Clinical study reports published by the European Medicines Agency 2016–2018: a cross-sectional analysis

David Byrne, Ciaran Prendergast, Tom Fahey, Frank Moriarty


STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This study characterises and describes the entire selection of clinical study reports made publicly available by the European Medicines Agency (EMA).
⇒ We use clinical study report (CSR) filenames to identify and navigate individual drug trials and document types.
⇒ This study examines the timeliness of publication of results of pivotal trials submitted to the EMA for regulatory drug approval, and compares the availability of their results across CSRs, journals and clinical trial registries.
⇒ Given the large volume of documents and trials, completing a thorough analysis of the characteristics and contents within the available CSRs was not feasible.
⇒ While the methods for identifying matched publications were rigorous and independently checked by three authors, some publications may not have been identified.

INTRODUCTION

Background and rationale

There is a growing body of evidence related to reporting biases in clinical trials which has been highlighted across a wide variety of treatments including antidepressant and oncological medications.1–7 Published studies of clinical trials are biased towards positive results.5,8 Considering only published studies has been shown to increase the apparent clinical effect size by one-third.5

Clinical study reports (CSRs) and other documents used to obtain regulatory approval for medications provide a rich and authoritative description of clinical trials, including details on study design, methods and results which may be absent or inaccurate in published papers.9–13 CSRs have been an increasing focus in research.14 The benefits of using such unpublished information in evidence synthesis and the subsequent impact on clinical efficacy estimates, clinical recommendations and decision-making have been demonstrated.10–13 15–17 Even when access to this more detailed information on clinical trials is possible, challenges have been highlighted, including inefficiencies in navigating vast amounts of documents and incomplete and delayed access to the information.18 There are ongoing calls for prompt, unrestricted access to CSRs to allow for a more robust appraisal of trials of new medications.18–21
EMA clinical trial publication and request policies

Since 2010, the EMA has released clinical trial reports, which include CSRs, on request under its ‘access to documents’ policy (EMA policy 0043). However, requesting CSRs by this method has been shown to be a long process with lengthy delays.

On 1 January 2015 the EMA introduced the first phase of a new policy involving the publication of clinical data for newly approved medications, entitled ‘EMA Policy 0070’. This policy requires that approval documents submitted to the EMA as part of a Marketing Authorisation Application (MAA) by a drug company, be made publicly available for non-commercial purposes. The first documents were published in October 2016 on the EMA's Clinical Data website; however, in December 2018, the EMA temporarily suspended the automatic publication of these regulatory documents, citing a need to prioritise implementation of their Business Continuation Plan, a contingency plan which relates to the UK’s withdrawal from the European Union (EU). To date this process remains suspended due to the impact of the COVID-19 pandemic, with the exception being an agreement to publish clinical trial information on COVID-19 medicines, including vaccinations.

Pivotal CSRs and drug approval pathways

It is generally recommended that drug approvals should be based on at least two clinical trials, often described as pivotal trials. The exact definition of a pivotal trial however is unclear, with a paucity of information available and no clear consensus between regulators. The increasing number of drug approvals based on a single pivotal trial, particularly for orphan drug designation, has been highlighted as a concern with the preference being for a more rigorous evidence base. CSRs of trials submitted to the EMA are labelled as pivotal during the submission process as advised by the regulator. Characteristics of the pivotal studies available on the EMA's Clinical Data website have yet to be summarised, with similar information previously described for such submissions to the Food and Drug Administration (FDA).

Publication of trial results

Despite the importance of accessing clinical trial results promptly, delays in the publication of clinical reports by the EMA under Policy 0070 have recently been highlighted. Paludan-Müller et al used the published information available from the EMA resource to describe the characteristics of the available trials, specifically the delay in publication of trial information relative to the EMA's planned timelines. They found that for only two medications, the EMA published the information within the recommended timelines, with a median delay of 446 days.

Overall, the EMA Clinical Data website represents a significant source of evidence for research purposes. Greater access to unpublished clinical trial information would allow for such rich sources of information to be used more widely in research, including evidence synthesis and decision making. Several studies have described the characteristics of unpublished trials using non-random samples of CSRs. To date, as far as the authors are aware, a comparison of the availability of clinical trial results from the EMA and from conventional published sources, including clinical trial registries and journal publication, has yet to be undertaken. This is relevant for those researchers hoping to include such additional trial information in their analyses, including evidence synthesis.

Objectives

The objectives of this study are:

- To describe the characteristics of the available CSRs and related regulatory documents, which have been published by the EMA on its Clinical Data website since 2016, summarised by medication characteristics, trial characteristics and document characteristics.
- To describe the characteristics of those trials labelled as ‘pivotal’, specifically document length and trial phase.
- To report the availability and timeliness of access to results from these pivotal trials across three sources; CSRs, trial registries and journal publications.

METHODS

Study design

This is a cross-sectional study of all documents submitted to the EMA by a Marketing Authorisation Applicant for regulatory drug approval, published the EMA. This study follows the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline for cross-sectional studies and the relevant STROBE checklist is available in online supplemental materials.

Data sources and data extraction

CSR documents

Summary information on each medication for all regulatory submissions published by the EMA under policy 0700 were downloaded in comma-separated value format from the EMA's Clinical Data website. Such information included authorisation type, approval pathway (including specific additional pathways such as submissions under orphan status), procedure type and Anatomical Therapeutic Chemical (ATC) classification code. Individual CSR documents for each of the medications were also downloaded and saved in PDF format.

The pdftools package in ‘R’ statistical software was used to extract individual CSR document filenames and page length for all documents for each medication. Using the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) structure and EMA guidance on the naming of CSR documents, filenames were used to identify individual trials with which the documents were associated. Filenames were screened for substrings which identified individual trials by their trial identification number, the specific document type
and whether the trial was classed as pivotal ‘p-CSR’ or supportive ‘s-CSR’ (eg, m5351-(trialid)-p-csr-body, m5351-(trialid)-p-app1612-crf).

For the subset of trials labelled as ‘pivotal’ CSRs, using a Microsoft Excel spreadsheet, further information was extracted from each document on date of completion of the trial’s CSR (the final completion date reported within the CSR) and date of CSR publication by the EMA. Some trials were listed as a combination of phase 1, 2 and 3 trials (eg, phase 1b/2, phase 2/3), however, in these cases for the purposes of analysis in this study, the higher of the two phases was used.

Journal publications and trial registries
For pivotal studies with an included CSR, literature databases (Medline (PubMed), Embase and Google Scholar) were searched for the primary journal publication for each trial. Searches were conducted initially using the trial ID. If no publication was found, searches were conducted using the ClinicalTrials.gov or European Union Drug Regulating Authorities Clinical Trials Database (EudraCT) trial registry numbers or a combination of the active drug substance and the number of randomised trial participants. The single earliest primary publication reporting results related to the individual trial was used where available. Pooled analyses were not included. For trials with no primary publication found, the first publication which included trial results was used (including where results were reported along with those from other related trials). For CSRs which were reports of interim analyses, the publication with the closest length of follow-up to the CSR was chosen. Length of journal publication and first publication date were extracted, using the main body of the publication excluding appendices.

Similarly, the clinical trial registry Clinicaltrials.gov was searched to identify entries corresponding to trials. Initially, the registry was searched using trial ID number and if not found, then searches were conducted using the drug substance, filtered by phase, with the registrations checked by matching criteria including trial dates and number of participants. If no registration was found then the WHO International Clinical Trials Registry Platform was searched using the trial ID number. The presence of a corresponding journal publication and/or trial results on clinical trial registries were documented, along with relevant publication date for these stages and the number of pages of the main text for the trial publication.

Initial searches of registry and publication sources were conducted by one author (DB). A second author (CP) verified the extracted information and repeated the search for trials where no registry entry or publication were identified. Outstanding trials at that stage were searched for by a third author (FM). Discrepancies were discussed between authors and resolved by consensus. All data analysed in our study are available in the Zenodo repository.32

Statistical methods
Using Stata software V.15,43 descriptive statistical analysis was carried out on the extracted data at the medication, trial and document levels, as well as for the pivotal trials. Median and IQRs as well as mean and SD were reported for relevant outcomes.

Analysis of medication, trial and document characteristics
All medications for which information were available were included in the initial overall descriptive characteristics. Medications were then excluded from further analysis if there were no documents provided, if none of the documents provided were related to individual trials, or if the documentation for a medication entirely duplicated those of another medication. These were then excluded from further analysis in order to obtain more accurate characteristics at the trial and document levels.

The number of documents, total number of pages and number of trials per medication were calculated, as well as the total number of pivotal and supportive trials and the total number of pages per document type. Differences in number of trials and documents per medication, by medication characteristics were assessed using the Wilcoxon rank-sum test (Mann-Whitney U) for two-sample comparisons and Kruskal-Wallis test for multisample comparisons.

Negative binomial regression analyses were used to investigate the association between medication characteristics and the number of pivotal trials per medication. This model was selected over a Poisson model due to over-dispersion in the variance compared with the mean count of pivotal trials. Univariate negative binomial regression analyses were initially conducted, followed by multivariable analysis including all factors, generating incident rate ratios (IRR) and 95% CIs. Medication characteristics included procedure type, approval pathway, generic status and therapeutic area (based on ATC code). These were chosen as they had shown significant associations with number of overall trials in previous non-parametric analyses.

The classification of each document by module type was extracted. A module is a component of a regulatory submission as outlined by the ICH. In addition, the number and proportion of trials which included separate documents for CSR body, statistical analysis plan (SAP), case report form (CRF) and protocol were reported, as well as those which included all four document types. At the document level, the total number of pages available from the EMA, as well as the number of pages per document type were reported.

Inclusion process for pivotal trials
Any pivotal trials which were also included in another submissions were excluded as duplicate trials. Non-randomised trials were also excluded.
Descriptive statistics of pivotal trials

For the subgroup of pivotal trials, the trial phase reported in the CSR was summarised. The proportion of all medications which were supported by a single pivotal trial was calculated and also the proportion of these studies by trial phase.

Publication availability and timeliness

The availability and timeliness of access to the results of these pivotal trials were also described, by calculating the length of time in days between the documented date of CSR completion, the date of release of documents by the EMA (herein defined as ‘EMA publication’) and publication of the clinical trial results in a journal publication or clinical trial registry. For those trials with both a journal publication and with clinical trial registry results, the earlier of the two dates was used. These time periods were plotted using both time from EMA publication and CSR completion dates as reference points.

CSR and journal publication length

We also calculated the ‘compression factor’ of included trials, which is a ratio of the page length of the pivotal CSR documents to that of the matched journal publications. This measure has been used previously to represent the compression and loss of information between a CSR and a typical journal publication.9 For trials whose CSRs bodies were in multiple parts, the total number of pages for all parts were used. For trials whose results were published in journals along with results of related trials for the same medication, the publication length was divided by the number of trials included in the paper.

Public and patient involvement

Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

RESULTS

Descriptive statistics of medication, trial and document characteristics

Medication characteristics

In total, there were regulatory drug approval submissions for 142 medications published on the EMA’s Clinical Data website between 2016 and 2018. The characteristics of these medications are shown in table 1. The majority of applications for these medications were for initial marketing authorisations and authorised products. There were 45 approvals via alternative or special pathways, the largest proportion of these (57.8%) were medications under orphan designation, which comprised 18.3% of all medications. By ATC classification, antineoplastic and immunomodulating agents accounted for the largest proportion (40.9%) of medications.

The process for including medications in further analysis is demonstrated in figure 1. Out of the 142 submissions, 14 were excluded as there were no documents provided. Four medications were excluded as no documents relating to trials were provided. Of these, the documents provided included only information on systematic

### Table 1 Characteristics of all medications (n=142)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Number of studies, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EMA publication year</strong></td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>6 (4.2)</td>
</tr>
<tr>
<td>2017</td>
<td>58 (40.9)</td>
</tr>
<tr>
<td>2018</td>
<td>78 (54.9)</td>
</tr>
<tr>
<td><strong>Product status</strong></td>
<td></td>
</tr>
<tr>
<td>Authorised</td>
<td>131 (92.2)</td>
</tr>
<tr>
<td>Withdrawn</td>
<td>11 (7.8)</td>
</tr>
<tr>
<td><strong>Procedure type</strong></td>
<td></td>
</tr>
<tr>
<td>Initial marketing authorisation</td>
<td>91 (64.1)</td>
</tr>
<tr>
<td>Extension of indication</td>
<td>44 (31)</td>
</tr>
<tr>
<td>Line extension</td>
<td>5 (3.5)</td>
</tr>
<tr>
<td>Workshare</td>
<td>2 (1.4)</td>
</tr>
<tr>
<td><strong>Generic status</strong></td>
<td></td>
</tr>
<tr>
<td>Non-generic/biosimilar</td>
<td>107 (75.4)</td>
</tr>
<tr>
<td>Generic</td>
<td>28 (19.7)</td>
</tr>
<tr>
<td>Biosimilar</td>
<td>7 (4.9)</td>
</tr>
<tr>
<td><strong>Additional pathway (n=45)</strong></td>
<td></td>
</tr>
<tr>
<td>Conditional approval</td>
<td>5 (3.5)</td>
</tr>
<tr>
<td>Exceptional circumstances</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td>Orphan status</td>
<td>26 (18.3)</td>
</tr>
<tr>
<td>Article 58</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td>Withdrawn procedure</td>
<td>12 (8.4)</td>
</tr>
<tr>
<td><strong>Therapeutic area (ATC code)</strong></td>
<td></td>
</tr>
<tr>
<td>Antineoplastic and immunomodulating agents (L)</td>
<td>58 (40.9)</td>
</tr>
<tr>
<td>Anti-infectives for systemic use (J)</td>
<td>23 (16.2)</td>
</tr>
<tr>
<td>Alimentary tract and metabolism (A)</td>
<td>18 (12.7)</td>
</tr>
<tr>
<td>Blood and blood forming organs (B)</td>
<td>9 (6.3)</td>
</tr>
<tr>
<td>Nervous system (N)</td>
<td>7 (5)</td>
</tr>
<tr>
<td>Cardiovascular system (C)</td>
<td>6 (4.2)</td>
</tr>
<tr>
<td>Genitourinary system and sex hormones (G)</td>
<td>4 (2.8)</td>
</tr>
<tr>
<td>Musculoskeletal system (M)</td>
<td>4 (2.8)</td>
</tr>
<tr>
<td>Various (V)</td>
<td>4 (2.8)</td>
</tr>
<tr>
<td>Systemic hormonal preparations, excl. sex hormones and insulins (H)</td>
<td>3 (2.1)</td>
</tr>
<tr>
<td>Respiratory system (R)</td>
<td>3 (2.1)</td>
</tr>
<tr>
<td>Sensory organs (S)</td>
<td>2 (1.4)</td>
</tr>
<tr>
<td>Dermatologicals (D)</td>
<td>1 (0.7)</td>
</tr>
</tbody>
</table>

*There were 45 submissions which were by additional pathway. ATC, Anatomical Therapeutic Chemical; EMA, European Medicines Agency.
literature reviews, meta-analyses, pooled analyses and immunogenicity reports and for one medication all available documents were exemption notices. There were five medications which were excluded as duplicate submissions, in that identical CSR documents were published for two product submissions, with a different product name used for each. After these exclusions, there were 119 unique medication submissions included in the further analysis. Table 2 summarises the trials and documents at the medication level. There was a median of 15 (IQR 5–46) documents, 5 (IQR 2–14) trials and 9629 (IQR 2711–26 673) pages per medication submission for drug approval. The median number of pivotal trials per medication was 2 (IQR 1–3). For 46.2% (55/119) of medications there was just a single pivotal trial submitted for regulatory drug approval. In contrast, one medication had 17 pivotal trials included in its submission.

There was a wide variation in the numbers of trials and pages per medication when stratified by different approval pathways and therapeutic areas (online supplemental table S1). Following Wilcoxon rank-sum and Kruskal-Wallis tests for comparisons, there was a significant difference in the median number of trials per medication for procedure types and conditional approval pathways. There was also a significant difference in the median number of pages per medication for conditional approval and for therapeutic class (online supplemental table S1).

In regression analysis of the number of pivotal trials per medication, there was an increased number of median trials per medication for conditional approval (online supplemental table S1).

---

**Figure 1** Inclusion and exclusion of medications and trials in analyses. CSRs, clinical study reports; EMA, European Medicines Agency; pCSR, pivotal CSR.
Pivotal trials associated with initial marketing authorisation (IRR 2.09 (95% CI 1.48 to 2.95)) compared with extension of indication. There was an increased number of pivotal trials associated with anti-infective therapeutic class (IRR 1.89 (95% CI 1.31 to 2.75)) compared with the therapeutic class reference category (antineoplastic and immunomodulating agents). Generic status was significantly associated with a lower number of pivotal trials per medication (IRR 0.31 (95% CI 0.18 to 0.55)). Results of regression analyses are shown in table 3.

**Trial characteristics**

In total, 2115 trials were included in the analysis. There were 1678 (79.3%) studies identified as supportive (sCSR) and 264 (12.5%) as pivotal (pCSR). The remaining 173 studies (8.2%) were neither tagged as supportive nor pivotal and they mostly comprised in-vitro studies and pharmacokinetic reports. There was a median of 1 (IQR 1–4) document per trial and a median of 336 (IQR 21–1192) total number of pages per trial. There was a significant difference in the median number of pages per trial for procedure types, orphan and biosimilar approvals and for therapeutic class (online supplemental table S2).

**Document characteristics**

Across all 142 drug regulatory submissions, there were 6317 documents published by the EMA with a total of 3155354 pages. Online supplemental table S3 summarises these documents by module type. There were 16 instances in which the protocol, SAP and CRF were included in one document and were therefore identified as individual documents in our analysis (online supplemental figure S1) meaning there were 6301 individual identified documents included in this analysis. Of these,

| Table 2 Characteristics of medications included in further analyses (n=119) |
|-----------------------------------------------------|-----------------|-----------------|-----------------|
| **Count of trials per medication** | **Number of medications (n=119)** | **%** |
| 1 | 24 | 20.2 |
| 2 | 16 | 13.4 |
| 3 | 8 | 6.7 |
| 4 | 7 | 5.9 |
| 5 | 7 | 5.9 |
| 6 | 5 | 4.2 |
| 7 | 6 | 5.0 |
| 8 | 4 | 3.4 |
| 9 | 6 | 5.0 |
| 10–19 | 12 | 10.1 |
| 19–50 | 12 | 10.1 |
| 50+ | 12 | 10.1 |

| **Count of pivotal trials per medication** | **Number of medications (n=119)** | **%** |
| 0 | 2 | 1.7 |
| 1 | 55 | 46.2 |
| 2 | 32 | 26.9 |
| 3 | 13 | 10.9 |
| 4 | 6 | 5.1 |
| 5 | 2 | 1.7 |
| 6 | 2 | 1.7 |
| 7 | 3 | 2.5 |
| 8 | 2 | 1.7 |
| 11 | 1 | 0.8 |
| 17 | 1 | 0.8 |
2465 (39.1%) were CSR body documents, 1182 (18.8%) were protocols, 840 (13.3%) were SAPs, 794 (12.6%) were sample CRFs, 118 (1.9%) were integrated summaries of safety, 53 (0.8%) were integrated summaries of efficacy and 849 (13.5%) were other documents (online supplemental figure S1). The median number of pages per document was 92 (IQR 23–321) which varied by document types (table 4).

### Table 4  Length of documents by document type

<table>
<thead>
<tr>
<th>Document type</th>
<th>Median (IQR)</th>
<th>Range</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All documents</td>
<td>92 (23–321)</td>
<td>1–42 084</td>
<td>503.73 (1938.2)</td>
</tr>
<tr>
<td>CSR body</td>
<td>181 (21–679)</td>
<td>2–33 943</td>
<td>744.6 (1725.7)</td>
</tr>
<tr>
<td>Protocol</td>
<td>101 (1–199)</td>
<td>1–2333</td>
<td>169.5 (239.4)</td>
</tr>
<tr>
<td>Statistical analysis plan (sap)</td>
<td>46 (23–95)</td>
<td>1–42 084</td>
<td>543.0 (3960.3)</td>
</tr>
<tr>
<td>Case report form (CRF)</td>
<td>123 (64–297)</td>
<td>1–11 174</td>
<td>273.6 (646.7)</td>
</tr>
<tr>
<td>Integrated summary of efficacy (ISE)</td>
<td>106 (36–321)</td>
<td>1–9964</td>
<td>826.8 (2179.1)</td>
</tr>
<tr>
<td>Integrated summary of safety (ISS)</td>
<td>622 (46–2104)</td>
<td>1–11 882</td>
<td>1480.5 (1980.5)</td>
</tr>
</tbody>
</table>

Of 2115 unique trials, the majority, 1945 (91.9%), included separate CSR body documents, however just 757 (35.8%) included all four of CSR body, CRF, SAP and protocol documents separately (online supplemental table S2).

Descriptive statistics of pivotal trials

Of the 119 unique submissions to the EMA for drug approval, there were a total 264 trials labelled as pivotal. Following exclusions of duplicates and of one trial which was excluded as it was not a randomised trial, there were 243 pivotal trials included in the analysis (figure 1).

Of these 243 pivotal trials, there were 148 (60.9%) phase 3 trials, 48 (19.8%) phase 2 trials, 45 (18.5%) phase 1 trials and two (0.8%) phase 4 trials (figure 2, online supplemental figures S2 and S3, table S4). Of the 119 included submissions, there were 55 medications (46.2%) for which their submission was supported by a single pivotal trial. One submission was based on a phase 4 trial, 31 were based on a single phase 3 trial and 6 were based on a single phase 2 trial. Sixteen submissions were based on a single pivotal phase 1 trial, which amounted to 13.4% of all 119 approvals. Of the 16 trials, 13 were related to initial marketing authorisation and 12 were for generic medications (online supplemental figure S4). Fifteen (93.8%) of these submissions based on a single pivotal phase 1 trial had no other supportive trials included in the application; one submission did include two further supportive phase 3 trials.

Publication availability and timeliness

Of the 243 pivotal trials, results were not identified on ClinicalTrials.gov for 68 (28.0%) trials, with four of these found on other clinical trials registries; three on the EU Clinical Trials Register and one on the Japanese Primary Registries Network. For 64 trials (26.3%), no trial registry entries or results were found. Journal publications were not located for 41 (16.9%) trials, of these 33 (80.5%) were phase 1, 2 (4.9%) were phase 2 and 6 (14.6%) were phase 3 (figure 2, online supplemental table S4). There were 33 trials (13.6%) for which neither clinical trial registry results, nor journal publications were available.

For 94.2% (195/207) of pivotal trial CSRs with a matched source, results were available in a publication or registry prior to EMA publication. However, for 5.8% (12/207) of matched pivotal trials, the EMA publication was the earliest source of results, available a median of 523 days (IQR 363–882 days) before the earliest corresponding publication in a journal or clinical trial registry (figure 3, table 5).

When dates of CSR completion were compared with matched published sources, CSRs were found to have been completed before journal or trial registry publication in 176 (85.0%) pivotal trials. These CSRs had been completed a median of 338 days (IQR 173–554 days, range 10–2367 days) before the date of journal or registry source publication (figure 3, table 5).

CSR and journal publication length

The median document length for pivotal CSR body documents was 1392 (IQR 545–3099) pages with a mean of 2426.5 (SD 3 016.8) pages. For journal trial publications, the median length was 10 (IQR 9–12) pages with a mean of 10.6 (SD 3.2) pages per publication. The compression...
factor ranged from 9.7 to 3361, and the median was 164.2 (IQR 76.3–316.7), with a mean of 277.4 (SD 358.2). Median compression factor was larger for extension of indication and was smaller for trials for medications under exceptional circumstances (online supplemental table S5).

**DISCUSSION**

**Key results**

There was a median of 15 (IQR 5–46) documents, 5 (IQR 2–14) trials and 9629 (IQR 2711–26,673) pages per submission for drug approval published by the EMA between 2016 and 2018. Out of 119 unique drug regulatory approval submissions, 46.2% were supported by a single pivotal trial and 13.4% were based on a single pivotal phase 1 trial, which were mostly for generic medications. There were 33 trials (13.6%) for which neither clinical trial registry results, nor journal publications were available. For 5.8% of pivotal trials, the EMA publication was the earliest information source for trials, with a median time difference of 523 (IQR 363–882) days between EMA and earliest journal or registry publication.

**Interpretation and comparison with existing literature**

We report a median of 15 (IQR 5–46) documents and 5 (IQR 2–14) trials per medication, which are similar to results reported by Paludan-Müller et al. However, we included all available study types in our analysis and reported 2115 unique trials, compared with 1005 reported by Paludan-Müller et al who did not include smaller clinical studies.

This study reports a median document length of 92 (IQR 23–321, range 1–42,084) pages across all documents and 181 (IQR 21–679) pages for CSR body documents. These estimates are shorter compared with those in a similar study by Doshi and Jefferson which reported a median document length of 644 pages (range 9–15 440). This may be due to the fact that Doshi and Jefferson used a non-random sample of CSRs requested from various sources. In addition, our study used all available trials, including supportive, phase 1 and undesignated trials which were shorter documents.

We reported an overall median compression factor of 164.2 (IQR 76.3–316.7, range 9–3361) using a sample of just pivotal trials for this estimation. We also report how such a measure varies between types of approval pathway. Doshi and Jefferson report a median compression factor range of 1–1221 for a non-random sample of CSRs, including all trial types, including supportive and undesignated trials. Schroll et al reported a compression factor of 71–270 in a sample which included just module 1 and 2 documents. These results highlight that compression factor in itself is a crude measure of the degree of condensing of trial information.

To our knowledge, this is the first study which summarises characteristics of all pivotal trials for which the EMA have published CSRs and related documentation since 2016. This study reports a median compression factor of 164.2 (IQR 76.3–316.7, range 9–3361) using a sample of just pivotal trials for this estimation. We also report how such a measure varies between types of approval pathway. Doshi and Jefferson report a median compression factor range of 1–1221 for a non-random sample of CSRs, including all trial types, including supportive and undesignated trials. Schroll et al reported a compression factor of 71–270 in a sample which included just module 1 and 2 documents. These results highlight that compression factor in itself is a crude measure of the degree of condensing of trial information.

To our knowledge, this is the first study which summarises characteristics of all pivotal trials for which the EMA have published CSRs and related documentation since 2016. This study reports a median compression factor of 164.2 (IQR 76.3–316.7, range 9–3361) using a sample of just pivotal trials for this estimation. We also report how such a measure varies between types of approval pathway. Doshi and Jefferson report a median compression factor range of 1–1221 for a non-random sample of CSRs, including all trial types, including supportive and undesignated trials. Schroll et al reported a compression factor of 71–270 in a sample which included just module 1 and 2 documents. These results highlight that compression factor in itself is a crude measure of the degree of condensing of trial information.

**Table 5** Time difference and CSR delays (days)

<table>
<thead>
<tr>
<th>Time difference from EMA CSR publication (n=207)</th>
<th>Median (IQR)</th>
<th>Range</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>To journal publication</td>
<td>–715 (–1128, –420)</td>
<td>–3563 to 1320</td>
<td>–851.2 (855.1)</td>
</tr>
<tr>
<td>To registry publication</td>
<td>–511.5 (–812.5, –289)</td>
<td>–2987 to 1267</td>
<td>–621.4 (759.5)</td>
</tr>
<tr>
<td>To earliest source</td>
<td>–737 (–1203, –529)</td>
<td>–3234 to 1267</td>
<td>–946.7 (843.5)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time difference from CSR completion (n=207)</th>
</tr>
</thead>
<tbody>
<tr>
<td>To journal publication</td>
</tr>
<tr>
<td>377 (114–654)</td>
</tr>
<tr>
<td>To registry publication</td>
</tr>
<tr>
<td>473 (288–684)</td>
</tr>
<tr>
<td>To earliest source</td>
</tr>
<tr>
<td>261 (91–480)</td>
</tr>
<tr>
<td>To EMA publication</td>
</tr>
<tr>
<td>1028 (898–1340)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Delay from EMA CSR publication (n=12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>To earliest source</td>
</tr>
<tr>
<td>523 (363–882)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Delay from CSR completion (n=176)</th>
</tr>
</thead>
<tbody>
<tr>
<td>To earliest source</td>
</tr>
<tr>
<td>338 (173–554)</td>
</tr>
</tbody>
</table>

CSR, clinical study report; EMA, European Medicines Agency.
pivotal trial, the EMA advise that the study should be ‘exceptionally compelling’ with a particularly high standard of trial design with respect to internal and external validity, clinical relevance, statistical significance, data quality and internal consistency. However, in this study, almost one in five studies (18.5%) identified as pivotal trials were phase 1 studies. Our results show that 46.2% of all EMA regulatory approvals between 2016 and 2018 were based on a single pivotal trial. These align with results published by Morant and Vestergaard who reported that 45% of newly approved medications by the EMA between 2012 and 2016 were based on a single pivotal trial and Downing et al who report a figure of 37% for approvals by the FDA between 2005 and 2012.

In addition, in our study, of those submissions supported by single trials which were highlighted as pivotal, almost one-third (29.6%) were phase 1 trials, meaning that 13.4% of all submissions to the EMA were based on a single pivotal phase 1 trial. This represents an increase in the proportion of approvals based solely on uncontrolled trials. Hatswell et al reported that between 1999 and 2014, 44 out of 795 (5%) approvals issued by the EMA were based on uncontrolled trials. While the concept of pivotal bioequivalence trials is described, this concept would be at odds with the more conventional definitions of pivotal trials and also seem to conflict with the EMA’s own guidance on pivotal trials that they should provide compelling evidence of efficacy and safety for the medications in question.

Our results show that for 16.9% of pivotal trials, no journal publication was found and for 13.6%, neither clinical trial registry results nor journal publications were available. While many of these were for phase 1 trials, it may be reasonable to expect that if a trial was defined as pivotal, such trial results should be available from conventional published sources. In addition, there were some phase 3 trials for which CSRs were the only available information source.

Our results also show that for 5.8% of pivotal trials, publication of CSRs by the EMA was the earliest source of trial information, available a median of 523 days (1.43 years) before the earliest publication or trial registry results. Paludan-Müller et al noted a median delay of 1.21 years between EMA publication of CSRs compared with the timeline stated under policy 0070, and so had CSRs been published on schedule, an even higher proportion may have been the earliest source of trial results.

We also showed a large time delay between CSR completion and availability of results from publication, trial registry or even the EMA’s Clinical Data website, with a median time difference of 338 days (IQR 173–554). A recent study by Lythgoe et al reported similar results, showing a median time to journal publication of 344 (IQR 149–544) days, however, this was from the date of oncological drug approval by the EMA as opposed to CSR completion. We reported overall delays of up to 2367 days. The EMA do report targets within which they aim to make CSRs available. Even accounting for these timelines, for many CSRs it would seem that there is still a sizeable delay in availability of results from either the EMA or Marketing Authorisation Holder following the approval process. However, the reasons for these delays are uncertain, potentially due to delays in the MAA submission to the EMA and delays in the EMA reviewing the submitted information and making a decision. We propose that earlier public access to such documents, for example, to clinical researchers and decision makers, represents a further potential opportunity improve access to clinical trial information.

**Study strengths**

This is one of the first studies which characterises and describes the entire selection of CSRs made publicly available by the EMA. We use a novel method of using CSR filenames to identify and navigate individual drug trials and document types. This allowed for a more detailed description of all included trials and documents, and specifically those trials identified as pivotal by the marketing authorisation applicant. An additional strength was our use of a systematic approach for searching and identifying corresponding published sources of trial results in journals and trial registries, in addition to considering the timeliness and availability of such information compared with the EMA. This study highlights the potential opportunity for earlier access to such information on clinical trials. As far as we are aware, this study is the first to examine the timeliness of publication of results of pivotal trials submitted to the EMA for regulatory drug approval, and to compare the availability of their results across CSRs, journals and clinical trial registries.

**Study limitations**

Given the large volume of documents and trials, completing a more thorough analysis of the characteristics and contents within the available CSRs was not feasible. Presence of individual CSR sections, such as CSR body, protocols, SAPs and CRFs within the documents were not explored. While the methods for identifying matched publications were rigorous and independently checked by three authors, some publications may not have been identified. Publication length, and hence compression factor, is a crude measure of the amount of information contained and lost across each source, and therefore does not provide an indication of what important information may be absent from publications relative to CSRs.

**CONCLUSIONS AND RECOMMENDATIONS**

Information on submissions for marketing authorisation approvals published by the EMA on its Clinical Data website includes large number of documents, however, the ICH naming structure can be used to identify individual trial and document types to navigate them more efficiently. We described the wide range of numbers of documents, pages and trials per approval application for each medication. We showed that most submissions...
included individual documents for the common recommended CSR document types (body, CRF, SAP and protocol), however, just one-third published all documents individually. EMA Policy 0700 mandates that these four specific sections, should be available for relevant trials.\textsuperscript{11 35 48-52} Such clinical trial information publicly available on the EMA Clinical Data website should be consistent across medications and trials. Absence of individual documents could be a potential barrier to full transparency and appraisal of information about particular trials. The marketing authorisation applicant should endeavour to include each of these document types individually in their submissions and the EMA should encourage this practice.

The definition of a ‘pivotal’ trial appears ambiguous, and there seems to be a disparity between those trials defined as pivotal for regulatory purposes and those which are compelling sources of evidence of efficacy and safety, with many non-randomised or preclinical trials being defined as such. More clarity on what constitutes a pivotal trial from drug regulators and adoption of a more uniform definition would be beneficial.

Researchers should consider using CSRs in evidence synthesis as they are often a timely and sometimes only source of information for trials. Search strategies for systematic reviewing should include strategies for identifying CSRs from relevant sources, including the EMA’s Clinical Data website. The standardised ICH naming structure could be used by, for example, researchers to more efficiently navigate submissions and identify relevant documents.

We have shown the lack of timeliness of access to CSRs for some trials, which can be a missed opportunity and barrier to their use in research. We would recommend the resumption of publication of CSRs by the EMA under policy 0070 as an important measure to improve transparency, and with more timely publication of CSRs. The Marketing Authorisation Applicant or the EMA also could potentially provide access to CSRs sooner following their completion to further enhance the availability of such evidence. Researchers and policy makers should advocate for unrestricted access to unpublished trial information, to support timely clinical decision-making.

**REFERENCES**


