

ORIGINAL ARTICLE

Retrieval of individual patient data depended on study characteristics: a randomized controlled trial

Areti Angeliki Veroniki^{a,b,c,*}, Huda M. Ashoor^b, Susan P.C. Le^b, Patricia Rios^b, Lesley A. Stewart^d, Mike Clarke^e, Dimitris Mavridis^{a,f}, Sharon E. Straus^{b,g}, Andrea C. Tricco^{b,h}

^aDepartment of Primary Education, School of Education, University of Ioannina, Ioannina, Greece

^bKnowledge Translation Program, Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Ontario, Canada

^cDepartment of Surgery & Cancer, Institute of Reproductive and Developmental Biology, Faculty of Medicine, Imperial College, London, UK

^dCentre for Reviews and Dissemination, University of York, York, UK

^eCentre for Public Health, Northern Ireland Hub for Trials Methodology Research, Queen's University Belfast, Belfast, UK

^fParis Descartes University, Sorbonne Paris Cité, Faculté de Médecine, Paris, France

^gDepartment of Geriatric Medicine, University of Toronto, Toronto, Ontario, Canada

^hEpidemiology Division, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada

Accepted 27 May 2019; Published online 30 May 2019

Abstract

Objectives: The aim of the study was to examine the effect of providing a financial incentive to authors of randomized clinical trials (RCTs) to obtain individual patient data (IPD).

Study Design and Setting: Parallel-group RCT with authors identified in the RCTs eligible for two systematic reviews. The authors were randomly allocated to the intervention (financial incentive with several contact approaches) or control group (using the same contact approaches). Studied outcomes are proportion of authors who provided IPD, time to obtain IPD, and completeness of IPD received.

Results: Of the 129 authors contacted, 37 authors suggested or contacted a person or funder providing relevant details or showed interest to collaborate, whereas 45 authors directed us to contact a person or funder, lacked resources or time, did not have ownership or approval to share the IPD, or claimed IPD was too old. None of the authors shared their IPD. We contacted 17 sponsors and received two complete IPD datasets from one sponsor. The time to obtain IPD was > 1 year after a sponsor's positive response. Common barriers included study identification, data ownership, limited data access, and required IPD licenses.

Conclusion: IPD sharing may depend on study characteristics, including funding type, study size, study risk of bias, and treatment effect, but not on providing a financial incentive. © 2019 Elsevier Inc. All rights reserved.

Keywords: Meta-analysis; Patient-level data; Individual participant data; Incentive; Data retrieval; Data sharing

Trial registration: Clinical Trials.gov (NCT02569411), registered on October 5, 2015.

Ethics approval: Ethical approval was obtained from the Research Ethics Board of St. Michael's Hospital on September 16, 2015, to conduct this randomized controlled trial (RCT; REB #15-240). The information generated during our RCT was kept confidential and limited to the study's purposes, as described in the protocol. The authors received anonymized IPD only as per their initial request, where each patient was linked to a specific identifier.

Consent for publication: Not applicable.

Availability of data and material: The full dataset is available from the corresponding author on reasonable request.

Funding: This work was supported by the Canadian Institutes of Health Research Knowledge Synthesis (No. 351143). A.A.V. was previously funded by the Canadian Institutes of Health Research (CIHR) Banting

Postdoctoral Fellowship Program (No. 139157). A.A.V. and D.M. are funded from the European Union's Horizon 2020 (No. 754936). S.E.S. is funded by a Tier 1 Canada Research Chair in Knowledge Translation. A.C.T. is funded by a Tier 2 Canada Research Chair in Knowledge Synthesis.

Conflicts of interest: A.C.T. and S.E.S. are on the editorial board for the journal but were not involved with the peer review process or decision to publish. All other authors declare that they have no competing interests.

Role of the funder: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the article; or decision to submit the article for publication.

* Corresponding author. Tel.: +30 26510 05694; fax: +30 26510 05854.

E-mail address: averonik@cc.uoi.gr (A.A. Veroniki).

What is new?**Key findings**

- Significant barriers were encountered in obtaining study individual participant data (IPD). These included identifying trial data based on published reports and other sources, negotiating data ownership (for both authors and sponsors), and limited data access (including time, ability to share data, and special software needed).

What this adds to what is known?

- Likelihood of sharing IPD may be associated with study-specific characteristics, such as funding type, study size, study risk of bias, and treatment effect. For example, authors of publicly sponsored studies with medium-large treatment effect (i.e., an estimated treatment effect above 0.2 on the standardized mean difference scale) tended to respond positively to IPD requests. Availability of IPD from sponsors tended to be positive for large studies with a low risk of bias.
- The time taken to obtain IPD was longer than a year after a sponsor's positive response. Data sharing agreements were required for all sponsors. Clarifications from sponsors regarding the agreements ranged between 0 and 24 days. Approval of data sharing agreements ranged between 86 and 168 days

What is the implication and what should change now?

- Sharing IPD has legal, ethical, and logistical constraints, which may deter researchers from embarking on these projects and may deter trial participants from participating. This may reinforce reliance on aggregate data (network) meta-analysis that may have inadequate statistical power and accuracy, reducing the quality of evidence available to health professionals, policymakers, and patients.
- Our findings show that obtaining study IPD can take longer than a year after a sponsor's positive response. Therefore, we recommend that future planning of IPD meta-analyses should provide sufficient time (e.g., at least 2 years) for the IPD retrieval process, particularly in clinical areas where the approach is not yet established.

analysis approach is pooling individual patient data (IPD) from RCTs of clinical interventions [1–7]. Several methods have been developed to use IPD in meta-analysis [6] and network meta-analysis (NMA) [8], and their use has increased over the years [8,9] because meta-analyses or NMAs based on aggregated data have limitations. In particular, IPD can be used to explain the variation of treatment effects between studies within pairwise comparisons (heterogeneity) and the variation of treatment effects between pairwise comparisons (inconsistency) [9–13]. For example, a pooled estimate based on aggregate data with substantially heterogeneous treatment effect estimates may not be informative and an exploration of whether the treatment effects differ across prespecified, clinically important subgroups may be needed. Although meta-regression aims to answer this, it has many limitations, such as “aggregation bias” [14]. Aggregation bias arises when one incorrectly assumes that relationships observed at the group level hold also at the patient level and is also known as ecological bias [15–17]. The use of IPD can result in greater statistical power to detect patient–treatment relationships and help individualize management for patients with certain characteristics. Hence, confidence in meta-analysis results can increase by including IPD on all randomized patients, irrespective of whether they were included in analyses of the primary RCT.

Technological advances, such as safeguarding confidential data through secured platforms, have potential to increase the feasibility of obtaining IPD, and there is a strong impetus to share anonymized IPD from RCTs [18–31]. However, it has been suggested that reluctance to share data is still the main obstacle for obtaining IPD and performing IPD meta-analysis [32,33]. Potential reasons for this include concerns about patient confidentiality, lack of time to share IPD, not “owning” the data, cost for deidentifying and formatting the data, or lack of access to data by primary study authors after study completion [32]. A scoping review of indirect comparisons with IPD showed that 67% of included studies obtained IPD through the establishment of a collaborative group [8]. Hence, the cooperation of the authors of the primary studies is crucial for providing IPD in a useable format and answering queries about their data. Because sharing IPD has legal, ethical, and logistical constraints, we need to understand how to optimize this process.

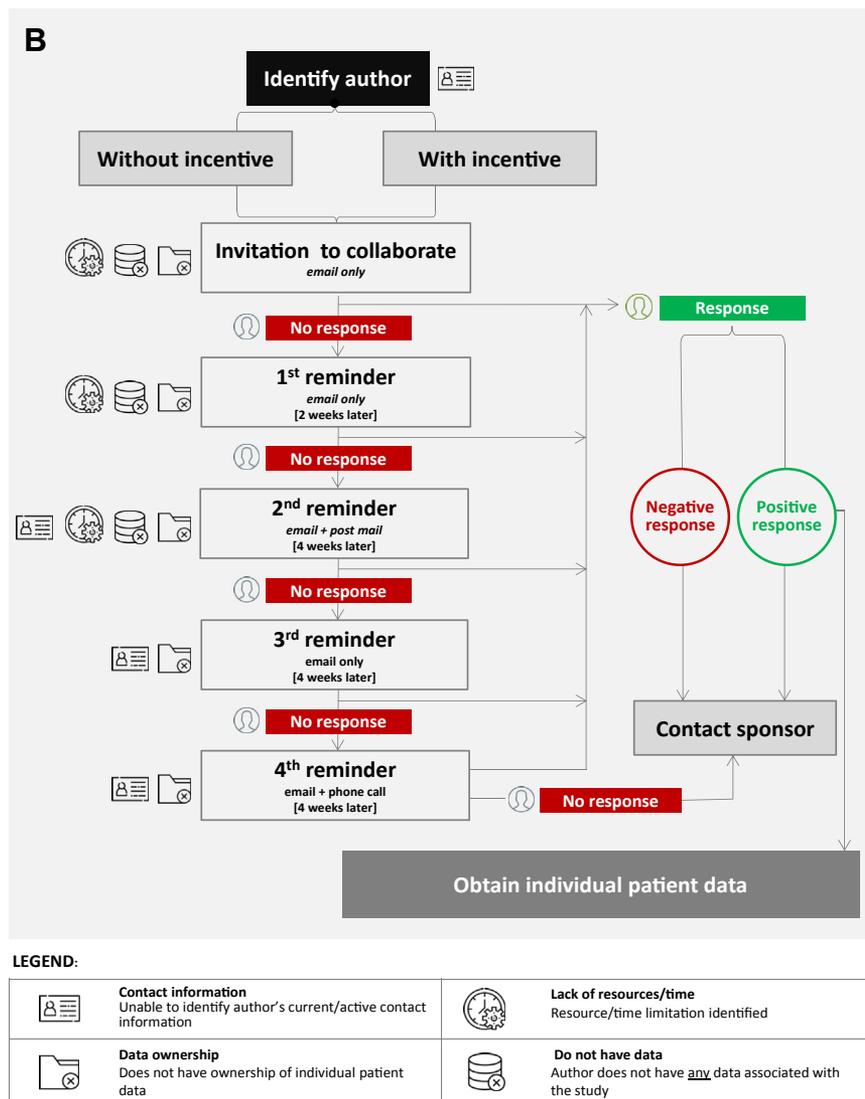
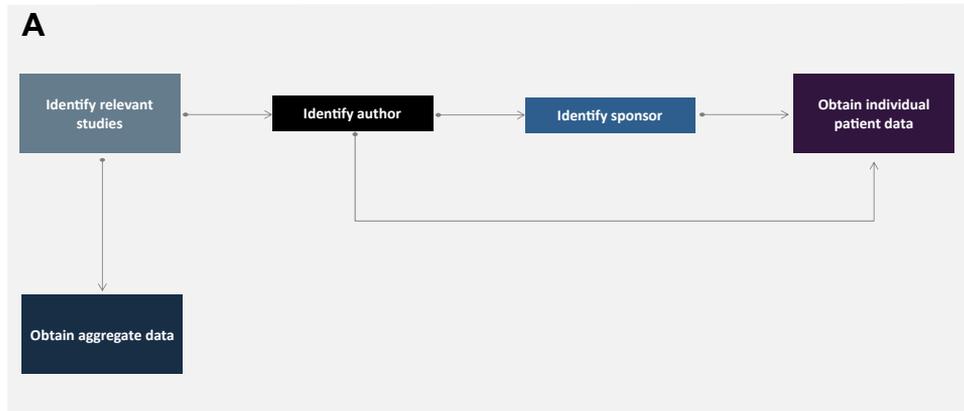
Our objective was to examine the impact of providing a small financial incentive to authors of RCTs that were eligible for a systematic review and NMA vs. usual contact strategies to obtain IPD. As a secondary objective, we aimed to describe potential barriers and facilitators associated with the data sharing process.

1. Introduction

The synthesis of data from multiple randomized clinical trials (RCTs) may strengthen scientific evidence used by health professionals and policymakers; the gold standard

2. Methods

The study protocol was registered with [ClinicalTrials.gov](https://www.clinicaltrials.gov) (NCT02569411; October 5, 2015). Our methods are



described briefly here; additional details can be found in the protocol publication and [Additional File 1: Appendices 1 and 2](#) [34]. Our RCT conforms to the Consolidated Standards to Reporting Trials (CONSORT) guidance [35] ([Additional File 2](#)).

We used RCTs identified through two systematic reviews and NMAs [36,37], and we followed the process as depicted in [Fig. 1A](#). Overall, we contacted both study authors and study sponsors ([Additional File 1: Appendix 2](#)) to obtain IPD. The process varied across sponsors ([Additional File 1: Appendix 3](#)).

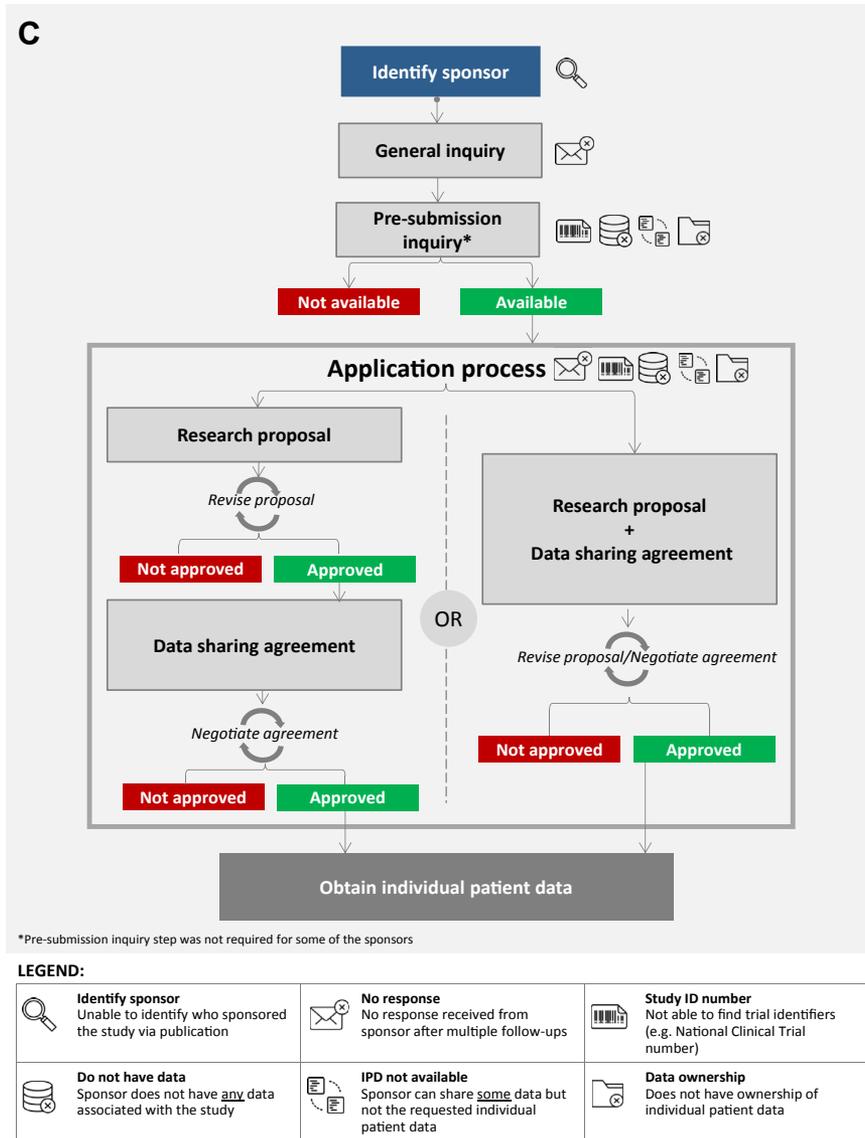


Fig. 1. Process of study for acquisition of IPD (A) through an author (B), and a sponsor (C), along with the barriers encountered at each step. IPD, individual patient data.

We performed a descriptive analysis using frequencies and percentages for all characteristics we either abstracted from trial publications or collected through the author and sponsor contacting process (Additional File 1: Appendix 2).

We compared author responses for which we received complete IPD, author response type (positive vs. negative), and response rate (response vs. no response) between experimental and control groups using the odds ratio (OR) and its corresponding 95% confidence interval (CI). On IPD receipt, we assessed data completion and time needed to share. Because only two IPD datasets were available across the intervention and control groups at the time these analyses were done, we could not compare the intervention group results according to the IPD characteristics. The OR and its corresponding 95% CI was used to compare author and sponsor response type and response rate in the

following groups: low vs. high/unclear risk of bias, industry/mixed-sponsored vs. publicly sponsored studies, large vs. small–moderate studies, statistically significant vs. nonstatistically significant treatment effects, small vs. medium–large effect studies. We assessed for a trend over publication years to respond using the Cox and Stuart trend test and the *trend* library in R [38]. We assessed whether a linear relationship existed between year of publication, absolute standardized mean difference (SMD) or sample size, and days to respond and calculated a Pearson correlation coefficient. The distribution of eligible studies by industry sponsor was plotted in a bubble plot using the *ggplot2* library in R [38]. Finally, we outlined barriers and resource requirements that prevented IPD from being obtained, challenges that delayed the process of obtaining IPD, as well as monetary costs and personnel resources required to obtain

Table 1. Author response per treatment

	Type I diabetes (N = 26 studies)			Alzheimer's dementia (N = 103 studies)			
	Detemir	Glargine	NPH	Donepezil	Galantamine	Memantine	Rivastigmine
Positive response	6	4	4	17	6	7	6
Negative response	6	8	10	17	9	7	6
Response ^a	12	12	14	34	15	14	12
No response	4	2	6	23	8	11	11
Total studies	16 (62%)	14 (54%)	20 (77%)	57 (55%)	23 (22%)	25 (24%)	23 (22%)

Abbreviations: NPH, neutral protamine Hagedorn.

^a Combined total of positive and negative responses.

IPD. We also describe the barriers encountered at the different levels of the author and sponsor contact process.

3. Results

We included 137 studies (29 RCTs for type I diabetes mellitus and 108 RCTs for Alzheimer's dementia) for which we attempted to obtain IPD by contacting the original authors and trial industry sponsors (Additional File 1: Appendix 4). The deadline for receiving IPD to be included in the analyses was February 28, 2018 (internal deadline set in our team only). In Additional File 1: Appendix 5, we present the number of eligible studies we requested from authors and sponsors separately and the number of studies we were able to acquire IPD. The individual study

characteristics are reported in Additional File 1: Appendices 6 and 7. Additional information on the results is reported in Additional File 1: Appendix 8.

3.1. Contacting authors for the RCT process

Of the 137 trials, we were unable to locate contact information for eight authors, and these were subsequently excluded. Of the eight trials, three were allocated to control, and five were allocated to intervention. These eight trials were published between 1998 and 2010 and had moderate to large sample size, low to unclear allocation concealment bias, and low to high incomplete outcome data bias. Of the eight trials, six were industry sponsored and two did not report funding, three compared neutral protamine Hagedorn (NPH) against glargine, and five compared galantamine, rivastigmine, or

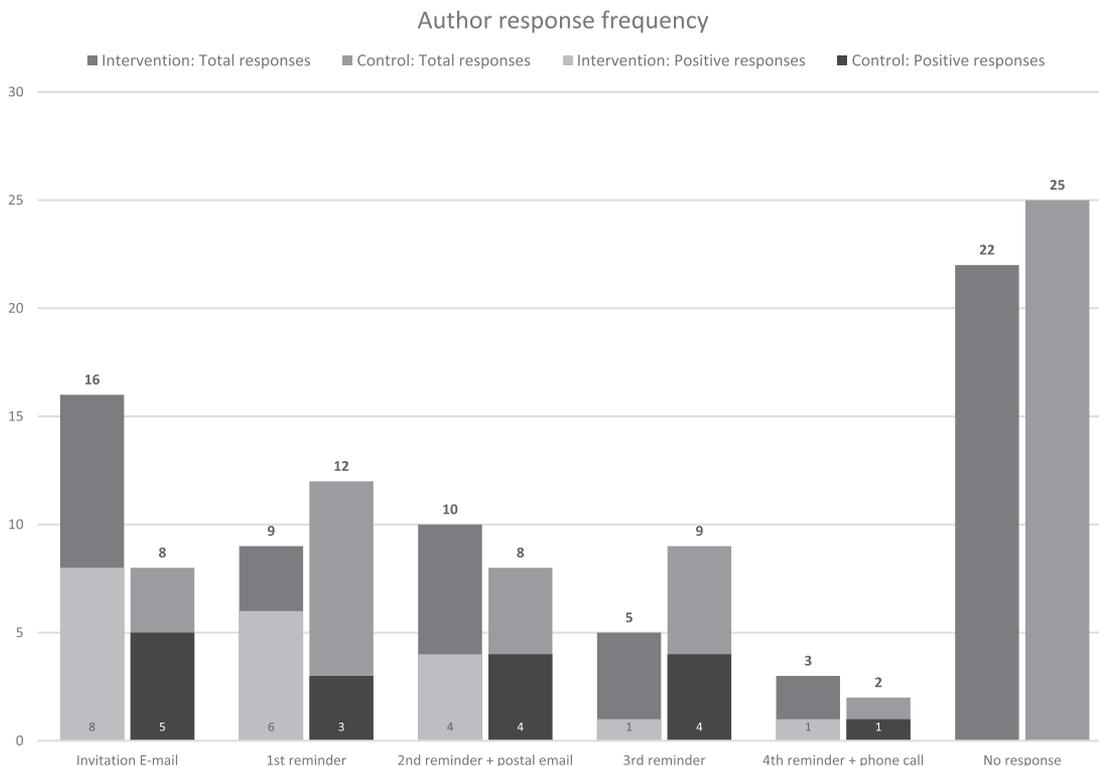


Fig. 2. Author response frequency by type of response and group author allocated per contact reminder.

Table 2. Author response summary

Number of authors contacted	129
No. of authors who did not respond	47 (36%) ^a
No. of authors who responded	82 (64%)
No. of authors who responded after first email	24 (0–15 d)
No. of authors who responded after second email	21 (20–48 d)
No. of authors who responded after third email	18 (50–83 d)
No. of authors who responded via post mail	0
No. of authors who responded after fourth email	14 (86–100 d)
No. of authors who responded after fifth email	5 (105–117 d)
No. of authors who responded via phone	5 ^b
Negative response	45
Contact funder/database	27 (60%)
Lack of resources/time	5 (11%)
Do not have approval/ownership	4 (9%)
Do not have data	3 (7%) ^c
Old data	3 (7%)
Not interested	2 (4%)
Contact corresponding author	1 (2%)
Positive response	37
Contact corresponding author/funder—provided contact person	20 (54%)
Contacted funder	5 (14%)
Interested but did not follow-up	12 (32%)
Time to respond (d)	0–117
Time to obtain data sharing approval (d)	467

^a Two email addresses were deactivated at the second reminder (6 wk after the initial email), one email address was deactivated at the third reminder (10 wk after the initial email), and one email address was deactivated at the fourth reminder (14 wk after the initial email).

^b Five calls were answered (Message left with admin [1]; language barrier [4]).

^c Two authors mentioned that they did not have the data available. One author mentioned that the data were destroyed.

donepezil against placebo or no treatment. In total, we included 26 type I diabetes studies, of which 20 (77%) compared NPH, 14 (54%) compared glargine, and 16 (62%) compared detemir to an alternative treatment. Of the 103 Alzheimer's dementia studies, 57 assessed donepezil (55%), 25 assessed memantine (24%), 23 assessed galantamine (22%), and 23 assessed rivastigmine (22%; Table 1). Additional File 1: Appendix 9 presents a CONSORT flow diagram depicting the process of the RCT and the IPD received across the two groups [35]. One hundred twenty-nine authors were contacted, and 82 (64%) responded (of which 37 [45%] responded positively and 45 [55%] responded negatively); 24 (19%) authors responded after the first invitation email, and the remainder responded across the four other reminders. Of the authors who did not respond (47 [36%]), two email addresses were deactivated at the second reminder (6 weeks after the initial email), one email address was deactivated at the third reminder (10 weeks after the initial email), and one email

address was deactivated at the fourth reminder (14 weeks after the initial email; Fig. 2 and Table 2). According to authors of the primary studies, 15 (33%) of the negative responses were due to lack of resources or time, lack of ownership or IPD, and old IPD that could not be retrieved. Of the positive responses, none of the authors shared their IPD.

The response type (OR = 1.13; 95% CI [0.47, 2.69]) and the response rate (OR = 1.25; 95% CI [0.61, 2.57]) were balanced in intervention and control groups (Additional File 1: Appendix 6). The response rates when the type I diabetes studies were categorized per treatment ranged between 70% (in NPH) and 85% (in glargine), in which the positive response rates ranged from 29% (in NPH) to 50% (in detemir). The response rates for the Alzheimer's dementia treatments ranged between 52% (in rivastigmine) and 65% (in galantamine), with positive responses ranging from 40% (in galantamine) to 50% (in donepezil; Table 2). The response type and response rate categorized per study characteristics are presented in Appendix 10.

The number of days required for an author to respond ranged from 0 to 117 days (mean days = 45, standard deviation [SD] = 39), irrespective of the response type (mean days for positive response = 40, SD = 38; mean days for negative response = 49, SD = 40). Similarly, the number of days required for an author in the intervention group to respond ranged from 0 to 117 days (mean days = 40, SD = 40), irrespective of the response type (mean days for positive response = 32, SD = 35; mean days for negative response = 48, SD = 43), and the number of days required for an author in the control group to respond ranged from 0 to 116 days (mean days = 49, SD = 38; mean days for positive response = 50, SD = 40, range [1,113]; mean days for negative response = 49, SD = 37, range [0, 116]).

3.2. Contacting sponsors and IPD databases

Of the 137 studies, 107 reported at least one industry-sponsored funder in their publication. In the remaining studies, 11 were publicly sponsored, and 19 did not report any information about funding. The 19 studies that did not report funding information were published in journals requiring disclosures for conflict of interest (COI; 1 [5%] study), COI/funding (2 [11%] studies), funding (2 [11%] studies), sponsor (9 [47%] studies), and not available (5 [26%] studies (Additional File 1: Appendices 6 and 7)). Across the 107 studies that reported sponsor information, 17 different industry sponsors were reported in the publications, and 24 (23%) studies reported at least two different sponsors (Additional File: Appendix 11). In total, we contacted 17 industry sponsors (four of which collaborate with two IPD databases) for 133 studies (or 107 unique studies because some studies reported multiple sponsors, where 83 studies were funded by one sponsor, 23 studies by two sponsors, and one study by five sponsors; see Additional File 1: Appendix 12 for list of cosponsored studies); and three sponsors (18%)

did not respond to any of our contact attempts. We contacted industry sponsors only, as we were not able to locate the contact information for most included public sponsors. In the following, we refer to the term “sponsor” to indicate an industry sponsor. Clinical Study Data Request and Yale University Open Data Access databases facilitated data sharing for 59 trials funded by four sponsors who made 18 trials available. In total, we included 25 type I diabetes studies, of which 20 (80%) compared NPH, 15 (60%) compared detemir, and 13 (52%) compared glargine with an alternative treatment. Of the 108 Alzheimer’s dementia studies, 69 assessed donepezil (64%), 26 assessed galantamine (24%), 23 assessed rivastigmine (21%), and 20 assessed memantine (19%; [Additional File 1: Appendix 13](#)). [Additional File 1: Appendix 14](#) shows the number of eligible studies per sponsor and per type of response regarding IPD availability. The response rate when the type I diabetes studies were categorized per treatment was 100% across all treatments. The positive response rates ranged from 15% (in glargine) to 80% (in detemir). The response rates for the Alzheimer’s dementia treatments ranged between 85% (in memantine) and 100% (in galantamine), with positive responses ranging from 19% (in galantamine) to 29% (in memantine). Of the total 133 studies, 38 (29%) unique studies were deemed available by the sponsors. However, the majority (89 studies; 67%) of the IPD were unavailable, and the reasons for refusal of providing IPD varied, including difficulty with study identification (46%), nonownership of IPD (26%), and the age of the study (too old, 12%; [Table 3](#), [Additional File: Appendix 15](#)).

Up until February 28, 2018, and within 318 days of contacting the sponsor, we received two complete IPD datasets from a single sponsor of 136 and 123 patients. We determined the dataset complete according to the study protocol. Allocation concealment was rated as low risk of bias for both studies; however, for incomplete outcome data, one study had low risk of bias, and one had a high risk of bias. Up until February 28, 2018, we also had another data sharing agreement signed by both parties for 12 unique studies. The time to clarify the data sharing agreement process ranged between 0 and 24 days, whereas the time to approve the data sharing agreement ranged between 86 and 168 days. All sponsors who agreed to share their IPD with us restricted its availability through a password-protected, software-restricted, and closed environment within a certain period of time ranging between 28 and 730 days. The exploration of response rate and response type across different study characteristics suggested effect sizes with wide CIs for most point estimates ([Additional File 1: Appendices 16 and 17](#)).

3.3. Barriers and resource requirements associated with the IPD acquisition

Several barriers and challenges were encountered during the IPD acquisition process. In [Table 4](#), we provide the barriers we encountered during the author and sponsor contact

process separately. The barriers and challenges are also depicted with different icons at the various levels of the author contact process in [Fig. 1B](#) and of the sponsor contact process in [Fig. 1C](#).

An important barrier in obtaining study IPD was the identification of the underlying trial dataset such as when an old study could not be easily located or when its data were lost by an author. Also, studies may not have been identifiable by sponsors when certain information was not available, such as the NCT number, due to the relatively recent widespread use of trial numbers (since 2005). In some cases, even when conducting exhaustive searches ([Additional File 1: Appendix 18](#)), an NCT number (or other related study ID) was difficult to find or did not exist, and hence, sponsors could not locate the study in their database solely based on the study citation details. Being unable to match study publication to the underlying studies when sponsors needed to be contacted was the most frequent reason for IPD being unavailable ([Additional File: Appendix 15](#)). Of the 98 studies with unavailable IPD, 41 studies (42%) could not be located by the study sponsor. Another important barrier was data ownership. It was often the case that study authors did not own the IPD, and the study funder had to be contacted to request the IPD (56 [43%]). This also applied to sponsors. For instance, because 18 of the eligible studies were cosponsored by Eisai and Pfizer ([Additional File 1: Appendix 11](#)), both sponsors were contacted to confirm ownership. Data ownership was the second most frequent reason (25 studies [26%]) for unavailable IPD ([Additional File: Appendix 15](#)). A barrier associated with the analysis once IPD was received was that IPD were only available through proprietary sponsor-specific platforms. This does not allow for IPD from different sponsor platforms to be combined (and could be a challenge for those who are unfamiliar with the software provided in the underlying platform). As the IPD could not be combined from all studies identified in a systematic review in a single place and model, a one-stage analysis was impossible. Also, the time that the platform permitted access to the IPD was often limited (e.g., 6 weeks), which is a significant constraint, given that IPD from different studies could be available at different time points. However, this required knowledge of the data items and times for access available from each sponsor.

4. Discussion

Our results showed that offering small financial incentives to study authors does not improve IPD retrieval. In our particular example, by the end of July 2017, we were unable to obtain any IPD datasets from trial authors and were only able to obtain two IPD datasets after contacting industry sponsors. We found that obtaining a response from authors to requests to access IPD may depend on study characteristics. Authors of publicly sponsored studies, those

Table 3. Sponsor response summary

Number of sponsors/databases ^a contacted	17
Number of sponsors who did not respond	3 (6 studies) ^b
Number of sponsors where data was unavailable	13 (89 studies) ^c
Cannot identify study	4 (41 studies)
Cannot share data	6 (22 studies)
Language	1 (1 study)
Old study	2 (11 studies)
Phase 4 study	1 (4 studies)
Potential business considerations under review	1 (1 study)
No details provided	1 (5 studies)
Do not own data	10 (23 studies)
IPD not available	1 (3 studies)
Number of sponsors who required a research proposal to be submitted first	7 (91 studies) ^d
Research proposal approved	5 (64 studies)
Research proposal not approved (no reason provided)	1 (5 studies)
Research proposal under review	1 (22 studies)
Number of sponsors who required a research proposal and data sharing agreement (DSA) to be submitted congruently	4 (24 studies) ^e
Research proposal approved and DSA approved	1 (15 studies)
Research proposal approved and DSA not approved (do not own data)	1 (2 studies)
Research proposal and DSA under review	2 (7 studies)
Number of studies where study identification number was required	62
Number of studies where author was contacted for study identification number	48
Number of studies where author provided study identification number	7
Number of studies where author did not provide study identification number:	41
No response	30
Does not have the information	4
Referred to sponsor	3
Does not have access to the information	2
Referred to lead PI of the study	1
Not registered	1
Number of studies where study identification number was found	13
Time to clarify data sharing process (d)	0–24 d
Time to approve research proposal (d)	22–121
Time to approve data sharing agreement (d)	86–168
Number of studies requested	137
Number of studies available	38
Number of studies shared	11

Abbreviations: IPD, individual patient data; PI, principal investigator.

^a ClinicalStudyDataRequest.com (CSDR); Yoda.Yale.edu (YODA).

^b Merz, ONO, Roivant.

^c Abbvie, Daiichi-Sankyo, Eisai, Forest Laboratories/Allergan, GlaxoSmithKline, Janssen, Lundbeck, Novartis, Novo Nordisk, Pfizer, Shire Pharmaceuticals, Takeda.

^d Abbvie, CSDR, Forest Laboratories/Allergan, Janssen, Pfizer.

^e AstraZeneca, Lundbeck, Novo Nordisk, Shire Pharmaceuticals.

that included between less than 50 and 150 patients, and those with a medium to large treatment effect (i.e., an estimated treatment above 0.2 on the SMD scale) tended to

respond positively to IPD requests. Increased odds of a positive response were also found in studies at high risk of bias. This is because small-to-moderate studies are

Table 4. Barriers and resource requirements

Activity/item	Contacting authors		Contacting sponsors		Resources
	Issue	Impact	Issue	Impact	
Initial contact	Cannot locate contact information (email address, mailing address, and/or phone number)	Spent extra time finding current contact information for authors Unable to invite authors to participate	Unable to locate an “obvious” contact for IPD requests Sponsors did not respond to initial contact	Spent extra time pursuing multiple avenues and contacting multiple sponsors (for cosponsored studies) before finding the correct one	Research staff: 1 research assistant 3 research coordinators 2 scientists Administrative staff: 2 administrative assistants
Ongoing communication	Emails become undeliverable Postal mail returned Initial contact directs to a coauthor that is already part of the RCT Authors do not respond (either to initial contact or later communication)	Loss to follow-up: unable to pursue any further	No direct avenues for communication with sponsors (e.g., general inquiry only) Multiple departments/teams involved in communication	Difficult to follow-up with sponsors when no response is received Extra time needed to relay updates to sponsors	Research staff: 1 research assistant 1 research coordinator Administrative staff: 2 administrative assistants Legal staff: 1 research contract specialist 1 research contract analyst Incentives/communication: Gift cards (incentives for intervention) Post mail (reminder) Long distance phone charges (reminder)
Application process	Not applicable	Not applicable	Differing requirements/processes between sponsors Additional items required by sponsors (e.g., additional training/agreements) Sponsors changed methods for application when application was in process	Significant delays to obtaining IPD	Research staff: 1 research assistant 1 research coordinator 2 scientists Administrative staff: 2 administrative assistants Legal staff: 1 research contract specialist 1 research contract analyst
Identifying studies	Study is “too old” to find/share data	Unable to obtain IPD from author	Could not identify studies with available information	Additional time/resources used to find trial identifiers/study information	Research staff: 1 research assistant 1 research coordinator Library staff: 1 information specialist
Legal agreements	Not applicable	Not applicable	Multiple revisions Regulations differ between countries Document formats (uneditable, need physical copies)	Time-consuming, lengthens agreement process Difficult to revise/edit documents	Research staff: 1 research assistant 1 research coordinator 2 scientists Administrative staff: 2 administrative assistants Legal staff: 1 research contract specialist 1 research contract analyst

(Continued)

Table 4. Continued

Activity/item	Contacting authors		Contacting sponsors		Resources
	Issue	Impact	Issue	Impact	
Data ownership	Study authors do not “own” data	Authors unable to share IPD	Sponsors do not own data	We cannot identify who owns the data and thus cannot obtain IPD Extra time and resources needed to identify “true” data owners	Research Staff 1 research assistant 1 research coordinator
Data Sharing/ Receiving Data	Lack of available resources Ethics restrictions on data sharing	Authors unable to share IPD	‘Out of scope’ of data sharing policy Unable to share all data from different studies at once Data only available through sponsor platforms Additional licenses required to access the data Data only available for a specified amount of time	Unable to obtain IPD Access to data is limited Unable to conduct one-stage analysis Additional cost or unable to obtain IPD	Research staff: 1 research assistant 1 research coordinator 1 scientist 2 research managers Licensing: WHO Drug Dictionary Approximately \$8,958.25 USD/ sponsor

Abbreviations: IPD, individual patient data; RCT, randomized clinical trial.

typically at high risk of bias in most domains and are associated with large treatment effects. In contrast, IPD availability from sponsors tended to be positive for large (> 150 patients) and low risk of bias studies, with small and nonstatistically significant treatment effects. This suggests that well-conducted industry-sponsored studies are more likely to be shared. It should be highlighted that there is a high risk of confounding in our results, as large studies are typically associated with small treatment effects and low risk of bias. In addition, these findings should be interpreted with caution, as our estimated ORs were associated with wide CIs. This high uncertainty in ORs may be associated with low power to detect the true effect. Similarly, the marginally nonstatistically significant trend of positive author or sponsor responses across publication years favoring newer RCTs may be associated with the low power of the test.

Sharing IPD may be constrained by a number of legal, ethical, and logistical factors, which may deter researchers from undertaking them and trial investigators participating in them. This may perpetuate reliance on the conduct of aggregate data meta-analysis and NMA that may reduce statistical power and accuracy of results. Significant barriers in obtaining study IPD from trial sponsors may include matching study publication to the underlying study, issues around data ownership, and acquiring of data dictionary licenses.

In general, time and cost may be a barrier to carrying out an IPD NMA. Costs include not only staff wages, including administrative, legal, library, and research staff, but also

license costs (when applicable, e.g., WHO Drug Dictionary license approximate cost \$8,958.25 USD per sponsor). We were surprised to encounter the license cost issue, as it has not been encountered previously in the context of collaborative group IPD meta-analyses and could be an isolated experience or an additional cost of obtaining data from trial sponsors and data repositories. The longer time required to conduct this type of research may be considered an additional barrier, especially when time-sensitive decisions need to be made. Our findings show that obtaining study IPD can take longer than a year after a sponsor’s positive response. Thus, accessing data via repositories may not be as rapid as was hoped, and therefore, we recommend that in accordance with customary practice in collaborative IPD meta-analyses, future planning of IPD meta-analyses that involve obtaining data from trial sponsors through data sharing platforms should include sufficient time for the IPD retrieval process (probably at least 2 years). For example, the IPD retrieval process for a recently published IPD NMA of 10 antiepileptic drugs required 4 years to obtain the 65% of the included participants and 38% of identified clinical trials [39]. Even if access to IPD is granted, additional barriers may arise, such as having to analyze IPD datasets using different sponsor data sharing platforms and software making one-stage analysis impossible. Being able to access IPD only for a limited time (e.g., 4–6 weeks) is also a serious limitation and constraint, as when analyzing IPD from different studies provided by different sponsors, the IPD datasets can be available at different time points, and different data variables may be available.

A limitation of our RCT is that we did not anticipate that the trial authors would not have authority to grant access to the data sets, and sponsors would need to be contacted instead. We contacted one author about a single study to avoid contamination bias in our RCT. Through this process, we avoided sending multiple requests to a single author. If an author directed us to another coauthor, then we discussed all articles with them. However, in a usual IPD project aiming to collect IPD from a number of studies, if multiple studies from the same research group are of interest, one would probably request IPD from all these studies at once to maximize the amount of data provided. Also, the time available to conduct the analysis in this study was probably another barrier in obtaining IPD. Another limitation is that blinding to treatment allocation was only possible for the statistician who conducted the analysis. Because of the nature of the intervention and the study design, blinding was impossible for research personnel and outcome assessors. However, given that neither group has contributed data, the unsuccessful blinding has not impacted our success rate. To reduce potential bias in the author responses received, we planned to send authors a debriefing letter at the end of the trial, informing them that they participated in an RCT. Also, to avoid misinterpretation of the small financial sum offered as compensations for the effort involved in the preparation of IPD, we proposed our IPD NMA as a collaborative project. If the authors met the International Committee of Medical Journal Editors criteria [40] and shared their IPD, they would be included in the collaborative group authorship of the final publication. The authors' reluctance to share their IPD could be because of the contact person's expertise outside the clinical field relevant to the trial. However, to increase author responses, an experienced scientist (A.C.T.) in the field of systematic reviews and meta-analyses contacted each author and provided citations of our published protocols, where researchers with significant reputations in the relevant fields were included as coauthors. In addition, in our communication with the trialists, we indicated our experience in the fields through our published systematic reviews in the area, which were funded through the Canadian Institutes of Health Research. An additional limitation is that we used different outcomes across studies to explore the association between response rate (and type) and the magnitude of treatment effect. In total, we used 62 studies with Mini-mental State Examination, 26 studies with serious adverse events, and 25 studies with A1c reported as the outcome measure. Although it is expected that the treatment effect will vary across outcomes, this was the only feasible approach to include the most data possible to explore this association. Another limitation is that we used studies examining response rates from surveys to inform our study size [41–44] because to the best of our knowledge, no studies assessing response rate in retrieving IPD from RCTs using a financial incentive is available. This may have underestimated the required sample size, producing imprecise

results. Our findings represent retrieving IPD from authors for two certain clinical areas (type I diabetes and Alzheimer's dementia), and these might not be well generalized to other drug trials. It should be noted although that the majority (77%) of the included RCTs were sponsored by pharmaceutical companies, and this may have affected the IPD retrieval.

Before deciding to conduct an IPD meta-analysis or NMA, one should consider and weigh up the benefits and limitations of the approach. Although the process of sharing IPD may vary according to the disease, treatment, and clinical question addressed [45], one should not only consider the cost and time needed to conduct an IPD meta-analysis or NMA but also the possibility of being unsuccessful in retrieving IPD [33]. This may be particularly important for NMAs that involved large numbers of studies. When IPD meta-analyses fail to obtain data, the impact of this on the analyses should be considered, especially when IPD are missing not at random (e.g., when acquirement of IPD depends on the RCT characteristics). This could threaten validity and, in turn, impact clinical decision-making, as the practice of evidence-based medicine relies on the availability of timely, relevant, and scientifically sound data on the risks and benefits associated with medical interventions. Important initiatives to reporting study results are currently being made by medical journals via encouraging authors to use the CONSORT checklist [35], as well as by study authors and organizations (<http://www.alltrials.net/>) [19–22]. However, as our results showed, IPD sharing is not yet well-established in the fields of type I diabetes and Alzheimer's dementia, and more efforts are required to achieve this goal.

CRediT authorship contribution statement

Areti Angeliki Veroniki: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Writing - original draft, Writing - review & editing. **Huda M. Ashoor:** Data curation, Project administration, Writing - review & editing. **Susan P.C. Le:** Data curation, Project administration, Writing - review & editing. **Patricia Rios:** Data curation, Project administration, Writing - review & editing. **Lesley A. Stewart:** Writing - review & editing. **Mike Clarke:** Writing - review & editing. **Dimitris Mavridis:** Writing - review & editing. **Sharon E. Straus:** Conceptualization, Methodology, Funding acquisition, Writing - review & editing. **Andrea C. Tricco:** Conceptualization, Methodology, Funding acquisition, Writing - review & editing.

Acknowledgments

The authors thank Ms. Becky Skidmore for updating the literature searches, John D. Ivory, Paul A. Khan, Fatehmeh

Yazdi, and Marco Ghassemi for data extraction and quality appraisal, and Fatiah De Matas for formatting the figures.

Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jclinepi.2019.05.031>.

References

- [1] Simmonds M, Stewart G, Stewart L. A decade of individual participant data meta-analyses: a review of current practice. *Contemp Clin Trials* 2015;45:76–83.
- [2] Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Eval Health Prof* 2002;25(1):76–97.
- [3] Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, et al. Preferred reporting items for systematic review and meta-analyses of individual participant data: the PRISMA-IPD statement. *JAMA* 2015;313:1657–65.
- [4] Tierney JF, Pignon JP, Gueffier F, Clarke M, Askie L, Vale CL, et al. How individual participant data meta-analyses have influenced trial design, conduct, and analysis. *J Clin Epidemiol* 2015;68:1325–35.
- [5] Oxman AD, Clarke MJ, Stewart LA. From science to practice. Meta-analyses using individual patient data are needed. *JAMA* 1995;274:845–6.
- [6] Debray TP, Moons KG, van Valkenhoef G, Efthimiou O, Hummel N, Groenwold RH, et al. Get real in individual participant data (IPD) meta-analysis: a review of the methodology. *Res Synth Methods* 2015;6(4):293–309.
- [7] Debray TP, Schuit E, Efthimiou O, Reitsma JB, Ioannidis JP, Salanti G, et al. An overview of methods for network meta-analysis using individual participant data: when do benefits arise? *Stat Methods Med Res* 2018;27(5):1351–64.
- [8] Veroniki AA, Straus SE, Soobiah C, Elliott MJ, Tricco AC. A scoping review of indirect comparison methods and applications using individual patient data. *BMC Med Res Methodol* 2016;16:47.
- [9] Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ* 2010;340:c221.
- [10] Stewart LA, Clarke MJ. Practical methodology of meta-analyses (overviews) using updated individual patient data. *Stat Med* 1995;14(19):2057–79.
- [11] Simmonds MC, Higgins JP, Stewart LA, Tierney JF, Clarke MJ, Thompson SG. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. *Clin Trials* 2005;2(3):209–17.
- [12] Donegan S, Williamson P, D’Alessandro U, Tudur Smith C. Assessing the consistency assumption by exploring treatment by covariate interactions in mixed treatment comparison meta-analysis: individual patient-level covariates versus aggregate trial-level covariates. *Stat Med* 2012;31:3840–57.
- [13] Schmid CH, Landa M, Jafar TH, Giatras I, Karim T, Reddy M, et al. Constructing a database of individual clinical trials for longitudinal analysis. *Control Clin Trials* 2003;24:324–40.
- [14] Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. *Stat Med* 2004;23:1663–82.
- [15] Berlin JA, Santanna J, Schmid CH, Szczech LA, Feldman HI. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Stat Med* 2002;21:371–87.
- [16] Cooper H, Patall EA. The relative benefits of meta-analysis conducted with individual participant data versus aggregated data. *Psychol Methods* 2009;14(2):165–76.
- [17] Higgins JP, Whitehead A, Turner RM, Omar RZ, Thompson SG. Meta-analysis of continuous outcome data from individual patients. *Stat Med* 2001;20:2219–41.
- [18] Sullivan PF. The psychiatric GWAS consortium: big science comes to psychiatry. *Neuron* 2010;68(2):182–6.
- [19] Drazen JM. Sharing individual patient data from clinical trials. *N Engl J Med* 2015;372:201–2.
- [20] El Emam K, Rodgers S, Malin B. Anonymising and sharing individual patient data. *BMJ* 2015;350:h1139.
- [21] Gotsche PC. Why we need easy access to all data from all clinical trials and how to accomplish it. *Trials* 2011;12:249.
- [22] Suvarna VR. Sharing individual patient data from clinical trials. *Perspect Clin Res* 2015;6(2):71–2.
- [23] Krleza-Jeric K, Chan AW, Dickersin K, Sim I, Grimshaw J, Gluud C. Principles for international registration of protocol information and results from human trials of health related interventions: ottawa statement (part 1). *BMJ* 2005;330(7497):956–8.
- [24] Goodyear MD, Krleza-Jeric K, Lemmens T. The declaration of helsinki. *BMJ* 2007;335(7621):624–5.
- [25] Krleza-Jeric K, Lemmens T. 7th revision of the Declaration of Helsinki: good news for the transparency of clinical trials. *Croat Med J* 2009;50(2):105–10.
- [26] Chan AW, Song F, Vickers A, Jefferson T, Dickersin K, Gotsche PC, et al. Increasing value and reducing waste: addressing inaccessible research. *Lancet* 2014;383(9913):257–66.
- [27] WHO statement on public disclosure of clinical trial results. World Health Organization; 2015. Available at: <http://www.who.int/ictrp/results/reporting/en/>. Accessed June 21, 2019.
- [28] Institute of Medicine (IOM). Sharing clinical trial data: maximizing benefits, minimizing risk. Washington, DC: The National Academies Press; 2015.
- [29] Office of the Independent Police Review Director [Internet]. Broken trust: indigenous people and the thunder bay police service. 2018. Available at <https://www.falconers.ca/wp-content/uploads/2018/12/Broken-Trust-Indigenous-People-and-the-Thunder-Bay-Police-Service.pdf>. Accessed June 21, 2019.
- [30] Smith CT, Hopkins C, Sydes M, Woolfall K, Clarke M, Murray G, et al. Good practice principles for sharing individual participant data from publicly funded clinical trials. *Trials* 2015;16(Suppl 2):O1.
- [31] Evangelou E, Ioannidis JP. Meta-analysis methods for genome-wide association studies and beyond. *Nat Rev Genet* 2013;14(6):379–89.
- [32] Villain B, Dechartres A, Boyer P, Ravaud P. Feasibility of individual patient data meta-analyses in orthopaedic surgery. *BMC Med* 2015;13:131.
- [33] Jaspers GJ, Degraeuwe PL. A failed attempt to conduct an individual patient data meta-analysis. *Syst Rev* 2014;3:97.
- [34] Veroniki AA, Straus SE, Ashoor H, Stewart LA, Clarke M, Tricco AC. Contacting authors to retrieve individual patient data: study protocol for a randomized controlled trial. *Trials* 2016;17(1):138.
- [35] Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med* 2010;152:726–32.
- [36] Veroniki AA, Straus SE, Ashoor HM, Hamid JS, Hemmelgarn BR, Holroyd-Leduc J, et al. Comparative safety and effectiveness of cognitive enhancers for Alzheimer’s dementia: protocol for a systematic review and individual patient data network meta-analysis. *BMJ Open* 2016;6(1):e010251.
- [37] Veroniki AA, Straus SE, Ashoor HM, Hamid JS, Yu C, Tricco AC. Safety and effectiveness of long-acting versus intermediate-acting insulin for patients with type 1 diabetes: protocol for a systematic review and individual patient data network meta-analysis. *BMJ Open* 2015;5(12):e010160.
- [38] T. Pohlert (2018). trend: Non-Parametric Trend Tests and Change-Point Detection. R package version 1.1.1. <https://CRAN.R-project.org/package=trend>. Accessed June 21, 2019.

- [39] Nevitt SJ, Marson AG, Davie B, Reynolds S, Williams L, Smith CT. Exploring changes over time and characteristics associated with data retrieval across individual participant data meta-analyses: systematic review. *BMJ* 2017;357:j1390.
- [40] Defining the role of authors and contributors. Int Committee Med J Editors. Available at <http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html>. Accessed September 6, 2018.
- [41] James KM, Ziegenfuss JY, Tilburt JC, Harris AM, Beebe TJ. Getting physicians to respond: the impact of incentive type and timing on physician survey response rates. *Health Serv Res* 2011;46:232–42.
- [42] Asch DA, Christakis NA, Ubel PA. Conducting physician mail surveys on a limited budget. A randomized trial comparing \$2 bill versus \$5 bill incentives. *Med Care* 1998;36:95–9.
- [43] Pit SW, Vo T, Pyakurel S. The effectiveness of recruitment strategies on general practitioner's survey response rates - a systematic review. *BMC Med Res Methodol* 2014;14:76.
- [44] Everett SA, Price JH, Bedell AW, Telljohann SK. The effect of a monetary incentive in increasing the return rate of a survey to family physicians. *Eval Health Prof* 1997;20(2):207–14.
- [45] Reaves TA, Heath JE. Interval coding of temperature by CNS neuroenes in thermoregulation. *Nature* 1975;257:688–90.