

第37回医療情報学連合大会シンポジウム
診療情報から「病態」を取り出そう！Phenotypingのススメ

診療情報による Phenotypingの現状・限界

河添 悦昌^{*1}、香川 璃奈^{*2}、今井 健^{*3}、大江 和彦^{*12}

^{*1}東京大学医学部附属病院 企画情報運営部

^{*2}東京大学大学院 医学系研究科 医療情報学分野

^{*3}東京大学大学院 医学系研究科 疾患生命工学センター

電子カルテデータの臨床研究への活用

- 臨床医学研究
 - 適格基準・除外基準を満たす候補症例を抽出
 - 仮説となるようなシーズを効率的にスクリーニング
- 薬剤疫学研究
 - ドラッグリパーパシング：既存薬の新たな効能を発見
 - がん患者でのDM治療薬（メトホルミン）によるOS延長^{*1}
- 遺伝学の研究：GWAS / PheWAS^{*2}
 - 表現型を有する群と対照群を集め、群間におけるSNPsの発現頻度の違いを調査
 - リバースGWAS：電子カルテの利用を前提とした方法

1. Xu H, et al. Validating drug repurposing signals using electronic health records: a case study of metformin associated with reduced cancer mortality.

2. Denny JC, et al. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations.

EHR Phenotyping：カルテからの表現型抽出

- 電子カルテから関心のある臨床的特徴を持つ症例を同定
 - 血圧／HbA1cなどの検査値歴
 - 降圧薬／血糖降下薬といった薬剤の服用歴
 - 喫煙や飲酒などの生活習慣
 - 本態性高血圧症／2型糖尿病といった疾患歴
- 何の病気で診療を受けているか、ということすら計算機で判断することが難しい
 - 診療録：自然言語による記載、記録が断片的、いつ時点のものか不明
 - レセプト：保険請求を目的とするため真の病名を反映していない
 - 検査値：結果がスパース、検査値だけで診断できない疾患も多い
 - 処方歴：同一薬が複数疾患に使われることも（ARB: 心不全、高血圧、CKD）

「真の病名」がカルテ記載されないため、これらを組合せ再現性良く同定する方法が必要

諸外国の事情

- 社会制度の異なる諸外国においても同様の事情
 - 心不全症例を同定するために、ICDコードのみを利用するだけでは不十分だったとのメタアナリシス結果*1
- 米国：eMERGEプロジェクト
 - NIH主導の遺伝学への応用を目的としたプロジェクト
 - 複数施設の電子カルテを情報源として大規模な症例コホートを作成
 - 施設間で相互利用可能なPhenotypingアルゴリズムを開発し、PheKBというカタログDBで公開
- 開発された13アルゴリズムの性能評価の結果*2
 - ケース例の陽性的中率がいずれも90%を越えた
 - アルゴリズムの開発と評価を繰り返して行うことが重要
 - 複数施設での評価が単施設での評価よりも重要

1. McCormick N, et al. Validity of heart failure diagnoses in administrative databases: a systematic review and meta-analysis.

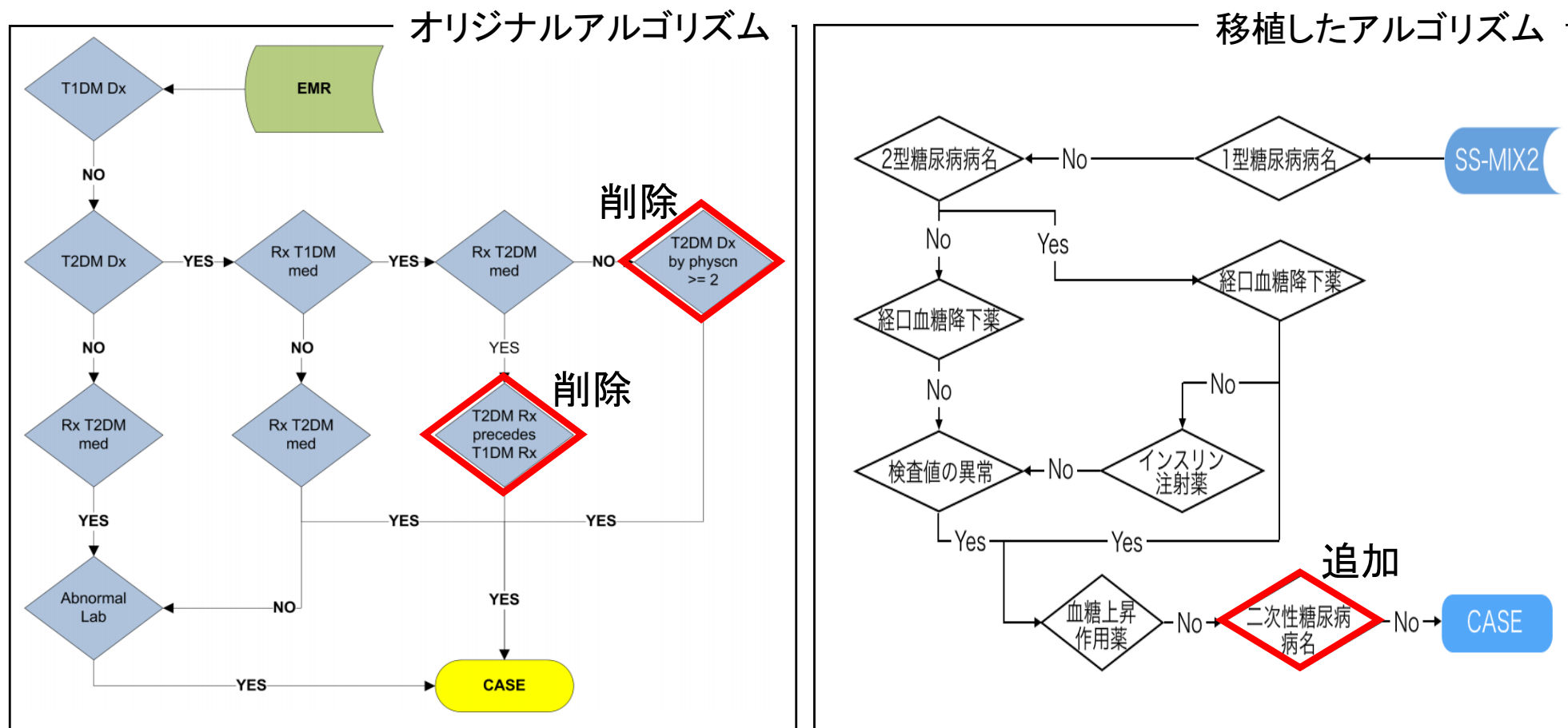
2. Newton KM, et al. Validation of electronic medical record-based phenotyping algorithms: results and lessons learned from the eMERGE network.

Phenotypingアルゴリズムの開発

- 本邦でのアルゴリズム開発の必要性
 - 疾患の診断基準の違い
 - 自然言語文章に対する解析精度の違い
 - 医療情報の種類の違い
 - などから、海外で開発されたアルゴリズムを利用する場合に、報告される精度を外挿することはできない。

ルール形式のアルゴリズム（方法）

- PheKBで公開される2型糖尿病症例の同定アルゴリズムを移植^{*1}
- データソースとしてSS-MIX2ストレージを利用



1. 香川 璃奈 他. SS-MIX2標準化ストレージを用いた2型糖尿病症例のEHR Phenotyping手法の開発と評価. 医療情報学35(Suppl.), pp.672-675, 2015.

ルール形式のアルゴリズム（結果）

2型糖尿病症例の同定アルゴリズム

手法	同定された症例数	陽性的中率	感度
先行研究	540 / 12,740	95.0%	67.1%
PheKB再現	1,662 / 115,039	94.0%	15.2% (推定値)
PheKB再現 + 二次性DM除外	877 / 115,039	100.0%	8.01% (推定値)

- 陽性的中率に関して：
 - PheKBを再現したのみでも 94 % と良好
 - 2 次性糖尿病を除外することでさらに向上
- 感度に関して：
 - 先行研究よりも低い。
 - アルゴリズムでは抽出の対象外となる、E14x（詳細不明の糖尿病）で登録されている症例が多い

ルール形式と機械学習の比較（方法）

- ルール形式のアルゴリズム
 - 人間がみて解釈できる
 - ルールを作るためにドメインの知識が必要
- 機械学習によるアルゴリズム
 - ドメインの知識はそれほど必要なし。
 - 解釈性が犠牲になる。入力データに対応する正解セットが必要。
- ルール形式と機械学習の方法の精度を比較*2
 - 本態性高血圧症、2型糖尿病
 - ルール形式、機械学習（勾配ブースティング決定木）
- データセット
 - 104,522人からランダムに抽出した650人に正解ラベルを付与
 - 本態性高血圧（150/650）、2型糖尿病（85/650）

ルール形式と機械学習の比較（結果）

本態性高血圧症

	陽性的中率	感度	正解率
機械学習	79.64%	66.86%	89.80%
ルール形式	75.50%	75.50%	71.56%

【ルール】

(A)かつ((B)または(C))

(A) 2次性高血圧に関する病名登録がない

(B) 2次性高血圧以外の高血圧の病名登録がある

(C) ARB/ACE阻害薬の処方がある

ルール形式と機械学習の比較（結果）

本態性高血圧症

	陽性的中率	感度	正解率
機械学習	79.64%	66.86%	89.80%
ルール形式	75.50%	75.50%	71.56%

2 型糖尿病

	陽性的中率	感度	正解率
機械学習	89.77%	67.86%	95.77%
ルール形式	93.94%	46.97%	87.54%

機械学習とルール形式をどのように使い分ければよいのだろうか？

ルール形式と機械学習の比較（結果）

ルール形式において変数を単独で用いた場合の陽性的中率を調査

【本態性高血圧】

変数	陽性的中率
高血圧の病名登録	77.54%
ARB/ACE阻害薬	71.26%

単独で高い陽性的中率を示す変数なし

【2型糖尿病】

変数	陽性的中率
2型糖尿病の病名登録	36.79%
インスリン処方	68.57%
異常検査値（注1）	90.63%
経口血糖降下薬	96.88%

単独で高い陽性的中率を示す変数あり

識別に有効な変数がわかっている場合は、ルール形式でも高いPPVを得られる
そのような変数が知られていない場合は機械学習が良い

正解ラベル付けの必要性

- 陽性的中率（PPV）だけを評価するだけでは..
 - 陽性と判定した症例のカルテを開き結果が正しいかどうか判定
 - 比較的簡便に行える。
- PPVだけでは十分でない理由
 - 確実にケース例といえる症例を1件だけ見つけるアルゴリズム
 - 有病率の異なる施設にアルゴリズムを適用する場合は、感度と特異度が指標としてより適切
- 正解ラベルを付きデータセットが必要
 - 感度、特異度の算出に必要
 - 開発例（2）では650人の正解ラベル付きセットを作成した
 - データセット作成に時間を要することがアルゴリズム開発の律速

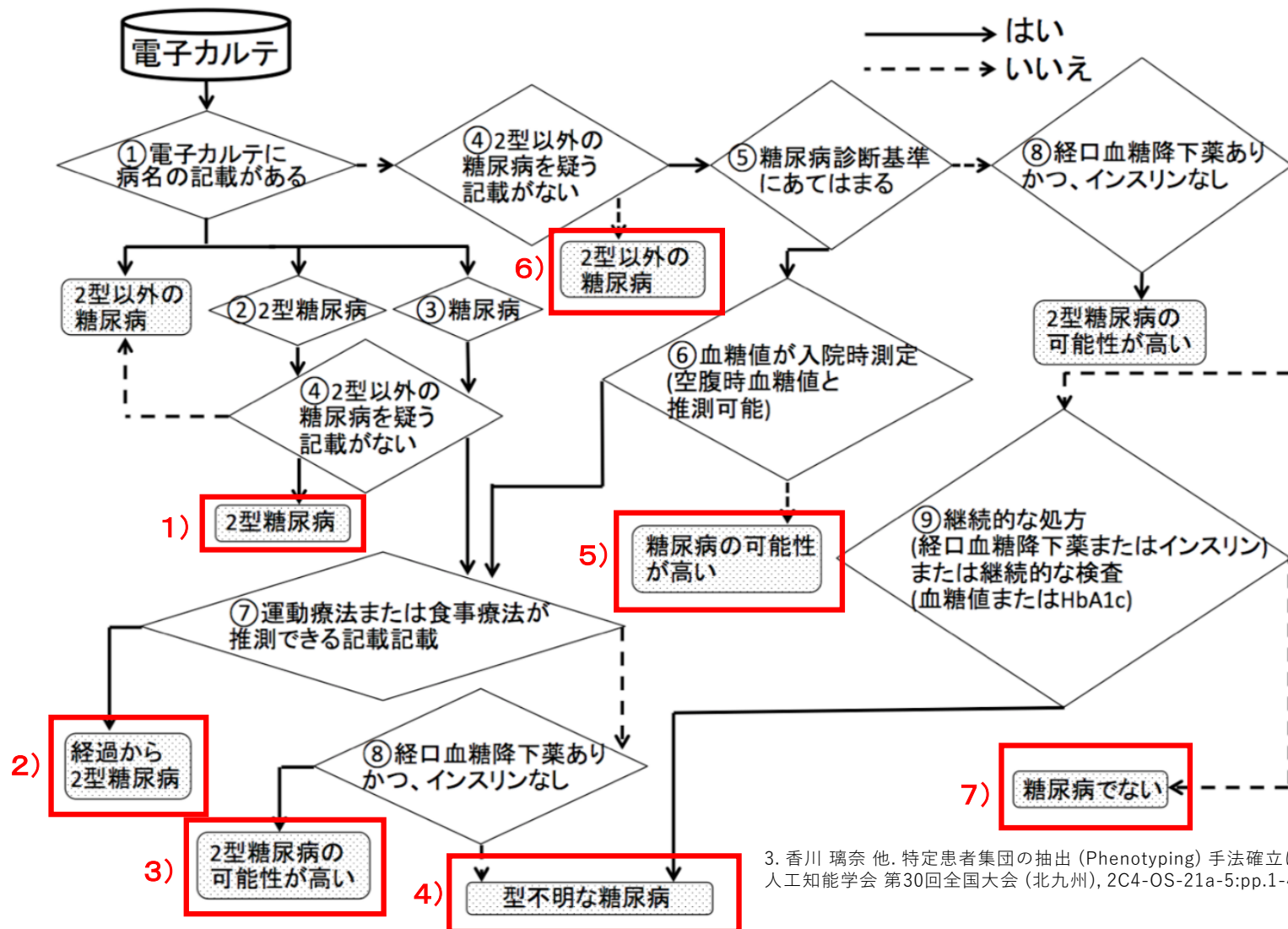
明確な基準による正解ラベル付けの必要性

- 2 型糖尿病の正解ラベル付けの際の問題
 - 診療録に2型糖尿病と記載されていなくとも、それに準じた治療行為から2型糖尿病が推察される場合。
 - 診療録に糖尿病とのみ記載があり、治療経過からほぼ2型糖尿病だろうが、1型を否定する明確な根拠がない。
 - 紹介状に糖尿病性網膜症と記載があるが、それ以上の情報がない。

このようなケースを2型糖尿病とするかどうかは、人によって判断が異なる
正解ラベルを付けるためにも明確な基準が必要だろう

明確な基準による正解ラベル付けの必要性

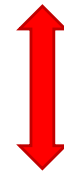
2型糖尿病の正解ラベル付けのフローチャートを作成^{*3}



3. 香川 璃奈 他. 特定患者集団の抽出 (Phenotyping) 手法確立に向けた技術的課題に関する考察. 人工知能学会 第30回全国大会 (北九州), 2C4-OS-21a-5:pp.1-4, 2016. 13

明確な基準による正解ラベル付けの必要性

- 1) 2型糖尿病の記載があるもの
- 2) 経過から2型糖尿病と判断できるもの
- 3) その他理由で2型糖尿病の可能性が高いもの



2型糖尿病に分類すると定義

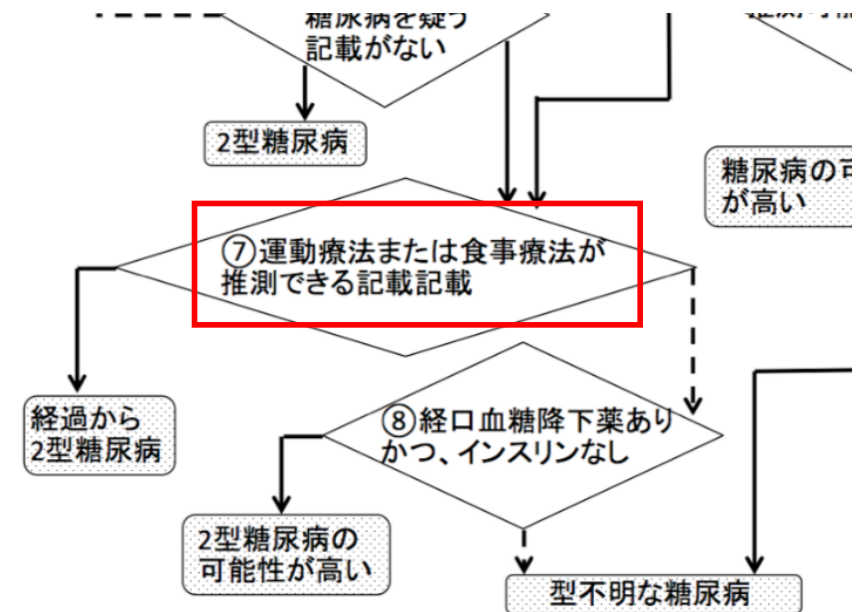
- 4) 型不明の糖尿病
- 5) 糖尿病の可能性が高いもの
- 6) 2型以外の糖尿病
- 7) 糖尿病でないもの

- 2型糖尿病に分類するとした症例のうちカルテに「2型糖尿病」が**明記されていたものは約57%**
- ICD-10のE11x（インスリン非依存性糖尿病）が登録されている症例のうち、**2型糖尿病に分類するとした患者は約半数**

- 診療録の記載に対する文字列のマッチングやレセプト病名だけでは不十分であることを改めて示した
- また、どの範囲を2型糖尿病に分類するかという閾値を変えることで、アルゴリズムの精度も容易に変わる
- e-Phenotypingにおいては、アルゴリズムの開発そのものよりも、質の高い正解ラベルを付与するための方法論が重要

計算機処理が困難である理由（１）

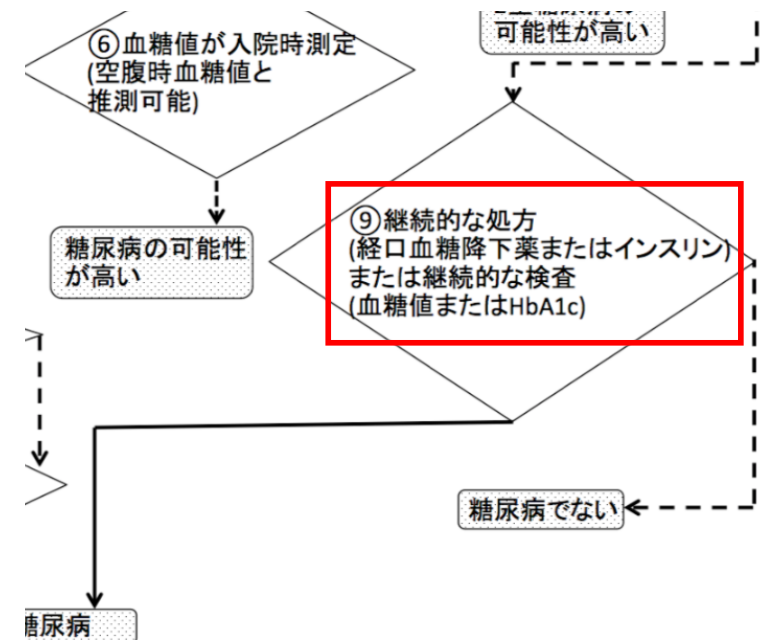
- 前述のフローチャートを計算機処理するうえでの難しさを検討した
- ⑦「運動療法または食事療法が推測できる記載」
 - 診療録に栄養指導を行ったことの記載はあるものの、2型糖尿病に対する指導であることが明示されないものがある。
 - 医療者であれば、栄養指導の対象となる疾患はいくつかに限定されることと、各疾患に対して行われるべき指導内容が想定されることから、2型糖尿病を対象としたものかどうかの判断は比較的容易。
 - そのような知識がない計算機にとって、記載される内容だけで判断することは難しい。



計算機処理が困難である理由（２）

- ⑨「経口血糖降下薬またはインスリンの継続的な処方」

- 一時的な血糖上昇に対して血糖を下げる薬剤が使われることがあるため、そのような症例を除外する目的。
- 医療者であれば一時的かどうかは、血糖上昇の前にどのような医療イベントが生じていたのかをみることで比較的明らかに判断できる。
- 計算機でこれを行うためには、血糖の上昇を伴う医療イベントの種類が何であるかということの知識が必要。
- また、一時的という表現を単純な期間に置き換えられないことから、計算機処理で判断することが難しい。



e-Phenotypingの現状とこれからの展望

- 現状の電子カルテの記録方式
 - 何の病気で診療を受けているか、ということすら計算機判断が難しい。
- e-Phenotypingアルゴリズムの必要性
 - 診療録の記載に対する単純な文字列のマッチングやレセプト病名だけは、対象症例が十分に抽出できない。
 - 異なる目的で記録された複数の診療情報と対象とする必要。
- e-Phenotypingアルゴリズム開発における問題
 - アルゴリズムの評価のために正解データセットが必要。
 - 正解かどうかの判定基準を明確にする必要がある
- 計算機処理がむずかしい理由
 - 記載される事実のみでは疾患有無の判定が難しい。
 - カルテは人間が読むこと（文脈を読む、知識で行間を補うこと）を前提とするため。

e-Phenotypingの現状とこれからの展望

- 過渡期においてはやむを得ないが、このような方法が続けることが得策とも思えない。
- また、臨床研究で使われるようなインテンシブなテンプレート入力を進めればよいという話でもない。
- 日常診療で医療者がスムーズに記録しつつ、Phenotype情報が正しく入力されるために必要なことを再考する必要。
- 紙カルテ時代の運用を基に電子化するのではなく、情報システムを前提とした運用をもとに電子カルテを設計する必要がある。
- 電子カルテに仕掛けを作る以外にも、医療者全体で問題を共有する必要がある。