

# Changing portrayals of medicine and patients in eighteenth-century medical writing

## Lexical bundles in public health, methods, and case studies

Anu Lehto

University of Helsinki

This chapter turns the attention to three-word lexical bundles and aims to reveal repeated linguistic elements and text-external concepts in medical texts. The study is carried out in the corpus categories of public health and methods and additionally in the genre of case studies that are found across the corpus. The analysis applies a corpus-driven approach and assesses the functions of the word strings by categorising them into referential and textual bundles and stance expressions. The survey reveals various differences and diachronic changes in each of the corpus categories. Texts on public health concern the wellbeing of citizens and discuss medical matters on the social level. Texts on methods, in contrast, concentrate on medicines and the diseases of the patients, and case studies similarly focus on the patient and employ narrative structures. The bundles indicate novel practices in eighteenth-century medical writing, i.e. many bundles point to the increasing importance of observation especially in methods and case studies, and bundles related to statistical methods and quantification are found among the repeated constructions. Further, the bundles display new themes in medicine, e.g. texts on public health contain bundles that refer to the hospital movement and extend an understanding of hygiene. At the same time, the distribution of the functional categories is rather similar in the material, as referential bundles prevail followed by textual bundles and stance expressions, indicating many continuities in medical writing that stem from earlier periods.

### 1. Introduction

This chapter assesses lexical bundles in the corpus of *Late Modern English Medical Texts* (LMEMT), focusing on their functions. These recurrent expressions are of interest in the corpus, as bundles have the potential to reveal pertinent features of eighteenth-century medical writing and the practising of medicine in the text-external reality. The analysis of bundles, hence, traces routine wordings in historical medical writing, but at the same time the attention is on diachronic changes and differences between the corpus categories. The survey encompasses three datasets: public health and the subcategory of methods in specific treatises as well as the subgenre of case studies distributed across the corpus. These three categories exhibit variation in the functions of lexical bundles, portraying differing views of the text-external reality, such as research methods and patients as well as showing variation in the authors' attitudes.

Lexical bundles reflect the idea that language use is often based on routine wordings rather than novel lexical combinations (see e.g. Sinclair 1991). A considerable portion of speaking and writing consists of formulaic expressions. Thus, Biber et al. (1999: 995) calculate that in conversation, approximately 30% of the wordings are prefabricated word strings and 21% in academic prose. Similar communicative situations with the same goals are repeated in society, and these settings strengthen the use of formulaic wordings. Therefore, speakers do not combine individual lexical items but unconsciously restate prefabricated strings of words, which are treated as single chunks

by language users (e.g. Sinclair 1991: 110; Ellis 1996; Wray 2002: 9). Although lexical bundles reveal routine wordings, genres are also dynamic and change across time to suit new communicative purposes (Taavitsainen 2001a: 139). In general, medical writing evolved towards empirical research and quantification in the Late Modern English era (see e.g. Tröhler 2005), and these developments can be observed in the repeated constructions in the data.

The three corpus categories chosen for scrutiny pose different goals and reveal new writing conventions in medicine. Thus, a novel development in Late Modern English medicine was the emergence of texts on public health, which forms a new category in the LMEMT corpus (see public health category description and Taavitsainen et al. 2014). The lexical bundles in this category reveal how medicine and public welfare are perceived in society and how the wellbeing of citizens is viewed in the texts. In contrast, specific treatises on methods centre on medical procedures, and they portray a growing interest in observation rather than relying heavily on the works of ancient authorities (see the methods category description in this volume). My analysis seeks to examine whether this shift in thought styles is visible in the functions of the lexical bundles. The case studies further provide material from a broad range of medical texts on one subgenre that narrate the medical cases of a single patient. Case studies have a long history in medical writing, as they originate from the medieval period and were already present in the Middle English period (see Taavitsainen 2011; Taavitsainen and Pahta 2000). The case studies traditionally use narration and follow the structure of natural narratives, but many innovative writing conventions emerged in the Late Modern era, e.g. quantification of case studies to form general conclusions (see Chapter 5 in this volume). Therefore, the growing emphasis on observation and quantification is present in the selected categories, and is reflected in the functions of the bundles.

Methodologically, the analysis of lexical bundles focuses on three-word bundles and applies a bottom-up and corpus-driven approach, as the word strings are retrieved from the LMEMT corpus automatically without preconceived ideas or pre-established word lists of bundles (see Tognini-Bonelli 2001: 65; Meyer 2015). The data is analysed with quantitative corpus methods that have been developed for this particular purpose, and the corpus queries make use of the AntConc software (Anthony 2014). Previous studies have identified various functions of lexical bundles in genres, and the bundles in the data are analysed against these categories, making use of both top-down and corpus-based approaches. Numerous novel functions emerge in the medical material, such as references to patients and body parts, while some functions typical of other genres, such as topic introduction are not found in the data.

## **2. Theoretical background of lexical bundles**

### **2.1 Lexical bundles in earlier studies**

Previous research has scrutinised lexical bundles mainly in present-day material, but there has been a growing interest concerning bundles in historical data as well; these assessments do not, however, cover bundles in medical writing from the Late Modern English period. The first quantitative studies on bundles concentrated on present-day spoken language and academic writing. Altenberg (1998) has analysed the grammatical structures of repeated word strings on the basis of frequency in spoken language in the *London-Lund Corpus*. Biber et al. (1999: 993) have explored lexical bundles in the registers of conversation and academic prose, forming an initial taxonomy of grammatical structures of bundles. Biber, Conrad, and Cortes (2004) have continued the analysis of lexical bundles in university classroom teaching and textbooks and launched a grammatical and functional

classification of the bundles (see also Biber 2006; Cortes 2004). They note that lexical bundles vary in spoken and written genres, as the discourses reflect different communicative functions and purposes, resulting in dissimilar writing conventions (Biber, Conrad and Cortes 2004: 400). Further, Hyland (2008) finds that bundles do not merely differ in genres but also within more specific discourses, such as in various disciplines of research writing including engineering and applied linguistics. In more recent studies, the research on lexical bundles has extended to a variety of other genres, e.g. lexical bundles have been studied in legal language, which has been generally noted to be more formulaic than other genres (Breeze 2013: 233; Lehto 2015).

Lexical bundles have been additionally assessed in earlier studies that concentrate on historical material, especially on historical legal language. Culpeper and Kytö (2010) consider the functions of lexical bundles in spoken language in Early Modern English trial proceedings and drama, contrasting these genres with their present-day counterparts. Formulaic language in written legal documents is analysed in the texts of Scottish burghs (Kopaczyk 2013a) and in Early Modern and present-day legal language (Lehto 2015, 2018). In relation to my own study, Kopaczyk (2013b) investigates lexical bundles diachronically in Early Modern English medical language by retrieving about 230 shared bundles from the corpus categories of EMEMT. The analysis divides the bundles into three categories, i.e. bundles sharing semantic areas, functional areas, and structural frames, and has discovered that a higher number of formulaic expressions stem from the later periods in EMEMT and that references, for instance, to quantities, body parts, and efficacy statements are found across the corpus categories.

## 2.2 Frequency and distribution of lexical bundles

Lexical bundles are continuous word strings of three or more words that are repeated in numerous texts in a corpus (see Biber, Conrad and Cortes 2004). The main criterion of regarding a string of words as a bundle is that the linkage of the words is more frequent than co-occurrence caused by chance (e.g. Biber et al. 1999; Hyland 2008: 5; Gries 2008:6). The bundles are hence recognised by their frequency and by their distribution over a set number of corpus texts, although some studies on repeated structures survey prefabricated lists of known idioms (e.g. Moon 1998). The earlier studies that rely on the frequency method have used different cut-off points for both frequency and distribution. The frequency has ranged at least from 10 to 40 instances per million words. The threshold of 10 instances per million words is used in the earliest studies by Biber et al. (1999) and Altenberg (1998), while the later studies on academic registers by Hyland (2008) and Cortes (2004) use the cut-off point of 20 per million words and Biber, Conrad, and Cortes (2004) and Biber (2006) rely on the normalised frequency of 40 per million words. For distribution, Hyland (2008) uses a dispersion that covers 10% of the corpus texts, while Biber et al. (2004) and Biber (2006) as well as the works on historical data by Culpeper and Kytö (2010) and Kopaczyk (2013b) study bundles that distribute over five or more corpus texts. Kopaczyk (2013a) includes bundles that stretch over at least 10 texts in the corpus of Scots burghs.

The corpora in these studies vary in size, and hence, the thresholds for distribution encompass different percentages of the corpus texts in each study. My scrutiny is in line with the frequencies and dispersion used in the earlier assessment, i.e. the normalised frequencies range from about 34 to 43 per million words and their distribution covers from approximately 9% to 18% of the corpus texts within one category (see Section 3).

Word strings of different lengths can be attested and earlier studies have explored word sequences extending at least from three to eight words. The earlier studies on various university genres concentrated on bundles of four words (Biber et al. 1999; Biber, Conrad and Cortes 2004, Biber 2006; Hyland 2008; Cortes 2004). It has been noted that shorter lexical strings often appear as part of longer bundles and that the shorter bundles are additionally more numerous than longer word strings (Biber et al. 1999: 992). This phenomenon can be seen in LMEMT as well. Table 1 lists the number of bundle types and tokens for bundles of different lengths in the three studied sets of data:

**Table 1.** Number of bundles of different lengths

Bundle length	Public welfare		Methods		Case studies	
	Types	Tokens	Types	Tokens	Types	Tokens
3-word bundles	859	9016	595	6515	563	7098
4-word bundles	71	746	59	689	66	757
5-word bundles	5	37	2	22	6	54

The bundles in each corpus category are not completely comparable, as different thresholds for frequency and distribution have been used in retrieving the bundles (see Section 3 on methods). However, it can be seen that the number of bundles considerably decreases as the bundle length increases, and there are no six-word bundles in any of the studied corpus categories. In public health, four-word bundles are twelve times less frequent than three-word bundles, while in methods, they are ten times less numerous; in case studies, the three-word bundles are over eight times more common. Similar differences within bundle lengths have been found in present-day conversation and academic prose, in which three-word bundles are nearly ten times more frequent than their four-word counterparts (Biber et al. 1999: 993). The difference is considerably less noticeable in historical legal language, where three-word bundles are only slightly over two times more numerous, showing the formulaic nature of legislative documents (Lehto 2015: 329). In my own analysis, the bundles are most numerous in public health, which is also the largest category of the three analysed data sets (see the discussion in Section 7).

The length of the bundle hence has an effect on the retrieved number of word strings. Further, Hyland (2008: 8) and Cortes (2004), among others, state that four-word bundles provide more complete word strings for analysis than three-word bundles. Biber et al. (1999: 990), nevertheless, note that lexical bundles are typically non-idiomatic and hence, do not form complete structures or meaning units (cf. Moon 1998). My analysis focuses on three-word bundles, since these word combinations offer enough bundles for investigation (see Table 1 above). Earlier studies on bundles in historical material have similarly chosen three-word bundles for analysis, i.e. Culpeper and Kytö (2010) have scrutinised three-word bundles in historical spoken legal language, while Kopaczyk (2013b) has investigated them in Early Modern English medical writing and Lehto (2015) in Early Modern English legal writing. Kopaczyk (2013a) further extends the analysis from three-word bundles to word strings of eight words, as the focus is on language standardisation (see also Breeze 2013).

### 2.3 Grammatical structures and functions of bundles

Lexical bundles are typically categorised according to their grammatical structure and function. Biber, Conrad, and Cortes (2004) establish three grammatical categories with various subcategories: noun and prepositional fragments, verb phrase fragments, and dependent clause fragments. Noun and prepositional fragments encompass five subcategories such as Noun phrases with *of*-phrase fragments and Noun phrases with other post-modifier fragments. Verb phrase fragments are divided into seven subcategories including Verb phrases with passive and non-passive verbs and Third-person pronoun + discourse marker. Dependent clause fragments combine five categories such as *Wh*-clause fragments, *That*-clause fragments and (verb/adjective +) *to*-clause fragments.

A functional classification in Biber, Conrad, and Cortes (2004) renders (1) referential bundles, (2) discourse organisers, (3) stance expressions, and (4) special conversational functions: (1) referential bundles denote concepts outside the text, such as time, objects, and people, (2) discourse or textual organisers allude to the text, (3) stance expressions convey attitude and focus on the relationship between the writer and the reader, and (4) special conversational functions express, for instance, politeness towards the discourse participants. The assessment by Biber, Conrad, and Cortes (2004) recognises several subcategories, with referential bundles introducing the highest number of functional types, including quantity specification and time/place/text references. Discourse organisers contain two subcategories, which are topic introduction/focus and topic elaboration/clarification (see examples of bundles below). Stance expressions are divided into epistemic bundles and attitudinal/modal bundles. Finally, special conversational functions include politeness, simple inquiry, and reporting.

Later studies after Biber, Conrad, and Cortes (2004) have followed the categorisation of bundles into the three main functions of referential bundles, discourse organisers, and stance expressions, but the classification has been altered by additional subcategories to suit different genres. Hyland (2008), for instance, classifies bundles into research-oriented (e.g. location, quantification), text-oriented (e.g. transition signals and resultative signals), and participant-oriented bundles (stance features and engagement features). The main classes hence reflect the three categories in Biber, Conrad, and Cortes (2004), while the subcategories are somewhat different and also overlap with the earlier study. A number of previous investigations further combine grammatical and functional categories, e.g. Breeze (2013) establishes classes for content noun and prepositional phrases and non-content noun and prepositional phrases. Further, Kopaczyk (2013b) divides bundles in EMENT mainly on semantic grounds, but also includes overlapping classes for grammatical bundles such as copula verbs and prepositional phrase fragments.

My examination addresses only the functional categories since the attention is, for instance, on the text-external realities that are referred to in medical writing rather than on the analysis of grammatical patterns. The analysis classifies the bundles into the three above-mentioned functional classes: (1) referential bundles, (2) discourse or textual organisers, and (3) stance expressions. The investigation encompasses subcategories from Biber, Conrad, and Cortes (2004), Hyland (2008), and Kopaczyk (2013a, 2013b), and additional functional classes are added that better correspond to the bundles in medical writing. Hence, referential bundles are divided into fourteen different subcategories including, for instance, word units of quantity, actions, and references to the patient. Textual organisers include, among other things, the subclasses of narration, framing signals, and cohesion. Finally, stance expressions encompass two main classes, epistemic bundles and attitudinal/modal bundles following Biber, Conrad, and Cortes (2004). The full list of functions used in the analysis is shown in the following list, and the functions which were not found in the data are listed in squared brackets:

### **Referential bundles**

1. Date and time (*e.g. at the morning*)
2. Collective reference to the authorities (*e.g. college of physicians*)
3. Collective reference to other community members (*of the public*)
4. Reference to named individuals (*of the king*)
5. Reference to unnamed individuals (*a person who*)
6. (Collective) reference(s) to patient(s) and body parts (*of the head, pain in his, years of age*)
7. Action (*the return of, to keep the*)
8. Location (*of the house*)
9. Object (*in a letter*)
10. Abstract concepts (*the honour of, the good of*)
11. Ambiguous cases (*and in the*)
12. Quantity (*many of the, an ounce of*)
13. Identification/focus (*one of the*)
14. Imprecision (*and the like*)

### **Textual bundles/discourse organisers**

1. Cohesion (*of the same*)
2. Narration (*she was found, I have been*)
3. Topic elaboration/clarification (*I mean the, the other hand*)
4. Framing signals (*with regard to*)
5. Resultative signals (*as a result*)
6. [Intertextuality (*in Jurin's work*)]
7. [Topic introduction/focus (*in this chapter*)]

### **Stance expressions/interactional bundles**

1. Epistemic stance (*it appears that*)
2. Attitudinal/modality stance
  - a. Desire (*was desired to*)
  - b. Obligation/directive (*ought to be, it must be*)
  - c. Intention/prediction (*they will not*)
  - d. Ability (*it is impossible*)

The group of referential expressions was supplemented by two classes to suit the bundles in the medical texts, i.e. bundles referring to abstract concepts and references to patients and body parts. The bundles on patients include references to individual or more numerous patients, body parts, and symptoms of diseases; these concepts are grouped together since the tokens of individual bundles can refer to a specific patient or to symptoms of diseases on people in general (e.g. *pain in the*) (cf. Kopaczyk 2013b). The bundles concerning patients are distinguished from references to other groups of people such as authorities and named individuals. Further, abstract concepts were included into the functions, as the texts refer to various concepts which differ from concrete objects such as *the nature of, the power of, good of the*. The bundles on actions include both verbs and noun phrases following Kopaczyk (2013a), such as *the use of* and *to keep the*. Ambiguous bundles are word patterns that fit into two categories or include only grammatical words (*and in a* and *of it in*). These are rather numerous in the material, as the three-word bundles do not always offer enough content words for analysis.

Textual bundles or discourse organisers encompass seven subcategories but only five of these functions are found in the data, i.e. bundles on cohesion, narration, topic elaboration, framing signals, and resultative signals. Hence, bundles on intertextuality were found in the earlier studies on legal language (Kopaczyk 2013a; Lehto 2015) and bundles on topic introduction in academic genres (e.g. Biber, Conrad and Cortes 2004) but not in the corpus categories in LMEMT studied

here. There are, however, general references to earlier writings in the bundles (e.g. *act of parliament*) in my material, but these instances are analysed as objects since they do not contain citations from specific parts of texts or names of authors. Further, Kopaczyk's (2013a) framework, which is applied to legal texts, establishes a specific category for textual bundles denoting cohesion that overlaps with referential bundles; bundles such as *the same person* are distinguished from other bundles on cohesion since they refer both to people outside the text and form an anaphoric link to earlier expressions in the text. In my analysis, these instances are marginal and they are simply listed under cohesive bundles.

Stance expressions include four categories, divided into two main classes of epistemic bundles, and attitudinal or modal bundles. The categorisation follows Biber, Conrad, and Cortes's (2004) listing, since it reflects well the bundles in the material. Kopaczyk's (2013a) class of interactional bundles overlaps with this classification with five subcategories of directive, representative, declarative, commissive, and modal bundles. Finally, the category of special conversational functions recognised in Biber, Conrad, and Cortes (2004) is left out from the above list. These bundles do not occur in my material, although conversational functions could be possible findings in letters in the corpus, e.g. expressions of politeness with *thank you*.

### 3. Material and methods

The material of the analysis covers two corpus categories and encompasses one text type within the whole LMEMT corpus. Lexical bundles are surveyed in the categories of public health and the subcategory methods of specific treatises. The public health category offers 30 texts for analysis and texts on methods include 17 works, and consequently the category of public health is larger in its word count (see the category descriptions in this volume). Additionally, the case studies were collected across the corpus categories by corpus linguistic searches and qualitative reading, e.g. the search words *case* and *mr* were used to locate the case reports (see Chapter 5 in this volume). The total number of case reports used in the analysis is 262 and they originate from 66 separate works. The highest numbers for case studies were found in the subcategories of specific treatises, surgical and anatomical texts, and scientific periodicals. Case studies are very common in periodicals, and the analysis includes only a sample from this category; all case studies found from the other corpus categories were included in the data. The word count for the case studies is approximately 156,000 words, and a higher proportion of case studies comes from the latter part of the eighteenth century.

Bundles consisting of three words are located by their frequency and distribution in the material. In order to analyse the three-word bundles in the categories of public health and methods the minimum frequency was set to 6, which means that only bundles that appear in the material at least 6 times are recognised. In public health, this frequency corresponds to the normalised frequency of 34.1 per million words (or 0.3 per 10,000 words), and in methods the frequency rises to 43.4 (0.4 per 10,000 words), as the word count is lower in this category. The range was set to 3, indicating that the bundles need to occur at least in 3 texts within the corpus category. This dispersion covers 17.6% of the texts in methods and 10% in public health; the latter category contains a higher number of (short) texts, lowering the percentage. These thresholds were chosen because they correspond to the frequencies used in earlier studies, and further they retrieve a manageable number of bundles for analysis. In public health, the analysis retrieves 859 bundles and 595 in methods. Repeated wordings are hence most common in public health, reflecting the higher word count and the number of texts in this category (see the discussion in Section 7). For case studies, the frequency was again set to 6 instances, and this raw frequency corresponds to the normalised frequency of 38.5 per

million words (or 0.4 per 10,000 words). The dispersion was set to 6, corresponding to 9.1% of the case studies within one work. Many of the texts contain multiple case studies, but instead of studying each case study individually, the case reports from the same work were kept together, as they are usually linked to each other in the main text.

The lexical bundles were retrieved by using the n-gram function in the corpus tool AntConc 3.4.4 (Anthony 2014). The method can be affected by differences in spelling, but orthography is already consistent in the eighteenth-century medical texts when compared to earlier periods. Thus, variation in spelling does not considerably affect the bundles. The source texts in the LMEMT corpus regularly use the long *s* (*s*) (e.g. Jurin, *The Patients are sometimes taken with flushing Heats in An account of the success of inoculating the small pox in Great Britain*, 1724), but these are replaced by the short *s* in the corpus texts. Further, the corpus texts often use capital letters for nouns in the middle of sentences, but capital and lower-case letters are considered to be the same characters in the analysis (e.g. the words *Patients* and *Heats* in the above example). Additionally, hyphens that divide words at the end of lines were deleted prior to analysis, e.g. the word *Physi-cians* was changed to *Physicians*. After the corpus queries, the bundles were categorised according to their functions following Biber et al. (2004), Hyland (2008), and Kopaczyk (2013a, 2013b), but also new functional categories were established to suit the bundles in medical texts (see Section 2.3). Normalised frequencies were further calculated in order to compare the functions in each category.

#### 4. Lexical bundles in public health

The analysis of lexical bundles in the category of public health retrieves 859 bundle types with 9,016 tokens. The most frequent repeated lexical strings are *as well as*, *of the college*, and *ought to be* of which the second string refers to the College of Physicians. The fifteen most common bundles are listed in the following Table 2:

**Table 2.** Fifteen most common bundles in the category of public health

Rank	Raw freq.	Range	Bundle
1	98	22	as well as
2	69	8	of the college
3	69	11	ought to be
4	53	9	the knowledge of
5	53	15	the practice of
6	51	23	part of the
7	50	21	the number of
8	47	9	knowledge of the
9	46	6	the college of
10	43	14	of the most
11	41	16	of the same
12	37	18	it may be
13	36	6	college of physicians
14	36	17	of all the
15	36	15	one of the

The most frequent bundles refer, for instance, to the practising of medicine (*the practice of (physic)*) as well as to the College of Physicians (*of the college*, *the college of*, *college of physicians*) and the



knowledge of the practitioners (*the knowledge of, knowledge of the*). These examples already indicate that many of the most repeated bundles in public health do not concentrate on patients but on the state of medicine. Further, it can be seen that many bundles originate from the same word combinations such as the word strings mentioning the College of Physicians.

All of the lexical bundles were analysed according to their function. This assessment indicates that most of the bundles are referential bundles that relate to concepts outside the texts (about 78%), while textual bundles (14%) and stance expressions (8%) are noticeably less common. The distribution of the percentages and normalised frequencies per 100,000 words are shown in Table 3:

**Table 3.** Distribution of lexical bundles in public health

Function	Raw freq.	Freq./100,000	
		words	%
<b>Referential bundles</b>	672	381.7	78.2
<b>Textual bundles</b>	116	65.9	13.5
<b>Stance expressions</b>	71	40.3	8.3
<b>Totals</b>	859	487.9	100

Within the category of referential bundles, the most frequent functions are actions (i.e. 23% of the referential bundles), quantities (17%), and abstract concepts (15%) (Table 4):

**Table 4.** Distribution of referential bundles in public health

Referential bundles	Raw freq.	Freq./100,000 words
action	153	86.9
quantity	111	63.0
abstract concepts	102	57.9
ambiguous cases	75	42.6
collective reference to the authorities	60	34.1
location	44	25.0
collective reference to other community members	43	24.4
date and time	37	21.0
(collective) reference(s) to patient(s) and body parts	28	15.9
object	9	5.1
reference to named individuals	6	3.4
identification/focus	2	1.1
imprecision	2	1.1
reference to unnamed individuals	0	0
<b>Total</b>	<b>672</b>	<b>381.7</b>

The word strings on actions include bundles on the practising of medicine such as *practice of physic* and *applied to the* and more general references such as *the use of*. Additionally, many of the actions indicate a concern for matters of public welfare, e.g. *the relief of, the care of, the admission of*, as in *the care of the sick, the care of public prosperity, the care of the poor*, and *admission of patients* (to hospitals). Many of the bundles on actions are related to the hospital movement and philanthropy in the eighteenth century: the texts underline the benevolent spirit of the nation in helping the sick poor and aim at appealing to possible donators to collect money for hospitals and dispensaries (e.g. Wilson 1990: 10–.; see category description for public health). The second most recurrent

functional bundles refer to quantities; these bundles, however, include only a few numbers (*one or two*), as they mainly indicate nonspecific amounts, e.g. *parts of the, some of them, of a few*. Exact numbers are less common, since the texts do not focus on describing cures with specific amounts of medicines but encompass a wider social view on medical matters. Further, abstract concepts refer to the field of medicine (*art of physic, the medical art, the healing art*), advancing the life of the nation (*good of the*) and the measures and opinions related to these matters (*the power of, the means of, the interest of, the knowledge of*). The emphasis on abstract concepts is again on the wellbeing of citizens. Additionally, the authors often support their arguments by citing legislation: the bundles on objects contain word strings, such as *of the charter, act of parliament, and of the legislature*. Moreover, medical authors in the eighteenth century relied increasingly on statistical methods, and the bundle *bills of mortality*, for instance, emerges in the category, originating among others from works on demographics, such as from Black's *An arithmetical and medical analysis of the diseases and mortality of the human species* (1789). Ambiguous cases are additionally common, such as *and in the*, as the focus is on rather short bundles. Further, the texts on public health often allude to authorities, who are mainly physicians and other medical practitioners and their societies (e.g. *college of physicians, of an apothecary, of the learned*). The authors thus regularly discuss the role of the practitioners and medical societies with repeated wordings. Other community members include references to citizens and the poor, such as *to the poor, of the public, of the inhabitants, of the nation, and wives and children*. References to patients include general concepts such as *to the patient and to the sick*, but instances of specific body parts are not encountered. Similarly, instead of proper nouns (*city of London*), the bundles centre on more general locations, such as *of the hospital, in the country, in the city*, and also references to the kingdom are found (*in this kingdom*).

Textual bundles are rather infrequent in the category when compared to referential bundles, and they distribute across five subtypes. Textual bundles focus mainly on narration, cohesion, framing signals and topic elaboration/clarification, as can be seen in Table 5:

**Table 5.** Distribution of textual bundles in public health

Textual bundles	Raw freq.	Freq./100,000 words
narration	62	35.2
cohesion	21	11.9
framing signals	19	10.8
topic elaboration/clarification	13	7.4
resultative signals	1	0.6
Totals	116	65.9

Narrative bundles include *he is a, there is a, and I am well*. Many of the texts in the category are written in the form of letters addressed to institutions, and hence, references to the authors with the first-person singular are common in the bundles. Topic elaboration or clarification encompasses bundles, including *as well as* and *I mean the*. The anonymous author of *Animadversions on the constitution of Physick* (1768), for instance, questions the role of the College of Physicians in admitting licences to physicians and shifts to the topic of apothecaries with the bundle *I mean the, i.e. And do they not constantly disregard those, who, from a different mode of practice, take no fees, and consequently do not immediately injure them, I mean the Apothecaries?*. Although many of the texts take part in ongoing controversies and comment on earlier works, these links do not emerge through intertextual bundles, although there are more general bundles that refer to previous legislation or books. Further, cohesion (*of the said*) is rather often found in texts on public welfare, but it is not a prevalent writing strategy, unlike in legal texts (Lehto 2018).

Stance expressions convey mainly bundles on the topics of obligation/directive, intention/prediction and epistemic stance, and there are also expressions of ability and desire, as illustrated in Table 6:

**Table 6.** Distribution of stance expressions in public health

Stance expressions	Raw freq.	Freq./100,000 words
a. epistemic stance	15	8.5
b. attitudinal/modality stance		
1. desire (want)	2	1.1
2. obligation/directive	26	14.8
3. intention/prediction	17	9.7
4. ability	11	6.2
Totals	71	40.3

The writers, hence, often express obligation such as *ought to be* and *it must be*, as the authors argue how the issues of public welfare and the medical profession should be organised, e.g. Borthwick underlines cleanliness with *it must be in*:

- (1) it is easy to see how essentially necessary it must be to take every step, that can tend to prevent the direful consequences arising from putrid substances laying in, and about a Town. (Borthwick, *The method of preventing and removing the causes of infectious diseases*, 1784: 11)

In general, directives are especially frequent in this treatise, as Borthwick lists numbered instructions on how to prevent infections in institutions such as in XXI. *Every room in jail, poorhouse, &c. ought to be swept carefully every morning*. The authors further often express predictions and possible threats that will face the profession and citizens in the future, e.g. *it will be*, *it would be*, and *they will not*. Hales, a Doctor of Divinity, warns citizens of the dangers of excessive drinking and intimidates drinkers by citing the Bible:

- (2) yet Men do not obey his Voice, therefore the Rechabites will be their Judges; and it will be more tolerable in the Day Of Judgment, for Mahometans and Idolaters, than for sottish Christians. (*A friendly admonition to the drinkers of brandy*, 1733: 18)

Epistemic stance is additionally found among the bundles, since the authors acknowledge their limits of knowledge, e.g. *it may be* and *it appears that*. The lexical bundles hence reflect various trends in the category of public health that are referred to by formulaic wordings, i.e. the philanthropic spirit of the era emerges in the referential bundles along with statistical methods, legislation, and the role of the medical practitioners and patients. Additionally, the increasing interest in instructions on hygiene is revealed through repetitive stance expressions.

## 5. Lexical bundles in methods

In the category of methods, there are a total of 595 bundles and these add up to 6,515 tokens. The most common bundle in methods is *the small pox*, followed by *on the th<sup>1</sup>* and *in order to* of which the former refers to different dates, as the numbers have been omitted (e.g. *on the 5th*) (see Table 7):

**Table 7.** Fifteen most common bundles in the category of methods

Rank	Raw freq.	Range	Bundle
1	86	3	the small pox
2	74	5	on the th
3	53	13	in order to
4	47	8	by means of
5	47	16	part of the
6	43	10	out of the
7	40	11	the use of
8	38	14	of the same
9	37	11	and in the
10	37	13	in the same
11	35	13	as well as
12	34	9	at the same
13	34	10	of the body
14	34	4	the th of
15	33	8	of the blood

Many of the most frequent bundles are the same as are found in the category of public health, i.e. *as well as*, *part of the*, *of the same*. The most repeated lexical strings, nevertheless, indicate that bundles in methods often refer to sicknesses (*the small pox*), time and date (*the 5th of*) and parts of the body (*of the blood*), while the top bundles in public health encompass the authorities (*of the college*), their knowledge (*the knowledge of*), and expression of opinions (*ought to be*).

The distribution of the functions of bundles indicates that referential bundles (78%) prevail, followed by textual bundles (16%), and stance expressions (6%), as illustrated in Table 8:

**Table 8.** Distribution of lexical bundles in methods

Function	Raw freq.	Freq./100,000	
		words	%
<b>Referential bundles</b>			
<b>textual bundles</b>	464	335.8	78.0
<b>Stance expressions</b>	93	67.3	15.6
<b>Total</b>	38	27.5	6.4
	595	430.6	100

The normalised frequencies are higher in the category of public health than in methods; the higher word count in public health adds to the number of bundles, but it can also suggest that the same topics and expressions are more repeated in this category (see the discussion in Section 7). The main distribution of the bundles is, nevertheless, quite similar to those in the category of public

<sup>1</sup> This marks an expression of date.

health. However, textual bundles are proportionally slightly more repeated in methods (cf. 13.5% in public health), and, further, stance expressions are rather recurrent in public health (8%). The authors of the works on public health, hence, express epistemic stance and attitude more often.

Within referential bundles, references to the body and patients are found most often, with 84 occurrences, followed by actions and date and time (Table 9):

**Table 9.** Distribution of referential bundles in methods

Referential	Raw freq.	Freq./100,000 words
(collective) reference(s) to patient(s) and body parts	84	60.8
action	77	55.7
date and time	71	51.4
quantity	64	46.3
ambiguous cases	54	39.1
abstract concepts	52	37.6
object	33	23.9
location	15	10.9
collective reference to other community members	4	2.9
collective reference to the authorities	3	2.2
reference to unnamed individual	3	2.2
identification/focus	2	1.4
imprecision	2	1.4
reference to named individual	0	0
Total	464	335.8

The works on methods often discuss patients and body parts such as *of the patient, in the head, vessels of the*. These references are rare in texts on public health, which only occasionally refer collectively to patients. The bundles on actions encompass different medical cures and observations of diseases such as *the application of* and *exposed to the*. The bundle *the care of* is found both in methods and in public health but in the former category it denotes caring for individual patients (e.g. *So I left him to the Care of a Surgeon* (Catherwood, *A new method of curing the apoplexy*, 1715)), while in public health these bundles refer to the state taking care of its citizens. Further, quantities in public health do not usually contain exact numbers, while in Methods there are more instances of specific quantities of medicines, for instance, *a pint of (water), an ounce of, and a few drops*. References to date and time are noticeably more common in methods than in public health, e.g. *an hour and, the next day, in the evening, the eight day*. These bundles are recurrent because the authors observe the patients and the methods of cure, and report on their effect. Expressions of date and time are also frequent in case studies, which are common in methods. References to abstract concepts include examples such as *the truth of* and *the state of*, while concrete objects are encountered less frequently (*a glass of* and *in a letter*). Further, references to authorities and community members are rare in the methodological works when compared to the works on public health. The works on methods hence revolve around patients, and the texts do not regularly refer to the ancient authorities by recurrent word strings but cite individual authors. The bundle *in his book* is, for instance, found in Smith's *The curiosities of common water* (1723), as he refers to various authors who recommend water for the cure of smallpox:

- (3) I find that Dr. Cook of Warwick, in his Book of Observations on English Bodies, does prescribe, for the Cure of Fevers, first a Vomit, and afterwards as much cold Water as the Patient can drink. (Smith, *The curiosities of common water*, 1723: 27)

The authors of public health often support their arguments or discuss the state of the medical profession by referring to the College of Physicians and learned authors with medical education. Additionally, the bundles on objects reflect the importance of correspondence between the medical authors, as the bundle *in a letter*, for instance, is found among the bundles.

Textual bundles centre on narrative bundles but there are also instances of framing signals, topic elaboration/clarification and cohesion (Table 10):

**Table 10.** Distribution of textual bundles in methods

Textual bundles	Raw freq.	Freq./100,000 words
narration	44	31.8
framing signals	21	15.1
topic elaboration/clarification	16	11.6
cohesion	12	8.7
resultative signals	0	0
Total	93	67.3

Narration is common in the texts, as the authors narrate past events on how patients were cured. The first-person singular and third-person singular are often encountered in these bundles such as in *I have been*, *he was taken*, *he complained of*, and *she was found*. Many of these bundles originate from case studies and the narrative bundles further display the importance of experiments in medical writing, as the authors often narrate their observations and mental processes including *I have observed* and *I have seen*. The emphasis on observation and case studies is additionally visible within bundles on topic elaboration/clarification, i.e. the bundle *by the following* appears with the combinations *by the following experiments/observations/cases/method*, since the authors support their arguments by experiments. If compared to the distribution of textual bundles in public health, framing signals and topic elaboration/clarification are more important in methods, because these texts often limit the topic, for instance, by such expressions as *in this case* and *with respect to*; the frequencies of other textual bundles are higher in public health, and the bundles are used for different purposes.

The most common stance expressions are bundles that express epistemic stance and attitudinal bundles of intention/prediction, and they both appear 13 times. These functional types are followed by bundles on obligation/directives and ability, as indicated in Table 11:

**Table 11.** Distribution of stance expressions in methods

Stance expressions	Raw freq.	Freq./100,000 words
a. epistemic stance	13	9.4
b. attitudinal/modality stance		
1. desire	0	0
2. obligation/directive	7	5.1
3. intention/prediction	13	9.4
4. ability	5	3.6

Epistemic stance is realised by bundles such as *it may be* and *it appears that*, the authors expressing uncertainty about their observations. Bundles on intention and prediction include *it will be* and *I shall only*: the latter bundle is often used before case studies when the authors intend to offer some illustrative examples. The anonymous author of *A practical display of the philosophical system called animal magnetism* (1790) uses the bundle *I shall only* to explain two methods of cures: *Several Modes have been practised, but I shall only mention Two, which will be quite sufficient for the Purpose*. If compared to the category of public health, bundles

posing obligations or directives are less common in methods. The texts on public welfare often argue for changes on social matters by using directives. Some of these bundles in methods include *ought to be* and *it is necessary*; in methods, these bundles do not impose obligations on society or the field of medicine, but dictate how specific methods should be carried out. Jurin, for instance, gives advice on inoculation: *3. The utmost Caution ought to be used in the Choice of proper Matter to communicate the Infection (An account of the success of inoculating the small pox in Great Britain, 1724)*.

## 6. Lexical bundles in case studies

The analysis of lexical bundles in case studies retrieved 563 bundle types with 7,098 tokens. The overall normalised frequency of lexical bundles is about 361 instances per 100,000 words. The frequencies are lower than the overall frequencies in public health (488) and in methods (431), although these numbers are not fully comparable due to different thresholds for frequency and distribution. Similar to lexical bundles in public health and methods, referential bundles prevail in case studies (75%), followed by textual bundles (21%) and stance expressions (4%) (Table 12):

**Table 12.** Distribution of lexical bundles in case studies

Function	Raw freq.	Freq./100,000 words	%
<b>referential bundles</b>	420	269.3	74.6
<b>textual bundles</b>	119	76.3	21.1
<b>stance expressions</b>	24	15.4	4.3
<b>Total</b>	563	360.9	100

The normalised frequency of textual bundles (76 or 21%) in case studies is notable when compared to the other analysed categories: approximately 73 instances (16%) in methods and 66 (14%) in public health. This difference originates from the high number of narrative bundles in the case studies, as will be discussed below.

Many of the top bundles refer to the patient (*years of age, cavity of the, the cavity of, of the abdomen*), who is the other main character in the case study in addition to the physician. Additionally, expressions of time are also frequent (*on the th, the th of, a few days*). The fifteen most common bundles in the case reports are shown in Table 13 below:

**Table 13.** Fifteen most common bundles in case studies

Rank	Raw freq.	Range	Word string
1	99	31	part of the
2	72	28	years of age
3	70	17	on the th
4	64	27	the use of
5	61	18	the th of
6	45	16	out of the
7	44	27	in the same
8	43	10	cavity of the
9	42	9	the cavity of
10	38	24	as well as
11	36	18	a few days
12	36	8	of the abdomen
13	35	22	in order to
14	35	22	the end of
15	34	21	one of the

Some of these word strings are also found among the top bundles in methods, including *part of the*, *out of the*, *the use of*, *in the same*, *on the th*, and *the th of*, which refer to dates. The bundles *part of the*, *one of the*, and *as well as* are further shared with the top bundles in public health, displaying repetition of these wordings across the corpus categories.

The distribution of referential bundles indicates that the case studies focus on the patient and body parts as well as on quantities and date and time (Table 14):

**Table 14.** Distribution of referential bundles in case studies

Referential bundles	Raw freq.	Freq./100,000 words
(collective) reference(s) to patient(s) and body parts	114	73.1
quantity	92	59.0
date and time	76	48.7
action	66	42.3
ambiguous cases	35	22.3
abstract concepts	27	17.3
object	3	1.9
reference to named individual	3	1.9
collective reference to the authorities	2	1.3
collective reference to other community members	0	0
identification/focus	0	0
imprecision	0	0
location	2	1.3
reference to unnamed individual	0	0
Total	420	269.3

References to patients are regular in the narratives and encompass bundles on body parts, symptoms, and descriptions of the patient: *the right side*, *neck of the*, and *years of age*, *a gentleman of*, and *pain in her*. The description of the patient's age and symptoms are usually found at the



beginning of the case studies, as the case reports follow the structure of natural narratives (see Chapter 5 in this volume). Bundles on quantity such as *two or three* and *about an inch* are frequent, the case studies reporting amounts of medicines and occurrences of symptoms. The bundle *three or four* often describes the number of times that medicines have been taken a day. However, in *Power of nature and art*, it is used to quantify numerous short case studies to make calculations on the effects of medicines, e.g. *Three also it vomited and purged, in one moderately, in the others, it vomited three or four times, and purged several ...* (*An essay on the power of nature and art, in curing diseases*, 1753: 48, in general treatises). Further, expressions of time give details on the onset of symptoms and treatment, such as *a few days*, *the next day*, *in the morning*, and *half an hour*. In his *An inquiry into the causes and effects of the variolæ vaccinae* (1798), Jenner describes the effects of inoculation chronologically, and the progress of the disease is narrated for each day such as in:

- (4) On the second day the incisions were inflamed and there was a pale inflammatory stain around them. On the third day these appearances were still increasing and their arms itched considerably. (Jenner, *An inquiry into the causes and effects of the variolæ vaccinae*, 1798: 43)

Further, descriptions of actions are rather frequent, e.g. *the use of* and *covered with a*; these elements again relate to the use of medicines and taking care of a patient or refer to the symptoms. The number of abstract concepts is lower in the case studies than in methods or public health, as the authors focus on narrating a single case of disease of a patient.

Textual bundles are common in the case studies because narrative bundles are the second most common subtype after references to patients. Other textual bundles are rather rare, but there are instances of cohesion, topic elaboration and framing signals (Table 15):

**Table 15.** Distribution of textual bundles in case studies

Textual bundles	Freq./100,000	
	Raw freq.	words
narration	95	60.9
cohesion	15	9.6
topic elaboration/clarification	5	3.2
framing signals	4	2.6
resultative signals	0	0
Total	119	76.3

The narrative bundles include fragments, such as *there was a*, *she had been*, *he began to*, *he complained of*, *I was called*, *I found the*, and *I gave her*. There are, hence, many bundles that portray the physician visiting and examining the patient. The bundles encompass both third-person singular and first-person singular forms, since the authors report the progress of the patient and the treatment performed by the physician. Further, correspondence emerges in the bundles: the bundle *I send you* is found in case studies in scientific periodicals, since the authors send illustrative cases for the journals. Narration is also common in methods, although the normalised frequency (31.8) is lower than in the case studies; in public health, narrative bundles add up to the frequency of 35.2. The texts on methods and case studies share a similar function: the authors explain their measures by narrative linguistic features. In public health, the narrative bundles critically discuss the actions of medical practitioners or refer to citizens.

Stance expressions convey mainly ability, and also epistemic stance and intention/prediction are expressed a few times in case reports, as can be seen in Table 16:

**Table 16.** Distribution of stance expressions in case studies

Stance expressions	Raw freq.	Freq./100,000 words
a. epistemic stance	6	3.8
b. attitudinal/modality stance		
1. desire	1	0.6
2. obligation/directive	1	0.6
3. intention/prediction	5	3.2
4. ability	11	7.1
Total	24	15.4

The expressions of ability regularly emerge to describe the limitations of patients' or physicians' actions, e.g. *could not be*, *he could not*, *I could not*, and *be able to*. The condition of the patient is reported in Colbatch's *A dissertation concerning misletoe* (1719, in therapeutic substances) as *he could not walk without difficulty*, and the physician explains his own standing in Catherwood's *A new method of curing the apoplexy* (1715, in methods) as *I could not promise much to raise her Expectations*. In general, stance expressions have a low frequency in case studies: the attention is on narration and on reporting the patient's progress, and even the expressions of ability, for instance, are part of the broader narrative structure.

## 7. Discussion and conclusion

The analysis has illustrated that many similar lexical bundles are found especially among the top bundles in public health, methods and case studies. The distribution of the main functional categories is rather similar in these datasets, as most of the recurrent word strings are referential bundles, followed by textual bundles and stance expressions. A more detailed analysis of the functions has, nevertheless, shown that the functions differ in various respects. The texts within a category contain different topics, portraying a discrete view of medicine, and patients and the authors additionally exhibit dissimilar attitudes towards external reality. In public health, the three most frequent functions are part of referential bundles, including bundles on actions, quantities, and abstract concepts. Similarly, in methods referential bundles prevail and the three most common functional types are references to patients and body, actions and quantities. In case studies, the references to patients and body are the most common, followed by narration, which is part of textual bundles, and the third most typical word strings are referential bundles of quantities. The prevailing functions point to the most important differences in the studied categories. In public health, the authors deal with medicine on the social level, arguing about medical authorities and the welfare of citizens. The texts on methods concentrate on healing patients' diseases and refer to quantities of medicines. In case studies, the attention is also on the patient and quantities of medicines, and these matters are discussed by means of narrative fragments.

The category of public health provided the highest number of lexical bundles for analysis. The bundles include many hypernyms when compared to the other studied categories: the bundles on quantities, for instance, do not contain exact numbers but more general references to amounts. The

texts focus on social issues, i.e. the bundles include, for instance, references to legislation, since the authors promote changes to regulations and appeal to the authorities by letters. In general, a number of the texts on public welfare are written in the form of a letter and consequently the category includes many narrative bundles that contain the first-person singular form. Bundles are additionally found when the authors state their opinions on public matters, while in methods and case studies narrative bundles are used to explain the patients' condition and cure. Further, the role of the College of Physicians is very visible in public health, and these bundles appear among the top word strings. The authors view the College of Physicians as an important association in regulating the profession and many texts also criticise the College as having too much control over licences and practitioners. Additionally, many hospitals were founded in the eighteenth century, and the hospital movement surfaces in the bundles on public welfare: the bundle *the admission of (patients)* among others is found especially in works that discuss the role of infirmaries. Along with the hospital movement, the bundles include many references to philanthropy, the authors often appealing to prospective donators. The importance of hygiene further emerges in stance expressions that are labelled as directives, i.e. various directives are used to give instructions on hygiene in hospitals and prisons, for instance. Generally, many bundles refer to the poor and to the wellbeing of citizens, i.e. word strings on actions, abstract concepts and references to society members include various bundles on aiding poor patients. Public welfare is seen as affecting the nation on a more general level (e.g. the bundle *the care of* in *the care of public prosperity*), although the idea of public health on the national level did not fully emerge until the nineteenth century. The role of the patient and the rights of citizens as patients in general are recurrent topics in this category. In addition, the bundles point to statistical methods, as there are, for instance, references to Bills of Mortality. Stance expressions further reflect the authors' attitude, and bundles on obligation are common; the authors state their arguments through directives and additionally they often make predictions on the future state of medicine in society.

In methods, the attention is mainly on the patient and healing procedures instead of social matters. Hence, referential bundles typically refer to patients and body parts, while bundles on authorities or other society members are infrequent. Additionally, referential bundles on actions suggest that the authors describe methods of cures, and further expressions of time and date and quantities are important when the texts discuss the progression of diseases and responses to medical treatment. Consequently, within textual bundles, narrative bundles are often used to narrate the progression of diseases and cures. The category of methods includes many case studies; these case studies surface through narrative bundles but also through bundles on topic elaboration/clarification, since these wordings often shift the attention to specific medical cases or methods of cure. In addition, the recurrent word strings indicate the importance of observation in methods. Observation emerges in the bundles on topic elaboration/clarification that are not only used to introduce illustrative cases but also observations. Similarly, narrative bundles that employ the first-person singular are used to reflect on observations and the mental processes of the author. The proportion of stance expressions is generally lower than in public health, and these bundles typically introduce observations (intention/prediction) or express epistemic stance, as opposed to directives that are common in public health.

The bundles in case studies illustrate that the reports clearly focus on narrating the cure of the patient. References to the patient and body are numerous as well as discussion of quantities and time, as the authors narrate diachronically the steps of cure and restate the amounts of medicines given to patients. This viewpoint is very different from texts in public health that concern the rights of patients and often refer to abstract concepts and authorities. In addition, narrative bundles are proportionally more numerous in case studies, reflecting the narrative structure of the reports. Stance expressions are not a common feature in case reports, but bundles on ability are found in the

texts. These bundles refer, for instance, to the bodily limitations that patients experience due to their illnesses or physicians stating their inability to aid the patient. Many case studies, nevertheless, seem to describe successful cures as the narrative bundle *he was perfectly (recovered/well)* is found among the recurrent word strings. The idea of observation and quantification are seen in case studies as well, as multiple cases are chained to strengthen the argument: the bundle *a great many* points to the numerous cases that the author has encountered. Some authors further refer to the case studies as “observations”, illustrating the importance of experiments in medical writing, and additionally the case studies often alternate with experiments (see Chapter 5 in this volume).

The bundles indicate that the medical texts are regularly based on prefabricated wordings. The highest number of bundles was found in the category of public health, followed by methods and case studies, i.e. the normalised frequency of bundles is 488 in public health, 431 in methods and 361 in case studies. The percentage of words that are part of the bundles is highest in public health, i.e. 15.4% of wordings in public health are covered by three-word bundles, while in methods, they constitute 14.1%, and 13.7% in case studies. This distribution can be linked to the word count in each dataset, as public health contains the highest number of words, followed by methods and case studies. In contrast, case studies encompass the highest numbers of texts, as there are 66 separate texts, while public health includes 30 texts and 17 come from methods. The number of bundles seems to reflect the word count in each category, and it is difficult to draw conclusions whether some of the categories are more formulaic than others. Further, the frequency threshold and dispersion differ in each category, making comparisons likewise problematic. The number of bundles can, nevertheless, be affected by the topics of the medical texts: works in public health concern rather general matters on society and they do not refer to single patients or dosages of medicines, while case studies name specific patients and discuss the progression of diseases on a detailed level, lowering the number of similar bundles in the reports.

The analysis of the bundles indicates that new trends in medical writing are found in LMEMT. Eighteenth-century medicine started to rely on statistical methods and quantification, and these developments can be encountered in the bundles, i.e. the texts in public health refer to Bills of Mortality to discuss diseases and numbers of deaths. Statistical information is also taken advantage of in methods and case studies, in which the authors, for instance, calculate effects of cures in various case studies that are used to make general conclusions. In addition, observation is a recurrent theme in the studied texts in methods and case studies. In methods, the bundle *an account of* among other repeated wordings refers to experiments and observations that are narrated to support the arguments. In case studies, the bundle *according to the (observations)* points to other case studies that are based on observation. When compared to medical texts in MEMT and EMEMT, developments can further be noticed in references to medical authorities. Early modern medical writing relied extensively on scholastic evidentiality, referring to ancient authorities, such as Galen, although a change towards empiricism was already taking place (see Taavitsainen 2009). The ancient authorities do not, however, surface in the bundles in LMEMT, since only references to contemporary authors are made, as shown by contextual analysis of the bundle *in his book*. Thus, instead of relying on ancient authorities, medical writing in the eighteenth century is increasingly based on observation and experiments. In addition, the humoral theory, which was still popular in the Early Modern English period, is not visible in the eighteenth-century texts. Bundles such as *hot and dry* and *cold and moist* are common in EMEMT across the corpus categories (Kopaczyk 2013b: 276), but these word strings are not encountered in my study in LMEMT. However, a comparison of repeated wordings in EMEMT and LMEMT suggests that many functions of the bundles and even exact word strings can be traced back to Early Modern medical writing. In EMEMT, most of the bundles that are found throughout the corpus timeline are references to quantification, measurement, and dosages (Kopaczyk 2013b: 268). Other common and persistent bundles in

EMEMT are references to the body, time and sequences, ingredients, quality descriptions, and references to diseases and cures (ibid.: 271–). The functions of quantification and references to the body, for instance, are frequent in LMEMT as well, and the bundles *part of the* and *of the body* among others are frequent in both corpora.

My analysis has concentrated on the functions of lexical bundles, and future studies could further examine the grammatical structures of bundles in LMEMT. It is worth noting here that *of*-constructions (including prepositional constructions and noun phrases with *of*-phrase fragment and post-modifiers) are common in the analysed datasets and especially in public health, i.e. 39% of the bundles in public health are *of*-constructions, while in methods they add up to 35%, and to 32% in case studies. These numbers are in line with broader changes in medical writing, since medical writing from the sixteenth to the twentieth century evolved towards a literary style (including less situation-dependent and non-involved features) with nouns and prepositional phrases, for instance (Biber 1995: 283–; Tyrkkö and Hiltunen 2009). However, *of*-constructions are lowest in case studies, as narrative constructions, including verb phrase fragments, are common in these texts. The lexical bundle method could hence reveal further developments in the grammatical structures of medical writing in different categories of LMEMT. All in all, the analysis of the functions of bundles has revealed various continuities and novel practices in eighteenth-century medical writing. The medical texts employ many specific wordings typical of the genre but the texts in public health and methods and case studies also portray a distinct view of medicine in the eighteenth century.

## References

- Altenberg, Bengt. 1998. On the phraseology of spoken English: The evidence of current word-combinations. In Cowie, A. P. (ed.), *Phraseology: Theory, Analysis and Applications*. Oxford: Oxford University Press. 100–122.
- Anthony, Laurence. 2014. AntConc 3.4.4. Corpus software. Tokyo, Japan: Waseda University. See <<http://www.laurenceanthony.net/>>
- Biber, Douglas. 2006. *University Language: A Corpus-based Study of Spoken and Written Registers*. Amsterdam/Philadelphia: John Benjamins.
- Biber, Douglas. 1995. *Dimensions of Register Variation*. Cambridge: Cambridge University Press.
- Biber, Douglas, Susan Conrad and Viviana Cortes. 2004. If you look at: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25.3. 371–405.
- Biber, Douglas, Stig Johanson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Breeze, Ruth. 2013. Lexical bundles across four legal genres. *International Journal of Corpus Linguistics*, 18.2. 229–253.
- Cortes, Viviana. 2004. Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23.4. 397–423.
- Culpeper, Jonathan and Merja Kytö. 2010. *Early Modern English Dialogues: Spoken Interaction as Writing*. Cambridge: Cambridge University Press.
- Ellis, Nick C. 1996. Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*, 18.1. 91–126.
- Gries, Stephan Th. 2008. Phraseology and linguistic theory: A brief survey. In Sylvaine Granger and Fanny Meunier (eds.), *Phraseology. An Interdisciplinary Perspective*. Amsterdam: Benjamins. 3–25.
- Hyland, Ken. 2008. *As can be seen*: Lexical bundles and disciplinary variation. In *English for Specific Purposes*, 27. 4–21.

- Kopaczyk, Joanna. 2013a. *The Legal Language of Scottish Burghs: Standardization and Lexical Bundles 1380–1560*. Oxford: Oxford University Press.
- Kopaczyk, Joanna. 2013b. Formulaic discourse across Early Modern English medical genres: Investigating shared lexical bundles. In Jucker, Andreas H., Daniela Landert, Annina Seiler and Nicole Studer-Joho (eds.), *Meaning in the History of English: Words and Texts in Context*. (Studies in Language Companion Series, 148.) Amsterdam: John Benjamins. 257–299.
- Lehto, Anu. 2018. Lexical bundles in early modern and Present-day English Acts of Parliament. In Kopaczyk, Joanna and Jukka Tyrkkö (eds.), *Applications of Pattern-driven Methods in Corpus Linguistics*. John Benjamins. 159–186.
- Lehto, Anu. 2015. *The Genre of Early Modern English Statutes: Complexity in Historical Legal Language*. Helsinki: Soci  t   N  ophilologique de Helsinki.
- Meyer, Charles F. 2015. Corpus-based and corpus-driven approaches to linguistic analysis: One and the same? In Taavitsainen, Irma, Merja Kyt  , Claudia Claridge and Jeremy Smith (eds.), *Developments in English: Expanding Electronic Evidence*. Cambridge: Cambridge University Press. 14–28.
- Moon, Rosamund. 1998. *Fixed Expressions and Idioms in English: A Corpus-Based Approach*. Oxford: Clarendon Press.
- Sinclair, John. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Taavitsainen, Irma. 2011. Medical case reports and scientific thought-styles. *Revista de Lenguas para Fines Espec  ficos*. (Special issue edited by Francisco Alonso Almeida.) 75–98.
- Taavitsainen, Irma. 2009. The pragmatics of knowledge and meaning: Corpus linguistic approaches to changing thought-styles in early modern medical discourse. In Jucker, Andreas H., Daniel Schreier, and Marianne Hundt (eds.), *Corpora: Pragmatics and Discourse Papers from the 29th International Conference on English Language Research on Computerized Corpora (ICAME 29)*. Amsterdam: Rodopi. 37–62.
- Taavitsainen, Irma. 2001. Changing conventions of writing: The dynamics of genres, text types, and text traditions. *European Journal of English Studies*, 5.2. 139–150.
- Taavitsainen, Irma, Turo Hiltunen, Anu Lehto, Ville Marttila, P  ivi Pahta, Maura Ratia, Carla Suhr and Jukka Tykk  . 2014. Late Modern English Medical Texts 1700–1800: A corpus for analysing eighteenth-century medical English. *ICAME Journal*, 38.1. 137–153.
- Taavitsainen, Irma and P  ivi Pahta. 2000. Conventions of professional writing: The medical case report in a historical perspective. *Journal of English Linguistics*, 28.1. 60–76.
- Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work*. Amsterdam/Philadelphia: John Benjamins.
- Tr  hler, Ulrich. 2005. Quantifying experience and beating biases: A new culture in eighteenth-century British clinical medicine. In Jorland, Gerard, Annick Opinel and George Weisz (eds.), *Body Counts: Medical Quantification in Historical and Sociological Perspective*. Montreal and Kingston: McGill–Queen’s University Press. 19–50.
- Tyrkk  , Jukka and Turo Hiltunen. 2009. Frequency of nominalization in Early Modern English medical writing. In Jucker, Andreas H., Daniel Schreier, and Marianne Hundt (eds.), *Corpora: Pragmatics and Discourse Papers from the 29th International Conference on English Language Research on Computerized Corpora (ICAME 29)*. Amsterdam: Rodopi. 293–316.
- Wilson, Adrian. 1990. The politics of medical improvement in Early Hanoverian London. In Cunningham, Andre and Roger French (eds.), *The Medical Enlightenment of the Eighteenth Century*. Cambridge: Cambridge University Press. 4–39.
- Wray, Alison. 2002. *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.