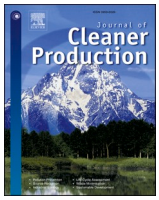
Journal of Cleaner Production 447 (2024) 141511

Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro



# Sustainability disclosure in the pharmaceutical and chemical industries: Results from bibliometric analysis and AI-based comparison of financial reports

Stefano Calciolari, Mirko Cesarini, Massimo Ruberti *

*Università Degli Studi di Milano Bicocca, Italy*

ARTICLE INFO

Handling Editor: Cecilia Maria Villas Bôas de Almeida

*Keywords:*
Environmental reporting
Pharmaceutical industry
Chemical industry
Content analysis
Non-financial disclosure
Corporate voluntary disclosure

ABSTRACT

Climate change is a pressing global challenge, and companies are under increasing pressure to adopt sustainable practices and transparently disclose their environmental performance. This study focuses on the Italian pharmaceutical industry to gain insights into the level of stakeholder and management attention to climate change and the growth of academic research in this critical area. We employ a bibliometric analysis of scientific literature and a content analysis of financial statements to examine sustainability practices and reporting in both pharmaceutical and chemical companies.

Our findings reveal a significant difference in stakeholder expectations and commitment to addressing climate change between the two industries. The chemical industry has demonstrated a stronger focus on sustainability, as evidenced by higher academic research and environmental disclosure levels. In contrast, the pharmaceutical sector has shown a comparatively lower level of environmental reporting and sustainability practices. However, our analysis shows a positive trend in pharmaceutical companies' voluntary disclosure of environmental information in their financial reports.

These results highlight the need for increased focus and effort in the pharmaceutical industry to address climate change and improve sustainability practices. Pharmaceutical companies can build stakeholder trust, improve environmental sustainability, increase transparency in reporting, align practices with global sustainability goals, enhance their reputation, and reduce their environmental impact. This research provides valuable guidance for policymakers, researchers and practitioners in promoting sustainable practices in the pharmaceutical industry and contributes to the broader understanding of ESG disclosure and climate change.

## 1. Introduction

Climate change has led to a call for concrete action to address the problem at both industrial and academic levels. While some initiatives have been underway for quite some time (Mitchell, 1989; Houghton, 1992), there is still much to do to mitigate and counter the adverse anthropogenic effects on our planet. On the background of this recent call for action, the concept of sustainability has gained significant attention, with increasing interest in various disciplines.

Research efforts in management and economics have focused on the need for companies to operate with environmental and social responsibility and disclose their environmental performance. Coherently with this purpose, investors and institutions (e.g., the European Commission) started asking companies for transparent and comparable information about their commitment and actions to reduce their carbon footprint (Venturelli et al., 2019; Lombardi et al., 2022). In this respect, an interesting question concerns whether research efforts are aligned with professionals' interests/priorities.

The aim of this paper is twofold. First, we want to assess the level of attention paid to climate change by stakeholders and managers of pharmaceutical companies. In addition, a parallel objective consists of assessing the growth of academic research on climate change with respect to the pharmaceutical sector. The analysis aimed to achieve the first objective assumed the contents of a financial report indicative of how corporate communication is oriented on the matter; and it applied a supervised AI-based Content Analysis on collected statements to evaluate the commitment to voluntary disclosures. The second objective was accomplished by evaluating the development of academic research

* Corresponding author. Via Bicocca degli Arcimboldi, 8, 20100, Milano MI, Italy
*E-mail addresses:* stefano.calciolari@unimib.it (S. Calciolari), mirko.cesarini@unimib.it (M. Cesarini), massimo.ruberti@unimib.it (M. Ruberti).

https://doi.org/10.1016/j.jclepro.2024.141511
Received 29 July 2023; Received in revised form 10 February 2024; Accepted 26 February 2024
Available online 27 February 2024

outputs on the topic via a bibliometric analysis of scientific literature.

Our focus on the pharmaceutical sector has two main intertwined motivations. The pharmaceutical industry is a major contributor to healthcare; therefore, its business activities have been essential in improving the quality of human life and are characterized by high political/social salience. However, this relevant social contribution has been accompanied by significant environmental impacts of increasing concern (Belkhir and Elmeligi, 2019). Compared to all the other sectors of chemical production, pharmaceutical manufacturing generates more waste and by-products and is the least efficient sector in terms of carbon emissions per production volume (Cue and Zhang, 2009). In addition, the negative environmental impacts of the sector are scattered across the entire life cycle of pharmaceuticals, as residues enter waterways after human consumption (Kümmerer, 2010). In this regard, Alajärvi et al. (2021) showed that the public wants better communication from pharmaceutical companies on environmental sustainability.

Based on such a line of argument, one may reasonably expect environmental sustainability to rank high in the strategic priorities of the pharmaceutical industry and the competent regulatory authorities. Yet, the pharmaceutical industry is not included in the EU taxonomy. This seems surprising because it may lead to underestimating the environmental impacts of such an industry and reduce the investors' and managers' incentives to invest in industry-level initiatives against climate change.

We did look for evidence regarding this potential issue by comparing the pharmaceutical and chemical industries because both perform similar industrial processes while differ in terms of political/social salience.

The remainder of this article is structured as follows: the next three sections provide, respectively, the conceptual background, the objectives, and the data and methods of the study. The results are then articulated into two research phases: Phase 1, regarding the bibliometric analysis of scientific literature about the academic interest in climate change, and Phase 2, concerning the content analysis on the voluntary environmental disclosure practices of two samples of pharmaceutical and chemical companies. Finally, we summarize our key findings, discuss their implications, and suggest potential avenues for future research.

## 2. Conceptual background

Recently, climate change has been at the core of the political and economic debate, putting pressure on businesses in this respect. Our study falls within the research strand that examines companies' voluntary disclosure of environmental sustainability information (such as Bewley and Li, 2000; Rezaee and Tuo, 2017; Zamil et al., 2023) because we investigate managers' behaviour when they are not constrained by compelling legislation.

Sustainability disclosure is important for several reasons in the pharmaceutical sector. First, the industry's products and services significantly impact the health and safety of people together with the environment. For example, pharmaceutical products can release hazardous chemicals into the air and water, thus negatively impacting the environment and, indirectly, public health. By disclosing their ESG performance, pharmaceutical companies can help stakeholders better understand the business's environmental impact and thus foster actions to mitigate such effects.

Second, sustainability disclosure can help pharmaceutical companies to build stakeholder trust and credibility. According to KPMG International (2021) , 93% of the 250 largest global companies by revenue disclosed sustainability information: the fact suggests that such practice has become a fundamental expectation of stakeholders. Therefore, pharmaceutical companies can demonstrate their commitment towards sustainability and build stakeholders' trust by providing transparent and reliable information about their ESG performances.

Third, sustainability disclosure can help pharmaceutical companies to identify improvement priorities and achieve sustainability goals. By reporting on their ESG performances, companies may identify areas of weakness and improve their sustainability practices with pragmatic initiatives. This may also lead to long-term benefits, such as increased operational efficiency, cost savings, and reduced ESG risks.

Stakeholder theory (Freeman, 1984) places corporate disclosure at the centre of academic debate. The dissemination of relevant information is beneficial for reducing information asymmetries between the firm and its close stakeholders and may also be of interest to outside parties. It could positively influence perceptions of the firm's future prospects for financial actors such as stock analysts, capital markets and institutional investors (Brammer, S., & Pavelin, S., 2008). In addition, stakeholder power, the firm's strategic posture and economic performance are significantly associated with the level of corporate social disclosure (Roberts, R. W., 1992).

Similarly, the level of CSR disclosure appears to be driven by characteristics such as company size, industry, profitability and corporate governance. In particular, socially visible companies appear to emphasize ESG issues more than the average business due to pressure from the media, NGOs and regulators. In this regard, the decision to disclose is influenced by political, social and cultural factors (Ali et al., 2017). As a result, the most frequently studied industries are environmentally sensitive ones (Fernandez-Feijoo et al., 2014).

Typically, studies focusing on corporate social responsibility (CSR) reporting rely on non-financial reports (NFRs), such as CSR/ESG statements, as their primary source of information (Ioannou and Serafeim, 2017; Baumüller and Sopp, 2022). This choice is driven by the desire to use a comprehensive data source to analyze the multifaceted nature of sustainability disclosure.

However, in the Italian pharmaceutical sector, only a minority of companies (three firms) are listed, and non-listed companies seldom produce an integrated or CSR report since it is not mandatory. Therefore, we relied on financial reports primarily in order to make our sample representative of the Italian market structure. In addition, the fact that a company includes ESG/CSR information in its financial statements may indicate that the matter ranks high in its strategic priorities. Lastly, the EU Directive 2014/95/EU is likely to have increased voluntary disclosure of environmental issues in official corporate documents (Doni et al., 2019; Raucci and Tarquinio, 2020).

## 3. Study objectives

As previously mentioned, the aim of our study is twofold. First, we want to analyze the variation over time in the level of attention given to climate change by academic research, in general and in the two selected industries. In this respect, we consider the relative number of published studies about climate change as a proxy of the academic attention level on the matter. The first phase of the research also aims to determine whether the climate change literature is equally distributed between the two sectors over time. The decision to focus on scientific research is based on the idea that academic literature can motivate corporate behaviour.

Indeed, knowledge transfer from academia to industry is an important driver of industrial innovation and economic growth (Bercovitz and Feldman, 2006; Mowery et al., 2015; De Wit-de Vries et al., 2019); further, the industry can orient scientific research via financial support, and evidence show that researchers with industry support tend to be more productive in terms of scientific publications (Blumenthal et al., 1986). Therefore, analyzing the trend of scientific production related to any industry can help to understand industrial strategic priorities.

Second, we aim to assess and compare the extent to which companies from the two industries care about environmental sustainability disclosure in their financial reports. In this regard, we consider the amount of selected types of information included in corporate documents as a measure of the companies' strategic attention on the matter. This second research phase aims to assess whether the two industries

showed different orientations towards environmental sustainability disclosure.

The study takes an exploratory approach and does not rely on relevant contributions from the extant literature to establish a plausible hypothesis on which industry might focus more on fighting climate change. However, we argue that the higher political/social salience of the pharmaceutical industry may have hidden its negative impact on the environment. Therefore, pharmaceutical companies may have experienced or perceived less pressure on sustainability practices and reporting. In addition, chemical companies have been under scrutiny for longer (Gillet, 1952) and thus may be more accustomed to sustainability practices. Following this reasoning, we hypothesized that sustainability has gained strategic importance for chemical companies and has been the object of academic attention in chemistry research earlier than in pharmacy.

## 4. Data and methods

This section focuses on the data and methods used in two phases of the study: a bibliometric analysis of scientific literature and an AI-based content analysis. Each phase is related to one of the two aforementioned objectives, respectively.

The bibliometric analysis of scientific literature assesses the academic research efforts regarding climate change in general and the two target sectors. The data source was the Web of Science database. The search considered all the peer-reviewed articles published until 2021, included, and used the most popular keywords evoking the selected topic:

Climate change [Title/Abstract/KeyWords] OR global warming [Title/Abstract/KeyWords] OR sustainability [Title/Abstract/KeyWords].

In the resulting list of 523,942 articles, we identified two subsets of publications, by including also the search term "pharma*" (subset 1: 1886 articles) and "chemi*" (subset 2: 29,387 articles).

Subset 1: (climate change [Title/Abstract/KeyWords] OR global warming [Title/Abstract/KeyWords] OR sustainability [Title/Abstract/KeyWords]) AND (pharma* [Title/Abstract/KeyWords]);

Subset 2: (climate change [Title/Abstract/KeyWords] OR global warming [Title/Abstract/KeyWords] OR sustainability [Title/Abstract/KeyWords]) AND (chemi* [Title/Abstract/KeyWords])

We show our results by adopting different classification perspectives. First, we display the geographic distribution of scholarly articles based on the authors' university affiliation. In addition, we identified the subgroups of articles published in the fields "business, economics, and management". Further, we show the ten disciplines associated with the highest number of articles according to the Web of Science scientific categories. Finally, we analyzed the growth of scholarly attention on the topic over time until 2021. Across the mentioned classifications, we propose a comparison between the two industries.

Content analysis is "a technique for gathering data that consists of codifying qualitative information in anecdotal and literary form into categories for deriving quantitative scales of varying levels of complexity" (Abbott and Monsen, 1979, 504). This method is widely used in major studies of social and environmental reporting (Gray et al., 1995; Parker, 2005, 2011) and relies on a fundamental axiom: the amount of information included in any conceptual category is a proxy of the importance of that category (Unerman, 2000, 667). In the literature, one can classify content analyses into two main groups (Pesci and Costa, 2014). The first group assesses the priority level of the information disclosed in documents (Deegan and Gordon, 1996; Neu et al., 1998). The second group uses quality assessment systems to evaluate, compare, and explain differences in the completeness and level of detail of the information included in documents (Al Tuwaijri et al., 2004), assuming that certain types of information (i.e., descriptive rather than numerical) may be perceived differently by different readers (Hooks and van Staden, 2011).

The second phase of our work falls into the first group of studies. We performed a content analysis on the financial reports of Italian non-SME and not-listed pharmaceutical companies over nine years (2012–2020). The sample selection identified a homogeneous group of companies not legally obliged to disclose information on CC issues but large enough to capture the attention of a relevant number of stakeholders. As a control group, we selected the top 100 not-listed and non-SME Italian chemical companies. This allowed us to compare the results of the analysis of the pharmaceutical sector with those of a similar sector, whose effects on the environment are well-known to the general public.

We downloaded the financial reports from the AIDA Bureau van Dijk database. These documents were analyzed by means of Artificial Intelligence (AI) algorithms developed by the authors to retrieve information traceable to environmental sustainability certifications (e.g., ISO 14001) and references to specific regulations dedicated to environmental sustainability (e.g., Legislative Decree 152/06). This tool can be logically divided into two components: 1) a component that uses the Tesseract Optical Character Recognition (OCR) library (Tesseract, 2022; Heggammer, 2022), based on an artificial neural network, which allows the extraction of text from pages of financial statements, automating the extraction of text from different financial statements contained in PDF files; 2) a component that uses classifiers to identify, among the extracted texts, those in which identified topics of interest are covered (e.g., presence of environmental disclosure or reporting, greenhouse gas emissions trading, etcetera). The second component generates a report in a spreadsheet that summarizes the retrieved topics and the parts of the text that contain the relevant passages in a table. This spreadsheet is the starting point for further analysis. Both components were built by developing ad hoc scripts in the Python language (Python, 2022) for the specific domain.

Below, we propose further details on the use of the two mentioned components:

- The use of OCR was decided after testing several methods for extracting text from financial statement PDFs. Although PDFs of financial statements contain much information already encoded in vector format, the PDFs contain several parts scanned from paper and stored as images (often with relevant information on environmental issues). The Tesseract library uses an artificial neural network (based on the LSTM, Long Short Term Memory Network, paradigm) that allows it to recognize both individual characters and the grouping of characters into words, as well as the grouping of words into lines. The results obtained were evaluated by randomly extracting parts of the households and comparing them with the original PDF. The results obtained were evaluated by randomly sampling parts of the information retrieved by the AI algorithm and comparing them with the text in the financial reports. The extracted text was mostly identical to the data source and showed minor typos in a few cases.
- Rule-based and supervised machine learning-based classifiers have been developed to assess whether a text deals with a specific topic. Rule based classifiers process texts using a (large) set of hand-crafted rules (an example of a single rule is: if sentence contains "ISO 14001" or sentence contains "Environmental Certification" then …) while supervised Machine Learning classifiers focus on algorithms which are trained by examples. Supervised ML algorithms are fed with both the sentences and the labels they should learn to guess (e.g., a binary classifier may focus on choosing between "Environmental Certification" or "No Environmental Certification"). After the training, the algorithms learn to produce the correct label for unseen texts (whose labels are unknown). Usually, training sets of hundreds of labelled documents are required to train those algorithms properly. Both the rule-based and the ML based approaches have their advantages and drawbacks. Focusing on rule-based classifiers, the set of rules is likely to become very large and difficult to maintain as people add more rules to cope with complex cases and to improve the performances, to

the point that the addition of a new rule can have unpredictable (catastrophic) effects. The NLP (Natural Language Processing) community focused on rule-based approaches as a first attempt to tackle the text classification problem. Eventually, the researchers switched to Machine Learning based algorithms since large set of rules can become unmanageable. ML classifiers during the training develop an internal classification model which can be hardly accessed and understood by people (in literature this is referred as the AI explainability issue), while rule-based classifiers can be more easily investigated by people. A simple rule-based classifier can be quickly built (few rules can reach sufficient classification performances) and can be used as baseline to compare the performances of more complex approaches e.g., ML based classifiers. For this reason, the two types of classifiers were used in the project. The two types of classifiers produced very similar results. Due to the simplicity of the studied vocabulary and the small number of texts available for training (in the range of hundreds), it is no surprise that machine learning-based classifiers did not perform significantly better than rule-based classifiers, as it is usually the case (Jurafsky and Martin, 2009). However, the fact that the two types of classifiers produced very similar results is a good indicator of the quality of the extracted information. The two classifiers were developed in Python using the Scikit-learn library (Pedregosa et al., 2011).

Extracting text from 500 financial statements required about 48 h of processing on a PC (Intel I7, 32Gbytes of RAM) with a Linux operating system. The developed tool allowed a precise content analysis of accounting documents, recognizing - among text and images - the presence of the information desired by the researcher. This tool, which is still under development to be refined with the needs of future research, will be released to the community in an open-source (or similar) version at the end of the project, allowing other researchers to use an efficient and customizable tool for the analysis of financial documents.

The AI content analysis examined four relevant categories of information identified by the authors after a direct preliminary screening of 20 randomly selected financial reports. Specifically, the information retrieved from the annual reports concerns:

(1) the existence of a specific section dedicated to environmental reporting (CA-ENV)
(2) references to environmental legislation enforced (CA-LAW);
(3) references to environmental certifications obtained by the company (CA-CER);
(4) disclosure of water discharge management (CA-WAT).

The four mentioned categories identify four dichotomous variables, assuming the value "1" when the information sought is retrieved and "0" otherwise. These four variables, together, determine the binary variable "CA-DIC", which is equal to "1" when any of the four variables above assume value "1" ("0" otherwise). In other words, this last variable indicates if a company is disclosing any environmental information.

## 5. Results

This section shows the results following the two study phases: the bibliometric analysis and the content analysis.

### 5.1. Phase 1: bibliometric analysis of scientific literature

Fig. 1 shows the annual production of scientific articles related to: CC topics (in any research field), pharmacy (any topic) and chemistry (any topic). A general growth in academic research is visible in all the identified areas of interest. However, among the three lines, CC is the steepest and shows an exponential increase in the total number of articles: the 1301 articles published in 1992 climbed to 80,123 in 2021, with an average 15.3% annual growth. In 2016, the CC articles overtook the number of contributions published in the pharmacy field: the growth of this last research area has been 6.7% over the three decades, reaching 55,698 articles in 2021. A similar growth (6.1%) has been observed for the papers published in the chemistry field, from 28,773 in 1992 to 158,985 in 2021. These data confirm the comparable solid interest in the pharmaceutical and chemical research fields, and the CC topic's bursting growth.

Fig. 2 displays the frequency of CC topic articles among the pharmacy and chemistry sets. Only a modest percentage of articles on

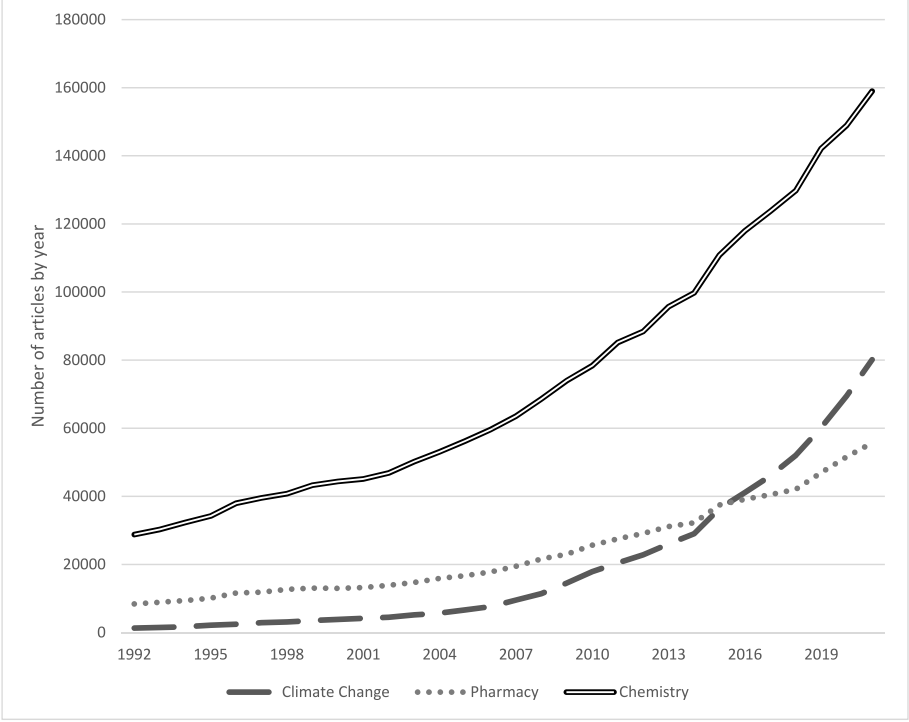



**Fig. 1.** Articles published annually in the research areas of Climate Change, Pharmacy and Chemistry. Source: Web of Science 2023.

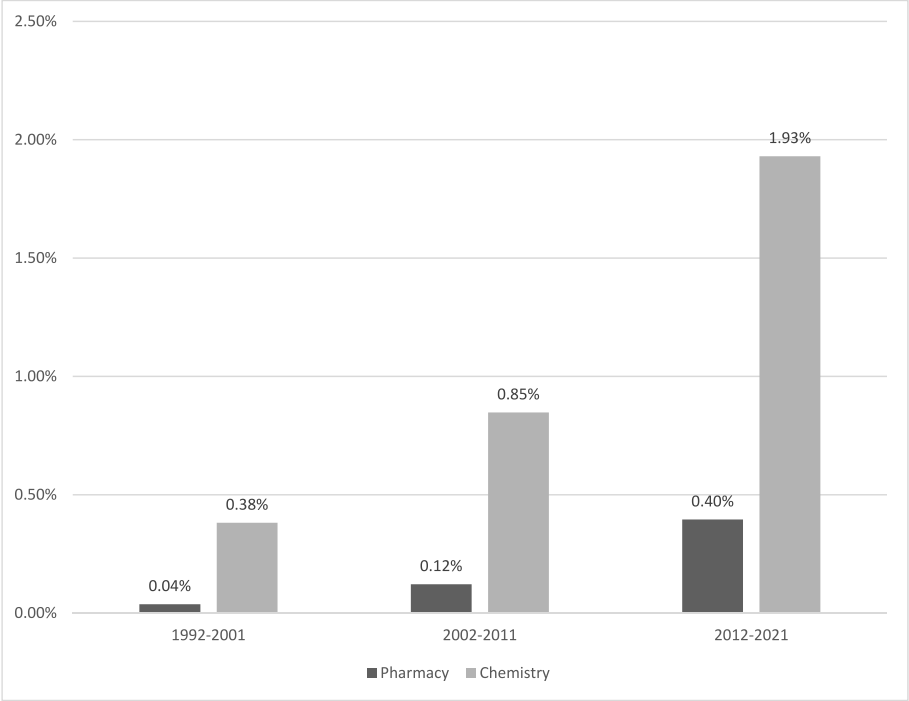



**Fig. 2.** Frequency of CC Articles by Research Area over three decades. Source: Web of Science 2023.

pharmacy deals with the subject of CC. However, while the frequency was as low as 0.04% in the first observed decade, it has increased tenfold (reaching 0.40%) in the last ten years of the observation period. It appears, instead, that chemistry articles have always been more sensitive to the emerging CC topic (0.38% in 1992–2001) and have grown up to about 2% of the total amount of articles in the last observed decade).

Fig. 3 shows the frequency of articles having "pharma*" or "chemi*" in the topic among the CC-related articles. This shows, from another point of view, the distribution of studies about CC in the two research fields. Indeed, the bar chart clearly displays the contribution of each field in dealing with CC: while over 5% of the articles were constantly published in chemistry research, less than one-tenth of the works in pharmacy are dedicated to the topic. Precisely, over 30 years, 5.11% of CC articles contained the keyword "chemi*", while only 0.32% contained "pharma*". This divergence (by a factor of about 16) is not explained by the different academic productivity in the two fields. The articles on the Web of Science focused on chemistry are 2,228,758, while the publications focused on the pharmaceutical industry were 714,114 (i.e., about one third) in the observation period.

These results confirm the perception of significantly lower attention to the fight against CC in the pharmaceutical research field compared with the chemical one. While the evidence shows a growth of the general academic attention on CC, such attention is stable in the chemical research field in the observation period. However, a recent increase in attention to CC can be observed in the pharmaceutical research field.

Table 1 displays the geographic origin of scholarly articles on CC in

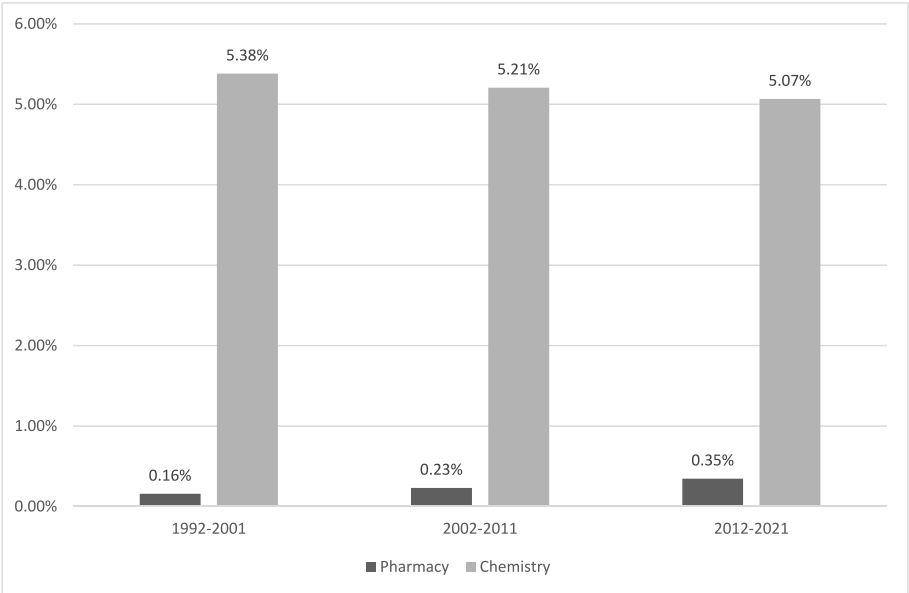



**Fig. 3.** Frequency of Pharmacy or Chemistry articles among CC articles. Source: Web of Science 2023.

**Table 1**
Countries contribution on the "CC + Pharma*" and CC + Chemi*" topics.

| Rank | Country | "CC + Pharma*" Articles *Number* | *Proportion* | Country | "CC + Chemi*" Articles *Number* | *Proportion* |
|---|---|---|---|---|---|---|
| #1 | USA | 476 | 16.8% | USA | 8927 | 18.2% |
| #2 | ENGLAND | 180 | 6.4% | CHINA | 4332 | 8.8% |
| #3 | AUSTRALIA | 138 | 4.9% | GERMANY | 2995 | 6.1% |
| #4 | INDIA | 134 | 4.7% | ENGLAND | 2910 | 5.9% |
| #5 | ITALY | 122 | 4.3% | CANADA | 2087 | 4.3% |
| #6 | SPAIN | 121 | 4.3% | FRANCE | 1854 | 3.8% |
| #7 | CHINA | 115 | 4.1% | ITALY | 1797 | 3.7% |
| #8 | GERMANY | 113 | 4.0% | AUSTRALIA | 1581 | 3.2% |
| #9 | CANADA | 98 | 3.5% | SPAIN | 1556 | 3.2% |
| #10 | NETHERLANDS | 77 | 2.7% | INDIA | 1539 | 3.1% |

Table 1. Source: Web of Science 2023

the two research fields. Looking at the first ten positions of these two rankings, we find almost the same (9 out of 10) countries, with the four Anglo-Saxon countries associated with 31.6% of the production in each category and the USA leading the two rankings. Interestingly, China and Germany are much more productive regarding the chemical industry than the pharmaceutical one, while the reverse holds for India and Australia.

Figs. 4 and 5 show the ten disciplines associated with the highest number of articles on the two topics. As expected, some categories have a broad industrial perspective (e.g., Environmental Science and Green Sustainable Science Technology), while others are close to the specific sector (e.g., Pharmacology/Pharmacy for "CC + Pharma*" and Chemistry Multidisciplinary for "CC + Chemi*").

The "Business, Economics, and Management" (BEM) research field has never been among the most active on the topic. However, while there are only 274 (0.9 per cent) BEM articles in the sub-sample focused on the chemical industry, the number of BEM articles is relatively high in the other sub-sample: 127 (6.7 per cent). Fig. 6 shows that the interest in the topic is increasing over time. The comparison of Fig. 6 with Fig. 1 suggests that BEM scholars have been investigating the topic more recently than their colleagues from other disciplinary fields: only eight articles classified as "CC + Pharma*" (41 as "CC + Chemi*") before 2010, compared to 185 (and 4942) when considering all the disciplines.

### 5.2. Phase 2: content analysis results

Tables 2 and 3 summarize the content analysis results performed on pharmaceutical and chemical companies with the AI-based tool.

Table 2 shows the mean proportions, expressed in percentage format, indicating how frequently the AI algorithms found each of the four predefined categories of information in the financial reports. The CA-DIC variable represents the percentage of companies disclosing at least a category of information. Results show that the annual reports include increased information in each information category over time, with the most relevant increases being certifications (CA-CER, +33.5%) and water discharge management (CA-WAT, +273.8%). The percentage of companies disclosing information (CA-DIC) increased by 9.3%.

Table 3 shows the same variables but for chemical companies. Remarkably, chemical companies had higher average values in each information category at the beginning of the observation period than pharmaceutical companies. Over time, however, this difference has shrunk until it disappeared due to a progressive decrease in two information disclosure categories (CA-ENV and CA-LAW) for chemical companies and a parallel increase in three disclosure categories (CA-LAW, CA-WAT, and CA-CER) for pharmaceutical companies. Interestingly, the variable CA-ENV showed a relevant decrease in both industries, suggesting that the environmental information is presented less in a specific section and more around the financial documents.

Figs. 7 and 8 show the disclosure patterns of pharmaceutical and chemical companies over the years. These visual representations, which build on the data from the previous tables, provide insights into the percentage of companies that refrain from ESG-related disclosure, those that disclose only one theme, and those that disclose at least two themes

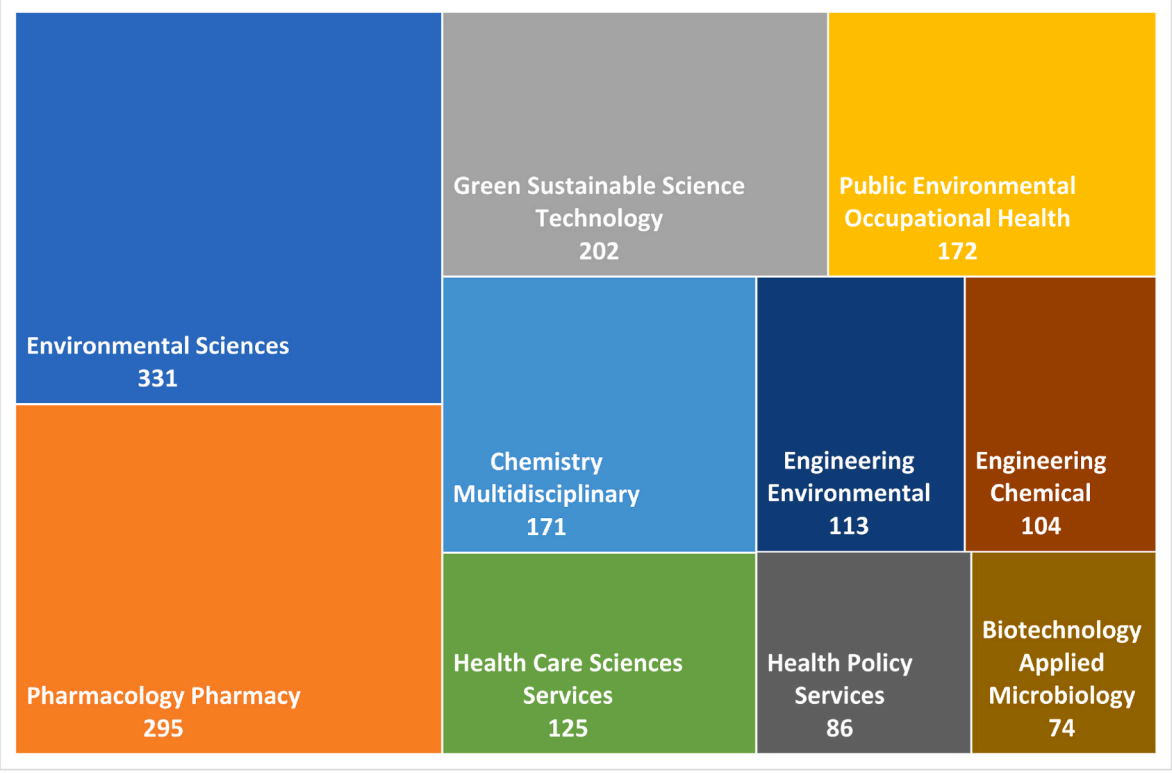



**Fig. 4.** Treemap chart of the top 10 science categories investigating the "CC + Pharma*" topic. Source: Web of Science 2023.

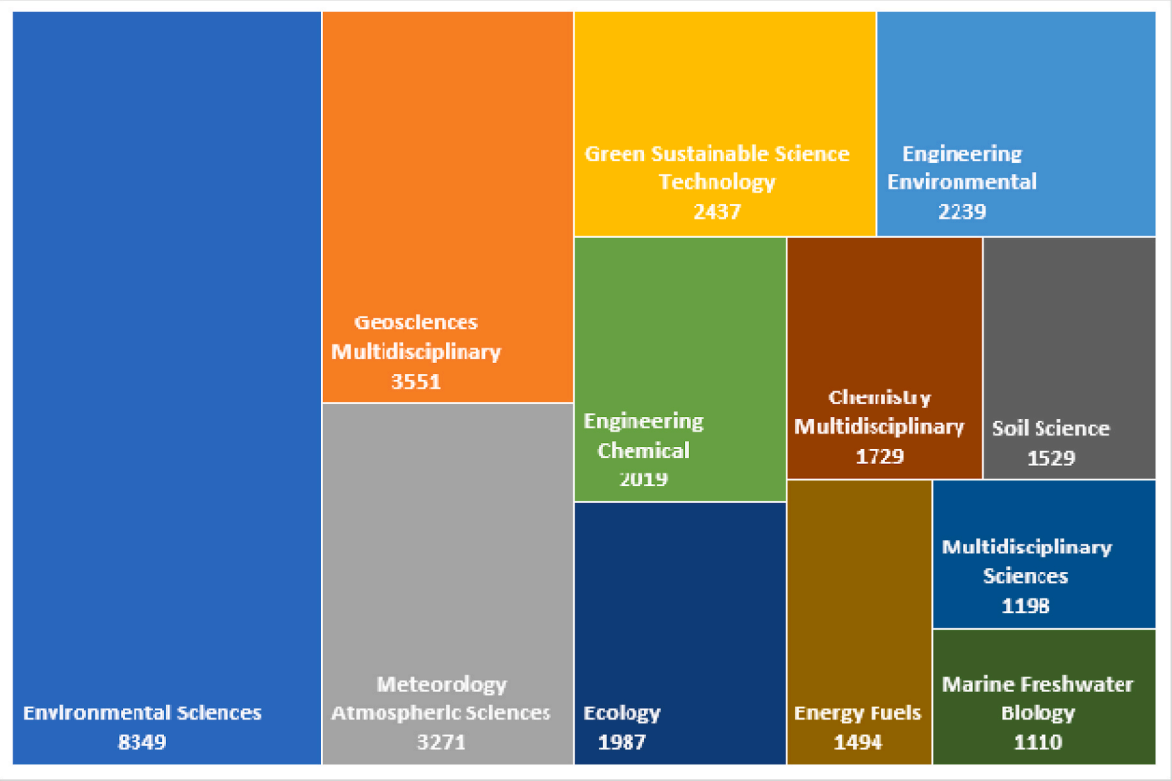



**Fig. 5.** Treemap chart of the top 10 science categories investigating the "CC + Chemi*" topic. Source: Web of Science 2023.

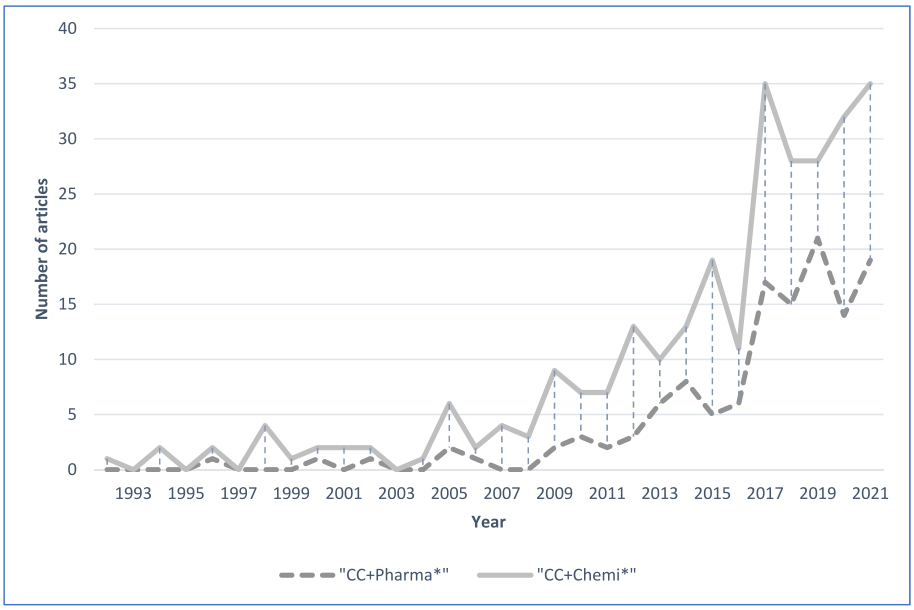



**Fig. 6.** The yearly number of articles in the field of BEM, by industry. Source: Web of Science 2023.

in the four categories mentioned above. Notably, the number of chemical companies that refrain from disclosing ESG topics in their financial reports is lower than that of pharmaceutical companies. In addition, pharmaceutical companies are less likely to report on two or more sustainability topics than their chemical counterparts.

Table 4 compares the CA results between the two industries by displaying the results of a *t*-test that determines whether there is a significant difference between the means of the two groups (i.e. chemical and pharmaceutical industries). All the searched categories displayed a higher value in the chemical case. The difference is statistically significant for every variable except for CA-ENV. These data confirm that pharmaceutical companies feel less urgent communicating with stakeholders about their environmental actions than chemical ones.

These results are consistent with those displayed by the bibliometric analysis of scientific literature and show a different path between the two industries regarding climate change. The chemical industry has been more active in academic research and corporate disclosure on this issue early on. Pharmaceutical companies, on the other hand, are

**Table 2**
Categories of contents found in the financial reports of major Italian pharmaceutical companies.

| YEAR | CA-ENV | CA-LAW | CA-WAT | CA-CER | CA-DIC |
|---|---|---|---|---|---|
| 2012 | 32.0% | 10.0% | 1.0% | 35.0% | 59.0% |
| 2013 | 25.3% | 8.1% | 5.1% | 42.4% | 58.6% |
| 2014 | 28.2% | 6.8% | 2.9% | 40.8% | 59.2% |
| 2015 | 26.7% | 5.7% | 1.9% | 40.0% | 58.1% |
| 2016 | 27.2% | 4.9% | 1.9% | 37.9% | 57.3% |
| 2017 | 26.4% | 7.5% | 3.8% | 39.6% | 60.4% |
| 2018 | 23.8% | 8.6% | 3.8% | 40.0% | 56.2% |
| 2019 | 26.4% | 10.4% | 3.8% | 43.4% | 63.2% |
| 2020 | 27.1% | 11.2% | 3.7% | 46.7% | 64.5% |
| OVERALL AVERAGE | **27.0%** | **8.1%** | **3.1%** | **40.6%** | **59.6%** |
| 2012-20 VARIATION | **−15.3%** | **12.1%** | **273.8%** | **33.5%** | **9.3%** |

Table 2. Source: Authors' content analysis applied to optical financial reports retrieved from AIDA BvD database

**Table 3**
Categories of contents found in the financial reports of major Italian chemical companies.

| YEAR | CA-ENV | CA-LAW | CA-WAT | CA-CER | CA-DIC |
|---|---|---|---|---|---|
| 2012 | 33.7% | 12.4% | 4.5% | 42.7% | 62.9% |
| 2013 | 30.8% | 11.0% | 3.3% | 42.9% | 61.5% |
| 2014 | 31.6% | 12.6% | 8.4% | 48.4% | 67.4% |
| 2015 | 36.1% | 13.4% | 6.2% | 46.4% | 68.0% |
| 2016 | 31.3% | 13.5% | 4.2% | 47.9% | 69.8% |
| 2017 | 26.0% | 9.0% | 5.0% | 51.0% | 67.0% |
| 2018 | 23.5% | 10.2% | 7.1% | 45.9% | 63.3% |
| 2019 | 27.0% | 12.0% | 7.0% | 47.0% | 67.0% |
| 2020 | 24.0% | 9.0% | 8.0% | 46.0% | 66.0% |
| OVERALL AVERAGE | **29.3%** | **11.5%** | **6.0%** | **46.5%** | **65.9%** |
| 2012-20 VARIATION | **−28.8%** | **−27.2%** | **78.0%** | **7.7%** | **4.9%** |

Table 3. Source: Authors' content analysis applied to optical financial reports retrieved from AIDA BvD database

catching up despite starting from a disadvantaged position. This can be seen in the relative increase in academic research on climate change in the pharmaceutical research sector and the recent increase in voluntary disclosure by companies.

## 6. Discussion and conclusions

This paper feeds into the ongoing debate in which the recent global climate crisis has prompted companies and institutions to prioritize their strategic objectives. In addition to recent regulatory interventions (notably the European Directive 2014/95/EU), professional and institutional working groups (e.g. EFRAG, IFRS and the EU Commission) are developing to regulate this area of corporate reporting. In this context, evaluating the level of attention given to climate change issues by stakeholders and management in different industries is crucial.

We focused on comparing the pharmaceutical and chemical industries because the higher political/social salience of the pharmaceutical industry may have hidden its negative environmental impacts from stakeholders, and, as a result, its companies may have felt less pressure on sustainability practices and reporting. The academic research and financial report analysis provided valuable insights into these two sectors' environmental sustainability reporting practices and stakeholder expectations.

The bibliometric analysis of scientific literature revealed a significant increase in academic research on climate change, indicating a growing recognition of its importance in both industries. However, the two sectors showed notable differences in attention and focus levels. The chemical industry, which has been under scrutiny for a longer time (Gillet, 1952), showed a stronger commitment to addressing the challenges of climate change. This suggests that stakeholders in the chemical industry had higher expectations and placed greater emphasis on sustainability practices and reporting.

In contrast, despite its significant contribution to healthcare and the improvement of human life, the pharmaceutical industry showed a comparatively lower level of evidence about environmental reporting and sustainability practices. This may be due to different stakeholder expectations of the pharmaceutical industry. The pharmaceutical sector's higher political and social salience may have overshadowed its negative environmental impacts, resulting in less pressure for sustainability practices and reporting. This finding suggests that stakeholders in the pharmaceutical industry may have had different priorities and expectations, possibly emphasizing other aspects of corporate responsibility.

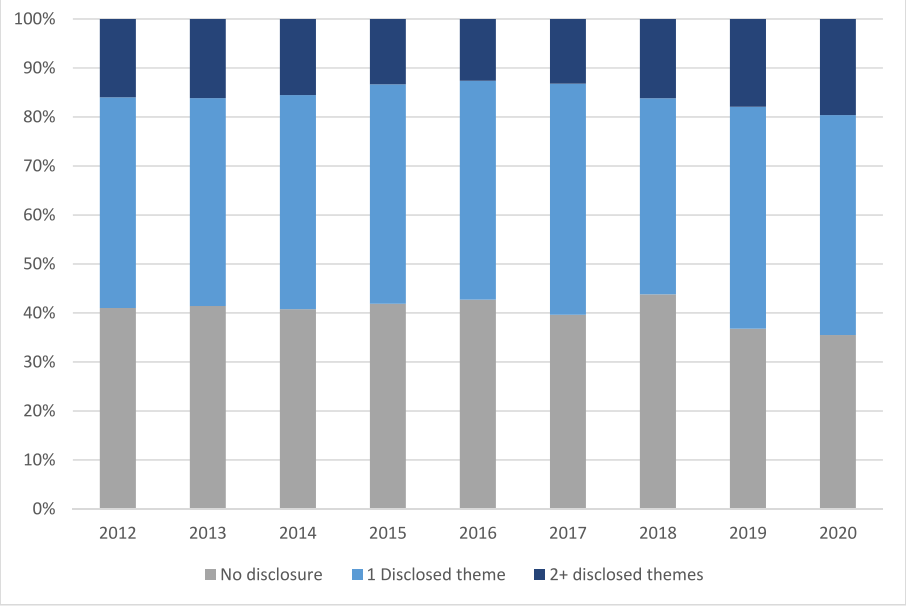



**Fig. 7.** Level of disclosure found in the financial reports of the pharmaceutical sample. Source: Authors' content analysis applied to optical financial reports retrieved from AIDA BvD database.

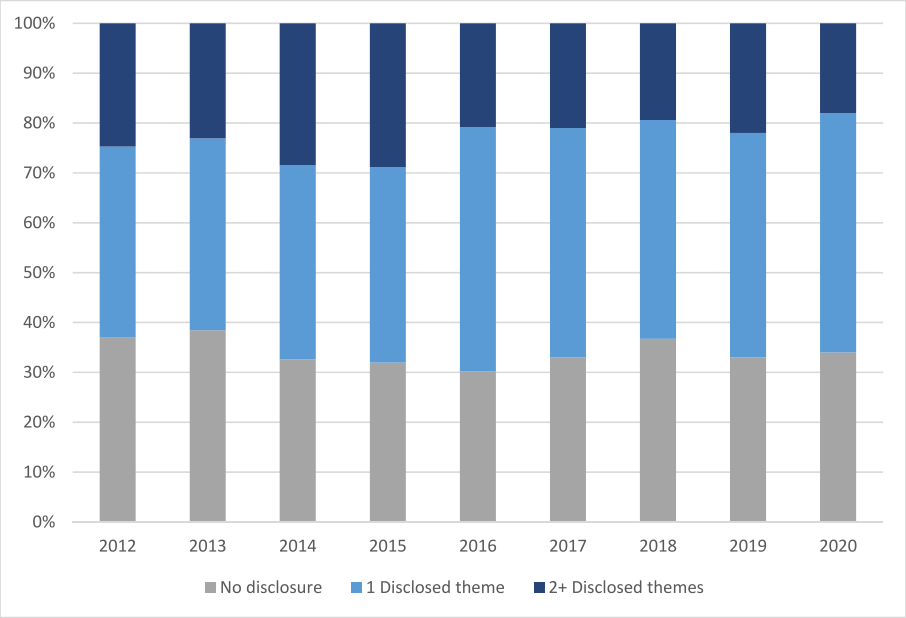



**Fig. 8.** Level of disclosure found in the financial reports of the chemical sample. Source: Authors' content analysis applied to optical financial reports retrieved from AIDA BvD database.

**Table 4**
Two-sample *t*-test with unequal variances.

| | Obs. Chemi | Obs. Pharma | Mean Chemi | Mean Pharma | Mean diff. | SE | t-value | p-value |
|---|---|---|---|---|---|---|---|---|
| **CA-ENV** | 866 | 934 | 0.292 | 0.270 | 0.022 | 0.021 | 1.05 | 0.292 |
| **CA-LAW** | 866 | 934 | 0.115 | 0.082 | 0.033 | 0.014 | 2.35 | 0.019 |
| **CA-WAT** | 866 | 934 | 0.060 | 0.031 | 0.029 | 0.010 | 2.95 | 0.004 |
| **CA-CER** | 866 | 934 | 0.466 | 0.407 | 0.059 | 0.024 | 2.50 | 0.013 |
| **CA-DIC** | 866 | 934 | 0.659 | 0.597 | 0.063 | 0.023 | 2.75 | 0.005 |

Table 4. Source: Authors' analysis applied to the Content Analysis results

Moreover, the lower attention paid to environmental change by the pharmaceutical vs. chemical sector may be related to a different social perception of the environmental risk posed by their industrial process. As of 2023, among 1335 major accidents registered by the UE Major Accident Reporting System (eMARS), eight were caused by the production of pharmaceuticals, while 694 are associated with the manufacturing, transportation or installation of chemical products (emars.jrc.ec.europa.eu, 2024). Such relevant difference may trigger a dissimilar social awareness of the risk associated with different industrial processes, with a consequent impact on policymakers' and stakeholders' attention.

The content analysis of financial reports further supports the differences in stakeholder expectations between the two industries. Compared to the pharmaceutical business, chemical companies had higher average scores in the environmental reporting categories at the beginning of the observation period. However, over time, the pharmaceutical industry showed improvements and a relative increase in voluntary disclosure on environmental sustainability. This suggests that pharmaceutical industry stakeholders showed increasing awareness of the importance of addressing climate change and increasingly aligned their expectations with those of the chemical industry.

Pharmaceutical companies showed a gradual increase in environmental disclosure, while chemical companies showed a slight decrease in specific categories, resulting in an overall decrease in environmental disclosure. These findings suggest that pharmaceutical companies started recognizing the higher importance of sustainability and took steps to address the environmental impacts of their operations. Indeed, the pharmaceutical industry needs to prioritize environmental sustainability, improve the transparency of its reporting, and align its practices with global sustainability goals. By doing so, the industry can build stakeholder trust, enhance its reputation, and help reduce the environmental impact of its operations. Stakeholders, including investors, regulators and the public, play a critical role in shaping industry expectations and driving change.

The present work's originality lies in studying a sector that seems to be underestimated in terms of environmental harm and using financial reports instead of sustainability reports as data sources. Indeed, at the time of writing, sustainability reports are the richest source of environmental information. However, this source of information is less consistent in structure and content than annual reports. Annual reports provided a homogeneous source of comparison between industries. This comparison was accelerated by creating an ad hoc algorithm to automate the content analysis of 1827 annual reports. This approach may have limitations compared to the critical eye of expert researchers. However, it has clear advantages in terms of efficiency and scalability, opening up prospects for extending further analyses to several industries and, with the necessary technical improvements, to other national contexts.

A relevant technical perspective of this study concerns the improvement of text search and interpretation of sustainability-related sentences through machine learning. In particular, by validating (with human experts) the output of the AI algorithms on a large volume of financial statements, it will be possible to train machine learning-based text classification algorithms with better and better performance.

The findings contribute to the growing body of knowledge in this area and provide guidance for policymakers, researchers, and practitioners in promoting sustainable practices and mitigating the environmental impact of the pharmaceutical industry. Further research and

initiatives are needed to motivate pharmaceutical companies to disclose climate change information and improve their environmental performance. In addition, this study could lead to comparative studies across countries and industries, providing valuable insights into the global landscape of ESG reporting.

Further research is needed to explore the reasons for the differences in stakeholder expectations and engagement in the pharmaceutical and chemical industries. Understanding these dynamics can inform targeted strategies and initiatives to promote sustainability and effectively address climate change challenges in both sectors. Collaboration between industry, academia and regulators is essential to drive positive change and ensure a sustainable future for all.

## Funding

This work was supported by the Italian Ministry of Education, Universities and Research (MIUR) within the Programma Operativo Nazionale "Ricerca e Innovazione 2014–2020" (PON R&I), through resources of the FSE REACT-EU.

## CRediT authorship contribution statement

**Stefano Calciolari:** Supervision, Validation, Writing – review & editing. **Mirko Cesarini:** Methodology, Software. **Massimo Ruberti:** Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## References

Abbott, W.F., Monsen, R.J., 1979. On the measurement of corporate social responsibility: self-reported disclosures as a method of measuring corporate social involvement. Acad. Manag. J. 22 (3), 501–515.
Alajärvi, L., Timonen, J., Lavikainen, P., Martikainen, J., 2021. Attitudes and considerations towards pharmaceuticals-related environmental issues among Finnish population. Sustainability 13 (22), 12930.
Ali, W., Frynas, J.G., Mahmood, Z., 2017. Determinants of corporate social responsibility (CSR) disclosure in developed and developing countries: a literature review. Corp. Soc. Responsib. Environ. Manag. 24 (4), 273–294.
Al-Tuwaijri, S.A., Christensen, T.E., Hughes Ii, K.E., 2004. The relations among environmental disclosure, environmental performance, and economic performance: a simultaneous equations approach. Account. Org. Soc. 29 (5–6), 447–471.
Baumüller, J., Sopp, K., 2022. Double materiality and the shift from non-financial to European sustainability reporting: review, outlook and implications. J. Appl. Account. Res. 23 (1), 8–28.
Belkhir, L., Elmeligi, A., 2019. Carbon footprint of the global pharmaceutical industry and relative impact of its major players. J. Clean. Prod. 214, 185–194.
Bercovitz, J., Feldman, M., 2006. Entpreprenerial universities and technology transfer: a conceptual framework for understanding knowledge-based economic development. J. Technol. Tran. 31, 175–188.
Bewley, K., Li, Y., 2000. Disclosure of environmental information by Canadian manufacturing companies: a voluntary disclosure perspective. In: Advances in Environmental Accounting & Management. Emerald Group Publishing Limited, pp. 201–226.
Blumenthal, D., Gluck, M., Louis, K.S., Stoto, M.A., Wise, D., 1986. University-industry research relationships in biotechnology: implications for the university. Science 232 (4756), 1361–1366.
Brammer, S., Pavelin, S., 2008. Factors influencing the quality of corporate environmental disclosure. Bus. Strat. Environ. 17 (2), 120–136.
Cue, B.W., Zhang, J., 2009. Green process chemistry in the pharmaceutical industry. Green Chem. Lett. Rev. 2 (4), 193–211.
De Wit-de Vries, E., Dolfsma, W.A., van der Windt, H.J., Gerkema, M.P., 2019. Knowledge transfer in university–industry research partnerships: a review. J. Technol. Tran. 44, 1236–1255.
Deegan, C., Gordon, B., 1996. A study of the environmental disclosure practices of Australian corporations. Account. Bus. Res. 26 (3), 187–199.
Doni, F., Martini, S.B., Corvino, A., Mazzoni, M., 2019. Voluntary versus Mandatory Non-financial Disclosure: EU Directive 95/2014 and Sustainability Reporting Practices Based on Empirical Evidence from Italy. Meditari Accountancy Research.
Fernandez-Feijoo, B., Romero, S., Ruiz, S., 2014. Effect of stakeholders' pressure on transparency of sustainability reports within the GRI framework. J. Bus. Ethics 122, 53–63.
Freeman, R.E., 1984. Strategic Management: A Stakeholder Approach. Cambridge University Press.
Gillet, J.M., 1952. Air pollution abatement by the chemical industry:—review and survey. Am. Ind. Hyg. Assoc. Q 13 (2), 59–61.
Gray, R., Kouhy, R., Lavers, S., 1995. Constructing a research database of social and environmentalreporting by UK companies. Account Audit. Account. J. 8 (2), 78–101.
Hegghammer, T., 2022. OCR with Tesseract, amazon textract, and google document AI: a benchmarking experiment. J. Comput. Soc. Sci. 5 (1), 861–882.
Hooks, J., van Staden, C.J., 2011. Evaluating environmental disclosures: the relationship between quality and extent measures. Br. Account. Rev. 43 (3), 200–213.
Houghton, J.T., 1992. Climate Change 1992, p. 212.
Ioannou, I., Serafeim, G., 2017. The Consequences of Mandatory Corporate Sustainability Reporting. Harvard Business School research working paper, pp. 11–100.
Jurafsky, D., Martin, J.H., 2009. Speech and Language Processing: an Introduction to Natural Language Processing, Speech Recognition, and Computational Linguistics, second ed. Prentice-Hall.
KPMG International, 2021. The KPMG Survey of Sustainability Reporting 2020. Retrieved from. https://kpmg.com/xx/en/home/insights/2020/11/the-time-has-come-survey-of-sustainability-reporting.html.
Kümmerer, K., 2010. Pharmaceuticals in the environment. Annu. Rev. Environ. Resour. 35, 57–75.
Lombardi, R., Schimperna, F., Paoloni, P., Galeotti, M., 2022. The climate-related information in the changing EU directive on non-financial reporting and disclosure: First evidence by Italian large companies. J. Appl. Account. Res. 23 (1), 250–273.
Mitchell, J.F., 1989. The "greenhouse" effect and climate change. Rev. Geophys. 27 (1), 115–139.
Mowery, D.C., Nelson, R.R., Sampat, B.N., Ziedonis, A.A., 2015. Ivory Tower and Industrial Innovation: University-Industry Technology Transfer before and after the Bayh-Dole Act. Stanford University Press.
Neu, D., Warsame, H., Pedwell, K., 1998. Managing public impressions: environmental disclosures in annual reports. Account. Org. Soc. 23 (3), 265–282.
Parker, L.D., 2005. Social and environmental accountability research: A view from the commentary box. Account Audit. Account. J. 18 (6), 842–860.
Parker, L.D., 2011. Twenty-one years of social and environmental accountability research: a coming of age. Account. Forum 35 (1), 1–10.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.
Pesci, C., Costa, E., 2014. Content analysis of social and environmental reports of Italian cooperative banks: methodological issues. Soc. Environ. Account. J. 34 (3), 157–171.
Python, 2022. Python Official Website. Retrieved from. https://www.python.org/. September 2022.
Raucci, D., Tarquinio, L., 2020. Sustainability performance indicators and non-financial information reporting. Evidence from the Italian case. Adm. Sci. 10 (1), 13.
Rezaee, Z., Tuo, L., 2017. Voluntary disclosure of non-financial information and its association with sustainability performance. Adv. Account. 39, 47–59.
Roberts, R.W., 1992. Determinants of corporate social responsibility disclosure: an application of stakeholder theory. Account. Org. Soc. 17 (6), 595–612.
Tesseract, 2022. Tesseract-OCR. Retrieved from. https://github.com/tesseract-ocr/tesseract. September 2022.
Unerman, J., 2000. Methodological issues-Reflections on quantification in corporate social reporting content analysis. Account Audit. Account. J. 13 (5), 667–681.
Venturelli, A., Caputo, F., Leopizzi, R., Pizzi, S., 2019. The state of art of corporate social disclosure before the introduction of non-financial reporting directive: a cross-country analysis. Soc. Responsib. J. 15 (4), 409–423.
Zamil, I.A., Ramakrishnan, S., Jamal, N.M., Hatif, M.A., Khatib, S.F., 2023. Drivers of corporate voluntary disclosure: a systematic review. J. Financ. Report. Account. 21 (2), 232–267.