についての既知情報を含み得るのであり、着目遺伝的バリエーションについての情報も含まれる。

【0207】

一部の実施形態では、医用画像と関連付けられているデータは、着目疾患に関する割り当て済み医療診断スコアを含み得る。割り当て済み医療診断スコアは、疾患状態を示し得る。一部の実施形態では、医療診断スコアは、一人以上の病理学者によって割り当てられた生検レベルスコアであることができ、これはそれらの病理学者の生検スライドの検討に基づいている。例えば、NASH疾患に関しては、NASHの重症度、及び肝線維症については、病理学者によってNASH CRN及びIshakステージ序数スコアによって組織学的に評価できる。一部の実施形態では、割り当て済み医療診断スコアはIshak線維症スコア等の生検レベル線維症スコアとされ得るのであり、これは離散値たる0, 1, 2, 3, 4, 5, 又は6とされる線維症の程度を示すものである。一部の実施形態では、割り当て済み医療診断スコアは生検レベル脂肪症スコアとされ得るのであり、これは離散値たる0, 1, 2, 及び3とされる脂肪症の程度を示すものである。一部の実施形態では、割り当て済み医療診断スコアは生検レベル小葉内炎症スコアとされ得るのであり、これは離散値たる0, 1, 及び2とされる小葉内炎症の程度を示すものである。一部の実施形態では、割り当て済み医療診断スコアは生検レベルルーニングスコアとされ得るのであり、これは離散値たる

40

50

0, 1, 及び2とされるバルーニングの程度を示すものである。これらのメトリックの量的分析は、その手法の疾患分類が解像度の低さ故に困難を含む。後述のように、線形モデルに関しては訓練をなして、画像データ（例えば、H & E 肝生検画像データ）から病理スコアについて予測をもたらし得る連続スコアを生成できる。連続スコアは、疾患進行についてより精密な定義付けを可能とし、縦型表現分析及び遺伝関連研究について促進し得る。

【0208】

一部の実施形態では、医用画像と関連付けられているデータは、生検サンプルの取得元の被験者の遺伝子データを含み得る。例えば、データ（例えば、関連データ304）は、複数の遺伝的バリエーションについての被験者遺伝子情報を含み得る（例えば、100,000種、100万種、1,000万種のバリエーション）。例えば、データは、被験者が複数の遺伝的バリエーションの各々を有するかを示し得る。例えば、母集団において2つのアレルを伴う遺伝的バリエーションについては、個体が最低頻度アレルのコピーを0, 1, 又は2個有しているかに応じて、医用画像は遺伝的バリエーション値（0, 1, 又は2）と関連付けられることができる。一部の実施形態では、被験者遺伝子データは、被験者が着目疾患に罹患する尤度を示す多遺伝子リスクスコアとされ得る。

【0209】

一部の実施形態では、医用画像と関連付けられているデータは（例えば、関連データ304、データ354等）、生検サンプルの取得元の被験者の人口統計データを含み得る。人口統計データは例えば、性別、年齢、及び／又は治療群（例えば、プラセボ、治療x、治療y）を含み得る。

【0210】

図3Aを参照するに、医用画像は複数の画像タイルに分割できる。例えば、生検1の医用画像302を分割して画像タイル306-1, 306-2, ..., 306-M1を取得でき；生検nの医用画像532を分割して画像タイル356-1, 356-2, ..., 356-Mnを取得できる。一部の実施形態では、画像タイルは、所定のグリッドを用いて医用画像から抽出されて、また、統一されたサイズの画像タイルとして記憶されることができる。1つの例示的実装例では、画像タイルはタイル寸法が $192\ \mu\text{m} \times 192\ \mu\text{m}$ とされる所定のグリッドを用いて抽出され、また、画像タイルは $224\text{画素} \times 224\text{画素}$ のサイズとされる画像として保存される。

【0211】

図2に戻るに、ブロック202では、システムは、臨床被験者群から取得された複数の医用画像を教師なし機械学習モデル内へと入力して、潜在空間内での複数の埋め込みを取得するように構成されていることができる。一部の実施形態では、システムは、複数の医用画像を上述のように複数の画像タイルに分割し、及び、各画像タイルを教師なし機械学習モデル内へと入力して、対応するタイル埋め込みを取得する。図3Aを参照するに、画像タイル306-1-306-M1の各画像タイルは教師なし機械学習モデルに入力されて、タイル埋め込みを取得できる。例えば、画像タイル306-1は（処理Aによって表されている）教師なし機械学習モデルに入力されて、タイル埋め込み308が取得される。したがって、生検1については、システムが、それぞれが画像タイル306-1-306-M1に対応するタイル埋め込み308-1 - 308-M1を取得する。同様に、生検nについては、システムが、画像タイル306-1-306-Mnに対応するタイル埋め込み358-1 358-Mnを取得できる。

【0212】

一部の実施形態では、システムは、複数の画像タイルのサブセットのみを、教師なし機械学習モデルによるさらなる処理に付すために選択する。例えば、システムは、生検試料を表す所与の画像タイルの部分を決定して、該部分が所定の閾値を超過する場合（例えば、 $>90\%$ ）にのみ画像タイルを入力できる。別の例を挙げるに、システムは、所与の生検の結果もたらされる画像タイルのカウントを決定して、カウントが所定の閾値を超過する場合（例えば、 >70 タイル）にのみ画像タイルを入力できる。

【0213】

図4Aは、ブロック202にて用いられる例示的な教師なし機械学習モデルについて示

す。図 4 A を参照するに、教師なし機械学習モデル 4 0 4 は、入力画像タイル 4 0 2（例えば、図 3 A の画像タイルの 1 つ）を受信して、出力タイル埋め込み 4 0 6 を提供するように構成されていることができる。タイル埋め込み 4 0 6 は入力画像タイル 4 0 2（例えば、タイル 3 0 6 - 1）の潜在空間内でのベクトル表現とされ得る。入力画像を埋め込みに変換することによって、元のデータのサイズ及び寸法を相当に減じることができる。例を挙げるに、224画素 × 224画素のサイズとされた画像タイルは、2048次元ベクトルに縮減されることができる。より低次元な埋め込みは、後述のように、下流側での処理に用いられることができる。

【0214】

一部の実施形態では、教師なし機械学習モデル 4 0 4 は、訓練済みの対照学習アルゴリズム（contrastive learning algorithm）である。対照学習とは、どのデータポイントが類似又は別異であるかをモデルに教える（teach）ことによってラベルを伴わずにデータセットの一般的特徴を学習（learn）するのに用いられる機械学習手法を、指す。対照学習モデル（contrastive learning model）は画像データから埋め込みを抽出できるのであって、これはそのようなデータにさもなくば割り当てられてしまいかねないラベルについて線形的に予測的である。適した対照学習モデルは、対照損失を最小化することによって訓練（train）されるのであって、これによって同じサンプル画像の異なる増強結果からの埋め込みの間の類似性が最大化され、且つ、異なるサンプル画像の埋め込みの間の類似性が最小化される。例えば、モデル（例えば、教師なし機械学習モデル 4 0 4）は、回転、反転、クロッピング、及び／又は色ジッタリングに関して不変となるタイル埋め込みをタイル画像（例えば、入力画像タイル 4 0 2）から抽出できる。例示的な対照学習モデルとしては SimCLR 及び SwAV が含まれるが、任意の対照学習アルゴリズムを教師なし機械学習モデル 4 0 4 として用い得ることに留意されたい。

【0215】

入力画像（例えば、入力画像タイル 4 0 2）を処理するために教師なし機械学習モデル 4 0 4 が用いられる前に、これを訓練しておくことを要する。図 4 B は、いくつかの実施形態による、例示的な対照学習アルゴリズムの訓練のためのデータアーキテクチャ 4 5 0 について示す。一部の実施形態では、図 4 A の教師なし機械学習モデル 4 0 4 は、図 4 B のエンコーダの 1 つとすることができる（例えば、エンコーダ 4 6 2 a, 4 6 2 b）。訓練に際しては、元の画像 4 5 2 を取得できる。データ変換又は増強 4 5 4 を元の画像 4 5 2 に対して適用して、ビューイングステージ 4 5 6 にて 2 つの強化画像 4 5 8 a, 4 5 8 b を取得できる。例えば、システムは、2 つの別個のデータ増強オペレータ（例えば、クロッピング、反転、色ジッタリング、グレースケール、ブラー）を適用して、増強画像 4 5 8 a, 4 5 8 b を取得できる。一部の実施形態では、2 つ以上の画像をもたらす得る。例えば、N 回の（類似し得る又は異なり得る）データ増強 4 5 4 を元の画像 4 5 2 に対して施して、N 個の増強画像を取得できる。

【0216】

データアーキテクチャ 4 5 0 はモデル訓練 4 5 0 内にエンコードステージ 4 6 0 を含み得るのであり、各増強画像（例えば、増強画像 4 5 8 a, 4 5 8 b）はそれぞれ 1 つのエンコーダ 4 6 2 a, 4 6 2 b によってエンコードされ得る。各増強画像 4 5 8 a, 4 5 8 b はエンコーダを通して、潜在空間にてそれぞれのベクトル表現 4 6 4 a, 4 6 4 b を取得できる。一部の実施形態では、エンコーダ 4 6 2 a, 4 6 2 b は共通の重み付けを有している。一部の実施形態では、各エンコーダ 4 6 2 a, 4 6 2 b はニューラルネットワークとして実装されている。例えば、エンコーダは、残差ニューラルネットワーク（「ResNet」）アーキテクチャのバリエーションを用いて実装できる。図示のように、エンコーダ 4 6 2 a, 4 6 2 b はそれぞれ、ベクトル表現 4 6 4 a（例えば、増強画像 4 5 8 a に基づいてエンコーダ 4 6 2 a によって出力された h_i ベクトル）及びベクトル表現 4 6 4 b（例えば、増強画像 4 5 8 b に基づいてエンコーダ 4 6 2 b によって出力された h_j ベクトル）を出力する。

【0217】

ベクトル表現 4 6 4 a, 4 6 4 b はそれぞれ、投影ヘッド 4 7 4 a, 4 7 4 b を通されて、2 つの

投影472a, 472bを取得できる。一部の実施形態では、投影ヘッド474a, 474bは一連の非線形層（例えば、Dense - Reu - Dense層）を備えるのであって、これはベクトル表現に対して非線形変換を適用して投影を取得する。例えば、投影ヘッド474aはdense層466aとReLU層468aとdense層470aとを含み得るのであり、また、投影ヘッド474bはdense層466bとReLU層468bとdense層470bとを含み得る。投影ヘッド474a, 474bのそれぞれは不変特徴を増幅するように構成されていることができ、また、同じ画像の異なる変換を識別することについてのネットワークの能力を最大化する。

【0218】

訓練に際しては、同じ入力画像（元の画像452）についての投影472a, 472bの間の類似性は最大化されることができる。例えば、投影472a, 472bに基づいて損失が算出され、また、損失に基づいて各エンコーダ462a, 462bを更新して、2つの潜在表現（例えば、表現464a, 464b）の間の類似性を最大化できる。同様に、異なる入力画像の投影間の類似性は、訓練に際して最小化できる。一部の例では、投影間の合致度（即ち、類似性）を最大化するためにシステムは類似性メトリックを余弦類似性として定義できる：

10

【数1】

$$\text{sim}(u, v) = \frac{u^T v}{\|u\| \|v\|}$$

20

【0219】

一部の例では、システムは、正規化温度スケールクロスエントロピー損失を最小化することによってネットワークを訓練する：

【数2】

$$\ell_{i,j} = -\log \frac{\exp\left(\frac{\text{sim}(z_i, z_j)}{\tau}\right)}{\sum_{k=1}^{2N} \mathbb{I}_{[k \neq i]} \exp\left(\frac{\text{sim}(z_i, z_j)}{\tau}\right)}$$

30

ここで、 τ は調整可能な温度パラメータを表し、また、 z_i, z_j はそれぞれ投影472a, 472bに対応する。したがって、訓練450を介して、データアーキテクチャ450のエンコーダ462a, 462bは、入力画像の不変特徴を保全しつつ画像特有特徴（例えば、撮像アングル、解像度、アーチファクト）を最小化するベクトル表現を出力することを学習するように構成されることができる。

40

【0220】

図2に戻るに、一部の実施形態では、教師なし機械学習モデル404は、非医用画像を用いて訓練でき、そして、医用画像を処理するのに用い得る（ブロック204）。一部の実施形態では、モデルはまず非医用画像を用いて訓練されて、そして、いくつかのエポックに亘って医用画像を用いて微調整（例えば、再訓練）されて、そして、入力医用画像を処理するのに用いられ得る（ブロック204）。一部の実施形態では、教師なし機械学習モデル404を微調整するのに用いられる医用画像は、生検1-nからの画像タイルから選択されることができる。換言するに、生検1-nからの画像タイル306-1 306-M₁, ..., 356-1 356-M_nは、まずモデルの訓練に用いることができ、そして、訓練済みモデルに入力してタイル埋め込みを取得できる。

50

【0221】

タイル埋め込み308-1 308- M_1 , ..., 358-1 358- M_n は、生検レベルで集約できる。集約は、生検内の全タイルに亘ってタイル埋め込みを平均することを伴い得る。図3Aを参照するに、生検1のタイル埋め込み308-1-308- M_1 を集約して生検埋め込み310を取得できる。同様に、生検nのタイル埋め込み358-1-358- M_n を集約して生検埋め込み360を取得できる。各生検埋め込みは、生検にて反映されている着目疾患との関係での表現型状態に対応する。例示的な実装例では、6,782件の生検は6,782件の生検埋め込みを結果としてもたらす。各生検埋め込みは2048次元ベクトルであり、これは複数の2048次元タイル埋め込みベクトルを平均することによって算出される。このデータは、 $6,782 \times 2,048$ の行列として表すことができる ($X \in \mathbb{R}^{N \times L}$, $N=6,782$ 及び $L=2,048$)。

10

【0222】

一部の実施形態では、さらなる処理前に、生検埋め込み310, 360は正規化されて、そして、埋め込み次元数の平方根の逆数によって再スケーリングされる。正規化によって、説明されるように、生検埋め込みに基づいて適合化された線形予測モデルの性能が向上し得る。

【0223】

ブロック204では、システムは、線形回帰モデルを用いて複数の埋め込みに対応する複数の予測連続医療診断スコアを取得できる。上述のように、各埋め込みは、画像データに反映されている着目疾患との関係での表現型状態に対応し得るのであり、また、画像データに反映されているリッチなセマンティック情報（例えば、組織の微細構造の特徴）を捉えている場合がある。埋め込みは、画像データについて細かい粒度の疾患関連ラベルを生成するために用いられ得る。例えば、各埋め込みは、着目疾患に関する連続医療診断スコアを予測するために線形モデルによって用いられ得る。予測連続スコアは、図3Aの関連データ304を参照して説明されている生検の割り当て済みスコアに比して優れている。具体的には、予測スコアは、離散的な値ではなく連続的な値をとるのであり（例えば、病理学者によって割り当てられた値（0, 1, 2, 3, 4, 5, 6））、故に病理学者によって割り当てられた離散スコアよりもニュアンスが多く捕捉されることとなる。埋め込み（及び画像データ）に連続的なスコアを割り当てる能力は、より高い精密性及び下流側の分析における向上した統計能力をもたらすのであり、例えば示される疾患状態の各々と遺伝的バリエーションとの間により緊密な関連を取得し得る。

20

30

【0224】

図3A及び図3Bにて示されている例について述べるに、生検1-nについての生検埋め込みは、線形回帰モデルを生成（例えば、適合化）するために用いられることができ（例えば、埋め込みスコア予測モデル312）、そしてこれは生検1-nについての予測医療診断スコアを生成するために用いられることができる。

【0225】

一部の実施形態では、線形回帰モデル（例えば、埋め込みスコア予測モデル312）は、入力として埋め込みを受信し及び予測医療診断スコアを出力するように構成されている。一部の実施形態では、線形回帰モデルは線形混合モデル（LMM、linear mixed model）として実装されており、これは固定型効果及びランダム型効果の両方を可能とする単純な線形モデルの拡張版である。線形混合モデルは、後述のように、共変量（例えば、性別、年齢、及び／又は臨床試験群等）について加味しつつ関連検定を行うことを可能とする。

40

【0226】

一部の実施形態では、線形回帰モデルは次のものとすることができる：

$$y = Fb + u + \psi$$

ここで：

$y \in \mathbb{R}^{N \times 1}$ はN個の個体についての医療診断スコア（例えば、生検レベル線維症スコア）を表す。

50

$F \in \mathbb{R}^{N \times K}$ はK個の共変量（例えば、性別、年齢、臨床試験群）についての行列を表す。

$b \in \mathbb{R}^{K \times 1}$ は共変量効果サイズベクトルを表す。該ベクトルは様々なモデルパラメータを含み、また、線形モデルにおける共変量の重みである。具体的には、K個の共変量についてK個の重みがそれぞれある。

$u \sim N(0, \sigma_x^2 XX^T)$ は組織学的埋め込みから寄与分をモデリングする。

$\psi \sim N(0, \sigma_e^2 I_N)$ は残差iidガウスノイズである。

$X \in \mathbb{R}^{N \times L}$ はL次元のN個の個体についての生検埋め込みの行列である（例えば、 $L=2048$ ）。

$I_N \in \mathbb{R}^{N \times N}$ は $N \times N$ 恒等行列を表す。

σ_x^2 及び σ_e^2 はスカラーモデルパラメータである。

【0227】

一部の実施形態では、システムは、医療診断スコアを予測するために線形回帰モデルについてパラメータを生成する（モデルに関して適合化をなす）。モデルについては、生検データを用いて適合化をなし得るのであり、生検埋め込み及び対応する医療診断スコアが含まれる。図3Aに示されている例では、生検1-nについての生検埋め込みは埋め込みスコア予測モデル312に関して適合化をなすために用いられる。

【0228】

図5は、医療診断スコアを予測するための例示的な機械学習モデル（例えば、埋め込みスコア予測モデル312）の適合化について示す。一部の実施形態では、機械学習モデルは線形回帰モデルとされ得る。図示のように埋め込みスコア予測モデル504は訓練データ510を用いて適合化され得る。訓練データ510は、生検1-nのデータを備えることができ、これには生検埋め込み（例えば、生検埋め込み310）及び対応する生検レベル割り当て済み医療スコア（medical score）が含まれる。例えば、埋め込みは図3Aを参照して説明されている生検1-nについての生検埋め込みであり、医療診断スコア（medical diagnosis score）は生検1-nについての割り当て済み線維症スコアである（例えば、図3Aの関連データ304, 354の一部として記憶されている）。

【0229】

例示的な実装例では、6,782件の生検は6,782件の生検埋め込みを結果としてもたらす。各生検埋め込みは2048次元ベクトルとすることができ、これは複数の2048次元タイル埋め込みベクトルを平均することによって算出される。このデータは、 $6,782 \times 2,048$ の行列として表すことができる（ $X \in \mathbb{R}^{N \times L}$, $N=6,782$ 及び $L=2,048$ ）。行列は機械学習モデル（例えば、線形回帰モデル）を適合化するための入力行列Xとして用いられる。共変量行列たる $F \in \mathbb{R}^{N \times K}$ は切片（即ち、全てが1とされる単一の列（ $K=1$ ））を含む。

【0230】

機械学習モデルの負対数限界尤度は次のように定めることができる：

$$f(b, \sigma_x^2, \sigma_e^2) = -\log N(y; Fb, \sigma_x^2 XX^T + \sigma_e^2 I_N),$$

パラメータは b 、 σ_x^2 及び σ_e^2 とされる。

【0231】

b 、 σ_x^2 、 σ_e^2 についての最大尤度推定器（MLE、Maximum likelihood estimator）は、類似の効率的な手順を用いることによって $f(b, \sigma_x^2, \sigma_e^2)$ を最小化することによって取得できる。一部の実施形態では、（i）共分散が対角的である空間にてデータを回転させること（例えば、 $f(b, \sigma_x^2, \sigma_e^2) = -\log N(U^T y; U^T F b, \sigma_x^2 S + \sigma_e^2 I_N)$ ）（ XX^T の固有値分解が USU^T として表される）、及び、（ii） $f(b, \sigma^2, \delta) = -\log N(U^T y; U^T F b, \sigma^2 \delta S + \sigma^2 (1 - \delta) I_N)$ としてモデルを再パラメータ化することによってこれを達成できる。新しい形式では、デルタ（ δ ）のそれぞれの値について b 、 σ^2 のMLEについて閉形式解を伴うデルタ（ δ ）に関するグリッド探索を行うことによって最適化を進めることができる。この最適化は演算的にも効率的である。最適化後においては、 σ_x^2 及び σ_e^2 のMLEは、MLEたる

10

20

30

40

50

【数 3】

$$\hat{\sigma}^2 \text{ 及び } \hat{\sigma}_x^2 \text{ からそれぞれ } \hat{\sigma}_x^2 = \hat{\sigma}^2 \hat{\delta} \text{ 及び } \hat{\sigma}^2 = \hat{\sigma}_x^2 (1 - \hat{\delta})$$

として算出できる。

【0232】

機械学習回帰モデル（例えば、埋め込みスコア予測モデル312, 504）が適合化された後に、予測連続医療診断スコアを取得できる。一部の実施形態では、予測連続医療診断スコアは、生検埋め込みを用いて単一除外（LOO、Leave-one-out）予測として取得できる。単一除外予測では、単一ポイントを除いて全てのデータに対して適合化されたモデル全体を用いるのであり、そして、該ポイントにて予測を立てる。LOOアプローチは演算的に低コストであり、また、予測をより良い結果へと寄せることできる。例えば、予測スコアたる y についての LOO 予測たる

10

【数 4】

$$\hat{y}_i$$

は適合化済み LMM から次のように取得できる。

$$\hat{y}_i = \frac{H_{ii} \tilde{y} - H_{ii} \tilde{y}}{1 - H_{ii}}$$

20

ここで、

$$\tilde{y} = y - F\hat{b}, H = \hat{\sigma}_x^2 XX^T (\hat{\sigma}_x^2 XX^T + \hat{\sigma}_e^2 I_N)^{-1} \text{ であり、}$$

また、 $\hat{b}$ 、 $\hat{\sigma}_x^2$ 及び $\hat{\sigma}_e^2$ は b 、 σ_x^2 及び σ_e^2 の MLE である。

【0233】

したがって、システムは、生検埋め込み（例えば、生検埋め込み310, 360）から予測された

30

【数 5】

予測医療診断スコア $\tilde{y} \in R^{N \times 1}$

を取得する。

【0234】

図3Bを参照するに、埋め込みスコア予測モデル312は、生検1-nについての予測スコアを生成するために用いられることができる。例えば、医療診断スコアは生検レベル線維症スコアとされ得る。生検についての予測スコアは、図3Aの関連データ304を参照して説明されている生検の割り当て済みスコアとは異なる。具体的には、予測スコアは、離散的な値ではなく連続的な値をとるのであり（例えば、病理学者によって割り当てられた値（0, 1, 2, 3, 4, 5, 6））、故により高い精密性及び向上した統計的パワーを下流側での分析にもたらす。

40

【0235】

図2に戻るに、ブロック206では、システムは、複数の医療診断スコアを、複数の医用画像の取得元の臨床被験者群によって表現されている複数の候補遺伝的バリエーションの候補遺伝的バリエーションと、関連付けることができる（例えば、関連について検定をなす）。これをなすに際して、システムは、特定の候補遺伝的バリエーションと着目疾患との間に統

50

計的に有意な関連があるかについて決定する。

【0236】

一部の実施形態では、複数の医療診断スコアを候補遺伝的バリエーションと関連付けるステップは、候補遺伝的バリエーションを受信し及び医療診断スコアを出力するように構成されたバリエーション特有モデルを生成するステップを含む。例えば、候補遺伝的バリエーションが遺伝的バリエーションAである場合、システムは、遺伝的バリエーションAについて示す値を受信し、及び、予測線維症スコアを出力するように構成されたモデルを生成できる。図3Bに示されている例では、バリエーション特有バリエーションスコア予測モデル316は、後述のように生成できる。

【0237】

バリエーション特有モデル（例えば、バリエーション特有バリエーションスコア予測モデル（variant-specific variant-score prediction model）316）は線形回帰モデルとすることができる。一部の実施形態では、線形回帰モデルは次のものとして行うことができる：

【数6】

$$y = Fb + x\beta + \psi$$

ここで、 $\psi \sim N(0, \sigma_e^2 I_N)$ である。

である。

【0238】

$y \in \mathbb{R}^{N \times 1}$ はN個の個体についての医療診断スコア（例えば、生検レベル線維症スコア）を表す。

【0239】

$X \in \mathbb{R}^{N \times 1}$ は検査されている遺伝的バリエーションの遺伝子型ベクトルを表す。例えば、母集団において2つのアレルを伴う遺伝的バリエーションについては、各個体が最低頻度アレルのコピーを0, 1,又は2個有しているかに応じて、遺伝子型ベクトルは0, 1,又は2の値をとる。

【0240】

$F \in \mathbb{R}^{N \times K}$ はK個の共変量（例えば、性別、年齢、臨床試験群）についての行列を表す。1つの例示的な実装例では、5つの共変量がある：性別、年齢、及び3つの治療群。Fの第1の列は患者の性別についての二値指標であり（例えば、XX染色体については0であり、XY染色体については1）、第2の列については患者の年齢が含まれ、また、残余の列は3つの治療に関する二値指標である（例えば、患者が特定の治療を受けた場合には1であり、その他の場合は0）。

【0241】

図6は、いくつかの実施形態による、医療診断スコアを予測するための例示的なバリエーション特有モデルの適合化について示す。図示のようにバリエーションスコア予測モデル604は訓練データ610を用いて適合化される。訓練データ610は生検1-nのデータを備えるのであり、これには被験者の遺伝的バリエーション値及び対応する予測医療スコアが含まれる。例えば、データには、生検1が行われる被験者の遺伝的バリエーション値、及び生検1についての予測線維症スコア等が含まれる。

【0242】

図3Bに戻るに、示されているように、バリエーション特有モデル（variant-specific model）316は、生検1-nについての予測線維症スコアを用いて適合化されることができる。一部の実施形態では、システムは、複数の候補遺伝的バリエーションに対応する例えばバリエーションスコア予測モデル（variant-score prediction model）316, 322等の複数のバリエーション特有モデルについて適合化をなす。例えば、バリエーション特有モデル（variant-specific model）316は候補遺伝的バリエーションAに関して特有のものとして行われることができ、候補遺伝的バリエーションAについて示す値を受信し及び医療診断スコアを

予測するように構成されていることができ、また、バリエーションスコア予測モデル (variant-score prediction model) 322 は候補遺伝的バリエーション B に関して特有のものとされることができ、候補遺伝的バリエーション B について示す値を受信し及び医療診断スコアを予測するように構成されていることができる。したがって、システムは、全ての着目遺伝的バリエーション (例えば、100,000 種、100 万種、1,000 万種のバリエーション、又はバリエーションについての他の量) についてバリエーション特有モデル (例えば、100,000 種、100 万種、1,000 万種のモデル、又はモデルについての他の量) を生成できる。

【0243】

ブロック 208 では、システムは、着目疾患と各候補遺伝的バリエーションとの間の相関メトリックに関連に基づいて決定して、少なくとも 1 つの着目遺伝的バリエーションを識別することができる。相関メトリックは、着目疾患に対して候補遺伝的バリエーションが及ぼすインパクトについて示す。一部の実施形態では、相関メトリックは、遺伝的バリエーションと着目疾患との間での関連について定量化する。図 3B を参照するに、バリエーション特有モデル 316 は、着目疾患と遺伝的バリエーション A との間の相関メトリック 318 を決定するために用いられることができ (医療診断スコアによって表される)、また、バリエーションスコア予測モデル 322 は、着目疾患と遺伝的バリエーション B との間の相関メトリック 324 を決定するために用いられること等ができる。したがって、各着目遺伝的バリエーションについて相関メトリックを算出できる。

【0244】

一部の実施形態では、相関メトリックは線形回帰モデル (例えば、バリエーションスコアモデル (variant-score model) 604) の P 値である。システムは $\beta \neq 0$ について検定する。P 値は標準対数尤度レシオ検定手順にて求めることができ、また、効果サイズ及び標準誤差は古典線形モデル理論にて求めることができる。手順では、被検定バリエーションと医療診断スコアとの間の関連についての P 値、バリエーションの効果サイズ (線形モデルにおける推定器の重み β)、標準誤差 (モデルからの効果サイズ推定についての誤差)、又は他の情報が戻される。

【0245】

一部の実施形態では、相関メトリックは 1 つ以上の所定の閾値と比較されて、着目遺伝的バリエーションと着目疾患との間で有意な関連があるかについて決定される。例えば、P 値が所定の閾値 (例えば、 5×10^{-8}) よりも小さい場合、システムは、有意な関連があると決定することができる。一部の実施形態では、システムは、比較に基づいて、着目遺伝的バリエーションと着目疾患との間の関係を識別する。一部の実施形態では、関係は因果関係である。識別された関係又は関連は、後述のように診断及び治療や薬剤の開発に活用できる。

【0246】

図 2 ~ 6 を参照して本明細書にて説明される手法は、単に例示的なものにすぎず、また、類似の手法を任意の着目表現型と共変量との間の関連を発見するために活用できる。図 2 ~ 6 を参照して本明細書にて説明される一部の例について述べるに：表現型データには医用画像が含まれ；着目表現型は着目疾患 (例えば、NASH) であり、これは医療診断スコア (例えば、線維症スコア) によって表されることができ；着目共変量は着目遺伝的バリエーションである。もっとも、本明細書にて説明される手法は別の着目表現型と別の共変量との間の関連を発見することに関して活用することもできるということに留意されたい。例示的な表現型データとしては次のものが含まれるがこれらには限定されない：医用画像 (例えば、MRI、X 線、CT スキャン)、病理組織学データ (例えば、H & E 染色、トリクローム染色)、臨床バイオマーカーデータ (例えば、プロテオミクスや cfDNA を含む血液検査測定値、認知／精神評価スコア、マイクロバイオーーム評価等)、又はゲノムバイオマーカーデータ (例えば、バルク RNA-seq、メチル化データ、ゲノム配列データ、エピジェネティック配列データ等)。例示的な表現型には次のものが含まれる：着目疾患、遺伝子発現、メタボロミクス、プロテオミクス、トランスクリプトミクス、又はリピドミクス等。例示的な共変量クラスには次のものが含まれる：人口統計情報 (例えば、

年齢や性別)、臨床共変量(例えば、疾患状態、臨床スコア又は血液バイオマーカ)、ゲノムデータ(例えば、遺伝子データ、発現データ、メチル化データ、等)等。

【0247】

一部の実施形態では、システムは、比較に基づいて、着目遺伝的バリエーションと着目疾患との間の関係を識別する。一部の実施形態では、関係は因果関係である。識別された関係は、新規被験者において着目疾患が発生する尤度について決定するために用いることができ、特定の症状又は他の疾患関連因子の存在をもとに新規被験者に疾患があるかをより高い信頼度を伴って診断できる。また、新規被験者にて遺伝的バリエーションが識別された場合、着目疾患についての診断又は予後診断を適宜提供でき、これには疾患がどのように進行することが期待され得るかについての予後診断が含まれ得る。例えば、NASHについて着目遺伝的バリエーションが発見された場合、新規被験者についてゲノム検査を行ってバリエーションを検出できる。バリエーションが存在する場合、システムは、疾患の発現を予測でき、診断を提供でき、及び／又は新規被験者においてどのように疾患が進行し得るかについて予後診断を提供できる。

10

【0248】

一部の実施形態では、識別済みの関係は、バイオマーカや疾患介入用の標的の識別に関して用いられ得る。例えば、発病又は進行との関係で因果関係にあると識別された特定の遺伝的バリエーションに関しては、治療的介入に関してそれらが適するかをさらに評価できる。また、遺伝的バリエーションが着目疾患に対して及ぼす機能的インパクトを考慮して、治療的介入のための追加的な生物学的標的を識別することができる。そのような生物学的標的は、例えば、着目遺伝的バリエーションが所在する遺伝子によって転写されるたんぱく質の場合がある。そのような生物学的標的は、着目疾患の少なくとも1つの遺伝的バリエーションの機能的インパクトに関して変更、オフセット、緩和、補足や補完するものと見込まれている他の遺伝子、たんぱく質又は代謝産物を含み得る。一部の実施形態では、識別済みの関係は、治療を開発するために用いられ得る。例えば、関連付けられている遺伝的バリエーションを有する被験者群に以前施された候補治療のインパクトを、候補治療に関しての変更型又は類似型の開発に際して、考慮できる。以前施された候補治療の効果に関しての縦型の及びクロスセクショナル型のデータの双方は、着目疾患の予測状態又は進行と比較でき、量的観点から、着目疾患の状態又は進行に関して着目遺伝的バリエーションがどの程度まで以前施された候補治療のインパクトに影響を与えるかを決定できる。そのような候補治療に関しての変更型又は類似型については、疾患及びその被験者群との関係での遺伝的バリエーションのインパクトを考慮して、強化された治療的効果又は減じられた有害副作用を有するように選択できる。また、遺伝的バリエーションを反映している疾患モデルは、治療的候補をスクリーニングするために用いられ得る。例えば、ゲノム特徴との関連付けの場合、統計的に有意な関連付けは新規な候補薬剤標的又は人間の疾患に関与する主要経路の発見につながり得る。同様に、関連付けられている遺伝的バリエーションを有する被験者群に以前施された候補治療(candidate therapies)のインパクト(疾患状態に対しての有害なインパクト又は副作用を含む)は、組合せ型治療の開発に用いられ得る。したがって、本明細書にて説明される手法は、異なる遺伝的背景を有する被験者間で治療的候補(therapeutic candidates)に対しての起こるであろう応答を予測するために、特定の治療を受けるべき適切な患者コホートを識別するために、及び結果を最適化するために臨床試験を一般的に設計するために用いることができる。

20

30

40

【0249】

一部の実施形態では、識別された関係は、治療を選択的に施し、調整し、又は適用するために用いられ得る。一部の実施形態では、識別済みの関係は医学的提案を提供するために用いられ得る。一部の実施形態では、識別済みの関係に基づいてレポートが生成され得る。人口統計及び臨床的な特徴と関連付けられている場合、有意な関連は、新たな関連付けの発見につながったり(例えば、性別、年齢、コレステロール水準等との関連)、及び／又は、データセット内のテクニカルバイアス(例えば、特定の臨床センタとの関連)の探知につながり得る。

50

【0250】

図7はいくつかの例による、着目疾患との関係で少なくとも1つの着目遺伝的バリエーションを識別するための処理700について示す。処理700は、例えば、ソフトウェアプラットフォームを実装する1つ以上の電子装置を使用して、実施される。いくつかの実施例では、処理700は、クライアント-サーバシステムを使用して実施され、処理700のブロックは、任意の様式において、サーバと1つ以上のクライアントデバイスとの間で分割される。したがって、処理700の部分がクライアントサーバシステムの特定の装置によって実行されると説明されているも、処理700はそのように限定されることは要さないことに留意されたい。他の例では、処理700はクライアント装置のみ又は複数のクライアント装置のみを用いて実行される。処理700では、いくつかのブロックは、随意に、組み合わせられ、いくつかのブロックの順序は、随意に、変更され、いくつかのブロックは、随意に、省略される。いくつかの実施例では、付加的ステップが、処理700と組み合わせて実施されてもよい。故に、図示（および下記により詳細に説明）されるような動作は、本質的に、例示的であって、したがって、限定として見なされるべきではない。

10

【0251】

図7を参照するに、ブロック702では、例示的なシステム（例えば、1つ以上の電子装置）は、臨床被験者群から取得された複数の医用画像を訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数の埋め込みを取得するステップであって、各埋め込みは複数の医用画像の1つ以上にて反映されている着目疾患との関係での表現型状態に対応する、ステップをなすように構成されていることができる。

20

【0252】

図8Aに示される例示的なワークフローでは、生検1-nがなされている。生検1-nは複数の被験者（例えば、がん患者）及び／又は同じ被験者の複数回の訪問（例えば、スクリーニング訪問回やフォローアップ訪問回）に対応し得る。示されている例では、着目疾患は非アルコール性脂肪性肝炎（NASH、non-alcoholic steatohepatitis）であり、また、生検はいくつかの臨床試験からのH&E染色肝生検からのものである。各生検（例えば、生検1）は、1つ以上の生検画像（例えば、生検1についての医用画像802）を結果としてもたらす。したがって、生検1-nは、医用画像802、医用画像852等を含む複数の医用画像を結果としてもたらす。

【0253】

医用画像は、様々なデータと関連付けられることができる。例えば、図8Aを参照するに、データ804は生検1についての医用画像802と関連付けられており、データ854は生検nについての医用画像852と関連付けられており以下同様。後述のように、データは着目疾患及び被験者についての既知情報を含み、着目遺伝的バリエーションについての情報も含まれる。

30

【0254】

一部の実施形態では、医用画像と関連付けられているデータは、着目疾患に関する割り当て済み医療診断スコアを含み得る。割り当て済み医療診断スコアは、疾患の状態又は進行具合を示し得る。一部の実施形態では、医療診断スコアは、一人以上の病理学者によって割り当てられた生検レベルスコアであることができ、これはそれらの病理学者の生検スライドの検討に基づいている。例えば、NASH疾患については、割り当て済み医療診断スコアはIshak線維症スコア等の生検レベル線維症スコアとされ得るのであり、これは離散値たる0, 1, 2, 3, 4, 5, 又は6とされる線維症の程度を示すものである。

40

【0255】

一部の実施形態では、医用画像と関連付けられているデータは、生検サンプルの取得元の被験者の遺伝子データを含み得る。例えば、データは、複数の遺伝的バリエーションについての被験者遺伝子情報を含み得る（例えば、100,000種、100万種、1,000万種のバリエーション）。例えば、データは、被験者が複数の遺伝的バリエーションの各々を有するかを示し得る。例えば、母集団において2つのアレルを伴う遺伝的バリエーションについては、個体が最低頻度アレルのコピーを0, 1, 又は2個有しているかに応じて、医用画像は遺伝的バリエ

50

アント値（0, 1, 又は2）と関連付けられることができる。一部の実施形態では、被験者遺伝子データは、被験者が着目疾患に罹患する尤度を示す多遺伝子リスクスコアとされ得る。

【0256】

一部の実施形態では、医用画像と関連付けられているデータは、生検サンプルの取得元の被験者の人口統計データを含み得る。人口統計データは例えば、性別、年齢、治療群（例えば、プラセボ、治療x、治療y）、又は他のデータを含み得る。

【0257】

図8Aを参照するに、医用画像は複数の画像タイルに分割できる。例えば、生検1の医用画像802を分割して画像タイル806-1-806-M₁を取得でき、また、生検nの医用画像852を分割して画像タイル856-1-856-M_nを取得できる。一部の実施形態では、画像タイルは、所定のグリッドを用いて医用画像から抽出されて、また、統一されたサイズの画像タイルとして記憶されることができる。1つの例示的実装例では、画像タイルはタイル寸法が192 μm $\times$ 192 μm とされる所定のグリッドを用いて抽出され、また、画像タイルは224画素 $\times$ 224画素のサイズとされる画像として保存される。一部の実施形態では、画像タイルサイズは動的に構成可能であり、事案毎に調整され得る。

【0258】

一部の実施形態では、システムは、複数の医用画像を複数の画像タイルに分割し、及び、各画像タイルを教師なし機械学習モデル内へと入力して、対応するタイル埋め込みを取得するように構成されている。図8Aを参照するに、画像タイル806-1-806-M₁のそれぞれは（処理Aによって表されている）教師なし機械学習モデルに入力されて、タイル埋め込み808-1-808-M₁が取得される。したがって、生検1については、システムが、それぞれが画像タイル806-1-806-M₁に対応するタイル埋め込み808-1-808-M₁を取得できる。同様に、生検nについては、システムが、画像タイル856-1-856-M_nに対応するタイル埋め込み858-1-858-M_nを取得できる。

【0259】

一部の実施形態では、システムは、複数の画像タイルのサブセットを、教師なし機械学習モデルによるさらなる処理に付すために選択する。例えば、システムは、生検試料を表す所与の画像タイルの部分を決定して、該部分が所定の閾値を超過する場合（例えば、>90%）にのみ画像タイルを入力できる。別の例を挙げるに、システムは、所与の生検の結果もたらされる画像タイルのカウントを決定して、カウントが所定の閾値を超過する場合（例えば、>70タイル）にのみ画像タイルを入力できる。

【0260】

ブロック702にて用いられる例示的な教師なし機械学習モデルが図4Aに示されており、これについては上述されている。例示的な対照学習アルゴリズムについては図4Bに示されており、これについては上述されている。一部の実施形態では、モデルは、非医用画像を用いて訓練でき、そして、医用画像を処理するのに用い得る（図7：ブロック702）。一部の実施形態では、モデルは先ず非医用画像を用いて訓練されて、そして、いくつかのエポック（例えば、5エポック、10エポック、50エポック、100エポック）に亘って医用画像を用いて微調整（例えば、再訓練）されて、そして、入力医用画像を処理するのに用いられ得る（図7：ブロック702）。一部の実施形態では、モデルを微調整するのに用いられる医用画像は、生検1-nからの画像タイルから選択されることができる。換言するに、生検1-nからの画像タイルは、先ずモデルの訓練に用いることができ、そして、訓練済みモデルに入力してタイル埋め込みを取得できる。

【0261】

タイル埋め込みは、生検レベルで集約できる。集約は、生検内の全タイルに亘ってタイル埋め込みを平均することを伴い得る。図8Aを参照するに、生検1のタイル埋め込み808-1-808-M₁を集約して生検埋め込み810を取得できる。同様に、生検nのタイル埋め込み858-1-858-M_nを集約して生検埋め込み860を取得できる。各生検埋め込みは、生検にて反映されている着目疾患との関係での表現型状態に対応する。例示的な実装

例では、6,782件の生検は6,782件の生検埋め込みを結果としてもたす。各生検埋め込みは2048次元ベクトルであり、これは複数の2048次元タイル埋め込みベクトルを平均することによって算出される。このデータは、 $6,782 \times 2,048$ の行列として表すことができる ($X \in \mathbb{R}^{N \times L}$, $N=6,782$ 及び $L=2,048$)。一部の実施形態では、生検埋め込みはさらなる処理前に正規化される。例えば、さらなる処理前に、生検埋め込みは正規化されて、そして、埋め込み次元数の平方根の逆数によって再スケーリングされ得る。正規化によって、説明されるように、生検埋め込みに基づいて適合化された予測モデルの性能が向上し得る。

【0262】

ブロック704では、システムは、複数の埋め込みを複数の候補遺伝的バリエーションの各候補遺伝的バリエーションと関連付けて、組織の微細構造と関連付けられている候補遺伝的バリエーションについてのサブセットを識別するように構成されていることができる。特に、遺伝的バリエーションと生検埋め込みとの間の関連について評価することによって、システムは、組織学的差異と関連付けられている複数の遺伝的バリエーションについてのサブセットを識別できる。本明細書にて説明される手法は、分析を特定の診断スコアに絞ることによって発見できないであろう組織学に影響を及ぼすバリエーションを識別することができる。

【0263】

一部の実施形態では、関連付けは、複数の候補遺伝的バリエーションの各候補遺伝的バリエーションについてバリエーション特有モデルを生成（例えば、適合化）することと、バリエーション特有モデルを評価して、1つ以上の閾値に基づいて遺伝的バリエーションと埋め込みとの間に 20
関連があるかを決定することとを伴う。関連がある場合、システムは、さらなる下流側での処理のために、候補遺伝的バリエーションを候補遺伝的バリエーションのサブセットに含めることができる。システムは、全ての着目遺伝的バリエーション（例えば、100,000種、100万種、1,000万種等のバリエーション）についてバリエーション特有モデル（例えば、100,000種、100万種、1,000万種等のモデル）を生成できる。

【0264】

図8Bにて示されている例では、生検埋め込みは、バリエーションAについて埋め込みバリエーションモデル812を、バリエーションBについて埋め込みバリエーションモデル814を、バリエーションZについて埋め込みバリエーションモデル816を、生成（例えば、適合化）するために用いられることができる。そして、システムは、各モデルについて相関メトリックを算 30
出することによって各モデルを評価することができる。例えば、システムは、埋め込みバリエーションモデル812について相関メトリック822を、埋め込みバリエーションモデル814について相関メトリック824を、埋め込みバリエーションモデル816について相関メトリック826を算出することができる。各相関メトリックについては、遺伝的バリエーションと埋め込みとの間に有意な関連があるかについて決定するために、評価をなすことができる（例えば、所定の閾値との比較）。示されている例では、システムは次のことを決定する：相関メトリック822に基づけばバリエーションAと埋め込みとの間には関連が存在すること、相関メトリック826に基づけばバリエーションZと埋め込みとの間には関連が存在すること、及び相関メトリック824に基づけばバリエーションBと埋め込みとの間には関連が存在していないこと。したがって、システムは、バリエーションA及びバリエーションZを更なる 40
処理のためにサブセットに含める一方で、バリエーションBをサブセットから排除することができる。遺伝的バリエーションのサブセット識別することによって、システムは更なる処理のために、医用画像（例えば、生検画像）に示されている組織学的特徴と関連付けられている遺伝的バリエーションを識別できる。遺伝的バリエーションについてのこのより小さなセットが下流側の分析にて探索されるであり、例えば後述のように各遺伝的バリエーションと着目疾患との間の関連を識別する。

【0265】

一部の実施形態では、バリエーション特有モデル（例えば、埋め込みバリエーションモデル812, 814, 816）は線形回帰モデルであり、これは、埋め込みを入力として受信し及び遺伝的バリエーションについて示す値を出力するように構成されている。一部の実施形態では、 50

前記線形回帰モデルは線形混合モデルとして実装されることができる。

【0266】

一部の実施形態では、線形回帰モデルは次のものとすることができる：

$g = Fb + u + \psi$ ここで：

【0267】

$g \in \mathbb{R}^{N \times 1}$ は検査されている遺伝的バリエーションの遺伝子型ベクトルを表す。N は生検画像データ及び遺伝子データの両方が利用可能となっている個体数である。

【0268】

$F \in \mathbb{R}^{N \times K}$ は K 個の共変量（例えば、性別、年齢、臨床試験群）についての行列を表す。例示的な実装例では、行列は性別、年齢及び 3 つの治療群についての情報を含む（ $K = 5$ ）。特に、F の第 1 の列は患者の性別についての二値指標であり（患者の染色体が X X の場合には 0 であり、患者の染色体が X Y の場合には 1）、第 2 の列については患者の年齢が含まれ、また、残余の列は 3 つの治療に関しての二値指標である（患者がその特定の治療を受けた場合には 1 であり、その余の場合は 0）。

【0269】

$X \in \mathbb{R}^{N \times L}$ は L 次元の N 個の個体についての生検埋め込みの入力行列である（例えば、 $L = 2048$ ）。

【0270】

$b \in \mathbb{R}^{K \times 1}$ は共変量効果サイズベクトルを表す。該ベクトルは様々なモデルパラメータを含み、また、線形モデルにおける共変量の重みである。具体的には、K 個の共変量について K 個の重みがそれぞれある。

【0271】

$u \sim N(0, \sigma_x^2 XX^T)$ は組織学的埋め込みから寄与分をモデリングする。

【0272】

$\psi \sim N(0, \sigma_e^2 I_N)$ は残差 iid ガウスノイズである。

【0273】

$I_N \in \mathbb{R}^{N \times N}$ は $N \times N$ 恒等行列を表す。

【0274】

σ_x^2 及び σ_e^2 はスカラーモデルパラメータである。

【0275】

一部の実施形態では、システムは、線形回帰モデルについてパラメータを生成する（モデルに関して適合化をなす）。モデルについては、生検データを用いて適合化をなし得るのであり、生検埋め込み（例えば、生検埋め込み 810, 860）及び対応する医療診断スコアが含まれる。図 8 B に示されている例では、生検 1-n についての生検埋め込み 810, 860 は埋め込みバリエーションモデル 812, 814, 816 の各々に関して適合化をなすために用いられる。

【0276】

図 9 は、例示的な線形回帰モデル（例えば、埋め込みバリエーションモデル 812）の適合化について示す図示のように埋め込みバリエーションモデル 904 は訓練データ 910 を用いて適合化され得る。一部の実施形態では、モデル 904 は線形回帰モデルとされ得る。訓練データ 910 は、生検 1-n のデータを備えることができ、これには生検埋め込み及び対応する遺伝的バリエーション値が含まれる。例えば、埋め込みは図 8 B を参照して説明されている生検 1-n についての生検埋め込み 810, 860 であり、また、遺伝的バリエーション値は生検 1-n にて特定の遺伝的バリエーション（例えば、バリエーション A）について示す値である（例えば、図 8 A の関連データ 804, 854 の一部として記憶されている）。

【0277】

例示的な実装例では、6,782 件の生検は 6,782 件の生検埋め込みを結果としてもたらす。各生検埋め込みは 2048 次元ベクトルとすることができ、これは複数の 2048 次元タイル埋め込みベクトルを平均することによって算出される。このデータは、 $6,782 \times 2,048$ の行列として表すことができる（ $X \in \mathbb{R}^{N \times L}$, $N = 6,782$ 及び $L = 2,048$ ）。行列

10

20

30

40

50

は、線形回帰モデルを適合化するための入力行列 X として用いられる。共変量行列たる $F \in \mathbb{R}^{N \times K}$ は切片（即ち、全てが 1 とされる単一の列（ $K=1$ ））を含む。

【0278】

線形回帰モデルの負対数限界尤度は次のように定めることができる：

【0279】

$f(b, \sigma_x^2, \sigma_e^2) = -\log N(y; Fb, \sigma_x^2 XX^T + \sigma_e^2 I_N)$ 、パラメータは b 、 σ_x^2 及び σ_e^2 とされる。

【0280】

$b, \sigma_x^2, \sigma_e^2$ についての最大尤度推定器（MLE、Maximum likelihood estimator）は、類似の効率的な手順を用いることによって $f(b, \sigma_x^2, \sigma_e^2)$ を最小化することによって取得できる。一部の実施形態では、（i）共分散が対角的である空間にてデータを回転させること（例えば、 $f(b, \sigma_x^2, \sigma_e^2) = -\log N(U^T y; U^T Fb, \sigma_x^2 S + \sigma_e^2 I_N)$ ）（ XX^T の固有値分解が USU^T として表される）、及び、（ii） $f(b, \sigma^2, \delta) = -\log N(U^T y; U^T Fb, \sigma^2 \delta S + \sigma^2 (1-\delta)I_N)$ としてモデルを再パラメータ化することによってこれを達成できる。新しい形式では、デルタ（ δ ）のそれぞれの値について b, σ^2 の MLE について閉形式解を伴うデルタ（ δ ）に関するグリッド探索を行うことによって最適化を進めることができる。この最適化は演算的にも効率的である。最適化後においては、 σ_x^2 及び σ_e^2 の MLE は、MLE たる

10

【数7】

$$\hat{\delta} \text{ 及び } \hat{\sigma}^2 \text{ からそれぞれ } \hat{\sigma}_x^2 = \hat{\sigma}^2 \hat{\delta} \text{ 及び } \hat{\sigma}_e^2 = \hat{\sigma}^2 (1 - \hat{\delta})$$

20

として算出できる。

【0281】

バリエント特有モデル（例えば、埋め込みバリエントモデル 812, 814, 816）が適合化された後、各適合済みモデルは、対応する遺伝的バリエントと埋め込みとの間に関連があるかについて決定するために評価される。一部の実施形態では、バリエント特有モデルを評価するステップは、バリエント特有モデルに基づいて相関メトリックを算出すること及び相関メトリックを所定の閾値と比較することを含む。

30

【0282】

一部の実施形態では、前記相関メトリックは前記バリエント特有モデルと関連付けられている P 値である。P 値は順列手順を介して取得できるのであって、実データからの対数尤度レシオ（LLR、log likelihood ratio）統計が、埋め込み行列内にて個体についての置換がなされる際に取得された LLR と、比較される（null モデルからの LLR）。一部の実施形態では、P 値は、実データ LLR より大きな並べ替え LLR の画分として定義される。

【0283】

例えば、各遺伝的バリエントについて、バリエント特有モデルは X の行についての K 個の順列に関して適合化できる。各遺伝的バリエントについて、この手順は、 K 個の LLR をもたらす。特に、実データを用いてバリエント特有モデルの適合化をなすと、1 つの LLR が結果としてもたらされ得る。バリエント特有モデルを、 X の行についての K 個の順列に関して適合化すると、 K 個の LLR が結果としてもたらされ得る。各遺伝的バリエントについて、P 値は、実データからの LLR が結果より大きな K 個の LLR の画分である。

40

【0284】

各遺伝的バリエントについての相関メトリックは所定の閾値に対して比較されて、遺伝的バリエントと埋め込みとの間に関連があるかを決定できる。一部の実施形態では、閾値は 5×10^{-8} に設定できる。例えば、閾値未満の P 値を伴うバリエントについては、埋め込みと関連付けられているものと決定される。示されている例では、システムは、次のこ

50

とについて決定するように構成されていることができる：相関メトリック 8 2 2 に基づけばバリエーション A と埋め込みとの間には関連が存在すること、相関メトリック 8 2 6 に基づけばバリエーション Z と埋め込みとの間には関連が存在すること、及び相関メトリック 8 2 4 に基づけばバリエーション B と埋め込みとの間には関連が存在していないこと。したがって、システムは、バリエーション A 及びバリエーション Z を更なる処理のためにサブセットに含める一方で、バリエーション B をサブセットから排除する。

【0285】

ブロック 706 では、システムは、候補遺伝的バリエーションのサブセットの各候補遺伝的バリエーションを着目疾患と関連付けて少なくとも 1 つの着目遺伝的バリエーションを識別する。これをなすに際して、システムは、特定の遺伝的バリエーションと着目疾患との間に統計的に有意な関連があるかについて決定する。識別された着目遺伝的バリエーションは、組織学及び着目疾患の両方と関連付けられる。

【0286】

一部の実施形態では、遺伝的バリエーションを着目疾患と関連付けることは、遺伝的バリエーション値を受信し、及び、着目疾患に関する医療診断スコアを出力するように構成されたバリエーション特有スコア予測モデルを生成（例えば、適合化）することを伴う。例えば、検査される遺伝的バリエーションが遺伝的バリエーション A である場合、システムは、遺伝的バリエーション A について示す値を受信し、及び、予測医療診断スコアを出力するように構成されたモデルを生成できる。図 8 B を参照するに：バリエーション A についてはバリエーション特有スコア予測モデル 8 3 2 が生成でき；バリエーション Z についてはバリエーション特有スコア予測モデル 8 3 6 が生成できる。バリエーション B についてはバリエーション特有スコア予測モデルが生成されないことができる。なぜならば、バリエーション B についての相関メトリック 8 2 4 が所定の閾値を超過する故にバリエーション B を更なる処理から排除できるからである。

【0287】

バリエーションスコアモデル（例えば、バリエーション特有スコア予測モデル 8 3 2）は線形回帰モデルとすることができる。一部の実施形態では、線形回帰モデルは次のものとしてすることができる：

$$y = Fb + x\beta + \psi$$

【0288】

ここで、 $\psi \sim N(0, \sigma_e^2 I_N)$ である。

【0289】

$y \in \mathbb{R}^{N \times 1}$ は N 個の個体についての医療診断スコア（例えば、生検レベル線維症スコア）を表す。

【0290】

$X \in \mathbb{R}^{N \times 1}$ は検査されている遺伝的バリエーションの遺伝子型ベクトルを表す。例えば、母集団において 2 つのアレルを伴う遺伝的バリエーションについては、各個体が最低頻度アレルのコピーを 0, 1, 又は 2 個有しているかに応じて、遺伝子型ベクトルは 0, 1, 又は 2 の値をとる。

【0291】

$F \in \mathbb{R}^{N \times K}$ は K 個の共変量（例えば、性別、年齢、臨床試験群）についての行列を表す。1 つの例示的な実装例では、5 つの共変量がある：性別、年齢、及び 3 つの治療群。F の第 1 の列は患者の性別についての二値指標であり（例えば、患者の染色体が XX の場合には 0 であり、また、患者の染色体が XY の場合には 1）、第 2 の列については患者の年齢が含まれ、また、残余の列は 3 つの治療に関する二値指標である（例えば、患者がその特定の治療を受けた場合には 1 であり、その余の場合は 0）。例示的なバリエーション特有スコア予測モデルについての適合化は図 6 を参照して説明されている。

【0292】

バリエーション特有スコア予測モデルが適合化された後、システムは、着目疾患と着目遺伝的バリエーションとの間の相関メトリックを決定できる。相関メトリックは、着目疾患に対して着目遺伝的バリエーションが及ぼすインパクトについて示す。図 8 B を参照するに、バリエ

10

20

30

40

50

ント特有スコア予測モデル (variant-specific score prediction model) 832は、着目疾患（医療診断スコアによって表される）と遺伝的バリエーションAとの間の相関メトリック842を決定するために用いられることができ、また、バリエーションスコアモデル (variant-score model) 836は、着目疾患と遺伝的バリエーションZとの間の相関メトリック846を決定するために用いられること等ができる。したがって、候補遺伝的バリエーションのサブセット内の各着目遺伝的バリエーションについて、相関メトリックが算出される。

【0293】

一部の実施形態では、相関メトリックは線形回帰モデルのP値である。システムは $\beta \neq 0$ について検定する。P値は標準対数尤度レシオ検定手順にて求めることができ、また、効果サイズ及び標準誤差は古典線形モデル理論にて求められる。手順では、被検定バリエーションと医療診断スコアとの間の関連についてのP値、バリエーションの効果サイズ（線形モデルにおける推定器の重み β ）及び標準誤差（モデルからの効果サイズ推定についての誤差）。

【0294】

一部の実施形態では、相関メトリックは1つ以上の所定の閾値と比較されて、着目遺伝的バリエーションと着目疾患との間で有意な関連があるかについて決定される。例えば、P値が所定の閾値（例えば、 5×10^{-8} ）よりも小さい場合、システムは、有意な関連があると決定する。一部の実施形態では、システムは、比較に基づいて、着目遺伝的バリエーションと着目疾患との間の関係を識別できる。一部の実施形態では、関係は因果関係である。識別された関係又は関連は、説明されるように診断及び治療や薬剤の開発に活用できる。例えば、図16～24を参照して説明されているように、システムが1つのデータセット（例えば、第1の臨床試験からの画像）から生成された埋め込みをもって、これを同じ予測的効果を伴って別のデータセットからの画像及び関連する遺伝的バリエーションデータ（例えば、第2の臨床試験からの画像及び関連する遺伝的バリエーション）に適用できるという意味でシステムは堅牢である。

【0295】

図7に戻るに、システムは、識別済み着目ジェネリックについての組織学的効果について可視化するためのシミュレーション画像を生成できる。この手順によって、これによらずにしては病理スコアについての関連研究によって検出ができないであろう関連する組織学的な変化についての可視化表示がもたらされる。特に、ブロック708では、システムは、着目疾患を表す複数のシミュレーション画像を生成するように構成されていることができる。複数のシミュレーション画像は、後述のように、少なくとも1つの着目遺伝的バリエーションの異なる値に対応する。

【0296】

一部の実施形態では、システムは、訓練済み生成的モデルを用いてシミュレーション画像の形式で予測画像を生成するように構成されていることができる。一部の実施形態では、訓練済み生成的モデルは敵対的生成ネットワーク (GAN) モデルである。図10は、いくつかの実施形態による、例示的なGANモデル1004について示す。一部の実施形態では、GANモデル1004は、訓練済み生成器コンポーネント1004aと訓練済み判別器コンポーネント1004bとを備え得る。訓練済み生成器コンポーネント1004aは、埋め込み（例えば、埋め込み1002-1 - 1002-k）とノイズ（不図示）とを受信し、及び、シミュレーション画像（例えば、シミュレーション画像1006-1 - 1006-k）を出力するように構成されていることができる。図10に示されているように、訓練済み生成器コンポーネント1004aは、埋め込み1002-1を受信してシミュレーション画像1006-1を出力し、埋め込み1002-2を受信してシミュレーション画像1006-2を出力し、以後同様にし、埋め込み1002-kを受信してシミュレーション画像1006-kを出力できる。

【0297】

一部の実施形態では、埋め込み1002-1-1002-kを生成するために、システムは、埋め込みを受信し及び着目疾患に関連する医療診断スコアを出力するように構成されたモデ

10

20

30

40

50

ルを生成（適合化）できる。一部の実施形態では、モデルは図8の埋め込みバリエーションモデル812等の線形回帰モデルである。そして、モデルは医療診断スコアを予測するために用いられ得る。例えば、各タイル埋め込み（例えば、タイル埋め込み808-1 - 808-M1）は、線形回帰モデル内へと入力されて（input into）、タイルレベル医療診断スコアを生成できる。生検1-nからのタイル埋め込みは、線形回帰モデルへと入力されて（input to）、タイルレベル医療診断スコアのセットを生成できる。システムは、第1～5パーセンタイルに属する対応する医療診断スコアを有する第1のタイル埋め込み群と、第95～99パーセンタイルに属する対応する医療診断スコアを有する第2のタイル埋め込み群とを識別することができる。第1の平均埋め込みは第1の埋め込み群を集約すること（併せてその平均をもとること）によって生成でき、また、第2の平均埋め込みは第2の埋め込み群を集約すること（併せてその平均をもとること）によって生成できる。そして、システムは、2つの平均埋め込み間で線形補間をなして埋め込み1-kを取得できる。埋め込み1002-1-1002-kに対応するシミュレーション画像1006-1 - 1006-kは、着目疾患の進行と関連付けられている組織学的効果について示すことができる。

10

20

30

40

50

【0298】

一部の実施形態では、埋め込み1002-1-1002-kを生成するために、システムは、埋め込みを受信し及び着目遺伝的バリエーション（例えば、図2及び図7にて識別されている着目遺伝的バリエーション）の値を出力するように構成されたモデルを生成（適合化）できる。遺伝的バリエーション値は着目遺伝的バリエーションについての遺伝子型情報について示す（例えば、個体がその遺伝子座にてマイナーアレルのコピーを0, 1,又は2個有している場合）。一部の実施形態では、モデルは、図3Aの埋め込みスコア予測モデル312と類似の線形回帰モデルである。そして、モデルは着目遺伝的バリエーションの値を予測するために用いられ得る。例えば、各タイル埋め込み（例えば、タイル埋め込み808）は、線形回帰モデルへと入力されて、タイルレベル遺伝的バリエーション値を生成できる。生検1-nからのタイル埋め込みは、線形回帰モデルへと入力されて、タイルレベル医療診断スコアのセットを生成できる。システムは、第1～5パーセンタイルに属する対応する遺伝的バリエーション値を有する第1のタイル埋め込み群と、第95～99パーセンタイルに属する対応する遺伝的バリエーション値を有する第2のタイル埋め込み群とを識別することができる。第1の平均埋め込みは第1の埋め込み群を集約すること（例えば、その平均をとること）によって生成でき、また、第2の平均埋め込みは第2の埋め込み群を集約すること（例えば、その平均をとること）によって生成できる。そして、システムは、2つの平均埋め込み間で線形補間をなして埋め込み1-kを取得できる。埋め込み1-kに対応するシミュレーション画像は、着目遺伝的バリエーションと関連付けられている組織学的効果について示すことができる。

【0299】

一部の実施形態では、GANモデルは条件付きGANモデルである。例えば、生成器は、埋め込みを条件として及びノイズを受信し、並びにシミュレーション画像を出力するように構成されていることができる。判別器は、シミュレーション画像又は現実画像とされ得る入力画像を受信し、及び入力画像をシミュレーション又は現実のものと分類するように構成されていることができる。訓練に際して、生成器はシミュレーション画像を生成し、また、シミュレーション画像及び現実の画像は判別器に分類のために提供される。判別器の出力に基づいて、生成器及び判別器は適宜更新して損失を最小化できる。一部の実施形態では、生成器及び判別器はニューラルネットワークである。

【0300】

一部の実施形態では、GANモデルは、プログレッシブGAN（pGAN）モデルを元に構築される。pGANモデルは次第に増加する画像解像度を伴ってプログレッシブな態様で訓練されて、訓練は安定化されまたモード崩壊が妨げられる。システムはプログレッシブ条件付きGAN（pcGAN）モデルにも拡張されるのであって、本明細書にて説明された医用画像に基づいて生成された埋め込みに対しての条件付けを可能とする。特に、生成器は、（条件としての）埋め込みx及び標準的な正規分布からサンプリングされたノ

イズベクトル u を受信するように構成されていることができる。また、判別器は生成された及び現実の画像並びに対応する埋め込みを受信する。図 11 は、いくつかの実施形態による、シミュレーション画像処理の生成について示す。示されているように、生成器 1104 は、埋め込み 1102 及びノイズベクトル 1108 を受信し及び画像 1106 を出力できる。さらに、同じ埋め込み及び異なるノイズが提供された場合、異なるシミュレーション画像が生成され得る。特に、各生成器 1104 が同じ埋め込み 1102 を受信するも異なるノイズベクトル 1108 を受信した場合、同じセマンティック生物学的内容を伴う異なるシミュレーション画像が生成され得る。したがって、埋め込みは複数の画像として可視化でき、埋め込みと関連付けられている組織学的特徴についての識別及び理解が促進される。

10

【0301】

1つの例示的な実装例では、埋め込み 1102 は 2048 次元の埋め込みであり、ノイズベクトル 1108 は標準的な正規分布からサンプリングされた 512 次元のベクトルである。埋め込み 1102 とノイズベクトル 1108 とは 512 次元ベクトル z に組み合わせることができることができ、これが pGAN 生成器の入力とされる。判別器は、画像 i に加えて、(条件としての) 埋め込み x (埋め込み 1102 等) を、入力として受信するように構成されていることができる。この追加によって、判別器は、入力埋め込みとの一貫性にも基づいて、現実の及び虚構の画像を分類できるようになる。例示的な実装例では、埋め込みは 2048 次元埋め込みであり、また、画像は 256x256 の画像である。

【0302】

例示的な実装例では、p cGAN モデルは、先ずタイル画像及び対応する埋め込みについて pGAN と同じパラメータが用いられて訓練される。訓練後は、p cGAN モデルの生成器は、入力として、2048 次元の埋め込みと標準的な正規分布からサンプリングされた 512 次元のベクトルを受信し及び 256 x 256 のシミュレーションタイル画像(その内容は入力埋め込みと一貫している)を出力できる。正規分布からサンプリングされた 512 次元のノイズベクトル 1108 を与えられた場合、システムは p cGAN を用いて k 個の画像を生成できるのであって、入力としてはそれぞれの k 個の補間埋め込み及びサンプリングされたノイズベクトル 1108 からのノイズが与えられる。したがって、この手順では、着目共変量(被分析生検についてのベクトル)、生検埋め込み(分析済み生検が行とされ、また、埋め込み次元が列とされる行列)、及びタイル埋め込み(対応するタイルが行とされ、また、埋め込み次元が列とされる行列)が入力として受け付けられて、また、一連の k 256x256 タイル画像が出力され得る。

20

30

【0303】

ブロック 710 では、システムは、ディスプレイ上で複数のシミュレーション画像 1106 を表示する(又は表示させる)ように構成されていることができる。ディスプレイは、例えば、コンピューティング装置等のユーザ装置(例えば、図 29 のコンピューティング装置 2900)のディスプレイスクリーンに対応し得る。一部の実施形態では、予測シミュレーション画像 1106 は、組織学的変化についてのアニメーションを作成するために用いられ得る。図 11 に示されるように GAN への入力として正規分布からの異なる 512 次元サンプルを提供することによって、複数のアニメーションが同じ着目共変量について作成され得る。一部の実施形態では、同じ予測シミュレーション画像 1106 のシリーズからの画像に関しては同じサンプルが用いられる。一部の実施形態では、予測シミュレーション画像はランキングされる。表示されるシミュレーション画像は予測シミュレーション画像の一部又は全部を含み得るのであり、また、ランキングに基づいて予測シミュレーション画像から選択されることができる。一部の実施形態では、ランキングは対応する医療診断スコアに基づいていることができる。例えば、予測シミュレーション画像は、その予測シミュレーション画像の演算医療診断スコアに基づいてランキングされることができる(埋め込みに基づいて決定できる)。一部の実施形態では、表示されるべき予測シミュレーション画像は、それらの対応する医療診断スコアに基づいて選択されることができる。例えば、低位線維症に対応する予測タイルを表し得る、第 1~5 パーセンタイルに

40

50

属する医療診断スコアを有する予測シミュレーション画像は、ユーザインタフェースでの表示のために選択されるかそこで表示されることができる。別の例を挙げるに、高位線維症に対応する予測タイルを表し得る、第95～99パーセンタイルに属する医療診断スコアを有する予測シミュレーション画像は、表示のために選択されるかそこで表示されることができる。予測シミュレーション画像を選択するためには別のパーセンタイル範囲を用いることもでき、先述の範囲は例示に過ぎない。

【0304】

図12Aは、いくつかの実施形態による、増大する線維症スコアを可視化するための予測画像タイルについての例示的なセット及び3つの予測シミュレーション画像を伴う例示的なシリーズについて示す。上述のように、システムは、第1～5パーセンタイルに属する対応する医療診断スコアを有する第1の画像タイル埋め込み群1202を識別することができるのであり、それによって低位線維症に対応する予測タイルを識別できる。また、システムは、第95～99パーセンタイルに属する対応する医療診断スコアを有する第2の画像タイル埋め込み群1204をも識別することができるのであり、それによって高位線維症に対応する予測タイルを識別できる。システムは、第1のタイル埋め込み群についての第1の平均埋め込み及び第2のタイル埋め込み群についての第2の平均埋め込みを取得することができる。線形補間を行ってk個の埋め込みを取得できる。例えば、 $k = 3$ である場合、システムは次のように3つの埋め込みを取得できる：第1の埋め込み（例えば、第1の平均埋め込み）、第2の埋め込み（例えば、第1の埋め込み及び第3の埋め込みの平均）、及び第3の埋め込み（例えば、第2の平均埋め込み）。3つの埋め込みは生成器に入力することができるのであり、線維症スコアの増大に関連する組織学的変化を可視化するための一連の画像1206を取得できる。

【0305】

図12Bは、いくつかの実施形態による、増大するPNPLA3タイルスコアを可視化するための予測画像タイル1252、1254についての例示的なセット及び3つのシミュレーション画像1256を伴う例示的なシリーズについて示す。画像1256は、一部の実施形態では、図12Aとの関連でせつめいされたのと類似の手法を用いて取得できる。

【0306】

一部の実施形態では、複数の予測シミュレーション画像はランキングされる。ランキングは、線維症スコアや他の尺度に基づいていることができる。一部の実施形態では、複数の予測シミュレーション画像の一部又は全部を表示することができる。複数の予測シミュレーション画像のサブセット（又は全部）はランキングに基づいて選択されることができる。また、予測シミュレーション画像はランキングに基づいて表示のために選択されることができる。

【0307】

図13では次のものが示されており、これらは本明細書にて説明される手法を用いて生成されている：脂肪症スコアと関連付けられている組織学的効果について可視化するための一連の例示的な予測シミュレーション画像1302；小葉内炎症スコアと関連付けられている組織学的効果について可視化するための一連の例示的な予測シミュレーション画像1304；バルーニングスコアと関連付けられている組織学的効果について可視化するための一連の例示的な予測シミュレーション画像1306；及び線維症スコアと関連付けられている組織学的効果について可視化するための一連の例示的な予測シミュレーション画像1308。一部の実施形態では、可視化画像タイル（例えば、画像1302-1308）は、特定の表現型について低スコアを有すると予測されるタイルに対応し得る。

【0308】

図14は、いくつかの実施形態による、医療診断スコアを予測するように構成された様々な線形モデルの性能について示す。示されているように、5つのモデルは次のことをなすように訓練されている：線維症スコアの予測；脂肪症スコアの予測；小葉内炎症スコアの予測；及び肝細胞バルーニングスコアの予測。各スコアタイプについて、バーの色は、凡例の色の順序と同じ順序にされる。

10

20

30

40

50

【0309】

初めの4つのモデルは図3Aの埋め込みスコア予測モデル312についての例である。事前訓練済みモデルは、事前訓練済みの対照モデル（例えば、SimCLR）によって生成された埋め込みに基づいて適合化された線形モデルを指す。事前訓練済み対照モデルは、同じ画像ドメイン内の画像（例えば、肝生検画像）を用いて微調整されていない（has not been fine-tuned）。事前訓練済み正規化済みモデルは、事前訓練済みの対照モデルによって生成され且つ正規化された埋め込みに基づいて適合化された線形モデルを指す。一部の実施形態では、埋め込みは正規化されて、そして、埋め込み次元数の平方根の逆数によって再スケーリングされる。微調整済みモデルは、微調整済みの対照モデル（例えば、SimCLR）によって生成された埋め込みに基づいて適合化された線形モデルを指す。微調整済み対照モデルは、同じ画像ドメイン内の画像（例えば、肝生検画像）を用いて再訓練されている（has been re-trained）。微調整済み正規化済みモデルは、微調整済みの対照モデルによって生成され且つ正規化された埋め込みに基づいて適合化された線形モデルを指す。

10

【0310】

他方で、教師ありモデル（例えば、Yr1）は、画像データを受信し及び医療診断スコアを予測するように構成された非線形機械学習モデル（例えば、ニューラルネットワーク）を指す。始めの4つのモデル等の線形回帰モデルは、訓練及び適用に関しては教師ありモデルに比してより演算的に効率的である。図14にて示されているように、埋め込みに基づいて生成された線形モデルは、教師あり機械学習モデルのそれと同様の又はそれより優れた予測力を提供できるのであり、その一方では訓練及び適用に関しては格段に少ないリソース及び時間要求で済む。

20

【0311】

図15Aは、研究、サイト、及び病理スコア効果を伴うバリエーションコンポーネントモデル1502について示す。示されているように、プロット1504によれば、バリエーションコンポーネントモデル1502によっては埋め込みの分散の34%しか説明できない。図15Bは、サイト及び研究影響について調整がなされる埋め込みについての全ゲノム的関連研究（GWAS、genome-wide association study）プロット1506について示す。示されているように、GWASプロット1506は、ベースライン埋め込みと関連付けられている3つのミスセンスバリエーションを識別する。バリエーションのサブセットについては、説明されているようにさらに分析できる。

30

【0312】

図15Bのプロット1508にて示されるように、組織学的分析を通じて優先順序が設けられた遺伝的バリエーションについては、利用可能なエンドポイントと関連付けて関連する生物学事項についての知見を得ることができる。例えば、PNPLA3 rs738409における主位バリエーションについてのPheWASによって、ベースライン血液バイオマーカー及びRNA-seqに対しての効果が識別されるも、病理又は連続NASHスコアについてはそうならない。図15Cでは、PNPLA3 rs738409における主位バリエーションについてのPheWASによっていくつかの血液バイオマーカー及び発現経路に対しての効果が識別されることを示すプロット1510が示されている。さらに、rs738409は組織学的疾患ラベルと関連付けられていないものとみられる。したがって、臨床試験データにおけるPheWASは、発見された遺伝的効果の解釈をサポートすることになる。

40

【0313】

図16、20、22、24は、縦型の研究（例えば、治療効果のそれ）に向けられた手法について示す。当業者ならば、本明細書にて説明される縦型の研究では、本明細書にて説明されたベースライン分析のための様々な手法（訓練済みのモデル及びシステムを含む）が活用されていることと理解されるであろう。例えば、入力として特定の被験者について縦型の進行埋め込みを受信し及びプラセボ対治療の決定を出力するシステムは（例えば、DRP分類モデル1730）、着目コバリエーション（例えば、遺伝的バリエーション）に基づいて疾患の状態又は進行について予測するように以前訓練されたシステムに基づいている

50

ことができる。

【0314】

図16は、いくつかの実施形態による、着目疾患との関係で治療を評価するための例示的な方法について示す。処理1600では、疾患の進行は、進行埋め込みを用いて定量化される。システムは、進行埋め込みを入力として受信しプラセボ又は治療を示す分類結果を出力するモデルからの予測として、薬剤応答表現型（DRP）として帰属させることができる。システムは、DRPと治療との間に有意な関連があるかについて決定できる。有意な関連がある場合、下流側の分析において治療をさらに分析できる（例えば、後述図20との関係でさらに説明されている）。

【0315】

処理1600は、例えば、ソフトウェアプラットフォームを実装する1つ以上の電子装置を使用して、実施される。いくつかの実施例では、処理1600は、クライアントサーバシステムを使用して実施され、処理1600のブロックは、任意の様式において、サーバと1つ以上のクライアントデバイスとの間で分割される。したがって、処理1600の部分がクライアントサーバシステムの特定の装置によって実行されると説明されているも、処理1600はそのように限定されることは要さないことに留意されたい。他の例では、処理1600はクライアント装置のみ又は複数のクライアント装置のみを用いて実行される。処理1600では、いくつかのブロックは、随意に、組み合わせられ、いくつかのブロックの順序は、随意に、変更され、いくつかのブロックは、随意に、省略される。いくつかの実施例では、付加的ステップが、処理1600と組み合わせて実施されてもよい。故に、図示（および下記により詳細に説明）されるような動作は、本質的に、例示的であって、したがって、限定として見なされるべきではない。

【0316】

ブロック1602では、例示的なシステム（例えば、1つ以上の電子装置）は、プラセボ群にプラセボが施される前に撮像された被験者プラセボ群についての複数のベースラインプラセボ画像と、プラセボ群にプラセボが施された後に撮像された被験者プラセボ群についての複数のフォローアッププラセボ画像とを取得するように構成されていることができる。図17は、いくつかの実施形態による、着目疾患との関係で治療を評価するための例示的な処理について示す。図示のように、システムは、プラセボ群にプラセボが施される前に撮像された被験者プラセボ群についての複数のベースラインプラセボ医用画像1702を取得する。システムはさらに、プラセボ群にプラセボが施された後に撮像された被験者プラセボ群についての複数のフォローアッププラセボ医用画像1704を取得する。

【0317】

ブロック1604では、システムは、複数のベースラインプラセボ画像及び複数のフォローアッププラセボ画像に基づいて複数のプラセボ進行埋め込みを取得するように構成されていることができる。図17を参照するに、システムは、複数のベースラインプラセボ埋め込み1712と複数のフォローアッププラセボ埋め込み1714とを取得できる。一部の実施形態では、ベースラインプラセボ埋め込み1712はベースラインプラセボ画像1702に基づいて取得でき、また、フォローアッププラセボ埋め込み1714はフォローアッププラセボ画像1704に基づいて取得できる。埋め込み1712、1714に基づいて、システムはプラセボ進行埋め込み1722を取得することができる。

【0318】

図18A～Bは、いくつかの実施形態による、例示的な進行埋め込み生成について示す。図18Aに示されているように、システムは、複数のベースラインプラセボ画像1804を訓練済み教師なし機械学習モデル（不図示）内へと入力して潜在空間内での複数のベースラインプラセボ埋め込み1812を取得できる。同様に、システムは、複数のフォローアッププラセボ画像1806を訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数のフォローアッププラセボ埋め込み1814を取得できる。一部の実施形態では、システムは、複数のベースライン治療画像1808を訓練済み教師なし機械学習モデル（不図示）内へと入力して潜在空間内での複数のベースライン治療埋め込み1816

10

20

30

40

50

を取得できる。さらに、一部の実施形態では、システムは、複数のフォローアップ治療画像 1810 を訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数のフォローアップ治療埋め込み 1818 を取得できる。一部の実施形態では、教師なし機械学習モデルは対照モデルであり、これは図 4A～B を参照して説明されたモデルと類似する。一部の実施形態では、前記対照モデルは SimCLR モデルである。一部の実施形態では、教師なし機械学習モデルは、図 8A～8B 及び関連図面との関係で説明されたものに沿う態様で設計及び訓練されたシステムである。

【0319】

そして、図 18B に示されているように、システムは、ベースラインプラセボ埋め込み 1812 を訓練済み機械学習モデル 1850 内へと入力して、潜在空間にて複数の予測フォローアッププラセボ埋め込み 1854 を取得できる。そして、システムは、フォローアッププラセボ埋め込み 1814 と予測フォローアッププラセボ埋め込み 1854 との間の差を算出することによって複数のプラセボ進行埋め込み 1856 を決定できる。一部の実施形態では、プラセボ群内の患者について、システムは、患者のフォローアッププラセボ埋め込みと患者の予測フォローアッププラセボ埋め込みとの間で減算を行って、患者のプラセボ進行埋め込みを取得する。

【0320】

図 18C は、いくつかの実施形態による、例示的な進行埋め込み生成について示す。一部の実施形態では、各予測フォローアップ埋め込みは、図 8A～8B 及び関連図面との関係で説明されたものに沿う態様で設計及び訓練されたシステムによって以前評価された着目コバリエーションに関して、患者がそれを有するかそれと関連付けられているかを考慮する。

【0321】

同様に、図 18D にて示されているように、システムは、複数のベースライン治療埋め込み 1816 を訓練済み機械学習モデル 1850 内へと入力して、潜在空間にて複数の予測フォローアップ治療埋め込み 1858 を取得できる。そして、システムは、フォローアップ治療埋め込み 1818 と予測フォローアップ治療埋め込み 1858 との間の差を算出することによって複数の治療進行埋め込み 1860 を決定できる。一部の実施形態では、治療群内の患者について、システムは、患者のフォローアップ治療埋め込みと患者の予測フォローアップ治療埋め込みとの間で減算を行って、患者の治療進行埋め込みを取得する。

【0322】

一部の実施形態では、訓練済み機械学習モデル 1850 は、ベースライン埋め込みを受信し及び予測フォローアップ埋め込みを出力するように構成されている。一部の実施形態では、訓練済み線形モデルは線形混合モデルであり、これは本明細書にて説明された他の線形混合モデルに類似している。図 19 は、いくつかの実施形態による、訓練済み機械学習モデル 1850 についての例示的な訓練処理について示す図である。示されているように、訓練済み機械学習モデル 1850 は、訓練データ 1960 を用いて訓練されるのであり、これはプラセボデータを含み得る。一部の実施形態では、訓練データ 1960 は、図 18A～B にて分析されているものとは異なるプラセボ群から取得される。

【0323】

ブロック 1606 では、システムは、治療群に治療が施される前に撮像された被験者治療群についての複数のベースライン治療画像と、治療群に治療が施された後に撮像された被験者治療群についての複数のフォローアップ治療画像とを取得できる。図 17 を参照するに、システムは、治療群に治療が施される前に撮像された被験者治療群についての複数のベースライン治療画像 1706 と、治療群に治療が施された後に撮像された被験者治療群についての複数のフォローアップ治療画像 1708 とを取得できる。

【0324】

ブロック 1608 では、システムは、複数のベースライン治療画像及び複数のフォローアップ治療画像に基づいて複数の治療進行埋め込みを取得できる。例えば、図 17 を参照

するに、システムは、複数のベースライン治療埋め込み 1716 と複数のフォローアップ治療埋め込み 1718 とを取得する。一部の実施形態では、ベースライン治療埋め込み 1716 はベースライン治療画像 1706 に基づいて取得され、また、フォローアップ治療埋め込み 1718 はフォローアップ治療画像 1708 に基づいて取得される。埋め込み 1716 及び 1718 に基づいて、システムは治療進行埋め込み 1726 を取得することができる。進行埋め込みの生成は、図 18A～C を参照して説明されている。

【0325】

ブロック 1610 では、システムは、複数の治療進行埋め込みに基づいて患者がプラセボ又は治療を受けたのかを決定するための分類モデルを生成できるのであって、分類モデルの出力は薬物反応組織学的表現型 (DRP) について示す。図 17 を参照するに、システムは、プラセボ進行埋め込み 1722 と治療進行埋め込み 1726 とに基づいて DRP 分類モデル 1730 を生成する。一部の実施形態では、分類モデル (例えば、DRP 分類モデル 1730) は、入力進行埋め込みを受信し、また、プラセボ又は治療を患者が受けたのかを示す分類結果を出力するように構成されている。分類モデルは、例えば、ロジスティック回帰モデル、人工ニューラルネットワークモデル、ランダムフォレストモデル、ナイーブなベイジアンモデル等として実装できる。

10

【0326】

ブロック 1612 では、システムは、分類モデルに基づいて治療と着目疾患との間の相関メトリック 1732 を決定するように構成されていることができる。相関メトリック 1732 は、治療と着目疾患の進行とが有意な程に関連付けられているかを示し得る。一部の実施形態では、相関メトリック 1732 は P 値である。

20

【0327】

一部の実施形態では、システムは相関メトリック 1732 を所定の閾値と比較するように構成されていることができる。一部の実施形態では、システムは、比較に基づいて、治療 A 1734 と関連付けられている特徴について関連分析をさらになすように構成されていることができる。

【0328】

一部の実施形態では、システムは、関連に基づいて新規被験者について治療 (例えば、画像 1706, 1708 と関連付けられている被験者についてなされた治療) を処方するように構成されていることができる。例えば、治療と着目疾患の進行とが有意な程に関連付けられている場合、同じ治療を、疾患を有する新規被験者に対して処方できる。別の例を挙げるに、治療と着目疾患の進行とが有意な程に関連付けられていない場合、同じ治療は、疾患を有する新規被験者に対して処方されないこととし得る。

30

【0329】

一部の実施形態では、システムは、関連に基づいて治療を施すように構成されていることができる。例えば、治療と着目疾患の進行とが有意な程に関連付けられている場合、同じ治療を、疾患を有する新規被験者に対して施すことができる。別の例を挙げるに、治療と着目疾患の進行とが有意な程に関連付けられていない場合、同じ治療は、疾患を有する新規被験者に対して施さないこととし得る。

【0330】

一部の実施形態では、システムは、関連に基づいて治療を調整するように構成されていることができる。例えば、治療と着目疾患の進行とが有意な程に関連付けられている場合、治療を増加させ得る。別の例を挙げるに、治療と着目疾患の進行とが有意な程に関連付けられていない場合、治療を減じるか停止できる。

40

【0331】

一部の実施形態では、システムは、関連に基づいて医学的提案を提供するように構成されていることができる。一部の実施形態では、システムは、関連に基づいてレポートを生成するように構成されていることができる。一部の実施形態では、治療と着目疾患の進行とが有意な程に関連付けられている場合、例えば図 21 にて説明されているように、システムでは治療についてさらに研究できる。

50

【0332】

一部の実施形態では、前記着目疾患は非アルコール性脂肪性肝炎（NASH、non-alcoholic steatohepatitis）である。

【0333】

例示的な実装例では、2つの臨床試験が行われている。より小規模な臨床試験では、ベースライン及びフォローアップ肝生検画像データ並びに画像データと関連付けられている病理学者によって割り当てられたスコアが取得される。特に、より小規模な臨床試験はNASH Ph2試験であり、サンプル数は約380であり、また、肝生検画像データはベースライン及び48週フォローアップにて取得される。より大規模な臨床試験では、アライン済みのベースライン及びフォローアップ肝生検画像データ並びに関連付けられている病理学スコアが取得される。特に、より大規模な臨床試験は2つのNASH Ph3試験であり、サンプル数は約1,600であり、また、肝生検画像データはベースライン及び48週フォローアップにて取得される。

10

【0334】

そして、システムは教師なし学習手順を介して全てのH&E染色肝生検画像（例えば、図17のベースラインプラセボ埋め込み1712、フォローアッププラセボ埋め込み1714、ベースライン治療埋め込み1716、及びフォローアップ治療埋め込み1718）から生検埋め込みを取得する。

【0335】

さらにシステムは、進行埋め込み（例えば、プラセボ進行埋め込み1722、治療進行埋め込み1726）を取得する。まず、システムは、表現型予測モデル（例えば、訓練済み機械学習モデル1850）を訓練して、ベースライン時の生検埋め込みからフォローアップ時（即ち、第48週）の生検埋め込みの全次元を予測するのであって、より大規模な臨床試験のプラセボ群内の患者のみを考慮してこれらがなされる。訓練済み機械学習モデル1850は、図3Aの埋め込みスコア予測モデル312（これも線形モデルとされ得る）と類似の態様で構成されることができる。一部の実施形態では、機械学習モデル1850は、線形混合モデル等の線形回帰モデルとされ得る。

20

【0336】

表現型予測モデルが訓練された後、システムは訓練済み表現型予測モデルを用いて、より小規模な臨床試験での患者のベースライン肝臓状態埋め込みを前提として、プラセボ下のフォローアップ肝臓状態埋め込みを予測できる。これらの予測フォローアップ埋め込みは、プラセボ患者のフォローアップ時の期待組織学的状態を捕捉する。システムは、進行埋め込みを、観測されたフォローアップ埋め込みと前のステップからの予測されたフォローアップ埋め込みとの間の差として定義する。定義されているように、進行埋め込みは、より小規模な臨床試験の患者の組織学的差異について表すのであり、それらのフォローアップ時の組織学的状態とプラセボ患者の期待組織学的状態との間が扱われる。

30

【0337】

各治療群について、システムは、治療特有モデル（例えば、DRP分類モデル1730）を訓練して、より小規模な臨床試験の進行埋め込みからのプラセボ及び治療群からの患者を分類する。このモデルからの予測は、フォローアップ時の薬剤応答組織学的表現型（DRP）の各患者についての尺度を提供する。一部の実施形態では、分類モデルの入力は、進行埋め込み（例えば、プラセボ進行埋め込み1856、1860）である。一部の実施形態では、分類モデルの出力は、患者が治療群又はプラセボ群のどちらに属するかを示す各患者についての二値指標である。DRP分類は、図3Aの埋め込みスコア予測モデル312と類似の態様で構成されることができる。

40

【0338】

予測DRPの統計的有意性は、例えば、本明細書にて説明される順列手順を用いて査定できる。例えば、システムは、モデルについて相関メトリックを算出できる。一部の実施形態では、前記相関メトリックはP値である。P値は順列手順を介して取得できるのであって、実データからの対数尤度レシオ（LLR、log likelihood ratio）統計が、埋め

50

込み行列内にて個体についての置換がなされる際に取得されたLLRと、比較される（nullモデルからのLLR）。一部の実施形態では、P値は、上述のように、実データLLRより大きな並べ替えLLRの画分として定義される。

【0339】

各遺伝的バリエーションについての相関メトリックは所定の閾値に対して比較されて、遺伝的バリエーションと埋め込みとの間に関連があるかを決定できる。1つの例示的な実装例では、システムは、0.05とされるBonferroni調整P値閾値を用いて、有意な治療組織学的効果を定義して治療が進行埋め込みに対して効果を及ぼすかを査定する。一部の実施形態では、有意なP値を有する治療のみがさらに研究されるのであり、例えば図20の処理2000を用いてこれらがなされ得る。

10

【0340】

一部の実施形態では、治療を分析する際には、治療の効果は、フォローアップ埋め込みを用いて直接的に分析できる。そのような場合では、ベースライン埋め込み／画像は、分析から除外され得る。

【0341】

図20では、治療についての薬物反応組織学的表現型（DRP）との関係で着目コバリエーションを識別する例示的な方法について示されている。DRPについての帰属決定は、進行埋め込みが利用可能である限り臨床試験データセットを用いることによって行うことができる。DRPと分子データ（例えば、表現及びジェネティクス）との間の有意な関連は関連検定を介して取り出すことができる。発現ありの関連付けでは、プラセボ対薬剤の差次的発現分析にて検出できなかった遺伝子を識別できる。遺伝的な関連付けによって着目疾患（例えば、NASH）についての候補標的遺伝子が識別される。

20

【0342】

処理2000は、例えば、ソフトウェアプラットフォームを実装する1つ以上の電子装置を使用して、実施される。いくつかの実施例では、処理2000は、クライアントサーバシステムを使用して実施され、処理2000のブロックは、任意の様式において、サーバと1つ以上のクライアントデバイスとの間で分割される。したがって、処理2000の部分がクライアントサーバシステムの特定の装置によって実行されると説明されているも、処理2000はそのように限定されることは要さないことに留意されたい。他の例では、処理2000はクライアント装置のみ又は複数のクライアント装置のみを用いて実行される。処理2000では、いくつかのブロックは、随意に、組み合わせられ、いくつかのブロックの順序は、随意に、変更され、いくつかのブロックは、随意に、省略される。いくつかの実施例では、付加的ステップが、処理2000と組み合わせて実施されてもよい。故に、図示（および下記により詳細に説明）されるような動作は、本質的に、例示的であって、したがって、限定として見なされるべきではない。

30

【0343】

ブロック2002では、例示的なシステム（例えば、1つ以上の電子装置）は、臨床被験者群から取得された共変量クラスのコバリエーション情報を受信するように構成されていることができる。一部の実施形態では、前記共変量クラスは人口統計情報、臨床共変量、又はゲノムデータを備える。ブロック2004では、システムは、臨床被験者群から複数のベースライン画像及び複数のフォローアップ画像を受信するように構成されていることができる。

40

【0344】

ブロック2002及びブロック2004で取得されたデータは、同じ治療群の被験者のデータとされることができる。治療群は、例えば、図16～19、23を参照して説明されるような態様で識別できる。図21を参照するに、システムは、被験者1について画像データ2102を、被験者2について画像データ2104を、...、及び被験者Nについて画像データ2106を受信する。被験者1-Nは同じ治療を受ける同じ治療群に属することができる。一部の実施形態では、治療群は識別された着目治療を受けており、例えば図16～19、23を参照して説明されている通りである（例えば、相関メトリック17

50

3 2 が所定の閾値を満たす場合には治療 A)。

【0345】

ブロック2006では、システムは、複数のベースライン画像及び複数のフォローアップ画像に基づいて複数の進行埋め込みを取得するように構成されていることができる。図21に示されているように、システムは、進行埋め込み2112, 2114, ..., 2116を取得できる。進行埋め込みの生成については、例えば図18A~Cを参照して説明されている。一部の実施形態では、前記複数のベースライン画像及び前記複数のフォローアップ画像に基づいて前記複数の進行埋め込みを取得するステップは次のステップを含む：前記複数のベースライン医用画像を訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数のベースライン埋め込みを取得するステップと、前記複数のフォローアップ医用画像を前記訓練済み教師なし機械学習モデル内へと入力して前記潜在空間内での複数のフォローアップ埋め込みを取得するステップと、前記複数のベースライン埋め込みを訓練済み線形モデル内へと入力して前記潜在空間内での複数の予測フォローアップ埋め込みを取得するステップと、前記複数のフォローアップ埋め込みと前記複数の予測フォローアップ埋め込みとの間の差を算出することによって前記複数の進行埋め込みを決定するステップ。

10

【0346】

一部の実施形態では、前記教師なし機械学習モデルは対照モデルである。

【0347】

一部の実施形態では、前記対照モデルはSimCLRモデルである。

【0348】

一部の実施形態では、前記訓練済み線形モデルはベースライン埋め込みを受信し及び予測フォローアップ埋め込みを出力するように構成されている。

20

【0349】

一部の実施形態では、前記訓練済み線形モデルは線形混合モデルである。

【0350】

ブロック2008では、システムは、複数の進行埋め込みを訓練済み分類モデル内へと入力して臨床被験者群のDRP値について示す複数の分類結果を取得するように構成されていることができる。図21に示されているように、各進行埋め込みは、DRP分類モデル2120に入力することができる。例えば、画像データ2102に基づいて生成された進行埋め込み2112は、DRP分類モデル2120に入力されて、被験者1についてのDRP値を取得でき；画像データ2104に基づいて生成された進行埋め込み2114は、DRP分類モデル2120に入力されて、被験者2についてのDRP値を取得でき；及び画像データ2106に基づいて生成された進行埋め込み2116は、DRP分類モデル2120に入力されて、被験者1-NについてのDRP値を取得できる。一部の実施形態では、前記訓練済み分類モデルは、入力進行埋め込みを受信し、また、患者がプラセボ又は前記治療を受けたかを決定するように構成されている。分類モデルは、図18B~Cの訓練済み機械学習モデル1850と同一又は類似のものとされ得る。したがって、システムは、薬剤応答表現型(DRP)を進行埋め込みから治療へのモデルからの予測として帰属させる(プラセボ対薬剤)。この帰属決定は、進行埋め込みが利用可能である限り他の臨床試験データセットにて行うことができる。

30

40

【0351】

ブロック2010では、システムは、臨床被験者群についてのコバリアント情報、複数の分類結果、及び1つ以上の線形回帰モデルに基づいて、複数の候補コバリアントの各候補コバリアントとDRP値(例えば、DRP値2122-1 - 2122-N)との間の関連を決定して、着目コバリアントを識別するように構成されていることができる。換言するに、システムはDRPと分子データとの間に有意な程の関連を識別できる(発現及びジェネティクス)。発現ありの関連付けでは、プラセボ対薬剤の差次的発現分析にて検出できなかった遺伝子を識別できる。一部の 경우에는、DRP分析では、真のプラセボ対薬剤(placebo-vs-drug)の差次的発現分析の症例対照として相関付けられている遺伝子セットを識別できる。一部の 경우에는、DRP分析では、より大きなコホートの分析故により大

50

きな遺伝子セットを識別できるのでありDRP相関の解釈を支援し得る。

【0352】

一部の実施形態では、着目コバリアントの識別は、DRP値と臨床被験者群のコバリアント情報とに基づいて複数の候補コバリアントの各候補コバリアントについてコバリアント特有モデル2130を生成することを伴い得る。一部の実施形態では、一部又は全部のコバリアント特有モデル2130は、説明されているのと類似の線形混合モデルとされることができる。システムは、全ての着目コバリアント（例えば、100,000種、100万種、1,000万種の候補）についてコバリアント特有モデル（例えば、100,000種、100万種、1,000万種のモデル）を生成できる。各モデルを評価して、各候補コバリアントとDRPとの間に有意な関連があるかについて決定して、1つ以上の着目コバリアントを識別できる。

10

【0353】

一部の実施形態では、システムは、モデルに基づいて相関メトリックを決定できる。相関メトリックは、候補コバリアント及びDRP値が有意な程に関連付けられているかを示す。一部の実施形態では、前記相関メトリックはP値である。一部の実施形態では、相関メトリックを所定の閾値に対して比較して候補コバリアントが着目コバリアントであるかを決定できる。

【0354】

一部の実施形態では、前記複数の候補コバリアントは複数の候補ミスセンスバリアントを備える。例示的な実装例では、臨床試験にてACCi+FXRa DRPの進行GWASが行われたのであり、27,270個のミスセンスバリアントについて着目した（MAF>1%）。分析ではミスセンス遺伝子座を識別できる。

20

【0355】

一部の実施形態では、前記複数の候補コバリアントは複数の候補遺伝子を備える。例示的な実装例では、様々な臨床試験においての発現とDRPとの間の関連によって、図27に示されているように何千もの関連遺伝子が識別された。

【0356】

一部の実施形態では、識別された着目コバリアントは、新規被験者に関して着目疾患を診断するために用いることができる。一部の実施形態では、識別された着目コバリアントに基づいて治療を開発できる。

30

【0357】

一部の実施形態では、識別された着目コバリアントに基づいて治療を施し、調整し、及び／又は適用できる。

【0358】

一部の実施形態では、識別された着目コバリアントに基づいて医学的提案を提供すること（又は求めること）ができる。

【0359】

一部の実施形態では、識別された着目コバリアントに基づいて生物学的標的を識別できる。

【0360】

一部の実施形態では、複数の医用画像は生検画像を含む。

40

【0361】

例示的な実装例では、2つの臨床試験が行われている。より小規模な臨床試験では、ベースライン及びフォローアップ肝生検画像データ並びに画像データと関連付けられている病理学者によって割り当てられたスコアが取得される。特に、より小規模な臨床試験はNASH Ph2治験であり、サンプル数は約380であり、また、肝生検画像データはベースライン及び48週フォローアップにて取得される。より大規模な臨床試験では、アライン済みのベースライン及びフォローアップ肝生検画像データ並びに関連付けられている病理学スコアが取得される。特に、より大規模な臨床試験は2つのNASH Ph3治験であり、サンプル数は約1,600であり、また、肝生検画像データはベースライン及び48週フォロ

50

ーアップにて取得される。

【0362】

上述のように、各治療群について、システムは、治療特有モデル（例えば、DRP分類モデル1730）を訓練して、より小規模な臨床試験の進行埋め込みからのプラセボ及び治療群からの患者を分類するように構成されていることができる。このモデルからの予測は、フォローアップ時の薬剤応答組織学的表現型（DRP）の各患者についての尺度を提供する。一部の実施形態では、分類モデルの入力は進行埋め込みである。分類モデルの出力は、各患者が治療群又はプラセボ群のどちらに属するかについての二値指標である。DRP分類は、図3Aの埋め込みスコア予測モデル312と類似の態様で構成されることができる。

10

【0363】

一部の実施形態では、システムは、訓練済みDRP分類モデルを用いて大規模臨床試験においてDRP値を予測するように構成されていることができる。そしてシステムは、より大規模な臨床試験を用いて、DRPを他のデータレイヤと併せて分析して具体的な問題に対応でき、以下の2つの例にて説明がある。

【0364】

1つの例では、システムは大規模臨床試験でのDRP及び遺伝子発現を分析して、治療と関連付けられている遺伝子及び経路を識別する。まず、システムは、図3Bのモデル316に関して説明されている線形モデル関連付け手順を用いて、DRP値と遺伝子発現との間の関連について検定を行う。例えば、システムは、遺伝子値を受信し及びDRP値を出力する遺伝子特有モデルについて適合化をなし得る。モデルについてP値を決定できる。そしてシステムは、既存のアプローチを用いて経路エンリッチメント分析を行う（例えば、前のステップにてDRPとの関連付けでのP値によってランキングされた遺伝子リスト及び外部ソース（例えば、遺伝子オントロジー）からの経路アノテーションを入力として受け付けるGSEA）。例示的な実装例では、この手順は、相関する関連統計を治療及びプラセボ群についての直接差次的発現分析にもたらす。また、治療ープラセボ差次的発現分析と関連付けられているものより遙かに大規模な遺伝子セットが大規模な研究のDRPと有意に関連付けられているものとして識別される。

20

【0365】

一部の実施形態では、遺伝子発現は帰属させることができる。例えば、訓練済み線形回帰モデル等の機械学習モデルは、生物学的サンプルについての画像埋め込みに対して適合化されることができる。機械学習モデルは、臨床試験からの患者組織サンプルについての画像埋め込みに基づいて、組織RNAシーケンス測定を予測するように訓練されていることができる。一部の実施形態では、別の臨床試験からの患者組織サンプルについての画像埋め込みに対して同一又は類似の機械学習モデル（例えば、同一又は類似の線形回帰モデル）が適用され得る。例えば、他の臨床試験の患者からの画像埋め込みは関連付けられているRNAシーケンスデータを有していない場合がある。予測RNAシーケンス測定と治療等の着目共変量との間の関連付けを査定して共変量の相関の解釈の一助とすることができる。

30

【0366】

別の例では、システムは、DRPについての遺伝子分析を行って、薬剤の影響を受けるのと同じ組織学的表現型に影響を及ぼす候補標的遺伝子を識別する（故に同様の経路に影響を及ぼすであろう）。特に、システムは、図3Bのモデル316に関して説明された線形モデル関連付け手順を用いてDRPとミスセンスバリエーションとの間の関連について検定を行うのであって、その際は年齢、性別、表現型PC、及び治療群を共変量として調整がなされている。例えば、バリエーション特有モデルは、ミスセンスバリエーション値を受信し及びDRP値を出力するように構成されている。この分析によって、DRPと関連付けられている遺伝子内のミスセンスバリエーションを識別することができる。一部の例では、複数の仮説検定は正をえた後に関連付けは有意となり得る。

40

【0367】

50

図 2 2 は、いくつかの実施形態による、着目疾患の進行との関係で治療を評価するための例示的な方法について示す。図 2 2 では、疾患進行は連続医療診断スコアによって定量化される。例示的な実装例では、着目共変量と進行スコアとの間の関連検定を行うことによって疾患進行について分析するのであり、その際は離散及び連続ベースライン疾患スコアの両方に加えて他の関連性を有する共変量について調整がなされる。連続スコアは、疾患進行についてより精密な定義付けを可能とし、例えば縦型表現分析（例えば、図 2 6 A）、遺伝関連研究（例えば、図 2 6 B）、及び治療応答についての関連研究に関して促進し得る。これらの例では、連続的疾患スコアの分析によって、離散的な病理学者によって割り当てられた疾患スコア（discrete pathologist-assigned disease score）の分析を通じて識別された関連が頻繁に反復され、また、病理スコアによっては識別されていない追加的な関連が識別されるということが見出される。

10

【0368】

この手順では、病理学者によって割り当てられた離散スコア（pathologist-assigned, discrete scores）を用いては検出できなかった、連続スコアに対しての薬剤の影響を取り出すことができる。処理 2 2 0 0 は、例えば、ソフトウェアプラットフォームを実装する 1 つ以上の電子装置を使用して、実施される。いくつかの実施例では、処理 2 2 0 0 は、クライアント-サーバシステムを使用して実施され、処理 2 2 0 0 のブロックは、任意の様式において、サーバと 1 つ以上のクライアントデバイスとの間で分割される。したがって、処理 2 2 0 0 の部分がクライアントサーバシステムの特定の装置によって実行されると説明されているも、処理 2 2 0 0 はそのように限定されることは要さないことに留意されたい。他の例では、処理 2 2 0 0 はクライアント装置のみ又は複数のクライアント装置のみを用いて実行される。処理 2 2 0 0 では、いくつかのブロックは、随意に、組み合わせられ、いくつかのブロックの順序は、随意に、変更され、いくつかのブロックは、随意に、省略される。いくつかの実施例では、付加的ステップが、処理 2 2 0 0 と組み合わせられて実施されてもよい。故に、図示（および下記により詳細に説明）されるような動作は、本質的に、例示的であって、したがって、限定として見なされるべきではない。

20

【0369】

ブロック 2 2 0 2 では、例示的なシステム（例えば、1 つ以上の電子装置）は医用画像を取得するように構成されていることができ、つぎのものが含まれる：（a）プラセボ群にプラセボが施される前に撮像された前記被験者プラセボ群についての複数のベースラインプラセボ画像と、（b）前記プラセボ群に前記プラセボが施された後に撮像された前記被験者プラセボ群についての複数のフォローアッププラセボ画像と、（c）治療群に前記治療が施される前に撮像された前記被験者治療群についての複数のベースライン治療画像と、（d）前記治療群に前記治療が施された後に撮像された前記被験者治療群についての複数のフォローアップ治療画像。図 2 3 を参照するに、システムは、ベースラインプラセボ画像 2 3 0 2 と、フォローアッププラセボ画像 2 3 0 4 と、ベースライン治療画像 2 3 0 6 と、フォローアップ治療画像 2 3 0 8 とを取得する。

30

【0370】

ブロック 2 2 0 4 では、システムは、医用画像を訓練済み教師なし機械学習モデル内へと入力して複数の埋め込みを取得するように構成されていることができ、各埋め込みは医用画像の 1 つ以上にて反映されている着目疾患との関係での表現型状態に対応する。一部の実施形態では、前記医用画像を訓練済み教師なし機械学習モデル内へと入力して前記複数の埋め込みを取得するステップは：（a）プラセボ群にプラセボが施される前に撮像された被験者プラセボ群についての複数のベースラインプラセボ画像を訓練済み教師なし機械学習モデル内へと入力して複数のベースラインプラセボ埋め込みを取得するステップと；（b）プラセボ群にプラセボが施された後に撮像された被験者プラセボ群についての複数のフォローアッププラセボ画像を訓練済み教師なし機械学習モデル内へと入力して複数のフォローアッププラセボ埋め込みを取得するステップと；（c）治療群に治療が施される前に撮像された被験者治療群についての複数のベースライン治療画像を訓練済み教師なし機械学習モデル内へと入力して複数のベースライン治療埋め込みを取得するステップと

40

50

；（d）治療群に治療が施された後に撮像された被験者治療群についての複数のフォローアップ治療画像を訓練済み教師なし機械学習モデル内へと入力して複数のフォローアップ治療埋め込みを取得するステップとを含む。

【0371】

図23を参照するに、システムは次のものを入力するように構成されていることができる：ベースラインプラセボ埋め込み2312を取得するためのベースラインプラセボ画像2302、フォローアッププラセボ埋め込み2314を取得するためのフォローアッププラセボ画像2304、ベースライン治療埋め込み2316を取得するためのベースライン治療画像2306、及びフォローアップ治療埋め込み2318を取得するためのフォローアップ治療画像2308。埋め込みを生成するために用いられる教師なし機械学習モデルは、図4A及び4Bを参照して説明されるモデルと類似していることができる。一部の実施形態では、前記教師なし機械学習モデルは対照モデルである。一部の実施形態では、前記対照モデルはSimCLRモデルである。

10

【0372】

ブロック2206では、システムは、複数の埋め込みを訓練済み線形回帰モデル内へと入力して複数の予測連続医療診断スコアを取得するように構成されていることができ、各予測連続医療診断スコアは着目疾患の状態について示し得る。一部の実施形態では、前記複数の埋め込みを前記訓練済み線形回帰モデル内へと入力するステップは：前記複数のベースラインプラセボ埋め込みを前記訓練済み線形モデル内へと入力して複数のベースラインプラセボスコア2322を取得することと、前記複数のフォローアッププラセボ埋め込みを前記訓練済み線形モデル内へと入力して複数のフォローアッププラセボスコア2324を取得することと、前記複数のベースライン治療埋め込みを前記訓練済み線形モデル内へと入力して複数のベースライン治療スコア2326を取得することと、前記複数のフォローアップ治療埋め込みを前記訓練済み線形モデル内へと入力して複数のフォローアップ治療スコア2328を取得することとを含む。

20

【0373】

一部の実施形態では、線形回帰モデルは線形混合モデルであり、これは図3Aを参照して説明されている埋め込みスコア予測モデル312に類似している。一部の実施形態では、前記訓練済み線形回帰モデルは複数の割り当て済み医療診断スコアに基づいて適合化されている。一部の実施形態では、前記複数の割り当て済み医療診断スコアは1人以上の医師によって提供される。一部の実施形態では、前記複数の割り当て済み医療診断スコアの各割り当て済み医療診断スコアは予め定義された値のセットから選択される。一部の実施形態では、前記複数の予測連続医療診断スコアは、複数の予測線維症スコア、複数の予測小葉内炎症スコア、又は複数の予測脂肪症スコアとされる。

30

【0374】

ブロック2208では、システムは、予測連続医療診断スコアに基づいて、複数のプラセボ進行スコア2332及び複数の治療進行スコア2334を決定する。一部の実施形態では、プラセボ進行スコア2332及び治療進行スコア2334を決定するステップは：ベースラインプラセボスコア2322とフォローアッププラセボスコア2324との間の差を決定してプラセボ進行スコア2332を決定することと、ベースライン治療スコア2326とフォローアップ治療スコア2328との間の差を決定して治療進行スコア2334を決定することとを含む。例えば、プラセボ群内の患者については、プラセボ進行スコアは、患者のベースラインプラセボスコアとフォローアッププラセボスコアとの間の差である。例えば、治療群内の患者については、治療進行スコアは、患者のベースライン治療スコアとフォローアップ治療スコアとの間の差である。

40

【0375】

一部の実施形態では、前記複数のプラセボ進行スコア及び前記複数の治療進行スコアを決定するステップは：前記プラセボ群内の各被験者について、前記プラセボ群内の前記被験者のベースラインプラセボスコア及びフォローアッププラセボスコアに少なくとも基づいて適合化された線形モデルの勾配を決定することと、前記治療群内の各被験者について

50

、前記治療群内の前記被験者のベースラインプラセボスコア及びフォローアッププラセボスコアに少なくとも基づいて適合化された線形モデルの勾配を決定することを含む。例えば、患者については、システムは、経時的に患者の医療診断スコア（ベースラインスコア及びフォローアップスコアを含む）を取得していくことができ、容量（又は治療時間）を受信し及び医療診断スコアを予測するように構成された線形モデルについて適合化をなし得る。患者についての進行スコアは、線形モデルの勾配とすることができる。

【0376】

ブロック2210では、システムは：複数のプラセボ進行スコア及び複数の治療進行スコアを治療と関連付けて；複数の疾患進行スコアと治療との間の相関メトリックを関連に基づいて決定するように構成されていることができる。一部の実施形態では、前記複数のプラセボ進行スコア及び前記複数の治療進行スコアを前記治療と関連付けるステップは、患者が前記治療を受けたかについての指示を受信し及び予測疾患進行スコアを出力するように構成されたモデルを生成するステップを含む。図23に示されているように、システムは、モデル2340を生成し及び相関メトリック2342を算出するように構成されていることができる。一部の実施形態では、モデルは本明細書にて説明されている線形混合モデルである。

【0377】

一部の実施形態では、前記相関メトリックはモデルのP値である。相関メトリックは、治療と疾患（disease）の進行との間に有意な関連があるかについて示す。

【0378】

一部の実施形態では、所与の治療についてのさらなる関連検定2344がなされ得る。例えば、相関メトリックは所定の閾値と比較され得る。一部の実施形態では、治療と着目疾患（disease of interest）との間の関連は、比較に基づいて識別され得る。一部の実施形態では、システムはさらに、関連に基づいて新規被験者について治療を処方するように構成されていることができる。例えば、治療と着目疾患の進行とが有意な程に関連付けられている場合、同じ治療を、疾患を有する新規被験者に対して処方できる。別の例を挙げるに、治療と着目疾患の進行とが有意な程に関連付けられていない場合、同じ治療は、疾患を有する新規被験者に対して処方されることはない。

【0379】

一部の実施形態では、システムは、関連に基づいて治療を施すようにさらに構成されていることができる。例えば、治療と着目疾患の進行とが有意な程に関連付けられている場合、同じ治療を、疾患を有する新規被験者に対して施すことができる。別の例を挙げるに、治療と着目疾患の進行とが有意な程に関連付けられていない場合、同じ治療は、疾患を有する新規被験者に対して施すことはない。

【0380】

一部の実施形態では、システムは、関連に基づいて治療を調整するようにさらに構成されていることができる。例えば、治療と着目疾患の進行とが有意な程に関連付けられている場合、治療を増加させ得る。別の例を挙げるに、治療と着目疾患の進行とが有意な程に関連付けられていない場合、治療を減じるか停止できる。

【0381】

一部の実施形態では、関連付けに基づいて医学的提案を提供することができる。一部の実施形態では、治療と着目疾患の進行とが有意な程に関連付けられている場合、例えば図21にて説明されているように、システムでは治療についてさらに研究できる。一部の実施形態では、前記着目疾患は非アルコール性脂肪性肝炎（NAFLD、non-alcoholic steatohepatitis）である。

【0382】

図24は、いくつかの実施形態による、着目患者サブグループを識別するための例示的な方法について示す。システムは、患者画像データから埋め込みを取得し及び埋め込みのクラスタを患者サブグループ（subgroup）として識別するように構成されていることができる。患者クラスタ（cluster）アイデンティティ及び疾患バイオマーカ、遺伝的バ

10

20

30

40

50

リアント及び発現レベル間の有意な関連が関連検定にて取り出される。この手順では、患者セグメント（segment）及び関連付けられている臨床ラベル及び分子ドライバが取り出されるのであり、それらは各患者セグメントを特徴付けることを支援する。

【0383】

処理2400は、例えば、ソフトウェアプラットフォームを実装する1つ以上の電子装置を使用して、実施される。いくつかの実施例では、処理2400は、クライアントサーバシステムを使用して実施され、処理2400のブロックは、任意の様式において、サーバと1つ以上のクライアントデバイスとの間で分割される。したがって、処理2400の部分がクライアントサーバシステムの特定の装置によって実行されると説明されているも、処理2400はそのように限定されることは要さないことに留意されたい。他の例では、処理2400はクライアント装置のみ又は複数のクライアント装置のみを用いて実行される。処理2400では、いくつかのブロックは、随意に、組み合わせられ、いくつかのブロックの順序は、随意に、変更され、いくつかのブロックは、随意に、省略される。いくつかの実施例では、付加的ステップが、処理2400と組み合わせて実施されてもよい。故に、図示（および下記により詳細に説明）されるような動作は、本質的に、例示的であって、したがって、限定として見なされるべきではない。

10

【0384】

ブロック2402では、システムは、臨床被験者群から取得された複数の医用画像を訓練済み教師なし機械学習モデル内へと入力して、潜在空間内での複数の埋め込みを取得するように構成されていることができる。ブロック2404では、システムは、複数の埋め込みをクラスタリングして1つ以上の埋め込みクラスタを生成するように構成されていることができる。ブロック2406では、システムは、1つ以上の埋め込みクラスタに対応する1つ以上の患者サブグループを識別するように構成されていることができる。

20

【0385】

ブロック2408では、システムは、1つ以上の患者サブグループの各患者サブグループをコバリエントと関連付けて前記着目患者サブグループを識別するように構成されていることができる。特に、システムは2つのタイプの分析をなし得る。第一に、システムは、関連検定を通じて患者サブグループと共変量（例えば、疾患バイオマーカー、遺伝的バリエーション及び発現レベル）との間に有意な関連があるかを決定することによって各患者サブグループについて特徴付けをなし得る。このようにして、システムは識別された患者サブグループと、関連付けられている臨床ラベル及び分子ドライバを取得でき、それらは患者サブグループを特徴付ける。一部の実施形態では、関連検定は次のことをなすモデルを生成することを伴う：患者が患者サブグループに属するかを示す入力を受信すること（例えば、患者が属さない場合には0とされ、患者がサブグループに属する場合には1とされる）、共変量値を出力すること。第二に、システムは、共変量（例えば、治療又は表現型）の効果を各患者サブグループ内にて特徴付けることができる。例えば、システムは、サブグループ内の患者のみを考慮しての分析（例えば、遺伝関連研究並びに治療と臨床的進行との間での関連付け）をなし得る。関連検定は、サブグループ内の患者のデータのみを用いてモデルを生成することを伴う。

30

【0386】

一部の実施形態では、前記教師なし機械学習モデルは対照モデルである。

40

【0387】

一部の実施形態では、前記対照モデルはSimCLRモデルである。

【0388】

一部の実施形態では、前記コバリエントは着目治療であり、また、前記着目患者サブグループは前記着目治療が相当なインパクトを及ぼすサブグループである。

【0389】

一部の実施形態では、前記1つ以上の患者サブグループの各患者サブグループを前記コバリエントと関連付けるステップは：患者サブグループについて、前記患者サブグループ内の患者が前記着目治療を受けたかについての指示を受信し及び予測疾患進行を出力する

50

ように構成されたモデルを、生成するステップと、前記モデルを評価して前記患者サブグループが前記着目患者サブグループであるかを決定するステップとを含む。

【0390】

一部の実施形態では、前記モデルを評価するステップは、前記モデルの相関メトリックを決定すること及び前記相関メトリックを所定の閾値に対して比較することを含む。

【0391】

一部の実施形態では、前記相関メトリックはP値である。

【0392】

一部の実施形態では、前記生成されたモデルは前記患者サブグループ内の被験者の疾患進行値によって訓練される。

【0393】

一部の実施形態では、前記疾患進行値は前記患者サブグループ内の前記被験者の医療診断スコアを含む。

【0394】

一部の実施形態では、前記疾患進行値は前記患者サブグループ内の前記被験者の進行スコアを含む。

【0395】

一部の実施形態では、前記疾患進行値は前記患者サブグループ内の前記被験者のDRP値を含む。

【0396】

一部の実施形態では、前記コバリエントは着目疾患の進行であり、また、前記着目患者サブグループは前記着目疾患の前記進行と有意な関連を有するサブグループである。

【0397】

一部の実施形態では、前記1つ以上の患者サブグループの各患者サブグループを前記コバリエントと関連付けるステップは：患者サブグループについて、患者が前記患者サブグループに属するかについての指示を受信し及び予測疾患進行を出力するように構成されたモデルを、生成するステップと、前記モデルを評価して前記患者サブグループが前記着目患者サブグループであるかを決定するステップとを含む。

【0398】

一部の実施形態では、前記モデルを評価するステップは、前記モデルの相関メトリックを決定すること及び前記相関メトリックを所定の閾値に対して比較することを含む。

【0399】

一部の実施形態では、前記相関メトリックはP値である。

【0400】

一部の実施形態では、前記生成されたモデルは前記臨床被験者群の疾患進行値によって訓練される。

【0401】

一部の実施形態では、前記疾患進行値は、前記患者サブグループ内の被験者の医療診断スコア、前記患者サブグループ内の被験者の進行スコア、又は前記患者サブグループ内の被験者のDRP値を含む。

【0402】

一部の実施形態では、前記コバリエントは有害副作用であり、また、前記着目患者サブグループは前記有害副作用と有意な関連を有するサブグループである。一部の実施形態では、前記1つ以上の患者サブグループの各患者サブグループを前記コバリエントと関連付けるステップは：患者サブグループについて、前記患者サブグループ内の患者が前記患者サブグループに属するかについての指示を受信し及び前記患者に前記有害副作用が生じるかを予測するように構成されたモデルを、生成するステップと、前記モデルを評価して前記患者サブグループが前記着目患者サブグループであるかを決定するステップとを含む。

【0403】

一部の実施形態では、前記コバリエントは有害副作用であり、また、前記着目患者サブ

10

20

30

40

50

グループは治療後に前記有害副作用を経験することについて有意な関連を有するサブグループである。一部の実施形態では、前記1つ以上の患者サブグループの各患者サブグループを前記コバリエントと関連付けるステップは：患者サブグループについて、前記患者サブグループ内の患者が前記治療を受けたかについての指示を受信し及び前記患者に前記有害副作用が生じるかを予測するように構成されたモデルを、生成するステップと、前記モデルを評価して前記患者サブグループが前記着目患者サブグループであるかを決定するステップとを含む。

【0404】

一部の実施形態では、前記モデルを評価するステップは、前記モデルの相関メトリックを決定すること及び前記相関メトリックを所定の閾値に対して比較することを含む。

10

【0405】

一部の実施形態では、前記相関メトリックはP値である。

【0406】

例示的な実装例では、システムは、臨床的に関連性を有する患者セグメントを識別して、効率を増大させASEを減少させるように構成されていることができる。まず、システムは、教師なし学習手順を介してH&E染色肝生検画像から生検埋め込み（一部又は全部）を取得することができる。そして、システムは、ベースライン生検埋め込みについての教師なし分析を行って、生検埋め込みに基づいて患者のクラスタを識別できる。各クラスタについて、システムは、患者規模での二値指標ベクトルを決定又は取得し、患者がクラスタに属するか否かを示すことができる。二値表現型は、下流側での分析に用い得ることができる。

20

【0407】

さらに、システムは、異なるクラスタの疾患進行との関連について査定するように構成されていることができ、例えば図3Bのモデル316にて説明されている線形モデル検定手順を用いる等してクラスタアイデンティティと臨床試験エンドポイントとの間での関連について検定することができる。例えば、システムはクラスタ特有モデルについて適合化を行うことができ、該モデルはクラスタ二値指標（クラスタに属する患者は1と符号化され、クラスタに属さない患者は0と符号化される）を受信し及び臨床スコア又はエンドポイント（例えば、疾患進行）を出力する。

【0408】

システムはまた、異なる患者クラスタ毎の治療効果についても査定できるのであり、例えば図3Bのモデル316にて説明されている線形モデル検定手順を用いる等して所定のクラスタ内の患者のみを考慮して治療と臨床エンドポイントとの間の関連について検定できる。例えば、システムは、治療対プラセボについての二値指標を受信し及び臨床エンドポイント（例えば、線維症進行）を出力するクラスタ特有モデルについて適合化を行うことができる。そして、分析を特定のクラスタ内の患者に限定することができる。

30

【0409】

システムはまた、異なるクラスタと関連付けられている有害副作用について査定することもできる（例えば、図3Bのモデル316にて説明されている線形モデル検定手順を用いる等してクラスタアイデンティティと有害副作用共変量との間での関連について検定することができる。）。例えば、システムはクラスタ特有モデルについて適合化を行うことができ、該モデルは二値指標（例えば、クラスタに属する患者は1と符号化され、クラスタに属さない患者は0と符号化される）を受信し及び副作用又は有害イベントを出力する。例えば、システムは分析を患者が特定の治療を受けている場合に限定でき、分析された出力は有害イベントによって臨床試験から脱落した患者であるかを（二値指標として）扱うことができる。これによって、治療からの有害副作用をより被りやすい患者を識別できることとなる。

40

【0410】

そして、特定のクラスタが進行と関連付けられている場合、システムはそのクラスタの遺伝子及び表現型バイオマーカーを識別できる（例えば、図3Bのモデル316にて説明さ

50

れている線形モデル検定手順を用いる等してクラスタアイデンティティとジェネティクスや発現やラボ値等との間での関連について検定することができる)。例えば、システムはモデルについて適合化を行うのであり、該モデルはクラスタ二値指標（例えば、クラスタに属する患者は1と符号化され、クラスタに属さない患者は0と符号化される）を受信し及び臨床スコア又はエンドポイント（例えば、疾患進行）を出力する。臨床エンドポイントは、進行スコア又は臨床試験内で監視される臨床エンドポイントとして定量化できる（例えば、患者に関して病理学評価に基づいた線維症スコアが高い又は低いこと）。

【0411】

一部の実施形態では、本明細書にて説明される手法は生検埋め込みではなく発現データに基づいていることができる。例えば、図3Bのモデル316に関して説明されている線形モデル検定手順を用いて、患者におけるベースライン発現レベルと疾患進行との間の関連について検定を行うことができる。例えば、線形混合モデルを生成して、入力としてベースライン発現レベルを受信し及び疾患進行予測を出力することができる。例を挙げるに、この手順では、進行と有意に関連付けられている130の遺伝子を識別できる。この例では、130の進行遺伝子の発現に基づいて患者についてのscanpyを用いてLeidenクラスタリングを行う（線維症ベースライン状態及び臨床試験指標を回帰で出した後）。

【0412】

システムは、異なるクラスタの疾患進行との関連について査定するのであり、図3Bのモデル316にて説明されている線形モデル検定手順を用いてクラスタアイデンティティと疾患進行との間での関連について検定することができる。例えば、線形混合モデルを生成して、入力としてクラスタアイデンティティ（例えば、患者がクラスタに属するか）を受信し及び疾患進行予測を出力することができる。例においては、分析では、進行と関連付けられているものとして2つのクラスタが、並びに、退行と関連付けられているものとして1つのクラスタが識別された。

【0413】

各遺伝子について、システムは、ベースライン状態及び臨床試験指標を入力として受信し及びその発現を出力する線形回帰モデルについて適合化を行うように構成されていることができる。システムは寄与分をベースライン状態及び臨床試験指標から減算し得るのであり、それは線形モデルを介して元の発現値から推定される。図25との関連で示されているように、分析例では3つの患者クラスタ（cluster）が示された。2つの進行クラスタは異なる遺伝子発現に対応する。換言するに、図25の例では、システムが、進行と関連付けられているベースライン発現レベルの遺伝子に基づいて患者をクラスタ化した後、3つの群（group）が認められる。これらのうち、2つの患者群は進行する傾向を示しているが、これらの群は異なる遺伝子シグネチャをベースラインにて有している。この不均一性は、これらの患者群間の基本的差異を示しているといえる。換言するに、これらの患者群は特定の治療に対して異なる態様で応答したり、遺伝的駆動因子が異なったり等し得る。

【0414】

システムは、クラスタアイデンティティと発現値との間での関連について検定するための図3Bのモデル316にて説明されている線形モデル検定手順を用いることによって、異なるクラスタと関連付けられている発現バイオマーカについてなされた探索について査定している。例えば、線形混合モデルを生成して、入力としてクラスタアイデンティティを受信し（例えば、患者がクラスタに属するか）並びに発現値を出力できる。この分析は、各クラスタと関連付けられている約10の発現バイオマーカを設定する。

【0415】

一部の実施形態では、画像ベースドバイオマーカを開発できる。例えば、予測画像を着目条件について可視化できる。例えば、着目条件は高い又は低い疾患スコアや、ある遺伝的シーケンスに対しての別の遺伝的シーケンス等を示し得る。関連付けられている画像特徴についての仮説を、予測画像の可視化に基づいて、生成できる。例えば、細胞内の細胞特徴は、着目条件のサンプル内においては、着目条件を有さないものとは違って、異なる

10

20

30

40

50

様相を呈するかもしれない。関連付けられている特徴を測定するように具体的に設計されているモデルを生成できるのであって、該特徴は画像ベースバイオマーカとされ得る。そして、モデルは新たなデータを伴って評価されることができ、新たな画像ベースバイオマーカとしてもちいられるかもしれない。

【0416】

図28は、いくつかの実施形態による、zスコアの比較について示す。一部の実施形態では、図28は、小規模臨床試験における治療対プラセボの分析のzスコアに対しての大規模な臨床試験における帰属DRP分析のそのの比較について示す。図28では、小規模サンプル内での真の治療効果についての分析と共に識別されるものと相関する遺伝子を識別できるDRP分析について示されている。さらに、図28では、先述の手法にはより良いパワーが付与されていることが示されている。さらに、このアプローチでは、治療分析が可能とされ得るのであり、着目共変量について測定がなされなかった研究に関して相関付けがなされる故にそうなる。

10

【0417】

図29は、一実施形態によるコンピューティング装置の一例について示す。装置2900は、ネットワークに接続されたホストコンピュータとすることができる。装置2900は、クライアントコンピュータ又はサーバとすることができる。図29に示されるように、装置2900は、パーソナルコンピュータ、ワークステーション、サーバ又は、電話若しくはタブレット等のハンドヘルド型コンピューティング装置（可搬電子装置）等の任意の適切な種類のマイクロプロセッサベースの装置とすることができる。装置2900は例えばプロセッサ2910、入力装置2920、出力装置2930、記憶装置2940、及び通信装置2960のうちの1つ以上を含むことができる。入力装置2920及び出力装置2930は概して上述のものに対応することができ、コンピュータに接続可能であるか又はコンピュータと一体化されるかのいずれかであってもよい。

20

【0418】

入力装置2920は、タッチスクリーン、キーボード若しくはキーパッド、マウス、又は音声認識装置等の、入力を提供する任意の適切な装置とすることができる。出力装置2930は、タッチスクリーン、触覚装置、又はスピーカ等の、出力を提供する任意の適切な装置とすることができる。

【0419】

記憶装置2940は、RAM、キャッシュメモリ、ハードドライブ、若しくはリムーバブル記憶ディスクを含む、電気、磁気、又は光学メモリ等の、記憶装置を提供する任意の適切な装置とすることができる。通信装置2960は、ネットワークインタフェースチップ又は装置等のネットワークを介して信号を送信及び受信することが可能な任意の適切な装置を含むことができる。コンピュータのコンポーネントは、物理的又は無線等を介して、任意の適切な手法で接続することができる。

30

【0420】

記憶装置2940に格納され、プロセッサ2910によって実行することができるソフトウェア2950は、例えば、本開示の機能を具現化する（例えば、上述したような装置において具現化される）プログラミングを含むことができる。

40

【0421】

ソフトウェア2950はまた、命令実行システム、機器、若しくは装置からソフトウェアに関連付けられた命令をフェッチし、その命令を実行することができる、上述したものなどの命令実行システム、機器、若しくは装置によって使用するための又はそれに関連して使用するための任意の非一時的コンピュータ可読記憶媒体内に格納及び／又は輸送することができる。本開示の文脈では、コンピュータ可読記憶媒体は、命令実行システム、機器、若しくは装置によって使用するための、若しくはそれに関連して使用するためのプログラムを含むか又は格納することができる、記憶装置2940などの任意の媒体とすることができる。

【0422】

50

ソフトウェア 2950 はまた、命令実行システム、機器、若しくは装置からソフトウェアに関連付けられた命令をフェッチし、その命令を実行することができる、上述したものなどの命令実行システム、機器、若しくは装置によって使用するための、又はそれに関連して使用するための、任意の輸送媒体内に伝播することができる。本開示の文脈では、輸送媒体は、命令実行システム、機器、若しくは装置によって使用するための、若しくはそれに関連して使用するためのプログラムを通信する、伝播する、又は輸送することができる任意の媒体とすることができる。輸送可読媒体は、電子、磁気、光学、電磁気、又は赤外線有線若しくは無線伝播媒体を含むことができるが、これらに限定されない。

【0423】

装置 2900 はネットワークに接続されていることができ、これは任意の適切な相互接続通信システムであることができる。ネットワークは、任意の適切な通信プロトコルを実装でき、また、任意の適切なセキュリティプロトコルによって保護できる。このネットワークは、無線ネットワーク接続、T1 若しくは T3 回線、ケーブルネットワーク、DSL、又は電話回線等の、ネットワーク信号の送信及び受信をなすことができる任意の適切な配置のネットワークリンクを備えることができる。

【0424】

装置 2900 は、ネットワーク上で動作するのに適した任意のオペレーティングシステムを実装できる。ソフトウェア 2950 は、C、C++、Java、又は Python 等の任意の適切なプログラミング言語で記述され得る。様々な実施形態において、本開示の機能を具現化するアプリケーションソフトウェアは、例えばクライアント／サーバ配置において又はウェブベースのアプリケーション若しくはウェブサービスとしてウェブブラウザを通して、異なる構成で展開できる。

【0425】

本開示及び実施例は添付の図面を参照して十分に説明されているも、様々な変更及び修正が当業者には明らかになることに留意されたい。そのような変更及び修正は、特許請求の範囲によって画定される開示及び実施例の範囲内に含まれているものと解されるべきである。

【0426】

前述の説明は、説明の目的で、特定の実施形態を参照して説明されている。もっとも、上記の例示的な議論は、網羅的であること、又は本発明を開示された正確な形態に限定することを意図するものではない。上記の教示に照らせば、多くの修正及び変形が可能である。実施形態は、技術の原理及びそれらの実際の応用を最も良く説明するために選択及び説明された。これにより、当業者であれば、想定される特定の用途に適した様々な修正を加えて、本技術及び様々な実施形態を最良に利用することができる。

【0427】

本手法は以下の列挙される実施形態を参照することによってより良く理解されよう：

【0428】

1. 表現型との関係で着目コバリエントを識別する方法であって：臨床被験者群から取得された共変量クラスについてのコバリエント情報及び表現型に関する対応する表現型データを受信するステップと；表現型データを訓練済み教師なし機械学習モデル内へと入力して潜在空間内で複数の埋め込みを取得するステップであって、各埋め込みは表現型データにて反映されている表現型状態に対応する、ステップと；(i) 臨床被験者群についてのコバリエント情報、(ii) 複数の埋め込み、及び(iii) 1つ以上の機械学習モデルに基づいて、複数の候補コバリエントの各々と表現型との間の関連を決定して、着目コバリエントを識別するステップを含む、方法。

【0429】

2. 実施例 1 に記載の方法において、前記 1つ以上の機械学習モデルは線形回帰モデルを備える、方法。

【0430】

3. 実施例 1～2 のいずれか 1つに記載の方法において、前記 1つ以上の機械学習モデル

は前記訓練済み教師なし機械学習モデルを含む、方法。

【0431】

4. 実施例1～3のいずれかに記載の方法において、前記表現型は着目疾患、遺伝子発現、メタボロミクス、プロテオミクス、又はリポミクスを含む、方法。

【0432】

5. 実施例1～4のいずれかに記載の方法において、前記表現型データは医用画像データ、生検データ、臨床バイオマーカーデータ、又はゲノミックバイオマーカーデータを含む、方法。

【0433】

6. 実施例1～5のいずれかに記載の方法において、前記共変量クラスは人口統計情報、臨床共変量、又はゲノムデータを備える、方法。

【0434】

7. 実施例1～6のいずれかに記載の方法において、各候補コバリアントと前記表現型との間の前記関連を決定するステップは：前記複数の埋め込みの各埋め込みを線形回帰モデル内へと入力して、複数の埋め込みの各埋め込みについて予測連続スコアを受信して、複数の予測連続スコアを取得するステップと、前記複数の予測連続スコアを、前記臨床被験者群によって発現されている候補コバリアントと関連付けるステップと、前記表現型と前記候補コバリアントとの間の相関メトリックを、前記関連に基づいて、決定するステップであって、前記相関メトリックは前記候補コバリアントが前記表現型に対して及ぼすインパクトについて示す、ステップとを含む、方法。

【0435】

8. 実施例1～6のいずれかに記載の方法において、前記候補コバリアントと前記表現型との間の前記関連を決定するステップは：前記複数の埋め込みを前記複数の候補コバリアントの各候補コバリアントと関連付けて、前記複数の候補コバリアントのサブセットを識別するステップと、前記サブセット内の各候補コバリアントを前記表現型と関連付けて、前記着目コバリアントを識別するステップとを含む、方法。

【0436】

9. 実施例1～8のいずれかに記載の方法において、さらに：前記表現型を表す複数の予測画像を、前記着目コバリアントに基づいて、生成するステップと、ディスプレイ上に前記複数の予測画像を表示するステップとを含む、方法。

【0437】

10. 実施例9に記載の方法において、さらに：前記複数の予測画像をランキングするステップを含む、方法。

【0438】

11. 実施例10に記載の方法において、前記複数の予測画像は前記ランキングに基づいて表示される、方法。

【0439】

12. 実施例1～11のいずれかに記載の方法において、さらに：前記コバリアントと前記表現型との間の関係を識別するステップを含む、方法。

【0440】

13. 実施例12に記載の方法において、前記関係は因果関係である、方法。

【0441】

14. 実施例12～13のいずれかに記載の方法において、さらに：前記関係に基づいて新規被験者について診断を提供するステップを含む、方法。

【0442】

15. 実施例12～14のいずれかに記載の方法において、さらに：前記関係に基づいて治療を開発するステップを含む、方法。

【0443】

16. 実施例12～15のいずれかに記載の方法において、さらに：前記関係に基づいて治療を施す、調整する、又は適用するステップを含む、方法。

【0444】

17. 実施例12～16のいずれかに記載の方法において、さらに：前記関係に基づいて医学的提案を提供するステップを含む、方法。

【0445】

18. 実施例12～17のいずれかに記載の方法において、さらに：前記関係に基づいて着目疾患の治療のための生物学的標的を識別するステップであって、前記表現型は前記着目疾患を含む、方法。

【0446】

19. 実施例18に記載の方法において、前記着目疾患は非アルコール性脂肪性肝炎（NASH、non-alcoholic steatohepatitis）である、方法。

10

【0447】

20. 着目疾患との関係で少なくとも1つの着目遺伝的バリエーションを識別する方法であって：臨床被験者群から取得された複数の医用画像を訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数の埋め込みを取得するステップであって、各埋め込みは前記複数の医用画像の1つ以上にて反映されている前記着目疾患との関係での表現型状態に対応する、ステップと、前記複数の埋め込みの各埋め込みを訓練済み機械学習モデル内へと入力して、前記複数の埋め込みの各埋め込みについて予測連続医療診断スコアを受信して、複数の予測医療診断スコアを取得するステップであって、各予測連続医療診断スコアは前記着目疾患の状態について示す、ステップと、前記複数の予測医療診断スコアを、複数の医用画像の取得元の臨床被験者群によって発現されている複数の候補遺伝的バリエーションの各候補遺伝的バリエーションと、関連付けるステップと、前記着目疾患と各候補遺伝的バリエーションとの間の相関メトリックを前記関連に基づいて決定し、前記少なくとも1つの着目遺伝的バリエーションを前記複数の候補遺伝的バリエーションから識別するステップであって、前記相関メトリックは各候補遺伝的バリエーションが前記着目疾患に対して及ぼすインパクトについて示すステップとを含む、方法。

20

【0448】

21. 実施例20に記載の方法において、さらに：前記相関メトリックを所定の閾値と比較するステップを含む、方法。

【0449】

22. 実施例21に記載の方法において、さらに：前記比較に基づいて、各着目候補遺伝的バリエーションと前記着目疾患との間の関係を識別するステップを含む、方法。

30

【0450】

23. 実施例22に記載の方法において、前記関係は因果関係である、方法。

【0451】

24. 実施例22～23のいずれかに記載の方法において、さらに：前記関係に基づいて新規被験者にて前記着目疾患を診断するステップを含む、方法。

【0452】

25. 実施例22～24のいずれかに記載の方法において、さらに：前記関係に基づいて治療を開発するステップを含む、方法。

【0453】

26. 実施例22～25のいずれかに記載の方法において、さらに：前記関係に基づいて治療を施す、調整する、又は適用するステップを含む、方法。

40

【0454】

27. 実施例22～26のいずれかに記載の方法において、さらに：前記関係に基づいて医学的提案を提供するステップを含む、方法。

【0455】

28. 実施例22～27のいずれかに記載の方法において、さらに：前記関係に基づいて前記着目疾患の治療のための生物学的標的を識別するステップを含む、方法。

【0456】

29. 実施例22～28のいずれかに記載の方法において、前記着目疾患は非アルコール

50

性脂肪性肝炎（NASH、non-alcoholic steatohepatitis）である、方法。

【0457】

30. 実施例20～29のいずれかに記載の方法において、前記複数の医用画像は生検画像を備える、方法。

【0458】

31. 実施例30に記載の方法において、前記生検画像は1つ以上の臨床試験に対応する、方法。

【0459】

32. 実施例20～31のいずれかに記載の方法において、さらに：前記複数の医用画像の医用画像を複数の画像タイルに分割するステップと、前記複数の画像タイルの各画像タイルを前記訓練済み教師なし機械学習モデル内へと入力して、各画像タイルについてタイル埋め込みを受信して、複数のタイル埋め込みを取得するステップと、前記複数のタイル埋め込みを集約して前記複数の埋め込みの埋め込みを取得するステップとを含む、方法。

10

【0460】

33. 実施例32に記載の方法において、前記複数のタイル埋め込みを集約するステップは前記複数のタイル埋め込みを平均化するステップを含む、方法。

【0461】

34. 実施例20～33のいずれかに記載の方法において、前記訓練済み教師なし機械学習モデルは対照モデルである、方法。

【0462】

20

35. 実施例34に記載の方法において、前記対照モデルはSimCLRモデルである、方法。

【0463】

36. 実施例20～34のいずれかに記載の方法において、前記訓練済み教師なし機械学習モデルは少なくとも部分的には前記複数の医用画像に基づいて訓練されている、方法。

【0464】

37. 実施例20～34のいずれかに記載の方法において、前記訓練済み教師なし機械学習モデルは前記複数の医用画像に基づいて微調整されている、方法。

【0465】

38. 実施例20～37のいずれか1つに記載の方法において、前記1つ以上の機械学習モデルは線形回帰モデルを備える、方法。

30

【0466】

39. 実施例38に記載の方法において、前記線形回帰モデルは訓練済み線形回帰モデルである、方法。

【0467】

40. 実施例39に記載の方法において、前記訓練済み線形回帰モデルは前記複数の埋め込み及び前記複数の埋め込みに対応する複数の割り当て済み医療診断スコアに基づいて適合化される、方法。

【0468】

41. 実施例20～37のいずれかに記載の方法において、前記1つ以上の機械学習モデルは線形混合モデルを備える、方法。

40

【0469】

42. 実施例20～41のいずれか1つに記載の方法において、前記1つ以上の機械学習モデルは前記訓練済み教師なし機械学習モデルを含む、方法。

【0470】

43. 実施例42に記載の方法において、前記複数の割り当て済み医療診断スコアは1人以上の医師によって提供される、方法。

【0471】

44. 実施例43に記載の方法において、前記複数の割り当て済み医療診断スコアの各割り当て済み医療診断スコアは予め定義された値のセットから選択される、方法。

【0472】

50

45. 実施例 20～44 のいずれかに記載の方法において、前記複数の予測連続医療診断スコアは、複数の予測線維症スコア、複数の予測小葉内炎症スコア、又は複数の予測脂肪症スコアとされる、方法。

【0473】

46. 実施例 20～45 のいずれかに記載の方法において、前記複数の予測医療診断スコアは、臨床試験中に取得された別個の測定を反映する予測医療診断スコア間の差として算出された疾患進行スコアを含む、方法。

【0474】

47. 実施例 20～45 のいずれかに記載の方法において、前記複数の予測医療診断スコアは、臨床試験中に各個人について取得された別個の測定を反映する予測医療診断スコアについて訓練された線形モデルによって決定された勾配として取得された疾患進行スコアを含む、方法。

【0475】

48. 実施例 18～47 のいずれか 1 つに記載の方法において、前記複数の予測医療診断スコアは、予測フォローアップスコアと観測フォローアップスコアとの間の差として対応するベースラインスコアについて調整がなされて算出された疾患進行スコアを含む、方法。

【0476】

49. 実施例 20～48 のいずれかに記載の方法において、前記複数の予測医療診断スコアを各候補遺伝的バリエーションと関連付けるステップは、前記候補遺伝的バリエーションについて示す値を受信し及び予測医療診断スコアを出力するように構成されたバリエーション特有モデルについて適合化をなすステップを含む、方法。

【0477】

50. 実施例 49 に記載の方法において、前記バリエーション特有モデルは線形モデルである、方法。

【0478】

51. 実施例 49 に記載の方法において、前記バリエーション特有モデルは、前記複数の予測医療診断スコアと前記候補遺伝的バリエーションについて示す複数の値とに基づいて適合化される、方法。

【0479】

52. 実施例 20～50 のいずれかに記載の方法において、前記相関メトリックを決定するステップはバリエーション特有モデルに基づいて P 値を決定するステップを含む、方法。

【0480】

53. 着目疾患との関係で少なくとも 1 つの着目遺伝的バリエーションを識別する方法であって：臨床被験者群から取得された複数の医用画像を訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数の埋め込みを取得するステップであって、各埋め込みは前記複数の医用画像の 1 つ以上に反映されている前記着目疾患との関係での表現型状態に対応する、ステップと、前記複数の埋め込みを複数の候補遺伝的バリエーションの各候補遺伝的バリエーションと関連付けて、前記複数の候補遺伝的バリエーションについてのサブセットを識別するステップであって、前記複数の候補遺伝的バリエーションのサブセットは前記複数の医用画像にて反映されている組織学的特徴と関連付けられている、ステップと、前記複数の候補遺伝的バリエーションの前記サブセットの各候補遺伝的バリエーションを前記着目疾患と関連付けて少なくとも 1 つの着目遺伝的バリエーションを前記サブセットから識別するステップとを含む、方法。

【0481】

54. 実施例 53 に記載の方法において、さらに：前記少なくとも 1 つの着目遺伝的バリエーションに基づいて、前記着目疾患を表す複数のシミュレーション画像を生成するステップと、ディスプレイ上に複数のシミュレーション画像を表示するステップとを含む、方法。

【0482】

55. 実施例 54 に記載の方法において、さらに：前記複数のシミュレーション画像をラ

10

20

30

40

50

ンキングするステップを含む、方法。

【0483】

56. 実施例55に記載の方法において、前記複数のシミュレーション画像は前記ランキングに基づいて表示される、方法。

【0484】

57. 実施例53～56のいずれかに記載の方法において、さらに：前記少なくとも1つの着目遺伝的バリエーションと前記着目疾患との間の関係を識別するステップを含む、方法。

【0485】

58. 実施例57に記載の方法において、前記関係は因果関係である、方法。

【0486】

59. 実施例57～58のいずれかに記載の方法において、さらに：前記関係に基づいて新規被験者にて前記着目疾患を診断するステップを含む、方法。

【0487】

60. 実施例57～59のいずれかに記載の方法において、さらに：前記関係に基づいて治療を開発するステップを含む、方法。

【0488】

61. 実施例57～60のいずれかに記載の方法において、さらに：前記関係に基づいて治療を施す、調整する、又は適用するステップを含む、方法。

【0489】

62. 実施例57～61のいずれかに記載の方法において、さらに：前記関係に基づいて医学的提案を提供するステップを含む、方法。

【0490】

63. 実施例57～62のいずれかに記載の方法において、さらに：前記関係に基づいて前記着目疾患の治療のための生物学的標的を識別するステップを含む、方法。

【0491】

64. 実施例53～63のいずれかに記載の方法において、前記着目疾患は非アルコール性脂肪性肝炎（NASH、non-alcoholic steatohepatitis）である、方法。

【0492】

65. 実施例53～64のいずれかに記載の方法において、前記複数の医用画像は生検画像を備える、方法。

【0493】

66. 実施例65に記載の方法において、前記生検画像は1つ以上の臨床試験に対応する、方法。

【0494】

67. 実施例53～66のいずれかに記載の方法において、さらに：前記複数の医用画像の医用画像を複数の画像タイルに分割するステップと、前記複数の画像タイルの各画像タイルを前記訓練済み教師なし機械学習モデル内へと入力して、各画像タイルについてタイル埋め込みを受信して、複数のタイル埋め込みを取得するステップと、前記複数のタイル埋め込みを集約して前記複数の埋め込みの埋め込みを取得するステップとを含む、方法。

【0495】

68. 実施例67に記載の方法において、前記複数のタイル埋め込みを集約するステップは前記複数のタイル埋め込みを平均化するステップを含む、方法。

【0496】

69. 実施例53～68のいずれかに記載の方法において、前記訓練済み教師なし機械学習モデルは対照モデルである、方法。

【0497】

70. 実施例69に記載の方法において、前記対照モデルはSimCLRモデルである、方法。

【0498】

71. 実施例53～70のいずれかに記載の方法において、前記訓練済み教師なし機械学

10

20

30

40

50

習モデルは少なくとも部分的には前記複数の医用画像に基づいて訓練されている、方法。

【0499】

72. 実施例53～70のいずれかに記載の方法において、前記訓練済み教師なし機械学習モデルは前記複数の医用画像に基づいて微調整されている、方法。

【0500】

73. 実施例53～72のいずれかに記載の方法において、前記複数の埋め込みを前記複数の候補遺伝的バリエーションの各遺伝的バリエーションと関連付けて前記複数の候補遺伝的バリエーションについての前記サブセットを識別するステップは：埋め込みを受信し及び前記候補遺伝的バリエーションの値を出力するように構成されたバリエーション特有モデルを、前記複数の候補遺伝的バリエーションの候補遺伝的バリエーションについて、生成するステップと、前記バリエーション特有モデルを評価して前記候補遺伝的バリエーションを前記サブセット内に含めるべきかを決定するステップとを含む、方法。

10

【0501】

74. 実施例73に記載の方法において、前記バリエーション特有モデルを評価するステップは：前記バリエーション特有モデルに基づいて相関メトリックを算出するステップと、前記相関メトリックを所定の閾値と比較するステップを含む、方法。

【0502】

75. 実施例74に記載の方法において、前記相関メトリックは前記バリエーション特有モデルと関連付けられているP値である、方法。

【0503】

20

76. 実施例53～75のいずれかに記載の方法において、前記複数の候補遺伝的バリエーションの前記サブセットの各遺伝的バリエーションを前記着目疾患と関連付けて前記少なくとも1つの着目遺伝的バリエーションを識別するステップは：前記遺伝的バリエーションについて示す値を受信し及び前記着目疾患に関する医療診断スコアを出力するように構成されたバリエーション特有モデルを、前記サブセット内の遺伝的バリエーションについて、生成するステップと、前記バリエーション特有モデルを評価して前記候補遺伝的バリエーションが前記少なくとも1つの着目遺伝的バリエーションであるかを決定するステップとを含む、方法。

【0504】

77. 実施例76に記載の方法において、前記バリエーション特有モデルを評価するステップは：前記バリエーション特有モデルに基づいて相関メトリックを算出するステップと、前記相関メトリックを所定の閾値と比較するステップを含む、方法。

30

【0505】

78. 実施例77に記載の方法において、前記相関メトリックは前記バリエーション特有モデルと関連付けられているP値である、方法。

【0506】

79. 着目疾患の進行に関して治療を評価する方法であって：被験者プラセボ群にプラセボが施される前に撮像された前記被験者プラセボ群についての複数のベースラインプラセボ医用画像と、前記被験者プラセボ群に前記プラセボが施された後に撮像された前記被験者プラセボ群についての複数のフォローアッププラセボ医用画像とを取得するステップと、前記複数のベースラインプラセボ医用画像及び前記複数のフォローアッププラセボ医用画像に基づいて複数のプラセボ進行埋め込みを取得するステップと、被験者治療群に前記治療が施される前に撮像された前記被験者治療群についての複数のベースライン治療医用画像と、前記被験者治療群に前記治療が施された後に撮像された前記被験者治療群についての複数のフォローアップ治療医用画像とを取得するステップと、前記複数のベースライン治療医用画像及び前記複数のフォローアップ治療医用画像に基づいて複数の治療進行埋め込みを取得するステップと、前記複数の治療進行埋め込みに基づいて患者が前記プラセボ又は前記治療を受けたのかを決定するための分類モデルを生成するステップとを含む、方法。

40

【0507】

80. 実施例79に記載の方法において、前記分類モデルの出力は薬物反応表現型につい

50

て示す、方法。

【0508】

81. 実施例80に記載の方法において、さらに：前記分類モデルに基づいて、前記治療と前記着目疾患の前記進行との間の相関メトリックを決定するステップを含む、方法。

【0509】

82. 実施例79～81のいずれか1つに記載の方法において、前記相関メトリックはP値である、方法。

【0510】

83. 実施例79～82のいずれか1つに記載の方法において、さらに：前記相関メトリックを所定の閾値と比較するステップを含む、方法。

【0511】

84. 実施例83に記載の方法において、さらに：前記比較に基づいて、前記治療と前記着目疾患の進行との間の関連を識別するステップを含む、方法。

【0512】

85. 実施例84に記載の方法において、さらに：前記関連に基づいて前記治療を新規被験者について処方するステップを含む、方法。

【0513】

86. 実施例84に記載の方法において、さらに：前記関連に基づいて前記治療を施すステップを含む、方法。

【0514】

87. 実施例84に記載の方法において、さらに：前記関連に基づいて前記治療を調整するステップを含む、方法。

【0515】

88. 実施例84に記載の方法において、さらに：前記関連に基づいて医学的提案を提供するステップを含む、方法。

【0516】

89. 実施例84に記載の方法において、さらに：前記関連に基づいてレポートを生成するステップを含む、方法。

【0517】

90. 実施例79～89のいずれかに記載の方法において、前記着目疾患は非アルコール性脂肪性肝炎（NAFLD、non-alcoholic steatohepatitis）である、方法。

【0518】

91. 実施例79～90のいずれかに記載の方法において、前記複数のプラセボ進行埋め込みを取得するステップは：前記複数のベースラインプラセボ医用画像を訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数のベースラインプラセボ埋め込みを取得するステップと、前記複数のフォローアッププラセボ医用画像を前記訓練済み教師なし機械学習モデル内へと入力して前記潜在空間内での複数のフォローアッププラセボ埋め込みを取得するステップと、前記複数のベースラインプラセボ埋め込みを1つ以上の機械学習モデル内へと入力して前記潜在空間内での複数の予測フォローアッププラセボ埋め込みを取得するステップと、前記複数のフォローアッププラセボ埋め込みと前記複数の予測フォローアッププラセボ埋め込みとの間の差を算出することによって前記複数のプラセボ進行埋め込みを決定するステップとを含む、方法。

【0519】

92. 実施例91に記載の方法において、前記1つ以上の機械学習モデルは訓練済み線形モデルを備える、方法。

【0520】

93. 実施例91～92のいずれかに記載の方法において、前記1つ以上の機械学習モデルは前記訓練済み教師なし機械学習モデルを含む、方法。

【0521】

94. 実施例91～93のいずれかに記載の方法において、前記複数の治療進行埋め込み

10

20

30

40

50

を取得するステップは：前記複数のベースライン治療医用画像を前記訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数のベースライン治療埋め込みを取得するステップと、前記複数のフォローアップ治療医用画像を前記訓練済み教師なし機械学習モデル内へと入力して前記潜在空間内での複数のフォローアップ治療埋め込みを取得するステップと、前記複数のベースライン治療埋め込みを前記訓練済み線形モデル内へと入力して前記潜在空間内での複数の予測フォローアップ治療埋め込みを取得するステップと、前記複数のフォローアップ治療埋め込みと前記複数の予測フォローアップ治療埋め込みとの間の差を算出することによって前記複数の治療進行埋め込みを決定するステップとを含む、方法。

【0522】

10

95. 実施例91～94のいずれかに記載の方法において、前記訓練済み教師なし機械学習モデルは対照モデルである、方法。

【0523】

96. 実施例95に記載の方法において、前記対照モデルはSimCLRモデルである、方法。

【0524】

97. 実施例91～96のいずれかに記載の方法において、前記訓練済み線形モデルはベースライン埋め込みを受信し及び予測フォローアップ埋め込みを出力するように構成されている、方法。

【0525】

98. 実施例97に記載の方法において、前記訓練済み線形モデルは線形混合モデルである、方法。

20

【0526】

99. 実施例97に記載の方法において、前記被験者プラセボ群は第1のプラセボ群であり、前記訓練済み線形モデルは前記第1のプラセボ群とは異なる第2のプラセボ群からの医用画像データから取得された埋め込みを用いて訓練されている、方法。

【0527】

100. 実施例79～99のいずれかに記載の方法において、前記分類モデルは、入力進行埋め込みを受信し、また、前記プラセボ又は前記治療を患者が受けたのかを示す分類結果を出力するように構成されている、方法。

【0528】

30

101. 実施例79～100のいずれかに記載の方法において、前記複数のベースラインプラセボ医用画像、前記複数のフォローアッププラセボ医用画像、前記複数のベースライン治療医用画像、及び前記複数のフォローアップ治療医用画像は、生検画像である、方法。

【0529】

102. 治療についての薬物反応表現型(DRP)との関係で着目コバリアントを識別する方法であって：臨床被験者群から取得された共変量クラスについてのコバリアント情報を受信するステップと、前記臨床被験者群から複数のベースライン画像及び複数のフォローアップ画像を受信するステップと、前記複数のベースライン画像及び前記複数のフォローアップ画像に基づいて複数の進行埋め込みを取得するステップと、前記複数の進行埋め込みを訓練済み分類モデル内へと入力して前記臨床被験者群のDRP値について示す複数の分類結果を取得するステップと、前記臨床被験者群についての前記コバリアント情報、前記複数の分類結果、及び1つ以上の機械学習モデルに基づいて、複数の候補コバリアントの各候補コバリアントと前記DRP値との間の関連を決定して、前記着目コバリアントを識別するステップとを含む、方法。

40

【0530】

103. 実施例102に記載の方法において、前記1つ以上の機械学習モデルは1つ以上の線形回帰モデルを備える、方法。

【0531】

104. 実施例102～103のいずれか1つに記載の方法において、前記複数の候補コバリアントは複数の候補ミスセンスバリエーションを備える、方法。

50

【0532】

105. 実施例102～103のいずれかに記載の方法において、前記複数の候補コバリエントは複数の候補遺伝子を備える、方法。

【0533】

106. 実施例102～105のいずれかに記載の方法において、前記共変量クラスは人口統計情報、臨床共変量、又はゲノムデータを備える、方法。

【0534】

107. 実施例102～106のいずれかに記載の方法において、さらに：前記識別された着目コバリエントに基づいて新規被験者にて着目疾患を診断するステップを含む、方法。

【0535】

108. 実施例102～107のいずれかに記載の方法において、さらに：前記識別された着目コバリエントに基づいて治療を開発するステップを含む、方法。

【0536】

109. 実施例102～108のいずれかに記載の方法において、さらに：前記識別された着目コバリエントに基づいて前記治療を施す、調整する、又は適用するステップを含む、方法。

【0537】

110. 実施例102～109のいずれかに記載の方法において、さらに：前記識別された着目コバリエントに基づいて医学的提案を提供するステップを含む、方法。

【0538】

111. 実施例102～110のいずれかに記載の方法において、さらに：前記識別された着目コバリエントに基づいて生物学的標的を識別するステップを含む、方法。

【0539】

112. 実施例102～111のいずれかに記載の方法において、前記複数のベースライン医用画像及び前記複数のフォローアップ医用画像は生検画像を備える、方法。

【0540】

113. 実施例102～112のいずれかに記載の方法において、前記複数のベースライン医用画像及び前記複数のフォローアップ医用画像に基づいて前記複数の進行埋め込みを取得するステップは：前記複数のベースライン医用画像を訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数のベースライン埋め込みを取得するステップと、前記複数のフォローアップ医用画像を前記訓練済み教師なし機械学習モデル内へと入力して前記潜在空間内での複数のフォローアップ埋め込みを取得するステップと、前記複数のベースライン埋め込みを訓練済み線形モデル内へと入力して前記潜在空間内での複数の予測フォローアップ埋め込みを取得するステップと、前記複数のフォローアップ埋め込みと前記複数の予測フォローアップ埋め込みとの間の差を算出することによって前記複数の進行埋め込みを決定するステップとを含む、方法。

【0541】

114. 実施例113に記載の方法において、前記訓練済み教師なし機械学習モデルは対照モデルである、方法。

【0542】

115. 実施例114に記載の方法において、前記対照モデルはSimCLRモデルである、方法。

【0543】

116. 実施例113に記載の方法において、前記訓練済み線形モデルはベースライン埋め込みを受信し及び予測フォローアップ埋め込みを出力するように構成されている、方法。

【0544】

117. 実施例113に記載の方法において、前記訓練済み線形モデルは線形混合モデルである、方法。

【0545】

118. 実施例102～117のいずれかに記載の方法において、前記訓練済み分類モデル

10

20

30

40

50

は、入力進行埋め込みを受信し、また、患者がプラセボ又は前記治療を受けたかを決定するように構成されている、方法。

【0546】

119. 実施例102～118のいずれかに記載の方法において、前記着目コバリエントを識別するステップは：前記複数の候補コバリエントの候補コバリエントについて：前記臨床被験者群のDRP値及び前記コバリエント情報に基づいたモデルを生成するステップと、前記モデルに基づいて相関メトリックを決定するステップとを含む、方法。

【0547】

120. 実施例119に記載の方法において、前記相関メトリックはP値である、方法。

【0548】

121. 実施例119に記載の方法において、さらに：前記相関メトリックを所定の閾値に対して比較して前記候補コバリエントが前記着目コバリエントであるかを決定するステップをさらに含む、方法。

【0549】

122. 着目疾患の進行に関して治療を評価する方法であって：医用画像を取得するステップであって、該医用画像は：(a)被験者プラセボ群にプラセボが施される前に撮像された前記被験者プラセボ群についての複数のベースラインプラセボ医用画像と、(b)前記被験者プラセボ群に前記プラセボが施された後に撮像された前記被験者プラセボ群についての複数のフォローアッププラセボ医用画像と、(c)被験者治療群に前記治療が施される前に撮像された前記被験者治療群についての複数のベースライン治療医用画像と、(d)前記被験者治療群に前記治療が施された後に撮像された前記被験者治療群についての複数のフォローアップ治療医用画像とを備える、ステップと、前記医用画像を訓練済み教師なし機械学習モデル内へと入力して複数の埋め込みを取得するステップであって、各埋め込みは前記医用画像の1つ以上にて反映されている前記着目疾患との関係での表現型状態に対応する、ステップと、前記複数の埋め込みを1つ以上の機械学習モデル内へと入力して複数の予測連続医療診断スコアを取得するステップであって、各予測連続医療診断スコアは前記着目疾患の状態について示す、ステップと、前記複数の予測連続医療診断スコアに基づいて、複数のプラセボ進行スコア及び複数の治療進行スコアを決定するステップと、前記複数のプラセボ進行スコア及び前記複数の治療進行スコアを前記治療と関連付けるステップと、前記関連に基づいて、前記複数のプラセボ進行スコアと前記複数の治療進行スコアとの間の相関メトリックを決定するステップとを含む、方法。

【0550】

123. 実施例122に記載の方法において、前記医用画像を訓練済み教師なし機械学習モデル内へと入力して前記複数の埋め込みを取得するステップは：前記(a)を前記訓練済み教師なし機械学習モデル内へと入力して複数のベースラインプラセボ埋め込みを取得することと、前記(b)を前記訓練済み教師なし機械学習モデル内へと入力して複数のフォローアッププラセボ埋め込みを取得することと、前記(c)を訓練済み教師なし機械学習モデル内へと入力して複数のベースライン治療埋め込みを取得することと、前記(d)を前記訓練済み教師なし機械学習モデル内へと入力して複数のフォローアップ治療埋め込みを取得することとを含む、方法。

【0551】

124. 実施例122～123のいずれかに記載の方法において、前記1つ以上の機械学習モデルは訓練済み線形回帰モデルを備える、方法。

【0552】

125. 実施例122～124のいずれか1つに記載の方法において、前記1つ以上の機械学習モデルは前記訓練済み教師なし機械学習モデルを含む、方法。

【0553】

126. 実施例124～125のいずれか1つに記載の方法において、前記複数の埋め込みを前記1つ以上の機械学習モデル内へと入力するステップは：前記複数のベースラインプラセボ埋め込みを前記訓練済み線形回帰モデル内へと入力して複数のベースラインプラセ

10

20

30

40

50

ボスコアを取得することと、前記複数のフォローアッププラセボ埋め込みを前記訓練済み線形回帰モデル内へと入力して複数のフォローアッププラセボスコアを取得することと、前記複数のベースライン治療埋め込みを前記訓練済み線形回帰モデル内へと入力して複数のベースライン治療スコアを取得することと、前記複数のフォローアップ治療埋め込みを前記訓練済み線形回帰モデル内へと入力して複数のフォローアップ治療スコアを取得することを含む、方法。

【0554】

127. 実施例126に記載の方法において、前記複数のプラセボ進行スコア及び前記複数の治療進行スコアを決定するステップは：前記複数のベースラインプラセボスコアと前記複数のフォローアッププラセボスコアとの間の差を決定して前記複数のプラセボ進行スコアを決定することと、前記複数のベースライン治療スコアと前記複数のフォローアップ治療スコアとの間の差を決定して前記複数の治療進行スコアを決定することを含む、方法。

10

【0555】

128. 実施例126に記載の方法において、前記複数のプラセボ進行スコア及び前記複数の治療進行スコアを決定するステップは：前記被験者プラセボ群内の各被験者について、前記被験者プラセボ群内の前記被験者のベースラインプラセボスコア及びフォローアッププラセボスコアに少なくとも基づいて適合化された線形モデルの勾配を決定することと、前記被験者治療群内の各被験者について、前記被験者治療群内の前記被験者のベースラインプラセボスコア及びフォローアッププラセボスコアに少なくとも基づいて適合化された線形モデルの勾配を決定することを含む、方法。

20

【0556】

129. 実施例122～128のいずれか1つに記載の方法において、前記複数の予測医療診断スコアは、予測フォローアップスコアと観測フォローアップスコアとの間の差として対応するベースラインスコアについて調整がなされて算出された疾患進行スコアを含む、方法。

【0557】

130. 実施例122～129のいずれかに記載の方法において、前記複数のプラセボ進行スコア及び前記複数の治療進行スコアを前記治療と関連付けるステップは、患者が前記治療を受けたかについての指示を受信し及び予測疾患進行スコアを出力するように構成されたモデルを生成するステップを含む、方法。

30

【0558】

131. 実施例130に記載の方法において、前記相関メトリックは前記モデルのP値である、方法。

【0559】

132. 実施例122～130のいずれかに記載の方法において、さらに：前記相関メトリックを所定の閾値と比較するステップを含む、方法。

【0560】

133. 実施例132に記載の方法において、さらに：前記比較に基づいて、前記治療と前記着目疾患との間の関連を識別するステップを含む、方法。

40

【0561】

134. 実施例133に記載の方法において、さらに：前記関連に基づいて前記治療を施す、調整する、又は適用するステップを含む、方法。

【0562】

135. 実施例133に記載の方法において、さらに：前記関連に基づいて医学的提案を提供するステップを含む、方法。

【0563】

136. 実施例122～135のいずれかに記載の方法において、前記着目疾患は非アルコール性脂肪性肝炎（NASH、non-alcoholic steatohepatitis）である、方法。

【0564】

50

137. 実施例 1 2 2 ~ 1 3 6 のいずれかに記載の方法において、前記訓練済み教師なし機械学習モデルは対照モデルである、方法。

【0565】

138. 実施例 1 3 7 に記載の方法において、前記対照モデルはSimCLRモデルである、方法。

【0566】

139. 実施例 1 2 4 ~ 1 3 8 のいずれかに記載の方法において、前記訓練済み線形回帰モデルは線形混合モデルである、方法。

【0567】

140. 実施例 1 2 4 ~ 1 3 9 のいずれかに記載の方法において、前記訓練済み線形回帰モデルは複数の割り当て済み医療診断スコアに基づいて適合化される、方法。

10

【0568】

141. 実施例 1 4 0 に記載の方法において、前記複数の割り当て済み医療診断スコアは 1 人以上の医師によって提供される、方法。

【0569】

142. 実施例 1 4 1 に記載の方法において、前記複数の割り当て済み医療診断スコアの各割り当て済み医療診断スコアは予め定義された値のセットから選択される、方法。

【0570】

143. 実施例 1 2 2 ~ 1 4 2 のいずれかに記載の方法において、前記複数の予測連続医療診断スコアは、複数の予測線維症スコア、複数の予測小葉内炎症スコア、又は複数の予測脂肪症スコアを含む、方法。

20

【0571】

144. 着目患者サブグループを識別する方法であって：臨床被験者群から取得された複数の医用画像を訓練済み教師なし機械学習モデル内へと入力して潜在空間内での複数の埋め込みを取得するステップと、前記複数の埋め込みをクラスタリングして 1 つ以上の埋め込みクラスタを生成するステップと、前記 1 つ以上の埋め込みクラスタに対応する 1 つ以上の患者サブグループを識別するステップと、前記 1 つ以上の患者サブグループの各患者サブグループをコバリエントと関連付けて前記着目患者サブグループを識別するステップとを含む、方法。

【0572】

30

145. 実施例 1 4 4 に記載の方法において、前記訓練済み教師なし機械学習モデルは対照モデルである、方法。

【0573】

146. 実施例 1 4 5 に記載の方法において、前記対照モデルはSimCLRモデルである、方法。

【0574】

147. 実施例 1 4 4 ~ 1 4 6 のいずれかに記載の方法において、前記コバリエントは着目治療であり、また、前記着目患者サブグループは前記着目治療が相当なインパクトを及ぼすサブグループである、方法。

【0575】

40

148. 実施例 1 4 7 に記載の方法において、前記 1 つ以上の患者サブグループの各患者サブグループを前記コバリエントと関連付けるステップは：患者サブグループについて、前記患者サブグループ内の患者が前記着目治療を受けたかについての指示を受信し及び予測疾患進行を出力するように構成されたモデルを、生成するステップと、前記モデルを評価して前記患者サブグループが前記着目患者サブグループであるかを決定するステップとを含む、方法。

【0576】

149. 実施例 1 4 8 に記載の方法において、前記モデルを評価するステップは、前記モデルの相関メトリックを決定すること及び前記相関メトリックを所定の閾値に対して比較することを含む、方法。

50

【0577】

150. 実施例148～149のいずれか1つに記載の方法において、前記相関メトリックはP値である、方法。

【0578】

151. 実施例148～150のいずれかに記載の方法において、前記生成されたモデルは前記患者サブグループ内の被験者の疾患進行値によって訓練される、方法。

【0579】

152. 実施例151に記載の方法において、前記疾患進行値は前記患者サブグループ内の前記被験者の医療診断スコアを含む、方法。

【0580】

153. 実施例151に記載の方法において、前記疾患進行値は前記患者サブグループ内の前記被験者の進行スコアを含む、方法。

【0581】

154. 実施例151に記載の方法において、前記疾患進行値は前記患者サブグループ内の前記被験者のDRP値を含む、方法。

【0582】

155. 実施例144～146のいずれかに記載の方法において、前記コバリエントは着目疾患の進行であり、また、前記着目患者サブグループは前記着目疾患の前記進行と有意な関連を有するサブグループである、方法。

【0583】

156. 実施例151に記載の方法において、前記1つ以上の患者サブグループの各患者サブグループを前記コバリエントと関連付けるステップは：患者サブグループについて、患者が前記患者サブグループに属するかについての指示を受信し及び予測疾患進行を出力するように構成されたモデルを、生成するステップと、前記モデルを評価して前記患者サブグループが前記着目患者サブグループであるかを決定するステップとを含む、方法。

【0584】

157. 実施例152に記載の方法において、前記モデルを評価するステップは、前記モデルの相関メトリックを決定すること及び前記相関メトリックを所定の閾値に対して比較することを含む、方法。

【0585】

158. 実施例157に記載の方法において、前記相関メトリックはP値である、方法。

【0586】

159. 実施例156～158のいずれかに記載の方法において、前記生成されたモデルは前記臨床被験者群の疾患進行値によって訓練される、方法。

【0587】

160. 実施例159に記載の方法において、前記疾患進行値は、前記患者サブグループ内の臨床被験者の医療診断スコア、前記患者サブグループ内の臨床被験者の進行スコア、又は前記患者サブグループ内の臨床被験者のDRP値を含む、方法。

【0588】

161. 実施例144～146のいずれかに記載の方法において、前記コバリエントは有害副作用であり、また、前記着目患者サブグループは前記有害副作用と有意な関連を有するサブグループである、方法。

【0589】

162. 実施例161に記載の方法において、前記1つ以上の患者サブグループの各患者サブグループを前記コバリエントと関連付けるステップは：患者サブグループについて、前記患者サブグループ内の患者が前記患者サブグループに属するかについての指示を受信し及び前記患者に前記有害副作用が生じるかを予測するように構成されたモデルを、生成するステップと、前記モデルを評価して前記患者サブグループが前記着目患者サブグループであるかを決定するステップとを含む、方法。

【0590】

10

20

30

40

50

163. 実施例 1 6 2 に記載の方法において、前記モデルを評価するステップは、前記モデルの相関メトリックを決定すること及び前記相関メトリックを所定の閾値に対して比較することを含む、方法。

【0591】

164. 実施例 1 6 3 に記載の方法において、前記相関メトリックは P 値である、方法。

【0592】

165. 実施例 1 4 4 ~ 1 4 6 のいずれかに記載の方法において、前記コバリエントは有害副作用であり、また、前記着目患者サブグループは治療後に前記有害副作用を経験することについて有意な関連を有するサブグループである、方法。

【0593】

166. 実施例 1 6 5 に記載の方法において、前記 1 つ以上の患者サブグループの各患者サブグループを前記コバリエントと関連付けるステップは：患者サブグループについて、前記患者サブグループ内の患者が前記治療を受けたかについての指示を受信し及び前記患者に前記有害副作用が生じるかを予測するように構成されたモデルを、生成するステップと、前記モデルを評価して前記患者サブグループが前記着目患者サブグループであるかを決定するステップとを含む、方法。

【0594】

167. システムであって：1 つ以上のプロセッサと、メモリと、1 つ以上のプログラムとを備えるのであって、前記 1 つ以上のプログラムは前記メモリ内に記憶されており且つ前記 1 つ以上のプロセッサによって実行されるように構成されており、前記 1 つ以上のプログラムは実施例 1 ~ 1 6 6 のいずれかの方法を行うための命令を含む。

【0595】

168. 1 つ以上のプログラムを記憶する非一時的コンピュータ可読記憶媒体であって、該 1 つ以上のプログラムは命令を備えるのであり、該命令は電子装置の 1 つ以上のプロセッサによって実行されると該電子装置に実施例 1 ~ 1 6 6 のいずれかの方法を行わせる、非一時的コンピュータ可読記憶媒体。

10

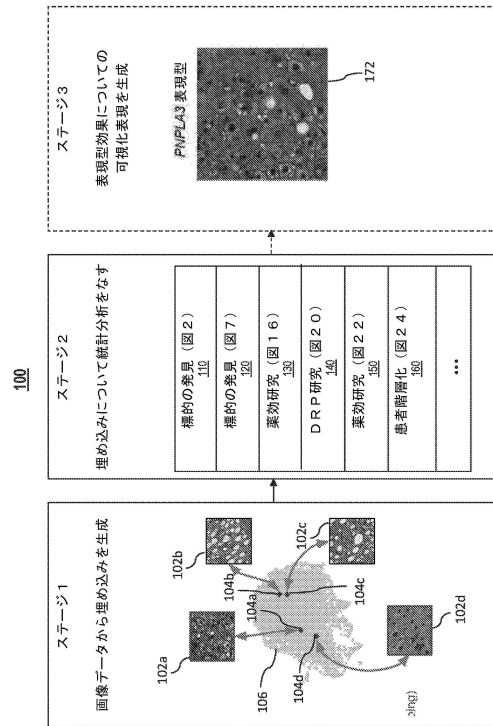
20

30

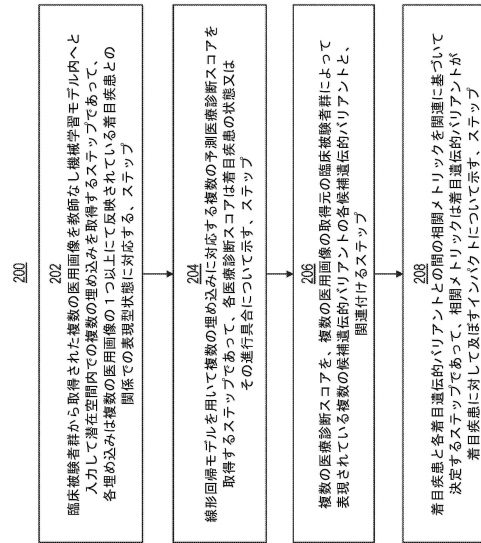
40

50

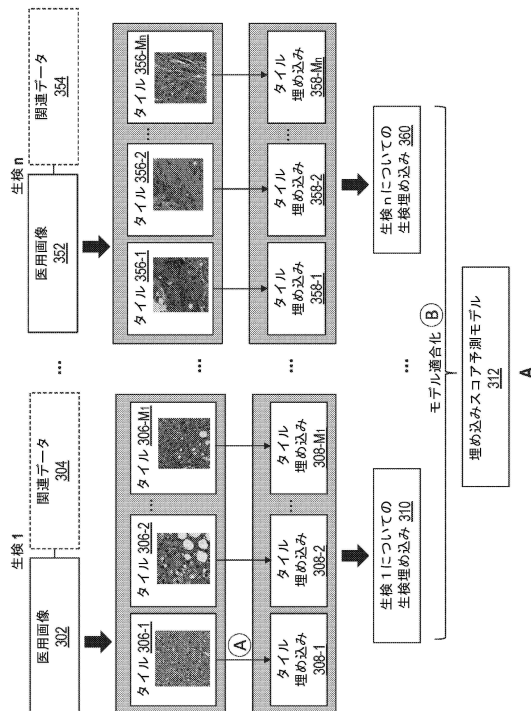
【図面】
【図 1】



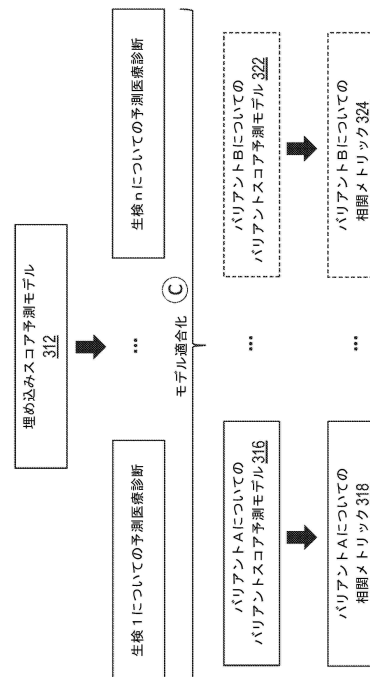
【図 2】



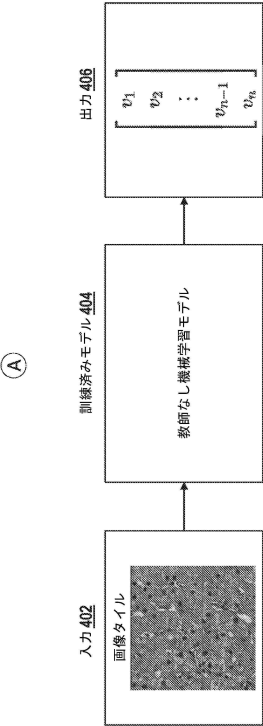
【図 3 A】



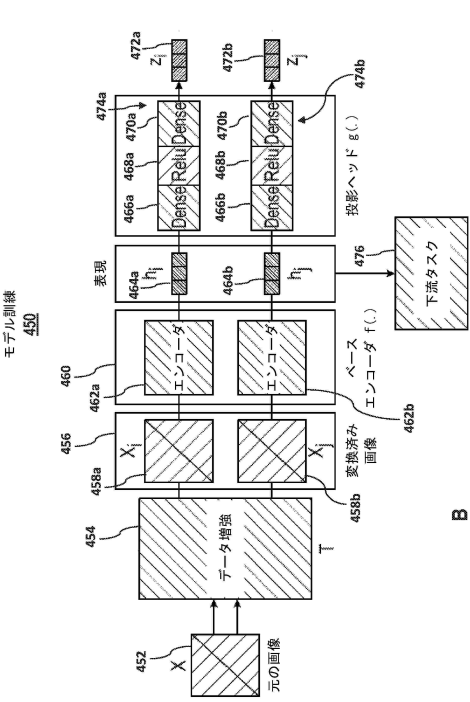
【図 3 B】



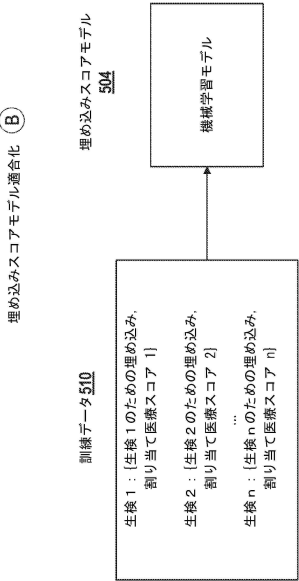
【図 4 A】



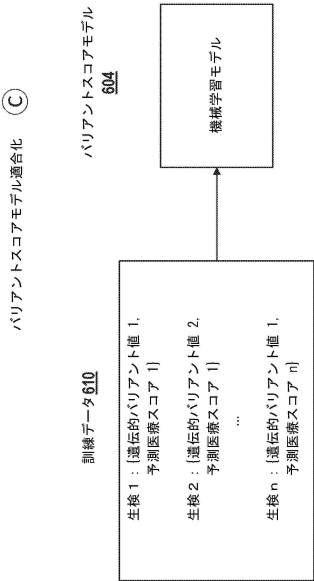
【図 4 B】



【図 5】



【図 6】



10

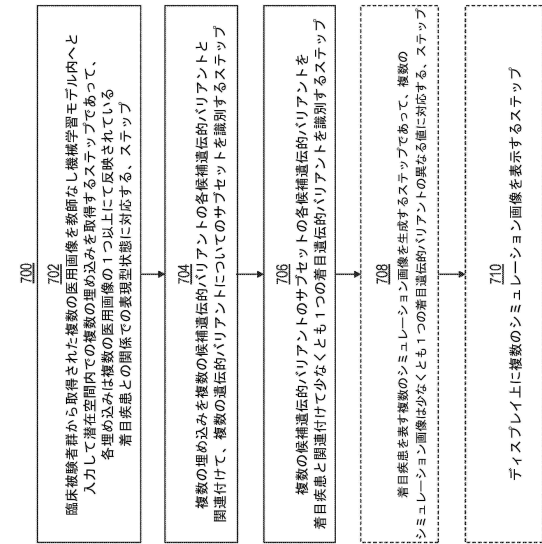
20

30

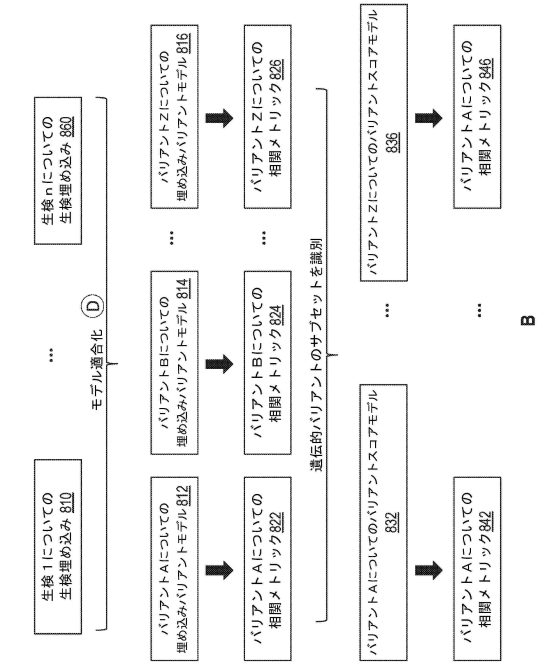
40

50

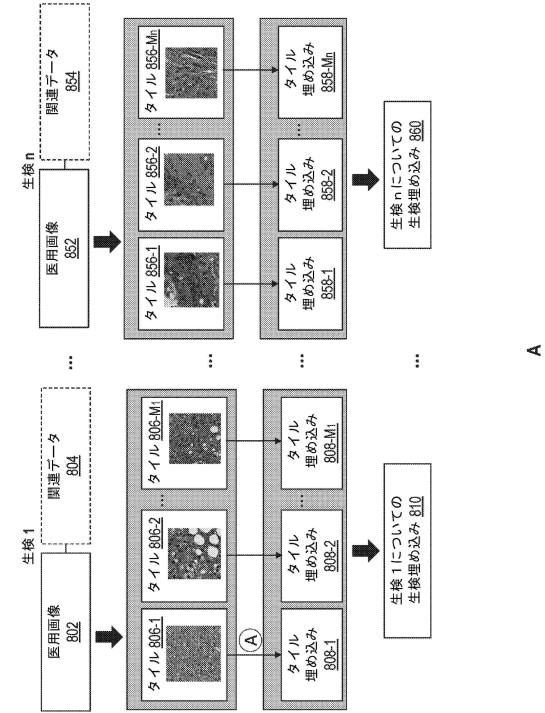
【図 7】



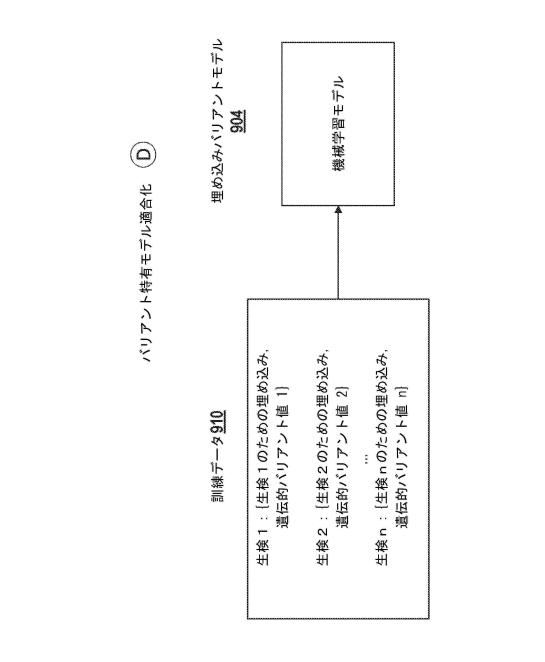
【図 8 B】



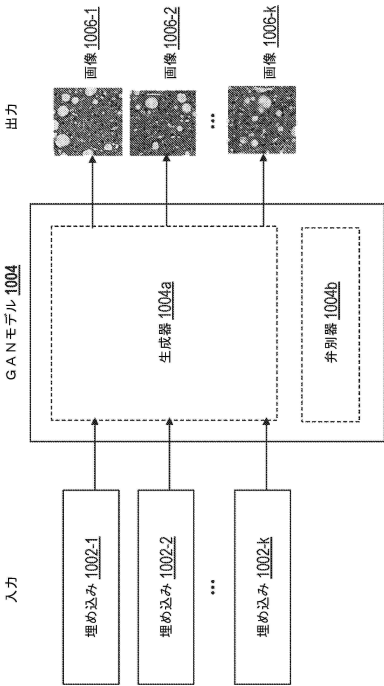
【図 8 A】



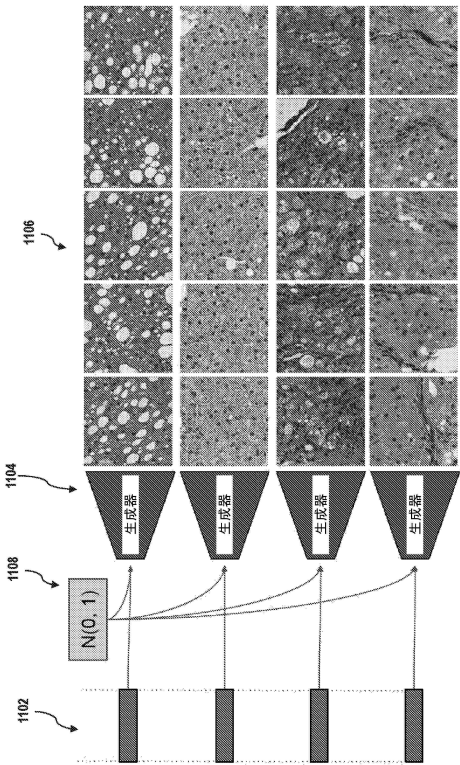
【図 9】



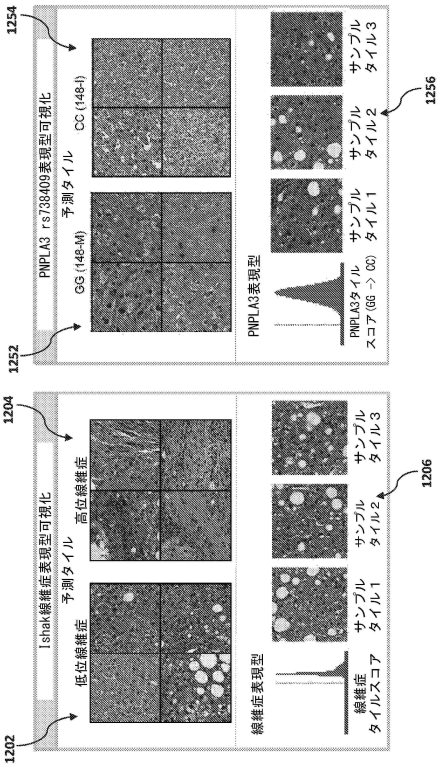
【図 1 0】



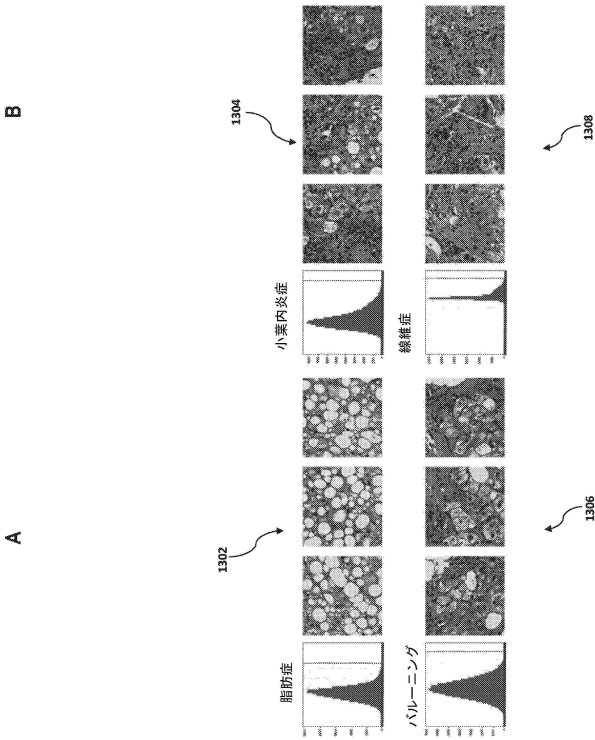
【図 1 1】



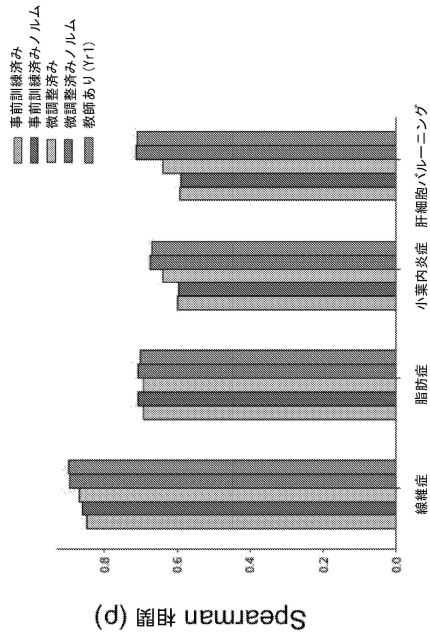
【図 1 2】



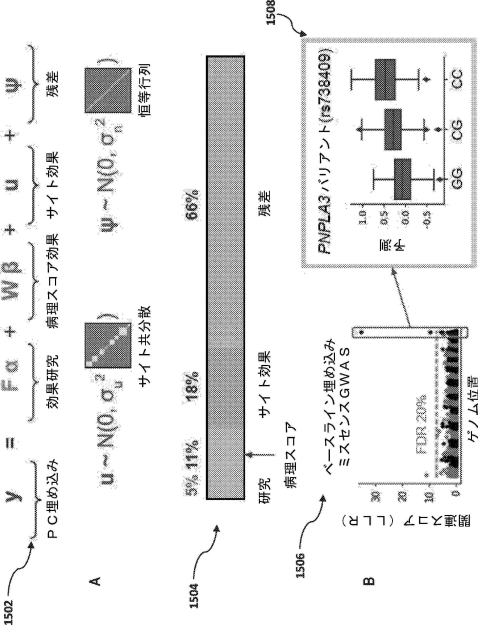
【図 1 3】



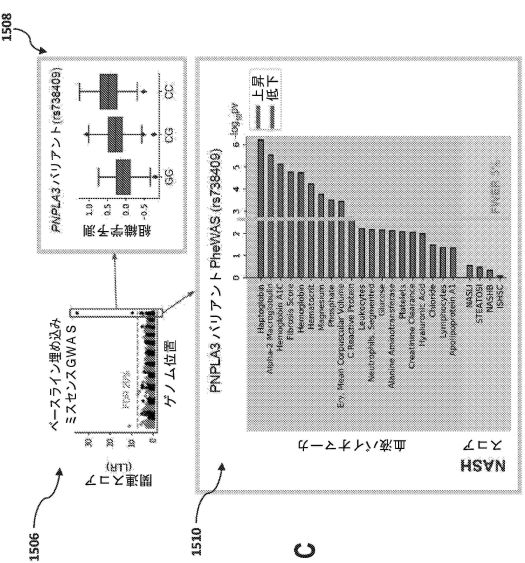
【図 1 4】



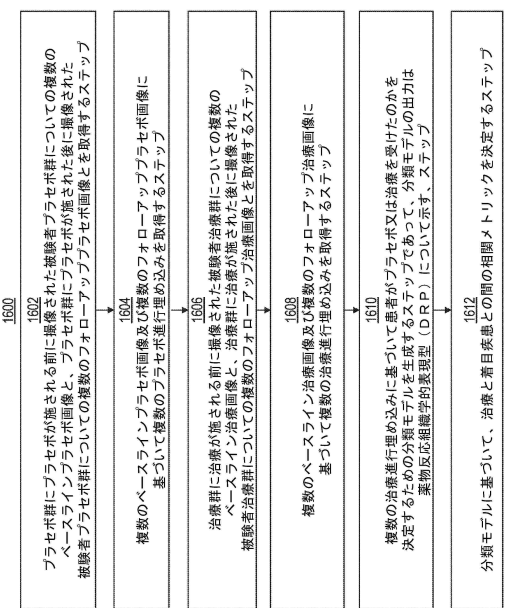
【図 1 5 - 1】



【図 1 5 - 2】



【図 1 6】



10

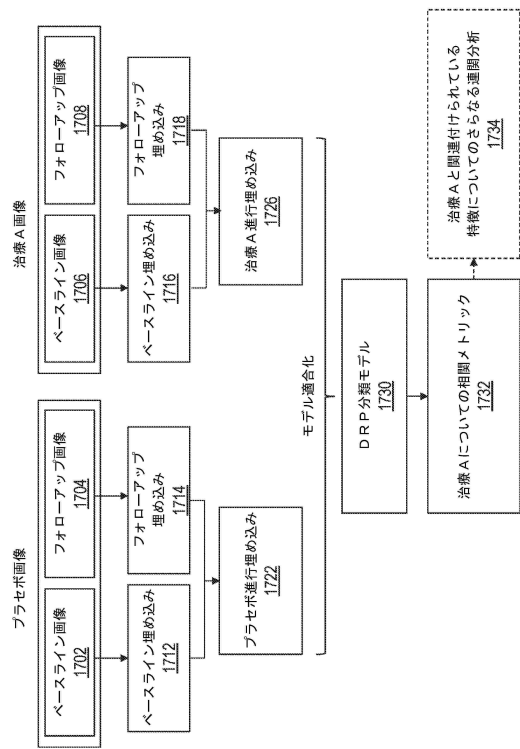
20

30

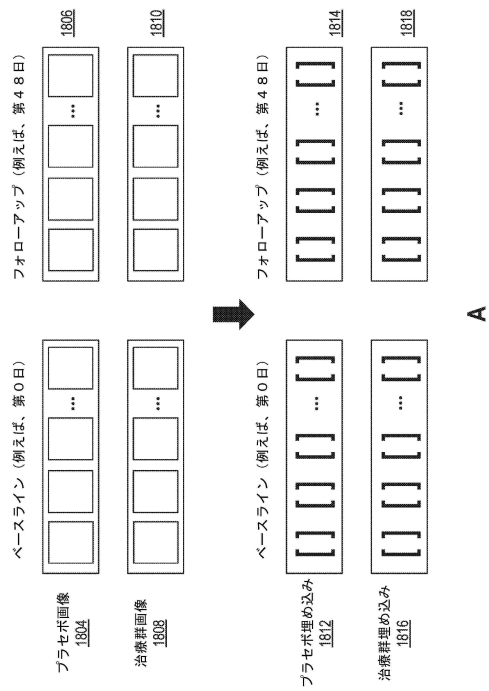
40

50

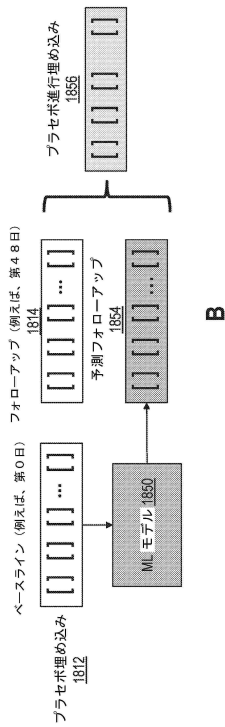
【図 1 7】



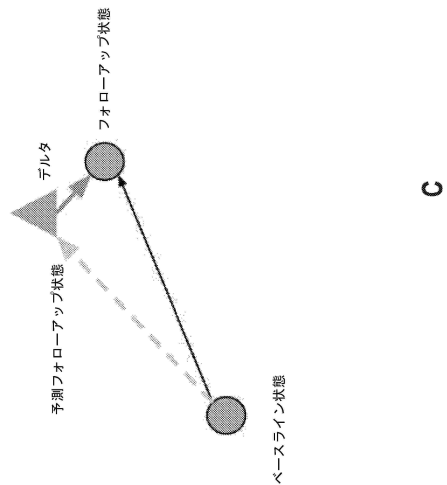
【図 1 8 A】



【図 1 8 B】



【図 1 8 C】



10

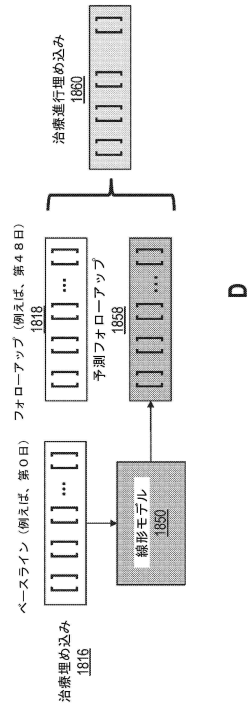
20

30

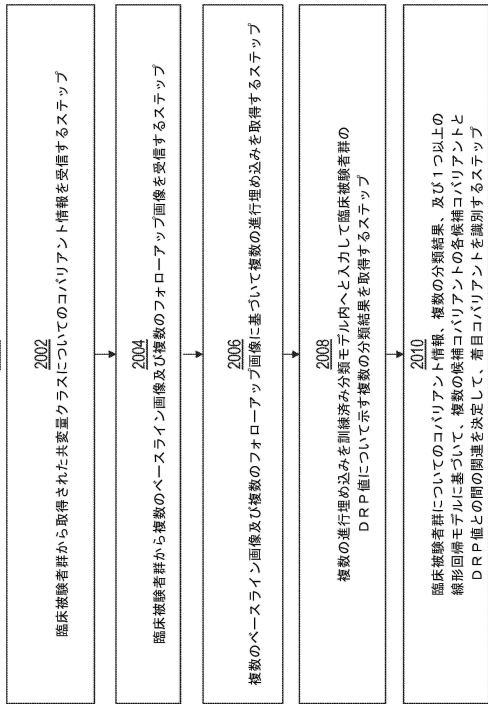
40

50

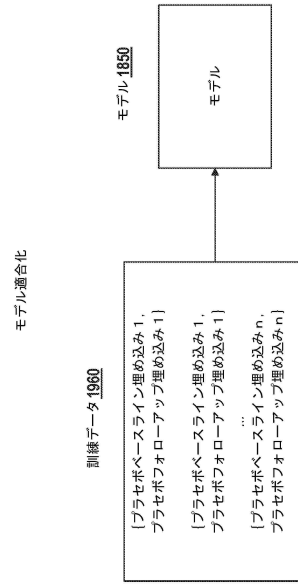
【図 18 D】



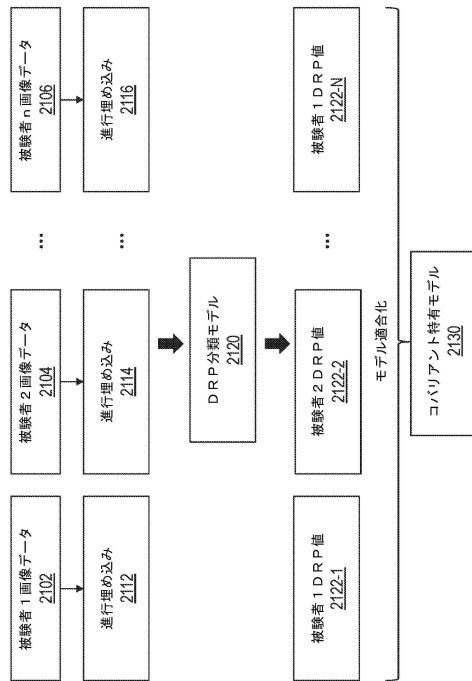
【図 20】



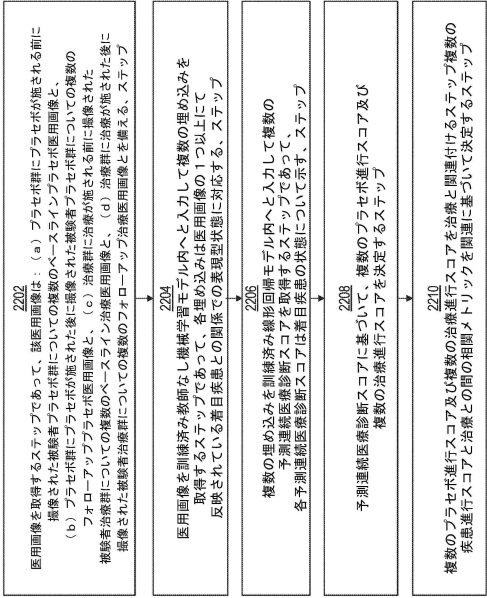
【図 19】



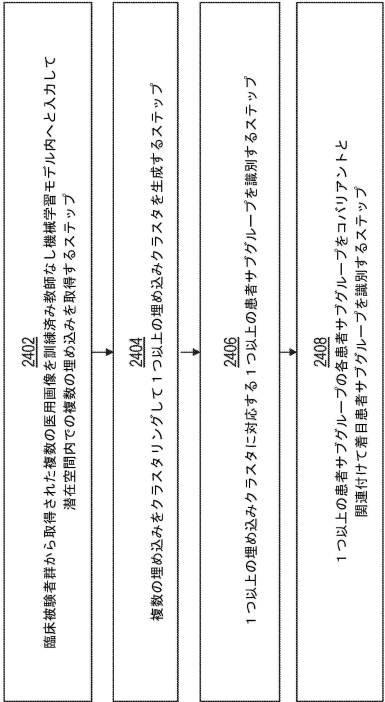
【図 21】



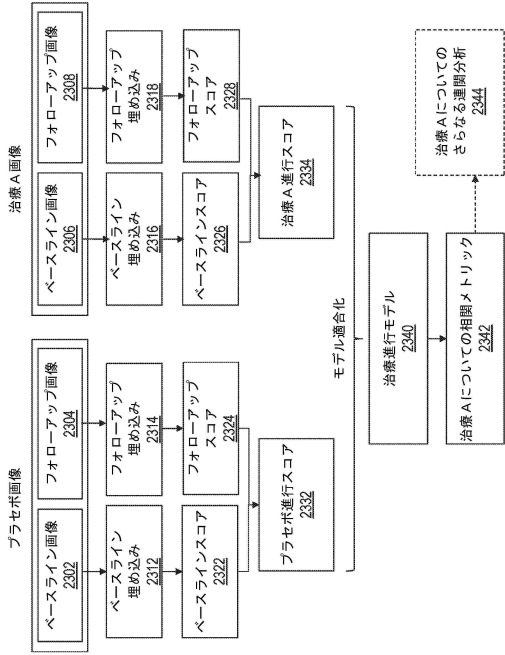
【図 2 2】



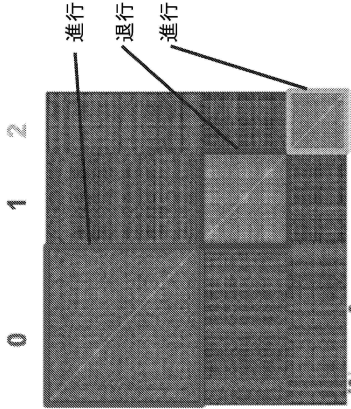
【図 2 4】



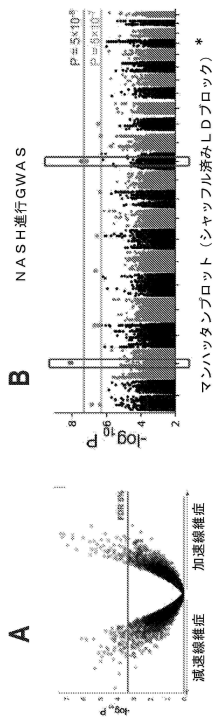
【図 2 3】



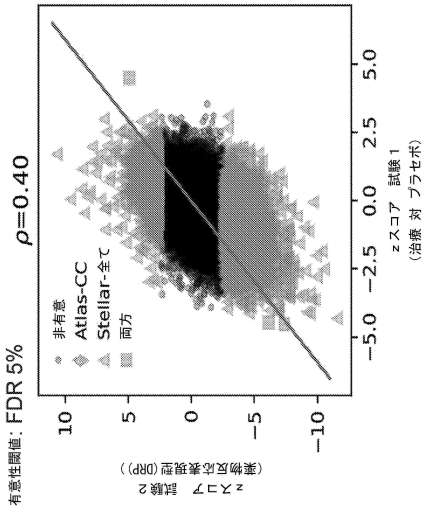
【図 2 5】



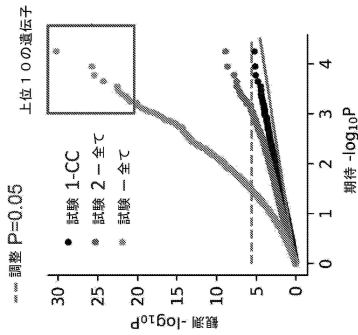
【図 2 6】



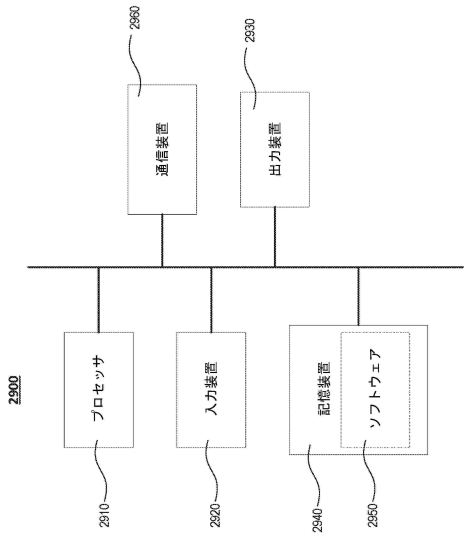
【図 2 8】



【図 2 7】



【図 2 9】



10

20

30

40

50

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2022/075006

A. CLASSIFICATION OF SUBJECT MATTER INV. G16H50/20 G16H50/50 ADD.		
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) G16H		
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	AMARO TAYLOR-WEINER ET AL: "A Machine Learning Approach Enables Quantitative Measurement of Liver Histology and Disease Monitoring in NASH", HEPATOLOGY, JOHN WILEY & SONS, INC, US, vol. 74, no. 1, 24 June 2021 (2021-06-24), pages 133-147, XP071565829, ISSN: 0270-9139, DOI: 10.1002/HEP.31750 the whole document -----	1-23, 44-65
A	WO 2010/082096 A2 (KONINKL PHILIPS ELECTRONICS NV [NL]; UNIV WASHINGTON [US] ET AL.) 22 July 2010 (2010-07-22) page 3, line 23 - page 9, line 16 ----- -/--	1-23, 44-65
☒ 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 23 November 2022		Date of mailing of the international search report 27/01/2023
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 Rinelli, Pietro

Form PCT/ISA/210 (second sheet) (April 2005)

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2022/075006

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	De Jong Johann ET AL: "Towards Realizing the Vision of Precision Medicine: AI Based Prediction of Clinical Drug Response Authors", , 28 July 2021 (2021-07-28), pages 1738-1750, XP093000008, Retrieved from the Internet: URL:https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8320273/pdf/awab108.pdf [retrieved on 2022-11-21] the whole document -----	1-23, 44-65

Form PCT/ISA/210 (continuation of second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2022/075006**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

10

20

Box No. III Observations where unity of invention is lacking (Continuation of Item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims.; it is covered by claims Nos.:
1-23, 44-65

30

40

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

50

International Application No. PCT/US2022 /075006

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-23, 44-65

Trial status determination

2. claims: 24-43, 66-88

Treatment status determination

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2022/075006

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 2010082096	A2	22-07-2010	AU 2009337516 A1	01-09-2011
			CN 102272795 A	07-12-2011
			EP 2387777 A2	23-11-2011
			JP 5775823 B2	09-09-2015
			JP 2012515010 A	05-07-2012
			US 2011263949 A1	27-10-2011
			US 2014328525 A1	06-11-2014
			WO 2010082096 A2	22-07-2010

10

20

30

40

50

フロントページの続き

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,N
E,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,
CV,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IQ,IR,IS,IT,J
M,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,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,ST,SV,SY,T
H,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW

(特許庁注：以下のものは登録商標)

1. J A V A

2. P Y T H O N

スト グランド アヴェニュー 279 スイート 200 インシトロ インコーポレイテッド内

(72)発明者 マイケル ベレケット

アメリカ合衆国 カリフォルニア州 94080 サウス サンフランシスコ イースト グランド ア
ヴェニュー 279 スイート 200 インシトロ インコーポレイテッド内

(72)発明者 マシュー アルバート

アメリカ合衆国 カリフォルニア州 94080 サウス サンフランシスコ イースト グランド ア
ヴェニュー 279 スイート 200 インシトロ インコーポレイテッド内

F ターム (参考) 5L099 AA26