

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Clinical Trial Data Sharing: Here's the Challenge

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-032334.R1
Article Type:	Communication article
Date Submitted by the Author:	20-Jun-2019
Complete List of Authors:	Kochhar, Sonali; Global Healthcare Consulting; Erasmus Medical Center, Dept Public Health Knoppers, Bartha; McGill University, Center of Genomics and Policy Gamble, Carrol; University of Liverpool, Biostatistics Chant, Alan; UK Clinical Research Collaboration Board Koplan, Jeffrey; Emory University, Emory Global Health Institute Humphreys, Georgina; Wellcome Trust, Open Research; University of Oxford Green Templeton College
Primary Subject Heading:	Communication
Secondary Subject Heading:	Health informatics
Keywords:	Data sharing, Clinical trials < THERAPEUTICS, Data re-use

SCHOLARONE™
Manuscripts

1 **Clinical Trial Data Sharing: Here's the Challenge**

2

3 Sonali Kochhar ^{a, b, c}, Bartha M Knoppers ^d, Carrol Gamble ^e, Alan Chant ^f, Jeffrey P Koplan ^g,
4 Georgina S Humphreys ^{h,i}

5

6 ^a Global Healthcare Consulting, New Delhi, India

7 ^b Department of Public Health, Erasmus MC, University Medical Center, Rotterdam, The
8 Netherlands

9 ^c Department of Global Health, University of Washington, Seattle, USA

10 ^d Center of Genomics and Policy, McGill University, Canada

11 ^e Department of Biostatistics, University of Liverpool, UK

12 ^f UK Clinical Research Collaboration Board, UK

13 ^g Emory Global Health Institute, Emory University, Atlanta, USA

14 ^h Wellcome Trust, London, UK

15 ⁱ Green Templeton College, University of Oxford, Oxford, UK

16

17 *Disclaimer:* The findings, opinions, conclusions, and assertions contained in this consensus
18 document are those of the individual members. They do not necessarily represent the official
19 positions of any participant's organization.

20

1
2
3 **21 Abstract**
4

5
6 **22** Anonymized patient-level data from clinical research is increasingly recognized as a fundamental
7
8 **23** and valuable resource. It has value beyond the original research project and can help drive
9
10 **24** scientific research and innovations and improve patient care. To support responsible data sharing
11
12 **25** we need to develop systems that work for all stakeholders. The members of the Independent
13
14 **26** Review Panel (IRP) for the data sharing platform Clinical Study Data Request (CSDR) describe
15
16 **27** here some summary metrics from the platform, and challenge the research community on why the
17
18 **28** promised demand for data has not been observed. While acknowledging there are areas for
19
20 **29** improvement in speed of access and promotion of the platform, the total number of applications
21
22 **30** for access and the resulting publications have been low, and challenge the sustainability of this
23
24 **31** model. What are the barriers for data contributors and secondary analysis researchers? If this model
25
26 **32** does not work for all, what needs to be changed? One thing is clear; that data access can realize
27
28 **33** new and unforeseen contributions to knowledge and improve patient health, but this will not be
29
30 **34** achieved unless we build sustainable models together that work for all.
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 37 CSDR is a consortium of 14 international pharmaceutical companies (GSK, Astellas Pharma,
4
5 38 Bayer, Chugai, Eisai, Eli Lilly, Novartis, ONO, Roche, Sanofi, Sunovion, Shionogi Inc., UCB,
6
7
8 39 and ViiV) and four academic research funders (The Wellcome Trust, The Bill & Melinda Gates
9
10 40 Foundation, The UK Medical Research Council and Cancer Research UK)¹. It was launched in
11
12 41 2013 and currently lists anonymized patient data from 3358 studies on the platform, including 10
13
14 42 studies from academic funders. The mandate is to reduce the barriers to access and re-use data,
15
16 43 thereby facilitating data sharing in an equitable, transparent, and independent manner.
17
18
19 44

20
21 45 Many global initiatives actively promote and enable sharing of research data and most funders
22
23 46 mandate researchers to plan for sharing their data globally^{2,3,4}. The European Medicines Agency
24
25 47 and US Food and Drug Administration have basic requirements in place for data disclosure and
26
27 48 clinical trial transparency. Trial participants' confidentiality and privacy need to be protected and
28
29 49 the terms of consent to participate in research respected. Managed access systems can help with
30
31 50 this, including an IRP's review of all requests for data access and having data sharing agreements
32
33 51 in place which place appropriate restrictions on data usage, though it should be recognized that
34
35 52 these systems add time to the process from application to data access.
36
37
38
39

40 53
41
42 54 CSDR's system allows researchers to request access to anonymized global clinical trial data from
43
44 55 multiple studies and sponsors. All data requests that pass initial administrative and data availability
45
46 56 checks are reviewed by an IRP, the secretariat for which is provided by Wellcome. Once access is
47
48 57 granted nearly all CSDR members restrict data access to a secure online analysis environment
49
50 58 which they say ensures patient privacy whilst maintaining the utility of the data for secondary
51
52 59 analysis. This system can limit the merging of data from other non-CSDR sources and the range
53
54
55
56
57
58
59

1
2
3 60 of software available for researchers. Free access to data is usually granted for 12 months, with the
4
5 61 possibility of extension, and CSDR requires researchers to report on findings, which are then listed
6
7
8 62 on the website.
9

10 63

11 64 **CSDR metrics**

12
13
14 65 From 2014 to the end of Jan 2019, there were a total of 473 research proposals (RPs) submitted to
15
16
17 66 CSDR. Of these, 364 met initial administrative and data availability checks. In reviewing
18
19 67 applications the IRP default position is to provide access, and they have approved 291 (84% of
20
21 68 those considered) and rejected 55. Thirty four of the rejected RPs were subsequently revised and
22
23
24 69 re-submitted. The remaining RPs are either still in process, withdrawn or no response has been
25
26 70 received from the researchers (Figure 1). The most common reasons for rejection were unsuitable
27
28 71 statistical methods, too technical lay summaries, and insufficient information presented.
29

30 72

31
32
33 73 Overall, the annual number of RPs submitted has remained fairly static with 70 submitted in 2014,
34
35 74 96(2015), 92(2016), 85(2017), and 97(2018). Researchers at institutions in over 30 different
36
37
38 75 countries have submitted RPs, although 168 RPs have been from researchers based in the USA
39
40 76 (35%) and very few have been submitted from researchers based in low and middle income
41
42 77 countries.
43

44 78

45
46
47 79 From January 2014 it became possible to request data from multiple sponsors and 17% (73/427)
48
49 80 of RPs have requested data from more than one sponsor. However, the median number of studies
50
51 81 per RP is 2 (1-5, 25th-75th percentile), with only a handful of RPs requesting a large number of
52
53
54 82 studies (the biggest request involved 192 studies and 11 sponsors).
55

1
2
3 83
4
5 84 Of the 90 research teams that completed their analysis more than 12 months ago, 41 (45%) have
6
7
8 85 at least one publication, 28 (31%) are publishing soon and the remaining 21 have either not
9
10 86 published or not responded to reminders from CSDR. Although these numbers appear encouraging
11
12 87 in terms of converting access to data into publications of new findings, of concern are the 54 RPs
13
14 88 (24%) whose researchers were granted access but did not log into the analysis environment. It is
15
16 89 puzzling why this would happen, given the significant investment on both the part of the sponsors
17
18
19 90 and the researchers in getting the RPs to this point in the process. CSDR is planning to contact
20
21 91 those researchers to understand their constraints and challenges.
22
23
24 92

25 26 93 **Lessons learned**

27
28 94 From an IRP's perspective, the lessons learned from the CSDR experience include the following:
29
30
31 95

32
33 96 CSDR is a valuable resource of data from pharmaceutical companies and academic research
34
35 97 funders which is available for free for researchers. However, it is an expensive and resource-
36
37 98 intensive task for trial sponsors to provide access to data through this managed access model and
38
39 99 this may challenge the long-term sustainability of such a platform.
40
41
42 100

43
44 101 Pharmaceutical companies and academic funders should pool resources to strengthen and
45
46 102 sustainably support data-sharing infrastructure and to develop and implement harmonized
47
48 103 principles, standards and best practices⁵. The portal cost for the sharing of data should not be
49
50 104 prohibitive for new data contributors.
51
52
53 105

1
2
3 106 It is important that there is a transparent system in place for data access decisions to maintain
4
5 107 equity for all those that want to re-use data. The IRP and the secretariat support provided by
6
7
8 108 Wellcome have been critical to ensuring a trusted and independent managed access system.
9

10 109
11
12 110 Having an experienced and multi-disciplinary IRP with expertise in ethics, statistics,
13
14 111 epidemiology, clinical research and a lay member has helped to ensure that feedback is provided
15
16 112 to the data requesters, including suggestions for improvement of proposals, and to ensure all
17
18 113 proposals receive consistent review. The IRP's perspective is to encourage and facilitate data
19
20 114 sharing (unless there are significant reasons not to do so). The respected quality of the IRP service
21
22 115 has been demonstrated by a newer data sharing platform (Vivli), launched in 2018, requesting the
23
24 116 Wellcome IRP also be available for data contributors to their system. The Wellcome IRP accepted
25
26 117 this request and has already considered several proposals through this platform too, applying the
27
28 118 same criteria as for CSDR.
29
30

31
32
33 119
34
35 120 It would be helpful if consent for clinical research could, as far as possible, include provision for
36
37 121 further data use beyond the original study.⁵ In the absence of specific legislation or professional
38
39 122 guidance, institutional ethics committees should also adopt consistent policies for the need (or not)
40
41 123 for ethics review for secondary use of anonymized data. This would clarify if ethics review is
42
43 124 needed (or not), increase the amount of data available for re-use, and decrease the time to access
44
45 125 the data.
46
47
48

49 126
50
51 127 The data access process should be easily discoverable with transparent metrics for potential data
52
53 128 users. A common data sharing agreement should be available for all the data providers and once
54
55
56
57
58
59
60

1
2
3 129 researchers from an institution have signed the agreement, it should be applicable for other
4
5 130 researchers from that institution (to decrease the time often taken by institutions negotiating
6
7
8 131 changes to the agreement).
9

10 132
11
12 133 The analysis environment must be easily usable by the researchers (including for merging multiple
13
14 134 datasets and statistical analysis) and not be expensive for the data providers. Increasing efficiency
15
16
17 135 in the process to ensure that data access happens as soon as possible benefits the researchers
18
19 136 (currently the median time from submission to data access is 190 days in CSDR). It is critical that
20
21 137 researchers that get access to the data, report their results in a timely manner (e.g. within six months
22
23 138 of completion of data access), so that it helps to move the research field forward, and reduce
24
25
26 139 research waste. Perhaps this should be mentioned in the data sharing agreement.
27

28 140
29
30
31 141 Some pharmaceutical companies had the fear that data might be accessed for competitive
32
33 142 advantage or to disprove trial results. In all the years of CSDR, these fears have been unfounded.
34
35 143 This should encourage other pharmaceutical companies to share their data.
36

37 144
38
39
40 145 Sharing of data on CSDR by academic funders is low and barriers to utilizing the strength of the
41
42 146 platform should be identified and addressed. The four academic funders who are CSDR members
43
44 147 are currently gathering feedback from their grantees about the challenges and support they need to
45
46 148 share clinical data. Other academic funders could encourage their grant holders to start sharing
47
48 149 their clinical research data utilizing this or similar platforms.
49

50
51 150

52 53 54 151 **Conclusions**

1
2
3 152 Sharing of anonymized, patient-level clinical trial data through platforms like CSDR advances
4
5 153 research and innovation. International pharmaceutical companies and academic funders are
6
7 154 making their data available for free to the research community in a transparent and equitable
8
9 155 manner. From our IRP perspective, while there are some areas for improvement for clinical data
10
11 156 sharing, to speed up the application process and to enable its value to be maximized, there are
12
13 157 numerous opportunities ahead. This is a readily available resource that we hope will continue to
14
15 158 develop to meet the needs of all stakeholders.
16
17
18
19
20
21
22 160

1
2
3 161 **Authors' contributions:** All authors were involved in the conception, drafting, and revision of
4
5 162 the article.

6
7
8 163 **Funding:** This research received no specific grant from any funding agency in the public,
9
10 164 commercial or not-for-profit sectors

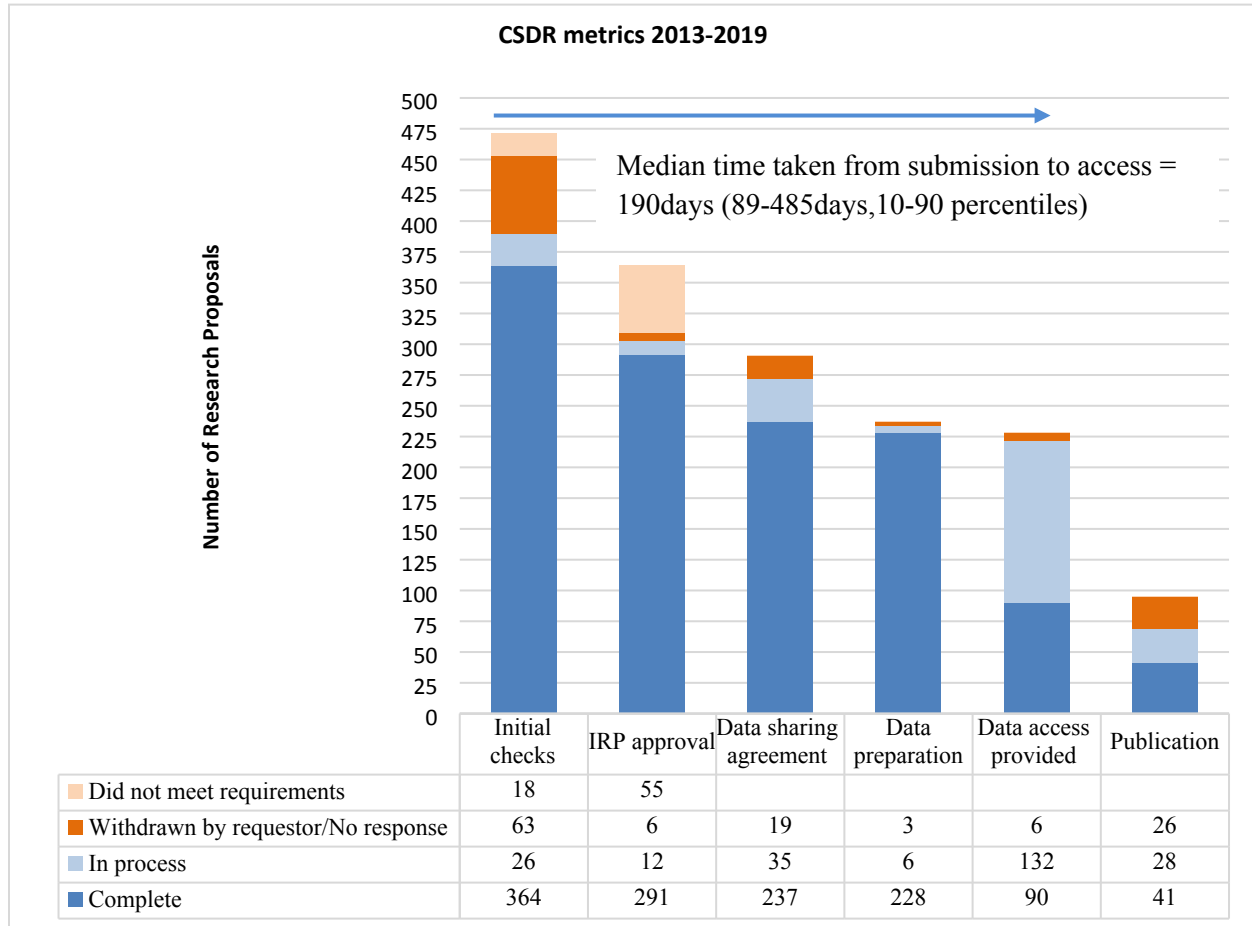
11
12 165 **Competing interests:** GH is the secretariat for the Independent Review Panel and all other authors
13
14
15 166 are members of the IRP.

16
17 167

18
19 168
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **169 References**
4

- 5
6 170 1. Clinical Study Data Request, Accessed on 12th June 19 at
7
8 171 <https://www.clinicalstudydatarequest.com/Default.aspx>
9
10 172 2. Institute of Medicine. Sharing Clinical Trial Data Maximizing Benefits, Minimizing Risk.
11
12 173 2015. Accessed on 12 June 19 at
13
14 174 [http://www.nationalacademies.org/hmd/~media/Files/Report%20Files/2015/SharingData/D](http://www.nationalacademies.org/hmd/~media/Files/Report%20Files/2015/SharingData/DataSharingReportBrief.pdf)
15
16 175 [ataSharingReportBrief.pdf](http://www.nationalacademies.org/hmd/~media/Files/Report%20Files/2015/SharingData/DataSharingReportBrief.pdf)
17
18
19 176 3. Kiley R, Peatfield T, Hansen J, Reddington F. Data Sharing from Clinical Trials -
20
21 A Research Funder's Perspective. *N Engl J Med*. 2017; 377(20):1990-1992. doi:
22 177 10.1056/NEJMs1708278.
23
24 178
25
26 179 4. Taichman DB, Backus J, Baethge C, et al. Sharing clinical trial data — a proposal from the
27
28 180 International Committee of Medical Journal Editors. *N Engl J Med* 2016; 374:384-386
29
30
31 181 5. Burton PR, Banner N, Elliot MJ, Knoppers BM, Banks J. Policies and strategies to facilitate
32
33 182 secondary use of research data in the health sciences. *Int J Epidemiol*. 2017; 46(6):1729-1733.
34
35 183 doi: 10.1093/ije/dyx195.
36
37
38 184
39
40 185
41
42 186
43
44 187
45
46 188
47
48 189
49
50 190
51
52 191
53
54
55
56
57
58
59
60

192 **Figure 1: Research Proposals progress through the CSDR system**

193

194 Of the 473 proposals submitted to CSDR (2013-Jan 2019), 123 were withdrawn by the requestor
 195 at some point through the process. The IRP rejected 55, but 34 of these went on to re-submit
 196 revised proposals following IRP suggestions for improvements. Of the 222 that gained access to
 197 the data (in process and completed) 41 have published at least one paper, with another 28
 198 expecting to publish soon.

199

200

BMJ Open

Clinical Trial Data Sharing: Here's the Challenge

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-032334.R2
Article Type:	Communication
Date Submitted by the Author:	25-Jul-2019
Complete List of Authors:	Kochhar, Sonali; Global Healthcare Consulting; Erasmus Medical Center, Dept Public Health Knoppers, Bartha; McGill University, Center of Genomics and Policy Gamble, Carrol; University of Liverpool, Biostatistics Chant, Alan; UK Clinical Research Collaboration Board Koplan, Jeffrey; Emory University, Emory Global Health Institute Humphreys, Georgina; Wellcome Trust, Open Research; University of Oxford Green Templeton College
Primary Subject Heading:	Communication
Secondary Subject Heading:	Health informatics
Keywords:	Data sharing, Clinical trials < THERAPEUTICS, Data re-use

SCHOLARONE™
Manuscripts

1 **Clinical Trial Data Sharing: Here's the Challenge**

2
3
4
5
6
7
8 3 Sonali Kochhar ^{a, b, c}, Bartha M Knoppers ^d, Carrol Gamble ^e, Alan Chant ^f, Jeffrey P Koplan ^g,
9
10 4 Georgina S Humphreys ^{h,i}

11
12
13
14
15 6 ^a Global Healthcare Consulting, New Delhi, India

16
17 7 ^b Department of Public Health, Erasmus MC, University Medical Center, Rotterdam, The
18
19 8 Netherlands

20
21
22 9 ^c Department of Global Health, University of Washington, Seattle, USA

23
24 10 ^d Center of Genomics and Policy, McGill University, Canada

25
26 11 ^e Department of Biostatistics, University of Liverpool, UK

27
28 12 ^f UK Clinical Research Collaboration Board, UK

29
30
31 13 ^g Emory Global Health Institute, Emory University, Atlanta, USA

32
33 14 ^h Wellcome Trust, London, UK

34
35 15 ⁱ Green Templeton College, University of Oxford, Oxford, UK: georginahumphreys@gmail.com

36
37
38
39
40
41
42 18 *Disclaimer:* The findings, opinions, conclusions, and assertions contained in this consensus
43
44 19 document are those of the individual members. They do not necessarily represent the official
45
46 20 positions of any participant's organization.

47
48
49
50
51 22 Word count 2,171

23 **Abstract**

24 Objective

25 Anonymized patient-level data from clinical research are increasingly recognized as a fundamental
26 and valuable resource. It has value beyond the original research project and can help drive
27 scientific research and innovations and improve patient care. To support responsible data sharing
28 we need to develop systems that work for all stakeholders. The members of the Independent
29 Review Panel (IRP) for the data sharing platform Clinical Study Data Request (CSDR) describe
30 here some summary metrics from the platform and challenge the research community on why the
31 promised demand for data has not been observed.

32 Summary of data

33 From 2014 to the end of Jan 2019, there were a total of 473 research proposals (RPs) submitted to
34 CSDR. Of these, 364 met initial administrative and data availability checks, and the IRP approved
35 291. Of the 90 research teams that had completed their analyses by Jan 2018, 41 reported at least
36 one resulting publication to CSDR. Less than half of the studies ever listed on CSDR have been
37 requested.

38 Conclusion

39 While acknowledging there are areas for improvement in speed of access and promotion of the
40 platform, the total number of applications for access and the resulting publications have been low
41 and challenge the sustainability of this model. What are the barriers for data contributors and
42 secondary analysis researchers? If this model does not work for all, what needs to be changed?
43 One thing is clear; that data access can realize new and unforeseen contributions to knowledge and
44 improve patient health, but this will not be achieved unless we build sustainable models together
45 that work for all.

1
2
3 46 CSDR is a consortium of 13 international pharmaceutical companies (GSK, Astellas Pharma,
4
5 47 Bayer, Chugai, Eisai, Novartis, ONO, Roche, Sanofi, Sunovion, Shionogi Inc., UCB, and ViiV)
6
7
8 48 and four academic research funders (The Wellcome Trust, The Bill & Melinda Gates Foundation,
9
10 49 The UK Medical Research Council and Cancer Research UK)¹. It was launched in 2013 and
11
12 50 currently lists anonymized patient data from 3374 studies on the platform, including 10 studies
13
14 51 from academic funders. The mandate is to reduce the barriers to access and re-use data, thereby
15
16 52 facilitating data sharing in an equitable, transparent, and independent manner.
17
18
19 53

20
21 54 Clinical trial data can be utilized beyond the original purpose for which it was generated, including
22
23 55 for analysis of new hypotheses, avoiding duplicative research, ensuring reproducibility, and to
24
25 56 drive scientific research and innovations to improve patient care. As the value of clinical data is
26
27 57 now widely recognized, many global initiatives actively promote and enable sharing of research
28
29 58 data and most funders mandate researchers to plan for sharing their data^{2,3,4}. The European
30
31 59 Medicines Agency and the National Institutes of Health have requirements in place for data
32
33 60 disclosure and clinical trial transparency⁵. Trial participants' confidentiality and privacy need to
34
35 61 be protected and the terms of consent to participate in research respected. Managed access systems
36
37 62 can help with this, including an IRP's review of all requests for data access and having data sharing
38
39 63 agreements in place which place appropriate restrictions on data usage, though it should be
40
41 64 recognized that these systems add time to the process from application to data access.
42
43
44
45
46
47
48

49 66 CSDR's system allows researchers to request access to anonymized global clinical trial data from
50
51 67 multiple studies and sponsors. All data requests that pass initial administrative and data availability
52
53 68 checks are reviewed by an IRP, the secretariat for which is provided by Wellcome. Once access is
54
55
56
57
58
59
60

1
2
3 69 granted nearly all CSDR members restrict data access to a secure online analysis environment
4
5 70 which they say ensures patient privacy whilst maintaining the utility of the data for secondary
6
7 71 analysis. This system can limit the merging of data from other non-CSDR sources and the range
8
9
10 72 of software available for researchers. Free access to data is usually granted for 12 months, with the
11
12 73 possibility of extension, and CSDR requires researchers to report on findings, which are then listed
13
14
15 74 on the website.
16
17 75

19 76 **CSDR metrics**

20
21 77 From 2014 to the end of Jan 2019, there were a total of 473 research proposals (RPs) submitted to
22
23 78 CSDR. Of these, 364 met initial administrative and feasibility checks from the sponsors. Although
24
25
26 79 the sponsors have a right to veto a request on the grounds of potential conflict of interest or
27
28 80 competitive risk, this veto has never been used. In reviewing applications, the IRP default position
29
30
31 81 is to provide access, and they have approved 291 (84% of those considered) and rejected 55.
32
33 82 Thirty-four of the rejected RPs were subsequently revised and re-submitted. The remaining RPs
34
35 83 are either still in process, withdrawn or no response has been received from the researchers (Figure
36
37
38 84 1). The most common reasons for rejection were unclear or unsuitable statistical methods (e.g. a
39
40 85 lack of detail on: the exact meta-analytic method proposed, how models will be validated, or how
41
42 86 data from different study designs and sites can be combined), too technical lay summaries, and
43
44
45 87 insufficient information presented.
46
47 88

48
49 89 Overall, the annual number of RPs submitted has remained fairly static with 70 submitted in 2014,
50
51 90 96(2015), 92(2016), 85(2017), and 97(2018). Researchers at institutions in over 30 different
52
53
54 91 countries have submitted RPs, although 168 RPs have been from researchers based in the USA
55
56
57
58
59
60

92 (35%) and very few have been submitted from researchers based in low- and middle-income
93 countries.

94
95 From January 2014 it became possible to request data from multiple sponsors and 17% (73/427)
96 of RPs have requested data from more than one sponsor. However, the median number of studies
97 per RP is 2 (1-5, 25th-75th percentile), with only a handful of RPs requesting a large number of
98 studies (the biggest request involved 192 studies and 11 sponsors).

99
100 Of the >4000 studies ever listed on CSDR (over the years studies have been listed and then
101 removed when sponsors leave the platform), 1457 have been requested. Interestingly the majority
102 (1157, 79%) have been requested only once or twice, but four have been requested more than 10
103 times (NCT00153062, NCT00268216, NCT00410384 which was always requested along with
104 NCT00424476).

105
106 Of the 90 research teams that completed their analysis by January 2018, 41 (45%) have at least
107 one publication, 28 (31%) are publishing soon and the remaining 21 have either not published or
108 not responded to reminders from CSDR. Although these numbers appear encouraging in terms of
109 converting access to data into publications of new findings, of concern are the 54 RPs (24%) whose
110 researchers were granted access but did not log into the analysis environment. It is puzzling why
111 this would happen, given the significant investment on both the part of the sponsors and the
112 researchers in getting the RPs to this point in the process. CSDR is planning to contact those
113 researchers to understand their constraints and challenges.

114

115 **Lessons learned**

116 From an IRP's perspective, the lessons learned from the CSDR experience include the following:

117
118 CSDR is a valuable resource of data from pharmaceutical companies and academic research
119 funders which is available for free for researchers. However, it is an expensive and resource-
120 intensive task for trial sponsors to provide access to data through this managed access model when
121 it involves secure analysis environments with licensed software, and this may challenge the long-
122 term sustainability of such a platform. With over 50% of studies never being requested, perhaps
123 more resources need to be focused on driving the re-use of data. Research agendas informed by
124 the whole community could drive the sharing and re-use of data for specific questions that are of
125 the highest priority for health practice, although this does limit the resource to current thinking on
126 what is most interesting. The research programmes of the Project Data Sphere cancer data sharing
127 platform is an example of how this model could work⁶.

128
129 Pharmaceutical companies and academic funders should pool resources to strengthen and
130 sustainably support data-sharing infrastructure and to develop and implement harmonized
131 principles, standards and best practices⁷. The portal cost for the sharing of data should not be
132 prohibitive for new data contributors.

133
134 It is important that there is a transparent system in place for data access decisions to maintain
135 equity for all those that want to re-use data. There is a minimum requirement for sufficient
136 statistical skills within the team requesting access to carry out the research proposed, and this may
137 mean there is currently a bias towards higher resourced settings. Funders should consider

1
2
3 138 supporting capacity building efforts to increase the data analysis expertise of teams based in low-
4
5 139 and middle-income settings.
6
7

8 140
9
10 141 The IRP and the secretariat support provided by Wellcome have been critical to ensuring a trusted
11
12 142 and independent managed access system. Having an experienced and multi-disciplinary IRP with
13
14 143 expertise in ethics, statistics, epidemiology, clinical research and a lay member has helped to
15
16 144 ensure that feedback is provided to the data requesters, including suggestions for how to improve
17
18 145 proposals with respect to the analysis methods and the clarity of how the research will benefit
19
20 146 patients, and to ensure all proposals receive consistent review. The IRP's perspective is to
21
22 147 encourage and facilitate data sharing (unless there are significant reasons not to do so). The
23
24 148 respected quality of the IRP service has been demonstrated by a newer data sharing platform
25
26 149 (Vivli), launched in 2018, requesting the Wellcome IRP also be available for data contributors to
27
28 150 their system. The IRP accepted this request and has already considered several proposals through
29
30 151 this platform too, applying the same criteria as for CSDR.
31
32
33
34
35
36

37 152
38 153 It would be helpful if consent for clinical research could, as far as possible, include provision for
39
40 154 re-use of their anonymized data beyond the original study⁷. In the absence of specific guidance,
41
42 155 institutional ethics committees should also adopt consistent policies for the need (or not) for ethics
43
44 156 review for secondary use of anonymized data. This would clarify if ethics review is needed (or
45
46 157 not), increase the amount of data available for re-use, and decrease the time to access the data.
47
48 158 Data generators might also feel reassured about using a file transfer model rather than restricting
49
50 159 to the use of a controlled analysis environment.
51
52
53

54 160
55
56
57
58
59
60

1
2
3 161 The data access process should be easily discoverable with transparent metrics for potential data
4
5 162 users. A common data sharing agreement should be available for all the data providers and once
6
7
8 163 researchers from an institution have signed the agreement, it should be applicable for other
9
10 164 researchers from that institution (to decrease the time often taken by institutions negotiating
11
12 165 changes to the agreement).

13
14 166
15
16
17 167 Merging multiple datasets is critical for finding small or subpopulation effects that could not have
18
19 168 been observed from any individual trial alone. However, the resource involved in pooling, or even
20
21 169 finding suitable data can be prohibitive if there are no common standards used. CSDR industry
22
23 170 sponsors mostly all use Clinical Data Interchange Standards Consortium (CDISC) standards⁸ but
24
25 171 the academic research community do not currently employ agreed standards which is barrier to
26
27 172 truly accessible and interoperable data.

28
29
30
31 173
32
33 174 The analysis environment must be easily usable by the researchers (including for merging multiple
34
35 175 datasets and statistical analysis) and not be expensive for the data providers. Increasing efficiency
36
37 176 in the process to ensure that data access happens as soon as possible benefits the researchers
38
39 177 (currently the median time from submission to data access is 190 days in CSDR). It is critical that
40
41 178 researchers that get access to the data, report their results in a timely manner (e.g. within six months
42
43 179 of completion of data access), so that it helps to move the research field forward, and reduce
44
45 180 research waste. This should be a requirement in the data sharing agreement.

46
47
48
49 181
50
51
52
53
54
55
56
57
58
59
60

1
2
3 182 Some pharmaceutical companies had the fear that data might be accessed for competitive
4
5 183 advantage or to disprove trial results. In all the years of CSDR, these fears have been unfounded.
6
7
8 184 This should encourage other pharmaceutical companies to share their data.
9

10 185
11
12 186 Sharing of data on CSDR by academic funders is low and barriers to utilizing the strength of the
13
14 187 platform should be identified and addressed. The four academic funders who are CSDR members
15
16 188 are currently gathering feedback from their grantees about the challenges and support they need to
17
18 189 share clinical data. Other academic funders could encourage their grant holders to start sharing
19
20 190 their clinical research data utilizing this or similar platforms.
21
22
23

24 191

25 26 192 **Conclusions**

27
28 193 Sharing of anonymized, patient-level clinical trial data through platforms like CSDR advances
29
30 194 research and innovation. International pharmaceutical companies and academic funders are
31
32 195 making their data available for free to the research community in a transparent and equitable
33
34 196 manner. Challenges still remain to speed up the process and enable data value to be maximized.
35
36 197 Researchers need incentives to share, such as citation of their data (which requires unique
37
38 198 identifiers to be embedded) being recognized by funders and institutions in decision making. The
39
40 199 costs of sharing and re-use need to decrease which will be helped by adoption of standards in the
41
42 200 creation of data, and reduced use of controlled analysis environments. Guidance from professional
43
44 201 bodies addressing, for example, consent issues and common data sharing agreements would help
45
46 202 promote data sharing. Despite these challenges great advances have already been made, and
47
48 203 models developed that mean there are more clinical trial data available for re-use than ever before.
49
50
51 204 We hope this field will continue to develop to meet the needs of all stakeholders.
52
53
54
55
56
57
58
59
60

1
2
3 205 **Authors' contributions:** SK, BK, CG, AC, JK, and GH were involved in the conception, drafting,
4
5 206 and revision of the article.
6

7
8 207 **Funding:** This research received no specific grant from any funding agency in the public,
9
10 208 commercial or not-for-profit sectors
11

12 209 **Competing interests:** GH is the secretariat for the Independent Review Panel and all other authors
13
14 210 are members of the IRP.
15
16 211

19 212 **References**

- 21 213 1. Clinical Study Data Request, Accessed on 20 July 19 at
22
23 214 <https://www.clinicalstudydatarequest.com/Default.aspx>
24
25 215 2. Institute of Medicine. Sharing Clinical Trial Data Maximizing Benefits, Minimizing Risk. 2015.
26
27 216 Accessed on 20 July 19 at
28
29 217 [http://www.nationalacademies.org/hmd/~media/Files/Report%20Files/2015/SharingData/DataShari](http://www.nationalacademies.org/hmd/~media/Files/Report%20Files/2015/SharingData/DataSharingReportBrief.pdf)
30
31 218 [ngReportBrief.pdf](http://www.nationalacademies.org/hmd/~media/Files/Report%20Files/2015/SharingData/DataSharingReportBrief.pdf)
32
33 219 3. Kiley R, Peatfield T, Hansen J, Reddington F. Data Sharing from Clinical Trials -
34
35 220 A Research Funder's Perspective. N Engl J Med. 2017; 377(20):1990-1992. doi:
36
37 221 10.1056/NEJMsbl708278.
38
39 222 4. Taichman DB, Backus J, Baethge C, et al. Sharing clinical trial data — a proposal from the
40
41 223 International Committee of Medical Journal Editors. N Engl J Med 2016; 374:384-386
42
43 224 5. National Institute for Health sharing policies. Accessed on 16 July 2019 at
44
45 225 <https://grants.nih.gov/policy/sharing.htm>
46
47 226 6. Project Data Sphere. Accessed 20 Jul 2019 at
48
49 227 <https://www.projectdatasphere.org/projectdatasphere/html/home>
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 228 7. Burton PR, Banner N, Elliot MJ, Knoppers BM, Banks J. Policies and strategies to facilitate
4
5 229 secondary use of research data in the health sciences. *Int J Epidemiol.* 2017; 46(6):1729-1733. doi:
6
7 230 10.1093/ije/dyx195
8
9 231 8. CDISC Standards in the Clinical Research Process. Accessed on 20 July 19 at
10
11 232 <https://www.cdisc.org/standards>
12
13
14
15 233

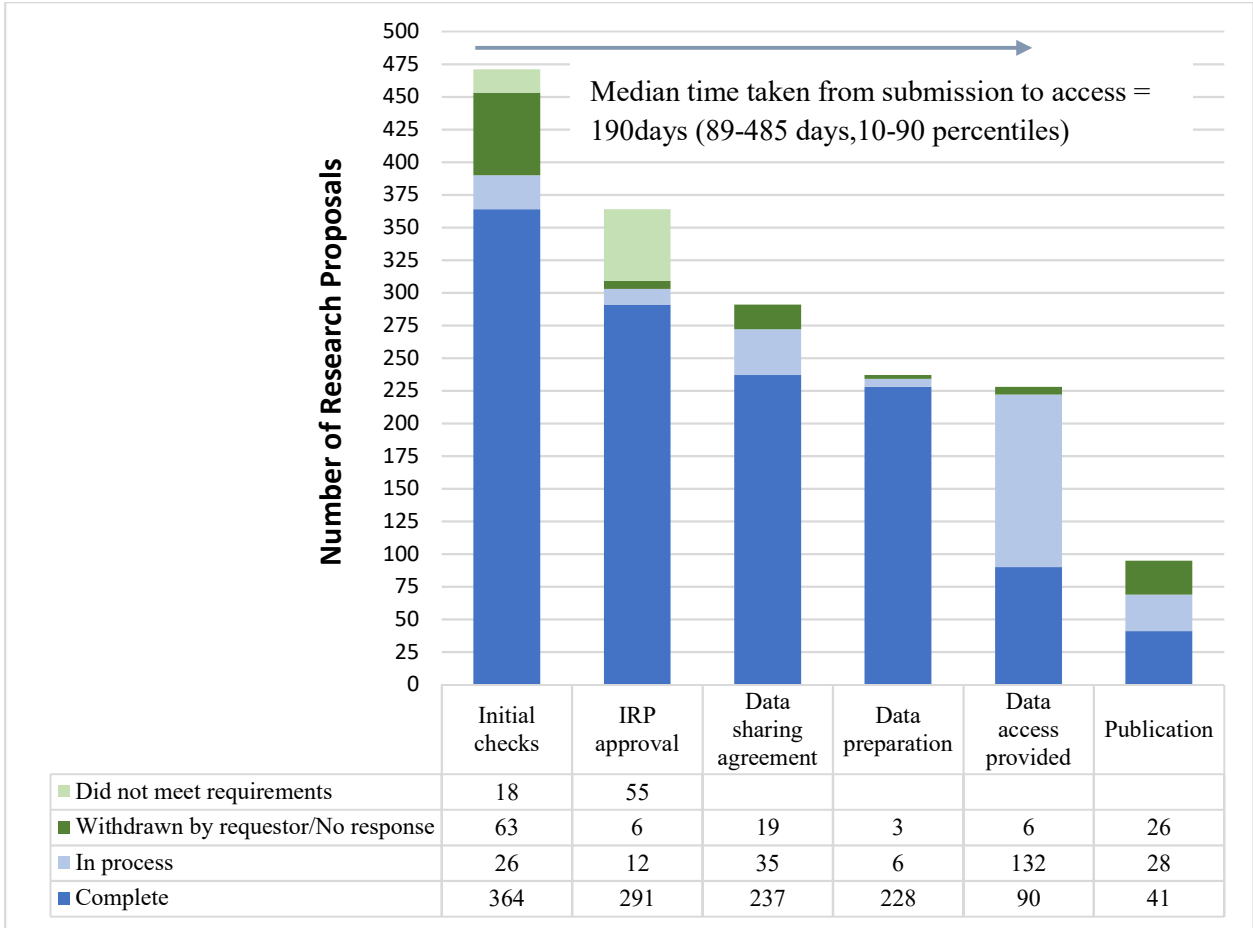
16
17 234 **Figure 1: The progress of Research Proposals through the CSDR system**

18
19 235 Of the 473 proposals submitted to CSDR (2013-Jan 2019), 123 were withdrawn by the requestor
20
21 236 at some point through the process. The IRP rejected 55, but 34 of these went on to re-submit
22
23 237 revised proposals following IRP suggestions for improvements. Of the 222 that gained access to
24
25 238 the data (in progress and completed) 41 have published at least one paper, with another 28
26
27 239 expecting to publish soon.
28
29
30

31 240

32
33
34 241
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



view only