

臨床医学テキストアノテーションガイドライン  
for Real-MedNLP

矢田 竣太郎<sup>†</sup> 荒牧 英治<sup>†</sup> 田中 リベカ<sup>‡</sup> Fei Cheng<sup>‡</sup> 黒橋 禎夫<sup>‡</sup>  
<sup>†</sup> 奈良先端科学技術大学院大学  
<sup>‡</sup> 京都大学

第 8 版  
2021 年 12 月 20 日

# 目次

<b>第 1 章</b>	<b>医療エンティティアノテーション</b>	<b>1</b>
1.1	概要 . . . . .	1
1.2	基本方針 . . . . .	1
1.2.1	入れ子構造の禁止 . . . . .	1
1.2.2	並列表現の一括タグ付け . . . . .	1
1.3	医療エンティティの種類 . . . . .	2
1.3.1	病変・症状 (Diseases and symptoms) . . . . .	2
1.3.1.1	エンティティ範囲の目安 . . . . .	2
1.3.1.2	確実性 (certainty) 属性に応じた手がかり表現 . . . . .	3
1.3.1.3	注意が必要な表現 . . . . .	4
1.3.2	臓器・部位 (Anatomical parts) . . . . .	5
1.3.2.1	注意が必要な表現 . . . . .	5
1.3.3	特徴・尺度 (Features and measurements) . . . . .	5
1.3.3.1	注意が必要な表現 . . . . .	6
1.3.4	変化 (Change) . . . . .	7
1.3.5	時間表現 (TIMEX3) . . . . .	7
1.3.6	検査・問診 (Test) . . . . .	9
1.3.7	薬品 (Medicine) . . . . .	10
1.3.8	治療 (Remedy) . . . . .	11
1.3.8.1	注意が必要な表現 . . . . .	12
1.3.9	クリニカルコンテキスト (Clinical Context) . . . . .	12
1.3.10	保留 (Pending) . . . . .	12
1.4	医療エンティティアノテーションの注意事項 . . . . .	13
1.4.1	長さ優先タグ付け . . . . .	13
1.4.2	病変・症状および時間表現タグの優先 . . . . .	13
1.4.3	誤字・脱字の訂正 (correction 属性) . . . . .	14
1.4.4	「術後」をめぐる表現 . . . . .	14
1.4.4.1	「(〇〇) 術後変化」の場合 . . . . .	14
1.4.4.2	「術後」の前後に医療エンティティがある場合 . . . . .	14
1.4.5	SOAP セクションと医療エンティティ属性 . . . . .	15
1.5	タグ付け文書例 . . . . .	15

# 第1章

## 医療エンティティアノテーション

### 1.1 概要

本章では、臨床テキストに出現する医療エンティティアノテーションについて説明する。医療エンティティアノテーションとは、電子カルテ (診療記録) や読影所見など実際の臨床テキストの記述から、病名、疾患名、部位名などの医学的な概念に相当する言語表現に対してタグを付与することで、それら言語表現をその他の記述から**医療エンティティ**として識別する作業である。

まずアノテーションの基本方針を述べ、医療エンティティアノテーションの種類について列挙する。続いて作業上の諸注意を述べ、最後に1文書全体で医療エンティティアノテーションを施した例を示す。

### 1.2 基本方針

ここでは、後述する医療エンティティアノテーションの種類によらず一般的に適用される基本的なアノテーション方針を説明する。なお、医療エンティティアノテーションの種類を踏まえた注意点を1.4節でさらにまとめてあるので、合わせて確認されたい。

#### 1.2.1 入れ子構造の禁止

本アノテーションでは、ある医療エンティティアノテーションの内部に別の医療エンティティアノテーションが入るような入れ子構造を認めない。例えば「慢性副鼻腔炎」という病名 (症状名) は医学的には「慢性」+「副鼻腔」+「炎 (症)」と分割でき、包含される「副鼻腔」は人体の部位の一つである。しかし、本ガイドラインでは `<d certainty="positive">慢性<a> 副鼻腔 </a> 炎 </d>` のようなアノテーションは認めないということである。全体として1種類の医療エンティティアノテーションであるか、または数種の医療エンティティアノテーションの連続という形式だけが許容される。

#### 1.2.2 並列表現の一括タグ付け

中黒 (・), スラッシュ (/), ハイフン (-), 読点 (,) のような記号や、「もしくは」「および」「ならびに」のような表現をはさんで同じ種類の医療エンティティアノテーションが並列されている場合は、全体をまとめて1つのタグとして良い。

## 例

- (1) <TIMEX3 type="DATE"> 前回 </TIMEX3> 指摘の<d certainty="positive"> すりガラス影や網状影, 牽引性気管支拡張, 蜂巣肺 </d> を認めます

## 1.3 医療エンティティの種類

### 1.3.1 病変・症状 (Diseases and symptoms)

病変・症状エンティティ<d>は具体的な病変や症状を指す表現に付与する。病変の存在を示す異常な見え(「すりガラス状」, 「網状影」など), 異常な聞こえ(「fine crackles」, 「coarse crackle」など)などの所見にもタグを付与する。

その病変・症状が実際に認められたかどうかの程度(確実性; certainty)を属性として付与し, 以下の値をとる。

**positive:** その病変・症状が患者に実際に認められた場合

**suspicious:** 患者がその病変・症状に該当するのではないかと疑われている場合(鑑別疾患<sup>\*1</sup>として提案されているなど)

**negative:** 患者についてその病変・症状の存在が否定された場合

**general:** 患者に関係ない一般的な病変・症状に関する記述や, 将来生じる可能性のある病変・症状を表現している場合

#### XML タグ

```
<d certainty="{ "positive" or "suspicious" or "negative" or "general" }>...</d>
```

#### 1.3.1.1 エンティティ範囲の目安

原則として「の」などの助詞等を含まない, ひとまとまりの複合名詞が単位の目安である。複合名詞として表現される病変や症状名の一部には, 特徴・変化的な表現(→ 1.3.3, 1.3.4 節)や部位名(→ 1.3.2 節)を含む場合もあるが, 全体として病変・症状タグを与える。また, 「〇〇の変化」のような表現の場合, 助詞「の」も含めて病変・症状エンティティとみなすことがある(→ 1.3.4 節)。他のタグの方が妥当な場合など, 適切なタグの選択にあたっては, 当該タグの説明も参照されたい。

## 例

- (2) 一方<a> 右肺下葉 </a> の<d certainty="positive"> 濃度上昇 </d> は<c> わずかに減少 </c> している
- (3) <d certainty="positive"> 間質性肺炎 </d> の<c> 増悪傾向 </c>  
(「間質性」は特徴 [<f>タグ→ 1.3.3 節] を表しているようだが, これ以上細かくは分割しない [→ 1.4.1 節])

\*1 症状がよく似ていて区別(鑑別)が必要な病気のこと。患者が呈する症状を考慮した場合にいくつかの病気が候補に挙がることもあり, 正確な診断のためにはそれらの中から識別する必要がある。このとき候補に挙がる病気を鑑別疾患と呼ぶ。

- (4) `<d certainty="positive"> 右肺 S2 結節 </d>`

(「右肺 S2」は部位名だが、「結節」と合わせて複合名詞を形成しているため、一括でタグを付与)

■**万病辞書・万病検索の活用** 病変・症状タグ<d>をつけるか否かや、タグを付与すべき表現の範囲に迷う場合は、幅広く病名を収集した**万病辞書**<sup>\*2</sup>が参考になる。万病辞書内において「標準病名」や「高頻度」と記載されているエントリは基本的に<d>タグを付与する対象であると考えて良い。

### 1.3.1.2 確実性 (certainty) 属性に応じた手がかり表現

以下、certainty 属性を判断する際に手がかりになる表現とその例(太字)を、値ごとに示す。

■**positive** 「認めます」「見られます」「散在」「(+)」など、存在していることを示す表現。

例

- (5) `<a> 右肺 </a>` に`<f> 一様 </f>` `<d certainty="positive"> すりガラス影 </d>` が`<f> 散在 </f>`。
- (6) `<d certainty="positive"> 感染の合併 </d>` が**見られます**。
- (7) `<a> 左第 1 肋骨 </a>` に`<d certainty="positive"> 骨折後の変化 </d>` が**認められます**。  
(「骨折後の変化」という表現が一体となって病状を表しているので、助詞「の」を含むものの、変化タグに切り分けず、全体として疾患・症状タグを付与)

■**suspicious** 「疑います」「否定はできません」「可能性があります」「鑑別にあがる」「としても」「とすれば」「思われる」「考えられる」

例

- (8) `<d certainty="suspicious"> 急性肺炎 </d>` の`<c> 増悪傾向 </c>` を**疑いますが**
- (9) `<d certainty="suspicious"> NSIP 感染の合併 </d>` は**否定できません**。
- (10) `<a> 両側肺上葉 </a>` に`<d certainty="suspicious"> 炎症後結節 </d>` と**思われる**  
`<d certainty="positive"> 石灰化 </d>` が認められます。
- (11) `<d certainty="suspicious"> サルコイド結節 </d>` **としても**、なんとも`<f> 非典型的 </f>`です。

**注意が必要な語句「D / D」** Differential Diagnosis の略語で、この略語以下に記載される病変・症状などと鑑別してくださいという、別の医師への指示を示す。そのため、D/D 以下の病変・症状には suspicious を付与する。

例

- (12) `<d certainty="suspicious"> IPF/UIP </d>` を疑います  
(**D / D** `<d certainty="suspicious"> 慢性過敏性肺臓炎 </d>`)

<sup>\*2</sup> <https://sociocom.naist.jp/manbyou-dic/>

■negative 「認めません」「〇〇なし」「〇〇は消失」「(ー)」「〇〇術後」

例

- (13) `<d certainty="negative">小リンパ節腫大</d>` を指摘できません.
- (14) `<a>左肺尖部</a>` の `<d certainty="negative">気胸嚢胞</d>` は `<c>消失</c>` しています.
- (15) `<d certainty="negative">咽頭癌</d>` `<TIMEX3 type="MED">術後</TIMEX3>`.  
(症状がなくなる変化を表す表現 [「術後」は摘出術後を表すことが多い] があった場合にもこのタグを付与.)

■general 「〇〇について説明」「〇〇(に)注意」「〇〇予防」「副作用(副作用の可能性の種類として病変・症状が記載されている場合に付与)」

例

- (16) `<d certainty="general">IP</d>` についてパンフを用いて説明
- (17) `<d certainty="general">感染</d>` に注意するよう指導した.
- (18) `<d certainty="general">間質性肺炎</d>` にはいくつかのパターンがあり,  
`<d certainty="general">特発性間質性肺炎</d>` は `<d certainty="general">膠原病</d>` などによって起こる `<d certainty="general">二次性間質性肺炎</d>`,  
`<d certainty="general">過敏性肺炎</d>` などがあります.

### 1.3.1.3 注意が必要な表現

■TNM分類 がんについて、その進行度合いに応じた分類として「TNM分類」が国際標準として定められている。臨床テキストでは T, N, M の前後に数字やアルファベットを伴うという形で文末に出現することが多い。

TNM 分類の記号にも病変・症状タグを付与するが、certainty 属性は付与しない。

例

- (19) `<d certainty="positive">原発性肺癌</d>` と考えます。 `<d> c T 1 c </d>`  
`<a>縦隔・肺門</a>` に `<f>病的有意</f>` な `<d certainty="negative">リンパ節腫大</d>`  
を認めない。 `<d> N 0 </d>`

■存在しないこと・少ないことが疾患である場合 例えば食欲がないという症状は「食思不振」, 「食欲不振」, 「食欲なし」, 「食欲低下:あり」などという形式で出現するが、あるべきものがない・少ないことが疾患である場合には、否定的表現を含めた全体を病変・症状エンティティとし、certainty="positive" を付与する。

例

- (20) `<TIMEX3 type="DATE">同年12月</TIMEX3>` に、 `<d certainty="positive">下痢・食欲不振</d>` を認めたため `<m-key state="negated">オフエブ</m-key>` 中止.

■「UIP パターンの間質性肺炎」 頻出表現の一つに「UIP パターンの間質性肺炎」があるが、「UIP パターン (の)」を区別せず全体に<d>タグを付与する。

### 1.3.2 臓器・部位 (Anatomical parts)

臓器・部位をはじめとした場所を表す表現に付与する。「辺縁」「内部」のように抽象的な表現も含む。

#### XML タグ

```
<a>...</a>
```

#### 例

- (21) <a> 右肺 </a> に<d certainty="positive"> 小リンパ節 </d> が<f> 散在 </f>.
- (22) <a> 両側肺下葉 </a> <f> 優位 </f> に<f> 不均一 </f> な<d certainty="positive"> 濃度上昇 </d> が認められます.
- (23) <a> 左第 3 肋骨 </a> に<d certainty="positive"> 骨折後の変化 </d> が認められます.
- (24) <a> 両側肺下葉胸膜直下 </a> の<d certainty="positive"> 結節 </d>

「辺縁に沿って」と「辺縁」は示す部位が異なるので、「辺縁に沿って」まで含めて部位タグとする。また、「辺縁」の対象が隣接する場合はそれも含める。

#### 例

- (25) <a> 内部空洞の辺縁に沿って </a> <d certainty="positive"> 充実部分 </d> が認められます

原則として臓器・部位タグは上記のような(複合)名詞を単位として付与することを想定しているが、「〇〇術後」の「〇〇」に部位を表す表現が含まれる場合は、その表現にも付与する。詳細は 1.4.4 節を参照。

#### 1.3.2.1 注意が必要な表現

■略語 何を指すかわからない略語に出会った場合は、キーワードを含めて web 検索すると手掛かりが見つかることがある。ただし、すぐにはわからなければ保留タグ(→ 1.3.10 節)を付与して良い。

#### 例

- (26) <a> SVC </a> には<d certainty="negative"> 浸潤所見 </d> なし。  
(SVC は superior vena cava の略で、上大静脈を指す)

### 1.3.3 特徴・尺度 (Features and measurements)

病変・症状の特徴及び尺度・値・範囲・程度に関する修飾語句や述部を構成する語幹(「散見」する、など)に付与する。程度表現(「軽度」など)も含めるが、変化を表す表現と連続している場合は特徴・尺度タグには含めない(後述する変化エンティティ<c>に含める[→ 1.3.4])。「名詞+な/に...」の場合、活用部分の「な/に」は含めず、語幹だけで良い。ただし「徐々に」のような副詞は活用がない(全体を含める)ことに注意す

る。「～に」の形をした副詞は「～の/な」などと言い換えられないことで判別できる（例えば、「徐々の/な」とは言わない）。

#### XML タグ

<f>...</f>

#### 例

- (27) <f> 境界明瞭・辺縁平滑 </f> な<d certainty="positive"> 結節影 </d> が認められます。
- (28) <a> 縦隔内 </a> に<d certainty="positive"> リンパ節 </d> が<f> 散見 </f> されます
- (29) <f> 不均一 </f> な<d certainty="positive"> 濃度上昇 </d>
- (30) <f> 局在性 </f> の<d certainty="positive"> 均等影 </d>
- (31) <a> 両肺 </a> に<f> びまん性 </f> に<d certainty="positive"> 気管支拡張影 </d> が出現
- (32) <f> 径 3cm 以下 </f> の<d certainty="positive"> リンパ節腫大 </d> を<f> 散見 </f> する
- (33) <f> 縦隔条件で 1. 6 × 1. 6 cm </f>
- (34) <f> 有意なサイズ </f> の<d certainty="negative"> リンパ節腫大 </d> を認めません。
- (35) <f> Positive size </f> の<d certainty="negative"> LN </d> は認められません。
- (36) <f> 少量 </f> の<d certainty="positive"> 両側肺胸水 </d> あり。
- (37) <f> 軽度 </f> の<d certainty="positive"> 陰影 </d> を認めます。

なお、病変は存在するが、ある特徴を**持たない**ことを表す記述は、その否定部分も含める。病変・症状エンティティと異なり、属性を用いないことに注意する\*3。

#### 例

- (38) <a> 左気管傍縦隔 </a> に<d certainty="positive"> リンパ節 </d> あり。<f> 病的有意な大きさではありません </f>。

### 1.3.3.1 注意が必要な表現

■○○**優位** 特に読影所見では「右側優位」「両肺上葉優位」などの表現が見られるが、「優位」は常に特徴・尺度エンティティとして扱う。

#### 例

- (39) <a> 両肺上葉 </a> <f> 優位 </f> に<d certainty="positive"> 網状影 </d> あり。

\*3 名詞的表現を中心とする病変・症状と異なり、特徴・尺度は活用する形容詞・動詞などの述語的表現で記述される。病変・症状タグと同様に特徴・尺度タグにも事実性を導入した場合、(1)「非特異的でない」など二重否定的表現が出現した場合にタグ付け範囲と事実性の判断に混乱が生じやすく、(2)「顕著ではないリンパ節腫大を認める」などと病変・症状エンティティと共起した場合にかえって事実性の解釈が複雑になってしまう。そこで、特徴・尺度タグではむしろ冗長に否定部分を含めることとした。なお、同様の性質を持つ変化タグ（後述）も同じ理由で否定表現を含めたタグ付けをする。



### 1.3.4 変化 (Change)

増悪，減少など，症状の変化，および検査値や治療にかかわる変化を表す表現に付与する．変化がないことを表す場合（「著変なし」など）にもタグを付与する．変化の程度を表す修飾語句と変化を表す表現は一括りにしてタグを付与する（「わずかに増強」など）．

ただし「変化」という語句を使っても，「骨折後の変化」のように全体を病変・症状タグ<d>として扱うものがある．また，「収縮性変化」「炎症後変化」のように病変・症状と強く結びついて複合名詞化している場合にはこの<c>タグを使わず，複合名詞（句）全体を病変・症状タグ<d>とする（この場合に限って助詞「の」で連結された名詞句も認める→1.3.1）．1.4.4.1節でも言及する「術後変化」はその1例である．

#### XML タグ

<c>...</c>

#### 例

- (40) <d certainty="positive">収縮性変化</d> が<c> 強くなっています</c>.
- (41) 全体的に<f> 中程度</f> の<c> 増強</c> がみられます
- (42) <a> 左肺尖部</a> の<d certainty="negative">びまん性気胸腔</d> は<c> 消失</c> しています.
- (43) <d certainty="suspicious">急性肺炎</d> の疑い  
<TIMEX3 type="DATE"> 前回</TIMEX3> より<c> 少し増悪</c>
- (44) <a> 右肺下葉</a> の<d certainty="positive">濃度上昇</d> は<c> わずかに増強</c>

なお，変化して**いない**ことを表す否定的記述は，否定部分も含めて全体を<c>タグで囲む（特徴タグ [→ 1.3.3 節] と同様）．

#### 例

- (45) <TIMEX3 type="DATE"> 前回</TIMEX3> とは<c> 著変なし</c>.
- (46) <a> 気管分岐下、左右下部気管支傍</a> に<d certainty="positive">小リンパ節</d> を認めるが、<c> 経時的变化に乏しい</c>.
- (47) <d certainty="suspicious">良性結節</d> が疑われるものの、<c> 縮小はみられない</c>

### 1.3.5 時間表現 (TIMEX3)

診療記録内に記載されている日付，期間，「前回」といった何らかのイベントが生じたタイミングなどを表す表現に対し，時間表現タグ<TIMEX3>\*4を付与する．時間の性質に応じて以下の type 属性を与える．

**DATE:** 日歴に焦点をあてた日付表現

\*4 本ガイドラインでは，時間表現に関する先行アノテーション仕様である TimeML (version 1.2.1) で定義された TIMEX3 を導入した．ただし，後述するように一部の属性しか用いず，医療文書向けに若干の拡張を施した．具体的には，MED と MISC が本ガイドラインで拡張した属性である．

**TIME:** 一日のうちのある時点に焦点をあてた表現や不定の現在を表す「今」「現在」などの時刻表現

**DURATION:** 時間軸上の両端ではなく期間全体を表すことに焦点をあてた期間表現

**SET:** 複数の日付・時刻・期間に焦点をあてた頻度集合表現

**AGE:** 年齢に関する表現

**MED:** 「術後」など医療に特徴的な時間表現

**MISC:** 以上のどれにも該当しない場合

以下に示す例文のように、「3ヶ月後」「5年前」は一括りにして<TIMEX3>タグを付与するのに対して、「H 3 5 8 / 3 から」「H 3 4 8 / 1 1 より」などの助詞には<TIMEX3>タグを付与しない。

#### XML タグ

```
<TIMEX3 type="{DATE, TIME, DURATION, SET, AGE, MED, MISC}">...</TIMEX3>
```

#### 例

- (48) <a>縦隔や肺門のリンパ節</a> は<TIMEX3 type="DATE">前回</TIMEX3> と<c>ほぼ同様</c> です。
- (49) <TIMEX3 type="TIME">朝</TIMEX3>, <a>右眼瞼</a> <d certainty="positive">開きにくい</d>
- (50) <t-key>喫煙</t-key> : <t-val>20本/day</t-val> × <TIMEX3 type="DURATION">40年</TIMEX3>  
<TIMEX3 type="AGE">59歳</TIMEX3> で<t-val>禁煙</t-val>  
(「喫煙」が問診項目で「:」以降は問診結果であるが、入れ子構造を認めない[→1.2.1節] ことと、時間表現タグを他のタグより優先する[→1.4.2節] ことから、時間表現タグを付与したあとに残った語句に検査結果タグ<t-val>を付与する)
- (51) <TIMEX3 type="SET">頻回</TIMEX3> の<d certainty="positive">発熱</d>
- (52) <TIMEX3 type="MED">術後</TIMEX3> 階段・小走り程度の運動で<d certainty="positive">息切れ</d> のある状態

なお、時間表現タグ<TIMEX3>は以下の例のように「後」だけにタグを付与することはせず、「術後」「切除後」のような複合語単位で付与する。

#### 例

- (53) <m-key state="executed">nasal</m-key> : <m-val>1L/min</m-val> で<t-key>SpO2</t-key> : <t-val>94-95%</t-val> で開始し、<TIMEX3 type="MED">鎮静後</TIMEX3>, <p>酸素化低下</p> あったため<m-val>6L/min</m-val> 併用。

### 1.3.6 検査・問診 (Test)

一式の検査・問診を表す表現には<t-test>を、検査・問診項目名（もしくは医学的指標名）には<t-key>を、検査・問診結果（もしくは医学的指標）の値には<t-val>を付与する。ただし<t-val>は、「陰性」「陽性」など、数値でないものもある。

検査名<t-test>についてはさらに、検査の適用状態として以下の state 属性を付与する。

**scheduled:** 今後予定している（記録に記載された時点では実施されていない）場合

**executed:** すでに実施済みの場合

**negated:** 中止など、実施しないことになった場合

**other:** 上記以外の状態の場合

なお state 属性は<t-key>と<t-val>には原則付与しないが、特定の検査項目に限って適用状態が記載されている場合には、<t-key>に state 属性を付与して良い（→例 (58)）。

#### XML タグ

```
<t-test state="{scheduled, executed, negated, other}">...</t-test> — 検査・問診名
<t-key (state="{scheduled, executed, negated, other}")>...</t-key> — 検査・問診項目
もしくは医学的指標名
<t-val>...</t-val> — 検査・問診結果もしくは医学的指標の値
```

#### 例

(54) <t-test state="executed">胸部C T </t-test>

(55) <t-key> F E V 1 </t-key> : <t-val> 1 . 9 7 L ( 8 0 . 0 % ) </t-val>

検査結果の値にタグを付与する際、値の左側にスペースを挟んで記載されている「前」「L」「H」などにはタグを付与しない。

#### 例

(56) <t-key> K L - 6 </t-key> □□□<t-val> 5 5 9 </t-val> □H

（この例ではスペースを□で示している。）

(57) <t-key> c o n s </t-key> : <t-val> a l e r t </t-val>

検査項目は基本的に検査値が後述されると期待できる表現である。ただし、まだ検査自体が実施されていないなどの理由で、検査結果が記載されていない場合もある。文意から検査項目が今後実施される予定だと判断できる場合には、state 属性 "scheduled" を付与する。

#### 例

(58) <t-key state="scheduled"> 6 分間歩行 </t-key>

（SOAP の P 欄に複数の検査項目とともに記載されていたため、次回実施予定の検査項目と判断する [→ 1.4.5 節]。また、本項目は5分でも10分でもなく「6分間歩行」というひとつの定まった検査項目であるため、専門用語として検査タグ一括りにし、時間表現タグを付与しない）

また、検査項目と検査結果が並記されている箇所などで、病変・症状タグや時間表現タグを付与できる場合は、検査結果タグよりも病変・症状タグや時間表現タグを優先して付与する（1.4.2節を参照）。

#### 例

- (59) `<d certainty="positive">痰</d>` (+)  
 (問診結果の羅列に上記のような記述が見られることがある)
- (60) `<t-key>喫煙を始めた年齢</t-key> : <TIMEX3 type="AGE"> 20歳</TIMEX3>`  
`<t-key>やめた年齢</t-key> : <TIMEX3 type="AGE"> 75歳</TIMEX3>`

### 1.3.7 薬品 (Medicine)

薬品名には`<m-key>`を、処方量など薬品にかかわる値には`<m-val>`を付与する。検査タグと同様に、薬品名`<m-key>`についてはさらに、投薬の適用状態として以下の `state` 属性を付与する。

**scheduled:** 今後投薬を予定している（記録に記載された時点で投薬されていない）場合

**executed:** すでに投薬中の場合、

**negated:** 中止など、投薬しないことになった場合

**other:** 上記以外の状態の場合

なお、`<m-val>`にも `state` 属性を付与する場合がある（→例 (63)）。

#### XML タグ

`<m-key state="{scheduled, executed, negated, other}">...</m-key>` — 薬品名  
`<m-val (state="{scheduled, executed, negated, other}")>...</m-val>` — 薬品の値

#### 例

- (61) `<TIMEX3 type="DATE"> 明日</TIMEX3>` から `<m-key state="scheduled"> プルモケア</m-key>` `<m-val> 100mL / 1hr</m-val>` を投与開始予定。

薬品に係わる投与表現が複数ある場合、最終状態を優先とし、文末までの表現をみてタグを付与する。

#### 例

- (62) `<TIMEX3 type="DATE"> 2027 / 7</TIMEX3>` `<m-key state="negated"> ピレスパ</m-key>` `<m-val> 1200mg</m-val>` で `<d certainty="positive"> 食思不振</d>` が出現、**導入断念**

原則として薬品名タグ (`<m-key>`) のみに `state` 属性を付与するが、実施中（予定）の薬品について値を変更したり中止する記述が見られる場合には、薬品値タグ (`<m-val>`) に付与できる。

#### 例

- (63) `<m-key> プレドニン</m-key>` `<m-val state="executed"> 10mg</m-val>` から `<m-val state="scheduled"> 7.5mg</m-val>` へ `<c> 減量</c>`

また、検査・問診値エンティティ (<t-val>) の場合と同様に、病変・症状ないし時間表現タグが薬品名エンティティ (<m-key>) の値として適当な場合は、薬品値タグ (<m-val>) より優先して付与する。

#### 例

- (64) <m-key state="executed">経腸栄養 </m-key> を<TIMEX3 type="DATE">本日 </TIMEX3> から<TIMEX3 type="SET">1日3回 </TIMEX3> の予定であったが、<TIMEX3 type="DATE">昨日 </TIMEX3> までのペースの<TIMEX3 type="SET">1日1回朝 </TIMEX3> だけに

### 1.3.8 治療 (Remedy)

療法や手術、処置など、治療方法を表す表現にタグ<r>を付与する。検査・問診名エンティティ (<t-test>) や薬品名エンティティ (<m-key>) と同様に、治療の適用状態によって state 属性を与える。

**scheduled:** 今後治療を予定している (記録に記載された時点で実施されていない) 場合

**executed:** すでに実施中の場合

**negated:** 中止など、実施しないことになった場合

**other:** 上記以外の状態の場合

#### XML タグ

```
<r state="{scheduled, executed, negated, other}">...</r>
```

#### 例

- (65) <t-test state="executed">CAG </t-test> 施行も<r state="negated">PCI </r> は行わず。
- (66) <r state="executed">除細動 </r> のため<TIMEX3 type="DATE">2026/3/6-7 </TIMEX3> 循環器内科<cc state="executed">入院 </cc>。

治療エンティティ (<r>) とみなす表現は医薬品の名称に似たものも多い。対応するなんらかの値や頻度が併記されず、実施状況が述べられているような表現に<r>タグを付与する。逆に以下の例文のように、対応する値が併記されている場合は (治療法のような表現であっても) 薬品エンティティ (<m-key>, <m-val>) とみなす。

#### 例

- (67) <d certainty="positive">NHL </d> に対して上記<TIMEX3 type="cc">ケモ時 </TIMEX3> <m-key state="executed">ステロイド </m-key> <m-val>2.5mg </m-val> が入っているせいか、現在<m-key state="negated">咳止め </m-key> も全く内服せずに <d certainty="negative">咳 </d> も出ず<d certainty="negative">DOE </d> もない。

### 1.3.8.1 注意が必要な表現

■**薬品名+「療法」** 薬品名の後に「療法」が続く表現は、一律に全体を治療エンティティとする。例えば「ステロイド」にはふつう薬品名タグ<m-key>が付与されるが、「ステロイド療法」と記述されている場合は全体を治療タグ<r>とする。

■**ペースメーカー** ペースメーカー自体は医療機器のことだが、臨床テキスト内ではペースメーカー治療のことを示すため、<r>タグを付与する。

### 1.3.9 クリニカルコンテキスト (Clinical Context)

入院、退院、再入院、転院、来院、初診、経過観察（フォローアップ、f/u）、など、患者が医療施設との関係でどういった状態にあるかを示す表現にクリニカルコンテキストタグ<cc>を付与する。診断、紹介といった医師の行為については付与しない。実施状態を state 属性で区別する。

**scheduled:** 今後予定している（記録に記載された時点では実施されていない）場合

**executed:** すでに実施済みの場合

**negated:** 中止など、実施しないことになった場合

**other:** 上記以外の状態の場合

#### XML タグ

```
<cc state="{scheduled, executed, negated, other}">...</cc>
```

#### 例

- (68) 循環器内科<cc state="executed"> 通院中 </cc> であったが、  
(69) 近医<cc state="executed"> 受診 </cc> し<TIMEX3 type="DATE"> 3月末 </TIMEX3> 当院  
免疫内科<cc state="executed"> 紹介受診 </cc> となる。  
(70) <TIMEX3 type="DATE"> 今週末（8 / 27） </TIMEX3> の<cc state="scheduled"> 退院  
</cc> を予定。

### 1.3.10 保留 (Pending)

医療エンティティに相当しそうな医学的専門用語と思わしいが、判断がつかないものに対し、積極的に付与する。

#### XML タグ

```
<p>...</p>
```

#### 例

- (71) <a> 胸膜直下 </a> は<p> spare </p> されており  
(72) <a> 縦隔内 </a> に<f> 短径4cm程度 </f> までの<p> LNs </p> が認められます。

(73) <p> Coronary </p> に<d certainty="positive"> 網状影 </d>

## 1.4 医療エンティティアノテーションの注意事項

### 1.4.1 長さ優先タグ付け

1つの医療エンティティの長さは原則として、病変・症状<sup>\*5</sup>及び臓器・部位エンティティの場合に複合語程度、特徴・尺度エンティティの場合は病変・症状エンティティの修飾部および述部を構成する語幹、変化エンティティの場合には動詞句程度を単位とする。

とくに複合語の場合、細分化してタグ付けせずに、主たる医学的意味を考慮して全体を1種類のタグとして扱うようにする。例えば「脳転移」は短く区切れれば「<a 脳 </a> <d certainty="positive"> 転移 </d>」と解釈できるが、本アノテーションにおいては「<d certainty="positive"> 脳転移 </d>」と全体を<d>タグとみなす。

文脈により様々な場合があるので、**作業者が迷った場合は長めの範囲でまとめてタグ付けを行う方針**とする。例えば、以下の例文に見られる「細菌・ウイルス・ニューモシスチスのような感染症の合併」も、全体を病変・症状エンティティとみなして良い（→1.2.2節も参照）。例外として、必ず決まった医療エンティティとしてアノテーションすべき表現がある（「優位」は必ず特徴・尺度エンティティとする [→1.3.3.1節]、など）ことには注意する。

#### 例

(74) <d certainty="negative"> 骨転移を疑う所見 </d> はありません。

(75) <a # 4 L縦隔 </a> に<d certainty="positive"> リンパ節 </d> あり、<f> 短径は1.9 cm </f> です。

（ここでの「リンパ節」は部位としてではなく「リンパ節腫大」という病変の短縮形と考えられるため<d>タグを付与）

(76) <d certainty="suspicious"> 間質性肺炎 </d> の<c> 増悪 </c> や  
<d certainty="suspicious"> 細菌・ウイルス・ニューモシスチスのような感染症の合併 </d>  
が疑われます。

### 1.4.2 病変・症状および時間表現タグの優先

ある語句に対して複数のタグを取りうる場合においては、病変・症状タグ<d>と時間表現タグ<TIMEX3>が優先される。

<sup>\*5</sup> 病変・症状タグを付与する範囲の判断には、1.3.1.1節でも示したように、万病辞書を参考にできる。

### 1.4.3 誤字・脱字の訂正 (correction 属性)

タグを付与した表現内に誤記と思われる表現が見られた場合は、タグに `correction` 属性を加え、正しいと思われる表記を `correction` 属性の値として記入する。タグを付与していない箇所にある誤記については対応しなくてもよい。また、高度な医学的知識がなければ判断できない誤記も対応不要である。

本文に「すりが荒らす影」という表現があり、これを「すりガラス影」の誤表記だと判断できた場合のタグ付け例を以下に示す。

例

```
(77) <d certainty="positive" correction="すりガラス影"> すりが荒らす影 </d>
```

### 1.4.4 「術後」をめぐる表現

臨床テキストでは「術後」を含む表現が頻出するが、場合により付与すべきタグの種類が異なる。

#### 1.4.4.1 「(〇〇) 術後変化」の場合

これ以降に紹介する場合に優先して、「術後変化」の部分に `<d>` タグを付与する。「胃癌術後変化」や「右腎切除術後変化」など、「〇〇術後変化」の形式で前置する表現（「〇〇」部分）は、後述する「〇〇術後」のケースと同様に扱う。

例

```
(78) <a> 直腸 </a> に <d certainty="positive"> 術後変化 </d> あり。
```

#### 1.4.4.2 「術後」の前後に医療エンティティがある場合

■〇〇術後 「〇〇術後」のように複合名詞の最後に「術後」が出現し、全体として「〇〇という手術の後」ないし「〇〇に対する手術の後」という意味になる場合は、「`<d certainty="negative"> 胃癌 </d> <TIMEX3 type="cc"> 術後 </TIMEX3>`」のように分割して、それぞれに適したタグ付けをする。特にこの例のように、胃癌の摘出術がなされた後の時点を「術後」が含意する場合、直前の病変・症状エンティティは `certainty="negative"` とすべきことに注意する。

例

```
(79) <r state="executed"> 胸部左肺上葉切除 </r> <TIMEX3 type="cc"> 術後 </TIMEX3>
```

■(△△) 術後〇〇 「術後」が複合名詞の最初または中間に出現する場合も、「術後」を時間表現エンティティとし、分割前後の表現にはそれぞれ適切なタグを付与する。ただし前述の通り、続く表現が「変化」の場合は「術後変化」を病変・症状エンティティとすることに注意する（→ 1.4.4.1 節）。



#### 例

- (80) `<TIMEX3 type="cc"> 術後 </TIMEX3> <d certainty="positive"> 左肺胸水 </d>`
- (81) `<a> 胃 </a> <TIMEX3 type="cc"> 術後 </TIMEX3> □ <d certainty="positive"> 局所再発 </d>`  
(全角スペース□が間に入る場合も同様)

### 1.4.5 SOAP セクションと医療エンティティ属性

対象テキストが診療記録である場合、SOAP というフォーマットにしたがって記述されていることが多い。SOAP とは

- S (Subject):** 主観的データ。患者の話など  
**O (Object):** 客観的データ。身体診察・検査データなど  
**A (Assessment):** データの評価  
**P (Plan):** データと評価をもとにした今後の治療方針

を意味し、これら 4 つのセクションに診療記録を分けて記述する方針のことである。この SOAP に加えて **I (Informed consent; 患者への直接の説明事項)** が追加されることもある。

例えば P 欄には今後の治療方針 (Plan) が記載されているため、基本的に state 属性を scheduled とする。病変・症状タグの場合は、予定を記載している文中に一般的な表現として出てくる場合に certainty="general" を付与する。時間表現タグを付与する際には、メタデータとしての記載日も確認しながら判断する。

以下に、SOAP セクションに応じたアノテーション例を示す。

#### 例

- (82) `<d certainty="positive"> I P ( I P F , N S I P , o t h e r s ) </d>`  
(この文が A 欄に出現した場合、「各種データから総合して IP=間質性肺炎と評価できる [ただし IPF=特発性肺線維症や NSIP=非特異性間質性肺炎、あるいはその他の可能性もあるが]」という意味になるため、certainty="positive" とする)
- (83) `<TIMEX3 type="DATE"> 9 / 2 7 </TIMEX3> <t-test state ="scheduled"> C T </t-test>`  
(この文が P 欄に出現した場合、この臨床テキストの記載日以降の 9 月 27 日に CT を実施予定という意味なので、state="scheduled" を付与する)
- (84) `<d certainty="general"> I P </d>` についてパンフを用いて説明  
(I 欄でよく見かけられる文で、患者ないし患者家族に IP (間質性肺炎) の一般的な説明をしたと解釈できることから、certainty="general" を付与する。)

## 1.5 タグ付け文書例

文書全体に本アノテーションを実施した例を示す。

例

- (85) <a> 両側胸膜直下 </a> の<a> 肺底部 </a> に<a> 右側 </a> <f> 優位 </f> に  
 <d certainty="positive"> 網状影や牽引性気管支拡張像 </d> を認め、  
 <d certainty="negative"> 蜂巣肺 </d> は認めません。 <p>Possible UIP pattern</p> を疑  
 います。 <a> 右肺 </a> に<d certainty="positive"> 斑状すりガラス影 </d> が<f> 散在  
 </f>。一連の<d certainty="positive"> 間質性肺炎に伴う変化 </d> で矛盾ありません。 <f>  
 有意 </f> な<d certainty="negative"> リンパ節腫大 </d> を指摘できません。  
 <d certainty="negative"> 胸水貯留 </d> を認めません。 <a> 大動脈 </a> の  
 <d certainty="positive"> 石灰化 </d> を認めます。 <a> 甲状腺両葉 </a> に  
 <d certainty="positive"> 小 LDA </d> を認めます。 <d certainty="positive"> 両腎囊  
 胞 </d> あり。  
 診断 <d certainty="suspicious"> 間質性肺炎 </d> の疑い
- (86) <a> 右側肺底部 </a> に<a> 右側 </a> <f> 優位 </f> に<d certainty="positive"> 網状影や  
 牽引性気管支拡張像 </d> を認めます。 <a> 胸膜直下 </a> は<p>spare</p> されており、  
 <d certainty="suspicious"> NSIP pattern </d> を疑います。 <a> 右肺 </a> に  
 <d certainty="positive"> 斑状すりガラス影 </d> が<f> 散在 </f>。一連の  
 <d certainty="positive"> 間質性肺炎に伴う変化 </d> で矛盾ありません。  
 <TIMEX3 type="DATE"> 前回 </TIMEX3> より少し<d certainty="positive"> 陰影増強 </d>  
 を認めます。  
 <a> 縦隔内 </a> に<d certainty="positive"> 小リンパ節 </d> を<f> 散見 </f> しますが、  
 <f> 有意なサイズ </f> の<d certainty="negative"> リンパ節腫大 </d> を指摘できません。  
 <d certainty="negative"> 胸水貯留 </d> を認めません。 <a> 大動脈 </a> の  
 <d certainty="positive"> 石灰化 </d> を認めます。 <a> 甲状腺両葉 </a> に  
 <d certainty="positive"> 小 LDA </d> を認めます。 <TIMEX3 type="DATE"> 前回  
 </TIMEX3> と<c> 著変なし </c>。  
 診断 <d certainty="suspicious"> 間質性肺炎 </d> の疑い <TIMEX3 type="DATE"> 前回  
 </TIMEX3> より<c> 少し増悪 </c>
- (87) S) <t-key>MRC</t-key> <t-val>grade 3</t-val>  
 O) <t-key>SpO2</t-key> <t-val>90% (RA)</t-val>  
 <t-test state="executed"> 在宅パルスオキシメーター </t-test> (  
 <TIMEX3 type="CC"> 睡眠時 </TIMEX3>)  
 <t-key> 平均 SpO2</t-key> <t-val>89%</t-val>  
 <t-key>SpO2 度数分布 (&< 90%) </t-key> <t-val>62.11%</t-val>  
 A) #<d certainty="positive"> 間質性肺炎 (UIP-like) </d>  
 <TIMEX3 type="DATE"> 2 年前頃 </TIMEX3> から<d certainty="positive"> 労作性呼  
 吸困難 </d> <c> 出現 </c>、<t-key>Hugh-Jones</t-key> <t-val>II</t-val>。

<TIMEX3 type="DATE">その後</TIMEX3> <f>徐々に</f> <c>悪化</c>、  
 <TIMEX3 type="TIME">現在</TIMEX3> は<t-key>MRC</t-key> <t-val>grade  
 3</t-val>  
 <t-test state="executed">精密呼吸機能</t-test>  
 <TIMEX3 type="DATE">2013.5.30</TIMEX3>  
 <t-key>VC</t-key> <t-val>2.27 (120.3%)</t-val>  
 <t-key>FEV1.0</t-key> <t-val>1.58 (120.2%)</t-val>  
 <t-key>FRC</t-key> <t-val>2.35 (159.8%)</t-val>  
 <t-key>DLCO</t-key> <t-val>3.53 (30.2%)</t-val>  
 <TIMEX3 type="DATE">2010.8.3</TIMEX3> <a>胸部</a>  
 <t-test state="executed">CT</t-test>  
 <d certainty="positive">肺気腫</d> (<a>上葉優位</a>) +  
 <d certainty="positive">UIP pattern</d> (<a>下葉優位</a>)  
 <TIMEX3 type="DATE">2010.8.16</TIMEX3>～<r state="executed">ピレスパ</r>  
 (<d certainty="positive">眼瞼皮膚炎</d> にて中止)  
 <TIMEX3 type="DATE">2010.12.27</TIMEX3> より<r state="executed">ピレスパ  
 </r>再開、<TIMEX3 type="DATE">2012.5.7</TIMEX3> 中止 (  
 <d certainty="positive">皮膚掻痒</d>)  
 <TIMEX3 type="DATE">2011.8.29</TIMEX3> 血清  
 #<d certainty="positive">高脂血症</d>  
 【<t-test state="executed">既往歴</t-test>】  
 <d certainty="positive">B型急性肝炎</d> (<TIMEX3 type="AGE">32歳  
 </TIMEX3>)  
 <t-key>喫煙</t-key> <t-val>20本</t-val> <TIMEX3 type="DATE">5年前  
 </TIMEX3> から<t-val>禁煙</t-val>  
**P)** <r state="scheduled">HOT</r> 導入<cc state="scheduled">入院</cc> 申し込み

# 謝辞

本ガイドラインの策定にあたっては、官民研究開発投資拡大プログラム PRISM の支援を受けた。

本ガイドラインの医療エンティティアノテーションの素案作成に携わった城綾実氏と伊藤薫氏に感謝申し上げます。また、医薬基盤研究所の伊藤真里氏と黒田正孝氏，国立がんセンターの小林和馬氏，並びに東京大学附属病院の中村優太氏には，医療専門家として多くの助言をいただき，品質向上に寄与していただいた。ここにお礼申し上げます。