

# Methods and computational techniques for predicting adherence to treatment: a scoping review

Beatriz Merino-Barbancho<sup>1\*</sup>, Ana Cipric<sup>2</sup>, Peña Arroyo<sup>1</sup>, Miguel Rujas<sup>1</sup>, Rodrigo Martín Gómez del Moral Herranz<sup>1</sup>, Torben Barev<sup>2</sup>, Nicholas Ciccone<sup>2</sup> and Giuseppe Fico<sup>1</sup>

<sup>1</sup>Universidad Politécnica de Madrid, Life Supporting Technologies Research Group, Madrid, Spain

<sup>2</sup>Novo Nordisk A/S, Vandtårnsvej 108-110, 2860, Søborg, Denmark

\*Corresponding author: [bmerino@lst.tfo.upm.es](mailto:bmerino@lst.tfo.upm.es)

## Abstract

**Background:** Treatment non-adherence of patients stands as a major barrier to effectively manage chronic conditions. Treatment adherence can be described as the extent to which a patient's behavior of taking medications follows the agreed recommendations from the healthcare provider. However, non-adherent behavior is estimated to affect up to 50% of patients with chronic conditions, leading to poorer health outcomes among patients, higher rates of hospitalization, and increased mortality. In fact, 200.000 premature deaths each year in the European Union are related to non-adherence. A promising approach to understand adherence behavior of patients represent artificial intelligence and computational techniques. These techniques can be especially useful in analyzing large amounts of heterogeneous patient data, identifying both inter and intra-relationships between factors and patterns associated with non-adherence.

**Objective:** This study offers a provision of a structured overview of the computational methods and techniques used to build predictive models of treatment adherence of patients.

**Methodology:** A scoping review was conducted, and the following databases were searched to identify relevant publications: PubMed, IEEE and Web of Science. The screening of publications consisted of two steps. First, the hits obtained from the search were independently screened and selected using an open-source machine learning (ML)-aided pipeline applying active learning: ASReview, Active learning for Systematic Reviews. Publications selected for further review were those highly prioritized by ASReview.

**Results:** 45 papers were selected into the second round of screening were reviewers performed the full-text screening. The final review included 29 papers. The findings suggest supervised learning (regression and classification) to be the most used analytical approach. Over 54% of adherence topics being related to chronic metabolic conditions such as diabetes, hypertension, and hyperlipidemia. Most assessed predictors were both treatment and socio-demographic and economic-related factors followed

by condition-related factors. The selection of a particular computational technique was based on the research question, the type of data available and the desired outcome. A limitation of the reviewed studies is the lack of accountancy for interrelationships between different determinants of adherence behavior. Adherence behavior is a complex phenomenon that is influenced by multiple factors, and it is likely that these factors interact with one another in complex ways.

**Conclusion:** The creation of systems to accurately predict treatment adherence can pave the way for improved therapeutic outcomes, reduced healthcare costs and enabling personalized treatment plans. This paper can support to understand the efforts made in the field of modeling adherence-related factors. In particular, the results provide a structured overview of the computational methods and techniques used to build predictive models of treatment adherence of patients in order to guide future advancements in healthcare.

## **Background**

The extent to which patients adhere to medical recommendations is a critical factor for successful disease prevention, management and intervention across a wide range of diseases and conditions (Vrijens et al., 2012). However, suboptimal adherence to recommendations is estimated to affect up to 50% of patients with chronic conditions, which results in 125 billion EUR in avoidable premature deaths, hospitalizations and emergency care visits in both EU and US yearly (Lehmann et al., 2014). Addressing the issues of non-adherence would significantly improve individual patient outcomes and reduce related costs. Although several solutions for non-adherence have been developed, evidence of the effects of these solutions is highly variable (Vervloet et al., 2012; Viswanathan et al., 2012), which might be an artifact of the paucity of generalizable knowledge on determinants of adherence behavior in general. In turn, knowledge about determinants of non-adherent behavior could enable healthcare strategies to target these factors in particular, before preventive or treatment measures failure occurs (Haynes et al., 2002).

The multifaceted nature of adherence behavior made an investigation into determinants of adherence behavior a complex task. There has been a vast of research conducted finding an ambiguous impact of a number of different factors. The ambiguity in findings could be due to a lack of accountancy for interrelationships between different determinants of adherence behavior. Artificial intelligence (AI) bears the potential to provide an understanding of adherence behavior by analyzing large amounts of heterogeneous patient data in identifying both inter and intra-relationships between factors and patterns associated with non-adherence. To understand the efforts made in the field of modeling adherence-related factors and in order to guide future advancements, the aim of this article is to provide an overview of computational techniques which have been employed so far.

## **Study objective and research question**

The main objective of this article is to conduct a scoping review to identify the computational techniques and their features that are implemented in treatment adherence models. Specifically, the objectives were to review the literature to identify features and evidence-based results that have demonstrated relevance to the implementation of these automated models for this purpose. Ultimately, the idea behind these scoping reviews was to have a structured overview of the key concepts in the research area, identify gaps in the existing research and succinctly summaries the key findings. This scoping review is based in the following research question:

- What are the computational methods and techniques used to build predictive models of treatment adherence?

## Materials and methods

### Scoping approach: sources and keywords

To conduct the search on computational methods and techniques employed in predictive modeling of treatment adherence, the preselected information sources were electronic databases of scientific publications, PubMed, IEEE and Web of Science in particular.

A series of keywords were selected and grouped into three groups:

- Theoretical models: Covering current theories and models on adherence behavior (e.g., “transtheoretical model”, “ABC Taxonomy”).
- Computational techniques: Covering the most relevant computational techniques related to artificial intelligence (e.g., “artificial intelligence”, “machine learning”, “classification”).
- Adherence field: Encompassing the most commonly used terms that represent the term adherence to treatment (e.g., “treatment adherence”, “patient compliance”).

### Search strategy

In order to identify relevant publications on computational methods/techniques for predicting adherence to treatment we performed a search across the three databases mentioned in the previous section. The search queries were adapted to each database to ensure comprehensive and right extraction of the literature, using both MeSH Terms and free-text keywords for PubMed, full-text and metadata searches for IEEE Xplore and topic searches for Web of Science. Table 1 includes the specific search queries used for each dataset, categorized into theoretical models, computational techniques and adherence-related terms.

Table 1. Databases search queries

Database	Search Query (per category)
PubMed	("ABC Taxonomy" OR "Andersen Model of Total Patient Delay" OR "ABM" OR "Andersen's Behavioural Model" OR "Behavioural Learning Theory" OR "BLT" OR "COM-B Model" OR "Common Sense Model of Self-Regulation" OR "CSM" OR "Comprehensive Medication Management" OR "CMM" OR "Ecological/Socio-Ecological Framework/Model" OR "Health Action Process Approach" OR "HAPA" OR "Information Motivation Behavioural Skills model of adherence" OR "Integrated Theory of Health Behaviour Change" OR "Integrated-Change Model" OR "I-Change Model" OR "Medication Adherence Model" OR "MAM" OR "Patient Health Engagement model" OR "PHE" OR "Protection Motivation Theory" OR "PMT" OR "Roy Adaptation Model" OR "RAM" OR "Self-Determination Theory" OR "SDT" OR "Self-Efficacy Theory" OR "Self-Management Theory" OR "Self-Regulation Theory" OR "SRT" OR "Situating IMB model of Care Initiation and Maintenance" OR "sIMB-CIM" OR "Subjective Experienced Health Methodology" OR "SEHM" OR "Theoretical Domains Framework" OR "TDF" OR "Theory of Planned Behaviour" OR "TPB" OR "Theory of Reasoned Action" OR "TRA" OR "models, biopsychosocial"[MeSH Terms] OR Biopsychosocial Model[Text Word] OR "health belief model"[MeSH Terms] OR Health Belief Model[Text Word] OR "medication therapy management"[MeSH Terms] OR Medication Therapy Management[Text Word] OR "psychological theory"[MeSH Terms] OR Social Cognitive Theory[Text Word] OR ("interdisciplinary studies"[MeSH Terms] OR Multidisciplinary[Text Word]) AND Team[All Fields]) OR "transtheoretical model"[MeSH Terms])
	("artificial intelligence"[MeSH Terms] OR ai artificial intelligence[Text Word] OR "machine learning"[MeSH Terms] OR transfer learning[Text Word] OR "deep learning"[MeSH Terms] OR deep learning[Text Word] OR "neural networks, computer"[MeSH Terms] OR neural networks[Text Word] OR "decision trees"[MeSH Terms] OR decision tree[Text Word] OR "support vector machine"[MeSH Terms] OR support vector machine[Text Word] OR regression[Text Word] OR "classification"[Subheading] OR "classification"[MeSH Terms] OR classification[Text Word] OR "cluster analysis"[MeSH Terms] OR clustering[Text Word] OR (supervised[All Fields] AND ("learning"[MeSH Terms] OR learning[Text Word])) OR (unsupervised[All Fields] AND ("learning"[MeSH Terms] OR learning[Text Word])) OR "decision support systems, clinical"[MeSH Terms] OR Clinical Decision Support[Text Word])

	("treatment adherence and compliance"[MeSH Terms] OR Treatment Adherence and Compliance[Text Word] OR "patient compliance"[MeSH Terms] OR Patient Compliance[Text Word] OR "patient dropouts"[MeSH Terms] OR Patient Dropouts[Text Word] OR Therapeutic Adherence[Text Word] OR Therapeutic Adherence and Compliance[Text Word] OR Treatment Adherence[Text Word] OR Non-Adherent Patient[Text Word] OR Patient Adherence[Text Word] OR Patient Non-Adherence[Text Word])
IEEE Xplore	<p>("Full Text &amp; Metadata": "ABC Taxonomy" OR "Full Text &amp; Metadata": "Andersen Model of Total Patient Delay" OR "Full Text &amp; Metadata": "ABM" OR "Full Text &amp; Metadata": "Andersen's Behavioural Model" OR "Full Text &amp; Metadata": "Behavioural Learning Theory" OR "Full Text &amp; Metadata": "BLT" OR "Full Text &amp; Metadata": "COM-B Model" OR "Full Text &amp; Metadata": "Common Sense Model of Self-Regulation" OR "Full Text &amp; Metadata": "CSM" OR "Full Text &amp; Metadata": "Comprehensive Medication Management" OR "Full Text &amp; Metadata": "CMM" OR "Full Text &amp; Metadata": "Ecological/Socio-Ecological Framework/Model" OR "Full Text &amp; Metadata": "Health Action Process Approach" OR "Full Text &amp; Metadata": "HAPA" OR "Full Text &amp; Metadata": "Information Motivation Behavioural Skills model of adherence" OR "Full Text &amp; Metadata": "Integrated Theory of Health Behaviour Change" OR "Full Text &amp; Metadata": "Integrated-Change Model" OR "Full Text &amp; Metadata": "I-Change Model" OR "Full Text &amp; Metadata": "Medication Adherence Model" OR "Full Text &amp; Metadata": "MAM" OR "Full Text &amp; Metadata": "Patient Health Engagement model" OR "Full Text &amp; Metadata": "PHE" OR "Full Text &amp; Metadata": "Protection Motivation Theory" OR "Full Text &amp; Metadata": "PMT" OR "Full Text &amp; Metadata": "Roy Adaptation Model" OR "Full Text &amp; Metadata": "RAM" OR "Full Text &amp; Metadata": "Self-Determination Theory" OR "Full Text &amp; Metadata": "SDT" OR "Full Text &amp; Metadata": "Self-Efficacy Theory" OR "Full Text &amp; Metadata": "Self-Management Theory" OR "Full Text &amp; Metadata": "Self-Regulation Theory" OR "Full Text &amp; Metadata": "SRT" OR "Full Text &amp; Metadata": "Situating IMB model of Care Initiation and Maintenance" OR "Full Text &amp; Metadata": "sIMB-CIM" OR "Full Text &amp; Metadata": "Subjective Experienced Health Methodology" OR "Full Text &amp; Metadata": "SEHM" OR "Full Text &amp; Metadata": "Theoretical Domains Framework" OR "Full Text &amp; Metadata": "TDF" OR "Full Text &amp; Metadata": "Theory of Planned Behaviour" OR "Full Text &amp; Metadata": "TPB" OR "Full Text &amp; Metadata": "Theory of Reasoned Action" OR "Full Text &amp; Metadata": "TRA" OR "Full Text &amp; Metadata": "Biopsychosocial model" OR "Full Text &amp; Metadata": "Health Belief Model" OR "Full Text &amp; Metadata": "Medication Therapy Management" OR "Full Text &amp; Metadata": "Social Cognitive Theory" OR "Full Text &amp; Metadata": "Multidisciplinary Team" OR "Full Text &amp; Metadata": "transtheoretical model")</p> <p>("Full Text &amp; Metadata": "artificial intelligence" OR "Full Text &amp; Metadata": "machine learning" OR "Full Text &amp; Metadata": "transfer learning" OR "Full Text &amp; Metadata": "deep learning" OR "Full Text &amp; Metadata": "neural network" OR "Full Text &amp; Metadata": "decision tree" OR "Full Text &amp; Metadata": "support vector machine" OR "Full Text &amp; Metadata": "cluster" OR "Full Text &amp; Metadata": "decision support system" OR ("Full Text &amp; Metadata": "classification NEAR/2 algorithm" OR ("Full Text &amp; Metadata": "supervised AND "Full Text &amp; Metadata": "learning" OR ("Full Text &amp; Metadata": "unsupervised AND "Full Text &amp; Metadata": "learning" OR ("Full Text &amp; Metadata": "techn* NEAR/2 "AI" OR ("Full Text &amp; Metadata": "technique NEAR/2 "ML" OR ("Full Text &amp; Metadata": "technique NEAR/2 "DL"))</p> <p>((("Full Text &amp; Metadata": "Adheren* NEAR/1 "therapeutic" ) OR ("Full Text &amp; Metadata": "Adheren* NEAR/1 "treatment") OR ("Full Text &amp; Metadata": "Patient NEAR/1 Non-Adheren*") OR ("Full Text &amp; Metadata": "Patient NEAR/1 Adheren*") OR ("Full Text &amp; Metadata": "Patient NEAR/1 "Non-Compliance") OR ("Full Text &amp; Metadata": "Patient NEAR/1 "Compliance" ) OR ("Full Text &amp; Metadata": "Patient NEAR/1 dropout"))</p>
Web Of Science	<p>ALL= ("ABC Taxonomy" OR "Andersen Model of Total Patient Delay" OR "ABM" OR "Andersen's Behavioural Model" OR "Behavioural Learning Theory" OR "BLT" OR "COM-B Model" OR "Common Sense Model of Self-Regulation" OR "CSM" OR "Comprehensive Medication Management" OR "CMM" OR "Ecological/Socio-Ecological Framework/Model" OR "Health Action Process Approach" OR "HAPA" OR "Information Motivation Behavioural Skills model of adherence" OR "Integrated Theory of Health Behaviour Change" OR "Integrated-Change Model" OR "I-Change Model" OR "Medication Adherence Model" OR "MAM" OR "Patient Health Engagement model" OR "PHE" OR "Protection Motivation Theory" OR "PMT" OR "Roy Adaptation Model" OR "RAM" OR "Self-Determination Theory" OR "SDT" OR "Self-Efficacy Theory" OR "Self-Management Theory" OR "Self-Regulation Theory" OR "SRT" OR "Situating IMB model of Care Initiation and Maintenance" OR "sIMB-CIM" OR "Subjective Experienced Health Methodology" OR "seem" OR "Theoretical Domains Framework" OR "TDF" OR "Theory of Planned Behaviour" OR "TPB" OR "Theory of Reasoned Action" OR "TRA" OR Biopsychosocial model* OR Health Belief Model OR Medication Therapy Management OR Social Cognitive Theory OR (Multidisciplinary AND Team) OR transtheoretical model*)</p> <p>TS= ("artificial intelligence" OR "machine learning" OR "transfer learning" OR "deep learning" OR neural network* OR decision tree* OR support vector machine OR (classification AND algorithm*) OR cluster* OR (supervised AND learning) OR (unsupervised AND learning) OR decision support system* OR ((AI OR ML OR DL) AND techn* ))</p> <p>TS= ((Adheren* AND ("treatment" OR "therapeutic")) OR (Patient* AND (Non-Adheren* OR Adheren* OR "Non-Compliance" OR "Compliance" OR "dropout*")))</p>

## Eligibility criteria

The eligibility criteria have been defined using JBI methods for a scoping review (Peters et al.,2020) in order to guide the selection process. The criteria were divided into the following areas:

- Population: Studies considering adult human subjects ( $\geq 16$  years old).
- Concept: Computational methods and techniques for treatment adherence prediction models.

- Timeframe: Within the last 11 years (2012-2023).
- Language: Articles only published in English.

### **Data extraction**

The data was extracted using Microsoft Excel. The search hits (including publication title, authors, abstract and DOI) are downloaded in .csv, .txt or .xlsx format, merged and cleared of duplicates. The title and abstract screening were performed using ASReview software (*ASReview: AI-Aided Open Source Systematic Review Software*, n.d.) and stored in the excel file for the text review. Reviewers, including biomedical engineers, behavioral data scientists, psychologists, pharmacologists and doctors, manually annotated the following publication details: the type of study/model and publication, the objective of the study, the type and the data used, the theoretical model behind, the disease that each article focuses, the type of methodology/technique used, the results obtained, the limitations of the study and proposed future research.

### **First screening: title & abstract**

The hits obtained from the search will be independently screened and selected using ASReview. The tool is initially trained with 10 relevant and 10 irrelevant publications selected by two independent researchers. After feeding the tool with the training publications, the tool returns the set of hits ordered according to relevance priority. These results are checked by an independent researcher. In case of several irrelevant results among the top priority hits, the tool is further trained by manually screening at least 1% of the total number of publications in the whole set. Publications selected for further full-text review were those highly prioritized by ASReview. After this, each resulting abstract is consulted by two independent reviewers who must decide whether or not to include the article for the full text review. Discrepancies between them were analyzed by a third independent reviewer who acted as a referee.

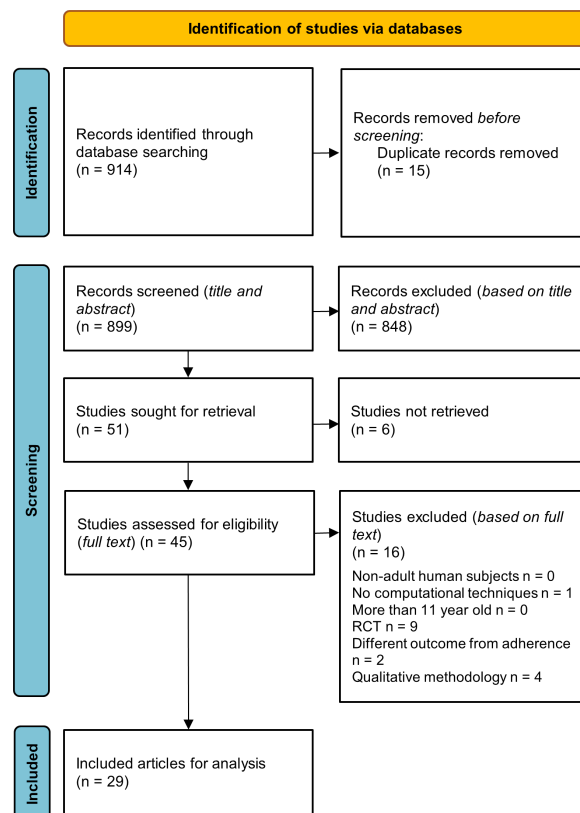
### **Second screening: full text**

Each reviewer performed the full-text screening of the assigned publications considering the eligibility criteria described above. For each publication assigned, the reviewer checked each eligibility criteria and consider for further inclusion only those publications meeting all the criteria. For the selected publications, reviewers will annotate additional publication details as requested on the spreadsheet. After the full-text screening, the final list of selected publications was ready for data extraction.

### **Outcomes**

The final list of included studies for the ScR (Scoping Review) was divided among several experts who worked in parallel to extract the relevant data from the assigned publications. To facilitate and homogenize the process of data extraction, a structured data collection sheet was shared among reviewers, in which the elements considered of interest in this scoping review were included. The articles for which data was extracted were articles that mentioned the architecture and computational techniques employed in the development of treatment adherence prediction models. Factors extracted from these articles were: medical condition, a measure of adherence as an outcome variable, sample size, adherence-related factors (e.g., socio-demographic, healthcare system-related, condition-related, treatment-related, and patient-related), computational technique and computational algorithm. Socio-demographic and economic factors refer to characteristics such as age, gender, income and education level. Healthcare system-related factors refer to descriptors of the healthcare environment, such as HC visits and interactions. Condition-related factors refer to aspects of the patient's health condition, such as symptom severity, which may impact adherence. Treatment-related factors refer to aspects of the treatment regimen such as complexity, dosing schedule, and side effects, which may impact adherence. Patient-related concern inter and intra-personal characteristics of the patient. Computational algorithms used were classified as clustering, classification and regression, depending on whether they were aimed at detecting association, classification or clusterisation.

Figure 1. Flow Diagram





## Materials

This study utilized a multifaceted approach integrating various technological tools. First, we employed the BEAMER project (*BEAMER HOME - BEAMER*, n.d.), funded through the Innovative Medicines Initiative (IMI2) and now under the Innovative Health Initiative (IHI), is a European research project (Grant agreement number: 101034369) dedicated to developing a behavioral and adherence model leveraging artificial intelligence (AI) techniques. BEAMER's primary aim is to enhance the quality, outcomes, and cost-effectiveness of healthcare services. Secondly, we incorporated the Active Learning for Systematic Reviews (ASReview) tool. ASReview is an open-source machine learning (ML)-aided pipeline that applies active learning methodologies. It significantly improves the efficiency of screening academic titles and abstracts by strategically prioritizing them through active learning algorithms. Lastly, Microsoft Excel, a widely recognized software program created by Microsoft, was utilized. Excel's powerful spreadsheet capabilities, which include organizing numbers and data with formulas and functions, were crucial in data analysis and management phases of the research.

## Results

Of 45 papers selected into the second round of screening, the final review included 29 papers (Figure 1)(Table 2). A total of twelve papers were excluded due to being randomized control trials (9), employing qualitative methodology (4), being a literature review of adherence factors without reference to analytical techniques (1) and/or for having different outcome variables in focus (2; Figure 1). With respect to analytical techniques, three major approaches were identified; generalized linear models accounted for 21,67% of employed techniques, followed by the family of logistic regressions (20,00%) and random forest techniques (18,33%; Table 3). The findings suggest supervised learning (regression and classification) to be the most used analytical approaches. The family of generalized linear models employed included multiple, multilevel, hierarchical, fixed-effects, OLS, mixed-effects and GEE (Table 3 and Table 4). Whether the selected algorithm was a regression or classification, was primarily determined by the data source and the scaling of the outcome and predictors measures (Table 2). The same applied to a selection of logistic regression or random forest techniques.

Table 2. Description of Articles \*

Authors, Year	Condition	Adherence (Outcome Variable)	Sample Size	Main Computational Technique	Computational Algorithm	ADHERENCE FACTORS: Socio-demographic	Healthcare System	Condition-related	Treatment-related	Patient-related
(Shiyanbola et al., 2018)	Type 2 Diabetes	Categorical (Self-reported) / The Morisky Medication Adherence Scale (MMAS-8)	174	Cluster Analysis	Clustering	+		+	+	+



(Colvin et al., 2018)	Diabetes, Hypertension, or Hyperlipidemia	Continuous / Pharmacy claims data - number of medication days	56	Linear Regression	Regression	+	+	+	+	+
(Fortuna et al., 2018)	Hypertension	Categorical (Self-reported) / The Morisky Medication Adherence Scale (MMAS-8)	2128	Logistic Regression	Classification	+			+	
(Wang et al., 2020)	Crohn's Disease	Categorical (Self-reported) / The Medication Adherence Report Scale (MARS)	446	Random Forest, Back-propagation Neural Network, Support Vector Machine, Logistic Regression	Classification, Regression	+	+	+	+	+
(Wawruch et al., 2019)	Peripheral Arterial Disease	Dichotomous (Adherent/Non-adherent) / 6-month period without statin prescription	8330	Cox Regression	Regression	+		+	+	
(Whittle et al., 2016)	Hypertension, Hyperlipidemia	Dichotomous (Adherent/Non-adherent) / <80 % of medication refilled	31250	Logistic Regression, Linear Regression (multivariate)	Classification, Regression	+				
(Wong et al., 2012)	Diabetes	Dichotomous (Adherent/Non-adherent) / <80 % of medication refilled	444418	Logistic Regression (hierarchical, multilevel)	Classification	+	+	+	+	+
(Wu et al., 2020)	Type 2 Diabetes	Dichotomous (Adherent/Non-adherent) / <80 % of medication refilled	401	Logistic Regression, C 5.0 model, Decision List, Bayesian Network, Discriminant model, KNN algorithm, LSVM, Random Forest, SVM, Tree-AS, CHAID, Quest, C&R Tree, Neural Net, and the Ensemble model	Classification	+		+	+	+
(Koulayev et al., 2017)	Hypertension, Diabetes, Dyslipidemia	Continuous / Number of refills ('medication possession ratio' (MPR))	15916	OLS Regression (fixed-effects)	Regression	+	+	+	+	
(Kozma et al., 2014)	Multiple sclerosis	Continuous / Number of refills	4606	Linear Regression (multiple)	Regression	+		+	+	
(Alatawi et al., 2016)	Type 2 Diabetes	Categorical (Self-reported) / three different adherence measures	222	Linear Regression (multiple)	Regression	+	+	+	+	+
(Berry et al., 2015)	Cancer	Categorical Dichotomized (Self-reported) / The Morisky Medication Adherence Scale (MMAS-8) - categorized	752	Logistic Regression (univariate), Classification Tree	Classification	+	+	+	+	
(Boland et al., 2016)	Chronic Obstructive Pulmonary Disease	Dichotomous (Adherent/Non-adherent) / <80 % of medication	511	Linear Regression (mixed-effect)	Regression	+		+	+	+

		refilled								
(Campbell et al., 2020)	Hypertension	Dichotomous (Adherent/Non-adherent) / <80 % of medication refilled	4842	Generalized Linear Models (with log link, gamma or negative binomial distribution)	Regression	+	+	+	+	
(Horberg et al., 2012)	HIV-positive	Categorical / prescription based treatment regimens	10801	Regression Tree, Linear Regression (mixed-effect), Logistic Regression (mixed-effect)	Regression, Classification	+	+	+	+	
(Javanmard ifard et al., 2020)	Type 2 Diabetes	Categorical (Self-reported) / The Morisky Medication Adherence Scale (MMAS-8)	227	Correlations	Description	+	+		+	
(Sleath et al., 2015)	Glaucoma	Continuous (Self-reported)	279	Generalizing Estimating Equations	Regression	+	+		+	+
(Tibble et al., 2020)	Asthma	Continuous / Doses taken (Percentage)	220	Principal Component Analysis, K-Means, Decision Trees	Classification, Regression, Clustering			+	+	
(Tiv et al., 2012)	Type 2 Diabetes	Categorical (Self-reported) / Six-items questionnaire - 'good', 'medium' and 'poor' (treated as nominal)	3637	Logistic Regression (multinomial polychotomous)	Classification	+	+	+	+	
(Verma et al., 2018)	Hypertension	Dichotomous (Adherent/Non-adherent Medication discontinuation)	13350	Cox proportional hazards regression (propensity score matched sample)	Classification, Regression				+	
(Hsu et al., 2022)	Cardiovascular disease	Dichotomous, Adherent/Non-adherent // Proportion of days covered (PDC) $\geq 80$ is used as the threshold for adherence	100096	LSTM, Simple RNN, MLP, Ridge Classifier, Logistic Regression	Regression, Classification	+	+	+	+	
(Senner et al., 2023)	Schizophrenia -spectrum and bipolar disorder	Categorical / The medication adherence was evaluated with a self-rating questionnaire with assemblance to BARS	862	Linear Regression (multiple), Cluster analysis (hierarchical)	Regression, Clustering	+		+	+	
(Innab et al., 2023)	Hypertension	Continuous / Correlations between adherence and some factors	114	Pearson correlation, Linear regression (hierarchical)	Regression	+		+	+	+
(Luo et al., 2023)	Breast Cancer	Dichotomous, Poor/Good functional exercise compliance // Based on the program of their hospitals	227	Classification Tree (CHAID)	Classification	+		+		+
(Carmody et al., 2023)	Sleep disordered breathing (SDB)	Continuous / Relationships between PAP frequency/duration and some factors	188	Linear Regression (multiple)	Regression	+		+	+	+
(Hu et al., 2022)	Breast Cancer	Dichotomous, Adherent/Non-	559	Linear Regression	Regression	+		+	+	+

		adherent / Adherence as the proportion of days covered (PDC) by AET		(multivariate)						
(Malo et al., 2022)	Cardiovascular disease	Continuous / Correlations between adherence and some factors	15332	Logistic Regression (multinomial)	Regression	+		+	+	
(Li et al., 2022)	Type 2 Diabetics	Dichotomous, Adherent/Non-adherent / Proportion of days covered (PDC) $\geq 80$ is used as the threshold for adherence	980	ANOVA	Regression	+		+	+	+
(Browne et al., 2023)	Prevention of Immunodeficiency Virus Infection	Dichotomous, Adherent/Non-adherent, Continuous / Taking and timing	63	Logistic Regression (mixed-effects)	Regression	+		+	+	+

\* Author, Year: Authors and year of publication for the study, Condition: disease under investigation, Adherence (outcome variable): feature measuring adherence, Sample Size: size of the sample in the study, Main Computational Technique: primary algorithm used in the analysis, ADHERENCE FACTORS: factors considered in the study, categorized into socio-demographic, healthcare system, condition-related, treatment-related, and patient-related, + means the presence of associated factors.

Table 3. Description of the Computational Techniques

Computational Technique	Number of Times Used	Percentage
1. Generalized Linear Model (multiple, multivariate, hierarchical, fixed-effects, OLS, mixed-effects, GEE)	14	21,67%
2. Logistic Regression (multilevel, univariate, multiple, hierarchical, mixed-effects, multinomial, polychotomous)	12	20,00%
3. Random Forest (decision list, decision tree, classification tree, CHAID, ensemble, C5.0)	11	18,33%
4. Neural Network (NN, BPNN, LSTM, Simple RNN, MLP)	5	8,33%
5. Cluster analysis	4	6,67%
6. Support Vector Machine (SVM and LSVM)	3	5,00%
7. Bayesian Network	3	5,00%
8. Correlational	3	5,00%
9. Cox Regression	2	3,33%
10. Principal Component Analysis	1	1,67%
11. K-nearest Neighbors (KNN)	1	1,67%
12. Discriminant Model	1	1,67%
13. Ridge classifier (RC)	1	1,67%
TOTAL	60	100,00%

Table 4 shows a distribution of treatments of interest in relation to adherence, with over 54% of adherence topics being related to chronic metabolic conditions such as diabetes, hypertension, and hyperlipidemia. The other conditions and equally distributed among each other, were HIV, Breast Cancer, Peripheral Arterial Disease (PAD), Chronic Obstructive Pulmonary Disease (COPD), Cancer (general), Asthma, Glaucoma, Multiple sclerosis, some mental disorders, Crohn's disease, some other cardiovascular diseases and Sleep disordered breathing (SDB). Most assessed predictors were both treatment and socio-demographic and economic-related factors, followed by condition-related factors. Somewhat less inspected are factors related to the healthcare system and patient-related individual

differences (see Table 1). The outcome adherence to the treatment variable was uniformly specified as dichotomous (12), categorical (9) and continuous (8) of the analyzed articles (Table 1). An example of categorical indicator for adherence was mostly The Morisky Medication Adherence Scale (MMAS-8), a self-reported measure of adherence. Some continuous and dichotomous indicators were based on the prescriptions regimens and the number of pharmacy claims. Dichotomous indicators classified adherence when a proportion of days covered by prescription/pharmacy claims equalled or exceeded 80%. MPR (medical possession ratio) and PDC (proportion of days covered by prescription) are almost identical and equally common formula-based calculations of adherence (and the 80% threshold).

Table 4. Adherence Models inspected related to the Conditions

Condition	Number of times researched	Percentage of times researched
Diabetes	9	25,71%
Hypertension	7	20,00%
Hyperlipidaemia/Dyslipidaemia	3	8,57%
HIV +	2	5,71%
Breast Cancer	2	5,71%
Crohn's disease	1	2,86%
Peripheral arterial disease (PAD)	1	2,86%
Chronic Obstructive Pulmonary Disease (COPD)	1	2,86%
Cancer (general)	1	2,86%
Asthma	1	2,86%
Glaucoma	1	2,86%
Multiple sclerosis	1	2,86%
Schizophrenia-spectrum disorder	1	2,86%
Bipolar disorder	1	2,86%
Primary cardiovascular disease	1	2,86%
Cardiovascular disease (general)	1	2,86%
Sleep disordered breathing (SBD)	1	2,86%
Total	35	100,00%

## Discussion

Based on the results of the review, to this date different computational techniques have been employed to model treatment adherence. The review included 29 papers of which over three quarters employed generalized linear models, logistic regressions and random forest techniques. Supervised learning (regression and classification) was the most used analytical approach. The review also revealed that chronic metabolic conditions such as diabetes, hypertension and hyperlipidaemia were the most common conditions for which adherence was modelled. The predictors assessed in the studies were treatment-related factors, socio-demographic and economic factors and condition-related factors, with

healthcare system and patient-related individual differences being somewhat less inspected.

The selection of a particular computational technique was based on the research question, the type of data available and the desired outcome. For example, regression techniques such as linear regression or logistic regression were commonly used when the outcome variable was continuous or dichotomous, respectively. Random forest techniques, on the other hand, were used when dealing with large, complex datasets that may contain many predictor variables. Indeed, the predictors assessed in the studies included several different factors.

However, one limitation of the reviewed studies is the lack of accountancy for interrelationships between different determinants of adherence behavior. Adherence behavior is a complex phenomenon that is influenced by multiple factors and it is likely that these factors interact with one another in complex ways. For example, a patient's socio-economic status may affect their ability to afford medication, which in turn may impact their adherence. Additionally, a patient's health condition may impact their ability to adhere to certain aspects of the treatment regimen, such as exercise or dietary restrictions. To address this limitation, future research could consider more comprehensive models that consider the interrelationships between different determinants of adherence behavior. This could involve the use of more advanced analytical techniques such as network analysis or systems modelling. Additionally, researchers could explore the use of machine learning algorithms that are better able to capture complex interactions between multiple predictors. This could ultimately lead to more effective interventions that are better tailored to individual patients' needs and circumstances. The results suggest that computational techniques have been useful in modelling treatment adherence for a variety of conditions. However, the review also highlights the need for further research to address the limitations and challenges in the current approaches and to refine and validate computational models of treatment adherence.

## **Conclusion**

Treatment adherence behaviour is a complex phenomenon that is influenced by multiple factors. This is critical, as treatment non-adherence stands as a major barrier to effectively manage chronic conditions. A promising approach to understand adherence behaviour of patients in detail and analyse large amounts of heterogeneous patient data represent artificial intelligence and computational techniques. These techniques can be especially fruitful in identifying both inter and intra-relationships between factors and patterns associated with non-adherence.

This article sheds light of the computational methods and techniques used to build predictive models of treatment adherence. In particular, the results of the conducted scoping review provide a structured

overview of the key concepts in the research area is provided, gaps in the existing research were identified and the key findings succinctly summarised.

These results of this paper contribute to understand the efforts made in the field of modelling treatment adherence-related factors and to guide future advancements. An advancement of knowledge of treatment adherence behaviour can make healthcare more efficient, can contribute to deliver high-value personalised care, increase the treatment adherence of patients, and result in a significant decrease in healthcare costs.

### **Limitations and Future Research**

To cover the dominant literature of peer-reviewed published papers in the selected field, the scoping review was limited to three electronic databases of scientific publications: PubMed, IEEE, and Web of Science. Only articles in English language have been included in the analysis. It must be noted that further studies can investigate additional data databases, data sources and articles published in different languages for their analysis. For structuring the results, different grouping and ranking strategies might have also added additional insights but were not attempted since they are outside the scope of this paper.

## References

- Alatawi, Y. M., Kavookjian, J., Ekong, G., & Alrayees, M. M. (2016). The association between health beliefs and medication adherence among patients with type 2 diabetes. *Research in Social & Administrative Pharmacy : RSAP*, 12(6), 914–925. <https://doi.org/10.1016/J.SAPHARM.2015.11.006>
- ASReview: AI-aided Open Source Systematic Review Software. (n.d.). Retrieved December 21, 2023, from <https://asreview.nl/>
- \_BEAMER HOME - BEAMER. (n.d.). Retrieved December 21, 2023, from <https://beamerproject.eu/>
- Berry, D. L., Blonquist, T. M., Hong, F., Halpenny, B., & Partridge, A. H. (2015). Self-reported adherence to oral cancer therapy: relationships with symptom distress, depression, and personal characteristics. *Patient Preference and Adherence*, 9, 1587–1592. <https://doi.org/10.2147/PPA.S91534>
- Boland, M. R. S., Van Boven, J. F. M., Kruis, A. L., Chavannes, N. H., Van Der Molen, T., Goossens, L. M. A., & Rutten-Van Mülken, M. P. M. H. (2016). Investigating the association between medication adherence and health-related quality of life in COPD: Methodological challenges when using a proxy measure of adherence. *Respiratory Medicine*, 110, 34–45. <https://doi.org/10.1016/J.RMED.2015.11.008>
- Browne, S. H., Vaida, F., Umlauf, A., Tucker, A. J., Blaschke, T. F., & Benson, C. A. (2023). Supporting the Art: Medication Adherence Patterns in Persons Prescribed Ingestible Sensor-enabled Oral Pre-Exposure Prophylaxis to Prevent Human Immunodeficiency Virus Infection. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, 76(1), 134–143. <https://doi.org/10.1093/CID/CIAC280>
- Campbell, P. J., Axon, D. R., Taylor, A. M., Pickering, M., Black, H., Warholak, T., & Chinthammit, C. (2020). Associations of Renin-Angiotensin System Antagonist Medication Adherence and Economic Outcomes Among Commercially Insured US Adults: A Retrospective Cohort Study. *Journal of the American Heart Association*, 9(17). <https://doi.org/10.1161/JAHA.119.016094>



Carmody, J. K., Duraccio, K. M. R., Krietsch, K. N., Simmons, D. M., & Byars, K. C. (2023). Factors affecting pediatric adherence to positive airway pressure: Patient- and caregiver-reported treatment barriers and sleep difficulties. *Sleep Medicine*, 101, 58–65. <https://doi.org/10.1016/J.SLEEP.2022.10.011>

Chapter 11: Scoping reviews. (2020). *JBIMES-20-12*. JBI Manual for Evidence Synthesis. <https://doi.org/10.46658/JBIMES-20-12>

Colvin, N. N., Mospan, C. M., Buxton, J. A., Waggett, J. “Davie,” & Gillette, C. (2018). Using Indian Health Service (IHS) counseling techniques in an independent community pharmacy to improve adherence rates among patients with diabetes, hypertension, or hyperlipidemia. *Journal of the American Pharmacists Association: JAPhA*, 58(4S), S59-S63.e2. <https://doi.org/10.1016/J.JAPH.2018.04.024>

Fortuna, R. J., Nagel, A. K., Rocco, T. A., Legette-Sobers, S., & Quigley, D. D. (2018). Patient Experience With Care and Its Association With Adherence to Hypertension Medications. *American Journal of Hypertension*, 31(3), 340–345. <https://doi.org/10.1093/AJH/HPX200>

Haynes, R. B., McDonald, H. P., & Garg, A. X. (2002). Helping patients follow prescribed treatment: clinical applications. *JAMA*, 288(22), 2880–2883. <https://doi.org/10.1001/JAMA.288.22.2880>

Horberg, M. A., Hurley, L. B., Towner, W. J., Allerton, M. W., Tang, B. T., Catz, S. L., Silverberg, M. J., & Quesenberry, C. P. (2012). Determination of optimized multidisciplinary care team for maximal antiretroviral therapy adherence. *Journal of Acquired Immune Deficiency Syndromes (1999)*, 60(2), 183–190. <https://doi.org/10.1097/QAI.0B013E31824BD605>

Hsu, W., Warren, J. R., & Riddle, P. J. (2022). Medication adherence prediction through temporal modelling in cardiovascular disease management. *BMC Medical Informatics and Decision Making*, 22(1), 1–21. <https://doi.org/10.1186/S12911-022-02052-9/TABLES/13>

Hu, X., Walker, M. S., Stepanski, E., Kaplan, C. M., Martin, M. Y., Vidal, G. A., Schwartzberg, L. S., & Graetz, I. (2022). Racial Differences in Patient-Reported Symptoms and Adherence to Adjuvant Endocrine Therapy Among Women With Early-Stage, Hormone Receptor-Positive Breast Cancer. *JAMA Network Open*, 5(8), E2225485. <https://doi.org/10.1001/JAMANETWORKOPEN.2022.25485>

- Innab, A., Kerari, A., Alqahtani, N., Albloushi, M., & Alshammari, A. (2023). Patient activation, adherence to hypertension treatment plans and blood pressure control in Saudi Arabia: a cross-sectional study. *BMJ Open*, 13(1). <https://doi.org/10.1136/BMJOPEN-2022-067862>
- Javanmardifard, S., Heidari, S., Sanjari, M., Yazdanmehr, M., & Shirazi, F. (2020). The relationship between spiritual well-being and hope, and adherence to treatment regimen in patients with diabetes. *Journal of Diabetes and Metabolic Disorders*, 19(2), 941–950. <https://doi.org/10.1007/S40200-020-00586-1>
- Koulayev, S., Simeonova, E., & Skipper, N. (2017). Can Physicians Affect Patient Adherence With Medication? *Health Economics*, 26(6), 779–794. <https://doi.org/10.1002/HEC.3357>
- Kozma, C. M., Phillips, A. L., & Meletiche, D. M. (2014). Use of an early disease-modifying drug adherence measure to predict future adherence in patients with multiple sclerosis. *Journal of Managed Care & Specialty Pharmacy*, 20(8), 800–807. <https://doi.org/10.18553/JMCP.2014.20.8.800>
- Lehmann, A., Aslani, P., Ahmed, R., Celio, J., Gauchet, A., Bedouch, P., Bugnon, O., Allenet, B., & Schneider, M. P. (2014). Assessing medication adherence: options to consider. *International Journal of Clinical Pharmacy*, 36(1), 55–69. <https://doi.org/10.1007/S11096-013-9865-X>
- Li, M., Lu, X., Yang, H. B., Yuan, R., Yang, Y., Tong, R., & Wu, X. (2022). Development and assessment of novel machine learning models to predict medication non-adherence risks in type 2 diabetics. *Frontiers in Public Health*, 10. <https://doi.org/10.3389/FPUBH.2022.1000622>
- Luo, Z., Luo, B., Wang, P., Wu, J., Chen, C., Guo, Z., & Wang, Y. (2023). Predictive Model of Functional Exercise Compliance of Patients with Breast Cancer Based on Decision Tree. *International Journal of Women’s Health*, 15, 397–410. <https://doi.org/10.2147/IJWH.S386405>
- Malo, S., Maldonado, L., Rabanaque, M. J., Gimeno-Miguel, A., Castel-Feced, S., Lallana, M. J., & Aguilar-Palacio, I. (2022). Patterns of statin adherence in primary cardiovascular disease prevention during the pandemic. *Frontiers in Pharmacology*, 13. <https://doi.org/10.3389/FPHAR.2022.980391>
- Senner, F., Hiendl, L., Bengesser, S., Adorjan, K., Angheliescu, I. G., Baune, B. T., Budde, M., Dannlowksi, U., Dietrich, D. E., Falkai, P., Fallgatter, A. J., Hasan, A., Heilbronner, M., Jäger, M., Juckel, G., Kalman, J. L., Konrad, C., Kohshour, M. O., Papiol, S., ... Dalkner, N. (2023).

Medication adherence and cognitive performance in schizophrenia-spectrum and bipolar disorder: results from the PsyCourse Study. *Translational Psychiatry*, 13(1). <https://doi.org/10.1038/S41398-023-02373-X>

Shiyanbola, O. O., Unni, E., Huang, Y. M., & Lanier, C. (2018). Using the extended self-regulatory model to characterise diabetes medication adherence: a cross-sectional study. *BMJ Open*, 8(11). <https://doi.org/10.1136/BMJOPEN-2018-022803>

Sleath, B., Carpenter, D. M., Blalock, S. J., Sayner, R., Muir, K. W., Slota, C., Giangiacomo, A. L., Hartnett, M. E., Tudor, G., & Robin, A. L. (2015). Applying the resources and supports in self-management framework to examine ophthalmologist-patient communication and glaucoma medication adherence. *Health Education Research*, 30(5), 693–705. <https://doi.org/10.1093/HER/CYV034>

Tibble, H., Chan, A., Mitchell, E. A., Horne, E., Doudesis, D., Horne, R., Mizani, M. A., Sheikh, A., & Tsanas, A. (2020). A data-driven typology of asthma medication adherence using cluster analysis. *Scientific Reports*, 10(1). <https://doi.org/10.1038/S41598-020-72060-0>

Tiv, M., Viel, J. F., Mauny, F., Eschwège, E., Weill, A., Fournier, C., Fagot-Campagna, A., & Penfornis, A. (2012). Medication adherence in type 2 diabetes: the ENTRED study 2007, a French Population-Based Study. *PloS One*, 7(3). <https://doi.org/10.1371/JOURNAL.PONE.0032412>

Verma, A. A., Khuu, W., Tadrous, M., Gomes, T., & Mamdani, M. M. (2018). Fixed-dose combination antihypertensive medications, adherence, and clinical outcomes: A population-based retrospective cohort study. *PLoS Medicine*, 15(6). <https://doi.org/10.1371/JOURNAL.PMED.1002584>

Vervloet, M., Linn, A. J., van Weert, J. C. M., de Bakker, D. H., Bouvy, M. L., & van Dijk, L. (2012). The effectiveness of interventions using electronic reminders to improve adherence to chronic medication: a systematic review of the literature. *Journal of the American Medical Informatics Association : JAMIA*, 19(5), 696–704. <https://doi.org/10.1136/AMIAJNL-2011-000748>

Viswanathan, M., Golin, C. E., Jones, C. D., Ashok, M., Blalock, S. J., Wines, R. C. M., Coker-Schwimmer, E. J. L., Rosen, D. L., Sista, P., & Lohr, K. N. (2012). Interventions to improve adherence to self-administered medications for chronic diseases in the United States: a systematic

review. *Annals of Internal Medicine*, 157(11), 785–795. <https://doi.org/10.7326/0003-4819-157-11-201212040-00538>

Vrijens, B., De Geest, S., Hughes, D. A., Przemyslaw, K., Demonceau, J., Ruppert, T., Dobbels, F., Fargher, E., Morrison, V., Lewek, P., Matyjaszczyk, M., Mshelia, C., Clyne, W., Aronson, J. K., & Urquhart, J. (2012). A new taxonomy for describing and defining adherence to medications. *British Journal of Clinical Pharmacology*, 73(5), 691–705. <https://doi.org/10.1111/J.1365-2125.2012.04167.X>

Wang, L., Fan, R., Zhang, C., Hong, L., Zhang, T., Chen, Y., Liu, K., Wang, Z., & Zhong, J. (2020). Applying Machine Learning Models to Predict Medication Nonadherence in Crohn's Disease Maintenance Therapy. *Patient Preference and Adherence*, 14, 917–926. <https://doi.org/10.2147/PPA.S253732>

Wawruch, M., Wimmer, G., Murin, J., Paduchova, M., Tesar, T., Hlinkova, L., Slavkovsky, P., Fabryova, L., & Aarnio, E. (2019). Patient-Associated Characteristics Influencing the Risk for Non-Persistence with Statins in Older Patients with Peripheral Arterial Disease. *Drugs & Aging*, 36(9), 863–873. <https://doi.org/10.1007/S40266-019-00689-2>

Whittle, J., Yamal, J. M., Williamson, J. D., Ford, C. E., Probstfield, J. L., Beard, B. L., Marginean, H., Hamilton, B. P., Suhan, P. S., & Davis, B. R. (2016). Clinical and demographic correlates of medication and visit adherence in a large randomized controlled trial. *BMC Health Services Research*, 16(1). <https://doi.org/10.1186/S12913-016-1471-X>

Wong, E. S., Piette, J. D., Liu, C. F., Perkins, M., MacIejewski, M. L., Jackson, G. L., Blough, D. K., Fihn, S. D., Au, D. H., & Bryson, C. L. (2012). Measures of adherence to oral hypoglycemic agents at the primary care clinic level: the role of risk adjustment. *Medical Care*, 50(7), 591–598. <https://doi.org/10.1097/MLR.0B013E318249CB74>

Wu, X. W., Yang, H. B., Yuan, R., Long, E. W., & Tong, R. S. (2020). Predictive models of medication non-adherence risks of patients with T2D based on multiple machine learning algorithms. *BMJ Open Diabetes Research & Care*, 8(1). <https://doi.org/10.1136/BMJDR-2019-001055>

## Acknowledgment

The project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 101034369. This joint undertaking receives support from the European Union's Horizon 2020 Research and Innovation Programme, the European Federation of Pharmaceutical Industries and Associations (EFPIA) and Link2Trials. This communication reflects the views of the authors and neither the IMI nor the European Union, EFPIA, or Link2Trials are liable for any use that may be made of the information contained herein. We would like to thank the consortium partners who actively participated in the scouting of articles in phases 1 and 2 and made this study possible.