



Factors Influencing the Generation of Evidence from Simple Data Held in International Rare Disease Patient Registries

Ravi Jandhyala¹ · Solomon Christopher¹

Published online: 5 December 2019
© The Author(s) 2019

Abstract

Background Rare diseases (defined as affecting < 1 in 2000 Europeans) may collectively affect up to approximately 8% of the population. The low prevalence of individual diseases limits patient studies and data collection is a key challenge; international rare disease patient registries are essential for optimal data collection and research. Registry data achieves value when research conducted on them are published—this is termed evidence generation.

Objective The aim of this study was to examine selected factors and their association with evidence generation, via scientific publication, from international rare disease patient registry data.

Methods All international rare disease patient registries listed in the Orphanet 2018 report were analysed. Rates of scientific publications were compared by funding stream, disease area and registry size using multivariable regression analyses. Publication characteristics, such as novelty of findings, were also compared by registry funding stream, disease area and duration of operation.

Results Privately funded registries had approximately two to four times higher rates of scientific publication compared with publically funded registries, with adjusted rate ratios of 1.85 (95% confidence interval [CI] 1.07–3.22) and 4.18 (95% CI 2.54–6.87) for private not-for-profit and private for-profit funding, respectively. The inclusion of outcomes, use of pharmaceutical medicines, novel findings and citation rate for publications generated from patient registries with any private funding was not significantly different from those attributed to only publicly funded registries.

Conclusion The results of this study indicate that privately funded international rare disease patient registries produce significantly more evidence than their publicly funded counterparts. Examination of the quality indicators of these publications showed they were of the same high quality as those generated from publicly funded patient registry data.

Key Points

Privately funded international rare disease patient registries generate more evidence (via scientific publications) than their publicly funded counterparts.

Privately funded registry publications are of the same quality as those from publicly funded registries.

1 Introduction

Individually, rare diseases affect fewer than 1 in 2000 Europeans, yet collectively they have been thought to affect as much as 8% of the population [1]. However, more recent estimates place the total burden of rare disease at only approximately 5.9% of the worldwide population; this correlates to around 446 million patients worldwide and 30 million patients across Europe [2].

Rare diseases can be difficult to study since the low prevalence of each disease often restricts studies of disease course and treatment response, and may also limit fundraising efforts for the study of these conditions. However, these challenges can be overcome with well-designed and supported rare disease patient registries (RDRs).

A patient registry is a prospective and/or retrospective observational, or ‘non-interventional’, cohort clinical study that can be local or international in scale [3, 4]. The collected data can include patient demographics, medical and treatment history, clinical measurements, and results for

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s40290-019-00316-w>) contains supplementary material, which is available to authorized users.

✉ Ravi Jandhyala
ravi@medialis.co.uk

¹ Medialis Ltd, 13 Horse Fair, Banbury, Oxford OX16 0AH, UK

biochemical/radiological tests [5]. International RDRs promote collaboration, facilitate standardisation of information and allow more powerful studies to be conducted on the acquired data than smaller, local registries [6, 7].

International RDRs are important and cost-effective resources for rare disease communities, collecting data and generating publications that complement the results of clinical trials and reflect actual clinical practice [1, 8]. These publications help regulators further evaluate the safety and efficacy of therapies, support continued reimbursement by payors, and contribute to the development of novel therapeutic approaches, collectively improving patients' quality of life [8, 9]. Furthermore, the publication of studies arising from an RDR raises awareness of the condition, which can contribute to future fundraising efforts that support all aspects of the disease community [5].

Patient registries can be publicly or privately funded, either on a for-profit or not-for-profit basis. Public patient registries are funded by government bodies, such as the Medical Research Council (MRC) in the UK, the European Research Council (ERC) in Europe, and the National Institutes of Health (NIH) in America, while private patient registries are funded by non-government sources, including philanthropic organisations, independently wealthy individuals and pharmaceutical companies. A recent review of patient registries described nine different funding models and concluded that the choice of funding model should be dictated by the needs of the registry [10], yet the relative benefits of different funding models for patient registries have only recently been studied; RDRs have not been specifically assessed [4].

The design of future registries could be guided by assessing and comparing the effects of funding model on the impact of existing registries. The transition from simple 'data' held in an RDR into 'evidence' that can inform clinical management decisions relies on the publication of good-quality research conducted on the registry population. Cumulative publication count is therefore a good indicator of the impact of a rare disease registry and can be used to compare the influence of public and private funding streams on the overall impact of RDRs. This study explores the transition of simple data entered in RDRs into published evidence via the analysis of publication count (as a marker of evidence generation) alongside publication quality, measured via inclusion of outcomes, pharmaceutical medicine inclusion, novelty of findings and citation rate.

2 Methods

2.1 Study Design and Data Sources

All international RDRs listed in the Orphanet 'Rare Disease Registries in Europe' report (May 2018) were included in

this study [11]. Patient registries were excluded from subsequent analyses if the duration of operation (start and end dates) could not be confirmed. Data were collected for all included patient registries up to December 2018. Further data were collected from other primary sources, specifically the registry owner, ClinicalTrials.gov and the registry's own website.

Registry-related publications were identified and quantified by searching for the included patient registry names in MEDLINE via PubMed (31 December 2018), before being analysed via full-text reviews to assess whether the results reported were unique or confirmed previously reported data. Publications were excluded from the analysis if they did not report on findings of studies conducted either wholly or in part on data contained directly within their respective patient registry.

2.2 Study Variables

The number of publications reporting data from the patient registry population was defined as an indicator of evidence generation. As both registry size and duration of registry operation are intrinsic factors influencing the amount of data accumulated, these were incorporated into the analyses. Duration of operation was defined as the duration between the registry start and end dates. For the registries that have ongoing operations, all information was censored on 31 December 2018. The main outcome of interest was the publication rate, or rate of evidence generation, defined as the number of publications over the period of registry operation.

To investigate the influence of funding stream on patient registry publication count, the main covariate was the type of funding for the operation of the registry (public, private for-profit, private not-for-profit, and mixed). Additional covariates of interest were disease/therapeutic area (genetic, congenital, oncology, other) and the target size of the registry (< 500, 500–1500, 1500–2500, ≥ 2500); the distinction between genetic and congenital disease/therapeutic areas was made based on whether there was a hereditary component to the disease/therapeutic area.

For publication-level analysis, several outcome variables were used in order to assess publication quality. In addition to the citation rate, which was calculated as the citation count from the date of publication to 31 December 2018, other factors were chosen for their informative nature for numerous stakeholders. Indeed, from empirical experience, the inclusion of patient outcomes provides important information about disease progression, both on and off treatment; this is a particularly relevant outcome for payor authorities. The inclusion of pharmaceutical medicines provides relevant information for regulatory bodies, and the reporting of new findings in the publication is valuable to both the prescriber and patients. The covariates were registry-level

characteristics, including funding, disease/therapeutic area and duration of operation.

2.3 Statistical Analyses

2.3.1 Registry-Level Analyses

The response variable was publication rate (number of publications containing data from a given patient registry) over the duration of operation of the patient registry. For the covariates hypothesised to be associated with the response, multivariable regression analysis was performed.

The excessive number of zeroes in the publication with registry data field gave rise to overdispersion. To account for this, in addition to fitting a Poisson regression, zero-inflated Poisson, negative binomial and zero-inflated negative binomial regression models were also fitted. Results are presented with data summaries (number and percentage of registries in each category, total number of publications and total duration of registry operation for each covariate category). The adjusted rate ratios (RRs) and their corresponding 95% confidence intervals (CIs) are also presented.

For the zero-inflated models, the model for the components (mixing model for zero inflation) does not include any covariates due to an inadequate number of registries for estimating the effect of the covariates in the mixing model. Mediation by an intermediate variable was empirically checked by specific models defining the pathways of supposed associations. A proportional odds logistic regression was used to assess the association between registry size and funding stream.

2.3.2 Publication-Level Analyses

Generalised linear mixed models were fitted with registries as random effects and covariates as fixed effects. Logistic mixed-effects models were used for the presence of outcomes, pharmaceutical medicines and new findings in publications; adjusted odds ratios are presented along with their 95% CI. A Poisson mixed-effects model was used for citation rate; adjusted rate ratios are presented along with their 95% CI.

2.3.3 Implementation of Statistical Methods

All statistical analyses were performed using R version 3.5.2 (The R Foundation for Statistical Computing, Vienna, Austria). Zero-inflated regressions were performed using the `pscl` library routines, while the proportional odds logistic regression was performed using the `MASS` library routines. The generalised linear mixed models were fitted using `lme4` library routines.

3 Results

3.1 Registry-Level Analyses

A total of 83 international RDRs from the Orphanet report were examined [11]. Table 1 provides the attrition in the data used for statistical analysis due to missing information in the collected variables. The number of publications reporting results on studies conducted on a patient's registry data was positively skewed with a large number of zeroes (Fig. 1). Of the 83 international RDRs included in this study, 26 (31%) were completely publicly funded, while of the 51 (61%) RDRs that had some form of private funding, 19 (23%) had private for-profit funding, 18 (22%) had private not-for-profit funding and 14 (17%) had mixed funding. For 6 (7%) RDRs, the correct funding stream could not be ascertained. Of all 83 RDRs, 33 (40%) did not have any publications, indicating that these RDRs have not contributed towards the evidence base.

Table 2 presents the adjusted rate ratios for the covariates (funding and disease area of the registry) using Poisson, zero-inflated Poisson, negative binomial and zero-inflated negative binomial regressions. When compared with registries with public funding, registries with private for-profit funding have an approximately four times higher rate of publication (RR 4.18, 95% CI 2.54–6.87 from the Poisson model), which was consistent across all four models used. Registries with private not-for-profit funding have an approximately two times higher rate of publication, however this effect was neither consistent nor statistically significant across the four models.

The disease/therapeutic area (genetic, congenital, oncology and other) did not have any significant association with the publication rate; however, this could be due to the small number of registries in the congenital and oncology categories.

In addition to funding and disease area, Table 3 includes the target registry size as a covariate, and details their adjusted rate ratios across all four described models. In this model, increased registry size (500–1500, 1500–2500, > 2500 vs < 500) is significantly associated with higher publication rates (RR 2.41, 95% CI 1.15–5.07; RR

Table 1 Attrition due to missing variable information

Attrition	Dataset	Sensitivity dataset
Total number of registries	83	83
Valid duration of operation	58	67
Valid number of publications	58	67
Valid therapeutic area	54	63
Valid registry size	36	41

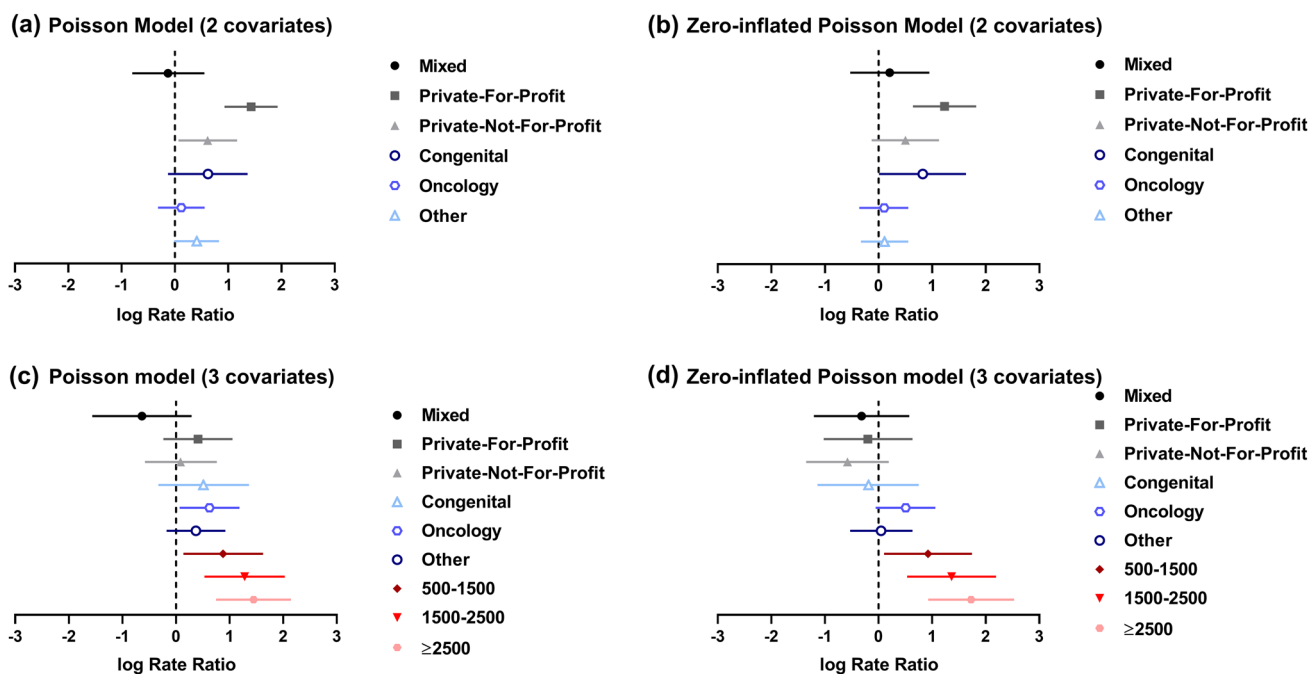


Fig. 1 Forest plots displaying the adjusted log-rate ratios of selected covariates. Estimates were obtained using Poisson and zero-inflated Poisson models. The symbols represent the adjusted log-rate ratios and the horizontal lines represent their corresponding 95% confidence intervals.

ence intervals. **a, b** Model estimates with funding and disease area as covariates. **c, d** Model estimates with funding, disease area and registry as covariates. The vertical dashed line is at value 0, which represents the null value for log-rate ratios

Table 2 Adjusted rate ratios and their 95% confidence intervals for variables associated with number of publications with registry data using Poisson, zero-inflated Poisson, negative binomial and zero-

inflated negative binomial regression models. Covariates include funding and disease area

Covariates	Categories	Summary		Adjusted rate ratio (95% CI)			
		N (%)	Events (p-y)	Poisson	ZIP	Negative Binomial	ZINB
Funding	Public	17 (32.08)	31 (146.96)	Reference	Reference	Reference	Reference
	Mixed	7 (13.21)	16 (86.87)	0.88 (0.45–1.74)	1.23 (0.59–2.58)	0.74 (0.20–2.75)	0.81 (0.23–2.83)
	Private for-profit	16 (30.19)	187 (283.52)	4.18 (2.54–6.87)	3.43 (1.90–6.18)	4.50 (1.65–12.27)	4.39 (1.66–11.63)
	Private not-for-profit	13 (24.53)	41 (128.26)	1.85 (1.07–3.22)	1.65 (0.88–3.09)	2.77 (0.97–7.89)	2.69 (1.00–7.24)
Disease area	Genetic	32 (60.38)	201 (417.92)	Reference	Reference	Reference	Reference
	Congenital	3 (5.66)	12 (41.92)	1.86 (0.88–3.91)	2.28 (1.01–5.12)	2.93 (0.59–14.58)	3.25 (0.65–16.34)
	Oncology	6 (11.32)	24 (61.87)	1.13 (0.73–1.75)	1.11 (0.70–1.74)	0.69 (0.20–2.32)	0.71 (0.25–2.01)
	Other	12 (22.64)	38 (123.90)	1.51 (0.99–2.29)	1.12 (0.72–1.74)	2.11 (0.82–5.45)	1.89 (0.73–4.90)
Total		53 (100.00)	275 (645.61)				

CI confidence interval, ZIP zero-inflated Poisson, ZINB zero-inflated negative binomial, p-y person-years

3.60, 95% CI 1.70–762; and RR 4.25, 95% CI 2.11–8.55, respectively); these values are based on the Poisson regression model. However, the association with funding stream is reduced in this model. The estimates and 95% CI of the adjusted rate ratios from Tables 2 and 3 are presented on a log scale using the forest plots seen in Fig. 1.

Electronic supplementary Tables 1 and 2 present the same analysis performed on the data found in Tables 2 and 3,

however the analysis has been conducted for the sensitivity dataset (registries with missing end dates of operation are assumed to be ongoing). Overall, the estimates from electronic supplementary Tables 1 and 2 are similar in direction and size to their corresponding estimates from Tables 2 and 3; a larger number of observations in the sensitivity dataset may have led to small changes in the estimates and/or widths of the 95% CI.

Table 3 Adjusted rate ratios and their 95% confidence interval for variables associated with number of publications with registry data using Poisson, zero-inflated Poisson, negative binomial and zero-inflated negative binomial regression models. Covariates include funding, disease area and registry size

Covariates	Categories	Summary		Adjusted rate ratio (95% CI)			
		N (%)	Events (p-y)	Poisson	ZIP	Negative Binomial	ZINB
Funding	Public	10 (27.78)	25 (81.13)	Reference	Reference	Reference	Reference
	Mixed	3 (8.33)	13 (47.71)	0.53 (0.21–1.34)	0.73 (0.30–1.78)	0.44 (0.06–3.00)	0.60 (0.11–3.23)
	Private for-profit	14 (38.89)	170 (264.62)	1.51 (0.79–2.87)	0.82 (0.36–1.89)	1.40 (0.41–4.74)	1.15 (0.30–4.43)
	Private not-for-profit	9 (25.00)	36 (89.05)	1.09 (0.56–2.14)	0.56 (0.26–1.21)	1.29 (0.36–4.61)	0.97 (0.27–3.51)
Disease area	Genetic	24 (66.67)	180 (347.85)	Reference	Reference	Reference	Reference
	Congenital	2 (5.56)	12 (22.42)	1.67 (0.72–3.90)	0.83 (0.32–2.12)	1.35 (0.21–8.88)	1.00 (0.18–5.42)
	Oncology	1 (2.78)	19 (17.08)	1.87 (1.07–3.27)	1.66 (0.95–2.89)	1.67 (0.19–14.59)	1.44 (0.26–7.98)
	Other	9 (25.00)	33 (95.16)	1.45 (0.84–2.51)	1.05 (0.59–1.89)	1.89 (0.62–5.79)	1.65 (0.62–4.40)
Registry size	< 500	9 (25.00)	10 (68.98)	Reference	Reference	Reference	Reference
	500–1500	8 (22.22)	35 (102.85)	2.41 (1.15–5.07)	2.52 (1.11–5.72)	3.41 (0.95–12.20)	3.01 (0.97–9.31)
	1500–2500	10 (27.78)	70 (122.78)	3.60 (1.70–7.62)	3.92 (1.71–8.98)	5.02 (1.33–19.01)	5.34 (1.64–17.44)
	> 2500	9 (25.00)	129 (187.90)	4.25 (2.11–8.55)	5.64 (2.53–12.56)	4.54 (1.30–15.88)	5.15 (1.52–17.43)
Total		36 (100.00)	244 (482.51)				

CI confidence interval, ZIP zero-inflated Poisson, ZINB zero-inflated negative binomial, p-y person-years

In order to fully understand the process of associations from Tables 2 and 3, an empirical check was performed to determine whether the effect of funding on publication rate is mediated by target registry size as an intermediate variable; the results from this mediation analysis are presented in Table 4. Three models were fitted to the data with non-missing values for all variables involved (36 data points): Model 1 (publication ~ funding), Model 2 (publication ~ funding + registry size) and Model 3 (registry size ~ funding). The estimates from the three models suggest a mediation effect according to the causal effect criteria [12]. Due to an inadequate number of data points and the non-Gaussian distributional assumptions of the variables involved, no formal methods were further employed to quantify the mediation effects.

3.2 Publication-Level Analyses

Publication data from included patient registries with at least one publication were assessed, and a total of 276 publications from 50 registries were included in the statistical analyses. The number of publications per registry utilising registry data ranged from 1 to 62, with a median of three publications. Table 5 presents the adjusted risk ratios for the covariates funding, disease/therapeutic area and duration of operation of the registry on publications to include outcomes, pharmaceutical drugs and new findings, along with their citation rate. The odds for publications to include outcomes are between two and four times higher for registries with any private funding (private for-profit,

Table 4 Mediation analysis

Covariates	Categories	Risk ratio	95% CI
Model 1: Funding	Publications ~ funding		
	Public	Reference	Reference
	Mixed	0.67	0.35–1.25
	Private for-profit	2.62	1.75–4.08
Model 2: Funding	Publications ~ funding + registry size		
	Public	Reference	Reference
	Mixed	0.56	0.29–1.05
	Private for-profit	1.28	0.82–2.08
Registry size	Private not-for-profit	0.96	0.57–1.65
	< 500	Reference	Reference
	500–1500	3.14	1.68–6.24
	1500–2500	5.29	2.93–10.28
Model 3: Funding	Registry size ~ funding		
	Public	Reference	Reference
	Mixed	1.80	0.25–12.60
	Private for-profit	9.64	2.05–51.90
	Private not-for-profit	4.29	0.84–24.00

CI confidence interval

private not-for-profit and mixed) in comparison with public funding only; the adjusted odds ratios after adjusting for disease area and duration of registry operation are 4.31 (95% CI 0.92–20.14), 2.44 (95% CI 0.63–9.38) and 4.43 (95% CI 0.60–32.50), respectively.

The odds for publications to include pharmaceutical drugs is at least two times higher for registries with any private funding (private for-profit, private not-for-profit and mixed) in comparison with public funding only. The adjusted odds ratios after adjusting for disease area and duration of registry operation are 6.76 (95% CI 0.55–83.44), 2.23 (95% CI 0.17–28.51) and 17.04 (95% CI 1.24–233.38), respectively.

The odds for publications to include new findings are not significantly different for registries with private funding (private for-profit, private not-for-profit and mixed) in comparison with public funding only. The adjusted odds ratios after adjusting for disease area and duration of registry operation are 1.50 (95% CI 0.19–11.90), 2.23 (95% CI 0.41–12.23) and 1.14 (95% CI 0.12–10.58), respectively.

The citation rate is not significantly different for registries with private funding (private for-profit, private not-for-profit and mixed) in comparison with public funding only. The adjusted rate ratios after adjusting for disease area and duration of registry operation are 0.78 (95% CI 0.39–1.55), 0.67 (95% CI 0.35–1.28) and 0.78 (95% CI 0.34–1.80), respectively.

The estimates and 95% CI of the adjusted risk ratios from Table 5 for funding stream on the described quality indicators are presented on a log scale using the forest plots in Fig. 2.

4 Discussion

RDRs play an important role in monitoring disease progression and outcomes of therapeutic interventions in a ‘real-world’ population. The impact of international RDRs

is driven by the amount of evidence generated and subsequently published in academic journals. Once in the public domain, a range of stakeholders are able to benefit from this knowledge, including payor authorities, regulatory bodies, patient groups and clinicians. In turn, this value increases the profiles of both the initial registry and the disease area [5, 13].

In order to account for the large number of RDRs with no publication results, zero-inflated models were applied. These are mixture models and are specified in two parts: a binary model for whether an RDR is amenable to generating evidence from data entered in it, and a count model for the research output in terms of number of publications. Due to the small number of RDRs included in the analysis, the binary model did not include any covariates.

The results show that, in general, privately funded RDRs have a higher rate of scientific publication and consequent evidence generation in comparison with publicly funded RDRs. In particular, private for-profit funding is associated with about a four times higher rate of publication in comparison with publicly funded international RDRs.

Additionally, the results show that the effect of funding on publication rate is mediated through the size of the registry. Private funding leads to larger-sized registries that can accumulate a greater volume of patient data, which, in turn, enables a threshold quality of research to be conducted on them, transitioning those data into evidence. From the descriptive summaries of the data analysed, it seems that publication rates may also be influenced by disease area (genetic, congenital, oncological and other). However, due to the small number of registries with non-missing data for these variables, this association cannot be appropriately assessed in this study.

Table 5 Adjusted odds ratios and their 95% confidence interval for variables associated with use of outcomes, use of pharmaceutical drugs, new findings in publications and citation counts using gener-

alized linear mixed models. Fixed-effects covariates include funding, disease area and duration of operation

Covariates	Categories	Outcomes (194/275)	Pharmaceutical drugs (106/275)	New findings (231/275)	Citations (276)
Funding	Public	Reference	Reference	Reference	Reference
	Mixed	4.43 (0.60–32.50)	17.04 (1.24–233.38)	1.14 (0.12–10.58)	0.78 (0.34–1.80)
	Private for-profit	4.31 (0.92–20.14)	6.76 (0.55–83.44)	1.50 (0.19–11.90)	0.78 (0.39–1.55)
	Private not-for-profit	2.44 (0.63–9.38)	2.23 (0.17–28.51)	2.23 (0.41–12.23)	0.67 (0.35–1.28)
Disease area	Genetic	Reference	Reference	Reference	Reference
	Congenital	1.00 (0.13–7.57)	3.37 (0.16–69.61)	0.53 (0.02–15.62)	2.12 (0.73–6.12)
	Oncology	1.09 (0.32–3.73)	1.41 (0.52–3.79)	0.30 (0.04–2.09)	0.68 (0.28–1.66)
	Other	3.93 (1.06–14.59)	0.20 (0.02–2.16)	2.28 (0.47–11.12)	1.14 (0.64–2.05)
Operation duration, years	< 7	Reference	Reference	Reference	Reference
	7–10	4.37 (0.83–22.99)	NE	55.55 (2.70–1141.85)	0.99 (0.46–2.13)
	10–16	1.49 (0.42–5.35)	NE	3.60 (0.72–18.00)	0.88 (0.46–1.71)
	> 16	1.65 (0.36–7.54)	NE	7.63 (0.84–69.62)	1.15 (0.52–2.55)

CI confidence interval, NE not estimated

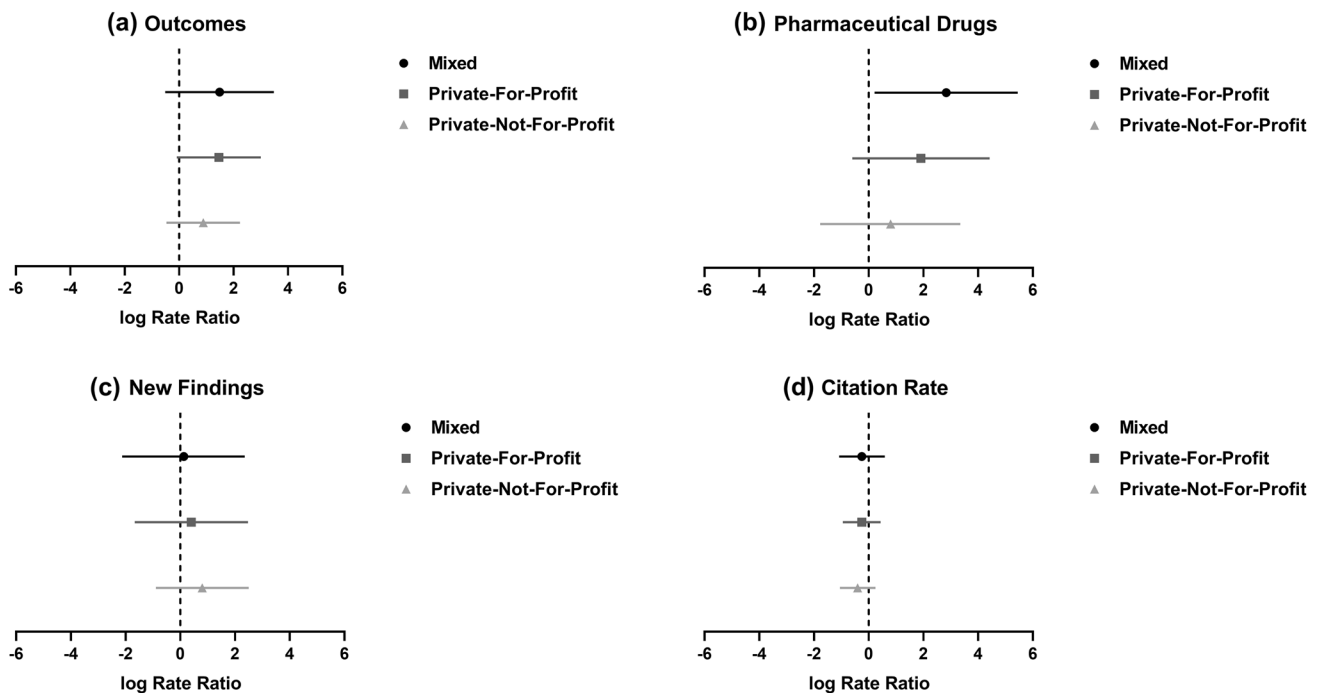


Fig. 2 Forest plots displaying the adjusted log-risk ratios for three private funding streams against only public funding for registries against four selected characteristics of registry publications. The estimates were obtained using generalised linear mixed models. The symbols represent the adjusted log-risk ratios and the horizontal lines

represent their corresponding 95% confidence intervals. **a–c** Log odds ratios for publications to include outcomes, pharmaceutical drugs and new findings. **d** Presents the log-rate ratios for citation rates. The vertical dashed line is at value 0, which represents the null value for log-risk ratios

From the publication-level analyses, the results showed that patient registries with any private funding (private for-profit, private not-for-profit and mixed) appeared more likely to include outcomes, pharmaceutical medicine data and new findings in associated publications than registries with only public funding, albeit no statistical significance was observed. This indicates that the quality of publications from registries with any private funding is at least as good as those from only publicly funded registries. In conjunction with the registry-level results, registries with private-funding had higher publication rates and the publication quality was found to be as good as, if not better than, publicly funded registries. The citation rates for publications from registries with any private funding were not significantly different from those attributed to only publicly funded registries.

The results of this study indicate that private for-profit registries may produce significantly more publications, and therefore evidence, than their publicly funded counterparts. However, industry funding of research projects has been recognised as a potential conflict of interest, and private for-profit registries are often viewed cautiously due to their dual obligations [14]; debate continues over the risk of industry-sponsored studies including market-driven research practices to promote therapeutic interventions versus the benefit of them being more rigorous due to higher

levels of resource and scrutiny [15]. However, a number of the concerns surrounding industry-sponsored research are being addressed, by increased clarity in financial sponsorship, improved reporting of conflicts of interest, increased publication of negative results and the running of larger and better-designed clinical trials [16]; compliance with transparency requirements for pharmaceutical industry-sponsored studies are also mandated by pharmaceutical industry codes of practice [17, 18]. Furthermore, it has also been shown that government-funded studies are significantly more likely to be missing data about study design and intervention, suggesting that the concerns around industry-sponsored studies may also be attributable to publicly funded research [19]. The interests of sponsors are also often aligned with those of the patients. For example, increased disease awareness generated by a registry may increase diagnosis rates, allowing more patients to receive appropriate treatment, which will also increase the market share of an industry partner distributing effective medicines. A ‘diversity of purpose’ in private for-profit registries has also recently been proposed to demonstrate the importance of their contribution to the real-world safety profile of therapeutic approaches [4].

There are some potential limitations of this study. This analysis utilised the Orphanet report to generate the list of patient registries of interest. The rationale for this was based

on the assumption that the report provided a complete listing of RDRs; should this not be the case, the results of this study may be skewed, particularly if a particular funding stream is found to be inappropriately under- or overrepresented. Furthermore, the assessment of publication findings was limited in scope solely to ascertain the uniqueness of each publication in relation to previous literature, not the level to which the findings have/will inform clinical decision making. This assessment would require clinicians with detailed experience of the management of patients and a degree of consensus among them, determined by existing practices, putting it beyond the scope of the current research project.

5 Conclusion

This study has highlighted a number of potential benefits, both to the rare disease evidence base and the scientific community, of private-funding streams, realised through facilitating the generation of a greater volume of evidence without compromising in quality when compared with publicly funded registries.

Acknowledgements Medical writing support for this manuscript was provided by Alexander T. Hardy and Cecile Limousin of Medialis Ltd. Christina Leal of Medialis Ltd contributed to data collection. Jack Lambshead and Áine Dillion contributed to study design and data collection for a previous version of this study.

Compliance with Ethical Standards

Funding No funding was received for the conduct or publication of this study. The open access fee was paid by Medialis Ltd.

Conflict of Interest RJ is Managing Director and SC is an employee of Medialis Ltd, a Medical Affairs consultancy that consults on clinical research as a paid service.

Ethical Approval As this study was a qualitative analysis of a public database, no ethical approval was required.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Julkowska D, Austin CP, Cutillo CM, Gancberg D, Hager C, Halftermeyer J, et al. The importance of international collaboration for rare diseases research: a European perspective. *Gene Ther.* 2017;24:562–71.
- Nguengang Wakap S, Lambert DM, Olry A, Rodwell C, Gueydan C, Lanneau V, et al. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *Eur J Hum Genet.* <https://doi.org/10.1038/s41431-019-0508-0>. (epub 16 Sep 2019).
- Jandhyala R. Comment on: “Patient registries: an underused resource for medicines evaluation: operational proposals for increasing the use of patient registries in regulatory assessments”. *Drug Saf.* 2019;42(12):1515–16.
- Travers K, Sallum RH, Burns MD, Barr CE, Beattie MS, Pashos CL, et al. Characteristics and temporal trends in patient registries: focus on the life sciences industry, 1981–2012. *Pharmacoepidemiol Drug Saf.* 2015;24:389–98.
- D’Agnolo HM, Kievit W, Andrade RJ, Karlsen TH, Wedemeyer H, Drenth JP. Creating an effective clinical registry for rare diseases. *United Eur Gastroenterol J.* 2016;4:333–8.
- Summar ML, Endo F, Kölker S. On the creation, utility and sustaining of rare diseases research networks: lessons learned from the Urea Cycle Disorders Consortium, the Japanese Urea Cycle Disorders Consortium and the European Registry and Network for Intoxication Type Metabolic Diseases. *Mol Genet Metab.* 2014;113:105–8.
- Baldovino S, Moliner AM, Taruscio D, Daina E, Roccatello D. Rare diseases in Europe: from a wide to a local perspective. *Isr Med Assoc J.* 2016;18:359–63.
- European Medicines Agency. Patient Registries Workshop: observations and recommendations arising from the workshop. London: European Medicines Agency; 2017.
- Committee on Comparative Effectiveness Research Prioritization, Institute of Medicine. Initial National Priorities for Comparative Effectiveness Research. Washington, DC: The National Academies Press; 2009.
- Gliklich RE, Dreyer NA, Leavy MB, editors. Registries for evaluating patient outcomes: a user’s guide. 3rd ed. Rockville: Agency for Healthcare Research and Quality; 2014.
- Orphanet. Rare Disease Registries in Europe 2019.
- MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. *Annu Rev Psychol.* 2007;58:593–614.
- Aljurf M, Rizzo JD, Mohty M, Hussain F, Madrigal A, Pasquini MC, et al. Challenges and opportunities for HSCT outcome registries: perspective from international HSCT registries experts. *Bone Marrow Transplant.* 2014;49:1016–21.
- Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ.* 2003;326:1167–70.
- Goswami ND, Tsalik EL, Naggie S, Miller WC, Horton JR, Pfeiffer CD, et al. A cross-sectional analysis of HIV and hepatitis C clinical trials 2007 to 2010: the relationship between industry sponsorship and randomized study design. *Trials.* 2014;15:31.
- Ivanov A, Kaczowska BA, Khan SA, Ho J, Tavakol M, Prasad A, et al. Review and analysis of publication trends over three decades in three high impact medicine journals. *PloS One.* 2017;12:e0170056.
- European Federation of Pharmaceutical Industries and Associations (EFPIA). EFPIA code on the promotion of prescription-only medicine to, and interactions with, healthcare professionals. Brussels: EFPIA; 2014.
- Prescription Medicines Code of Practice Authority (PMCPA). Code of practice for the pharmaceutical industry. London: PMCPA; 2016.
- Roumiantseva D, Carini S, Sim I, Wagner TH. Sponsorship and design characteristics of trials registered in ClinicalTrials.gov. *Contemp Clin Trials.* 2013;34:348–55.