The importance of effective registries in pulmonary diseases and how to optimize their output

Joanna Chorostowska-Wynimko[1] ,
Marion Wencker[2] and Ildikó Horváth[3,4]

Abstract
Randomized controlled trials (RCTs) are essential for the approval of new therapies; however, because of their design, they provide little insight concerning disease epidemiology/etiology and current clinical practice. Particularly, in lung disease, rigid inclusion/exclusion criteria can limit the generalizability of pivotal trial data. Noninterventional studies (NIS), conducted through the well-established mechanism of patient registries, are undervalued as a means to close data gaps left by RCTs by providing essential data that can guide patient care at different levels from clinical decision-making to health-care policy. While NIS contribute valuable data in all disease areas, their importance in rare diseases cannot be underestimated. In respiratory disease, registries have been essential in understanding the natural history and different phenotypes of rare conditions, such as alpha 1 antitrypsin deficiency, cystic fibrosis, and idiopathic pulmonary fibrosis. Importantly, additional therapeutic outcome data were generated. While measures for enhancing data quality in RCTs have evolved significantly, the approach and effectiveness of registries is variable. Within this article, we review the contribution of registries to pulmonary disease and make recommendations for their effective management. Additionally, we assess limitations of registry data as well as challenges to registry operation, including the impact of the European Union General Data Protection Regulation.

Keywords
Alpha 1 antitrypsin deficiency, lung diseases, rare diseases, registries, quality improvement

Date received: 24 May 2019; accepted: 17 September 2019

Introduction

Data from randomized controlled trials (RCTs) are considered the gold standard for evaluating the efficacy of an intervention and building evidence to support regulatory approval and drug licensing. Noninterventional studies (NIS), which comprise a group of studies with a variety of methodologies, can be used to collect real-word clinical data and to support findings of RCTs. In addition, NIS, often conducted through the well-established mechanism of patient registries, provide longitudinal outcomes, context in clinical practice, and help to shape future study initiatives. NIS are also frequently used to investigate the effects of drugs and other interventions but have been criticized for their potential to overestimate treatment effects, particularly if treatment is not concealed and a control group is lacking.[1] However, when analyzing the outcomes of single RCTs and

[1] Department of Genetics and Clinical Immunology, National Institute of Tuberculosis and Lung Diseases, Warsaw, Poland
[2] Conresp, Lörzweiler, Germany
[3] Department of XIV, National Koranyi Institute for Pulmonology, Budapest, Hungary
[4] Department of Public Health, Semmelweis University, Budapest, Hungary

Corresponding author:
Joanna Chorostowska-Wynimko, Department of Genetics and Clinical Immunology, National Institute of Tuberculosis and Lung Diseases, ul. Płocka 26 01-138 Warsaw, Poland.
Email: j.chorostowska@igichp.edu.pl

observational studies or the conclusions of meta-analyses, the treatment effects reported in RCTs and observational studies were found to be very similar.[1,2] Therefore, NIS, particularly registry studies, provide an alternative source of data that can support both regulatory and routine clinical decision-making but are currently underutilized for this function. The purpose of this article is to assess the contribution of registry studies to the understanding of pulmonary diseases and outline requirements for the effective management and continuous improvement of registries in respiratory medicine.

Requirements for drug registration: The necessity of RCTs

Data from RCTs—particularly double-blind RCTs—are required to prove the efficacy of an intervention and are an integral part of the approval and licensing process for new therapies. RCTs are designed for success. This is achieved by reducing bias and confounding factors through precise study design and use of narrow inclusion and exclusion criteria, in addition to, the randomized allocation of study interventions. These measures increase internal validity[3] but come at a price: factors that create high internal validity in RCTs may reduce the generalizability of the data, with data on effects in broader populations often obtained postapproval. Low generalizability of RCT data has been particularly noted for common respiratory conditions. In a survey of 749 asthma patients, only 6% of those receiving asthma treatments were eligible to participate in RCTs of asthma medications, based on the Global Initiative for Asthma guidelines.[4] A further study found that by utilizing common inclusion criteria for clinical trials in chronic obstructive pulmonary disease (COPD), including nonatopic, forced expiratory volume in one second <70% predicted and smoking history >15 pack-years, fewer than 17% of the COPD population tested were eligible.[5]

Pragmatic RCTs (pRCTs) utilize patient populations that closely mirror that of the general patient population and may increase the generalizability of clinical trials.[6] The differences in the approaches of RCTs and pRCTs, in addition to details of a recent high-profile example of a pRCT in respiratory medicine—The Salford Lung Study—are shown in Table 1.[7,8] Despite the advantages of pRCTs, they cannot be applied to all scenarios. A less homogenous study population and poor medication compliance contribute to increased variance in treatment response

and a lower signal-to-noise ratio.[9,10] This can make it more difficult to determine significant differences between groups and may result in incorrect conclusions about the intervention.[9] Therefore, pRCTs do not fulfill regulatory guidelines and thus are not suitable for drug development.

Challenges for RCTs in rare diseases

Regulatory agencies require data from RCTs before licensing a new compound in a rare disease. Although orphan drug legislation helps to facilitate drug development in rare diseases, conducting RCTs in this area poses additional challenges.[11,12] First, it is often difficult to recruit a large enough sample of comparable patients.[13] Eligible patients can be geographically dispersed worldwide. Second, trial recruitment can be hampered when a licensed drug is already on the market. Alpha 1 antitrypsin (AAT) therapy for the treatment of alpha 1 antitrypsin deficiency (AATD) has been on the market in some countries for a number of years, with US Food and Drug Administration (FDA) approval originally based on biochemical efficacy.[14,15] Consequently, a placebo-controlled trial to determine the clinical efficacy of AAT therapy in AATD (the RAPID clinical trial program) needed 4.5 years to recruit 180 patients worldwide.[16] In general, long recruitment and study periods provide further challenges. In the case of placebo-controlled studies, it may be ethically questionable to withhold treatment from patients for an extended period if there is already a good body of evidence for efficacy. In addition, long recruitment periods may jeopardize the study if advances in the quantification of disease severity have been made, for example, advancements in imaging technologies, or other drugs have become available. Patients may also choose to withdraw consent once a treatment becomes available in their country. Moreover, patient dispersal among treatment centers worldwide creates further issues. The inclusion of smaller centers with few patients and less experience of the disease within RCTs may add variability and bias, and RCTs could also be sensitive to patient variability because of environmental influence. Environmental factors may be particularly pertinent in respiratory disease, as different concentrations of ambient air pollutants could contribute to variations in baseline lung function and disease status/progression.[17] These factors may limit the utility of the outcome data recorded and can contribute to gaps within the final trial data.

Table 1. Key differences between efficacy and effectiveness studies in comparison to the Salford Lung Study.[8]

	Efficacy—Can the intervention work? (RCTS)	Effectiveness—Does the intervention work when used in normal practice? (pRCTS)	Example of effectiveness study—Salford Lung Study
Setting	Well resourced, "ideal" setting.	Normal practice.	Seventy-five general practice (primary care) surgeries located in Salford and South Manchester, UK.
Participants	Highly selected. Poorly adherent participants and those with conditions that might dilute the effect are often excluded.	Little or no selection beyond the clinical indication of interest.	Inclusion criteria • ≥40 years • Documented COPD diagnosis from a general practitioner • ≥1 exacerbation in the past 3 years • Use of maintenance LABA/LAMA (alone or in combination) Exclusion criteria Of those listed, two of note were: • Occurrence of an exacerbation within the previous two weeks • Long-term corticosteroid use No restrictions were applied based on spirometric values or smoking history.
Intervention	Strictly enforced and adherence is monitored closely.	Applied flexibility as it would be in normal practice.	Patients receiving the intervention (fluticasone furoate–vilanterol group) could switch to usual care; however, patients receiving usual care (control group) could not switch to the intervention. Utilization of established electronic health records enabled nonintrusive monitoring of safety and efficacy in routine clinical practice.
Outcomes	Often short-term surrogates or process measures.	Directly relevant to participants, funders, communities, and health-care practitioners.	Primary Annual rate of moderate or severe exacerbations that resulted in the use of systemic corticosteroids/antibiotics, or led to hospital visits or admission. Secondary COPD symptoms and quality of life measures: CAT and EQ-5D.
Relevance to practice	Indirect—little effort made to match design to decision-making needs of those in usual setting in which intervention will be implemented.	Direct—trial is designed to meet needs of those making decisions about treatment options in settings in which intervention will be implemented.	In an environment where inhaler adherence was highly variable, and thereby closer to the realities of clinical practice, the study demonstrated a significant reduction in rates of moderate or severe exacerbations (8.4%; $p = 0.02$) following fluticasone furoate-vilanterol therapy versus usual care.

Source: Republished with permission of British Medical Journal, from Improving the reporting of pragmatic trials: an extension of the CONSORT statement, Zwarenstein M, et al. 337, 2008[7]; permission conveyed through Copyright Clearance Center, Inc.
CAT: COPD assessment test; COPD: chronic obstructive pulmonary disease; EQ-5D: European Quality of Life–5 dimensions; LABA: long-acting beta agonist; LAMA: long-acting muscarinic agonists; pRCT: pragmatic randomized controlled trial; RCT: randomized controlled trial.

A further complication for RCTs in rare disease arises when the natural history of a disease and the diversity of disease presentation are poorly understood.[18] This may mean that it is difficult to design the trial with regards to patient selection and duration of treatment/follow-up. In addition, where

information concerning disease activity or disease progression is lacking, it can be difficult to select the most suitable clinical endpoint, potentially necessitating the use of a nonspecific outcome measure.[18,19] As a result, larger patient numbers and/or longer follow-up times may be required. This is particularly the case with patient-reported outcomes (PROs), for example, quality of life assessments. PROs are often less sensitive than other clinical outcomes and, due to their subjective nature, can be associated with higher inter-patient variability.[20] For example, using St George's Respiratory Questionnaire as an outcome parameter in AATD would require more than 8500 patients per study arm to be followed for more than 3 years to detect a 25% reduction in decline.[21] This is at a time when regulatory agencies are recommending that more pivotal studies provide data on PROs—the FDA recently set out to dramatically increase the level of patient input in clinical trials.[22]

These factors mean that clinical trials in rare disease must be extremely well designed and highly controlled to maximize the potential to demonstrate significant differences between groups, and the difficulty in accurately assessing PROs in rare diseases represents a significant data gap. However, a recent article in the *New England Journal of Medicine* by Frieden et al. highlights that it may not always be the best course of action to wait for key data to be produced by an RCT when these data could be generated through well-designed and implemented NIS.[23] Therefore, in an ideal situation, results from both RCTs and NIS would be available for a disease, to provide complementary and unique insights into whom to treat, and with which treatment to obtain the best results.

Bridging the data gap with NIS

Clinical trials are characterized by an experimental design that randomly assigns patients to treatment arms or subgroups. In contrast, NIS observe and record the disease course and are designed to answer questions reflecting routine clinical practice and the impact of disease and treatments on patients' everyday lives. Consequently, only treatments and diagnostic methods that are the normal standard-of-care can be used, and only marketed medical interventions can be utilized in accordance with their license restrictions.[24]

NIS, including population, case-control, and cohort studies, each have their own unique features, aims, and study design (Figure 1).[25,26] Potentially, the NIS most frequently encountered in everyday clinical

practice is the clinical audit, which is an essential tool for capturing a snapshot of health-care provision and highlighting areas for improvement. Although most frequently associated with local implementation, large-scale audits involving mass data collection that capture the status of service provision across countries and regions are becoming more frequent, the European COPD Audit being a key example.[27] However, for clinical research, and for investigating disease etiology and the effect of treatments in particular, longitudinal follow-up is required. An important type of NIS in this regard, and focus of the current review, is registry studies (Figure 2).

Bridging the gap: Focus on registry studies

Registry studies are inherently observational in nature, although some degree of control can be implemented in that any planned data analyses can drive the manner of the data collection.[28] The implementation of NIS provides an effective toolset for successful registries by facilitating longitudinal patient follow-up and uniform data collection. These data can be used to answer a broad range of questions that can contribute to strategic health-care planning, the development of disease prevention activities, improving health care quality, planning of clinical trials, regulatory policy, and clinical decision-making.[29] Although registries can contribute useful data in all disease areas, their use is essential in understanding the nature of rare diseases. Of particular importance are registries that include patients treated with Orphan drugs, as these databases can inform on the effectiveness and safety of therapies that received marketing authorization at a time when the clinical evidence was credible but not conclusive.[30] A summary of several important registries in respiratory medicine and their contributions to their respective fields are presented in Table 2.[31–51]

Considerations for the effective management of registries
Roadmap to the continuous improvement of registries

To maximize the usefulness of data produced by registries, high-quality data entry is essential. However, operating patient registries requires significant resources and organizational collaboration between many stakeholders;[13] yet independent rare disease registries, in particular, operate with few resources
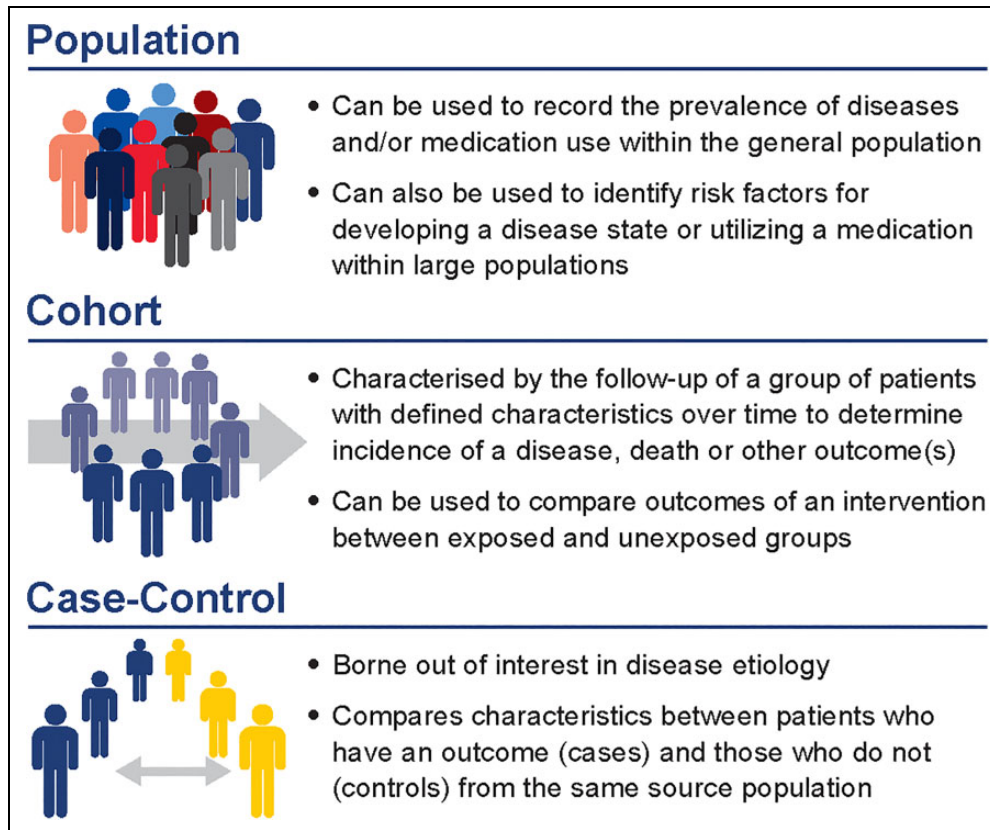
Figure 1. Key features of different types of NIS.[24,25] COPD: chronic obstructive pulmonary disease; NIS: noninterventional studies.
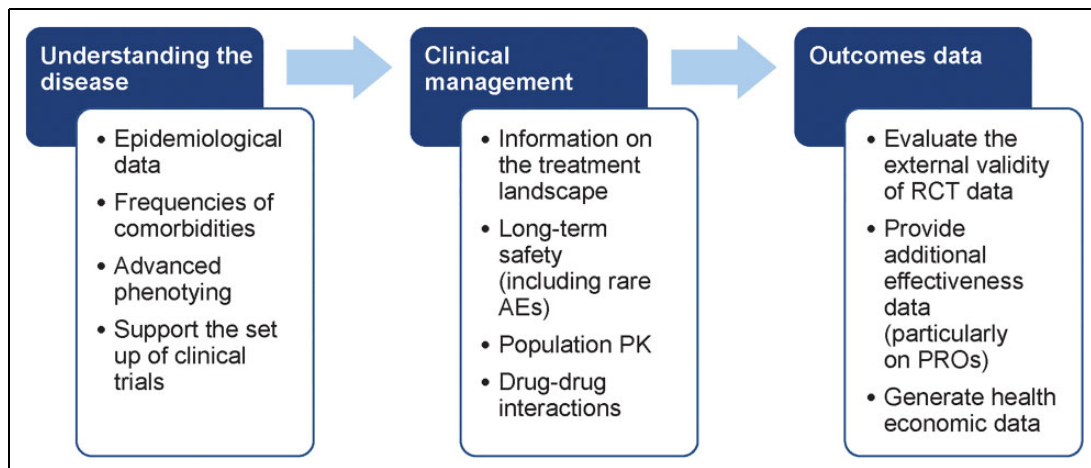

Figure 2. The utility of registry studies. AE: adverse event; PK: pharmacokinetics; PRO: patient-reported outcome; RCT: randomized clinical trial.

and often rely on the dedicated work of a few physicians.

Consequently, data quality in registries may be reduced due to inaccurate/incomplete data recording or incorrectly registered patients.[52] As a result, important disease outcomes can be inaccurately reported

and the magnitude of any treatment effect can be under- or overestimated. To obtain good quality data and reduce the amount of bias introduced in the data collection, quality control should be similar to that employed in RCTs. Best practice for quality control in RCTs is based on the principles outlined in the

Table 2. Examples of important registries in respiratory medicine and their contributions.

Name of registry	Knowledge contributed to the field/Impact on treatment guidelines
AATD	
National Heart, Lung and Blood Institute (NHLBI) registry[31]	• Provided some of the first evidence for the clinical efficacy of AAT therapy on spirometric decline and mortality, which was subsequently reflected in treatment guidelines.
Danish AATD registry[32,33]	• Helped to define differences in disease course between index and nonindex cases • Demonstrated the effect of primary prophylaxis (i.e., a smoking-free lifestyle) on improving disease outcomes.
Swedish AATD registry[34]	• Enabled characterization of lung function and clinical status changes through the first four decades of life, as a result of identification by neonatal screening undertaken in the 1970s. • Also demonstrated the effect of smoking-free lifestyle on disease outcomes.
Italian/Polish AATD registry[35,36]	• Revealed the diversity of genetic variants associated with AATD.
German AATD registry[37]	• The first registry to identify different phenotypes of disease progression in AATD, that is, individuals with very fast decline in lung function.
Irish AATD registry[38]	• Enabled characterization of disease risk in carriers (heterozygotes) of the key deficiency variant in AATD, which has shaped recent guideline recommendations on the management of these individuals.
Asthma	
The Belgian Severe Asthma Registry (BSAR)[39]	• Demonstrated that the majority of patients with severe asthma present with persistent airflow limitation and eosinophilic inflammation despite high-dose corticosteroid use, highlighting an unmet need in asthma therapeutics.
The Spanish Severe Asthma Registry[40]	• Collected data that could help to expand the indication for omalizumab to patients with IgE levels outside the normal range.
The eXpeRience registry[41,42]	• Demonstrated improvements in PROs and health-care utilization with omalizumab therapy that could not have been established through RCTs.
COPD	
COPD and SYstemic consequences-COmorbidities NETwork (COSYCONET)[43,44]	• Enabled characterization of the health-care utilization costs associated with progression of COPD through GOLD stages and the impact of COPD comorbidities on QoL.
Cystic fibrosis	
Cystic Fibrosis Foundation registry[45-47]	• Enabled characterization of trends in survival in patients with CF in the United States and reduced survival in the United States compared with Canada. • Publishes extensive annual reports on the status of CF management in the United States.
European Cystic Fibrosis Society patient registry (ECFSPR)[48,49]	• Enabled characterization of factors associated with lung function decline in patients with CF. • Publishes annual reports on patient demographics and treatment of CF in Europe.
Bronchiectasis	
Bronchiectasis Audit and Research Collaboration (EMBARC)[50]	• Provided a platform for research that was previously greatly lacking, enabling the publication of several important analyses, mostly focused on characterizing the etiology of non-CF bronchiectasis.
Idiopathic pulmonary fibrosis	
Australian Idiopathic Pulmonary Fibrosis Registry (AIPFR)[51]	• Contributed valuable data regarding the survival advantage with antifibrotic therapy in patients with IPF.

AAT: alpha 1 antitrypsin; AATD: alpha 1 antitrypsin deficiency; CF: cystic fibrosis; COPD: chronic obstructive pulmonary disease; GOLD: global initiative for chronic obstructive lung disease; IPF: idiopathic pulmonary fibrosis; PRO: patient reported outcome; QoL: quality-of-life; RCT: randomized controlled trial.

International Council for Harmonization guidelines.[53] While there are guidelines on Good Epidemiological Practice,[54] a framework for data quality in registries, and to help ensure that data collection between centers is controlled and standardized, particularly when collecting data on PROs, is required. These principals are reflected in a recent discussion article by the European Medicines Agency's Cross-Committee Task Force on Patient Registries, which has the objective of facilitating the use of patient registries in supporting regulatory decision-making.[55]

When looking for guidance on the optimal operation of registries, it is also possible to look toward examples of effectively run registries. The US Cystic Fibrosis Foundation Patient Registry (CFFPR) has been invaluable in identifying the epidemiology of the genetic variants associated with the disease and monitoring the microbiology of chronic lung infections and the effectiveness of prophylactic antibiotics. The registry has been operational since 1986; in this time, it has accumulated data from over 48,463 CF patients.[56] Participation among patients assessed at CFFPR-accredited care centers is high, and few patients are lost to follow-up. A significant feature of the registry is the publication of annual summary reports that are freely accessible to the general public. These reports comprehensively review the status of CF diagnosis and management in the US, covering everything from patient demographics, diagnosis, and genetic testing to guidelines for screening and patient management, microbiology, current treatment modalities, and clinical outcome data.[47] The registry utilizes stringent quality assurance protocols to ensure that the data collected are comprehensive and highly accurate. A 2012 audit of the CFFPR demonstrated the rigorousness of data collection. Of patients included in the registry, data were recorded for 95% of clinic visits and 90% of hospitalizations.[56] In addition, all audited fields were found to be highly accurate, with inaccurate reporting ranging from 0.8% to 17.4%. The European equivalent to the CFFPR, European Cystic Fibrosis Society Patient Registry (ECFSPR),[49] has been similarly successful and is widely used for both academic and industry research purposes. Overall, the success of registries such as the CFFPR and ECFSPR is dependent on three criteria: quality of structure, quality of process, and quality of data (Table 3).

Quality of structure. A professional structure underpins the success of any registry. Many registries, particularly in rare diseases, are dependent upon the dedicated work of a small group of physicians. To generate the data output required to perform effective analyses in a timely and efficient manner, dedicated staff are required. In addition, oversight by a subcommittee and defined roles and responsibilities can ensure that funding structures are in place, research areas of interest are identified and focused on, and all regulations, such as data protection and good clinical practice guidelines, are adhered to. Quality of structure undoubtedly requires increased resources and funding; however, it can be argued that an efficiently run, well-resourced registry can contribute far more valuable data in the long term.

Quality of process. The processes of data collection and handling in registries are multifaceted. Standardization is an essential pillar that should be implemented within registries and agreed between international counterparts. This requires defined research protocols to ensure that methodologies do not deviate between centers contributing data. In addition, utilizing accredited centers (as in the CFFPR) can also aid the collection of a homogenous/comparable data set. A quality control plan and regular auditing can ensure that targets for data collection are met and that any deficiencies in standardization are addressed. Furthermore, effective design of the registry's information technology infrastructure can contribute to quality assurance. The Alpha-1 Foundation research registry utilizes customized data archiving systems and portals for participating centers to achieve this. Another consideration is the ethical framework of the data collection, an extremely important aspect with potential legal ramifications if regulations are not adhered to. An ethics committee would usually be required to approve the conduct of any research studies and to ensure that all patient consent requirements are followed.

Quality of data. Inaccuracies in data quality, which can be due to a range of factors, including programing errors, transcription/typing errors, and deviations from data collection protocols,[57,58] can ultimately affect the outcome of any analyses performed. Procedures are therefore required for identifying outlying data and for recording and following-up on missing or incomplete data. To reduce imputation errors, databases can utilize automated logic check algorithms that prevent out-of-range values being entered and reject patient cases with missing/incomplete data—a system utilized by the European Multicentre

Bronchiectasis Audit and Research Collaboration.[59] Overall, ensuring high-quality audited data is essential to secure stakeholder funding, particularly from industry.

Recognizing the limitations of registries and current challenges

Despite the advantages of registry studies described above, a number of limitations should be noted. A key strength of registry data is that it is more generalizable, as data input into registries is often not strictly monitored; however, the inclusion of ineligible patients, for example, those with less severe disease, can occur.[60] Furthermore, bias can be included at several stages, in part due to a lack of randomization in observational research (Figure 3).[61] Researchers should seek to limit the influence of biases wherever possible, for example, by ensuring accurate data collection and limiting selection bias. However, bias through confounding (both measured and unmeasured), which is largely mitigated by randomization in RCTs, is a particular issue for NIS.[62] It is therefore important for researchers to be aware of the statistical analyses that can be used to help adjust for confounders. The influence of measured confounders can be assessed through stratification and multivariate analyses (including logistic regression, linear regression, and analysis of covariance).[62] Unmeasured confounders, as would be expected, are more difficult to adjust for; however, several methods, such as the confounding function approach, have been developed in this regard.[63] The potential influence of unmeasured confounders in NIS highlights the importance of compressive data collection in the design of registries, and the level of evidence provided by registry studies is influenced to a large extent by the ability of analyses to adjust for confounders.

In addition to known limitations, open questions remain on how registry data can be utilized. The introduction of the new European Union General Data Protection Regulation (GDPR) may create significant issues for EU patient registries, particularly in rare diseases. In addition, although these directions only apply to EU citizens, any collaboration of US registries with EU counterparts would necessitate alignment with GDPR. GDPR lays down strict regulations that go beyond the current US confidentiality laws, that is, Health Insurance Portability and Accountability Act of 1996.[64] First, institutions involved in systematic data collection on a large
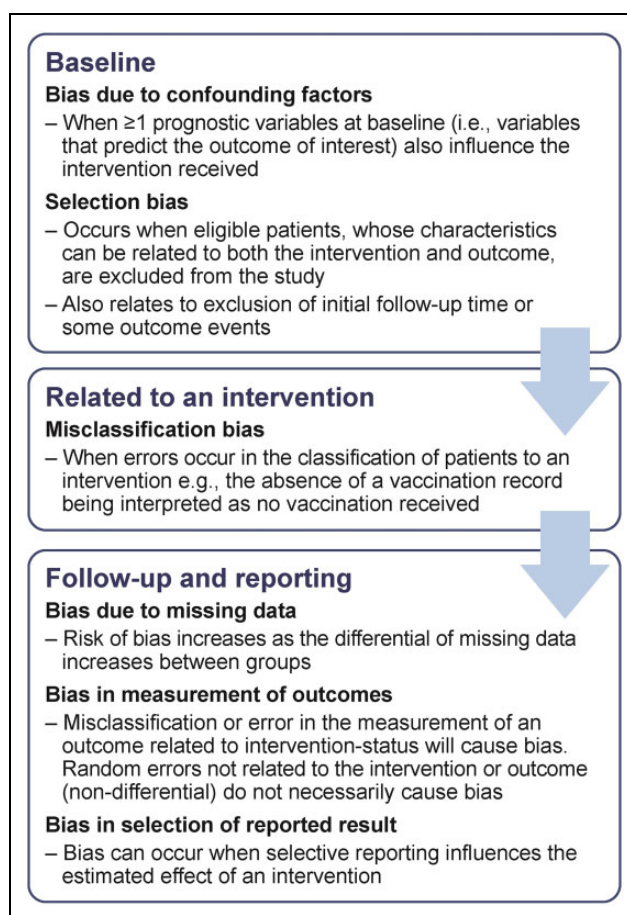

Baseline
Bias due to confounding factors
– When ≥1 prognostic variables at baseline (i.e., variables that predict the outcome of interest) also influence the intervention received

Selection bias
– Occurs when eligible patients, whose characteristics can be related to both the intervention and outcome, are excluded from the study
– Also relates to exclusion of initial follow-up time or some outcome events

Related to an intervention
Misclassification bias
– When errors occur in the classification of patients to an intervention e.g., the absence of a vaccination record being interpreted as no vaccination received

Follow-up and reporting
Bias due to missing data
– Risk of bias increases as the differential of missing data increases between groups

Bias in measurement of outcomes
– Misclassification or error in the measurement of an outcome related to intervention-status will cause bias. Random errors not related to the intervention or outcome (non-differential) do not necessarily cause bias

Bias in selection of reported result
– Bias can occur when selective reporting influences the estimated effect of an intervention

Figure 3. Types of bias in observational studies.[61]

scale that requires regular monitoring, such as patient registries, will be required to appoint a Data Protection Officer.[65] Second, there are further obligations to maintain documentation, implement data protection protocols, and conduct data protection impact assessments.[65] Third, the GDPR requires explicit informed consent separate from any other agreements; although existing consent agreements may still apply, it is likely that many will not. This has been a source of much controversy in the medical community,[66] particularly with respect to rare diseases. EPIRARE, the foremost European rare disease society, previously set up a petition to protest the requirements for new levels of consent that may have an adverse effect on the ability of registries to collect essential data.[67] The workability of these regulations within existing medical research apparatuses has yet to be determined; however, concessions for medical research may be required down the line to avoid these regulations having an adverse impact on essential data collection efforts.

Table 3. Prerequisites for a successful disease registry.

Quality of structure	Quality of process	Quality of data
- Oversight by a subcommittee - Definition of roles and responsibilities within the registry - Dedicated staff - Sufficient resources available, including funding	- Centre accreditation - One centralized database - Protocol and defined research question(s) - Standardized policies and procedures - Ethics committee and patient consent - Quality control plan – Review criteria – Timetable for monthly, biannual or annual activities - Monitoring/external audits - Recommendation of corrective action if needed - Definition of how to handle missing data (data imputation)	- Monitoring of accuracy and completeness of data - Source data verification - Query process for missing, incomplete or outlying data/inconsistent values - Defined regular data updates – % data missing – % data not known

In addition to data protection, data ownership is a major issue, which may be particularly pertinent within consortia-led registries. Public funding for patient registries is fiercely competed, and therefore, many rely upon fundraising activities or support from industry to fill this gap. Consequently, the funding of registries varies widely and there are often several stakeholders involved, including patients, patient advocacy groups, clinicians, academics, pharmaceutical companies, and government regulatory agencies and payers.[13] Some registries are fully maintained by government agencies—mandatory registries exist for significant threats to public health (e.g. tuberculosis),[68] and others are set up to encourage research, collaboration and increase knowledge and awareness on less-understood topics, the US Rare Disease Network being an example of this.[69] Alternatively, some groups receive funding from both industry and public bodies. The International Rare Disease Research Consortium—an extremely large worldwide network—is an example of this approach.[70] However, in scenarios where funding sources are mixed, it can be difficult to manage the expectations of all parties involved with regard to data collected and/or reported. To avoid controversy, detailed consortium agreements should be developed at the outset to ensure that the interests of each collaborating party are met. In addition, to maintain transparency, concerning any potential or perceived financial conflicts of interest, there is a need for protocols that define the scope of industry input and funding when setting up registries.

Conclusion

RCTs are essential for the approval of drugs; however, due to stringent control on trial recruitment and limited follow-up periods, the data often provide limited epidemiological/etiological insight, can lack generalizability, and can provide limited information on PROs. NIS, all of which can be conducted through patient registries, can be utilized to fill these data gaps. However, there are also limitations to the data generated by NIS, principally due to the lack of control for confounding factors. Therefore, RCTs and registry studies should be viewed in parallel, with both providing distinct but complementary data that helps our understanding of disease areas and treatment modalities. To obtain quality data from registries, there should be the same rigor in data collection and validation as applied to clinical trials. A well-designed registry should ensure the quality of structure, process, and data; this will undoubtedly require additional resources and dedicated staff. In addition, procedural requirements, including the new EU data protection regulations, should be carefully considered when establishing new patient registries. Overall, registry studies contribute to the understanding of disease epidemiology, clinical utility of new drugs, and policy making.

Author contributions

All authors contributed to the writing of the manuscript, reviewed the manuscript, and approved the manuscript for submission.

ORCID iD

Joanna Chorostowska-Wynimko https://orcid.org/0000-0003-1743-2961

References

1. Concato J, Shah N, and Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med* 2000; 342(25): 1887–1892.

2. Benson K and Hartz AJ. A comparison of observational studies and randomized, controlled trials. *Am J Ophthalmol* 2000; 130(5): 688.

3. Rothwell PM. Factors that can affect the external validity of randomised controlled trials. *PLoS Clin Trials* 2006; 1(1): e9.

4. Travers J, Marsh S, Williams M, et al. External validity of randomised controlled trials in asthma: to whom do the results of the trials apply? *Thorax* 2007; 62(3): 219–223.

5. Herland K, Akselsen JP, Skjonsberg OH, et al. How representative are clinical study patients with asthma or COPD for a larger "real life" population of patients with obstructive lung disease? *Respir Med* 2005; 99(1): 11–19.

6. Roland M and Torgerson DJ. Understanding controlled trials: what are pragmatic trials? *Br Med J* 1998; 316(7127): 285.

7. Zwarenstein M, Treweek S, Gagnier JJ, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *Br Med J* 2008; 337: a2390.

8. Vestbo J, Leather D, Diar Bakerly N, et al. Effectiveness of fluticasone furoate-vilanterol for COPD in clinical practice. *N Engl J Med* 2016; 375(13): 1253–1260.

9. Moller HJ. Effectiveness studies: advantages and disadvantages. *Dialogues Clin Neurosci* 2011; 13(2): 199–207.

10. Treweek S and Zwarenstein M. Making trials matter: pragmatic and explanatory trials and the problem of applicability. *Trials* 2009; 10: 37.

11. European Medicines Agency. Human regulatory. Research and development: orphan incentives, https://www.ema.europa.eu/en/human-regulatory/research-development/orphan-designation/orphan-incentives (accessed 14 May 2019).

12. Federal Drug Administration. Rare diseases: common issues in drug development; draft guidance for industry; availability. Federal Register / Vol. 84, No. 22 / Friday, February 1, 2019/Notices, https://www.govinfo.gov/content/pkg/FR-2019-02-01/pdf/2019-00677.pdf (2019, accessed 14 May 2019).

13. Gliklich RE, Dreyer NA, Leavy MB, et al. 20 Rare disease registries. *Registries for evaluating patient outcomes: a user's guide [internet]*, 3rd ed. Rockville: Agency for Healthcare Research and Quality (US), 2014.

14. Wewers MD, Casolaro MA, Sellers SE, et al. Replacement therapy for alpha 1-antitrypsin deficiency associated with emphysema. *N Engl J Med* 1987; 316(17): 1055–1062.

15. Wewers MD, Casolaro MA, and Crystal RG. Comparison of alpha-1-antitrypsin levels and antineutrophil elastase capacity of blood and lung in a patient with the alpha-1-antitrypsin phenotype null-null before and during alpha-1-antitrypsin augmentation therapy. *Am Rev Respir Dis* 1987; 135(3): 539–543.

16. Chapman KR, Burdon JG, Piitulainen E, et al. Intravenous augmentation treatment and lung density in severe alpha1 antitrypsin deficiency (RAPID): a randomised, double-blind, placebo-controlled trial. *Lancet* 2015; 386(9991): 360–368.

17. Kelly FJ and Fussell JC. Air pollution and airway disease. *Clin Exp Allergy* 2011; 41(8): 1059–1071.

18. Pariser A. Rare disease and clinical trials. Office of translational sciences. Centre for drug evaluation and research. Federal Drug Administration. https://www.fda.gov/downloads/Drugs/NewsEvents/UCM440797.pdf (2014, accessed 26 July 2018).

19. Augustine EF, Adams HR, and Mink JW. Clinical trials in rare disease: challenges and opportunities. *J Child Neurol* 2013; 28(9): 1142–1150.

20. Ali S, Littlewood E, McMillan D, et al. Heterogeneity in patient-reported outcomes following low-intensity mental health interventions: a multilevel analysis. *PLoS One* 2014; 9(9): e99658.

21. Stockley RA, Edgar RG, Starkey S, et al. Health status decline in α-1 antitrypsin deficiency: a feasible outcome for disease modifying therapies? *Respir Res* 2018; 19(1): 137.

22. Federal Drug Administration. Percentage increase in the number of approved IDEs (pivotal studies only) with patient reported outcomes (PRO), https://www.accessdata.fda.gov/scripts/fdatrack/view/track.cfm?program=cdrh&status=public&id=CDRH-Number-of-patient-perspective-studies-used-in-support-of-pre market-and-postmarket-regulatory-decisions&fy=All (2017, accessed 26 July 2018).

23. Frieden TR. Evidence for health decision making - beyond randomized, controlled trials. *N Engl J Med* 2017; 377(5): 465–475.

24. Theobald K, Capan M, Herbold M, et al. Quality assurance in non-interventional studies. *Ger Med Sci* 2009; 7: Doc29. [Article in English, German]

25. Song JW and Chung KC. Observational studies: cohort and case-control studies. *Plast Reconstr Surg* 2010; 126(6): 2234–2242.

26. Gershon AS, Warner L, Cascagnette P, et al. Lifetime risk of developing chronic obstructive pulmonary disease: a longitudinal population study. *Lancet* 2011; 378(9795): 991–996.

27. Lopez-Campos JL, Hartl S, Pozo-Rodriguez F, et al. European COPD audit: design, organisation of work and methodology. *Eur Respir J* 2013; 41(2): 270–276.

28. Gliklich RE, Dreyer NA, Leavy MB, et al. 3 Registry design. *Registries for evaluating patient outcomes: a user's guide [Internet]*. 3rd ed. Rockville: Agency for Healthcare Research and Quality (US), 2014.

29. Kissne Horvath I.[Patient registries from the view of health policy]. *Orv Hetil* 2014; 155(19): 729–731.

30. Orphanet Report Series. Rare diseases collection. Rare disease registries in Europe, http://www.orpha.net/orphacom/cahiers/docs/GB/Registries.pdf (2019, accessed 26 July 2018).

31. Survival and FEV1 decline in individuals with severe deficiency of alpha1-antitrypsin. The alpha-1-antitrypsin deficiency registry study group. *Am J Respir Crit Care Med* 1998; 158(1): 49–59.

32. Seersholm N, Kok-Jensen A, and Dirksen A. Survival of patients with severe alpha 1-antitrypsin deficiency with special reference to non-index cases. *Thorax* 1994; 49(7): 695–698.

33. Seersholm N and Kok-Jensen A. Clinical features and prognosis of life time non-smokers with severe alpha 1-antitrypsin deficiency. *Thorax* 1998; 53(4): 265–268.

34. Piitulainen E, Mostafavi B, and Tanash HA. Health status and lung function in the Swedish alpha 1-antitrypsin deficient cohort, identified by neonatal screening, at the age of 37-40 years. *Int J Chron Obstruct Pulmon Dis* 2017; 12: 495–500.

35. Luisetti M, Ferrarotti I, Corda L, et al. Italian registry of patients with alpha-1 antitrypsin deficiency: general data and quality of life evaluation. *Chronic Obstruct Pulm Dis* 2015; 12(Suppl 1): 52–57.

36. Duk K, Zdral A, Szumna B, et al. Frequency of rare alpha-1 antitrypsin variants in Polish patients with chronic respiratory disorders. *Adv Exp Med Biol* 2016; 910: 47–53.

37. Wencker M, Fuhrmann B, Banik N, et al. Longitudinal follow-up of patients with alpha(1)-protease inhibitor deficiency before and during therapy with IV alpha(1)-protease inhibitor. *Chest* 2001; 119(3): 737–744.

38. Molloy K, Hersh CP, Morris VB, et al. Clarification of the risk of chronic obstructive pulmonary disease in alpha1-antitrypsin deficiency PiMZ heterozygotes. *Am J Respir Crit Care Med* 2014; 189(4): 419–427.

39. Schleich F, Brusselle G, Louis R, et al. Heterogeneity of phenotypes in severe asthmatics. The Belgian Severe Asthma Registry (BSAR). *Respir Med* 2014; 108(12): 1723–1732.

40. Vennera Mdel C, Perez De Llano L, Bardagi S, et al. Omalizumab therapy in severe asthma: experience from the Spanish registry–some new approaches. *J Asthma* 2012; 49(4): 416–422.

41. Braunstahl GJ, Chen CW, Maykut R, et al. The experience registry: the 'real-world' effectiveness of omalizumab in allergic asthma. *Respir Med* 2013; 107(8): 1141–1151.

42. Braunstahl GJ, Canvin J, Peachey G, et al. Healthcare resource utilization in patients receiving omalizumab for allergic asthma in a real-world setting. *Biol Ther* 2014; 4(1-2): 57–67.

43. Wacker ME, Jorres RA, Schulz H, et al. Direct and indirect costs of COPD and its comorbidities: results from the German COSYCONET study. *Respir Med* 2016; 111: 39–46.

44. Wacker ME, Jorres RA, Karch A, et al. Relative impact of COPD and comorbidities on generic health-related quality of life: a pooled analysis of the COSYCONET patient cohort and control subjects from the KORA and SHIP studies. *Respir Res* 2016; 17(1): 81.

45. MacKenzie T, Gifford AH, Sabadosa KA, et al. Longevity of patients with cystic fibrosis in 2000 to 2010 and beyond: survival analysis of the cystic fibrosis foundation patient registry. *Ann Intern Med* 2014; 161(4): 233–241.

46. Stephenson AL, Sykes J, Stanojevic S, et al. Survival comparison of patients with cystic fibrosis in Canada and the United States: a population-based cohort study. *Ann Intern Med* 2017; 166(8): 537–546.

47. Cystic Fibrosis and Foundation. Cystic fibrosis foundation patient registry 2015 annual data report, https://www.cff.org/Our-Research/CF-Patient-Registry/2015-Patient-Registry-Annual-Data-Report.pdf (2015, accessed 27 July 2018).

48. Kerem E, Viviani L, Zolin A, et al. Factors associated with FEV1 decline in cystic fibrosis: analysis of the ECFS patient registry. *Eur Respir J* 2014; 43(1): 125–133.

49. European Cystic Fibrosis Society (ECFS). ECFS patient registry annual data report 2015, https://www.ecfs.eu/sites/default/files/general-content-images/working-groups/ecfs-patient-registry/ECFSPR_Report2015_Nov2017.pdf (2015, accessed 26 July 2018.

50. Aliberti S, Polverino E, Chalmers JD, et al. The European multicentre bronchiectasis audit and research collaboration (EMBARC) ERS clinical research collaboration. *Eur Respir J* 2018; 52(5): 1802074.

51. Jo HE, Glaspole I, Grainge C, et al. Baseline characteristics of idiopathic pulmonary fibrosis: analysis from the Australian Idiopathic Pulmonary Fibrosis Registry. *Eur Respir J* 2017; 49(2): 1601592.

52. Arts DG, De Keizer NF, and Scheffer GJ. Defining and improving data quality in medical registries: a literature review, case study, and generic framework. *J Am Med Inform Assoc* 2002; 9(6): 600–611.

53. International Council for Harmonisation. Good clinical practice guidelines, http://www.ich.org/products/guidelines/efficacy/efficacy-single/article/good-clinical-practice.html (2016, accessed 27 July 2018).

54. International Epidemiology Association. Good epidemiological practice. IEA guidelines for proper conduct in epidemiological research, http://ieaweb.org/good-epidemiological-practice-gep/ (2017, accessed 27 July 2018).

55. European Medicines Agency. Discussion Paper. Use of patient disease registries for regulatory purposes – methodological and operational considerations. The cross-committee task force on patient registries, https://www.ema.europa.eu/documents/other/discussion-paper-use-patient-disease-registries-regulatory-purposes-methodological-operational_en.docx (2018, accessed 18 December 2018).

56. Knapp EA, Fink AK, Goss CH, et al. The cystic fibrosis foundation patient registry. design and methods of a national observational disease registry. *Ann Am Thorac Soc* 2016; 13(7): 1173–1179.

57. Hilner JE, McDonald A, Van Horn L, et al. Quality control of dietary data collection in the CARDIA study. *Control Clin Trials* 1992; 13(2): 156–169.

58. Gassman JJ, Owen WW, Kuntz TE, et al. Data quality assurance, monitoring, and reporting. *Control Clin Trials* 1995; 16(Suppl 2): 104S–136S.

59. Chalmers JD, Aliberti S, Polverino E, et al. The EMBARC European Bronchiectasis Registry: protocol for an international observational study. *ERJ Open Res* 2016; 2(1): 00081–2015.

60. Levine MN and Julian JA. Registries that show efficacy: good, but not good enough. *J Clin Oncol* 2008; 26(33): 5316–5319.

61. Sterne JA, Hernan MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *Br Med J* 2016; 355: i4919.

62. Pourhoseingholi MA, Baghestani AR, and Vahedi M. How to control confounding effects by statistical analysis. *Gastroenterol Hepatol Bed Bench* 2012; 5(2): 79–83.

63. Kasza J, Wolfe R, and Schuster T. Assessing the impact of unmeasured confounding for binary outcomes using confounding functions. *Int J Epidemiol* 2017; 46(4): 1303–1311.

64. Beaver L.Digital health briefing: how GDPR will impact the US health system — Digital health deals hit $1.6 B in Q1 — Ascension sells hospital, pivots to virtual medicine. Business Insider UK, http://uk.businessinsider.com/digital-health-briefing-how-gdpr-will-impact-the-us-health-system-digital-health-deals-hit-16-b-in-q1-ascension-sells-hospital-pivots-to-virtual-medicine-2018-4 (2018, accessed 26 July 2018).

65. The European Union. General Data Protection Regulation. A new data Potection Landscape. Allen & Overy LLP, http://www.allenovery.com/SiteCollectionDocuments/Radical%20changes%20to%20European%20data%20protection%20legislation.pdf (2017, accessed 27 July 2018).

66. Ploem MC, Essink-Bot ML, and Stronks K. Proposed EU data protection regulation is a threat to medical research. *Br Med J* 2013; 346: f3534.

67. Vittozzi L.Re: Proposed EU data protection regulation is a threat to medical research, http://www.bmj.com/content/346/bmj.f3534/rr/649040 (2013, accessed 26 July 2018).

68. Ronald LA, Campbell JR, Balshaw RF, et al. Predicting tuberculosis risk in the foreign-born population of British Columbia, Canada: study protocol for a retrospective population-based cohort study. *BMJ Open* 2016; 6(11): e013488.

69. Rare Disease Clinical Research Network. https://www.rarediseasesnetwork.org/. (accessed 10 April 2018).

70. International Rare Diseases Research Consortium. Goals and vision, http://www.irdirc.org/about-us/vision-goals/ (accessed 10 April 2018)..