Extracting Relations between Radiotherapy Treatment Details

Danielle S. Bitterman[1,2,3], Timothy A. Miller[3], David Harris[3], Chen Lin[3],
Sean Finan[3], Jeremy L. Warner[4,5], Raymond H. Mak[2], Guergana K. Savova[3]

[1]Harvard Radiation Oncology Program, Boston, MA
[2]Department of Radiation Oncology, Brigham and Women's Hospital/Dana-Farber Cancer Institute, Boston MA
[3]Boston Children's Hospital Informatics Program, and Harvard Medical School, Boston, MA
[4]Division of Hematology/Oncology, Vanderbilt University, Nashville, TN
5Department of Biomedical Informatics, Vanderbilt University, Nashville, TN

Abstract

We present work on extraction of radiotherapy treatment information from the clinical narrative in the electronic medical records. Radiotherapy is a central component of the treatment of most solid cancers. Its details are described in non-standardized fashions using jargon not found in other medical specialties, complicating the already difficult task of manual data extraction. We examine the performance of several state-of-the-art neural methods for relation extraction of radiotherapy treatment details, with a goal of automating detailed information extraction. The neural systems perform at 0.82-0.88 macro-average F1, which approximates or in some cases exceeds the inter-annotator agreement. To the best of our knowledge, this is the first effort to develop models for radiotherapy relation extraction and one of the few efforts for relation extraction to describe cancer treatment in general.

1 Introduction

Radiotherapy is the use of ionizing radiation, which is radiation with enough energy to remove electrons from atoms and molecules, to treat disease (Gunderson and Tepper, 2020). The predominant indication for radiotherapy is the treatment of cancer, where it exerts its antineoplastic effect via DNA damage, which preferentially kills cancer cells over health tissue cells (McDermott and Orton, 2010). Radiotherapy plays a central role in the curative and palliative treatment of many cancers. It is estimated that up to 30% of cancer patients receive radiotherapy as a part of their first-line treatment, and approximately 50% of all cancer patients receive radiotherapy during the course of their cancer care (Delaney et al., 2005; Smith et al., JCO 2010).

Despite its importance in cancer treatment, radiotherapy is included in cancer registries in only high-level, often cursory detail, if at all. For example, radiotherapy details are only available by custom request in the publicly available Surveillance, Epidemiology, and End Results Program (SEER) cancer registry, acknowledging incompleteness and errors in this manually extracted data (Surveillance, Epidemiology, and End Results Program). The reason for this is multifold. First, radiotherapy is a highly technical field not extensively taught in medical school, and uses its own jargon not found in other medical texts. Additionally, radiation treatment details are frequently not entered into the electronic medical records (EMR) as structured data. Instead, radiotherapy is described in clinical free text using descriptive and very non-standardized language. Radiotherapy treatment descriptions are more similar to the documentation of operative procedures than to documentation of medication regimens. Because radiation is personalized to each patient's disease and anatomy, it cannot be described with standard reporting of the type of radiation, dose, and frequency. Additionally, radiotherapy is often delivered in multiple phases, each treating a different anatomical site to different doses and with different types of radiation, yielding complex descriptions of treatment courses. These features, in concert with a lack of widely used standardized nomenclatures (Mir et al., 2020; Phillips et al., 2020; Traverso et al., 2018), limit manual data extraction, hindering the potential of

big data to improve cancer research and clinical care.

While algorithms for named entity recognition have previously been reported for radiotherapy details (Bitterman et al., 2020) and other cancer therapies (Yin et al., 2018; Wang et al., 2019; Yim et al, 2016; Savova et al., 2019), relation extraction remains a relatively underexplored task in clinical NLP (Sheikhalishahi et al., 2019). There are few examples of relation extraction models for cancer characteristics in general (Bozkurt et al., 2016; Savova et al., 2017; Sheikhalishahi et al., 2019), and to the best of our knowledge none for cancer treatment, including radiotherapy. Identifying treatment entities in isolation and not linking them to a specific treatment instance is insufficient to coherently describe cancer therapies, especially because concurrent and serial treatments are often described together in the same note. For example, extracting frequency without linking it to a specific treatment is not informative by itself. Relation extraction is a critical component of information extraction for radiotherapy as this treatment is often given in multiple sequential or nested phases. Linking relevant treatment entities with the same phase is necessary to summarize how and why a treatment was delivered, necessary components to understanding treatment outcomes and quality. However, this is a very challenging task even for expert human annotators, as demonstrated by the challenges in accurate extraction of such data for SEER. Therefore, there is a need for more reliable relation extraction methods for radiotherapy.

Relations can be modeled in various ways in neural networks, including inserting special tokens around the arguments of interest and using this augmented text as input into the model (Dligach et al., 2017), and using token position embeddings to encode the relative distance of words to the arguments (Zeng et al., 2014; Nguyen and Grishman et al., 2015; Shi and Lin et al., 2019; Wang et al., 2019). Using the former approach, we aimed to investigate several approaches for relation extraction from clinical texts describing radiotherapy, with a goal of augmenting reporting of cancer treatment details for research and clinical purposes.

The contributions of the work described in the paper are (1) the definition of the task of radiotherapy information extraction from the EMR clinical narrative, (2) the creation of resources for the task (annotation guidelines and corpus), (3) the exploration of state-of-the-art neural methods to this highly impactful clinical task, and (4) establishing a baseline for the task.

2 Data

Data for this work consisted of texts describing radiotherapy from three complementary sources. First, we included 270 clinical descriptions of radiotherapy regimens from HemOnc.org, which is a publicly available wiki of cancer and blood disorder treatment regimens and interventions. Second, we included 73 radiotherapy descriptions from a state cancer registry. These are abstractions from patients' EMR, often copied and pasted from clinician notes, describing details of cancer treatment for use by cancer registrars within a patient-level XML. We used the entire text of the XML categories that contained radiotherapy details as model input. Third, 79 completed breast and colorectal cancer clinician notes that contained radiotherapy details from the THYME corpus from Clinical TempEval (Bethard et al., 2015) and an internal corpus with breast cancer notes.

Annotation guidelines for radiotherapy properties and treatment instances were developed (Bitterman et al., 2020).[1] If an overall radiotherapy treatment course was delivered in more than one

---

Dose-Treatment Site relationship is represented as:

…She presented after a screening mammogram showed a nodule in the *left breast upper outer quadrant*. After lumpectomy, she was treated with radiation to a dose of 50 Gy in 25 fractions to the *left breast*, followed by a boost of RT_DosageStart **10 Gy** RTDosage_End in 5 fractions to the TxSiteStart **tumor bed** TxSiteEnd…

Figure 1: Mock text segment describing radiotherapy with start and stop tokens, illustrating the Dose-Treatment Site relationship between the two bolded entities: 10 Gy (Dose) and tumor bed (Treatment Site). The zigzag and dashed lines indicate adjacent spans describing two different radiotherapy instances, and italicized entities are non-related anatomical/treatment sites close to the anchor dose mention. Both of these

---

[1]https://github.com/RTParse/RTAnnot

Relation	IAA
Dose-Dose	0.94
Dose-Treatment Site	0.90
Dose-Frequency	1.00
Dose-Fraction Number	0.98
Dose-Boost	0.67
None	0.74

Table 1: Inter-Annotator Agreement (IAA) for the six relation categories.

phase, as described above, each phase was considered a separate radiotherapy instance (Figure 1). Gold annotations for relations between the following key properties were created: Dose, Fraction Number, Fraction Frequency, Treatment Site, and Boost. Radiotherapy is most often delivered in many small doses, or fractions, over a given period of time. Dose is any description of radiation dosage in the text, either the total dose or fractional doses, generally described using the unit Gray (Gy). Fraction Number is any mention of the number of fractions delivered, and Fraction Frequency is the frequency of fraction delivery. Treatment site is the actual or relative anatomical site that is targeted with radiotherapy. Boost is a mention that conveys the treatment instance is a second phase of radiotherapy that brings a smaller treatment site to a higher dose. Properties that described the same treatment instance were linked together as a relation.

Two expert human annotators completed gold annotations for 47 radiotherapy instances containing 310 relation instances to calculate inter-annotator agreement, after which a single human annotator completed the gold annotations. As most radiotherapy instances included Dose, we chose to classify the relations between each Dose mention and every other property mentioned in the radiation instance. Thus, Dose mentions served as the anchor for the relations within a radiotherapy instance and were labeled as Dose-Mention. Of note, Dose-Dose refers to a relation of two different Dose mentions in the same radiotherapy instance. In our dataset, there were on average 1.4 radiation instances per document. Documents were split into train, development, and test sets. The gold annotated HemOnc.org corpus will be made publicly available for research purposes.

3 Methods

We explored two state-of-the-art neural network methods for this task. First, we used Flair (Akbik

et al,, 2018), which is a pre-trained character language model trained on one billion words of text (Chelba et al., 2013) to train a multi-layer Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997) to generate contextual embeddings. Text is then passed into this model, put into an LSTM to obtain a text representation, and this is passed into a final linear layer for classification. We were motivated to explore this approach as we hypothesized that the contextual character embedding may be better at handling the rare and misspelled words in cancer texts, as well as numbers and short abbreviations common in radiation descriptions. Models had a hidden state of 128 and a dropout of 0.15-0.24, and were trained for 100 epochs using a mini-batch size of 8 and learning rate starting at 0.2 with an anneal factor of 0.5, using an SGD optimizer (Robbins and Monro, 1951; Kiefer and Wolfowitz, 1952; Bottou et al., 2016). Second, we assessed the performance of bidirectional encoder representations from transformers (BERT) base uncased model, fine-tuned on this relation task using a recurrent neural network to predict the class label (Devlin et al.; 2019). We chose to explore this method given the excellent relation classification performance of attention-based models in biomedical texts (Verga et al., 2018; Wei et al., 2019; Lee et al., 2020). These models had a hidden size of 128 and dropout of 0.5, and were trained for 30 epochs using a mini-batch size of 8 and learning rate starting at 0.00003 with an anneal factor of 0.5, using an Adam optimizer (Kingma and Ba, 2014). For all, the model that performed best on the development set was evaluated on the held-out test set. Rule-based methods were considered, but as there are often more than one radiotherapy instance mentioned in a clinical text, frequently in close proximity and described in nested fashion, we did not feel there was a straightforward approach (Figure 1).

To generate candidate relations, we extracted text windows encompassing two different token lengths on each side of the gold annotated Dose mention anchor: 46 tokens and 90 tokens (95th and 99th percentile of token span lengths between Dose-Mentions in the same gold radiotherapy instances in the train and development sets, respectively). Every Dose-Mention pair in the text window was considered a candidate relation. For each relation candidate, start and stop tokens were inserted around the Dose and candidate property (Figure 1). The text segment was labeled with a

Flair/BERT Models	Precision		Recall		F1	
	92 Token Windows	180 Token Windows	92 Token Windows	180 Token Windows	92 Token Windows	180 Token Windows
Dose-Dose	0.77/0.88	0.79/0.68	0.75/0.82	0.83/0.81	0.76/0.85	0.81/0.74
Dose-Treatment Site	0.84/0.79	0.61/0.76	0.86/0.88	0.92/0.89	0.85/0.83	0.73/0.82
Dose-Fraction Frequency	0.79/0.78	0.84/0.88	0.95/0.90	1.00/1.00	0.86/0.84	0.91/0.93
Dose-Fraction Number	0.95/0.94	0.90/0.87	0.93/0.96	0.92/0.95	0.94/0.95	0.91/0.91
Dose-Boost	1.00/0.85	0.56/0.60	0.69/0.85	0.69/0.92	0.82/0.85	0.62/0.73
None	0.95/0.95	0.98/0.98	0.95/0.94	0.94/0.95	0.95/0.95	0.96/0.96
Average	0.88/0.87	0.78/0.80	0.86/0.89	0.88/0.92	0.86/0.88	0.82/0.85

Table 2: Results of the two Flair and two BERT relation classification models, developed using two different text windows around the anchor Dose entity. The first result is from Flair, the second – from BERT (Flair/BERT)

positive Dose-Property relation if the Dose and candidate property were in the same radiotherapy instance.

The precision (True Positives/Predictions), recall (True Positives/Gold Positives), and F1 score ((2*Precision*Recall)/(precision+recall)) are reported for each model. Error analysis via manual inspection was carried out to better understand how and where the models performed poorly.

4 Evaluation

Table 1 shows the IAA for each of the relation categories. IAA was ≥ 0.9 for all categories except for Dose-Boost (0.67) and None (0.74). Of note, there were only 6 total Boost mentions in the pilot dataset, limiting interpretation of this IAA.

Table 2 shows the performance of the Flair and BERT relation classification models. Overall, the best performing model was the BERT model fine-tuned on the 92 token text windows (macro-average F1: 0.88), followed by the Flair model fine-tuned on the 92 token text windows (macro-average F1: 0.86). Precision for the Flair model with this window was on average slightly higher than that for the BERT model. The models fine-tuned on the 180 text windows had slightly worse performance (BERT model macro-F1: 0.85, Flair model macro-F1: 0.82).

Qualitative error analysis of each model revealed five main categories of failure: 1) human errors in gold labeling, 2) false positives due to two mentions being in close proximity despite relating to different treatment instances, 3) false negatives

Error Type	Text	Gold Label	Predicted Label
Human error in gold labeling	"…2 Gy fractions **once per da**y x 15 fractions (total dose: **30 Gy**)…"	None	Dose-Frequency
False positive due to two mentions in close proximity	"…Synchronously, PTV1 and PTV2 received **45** and **50.4 Gy**, respectively..."	None	Dose-Dose
False negative due to distant related mentions	"…**1.5 Gy** fractions x 16 fractions, given twice per day (4 to 6 hour interval between treatments) on days 1 to 5, 8 to 10. Total dose during consolidation is **24 Gy**..."	Dose-Dose	None
Incorrect label in texts with atypical treatment description	"...additional treatment up to a total of **55.8 Gy** was administered to original bony tumors and the postinduction chemotherapy soft tissue volumes plus a **2 cm margin**..."	Dose-Treatment Site	None

Table 3: Examples of error types identified in qualitative error analysis. The mentions being classified are bolded in the text.

because related mentions were distant and/or crossing sentence boundaries 4) incorrect labeling in texts with very atypical descriptions of treatment, 5) incorrect labeling in tabular text, and 6) other/unknown. Examples of errors are shown in Table 3. All models suffered from similar methods of failure, and texts describing radiotherapy courses with several phases appeared to be particularly challenging. Interestingly, the BERT model fine-tuned on the 180 token text windows was better able to correctly label relations in tables than the other models, although there were only rare examples of tables in these corpora.

Additionally, in the Flair models using the 92 and 180 token window texts, there were 3 and 8 true positive relations, respectively, that were labeled with an incorrect label other than "None". All incorrect labels in the BERT models were either a "None" label assigned to a true relation, or a false positive relation assigned to a true "None" relation.

5   Discussion and Conclusion

The neural models had very good performance on the radiotherapy relations explored in these experiments, often approaching or exceeding IAA. The performance of the BERT and Flair models were overall comparable, with the best BERT model outperforming the best Flair model. Interestingly, the size of the text windows used for fine-tuning appeared to have a larger impact on performance than the type of model itself, with shorter texts yielding better results. This may be due to the frequent repetition of similar but unrelated entities and treatments in clinical texts, and optimizing this parameter should be explored when developing models for clinical relation extraction.

These findings suggest that neural methods may be a good avenue for clinical relation extraction for complex, highly specialized treatments such as radiotherapy. Future work will develop models to extract relations between Dose and additional relevant entities, and will investigate end-to-end entity and relation extraction systems for robust information extraction pipelines.

References

Alan Akbik, Duncan Blythe, and Roland Vollgraf. 2018. Contextual string embeddings for sequence labeling. *In Proceedings of the 27th International Conference on Computational Linguistics*, pages 1638–1649. Retrieved from https://www.aclweb.org/anthology/C18-1139/

Danielle Bitterman, Timothy Miller, David Harris, Chen Lin, Sean Finan, Jeremy Warner, Ray Mak, Guergana Savova. 2020. Extracting radiotherapy treatment details using neural network-based natural language processing. *Presented at the Poster to be presented at the 62nd Annual Meeting of the American Society for Radiation Oncology*.

Steven Bethard, Leon Derczynski, Guergana Savova, James Pustejovsky, and Marc Verhagen. 2015. Semeval-2015 task 6: Clinical tempeval. *In Proceedings of the 9th international workshop on semantic evaluation (SemEval 2015)*, pages 806–814. Retrieved from https://www.aclweb.org/anthology/S15-2136.pdf

Léon Bottou, Frank E. Curtis, and Jorge Nocedal. 2016. Optimization Methods for Large-Scale Machine Learning. June. *arXiv [stat.ML].* Retrieved from arXiv.

Selen Bozkurt, Jafi A. Lipson, Utku Senol, and Daniel L. Rubin. 2015. Automatic abstraction of imaging observations with their characteristics from mammography reports. *Journal of the American Medical Informatics Association: JAMIA*, 22(e1):e81–92, April. Retrieved from http://dx.doi.org/10.1136/amiajnl-2014-003009

Ciprian Chelba, Tomas Mikolov, Mike Schuster, Qi Ge, Thorsten Brants, Phillipp Koehn, and Tony Robinson. 2013. One Billion Word Benchmark for Measuring Progress in Statistical Language Modeling. December. *arXiv [cs.CL].* Retrieved from arXiv.

Geoff Delaney, Susannah Jacob, Carolyn Featherstone, and Michael Barton. 2005. The role of radiotherapy in cancer treatment: estimating optimal utilization from a review of evidence-based clinical guidelines. *Cancer*,

104(6):1129–1137, September. Retrieved from http://dx.doi.org/10.1002/cncr.21324

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. October. *arXiv [cs.CL]*. Retrieved from arXiv.

Dmitriy Dligach, Timothy Miller, Chen Lin, Steven Bethard, and Guergana Savova. 2017. Neural temporal relation extraction. *In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pages 746–751. Retrieved from https://www.aclweb.org/anthology/E17-2118.pdf

Leonard Gunderson MD MS and Joel E. Tepper. 2015. *Clinical Radiation Oncology*. Elsevier Health Sciences, August. Retrieved from https://play.google.com/store/books/details?id=NmFyCgAAQBAJ

HemOnc.org - A Hematology Oncology Wiki. Retrieved June 18, 2020, from http://hemonc.org

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long Short-Term Memory. *Neural Computation*. Retrieved from http://dx.doi.org/10.1162/neco.1997.9.8.1735

J. Kiefer and J. Wolfowitz. 1952. Stochastic Estimation of the Maximum of a Regression Function. *Annals of Mathematical Statistics*, 23(3):462–466, September. Retrieved August 10, 2020, from https://projecteuclid.org/euclid.aoms/1177729392

Diederik P. Kingma and Jimmy Ba. 2014. Adam: A Method for Stochastic Optimization. December. *arXiv [cs.LG]*. Retrieved from arXiv.

Jinhyuk Lee, Wonjin Yoon, Sungdong Kim, Donghyeon Kim, Sunkyu Kim, Chan Ho So, and Jaewoo Kang. 2020. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, 36(4):1234–1240, February. Retrieved from https://doi.org/10.1093/bioinformatics/btz682

THYME. Retrieved August 9, 2020, from http://thyme.healthnlp.org

Patrick N. McDermott and Colin G. Orton. 2010. *The physics & technology of radiation therapy.* Medical Physics Publishing.

Romaana Mir, Sarah M. Kelly, Ying Xiao, Alisha Moore, Catharine H. Clark, Enrico Clementel, Coreen Corning, Martin Ebert, Peter Hoskin, Coen W. Hurkmans, Satoshi Ishikura, Ingrid Kristensen, Stephen F. Kry, Joerg Lehmann, Jeff M. Michalski, Angelo F. Monti, Mitsuhiro Nakamura, Kenton Thompson, Huiqi Yang, et al. 2020. Organ at risk delineation for radiation therapy clinical trials: Global Harmonization Group consensus guidelines. *Radiotherapy and oncology: journal of the European Society for Therapeutic Radiology and Oncology*, June. http://dx.doi.org/10.1016/j.radonc.2020.05.038

Thien Huu Nguyen and Ralph Grishman. 2015. Relation extraction: Perspective from convolutional neural networks. *In Proceedings of the 1st Workshop on Vector Space Modeling for Natural Language Processing*, pages 39–48. Retrieved from https://www.aclweb.org/anthology/W15-1506.pdf

Mark H. Phillips, Lucas M. Serra, Andre Dekker, Preetam Ghosh, Samuel M. H. Luk, Alan Kalet, and Charles Mayo. 2020. Ontologies in radiation oncology. *Physica medica: PM: an international journal devoted to the applications of physics to medicine and biology: official journal of the Italian Association of Biomedical Physics* , 72:103–113, April. Retrieved from http://dx.doi.org/10.1016/j.ejmp.2020.03.017

Herbert Robbins and Sutton Monro. 1951. A Stochastic Approximation Method. *Annals of Mathematical Statistics*, 22(3):400–407, September. Retrieved August 10, 2020, from http://projecteuclid.org/euclid.aoms/1177729586

Marta Recasens, Christian Danescu-Niculescu-Mizil, and Dan Jurafsky. 2013. Linguistic Models for Analyzing and Detecting Biased Language. *In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, pages 1650–1659, Sofia, Bulgaria.

Guergana K. Savova, Eugene Tseytlin, Sean Finan, Melissa Castine, Timothy Miller, Olga Medvedeva, David Harris, Harry Hochheiser, Chen Lin, Girish Chavan, and Rebecca S. Jacobson. 2017. DeepPhe: A Natural Language

Processing System for Extracting Cancer Phenotypes from Clinical Records. *Cancer research*, 77(21):e115–e118, November. Retrieved from http://dx.doi.org/10.1158/0008-5472.CAN-19-0579

Seyedmostafa Sheikhalishahi, Riccardo Miotto, Joel T. Dudley, Alberto Lavelli, Fabio Rinaldi, and Venet Osmani. 2019. Natural Language Processing of Clinical Notes on Chronic Diseases: Systematic Review. *JMIR medical informatics*, 7(2):e12239, April. Retrieved from http://dx.doi.org/10.2196/12239

Peng Shi and Jimmy Lin. 2019. Simple BERT Models for Relation Extraction and Semantic Role Labeling. April. *arXiv [cs.CL]*. Retrieved from arXiv.

Benjamin D. Smith, Bruce G. Haffty, Lynn D. Wilson, Grace L. Smith, Akshar N. Patel, and Thomas A. Buchholz. 2010. The future of radiation oncology in the United States from 2010 to 2020: will supply keep pace with demand? *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*, 28(35):5160–5165, December. Retrieved from http://dx.doi.org/10.1200/JCO.2010.31.2520

Surveillance, Epidemiology, and End Results Program. (n.d.). . Retrieved August 9, 2020, from https://seer.cancer.gov/

Khajamoinuddin Syed, William Sleeman Iv, Kevin Ivey, Michael Hagan, Jatinder Palta, Rishabh Kapoor, and Preetam Ghosh. 2020. Integrated Natural Language Processing and Machine Learning Models for Standardizing Radiotherapy Structure Names. *Healthcare (Basel, Switzerland)*, 8(2), April. Retrieved from http://dx.doi.org/10.3390/healthcare8020120

Alberto Traverso, Johan van Soest, Leonard Wee, and Andre Dekker. 2018. The radiation oncology ontology (ROO): Publishing linked data in radiation oncology using semantic web and ontology techniques. *Medical physics*, 45(10):e854–e862, October. Retrieved from http://dx.doi.org/10.1002/mp.12879

Patrick Verga, Emma Strubell, and Andrew McCallum. 2018. Simultaneously Self-Attending to All Mentions for Full-Abstract Biological Relation Extraction. February. *arXiv [cs.CL]*. Retrieved from arXiv.

Gary Walker, Ergin Soysal, and Hua Xu. 2019. Development of a Natural Language Processing Tool to Extract Radiation Treatment Sites. *Cureus*, 11(10):e6010, October. Retrieved from http://dx.doi.org/10.7759/cureus.6010

Haoyu Wang, Ming Tan, Mo Yu, Shiyu Chang, Dakuo Wang, Kun Xu, Xiaoxiao Guo, and Saloni Potdar. 2019. Extracting Multiple-Relations in One-Pass with Pre-Trained Transformers. February. *arXiv [cs.CL]*. Retrieved from arXiv.

Liwei Wang, Lei Luo, Yanshan Wang, Jason Wampfler, Ping Yang, and Hongfang Liu. 2019. Natural language processing for populating lung cancer clinical research data. *BMC medical informatics and decision making*, 19(Suppl 5):239, December. Retrieved from http://dx.doi.org/10.1186/s12911-019-0931-8

Qiang Wei, Zongcheng Ji, Yuqi Si, Jingcheng Du, Jingqi Wang, Firat Tiryaki, Stephen Wu, Cui Tao, Kirk Roberts, and Hua Xu. 2019. Relation Extraction from Clinical Narratives Using Pre-trained Language Models. *AMIA … Annual Symposium proceedings / AMIA Symposium. AMIA Symposium*, 2019:1236–1245.

Wen-Wai Yim, Meliha Yetisgen, William P. Harris, and Sharon W. Kwan. 2016. Natural Language Processing in Oncology: A Review. *JAMA oncology*, 2(6):797–804, June. Retrieved from http://dx.doi.org/10.1001/jamaoncol.2016.0213

Zhijun Yin, Morgan Harrell, Jeremy L. Warner, Qingxia Chen, Daniel Fabbri, and Bradley A. Malin. 2018. The therapy is making me sick: how online portal communications between breast cancer patients and physicians indicate medication discontinuation. *Journal of the American Medical Informatics Association: JAMIA*, 25(11):1444–1451, November. Retrieved from http://dx.doi.org/10.1093/jamia/ocy118

Daojian Zeng, Kang Liu, Siwei Lai, Guangyou Zhou, and Jun Zhao. 2014. Relation classification via convolutional deep neural network. *In Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pages 2335–2344. Retrieved from https://www.aclweb.org/anthology/C14-1220.pdf