

CURRICULUM VITAE

DAVID S. CARRELL, PHD

EDUCATIONAL AND BIOGRAPHICAL INFORMATION

Email	david.s.carrell@kp.org
Desk phone	206-287-2705
Fax	206-287-2871
Website	https://www.kp.washingtonresearch.org/our-research/our-scientists/carrell-david
Office	<i>Kaiser Permanente Washington Health Research Institute</i> 1730 Minor Avenue, Suite 1600 Seattle, WA 98101
1990	PhD, Political Science, University of Washington, Seattle, WA
1983	MS, Political Science, State University of New York, Binghamton, NY
1979	BA, Philosophy and Religion, Northwest Nazarene College, Nampa, ID

PROFESSIONAL POSITIONS

2021-present	Affiliate Associate Professor, University of Washington School of Medicine, Department of Biomedical Informatics and Medical Education (BIME), Division of Biomedical Health Informatics, Seattle, WA
2020-present	Associate Investigator, *Kaiser Permanente Washington Health Research Institute, Seattle, WA
2017-2020	Assistant Investigator, *Kaiser Permanente Washington Health Research Institute, Seattle, WA
2004-2017	Clinical Data Scientist, *Kaiser Permanente Washington Health Research Institute, Seattle, WA
2000-2004	Research Scientist, Virology Research Clinic, University of Washington, Seattle, WA
1997-2000	President and owner, Carrell DB Inc., Seattle, WA (private contractor)
1993-1996	Research Assistant, Professor, School of Nursing, University of Washington, Seattle, WA
1990-1992	National Institute on Aging Fellow, Institute for Health and Aging, University of California, San Francisco
1990-1992	Pew Health Policy Fellow, Institute for Health Policy Studies, University of California, San Francisco <i>*Formerly Group Health Research Institute</i>

PROFESSIONAL HONORS

1983-1987 J. Allen Smith Graduate Student Scholarship, Department of Political Science, University of Washington, Seattle, WA

MEMBERSHIPS

2005-present **Member**, American Medical Informatics Association

ORGANIZATIONAL SERVICE

2021-present **Coordinator**, KPWHRI scientific seminar series
2021-present **Member**, hiring committee, KPWHRI Computational Linguist
2021 **Member**, hiring committee, KPWHRI IT Senior Manager
2020 **Member**, KPWHRI Equitable Evaluation Hiring Workgroup
2015-present **Member**, KPWHRI Data & Informatics Strategy Committee
2017-2018 **Member**, KPWHRI biostatistician search committee
2012-2015 **Member**, KP Washington, Information Technology Strategy Council

OTHER PROFESSIONAL SERVICE

2019-2020 **Participant**, NLP and Machine Learning Workgroup and CTN Data Science Summary Institute Workgroup, National Institute of Drug Abuse, Clinical Trials Network
2016-2019 **Member**, Women in AMIA Taskforce, American Medical Informatics Association
2014-2015 **Member**, Epic® Research Data Network, Scientific/Clinical Usefulness Workgroup
2013 **Panelist**, Workshop on NLP Methods for Pharmacoepidemiology, ISPE, Montreal, Canada, August 25-28, 2013
2011 **Conference organizer**, "Security Roundtable for Cloud-Deployed Clinical Natural Language Processing," sponsored by the Office of the National Coordinator for Health Information Technology, Seattle, WA, May 23-24, 2011 (attended by representatives of 14 institutions nationwide, including researchers and private sector business partners)
2008 **Conference organizer**, "Developing Clinical Natural Language Processing (NLP) Capacity in Applied Research Settings," Group Health Research Institute, Seattle, WA, December 17-18, 2008 (attended by representatives of nine institutions in the Health Care Systems Research Network)

SPECIAL NATIONAL RESPONSIBILITIES

2021	Reviewer , National Cancer Institute (NCI), Small Business Innovation Research (SBIR) Contract Review Meetings: Software to Address Social Determinants of Health in Oncology Practices (TEP-8B, Topic 423), and Digital Healthcare Platform to Reduce Financial Hardship for Cancer Patients (TEP-8B, Topic 384). March 18, 2021.
2020	Reviewer , National Library of Medicine (NLM), Special Emphasis Panel, Computational Approaches to Curation at Scale for Biomedical Research Assets. May 21, 2020.
2020	Reviewer , National Heart, Lung and Blood Institute (NHLBI), Special Emphasis Panel, Phenotype Finder in Data Resources. March 23, 2020.
2019	Reviewer , NIH Small Business Innovation Research (SBIR) Biotechnology Review Panel.
2015-2017	Standing member , Veteran's Administration, Health Services Research and Development (HSR&D), Healthcare Informatics Review Panel (3-year term).
2012-2015	Ad hoc member , Veteran's Administration, Health Services Research and Development (HSR&D), Healthcare Informatics Review Panel.
2014	Panelist , Patient Centered Outcomes Research Institute Review Panel.
2013	Reviewer , Cancer Research Network, Pilot Fund Proposals.
2012	Panelist , National Institute of Environmental Health Sciences Review Panel.
2011	Reviewer , National Heart, Lung and Blood Institute (NHLBI), Data Discovery Among NHLBI Genomic Studies (UG2/UG3).

EDITORIAL RESPONSIBILITIES

2022-present	Nature, reviewer
2022-present	PLOS Digital Health, reviewer
2022-present	Drug Safety, reviewer
2022-present	Clinical Pharmacology & Therapeutics, reviewer
2021-present	Pharmacoepidemiology and Drug Safety, reviewer
2021-present	Frontiers in Digital Health, reviewer
2021-present	Gates open research, reviewer
2021-present	Nature, reviewer
2021-present	Vaccine, reviewer
2020-present	Editorial Board Member, Journal of the American Medical Informatics Association (JAMIA)
2019-present	Applied Clinical Informatics, reviewer
2019-present	Substance Abuse, reviewer
2017-present	International Journal of Medical Informatics, reviewer
2017-present	Medical Informatics and Decision Making, reviewer
2012-present	Journal of the American Medical Informatics Association, reviewer

2013-2020 Journal of Biomedical Informatics, reviewer
 2017 Journal of Medical Internet Research, reviewer
 2015 Methods of Information in Medicine, reviewer
 2014 American Journal of Epidemiology, reviewer

RESEARCH FUNDING (last updated 12/1/2020)

Project Title	PI	Funder	Direct Costs	Role	Dates	Effort
Completed						
Programmer Support for NLP Software Deployment	Carrell	Internal	\$14,658	PI	8/16/09 - 12/31/09	2%
Natural Language Processing for Cancer Research Network Surveillance Studies (ARRA Challenge Grant) RC1CA146917	Carrell	NCI	\$688,161	PI	9/30/09 - 8/31/12	25%
SHARP Area 4: Secondary Use of EHR Data (Office of the National Coordinator) 90TR0002/01	Chute	ONC	\$373,665	Site-PI	4/1/10 - 12/18/13	20%
Consolidation of Text Mining Resources and Tools	Carrell	Internal	\$81,416	PI	8/1/10 - 12/31/11	5%
eMERGE Network - Genetic Discovery and Application U01HG006375	Larson/ Jarvik	NHGRI	\$2,912,449	Co-I	8/15/11 - 8/31/15	10%
Patient Outcomes for High Tech Diagnostic Imaging Studies (PCORI)	Solberg	PCORI	\$44,530	Site-PI	7/1/12 - 6/30/15	5%
Scalable and Robust Clinical Text De-Identification Tools (R01) R01LM011366	Carrell	NLM	\$703,238	PI	9/1/12 - 8/31/16	20%
Identifying Opioid Misuse and Abuse using EHR data and NLP 8500617525	Von Korff	Pfizer	\$156,149	Co-I	5/1/13 - 6/30/14	15%
Precursors of First-Episode Psychosis in a Population-Based Sample R01MH099666	Simon	NIMH	\$2,056,240	Co-I	7/19/13 - 6/30/17	2%
Measuring and Improving Colonoscopy Quality Using NLP (R01) R01CA168959	Mehrotra	NCI	\$165,402	Site-PI	8/1/13 - 2/28/17	20%
NLP Algorithms to Capture Molecular Markers from Tumors (CRN Pilot) U24CA171524	Kushi	NCI	\$40,441	Site-PI	9/1/13 - 10/31/14	10%
NLP for Identifying Clinical Risk Factors in Pulmonary Nodule Surveillance (CRN Pilot) U24CA171524	Farjah	NCI	\$45,619	Site-PI	9/1/13 - 10/31/14	10%
Identification of Opioid Abuse Using EHR Data and NLP Algorithms 8500617525	Carrell	Pfizer	\$156,166	PI	9/1/14 - 12/31/15	25%
Sustained Implementation of Patient-Centered Care for Alcohol Misuse R18HS023173	Bradley	AHRQ	\$480,400	Co-I	9/1/14 - 8/31/19	2%
U.S. Study of Uterine Fibroids: Case-Finding, Incidence and Patterns of Care	Scholes	Bayer	\$197,132	Co-I	9/15/14 - 9/30/16	2%
Computable Algorithms for Identifying Opioid Abuse and Addiction, Part I STUDY 3B	Carrell	Syneos	\$14,363	PI	4/1/15 - 5/31/15	20%
Computable Algorithms for Identifying Opioid Abuse and Addiction, Part II	Carrell	Syneos	\$1,282,991	PI	5/4/15 - 12/31/17	20%
eMERGE Network - Phase III U01HG008657	Larson/ Jarvik	NHGRI	\$2,725,300	Co-I	9/1/15 - 5/31/19	10%

Project Title	PI	Funder	Direct Costs	Role	Dates	Effort
Incidence and Predictors of Opioid Overdose and Death Using Coded Data and Natural Language Processing	Carrell	Syneos	\$531,100	Site-PI	2/1/16 - 12/31/18	20%
Incidence and Predictors of Opioid Overdose and Death Among ER/LA Opioid Analgesics Users as Measured by Diagnoses and Death Records SOW#5 2065-1B	Carrell	Syneos	\$415,000	Site-PI	2/1/16 - 4/16/19	20%
Endometriosis and Adenomyosis: Incidence, Case-finding and Patterns of Care 439 BAYER PHARMA AG	Scholes	Bayer	\$228,233	Co-I	3/21/16 - 6/30/17	2%
Family Network Approach to Assess Trickle-Down Effects of Genetic Testing	Larson/ Jarvik	NHGRI	\$100,000	Co-I	6/1/16 - 5/31/17	2%
Risk of Cognitive Dysfunction and Related Dementias Following General Anesthesia	Penfold	J&J	\$134,545	Co-I	7/1/16 - 10/31/17	10%
ER/LA Opioid PMR Consortium Data Analysis Coordinating Center	Nelson	FDA	\$1,260,000	Co-I	7/1/16 - 6/30/19	20%
PMI Recruitment Partnership in the Urban and Rural Northwest	Simon	NIH	\$621,091	Co-I	9/30/16 - 9/29/17	1%
eMERGE Phenotyping Improvement Supplement	Larson/ Jarvik	NHGRI	\$95,000	Co-I	9/30/16 - 9/29/18	20%
Geocoding the eMERGE Consortium U01HG008657	Larson/ Jarvik	NHGRI	\$45,000	Co-I	9/30/16 - 9/29/19	10%
Development and Validation of a Tool for Automated Monitoring of Medication Allergies and Side Effects (AMMASE)	Carrell	Internal	\$56,000	PI	1/1/17- 12/31/18	7%
Factors Associated with Sustained Opioid Taper 3110021954	Carrell	Purdue	\$245,500	PI	12/1/17- 2/28/19	20%
Analysis of the Risk of Motor Vehicle Accidents in Patients with Obstructive Sleep Apnea JA #32369	Hansen	Jazz Pharma	\$214,800	Co-I	2/1/18 - 7/30/19	5%
Current						
Integrating Addiction Research in Health Systems (CTN) UG1DA040314	Campbell/ Bradley	NIDA	\$684,900	Co-I	9/1/15 - 5/31/20	5%
Integrating Addiction Research in Health Systems (CTN Supplement) UG1DA040314-0651	Campbell/ Bradley	NIDA	\$895,00	Co-I	9/1/15 - 5/31/20	2%
The Electronic Medical Records and Genomics Network (eMERGE III) U01HG008657	Larson/ Jarvik	NHGRI	\$568,500	Co-I	9/1/15 - 3/31/20	10%
Seeing and Sharing What Matters: Personal Health Libraries for Individuals with Multiple Chronic Conditions R01LM012813	Ralston	NLM	\$880,500	Co-I	9/7/17 - 8/31/21	10%
Randomized Trial of Pain Self-Management Training to Support Prescription Opioid Taper R01DA044970	Sullivan	NIDA	\$1,169,500	Co-I	3/15/18 - 12/31/21	5%
Sentinel Operations Center Learning Phenotype Laboratory – Phase I Task Order	Platt/Nelson/ Carrell	FDA	\$402,000	Co-PI	7/1/18 - 8/31/20	15%
Assisted Identification and Navigation of Early Mental Health Symptoms in Youth R56MH117767	Penfold	NIMH	\$583,800	Co-I	8/3/18 - 12/31/19	7%
Susceptibility Gene for Erectile Dysfunction R01DK104764	Wessells	NIDDKD	\$57,200	Co-I	12/1/18 - 5/31/20	7%
Detecting Cost-Related Cancer Treatment Conversations in the Electronic Medical Record	Yeung	Internal	\$30,000	Co-I	1/1/19 - 12/31/19	2%
Sentinel Operations Center Learning Phenotype Laboratory – Phase II Task Order 75F40119D10037	Platt/Nelson/ Carrell	FDA		Co-PI	4/1/19 - 3/31/21	10%

Project Title	PI	Funder	Direct Costs	Role	Dates	Effort
Predicting Self-Harm, Suicide Attempt and Suicidal Death Using Longitudinal EHR, Claims and Mortality Data	Pathak	NIMH	\$186,300	Co-I	4/1/19 - 3/31/23	2%
The Electronic Medical Records and Genomics Network (eMERGE III), Natural Language Processing Supplement U01HG008657	Larson/ Jarvik	NHGRI	\$2,724,000	Co-I	4/25/19 - 3/31/20	20%
FDA Sentinel Innovation Center	Schneeweiss: Contact PI	FDA	\$249,000	Co-I	9/28/19 - 9/27/24	5%
Feasibility Assessment: Post-Marketing Safety Surveillance Among Patients using Tanezumab	Carrell	RUF	\$26,608	Site-PI	2/1/20 - 9/30/20	7%
An Electronic Health Record-Based Tool to Identify Newly Diagnosed Breast Cancer Patients at Risk of Low Social Support	Campbell	NCI	\$456,949	Co-I	7/1/20 - 6/30/24	10%
Implementing Routine Screening for Cannabis and Other Drug Use Disorders in Primary Care R01DA047312-01A1	Carrell	NIDA	\$649,000	PI	8/15/20 - 7/31/23	30%
Pending						
Innovation Project 2: Unstructured EHR Data and Analytic Methods to Improve Signal Identification (Task Order, FDA Sentinel Innovation Center)	Schneeweiss: Contact PI	FDA	TBD	Co-PI	11/1/20 - 10/30/22	20%
Improving Suicide Risk Prediction by Natural Language Processing of Clinical Notes	Carrell	NIMH	\$1,999,082	PI	7/1/21 - 6/30/25	30%

TEACHING AND MENTORING RESPONSIBILITIES

Course Responsibilities

- 1988-1989 **Lecturer**, Introduction to Statistics for Business Majors, University of Washington School of Business, Seattle, WA (lecturer for three consecutive quarters)
- 2014 **Guest lecturer**, “Risks and Opportunities in Secondary Use of Electronic Health Record Data,” Information School, Master’s degree in cyber security students, University of Washington, Seattle, WA

Formal Mentoring of Junior Faculty and Professional Staff

- 2018-present Mentor to Kai Yeung, PharmD, PhD, Assistant Investigator, Kaiser Permanente Washington Health Research Institute, on incorporating clinical natural language processing (NLP) methods into their program of research
- 2017-present Mentor to Vina Graham, Programmer/Analyst, Kaiser Permanente Washington Health Research Institute, on professional and career development
- 2017-2019 Mentor to Lu Chen, PhD, Research Associate I, Kaiser Permanente Washington Health Research Institute, on incorporating clinical natural language processing (NLP) methods into their program of research
- 2016-present Mentor to Arvind Ramaprasan, Programmer/Analyst, Kaiser Permanente Washington Health Research Institute, on professional and career development

2015-2017	Recruited and mentored Andrew Baer, MS, Computational Linguist, on career development
2012-present	Recruited and mentored David Cronkite, MS, Computational Linguist, on development of skills needed in application of clinical NLP in health research, contributing to scientific manuscripts

PUBLICATIONS

Peer-reviewed Publications of Original Work

(** = senior author)

1. **Carrell DS**. Whither the Revolution? *The Tocqueville Review*. 1987;8:39-92.
2. Bowen DJ, Urban N, **Carrell DS**, Kinne S. Comparisons of strategies to prevent breast cancer mortality. *J Soc Issues*. 1993 Summer;49(2):35-60. PMID: 17165234.
3. Bowen DJ, Kestin M, McTiernan A, **Carrell DS**, Green P. Effects of Dietary Fat Intervention on mental health in women. *Cancer Epidemiol Biomarkers Prev*. 1995 Jul-Aug;4(5):555-9. PMID: 7549814.
4. Wilkerson JD, **Carrell DS**. Money, Politics, and Medicine: The American medical PAC's strategy of giving in U.S. house races. *J Health Polit Policy Law*. 1999 Apr;24(2):335-55. PMID: 10321360.
5. Wald A, **Carrell DS**, Remington M, Kexel E, Zeh J, Corey L. Two-day Regimen of Acyclovir for treatment of recurrent genital herpes simplex virus type 2 infection. *Clin Infect Dis*. 2002 Apr;34(7):944-8. PMID: 11880960.
6. Engelberg R, **Carrell DS**, Krantz E, Corey L, Wald A. natural history of genital herpes simplex virus type 1 infection. *Sex Transm Dis*. 2003 Feb;30(2):174-7. PMID: 12567178.
7. Wald A, Huang ML, **Carrell DS**, Selke S, Corey L. Polymerase chain reaction for detection of herpes simplex virus (HSV) DNA on mucosal surfaces: Comparison with HSV isolation in cell culture. *J Infect Dis*. 2003 Nov;188(9):1345-51. PMID: 14593592.
8. Meyer JL, Crosby RA, Whittington WL, **Carrell DS**, Ashley-Morrow R, Meier AS, Harrington RD, DiClemente R, Wald A. The psychosocial impact of serological herpes simplex type 2 testing in an urban HIV clinic. *Sex Transm Infect*. 2005 Aug;81(4):309-15. PMC1745021.
9. Bonomi AE, Thompson RS, Anderson M, Rivara FP, Holt VL, **Carrell DS**, Martin DP. Ascertainment of intimate partner violence using two abuse measurement frameworks. *Inj Prev*. 2006 Apr;12(2):121-4. PMC2564437.
10. Bonomi AE, Thompson RS, Anderson M, Reid RJ, **Carrell DS**, Dimer JA, Rivara FP. Intimate partner violence and women's physical, mental, and social functioning. *Am J Prev Med*. 2006 Jun;30(6):458-66. PMID: 16704938.

11. Rivara FP, Anderson ML, Fishman P, Bonomi AE, Reid RJ, **Carrell DS**, Thompson RS. Intimate partner violence and health care costs and utilization for children living in the home. *Pediatrics*. 2007 Dec;120(6):1270-7. PMID: 18055676.
12. Bonomi AE, Anderson ML, Rivara FP, Cannon EA, Fishman PA, **Carrell DS**, Reid RJ, Thompson RS. Health care utilization and costs associated with childhood abuse. *J Gen Intern Med*. 2008 Mar;23(3):294-9. PMC2359481.
13. Ralston JD, Rutter CM, **Carrell DS**, Hecht J, Rubanowice D, Simon GE. Patient use of secure electronic messaging within a shared medical record: A cross-sectional study. *J Gen Intern Med*. 2009 Mar;24(3):349-55. PMC2642567.
14. Green BB, Ralston JD, Fishman PA, Catz SL, Cook A, Carlson J, Tyll L, **Carrell DS**, Thompson RS. Electronic communications and home blood pressure monitoring (e-BP) study: design, delivery, and evaluation framework. *Contemp Clin Trials*. 2008 May;29(3):376-95. PMC2645352.
15. Green BB, Cook AJ, Ralston JD, Fishman PA, Catz SL, Carlson J, **Carrell DS**, Tyll L, Larson EB, Thompson RS. Effectiveness of home blood pressure monitoring, web communication, and pharmacist care on hypertension control: A randomized controlled trial. *JAMA*. 2008 Jun;299(24):2857-67. PMC2715866.
16. Reid RJ, Bonomi AE, Rivara FP, Anderson ML, Fishman PA, **Carrell DS**, Thompson RS. Intimate partner violence among men prevalence, chronicity, and health effects. *Am J Prev Med*. 2008 Jun;34(6):478-85. PMID: 18471583.
17. Grembowski D, Anderson ML, Conrad DA, Fishman PA, Larson EB, Martin DP, Ralston JD, **Carrell DS**, Hecht J. Evaluation of the Group Health Cooperative access initiative: Study design challenges in estimating the impact of a large-scale organizational transformation. *Qual Manag Health Care*. 2008 Oct-Dec;17(4):292-303. PMID: 19020399.
18. Bonomi AE, Anderson ML, Reid RJ, Rivara FP, **Carrell DS**, Thompson RS. Medical and psychosocial diagnoses in women with a history of intimate partner violence. *Arch Intern Med*. 2009 Oct;169:1692-7. PMID: 19822826.
19. Rivara FP, Anderson ML, Fishman P, Reid RJ, Bonomi AE, **Carrell DS**, Thompson RS. Age, period, and cohort effects on intimate partner violence. *Violence Vict*. 2009;24(5):627-38. PMID: 19852403.
20. Green BB, Wang CY, Horner K, Catz S, Meenan RT, Vernon SW, **Carrell DS**, Chubak J, Ko C, Laing S, Bogart A. Systems of support to increase colorectal cancer screening and follow-up rates (SOS): Design, challenges, and baseline characteristics of trial participants. *Contemp Clin Trials*. 2010 Nov;31(6):589-603. PMC2956868.
21. Schildcrout JS, Basford MA, Pulley JM, Masys DR, Roden DM, Wang D, Chute CG, Kullo IJ, **Carrell DS**, Peissig P, Kho A, Denny JC. An analytical approach to characterize morbidity profile dissimilarity between distinct cohorts using electronic medical records. *J Biomed Inform*. 2010 Dec;43(6):914-23. PMC2991387.

22. Denny JC, Crawford DC, Ritchie MD, Bielinski SJ, Basford MA, Bradford Y, Chai HS, Bastarache L, Zuvich R, Peissig P, **Carrell DS**, Ramirez AH, Pathak J, Wilke RA, Rasmussen L, Wang X, Pacheco JA, Kho AN, Hayes MG, Weston N, Matsumoto M, Kopp PA, Newton KM, Jarvik GP, Li R, Manolio TA, Kullo IJ, Chute CG, Chisholm RL, Larson EB, McCarty CA, Masys DR, Roden DM, de Andrade M. Variants near FOXE1 are associated with hypothyroidism and other thyroid conditions: Using electronic medical records for genome-and phenome-wide studies. *Am J Hum Genet.* 2011 Oct;89(4):529-42. PMC3188836.
23. **Carrell DS**. A strategy for deploying secure cloud-based natural language processing systems for applied research involving clinical text. Proceedings of the 44th Hawaii International Conference on System Sciences (HICSS). 2011 Jan 4-7: Kauai, HI. 11.
24. Floyd JS, Heckbert SR, Weiss NS, **Carrell DS**, Psaty BM. Use of administrative data to estimate the incidence of statin-related rhabdomyolysis. *JAMA.* 2012 Apr;307(15):1580-1582. PMC3483067.
25. Peissig PL, Rasmussen LV, Berg RL, Linneman JG, McCarty CA, Waudby C, Chen L, Denny JC, Wilke RA, Pathak J, **Carrell DS**, Kho AN, Starren JB. Importance of multi-modal approaches to effectively identify cataract cases from electronic health records. *J Am Med Inform Assoc.* 2012 Mar-Apr;19(2):225-234. PMC3277618.
26. Mosley JD, Van Driest SL, Larkin EK, Weeke PE, Witte JS, Wells QS, Karnes JH, Guo Y, Bastarache L, Olson LM, McCarty CA, Pacheco JA, Jarvik GP, **Carrell DS**, Larson EB, Crosslin DR, Kullo IJ, Tromp G, Kuivaniemi H, Carey DJ, Ritchie MD, Denny JC, Roden DM. Mechanistic phenotypes: An aggregative phenotyping strategy to identify disease mechanisms using GWAS data. *PLoS ONE.* 2012 Dec 12;8(12): e81503. PMC3861317.
27. **Carrell DS**, Malin B, Aberdeen J, Bayer S, Clark C, Wellner B, Hirschman L. Hiding in plain sight: Use of realistic surrogates to reduce exposure of protected health information in clinical text. *J Am Med Inform Assoc.* 2013 Mar-Apr;20(2):342-8. PMC3638183.
28. Crosslin DR, McDavid A, Weston N, Zheng X, Hart E, de Andrade M, Kullo IJ, McCarty CA, Doheny KF, Pugh E, Kho A, Hayes MG, Ritchie MD, Saip A, Crawford DC, Crane PK, Newton K, **Carrell DS**, Gallego CJ, Nalls MA, Li R, Mirel DB, Crenshaw A, Couper DJ, Tanaka T, van Rooij FJ, Chen MH, Smith AV, Zakai NA, Yango Q, Garcia M, Liu Y, Lumley T, Folsom AR, Reiner AP, Felix JF, Dehghan A, Wilson JG, Bis JC, Fox CS, Glazer NL, Cupples LA, Coresh J, Eiriksdottir G, Gudnason V, Bandinelli S, Frayling TM, Chakravarti A, van Duijn CM, Melzer D, Levy D, Boerwinkle E, Singleton AB, Hernandez DG, Longo DL, Witteman JC, Psaty BM, Ferrucci L, Harris TB, O'Donnell CJ, Ganesh SK, Larson EB, Carlson CS, Jarvik GP. Genetic variation associated with circulating monocyte count in the eMERGE Network. *Hum Mol Genet.* 2013 May;22(10):2119-27. PMC3633369.
29. Dublin S, Baldwin E, Walker RL, Christensen LM, Haug PJ, Jackson ML, Nelson JC, Ferraro J, **Carrell DS**, Chapman WW. Natural language processing to identify pneumonia from radiology reports. *Pharmacoepidemiol Drug Saf.* 2013 Aug;22(8):834-841. PMC3811072.

30. Fishman PA, Cook AJ, Anderson ML, Ralston JD, Catz SL, **Carrell DS**, Carlson J, Green BB. Improving BP control through electronic communications: An economic evaluation. *Am J Manag Care*. 2013 Sep;19(9):709-16. PMC3938103.
31. Gao H, Bowles EA, **Carrell DS**, Buist DSM. PS2-25: Using natural language processing to extract findings from mammography reports. *Clin Med Res*. 2013 Sep;11(3):154. PMC3788501.
32. Denny JC, Bastarache L, Ritchie MD, Carroll RJ, Zink R, Mosley JD, Field JR, Pulley JM, Ramirez AH, Bowton E, Basford MA, **Carrell DS**, Peissig PL, Kho AN, Pacheco JA, Rasmussen LV, Crosslin DR, Crane PK, Pathak J, Bielinski SJ, Pendergrass SA, Xu H, Hindorff LA, Li R, Manolio TA, Chute CG, Chisholm RL, Larson EB, Jarvik GP, Brilliant MH, McCarty CA, Kullo IJ, Haines JL, Crawford DC, Masys DR, Roden DM. Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat Biotechnol*. 2013;31(12):1102-1110. PMC3969265.
33. **Carrell DS**, Halgrim S, Tran DT, Buist DSM, Chubak J, Chapman WW, Savova G. Using natural language processing to improve efficiency of manual chart abstraction in research: The case of breast cancer recurrence. *Am J Epidemiol*. 2014 Mar 15;179(6):749-758. PMC3939853.
34. Rasmussen LV, Thompson WK, Pacheco JA, Kho AN, **Carrell DS**, Pathak J, Peissig PL, Tromp G, Denny JC, Starren JB. Design patterns for the development of electronic health record-driven phenotype extraction algorithms. *J Biomed Inform*. 2014 Oct;51:280-6. PMC4194216.
35. Rasmussen-Torvik LJ, Stallings SC, Gordon AS, Almoguera B, Basford MA, Bielinski SJ, Brautbar A, Brilliant MH, **Carrell DS**, Connolly JJ, Crosslin DR, Doheny KF, Gallego CJ, Gottesman O, Kim DS, Leppig KA, Li R, Lin S, Manzi S, Mejia AR, Pacheco JA, Pan V, Pathak J, Perry CL, Peterson JF, Prows CA, Ralston J, Rasmussen LV, Ritchie MD, Sadhasivam S, Scott SA, Smith M, Vega A, Vinks AA, Volpi S, Wolf WA, Bottinger E, Chisholm RL, Chute CG, Haines JL, Harley JB, Keating B, Holm IA, Kullo IJ, Jarvik GP, Larson EB, Manolio T, McCarty CA, Nickerson DA, Scherer SE, Williams MS, Roden DM, Denny JC. Design and anticipated outcomes of the eMERGE-PGx project: A multicenter pilot for preemptive pharmacogenomics in electronic health record systems. *Clin Pharmacol Ther*. 2014 Oct;96(4):482-9. PMC4169732.
36. Li M, **Carrell DS**, Aberdeen J, Hirschman L, Malin BA. De-identification of clinical narratives through writing complexity measures. *Int J Med Inform*. 2014 Oct;83(10):750-67. PMC4215974.
37. Cronin RM, Field JR, Bradford Y, Shaffer CM, Carroll RJ, Mosley JD, Bastarache L, Edwards TL, Hebring SJ, Lin S, Hindorff LA, Crane PK, Pendergrass SA, Ritchie MD, Crawford DC, Pathak J, Bielinski SJ, **Carrell DS**, Crosslin DR, Ledbetter DH, Carey DJ, Tromp G, Williams MS, Larson EB, Jarvik GP, Peissig PL, Brilliant MH, McCarty CA, Chute

- CG, Kullo IJ, Bottinger E, Chisholm R, Smith ME, Roden DM, Denny JC. Phenome-wide association studies demonstrating pleiotropy of genetic variants within FTO with and without adjustment for body mass index. *Front Genet.* 2014 Aug;5:250. PMC4134007.
38. Ritchie MD, Verma SS, Hall MA, Goodloe RJ, Berg RL, **Carrell DS**, Carlson CS, Chen L, Crosslin DR, Denny JC, Jarvik G, Li R, Linneman JG, Pathak J, Peissig P, Rasmussen LV, Ramirez AH, Wang X, Wilke RA, Wolf WA, Torstenson ES, Turner SD, McCarty CA. Electronic medical records and genomics (eMERGE) network exploration in cataract: Several new potential susceptibility loci. *Mol Vis.* 2014 Sep; 20:1281-95. PMC4168835.
39. Wu S, Miller T, Masanz J, Coarr M, Halgrim S, **Carrell DS**, Clark C. Negation's not solved: generalizability versus optimizability in clinical natural language processing. *PLoS ONE.* 2014 Nov 13;9(11): e112774. PMC4231086.
40. Huang SH, LePendu P, Iyer SV, Tai-Seale M, **Carrell DS**, Shah NH. Toward personalizing treatment for depression: Predicting diagnosis and severity. *J Am Med Inform Assoc.* 2014;21(6):1069-75. PMC4215055.
41. Crosslin DR, **Carrell DS**, Burt A, Kim DS, Underwood JG, Hanna DS, Comstock BA, Baldwin E, deAndrade M, Kullo IJ, Tromp G, Kuivaniemi H, Borthwick KM, McCarty CA, Peissig PL, Doheny KF, Pugh E, Kho A, Pacheco J, Hayes MG, Ritchie MD, Verma SS, Armstrong G, Stallings S, Denny JC, Carroll RJ, Crawford DC, Crane PK, Mukherjee S, Bottinger E, Li R, Keating B, Mirel DB, Carlson CS, Harley JB, Larson EB, Jarvik GP. Genetic variation in the HLA region is associated with susceptibility to herpes zoster. *Genes Immun.* 2015 Jan-Feb;16(1):1-7. PMC4308645.
42. Hylan TR, Von Korff M, Saunders K, Masters E, Palmer RE, **Carrell DS**, Cronkite D, Mardekian J, Gross D. Automated prediction of risk for problem opioid use in a primary care setting. *J Pain.* 2015 Apr;16(4):380-387. PMID: 25640294.
43. Gao H, Bowles EJA, **Carrell DS**, Buist DS. Using natural language processing to extract mammographic findings. *J Biomed Inform.* 2015 Apr;54:77-84. PMC4408241.
44. Wieneke AE, Bowles EJA, Cronkite D, Wernli KJ, Gao H, **Carrell DS**, Buist DSM. Validation of natural language processing to extract breast cancer pathology procedures and results. *J Pathol Inform.* 2015 Jun;6:38. PMC4485196.
45. Palmer RE, **Carrell DS**, Cronkite D, Saunders K, Gross DE, Masters E, Donevan S, Hylan TR, Von Korff M. The prevalence of problem opioid use in patients receiving chronic opioid therapy: Computer-assisted review of electronic health record clinical notes. *Pain.* 2015 Jul;156(7):1208-1214. PMID: 25760471.
46. Abdul-Baki H, Schoen RE, Dean K, Rose S, Leffler DA, Kuganeswaran E, **Carrell DS**, Mehrotra, A. Public reporting of colonoscopy quality is associated with an increase in endoscopist adenoma detection rate. *Gastrointest Endosc.* 2015 Oct;82(4):676-82. PMC4575767.
47. Gallego CJ, Burt A, Sundaresan AS, Ye Z, Shaw C, Crosslin DR, Crane PK, Fullerton SM, Hansen K, **Carrell DS**, Kuivaniemi H, Derr K, de Andrade M, McCarty CA, Kitchner TE,

- Ragon BK, Stallings SC, Papa G, Bochenek J, Smith ME, Aufox SA, Pacheco JA, Patel V, Friesema EM, Erwin AL, Gottesman O, Gerhard GS, Ritchie M, Motulsky AG, Kullo IJ, Larson EB, Tromp G, Brilliant MH, Bottinger E, Denny JC, Roden DM, Williams MS, Jarvik GP. Penetrance of hemochromatosis in HFE genotypes resulting in p.Cys282Tyr and p.[Cys282Tyr];[His63Asp] in the eMERGE network. *Am J Hum Genet.* 2015 Oct;97(4):512-20. PMC4596892.
48. Bielinski SJ, Pathak J, **Carrell DS**, Takahashi PY, Olson JE, Larson NB, et al. Liu H, Sohn S, Wells QS, Denny JC, Rasmussen-Torvik LJ, Pacheco JA, Jackson KL, Lesnick TG, Gullerud RE, Decker PA, Pereira NL, Ryu E, Dart RA, Peissig P, Linneman JG, Jarvik GP, Larson EB, Bock JA, Tromp GC, de Andrade M, Roger VL. A robust e-Epidemiology tool in phenotyping heart failure with differentiation for preserved and reduced ejection fraction: The Electronic Medical Records and Genomics (eMERGE Network). *J Cardiovasc Trans Res.* 2015 Nov;8(8):475-83. PMC4651838.
49. Mo H, Thompson WK, Rasmussen LV, Pacheco JA, Jiang G, Kiefer R, Zhu Q, Xu J, Montague E, **Carrell DS**, Lingren T, Mentch FD, Ni Y, Wehbe FH, Peissig PL1, Tromp G, Larson EB, Chute CG, Pathak J, Denny JC, Speltz P, Kho AN, Jarvik GP, Bejan CA, Williams MS17, Borthwick K, Kitchner TE, Roden DM, Harris PA. Desiderata for computable representations of electronic health records-driven phenotype algorithms. *J Am Med Inform Assoc.* 2015 Nov;22(6):1220-30. PMC4639716.
50. Solberg LI, Asche SE, Butler J, **Carrell DS**, Norton CK, Jarvik JG, Smith-Bindman, R, Tillema, JO, Whitebird, RR, Ziegenfuss, JY. The effect of achieving patient-reported outcome measures on satisfaction. *J Am Board Fam Med.* 2015 Nov-Dec;28(6):785-92. PMID: 26546655.
51. **Carrell DS**, Cronkite D, Palmer RE, Saunders K, Gross DE, Masters ET, Hylan TR, Von Korff M. Using natural language processing to identify problem usage of prescription opioids. *Int J Med Inform.* 2015 Dec;84(12):1057-1064. PMID: 26456569.
52. Borthwick KM, Smelser DT, Bock JA, Elmore JR, Ryer EJ, Ye Z, Pacheco JA, **Carrell DS**, Michalkiewicz M, Thompson WK, Pathak J, Bielinski SJ, Denny JC, Linneman JG, Peissig PL, Kho AN, Gottesman O, Parmar H, Kullo IJ, McCarty CA, Böttinger EP, Larson EB, Jarvik GP, Harley JB, Bajwa T, Franklin DP, Carey DJ, Kuivaniemi H, Tromp G. ePhenotyping for abdominal aortic aneurysm in the electronic medical records and genomics (eMERGE) network: Algorithm development and Konstanz information miner workflow. *Int J Biomed Data Min.* 2015 Dec;4(1):113. PMC4820287.
53. Crosslin DR, Robertson PD, **Carrell DS**, Gordon AS, Hanna DS, Burt A, Fullerton SM, Scrol A, Ralston J, Leppig K, Hartzler A, Baldwin E, Andrade Md, Kullo IJ, Tromp G, Doheny KF, Ritchie MD, Crane PK, Nickerson DA, Larson EB, Jarvik GP. Prospective participant selection and ranking to maximize actionable pharmacogenetic variants and discovery in the eMERGE Network. *Genome Med.* 2015 Jul;7(1):67. PMC4517371.

54. Hall MA, Verma SS, Wallace J, Lucas A, Berg RL, Connolly J, Crawford DC, Crosslin D, de Andrade M, Doheny KF, Haines JL, Harley JB, Jarvik GP, Kitchner T, Kuivaniemi H, Larson EB, **Carrell DS**, Tromp G, Vrabec TR, Pendergrass SA, McCarty CA, Ritchie MD. Biology-driven gene-gene interaction analysis of age-related cataract in the eMERGE network. *Genet Epidemiol*. 2015 Jul;39(5):376-84. PMC4550090.
55. Van Driest SL, Wells QS, Stallings S, Bush WS, Gordon A, Nickerson DA, Kim JH, Crosslin DR, Jarvik GP, **Carrell DS**, Ralston JD, Larson EB, Bielinski SJ, Olson JE, Ye Z, Kullo IJ, Abul-Husn NS, Scott SA, Bottinger E, Almoguera B, Connolly J, Chiavacci R, Hakonarson H, Rasmussen-Torvik LJ, Pan V, Persell SD, Smith M, Chisholm RL, Kitchner TE, He MM, Brilliant MH, Wallace JR, Doheny KF, Shoemaker MB, Li R, Manolio TA, Callis TE, Macaya D, Williams MS, Carey D, Kapplinger JD, Ackerman MJ, Ritchie MD, Denny JC, Roden DM. Association of arrhythmia-related genetic variants with phenotypes documented in electronic medical records. *JAMA*. 2016 Jan;315(1):47-57. PMC4758131.
56. Simonti CN, Vernot B, Bastarache L, Bottinger E, **Carrell DS**, Chisholm RL, Crosslin DR, Hebring SJ, Jarvik GP, Kullo IJ, Li R, Pathak J, Ritchie MD, Roden DM, Verma SS, Tromp G, Prato JD, Bush WS, Akey JM, Denny JC, Capra JA. The phenotypic legacy of admixture between modern humans and Neandertals. *Science*. 2016 Feb;351(6274):737-41. PMC4849557.
57. Kirby JC, Speltz P, Rasmussen LV, Basford M, Gottesman O, Peissig PL, Pacheco, JA, Tromp, G, Pathak J, **Carrell DS**, Ellis SB, Lingren T, Thompson WK, Savova G, Haines J, Roden DM, Harris PA, Denny JC. PheKB: A catalog and workflow for creating electronic phenotype algorithms for transportability. *J Am Med Inform Assoc*. 2016 Mar;23(6):1046-1052. PMC5070514.
58. Mosley JD, Shaffer CM, Van Driest SL, Weeke PE, Wells QS, Karnes JH, Velez Edwards DR, Wei WQ, Teixeira PL, Bastarache L, Crawford DC, Li R, Manolio TA, Bottinger EP, McCarty CA, Linneman JG, Brilliant MH, Pacheco JA, Thompson W, Chisholm RL11, Jarvik GP, Crosslin DR, **Carrell DS**, Baldwin E, Ralston J, Larson EB, Grafton J, Scrol A, Jouni H, Kullo IJ, Tromp G, Borthwick KM, Kuivaniemi H, Carey DJ, Ritchie MD, Bradford Y, Verma SS, Chute CG, Veluchamy A, Siddiqui MK, Palmer CN, Doney A, MahmoudPour SH, Maitland-van der Zee AH1 Morris AD, Denny JC, Roden DM. A genome-wide association study identifies variants in KCNIP4 associated with ACE inhibitor-induced cough. *Pharmacogenomics J*. 2016 Jun;16(3):231-7. PMC4713364.
59. Li M, **Carrell DS**, Aberdeen J, Hirschman L, Kirby J, Li B, Vorobeychik Y, Malin BA. Optimizing annotation resources for natural language de-identification via a game theoretic framework. *J Biomed Inform*. 2016 Jun;61:97-109. PMC4996128.
60. Bush WS, Crosslin DR, Owusu-Obeng A, Wallace J, Almoguera B, Basford MA, Bielinski SJ, **Carrell DS**, Connolly JJ, Crawford D, Doheny KF, Gallego CJ, Gordon AS, Keating B, Kirby J, Kitchner T, Manzi S, Mejia AR, Pan V, Perry CL, Peterson JF, Prows CA, Ralston J, Scott SA, Scrol A, Smith M, Stallings SC, Veldhuizen T, Wolf W, Volpi S, Wiley K, Li R, Manolio T,

- Bottinger E, Brilliant MH, Carey D, Chisholm RL, Chute CG, Haines JL, Hakonarson H, Harley JB, Holm IA, Kullo IJ, Jarvik GP, Larson EB, McCarty CA, Williams MS, Denny JC, Rasmussen-Torvik LJ, Roden DM, Ritchie MD. Genetic variation among 82 pharmacogenes: The PGRN-Seq data from the eMERGE network. *Clin Pharmacol Ther*. 2016 Aug;100(2):160-9. PMC5010878.
61. **Carrell DS**, Cronkite DJ, Malin BA, Aberdeen JS, Hirschman L. Is the juice worth the squeeze? Costs and benefits of multiple human annotators for clinical text de-identification. *Methods Inf Med*. 2016 Aug;55(4):356-64. PMC5194214.
 62. Farjah F, Halgrim S, Buist DS, Gould MK, Zeliadt SB, Loggers ET, **Carrell DS**. An Automated method for identifying individuals with a lung nodule can be feasibly implemented across health systems. *eGEMs*. 2016 Aug;4(1):1254. PMC5013935.
 63. van 't Hof FN, Ruigrok YM, Lee CH, Ripke S, Anderson G, de Andrade M, Baas AF, Blankensteijn JD, Böttlinger EP, Bown MJ, Broderick J, Bijlenga P, **Carrell DS**, Crawford DC, Crosslin DR, Ebeling C, Eriksson JG, Fornage M, Foroud T, von Und Zu Fraunberg M, Friedrich CM, Gaál EI, Gottesman O, Guo DC, Harrison SC, Hernesniemi J, Hofman A, Inoue I, Jääskeläinen JE, Jones GT, Kiemeny LA, Kivisaari R, Ko N, Koskinen S, Kubo M, Kullo IJ, Kuivaniemi H, Kurki MI, Laakso A, Lai D, Leal SM, Lehto H, LeMaire SA, Low SK, Malinowski J, McCarty CA, Milewicz DM, Mosley TH, Nakamura Y, Nakaoka H, Niemelä M, Pacheco J, Peissig PL, Pera J, Rasmussen-Torvik L, Ritchie MD, Rivadeneira F, van Rij AM, Santos-Cortez RL, Saratzis A, Slowik A, Takahashi A, Tromp G, Uitterlinden AG, Verma SS, Vermeulen SH, Wang GT; Aneurysm Consortium; Vascular Research Consortium of New Zealand, Han B, Rinkel GJ, de Bakker PI. Shared genetic risk factors of intracranial, abdominal, and thoracic aneurysms. *J Am Heart Assoc*. 2016 Jul 14;5(7):e002603. PMC5015357.
 64. Jones GT, Tromp G, Kuivaniemi H, Gretarsdottir S, Baas AF, Giusti B, Strauss E, van 't Hof FN, Webb T, Erdman R, Ritchie MD, Elmore JR, Verma A, Pendergrass S, Kullo IJ, Ye Z, Peissig PL, Gottesman O, Verma SS, Malinowski J, Rasmussen-Torvik LJ, Borthwick K, Smelser DT, Crosslin DR, de Andrade M, Ryer EJ, McCarty CA, Bottinger EP, Pacheco JA, Crawford DC, **Carrell DS**, Gerhard GS, Franklin DP, Carey DJ, Phillips VL, Williams MJ, Wei W, Blair R, Hill AA, Vasudevan TM, Lewis DR, Thomson IA, Krysa J, Hill GB, Roake J, Merriman TR, Oszkinis G, Galora S, Saracini C, Abbate R, Pulli R, Pratesi C, Saratzis A, Verissimo A, Bumpstead SJ, Badger SA, Clough RE, Cockerill GW, Hafez H, Scott DJ, Futers TS, Romaine SP, Bridge K, Griffin KJ, Bailey MA, Smith A, Thompson MM, van Bockxmeer F, Matthiasson SE, Thorleifsson G, Thorsteinsdottir U, Blankensteijn JD, Teijink JA, Wijmenga C, de Graaf J, Kiemeny LA, Lindholt JS, Hughes AE, Bradley DT, Stirrups K, Golledge J, Norman PE, Powell JT, Humphries SE, Hamby SE, Goodall AH, Nelson CP, Sakalihasan N, Courtois A, Ferrell RE, Eriksson P, Folkersen L, Franco-Cereceda A, Eicher JD, Johnson AD, Betsholtz C, Ruusalepp A, Franzén O, Schadt E, Björkegren JL, Lipovich L, Drolet AM, verhoeven E, Zebregts CJ, Geelkerken RH, van Sambeek MR, van Sterkenburg SM, Jean-Paul P de Vries JP, Stefansson K, Thompson JR,

- de Bakker PI, Deloukas P, Sayers RD, Harrison S, van Rij AM, Samani NJ, Bown MJ. Meta-analysis of genome-wide association studies for abdominal aortic aneurysm identifies four new disease-specific risk loci. *Circ Res*. 2017 Jan;120(2):341-353. PMC5253231.
65. De R, Verma SS, Holzinger E, Hall M, Burt A, **Carrell DS**, Crosslin DR, Jarvik GP, Kuivaniemi H, Kullo IJ, Lange LA, Lanktree MB, Larson EB, North KE, Reiner AP, Tragante V, Tromp G, Wilson JG, Asselbergs FW, Drenos F, Moore JH, Ritchie MD, Keating B, Gilbert-Diamond D. Identifying gene-gene interactions that are highly associated with four quantitative lipid traits across multiple cohorts. *Hum Genet*. 2017 Feb;136(2):165-178. PMID: 27848076.
66. Schmidt AF, Swerdlow DI, Holmes MV, Patel RS, Fairhurst-Hunter Z, Lyall DM, Hartwig FP, Horta BL, Hyppönen E, Power C, Moldovan M, van Iperen E, Hovingh GK, Demuth I, Norman K, Steinhagen-Thiessen E, Demuth J, Bertram L, Liu T, Coassin S, Willeit J, Kiechl S, Willeit K, Mason D, Wright J, Morris R, Wanamethee G, Whincup P, Ben-Shlomo Y, McLachlan S, Price JF, Kivimaki M, Welch C, Sanchez-Galvez A, Marques-Vidal P, Nicolaides A, Panayiotou AG, Onland-Moret NC, van der Schouw YT, Matullo G, Fiorito G, Guarrera S, Sacerdote C, Wareham NJ, Langenberg C, Scott R, Luan J, Bobak M, Malyutina S, Pajak A, Kubinova R, Tamosiunas A, Pikhart H, Husemoen LL, Grarup N, Pedersen O, Hansen T, Linneberg A, Simonsen KS, Cooper J, Humphries SE, Brilliant M, Kitchner T, Hakonarson H, **Carrell DS**, McCarty CA, Kirchner HL, Larson EB, Crosslin DR, de Andrade M, Roden DM, Denny JC, Carty C, Hancock S, Attia J, Holliday E, O'Donnell M, Yusuf S, Chong M, Pare G, van der Harst P, Said MA, Eppinga RN, Verweij N, Snieder H; LifeLines Cohort study group, Christen T, Mook-Kanamori DO, Gustafsson S, Lind L, Ingelsson E, Pazoki R, Franco O, Hofman A, Uitterlinden A, Dehghan A, Teumer A, Baumeister S, Dörr M, Lerch MM, Völker U, Völzke H, Ward J, Pell JP, Smith DJ, Meade T, Maitland-van der Zee AH, Baranova EV, Young R, Ford I, Campbell A, Padmanabhan S, Bots ML, Grobbee DE, Froguel P, Thuillier D, Balkau B, Bonnefond A, Cariou B, Smart M, Bao Y, Kumari M, Mahajan A, Ridker PM, Chasman DI, Reiner AP, Lange LA, Ritchie MD, Asselbergs FW, Casas JP, Keating BJ, Preiss D, Hingorani AD; UCLEB consortium, Sattar N. PCSK9 genetic variants and risk of type 2 diabetes: a mendelian randomisation study. *Lancet Diabetes Endocrinol*. 2017 Feb;5(2):97-105. PMC5266795.
67. Almoguera B, Vazquez L, Mentch F, Connolly J, Pacheco JA, Sundaresan AS, Peissig PL, Linneman JG, McCarty CA, Crosslin D, **Carrell DS**, Lingren T, Namjou-Khales B, Harley JB, Larson E, Jarvik GP, Brilliant M, Williams MS, Kullo IJ, Hysinger EB, Sleiman PM, Hakonarson H. Identification of four novel loci in asthma in European American and African American populations. *Am J Respir Crit Care Med*. 2017 Feb;195(4):456-463. PMC5378422.
68. Dumitrescu L, Ritchie MD, Denny JC, El Rouby NM, McDonough CW, Bradford Y, Ramirez AH, Bielinski SJ, Basford MA, Chai HS, Peissig P, **Carrell D**, Pathak J, Rasmussen LV, Wang X, Pacheco JA, Kho AN, Hayes MG, Matsumoto M, Smith ME, Li R, Cooper-DeHoff RM, Kullo IJ, Chute CG, Chisholm RL, Jarvik GP, Larson EB, Carey D, McCarty CA, Williams MS,

- Roden DM, Bottinger E, Johnson JA, de Andrade M, Crawford DC.; Genome-wide study of resistant hypertension identified from electronic health records. *PLoS One*. 2017 Feb;12(2):e0171745. PMC5319785.
69. Sanderson SC, Brothers KB, Mercaldo ND, Clayton EW, Antommaria AHM, Aufox SA, Brilliant MH, Campos D, **Carrell DS**, Connolly J, Conway P, Fullerton SM, Garrison NA, Horowitz CR, Jarvik GP, Kaufman D, Kitchner TE, Li R, Ludman EJ, McCarty CA, McCormick JB, McManus VD, Myers MF, Scrol A, Williams JL, Shrubsole MJ, Schildcrout JS, Smith ME, Holm IA. Public attitudes toward consent and data sharing in biobank research: A large multi-site experimental survey in the US. *Am J Hum Genet*. 2017 Mar;100(3):414-427. PMC5339111.
70. Simon GE, Coleman KJ, Yarborough BJH, Operskalski B, Stewart C, Hunkeler EM, Lynch F, **Carrel DS**, Beck A. First presentation with psychotic symptoms in a population-based sample. *Psychiatr Serv*. 2017 May;68(5):456-461. PMC5811263.
71. Sims R, van der Lee SJ, Naj AC, Bellenguez C, Badarinarayan N, Jakobsdottir J, Kunkle BW, Boland A, Raybould R, Bis JC, Martin ER, Grenier-Boley B, Heilmann-Heimbach S, Chouraki V, Kuzma AB, Sleegers K, Vronskaya M, Ruiz A, Graham RR, Olaso R, Hoffmann P, Grove ML, Vardarajan BN, Hiltunen M, Nöthen MM, White CC, Hamilton-Nelson KL, Epelbaum J, Maier W, Choi SH, Beecham GW, Dulary C, Herms S, Smith AV, Funk CC, Derbois C, Forstner AJ, Ahmad S, Li H, Bacq D, Harold D, Satizabal CL, Valladares O, Squassina A, Thomas R, Brody JA, Qu L, Sánchez-Juan P, Morgan T, Wolters FJ, Zhao Y, Garcia FS, Denning N, Fornage M, Malamon J, Naranjo MCD, Majounie E, Mosley TH, Dombroski B, Wallon D, Lupton MK, Dupuis J, Whitehead P, Fratiglioni L, Medway C, Jian X, Mukherjee S, Keller L, Brown K, Lin H, Cantwell LB, Panza F, McGuinness B, Moreno-Grau S, Burgess JD, Solfrizzi V, Proitsi P, Adams HH, Allen M, Seripa D, Pastor P, Cupples LA, Price ND, Hannequin D, Frank-García A, Levy D, Chakrabarty P, Caffarra P, Giegling I, Beiser AS, Giedraitis V, Hampel H, Garcia ME, Wang X, Lannfelt L, Mecocci P, Eiriksdottir G, Crane PK, Pasquier F, Boccardi V, Henández I, Barber RC, Scherer M, Tarraga L, Adams PM, Leber M, Chen Y, Albert MS, Riedel-Heller S, Emilsson V, Beekly D, Braae A, Schmidt R, Blacker D, Masullo C, Schmidt H, Doody RS, Spalletta G, Longstreth WT Jr, Fairchild TJ, Bossù P, Lopez OL, Frosch MP, Sacchinelli E, Ghetti B, Yang Q, Huebinger RM, Jessen F, Li S, Kamboh MI, Morris J, Sotolongo-Grau O, Katz MJ, Corcoran C, Dunstan M, Braddel A, Thomas C, Meggy A, Marshall R, Gerrish A, Chapman J, Aguilar M, Taylor S, Hill M, Fairén MD, Hodges A, Vellas B, Soininen H, Kloszewska I, Daniilidou M, Uphill J, Patel Y, Hughes JT, Lord J, Turton J, Hartmann AM, Cecchetti R, Fenoglio C, Serpente M, Arcaro M, Caltagirone C, Orfei MD, Ciamarella A, Pichler S, Mayhaus M, Gu W, Lleó A, Fortea J, Blesa R, Barber IS, Brookes K, Cupidi C, Maletta RG, **Carrell DS**, Sorbi S, Moebus S, Urbano M, Pilotto A, Kornhuber J, Bosco P, Todd S, Craig D, Johnston J, Gill M, Lawlor B, Lynch A, Fox NC, Hardy J; ARUK Consortium, Albin RL, Apostolova LG, Arnold SE, Asthana S, Atwood CS, Baldwin CT, Barnes LL, Barral S, Beach TG, Becker JT, Bigio EH, Bird TD, Boeve BF, Bowen JD, Boxer A, Burke JR, Burns JM, Buxbaum JD, Cairns NJ, Cao C, Carlson

CS, Carlsson CM, Carney RM, Carrasquillo MM, Carroll SL, Diaz CC, Chui HC, Clark DG, Cribbs DH, Crocco EA, DeCarli C, Dick M, Duara R, Evans DA, Faber KM, Fallon KB, Fardo DW, Farlow MR, Ferris S, Foroud TM, Galasko DR, Gearing M, Geschwind DH, Gilbert JR, Graff-Radford NR, Green RC, Growdon JH, Hamilton RL, Harrell LE, Honig LS, Huentelman MJ, Hulette CM, Hyman BT, Jarvik GP, Abner E, Jin LW, Jun G, Karydas A, Kaye JA, Kim R, Kowall NW, Kramer JH, LaFerla FM, Lah JJ, Leverenz JB, Levey AI, Li G, Lieberman AP, Lunetta KL, Lyketsos CG, Marson DC, Martiniuk F, Mash DC, Masliah E, McCormick WC, McCurry SM, McDavid AN, McKee AC, Mesulam M, Miller BL, Miller CA, Miller JW, Morris JC, Murrell JR, Myers AJ, O'Bryant S, Olichney JM, Pankratz VS, Parisi JE, Paulson HL, Perry W, Peskind E, Pierce A, Poon WW, Potter H, Quinn JF, Raj A, Raskind M, Reisberg B, Reitz C, Ringman JM, Roberson ED, Rogava E, Rosen HJ, Rosenberg RN, Sager MA, Saykin AJ, Schneider JA, Schneider LS, Seeley WW, Smith AG, Sonnen JA, Spina S, Stern RA, Swerdlow RH, Tanzi RE, Thornton-Wells TA, Trojanowski JQ, Troncoso JC, Van Deerlin VM, Van Eldik LJ, Vinters HV, Vonsattel JP, Weintraub S, Welsh-Bohmer KA, Wilhelmsen KC, Williamson J, Wingo TS, Woltjer RL, Wright CB, Yu CE, Yu L, Garzia F, Golamaully F, Septier G, Engelborghs S, Vandenberghe R, De Deyn PP, Fernandez CM, Benito YA, Thonberg H, Forsell C, Lilius L, Kinhult-Stählbom A, Kilander L, Brundin R, Concari L, Helisalmi S, Koivisto AM, Haapasalo A, Dermecourt V, Fievet N, Hanon O, Dufouil C, Brice A, Ritchie K, Dubois B, Himali JJ, Keene CD, Tschanz J, Fitzpatrick AL, Kukull WA, Norton M, Aspelund T, Larson EB, Munger R, Rotter JJ, Lipton RB, Bullido MJ, Hofman A, Montine TJ, Coto E, Boerwinkle E, Petersen RC, Alvarez V, Rivadeneira F, Reiman EM, Gallo M, O'Donnell CJ, Reisch JS, Bruni AC, Royall DR, Dichgans M, Sano M, Galimberti D, St George-Hyslop P, Scarpini E, Tsuang DW, Mancuso M, Bonuccelli U, Winslow AR, Daniele A, Wu CK; GERAD/PERADES, CHARGE, ADGC, EADI, Peters O, Nacmias B, Riemenschneider M, Heun R, Brayne C, Rubinsztein DC, Bras J, Guerreiro R, Al-Chalabi A, Shaw CE, Collinge J, Mann D, Tsolaki M, Clarimón J, Sussams R, Lovestone S, O'Donovan MC, Owen MJ, Behrens TW, Mead S, Goate AM, Uitterlinden AG, Holmes C, Cruchaga C, Ingelsson M, Bennett DA, Powell J, Golde TE, Graff C, De Jager PL, Morgan K, Ertekin-Taner N, Combarros O, Psaty BM, Passmore P, Younkin SG, Berr C, Gudnason V, Rujescu D, Dickson DW, Dartigues JF, DeStefano AL, Ortega-Cubero S, Hakonarson H, Campion D, Boada M, Kauwe JK, Farrer LA, Van Broeckhoven C, Ikram MA, Jones L, Haines JL, Tzourio C, Launer LJ, Escott-Price V, Mayeux R, Deleuze JF, Amin N, Holmans PA, Pericak-Vance MA, Amouyel P, van Duijn CM, Ramirez A, Wang LS, Lambert JC, Seshadri S, Williams J, Schellenberg GD. Rare coding variants in PLCG2, ABI3, and TREM2 implicate microglial-mediated innate immunity in Alzheimer's disease. *Nat Genet.* 2017 Sep;49(9):1373-1384. PMC5669039.

72. Holzinger ER, Verma SS, Moore CB, Hall M, De R, Gilbert-Diamond D, Lanktree MB, Pankratz N, Amuzu A, Burt A, Dale C, Dudek S, Furlong CE, Gaunt TR, Kim DS, Riess H, Sivapalaratnam S, Tragante V, van Iperen EPA, Brautbar A, **Carrell DS**, Crosslin DR, Jarvik GP, Kuivaniemi H, Kullo IJ, Larson EB, Rasmussen-Torvik LJ, Tromp G, Baumert J,

- Cruickshanks KJ, Farrall M, Hingorani AD, Hovingh GK, Kleber ME, Klein BE, Klein R, Koenig W, Lange LA, März W, North KE, Charlotte Onland-Moret N, Reiner AP, Talmud PJ, van der Schouw YT, Wilson JG, Kivimaki M, Kumari M, Moore JH, Drenos F, Asselbergs FW, Keating BJ, Ritchie MD. Discovery and replication of SNP-SNP interactions for quantitative lipid traits in over 60,000 individuals. *BioData Min.* 2017 Jul 24;10:25. PMC5525436.
73. Huhdanpaa HT, Tan WK, Rundell SD, Suri P, Chokshi FH, Comstock BA, Heagerty PJ, James KT, Avins AL, Nedeljkovic SS, Nerenz DR, Kallmes DF, Luetmer PH, Sherman KJ, Organ NL, Griffith B, Langlotz CP, **Carrell DS**, Hassanpour S, Jarvik, JG. Using natural language processing of free-text radiology reports to identify type 1 modic endplate changes. *J Digit Imaging.* 2018 Feb;(31):84-90. PMC5788819.
 74. **Carrell DS**, Schoen RE, Leffler DA, Morris M, Rose S, Baer A, Crockett SD, Gourevitch RA, Dean KM, Mehrotra A. Challenges in adapting existing clinical natural language processing systems to multiple, diverse health care settings. *J Am Med Inform Assoc.* 2017 Sep;24(5):986-991. PMC6080843.
 75. Mehrotra A, Morris M, Gourevitch RA, **Carrell DS**, Leffler DA, Rose S, Greer JB, Crockett SD, Baer A, Schoen RE. Physician characteristics associated with higher adenoma detection rate. *Gastrointest Endosc.* 2018 Mar;87(3):778-786.e5. PMC5817032.
 76. Solberg LI, Asche SE, Butler JC, **Carrell DS**, Norton CK, Jarvik JG, Smith-Bindman R, Tillema JO, Whitebird RR, Werner AM, Ziegenfuss JY. Patient centered outcomes measurement: Does it require report from patients? *Patient Cent Res Rev.* 2017 Nov;4(4):221-229. PMC6664353.
 77. Simon GE, Stewart C, Hunkeler EM, Yarborough BJ, Lynch F, Coleman KJ, Beck A, Operskalsk BH, Penfold RS, **Carrell DS**. Care pathways prior to first diagnosis of psychotic disorder in adolescents and young adults. *Am J Psychiatry.* 2018 May;175(5): 434-442. PMC5930077.
 78. Gourevitch RA, Rose S, Crockett SD, Morris M, **Carrell DS**, Greer JB, Pai RK, Schoen RE, Mehrotra A. Variation in pathologist classification of colorectal adenomas and serrated polyps. *Am J Gastroenterol.* 2018 Mar;113(3):431-439. PMC6049074.
 79. Lapham GT, Lee A, Caldeiro RM, Glass JE, **Carrell DS**, Richards JE, Bradley KA. Prevalence of behavioral health conditions across frequency of cannabis use among adult primary care patients in Washington state. *J Gen Intern Med.* 2018 Nov;33(11):1833-1835. PMC6206344.
 80. Mosley JD, Feng Q, Wells QS, Van Driest SL, Shaffer CM, Edwards TL, Bastarache L, Wei W, Davis LK, McCarty CA, Thompson W, Chute CG, Jarvik GP, Gordon AS, Palmer MR, Crosslin DR, Larson EB, **Carrell DS**, Kullo IJ, Pacheco JA, Peissig PL, Brilliant MH, Linneman JG, Namjou B, Williams MS, Ritchie MD, Borthwick KM, Verma SS, Karnes JH, Weiss ST, Wang TJ, Stein CM, Denny JC, Roden DM. A study paradigm to integrate prospective

- epidemiologic cohorts and electronic health records to identify new disease biomarkers. *Nat Commun*. 2018 Aug;9(1):3522. PMC6117367.
81. Stanaway IB, Hall TO, Rosenthal EA, Palmer M, Naranbhai V, Knevel R, Namjou-Khales B, Carroll RJ, Kiryluk K, Gordon AS, Linder J, Howell KM, Mapes BM, Lin FTJ, Joo YY, Hayes MG, Gharavi AG, Pendergrass SA, Ritchie MD, de Andrade M, Croteau-Chonka DC, Raychaudhuri S, Weiss ST, Lebo M, Amr SS, **Carrell D**, Larson EB, Chute CG, Rasmussen-Torvik LJ, Roy-Puckelwartz MJ, Sleiman P, Hakonarson H, Li R, Karlson EW, Peterson JF, Kullo IJ, Chisholm R, Denny JC, Jarvik GP; eMERGE Network, Crosslin DR. The eMERGE genotype set of 83,717 subjects imputed to ~40 million variants genome wide and association with the herpes zoster medical record phenotype. *Genet Epidemiol*. 2019 Feb;43(1):63-81. PMC6375696.
 82. Mercaldo ND, Brothers KB, **Carrell DS**, Clayton EW, Connolly JJ, Holm IA, Horowitz CR, Jarvik GP, Kitchner T, Li R, McCarty CA, McMormick JB, McManus VD, Myers MF, Pankratz JJ, Shrubsole MJ, Smith ME, Stallings SC, Williams JL, Schildcrout JS. Enrichment sampling for a multi-site patient survey using electronic health records and census data. *Am J Am Med Inform Assoc*. 2019 Mar;26(3):219-227. PMC6351976.
 83. Crockett DS, Gourevitch RA, Morris M, **Carrell DS**, Rose S, Shi Z, Greer JB, Schoen RE, Mehrotra A. Serrated polyp detection is related to training and colonoscopy volume: Results from a multicenter study. *Endoscopy*. 2018 Oct;50(10):984-992. PMC6160341.
 84. Liu Y, Wan Z, Xia W, Kantarcioglu M, Vorobeychik Y, Clayton EW, Kho A, **Carrell D**, Malin BA. Detecting the presence of an individual in phenotypic summary data. Paper accepted at: AMIA 2018 Annual Symposium. 2018 Dec;2018:760-769. PMC6371366.
 85. Mosley JD, Benson MD, Smith JG, Melander O, Ngo D, Shaffer CM, Ferguson JF, McCarty CA, Thompson W, Chute CG, Jarvik GP, Gordon AS, Palmer MR, Crosslin DR, Larson EB, Scrol A, **Carrell DS**, Kullo IJ, Pacheco JA, Peissig PL, Brilliant MH, Kitchner TE, Linneman JG, Namjou B, Williams MS, Ritchie MD, Borthwick KM, Verma SS, Pendergrass SA, Kiryluk K, Mentch FD, Sleiman PM, Karlson EW, Zhu Y, Ramachandran VS, Yang Q, Denny JC, Roden DM, Gerszten RE, Wang TJ. Probing the virtual proteome to identify novel disease biomarkers. *Circulation*. 2018 Nov 27;138(22):2469-2481. PMID: 30571344.
 86. Tan WK, Hassanpour S, Heagerty PJ, Rundell SD, Suri P, Huhdanpaa HT, James K, **Carrell DS**, Langlotz CP, Organ NL, Meier EN, Sherman KJ, Kallmes DF, Luetmer PH, Griffith B, Nerenz DR, Jarvik JG. A natural language processing system to identify lumbar spine imaging findings related to low back pain. *Acad Radiol*. 2018 Nov;25(11):1422-1432. PMC6162177.
 87. Ezaz G, Leffler DA, Beach S, Schoen RE, Crockett SD, Gourevitch RA, Rose S, Morris M, **Carrell DS**, Greer JB, Mehrotra A. Association between endoscopist personality and rate of adenoma detection. *Clin Gastroenterol Hepatol*. 2019 Jul;17(8):1571-1579.e7. PMID: 30326300.

88. **Masters ET, Ramaprasan A, Mardekian J, Saunders K, Palmer RE, Gross DE, Cronkite D, Von Korff M, **Carrell DS**. Natural language processing-identified problem opioid use and its associated health care costs. *J Pain Palliat Care Pharmacother*. 2018 Jun - Sep;32(2-3):106-115. PMID: 30702378.
89. Chiang T, Liu X, Wu T-J, Hu J, Sedlazeck FJ, White S, Schaid DJ, de Andrade M, Jarvik G, Crosslin D, Stanaway IB, **Carrell DS**, Connolly JJ, Hakon Hakonarson H, Groopman EE, Gharavi AG, Fedotov A, Bi W, Leduc MS, Murdock DR, Jiang Y, Meng L, Eng CM, Wen S, Yang Y, Muzny DM, Boerwinkle E, William Salerno E, Venner E, Gibbs RA. Atlas-CNV: A validated approach to call Single-Exon CNVs in the eMERGESeq gene panel. *Genet Med*. 2019 Sep;21(9):2135-2144. PMC6752313.
90. **Carrell DS**, Cronkite DJ, Li R, Aberdeen J, Malin BA, Hirschman L. The machine giveth and the machine taketh away: A parrot attack on clinical text de-identified with hiding in plain sight. *J Am Med Inform Assoc*. 2019 Dec ;26(12):1536-1544. PMC6857511.
91. Gordon AS, Rosenthal ER, **Carrell DS**, Amendola LM, Dorschner MO, Scrol A, Stanaway IB, Ralston JD, Zouk H, Rehm HL, Larson E, Crosslin D, Leppig KA, Jarvik GP. Rates of actionable genetic findings in individuals with colorectal cancer or polyps ascertained from a community medical setting. *Am J Hum Genet*. 2019 Sep;105(3):526-533. PMC6731361.
92. Taylor CO, Lemke KW, Richards TM, Roe KD, He T, Arruda-Olson A, **Carrell DS**, Denny JC, Hripcsak G, Kiryluk K, Kullo I, Larson EB, Peissig P, Walton NA, Wei-Qi W, Ye Z, Chute CG, Weiner JP. Comorbidity characterization among eMERGE institutions: A pilot evaluation with the Johns Hopkins Adjusted Clinical Groups® system. *AMIA Jt Summits Transl Sci Proc*. 2019 May;2019:145-152. PMC6568092.
93. Namjou B, Lingren T, Huang Y, Parameswaran S, Cobb BL, Stanaway IB, Connolly JJ, Mentch FD, Benoit B, Niu X, Wei WQ, Carroll RJ, Pacheco JA, Harley ITW, Divanovic S, **Carrell DS**, Larson EB, Carey DJ, Verma S, Ritchie MD, Gharavi AG, Murphy S, Williams MS, Crosslin DR, Jarvik GP, Kullo IJ, Hakonarson H, Li R; eMERGE Network, Xanthakos SA, Harley JB. GWAS and enrichment analyses of non-alcoholic fatty liver disease identify new trait-associated genes and pathways across eMERGE Network. *BMC Med*. 2019 Jul;17(1):135. PMC6636057.
94. Hripcsak G, Shang N, Peissig PL, Rasmussen LV, Liu C, Benoit B, Carroll RJ, **Carrell DS**, Denny JC, Dikilitas O, Gainer VS, Marie Howell K, Klann JG, Kullo IJ, Lingren T, Mentch FD, Murphy SN, Natarajan K, Pacheco JA, Wei WQ, Wiley K, Weng C. Facilitating phenotype transfer using a common data model. *J Biomed Inform*. 2019 Aug;96: 103253. PMC6697565.
95. Shang N, Liu C, Rasmussen LV, Ta CN, Carroll RJ, Benoit B, Lingren T, Dikilitas O, Mentch FD, **Carrell DS**, Wei WQ, Luo Y, Gainer VS, Kullo IJ, Pacheco JA, Hakonarson H, Walunas TL, Denny JC, Wiley K, Murphy SN, Hripcsak G, Weng C. Making work visible for

- electronic phenotype implementation: Lessons learned from the eMERGE network. *J Biomed Inform.* 2019 Nov;99:103293. PMC6894517.
96. Green CA, Perrin NA, Hazlehurst B, Janoff SL, DeVeugh-Geiss A, **Carrell DS**, Grijalva CG, Liang C, Enger CL, Coplan PM. Identifying and classifying opioid-related overdoses: A validation study. *Pharmacoepidemiol Drug Saf.* 2019 Aug;28(8): 1127-1137. PMC6767606.
 97. Hazlehurst B, Green CA, Perrin NA, Brandes J, **Carrell DS**, Baer A, DeVeugh-Geiss A, Coplan PM. Using natural language processing of clinical text to enhance identification of opioid-related overdoses in electronic health records data. *Pharmacoepidemiol Drug Saf.* 2019 Aug;28(8): 1143-1151. PMC6772185.
 98. Schmidt AF, Holmes MV, Preiss D, Swerdlow DI, Denaxas S, Fatemifar G, Faraway R, Finan C, Valentine D, Fairhurst-Hunter Z, Hartwig FP, Horta BL, Hypponen E, Power C, Moldovan M, van Iperen E, Hovingh K, Demuth I, Norman K, Steinhagen-Thiessen E, Demuth J, Bertram L, Lill CM, Coassin S, Willeit J, Kiechl S, Willeit K, Mason D, Wright J, Morris R, Wanamethee G, Whincup P, Ben-Shlomo Y, McLachlan S, Price JF, Kivimaki M, Welch C, Sanchez-Galvez A, Marques-Vidal P, Nicolaides A, Panayiotou AG, Onland-Moret NC, van der Schouw YT, Matullo G, Fiorito G, Guarrera S, Sacerdote C, Wareham NJ, Langenberg C, Scott RA, Luan J, Bobak M, Malyutina S, Pajak A, Kubinova R, Tamosiunas A, Pikhart H, Grarup N, Pedersen O, Hansen T, Linneberg A, Jess T, Cooper J, Humphries SE, Brilliant M, Kitchner T, Hakonarson H, **Carrell DS**, McCarty CA, Lester KH, Larson EB, Crosslin DR, de Andrade M, Roden DM, Denny JC, Carty C, Hancock S, Attia J, Holliday E, Scott R, Schofield P, O'Donnell M, Yusuf S, Chong M, Pare G, van der Harst P, Said MA, Eppinga RN, Verweij N, Snieder H; Lifelines Cohort authors, Christen T, Mook-Kanamori DO; ICBP Consortium, Gustafsson S, Lind L, Ingelsson E, Pazoki R, Franco O, Hofman A, Uitterlinden A, Dehghan A, Teumer A, Baumeister S, Dörr M, Lerch MM, Völker U, Völzke H, Ward J, Pell JP, Meade T, Christophersen IE, Maitland-van der Zee AH, Baranova EV, Young R, Ford I, Campbell A, Padmanabhan S, Bots ML, Grobbee DE, Froguel P, Thuillier D, Roussel R, Bonnefond A, Cariou B, Smart M, Bao Y, Kumari M, Mahajan A, Hopewell JC, Seshadri S; METASTROKE Consortium of the ISGC, Dale C, Costa RPE, Ridker PM, Chasman DI, Reiner AP, Ritchie MD, Lange LA, Cornish AJ, Dobbins SE, Hemminki K, Kinnersley B, Sanson M, Labreche K, Simon M, Bondy M, Law P, Speedy H, Allan J, Li N, Went M, Weinhold N, Morgan G, Sonneveld P, Nilsson B, Goldschmidt H, Sud A, Engert A, Hansson M, Hemingway H, Asselbergs FW, Patel RS, Keating BJ, Sattar N, Houlston R, Casas JP, Hingorani AD. Phenome-wide association analysis of LDL-cholesterol lowering genetic variants in PCSK9. *BMC Cardiovasc Disord.* 2019 Oct 29;19(1):240. PMC6820948.
 99. Hall TO, Stanaway IB, **Carrell DS**, Carroll RJ, Denny JC, Hakonarson H, Larson EB, Mentch FD, Peissig PL, Pendergrass SA, Rosenthal EA, Jarvik GP, Crosslin DR. Unfolding of hidden white blood cell count phenotypes for gene discovery using latent class mixed modeling. *Genes Immun.* 2019 Sep;20(7):555-565. PMC6541537.

100. Robinson JR, Carroll RJ, Bastarache L, Chen Q, Mou Z, Wei WQ, Connolly JJ, Mentch F, Sleiman P, Crane PK, Hebring SJ, Stanaway IB, Crosslin DR, Gordon AS, Rosenthal EA, **Carrell DS**, Hayes MG, Wei W, Petukhova L, Namjou B, Zhang G, Safarova MS, Walton NA, Still C, Bottinger EP, Loos RJF, Murphy SN, Jackson GP, Kullo IJ, Hakonarson H, Jarvik GP, Larson EB, Weng C, Roden DM, Denny JC. Association of genetic risk of obesity with postoperative complications using mendelian randomization. *World J Surg*. 2020 Jan;44(1):84-94. PMID: 31605180; PMC6925615
101. Sullivan M, Boudreau D, Ichikawa L, Cronkite DJ, Albertson-Junkans L, Salgado G, VonKorff M, **Carrell DS**. Primary care opioid taper plans are associated with sustained opioid dose reduction. *J Gen Intern Med*. 2020 Mar; 35(3):687-695. PMC7080895.
102. **Carrell DS**, Malin BA, Cronkite DJ, Aberdeen JS, Clark C, Li M, Bastakoty D, Nyemba S, Hirschman L. Resilience of clinical text de-identified with hiding in plain sight to hostile re-identification attacks by human readers. *J Am Med Inform Assoc*. 2020 Jul 1;27(9):1374-1382. PMC7647331.
103. **Carrell DS**, Albertson-Junkans L, Ramaprasan A, Scull G, Mackwood M, Cronkite D, Baer A, Hansen K, Coplan P, DeVeaugh-Geiss A, Johnson E, Lange J, Shortreed S, Von Korff M. Measuring problem prescription opioid use among patients receiving long-term opioid analgesic treatment: Development and evaluation of an algorithm for use in EHR and claims data. *J Drug Assess*. 2020;9(1):97–105. PMC7241518.
104. Gordon AS, Zouk H, Venner E, Eng CM, Funke BH, Amendola LM, **Carrell DS**, Chisholm RL, Chung WL, Denny JC, Hakonarson H, Kullo IJ, Larson EB, Leduc MS, Leppig KA, Lennon NJ, Linder JE, Muzny DM, Prows CA, Rasmussen-Torvik LJ, Rasouly HM, Roden DM, Rosenthal EA, Smith ME, Stanaway IB, Van Driest SL, Walker K, Wiesner GL, Williams MS, Witkowski L, Crosslin DR, Gibbs RA, Rehm HL, The eMERGE Clinical Annotation Working Group, Jarvik GP. Frequency of genomic incidental findings among 21,915 eMERGE network participants. *Genetics in Medicine*. 2020 Sep;22(9):1470-1477. PMC7713503.
105. Joo YY, Lin FT, Pacheco JA, Thompson WK, Rasmussen-Torvik LJ, Gawron AJ, Rasmussen LV, Peissig PL, Veerappan A, Tromp GC, Borthwick KM, Denny JC, Carrell DS, de Andrade M, Pathak J, Bottinger E, Ellis SB, Gottesman O, Smith M, Chisholm RL, Hayes MG, Kho AN. A Genome-wide and Phenome-wide Association Study of Diverticular Disease Using Electronic Health Records. *BioRxiv*. June 9, 2020. DOI:10.1101/2020.06.08.138735.
106. The eMERGE Consortium (including **Carrell DS**). Lessons learned from the eMERGE Network: Balancing genomics in discovery and practice. *Human Genetics and Genomics Advances* 2, 100018, January 14, 2021.

107. Palmer MR, Kim DS, Crosslin DR, Stanaway IB, Rosenthal EA, **Carrell DS**, Cronkite DJ, Gordon A, Du X, Li YK, Williams MS, Weng C, Feng Q, Li R, Pendergrass SA, Hakonarson H, Fasel D, Sohn S, Sleiman P, Handelman SK, Speliotes E, Kullo IJ, Larson EB; eMERGE Consortium, Jarvik GP. Loci identified by a genome-wide association study of carotid artery stenosis in the eMERGE network. *Genet Epidemiol*. 2021 Feb;45(1):4-15. PMC7891640.
108. Li R, Duan R, Zhang X, Lumley T, Pendergrass S, Bauer C, Hakonarson H, Carrell DS, Smoller J, Wei W, Carroll R, Edwards DV, Wiesner G, Sleiman P, Denny J, Mosley J, Ritchie M, Chen Y, Moore J. Lossless integration of multiple electronic health records for identifying pleiotropy using summary statistics. *Nat Commun*. 2021 Jan 8;12(1):168. PMC7794298.
109. Rosenthal E, Gordon A, Crosslin DR, **Carrell DS**, Stanaway I, Larsen E, Grafton J, Wei W, Feng Q, Denny JC, Shah A, Sturm A, Ritchie M, Pacheco J, Hakonarson H, Rasmussen-Torvik L, Connolly J, Jarvik G. Association between triglycerides, known risk SNVs, and conserved rare variation in SLC25A40 in a multi-ancestry cohort. *BMC Med Genomics*. 2021 Jan 6;14(1):11. PMC7789246.
110. Suri, P, Stanaway, I B, Zhang, Y, Freidin, MB, Tsepilov, YA, **Carrell, DS**, Williams, FMK, Aulchenko, YS, Hakonarson, H, Namjou, B, Crosslin, DR, Jarvik, GP, Lee, MT. (2021). "Genome-wide association studies of low back pain and lumbar spinal disorders using electronic health record data identify a locus associated with lumbar spinal stenosis." *Pain*. 2021 Aug 1;162(8):2263-2272. PMC8277660.
111. Shang N Khan A Polubriaginof F, Zanoni F, Mehl K, Fasel D, Drawz PE, Carrol RJ, Denny JC, Hathcock MA, Arruda-Olson AM, Peissig PL, Dart RA, Brilliant MH, Larson EB, **Carrell DS**, Pendergrass S, Verma SS, Ritchie MD, Benoit B, Gainer VS, Karlson EW, Gordon AS, Jarvik GP, Stanaway IB, Crosslin DR, Mohan S, Ionita-Laza I, Tatonetti NP, Gharavi AG, Hripcsak G, Weng C, Kiryluk K. Medical records-based chronic kidney disease phenotype for clinical care and "big data" observational and genetic studies. *NPJ Digit Med*. 2021 Apr 13;4(1):70. PMC8044136.
112. Gruber S, **Carrell DS**, Floyd JS, Nelson JC, Hazlehurst BL, Heagerty PJ. Letter to the Editor re Beachler, et al, 2021. *Pharmacoepidemiol Drug Saf*. 2021;30(12):1735-36 PMID: 34409675.
113. Matson TE, **Carrell DS**, Bobb JF, Cronkite DJ, Oliver MM, Luce C, Ghitza UE, Hsu CW, Campbell CI, Browne KC, Binswanger IA, Saxon AJ, Bradley KA, Lapham GT. Prevalence of medical cannabis use and associated health conditions documented in electronic health records among primary care patients in Washington state. *JAMA Netw Open* 2021 May 3;4(5):e219375. PMC8103224.
114. Bann MA, **Carrell DS**, Gruber S, Shinde M, Ball R, Nelson JC, Floyd JS. Identification and Validation of Anaphylaxis Using Electronic Health Data in a Population-based Setting. *Epidemiology*. 2021 May 1;32(3):439-443. PMID: 33591057.

115. **Carrell DS**, Cronkite DJ, Shea M, Oliver MM, Luce C, Matson TE, Bobb JF, Hsu CW, Binswanger I, Brown K, Saxon A, McCormack J, Jelstrom E, Ghitza U, Campbell CI, Bradley KA, Lapham G. Clinical documentation of patient-reported medical cannabis use in primary care: Towards scalable extraction using natural language processing methods. *Subst Abus.* 2022;43(1):917-924. doi: 10.1080/08897077.2021.1986767.
116. Yu J, Pacheco JA, Ghosh A, Luo Y, Weng C, Shang N, Benoit B, **Carrell DS**, Carroll R, Dikilitas O, Freimuth R, Gainer V, Hakonarson H, Hripcsak G, Kullo IJ, Mentch F, Murphy S, Peissig P, Ramirez A, Walton N, Wei W, Rasmussen L. Under-specification as the source of ambiguity and vagueness in narrative phenotype algorithm definitions. *BMC Med Inform Decis Mak.* 2022;22:23. <https://doi.org/10.1186/s12911-022-01759-z>.
117. Glazer AM, Davogustto D, Shaffer CM, Vanoye CG, Desai RR, Farber-Eger EH, Dikilitas O, Shang N, Pacheco JA, Yang T, Muhammad A, Mosley JD, Van Driest SL, Wells QS, Rinke LL, Kalash OR, Wada Y, Bland S, Yoneda ZT, Mitchell DW, Kroncke BM, Kullo IJ, Jarvik GP, Gordon AS, Larson EB, Manolio TA, Mirshahi T, Luo JZ, Schaid D, Namjou B, Alsaied T, Singh R, Singhal A, Liu C, Weng C, Hripcsak G, Ralston JD, McNally EM, Chung WK, **Carrell DS**, Leppig KA, Hakonarson H, Sleiman P, Sohn S, Glessner J, Denny J, Wei J, George Jr. AL, Shoemaker MB, Roden DM. Arrhythmia variant associations and reclassifications in the eMERGE-III sequencing study. *Circulation.* 2022 Mar 22;145(12):877-891. doi: 10.1161/CIRCULATIONAHA.121.055562.
118. Lapham GT, Matson TE, **Carrell DS**, Bobb JF, Luce C, Oliver MM, Ghitza UE, Hsu C, Browne KC, Binswanger IA, Campbell CI, Saxon AJ, Vandrey R, Schauer G, Pacula RL, Horberg MA, Bailey SR, McClure EA, Bradley KA. Comparison of Medical Cannabis Use Reported on a Confidential Survey vs Documented in the Electronic Health Record Among Primary Care Patients. *JAMA Netw Open.* 2022 May 2;5(5):e2211677. doi: 10.1001/jamanetworkopen.2022.11677.
119. Penfold RB, **Carrell DS**, Cronkite D, Pabiniak C, Dodd TS, Glass A, Johnson E, Thompson E, Arrighi HM, Stang PE. Development of a machine learning model to predict mild cognitive impairment using natural language processing in the absence of screening. *BMC Med Inform Decis Mak.* 2022 May 12;22(1):129. doi: 10.1186/s12911-022-01864-z.
120. Floyd JS, Bann MA, Felcher AB, Sapp D, Nguyen MD, Ajao, A, **Carrell DS**, Nelson JC, Hazlehurst BL. Validation of acute pancreatitis among adults in an integrated healthcare system. *Epidemiology.* 2023 Jan 1;34(1):33-37. doi: 10.1097/EDE.0000000000001541. Epub 2022 Aug 25.
121. Shinde M, Rodriguez-Watson C, Zhang T, **Carrell DS**, Mendelsohn AB, Nam YH, Carruth A, Petronis KR, McMahaill-Walraven CN, Jamal-Allial A, Nair V, Pawloski PA, Hickman A, Brown MT, Francis J, Hornbuckle K, Brown JS, Mo J. Patient characteristics, pain treatment patterns, and incidence of total joint replacement in a US population with osteoarthritis. *BMC Musculoskelet Disord.* 2022 Sep 23;23(1):883. doi: 10.1186/s12891-022-05823-7.

122. **Carrel DS**, Gruber S, Floyd JS, Bann MA, Cushing-Haugen KL, Johnson RL, Graham V, Cronkite DJ, Hazlehurst BL, Felcher AH, Bejan CA, Kennedy A, Shinde M, Karami S, Ma Y, Stojanovic D, Zhao Y, Ball R, Nelson J. Improving methods of identifying anaphylaxis for medical product safety surveillance using natural language processing and machine learning. *Am J Epidemiol*. 2022 Nov 4;kwac182. doi: 10.1093/aje/kwac182. Online ahead of print.

Accepted peer-reviewed papers in press

123. Pacheco JA, Rasmussen LV, Wiley K, Person TN, Cronkite DJ, Sohn S, Gundelach JH, Gainer V, Castro VM, Liu C, Mentch F, Lingren T, Sundaresan AS, Eickelberg G, Willis V, Patel R, **Carrell DS**, Deng Y, Walton N, Satterfield BA, Kullo IJ, Dikilitas O, Smith JC, Peterson JF, Shang N, Kiryluk K, Ni Y, Li Y, Nadkarni, GN, Rosenthal EA, Walunas TL, Williams M, Karlson EW, Linder JE, Luo Y, Weng C, Wei W. Evaluation of the Portability of Computable Phenotypes with Natural Language Processing in the eMERGE Network. *Nature Sci Rpt*. Accepted 2022 Nov.

Submitted papers under review (as of 12/5/2022)

124. Joo YY, Lin FT, Pacheco JA, Thompson WK, Rasmussen-Torvik LJ, Gawron AJ, Rasmussen LV, Peissig PL, Veerappan A, Tromp GC, Borthwick KM, Denny JC, **Carrell DS**, de Andrade M, Pathak J, Bottinger E, Ellis SB, Gottesman O, Smith M, Chisholm RL, Hayes MG, Kho AN. Multi-ancestry genome- and phenome-wide association studies of diverticular disease in electronic health records with natural language processing enriched phenotyping algorithm. *PLOS ONE*. Under review 2022.
125. Brandt PS, Kho A, Luo Y, Pacheco JA, Walunas TL, Hakonarson H, Hripcsak G, Liu C, Shang N, Weng C, Walton N, **Carrell DS**, Crane PK, Larson E, Chute CG, Kullo I, Carroll R, Denny J, Ramirez A, Wei WQ, Pathak J, Wiley LK, Richesson R, Starren JB, Rasmussen LV. Characterizing Variability of EHR-Driven Phenotype Definitions. *J Am Med Inform Assoc*. Under review 2022.

Non-refereed papers/articles

126. **Carrell DS**, Ralston J. Messages, strands and threads: Measuring use of electronic patient-provider messaging. *AMIA Annu Symp Proc*. 2005:913. PMC1560601.
127. **Carrell DS**, Ralston JD. Variation in adoption rates of a patient web portal with a shared medical record by age, gender, and morbidity level. *AMIA Annu Symp Proc*. 2006:871. PMC1839457.
128. Ralston JD, Hereford J, **Carrell DS**. Use and satisfaction of a patient Web portal with a shared medical record between patients and providers. *AMIA Annu Symp Proc*. 2006:1070. PMC1839721.
129. **Carrell DS**, Miglioretti D, Smith-Bindman R. Coding free text radiology reports using the Cancer Text Information Extraction System (caTIES). *AMIA Annu Symp Proc*. 2007 Oct 11;889. PMID 18693990.

130. Ralston JD, **Carrell DS**, Reid R, Anderson M, Moran M, Hereford J. Patient web services integrated with a shared medical record: Patient use and satisfaction. *J Am Med Inform Assoc*. 2007 Nov-Dec;14(6):798-806. PMC2213480.
131. **Carrell DS**, Baldwin D. A method for discovering variant spellings of terms of interest in clinical text. *Clin Med Res*. 2010 Dec;8(3-4):183-198.
132. Conway M, Berg RL, **Carrell DS**, Denny JC, Kho AN, Kullo IJ, Linneman JG, Pacheco JA, Peissig P, Rasmussen L, Weston N, Chute CG, Pathak J. Analyzing the heterogeneity and complexity of electronic health record oriented phenotyping algorithms. *AMIA Annu Symp Proc*. 2011;2011:274-83. PMC3243189.
133. **Carrell DS**, Halgrim S, Buist DSM, Chubak, Tran DT. Carrell et al. respond to "Observational research and the EHR." *Am J Epidemiol*. 2014 Mar 15;179(6):762-763. PMC4081638.
134. **Carrell DS**, Mardekian J, Cronkite DJ, Ramaprasan A, Hansen K, Gross DE, Palmer RE, Masters E, Von Korff M. A fully automated algorithm for identifying patients with problem prescription opioid use using electronic health record data. *Drug Alcohol Depend*. 2017; 100.171: e36

PRESENTATIONS

(* = invited presentation)

1. * National Cancer Institute (NCI), Cancer Surveillance Planning Workshop. Member of expert panel on future directions for cancer surveillance research. Bethesda, MD. October 22-23, 2009.
2. * University of Washington, Computational Linguistics Seminar, "Applied Natural Language Processing for Clinical, Epidemiological and Health Systems Research." Seattle, WA, April 13, 2010.
3. * University of Pittsburgh Medical Center, Cancer Text Information Extraction System (caTIES) User Meeting. Invited presentation "Strategies for Implementing NLP Software in Applied Healthcare Research Settings," Pittsburgh, PA, January 26-28, 2010.
4. * National Academy of Medicine and the Office of the National Coordinator (ONC) for Health Information Technology sponsored workshop on "Electronic Infrastructure for the Learning Healthcare System," expert panel member addressing issues of information extraction from electronic medical record systems. Washington, DC, July 27-28, 2011.
5. * University of California San Diego, workshop sponsored by iDASH (Integrating Data for Analysis, Anonymization and Sharing) on "Overcoming barriers to secondary use of EHR data: Cloud deployed NLP," invited presentation. La Jolla, CA May 2-3, 2011.
6. * Johns Hopkins University, Institute for Computational Medicine, Cardiovascular Research Grid, Webinar presentation on "Automated Extraction of Cardiovascular Findings from Clinical Reports: The Promise, the Challenges, and Ideas for the Future," co-presenter Wendy W. Chapman. October 13, 2011.

7. * Stanford Center For Biomedical Informatics Research, Research Colloquium, “Resources and Tools to Support Collaborative Development and Application of Clinical Natural Language Processing Technologies.” Stanford, CA, April 12, 2012.
8. Epic Spring Research Advisory Council, “High-throughput Clinical Phenotyping using Epic Data.” Verona, WI. March 2, 2014.
9. *National Heart, Lung, and Blood Institute (NHLBI), “Using Electronic Health Record Data for Large-Scale Research: Successes and Challenges.” Workshop on The Future of Genomic Research for the Prevention and Treatment of Heart, Lung and Blood Diseases in Populations and Patients. Bethesda, MD. April 9-10, 2014
10. Developing and Validating Disease-Specific Algorithms Using Electronic Health Data. Health Care Systems Research Network Annual Meeting, San Diego, CA March 21, 2017.
11. Podium presentation: “Is an Algorithm Just an Algorithm? How to Define Clinical Concepts for Use in Comparative Safety and Effectiveness Distributed Research Networks.” American Medical Informatics Association (AMIA) Joint Summits. San Francisco, CA, March 28, 2017.
12. * U.S. Food and Drug Administration workshop. “Use of Natural Language Processing to Enhance Clinical Trials and Safety Surveillance of Medical Products” (co-hosted by the UCSF-Stanford Center of Excellence in Regulatory Science and Innovation, and San Francisco State University). Crown Hill, MD, June 15, 2017.
13. Podium presentation: “Automatic Text De-Identification: How and When is it Acceptable?” American Medical Informatics Association (AMIA) Annual Symposium, San Francisco, CA, November 5, 2018.
14. * Margolis Center for Health Policy (Duke University) workshop on Improving the Efficiency of Outcome Validation in the Sentinel System, “Identifying Best Practices for Extracting Standardized Data from Medical Records.” Margolis Center for Health Policy, Washington, DC, May 17, 2018.
15. * NIDA Clinical Trials Network workshop on Data Science: “Using machine learning and natural language processing (NLP) to assess substance use disorders.” Bethesda, MD, September 18, 2019.
16. * National Cancer Institute workshop on “De-identification of narrative text clinical documents: stakeholders’ perspectives and acceptance of machine-based de-identification.” Bethesda, MD, February 24-25, 2020.
17. * Sentinel Innovation Center on “Health Outcomes of Interest (HOI) 2.0: Aims, Approach and (Early) Lessons from the Learning Phenotype Lab Pilot Work” (with Jennifer Nelson). March 13, 2020.

18. * US Food and Drug Administration Monthly Seminar Series on “Electronic Health Records (EHR) + Natural Language Processing + Machine Learning = Improved Sentinel Outcome Detection Algorithms” (With Jennifer Nelson and Susan Gruber). November 19, 2020.
19. * International Society for Pharmacoepidemiology (ICPE) 2021 ICPE Skills Course, “Pharmacoepidemiologic Methods used in COVID-19 Research, 1.5 Years into the Pandemic,” presentation “Scalable, computable algorithms for identifying patients with health outcomes of interest such as COVID-19 using electronic health record data.” August 16, 2021.
20. * International Society for Pharmacoepidemiology (ICPE) 2021 presentation “Improving Outcome Ascertainment by Applying Natural Language Processing and Machine Learning to Electronic Health Record Data: Identifying Anaphylaxis” (Symposium on New frontiers in computable phenotyping for medical product safety evaluation). ICPE 2021 All Access Virtual international meeting. August 24, 2021.
21. * International Society for Pharmacoepidemiology (ICPE) 2021, Methods in Pharmacoepidemiology — Informatics, presentation “Improving Methods of Identifying Anaphylaxis for Medical Product Safety Surveillance Using Natural Language Processing and Machine Learning.” ICPE 2021 All Access Virtual international meeting. August 25, 2021.
22. * Great Lakes Pharmacovigilance and Drug Safety Forum invited presentation “Electronic Health Records + Natural Language Processing + Machine Learning = Improved Sentinel Outcome Detection Algorithms” (With Susan Gruber, PhD. October 1, 2021.
23. 2022 International Conference on Pharmacoepidemiology (ICPE), presentation (with Joshua C. Smith, PhD, Vanderbilt University; presented in our absence by Sengwee Darren Toh, ScD, Harvard Medical School) on “Data-Driven Approaches to Improve Phenotype Sensitivity Using EHR Data.” Copenhagen, Denmark. August 24-28, 2022.
24. * Fourteenth Annual Sentinel Initiative Public Workshop (US, Food and Drug Administration) presentation on “Automated approaches to developing computable phenotype algorithms using scalable natural language processing (NLP) and machine learning methods with COVID-19 disease as the example.”. November 16, 2022.
25. American Medical Informatics Association (AMIA) 2022 Annual Symposium, presentation (with Joshua C. Smith, PhD, Vanderbilt University) on “Data-Driven Automated Classification Algorithms for Acute Health Conditions: Applying PheNorm to COVID-19 Disease.” Session S39, Oral Presentations, Innovations in Informatics. Washington, DC, Nov. 7, 2022.