In this paper, we present a multidisciplinary project, IMPRESS, which combines the digitization of three major nineteenth-century Belgian medical journals with a historical research project on the role of ideology in nineteenth-century Belgian medicine. We focus on the extent to which text mining has allowed us to identify and evaluate expressions of ideology in a corpus of medical texts. In showing how we used the digital tool AntConc to answer typically historical research questions, we intend to contribute to current debate on the gains and limitations of digital methods in the humanities. We conclude that, while acknowledging the many interpretative interventions in preparing searches and qualifying outcomes, the use of the tool has enabled us to shed new light on the role of ideology in scientific exchange. Text-mining operations have offered a fresh insight into the chronology of ideological vocabulary, the used language, and the distribution of ideological patterns across journal sections.
1. Introduction

In this paper, we present results from IMPRESS, the acronym of the multidisciplinary research project “Beyond ideological conflict: religion and free-thought in the Belgian medical press,” which was finished in 2021. Its first aim was to highlight the scientific potential of nineteenth-century Belgian medical journals as an undervalued type of heritage and to stimulate digital historical research into scientific journals. To this end, three major medical journals published between 1840 and 1914 were digitized and made publicly available. The second aim was to use these journals as the sources for a primary research project on the relation between medicine, ideology, and religion in nineteenth-century Belgium, testing in this way their potential for (digital) historical research.

While we will briefly present the process of digitization and the digital tools used in the first part of this article, the paper's main focus is on the results of the primary research. We will, more specifically, demonstrate the extent to which text mining has allowed us to identify and evaluate expressions of nineteenth-century worldviews in a corpus of scientific texts and the ways in which digital analysis was informed and complemented by contextual close reading of sources. Taking into account the current state of digital humanities scholarship, we believe that the field needs case studies that actually demonstrate the gains and limitations of using a specific digital tool to answer concrete research questions within an established humanities discipline such as (medical) history.

Literature on the promises and pitfalls of the use of digital tools in the humanities is abundant, and characterized by disagreement, as is testified by the term “The Digital Humanities Wars” (Da 2019). Within the discipline of history as well, enthusiasm about the opportunities of large datasets (e.g., Eijnatten, Pieters, and Verheul 2013; Laubichler, Maienschein, and Renn 2019) is paralleled by more critical voices that point to the need to “demystify” digital methodologies and increase debate (e.g., Piersma and Ribbens 2013). With regard to medical history and the study of historical periodicals, some pioneering empirical work has been done (e.g., Daems et al. 2019). Alongside research on the presence of medical discourse in the daily press (e.g., Huistra 2017), medical periodicals have been digitized and made accessible. This has in particular been the case for American and British sources. The American Medical Heritage Library digitized over 3000 volumes of mainly nineteenth-century medical journals. Important British journals such as the British Medical Journal and The Lancet have also been digitized and made (semi) accessible. Apart from recent work such as Eva Andersen’s dissertation on transnational psychiatric knowledge circulation, most empirical research based on these digitized sources has focused on the history of concepts, studying the changing

Even such seemingly uncomplicated types of questions have confronted historians with the need to always triangulate digital results with textual analysis and contextualization (Aiello and Simeone 2019). As medical historians Alex Mold and Virginia Berridge concluded in a comparative study of the use of terms related to addiction, “focusing only on terms and not the context in which they existed both on the page and more broadly risks missing the bigger picture” (Mold and Berridge 2019). Moreover, in research such as ours, that goes beyond the scope of a mere conceptual history and aims to understand the role of religion or ideology in discussions between doctors, more sophisticated forms of textual analysis are needed. Just like historians who study political debate, we have to take into account typical discursive aspects of debate such as quoting and paraphrasing, the use of irony or sarcasm, or the role of silence (Beyen 2013). Moreover, working with scientific journals also entails taking into account the role of different journal sections such as case reports, reviews, correspondence, meeting reports, and “miscellaneous” sections. With our study, we hope to demonstrate how indeed the combination of digital analysis and sophisticated textual analysis can help historians of medicine to answer such typically historical research questions.

The study also engages with a new historiography of medicine and religion. Until recently, the relation between these fields has often been imagined as one of conflict, particularly in the case of Catholicism with its strict teachings on questions of body and soul. Research into the relations between modern medicine and religion often assumed an opposition between both fields. Recent studies, however, have complicated this view by pointing to interactions and crossovers between religious and medical actors, beliefs, and practices (Guillemain 2003; Guillemain 2006; Donato et al. 2013; Ferngren 2014; Van Osselaer, De Smaele, and Wils 2017; Wils et al. 2022). These studies stimulate awareness of the diversity of interactions according to specific political contexts and medical disciplines. This implies, for instance, that the well-studied case of France, where sentiments of anti-Catholicism and radical liberalism were prominently present in the medical press, might not be representative for Catholic Europe as a whole (Goldstein 1982; Ellis 1990; Verhoeven 2012). This project sets out to challenge the (French–based) image of the nineteenth–century medical profession as an almost entirely liberal and anticlerical body by scrutinizing the role of religion and ideology within the scientific medical field in Belgium.

The Belgian case is rather unique. The country had no state church and an exceptionally liberal constitution (1831) that was supported by both liberals and Catholics, albeit for different reasons. Yet Belgium’s political culture of negotiation
could not prevent the rise of ideological tensions concerning specific issues, such as the role of religion in primary education—giving rise to a so-called school war (1878–1884)—or in institutions of medical care (Witte 2003). As the liberal state was barely involved in medical care provisions, religious orders had ample space to strengthen their position in the field of health care, to the frustration of free-thinking groups of liberals. Whether such a dynamic of compromise and incidental ideological conflict also marked scientific debate among physicians is what we wanted to research. In doing so, it is important to keep in mind that Catholic and liberal doctors were anything but well-defined groups. Only a minority of Catholic physicians overtly stated that they based their conduct on official clerical guidelines. Most Catholic physicians were more pragmatic and less inclined to refer to Catholic doctrine. Liberals too were divided along religious lines. Belgian liberal physicians could adopt a moderate, merely anticlerical position regarding the entanglement of Catholicism and Belgian society. More radical, free-thinking liberal doctors also contested the supernatural claims of Catholicism.

We inquire into the attitudes and professional identities of the physicians involved by answering questions such as: What was the importance of ideological views within scientific exchange? Were doctors willing to bypass ideological disagreements? How were professional and/or scholarly identities formed? Were they common for both groups, or ideologically specific? We treat these questions in three subsequent steps. We first sketch the general chronology of the presence and absence of explicit ideological language. We then zoom in on the role of specific concepts and the challenges we encountered in researching them. We end with an analysis of the role of different journal sections in creating space for ideology in a scientific setting. But before turning to these research results, we present the corpus, its digitization, and the text mining techniques we used to analyze the corpus.

2. The text corpus

The digitized text corpus of this study consists of three nineteenth-century Belgian medical journals: the Bulletin de l’Académie Royale de Médecine de Belgique (BARMB), the Journal de Médecine, de Chirurgie et de Pharmacologie (JMCP) and the Journal des Sciences Médicales de Louvain (JSML) and its successors. Together, they comprise 239 digitized volumes, counting up to 186,000 pages. Each of these journals has a distinct ideological profile, turning them into suitable sources to scrutinize the role of ideology and religion in the Belgian medical sciences. The BARMB (1841–1914) assembled the verbatim—yet embellished—meeting reports of the Royal Belgian Academy of Medicine (1841). The academy was an important meeting space for Belgium’s medical elite and consisted of—alongside other members—Catholic professors of the Catholic University of
Leuven and their liberal colleagues from the Free University of Brussels. Its meeting reports were widely read and commented upon in both the medical and the general press. The journal therefore formed an influential means for intellectual exchange and ideological profiling: a space where ideological divides (in particular between Catholics and liberals) materialized most clearly. The JMCP (1843–1896) was published by the Society of Medical and Natural Sciences of Brussels (1822). Since the latter society was run by the professors at the Free University of Brussels, the journal had a clear “liberal” profile. It aimed for an audience of private practitioners and published original articles, summaries of articles in the foreign medical press, book reviews, meeting reports of medical societies, and all sorts of professional and scientific news. The JSML was established in 1876 and succeeded by the *Revue Médicale* and the *Revue Médicale de Louvain*. In what follows, we will refer to these journals as “the Leuven journals” (1876–1913). This publication had a strong Catholic label, as it was edited by the professors of the Faculty of Medicine at the Catholic University of Leuven. While it aimed for a broad readership, it was particularly intended for Leuven graduates, providing them with a form of postgraduate education. It consisted primarily of original articles, with particular attention to those topics that might interest “Catholic” physicians.

The digitization of the bound volumes of the journals was carried out by the Royal Library of Belgium (KBR). This step implied the preparation of the documents, a metadata check, copyright clarification, image capture of each page, and OCR processing, followed by a quality control, then the production of XML METS, of permalinks (UURLs), and of special deliverables for the needs of research (PDF, TXT Files, etc.). For ingestion in the KBR long-term preservation, data had to be cleared of all project deliverables: the journals are published online and made available to the general public, on BELGICA (KBR Digital Library) and its thematic dissemination channel dedicated to periodicals (BELGICA Periodicals, https://www.belgicaperiodicals.be/?lang=EN). The KBR plans to offer these and other datasets to the public through a IIIF display infrastructure. To this end, it will launch a new website (www.data.kbr.be).

The quality of the journals published in the second half of the nineteenth and in the beginning of the twentieth century caused specific difficulties with OCR, a problem that is well known by researchers working with digitized historical texts (see for example Thompson, McNaught, and Ananiadou 2015). Specific issues with certain fonts, the use of capitals, and hyphenation proved too hard to handle for the OCR software and led to non-existing words in the resulting text, but also tables and schemes at times yielded bad results, when for instance words from the table erroneously ended up in running text. In order to deal with those problems, we developed a series of scripts, with the objective of eliminating as much noise as possible.
For example, the problem of hyphenated words, usually generating two words after OCR, is signalled as shown below in the ALTO-output of the OCR software. We developed a simple R script (i.e., a script written in the programming language R) to automatically substitute, for instance, “intéres” and “sante” with “intéressante” in our final digital corpus.

\[
\text{...}
\text{<String CONTENT="à" /> <String CONTENT="cette" />}
\text{<String CONTENT="intéres" SUBS_TYPE="HypPart1"}
\text{ SUBS_CONTENT="intéressante" />}
\text{<HYP CONTENT="" />}
\text{</TextLine>}
\text{<TextLine>}
\text{<String CONTENT="sante" SUBS_TYPE="HypPart2"

SUBS_CONTENT="intéressante" />}
\text{<String CONTENT="et" /> <String CONTENT="difficile" />}
\text{...}
\text{</TextLine>}
\]

We also took other steps to clean up the corpus. There was a need for normalization for the purpose of collation, due to the existence of different conventions across OCR tools with respect to the recognition of specific characters: for example, the preposition à (“to”) may be recognized as the entity \&agrave; or as a single character à, or as character a plus accent ‘; another example is the existence of “protected characters” in ALTO’s XML context, such as ‘. Again, we used an R script to perform this normalization.

Further possible steps, which we did not implement, would consist of using the confidence values reported by the OCR software to identify and correct low confidence output. The OCR software adds confidence values to the words it recognizes. For instance, the ALTO snippet \text{<String WC="0.150000006" CONTENT="Vexamen">} assigns low confidence. The attribute \text{WC} stands for “Word Confidence,” with values ranging from 0.0 [worst] to 1.0 [best]. In this example, \text{Vexamen} is rightfully assigned a low word confidence. The correct output would have been “l’examen.” Eventually, all these steps can offer a better text, but without manual verification, it is not possible to obtain a completely correct text.

As is shown by Thompson, McNaught, and Ananiadou (2015), who work on historical medical text archives of the \text{British Medical Journal} of the mid-nineteenth to the late
In the twentieth centuries, more measures can be taken to automatically correct large text corpora. They experimented with different spellcheckers to correct OCR errors (ASpell, Hunspell, Microsoft Word, and MAC OS), and decided to use Hunspell as the basis of their correction strategy. They needed to modify the tool and enrich it with archive frequency based re-ranking of Hunspell suggestions, using frequency lists obtained per decade of the journal issues they analyzed, and also with a medically enhanced Hunspell dictionary. Also, rule-based OCR correction was experimented with. In the end, a combination of a rule-based pre-processing of the corpus by the authors, in combination with the enhanced spellchecker, gave way to a corpus with a high degree of correct spellings. Likewise, in their experiment with Belgian socialist newspapers of the end of the nineteenth century, Daems et al. (2019) have used Transkribus for re-OCRing their corpus, which gives a high rate of correct digitization (up to 77%, or even 95% accuracy, depending upon the newspaper that was digitized). Unfortunately, we did not yet have the opportunity to execute this type of time-consuming treatment of the corpus.

With a view to creating new research possibilities, we did investigate the possibility of breaking down the issues of the journals into sections corresponding to the original content division of the journals. Making this organization by journal section explicit in the annotation, for instance, could enable the investigation of one single section within a journal (e.g., all book reviews, all research articles) as well as comparative research of one single section in different journals (e.g., all meeting reports in the three journals). The creation of sub-corpora can be automated by detecting certain boundaries between journal sections (i.e., training a computer program to detect the beginning of a section based on a different font size, capitalized titles, or white spaces between sections). An experiment with the automation of the Leuven journals, however, showed that such operations are quite time-consuming. First of all, the digital corpora did not have easily detectable boundaries between journal sections. The tables of contents did not convey information on the sections in a volume either, which implies that one should go through each volume to detect the starting point of each section. Secondly, the automation of sub-corpora could not overcome problems such as the absence of names of journal sections at the start of a significant number of sections (with the result that sections without names were not incorporated in the sub-corpora) or the problem of changing names of journal sections over time. The miscellaneous section in Belgian medical journals had, for instance, names such as Variétés, Nouvelles de la semaine, Mélanges scientifiques, Faits divers, etc. Creating workable sections hence implies the active intervention of the researcher, who has to identify sections and assign uniform headings to them (e.g., miscellanea for all types of miscellaneous sections). In other words, breaking down the volumes into sections would require a manual intervention for each of the 239 volumes, which we decided not to undertake.
Instead, we focused on text-mining techniques to study the text corpus in its entirety. We looked for digital research strategies that would enable an understanding of the place of religion and ideology in medical debate by studying the text across the now digitized journals, and which could not be achieved by close reading alone. A first, quite obvious, operation was the detection of relevant passages with respect to our research questions regarding ideological conflict in Belgian medicine, by distant reading via concordance tools. We also wanted tools to reveal language strategies by analyzing word frequencies, word combinations, and collocates, that is, words in the vicinity of a search term, in the context to its left and/or its right. Pre-existing ideas about historical developments can be challenged by comparing concordances and keywords for different time periods and journals. Finally, specific informed searches can test research results related to one particular medical debate in a larger corpus.

After having tested several tools, we chose to work with AntConc, developed by Laurence Anthony (Anthony 2005; 2013), that allows all these searches. We tested “Tropes” (version 8.4), software developed by Pierre Molette and Agnès Landré, based on Rodolphe Ghiglione’s work; cf. (https://www.tropes.fr/index.htm; see also Ghiglione et al. 1998; Molette 2009). The size of our texts and their specific nature, however, caused problems. We also tried out “LancsBox©,” developed by V. Brezina, T. McEnery, and M. Timperley (Brezina, McEnery, and Wattam 2015; Brezina, Timperley, and McEnery 2018); it allows, among other things, the comparison of concordances and visualization via collocation networks, but those tend to be overcrowded. The topic modelling tool “Mallet” (e.g., Graham, Weingart, and Milligan 2012) was also experimented with, but the first attempts did not generate meaningful results for the specific research questions of the project, because explicit discourses related to ideology and religion are rather scarce in the medical journals under study.

AntConc is very user-friendly and gives good results, as it allows large text corpora to be uploaded and analyzed. This software can select meaningful passages, using the concordance function, and investigate co-occurring words with the collocate function. The clusters tool within AntConc allows you to search for combinations formed by your search term with other words surrounding it, to see, for instance, what the most frequent word combinations are. The collocate tool, on the other hand, looks for non-sequential patterns in the journals. Unfortunately, AntConc analyzes only collocations in pairs. It does not allow semantic analyses. By means of the concordance plot tool within AntConc, it becomes possible to compare the frequency and the distribution of specific search terms for the three journals, as well as for different time slots (e.g., 1840–1850 and 1880–1890). It is an application that quickly visualizes the frequencies of a word in a number of files.
3. Chronology of ideological vocabulary

We used the tools in AntConc to provide answers on our research questions about the place of ideology in the scientific medical press in nineteenth-century Belgium. Both the concordance plot tool and the concordance function were useful to study the periods in which ideological conflict became most tangible in the three selected medical journals. We searched for queries, such as “Catholic*” (catholi*), “Christian*” (chr*tien*), “liberal*” (lib*ral*) and “freedom” (liberté), after which we contextualized the search results in relation to the date of publication. In doing so, this project has been able to distinguish the periods in which ideological discourse was most visible in the scientific medical press (see also: Andersen and Gijbels 2020). Until the 1860s, it was highly uncommon to refer to political statements, ideas, or events in scientific exchange. As the query “liberal*” (lib*ral*) in AntConc shows, words related to liberalism do not appear in a political context in the 1840s and 1850s, not even in a medical journal with a predominantly liberal membership such as the Journal de médecine. From the thirty hits found for the period 1843–1850 in this journal, for instance, not a single hit refers to political events or opinions of political liberals. If the search string “liberal*” did appear, it was often in paragraphs where the liberal profession of medicine is addressed. (Figure 1: “lib*ral*”)

Figure 1: The first twenty-four results of the search query “liberal*” (lib*ral*) in the concordance function of AntConc. The results are found for the years 1843–1850 in Journal de médecine, de chirurgie et de pharmacologie. They are sorted according to one neighbouring word to the left and one to the right.
Findings such as these make clear that numbers of hits never speak for themselves. Interesting findings resulting from distant reading cannot be properly explained without close reading. For each of the hits, we verified the relevance of the results with a view to our specific research questions. We did this by going back and forth between the different functionalities in AntConc: from the concordance tool (where the word appears in its immediate context), to the file view (in which the word appears in the text file). If those functionalities did not provide enough clarity about the context of the keyword, the PDF-format of the journal volumes was consulted.

The lack of relevant hits for ideologically relating strings such as “liberal*” makes one thing clear: the editors of the Journal de médecine did not attempt to give the journal a liberal label, nor to emphasize oppositions between liberals and Catholics in mid-nineteenth century Belgium. This finding confirms a hypothesis about the scientific ideals of Belgian physicians. The expression of political views was seen as incompatible with the scientific aims of physicians engaging in medical societies and scientific publishing. As research on the scientific codes of conduct within Belgian medical societies has shown, both political questions and conflict were avoided during society meetings. Physicians even feared the polemical tone that often characterized parliamentary debates between Catholics and liberals (Vandendriessche 2018).

A comparison of hits found for the period 1840–1900 in the three selected medical journals has shown that ideological polarization marked the scientific medical press during the school war. More than in previous periods, brief comments on political events revealed the political colour of a member. In 1878, for instance, a Brussels professor who had been elected as a liberal member of the Belgian senate emphasized the importance of the liberal electoral triumph (Pigeolet 1878). The increased ideological tensions were also noticeable in physicians’ discourse about science and religion. For the first time, statements were published about the incompatibility of science and religion. A few anticlerical physicians portrayed science as a “rational” and independent field that had to be safeguarded from religious influences. Religion, on the contrary, was associated with dogmas and a lack of freedom. This type of conflict model mainly marked liberal reflections on the relationship between the Free University of Brussels and the Catholic University of Leuven (Van den Corput 1876).

The majority of scientific discussions, however, lacked visible traces of the increased ideological tensions on the Belgian political scene in the 1870s. Even within those medical disciplines that have a history of ideologically charged discussions, the political and religious ideas of physicians were remarkably absent. Obstetric debates are a clear example of this. More than any other medical domain, obstetrics confronted nineteenth-century physicians with ethical and religious questions relating to the
physical and spiritual life of fetuses in peril of death. Our text-mining operations in AntConc have helped to identify large debates about difficult births in the 1840s and 1850s. These debates focused on procedures to baptize the unborn in the uterus on the one hand, and on the morality of obstetric interventions ending in the death of the fetus on the other hand (Gijbels 2019; Gijbels 2020). Central to the latter discussion was the question of whether or not physicians could give priority to the life of a woman when it was impossible to save both her life and the life of her unborn child. For the most part of the nineteenth century, the caesarean section had an extremely high risk of maternal mortality, as a consequence of which physicians turned to operations that jeopardized or ended the life of the fetus in an attempt to save the woman. In the 1860s, however, the morality of lethal interventions ceased to elicit ideologically charged discussions. While Catholic and liberal doctors were still divided over procedures with a fatal outcome for the fetus, they stopped justifying their conduct in medical publications. Doctors now seemed to avoid open confrontations over ideologically charged medical debates.

Interestingly, the school war is characterized by historians as the most ideologically tense period of the nineteenth century in Belgium, but in obstetric debates, it was a period with remarkably few ideological statements. After the school war, around 1890, ideology again entered the scholarly meeting room in a context of obstetric innovations. At the moment that surgical improvements resulted in better survival chances for women who underwent a caesarean section, Catholic and liberal doctors renewed their discussion on the morality of procedures that sacrificed fetuses. The more the caesarean section came to be seen as a rational choice that could save two lives, the more lethal operations on fetuses were contested. In fact, close reading in medical journals has revealed that ideological conflicts were caused by “internal” medical developments, rather than by “external” political tensions. The two periods with most ideological confrontations in obstetric debate took place much earlier (ca. 1840–1860) and much later (1892) than the school war. In both cases, obstetric developments led Catholic physicians to raise objections against obstetric interventions that sacrificed fetuses (Gijbels 2021).

The only setting where the controversial issue of lethal obstetric interventions was regularly brought up during the school war was the Leuven journals. Distant reading, by means of the concordance tool, and close reading have revealed that most references to “feticide” were made by Catholic professors. Further analysis of the context of this search query has shown that it was used to denounce obstetric operations resulting in the death of fetuses. Similar to “feticide,” variations to the verb “to kill” were central to the discourse of Catholic doctors. They served to underline the injustice that was
done to fetuses. Such recurring statements about the immorality of lethal interventions served the purpose of identity-building amongst Catholic doctors. After all, the journal was mainly targeted at a homogenous audience of doctors who (had) studied in Leuven. Sometimes medical articles were limited to the religious point of view with this audience in mind (Gijbels 2021).

4. Concepts with an ideological connotation

Text-mining operations in AntConc have also supported our analysis of the rhetorical devices of physicians. In ideologically charged interactions, Catholics and liberals used particular figures of speech that can be identified as typically “Catholic” or “liberal.” Starting from our knowledge of the historiography of ideological conflict, we studied the occurrence of such rhetoric in our digitized medical corpus using the concordance tool and the concordance plot tool. The topos of “free enquiry,” for instance, characterized liberal discourse about the Free University of Brussels. Historians have shown that, in the decades following the foundation of the Brussels university (1834), it became an integral part of the public self-representation of the university. University representatives, in particular, used it to emphasize the rejection of authority—including religious authority—and the fundamental importance of scientific freedom at the university. In the beginning, this discourse evolved around the concept of “freedom of education.” It was used to underline that professors at the Brussels university were free to teach scientific doctrines in complete independence. In 1854, the concept of “free examination”—a concept that is still in use up to today—made its entrance in a speech by Pierre-Théodore Verhaegen, the founder of the university. From then on, the emphasis was put on the freedom of scientists to develop scientific theories without external influences (Stengers 1959).

However, as queries in AntConc show, the concept figured much less in scientific medical journals than in public performances. A query in the *Journal de médecine*, for instance, shows that the concept appeared only fourteen times in relevant contexts in the 1850s, 1860s, and 1870s. In the journal it was for the first time introduced in 1855, after which it was sporadically used to underline the importance of independent scientific discussions and research. By the 1870s, the period building up to the school war, the liberal topos of “free inquiry” became associated with ideas about the incompatibility of science and religion. Brussels academics and liberals started using the topos to make a distinction between the religiously inspired research at the Catholic University of Leuven and the so-called independent rational research at the Free University of Brussels. In *Journal de médecine*, this new meaning of the topos was also, though to a lesser extent, noticeable. (Figure 2: “libre examen”)
The liberal topos of free enquiry is an example of a rhetorical device used by liberal physicians to define what it meant to be a liberal scientist. Yet, there are also examples of words that served to put the “politically other” in a negative light. The word “materialist,” for instance, functioned as a powerful label with a pejorative charge to denote anticlericals. The word was mostly used by Catholic doctors. It suggested that liberal doctors did not acknowledge any reality outside material reality, implying an anti-religious and amoral stance which denied the free will and the immortality of the soul. Within medicine, the term could also refer, more specifically, to an Enlightenment-inspired current of medical thought that aimed at explaining both bodily diseases and human consciousness and conduct in terms of physical rather than immaterial substances. Radical materialism had its origins in France around 1800. At that moment, a number of French doctors, including Pierre Jean Georges Cabanis and François-Joseph–Victor Broussais, started explaining medical diseases in terms of affected organs and physiological processes. Their approach was highly controversial at the time. After all, theories of materialism clashed with prevailing spiritualist conceptions casting diseases in more abstract religious ways (Goldstein 1987; Daled 1998; Albury 1998; Clark and Kaiser 2003; Wils 2005).

Informed by historical scholarship on materialism, we wanted to verify to what extent materialism penetrated the Belgian medical press. Strikingly, we found almost no substantive discussions in which philosophical materialism was embraced. In fact, in the second half of the nineteenth century, “materialist” functioned as a label with
self-explanatory power. The word seemed hardly ever to be associated with particular content, let alone a specifically medical content. This suggests that its significance in scientific medical journals was very similar to its meaning in public debates, in the wider context of discussions on the role of religion in the organization of society. It also suggests that medical materialism, which had been so central in French medical debate of the early nineteenth century, was not present in the Belgian medical press of the second half of the century.

In the digitized corpus of the Leuven journals, the tool AntConc delivered 21 hits for the truncation “materialis” (*matérialis*). In order to analyze the role and meaning of these references, we distinguished three types of connotations, positive, negative, and neutral, based on our interpretation of the sentence in which the words occur (in Figure 3). In other words, a sentence was marked as “neutral” when it referred to a person, group, or idea without a particular connotation. This approach obviously has limitations. A sentence can appear as neutral, whereas the paragraph or the article of which it is a part can convey a sense of conflict. As it turns out, the string “materialis” was only used in a negative (e.g., as a swear word) or a neutral way (e.g., as part of a dissertation title). Not surprisingly, the bulk of hits with negative connotations were found for the period that is known as the school war and in journal sections dealing with the question of difficult births. This finding supports our previous conclusions about the occurrence of ideologically charged words in scientific discussions. Compared to their liberal colleagues, Catholic physicians were more tempted to speak in negative terms about physicians with different ideological views during the school war. (Figure 3: *matérialis* in Leuven journals)

![Figure 3: Number of positive, negative and neutral occurrences of the string "materialis" (*matérialis*) generated by AntConc in the Leuven journals between 1876 and 1913.](image-url)
Our analysis of the immediate context of the hits revealed that the string “materialis*” most often appeared in negative contexts of accusation, opposition, and disapproval. One of the editors of the Leuven journals, for instance, often evoked the idea of a struggle between what he called the spiritualist school and the materialist school. According to the medical professor, these were two irreconcilable camps:

On the one hand, the legion of positivist or materialist doctors, among whom certain polemicists who willingly represent the opponents of feticide as retrogrades and the partisans of the caesarean operation as fanatics or savages cutting down the tree to get the fruit—on the other hand, a small church, those who could be called the faithful, refusing to carry out, no matter how advantageous it may be, an operation that is condemned by natural law, morality, and the Catholic religion. (Hubert 1884, our own translation)

In the example of the quote, the label “materialist doctors” refers to liberal physicians who carried out obstetric interventions that sacrificed fetuses in order to save the lives of birthing women. “The faithful” are mostly Catholic doctors who contested this type of procedure. By distinguishing “materialist” or “positivist” doctors from Catholic doctors, the author painted a picture of a highly polarized medical profession.

As regards the pejorative charge of the label “materialist,” it does not come as a surprise to note that physicians did not self-identify as materialist. Search results in the Journal de médecine have shown that even radical free-thinkers considered the qualification of materialist to be a real insult and did not use it. This conclusion fits in with earlier findings on the nature and tone in public debates within broader intellectual circles and the daily press in the nineteenth century. Here too, liberals distanced themselves from the label materialist in ideologically charged discussions. Yet, compared to public debate, medical interactions in scientific exchange were less harsh. This result confirms earlier research on learned medical societies, where written and unwritten rules emphasized the need for a dignified debate, which meant that politics had to be left outside societies’ meetings halls (Vandendriessche 2018). It also shows a certain continuity with a heritage of the rather literary style of eighteenth-century physicians (Porter and Roberts 1993; Wenger 2014).

5. Ideology and journal sections

Finally, the use of the digital tool AntConc has offered invaluable insights into the genre of nineteenth-century medical journals and the practice of medical science. To assess the weight of ideological division in the medical press, the sections where physicians made their politically charged statements matter. In the Leuven journals, traces of religion and ideology did occur most often in the miscellaneous sections (Nouvelles,
Nouvelles et faits divers, Mélanges scientifiques). These sections did address politics and university related topics, such as awarded prizes, anniversaries and educational laws, besides medical themes. In contrast with these short articles, the “scientific” core of the journal rarely contained references to Catholicism and liberalism. This core was formed by thematic papers that discussed medical observations or scientific findings. They not only constituted the largest parts of medical journals, they were also the first sections that physicians read. As the distribution of words containing the string “catholi*” (catholi*) shows, digital techniques can support historians in assessing the function of the journal sections as “areas” where ideology was either embraced or left out. In the midst of the school war, medical journals apparently shaped a community in which some disagreement and identity formation was possible in the margins, but also with clear “scientific” spaces that exceeded ideological preoccupations. (Figure 4: “catholi*” in Leuven journals)

![Figure 4: Number of occurrences of the string catholi*, according to sections, generated by AntConc in the Leuven journals between 1876–1895.](image)

As regards the Brussels journal, it is harder to draw conclusions about the function of journal sections, considering the very limited number of relevant hits for the search string “catholi*” (catholi*). For the period 1870–1895, we found twenty-three results in total. Overall, the sections with meeting reports contained most hits (fifteen). The
Journal de médecine usually published summarized reports of the Society of Medical and Natural Sciences of Brussels, those of the Belgian Academy of Medicine, and the reports of the French Academy of Medicine. Compared to other journal sections, such as sections with academic articles, necrologies, reviews, and the miscellaneous sections, meeting reports were the most performative genre. Proceedings evoke the atmosphere of the oral discussion in the meeting room. They constituted a genre that left room for bold statements, accusations, and the expression of ideological identity. (Figure 5: catholi* in Journal de médecine)

![Occurrence of catholi* according to sections in Journal de médecine 1870-1895](chart)

**Figure 5:** Number of occurrences of the string catholi* generated by AntConc in the Journal de médecine between 1870 and 1895.

The difference in the distribution of hits in the meeting reports between the Brussels journal and the Leuven journals can be explained by the different position of this section. Unlike their Brussels colleagues, Leuven professors did not systematically report on meetings of medical societies, which explains the limited amount of hits (two on a total of sixty-nine hits) within the section of meeting reports. Yet, the few summarized meeting reports that appeared in the journal support the view that ideological views could thrive within performative genres such as the meeting report. Some of the summaries of meeting reports in the Leuven journals explicitly responded to ideological provocations in the meeting rooms of medical societies. In 1878, for instance, the Leuven professor Ernest Masoin denounced the speech
of the anticlerical Brussels professor Jean-Hubert Thiry in the Belgian Academy of Medicine:

[B]ut how can Mr. Thiry seriously say, in the middle of the Academy, that the foundation of the existing University of Leuven was a danger, while the creation of the University of Brussels, “which removed Belgian higher education from political and religious constraints, was a subject of astonishment and admiration for intelligent Europe?” Ah! if someone were to take the liberty of repeating these expressions in reverse before the Academy—which we could, by the way, not approve of—and of representing the University of Brussels as a danger to the fatherland, to civilization, etc., what do I know?—Would we not see the honorable Mr. Thiry leaping to the tribune as one of the first to protest against such reckless assertions? (Masoin 1878, our own translation)

However, the weight of ideology in the meeting reports of the Leuven and Brussels journals was modest in comparison with its presence in the proceedings of the ideologically mixed Belgian Academy in the period leading up to the school war. Between 1873 and 1875, two major debates were held that touched on fundamental ideological principles. The first discussion focused on “general paralysis,” a disease that is today known under the name neurosyphilis. People suffering from it had a wide range of physical symptoms including paralysis, hallucinations, and dementia. According to Catholic doctors, the disease was caused by the declining influence of religion in society. Liberal physicians, for their part, contested this religion-based condemnation of modernity. The second debate centred on the explanations of the religious exaltations and wounds of the Belgian stigmatic Louise Lateau. According to Catholic doctors this corporeal phenomenon was inexplicable from a scientific perspective. Liberal physicians contested such supernatural interpretations and attributed Lateau’s symptoms to physiological causes. Both debates were marked by bold statements about the Catholic religion and its relationship with medicine (Lachapelle 2004; Andersen 2021). (Figure 6: catholi* in Bulletin)

Findings about the distribution of traces of ideology shed new light on the history of ideological conflict. Apparently, the expression of ideological views was limited to sections that reflected an oral culture of debate (meeting reports) or to less important sections that existed at the margins of medical journals (miscellaneous sections). These results urge historians to take the place of ideology in medical publications into consideration. Whether religious concerns and moral reflections were addressed in medical journals largely depended on the genre of the journal section.
6. Conclusion

With the digitization of three important nineteenth-century medical journals, corresponding to up to 186,000 published pages, an important step has been taken in the process of making Belgium’s scientific medical heritage publicly accessible in a digital format. In this article, we have offered a view of the process of digitization and the used digital tools. The main focus of the article was, however, on the use of text mining to answer historical research questions. We have discussed the different ways in which digital analysis has allowed us to gain insights in the role of religion and ideology in nineteenth-century medical journals in Belgium. In what follows, we reflect briefly on the added value and the limitations of AntConc, the text-mining program that seemed most suitable for our research.

It is quite obvious, in the first place, that AntConc has been a tremendous help in mastering a corpus of so many pages. It has allowed us to detect general tendencies that we would never have detected with such certainty using traditional research methods. Without the availability of a digital corpus, historians are mostly dependent on extensive medical debates, in which ideology figured prominently. These discussions are easy to detect by browsing the indices of the physical journal copies. The employment of text mining techniques has, however, led us to a considerable amount of small passages of text that we would not have been able to find without distant reading. They provided

![Occurrences of catholi* according to sections in The Bulletin de l'Académie](image)

**Figure 6:** Number of occurrences of the string *catholi* generated by AntConc in the *Bulletin de l'Académie* between 1870 and 1895.
additional evidence to make claims about the discursive strategies of Catholic and liberal physicians. By searching the digital corpus, we came to the conclusion that language referring to religion and ideology was much less present in the medical journals than we might have expected on the basis of existing secondary literature.

The detection of these small passages of text also advanced our historical understandings of the role of ideology in Belgian medical journals in several ways. We found, firstly, that there was a clear rise in religious and strong ideological vocabulary in the time period of the so-called school war, an era of polarized public debate on the position of religion in education. However, this rise was less visible in scientific discussions. While small articles dealing with political developments in Belgium were likely to reveal the ideological convictions of the author, scientific articles about medical issues generally lacked religious and ideological statements. In this respect, our digital research has shown that if such statements were made in the context of scientific debate, they occurred more often in the Leuven journals. Apparently, Catholic physicians were more tempted to create a polarizing ideological climate in scientific debate during the school war.

Our findings with regard to the ideologically charged figures of speech are similar to those about the chronology of ideological figures of speech. Concepts with an ideological connotation that often appeared in public debate, such as “free enquiry” and “materialism,” were also found in the digitized medical corpus. Yet, all in all, the use of these figures of speech in medical settings of scientific exchange was rather limited. Another similarity concerns the combative language of Catholic physicians. More than liberal physicians, they used labels with a pejorative charge such as ”materialist” to denote their ideological counterparts.

Finally, distant reading has led to insights into the “areas” in medical journals where expressions of ideology appeared most. By taking into account the variety of text genres (the different types of sections) within the journals and linking the occurrence of specific words to these different genres, we were able to show how religion and ideological conflict were quite absent within medical articles, but could nevertheless play a role in miscellaneous sections and meeting reports. Such a distribution of ideological patterns allowed physicians to live up to an ethos of scientific purity without therefore having to be completely cut off of larger social and political identifications and tensions.

Besides the advantages of digital analysis for historical research, we were also confronted with its limits. When we tried to qualify general quantitative tendencies, we encountered a number of difficulties. Close reading revealed that caution remains
necessary when interpreting the frequency of hits generated by AntConc. After all, the amount of relevant search results relating to religion and ideology in the total corpus remained quite low. In the whole period under study, it was rather exceptional that ideology became visible in medical debates. Closer analysis, moreover, revealed that many found hits turned out to be not so meaningful in relation to our specific questions. The word Catholic is for instance often mentioned as part of the institutional name of the Leuven university (*Université Catholique de Louvain*). Such examples show that historical explanations about the importance of debates cannot be derived from the mere statistical occurrence of particular words. Hermeneutical text analysis is always needed to interpret search results within the discursive context in which they have been used. Moreover, close reading involves many invisible interpretative choices. We were obliged to reflect on these choices when trying to categorize particular phrases as “positive,” “negative,” or “neutral.” In short, we came to realize that working with a text-mining tool in no way releases historians from the task of close reading.
Competing Interests
The authors have no competing interests to declare.

Contributions
Authorial
Authorship is alphabetical after the drafting author and principal technical lead. Author contributions, described using the CASRAI CredIT typology, are as follows:

Author name and initials:
Jolien Gijbels – jg
Michèle Goyens – mg
Frédéric Lemmers – fl
Dirk Speelman – ds
Joris Vandendriessche – jv
Cécile Vanderpelen – cv
Kaat Wils – kw

Authors are listed in descending order by significance of contribution. The corresponding author is Kaat Wils

Conceptualization: kw, mg, jv, jg
Methodology: jg, mg, ds, kw, jv
Software: fl, mg, ds
Formal Analysis: jg
Investigation: jg
Data Curation: fl, mg, ds
Writing – Original Draft Preparation: jg, kw, mg, jv, ds
Writing – Review & Editing: jg, mg, ds, kw, jv, cv

Editorial
Recommending Editor
Dominic Forest, École de bibliothéconomie et des sciences de l’information, Université de Montréal, Canada

Section/Review Editor
Iftekhar Khalid, The Journal Incubator, University of Lethbridge, Canada

Bibliography and Citations Editor
Morgan Pearce, The Journal Incubator, University of Lethbridge, Canada
References


