Implementation of patient-reported outcome measures in real-world evidence studies

Maruszczyk, Konrad; Aiyegbusi, Olalekan Lee; Roth Cardoso, Victor; Gkoutos, Georgios; Slater, Luke; Collis, Philip ; Keeley, Thomas; Calvert, Melanie

DOI:
10.1016/j.cct.2022.106882

Document Version
Publisher's PDF, also known as Version of record

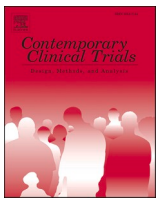
Citation for published version (Harvard):
Maruszczyk, K, Aiyegbusi, OL, Roth Cardoso, V, Gkoutos, G, Slater, L, Collis, P, Keeley, T & Calvert, M 2022, 'Implementation of patient-reported outcome measures in real-world evidence studies: analysis of ClinicalTrials.gov records (1999-2021)', *Contemporary Clinical Trials*, vol. 120, 106882. https://doi.org/10.1016/j.cct.2022.106882

Contemporary Clinical Trials 120 (2022) 106882

Contents lists available at ScienceDirect

Contemporary Clinical Trials

journal homepage: www.elsevier.com/locate/conclintrial



Implementation of patient-reported outcome measures in real-world evidence studies: Analysis of ClinicalTrials.gov records (1999–2021)

Konrad Maruszczyk [a,b], Olalekan Lee Aiyegbusi [a,b,c,d,e,*], Victor Roth Cardoso [f], Georgios V. Gkoutos [f], Luke T. Slater [f], Philip Collis [a,e,g], Thomas Keeley [a,h], Melanie J. Calvert [a,b,c,d,e]

[a] Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, Birmingham, UK
[b] Birmingham Health Partners Centre for Regulatory Science and Innovation, University of Birmingham, Birmingham, UK
[c] National Institute for Health and Care Research (NIHR) Birmingham Biomedical Research Centre, University of Birmingham, Birmingham, UK
[d] NIHR Surgical Reconstruction and Microbiology Research Centre, University of Birmingham, Birmingham, UK
[e] NIHR Applied Research Collaboration (ARC) West Midlands, University of Birmingham, Birmingham, UK
[f] Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, UK
[g] Patient partner, UK
[h] GlaxoSmithKline (GSK), Patient Centered Outcome, Value Evidence and Outcomes, Brentford, UK

ARTICLE INFO

Keywords:
Clinical trial
RWE
PRO
PROM
Phase IV
RWD

ABSTRACT

Background: Real-world evidence (RWE) plays an increasingly important role within global regulatory and reimbursement processes. RWE generation can be enhanced by collecting and using patient-reported outcomes (PROs), which can provide valuable information on the effectiveness, safety, and tolerability of health interventions from the patient perspective. This analysis aims to examine and summarise the utilisation of patient-reported outcomes measures (PROMs) in real-world studies.

Methods: Descriptions of phase IV trials were downloaded on July 22, 2021 from the Clinicaltrials.gov database since its inception. An automated algorithm was built to detect trials utilising PROMs and composite measures including patient-reported components. Search terms were developed based on the PROQOLID database.

Results: Of 27,976 phase IV clinical trials posted on Clinicaltrials.gov between 1999 and July 2021, 21% and 4% used PROMs and composite measures, respectively. Recent years demonstrated a steady increase in the utilisation of PROMs in phase IV trials.

Conclusions: The use of PROMs in phase IV trials seems to be lower than its use in earlier phases of clinical research. Increased uptake of PROMs in RWE studies can be facilitated in a number of ways including the development of standards for their collection, analysis and use.

1. Introduction

Real-world evidence (RWE) is increasingly used to support regulatory and reimbursement decision-making processes globally [1,2]. The U.S. Food and Drug Administration (FDA) published a framework for Real-World Evidence [3], which was recently supplemented by four real-world data (RWD) draft guidelines on data sources, data standards, and regulatory considerations [4–7]. In the UK, the Medicines & Healthcare products Regulatory Agency (MHRA) recently issued two guideline documents focusing on the utilisation of RWD to support regulatory decisions [8,9]. Moreover, the National Institute for Health and Care Excellence draft real-world evidence framework is currently available for public consultation [10]. The European Medicines Agency (EMA) currently uses RWE for safety monitoring and recently announced that the use of RWE will be established across its spectrum of regulatory use cases by 2025 [11].

Contrary to the highly controlled environment of phase III

Abbreviations: COA, clinical outcome assessments; EMA, The European Medicines Agency; ePRO, electronic patient-reported outcome; FDA, U.S. Food and Drug Administration; MHRA, The Medicines & Healthcare products Regulatory Agency; PRO, patient-reported outcome; PROM, patient-reported outcome measure; PROQOLID, Patient-Reported Outcome and Quality of Life Instruments Database; RWD, real-world data; RWE, real-world evidence; XML, Extensible Markup Language.

* Corresponding author at: Murray Learning Centre, Edgbaston, Birmingham B15 2TT, UK.
E-mail address: o.l.aiyegbusi@bham.ac.uk (O.L. Aiyegbusi).

https://doi.org/10.1016/j.cct.2022.106882
Received 30 May 2022; Received in revised form 28 July 2022; Accepted 10 August 2022
Available online 13 August 2022

Box 1
Lay summary.

One way to assess the impact of medical treatments is to assess the impact they have on patient symptoms and quality of life. Patient symptom and quality of life data are increasingly collected in clinical trials to assess whether treatments are safe and work. Once treatments have received regulatory approval for use it is important to assess longer term patient outcomes. This could include real-world impact on symptoms and quality of life.

Using data from an international registry this research paper investigates the evidence of the use of patient-reported outcomes, such as symptom and quality of life data, to provide real-world evidence of the safety and effectiveness of therapies. The research shows an increase in use over time from 1999 to 2021. However, the research shows that collection of data in this setting is still quite low, suggesting the need to further develop the methods of data collection.

registration trials (i.e. to establish an acceptable benefit/safety profile in order to seek regulatory approval for a precisely defined indication), which are usually characterised by close patient monitoring and artificially high patient compliance, RWE studies are characterised by less constrained inclusion criteria and usually, involve a greater number of diverse participants [12]. By evaluating health interventions among diverse, large, and heterogeneous patient populations, RWE studies provide a better understanding of their real-world effectiveness, safety, and tolerability. RWE can therefore inform regulatory decisions, reimbursement, and health policy-making purposes. RWE can be generated through various study designs by analysing real-world data (RWD). Both prospective and retrospective data collection can be utilised to inform RWE generation. The most common sources of RWD are electronic health records, claims databases, registries, and patient-generated data [13].

Patient-reported outcomes (PROs) are reports of health status or quality of life directly provided by patients, without interpretation by a clinician or anyone else [14]. Therefore, PROs provide a unique and valuable source of information and are usually assessed using patient-reported outcome measures (PROMs) – validated measurement tools mainly in a form of questionnaires. Moreover, composite measures are used, incorporating multiple clinical outcome assessments (COAs), including patient-reported ones. PROMs and composite measures are routinely captured within RCTs, primarily to inform regulatory and reimbursement processes [15]. Recently, global regulators, payers, and policymakers have increasingly recognised that PROs can provide valuable information on the effectiveness, safety, and tolerability of drugs from the patient perspective [15–18]. The Framework for FDA's Real-World Evidence Program [3] highlights the use of PROs in RWE generation by acknowledging that PROs provide unique and valuable information which may complement the evidence obtained using traditional clinician-focused parameters [3]. This increased interest in collecting PROs to enrich RWD can be seen as part of a commitment to strengthen patient-centricity in drug development processes.

Recent developments in health informatics infrastructure allow for the use of electronic PROs (ePROs). PROs are being collected in routine care to facilitate individual-level treatment decisions and to support disease progression monitoring. Despite this, PRO data collection in real-world settings remains limited. It seems that broader adoption of PROs for RWE generation could be facilitated by setting up standards for data collection, analysis, and use. In a commentary article, Calvert and colleagues [15] pointed out several priorities. Addressing them would make it possible to fully benefit from the use of PROs in RWE generation. They also called for efforts to advance the understanding of successful PRO implementation in RWE studies. The lack of international guidelines to facilitate the use of PROs in RWE studies was highlighted by a recent systematic review [19].

To inform the development of best practice guidance for PRO data utilisation in RWE generation, it is crucial to understand better how PROMs are currently being used and how this has evolved over time. One possible approach to determine PROMs utilisation is to scrutinise trial registers available in the public domain. Most of the journals require authors to register their studies in publicly available databases prior to the publication of study results. Clinicltrials.gov is a commonly used database for the registration of trials. This database was previously used to assess PROMs utilisation at two time periods: 2004–2007 and 2007–2013 by Scoggins and Patrick [20], and Vodicka et al. [21], respectively. Both studies investigated the use of PROMs in all registered clinical trials, but the latter one focused on utilisation of PROMs in oncological trials. Both forementioned studies are now outdated, and they did not explicitly focus on RWE studies. Thus, there is a need to conduct an up-to-date analysis of Clinicaltrials.gov records, focusing on RWE studies. This will provide an understanding of the current picture of PROMs' use and support future endeavours to facilitate the broader implementation of PROMs in RWE generation.

The research objectives were: (1) quantify the usage of PROMs and composite measures in RWE studies (phase IV trials), (2) describe their utilisation patterns over time and (3) investigate the use of PROMs and composite measures across different disease areas in phase IV trials. An automated searching algorithm was used to identify phase IV studies registered in ClinicalTrial.gov, which report PROMs and composite measures.

2. Materials and methods

The methodology adopted by this study built on the previous analyses of the Clinicaltrials.gov database by Vodicka et al. [21]. Nevertheless, the searching algorithm was developed de novo. The search term list was constructed based on the PROQOLID (Patient-Reported Outcome and Quality of Life Instruments Database), including PROMs and composite measures, including a patient-reported component.

2.1. *Clinicaltrials.gov database*

The Clinicaltrials.gov database holds information provided by researchers about the studies they plan to conduct. High-level trial characteristics, along with details about outcomes assessed within the studies, are stored on the database. Clinicaltrials.gov website allows users to download a complete record of all trials registered on the database. Records are made available in the form of the Extensible Markup Language (XML) files. On July 22, 2021 Clinicaltrials.gov database snapshot, since its inception, was downloaded. The scope of this paper is solely on RWE studies. The database allows filtering records by stage of a clinical trial, based on definitions developed by the FDA. Our search was restricted to phase IV studies only. This filter was deemed the most appropriate to use, although RWE can be generated by multiple study designs and might be considered as a broader term than "phase IV clinical trials". As a result, records of 27,976 trials were made available for further analysis. Studies included in the analysis reported 159,386 outcomes, as a single study can assess multiple outcomes. The following outcome types are differentiated in the Clinicaltrial.gov database: "primary", "secondary", and "other". Apart from trial

outcomes, high-level trial characteristics were also extracted, including trial ID, first posted date, condition, intervention type, lead sponsor, and country information.

2.2. Search terms lists

The PROQOLID database, part of the ePROVIDE platform, that gathers information about COAs available for use in medical research [22], was used to create the list of search terms. PROQOLID database since its inception in 2002 gathered information about more than five thousand COAs. It was created to facilitate the search, evaluation and selection of appropriate COAs. PROQOLID is the most comprehensive database of PROMs and composite measures and also holds their descriptive information.

Filters embedded in the PROQOLID database allow searching for specific types of outcomes. The database distinguishes the following types of COAs: patient-reported, clinician-reported, observer-reported and performance outcome assessments. Additionally, a composite measure category is available, containing instruments that fall under more than one of the above categories. For this study, two separate search terms lists were created. The first one was constructed using the "PRO" filter, while the second used the "Composite measure" filter to identify measures with patient-reported component.

On July 19, 2021 search term lists were manually copied from the PROQOLID website resulting in 2806 PROMs and 182 composite measure records. For each instrument, the full and abbreviated names were captured as they appeared in the PROQOLID database. To ensure that all relevant trials were identified, even when ClinicalTrials.gov record does not mention exact PROM's name in the outcome description or mentioned name differs from the one in the search term list, the following phrases were added to the PROM search term list: "Quality of life" and "eq5d". Moreover, some abbreviated names of instruments were manually removed from the lists while retaining the full instrument names to increase the searching algorithm specificity. The terms that most frequently resulted in false-positive instrument identification were removed – 33 terms from the PRO list and three from the composite measure list. A list of removed terms is available in the Appendix 1. Complete lists of PROMs and Composite measures search terms are available in Appendixs 2 and 3, respectively.

2.3. Trial characteristics grouping

Conditions investigated in trials are reported as free-text information on ClinicalTrials.gov. Additionally, a list of all conditions grouped into 23 categories is available on the ClinicalTrials.gov website [23]. For this analysis, we adopted the ClinicalTrials.gov disease area grouping. A newly created category "Multimorbidity" was assigned to trials investigating conditions included in more than one group.

Similarly, ClinicalTrials.gov grouping was utilised for intervention type, lead sponsor and region. For intervention type and lead sponsor, additional categories – "Multiple interventions/sponsors" – were created in case more than one intervention/sponsor type was reported for the study. Indexing on ClinicalTrials.gov was not complete; in such instances when a missing value for a trial characteristic was present or the algorithm developed to assign groups to free-text fields was unable to do so based on the information provided on the ClinicalTrials.gov website, "N/A" value was assigned for that variable.

2.4. Searching algorithm development and validation

A computer algorithm was developed de novo to search the ClinicalTrials.gov database snapshot against the full and abbreviated names of instruments stored in search terms lists from PROQOLID. Data compilation and processing were done in Python version 3.8.8 using exact matching. Although alternative approaches were tested including: fuzzy string matching algorithm [24] (matches the sentences using Levenshtein Distance [25]), word ratio (calculates ratio of words that are similar between the compared terms), word2vec [26] (counts words for each term into a vector), and TF-IDF [27] (counts words for each term into a vector but the most important words weight more). Each outcome and its description reported in the ClinicalTrials.gov were matched against up to five terms from search term lists (to capture multiple instruments reported within a single trial outcome).

The Python algorithm was iteratively revised to increase its accuracy by altering algorithm settings. Those settings pertained various approaches to text transformation, length of compared text strings and inclusion of outcome description in searching. For the final analysis, the searching algorithm ignored capitalisation, removed any punctuation, and added spaces before and after searched terms to avoid finding phrases of interest within some other words (e.g. "SOC", which often was identified within the word "social").

Table 1
Search algorithm validation parameters.

	PROMs (%)	Composite measures (%)
Sensitivity	88.3	83.3
Specificity	98.6	98.8
Accuracy	97.6	98.6
PPV	86.5	45.5
NPV	98.8	99.8

PPV, positive predictive value; NPV, negative predictive value.

3. Results

Records of 27,976 phase IV trials were downloaded for analysis. The trials assessed 159,386 outcomes, of which 43,150 were primary and 109,410 secondary outcomes. The remaining 6826 were classified as other outcomes.

The performance of the searching algorithm was evaluated by manual cross-checking by one researcher (KM). A sample of trial records (108 most recently published and 31 oldest records) was screened for the existence of outcomes utilising PROMs or composite measures present in search terms lists. KM evaluated 1003 (0.6%) outcomes from 139 (0.5%) trials. Outcomes flagged by the algorithm as containing at least one instrument of interest were compared to the manual screening conducted by the researcher. The sensitivity and specificity of searching algorithm were calculated. For the PROMs search, sensitivity was 88.3% and specificity 98.6%. Sensitivity and specificity yielded 83.3% and 98.8% respectively for composite measure search (Table 1). Accuracy of the algorithm for both outcome types was higher than the one obtained by Vodicka et al. [21] and was deemed satisfactory. Outcomes incorrectly identified as PROMs or composite measures (false positives) were mainly picked up in two ways: 1) the outcome name or description included on ClinicalTrials.gov matched a term from the PROQOLID list, but in fact did not refer to that measure or instrument listed in PROQOLID. Instead, this referred to a measure with the same name (e.g. National Comorbidity Survey and Nerve Conduction Studies are both written as NCS in the abbreviated form); 2) an outcome reported on clinicaltrials.gov matched a term in PROQOLID, but additional information provided in the ClinicalTrials.gov record indicated that this had been completed by a proxy (parent or teacher of a child). In turn, PROMs or composite measures that algorithm failed to identify (false negatives) were mostly caused by differences in how instrument full name was written and lack of abbreviated name in outcome description.

Out of 159,386 outcomes analysed, 8% assessed at least one PROM. Slightly more than 1% of outcomes were composite measures, including patient-reported component. PROMs were mostly investigated as secondary outcomes, and almost 9% of secondary outcomes utilised PROMs. Counts of trials outcomes utilising PROMs and composite measures are available in Appendix 4.

Out of 27,976 phase IV trials analysed, almost 21% collected at least

Table 2
Use of PROs and composite measures in phase IV trials.

	Number of trials reporting instrument (%)		Number of trials
	PROMs	Composite measures	
Trials reporting at least one instrument	5812 (20.77)	1105 (3.95)	27,976
Trials reporting at least one instrument as primary outcome	1906 (6.81)	436 (1.56)	27,969*
Trials reporting at least one instrument as secondary outcome	4561 (19.94)	797 (3.48)	22,870#
Intervention			
Behavioral	61 (27.6)	8 (3.62)	221
Biological	117 (11.54)	35 (3.45)	1014
Combination Product	14 (22.95)	3 (4.92)	61
Device	355 (22.54)	37 (2.35)	1575
Diagnostic Test	5 (19.23)	0 (0)	26
Dietary Supplement	63 (18.1)	9 (2.59)	348
Drug	3885 (20)	789 (4.06)	19,427
Genetic	3 (33.33)	1 (11.11)	9
Multiple interventions	942 (26.54)	181 (5.1)	3549
Other	152 (22.96)	19 (2.87)	662
Procedure	196 (19.72)	20 (2.01)	994
Radiation	3 (12.5)	0 (0)	24
N/A	16 (24.24)	3 (4.55)	66
Lead sponsor type			
Clinical Research Network	26 (18.98)	6 (4.38)	137
Government, excluding U.S. Federal	125 (15.66)	24 (3.01)	798
Industry	1650 (26.46)	292 (4.68)	6235
National Institute of Health	17 (16.67)	7 (6.86)	102
U.S. Federal Agency, excluding NIH	55 (21.74)	12 (4.74)	253
University/Organization	3503 (18.94)	673 (3.64)	18,497
N/A	436 (22.31)	91 (4.66)	1954
Region			
Africa	67 (7)	8 (0.84)	957
Central America	2 (9.52)	2 (9.52)	21
East Asia	688 (17.11)	159 (3.96)	4020
Europe	1544 (22.36)	287 (4.16)	6906
Middle East	104 (10.77)	22 (2.28)	966
Multiple regions	408 (32.61)	104 (8.31)	1251
North America	2137 (23.19)	380 (4.12)	9214
North Asia	42 (25.45)	7 (4.24)	165
Pacifica	56 (25.93)	9 (4.17)	216
South America	153 (19.01)	21 (2.61)	805
South Asia	63 (14.96)	13 (3.09)	421
Southeast Asia	98 (19.56)	8 (1.6)	501
N/A	450 (17.77)	8 (3.36)	2533
Disease group			
Behaviors and Mental Disorders	597 (43.8)	123 (9.02)	1363
Blood and Lymph Conditions	63 (22.26)	1 (0.35)	283
Digestive System Diseases	162 (15.07)	52 (4.84)	1075
Diseases and Abnormalities at or Before Birth	55 (23.11)	40 (16.81)	238
Ear, Nose, and Throat Diseases	62 (36.26)	6 (3.51)	171
Eye Diseases	80 (13.18)	1 (0.16)	607
Gland and Hormone Related Diseases	138 (14.84)	3 (0.32)	930
Heart and Blood Diseases	323 (14.76)	75 (3.43)	2189
Immune System Diseases	313 (32.95)	181 (19.05)	950
Infections	80 (8.94)	16 (1.79)	895
Mouth and Tooth Diseases	28 (14.74)	2 (1.05)	190
Multimorbidity	337 (23.42)	46 (3.2)	1439
Musculoskeletal Diseases	268 (45.58)	36 (6.12)	588
Neoplasms	41 (22.78)	1 (0.56)	180
Nervous System Diseases	126 (36.21)	35 (10.06)	348
Nutritional and Metabolic Diseases	53 (10.1)	5 (0.95)	525
Respiratory Tract (Lung and Bronchial) Diseases	117 (37.26)	13 (4.14)	314
Skin and Connective Tissue Diseases	69 (24.56)	4 (1.42)	281
Substance Related Disorders	2 (15.38)	0 (0)	13
Symptoms and General Pathology	357 (19.91)	24 (1.34)	1793
Urinary Tract, Sexual Organs, and Pregnancy Conditions	45 (9.43)	1 (0.21)	477
Wounds and Injuries	21 (21)	3 (3)	100
N/A	2475 (19)	437 (3.35)	13,027

* Represents the total number of trials which assessed primary outcomes. Descriptions of seven trials did not contain information about the primary endpoint. It was most likely caused by data errors in the Clinicaltrials.gov database.

Represents a total number of trials which assessed secondary outcomes.

one PROM. At least one composite measure was assessed by nearly 4% of investigated trials (Table 2). Both PROMs and composite measures tended to be assessed as secondary outcomes. The utilisation of PROMs among phase IV trials did not vary greatly between different types of interventions being assessed. The greatest variation was observed in trials investigating biological, genetic and radiation intervention types, but this might be due to a relatively small number of trials grouped in these categories. Trials focusing on biological and radiological interventions assessed PROMs significantly less frequently than on average. On the other hand, PROMs were often collected to investigate genetic treatments. PROMs were most often collected as part of industry-sponsored phase IV trials when compared with other types of lead sponsors. The lowest penetration of PROMs in the phase IV trials was observed in Africa, Central America, and the Middle East.

Despite a substantial quantity of missing data that prevented the identification of disease categories for almost half of the trials included in the analysis, some areas of the most extensive use of PROMs can be described (Table 2). Trials focusing on: Behaviors and Mental Disorders, Ear, Nose, and Throat Diseases, Nervous System Diseases, and

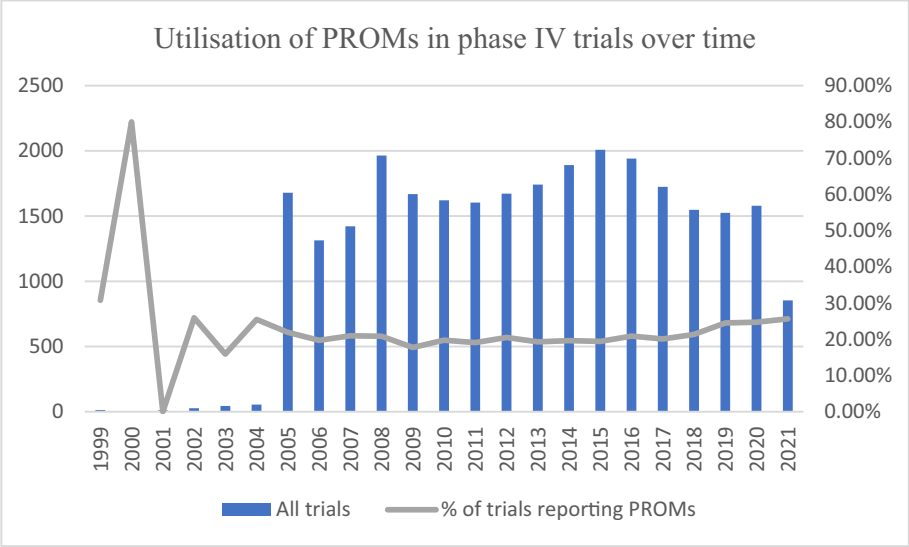


Fig. 1. The utilisation of PROMs in phase IV trials over time.

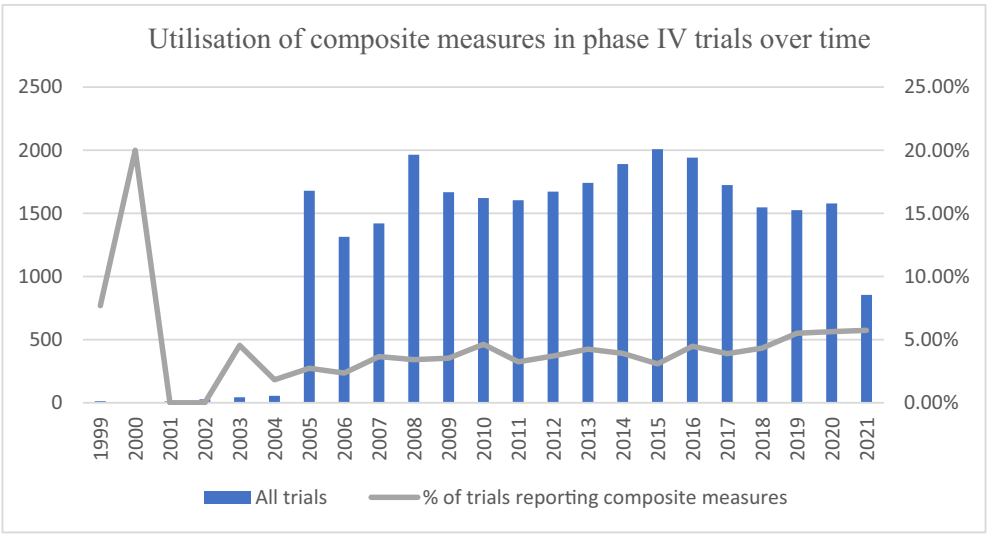


Fig. 2. The utilisation of composite measures in phase IV trials over time.

Respiratory Tract (Lung and Bronchial) Diseases were more likely to collect this type of data. Trials focusing on Infections, Nutritional and Metabolic Diseases, and Urinary Tract, Sexual Organs, and Pregnancy Conditions collected PROMs least often.

A stable level of utilisation for both PROMs and composite measures can be observed since 2005 (Appendix 5). A considerable variation in instruments utilisation was observed before 2005, which might be caused by a low number of trials posted in this period. An increase in PROMs uptake in phase IV trials can be observed since 2019 (*Fig.* 1). Similarly, increase in the utilisation of composite measures was captured since 2019 (Fig. *2*).

Table 3 presents 30 of the most frequently used PROMs. "Quality of life" search term is an umbrella term that picked out PROs that were not specified with exact instrument names but mentioned assessing patients' quality of life. Additionally, trials utilising different types of EQ-5D questionnaires (e.g. EQ-5D-3 L, EQ-5D-5L) were aggregated into a common category. The top five most frequently utilised composite measures included: Pittsburgh Sleep Quality Index, Bleeding Academic Research Consortium Scale, American College of Rheumatology, Diagnostic and Statistical Manual of Mental Disorders, Unified Parkinson's Disease Rating Scale (Appendix 6).

4. Discussion

Of phase IV clinical trials posted on Clinicaltrials.gov between 1999 and July 2021, 21% and 4% used PROMs and composite measures, respectively. Our findings imply a slightly lower utilisation of PROMs than one described by Vodicka et al. [21] (27%). Their analysis covered 2007–2013 and was not restricted only to phase IV studies. These results might suggest lower penetration of PROMs among phase IV studies when compared to earlier phases. The reason for limited widespread of PROMs among phase IV trials is unclear but may be associated with greater difficulties encountered in PRO data collection in a real-world setting and a lack of consensus for optimal data collection and analyses. Collecting PROs is related to additional burden on healthcare professionals, require adjustments to clinical pathways and generate additional costs. Moreover, especially remote utilisation of PROMs is based on patients' compliance and their willingness to provide data which sometimes might be challenging. Mentioned examples offer just a few possible hurdles associated with the use of PROMs in real-world

Table 3
The 30 most frequently used PROMs.

Measure	Number of trials (%)
Quality of Life (umbrella term)	1297 (4.64)
SF-36 Health Survey	507 (1.81)
EQ-5D (sum for different questionnaire versions)	429 (1.53)
Montgomery-Asberg Depression Rating Scale	195 (0.7)
Western Ontario and McMaster Universities Arthritis Index	148 (0.53)
Brief Pain Inventory	138 (0.49)
Health Assessment Questionnaire	127 (0.45)
Hospital Anxiety and Depression Scale	126 (0.45)
SF-12 Health Survey	124 (0.44)
Dermatology Life Quality Index	117 (0.42)
Life Quality Index	103 (0.37)
Epworth Sleepiness Scale	95 (0.34)
Asthma Control Test	95 (0.34)
Pain Catastrophizing Scale	92 (0.33)
International Index of Erectile Function	86 (0.31)
COPD Assessment Test	77 (0.28)
Oswestry Disability Index	74 (0.26)
Balanced Inventory for Spinal disorders	72 (0.26)
International Prostate Symptom Score	71 (0.25)
Quality of Life Scale	64 (0.23)
Ocular Surface Disease Index	63 (0.23)
Severity of Dependence Scale	62 (0.22)
Knee Injury and Osteoarthritis Outcome Score	57 (0.2)
Sheehan Disability Scale	55 (0.2)
Beck Depression Inventory - Second Edition	53 (0.19)
Kansas City Cardiomyopathy Questionnaire	51 (0.18)
Total Symptom Score	51 (0.18)
St George's Respiratory Questionnaire	49 (0.18)
Patient Health Questionnaire	47 (0.17)
Total Nasal Symptom Score	46 (0.16)

settings, which holds back their full implementation. Undoubtedly, more issues need to be resolved, and additional guidance how to tackle these is required. The uptake of PROMs in RWE generation can be stimulated by initiatives aiming to produce guidance on methodologies for data collection, analysis and PRO data use. International agreement upon standards for PROMs' utilisation should facilitate its uptake in RWE generation. The regulators (MHRA, FDA or EMA) or international societies (The International Society for Health Economics and Outcomes Research or International Society for Quality of Life Research) have an essential role in promoting PROs for RWE generation. Guidelines for PRO data collection and utilisation should increase their use in real-world studies.

Our findings depicted a relatively steady uptake of PROMs in phase IV clinical trials from 2005 to 2019. A gradual increase in the utilisation of this type of outcome was observed from 2019. An increase in the utilisation of PROMs over time was also captured by previous studies, which were not restricted to phase IV trials. The earlier analysis of Clinicaltrials.gov records by Scoggins and Patrick [20], which spanned between 2004 and 2007, reported that 14% of trials used at least one PROM. This constitutes a significant increase in PROMs' utilisation since 1997, when Sanders [23] observed that only 4.2% of studies used it. A similar percentage (4.4%) was observed by Naito et al. [23] among Japanese trials between 2000 and 2003.

Several important limitations merit discussion. Although our primary interest is in RWE, we were forced to focus on phase IV trials only in this analysis. Due to the indexing of Clinicaltrials.gov database, the trial phase was applied to filter records. RWE can be generated using different study designs and is undoubtedly a broader term than the phase IV trial. This can be seen as one of the limitations of this study. Nevertheless, in our opinion, the main observations - limited use of PROMs when compared with earlier phases trials - can be extrapolated to the entire body of RWE. Another limitation of this study is the US-focused nature of Clinicaltrials.gov database. Thus, studies conducted in some geographies might be overlooked. Nevertheless, as already presented in Table 2, our approach allowed for international coverage of trials included into analysis. Moreover, missing field completion on the database hampered analysis of some of the trial characteristics of interest. This was particularly visible when summarising conditions targeted by individual studies.

Additionally, the use of a searching algorithm imposed some challenges and although this may not be as accurate as manual records screening, this approach allowed for analysis of the large sample size, which would have been difficult manually. In addition, the algorithm can be easily replicated on other data sets. The method utilised in this study allows for identifying only these measures, which are indexed in PROQOLID. Thus, our results might slightly underestimate the actual uptake of PROMs and composite measures in phase IV trials, mainly when investigators have used non-specific terminology around symptom assessment and measurement scales. Improved labelling of trial outcomes by clear defining the PROMs would facilitate indexing and registration on Clinicaltrials.gov. This would certainly enhance the execution of similar research in the future. Another limitation of this study is associated with the fact that trials' outcomes captured in the Clinicaltrials.gov database might not accurately represent clinical trial protocols. Again, improvements in reporting to the database should allow for more robust conclusions drawn from this type of research in the future.

5. Conclusions

In conclusion, the use of PROMs in phase IV trials seems to be lower than its use in earlier phases of clinical research. Recent years demonstrated a steady increase in the utilisation of PROMs in phase IV trials. A number of initiatives can be developed to improve the incorporation of PROMs in RWE studies including the development of best practices for their use and highlighting needs of regulators and payers.

Financial disclosure statement

This research was conducted as part of a PhD programme funded by GSK. TK is a co-supervisor of KM (the holder of the GSK PhD grant) and is a Director at GSK Ltd. In his role as co-supervisor TK inputted to all stages of this research.

CRediT authorship contribution statement

Konrad Maruszczyk: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing – original draft. **Olalekan Lee Aiyegbusi:** Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing. **Victor Roth Cardoso:** Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Writing – review & editing. **Georgios V. Gkoutos:** Formal analysis, Methodology, Resources, Software, Writing – review & editing. **Luke T. Slater:** Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Writing – review & editing. **Philip Collis:** Methodology, Writing – review & editing. **Thomas Keeley:** Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing. **Melanie J. Calvert:** Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

MC is Director of the Birmingham Health Partners Centre for Regulatory Science and Innovation, Director of the Centre for Patient-Reported Outcomes Research and is a National Institute for Health and Care Research (NIHR) Senior Investigator. She receives funding from the NIHR Birmingham Biomedical Research Centre, the NIHR Surgical Reconstruction and Microbiology Research Centre and NIHR Applied Research Collaboration (ARC) West Midlands at the University

of Birmingham and University Hospitals Birmingham NHS Foundation Trust, Health Data Research UK, Innovate UK (part of UK Research and Innovation), Macmillan Cancer Support, UCB and GSK. MC has received personal fees from Astellas, Takeda, Merck, Daiichi Sankyo, Glaukos, CIS Oncology, Aparito Ltd, GSK, Genentech and the Patient-Centered Outcomes Research Institute (PCORI) outside the submitted work.

OLA receives funding from the NIHR Birmingham Biomedical Research Centre (BRC), NIHR Applied Research Collaboration (ARC) West Midlands at the University of Birmingham and University Hospitals Birmingham NHS Foundation, Innovate UK (part of UK Research and Innovation), Gilead Sciences Ltd, Janssen Pharmaceuticals, Inc., and Sarcoma UK. OLA declares personal fees from Gilead Sciences Ltd, GSK and Merck outside the submitted work.

TK is an employee and shareholder of GSK Ltd.

KM is the holder of the GSK PhD grant.

Other authors declare no competing interests.

The views expressed in this article are those of the authors and not necessarily those of the NIHR, or the Department of Health and Social Care.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cct.2022.106882.

References

[1] B.K. Beaulieu-Jones, S.G. Finlayson, W. Yuan, R.B. Altman, I.S. Kohane, V. Prasad, et al., Examining the use of real-world evidence in the regulatory process, Clin. Pharmacol. Ther. 107 (4) (2020) 843–852. Epub 2019/09/29, https://doi.org/10.1002/cpt.1658. PubMed PMID: 31562770; PubMed Central PMCID: PMCPMC7093234.

[2] A.A. Pulini, G.M. Caetano, H. Clautiaux, L. Vergeron, P.J. Pitts, G. Katz, Impact of real-world data on market authorization, reimbursement decision & price negotiation, Therap. Innovat. Regulat. Sci. 55 (1) (2021) 228–238. Epub 2020/08/29.

[3] FDA, US Food & Drug Administration. Framework for FDA's Real-World Evidence Program, 2018.

[4] FDA, Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products, 2021.

[5] FDA, Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products Guidance for Industry, 2021.

[6] FDA, Data Standards for Drug and Biological Product Submissions Containing Real-World Data, 2021.

[7] FDA, Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products, 2021.

[8] MHRA, MHRA Guideline on Randomised Controlled Trials using Real-World Data to Support Regulatory Decisions, 2021.

[9] MHRA, MHRA Guidance on the Use of Real-World Data in Clinical Studies to Support Regulatory Decisions, 2021.

[10] NICE. Real-world evidence framework feedback 2022 [13/04/2022]. Available from: https://www.nice.org.uk/about/what-we-do/real-world-evidence-framework-feedback.

[11] P. Arlett, J. Kjær, K. Broich, E. Cooke, Real-world evidence in eu medicines regulation: enabling use and establishing value 111 (1) (2022) 21–23, https://doi.org/10.1002/cpt.2479.

[12] Kim HS, Lee S, Kim JH. Real-world evidence versus randomized controlled trial: clinical research based on electronic medical records. J. Korean Med. Sci. 2018;33 (34):e213. Epub 2018/08/22. doi: https://doi.org/10.3346/jkms.2018.33.e213. PubMed PMID: 30127705; PubMed Central PMCID: PMCPMC6097073.

[13] FDA. FDA Real-world evidence 2021 [01/11/2021]. Available from: https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence.

[14] P.G. Kluetz, D.J. O'Connor, K. Soltys, Incorporating the patient experience into regulatory decision making in the USA, Europe, and Canada, Lancet Oncol. 19 (5) (2018), https://doi.org/10.1016/s1470-2045(18)30097-4 e267-e74. Epub 2018/05/05. PubMed PMID: 29726391.

[15] M.J. Calvert, D.J. O'Connor, E.M. Basch, Harnessing the patient voice in real-world evidence: the essential role of patient-reported outcomes, Nat. Rev. Drug Discov. 18 (10) (2019) 731–732. Epub 2019/10/02, https://doi.org/10.1038/d41573-019-00088-7. PubMed PMID: 31570837.

[16] FDA. FDA Patient-Focused Drug Development Guidance Series for Enhancing the Incorporation of the Patient's Voice in Medical Product Development and Regulatory Decision Making 2020 [01/11/2021]. Available from: https://www.fda.gov/drugs/development-approval-process-drugs/fda-patient-focused-drug-development-guidance-series-enhancing-incorporation-patients-voice-medical.

[17] MAHRA. Guidance. Innovative Licensing and Access Pathway 2021 [01/11/2021]. Available from: https://www.gov.uk/guidance/innovative-licensing-and-access-pathway.

[18] European Medicines Agency, Appendix 2 to the Guideline on the Evaluation of Anticancer Medicinal Products in Man: The Use of Patient-Reported Outcome (PRO) Measures in Oncology Studies, 2016.

[19] K. Maruszczyk, O.L. Aiyegbusi, B. Torlinska, P. Collis, T. Keeley, M.J. Calvert, Systematic review of guidance for the collection and use of patient-reported outcomes in real-world evidence generation to support regulation, reimbursement and health policy, J. Patient-Reported Outcomes. 6 (1) (2022) 57, https://doi.org/10.1186/s41687-022-00466-7.

[20] J.F. Scoggins, D.L. Patrick, The use of patient-reported outcomes instruments in registered clinical trials: evidence from ClinicalTrials.gov, Contemp. Clin. Trials. 30 (4) (2009) 289–292, https://doi.org/10.1016/j.cct.2009.02.005.

[21] E. Vodicka, K. Kim, E.B. Devine, A. Gnanasakthy, J.F. Scoggins, D.L. Patrick, Inclusion of patient-reported outcome measures in registered clinical trials: Evidence from ClinicalTrials.gov (2007–2013), Contemp. Clin. Trials. 43 (2015) 1–9, https://doi.org/10.1016/j.cct.2015.04.004.

[22] Mapi Research Trust, 2022 [20/03/2022]. Available from: https://eprovide.mapi-trust.org/about/about-proqolid.

[23] National Library of Medicine, ClinicalTrials.gov 2022 [cited 2022 25/02/2022]. Available from: https://clinicaltrials.gov/ct2/search/browse?brwse=cond_cat.

[24] SeatGeek Inc, fuzzywuzzy: Fuzzy String Matching in Python, 2014.

[25] V.I. Levenshtein, Binary codes capable of correcting deletions, insertions and reversals, Soviet Phys. Doklady. 10 (1966) 707.

[26] Q. Le, T. Mikolov, Distributed Representations of Sentences and Documents, Available from: https://arxiv.org/pdf/1405.4053.pdf, 2014.

[27] Jones K. Sparck, A statistical interpretation of term specificity and its application in retrieval, J. Doc. 28 (1) (1972) 11–21, https://doi.org/10.1108/eb026526.