Named Entity Recognition for Cancer Immunology Research Using Distant Supervision

Hai-Long Trieu [1,3]**, Makoto Miwa** [1,2] and **Sophia Ananiadou** [3]
[1]Artificial Intelligence Research Center (AIRC),
National Institute of Advanced Industrial Science and Technology (AIST), Japan
[2]Toyota Technological Institute, Japan
[3]National Centre for Text Mining, University of Manchester, United Kingdom
long.trieu@aist.go.jp, makoto-miwa@toyota-ti.ac.jp,
sophia.ananiadou@manchester.ac.uk

Abstract

Cancer immunology research involves several important cell and protein factors. Extracting the information of such cells and proteins and the interactions between them from text are crucial in text mining for cancer immunology research. However, there are few available datasets for these entities, and the amount of annotated documents is not sufficient compared with other major named entity types. In this work, we introduce our automatically annotated dataset of key named entities, i.e., T-cells, cytokines, and transcription factors, which engages the recent cancer immunotherapy. The entities are annotated based on the UniProtKB knowledge base using dictionary matching. We build a neural named entity recognition (NER) model to be trained on this dataset and evaluate it on a manually-annotated data. Experimental results show that we can achieve a promising NER performance even though our data is automatically annotated. Our dataset also enhances the NER performance when combined with existing data, especially gaining improvement in yet investigated named entities such as cytokines and transcription factors.

1 Introduction

Cancer immunology research has a central focus on T lymphocytes (*T-cells*), which engage the immune system in fighting against cancer (Luckheeram et al., 2012; Waldman et al., 2020; Kim et al., 2021). The development of T-cells can be guided by *cytokines* and *transcription factors* (Hosokawa and Rothenberg, 2018). *Transcription factors* (TF) are nuclear proteins that bind specific gene sequences and involved in decision-making processes during T-cell differentiation (Naito et al., 2011; Xia et al., 2019). Meanwhile, *cytokines* are signaling molecules secreted and sensed by immune and other cell types (Kveler et al., 2018). Extracting *T-cell*, *cytokine*, and *TF* entities and the interactions between them can be crucial for text mining in cancer immunology research.

However, there are few existing datasets containing these entities to train text mining models. At the core of text mining tasks, the named entity recognition (NER) task also lacks such datasets for training NER models to detect these named entities, which may limit the development of text mining systems in this cancer immunology research field. There is an existing T-cell related named entity dataset called TCRE (Czech and Hammerbacher, 2019), but the amount of annotated data is also limited to only 89 documents. Several knowledge bases related to immune system have been proposed such as immuneXpresso (Kveler et al., 2018) and DES-Tcell (AlSaieedi et al., 2021), which contain cell type and cytokine information, but they lack utilizing and evaluating with modern NER models on these named entities.

In this paper, as a step to fill these gaps and promote the development of text mining systems on these named entities in cancer immunology research articles, we present our automatically annotated dataset containing named entities of *T-cell*, *cytokine* and *TF*, which are important for mining and understanding cancer immunology research articles. The entities in the dataset are automatically annotated using dictionary matching based on the UniProtKB (UniProt-Consortium, 2021), a knowledgebase of protein sequences with functional information.[1] From the annotations of cytokine and TF entries in UniProtKB, a dictionary is constructed to annotate cytokine and TF named entities in their referenced PubMed articles. Additionally, we utilized the existing JNLPBA corpus, which contains manually annotated protein named entities, to annotate cytokine and TF entities. We build a NER model based on the span-based model with pre-trained BERT. We trained the NER model on our automatically annotated dataset and evaluated the model on an existing manually annotated T-cell related named entity TCRE dataset (Czech

---

[1]https://www.uniprot.org/uniprot/

171

Item	cytokine	TF
# UniProtKB entries	1,001	3,418
# Dictionary size	6,859	20,055
# Collected articles	585	1,903

Table 1: UniProtKB entries and annotated data

Data	#Docs.	#Entities		
		CT	CY	TF
KB-T-cell	386	340	744	2,891
Dic-T-cell	761	2,686	1,752	2,686
TCRE	89	1,006	235	114

Table 2: Statistics of the datasets (*Docs*: documents; CT (cell type), CY (cytokine), TF (transcription factor))

and Hammerbacher, 2019). We achieve a promising result that the NER model trained on our automatically annotated data gains a slightly lower performance than a supervised NER model trained on a manually annotated data, although our data is automatically annotated. Furthermore, our data enhances NER performance when combined with the existing manually annotated data.

2 Approach

We present our datasets containing three named entity types: cell_type, cytokine, and transcription factor (TF). The datasets are automatically annotated using dictionary matching with the entries in the UniProtKB in two different ways.

2.1 UniProtKB

Cytokine and TF queries From the UniProtKB, we obtain entries by querying *cytokine*. We filtered the options to keep only *Reviewed* annotations (manually annotated, added by expert biocuration team) and for *Human* organism. Similarly, we conducted for *transcription factor*. They are equivalent to the following queries.

- *cytokine AND reviewed:yes AND organism:"Homo sapiens (Human) [9606]"*.

- *transcription factor AND reviewed:yes AND organism:"Homo sapiens (Human) [9606]"*

UniProtKB entries We obtained 1,001 entries for cytokine and 3,418 entries for TF from UniProtKB. Each entry contains protein names, gene names, and referenced PubMed articles, etc.

UniProtKB-dictionary We built a dictionary containing protein and gene names of the cytokine and TF entries in UniProtKB, which we named *UniProtKB-dictionary*.

Collecting PubMed references For each UniProtKB entry, there is a list of referenced PubMed articles. We collect the referenced articles' abstract texts from PubMed for each entry. Since there is a large number of references, we only collect the

abstracts that contain a large number ($\geq k$) of cytokine/TF protein and gene names (we set $k = 20$, which we based on several preliminary experiments to remove abstracts containing few annotations). We present the statistics of UniProtKB entries and related annotated data in Table 1.

2.2 Automatically Annotated Datasets

We constructed two automatically annotated datasets using the *UniProtKB-dictionary*. The statistics for automatically annotated datasets are presented in Table 2.

2.2.1 Knowledge-based Annotation (KB-T-cell)

Annotating cytokine and TF From the *UniProtKB dictionary*, we identify the position of each name in the collected articles by strict text matching to annotate cytokine and TF named entities.

Annotating cell_type We found that JNLPBA (Collier and Kim, 2004) is a large manually annotated dataset for NER, which contains named entities of cell_type, protein, etc. Therefore, we utilized the JNLPBA data to train a NER model to predict cell_type named entities in the collected articles. We build a neural-based NER method with span-based and pre-trained BERT model, which we present in §3. These cell_type entities are combined with the cytokine and TF named entities, and we named *KB-T-cell*.

2.2.2 Dictionary-based Re-annotation (Dic-T-cell)

Since the JNLPBA dataset contains protein entities while CT and TF are proteins, we utilized the annotated protein names in the JNLPBA to annotate cytokine and TF entities. Specifically, if an annotated protein name in the JNLPBA is included in the *UniProtKB-dictionary*, we re-annotate it as cytokine or TF, correspondingly. We ignored documents which do not contain any matched CT/TF entity. We named this dataset as *Dic-T-cell*.

3 NER model

We explain the NER model to be trained on the annotated datasets. We build a neural-based NER model using a span-based method (Lee et al., 2017; Luan et al., 2018) and finetuned pre-trained BERT (Devlin et al., 2019). Specifically, each sentence is split into sub-word sequences, which are passed through the BERT layer for contextual representations. Then, for each span (i.e., a sequence of continuous words in a sentence), its representation is calculated by concatenating the representations of the first, last, and averaged sub-words of the span, which follows (Sohrab and Miwa, 2018a; Trieu et al., 2020). Finally, each span representation is passed to classifiers to predict named entity types for each span.

4 Experiments

4.1 Data

We used our datasets *KB-T-cell* and *Dic-T-cell* to train NER models using the NER model introduced in §3 and evaluated NER performance.

TCRE   For evaluation data, we employed the TCRE (Czech and Hammerbacher, 2019), an existing manually annotated data which contains 89 documents of cell_type, cytokine, and TF named entities. We utilized this data for training supervised NER models and for evaluation. The original TCRE dataset contains a mixture of both abstract and full-text documents. For the scope of this paper, we aim at utilizing only abstracts from both UniProtKB's references and JNLPBA data. Therefore, we used only the abstract documents and the abstract section of full-text documents from the TCRE data.

The data statistics of the datasets are presented in Table 2.

4.2 Settings

Cross validation   We conducted k-fold cross validation evaluation on the TCRE dataset. Since the TCRE data size is quite small, we set $k = 3$ to ensure a reasonable amount of data in the test set. For each fold, we further randomly split the training set into train/development sets so that we can tune hyper-parameters to get the best models on the development set. Finally, all of our reported results are based on the TCRE test set in each fold.

NER training settings   Our model was implemented on PyTorch (Paszke et al., 2017). We used the BERT model from the PyTorch Pretrained BERT repository[2] as our BERT layer. We employed the pre-trained SciBERT model (Beltagy et al., 2019) trained on large-scale biomedical texts. The model is trained on multiple GPUs in the AI Bridging Cloud Infrastructure (ABCI)[3]. We train the model with the Adam optimizer (Kingma and Ba, 2015), gradient clipping, dropout, and L2 regularization. The model is trained with early-stopping, and the training mini-batch size is set as 16.

Evaluation settings   We compared the following NER models, which mostly differ in the training data settings.

1. **Matching-NER**: we created a baseline using dictionary matching. The dictionary is built from the entity's texts of the JNLPBA training data (for cell_type) and the UniProtKB-dictionary for cytokine and TF.

2. **Supervised-NER**: we used the training set of the TCRE data to train the NER model.

3. **KB-NER**, **Dic-NER**, **KB-Dic-NER**: we train the NER models on our annotated datasets: KB-T-cell, Dic-T-cell, and merged the KB-T-cell and Dic-T-cell, respectively.

4. **Enhanced-KB-NER**, **Enhanced-Dic-NER**, **Enhanced-KB-Dic-NER**: we merge the training set of the TCRE with the KB-T-cell, Dic-T-cell, and merged KB-T-cell and Dic-T-cell, respectively, to train NER models.

The results are reported based on the commonly used micro-averaged precision (P), recall (R), and F-score (F) metrics at entity level.

4.3 Results

We compare the results of different NER models on each data fold in Table 3.

Enhancement   Using our automatically annotated dataset, we achieved the best performance with 2-5% point improvements in F-score (Enhanced-KB-NER) in comparison with the Supervised-NER in all of the data folds.

---

[2] https://github.com/huggingface/pytorch-pretrained-BERT/tree/34cf67fd6c
[3] https://abci.ai/

Model	Fold-1			Fold-2			Fold-3		
	P	R	F	P	R	F	P	R	F
Matching-NER	39.88	66.16	49.76	39.54	68.63	50.17	38.05	69.27	49.12
Supervised-NER	68.67	66.92	67.78	70.92	70.75	70.84	73.36	74.23	73.80
KB-NER	64.55	62.09	63.29	71.34	54.01	61.48	63.85	57.21	60.35
Dic-NER	63.19	61.58	62.37	66.67	60.38	63.37	67.00	64.30	65.62
KB-Dic-NER	65.33	66.16	65.74	71.74	62.26	66.67	67.07	65.48	66.27
Enhanced-KB-NER	**72.98**	**73.54**	**73.26**	**75.12**	**76.89**	**75.99**	**75.71**	**75.89**	**75.80**
Enhanced-Dic-NER	71.11	72.02	71.55	70.14	73.11	71.59	73.23	75.65	74.42
Enhanced-KB-Dic-NER	72.18	73.28	72.73	72.86	72.17	72.51	74.13	75.18	74.65

Table 3: Comparison NER results of the models (the best scores are in bold)

Model	Fold-1			Fold-2			Fold-3		
	CT	CY	TF	CT	CY	TF	CT	CY	TF
Matching-NER	65.18	1.45	15.07	66.42	6.00	18.44	65.96	6.86	5.97
Supervised-NER	71.22	56.64	41.18	76.36	56.36	32.14	76.15	65.45	57.78
KB-NER	69.57	31.46	<u>52.38</u>	73.70	18.95	0.00	70.79	20.95	8.00
Dic-NER	<u>72.81</u>	3.33	0.00	<u>79.50</u>	5.56	3.03	76.00	13.19	0.00
KB-Dic-NER	<u>73.62</u>	22.54	35.29	<u>79.21</u>	8.33	0.00	<u>78.06</u>	18.69	8.00
Enhanced-KB-NER	<u>76.32</u>	<u>62.50</u>	<u>63.77</u>	<u>82.16</u>	<u>60.66</u>	<u>43.48</u>	<u>80.65</u>	<u>68.91</u>	21.74
Enhanced-Dic-NER	<u>77.49</u>	55.32	18.18	<u>81.61</u>	36.51	<u>39.44</u>	<u>79.21</u>	60.34	27.03
Enhanced-KB-Dic-NER	<u>77.55</u>	<u>64.08</u>	30.77	<u>81.33</u>	41.44	<u>37.68</u>	<u>80.06</u>	63.64	15.79

Table 4: Results on each entity type in F-score (%). The underline scores are higher than the Supervised-NER's.

Supervised vs. unsupervised When training NER models on our automatically annotated datasets (KB-NER, Dic-NER, KB-Dic-NER), the performance is lower than the Supervised-NER, which is trained on a time-consuming manually annotated data. The degraded performance is about 5-7% points in F-score, which are acceptable considering that our datasets are automatically annotated. We can further improve the quality of our datasets in future work, such as filtering noisy annotations.

Dictionary matching Since our automatically annotated data is based on the dictionary built from the UniProtKB and JNLPBA, we may raise a question whether using only the dictionary with the same vocabulary is still enough. The results of KB-NER and Dic-NER show that our automatically annotated data can improve from 11-15% in comparison with the Matching-NER.

KB vs Dic Table 3 also shows that the NER models based on the KB-T-cell (KB-NER, Enhanced-KB-NER) obtain higher performance than those based on the Dic-T-cell (Dic-NER, Enhanced-Dic-NER). When combining these two datasets, the performance decreased even though the data size

of the Dic-T-cell is mostly double of the KB-T-cell, which indicates that we need to investigate a better combination. Another possible direction can be filtering noisy annotations of the Dic-T-cell.

4.4 Analyses and Discussions

We further investigate the detailed performance on each entity type: cell_type, cytokine, and TF. The results from Table 4 show that the Enhanced-KB-NER achieves improvements on all entity types except for the TF entity type in Fold-3.

Comparing the performance among the entity types between the Supervised-NER and the enhanced models, the CT type performance gains improvement (3-5% points) in most cases. The reason may come from the quality of the CT type in the large manually annotated JNLPBA data. Meanwhile, the improvement of the CY type is 3-6% points, and the improvement of TF is 11-22% points. When training only on our automatically annotated datasets (KB-NER, Dic-NER), we still obtain the higher performance for the CT type. We obtain some reasonable performance in cytokines (lower than the Supervised-NER but much better than the Matching-NER).

Limitation The performance of CY and TF from KB-NER and Dic-NER is low in most cases. There is no correct TF prediction (Dic-NER in Fold-1 and Fold-3, KB-NER in Fold-2). For CY, the performance is also low from Dic-NER (3% to 13% F-score), but it is slightly better in KB-NER (18% to 31% F-score). These results show a challenge to extract CY and TF entities based on only our automatically annotated corpus. This work is our first investigation in utilizing the UniProtKB and the existing JNLPBA corpus for our research goal in extracting T-cell related entities, and we accept this limitation in this first version. It is required to conduct further investigation and improvement especially for these CY and TF types in future work.

Future work We would like to improve the performance of CY and TF. We also plan to conduct the evaluation not only on the TCRE task but other NER tasks such as JNLPBA (Collier and Kim, 2004), NCBI (Doğan et al., 2014), and BC5CDR (Li et al., 2016). Additionally, we intend to extend our corpus for other tasks such as relation and event extraction on these T-cell named entities.

5 Related Work

Distant supervision methods for NER have been investigated in several previous works. (Shang et al., 2018) revised the LSTM-CRF NER model (Lample et al., 2016) and utilized the MeSH database for chemical and disease entities. Some methods are proposed to reduce noisy annotations for Chinese NER(Yang et al., 2018), or general domain OntoNotes (Liang et al., 2020; Meng et al., 2021).

The span-based method has been used to build our NER model in this work. The method was proposed and employed in previous work (Lee et al., 2017; Luan et al., 2018; Sohrab and Miwa, 2018b; Trieu et al., 2020), which have shown the advantages in extracting nested or continuous text sequences and successful in many sequence labeling tasks such as NER or coreference resolution.

Immunotherapy has achieved remarkable advances in recent years and can be important cancer treatment in future (Falzone et al., 2018; Zhang and Chen, 2018; Kruger et al., 2019). However, there are few related work or annotated datasets in text mining on this domain. immuneXpresso (Kveler et al., 2018) is a text mining engine related to mammalian immune system, and NER is evaluated on cells and cytokine using dictionary matching. DES-Tcell (AlSaieedi et al., 2021) is a knowledgebase containing concepts of T-cell and other types of drugs, diseases, genes, etc in PubMed documents. However, it lacks utilizing novel text mining methods in the creation and evaluation the extracted data including NER tasks.

For the datasets used in our work, TCRE is manually annotated by Czech and Hammerbacher (2019) containing cell_type, cytokine, and TF entities, which are closed to our goal, and we used for our evaluation. A limitation of the TCRE is that it contains only 89 documents, which is insufficient to train powerful NER models. Therefore, our annotation method in this work can advance the task in extracting T-cell named entities. JNLPBA (Collier and Kim, 2004) contains manually annotated cell_type and protein entities. Meanwhile, UniProtKB (UniProt-Consortium, 2021) is a large and useful knowledgebase containing protein sequences annotated by experts with corresponding PubMed references. The UniProtKB and JNLPBA are leveraged to build our corpus.

6 Conclusion

We introduce our automatically annotated dataset for NER containing cell_type, cytokine, and TF entities, which are important in cancer immunology research, using a distant supervision method. The dataset is automatically annotated based on the entries in the UniProtKB knowledge base. We built a dictionary of the protein and gene names of cytokines and TF from the UniProtKB annotations. We then collected referenced PubMed articles and annotated these names in the texts using text matching with the dictionary entries. Additionally, we utilized the large manually annotated JNLPBA dataset, which contains cell_type and protein named entities to build our dataset. We trained NER models on our automatically annotated dataset and evaluated them on a manually annotated T-cell corpus. The results show that our automatically annotated dataset helps to improve the NER performance by extracting more named entities of cytokines and TF accurately. For future work, we plan to improve and extend our dataset to extract interactions or events related to these entities for text mining in cancer immunology research.

Acknowledgements

References

Ahdab AlSaieedi, Adil Salhi, Faroug Tifratene, Arwa Bin Raies, Arnaud Hungler, Mahmut Uludag, Christophe Van Neste, Vladimir B Bajic, Takashi Gojobori, and Magbubah Essack. 2021. Des-tcell is a knowledgebase for exploring immunology-related literature. *Scientific reports*, 11(1):1–11.

Iz Beltagy, Kyle Lo, and Arman Cohan. 2019. SciBERT: A pretrained language model for scientific text. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3606–3611, Hong Kong, China. Association for Computational Linguistics.

Nigel Collier and Jin-Dong Kim. 2004. Introduction to the bio-entity recognition task at JNLPBA. In *Proceedings of the International Joint Workshop on Natural Language Processing in Biomedicine and its Applications (NLPBA/BioNLP)*, pages 73–78, Geneva, Switzerland. COLING.

Eric Czech and Jeff Hammerbacher. 2019. Extracting T cell function and differentiation characteristics from the biomedical literature. *bioRxiv*, page 643767.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL*, pages 4171–4186.

Rezarta Islamaj Doğan, Robert Leaman, and Zhiyong Lu. 2014. Ncbi disease corpus: a resource for disease name recognition and concept normalization. *Journal of biomedical informatics*, 47:1–10.

Luca Falzone, Salvatore Salomone, and Massimo Libra. 2018. Evolution of cancer pharmacological treatments at the turn of the third millennium. *Frontiers in pharmacology*, page 1300.

Hiroyuki Hosokawa and Ellen V Rothenberg. 2018. Cytokines, transcription factors, and the initiation of T-cell development. *Cold Spring Harbor perspectives in biology*, 10(5):a028621.

Ji-Hae Kim, Kun-Joo Lee, and Seung-Woo Lee. 2021. Cancer immunotherapy with T-cell targeting cytokines: IL-2 and IL-7. *BMB reports*, 54(1):21.

Diederik P Kingma and Jimmy Lei Ba. 2015. Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR2015)*.

Stephan Kruger, Matthias Ilmer, Sebastian Kobold, Bruno L Cadilha, Stefan Endres, Steffen Ormanns, Gesa Schuebbe, Bernhard W Renz, Jan G D'Haese, Hans Schloesser, et al. 2019. Advances in cancer immunotherapy 2019–latest trends. *Journal of Experimental & Clinical Cancer Research*, 38(1):1–11.

Ksenya Kveler, Elina Starosvetsky, Amit Ziv-Kenet, Yuval Kalugny, Yuri Gorelik, Gali Shalev-Malul, Netta Aizenbud-Reshef, Tania Dubovik, Mayan Briller, John Campbell, et al. 2018. Immune-centric network of cytokines and cells in disease context identified by computational mining of PubMed. *Nature biotechnology*, 36(7):651–659.

Guillaume Lample, Miguel Ballesteros, Sandeep Subramanian, Kazuya Kawakami, and Chris Dyer. 2016. Neural architectures for named entity recognition. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 260–270.

Kenton Lee, Luheng He, Mike Lewis, and Luke Zettlemoyer. 2017. End-to-end neural coreference resolution. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 188–197.

Jiao Li, Yueping Sun, Robin J Johnson, Daniela Sciaky, Chih-Hsuan Wei, Robert Leaman, Allan Peter Davis, Carolyn J Mattingly, Thomas C Wiegers, and Zhiyong Lu. 2016. Biocreative v cdr task corpus: a resource for chemical disease relation extraction. *Database*, 2016.

Chen Liang, Yue Yu, Haoming Jiang, Siawpeng Er, Ruijia Wang, Tuo Zhao, and Chao Zhang. 2020. Bond: Bert-assisted open-domain named entity recognition with distant supervision. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1054–1064.

Yi Luan, Luheng He, Mari Ostendorf, and Hannaneh Hajishirzi. 2018. Multi-task identification of entities, relations, and coreference for scientific knowledge graph construction. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3219–3232.

Rishi Vishal Luckheeram, Rui Zhou, Asha Devi Verma, and Bing Xia. 2012. CD4+ T cells: differentiation and functions. *Clinical and developmental immunology*, 2012.

Yu Meng, Yunyi Zhang, Jiaxin Huang, Xuan Wang, Yu Zhang, Heng Ji, and Jiawei Han. 2021. Distantly-supervised named entity recognition with noise-robust learning and language model augmented self-training. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 10367–10378.

Taku Naito, Hirokazu Tanaka, Yoshinori Naoe, and Ichiro Taniuchi. 2011. Transcriptional control of T-cell development. *International immunology*, 23(11):661–668.

Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. 2017. Automatic differentiation in pytorch. In *NIPS-W*.

Jingbo Shang, Liyuan Liu, Xiaotao Gu, Xiang Ren, Teng Ren, and Jiawei Han. 2018. Learning named entity tagger using domain-specific dictionary. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2054–2064, Brussels, Belgium. Association for Computational Linguistics.

Mohammad Golam Sohrab and Makoto Miwa. 2018a. Deep exhaustive model for nested named entity recognition. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2843–2849.

Mohammad Golam Sohrab and Makoto Miwa. 2018b. Deep exhaustive model for nested named entity recognition. In *Proceedings of EMNLP*, pages 2843–2849. ACL.

Hai-Long Trieu, Thy Thy Tran, Khoa NA Duong, Anh Nguyen, Makoto Miwa, and Sophia Ananiadou. 2020. Deepeventmine: end-to-end neural nested event extraction from biomedical texts. *Bioinformatics*, 36(19):4910–4917.

UniProt-Consortium. 2021. UniProt: the universal protein knowledgebase in 2021. *Nucleic acids research*, 49(D1):D480–D489.

Alex D Waldman, Jill M Fritz, and Michael J Lenardo. 2020. A guide to cancer immunotherapy: from T cell basic science to clinical practice. *Nature Reviews Immunology*, 20(11):651–668.

Anliang Xia, Yan Zhang, Jiang Xu, Tailang Yin, and Xiao-Jie Lu. 2019. T cell dysfunction in cancer immunity and immunotherapy. *Frontiers in immunology*, 10:1719.

Yaosheng Yang, Wenliang Chen, Zhenghua Li, Zhengqiu He, and Min Zhang. 2018. Distantly supervised ner with partial annotation learning and reinforcement learning. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 2159–2169.

Hongming Zhang and Jibei Chen. 2018. Current status and future directions of cancer immunotherapy. *Journal of Cancer*, 9(10):1773.