

臨床テキストアノテーションガイドライン

矢田 竣太郎 荒牧 英治 田中 リベカ Fei Cheng 黒橋 禎夫
NAIST 京都大学

Ver. 8 β
2021年6月29日

目次

第 1 章	概要	1
1.1	目的	1
1.2	表記方法	1
1.3	注意事項	1
第 2 章	医療エンティティアノテーション	2
2.1	概要	2
2.2	基本方針	2
2.2.1	入れ子構造の禁止	2
2.2.2	並列表現の一括タグ付け	2
2.3	医療エンティティの種類	3
2.3.1	病変・症状 (Diseases and symptoms)	3
2.3.1.1	エンティティ範囲の目安	3
2.3.1.2	確実性 (certainty) 属性に応じた手がかり表現	4
2.3.1.3	注意が必要な表現	5
2.3.2	臓器・部位 (Anatomical parts)	6
2.3.2.1	注意が必要な表現	6
2.3.3	特徴・尺度 (Features and measurements)	6
2.3.3.1	注意が必要な表現	7
2.3.4	変化 (Change)	8
2.3.5	時間表現 (TIMEX3)	8
2.3.6	検査・問診 (Test)	10
2.3.7	薬品 (Medicine)	11
2.3.8	治療 (Remedy)	12
2.3.8.1	注意が必要な表現	13
2.3.9	クリニカルコンテキスト (Clinical Context)	13
2.3.10	保留 (Pending)	13
2.4	医療エンティティアノテーションの注意事項	14
2.4.1	長さ優先タグ付け	14
2.4.2	病変・症状および時間表現タグの優先	14
2.4.3	誤字・脱字の訂正 (correction 属性)	15
2.4.4	「術後」をめぐる表現	15

2.4.4.1	「(〇〇) 術後変化」の場合	15
2.4.4.2	「術後」の前後に医療エンティティがある場合	15
2.4.5	SOAP セクションと医療エンティティ属性	16
2.5	タグ付け文書例	16
第 3 章	医療エンティティ間の関係アノテーション	19
3.1	概要	19
3.2	基本関係	19
3.2.1	changeSbj 関係	20
3.2.2	changeRef 関係	21
3.2.3	featureSbj 関係	22
3.2.4	subRegion 関係	24
3.2.5	keyValue 関係	25
3.2.6	pending 関係	26
3.2.7	医療エンティティと基本関係の対応	26
3.3	時間関係	27
3.3.1	timeOn 関係	27
3.3.2	timeBefore 関係	30
3.3.3	timeAfter 関係	30
3.3.4	timeBegin 関係	31
3.3.5	timeEnd 関係	31
3.3.6	時間表現間の時間関係	32
3.3.7	医療エンティティの state 属性と時間関係	33
3.3.8	臨床テキストの種類と時間関係アノテーション	35
3.4	関係アノテーションの省略	35
3.4.1	基本関係 keyValue の自明なケース	35
3.4.2	推論関係が自明なケース	36
3.4.3	時間関係 timeOn の省略	37

第 1 章

概要

1.1 目的

本文書は、電子カルテや読影所見といった臨床現場で生産されるテキスト（**臨床テキスト**）から情報を抽出するために、テキストへの情報付与（**アノテーション**）について仕様を定めたガイドラインである。臨床テキストの種類によらず汎用的に適用できる標準ガイドラインを目指す。

本アノテーションの対象は、病名・疾患名や部位名といった医学的概念を表す言語表現であり、簡単に**医療表現**と呼ぶ。アノテーションは大きく 2 つに分かれる。

医療エンティティアノテーション 医療表現の範囲に対して、病名・疾患名、部位名といった医学的概念の種類に応じてタグを付与することで、**医療エンティティ**（医療表現の一事例）として識別する

医療エンティティ間の関係アノテーション 病名・疾患名と部位の関係、検査項目と検査値の関係といった、医療エンティティ同士の関係を付与する

定義から分かる通り、エンティティ間関係アノテーションはエンティティアノテーションを前提とする。したがって、本ガイドラインは第 2 章で医療エンティティアノテーションを、第 3 章で医療エンティティ間の関係アノテーションを説明する。

1.2 表記方法

本ガイドラインでは、アノテーション仕様の表現として、便宜的に XML タグ形式（<Tag>...</Tag>）を用いる。仕様として定義する事項を表現できる記述形式であれば、実際のアノテーション作業は XML 形式以外で実施しても良い。

1.3 注意事項

本ガイドラインで使用している例文は、実際の医学的現象や医療行為に正確には合致しないような表現を含む。

第2章

医療エンティティアノテーション

2.1 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。

2.2 基本方針

ここでは、後述する医療エンティティ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、医療エンティティの種類を踏まえた注意点を2.4節でさらにまとめてあるので、合わせて確認されたい。

2.2.1 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の内部に別の医療エンティティ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 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であるか、または数種の医療エンティティの連続という形式だけが許容される。

2.2.2 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

2.3 医療エンティティの種類

2.3.1 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^{*1}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>
```

2.3.1.1 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(→ 2.3.3, 2.3.4 節)や部位名(→ 2.3.2 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(→ 2.3.4 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 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→ 2.3.3 節] を表しているようだが, これ以上細かくは分割しない [→ 2.4.1 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**^{*2}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。

2.3.1.2 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 ` 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 ` 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 ` 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>`)

^{*2} <https://sociocom.naist.jp/manbyou-dic/>

■negative 「認めません」「〇〇なし」「〇〇は消失」「(-)」「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` 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.

2.3.1.3 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。また、後述する関係アノテーション(第3章)では TNM 分類に関係を付与しない。

例

- (19) `<d certainty="positive">原発性肺癌</d>` と考えます。 `<d>CT1c</d>`
`<a>縦隔・肺門` に`<f>病的有意</f>` な`<d certainty="negative">リンパ節腫大</d>`
を認めない。 `<d>NO</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。

2.3.2 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 に<d certainty="positive"> 小リンパ節 </d> が<f> 散在 </f>.
- (22) <a> 両側肺下葉 <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 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 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 <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2.4.4節を参照。

2.3.2.1 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(→2.3.10節)を付与して良い。

例

- (26) <a> SVC 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)

2.3.3 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[→2.3.4])。「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 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 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 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>。

2.3.3.1 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 <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。

2.3.4 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→2.3.1）。2.4.4.1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 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 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 [→2.3.3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 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>

2.3.5 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 1.21) 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 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 <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 [→ 2.2.1 節] ことと、時間表現タグを他のタグより優先する [→ 2.4.2 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">n a s a l</m-key> : <m-val>1 L/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>6 L/min</m-val> 併用。

2.3.6 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 [→ 2.4.5 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（2.4.2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>`

2.3.7 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

2.3.8 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。

2.3.8.1 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。

2.3.9 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。

2.3.10 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 は<p> spare </p> されており

(72) <a> 縦隔内 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>

2.4 医療エンティティアノテーションの注意事項

2.4.1 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^{*5}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 <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（→2.2.2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 [→2.3.3.1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 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。

2.4.2 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。

^{*5} 病変・症状タグを付与する範囲の判断には、2.3.1.1節でも示したように、万病辞書を参考にできる。

2.4.3 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>
```

2.4.4 「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。

2.4.4.1 「(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。
```

2.4.4.2 「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（→ 2.4.4.1 節）。

例

- (80) `<TIMEX3 type="cc"> 術後 </TIMEX3> <d certainty="positive"> 左肺胸水 </d>`
- (81) `<a> 胃 <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)

2.4.5 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。)

2.5 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 の<a> 肺底部 に<a> 右側 <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 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 の
<d certainty="positive"> 石灰化 </d> を認めます。 <a> 甲状腺両葉 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 に<a> 右側 <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 は<p>spare</p> されており、
<d certainty="suspicious"> NSIP pattern </d> を疑います。 <a> 右肺 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 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 の
<d certainty="positive"> 石灰化 </d> を認めます。 <a> 甲状腺両葉 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
 <t-test state="executed">CT</t-test>
 <d certainty="positive">肺気腫</d> (<a>上葉優位) +
 <d certainty="positive">UIP pattern</d> (<a>下葉優位)
 <TIMEX3 type="DATE">2010.8.16</TIMEX3>～<r state="negated">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

第3章

医療エンティティ間の関係アノテーション

3.1 概要

この章では、医療エンティティ間の関係アノテーションについて説明する。関係を考慮する医療エンティティは第2章で定義した通りである。これら医療エンティティは、テキスト中で相互に関係を持つ。たとえば、以下の文中には「骨折後の変化」という病変・症状エンティティ (<d>) と、「左第1肋骨」という臓器・部位エンティティ (<a>) が出現する。これらは、「骨折後の変化」が「左第1肋骨」という場所に位置する、という関係にある。

`<a> 左第1肋骨 ` に `<d certainty="positive"> 骨折後の変化 </d>` が認められます。

このように、臨床テキスト中の医療エンティティについて、その間に成立する関係をアノテーションする。

1つの関係は2つの医療エンティティの間で成立し、方向によって意味が異なる。そこで、ある医療エンティティ E_{from} からもう一つの医療エンティティ E_{to} への何らかの関係 “rel” を “rel(E_{from} , E_{to})” のように記述する。

本ガイドラインでは、医療エンティティ間の関係を**基本関係**と**時間関係**に分けて定義する。次節より、それぞれにおいて付与する関係の種類を説明する。

3.2 基本関係

医療エンティティ間の関係のうち、時間表現以外の医療エンティティが中心となるものを、時間関係と区別して**基本関係**と呼ぶ。基本関係は changeSbj, changeRef, featureSbj, subRegion, keyValue の5種類のタイプに分けられる。以下では、基本関係のXML表記に<brel>タグ(空要素)を使用する。

XML タグ

```
<brel id="rk" from="ei" to="ej" reltype="{types}" />
```

ここで、idは基本関係のID(識別子)、fromとtoはそれぞれ E_{from} と E_{to} のエンティティIDを明示するための属性である^{*1}。エンティティIDは、医療エンティティのみを扱う範囲においては必須ではないが、

^{*1} IDの形式は、個々の関係とエンティティを一意に識別できればどのような形式でも良い。本ガイドラインでは、関係IDを“rk”，エンティティIDを“ei”のような形式(k, i は自然数)で表す。数字は便宜的に、文頭からの出現順とした。

エンティティ間関係の付与においては必要となる。本章冒頭の例では、以下のようにエンティティ ID が付与される。

`左第1肋骨` に `<d id="e2" certainty="positive">骨折後の変化</d>` が認められます。

次節から各基本関係タイプを詳述するが、その中ではアノテーション例を XML タグ形式および矢印を用いた視覚的な図でそれぞれ示す。図の形式では、枠で囲まれた表現が医療エンティティ*2 を、 E_{from} から E_{to} への矢印が関係を表現する。ただしこれらの例では、見やすさのために、着目している以外の関係タイプについては省略する。実際には表示されている以上の関係を付与すべき事例も含まれていることに注意されたい。

3.2.1 changeSbj 関係

基本関係 $\text{changeSbj}(E_{\text{from}}, E_{\text{to}})$ は、 E_{from} が変化エンティティ (`<c>`) であるときに成立し、対象 E_{to} が E_{from} のように変化することを表す。

XML タグ

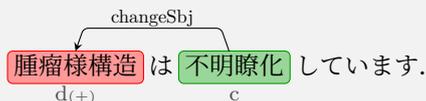
```
<brel id="rk" from="ei" to="ej" reltype="changeSbj" />
```

下記の例 (90) のように、変化の対象 E_{to} は文をまたいで出現しうるので、注意する。

例

(88) `<d id="e1" certainty="positive">腫瘍様構造</d>` は `<c id="e2">不明瞭化</c>` しています。

```
<brel id="r1" from="e2" to="e1" reltype="changeSbj" />
```



(89) `肝内胆管` の `<c id="e2">拡張</c>` がみられます。

```
<brel id="r1" from="e2" to="e1" reltype="changeSbj" />
```

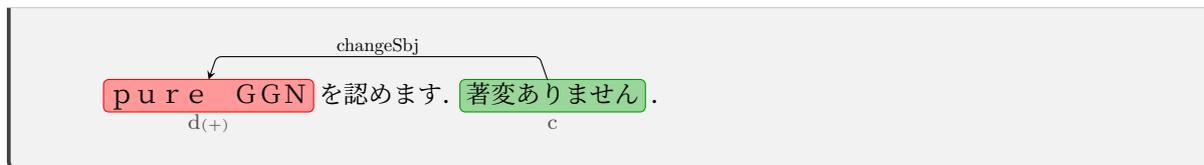


(90) `<d id="e1" certainty="positive">pure GGN</d>` を認めます。 `<c id="e2">著変ありません</c>`。

```
<brel id="r1" from="e2" to="e1" reltype="changeSbj" />
```

*2 医療エンティティタグは枠下に付記する。また、その属性を括弧内に以下の記号を用いて表記する。

	+	?	-	*
certainty 属性	positive	suspicious	negative	general
state 属性	executed	scheduled	negated	other

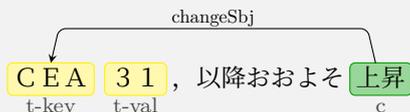


なお、検査に関する変化が述べられている場合には、検査・問診結果エンティティ (<t-val>)ではなく検査・問診名エンティティ (<t-test>) や検査・問診項目エンティティ (<t-key>) を変化の対象 E_{to} とする。検査・問診結果は変化の前後で計測されるものであり、変化の主体そのものではないためである。ただし、該当する検査・問診名エンティティ (<t-test>) や検査・問診項目エンティティ (<t-key>) が存在しない場合には、検査・問診結果エンティティ (<t-val>) を E_{to} とする。

例

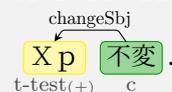
(91) <t-key id="e1"> C E A </t-key> <t-val id="e2"> 3 1 </t-val>, 以降おおよそ
<c id="e3"> 上昇 </c>

<brel id="r1" from="e3" to="e1" reltype="changeSbj" />



(92) <t-test id="e1" state="executed"> X p </t-test> <c id="e2"> 不変 </c>.

<brel id="r1" from="e2" to="e1" reltype="changeSbj" />



3.2.2 changeRef 関係

基本関係 $\text{changeRef}(E_{\text{from}}, E_{\text{to}})$ は、 E_{from} が変化エンティティ (<c>) であるときに成立する関係であり、 E_{from} が表す変化の比較対象・基準点が E_{to} であることを表す。

XML タグ

```
<brel id="rk" from="ei" to="ej" reltype="changeRef" />
```

E_{from} と E_{to} が同一文中に存在し、比較対象・基準点が明示される場合にのみ関係を付与する。比較を明示する表現には、「～より」や「～と比較すると」などがある（手掛かり表現を太字で示す）。

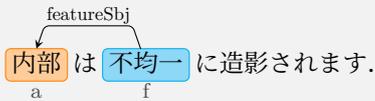
例

(93) <TIMEX3 id="e1" type="DATE"> 2003年9月 </TIMEX3> **よりは** <c id="e2"> 著変ありません </c>.

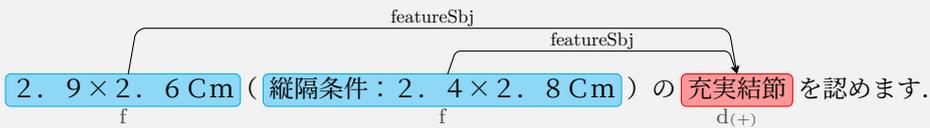
<brel id="r1" from="e2" to="e1" reltype="changeRef" />



(98)



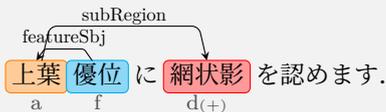
(99)



「左肺優位 (に)」などの表現にみられる「優位」という特徴・尺度エンティティは、必ず直前の臓器・部位エンティティ (<a>) に対して基本関係 featureSbj をもつ (→例 (100)).

例

(100)



特に読影所見において「境界が明瞭である」という言い回しが頻出する。「境界が明瞭 (な)」という構文の場合は、例 (101) の R3 のように「明瞭」を「正常肺野との境界」と関係づける*4。

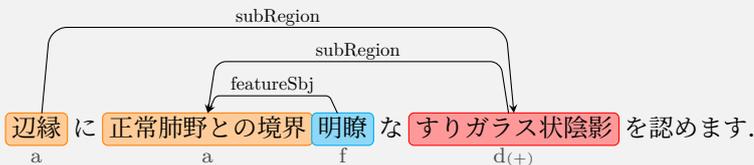
一方、「境界明瞭 (な)」という一語化した表現も存在する。例 (102) の featureSbj 関係のように「正常肺野との境界 (が) 明瞭」と助詞を補って解釈できる場合は「明瞭」を「正常肺野との境界」と関係づける一方、例 (103) のように「境界明瞭」全体が特徴・尺度エンティティ (<f>) の場合は直後の「結節」と関係づける。

例

(101)

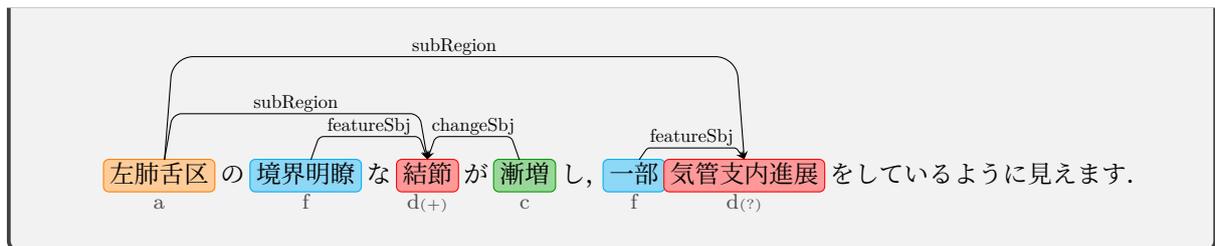


(102)



(103)

*4 「境界」は「AとBの境界」というように2つの領域の間を表すため、意味的には2つの項を要求する表現である。そのため、「正常肺野との境界」は、後述する基本関係 subRegion を付与すべき「相対的な部位表現」である。ここでは「すりガラス影」と「正常肺野との境界」を意味するので、「すりガラス影」から「正常肺野との境界」に基本関係 subRegion を付与する。



3.2.4 subRegion 関係

基本関係 subRegion(E_{from} , E_{to}) は、 E_{from} が表す場所（位置・範囲）に E_{to} が含まれることを表す。

XML タグ

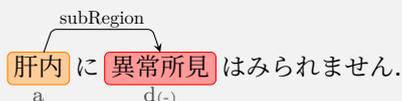
```
<brel id="rk" from="ei" to="ej" reltype="subRegion" />
```

E_{from} と E_{to} の内容によって、基本関係 subRegion は以下のようなケースを含む。

- ある部位に存在する病変への言及 (E_{from} = 臓器・部位エンティティ (<a>), E_{to} = 病変・症状エンティティ (<d>))
 - ある部位の中に存在する、より下位の部位への言及 (E_{from} と E_{to} はともに臓器・部位エンティティ)
 - ある病変の特定の部位への言及 (E_{from} = 病変・症状エンティティ, E_{to} = 臓器・部位エンティティ)
- いずれの場合も「 E_{from} が E_{to} を含む」という包含関係に注意が必要である。

例

(104)



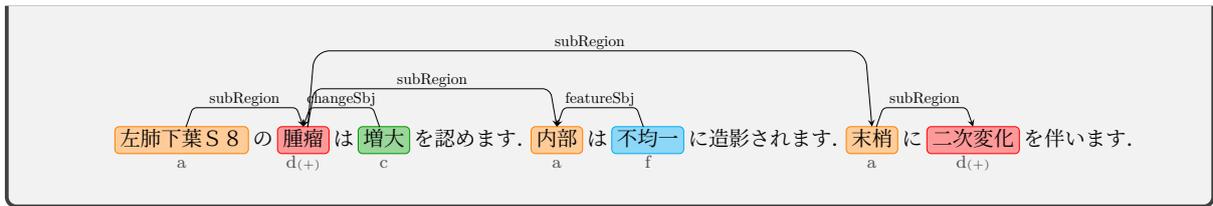
(105)



臓器・部位エンティティには「左肺上葉」といった具体的な臓器・部位のほかに、相対的な表現が含まれる。「内部」「辺縁」「境界」「末梢」「左」といった相対的部位表現は「何の内部なのか」「何の辺縁なのか」といった情報に依存することから、「何の」にあたる表現を E_{from} として相対表現 E_{to} に向けて基本関係 subRegion を付与する。このとき、 E_{from} が同一文内になくても文間をまたいで関係を付与することとする。例 (106) の関係 ID r3 と r6 は、相対表現が「内部」「末梢」の場合である。

例

(106)



特に「境界」という相対表現は、読影所見において「～との境界明瞭な」という形式で頻出する。その場合のアノテーションについては featureSbj 関係と合わせて注意が必要なため、前節の例 (101) や例 (102) を参照されたい。

3.2.5 keyValue 関係

基本関係 keyValue(E_{from} , E_{to}) は、 E_{from} が検査エンティティ (<t-test>), 検査項目エンティティ (<t-key>), 薬品エンティティ (<m-key>) のいずれかであるときに成立する関係であり、 E_{to} が E_{from} の値 (検査値や検査結果, 量) であることを表す。 E_{to} は、典型的には対応する <t-val> タグや <m-val> タグである。

XML タグ

```
<brel id="rk" from="ei" to="ej" reltype="keyValue" />
```

E_{from} と E_{to} は、下記の例 (107) のように検査エンティティないし薬品エンティティの 1 対 1 ペアで出現することが多い (<t-key>→<t-val>, <m-key>→<m-val>). その場合に限り、関係付与を省略できる (→ 3.4.1 節).

例

(107)

2027年03月10日
TIMEX3(DATE)

keyValue

WBC (白血球数) 10.72 H 前
t-key t-val

keyValue

Hb (ヘモグロビン) 10.4 L 前
t-key t-val

(108)

keyValue

喫煙を始めた年齢 : 20歳
t-key TIMEX3(AGE)

keyValue

やめた年齢 : 75歳
t-key TIMEX3(AGE)

ただし、1つの検査項目エンティティ (<t-key>) に対して、対応する検査値エンティティ (<t-val>) が2つ以上存在する場合、keyValue 関係を必ず付与する。

例

(109)



また E_{to} に病変・症状エンティティ (<d>) や時間表現エンティティ (TIMEEX3) などが対応する場合もあり、このような場合も keyValue 関係付与は必須である。

例

(110)



3.2.6 pending 関係

基本関係 pending(E_{from} , E_{to}) は、2つの医療表現間に何らかの関係があると考えられるものの、判断を保留する場合に用いる。

XML タグ

```
<brel id="rk" from="ez" to="ej" reltype="pending" />
```

3.2.7 医療エンティティと基本関係の対応

一部の医療エンティティには、ほとんどの場合に E_{from} となって何らかの関係を付与することとなる。 E_{from} にあたる医療エンティティの種類ごとにまとめる。

変化エンティティ (<c>) 変化の主体となる医療エンティティに対して基本関係 changeSbj を付与する。

この関係は文間をまたぐ場合にも付与する。また、同一文内に明示的な比較表現が出現する場合には、比較の対象・基準点に対して基本関係 changeRef も付与する。

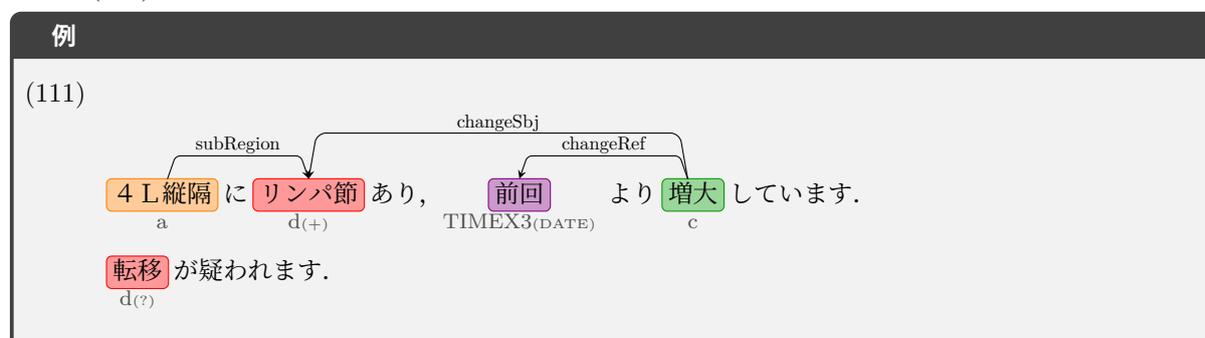
特徴・尺度エンティティ (<f>) 修飾先となる医療エンティティに対して基本関係 featureSbj を付与する。

この関係は文間をまたぐ場合にも付与する。

臓器・部位エンティティ (<a>) 多くの場合、包含する病変・症状エンティティや臓器・部位エンティティに対して基本関係 subRegion を付与する。特に E_{to} の臓器・部位エンティティが「内部」「辺縁」などの相対的な部位表現であれば文をまたいで積極的に付与する。

検査項目エンティティ (<t-key>)・薬品名エンティティ (<m-key>) 対の関係にある検査値エンティティ (<t-val>) や薬品値エンティティ (<m-val>) に向けて基本関係 keyValue を付与する (が、多くの場合は省略可能である→ 3.4.1 節)。

ただし、全ての医療エンティティに何らかの関係を付与しなければならないわけではない。例えば鑑別疾患や一般的な表現など、特定の部位や特徴・変化に関する言及が認められないことがあるためである。実際、以下の例(111)の最終行に出現する「転移」には、前節までに述べたどの基本関係も付与されない。



3.3 時間関係

医療エンティティ間の関係のうち、ある事態がいつ起きたのか、ある状態がいつ存在したのか、といった時間的な情報を表す関係を**時間関係**と呼ぶ。そのXML表記に空要素タグ<trel>を用いるものとする。

XML タグ

```
<trel id="rk" from="ei" to="ej" reltype="{types}" />
```

時間関係 $trel(E_{from}, E_{to})$ は、 E_{to} が時間表現であるときに成立する。ただし、ここでいう時間表現とは、テキスト中の時間表現 (<TIMEX3>タグ) エンティティが付与された表現に加え、概念としての**文書作成日時** (Document Creation Time; DCT) も含む。一般に DCT はテキスト中に明示的に出現せず、臨床テキストのメタデータとして与えられる。以降、エンティティ ID 「e0」は DCT を表す。

E_{from} となる医療エンティティは、

- 時間表現エンティティ (<TIMEX3>+ DCT)
- certainty 属性や state 属性を持つ医療エンティティ (<d>や<m-key>タグのエンティティなど)
- changeRef 関係を持たない変化エンティティ (<c>)

であり、基本的にはこれらの医療エンティティ全てに時間関係を付与する。臓器・部位エンティティ (<a>) や特徴・尺度エンティティ (<f>) については時間関係を付与せず、それらと基本関係で関係づけられた病変・症状エンティティ等に付与することを基本とする。ただし、certainty 属性や state 属性を持たない医療エンティティであっても、必要に応じて時間関係を付与することは禁止しない。また、certainty 属性が general であったり state 属性が other である医療エンティティには時間関係を付与しない。

時間関係には、timeOn, timeBefore, timeAfter, timeBegin, timeEnd の5タイプがある。図3.1に、5つの時間関係タイプが表現する時間概念を模式化する。次節より、それぞれについて詳細に説明する。

3.3.1 timeOn 関係

時間関係 $timeOn(E_{from}, E_{to})$ は、 E_{from} が E_{to} に相当する時間で起こったことを表す時間関係である。

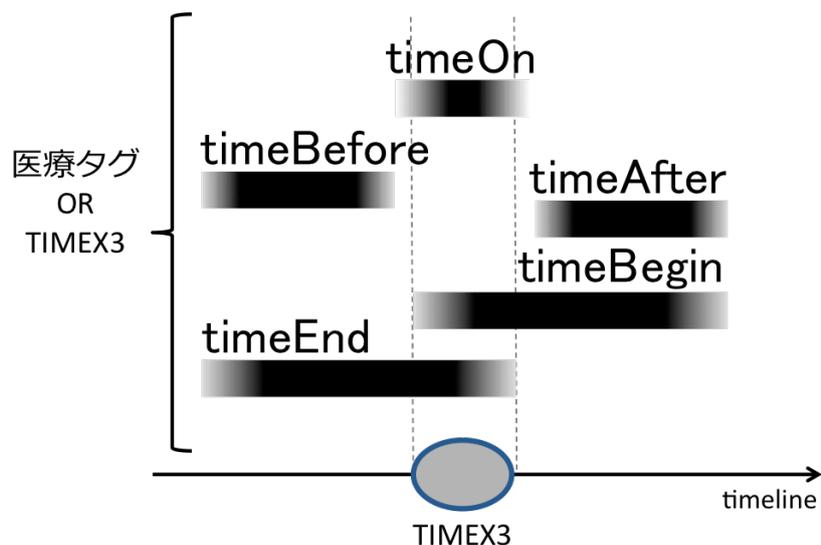


図 3.1 時間表現エンティティと医療エンティティの時間軸上の関係と、timeOn, timeBefore, timeAfter, timeBegin, timeEnd の対応。

XML タグ

```
<trel id="rk" from="ei" to="ej" reltype="timeOn" />
```

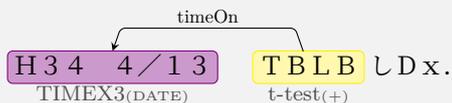
図 3.1 が示す通り、 E_{from} の開始・終了時点が厳密に E_{to} と一致している必要はない。正確な開始・終了時点が不明だが「およそ E_{to} ごろ」である場合、 E_{from} が E_{to} に相当する時間に包含されていることが読み取れる場合も含め、ひろく timeOn 関係とする。

ただし、 E_{from} と E_{to} の開始・終了時点に関して、以下のような時間的な前後関係がテキスト中に明示されている場合には、他の 4 つの時間関係を優先するものとする。

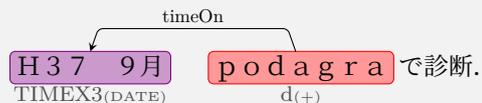
- E_{from} の終了時点が E_{to} より前である → timeBefore (3.3.2 節)
- E_{from} の開始時点が E_{to} より後である → timeAfter (3.3.3 節)
- E_{from} の開始時点が E_{to} である → timeBegin (3.3.4 節)
- E_{from} の終了時点が E_{to} である → timeEnd (3.3.5 節)

例

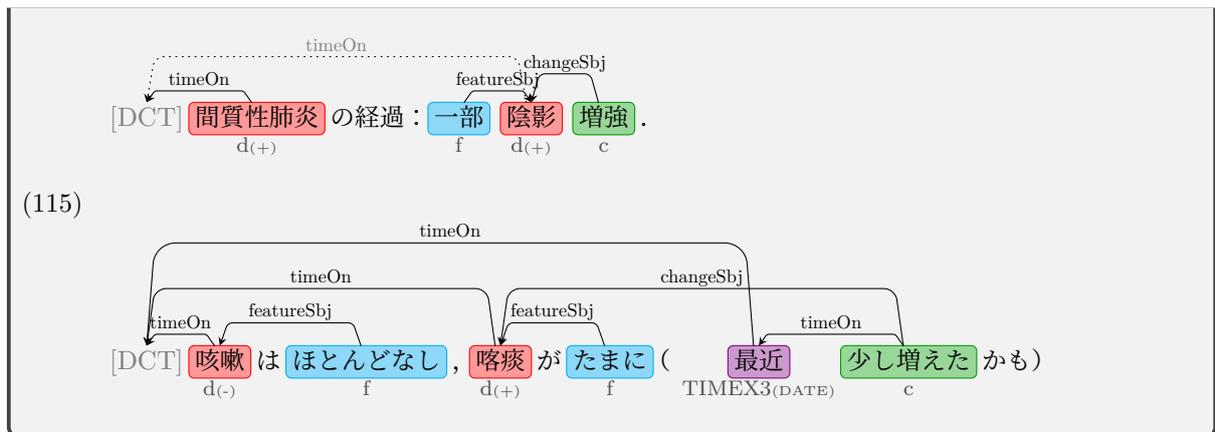
(112)



(113)



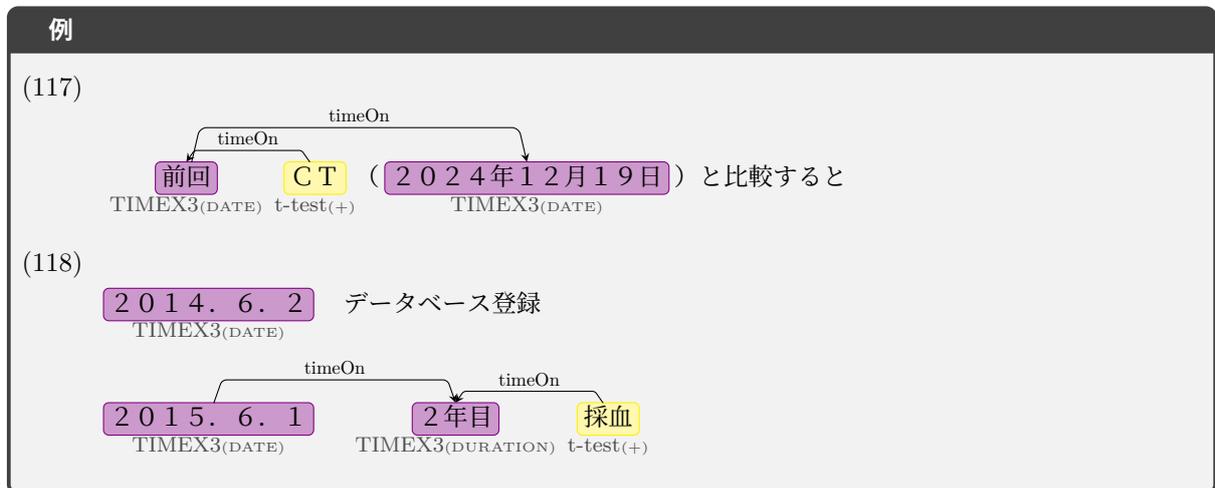
(114)



時間関係 timeOn によって、 E_{from} は時間軸上の E_{to} ごろに位置づけられるが、実際には単一の出来事であるとは限らない。たとえば例 (116) では、「クラビット」の使用は「1週間」の間に複数回発生していると考えられる。



時間関係 timeOn では、 E_{from} と E_{to} がともに時間表現であるとき、2項を入れ替えても意味は変わらないものとする。付与時には、便宜上、テキスト中で先に出現した方を E_{from} とする。



例からわかるように、基本関係 timeOn は多くの事例に適用され、時間的な前後関係や期間が必ずしも明確でない場合も多い。例 (113) では、病変・症状である「podagra」は、「H37 9月」より前から存在していた可能性もあり、また10月以降も完治せず継続する可能性がある。このように、開始・終了時点は定かではなく、「H37 9月」に存在していたことのみがテキストから明らかである場合も、時間関係 timeOn を付与する。

3.3.2 timeBefore 関係

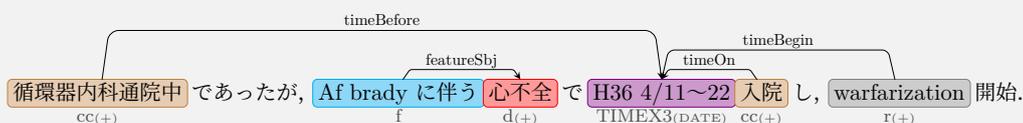
時間関係 $\text{timeBefore}(E_{\text{from}}, E_{\text{to}})$ は、 E_{from} が時間軸上において E_{to} より前の出来事である場合、すなわち、 E_{from} の終了時点が E_{to} の開始時点より前である場合に付与する。

XML タグ

```
<trel id="rk" from="ei" to="ej" reltype="timeBefore" />
```

例

(119)



3.3.3 timeAfter 関係

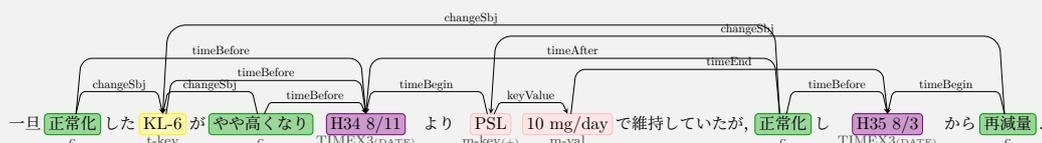
時間関係 $\text{timeAfter}(E_{\text{from}}, E_{\text{to}})$ は、 E_{from} が時間軸上において E_{to} より後の出来事である場合、すなわち、 E_{from} の開始時点が E_{to} の終了時点より後である場合に付与する。

XML タグ

```
<trel id="rk" from="ei" to="ej" reltype="timeAfter" />
```

例

(120)

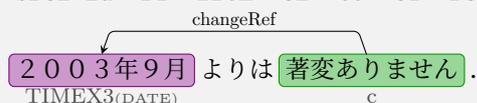


基本関係 changeRef (→ 3.2.2 節) が付与できる状況では、時間関係 timeAfter にも該当するよう思える場合が多い。以下に例 (93) を再掲する。

例 (再掲)

```
<TIMEX3 id="e1" type="DATE"> 2 0 0 3年9月 </TIMEX3> よりは<c id="e2"> 著変ありません </c>.
```

```
<brel id="r1" from="e2" to="e1" reltype="changeRef" />
```



この場合、「著変ありません」の時点は「2003年9月」より後に位置し、timeAfter(T2, T1)という時間関係も成立する。

ここでは、両方の関係を付与することはせず、基本関係 changeRef が付与される場合には、時間関係については付与しないものとする。

3.3.4 timeBegin 関係

時間関係 timeBegin(E_{from} , E_{to}) は、 E_{from} が時間軸上において E_{to} に始まった出来事である場合、すなわち、 E_{from} の開始時点が E_{to} である場合に付与する。特に、「開始」や「～から」といった表現が手掛かりになる。

XML タグ

```
<trel id="rk" from="ei" to="ej" reltype="timeBegin" />
```

例

(121)



3.3.5 timeEnd 関係

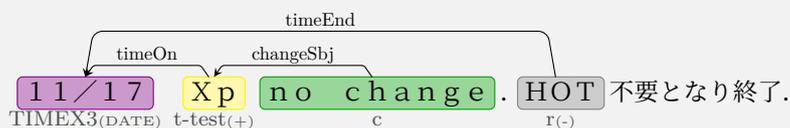
時間関係 timeEnd(E_{from} , E_{to}) は、 E_{from} が時間軸上において E_{to} に終了した出来事である場合、すなわち、 E_{from} の終了時点が E_{to} である場合に付与する。特に、「終了」や「～までに」といった表現が手掛かりになる。

XML タグ

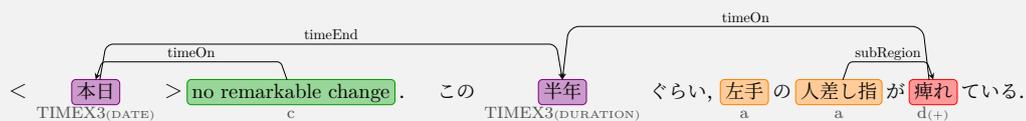
```
<trel id="rk" from="ei" to="ej" reltype="timeEnd" />
```

例

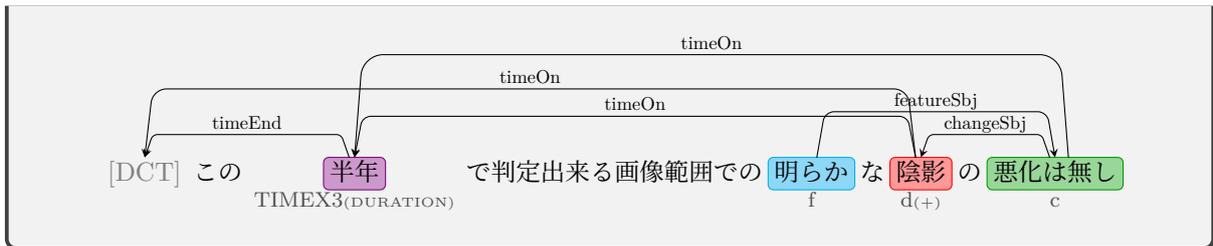
(122)



(123)



(124)



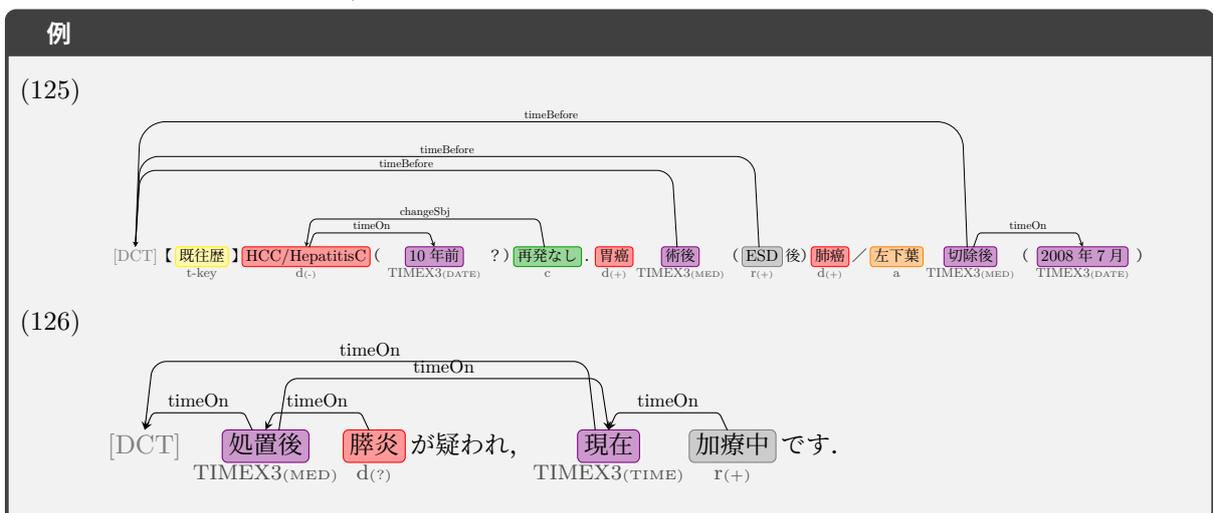
3.3.6 時間表現間の時間関係

先述したように、時間関係は時間表現同士の間付与される場合もある。時間表現は、単独で時間軸上のある時点に一意に対応するを**絶対的な時間表現**と、そうでない**相対的な時間表現**に分けられる。絶対的な時間表現は「2024年12月19日」「2003年9月」といった具体的な日付である。DCTも絶対的な時間表現である。これに対し、相対的な時間表現は「前回」「術後」と言った、それ自体では正確な日時を決められない表現である。

相対的な時間表現については、単独では時間軸上の位置が決まらないため、時間関係を付与し、絶対的な時間表現と関係づける必要がある。適切な絶対表現が存在しない場合には、相対的な時間表現同士の時間関係を可能なかぎり付与することで、時間的な前後関係が部分的に表現できる場合もある。

さらに、相対的な時間表現のうち、**type** 属性が MED である表現は臨床テキストに特有であり、時間関係の付与にあたっては注意が必要である。典型的には、「術後」や「処置後」「切除後」といった表現が該当する。

たとえば「術後」の場合、字義通りには手術終了以降の期間を指すが、ここでは「手術の後の一定期間」と解釈する。言い換えると、手術終了から現在に至るまでの長い期間だとは考えず、手術が影響を及ぼした限定的な期間を指していると考える。その上で、時間関係を付与する際には、「術後」の影響が続いている期間内であるかどうかを文脈で判断し、それによって付与する時間関係を選ぶ。以下に例を示す。



例(125)では、「術後」「ESD後」「切除後」をいずれも DCT より前の出来事と解釈^{*5}、時間関係 **timeBefore** を付与する。これらは、過去の既往歴として書かれているだけで、現在の文脈ではその影響は特に見られないと解釈できるためである。

*5 なお、「術後」「ESD後」「切除後」同士の関係はこの記述からはわからないので関係付与しなくて良い。

一方、例 (126) では、「処置後」の影響が「現在」に続いていると解釈する。この場合、「現在」は「処置後」に時間的に包含されるため、5つの時間関係のうち時間関係 on が該当する。

3.3.7 医療エンティティの state 属性と時間関係

検査、治療といった医療エンティティには、state 属性 (scheduled, executed, negated) が付与されている。state 属性の値は、その周囲の文脈における特定の時点を基準としたときの状態を表したものであり、時間に相対的に決定される。そのため、同じ文書中・文中であっても、ある時点では「予定」の状態である検査が、それより後のある時点では「実施済み」の状態である可能性がある。

state 属性をもつ医療エンティティに対する時間関係アノテーションでは、state 属性の値を踏まえた上で、実際の時間軸上での順序関係と整合するように時間関係の種類を決定する必要がある。

以下の表は、医療エンティティの state 属性の値と 5 種類の時間関係の組み合わせに対し、時間的解釈を示したものである。特に、state 属性が scheduled や negated である場合には注意が必要である。

state 属性	時間関係	解釈
scheduled	on	その時点で予定されている (実施予定日など) 例) TIMEX3(DATE) t-test(?)
	timeBefore	その時点より前のある時点での実施が予定されている 例) TIMEX3(DATE) t-test(?)
	timeAfter	その時点より後のある時点での実施が予定されている 例) TIMEX3(DATE) d(+) c cc(?)
	timeBegin	その時点から開始することが予定されている 例) TIMEX3(DATE) m-key(?) m-key
	timeEnd	その時点で終了することが予定されている 例) TIMEX3(TIME) r(?)
executed	on	その時点で実施された 例) TIMEX3(DATE) cc(+)

3.3.8 臨床テキストの種類と時間関係アノテーション

時間関係アノテーションにおいては、対象とする臨床テキストの種類とその特徴が、付与する関係の判断に役立つことがある。

例えば読影所見は、画像から読み取れる情報を記載するものであり、検査日に撮影した画像についての説明が主となる。前回や前々回の検査で撮影した画像との比較は頻出するが、原則として、DCT を基準に書かれている。

一方、診療記録には、過去から現在に至るまでの患者の既往歴や治療の履歴など様々な情報が記載されている。1 文書の中に出現する時間表現の数も多く、参照される時点も多い。それらの時点は、文書内で順を追って言及されるわけではないため、文脈に注意が必要である。

診療記録では、加えて、SOAP セクション構造も考慮する必要がある。各セクションには、主に以下の内容が記載されており、前提とされる時間も異なる場合がある。

- S (subjective)：主観的情報
- O (objective)：客観的情報
- A (assessment)：評価
- P (plan)：計画（治療）

たとえば、P（計画）のセクションに記載されるのは、原則として、治療の予定等、次回以降に実施される事柄である。そのため、P セクションに治療エンティティ (<r>) や検査エンティティ (<t-test/key/val>) が出現し、周辺に時間表現が明示的に出現しない場合は、DCT より後に位置づけるのが妥当である。以下は、ある診療記録の P セクションの記載の例である。

P S L 中止後も再燃なさそうであるので、ボナロンも中止。

「ボナロン」の中止が過去の出来事であるか否かはこの文だけでは判断できないが、P セクションの記載であることを考慮すれば、「ボナロン」の中止はこの時点での“予定”であると解釈できる。

SOAP のセクションをまたぐ際には、記載内容が切り替わることで異なる時点を参照する機会が多いので注意が必要である。

3.4 関係アノテーションの省略

一部の形式的に明らかなケースでは、アノテータは関係の付与を省略することができる。省略された関係はその後の工程で自動処理によって復元される。アノテーションの省略は、あくまでも作業を簡便にするためのものであり、付与を禁止するものではない。また、一部だけ省略することも許される。以下に、省略可能な場合を述べる。

3.4.1 基本関係 keyValue の自明なケース

検査項目エンティティ (<t-key>) と検査値エンティティ (<t-val>) が 1 対 1 ペアで出現する場合、すなわち、これらのエンティティ間に他のエンティティが出現しない場合、keyValue 関係の付与を省略することができる。薬品名エンティティ (<m-key>) と薬品値エンティティ (<m-val>) のペアの場合も同様である。

たとえば、例 (107) では、検査項目エンティティと検査値エンティティが 1 対 1 の対で出現している。以下に再掲する。



このとき、r1, r2 ともに省略可能である。

3.4.2 推論関係が自明なケース

2つの医療エンティティ間で成立する関係が、周囲のエンティティ間関係から容易に推論ができる場合がある。その中でももっとも自明なパターンの1つを以下に示す。

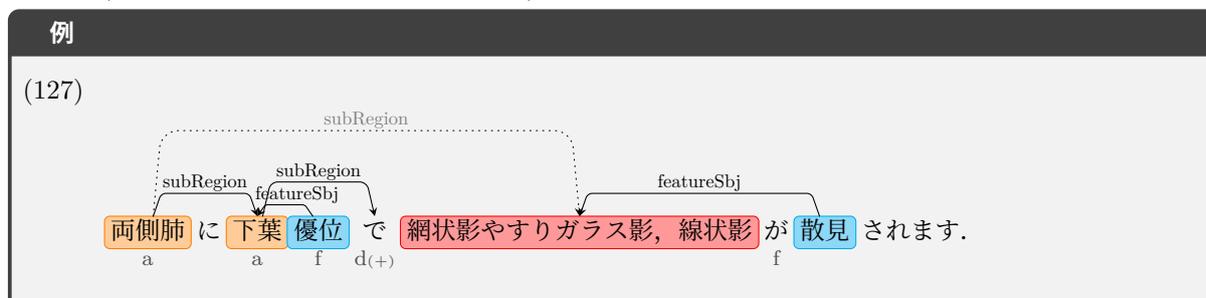
パターン：subRegion(A, B), subRegion(B, C) であるとき、subRegion(A, C) も成立

基本関係 subRegion(E_{from} , E_{to}) は、空間的に E_{from} に E_{to} が存在することを表す (3.2.4 参照)。従って、A に B が存在し、B に C が存在するとき、A に C が存在するのは必然である。このように、エンティティ A・B の間とエンティティ B・C の間で成立する subRegion 関係は、エンティティ A・C の間にも自明に成立する (このことを、基本関係 subRegion は推移律を満たすともいう)。このとき、エンティティ A・C の間の関係付与を省略できるものとする。

以下に具体例を示す。

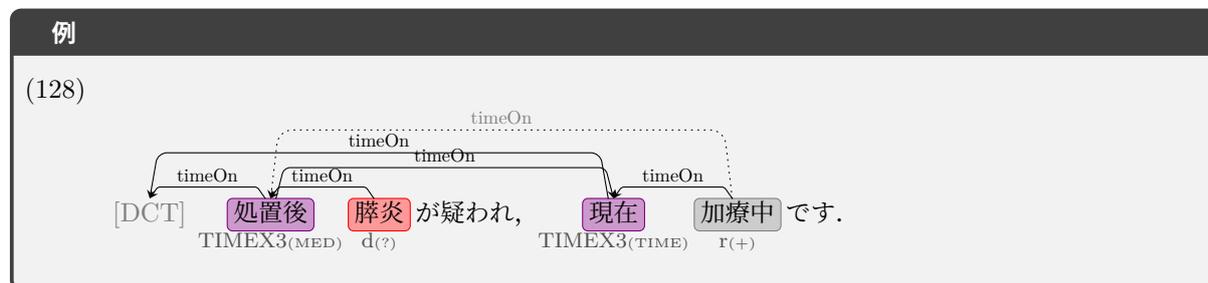
両側肺に下葉優位で網状影やすりガラス影、線状影が散見されます。

この例では、「両側肺」に「下葉」が位置し、「下葉」に「網状影やすりガラス影、線状影」が位置すると読み取れる。よって、「両側肺」に「網状影やすりガラス影、線状影」が位置すると推論され、実際そのように読み取れるが、この関係は省略の対象である。結果、上記の例は以下のようにアノテーションされる。



時間関係 timeOn, timeBefore, timeAfter に関しても同様の推論が成り立つ場合がある。以下の例からは、処置後に加療中であることが読み取れる。つまり、timeOn(加療中, 処置後) という関係が成立するが、この

関係は (R4 の 2 項が前後入れ替え可能であることを考慮すると) R4 と R5 から推論することができ、あえて関係を付与する必要はない。



3.4.3 時間関係 timeOn の省略

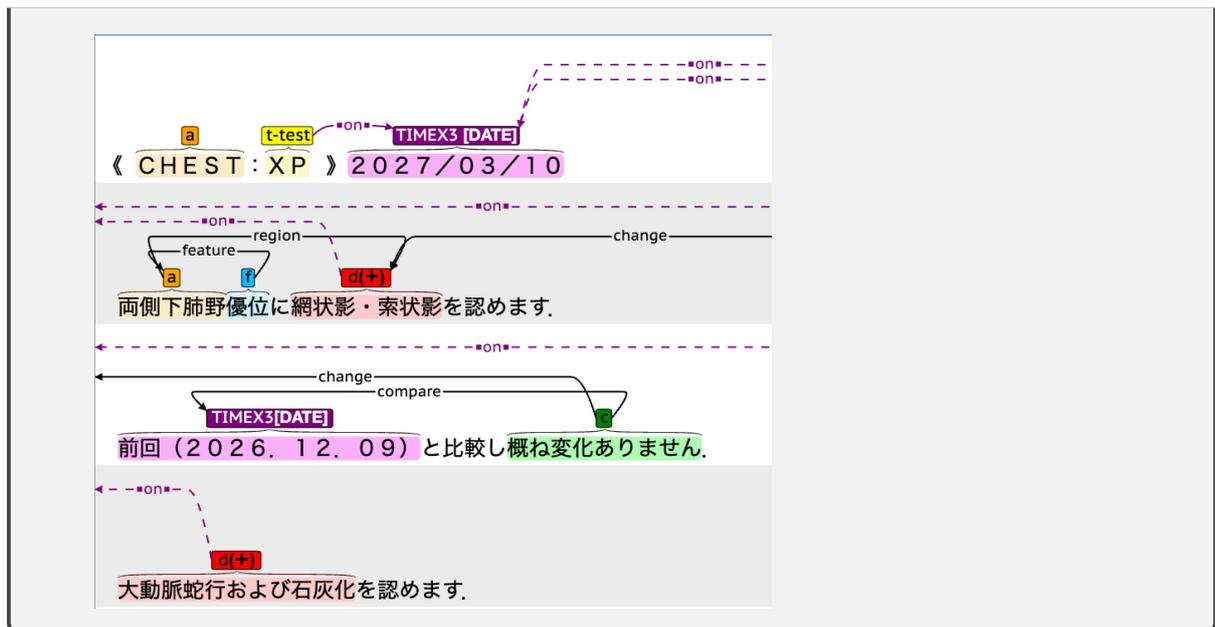
次の場合、時間関係 timeOn の付与を省略できるものとする。

あるエンティティ **X** と、ある時間表現エンティティ **T** との間の **timeOn** 関係は、**X** の直前に出現するエンティティ (時間関係の E_{from} となるもの) が同じ時間エンティティ **T** と timeOn 関係を持っている場合には省略してよい。また、この省略条件は連鎖的に適用できる。

例えば、ある時点について記述されたセクションの中では、多数の医療エンティティが1つの時点に属することになり、ある同一の時間表現に対して多数の timeOn 関係が付与されることになる。長い文書の場合などでこの作業は煩雑になるため、次の時点に関する記述が始まるまでの間、最初に付与した timeOn 関係以外を省略できるということである。

逆に、ある医療エンティティ **X** についての時間関係が省略されている場合、その時間関係は timeOn 関係であり、関係を付与する先の時間は直前の timeOn 関係と同じであると解釈する。この省略の条件は、テキスト中に出現する時間表現エンティティ (TIMEX3) に限定せず、DCT の場合にも同様に適用される。





この例では、ID r5 以降の関係を省略できる。まず、時間関係を付与すべき医療エンティティとして最初に出現する「XP」(e2) が時間表現エンティティ「2027/03/10」(e3) に対して timeOn 関係を持つ (関係 r4)。この次に時間関係を付与すべき医療エンティティは e6 の「網状影・索状影」であるが*6、関係 r4 と同じく E_{to} が e3 であることから、関係 r5 は省略可能となる。さらに、次の「大動脈蛇行および石灰化」(e9) も、直前の「網状影・索状影」(e6) と同じく時間表現 e3 に対して timeOn 関係にある (関係 r6) ため、関係 r5 と同じく省略できる (連鎖適用)。これ以降も、同じ時間点「2027/03/10」(e3) に対して timeOn 関係にある医療エンティティが連続して出現している限り、最初の 1 つである関係 r4 を除いて timeOn 関係の付与を省略してよいということである。なお繰り返しになるが、省略は任意であり、冗長に付与しても良い。

*6 時間関係の E_{from} となる医療エンティティは、時間表現、certainty 属性や state 属性を持つ医療エンティティ、および changeRef 関係を持たない変化エンティティ (<c>) であることに注意する。

謝辞

本ガイドライン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。