Publication Landscape on Cushing's Disease: From the Past 40 Years to Future Directions

Ming-Jie Luo
Tsinghua University School of Medicine

Kelei Du
Tsinghua University School of Medicine

Xiaopeng Guo
Peking Union Medical College Hospital

Zihao Wang
Peking Union Medical College Hospital

Bing Xing (xingbingemail@aliyun.com)
Peking Union Medical College Hospital
https://orcid.org/0000-0002-3864-5168

Research Article

Keywords: Cushing's disease, bibliometrics, machine learning, collaboration network

Posted Date: September 9th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-834733/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Purpose

Though literature related to Cushing's disease (CD) has grown significantly, previous reviews exclusively focused on specific research areas and were biased towards highly cited articles. This study aims to systemically analyze the research landscapes and trends using unbiased methods.

Methods

We queried all the CD-related publications in PubMed and clinical trials registered on clinicaltrials.gov. Latent Dirichlet allocation (LDA), a machine learning method, was used to derive research hotspots from article texts. The research topic clusters and country-level collaboration were revealed by network analysis.

Results

5015 articles were published since 1981, currently growing at 155 per year, with more retrospective studies but fewer prospective studies. Interestingly, the most popular LDA research topics were complications and comorbidities, endocrine hormone tests and surgical therapy, and they formed a remarkable triangle relationship in the research topic network. These topics had numerous international studies and were supported by most funding. In addition, many topics in the basic research domain were proliferating, including mutation, biomarkers, endopeptidases, and other molecular genetics and pathology of CD. Out of 63 registered clinical trials, over 25% were withdrawn due to inadequate patient recruitment or lack of funding.

Conclusions

This publication landscape analysis provided a systemic representation of CD literature regarding the history, current challenges, and future directions, enabling clinicians a rapid and comprehensive insight into the disease.

Introduction

Cushing's disease (CD), also known as Pituitary-dependent Cushing's syndrome, is a chronic multisystem disease due to metabolic disorders caused by high cortisol levels in the serum [1]. ACTH secretion from the hypothalamus pituitary stimulates blood cortisol secretion from the adrenal glands, accounting for 60-80% of Cushing's syndrome [2, 3]. Clinical characteristics of Cushing's disease include weight gain, central obesity, hypertrichosis, hirsutism, and proximal muscle atrophy, as well as hypertension, osteoporosis, and diabetes mellitus [3]. CD incidence and prevalence are estimated to be approximately 2.4 and 39 per million, respectively [4]. The corticotroph adenoma of CD are mostly microadenomas (i.e., <10 mm in diameter), with fewer macroadenomas (5-10%) [5].

In the past decades, increasing interest has been given to the researches of CD regarding the molecular basis, pathophysiology, hormone testing, improved surgical techniques and novel drug treatment strategies. However, patients often suffer from multisystem involvement before present at the clinic due to delays in diagnosing and treating CD [6, 7]. Patients in late-stage disease have a much higher mortality rate and poorer prognosis compared to early-stage patients or healthy individuals [7]. This is because patients with CD disease are prone to infectious diseases and cardiovascular and cerebrovascular events due to poor immunity [8]. The clinical management of patients with CD remains a challenge for clinicians due to small adenomas or negative MRI images, atypical hormone test results, periodic cyclic Cushing's syndrome, and lack of experience in treating CD.

Although there has been a remarkable increase in the amount of literature related to Cushing's disease, the research trends and highlights have not yet been systemically reviewed and adequately addressed. Previous literature reviews have almost exclusively focused on specific research areas and considered only highly cited articles [9, 10]. These reviews include clarification of published evidence, expert opinions, and research consensus on clinical trials [11–13]. Therefore, a comprehensive landscape study is needed.

Scholars have recently used computational tools to understand biomedical publications by applying classic statistical and machine learning methods to large-scale publication databases [14, 15]. Computational methods have been applied to large publication datasets to identify specific themes and research trends [16, 17]. In addition, Medical Subject Headings (MeSH) is a vocabulary used to index articles for articles in PubMed. MeSH covers a broad range of biological topics and medical terms, now widely used to perform bioinformatic analysis [18].

This study utilizes a bibliometric approach to systemically investigate all CD-related literature published in PubMed and clinical trials on clinicaltrials.gov. Using both classic bibliometric methods and machine learning algorithms, we attempted to comprehensively explore the historical and present state of CD researches and provide direction for new discoveries and technologies. From the perspective of research collaboration and funding, we also proposed a way to better tailor researches of the greatest opportunities.

Methods

Databases and Data Extraction

We searched for all publications related to CD in the English language using the public version of PubMed. The search keyword is included in the supplementary text (Online Resource 1). The search included title, abstract, and MeSH terms and was limited to articles published before January 1, 2021. The complete records of the search results were downloaded in XML format using the E-Utilities API tools of PubMed
Comorbidities has been the most popular international research of interest and had considerable funding support. The number of international multicenter studies for each study topic versus the number of their corresponding funding support. By far, Complications & Surgery (Fig. 4B). There were sporadic collaborations between many other countries as well. Out of 5015 PubMed articles, we retrieved funding information for 401 articles, while the rest of the studies did not provide funding information. Among them, 90% were supported by the National Institutes of Health (NIH).

CD (Fig. 4B). There were sporadic collaborations between many other countries as well. Out of 5015 PubMed articles, we retrieved funding information for 401 articles, while the rest of the studies did not provide funding information. Among them, 90% were supported by the National Institutes of Health (NIH).

Collaboration network and funding analysis

We then investigated the trend of multicenter international researches. In the research field of CD, the United States is far ahead in terms of the number of papers published, followed by Italy and the U.K. (Fig. 4A). The international cooperation between multiple research centers is becoming a rising trend, with an average of 50 international studies in the recent five years. The U.S., U.K. and European countries, as a triad, are cooperating closely in the research of CD (Fig. 4B). There were sporadic collaborations between many other countries as well. Out of 5015 PubMed articles, we retrieved funding information for 401 articles, while the rest of the studies did not provide funding information. Among them, 90% were supported by the National Institutes of Health (NIH).

The Medical Research Council and the Wellcome Trust in the U.K. are the other main providers of research funds for the Cushing disease. Figure 4C shows the number of international multicenter studies for each study topic versus the number of their corresponding funding support. By far, Complications & Comorbidities has been the most popular international research of interest and had considerable funding support.
The clinical trials for Cushing's disease

We have queried clinicaltrials.gov for all clinical trials with CD patients initiated since 1981. We summarized the clinical trial type, phase, status, location, sample size, start and end dates. The trials are summarized in Table 2. These trials were either focusing primarily on CD (41 trials) or together with its complications and comorbidities and other diseases (22 trials), including acromegaly, prolactinoma, Addison disease, adrenal neoplasm, and ACTH-independent macronodular adrenal hyperplasia (MIMAH). There were 20 observational and 43 interventional trials, with the median trial duration being 43 and 49 months, respectively. The median enrollment number of patients in the observational clinical trial was 200, and the interventional was 34. Among the interventional trials, 19 trials were at or prior to phase II, 16 trials in phase II and beyond, and eight trials of unknown status. The median duration was 49 months for phase II, 54 months for phase III and 47 for phase IV. The majority of trials were at the intermediate stage, 27% being phase II and 19% being phase III. The median number of patients was 26 for phase II, 73 for phase III and 249 for phase IV. The number of clinical trials for terminated, withdrawn and unknown status is 7, 1 and 8. The recorded reasons for clinical trial termination were lack of accrual or inadequate patient recruitment in (4 trials), lack of funding (2 trials) and unknown (2 trials). The median trial duration for all terminated trials was 55 months. The withdrawn clinical trials were documented as lack of patient enrollment, and the results were not released.

Discussion

CD is a complex disease, with its diagnosis and treatment still being a challenge. No study has summarized all the studies related to CD in recent decades, as well as the intrinsic connections between different studies and the transition of research hotspots. We aimed to present this information to clinicians through an overview of the literature to help them better understand and appreciate the changes in CD-related researches. Although the overall trend of clinical studies is rising, our detailed MeSH analysis showed that clinical studies were more focused on retrospective studies but less on prospective studies (Fig. 3C). This might be because that CD was a rare disease with a small number of patients in individual centers and retrospective studies are now possible with the accumulated cases in these centers. With an increasing number of articles published, there are great advances in diagnosis and treatment methods for CD, as shown in Figures 1 and 2. Previous studies focused on specific areas such as complications, comorbidities, surgical outcomes, hormonal assays, and pathology studies among the various research categories. These studies include the summary of clinical cases from various medical centers and some of the advances in basic research in the last 20 years, as seen from Figure 2C.

Transsphenoidal pituitary tumor resection is currently the first-line therapy for CD, achieving a cure rate of about 80% [21]. Nevertheless, the recurrence rate after surgery is about one quarter after a long-term follow-up [22]. Moreover, some patients are not suitable for surgery or reoperation. Therefore, these patients require adjuvant therapy (e.g., pharmacological, radiation therapy) as a supplement. Before this era, few effective medical treatments directly target ACTH-secreting pituitary adenomas. With a tremendous increase of basic biomedical research in the last decade (Fig. 2 and 3), many breakthroughs in fundamental researches, particularly the molecular genetics and pathology of CD, have led to new potential therapeutic discoveries, including cabergoline, pasireotide, ketoconazole, metyrapone, mitotane and mifepristone [23–27]. These medical treatments act at different levels of the hypothalamic-pituitary-adrenal axis, including the pituitary gland (inhibiting ACTH secretion), the adrenal gland (inhibiting steroidogenesis), the target tissue (blocking the glucocorticoid receptor) [28]. In addition, as one of the hotspot genes of CD, the USP8 gene was automatically identified as an important keyword in the vast literature by LDA analysis. The mutation of USP8 causes sustained EGF signaling and enhanced propiomelanocortin, leading to CD [29]. Clinical trials currently test Gefitinib as targeted therapy in patients with USP8-mutated CD (clinicaltrials.gov ID: NCT02484755).

As CD is a neuroendocrine tumor that affects systemic metabolism, clinicians are mainly concerned with clinical issues such as complications, comorbidities, surgical outcomes, and endocrine hormone tests [1]. By virtue of technological innovations and the discovery of new drug targets, researches related to the treatment of CD has increased considerably, including surgical and pharmacological, and adjuvant therapy. Through LDA analysis (Fig. 2), we found the predominant topics for the literature over these 40 years included Endocrine hormone tests, Surgical treatment and Complications & Comorbidities, which are consistent with the three crucial aspects of CD. Among these, Complications & Comorbidities has been the fastest-growing topic over the years, indicating the researchers continuously focus on the complications and comorbidities of CD. The sub-topics of Complications & Comorbidities include Bone Metabolism, Pregnancy Complications, Rare Comorbidities and Multiple Complications, while Endocrine Hormone Tests include Hormone Circadian Periodicity, Hydrocortisone & Derivatives, Inferior Petrosal Sinus Sampling, and LDDST & HDDST Low/High-Dose Dexamethasone-Suppression Test). The number of studies focusing on these topics is increasing. Researchers will likely achieve fundamental breakthroughs in these areas that can be leveraged to direct future translational researches.

From the country-level collaboration network and research funding analysis (Fig. 4), we found that research on CD is mainly conducted by the U.S., U.K. and European countries, as well as by their joint efforts. A portion of the research topics is supported by most research grants and has many corresponding international multicenter studies. These include complication and comorbidity studies, basic researches such as hormone measurements and protein expressions, and clinical studies such as surgical treatments and drug efficacy (Fig. 4C). We also found that the number of formal clinical trials is small compared to the numerous clinical studies, possibly because some trials have not yet been published or included in PubMed. As a supplement, we analyzed all CD-related trials registered in clinicaltrials.gov (Table 2). 40% of trials are in completion status, but only one quarter have published the trial results. Over 25% were withdrawn or in unknown status, with reasons for withdrawal being inadequate patient recruitment or lack of funding. This further suggests that we need a deeper understanding of the most urgently needed patients for CD and which research directions are more attractive to funding agencies.

There are several limitations to this study. First, to define the search keywords for PubMed, we used MeSH terms and "Title/Abstract" as keywords. Search results were incomplete and contained many irrelevant results if just "Pituitary ACTH Hypersecretion" (the equivalent MeSH term to CD) was used as the MeSH search keyword. This is because this term, together with "ACTH-Secreting Pituitary Adenoma" was not included as MeSH terms in PubMed until
2001. These limitations of MeSH terms have been reported [17, 30]. Second, there are advantages and disadvantages of using LDA and MeSH for literature content research. MeSH terms are accurate and do not require manual efforts. However, MeSH provides outlying information as indexed by PubMed and cannot perfectly express the article content due to the limited number of MeSH. On the contrary, the LDA method directly extracts information beyond what is covered by MeSH from the context of the article itself. However, the representation created using the LDA method suffers the drawbacks of computational algorithms [19]. Therefore, we used both LDA and MeSH as the analysis tool. Third, since we only used the literature abstract in PubMed as input documents, the literature representations created by the LDA method did not have access to the full text of the article. As a result, the literature representations may be biased, e.g., missing details of experimental techniques, description of statistical methods. Investigators might analyze other biomedical databases in future studies.

**Conclusion**

The publication landscape analysis provided a systemic representation of CD, enabling clinicians a rapid and comprehensive insight into the disease. Researchers can apply the methods to other research directions. The most common research directions of CD are complications and comorbidities, surgical treatment and endocrine hormone tests. Future researches should be focused on molecular genetics and pathology and targeted treatment of the disease.

**Declarations**

**Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

**Conflicts of interest**

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Availability of data and material**

The datasets used in the study are available from Zenodo https://doi.org/10.5281/zenodo.5154935.

**Code availability**

The programming codes used in the study are available from Zenodo https://doi.org/10.5281/zenodo.5154935.

**Ethics approval**

Not applicable

**Consent to participate**

Not applicable

**Consent for publication**

Not applicable

**Acknowledgments**

Not applicable

**References**


**Tables**

| Table 1 Top 10 LDA topics in the CD related researched in each 5-year period. The number and fraction of articles with specified LDA terms are shown in parentheses |

---

**Note:** The table content is not provided in the text and would require further analysis or browsing through the full document for accurate representation.
Table 2 Clinical trials for Cushing's disease found from clinicaltrials.gov published before January 1, 2021

<table>
<thead>
<tr>
<th>Status</th>
<th>Observational</th>
<th>Intervventional*</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Completed</td>
<td>7</td>
<td>1</td>
<td>27 (42.9%)</td>
</tr>
<tr>
<td>Active, not recruiting</td>
<td>-</td>
<td>-</td>
<td>1 (1.6%)</td>
</tr>
<tr>
<td>Recruiting</td>
<td>7</td>
<td>3</td>
<td>16 (25.4%)</td>
</tr>
<tr>
<td>Not yet recruiting</td>
<td>1</td>
<td>1</td>
<td>2 (3.2%)</td>
</tr>
<tr>
<td>Enrolling by invitation</td>
<td>-</td>
<td>-</td>
<td>1 (1.6%)</td>
</tr>
<tr>
<td>Terminated</td>
<td>2</td>
<td>-</td>
<td>7 (11.1%)</td>
</tr>
<tr>
<td>Withdrawn</td>
<td>-</td>
<td>-</td>
<td>1 (1.6%)</td>
</tr>
<tr>
<td>Unknown status</td>
<td>3</td>
<td>3</td>
<td>8 (12.7%)</td>
</tr>
<tr>
<td>Has Results</td>
<td>-</td>
<td>-</td>
<td>16 (25.4%)</td>
</tr>
<tr>
<td>No Results Available</td>
<td>20</td>
<td>8</td>
<td>47 (74.6%)</td>
</tr>
<tr>
<td>Number of trials</td>
<td>20 (31.7%)</td>
<td>8 (12.7%)</td>
<td>63 (100%)</td>
</tr>
</tbody>
</table>

Table 2: Clinical trials for Cushing's disease found from clinicaltrials.gov published before January 1, 2021.
Interventions include medical, surgical treatment, new tests, tests, questionnaire evaluation, etc.

**Figures**

Figure 1

The number of articles published per year queried using the PubMed database.
Figure 2

The historical changes and interrelationships in the CD-related research topics were revealed using Latent Dirichlet Allocation (LDA) analysis. The number of articles published per five-year period is shown for the indicated LDA topics. The top 10 LDA topics with greatest volumes (a) and greatest changes (b) over the years are shown, respectively. The interrelationships of the LDA topics are shown using cluster network (c). A different color represents each cluster: "clinical study" (green), "treatment" (red), "hormones" (blue), "Basic research" (dark blue), "pathology" (purple), "study type" (pale red).

Figure 3

The trend of Medical Subject Headings (MeSH) indexed in PubMed. The top 10 MeSH terms with the greatest changes related to basic research (a) and to clinical research (b) over the years are shown, respectively. (c) The changes in MeSH terms per five years related to study types.
Figure 4

The country-level collaboration network and research funding analysis. (a) Top 20 countries with the highest total number of published articles. (b) The country-level collaboration network based on the international multicenter studies. (c) The scatter plot of LDA topics in terms of the number of funding support (x-axis) and international multicenter studies (y-axis).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ESM1.pdf
- ESM2.pdf