

Development of a Natural Language Processing Algorithm to Identify Helmet Use Among Injured Bicyclists in Syndromic Surveillance Data

Report Prepared by:

Weipeng Zhou

Xinyao deGrauw

Kushang V. Patel

Stephen J. Mooney

WTSC Report Number 2024-Sub-grants-4983

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YVYY) 2. REPORT TYPE 3. DATES COVERED (From - To) 12/31/2024 Final Research Project Report 3. DATES COVERED (From - To) 12/31/2024 Sig2023-12/31/2024 5. CANT ACT NUMBER Development of a Natural Language Processing Algorithm to Identify Helmet 5. CANT NUMBER Use Armong Injured Bicyclists in Syndromic Surveillance Data: 5. CANT NUMBER Veiloping Tou 5. CRATN NUMBER Xinyao deGrauw 5. CRATN NUMBER Kushang V. Patel 5. CRATN NUMBER Stephen J. Mooney 6. TASK NUMBER 7. PERFORMING ORGANIZZATION NAME(S) AND ADDRESS(ES) 10. SPONSORIMONITORING AGENCY NAME(S) AND ADDRESS(ES) University of Washington Traffic Safety Commission 8. PERFORMING ORGANIZZATION RAME(S) AND ADDRESS(ES) Vomplate Report To public release. 10. SPONSORIMONITOR'S ACRONYM(S) 12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release. 13. SUPPLEMENTARY NOTES 11. SPONSOR/MONITOR'S ACCONYM(S) 14. ABSTRACT 17. UNITATION OF TAME 15. SUPLEMENTARY NOTES 14. ADSTRACT 16. SECURITY CLASSIFICATION OF: 17. UNITATION OF TAME 16. SUPLEMENTARY MOLES 10. SPONSOR/MONITOR'S CARCHONYM(S) 17. SUPLEMENTARY N	The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.							
12/31/2024 Final Research Project Report 5/8/2023-12/31/2024 4. TITLE AND SUBTITLE Development of a Natural Language Processing Algorithm to Identify Heimet 56. ORTACX NUMBER 2024-Sub-grants-4983 56. PROGRAM ELEMENT NUMBER 2024-Sub-grants-4983 56. AUTHOR(5) 56. PROGRAM ELEMENT NUMBER 2024-Sub-grants-4983 Weipeng Zhou 56. PROGRAM ELEMENT NUMBER 2024-Sub-grants-4983 57. PROFORMANDY CONTROL AND ADDRESS(ES) 56. PROJECT NUMBER 2024-Sub-grants-4983 Viryao deCrauw 56. PROJECT NUMBER 2024-Sub-grants-4983 Kushang V. Patel 56. PROJECT NUMBER 2024-Sub-grants-4983 Stephen J. Mooney 7. FERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) N/a University of Washington 8. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) N/a 9. SPONSOR/MONITOR'S ACRONYM(S) Vashington Traffic Safety Commission 9. SPONSOR/MONITOR'S ACRONYM(S) VTSC 11. SPONSOR/MONITOR'S ACRONYM(S) Vashington 98504-0944 11. SPONSOR/MONITOR'S REPORT NUMBER(S) 2024-Sub-grants-4983 12. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT 14.	1. REPORT DA	TE (DD-MM-YYYY) 2. REPOR	ГТҮРЕ			3. DATES COVERED (From - To)	
4. TTLE AND SUBTITLE Development of a Natural Language Processing Algorithm to Identify Helmett Se. CONTRACT NUMBER Use Among Injured Bicyclists in Syndromic Surveillance Data: 2024-Sub-grants-4983 2024-Sub-grants-4983 6. AUTHOR(5) 5c. PROGRAM ELEMENT NUMBER 2024-Sub-grants-4983 5c. PROGRAM ELEMENT NUMBER Velipting Zhou Xinyao deGrauw 5d. GRANT NUMBER 2024-Sub-grants-4983 5c. PROGRAM ELEMENT NUMBER Xinyao deGrauw Kushang V. Patel 5d. RORK UNIT NUMBER 2024-Sub-grants-4983 5d. RORK UNIT NUMBER Stephen J. Mooney 5d. FASK NUMBER No 5d. RORK UNIT NUMBER 2024-Sub-grants-4983 Stephen J. Mooney 5d. FASK NUMBER No 5d. RORK UNIT NUMBER 7d. Stephen J. Mooney 5d. GRANT AUMBER 7d. 7d. Stephen J. Mooney 5d. GRANT AUMBER 7d. Stephen J. Mooney 5d. GRANT AUMBER 7d. 9. SPONSORIMOMOTORING AGENCY NAME(S) AND ADDRESS(ES) No 8. PERFORTINUMER Virisolution 7d. Steport Number 7d. 1d. SpONSOR/MONITOR'S ACRONYM(S) WTSC 11. SPONSOR/MONITORING AGENCY NAME(S) AND ADDRESS(ES) 11. SPONSOR/MONITOR'S ACRONYM(S) WTSC 12. D	12/31/2024		Final Res	earch Project Repo	rt		5/8/2023-12/31/2024	
Development of a Natural Language Processing Algorithm to Identity Heimet 2024-Sub-grants-4983 Use Among Injured Bicyclists in Syndromic Surveillance Data: 56. GRANT NUMBER 2024-Sub-grants-4983 56. PROGRAM ELEMENT NUMBER 2024-Sub-grants-4983 56. PROGRAM ELEMENT NUMBER Virga deGrauw 56. GRANT NUMBER Kushang V. Patel 56. TASK NUMBER Stephen J. Mooney 66. TASK NUMBER 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) n/a Iniversity of Washington 8. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) 10. SPONSOR/MONITOR'S ACRONYM(S) Washington Traffic Safety Commission 90. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S ACRONYM(S) Washington Traffic Safety Commission 90. SPONSOR/MONITOR'S ACRONYM(S) VTSC VUmber Qiving Washington Stafe Safety Commission 90. SPONSOR/MONITOR'S ACRONYM(S) VTSC 11. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT 10. SPONSOR/MONITOR'S ACRONYM(S) VUmber Qiving as was associate to bike injury cases from the exponsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT 14. ABSTRACT This Study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases	4. TITLE AND S	SUBTITLE	-			5a. CC	5a. CONTRACT NUMBER	
Observed for public relation of yndrolining of weining of we		: of a Natural La niured Biovolist	inguage Proc	essing Algorithm to ic Surveillance Data	Identify Hein	1et 2024-	2024-Sub-grants-4983	
2024-Sub-grants-4983 5c. PROGRAM ELEMENT NUMBER 1n/a 5d. PROJECT NUMBER 1n/a 5d. PROJECT NUMBER 1n/a 5d. PROJECT NUMBER 1n/a 5d. PROJECT NUMBER 1n/a 7d	Use Among I		s in Gynarolli		a.	5b. GI	RANT NUMBER	
6. AUTHOR(S) 5c. PROGRAM ELEMENT NUMBER Weipeng Zhou 5d. PROJECT NUMBER Kushang V. Patel 5d. PROJECT NUMBER Stephen J. Mooney 5d. STASK NUMBER 1/a 5f. WORK UNIT NUMBER						2024-	Sub-grants-4983	
6. AUTHOR(s) Weipeng Zhou Xinyao deGrauw Kushang V, Patel 5d. PROJECT NUMBER n/a Stephen J. Mooney 9. FORCHURCH NUMBER n/a 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Washington 8. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Washington 8. PERFORMING ORGANIZATION REPORT NUMBER n/a 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Washington Traffic Safety Commission PO Box 40944 10. SPONSOR/MONITOR'S ACRONYM(S) WTSC 10. Sponsor/MONITOR'S REPORT NUMBER(S) 2024-Sub-grants-4983 10. SPONSOR/MONITOR'S REPORT NUMBER(S) 2024-Sub-grants-4983 12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release. 11. SPONSOR/MONITOR'S REPORT NUMBER(S) 2024-Sub-grants-4983 13. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of blike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0. 45, Fr el - 0.48). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2023, and further investigation is neceded. 15. SUBJECT TERMS natural language processing, helmet use, bicyclist 18. NUMBER Max Roberts, Ph.D., Sr. Research Associate PAGES None 18. NUMBER None 19. NUMBER None						5c. PF	5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(s) Weipeng Zhou Xinyao deGrauw Kushang V. Patel Stephen J. Mooney Stephen J. Mooney						n/a	n/a	
Weipeng Zhou Xinya de Graww n/a Stephen J. Mooney is. TASK NUMBER n/a Stephen J. Mooney is. CASK NUMBER n/a Stephen J. Mooney is. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Washington is. PERFORMING ORGANIZATION REPORT NUMBER n/a S. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Washington Traffic Safety Commission PO Box 40944 10. SPONSOR/MONITOR'S ACRONYM(S) WTSC Olympia, Washington 98504-0944 11. SPONSOR/MONITOR'S ACRONYM(S) WTSC 12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release. 30/24-Sub-grants-4983 13. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of heimet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0. 47, precision = 0.37, precision = 0.37, precision = 0.34, FC = 0.38), Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2021, and further investigation is needed. 15. SUBJECT TERMS natural language processing, helmet use, bicyclist 16. NUMBER [findude aree code] Max Roberts, Ph.D., Sr. Research Associate PACES None 18. NUMBER [findude aree code] 360-725-9894	6. AUTHOR(S)					5d. PF	OJECT NUMBER	
Xinyao deGrauw Kushag V. Patel Stephen J. Mooney 5e. TASK NUMBER IVa Stephen J. Mooney 5e. TASK NUMBER IVa 5. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Washington 8. PERFORMING ORGANIZATION REPORT NUMBER IVa 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Washington Traffic Safety Commission PO Box 40944 Olympia, Washington 98504-0944 10. SPONSOR/MONITOR'S ACRONYM(S) WTSC 12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release. 11. SPONSOR/MONITOR'S REPORT NUMBER(S) 2024-Sub-grants-4983 13. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2011 and 2023, and further investigation is needed. 15. SUBJECT TERMS natural language processing, helmet use, bicyclist 18. NUMBER Max Roberts, Ph.D., Sr. Research Associate 19. TELEPHONE NUMBER (include area code) 360-725-9894	Weipeng Zho	u				n/a		
Number n/a Stephen J. Mooney n/a 6f. WORK UNIT NUMBER n/a 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Image: stephen J. Mooney University of Washington 8. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Washington Traffic Safety Commission Na PO Box 40944 10. SPONSOR/MONITOR'S ACRONYM(S) WTSC 11. SPONSOR/MONITOR'S REPORT Number(s) 2024-Sub-grants-4983 12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release. 13. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This subset of bik injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2018 and 2023, and further investigation is needed. 16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT None None 18. NUMBER 19. ABSTRACT	Xinyao deGra	auw				5e. TA	SK NUMBER	
6f. WORK UNIT NUMBER n/a 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Washington 8. PERFORMING ORGANIZATION REPORT NUMBER n/a 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Washington Traffic Safety Commission PO Box 40944 Olympia, Washington 98504-0944 10. SPONSOR/MONITOR'S ACRONYM(S) WTSC 12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release. VITSC 13. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recail = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 16. SUBJECT TERMS natural language processing, helmet use, bicyclist 16. SECURITY CLASSIFICATION OF: A. REPORT None 17. LIMITATION OF ABSTRACT 18. NUMBER Max Roberts, Ph.D., Sr. Research Associate 19. TELEPHONE NUMBER (<i>include area code</i>) 360-725-9894	Stephen I M					n/a		
n/a 7. PERFORMING ORGANIZATION NAME(\$) AND ADDRESS(E\$) 8. PERFORMING ORGANIZATION REPORT NUMBER University of Washington n/a 9. SPONSORING/MONITORING AGENCY NAME(\$) AND ADDRESS(E\$) 10. SPONSOR/MONITOR'S ACRONYM(\$) Washington Traffic Safety Commission PO Box 40944 Olympia, Washington 98504-0944 11. SPONSOR/MONITOR'S ACRONYM(\$) VTSC 11. SPONSOR/MONITOR'S REPORT NUMBER(\$) 2024-Sub-grants-4983 2024-Sub-grants-4983 12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release. 13. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injure cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2018 and 2023, and further investigation is needed. 15. SUBJECT TERMS Interpret TERMS natural language processing, helmet use, bicyclist Max Roberts, Ph.D., Sr. Research Associate 196. TELEPHONE NUMBER (include area code) 360-725-9894	Stephen 5. M	ooney				5f. WC	ORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Image: Content of the second of						n/a		
University of Washington REPORT NUMBER 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Internition Washington Traffic Safety Commission WTSC PO Box 40944 Internition Olympia, Washington 98504-0944 WTSC 10. SPONSOR/MONITOR'S ACRONYM(S) WTSC PO Box 40944 Internition Olympia, Washington 98504-0944 WTSC 11. SPONSOR/MONITOR'S REPORT NumbeR(S) 2024-Sub-grants-4983 12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release. 13. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.80. Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2018 and 2023, and further investigation is needed. 15. SUBJECT TEMMS Internition OF: 16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF 18. N	7 PERFORMIN					1,74	8 PERFORMING ORGANIZATION	
n/a n/a s. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) 10. SPONSOR/MONITOR'S ACRONYM(S) Washington Traffic Safety Commission PO Box 40944 Olympia, Washington 98504-0944 11. SPONSOR/MONITOR'S ACRONYM(S) 2024-Sub-grants-4983 2024-Sub-grants-4983 12. DISTRIBUTION/AVAILABILITY STATEMENT 2024-Sub-grants-4983 Approved for public release. 2024-Sub-grants-4983 14. ABSTRACT The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 15. SUBJECT TERMS None 18. NUMBER PAGES None None 18. 19a. NAME OF RESPONSIBLE PERSON Max Roberts, Ph.D., Sr. Research Associate 19b. TELEPHONE NUMBER (include area code)	University of	Washington	(0)				REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) 10. SPONSOR/MONITOR'S ACRONYM(S) Washington Traffic Safety Commission WTSC PO Box 40944 11. SPONSOR/MONITOR'S ACRONYM(S) Olympia, Washington 98504-0944 11. SPONSOR/MONITOR'S REPORT NUMBER(S) 2024-Sub-grants-4983 12. DISTRIBUTION/AVAILABILITY STATEMENT 2024-Sub-grants-4983 Approved for public release. 11. SPONSOR/MONITOR'S REPORT 14. ABSTRACT The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 15. SUBJECT TERMS natural language processing, helmet use, bicyclist 16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT a. REPORT b. ABSTRACT None None 18	-	-					n/a	
s. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) 10. SPONSOR/MONITOR'S ACRONYM(S) Washington Traffic Safety Commission PO Box 40944 11. SPONSOR/MONITOR'S ACRONYM(S) PO Box 40944 11. SPONSOR/MONITOR'S ACRONYM(S) WTSC 11. SPONSOR/MONITOR'S REPORT NUMBER(S) 2024-Sub-grants-4983 2024-Sub-grants-4983 12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release. 13. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 15. SUBJECT TERMS natural language processing, helmet use, bicyclist 16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT 18. NUMBER (Include area code) None None 18 360-725-9894								
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) 10. SPONSOR/MONITOR'S ACRONYM(S) Washington Traffic Safety Commission WTSC PO Box 40944 11. SPONSOR/MONITOR'S REPORT Olympia, Washington 98504-0944 11. SPONSOR/MONITOR'S REPORT NUMBER(S) 2024-Sub-grants-4983 12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release. 13. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 15. SUBJECT TERMS 17. LIMITATION OF ABSTRACT 16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT 17. DNone None								
Virise With Sc PO Box 40944 Olympia, Washington 98504-0944 11. SPONSOR/MONITOR'S REPORT NUMBER(S) 2024-Sub-grants-4983 12. DISTRIBUTION/AVAILABILITY STATEMENT 2024-Sub-grants-4983 Approved for public release. 2024-Sub-grants-4983 13. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.33, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 16. SUBJECT TERMS natural language processing, helmet use, bicyclist natural language processing, helmet use, bicyclist 16. SUBJECT TERMS natural language processing, helmet use, bicyclist 17. LIMITATION OF a. REPORT b. ABSTRACT None 18	9. SPONSORIN	G/MONITORING A		(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)	
Olympia, Washington 98504-0944 11. SPONSOR/MONITOR'S REPORT NUMBER(S) 2024-Sub-grants-4963 12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release. 2024-Sub-grants-4963 13. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR).The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 16. SUBJECT TERMS natural language processing, helmet use, bicyclist 11. SUMBER PAGES 19a. NAME OF RESPONSIBLE PERSON Max Roberts, Ph.D., Sr. Research Associate 19b. TELEPHONE NUMBER (Include area code) 360-725-9894	PO Box 4094	A	511111551011				WISC .	
2024-Sub-grants-4983 12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release. 2024-Sub-grants-4983 13. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 15. SUBJECT TERMS natural language processing, helmet use, bicyclist 16. SECURITY CLASSIFICATION OF: a. REPORT b. ABSTRACT c. THIS PAGE 18. NUMBER 19a. NAME OF RESPONSIBLE PERSON Max Roberts, Ph.D., Sr. Research Associate None None None 18 360-725-9894	Olympia, Washington 98504-0944 NUMBER(S)					11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release. 13. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 15. SUBJECT TERMS natural language processing, helmet use, bicyclist 11. LIMITATION OF ABSTRACT 18. NUMBER OF PAGES 19a. NAME OF RESPONSIBLE PERSON Max Roberts, Ph.D., Sr. Research Associate 19b. TELEPHONE NUMBER (Include area code) 360-725-9894							2024-Sub-grants-4983	
Approved for public release. 13. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helme use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 15. SUBJECT TERMS natural language processing, helmet use, bicyclist 16. SECURITY CLASSIFICATION OF: a. REPORT b. ABSTRACT c. THIS PAGE None None None None	12. DISTRIBUT	ION/AVAILABILIT	Y STATEMENT				, i i i i i i i i i i i i i i i i i i i	
13. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 15. SUBJECT TERMS natural language processing, helmet use, bicyclist 16. SECURITY CLASSIFICATION OF: a. REPORT b. ABSTRACT None None None None	Approved for public release.							
13. SUPPLEMENTARY NOTES The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT a. REPORT b. ABSTRACT c. THIS PAGE 17. LIMITATION OF ABSTRACT None None								
The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 15. SUBJECT TERMS natural language processing, helmet use, bicyclist 14. ABSTRACT In LIMITATION OF: ABSTRACT a. REPORT b. ABSTRACT None None None None None None None None	13. SUPPLEMENTARY NOTES							
the funding agency. 14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 15. SUBJECT TERMS natural language processing, helmet use, bicyclist 16. SECURITY CLASSIFICATION OF: a. REPORT b. ABSTRACT c. THIS PAGE 17. LIMITATION OF ABSTRACT None None None None	The contents of the manuscript are solely the responsibility of the authors and do not necessarily reflect the official views of							
14. ABSTRACT This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 15. SUBJECT TERMS natural language processing, helmet use, bicyclist 16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT a. REPORT b. ABSTRACT c. THIS PAGE 17. LIMITATION OF ABSTRACT None None	the funding agency.							
This study developed a natural language processing (NLP) algorithm to find mentions of helmet use in hospital triage notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 15. SUBJECT TERMS natural language processing, helmet use, bicyclist 16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT a. REPORT b. ABSTRACT c. THIS PAGE 17. LIMITATION OF ABSTRACT None None None None	14. ABSTRACT							
using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 15. SUBJECT TERMS natural language processing, helmet use, bicyclist 16. SECURITY CLASSIFICATION OF: a. REPORT b. ABSTRACT c. THIS PAGE 17. LIMITATION OF ABSTRACT 18. NUMBER OF RESPONSIBLE PERSON Mone None None 18 19a. NAME OF RESPONSIBLE PERSON	This study de	veloped a natu	ral language	processing (NLP) a	lgorithm to fir	nd mentions	of helmet use in hospital triage notes,	
achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 15. SUBJECT TERMS natural language processing, helmet use, bicyclist 16. SECURITY CLASSIFICATION OF: a. REPORT b. ABSTRACT c. THIS PAGE None None None None None None	using a subse	et of bike injury	cases from th	ne UW Medicine ele	ectronic medi	cal records	(UW EMR).The trained model	
Intart relimet use anong injured cyclists seen in OW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed. 15. SUBJECT TERMS natural language processing, helmet use, bicyclist 16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT a. REPORT b. ABSTRACT c. THIS PAGE 17. LIMITATION OF ABSTRACT None None None None	achieved high	n performance	(accuracy = 0	0.87, precision = 0.8	3, recall = 0.9	94, ⊢1 = 0.8 m. around :	38). Applying the model, we observed	
16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT 18. NUMBER OF RESPONSIBLE PERSON Max Roberts, Ph.D., Sr. Research Associate 16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT 18. NUMBER OF RESPONSIBLE PERSON Max Roberts, Ph.D., Sr. Research Associate 17. None None None None 18 360-725-9894	Ear RHINO d	se among injur ata, we found a	ed cyclists se	trend in the missing	Increased inc	from bike i	15% to 20% between 2021 and 2024.	
15. SUBJECT TERMS natural language processing, helmet use, bicyclist 16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT a. REPORT b. ABSTRACT c. THIS PAGE 17. LIMITATION OF ABSTRACT None None None None	further investigation is needed							
16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT 18. NUMBER 19a. NAME OF RESPONSIBLE PERSON a. REPORT b. ABSTRACT c. THIS PAGE 17. LIMITATION OF ABSTRACT 0F PAGES None None None None None 18. NUMBER 19a. NAME OF RESPONSIBLE PERSON Max Roberts, Ph.D., Sr. Research Associate 0F PAGES 19b. TELEPHONE NUMBER (Include area code) None None 18 360-725-9894								
16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT 18. NUMBER OF RESPONSIBLE PERSON a. REPORT b. ABSTRACT c. THIS PAGE 17. LIMITATION OF ABSTRACT 18. NUMBER OF RESPONSIBLE PERSON None None None None 18. NONE 18. NUMBER OF RESPONSIBLE PERSON Max Roberts, Ph.D., Sr. Research Associate 19b. TELEPHONE NUMBER (Include area code) None None 18 360-725-9894	natural language processing helmet use hisvelist							
16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT 18. NUMBER OF RESPONSIBLE PERSON a. REPORT b. ABSTRACT c. THIS PAGE OF PAGES Max Roberts, Ph.D., Sr. Research Associate None None None None 18 360-725-9894								
16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT 18. NUMBER 19a. NAME OF RESPONSIBLE PERSON a. REPORT b. ABSTRACT c. THIS PAGE 17. LIMITATION OF ABSTRACT OF PAGES Max Roberts, Ph.D., Sr. Research Associate None None None None None 18 18. 19a. NAME OF RESPONSIBLE PERSON								
16. SECURITY CLASSIFICATION OF: 17. LIMITATION OF ABSTRACT 18. NUMBER 19a. NAME OF RESPONSIBLE PERSON a. REPORT b. ABSTRACT c. THIS PAGE 0F 0F Max Roberts, Ph.D., Sr. Research Associate None None None None 18 360-725-9894								
a. REPORT b. ABSTRACT c. THIS PAGE PAGES Max Roberts, Ph.D., Sr. Research Associate None None None None 18 360-725-9894								
NoneNoneNone1819b. TELEPHONE NUMBER (Include area code)NoneNone18360-725-9894	a. REPORT	b. ABSTRACT	c. THIS PAGE		PAGES	Max Robe	rts, Ph.D., Sr. Research Associate	
None None None 18 360-725-9894						19b. TELEP	HONE NUMBER (Include area code)	
	None None None 18 360-725-9894			394				

Development of a Natural Language Processing Algorithm to Identify Helmet Use Among Injured Bicyclists in Syndromic Surveillance Data:

A Technical Report for the Washington Traffic Safety Commission

Report

Prepared by:

Weipeng Zhou

Xinyao deGrauw

Kushang V. Patel

Stephen J. Mooney

Contents

Abstract	3
Introduction	4
Methods	4
Dataset	4
Model development	5
Results	5
UW	5
RHINO	9
Conclusions	
References	

Abstract

The repeal of King County's adult bicycle helmet law in March 2022 raised concerns that fewer cyclists might wear helmets, resulting in more head injuries. Data from the Rapid Health Information NetwOrk (RHINO) may be used to track changes in helmet use over time, but helmet use is mentioned only in triage notes, which are difficult to review manually. We developed a natural language processing (NLP) algorithm to find mentions of helmet use in these notes, using a subset of bike injury cases from the UW Medicine electronic medical records (UW EMR). We then applied this algorithm to the full UW EMR dataset and the RHINO dataset. The trained model achieved high performance (accuracy = 0.87, precision = 0.83, recall = 0.94, F1 = 0.88). Applying the model, we observed that helmet use among injured cyclists seen in UW Medicine increased from around 15% to 20% between 2021 and 2024. For RHINO data, we found an increasing trend in the missing triage notes from bike injuries between 2018 and 2023, and further investigation is needed.

Introduction

For many years, King County was the largest county in the US with a law requiring adults to wear helmets while riding bicycles. However, after controversy regarding disproportionate enforcement of the law, King County's adult helmet law was repealed in March 2022 [1]. This repeal raised concerns that a resulting decrease in helmet use might in turn lead to an increase in head injuries.

Evidence of such an increase might be available in syndromic surveillance emergency department records currently captured in the Rapid Health Information NetwOrk (RHINO) system run by the Washington State Department of Health [2]. However, patient helmet use is recorded only in triage notes – not in structured data that could counted algorithmically – and so such a study would require prohibitive amounts of manual chart review.

In this study, we created a natural language processing (NLP) algorithm [3] that can identify helmet use among triage notes for injured bicyclists who received medical care at UW Medicine facilities and applied it to triage notes from RHINO. This prepares us for future work that can apply this algorithm to all RHINO data to estimate the impact of the repeals. Specifically, we:

(1) Developed an NLP model by creating an annotated dataset from a subset of UW electronic health records (EHR) bike injury triage notes.

(2) Applied the model to UW EHR and RHINO bike injury-related triage notes to estimate the effect of helmet law repeal on bike travel safety.

Methods

Dataset

Bike-related medical encounters were identified in UW Medicine's electronic health records (EHR) using ICD-10-CM codes (Appendix A). Between April 2021 and September 2024, 2,304 bike injury-related medical visits were identified in UW Medicine's EHR; only 887 cases had triage notes (Table 1).

Between January 2018 and December 2023, 5,566 bike injury-related medical visits were identified using the RHINO bike injury surveillance query applied to RHINO data; only 2,097 of them had triage notes (Table 2).

We used the 887 UW cases and 2097 RHINO cases (named as UW dataset and RHINO dataset, respectively) to develop an automated helmet use identification mode. Specifically, we annotated cases between 2021 and 2022 (335 cases) from the UW dataset and used these to develop and train a machine learning model to algorithmically identify whether triage notes confirm helmet use. The trained model was then applied to the full UW dataset (887 cases) and the RHINO dataset. When annotating the triage note subsets from the UW dataset, we classified triage notes into two categories: (1) the triage note indicated that the patient was wearing a helmet, and (2) the triage note asserted that the patient was not wearing helmet OR the information in the note was insufficient for inferring patients helmet status. We named these two categories as Yes and No/Not Mentioned in the result section.

Model development

In developing the model, we split the annotated dataset into a training and test set with an 8:2 ratio. We created a machine learning pipeline that first converted the input triage notes into TF-IDF (term frequency-inverse document frequency) word-vector features [4] and the outputs were provided to an XGBoost (eXtreme Gradient Boosting) classifier [5] with the training objective of determining whether the triage notes mentioned that the patient was wearing a helmet. XGBoost is a powerful machine learning algorithm based on gradient boosting that builds decision tree models sequentially to optimize performance and accuracy. We used Python 3.8 [6] as the programming language and Scikit-learn 1.6.0 for the TF-IDF implementation [7]. The XGBoost algorithm was provided by the xgboost package version 2.1.3. The data preprocessing was performed using Pandas 2.2.3 [6].

We created the pipeline using the training set and applied it to the test set to evaluate the model performance by comparing the model's prediction against human annotation. We used scikit-learn 1.6.0 to compute multiple metrics (accuracy, precision, recall, and F1, with values ranging from 0 [all failed] to 1 [perfect]) by providing it with the model's predictions and human annotations.

Finally, we applied the trained model to the full UW dataset and the RHINO datasets to investigate trends in triage note mention of helmet use over time.

Results

On the UW test dataset, the model achieved accuracy=0.87, precision=0.83, recall=0.94, and F1=0.88, which indicated that our trained model produced predictions aligned well with human annotations. We then applied the model to the UW and RHINO dataset.

Bike injury-related medical visits in UW EMR

Figure 1 shows helmet use documented in triage notes from 2021 to 2024 (Q1-Q3) in UW Medicine.

Triage notes became available in more bike-related visits in UW Medicine over time but still remained in the minority of records. Cases reporting helmet use in bike incidents (Yes) gradually increased from around 15% to above 20%, while cases reporting no helmet use or not reporting anything about helmet use also increased.



Figure 1. Overall trends of helmet status between 2021 (Q2-Q4) and 2024 (Q1-Q3) in bike-related medical visits at UW Medicine facilities. "Yes" indicates that the triage notes mentioned patients wearing a helmet, "No/Not Mentioned" indicates that triage notes mentioned patients not wearing a helmet or did not mention helmet use at all, and "No Triage Note Available" indicates that triage notes of the bike-related medical visits were not available.

Year	Helmet status	Count (Percent)	
	No/Not Mentioned	71 (13.4%)	
2021	No Triage Note Available	377 (71.4%)	
	Yes	80 (15.2%)	
	No/Not Mentioned	114 (18.9%)	
2022	No Triage Note Available	401 (66.6%)	
	Yes	87 (14.5%)	
	No/Not Mentioned	135 (22.0%)	
2023	No Triage Note Available	340 (55.5%)	
	Yes	138 (22.5%)	
2024	No/Not Mentioned	126 (22.5%)	
	No Triage Note Available	299 (53.3%)	
	Yes	136 (24.2%)	

Table 1. Helmet status mentioned in triage notes of bike-related medica	l visits	between
2021 (Q2-Q4) and 2024 (Q1-Q3) in UW Medicine.		

The percentage of incidents involving female riders in which helmet use was documented more than doubled over the years (11% to 26.2%). For male riders, the percentage of notes with documented helmet use decreased in 2022 but increased in 2023 (Figure 2).



Figure 2. Percentage of notes documenting helmet use in bike-related injuries between 2021 (Q2-Q4) and 2024 (Q1-Q3) in UW Medicine by gender. The count and percentage are listed in the Appendix Tables.

The percentage of notes documenting helmet use varied among age groups. For patients aged ≤ 14 , 25–44, and 45–64, the helmet use percentage decreased in 2022 and increased afterward. The documented helmet use percentage for patients in other age groups increased over the years (Figure 3).



Figure 3. Percentage of notes documenting helmet use in bike-related injuries between 2021 (Q2-Q4) and 2024 (Q1-Q3) in UW Medicine by age groups. The count and percentage are listed in the Appendix Tables.

The percentage of notes documenting helmet use among bike-related ED visits decreased slightly in 2022 and increased afterward. The percentage of helmet use fluctuated among inpatient/observation visits (Figure 4).



Figure 4. Percentage of notes documenting helmet use in bike-related injuries between 2021 (Q2-Q4) and 2024 (Q1-Q3) in UW Medicine by admission type. The count and percentage are listed in the Appendix Tables.

The percentage of notes documenting helmet use among bike-related visits in Harborview Medical Center and Northwest Medical Center decreased in 2022 and increased afterward. The percentage of notes documenting helmet use in bike-related visits in Montlake Medical Center has tripled over the years (4.6% to 14.5%) (Figure 5).



Figure 5. Percentage of notes documenting helmet use in bike-related injuries between 2021 (Q2-Q4) and 2024 (Q1-Q3) in UW Medicine by facilities. The count and percentage are listed in the Appendix Tables.

Bike injury-related medical visits in RHINO

Figure 6 shows the results of our algorithm applied to RHINO data. Notably, triage notes were less often available for more recent visits. Future studies should investigate why notes were less often available.



Figure 6. Overall trends of helmet status in bike-related medical visits between 2018 and 2023 statewide in WA (data source: RHINO). "Yes" represented triage notes mentioning helmet use, "No/Not Mentioned" represented triage notes mentioning helmet non-use or not mentioning helmet, and "No Triage Note Available" indicated that triage notes if the medical visits were not available.

Year	Helmet status	elmet status Count (Percent)	
	No/Not Mentioned	398 (52.9%)	
2018	No Triage Note Available	127 (16.9%)	
	Yes	227 (30.2%)	
	No/Not Mentioned	433 (46.6%)	
2019	No Triage Note Available	224 (24.1%)	
	Yes	273 (29.4%)	
	No/Not Mentioned	389 (48.0%)	
2020	No Triage Note Available	228 (28.1%)	
	Yes	193 (23.8%)	
	No/Not Mentioned	270 (28.1%)	
2021	No Triage Note Available	593 (61.8%)	
	Yes	97 (10.1%)	
	No/Not Mentioned	326 (32.1%)	
2022	No Triage Note Available	610 (60.0%)	
	Yes	80 (7.9%)	
	No/Not Mentioned	343 (31.2%)	
2023	No Triage Note Available	687 (62.6%)	
	Yes	68 (6.2%)	

Table 2. Counts and percentages of helmet status between 2018 and 2023 for RHINO.

Conclusions

We developed a model to monitor helmet use in bike injuries over time in UW EMR and RHINO triage note data. While our model performed well on our test data, triage notes were missing for too many patient visits in RHINO to draw surveillance conclusions over time.

References

- [1] S. Kasakove, "Seattle Bike Helmet Rule Is Dropped Amid Racial Justice Concerns," *The New York Times*, Feb. 19, 2022. Accessed: Dec. 16, 2024. [Online]. Available: https://www.nytimes.com/2022/02/18/us/seattle-bicyclehelmet.html
- [2] D.-P.-P. O. C. Surv--5700 Info and, "Syndromic Surveillance (RHINO) | Washington State Department of Health." Accessed: Dec. 16, 2024. [Online]. Available: https://doh.wa.gov/public-health-provider-resources/healthcareprofessions-and-facilities/data-exchange/syndromic-surveillance-rhino

- [3] S. Vajjala, B. Majumder, A. Gupta, and H. Surana, *Practical Natural Language Processing: A Comprehensive Guide to Building Real-World NLP Systems*. O'Reilly Media, Inc., 2020.
- [4] G. Zhao, Y. Liu, W. Zhang, and Y. Wang, "TFIDF based Feature Words Extraction and Topic Modeling for Short Text," in *Proceedings of the 2018 2nd International Conference on Management Engineering, Software Engineering and Service Sciences*, in ICMSS 2018. New York, NY, USA: Association for Computing Machinery, Jan. 2018, pp. 188–191. doi: 10.1145/3180374.3181354.
- [5] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, in KDD '16. New York, NY, USA: Association for Computing Machinery, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.
- [6] W. McKinney, *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython.* O'Reilly Media, Inc., 2012.
- [7] *scikit-learn: A set of python modules for machine learning and data mining*. C, Python. Accessed: Nov. 03, 2024. [MacOS, Microsoft :: Windows, POSIX, Unix]. Available: https://scikit-learn.org

Appendices

Appendix A, ICD-10-CM codes of bike-related inj	ury
---	-----

ICD-10-CM	Description	Start
		year
V12.3XX	Person boarding or alighting a pedal cycle injured in collision with two- or three- wheeled motor vehicle	2015
V12.4XX	Pedal cycle driver injured in collision with two- or three-wheeled motor vehicle in traffic accident	2015
V12.5XX	Pedal cycle passenger injured in collision with two- or three-wheeled motor vehicle in traffic accident	2015
V12.9XX	Unspecified pedal cyclist injured in collision with two- or three-wheeled motor vehicle in traffic accident	2015
V13.3XX	Person boarding or alighting a pedal cycle injured in collision with car, pick-up truck or van	2015
V13.4XX	Pedal cycle driver injured in collision with car, pick-up truck or van in traffic accident	2015
V13.5XX	Pedal cycle passenger injured in collision with car, pick-up truck or van in traffic accident	2015

V13.9XX	Unspecified pedal cyclist injured in collision with car, pick-up truck or van in traffic accident	2015
V14.3XX	Person boarding or alighting a pedal cycle injured in collision with heavy transport vehicle or bus	2015
V14.4XX	Pedal cycle driver injured in collision with heavy transport vehicle or bus in traffic accident	2015
V14.5XX	Pedal cycle passenger injured in collision with heavy transport vehicle or bus in traffic accident	2015
V14.9XX	Unspecified pedal cyclist injured in collision with heavy transport vehicle or bus in traffic accident	2015
V19.40X	Pedal cycle driver injured in collision with unspecified motor vehicles in traffic accident	2015
V19.49X	Pedal cycle driver injured in collision with other motor vehicles in traffic accident	2015
V19.50X	Pedal cycle passenger injured in collision with unspecified motor vehicles in traffic accident	2015
V19.59X	Pedal cycle passenger injured in collision with other motor vehicles in traffic accident	2015
V19.60X	Unspecified pedal cyclist injured in collision with unspecified motor vehicles in traffic accident	2015
V19.69X	Unspecified pedal cyclist injured in collision with other motor vehicles in traffic accident	2015
V19.9XX	Pedal cyclist (driver) (passenger) injured in unspecified traffic accident	2015
V10.0XX	Pedal cycle driver injured in collision with pedestrian or animal in nontraffic accident	2015
V10.1XX	Pedal cycle passenger injured in collision with pedestrian or animal in nontraffic accident	2015
V10.2XX	Unspecified pedal cyclist injured in collision with pedestrian or animal in nontraffic accident	2015
V10.3XX	Person boarding or alighting a pedal cycle injured in collision with pedestrian or animal	2015
V10.4XX	Pedal cycle driver injured in collision with pedestrian or animal in traffic accident	2015
V10.5XX	Pedal cycle passenger injured in collision with pedestrian or animal in traffic accident	2015
V10.9XX	Unspecified pedal cyclist injured in collision with pedestrian or animal in traffic accident	2015
V11.0XX	Pedal cycle driver injured in collision with other pedal cycle in nontraffic accident	2015
V11.1XX	Pedal cycle passenger injured in collision with other pedal cycle in nontraffic accident	2015

V11.2XX	Unspecified pedal cyclist injured in collision with other pedal cycle in nontraffic accident		
V11.3XX	Person boarding or alighting a pedal cycle injured in collision with other pedal cycle	2015	
V11.4XX	Pedal cycle driver injured in collision with other pedal cycle in traffic accident	2015	
V11.5XX	Pedal cycle passenger injured in collision with other pedal cycle in traffic accident	2015	
V11.9XX	Unspecified pedal cyclist injured in collision with other pedal cycle in traffic accident	2015	
V12.0XX	Pedal cycle driver injured in collision with two- or three-wheeled motor vehicle in nontraffic accident	2015	
V12.1XX	Pedal cycle passenger injured in collision with two- or three-wheeled motor vehicle in nontraffic accident	2015	
V12.2XX	Unspecified pedal cyclist injured in collision with two- or three-wheeled motor vehicle in nontraffic accident	2015	
V13.0XX	Pedal cycle driver injured in collision with car, pick-up truck or van in nontraffic accident	2015	
V13.1XX	Pedal cycle passenger injured in collision with car, pick-up truck or van in nontraffic accident	2015	
V13.2XX	Unspecified pedal cyclist injured in collision with car, pick-up truck or van in nontraffic accident	2015	
V14.0XX	Pedal cycle driver injured in collision with heavy transport vehicle or bus in nontraffic accident	2015	
V14.1XX	Pedal cycle passenger injured in collision with heavy transport vehicle or bus in nontraffic accident	2015	
V14.2XX	Unspecified pedal cyclist injured in collision with heavy transport vehicle or bus in nontraffic accident	2015	
V15.0XX	Pedal cycle driver injured in collision with railway train or railway vehicle in nontraffic accident	2015	
V15.1XX	Pedal cycle passenger injured in collision with railway train or railway vehicle in nontraffic accident	2015	
V15.2XX	Unspecified pedal cyclist injured in collision with railway train or railway vehicle in nontraffic accident	2015	
V15.3XX	Person boarding or alighting a pedal cycle injured in collision with railway train or railway vehicle	2015	
V15.4XX	Pedal cycle driver injured in collision with railway train or railway vehicle in traffic accident	2015	
V15.5XX	Pedal cycle passenger injured in collision with railway train or railway vehicle in traffic accident	2015	
V15.9XX	Unspecified pedal cyclist injured in collision with railway train or railway vehicle in traffic accident	2015	

V16.0XX	Pedal cycle driver injured in collision with other nonmotor vehicle in nontraffic accident	
V16.1XX	Pedal cycle passenger injured in collision with other nonmotor vehicle in nontraffic accident	2015
V16.2XX	Unspecified pedal cyclist injured in collision with other nonmotor vehicle in nontraffic accident	2015
V16.3XX	Person boarding or alighting a pedal cycle injured in collision with other nonmotor vehicle in nontraffic accident	2015
V16.4XX	Pedal cycle driver injured in collision with other nonmotor vehicle in traffic accident	2015
V16.5XX	Pedal cycle passenger injured in collision with other nonmotor vehicle in traffic accident	2015
V16.9XX	Unspecified pedal cyclist injured in collision with other nonmotor vehicle in traffic accident	2015
V17.0XX	Pedal cycle driver injured in collision with fixed or stationary object in nontraffic accident	2015
V17.1XX	Pedal cycle passenger injured in collision with fixed or stationary object in nontraffic accident	2015
V17.2XX	Unspecified pedal cyclist injured in collision with fixed or stationary object in nontraffic accident	2015
V17.3XX	Person boarding or alighting a pedal cycle injured in collision with fixed or stationary object	2015
V17.4XX	Pedal cycle driver injured in collision with fixed or stationary object in traffic accident	2015
V17.5XX	Pedal cycle passenger injured in collision with fixed or stationary object in traffic accident	2015
V17.9XX	Unspecified pedal cyclist injured in collision with fixed or stationary object in traffic accident	2015
V4.0.0VV		2015
V18.0XX	Pedal cycle driver injured in noncollision transport accident in nontraffic accident Pedal cycle passenger injured in noncollision transport accident in nontraffic	2015
V18.1XX	accident	2015
V18.2XX	Unspecified pedal cyclist injured in noncollision transport accident in nontraffic accident	2015
V18.3XX	Person boarding or alighting a pedal cycle injured in noncollision transport accident	2015
V18.4XX	Pedal cycle driver injured in noncollision transport accident in traffic accident	2015
V18.5XX	Pedal cycle passenger injured in noncollision transport accident in traffic accident	2015
V18.9XX	Unspecified pedal cyclist injured in noncollision transport accident in traffic accident	2015

	Pedal cycle driver injured in collision with unspecified motor vehicles in nontraffic	
V19.00X	accident	2015
V19.09X	Pedal cycle driver injured in collision with other motor vehicles in nontraffic accident	2015
V19.10X	Pedal cycle passenger injured in collision with unspecified motor vehicles in nontraffic accident	2015
V19.19X	Pedal cycle passenger injured in collision with other motor vehicles in nontraffic accident	2015
V19.20X	Unspecified pedal cyclist injured in collision with unspecified motor vehicles in nontraffic accident	2015
V19.29X	Unspecified pedal cyclist injured in collision with other motor vehicles in nontraffic accident	2015
V19.3XX	Pedal cyclist (driver) (passenger) injured in unspecified nontraffic accident	2015
V19.81X	Pedal cyclist (driver) (passenger) injured in transport accident with military vehicle	2015
V19.88X	Pedal cyclist (driver) (passenger) injured in other specified transport accidents	2015

Table 1. Helmet use mentioned in the triage notes of bike-related medical visits in UW Electronic Medical Records 2021 (Q2-Q4) and 2024 (Q1-Q3) by gender.

Year	Gender	Female	Male
2021	No/Not Mentioned	18 (11.0%)	53 (14.5%)
	No Triage Note Available	127 (77.9%)	250 (68.5%)
	Yes	18 (11.0%)	62 (17.0%)
	No/Not Mentioned	24 (15.4%)	90 (20.4%)
2022	No Triage Note Available	107 (68.6%)	291 (66.0%)
	Yes	25 (16.0%)	60 (13.6%)
2023	No/Not Mentioned	33 (20.8%)	102 (22.5%)
	No Triage Note Available	97 (61.0%)	243 (53.6%)
	Yes	29 (18.2%)	108 (23.8%)
2024	No/Not Mentioned	21 (14.9%)	105 (25.0%)
	No Triage Note Available	83 (58.9%)	216 (51.4%)
	Yes	37 (26.2%)	99 (23.6%)

	Electionic inculcar records 2021 (Q2-Q4) and 2024 (Q1-Q3) by age.						
Year	Age	<=14	15-18	19-24	25-44	45-64	65+
2021	No/Not Mentioned	3 (11.5%)	2 (15.4%)	7 (20.0%)	30 (14.2%)	19 (11.0%)	10 (14.3%)
	No Triage Note Available	19 (73.1%)	9 (69.2%)	24 (68.6%)	153 (72.2%)	124 (72.1%)	48 (68.6%)
	Yes	4 (15.4%)	2 (15.4%)	4 (11.4%)	29 (13.7%)	29 (16.9%)	12 (17.1%)
2022	No/Not Mentioned	7 (41.2%)	3 (27.3%)	9 (22.5%)	51 (18.5%)	35 (19.2%)	8 (10.5%)
	No Triage Note Available	8 (47.1%)	8 (72.7%)	20 (50.0%)	190 (69.1%)	121 (66.5%)	54 (71.1%)
	Yes	2 (11.8%)	0 (0.0%)	11 (27.5%)	34 (12.4%)	26 (14.3%)	14 (18.4%)
2023	No/Not Mentioned	12 (38.7%)	5 (31.2%)	13 (23.6%)	63 (23.2%)	26 (14.8%)	15 (23.8%)
	No Triage Note Available	15 (48.4%)	7 (43.8%)	28 (50.9%)	148 (54.6%)	108 (61.4%)	34 (54.0%)
	Yes	4 (12.9%)	4 (25.0%)	14 (25.5%)	60 (22.1%)	42 (23.9%)	14 (22.2%)
2024	No/Not Mentioned	10 (35.7%)	3 (20.0%)	16 (32.0%)	53 (21.8%)	35 (22.6%)	9 (12.9%)
	No Triage Note Available	13 (46.4%)	8 (53.3%)	20 (40.0%)	131 (53.9%)	88 (56.8%)	39 (55.7%)
	Yes	5 (17.9%)	4 (26.7%)	14 (28.0%)	59 (24.3%)	32 (20.6%)	22 (31.4%)

Table 2. Helmet use mentioned in the triage notes of bike-related medical visits in UW Electronic Medical Records 2021 (Q2-Q4) and 2024 (Q1-Q3) by age.

Table 3. Helmet use mentioned in the triage notes of bike-related medical visits in UW Electronic Medical Records 2021 (Q2-Q4) and 2024 (Q1-Q3) care type.

Vear	Care type	Emergency	Inpatient/Observatio	Outpatient/urgen	
rear	care type	Linergency	n	t care	
	No/Not	51 (16.8%)	20 (27 0%)	0 (0.0%)	
	Mentioned	51 (10.870)	20 (27.070)		
2021	No Triage				
2021	Note	196 (64.5%)	31 (41.9%)	150 (100.0%)	
	Available				
	Yes	57 (18.8%)	23 (31.1%)	0 (0.0%)	
	No/Not	94 (DD E0/)		0 (0 00/)	
	Mentioned 84 (23.5%)		50 (55.7%)	0 (0.0%)	
2022	No Triage				
	Note 209 (58.4%)		32 (38.1%)	160 (100.0%)	
	Available				

	Yes	65 (18.2%)	22 (26.2%)	0 (0.0%)
2023	No/Not Mentioned	102 (26.3%)	32 (35.6%)	1 (0.7%)
	No Triage Note Available	185 (47.7%)	23 (25.6%)	132 (97.8%)
	Yes	101 (26.0%)	35 (38.9%)	2 (1.5%)
	No/Not Mentioned	92 (25.3%)	33 (38.8%)	1 (0.9%)
2024	No Triage Note Available	162 (44.6%)	25 (29.4%)	112 (99.1%)
	Yes	109 (30.0%)	27 (31.8%)	0 (0.0%)

Table 4. Helmet use mentioned in the triage notes of bike-related medical visits in UW Electronic Medical Records 2021 (Q2-Q4) and 2024 (Q1-Q3) by facility.

					UW primary
		Harborview		Northwest	care/urgent
Year	Facility	MC	Montlake MC	MC	care
	No/Not				
	Mentioned	49 (37.4%)	15 (5.8%)	7 (26.9%)	0 (0.0%)
	No Triage				
	Note				
	Available	23 (17.6%)	233 (89.6%)	10 (38.5%)	111 (100.0%)
2021	Yes	59 (45.0%)	12 (4.6%)	9 (34.6%)	0 (0.0%)
	No/Not				
	Mentioned	81 (44.3%)	23 (8.0%)	10 (21.7%)	0 (0.0%)
	No Triage				
	Note				
	Available	45 (24.6%)	244 (84.7%)	27 (58.7%)	85 (100.0%)
2022	Yes	57 (31.1%)	21 (7.3%)	9 (19.6%)	0 (0.0%)
	No/Not				
	Mentioned	86 (43.9%)	24 (8.9%)	24 (57.1%)	1 (1.0%)
	No Triage				
	Note				
	Available	18 (9.2%)	211 (77.9%)	8 (19.0%)	103 (99.0%)
2023	Yes	92 (46.9%)	36 (13.3%)	10 (23.8%)	0 (0.0%)
	No/Not				
	Mentioned	91 (48.4%)	23 (9.8%)	12 (34.3%)	0 (0.0%)
	No Triage				
	Note				
2024	Available	7 (3.7%)	178 (75.7%)	11 (31.4%)	103 (100.0%)

Yes 90 (47.9%) 34 (14.5	5%) 12 (34.3%) 0 (0.0%)
-------------------------	-------------------------