

# AI & RWD IN THE JOURNALS

**November 2020** - Roundup & links below.

## Life-Science:

**Merck** – lever Beijing hospital EMR to profile treatment and outcomes of 300 patients with advanced melanoma, highlighting poor overall outcomes

**Actelion** – lever the Optum database to suggest, across 100 patients, that selexipag is associated with lower hospitalization risk and rate in PAH

**Janssen/OHDSI** – on COVID “Fake” news or retracted publications, noting a number of open source tools (including OHDSI ones) to prevent errors leading to retraction

**Pfizer** – using Vendor Oncology EMR (ConcertAI), profile treatment patterns and clinical outcomes for 200+ patient on palbociclib-based therapy

**Roche** – pull EMR data from 120 French hospitals to describe treatment patterns across 28k patients with HER2+ BC

**Novo** – looks at data from Novo Nordisk’s Cornerstones4Care Powered by Glooko diabetes support app across 10k patients to uncover strategies for app engagement

**Novartis** – work with National Marrow Donor Program/Be The Match program’s infra for data collection on non-HCT cellular therapies to show outcomes with Tisagenlecleucel (CAR T-cell) across 400 patients

**EME/Pharma** – multiple organizations co-publish to profile utility of RWE in the cycle of medicinal product benefit-risk (BR) assessment, communication/risk minimization and evaluation (“BRACE”).

**Pharma Mar** – sponsor a study using IQVIA oncology advantage to profile treatment patterns in aSTS and use of trabectebin across 5k+ patients in Europe

**Sunovion** – levers linked IQVIA open medical claims and hospital charge master to profile HC utilization for 800 COPD patients using the hGLY eFlow nebulizer

## Vendors & networks:

**IQVIA** – developed a deep-learning model to predict patients developing NASH as progression from NAFLD, leveraging LTSM and liver biopsy tests (ACoE/AA)

**IQVIA** – on including CVX Vaccines within the overall OMOP vocabulary and mapping

**EMIF-IMI** – ongoing publications of predecessor to EHDEN EU network, using 6 databases and data on 300k risk factors to explore CV and dementia associations

**Eye Smart** - the Indian optha EMR vendor profiles the possibilities of RWE on a database with over 1m optha patients, in profiling risk factors to cataracts

**PHE/Nuffield** – use the UK GP network to study household ILI for last 5 years, showing that young children, female gender and household size as key association with increased ILI rates

**HDRUK** – use the cogstack NLP engine to analyse 1.2k COVID patient notes and associations with severe COVID, and did not find link of common medications (ACE-Is, AR2Bs) and severity

**AFDLI** – reflects on overall drug development, and that despite limitations in progress to date RWD remains the key approach to driving productivity

**X2AI Inc** – provides a descriptive of usage across 400 patients of TESS, their automated mental health chatbot powered by AI, and note future direction of evidence development

**FIRE** – use of the Swiss network of 400 GPs and FIRE project to characterize risk patterns in 1.3k patients with HF

#### **Academic medical science:**

**Research consortia** – a reflection on how public-private partnerships form, typical governance mechanisms, and strengths and weaknesses thereof (e.g., in upstream negotiation of funding terms)

**AI in AMD** – a clinical perspective on AI tools in optha, noting the potential, but that practical only one tool (IDx-DR) for diabetic retinopathy is both FDA approved and generally commercially available

**Brazil** – use of DataSUS government data to profile treatment patterns of 70k lymphoma cases, and how late diagnosis/treatment initiation drives higher mortality

**AI COVID** – an overview earlier in the pandemic, including predictions for ML SIR models that suggested the diseases would be suppressed (that did not turn out to be true as external conditions changed)

**Patient portals** – a systematic to review to characterize types of portal, purpose and adoption dynamics (including a clinical adoption meta-model of 4 phases of use)

**Data sharing** – US patient's attitudes on, showing ongoing support for use to support care delivery or research, but low support for financial related purposes

**RWD in trials** – a systematic review, and across 90 articles explicitly mentioning specific RWD databases, a focus on design aspects and baseline statistic comparisons

**EMR for RCT** – and editorial and overview aimed more at clinicians to derive potential leverage of EMRs for trials

**Germany** – use AOK Bayern claims data in a conjunct analysis with survey data of 12k patients to examine factors in COPD health-related quality of life (HRQoL)

**RWE in NSCLC** – a review of real-world evidence in EGFR mutation-positive non-small-cell lung cancer, including to inform clinical decisions in populations poorly covered by RCTs

#### **Genomics & Informatics:**

**Insilco genomics** – and example of using RNA-seq data and Monocle3 or PANTHER tools to try to better understand AP-2α/γ roles

**Genomics China** – a review of the national centres available databases (Aging Atlas, BrainBase, LncExpDB, TransCirc, BioProject, BioSample, GVM (Genome Variation Map), GEN (Gene Expression Nebulas), etc..)

**Clinical genomics** – a nice overview on biomarker to full genomic sequencing, and reflection that ML and CNNs might solve some of the issues in test accuracy

**Genomics in the Era of COVID** – a review of potential applications, but also concern that discovery of COVID resistance in

**Rare disease genomics** – a discussion on how to translate functional genomics of rare diseases into clinical space to accelerate rare disease diagnosis

**Transparency by design** – an evolution in the reflection on AI explain-ability to ensuring that a principles-based approach achieve transparency on AI

**NLP for COVID** – the NIH reviews various tools and algorithms used during the pandemic for training, medical information and extraction, as well as NLQ (chatbots)

**NER** – a nice overview on named entity recognition, typical techniques (ME, CRF, SSVM, RNN, LSTM), and example using pretraining with BiLSTM and Transformer

**NER** – a Chinese team look at “noisy label” or “uncertainty Strategy” in NER NLP, as form of adversarial training and minor improvements on CRF/BiLSTM/BERT models

**DQ frameworks** – a systematic review of data quality (DQ) frameworks, and suggestion of having a min set of DQ elements for the production of “FAIR” datasets

#### Life-Science:

[HTML] [Real-world clinical outcomes of anticancer treatments and prognostic factors in patients with advanced melanoma in China](#) C Cui, X Yan, B Li, L Si, C Zhihong, X Sheng, B Lian... - *IJS Oncology*, 2020 China has much lower 5-year survival rates among melanoma

[HTML] [Comparative effectiveness of oral prostacyclin pathway drugs on hospitalization in patients with pulmonary hypertension in the United States: a retrospective database ...](#) JW McConnell, Y Tsang, ... 12,13 Furthermore, hospitalizations in patients with PAH are costly, and readmission is common following discharge. 23. Hospitalization data can be extracted from medical administrative claims data, a source of real-world evidence

[PDF] [“Fake News” in the Rush to Inform on COVID-19: How Research Retractions Can Inform Best Practices for Real-World Data Quality and Reporting](#) RL DiSantostefano, MPH Clair Blacketer ... *BMJ* 2018;363:k3532 8. Wang SV, Pinheiro S, Hua W, Arlett P, Uyama Y, Berlin JA, Bartels D, Kahler K, Bessette LG, Schneeweiss S, 2020, A structured template and reporting tool for real-world evidence studies to facilitate

[PDF] [Treatment patterns and clinical outcomes with palbociclib-based therapy received in US community oncology practices](#) J Lin, L McRoy, MD Fisher, N Hu, C Davis, D Mitra... - *Future Oncology*, 2020 ... evidence of overall survival benefit. The present study contributes to a growing body of real-world evidence of treatment patterns and clinical effectiveness of palbociclib in real-world settings [9–18]. The objective of this study ...

[HTML] [Evolution in the real-world therapeutic strategies in more than 20,000 women with breast cancer having received human epidermal growth factor receptor 2–targeted ...](#) P Cottu, B Coudert, D Perol, A Doly, J Manson... - *European Journal of Cancer*, 2020 ...

Highlights. • There is a growing need for real-world data on how cancer treatments are used.

• ... We provide real-world data on 2011–2018 evolution in injectable HER2 treatments.

[HTML] [Real-World Evidence of User Engagement With Mobile Health for Diabetes Management: Longitudinal Observational Study](#) AK Böhm, ML Jensen, MR Sørensen, T Stargardt - *JMIR mHealth and uHealth*, 2020 ... Methods: The analysis is based on real-world data obtained by Novo Nordisk's Cornerstones4Care Powered by Glooko diabetes support app ... explores real-world data obtained by Novo Nordisk's Cornerstones4Care

[HTML] [Real-world evidence of tisagenlecleucel for pediatric acute lymphoblastic leukemia and non-Hodgkin lymphoma](#) MC Pasquini, ZH Hu, K Curran, T Laetsch, F Locke... - *Blood Advances*, 2020 Key Points. Represents the largest set of safety and efficacy data for tisagenlecleucel. Outcomes in the real-world setting are similar to results in the pivota.

[The Utility of Real-World Evidence for Benefit Risk Assessment, Communication, and Evaluation of Pharmaceuticals: Case studies](#) CA Radawski, TA Hammad, ... and Drug Safety, 2020 Purpose In recent years, novel types of real-world evidence (RWE) have played a role in various decision-making processes relating to medicinal products, including regulatory approval, patient access, health technology assessment, safety

[PDF] [Current treatment patterns in advanced soft tissue sarcoma: real-world evidence of over 5,000 European patients](#) A Le Cesne, J Martín-Broto, P Reichardt, P Picci... No part of this publication may be reproduced or photocopying without the prior written permission of

[HTML] [... resource utilization and exacerbations in patients with chronic obstructive pulmonary disease treated with glycopyrrolate/eFlow® in the US: A real-world data analysis](#) X Niu, V Divino, S Sharma, M Dekoven, VR Anupindi... - Journal of Medical ..., 2020 Aims This study compared medication use, healthcare resource utilization (HRU), and exacerbations among individuals with chronic obstructive pulmonary disease (COPD) who initiated glycopyrrolate/eFlow® Closed System nebulizer 25 mcg/m

### **Vendors & networks**

[A deep learning approach for classifying nonalcoholic steatohepatitis patients from nonalcoholic fatty liver disease patients using electronic medical records](#) PB Suresha, Y Wang, C Xiao, L Glass, Y Yuan... - Explainable AI in Healthcare ..., 2020 ... NAFLD. Specifically, we used the electronic medical record (EMR) data of patients to train a classifier that yields the probability of a patient developing NASH in the next month, given that the patient is already

[PDF] [Incorporation of CVX Vaccines into OMOP Drug Hierarchy](#) MD Violetta Komar, MD Dmitry Dymshits... ... 2IQVIA, Cambridge, MA, USA Page 2. Abstract To generate real-world evidence about the effectiveness of vaccination, we need to handle electronic health records and their inevitable processing in terms of ETL. Thus, there is an ...

[HTML] [Vascular and metabolic risk factor differences prior to dementia diagnosis: a multidatabase case-control study using European electronic health records](#) G Perera, PR Rijnbeek, M Alexander, D Ansell... - BMJ Open, 2020

[Biogeographical and Altitudinal Distribution of Cataract: A Nine-Year Experience Using Electronic Medical Record-Driven Big Data Analytics in India](#) H Garrigan, C Ifantides, GS Prashanthi, AV Das - Ophthalmic Epidemiology, 2020 Purpose: To use electronic medical record data to study the altitude, UV exposure, and biogeographical distribution of senile cataract in India. Methods: This is a hospital-based, cross-sectional study of patients over 40

[HTML] [Household presentation of influenza and acute respiratory illnesses to a primary care sentinel network: retrospective database studies \(2013–2018\)](#) S de Lusignan, J Sherlock, O Akinyemi, R Pebody... - BMC Public Health, 2020 Direct observation of the household spread of influenza and respiratory infections is limited; much of our understanding comes from mathematical models. The study aims to determine household incidence of influenza-like i

[HTML] [COVID-19 Analyses COVID-19 Analyses](#) S Sun, A Folarin, Y Ranjan, Z Rashid, P Conde... ... DMB is funded by a UKRI Innovation Fellowship (Health Data Research UK MR/S00310X/1). RB is funded in part by ... RJBD is supported by Health Data Research UK; the BigData@Heart Consortium, funded by the Innovativ

[HTML] [Real World Evidence: A Potential Solution](#) CT Fulmer ... "Real world evidence is the clinical evidence regarding the usage and potential benefits or risks of a medical product derived from analysis of RWD." [38] Real-world evidence can be derived from hybrid design trials, pragmatic trials, and observation studies

[HTML] [Artificial Intelligence Chatbot for Depression: Descriptive Study of Usage](#) G Dosovitsky, BS Pineda, NC Jacobson, C Chang... - JMIR Formative Research, 2020 Chatbots could be a scalable solution that provides an interactive means of engaging users in behavioral health interventions driven by artificial intelligence. Although some chatbots have shown promising early efficacy results

[HTML] [Heart failure epidemiology and treatment in primary care: a retrospective cross-sectional study](#) Y Rachamin, R Meier, T Rosemann, AJ Flammer... - ESC Heart Failure, 2020 ... namely sacubitril/valsartan. Methods and results. This was a retrospective cross-sectional study using data from an electronic medical record database of Swiss general practitioners from 2016 to 2019. Multilevel logistic regression

#### **Academic medical informatics:**

[HTML] [Governance of research consortia: challenges of implementing Responsible Research and Innovation within Europe](#) M Morrison, M Mourby, H Gowans, S Coy, J Kaye - Life Sciences, Society and Policy, 2020 ... Governance of research consortia: challenges of implementing Responsible Research and Innovation within Europe ... Metrics details. Abstract. Responsible Research and Innovation ('RRI') is a cross-cutting priority for ...

[PDF] [Artificial Intelligence in Age-related Macular Degeneration](#) LSM Mendonça, CR Bauml - Ophthalmology, 2020 The world as we knew it before the COVID-19 pandemic has changed. The year 2020 will be thought of as a turning point when people began to reconsider the need for professionals to be physically present to practice their professions;

[HTML] [Overview of lymphoma diagnosis in Brazilian public health system patients: Open data analysis for health care planning](#) DP Martins, NF Correa-Netto, N Melo, SR Loggett, Transfusion and Cell ..., 2020

[PDF] [Evolution of Artificial Intelligence during COVID 19 to Predict the Future](#) M Roy The outbreak of COVID-19 Coronavirus or a SARS-CoV-2 virus, has created a global catastrophe. The overall incident of COVID-19 is rapidly increasing day by day. Artificial Intelligence (AI), Machine learning (ML) and Cloud Computing can be more

[HTML] [The State of Evidence in Patient Portals: Umbrella Review](#) MG Antonio, O Petrovskaya, F Lau - Journal of Medical Internet Research, 2020 ... It is also known as a tethered personal health record, which is a web-based interface linked to an electronic health record (EHR) where patients can view and interact with their health care data and providers [11]. Comparison With Prior Work ...

[HTML] [The public's comfort with sharing health data with third-party commercial companies](#) MG Trinidad, J Platt, SLR Kardia - Humanities and Social Sciences Communications, 2020 Healthcare systems are using big data-driven methods to realize the vision of learning health systems and improve care quality. In so doing, many are partnering with third-party commercial companies to provide novel data

[Contemporary use of real-world data for clinical trial conduct in the United States: a scoping review](#) JR Rogers, J Lee, Z Zhou, YK Cheung, G Hripcsak... - Journal of the American ..., 2020.

Real-world data (RWD), defined as routinely collected healthcare data, can be a potential catalyst for addressing challenges faced in clinic.

[HTML] [Utility of Electronic Health Records in Randomized Controlled Trials](#) M Dotzert - Technology, 2020 ... In addition to their utility in intervention delivery, EHRs are increasingly used for recruitment and outcome assessment. EHRs also support pragmatic trials and may be incorporated into a hybrid design with both traditional and pragmatic elements ...

[HTML] [The relationship between body mass index and health-related quality of life in COPD: real-world evidence based on claims and survey data](#) MB Huber, C Kurz, F Kirsch, L

Schwarzkopf... - Respiratory Research, 2020 Body mass index (BMI) is an important parameter associated with mortality and health-related quality of life (HRQoL) in chronic obstructive pulmonary disease (COPD). However, informed guidance on stratified weight

[PDF] [Benefits and limitations of real-world evidence: lessons from EGFR mutation-positive non-small-cell lung cancer](#) B Nazha, JCH Yang, TK Owonikoko - Future Oncology, 2020 While randomized controlled trials (RCTs) are the gold standard for evidence-based medicine, they do not always reflect real-world patient populations, limiting their generalizability and external validity. Real-world evidence (RWE), generated during ...

#### **Bioinformatics & Genomics:**

[HTML] [Functional genomics of AP-2 \$\alpha\$  and AP-2 \$\gamma\$  in cancers: in silico study](#) D Kołat, Ż Kałuzińska, M Orzechowska, AK Bednarek... - BMC Medical Genomics, 2020 Among all causes of death, cancer is the most prevalent and is only outpaced by cardiovascular diseases. Molecular theory of carcinogenesis states that apoptosis and proliferation are regulated by groups of tumor suppressors or oncogenes ...

[PDF] [Database Resources of the National Genomics Data Center, China National Center for Bioinformation in 2021](#) Y Xue, Y Bao, Z Zhang, W Zhao, J Xiao, S He, G Zhang... - Nucleic Acids Research, 2020 The National Genomics Data Center (NGDC), part of the China National Center for Bioinformation (CNCB), provides a suite of database resources to support worldwide research activities in both academia and industry.

[Bioinformatics, Sequencing Accuracy, and the Credibility of Clinical Genomics](#) WG Feero - JAMA, 2020 The adoption of clinical exome and whole-genome sequencing based on next-generation sequencing technologies has increased rapidly over the last decade; this has been accelerated by increasing coverage of these services by private and public ...

[HTML] [Expediting rare disease diagnosis: a call to bridge the gap between clinical and functional genomics](#) SN Hartin, JC Means, JT Alaimo, ST Younger - Molecular Medicine, 2020 Approximately 400 million people throughout the world suffer from a rare disease. Although advances in whole exome and whole genome sequencing have greatly facilitated rare disease diagnosis, overall diagnostic rates remain below 50 ...

[HTML] [Genomics in the era of COVID-19: ethical implications for clinical practice and public health](#) G Geller, P Duggal, CL Thio, D Mathews, JP Kahn... - Genome Medicine, 2020 Genomic studies of patients with COVID-19, or exposed to it, are underway to delineate host factors associated with variability in susceptibility, infectivity, and disease severity.

[HTML] [Towards Transparency by Design for Artificial Intelligence](#) H Felzmann, E Fosch-Villaronga, C Lutz... - Science and Engineering ..., 2020 In this article, we develop the concept of Transparency by Design that serves as practical guidance in helping promote the beneficial functions of transparency while mitigating its challenges in automated-decision making

[PDF] [Artificial Intelligence \(AI\) in Action: Addressing the COVID-19 Pandemic with Natural Language Processing \(NLP\)](#) Q Chen, R Leaman, A Allot, L Luo, CH Wei, S Yan, Z Lu - arXiv preprint arXiv ..., 2020 The COVID-19 pandemic has had a significant impact on society, both because of the serious health effects of COVID-19 and because of public health measures implemented to slow its spread. Many of these difficulties are fundamentally ...

[HTML] [Clinical Named Entity Recognition from Chinese Electronic Medical Records Based on Deep Learning Pretraining](#) L Gong, Z Zhang, S Chen - Journal of Healthcare Engineering, 2020 . Clinical named entity recognition is the basic task of mining electronic medical records text, which are with some challenges containing the language features of Chinese electronic medical records text with many ...

[PDF] [Noisy Label Learning for Chinese Medical Named Entity Recognition Based on Uncertainty Strategy](#) Z Li, Z Gan, B Zhang, Y Chen, J Wan, J Zhao ... In this task, we need to identify the entity boundary and category labels of six entities, including disease, imaging examination, laboratory examination, drug, operation, and anatomy from Chinese electronic medical record(EMR) ...

[HTML] [Assessing the practice of data quality evaluation in a national clinical data research network through a systematic scoping review in the era of real-world data](#) J Bian, T Lyu, A Loiacono, ..., 2020 ... Rising from the US Food and Drug Administration (FDA) Real-world Evidence (RWE) program, RWD such as those in the OneFlorida are increasingly important to support a wide range of healthcare and regulatory decisions.

*Any opinions expressed are solely my own.*