

University of Groningen

## Capturing Data in Rare Disease Registries to Support Regulatory Decision Making

Jonker, Carla J.; de Vries, Sieta T.; van den Berg, H. Marijke; McGettigan, Patricia; Hoes, Arno W.; Mol, Peter G. M.

*Published in:*  
Drug Safety

*DOI:*  
[10.1007/s40264-021-01081-z](https://doi.org/10.1007/s40264-021-01081-z)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

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

*Publication date:*  
2021

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Jonker, C. J., de Vries, S. T., van den Berg, H. M., McGettigan, P., Hoes, A. W., & Mol, P. G. M. (2021). Capturing Data in Rare Disease Registries to Support Regulatory Decision Making: A Survey Study Among Industry and Other Stakeholders. *Drug Safety*. <https://doi.org/10.1007/s40264-021-01081-z>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*



# Capturing Data in Rare Disease Registries to Support Regulatory Decision Making: A Survey Study Among Industry and Other Stakeholders

Carla J. Jonker<sup>1,2</sup> · Sieta T. de Vries<sup>2,3</sup> · H. Marijke van den Berg<sup>4</sup> · Patricia McGettigan<sup>5</sup> · Arno W. Hoes<sup>1</sup> · Peter G. M. Mol<sup>2,3</sup>

Accepted: 16 May 2021  
© The Author(s) 2021

## Abstract

**Introduction** In rare diseases, registry-based studies can be used to provide natural history data pre-approval and complement drug efficacy and/or safety knowledge post-approval.

**Objective** The objective of this study was to investigate the opinion of stakeholders about key aspects of rare disease registries that are used to support regulatory decision making and to compare the responses of employees from industry to other stakeholders.

**Methods** A web-based survey was used to gauge the importance of (1) common data elements (including safety outcomes), (2) data quality and (3) governance aspects that are generic across different rare diseases. The survey included 47 questions. The data were collected in the period April–October 2019.

**Results** Seventy-three respondents completed  $\geq 80\%$  of the survey. Most of the respondents were from the industry ( $n = 42$ , 57%). For safety data, 31 (42%) respondents were in favour of collecting all adverse events. For data quality, the respondents found a level of 30% reasonable for source data verification. For missing data, a level of 20% was considered acceptable. Compared to responders from industry, the other stakeholders found it less relevant to share data with industry and found it less acceptable if the registry is financed by industry.

**Conclusions** This study showed that the opinion towards data and governance is well aligned across parties, and issues of data and governance on their own should not pose a barrier to collaboration. This finding is supportive of the European Medicines Agency's efforts to encourage stakeholders to work with existing registries when collecting data to support regulatory decision making.

✉ Carla J. Jonker  
c.jonker@cbg-meb.nl; c.j.jonker-2@umcutrecht.nl

- <sup>1</sup> Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, University of Utrecht, Utrecht, The Netherlands
- <sup>2</sup> Dutch Medicines Evaluation Board (CBG-MEB), Graadt van Roggenweg 500, 3531 AH Utrecht, The Netherlands
- <sup>3</sup> Department Clinical Pharmacy and Pharmacology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
- <sup>4</sup> PedNet Haemophilia Research Foundation, Baarn, The Netherlands
- <sup>5</sup> William Harvey Research Institute, Queen Mary University of London, London, UK

## Key Points

A survey among industry and other stakeholders was used to investigate the key aspects of rare disease registries to support regulatory decision making.

A set of demographics, clinical and medication-related data were identified that focused primarily on the disease of interest with much less emphasis on co-morbidities or adverse events.

Compared to responders from industry, the other stakeholders found it less relevant to share data with industry and found it less acceptable if the registry is financed by industry.

## 1 Introduction

There is a large unmet medical need for effective treatments for the 30 million patients in Europe who suffer from one of 6000–8000 rare diseases [1]. The search for new treatments in rare diseases is challenging, owing to the small and heterogeneous patient populations and often limited knowledge of the diseases' natural history [2]. Because of the scarcity of patients, and ethical concerns with denying beneficial active treatment, well-designed controlled clinical trials to assess the efficacy of a medicinal product and to detect serious adverse events can be difficult to perform [3]. Typically, marketing authorisation of therapies in rare diseases is therefore based on less evidence than for more common disorders. Consequently, post-marketing activities, such as registry-based studies, to further evaluate the effectiveness and/or safety of a new medicinal product are crucial in this patient population. Moreover, disease registries could provide data on the natural course of a disease, which may be used as a historical or external control to support a marketing application based on single-arm trial data. However, the contribution of registries to the knowledge of a new medicinal product is currently limited because of problems with the patient accrual rate, delayed start of patient inclusion, low data quality and missing data [4–6].

In Europe, governance organisations are aware of these problems leading to the suboptimal use of registries and introduced efforts to improve the contribution of registries to assess risks and benefits of treatments. The European Medicines Agency (EMA) initiated the Patient Registry Initiative [7], with the main goal to make better use of existing registries to support regulatory decision making. Among other things, the Patient Registry Initiative organised meetings with stakeholders to discuss the importance of key elements of registries, including common data elements to be collected, data quality and governance aspects [8]. Attendees of these meetings were employees working in the pharmaceutical industry, regulators, academia, registry owners and patient representatives. Recently, the European Network for Health Technology Assessment (EUnetHTA) developed the Registry Evaluation and Quality Standards Tool (REQueST) [9]. This tool assesses the quality of patient registries to support more systematic and widespread use of registry data in health technology assessments. In parallel, the European Reference Networks on rare diseases made recommendations to improve the quality of registries [10] and included similar key elements as identified by the Patient Registry Initiative.

The primary aim of this study was to quantify the opinion of stakeholders about key elements of registries as data sources for studies that support regulatory decision making in the field of rare diseases. The secondary aim was to assess whether the importance attached to these key elements differed between industry stakeholders vs others.

## 2 Methods

### 2.1 Study Design and Participants

We conducted a web-based survey among stakeholders familiar with the use of registries in a regulatory context. People known via the Patient Registry Initiative of the EMA and/or owners of registries who were identified via the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance were contacted. These people received a link to the survey via an e-mail and a reminder and could further disseminate the e-mail to people in their network. The survey data were collected during April–October 2019. Ethical approval was not necessary because of the nature of the study.

### 2.2 Outcome Assessment

The survey was constructed in Qualtrics<sup>XM</sup> (a tool for online surveys, <https://www.qualtrics.com>) and included in total 47 questions (Electronic Supplementary Material [ESM]). The survey was created in an iterative process by members of the project team, and members of the EMA's Patient Registry Initiative. The survey questions were based on the themes of the workshops held at the EMA, i.e. common data elements, data quality and governance, and recurrent issues as described elsewhere [8]. In the survey, a registry was defined as “an organized system that uses observational methods to collect uniform data on specified outcomes in a population defined by a particular disease, condition or exposure” [11].

In the survey, two questions were used to assess the characteristics of the responder and two questions were included about *registries in general*. The other questions concerned the three key elements of registries; i.e. 24 questions on *common data elements* that included aspects considered to be essential for the collection of demographic and baseline data, treatment and safety outcomes, and duration of follow-up; ten questions about *data quality* covering data entry, optimising and improving data quality, source data verification and missing data; and four questions about aspects of *governance*. As the intention of this study was to assess the importance attached to these key elements for the use of post-marketing studies, five questions were added about *registry-based studies*. A registry-based study uses a registry infrastructure for patient recruitment and data collection [11].

The survey contained three types of questions: (1) multiple-choice questions where the respondent could choose one or more of the answer options; (2) Likert scale questions with answer options from (1) very unimportant to (5) very important; and (3) a visual analogue scale (VAS). The web-based version of the survey was pretested on functioning and content by seven persons working at the

EMA and/or the Dutch Medicines Evaluation Board. After minor adaptations, the survey was distributed.

## 2.3 Analyses

Respondents who completed  $\geq 80\%$  of the survey were included in the analyses. Descriptive analyses were conducted and results of the multiple-choice questions and Likert scale questions are presented as number and percentage for all included responders and per stakeholder group (i.e. industry vs other stakeholders). The results of the VAS are presented as medians with the interquartile range (IQR).

Differences in responses between industry and the other stakeholders were tested using Pearson  $\chi^2$  tests or Mann–Whitney U tests for the outcomes measured using respectively multiple-choice and Likert scale questions/VAS scale questions. Given the large number of questions and subsequent tests performed,  $p$ -values  $< 0.01$  were considered statistically significant. An element was considered important if  $\geq 80\%$  of the respondents gave it a score of important or very important. This cut-off has been used previously [12] and was used to separate “need to know” elements from “nice to know” elements. The same cut-off was used for the multiple-choice questions and the VAS. Data were analysed using IBM SPSS Statistics 20 (IBM, Armond, New York, United States). Microsoft Excel 2010 (Microsoft, Redmond, Washington, United States) was used for graphical presentation of results.

## 3 Results

There were 201 persons who opened the survey of whom 73 respondents (36%) completed  $\geq 80\%$  of the survey. The median time to complete the survey was 26 minutes (IQR 18–59). Most of them were employees of the pharmaceutical industry ( $n = 42$ ; 57%). The other 31 respondents were employees of European regulatory authorities ( $n = 9$ ; 12%), employees of health technology assessment agencies ( $n = 5$ ; 7%), owners of registries ( $n = 5$ ; 7%), patient representatives ( $n = 3$ ; 4%), physicians ( $n = 2$ ; 3%) or they did not specify their role ( $n = 7$ ; 10%).

### 3.1 Overall Results

#### 3.1.1 General Questions About Registries

A minimal coverage of patients that in the respondents’ view was needed to guarantee a minimal representation of the disease population for use of the registry to support

regulatory decision making was 40% (median, IQR 28–60) (Fig. 1). For the geographical spread of the centres, most respondents considered it important to have centres within more than one country in Europe (92%) and to have at least more than one clinical centre (90%) that collects data (ESM).

#### 3.1.2 Common Data Elements

*Demographic data* considered important to be collected in registries were sex (99%), vital status (93%), age (88%) and current pregnancy (86%) (Fig. 2 and ESM).

Data that need to be collected were clinical data (96% of respondents), treatment data (96%), laboratory data (90%) and patient-reported outcomes (PROs; 82%). For clinical data, the elements considered important were (first date of) diagnosis, severity of the disease, physical function, organ damage and confirmation of the diagnosis; for treatment data, the medical product (93%) and the intervention (86%); for laboratory, blood tests (88%) and biomarkers; (86%); and for PROs, the validated disease questionnaires (82%). Sixty-one percent of the respondents would collect a limited set of baseline data only or found co-morbidity data collection not necessary. Most respondents reported to base the diagnosis on clinical practice guidelines (73%) and the confirmation on a doctor’s recorded diagnosis (80%) [ESM].

Respondents indicated that for the *medication* used to treat the disease of interest, the dosage (96%), the substance (90%), the reason to stop or to switch to another product (89%), and the start and stop date (84%) should be captured (Table 1). Information on non-pharmacological interventions should be captured according to 89% of the respondents. Seventy-four percent of the respondents would collect

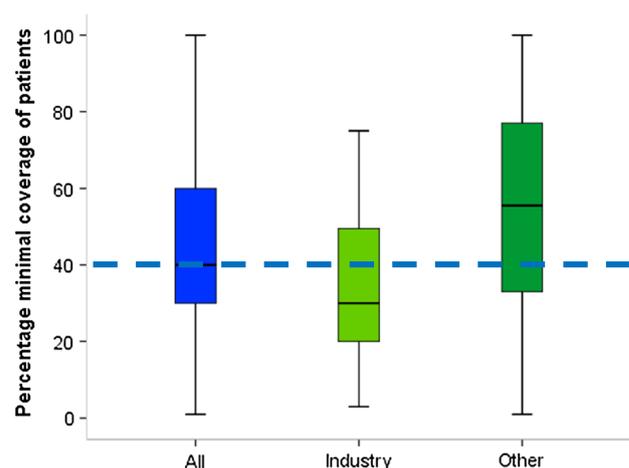
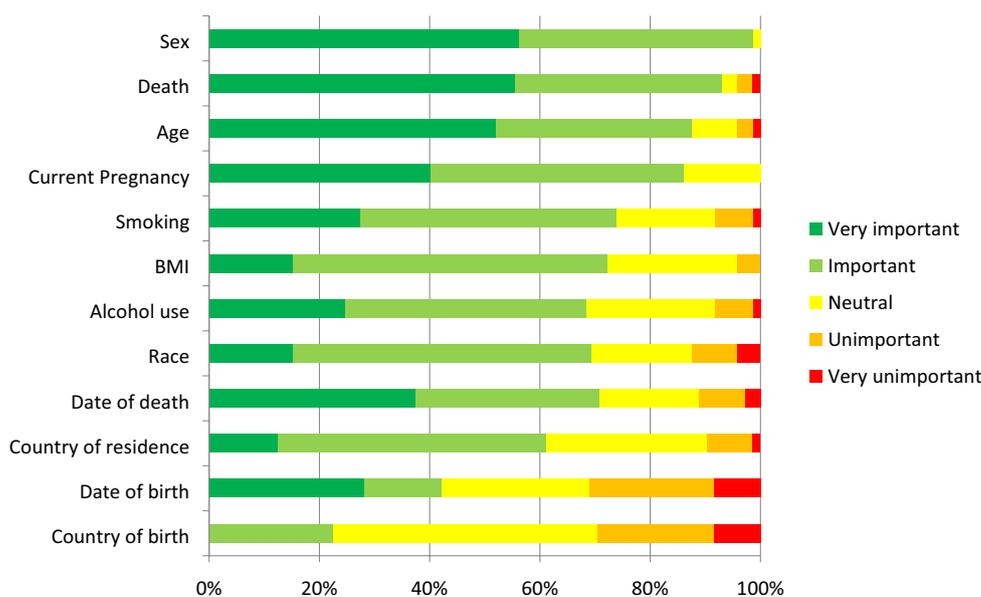


Fig. 1 Percentage of *minimal coverage of patients* that in the respondents’ view is needed to represent the disease population by stakeholder group (all, industry, other stakeholders)

**Fig. 2** Results of the importance attached to the collection of various demographic data. BMI body mass index



no or a limited set of data only for medicinal products used to treat co-morbidities (ESM).

Respondents considered the following data about *pregnant women* to be of key importance; the exposure to any medication during pregnancy (90%), the outcome of the pregnancy (90%), the trimester during exposure (84%) and the follow-up of teratogenic events (84%) (Table 1). Less than half of the participants (44%) would collect details of medication use of the partner, either before or during pregnancy (ESM).

*Treatment outcomes* that the respondents found important to collect pertained to clinical (97%), treatment (96%), laboratory (89%) and PRO (86%) data. Respondents would base the efficacy outcomes primarily on clinical practice (73%), EMA guidelines (69%) and evidence-based literature (69%), but these scores did not reach the predefined 80% threshold. A doctor's recorded diagnosis (81%) was considered most relevant for confirmation of endpoints. Seventy-eight percent of the respondents indicated that only validated endpoints should be used. The majority of respondents thought that endpoints measuring disease progression should be monitored at least quarterly (43%) or twice a year (31%) [ESM].

For *safety outcomes*, 64% of the respondents indicated to collect adverse events of special interest, 62% serious adverse events, and 42% all adverse events. In the case of a reported adverse event, all elements were considered important to record; severity (97%), duration (85%) and, if applicable, causality assessment (85%) (Table 1).

Fifty-four percent of the respondents indicated that the duration of the *follow-up* of a patient should be 1–5 years. During follow-up, use of a medicinal product (89%), whether a patient was lost to follow-up (88%) and the underlying reason (82%) should be captured (ESM).

### 3.1.3 Data Quality

According to the respondents, data were currently mostly entered into registries through web-based platforms (46%) and imported from electronic health records (34%). Data should be entered, in order of preference, by trained staff (75%), treating physicians (60%), patients themselves (58%) or a study coordinator (42%). Data entry was preferred at the time of the actual patient visit (75%) [ESM]. To improve data quality, it was considered of prime importance to have collection instructions (96%), use appropriate software (94%), have well-trained staff (94%) and use standard terminology (90%). To minimise missing data, respondents indicated the value of automated queries (97%), maximising data import from electronic health records (92%) and the use of mandatory fields (90%) to be important. For the improvement of consistency and/or accuracy, alerts (94%) and missing fields over time (80%) were considered important. Annual regular checks (78%) were strongly preferred over random checks (17%) [ESM]. Respondents indicated that to ensure data quality, 30% (IQR 10–54) of source data should be verified and up to 20% (IQR 10–25) of missing data for the key values could be acceptable (Figs. 3 and 4).

### 3.1.4 Governance

The availability of a central contact point (96%) and data sharing across countries (86%) were considered important. Most respondents considered it relevant for regulatory decision making that registry data are shared with regulatory authorities (94%) and academic centres (85%). Additionally, most respondents found it acceptable that registry data are financed by regulatory authorities (92%) or academia (83%),

**Table 1** Number (percentage) of respondents (all, industry and other) that considered the *common data elements*-related questions important<sup>a</sup> with *p*-values of Pearson  $\chi^2$  tests for differences between industry and the other stakeholders

	All (N = 73)	Industry (N = 42)	Other (N = 31)	P-value
<i>Medication</i> : which of the following details with regard to the medicinal product to treat the disease of interest should be captured?				
Dosage	70 (96)	40 (95)	30 (97)	0.74
Substance name	66 (90)	40 (95)	26 (84)	0.10
Reason for stop/switch to other product registered	65 (89)	39 (93)	26 (84)	0.22
Start and stop date	61 (84)	35 (83)	26 (84)	0.95
Duration of the treatment	49 (67)	25 (60)	24 (77)	0.11
ATC <sup>b</sup> classification	33 (45)	14 (33)	19 (61)	0.02
<i>Pregnancy</i> : if women of childbearing potential are captured in the registry, how important is the collection of the following data if a woman becomes pregnant? <sup>c</sup>				
Exposure during pregnancy <sup>d</sup>	64 (90)	42 (100)	22 (76)	<0.01
Outcome of pregnancy <sup>d</sup>	64 (90)	40 (95)	24 (83)	0.20
Trimester during exposure <sup>e,f</sup>	58 (84)	36 (88)	22 (79)	0.61
Follow-up teratogenic events <sup>d,g</sup>	58 (84)	35 (88)	23 (79)	0.26
Follow-up child <sup>d,g</sup>	55 (80)	34 (85)	21 (72)	0.12
Follow-up mother <sup>e,h</sup>	51 (75)	33 (80)	18 (67)	0.25
Birth weight <sup>f,g</sup>	43 (63)	23 (58)	20 (71)	0.06
<i>Safety outcomes</i> : which adverse drug events should be collected?				
Adverse events of special interest	47 (64)	27 (64)	20 (65)	0.98
Serious adverse events	45 (62)	24 (57)	21 (68)	0.36
All adverse events	31 (42)	18 (43)	13 (42)	0.94
For <i>adverse events potentially related</i> to medicinal products the following should be recorded <sup>c</sup>				
Severity of the event	71 (97)	41 (98)	30 (97)	0.97
Duration of the event	62 (85)	34 (81)	28 (90)	0.74
A causality assessment <sup>e</sup>	61 (85)	38 (90)	23 (77)	0.30

<sup>a</sup>An element was considered important if  $\geq 80\%$  of the respondents gave it a score of important or very important

<sup>b</sup>ATC classification: Anatomical Therapeutic Chemical classification

<sup>c</sup>Question assessed using a Likert scale. See ESM, for additional figures of all answers to the Likert scales

<sup>d</sup>Two missing from the respondents from the other stakeholders

<sup>e</sup>One missing from the respondents from industry

<sup>f</sup>Three missing from the respondents from the other stakeholders

<sup>g</sup>Two missing from the respondents from industry

<sup>h</sup>Four missing from the respondents from the other stakeholders

but less so by pharmaceutical companies (78%) or patients (53%) (Table 2). Moreover, respondents indicated that 92% (IQR 81–100) of the registry data should be FAIR (Findable, Accessible, Interoperable, and Reusable; Table 2).

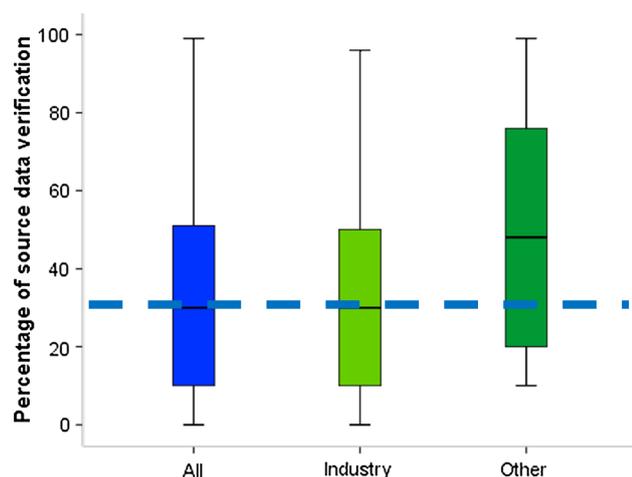
### 3.1.5 Registry-Based Studies

Regarding registry-based studies, respondents found it important that a common study protocol is available in the case of a multi-centre registry (92%) and to have the registry-based study protocol recorded in the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance database (84%). The primary objective (97%), the secondary objective (90%) and a statistical

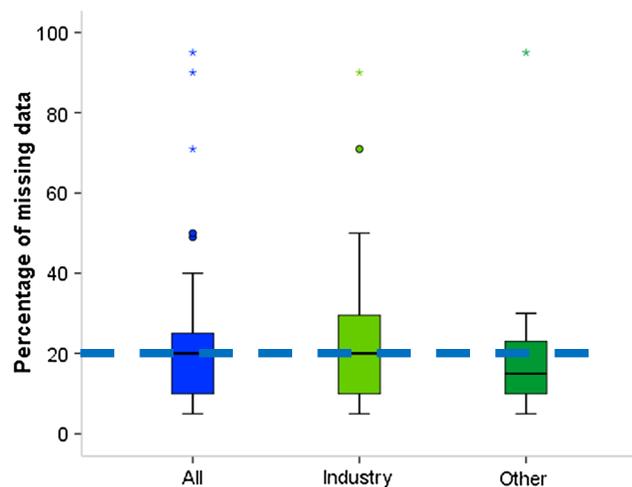
analysis plan (85%) should be predefined in a study protocol. The respondents suggested that missing data, analysis strategy, bias, treatment discontinuation, confounders and effect modifiers should all be addressed in the statistical analysis plan. To perform a randomised study in the context of registry-based studies was considered less essential (i.e. selected by 61% of the respondents) [ESM].

### 3.2 Industry vs Other Stakeholders

In six of the 47 questions, a statistically significant difference was observed between industry and the other stakeholders. The coverage of patients needed to guarantee a minimally acceptable representation of the disease population



**Fig. 3** Percentage of source verification needed that is acceptable by the respondents by stakeholder group (all, industry, other stakeholders)



**Fig. 4** Percentage of missing data that is acceptable by the respondents by stakeholder group (all, industry, other stakeholders)

within the registry was lower for respondents from industry than for respondents from other stakeholders (30% [IQR 20–50] vs 56% [IQR 32–78],  $p < 0.01$ ) (Fig. 1). Compared with the group of other stakeholders, respondents from the industry found the exposure to any medication during pregnancy more important to register (100% vs 76%,  $p < 0.01$ ) (Table 1); would use evidence-based literature less often for selecting the common data elements about the disease; i.e. 52% vs 84%,  $p < 0.01$  (ESM); rated the possibility to request additional information from a treating physician as more important (86% vs 52%,  $p < 0.01$ ); found it more relevant to share data with pharmaceutical companies (90% vs 45%,  $p < 0.01$ ); and found it more acceptable if the registry is

financed by pharmaceutical companies (95% vs 53%,  $p < 0.01$ ) (Table 2).

## 4 Discussion

This study indicated the key aspects in terms of common data elements, data quality and governance for rare disease registries that were important (rated important or very important by  $\geq 80\%$  respondents) to stakeholders. A set of demographics, clinical and medication-related data were identified that focused primarily on the disease of interest with much less emphasis on co-morbidities or adverse events. Respondents considered that 30% of source data verification and 20% of missing data would provide acceptable levels of data quality. Regarding governance, availability of a central contact point and the ability to share data with regulatory authorities was considered important for disease registries to support regulatory decision making in the setting of rare diseases. Regarding registry-based studies, thorough epidemiological and predefined research protocols were expected, with less emphasis on the need for randomised designs. There were few differences between the industry and the other stakeholders. With regard to governance aspects, the other stakeholders found it less relevant to share data with industry and found it less acceptable when a registry is financed by industry.

A core common data set is essential for the interoperability of registries to allow the exchange of data [13]. Previously, a set of common data elements was released by the European Platform on Rare Disease Registration [14]. This set included date of birth, sex and vital status. Our study confirmed the importance of collecting these elements and additionally identified that the stakeholders find it important to collect data on pregnancy. Depending on the therapeutic area or patient population, the choice of key disease-related data elements may, however, differ [15]. To prevent inconsistency in the capturing of the elements, clear definitions need to be formulated [13]. In this context, the defined core data sets by the European Society for Blood and Marrow Transplantation Registry and the European Cystic Fibrosis Society Patient Registry can be used, which have been shown to support regulatory decision making [16–19].

In our study, less than 50% of the respondents indicated that all adverse events should be collected. During one of the previous meetings organised by the EMA's Patient Registry Initiative, routine collection of adverse events was indicated to be a burden [20]. Most registries focus on the collection of serious adverse events and/or adverse events of special interest [20]. An example is the TREatment of ATopic eczema (TREAT) Registry that has the secondary objective to collect eye disorders and eosinophilia as specific types of

**Table 2** Number (percentage) of respondents (all, industry and other) that considered the *governance*-related questions important<sup>a</sup> with *p*-values of Pearson  $\chi^2$  tests for differences between industry and the other stakeholders

	All ( <i>N</i> = 73)	Industry ( <i>N</i> = 42)	Other ( <i>N</i> = 31)	<i>P</i> -value
How important are the following statements about data sharing in relation to governance of a registry? <sup>b</sup>				
Availability of a central contact point	70 (96)	41 (98)	29 (94)	0.25
Data sharing across countries	63 (86)	36 (86)	27 (87)	0.75
Data linkage to other data sources	55 (75)	33 (79)	22 (71)	0.07
Request for additional information, if needed by external stakeholders	52 (71)	36 (86)	16 (52)	<0.01
How relevant is it for regulatory decision-making that registry data are shared with ...? <sup>b</sup>				
Regulatory authorities <sup>c</sup>	68 (94)	39 (95)	29 (94)	0.69
Academic centres <sup>c</sup>	61 (85)	36 (88)	25 (81)	0.18
Pharmaceutical companies <sup>c</sup>	51 (71)	37 (90)	14 (45)	<0.01
How acceptable is it that a registry is (partly) financed by ... if the data are to be used by regulators: <sup>b</sup>				
Regulatory authorities <sup>d</sup>	66 (92)	37 (88)	29 (97)	0.20
Academia <sup>d</sup>	60 (83)	38 (90)	22 (73)	0.12
Independent stakeholders <sup>d</sup>	57 (79)	32 (76)	25 (83)	0.29
Pharmaceutical companies <sup>d</sup>	56 (78)	40 (95)	16 (53)	<0.01
Patients <sup>d</sup>	38 (53)	26 (62)	12 (40)	0.11
FAIR: Findable, Accessible, Interoperable, and Reusable - How important is it that the data of the registry should be FAIR? <sup>c</sup>				
(median, IQR)	92 (81–100)	92 (81–100)	95 (82–100)	0.89

*IQR* interquartile range

<sup>a</sup>An element was considered important if  $\geq 80\%$  of the respondents gave it a score of important or very important

<sup>b</sup>See ESM, for additional figures of all answers to the Likert scales

<sup>c</sup>One missing from the respondents from industry

<sup>d</sup>One missing from the respondents from the other stakeholders

adverse events [21]. Another example is the REGIMS Registry that aims to assess the incidence, type and consequences of side effects of multiple sclerosis immunotherapies [22]. Requirements for post-approval safety data management are described in the International Council on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use guideline E2D and good pharmacovigilance practices module VI [23, 24]. Companies have the obligation for solicited cases to perform an assessment of causality and submit the causally related adverse drug reactions to the relevant authorities [25, 26]. While registries could be an important source to evaluate long-term drug effects including safety outcomes that are often incomplete at the time of drug approval, the collection of such data within most disease registries does not allow a causality assessment in line with International Council on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use guidance [27]. If registries are used as a data source for post-authorisation safety studies, it should be clear what the expectations are with respect to adverse event collection and managing follow-up information, causality assessment and, where appropriate, reporting timelines. Registries should provide accurate, timely and follow-up data on serious adverse events to enable a causality assessment.

Good quality data is crucial for a thorough benefit-risk evaluation of medicinal products. According to the International Council on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use Guidance for good clinical practice, clinical trials should be monitored to verify that reported data are accurate, complete and accounted for by source records [28]. Our survey study showed that responders considered a level of 20% acceptable for missing data. A threshold of 10% of missing data based on a sample of 200 patients of the registry data was suggested by participants of one of the EMA disease-specific registry workshops [20]. It is recognised that most patient registries will have at least some missing data. Approaches to minimise the amount of missing data should be considered as part of the registry protocol and analysis plan. However, no guidance is given on what proportion of data should be verified. It is acknowledged that even 100% source data verification does not guarantee that a 0% error rate can be achieved [29]. A risk-based approach in combination with reduced source data verification could be a good solution to verify the data [30–32]. Source data verification for 10% of the registry data were suggested by participants of one of the EMA disease-specific registry workshops [33]. This implies that the outcome of our study, laying the bar at

30% for data verification, may not be feasible in practice. The results of our survey could provide a starting point to discuss which and how much data should be verified to guarantee validity of the data to be acceptable to all potential stakeholders.

To our knowledge, this is the first study assessing on a larger scale the importance that stakeholders attach to key aspects of registries in the field of rare diseases for regulatory decision making. A limitation of the study is that of the 201 participants who received and opened the survey, 82 (41%) did not respond to any question. Forty-six participants (23%) finished after completing only a few questions and 73 (36%) completed  $\geq 80\%$  of the survey and were included in the analyses. Reasons for the drop-out are regrettably unknown, but 82 (41%) did not respond to any question. This suggests that problems related to navigating through the survey and its content are unlikely. Related to this is that it should be noted that the survey used in this study was pretested on functioning and content among a small number ( $n = 7$ ) of regulators only and that no formal validation procedures were applied. Although respondents were allowed to skip questions in the survey, for instance, if a question was unclear to the respondent or not applicable, it is possible that some questions have been interpreted differently between respondents. Additionally, the generalisability of our findings to a wider population should be assessed in future studies because the individuals included in our study probably had a particular interest in registries having participated in the EMA workshops, or were connected to the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance, and only a small and heterogenous number of other stakeholders was included. Furthermore, we used a cut-off of 80% for the responses to indicate importance, both for the Likert scale and the multiple-choice questions. Although this cut-off level has been used previously [12], this could still be considered rather arbitrary.

## 5 Conclusions

This study showed that the opinion towards data and governance is well aligned across parties, and issues of data and governance on their own should not pose a barrier to collaboration. This finding is supportive of the EMA's efforts to encourage stakeholders to work with existing registries when collecting data to support regulatory decision making.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s40264-021-01081-z>.

**Acknowledgements** We thank Carla Alonso Olmo, Anja van Haren and the members of the Patient Registry Initiative for their useful comments on the survey and/or the manuscript.

## Declarations

**Funding** No funding for the conduct of this study or the preparation of this article was received by any of the authors.

**Conflicts of interest/Competing interests** Carla J. Jonker, Sieta T. de Vries, H. Marijke van den Berg, Patricia McGettigan, Arno W. Hoes and Peter G.M. Mol have no conflicts of interest that are directly relevant to the content of this article.

**Ethics approval** Not applicable.

**Consent to participate** Not applicable.

**Consent for publication** Not applicable.

**Availability of data and material** The data that support the findings of this work are available from the corresponding author on reasonable request.

**Code availability** Not applicable.

**Authors' contributions** CJJ, STdV, PMcG and PGMM were involved in the conception and design of the study; CJJ, STdV and PGMM conducted the analyses; and all authors contributed to the interpretation of the data. All authors reviewed and edited the manuscript. All authors have read and approved the final manuscript. All authors attest they meet the International Committee of Medical Journal Editors criteria for authorship.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

## References

1. Eurodis. What is a rare disease? Last updated 21 July 2020. [https://www.eurodis.org/sites/default/files/publications/Fact\\_Sheet\\_RD.pdf](https://www.eurodis.org/sites/default/files/publications/Fact_Sheet_RD.pdf). Accessed 22 Jan 2021.
2. Fonseca DA, Amaral I, Pinto AC, Cotrim MD. Orphan drugs: major development challenges at the clinical stage. *Drug Discov Today*. 2019;24(3):867–72.
3. Cremers S, Aronson JK. Drugs for rare disorders. *Br J Clin Pharmacol*. 2017;83:1607–13.
4. Pacurariu A, Plueschke K, Olmo CA, Kurz X. Imposed registries within the European postmarketing surveillance system: extended analysis and lessons learned for regulators. *Pharmacoepidemiol Drug Saf*. 2018;27(7):823–6.
5. Jonker CJ, Kwa MSG, van den Berg HM, Hoes AW, Mol PGM. Drug registries and approval of drugs: promises, placebo, or a real success? *Clin Ther*. 2018;40(5):768–73.

6. Bouvy JC, Blake K, Slattery J, De Bruin ML, Arlett P, Kurz X. Registries in European post-marketing surveillance: a retrospective analysis of centrally approved products, 2005–2013. *Pharmacoepidemiol Drug Saf.* 2017;26(12):1442–50.
7. European Medicines Agency. Patient registries. <https://www.ema.europa.eu/en/human-regulatory/post-authorisation/patient-registries>. Accessed 22 Jan 2021.
8. McGettigan P, Alonso Olmo C, Plueschke K, Castillon M, Noguera Zondag D, Bahri P, et al. Patient registries: an under-used resource for medicines evaluation: operational proposals for increasing the use of patient registries in regulatory assessments. *Drug Saf.* 2019;42(11):1343–51.
9. Eunetha. The Registry Evaluation and Quality Standards Tool (REQueST). <https://eunetha.eu/request-tool-and-its-vision-paper>. Accessed 22 Jan 2021.
10. Kodra Y, Weinbach J, Posada-de-la-Paz M, Coi A, Lemonnier SL, van Enckevort D, et al. Recommendations for improving the quality of rare disease registries. *Int J Environ Res Public Health.* 2018;15(8):1644.
11. European Medicines Agency. Guideline on registry-based studies, draft. Dated 24 September 2020. [https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-registry-based-studies\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-registry-based-studies_en.pdf). Accessed 22 Jan 2021.
12. Brinkman DJ, Tichelaar J, Mokkink LB, Christiaens T, Likic R, Maciulaitis R, et al. Key learning outcomes for clinical pharmacology and therapeutics education in Europe: a modified Delphi study. *Clin Pharmacol Ther.* 2018;104(2):317–25.
13. Tcheng JE, Drozda JP, Gabriel D, Heath A, Wilgus RW, Williams M, et al. Achieving data liquidity: lessons learned from analysis of 38 clinical registries (the Duke-Pew Data Interoperability Project). *AMIA Annu Symp Proc.* 2019;2019:864–73.
14. European Platform on Rare Disease Registration (EU RD Platform). Set of common data elements for rare diseases registration. [https://eu-rd-platform.jrc.ec.europa.eu/sites/default/files/CDS/EU\\_RD\\_Platform\\_CDS\\_Final.pdf](https://eu-rd-platform.jrc.ec.europa.eu/sites/default/files/CDS/EU_RD_Platform_CDS_Final.pdf). Accessed 22 Jan 2021.
15. Gliklich RE, Dreyer NA, Leavy MB. Registries for evaluating patient outcomes: a user's guide. 3rd ed. 2014. <https://www.ncbi.nlm.nih.gov/books/NBK208643/>. Accessed 22 Jan 2021.
16. McGrath E, Chabannon C, Terwel S, Bonini C, Kuball J. Opportunities and challenges associated with the evaluation of chimeric antigen receptor T cells in real-life. *Curr Opin Oncol.* 2020;32(5):427–33.
17. Viviani L, Zolin A, Mehta A, Olesen HV. The European Cystic Fibrosis Society Patient Registry: valuable lessons learned on how to sustain a disease registry. *Orphanet J Rare Dis.* 2014;9:81.
18. European Medicines Agency. Qualification opinion on cellular therapy module of the European Society for Blood & Marrow Transplantation (EBMT) Registry. 2019. [https://www.ema.europa.eu/en/documents/scientific-guideline/qualification-opinion-cellular-therapy-module-european-society-blood-marrow-transplantation-ebmt\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/qualification-opinion-cellular-therapy-module-european-society-blood-marrow-transplantation-ebmt_en.pdf). Accessed 22 Jan 2021.
19. European Medicines Agency. Qualification opinion on the European Cystic Fibrosis Society Patient Registry (ECFSPR) and CF pharmaco-epidemiology studies. [https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-opinion-european-cystic-fibrosis-society-patient-registry-ecfspr-cf-pharmaco\\_en.pdf](https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-opinion-european-cystic-fibrosis-society-patient-registry-ecfspr-cf-pharmaco_en.pdf). Accessed 22 Jan 2021.
20. European Medicines Agency. Report of the workshop on the use of registries in the monitoring of cancer therapies based on tumours' genetic and molecular features. 2020. [https://www.ema.europa.eu/en/documents/report/report-workshop-use-registries-monitoring-cancer-therapies-based-tumours-genetic-molecular-features\\_en.pdf](https://www.ema.europa.eu/en/documents/report/report-workshop-use-registries-monitoring-cancer-therapies-based-tumours-genetic-molecular-features_en.pdf). Accessed 22 Jan 2021.
21. Bosma AL, Spuls PI, Garcia-Doval I, Naldi L, Prieto-Merino D, Tesch F, et al. TREATment of ATopic eczema (TREAT) Registry Taskforce: protocol for a European safety study of dupilumab and other systemic therapies in patients with atopic eczema. *Br J Dermatol.* 2020;182(6):1423–9.
22. Simbrich A, Thibaut J, Khil L, Maximov S, Wiendl H, Berger K, REGIMS Investigators. Chances and challenges of registry-based pharmacovigilance in multiple sclerosis: lessons learnt from the implementation of the multicenter REGIMS Registry. *Drug Saf.* 2021;44(1):7–15.
23. European Medicines Agency. ICH topic E 2 D post approval safety data management. 2004. [https://www.ema.europa.eu/en/documents/scientific-guideline/international-conference-harmonisation-technical-requirements-registration-pharmaceuticals-human-use\\_en-12.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/international-conference-harmonisation-technical-requirements-registration-pharmaceuticals-human-use_en-12.pdf). Accessed 22 Jan 2021.
24. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP). Module VI: collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2). 2017. [https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-good-pharmacovigilance-practices-gvp-module-vi-collection-management-submission-reports\\_en.pdf](https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-good-pharmacovigilance-practices-gvp-module-vi-collection-management-submission-reports_en.pdf). Accessed 22 Jan 2021.
25. Official Journal of the European Union. Directive 2010/84/EU of the European Parliament and of the Council. 2010. Available from: <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:348:0074:0099:EN:PDF>. [Accessed 22 Jan 2021].
26. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP). Module VIII: post-authorisation safety studies (Rev 3). 2017. [https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-module-viii-post-authorisation-safety-studies-rev-3\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-module-viii-post-authorisation-safety-studies-rev-3_en.pdf). Accessed 22 Jan 2021.
27. European Medicines Agency. Guideline for good clinical practice E6(R2). 2016. [https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-6-r2-guideline-good-clinical-practice-step-5\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-6-r2-guideline-good-clinical-practice-step-5_en.pdf). Accessed 22 Jan 2021.
28. European Medicines Agency. Report on CAR T-cell therapy registries. Workshop 9 February 2018. [https://www.ema.europa.eu/documents/report/report-car-t-cell-therapy-registries-workshop\\_en.pdf](https://www.ema.europa.eu/documents/report/report-car-t-cell-therapy-registries-workshop_en.pdf). Accessed 4 May 2021.
29. Andersen JR, Byrjalsen I, Bihlet A, Kalakou F, Hoeck HC, Hansen G, et al. Impact of source data verification on data quality in clinical trials: an empirical post hoc analysis of three phase 3 randomized clinical trials. *Br J Clin Pharmacol.* 2015;79(4):660–8.
30. Olsen R, Bihlet AR, Kalakou F, Andersen JR. The impact of clinical trial monitoring approaches on data integrity and cost: a review of current literature. *Eur J Clin Pharmacol.* 2016;72(4):399–412.
31. European Medicines Agency. Reflection paper on risk based quality management in clinical trials. 2013. [https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-risk-based-quality-management-clinical-trials\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-risk-based-quality-management-clinical-trials_en.pdf). Accessed 22 Jan 2021.
32. European Commission. Risk proportionate approaches in clinical trials. 2017. [https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/2017\\_04\\_25\\_risk\\_proportionate\\_approaches\\_in\\_ct.pdf](https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/2017_04_25_risk_proportionate_approaches_in_ct.pdf). Accessed 22 Jan 2021.
33. European Medicines Agency. Report on cystic fibrosis registries: workshop 14 June 2017. 2017. [https://www.ema.europa.eu/en/documents/report/report-cystic-fibrosis-registries\\_en.pdf](https://www.ema.europa.eu/en/documents/report/report-cystic-fibrosis-registries_en.pdf). Accessed 22 Jan 2021.