

合成データを用いた日本語電子カルテ文書の構造化

水見 和¹ 清水 聖司¹ 浅山 能久² 石川 千里² 内田 大輔² 中川 章²西山 智弘¹ 若宮 翔子¹ 荒牧 英治¹¹ 奈良先端科学技術大学院大学² 富士通株式会社 富士通研究所 人工知能研究所

himi.yamato.iz5@naist.ac.jp, shimizu.seiji.so8@is.naist.jp

{y.asayama,ishikawa.chisat,uchida.daisuke,anaka}@fujitsu.com

{nishiyama.tomohiro.ns5,wakamiya,aramaki}@is.naist.ac.jp

概要

電子カルテ文書から事前に定められた項目を情報抽出する**構造化**技術は、テキストデータの二次利用に不可欠である。近年は大規模言語モデル (LLM) による構造化が期待されているが、医療現場においては、外部 API 型 LLM の利用は制約される。さらに、学習データ作成では実データの共有が困難であることに加え、人手によるアノテーションは高コストであり、判断基準の揺れも伴うという課題がある。本研究では、公開症例報告から LLM で合成データを生成し、その合成データで Small Language Model (SLM) をファインチューニングして構造化モデルを構築する。肺がん診療録の構造化タスクでの実験により、合成データが学習用データとして網羅性を満たし、構造化モデル構築に有効であることを示した。

1 はじめに

電子カルテ文書には、診断や検査結果といった患者の臨床情報が自由記述で記載されており、これらを一定のフォーマットや表項目に従って情報抽出できれば、予後の予測や医薬品有害事象などの臨床研究といった多様な二次利用につながる [1, 2]。こうしたテキストからの情報抽出については、臨床概念を抜き出す固有表現抽出や、概念同士の関係性を特定する関係抽出などが研究されてきた [3]。しかし、臨床現場での実用性を高めるためには、個別の概念や関係性の抽出にとどまらず、抽出した情報を診療の文脈に即して体系的に再構成し、**構造化**して出力する必要がある。そこで本研究では、診療録から診断名や検査値などの事前に定義した項目に対応する情報を抽出し、日付を軸に診断名と検査値や有害事

1. 右上葉肺癌 (Ad) cT1 bN0M0 Stage IA2
S)特に自覚症状はありません。
O)
2025/02/05
体幹部造影CT：右肺上葉S2に最大径20mmの結節性病変あり。
病理結果：RET遺伝子変異陽性。
A)
2025/02/05
Grade 2相当の悪心を認めたが、対症療法でコントロール可能であった。
:

構造化

日時	診断名	腫瘍部位	腫瘍サイズ	RET	患者の反応	毒性グレード
2025/02/05	右上葉肺癌	右肺上葉	20	陽性 (+)	—	—
2025/02/05	—	—	—	—	悪心	Grade2
:	:	:	:	:	:	:

象などの内容を**医療イベント**単位で整理するタスクを**構造化**と定義し、これを情報抽出の応用形態と位置づける。さらに、1文書から得られる医療イベントの集合を**構造化結果**と呼ぶ。例えば、図1のように、自由記述の診療録を構造化することで、腫瘍サイズやステージといった検査値が2025年2月5日の右肺上葉に対するものであると整理できる。

構造化は従来、固有表現認識、正規化、事実性判定などの複数の処理のパイプラインからなる複雑なタスクであったが、近年では、LLMを情報抽出器として用いることで、完結かつ高精度に構造化が実現できる可能性が示されている [4, 5]。しかし、医療現場では、電子カルテが個人情報を含むため外部 API 型 LLM の利用は制約される [6]。さらに、大規模なモデルの院内等のオフライン環境での運用は、高い計算コストを要するため、困難となる。

この問題に対処するため、計算コストを抑え、か

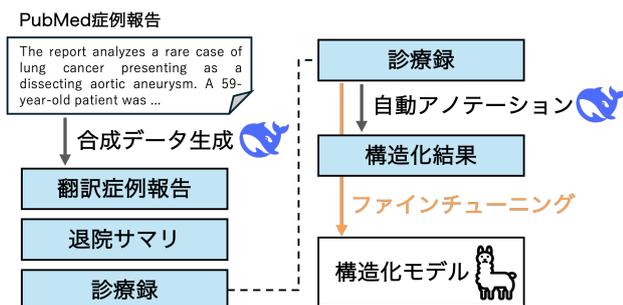


図2 提案手法の概要. PubMed 症例報告から, DeepSeek により翻訳症例報告, 退院サマリ, 診療録を生成する. 次に, 診療録から, DeepSeek により構造化の自動アノテーションを行う. 得られた診療録と構造化結果で SLM をファインチューニングし, 構造化モデルへ知識蒸留する.

高性能な構造化を実現する方法として, SLM のファインチューニングが有効であることが示されている [7, 8]. しかし, 学習データ作成の人手アノテーションは実用上困難であり, 専門性の高さや高コストに加えて, 作業者の知識量に応じて付与基準の判断の揺れも生じやすい.

そこで, 本研究では, LLM を用いて (i) 電子カルテ文書を合成し, (ii) 構造化の自動アノテーションを行うことで, SLM の学習用データを自動的に構築する [9]. 提案手法は 2 段階からなる (図 2). まず, 公開症例報告から退院サマリと診療録を段階的に生成する. 次に, 生成した診療録に対して, LLM で構造化の自動アノテーションを行い, 得られた診療録と構造化結果で SLM をファインチューニングする.

実験では, **肺がん患者に対する診療録の構造化**を想定し, 学習データを作成した. この結果, 作成した 3,529 件のデータにおいて, 本タスクにおける全ての抽出項目が少なくとも一度は出現し, 学習データとしての網羅性を満たすことを確認した. さらに, 模擬診療録 98 件を用いた構造化性能の評価において, ファインチューニングした SLM は In-Context Learning (ICL) ベースラインと比較して micro F1 で 0.556 から 0.650 へ性能が向上した. 以上より, 公開症例報告からの文書合成, 自動アノテーション, 知識蒸留という枠組みが, 電子カルテの構造化モデル構築に有効である可能性を示した.

2 関連研究

2.1 電子カルテ文書の構造化タスク

本研究における電子カルテ文書の構造化とは, 診療録からあらかじめ定義した抽出項目を取り出

し, 日付を軸に医療イベントという単位で集約するタスクを指す. このタスクの基盤となる技術は情報抽出であり, 臨床分野においても固有表現抽出や関係抽出といった情報抽出の手法が研究されてきた [10, 11]. 近年ではこれらを一貫して行う End-to-end のアプローチが注目されており, n2c2 2018 においては, 概念抽出と関係抽出を同時に行う End-to-end のタスクが設定されている [1]. 本研究もこの流れに沿い, 診療録から End-to-end の生成タスクとして構造化を行う. ただし, 従来の情報抽出は主に個別の概念や局所的な属性のペアを対象としたのに対し, 本研究の構造化は, 抽出項目を医療イベントとしてまとめる点に特徴がある. これにより, 患者の経過を反映した体系的な情報を抽出する.

2.2 公開症例報告に基づく合成データ生成と知識蒸留

公開症例報告を用いた電子カルテ文書の合成と, その合成データを使った知識蒸留は, 電子カルテデータの共有制約を回避しつつ, 臨床 NLP モデルを構築するための現実的なアプローチとして位置づけられる. 先行研究では, 公開症例報告から退院サマリ形式の合成文書を生成し, さらに NER や QA などの汎用タスクに対する指示文と回答のペアを生成して学習に用いる枠組みが提案されている [7]. さらに, 合成データを介して大規模モデルの知識を小規模モデルへ移す方法として, 教師モデルが生成した学習例で生徒モデルをファインチューニングする知識蒸留の枠組みも示されている [12]. 本研究では, これらの先行研究に着想を得て, 英語で書かれた公開症例報告を基に日本語診療録を合成し, 肺がん構造化という特化タスクに対する学習データとして用いる. これにより, SLM を学習し, オフライン環境での運用が可能な構造化モデルの構築を目指す.

3 手法

3.1 電子カルテ文書の合成

本研究では, 公開症例報告を基に, 日本語の電子カルテ文書を段階的に合成し, 構造化学習に利用可能なデータを構築する. 具体的には, PubMed に掲載された症例報告の Abstract を収集し, 日本語翻訳後, 退院サマリ, 診療録の順に文書を生成する. 各文書の生成では, (i) 臨床情報の補完および (ii) 文書フォーマットの変換により, 実電子カルテ文書に近い文書を生成する. (i) について, 症例報告

の Abstract は簡潔に記述されるため、検査値、診断、治療経過などの詳細な臨床情報が省略されることが多い。そこで、退院サマリおよび診療録の生成時には、Abstract の記述と整合する範囲で不足する臨床情報を LLM で補完する。(ii) については、実際の電子カルテ文書のフォーマットを再現するため、模擬電子カルテ文書を ICL の例として与え、文書フォーマットを制御する。

3.2 知識蒸留

計算コストを抑えつつ、高性能な構造化モデルを構築するため、LLM から SLM への知識蒸留を行う。まず、生成した診療録に対して、LLM により構造化の自動アノテーションを行う。具体的には、生成された診療録の中から抽出項目を取り出し、図 1 のように、構造化結果を医療イベントのまとまりとして出力する。ここで、各抽出項目の値は、ステージ分類のようなカテゴリ値、検査値のような数値や陽性・陰性などの 2 値、診断名のような文字列など、項目に応じて異なる。この際、出力形式を制御するため、人手でアノテーションした模擬診療録と構造化結果のペアを ICL 例として与える。次に、自動アノテーションで得られた“生成された診療録”と“自動アノテーションされた構造化結果”のペア（合成学習データ）を用いて SLM を QLoRA によりファインチューニングする [13]。これにより、外部 API 型 LLM の利用が難しい環境でも動作可能な、構造化タスク特化の小規模モデルを得る。

4 実験

4.1 材料

PubMed 収録文献の中から、「肺がん」と「患者の年齢」を含む症例報告を文字列一致により抽出し、得られた合計 3,787 件を用いた。フォーマット変換のための ICL 例には、模擬退院サマリおよび模擬診療録を 3 文書 (3-shot)、構造化の自動アノテーションには、人手でアノテーションされた 2 文書 (2-shot) を用いた。自動アノテーションの結果、抽出項目が全く検出されなかった文書を除外し、最終的に 3,529 件の合成学習データが得られた。ファインチューニングの評価には、LLM による合成データとは別に、人手により作成およびアノテーションを行った計 100 文書の模擬診療録と構造化結果のペアを用意した。そのうち、ICL 例に用いた 2 文書を

除く 98 文書を評価に用いた。

4.2 実験設定

対象タスク：肺がん診療録の構造化 本研究で検証する具体的なタスクは肺がん診療録の構造化である。本タスクは、自由記述で書かれた診療録から肺がんに関わる重要情報を抽出し、日付を軸に医療イベントをまとめて出力することを目的とする。抽出項目は 41 項目あり、基本情報（例：日付、診断名）、腫瘍情報（例：発生部位、腫瘍サイズ等）、有害事象（例：患者の反応、毒性グレード）、検査所見（例：TTF1, NapsinA 等）、治療（例：治療内容、放射線総量、放射線照射回数）から成る。検査所見には、肺がんの診断や治療方針の決定で使われる項目が多く含まれる。全項目の一覧は付録の表 A.1 に示す。

合成データ生成（症例報告 Abstract の日本語翻訳、退院サマリ、診療録）および自動アノテーションには DeepSeek-V3.2 (Non-thinking Mode) を使う。このモデルは、生成出力を学習データとして二次利用可能であることを考慮して選定した。構造化モデルには Llama-3.1-Swallow-8B-Instruct-v0.5 を採用し、知識蒸留は、QLoRA によりファインチューニングする。ハイパーパラメータは、LoRA を $r = 16$, $\alpha = 64$, dropout=0.05, 学習を epoch 数 3, 学習率 2×10^{-4} , 実効バッチサイズ 16 とした。ベースラインは、同一の基盤モデルをファインチューニングせずに、自動アノテーション時と同じプロンプトと ICL 例を付与して構造化を行う。これにより、知識蒸留による精度の変化を比較する。

4.3 評価方法

本研究では、合成学習データの網羅性と、知識蒸留により得られた構造化モデルの性能を評価する。

合成学習データの網羅性評価 作成した合成学習データが事前に用意した抽出項目をどの程度含むかを網羅性として確認するため、各項目の出現数を集計する。具体的には、各診療録に当該抽出項目が 1 回以上出現するかを判定し、出現した診療録の件数をカウントする。

構造化性能の評価 構造化性能は、SLM の出力と正解の構造化結果に含まれる医療イベントを 1 対 1 に対応付け、micro 平均の Precision, Recall, F1 で評価する。対応付けの方法は、抽出項目と値が一致した数の合計が最大となる医療イベントの組み合わせを採用する。具体的には、対応付けられた出力と正

表 1 構造化性能の結果. 提案手法が全ての指標でベースラインを上回った.

手法	Precision	Recall	F1
ベースライン	0.606	0.514	0.556
提案手法	0.652	0.648	0.650

解の医療イベントにおいて、同一の抽出項目と値の両方が一致した場合のみ正解とした（例：値が同じリンパ節でも、転移部位=リンパ節と再発部位=リンパ節は別項目のため一致しない）。

5 結果と考察

合成学習データの網羅性 合成データにおける各項目を含む文書数について、作成した合成学習データでは、全 41 の抽出項目が少なくとも一度は含まれており、提案手法により網羅的な合成データ生成が可能であることが示された。詳細は、付録の表 A.1 に示す。一方で、項目間の出現文書数には偏りが見られ、「日付」、「診断名」、「腫瘍関連の基本情報」、「治療」など、多くの症例で言及されやすい項目は高頻度出現するのに対し、「超音波気管支鏡下穿刺吸引」、「オンコマイン」、「NGS」、「ROSE」の検査項目では、出現文書数が 100 件未満にとどまった。

構造化性能 表 1 に、micro 平均の Precision, Recall, F1 を示す。ファインチューニングされた SLM の性能はベースラインと比較して、Precision, Recall, F1 がそれぞれ、0.046, 0.134, 0.094 向上し、特に Recall が大きく向上した。

41 項目の項目別の指標について、評価データに出現しない 2 項目を除く、39 項目で比較すると、向上した項目は 27 件、低下した項目は 6 件、同等は 6 件であった。このことから、精度の向上は幅広い項目に及んでいることが分かる。結果の詳細は、付録の表 A.2 に示す。その中でも特に性能変化の大きい項目を図 3 に示す。差分上位の抽出では、出現数が少ない項目は推定値が不安定になるため、出現数が 10 以上の項目に限定した。複数の項目で大きな上昇が見られた一方、性能が低下した項目はいずれも低下幅が 0.1 未満にとどまった。したがって、提案手法は一部項目での小さな低下を伴いながらも、全体として構造化性能を底上げすることに成功した。

6 おわりに

本研究では、公開症例報告から日本語退院サマリと診療録を合成し、肺がんを対象として、事前に定

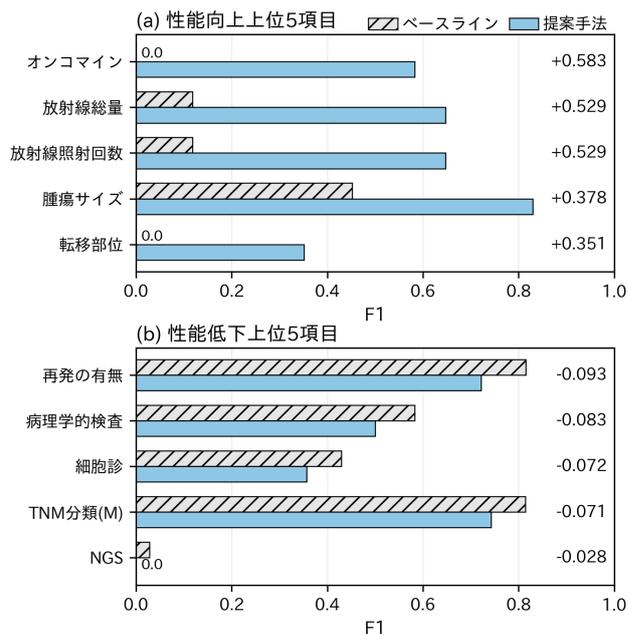


図 3 構造化性能の向上 (a) および低下 (b) における上位 5 項目の比較. 全体として性能向上への寄与は大きく、性能低下の影響は小さかった。

義した抽出項目に基づく LLM での構造化の自動アノテーションを行った。これにより、診療録テキストを入力、自動アノテーションの構造化結果を出力とした合成学習データを作成した。さらに、作成した合成学習データを用いて SLM をファインチューニングし、オフライン環境で運用可能な構造化モデルの構築を試みた。

実験の結果、合成データ 3,529 件において抽出項目の全てが少なくとも一度は観測され、項目の観点からは構造化学習に必要な網羅性を満たすことを確認した。加えて、構造化モデルは ICL ベースラインに対して micro F1 を 0.556 から 0.650 へ向上し、合成データ生成と自動アノテーションに基づく知識蒸留が、日本語電子カルテ文書の構造化性能向上に寄与する可能性を示した。

一方で、項目の出現頻度には依然として偏りが見られ、低頻度項目において学習や評価の安定性に欠ける点が課題である。さらに、今回は項目の出現数のみに注目したが、各項目が取り得る値の範囲を十分に網羅できているかについては検証の余地がある。今後は、低頻度項目を意図的に含め、かつ、値の多様性を満たす生成制御やサンプリング戦略の改善を進める。これらを通じて、オフライン環境で運用可能な SLM による日本語電子カルテ文書構造化の社会実装を目指す。

謝辞

本研究の一部は、「戦略的イノベーション創造プログラム (SIP)」「統合型ヘルスケアシステムの構築」JPJ012425, JSPS 研究スタート支援 JP25K24412 の補助を受けて行った。

参考文献

- [1] Sam Henry, Kevin Buchan, Michele Filannino, Amber Stubbs, and Ozlem Uzuner. 2018 n2c2 shared task on adverse drug events and medication extraction in electronic health records. **Journal of the American Medical Informatics Association**, Vol. 27, No. 1, pp. 3–12, 10 2019.
- [2] Yanshan Wang, Liwei Wang, Majid Rastegar-Mojarad, Sungnim Moon, Feichen Shen, Naveed Afzal, Sijia Liu, Yuqun Zeng, Saeed Mehrabi, Sunghwan Sohn, and Hongfang Liu. Clinical information extraction applications: A literature review. **Journal of Biomedical Informatics**, Vol. 77, pp. 34–49, 2018.
- [3] Wendy W Chapman, Prakash M Nadkarni, Lynette Hirschman, Leonard W D’Avolio, Guergana K Savova, and Ozlem Uzuner. Overcoming barriers to nlp for clinical text: the role of shared tasks and the need for additional creative solutions. **Journal of the American Medical Informatics Association**, Vol. 18, No. 5, pp. 540–543, 2011. Sep–Oct 2011.
- [4] Monica Agrawal, Stefan Hegselmann, Hunter Lang, Yoon Kim, and David Sontag. Large language models are few-shot clinical information extractors. In Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang, editors, **Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing**, pp. 1998–2022, Abu Dhabi, United Arab Emirates, December 2022. Association for Computational Linguistics.
- [5] Yaojie Lu, Qing Liu, Dai Dai, Xinyan Xiao, Hongyu Lin, Xianpei Han, Le Sun, and Hua Wu. Unified structure generation for universal information extraction. In Smaranda Muresan, Preslav Nakov, and Aline Villavicencio, editors, **Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)**, pp. 5755–5772, Dublin, Ireland, May 2022. Association for Computational Linguistics.
- [6] Isabella Catharina Wiest, Dyke Ferber, Jiefu Zhu, Marko van Treeck, Sonja K. Meyer, Radhika Juglan, Zunamys I. Carrero, Daniel Paech, Jens Kleesiek, Matthias P. Ebert, Daniel Truhn, and Jakob Nikolas Kather. Privacy-preserving large language models for structured medical information retrieval. **npj Digital Medicine**, Vol. 7, No. 1, p. 257, 2024. Published 2024-09-20.
- [7] Sunjun Kweon, Junu Kim, Jiyou Kim, Sujeong Im, Eunbyeol Cho, Seongsu Bae, Jungwoo Oh, Gyubok Lee, Jong Hak Moon, Seng Chan You, Seungjin Baek, Chang Hoon Han, Yoon Bin Jung, Yohan Jo, and Edward Choi. Publicly shareable clinical large language model built on synthetic clinical notes. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar, editors, **Findings of the Association for Computational Linguistics: ACL 2024**, pp. 5148–5168, Bangkok, Thailand, August 2024. Association for Computational Linguistics.
- [8] Chaoyi Wu, Weixiong Lin, Xiaoman Zhang, Ya Zhang, Weidi Xie, and Yanfeng Wang. PMC-LLaMA: toward building open-source language models for medicine. **Journal of the American Medical Informatics Association**, Vol. 31, No. 9, pp. 1833–1843, 09 2024.
- [9] Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, and Hannaneh Hajishirzi. Self-instruct: Aligning language models with self-generated instructions. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, **Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)**, pp. 13484–13508, Toronto, Canada, July 2023. Association for Computational Linguistics.
- [10] 西山智弘, 柴田大作, 宇野裕, 辻川剛範, 北出祐, 久保雅洋, 矢田竣太郎, 若宮翔子, 荒牧英治. 生成モデルは医療テキストの固有表現抽出に使えるか? 言語処理学会 第 30 回年次大会 発表論文集, pp. 3072–3077. 言語処理学会, March 2024.
- [11] 風間正弘, 太田満久, 西林孝. プロンプト最適化を用いた大規模言語モデルによる医療文章からの固有表現抽出. 人工知能学会全国大会論文集, Vol. JSAI2025, pp. 1Q3OS3504–1Q3OS3504, 2025.
- [12] Elizabeth Geena Woo, Michael C. Burkhart, Emily Alsentzer, and Brett K. Beaulieu-Jones. Synthetic data distillation enables the extraction of clinical information at scale. **npj Digital Medicine**, Vol. 8, No. 1, p. 267, 2025.
- [13] Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, and Luke Zettlemoyer. Qlora: Efficient finetuning of quantized llms. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, **Advances in Neural Information Processing Systems**, Vol. 36, pp. 10088–10115. Curran Associates, Inc., 2023.

A 評価結果の詳細

表 A.1: 合成データにおける各項目を含む文書数. 全ての抽出項目で一度は出現したことより構造化タスク学習用データとしての網羅性を満たすことを示す.

カテゴリ	抽出項目	出現数	カテゴリ	抽出項目	出現数
日付	日付	3529	所見 (検査)	TTF1	2087
診断名	診断名	3521	所見 (検査)	p40	1778
腫瘍	腫瘍が発生した部位	2827	所見 (検査)	p63	1411
腫瘍	腫瘍サイズ	2248	所見 (検査)	CK5_6	1546
腫瘍	TNM 分類 (T)	2435	所見 (検査)	NapsinA	2009
腫瘍	TNM 分類 (N)	2387	所見 (検査)	EGFR	1748
腫瘍	転移部位	1513	所見 (検査)	ALK	1661
腫瘍	TNM 分類 (M)	2513	所見 (検査)	ROS1	1519
腫瘍	ステージ分類	2674	所見 (検査)	RET	538
腫瘍	再発の有無	1804	所見 (検査)	BRAF	549
腫瘍	再発部位	426	所見 (検査)	KRAS	683
腫瘍	組織学的グレード名	648	所見 (検査)	HER2	532
有害事象	患者の反応	1343	所見 (検査)	病理学的検査	1693
有害事象	毒性グレード	853	所見 (検査)	評価部位	630
所見 (検査)	細胞診	225	所見 (検査)	NTRK	436
所見 (検査)	経気管支肺生検 (TBB)	102	所見 (検査)	NGS	92
所見 (検査)	経気管支肺生検 (TBLB)	1097	所見 (検査)	ROSE	36
所見 (検査)	超音波気管支鏡下穿刺吸引	29	治療	治療	2723
所見 (検査)	PD-L1	1646	治療	放射線総量	573
所見 (検査)	PET-CT の SUVmax	1594	治療	放射線照射回数	567
所見 (検査)	オンコマイン	16			

表 A.2: 提案手法の抽出項目別の構造化性能. 出現数は正解データにおける抽出項目の出現数を示す.

抽出項目	出現数	ベースライン	提案手法	抽出項目	出現数	ベースライン	提案手法
日付	978	0.629	0.731	TTF1	42	0.810	0.976
診断名	552	0.196	0.324	p40	29	0.897	0.931
腫瘍が発生した部位	567	0.730	0.854	p63	0	N/A	N/A
腫瘍サイズ	230	0.452	0.830	CK5_6	6	0.833	1.000
TNM 分類 (T)	140	0.850	0.864	NapsinA	11	0.727	1.000
TNM 分類 (N)	140	0.857	0.857	EGFR	29	0.724	0.897
転移部位	57	0.000	0.351	ALK	12	0.667	0.833
TNM 分類 (M)	140	0.814	0.743	ROS1	15	0.733	0.800
ステージ分類	138	0.623	0.841	RET	2	1.000	1.000
再発の有無	54	0.815	0.722	BRAF	0	N/A	N/A
再発部位	2	0.000	0.000	KRAS	4	0.750	0.500
組織学的グレード名	17	0.000	0.294	HER2	3	1.000	1.000
患者の反応	299	0.415	0.585	病理学的検査	12	0.583	0.500
毒性グレード	289	0.540	0.616	評価部位	107	0.019	0.196
細胞診	14	0.429	0.357	NTRK	1	1.000	1.000
経気管支肺生検 (TBB)	3	0.000	0.000	NGS	36	0.028	0.000
経気管支肺生検 (TBLB)	79	0.582	0.911	ROSE	8	0.125	0.750
超音波気管支鏡下穿刺吸引	9	0.778	0.889	治療	284	0.236	0.303
PD-L1	90	0.378	0.667	放射線総量	17	0.118	0.647
PET-CT の SUVmax	152	0.401	0.651	放射線照射回数	17	0.118	0.647
オンコマイン	12	0.000	0.583				