Åbo Akademi University

Jonas Tana

Infodemiology

Studying rhythmicity in online health information behaviour

Jonas Tana (born 1979)

Degrees:

Bachelor in Emergency Care / Registered Nurse, Arcada University of Applied Sciences, Finland, 2004

Master of Arts, University of Helsinki, 2012

Infodemiology

Studying rhythmicity in online health information behaviour

Jonas Tana

Information Studies
Faculty of Social Sciences, Business and Economics Åbo Akademi University
Åbo, Finland, 2019

ISBN 978-952-12-3876-5 (printed)
ISBN 978-952-12-3877-2 (digital)
Painosalama Oy, Turku, Finland 2019

Acknowledgements

I would like to thank my supervisors, professor Gunilla Widén and university teacher Kristina Eriksson-Backa for their support and encouragement during this process. I would also like to thank my reviewers, professors Peter Bath and Reijo Savolainen, for their insightful and valuable comments. Thank you Emil, Kristina, Jyrki, and Heikki for your contributions and support as co-authors. I stand in gratitude to Fonden för teknisk undervisning & forskning for financing my research. I also want to thank all my dear colleagues at Arcada University of Applied Sciences for their support as well as everyone at the Department of Information Studies at Åbo Akademi University for their comments and ideas during the research seminars.

To Beata, Amos and Elin. Thank you for putting things in perspective.

Helsingfors, 1.10.2019

_____

List of original publications

I. Tana, J. (2018). An infodemiological study using search engine query data to explore the temporal variations of depression in Finland. *Finnish Journal of eHealth and eWelfare, 10*(1), 133-142. https://doi.org/10.23996/fjhw.60778

II. Tana, J. C., Kettunen, J., Eirola, E. and Paakkonen, H. (2018). Diurnal variations of depression-related health information seeking: case study in Finland using Google Trends data. *JMIR Mental Health, 5*(2), e43. https://doi.org/10.2196/mental.9152

III. Tana, J., Eirola, E. and Eriksson-Backa, K. Exploring temporal variations of depression-related health information behaviour in a discussion forum - case Suomi24. Accepted to *Information Research*.

IV. Tana, J., Eirola, E. and Eriksson-Backa, K. (2019). Rhythmicity of health information behaviour: utilizing the infodemiology approach to study temporal patterns and variations. *Aslib Journal of Information Management*, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/AJIM-01-2019-0029

V. Tana, J., Eirola, E. and Eriksson-Backa, K. (2019). The aspect of time in online health information behaviour: a longitudinal extensive analysis of the Suomi24 discussion forum. *Informaatiotutkimus, 38*(2). https://doi.org/10.23978/inf.83315

All articles are reprinted with permission.

Author contributions

Study I has a sole contribution by the author. In Study II, the author conceived the study, drafted the manuscript and was in charge of the overall research design, planning and management of the study. Data and statistical analyses and computations were conducted by Kettunen and Eirola, and Paakkonen functioned as an advisor. All authors contributed to the final version of the manuscript. In Study III-V, the author was in charge of the general work, conceived the original idea, planned the data gathering and analysis, and wrote the manuscript with support from Eriksson-Backa. Eriksson-Backa also had a supervising and advising role. Eirola conducted the data-analysis, as especially in studies III-V, methods and knowledge from big data analytics were needed to be able to handle large data sets. All authors discussed and commented on the results and helped to shape the final manuscript.

Abstract

The aim of this thesis is to examine the rhythmicity and temporal variations of online health information behaviour. Time and temporal aspects have previously been neglected in information science, as well as in research on health information behaviour. Yet, these aspects are of great significance, as health is a dynamic process bound by rhythmicity and temporal structures. The health related issues and threats, or knowledge gaps, that individuals face, might rise, change and intensify at any time of the day, week, month or year, and have been suggested to trigger health information behaviour.

Health information behaviour is defined as how people seek, obtain, evaluate, categorise and use health-related information in relation to their health. This behaviour is rather extensive, and today individuals are able to make sense of their situation and engage in health information behaviour online, anywhere and at any time. People frequently try to find out about and treat symptoms, diagnose themselves, or manage illness before or even without the involvement of health professionals. This means that much of this behaviour fall outside the scope of traditional healthcare systems, and leaves a substantial amount of relevant data about different health behaviours unexplored. Today however, a majority of this health information behaviour happens online, leaving vast amounts of digital traces behind. Aggregating and analysing these digital traces to study online health behaviours has been dubbed infodemiology, and is rooted in the idea that there is a relationship between population health on one hand, and information and communication patterns on the internet, on the other.

This thesis aims to answer if the infodemiology approach and metrics can be utilised to study rhythmicity and temporal aspects of online health information behaviour in relation to different health issues. Answering these questions not only allows us to gain insights into a new contextual aspect of health information behaviour, it also allows us to understand the multidimensional relationship between time and health. Individuals and their health status are different at different times, and health related behaviours follow these rhythmical and cyclical patterns and variations, which can have practical implications for many disciplines. This is especially relevant in an era where individuals are expected to take greater responsibility for their health with the help of an abundance of digital means. Turning the focus to time and studying when health information behaviour happens can have far-reaching consequences.

Abstrakt

Syftet med denna avhandling är att undersöka rytmiska och temporala variationer av hälsoinformationsbeteende på nätet. Tid och temporala aspekter har i tidigare forskning inom såväl hälsoinformationsbeteende som informationsvetenskap försummats. Ändå är dessa aspekter betydande, speciellt för hälsoinformationsbeteendet eftersom hälsa i sig är en dynamisk process, bunden av rytmiska och temporära strukturer. Det betyder att de hälsorelaterade utmaningar som individer möter och som har föreslagits leda till hälsoinformationsbeteende, kan uppenbara sig, förändras eller intensifieras när som helst, under dygnet, veckan, månaden eller året.

Hälsoinformationsbeteende definieras som hur människor söker, införskaffar, utvärderar, kategoriserar och använder hälsorelaterad information i relation till sin hälsa. Hälsoinformationsbeteendet är idag omfattande, eftersom individerna har möjlighet att identifiera hälsorelaterade problem och engagera sig i hälsoinformationsbeteende på nätet, oberoende av tid och plats. Individerna försöker ofta med hjälp av detta beteende utreda och behandla symtom, diagnostisera sig själva eller sköta sjukdomar eller utmaningar med hälsan innan besök hos, eller till och med helt utan, hälso- och sjukvårdspersonal. Detta innebär att majoriteten av hälsoinformationsbeteendet faller utanför räckvidden för den traditionella sjukvården och lämnar således en betydande mängd relevant information om olika hälsobeteenden outforskade. Hälsoinformationsbeteende på nätet efterlämnar stora mängder digitala spår. Sammanställningen och analysen av dessa digitala spår för att studera hälsobeteenden kallas infodemiologi, som utgår från att det finns en nära relation mellan informations- och kommunikationsmönster på internet å ena sidan och folkhälsa å andra sidan.

Målet med denna avhandling är att svara på om infodemiologi som angreppssätt och infodemiologisk data kan användas för att studera rytmicitet och temporala aspekter av hälsoinformationsbeteende på nätet. Att svara på dessa frågor tillåter oss inte enbart att få insikter i en ny kontextuell aspekt av hälsoinformationsbeteende, utan även i det flerdimensionella förhållandet mellan hälsa och tid och. Individernas hälsa varierar vid olika tidpunkter, och hälsorelaterade beteenden följer dessa rytmiska och cykliska mönster och variationer, vilket kan ha praktiska konsekvenser för flera olika discipliner. Detta är särskilt relevant i en tid där individerna förväntas ta ett större ansvar för sin egen hälsa med hjälp av ett större utbud av digitala tjänster och medel. Att fokusera på tid och studera när man ägnar sig åt hälsoinformationsbeteende kan därför få långtgående konsekvenser

Table of contents

List of Figures

Figure 21. Daily distribution over a year of all messages (green), messages in the health (fi. Terveys) category (blue) and messages in the depression (fi. Masennus) subcategory (orange). Filled diamonds denote cases where the difference of the relative share from the overall average is statistically significant. For clarity, this is only shown for the depression category and when it is significant.

List of tables

1. Introduction

Every disease occurs at all seasons of the year but some of them more frequently occur and are of greater severity at certain times. –Hippocrates (Aphorisms, III/19, Lloyd, 1983, p. 215)

Health, well-being and illness are all dynamic processes that change over time. This places individuals on a continuous and cyclical spectrum ranging from health, well-being and optimal function on one end to illness and death, at the other (Eriksson, 1993, pp. 54-56, 76-77; Henly et al., 2011; Svalastog et al., 2017). This cyclical nature of health is also affected by nature itself, as time and the temporal structures and rhythms of days, weeks, months, and seasons affect nature and all living things, from single cells to human beings. Night follows day, tides ebb and flow and seasons change. Nature's rhythmicity and the temporal variations and patterns that relate to it govern most human behaviours from sleeping, breathing, eating and thinking to issues of health and illness (Adam, 1990, p. 72; Mogilner et al., 2018; Reinberg et al., 2017; Roenneberg, 2012, pp. 1-2).

Health information behaviour, or how people seek, obtain, evaluate, categorise and use health-related information in relation to their health and health threats, is closely associated with the discontinuous and cyclical nature of health and illness (Ek, 2013). Within the health and illness continuum, the health issues and threats that individuals face have been suggested to trigger health information behaviour, which is essential in regards to coping with illness and maintaining proper health and health behaviours (Ek, 2013; Lambert and Loiselle, 2007). These issues and threats that individuals face, and that individuals either face or avoid (Wilson, 2000), might rise and change at any time during the day, week, month or year.

Modern healthcare is reliant on the individual in management and self-care of health and illness related issues. This self-care encompasses how individuals are encouraged to manage parts of their health, including monitoring symptoms and conditions as well as managing treatments (Erikainen et al., 2019; Petrakaki et al., 2018). More and more people also take bigger responsibility in maintaining their health by different health related behaviours, and laypersons become experts of their own health projects, keen to find out more about their health, share their experiences and expertise. This empowerment is highly associated with the

responsibility and accountability of individuals in relation to their health order to reduce unnecessary health expenses (Petrakaki et al., 2018).

This behaviour is rather extensive, as already mentioned, and a majority of people even try to treat or diagnose themselves before seeking medical advice from professionals or without being directly managed by others such as doctors or nurses (Erikainen et al., 2019; McGowan, 2002; Oliphant, 2010, p. 1; Petrakaki et al., 2018; Siepmann, 2008, p. 68). Research also suggests that people prefer to interact and ask questions online, since members of online communities might be more knowledgeable about the condition, or dealt with or survived a similar experience (Brady et al., 2016; Gustafson et al., 1999; Halder, Poddar and Kan, 2017; Naslund et al., 2016; Walther and Boyd, 2002, p. 153). This consists of everything from finding out about and treating symptoms to peer-support in managing illness, producing a complex information context of digital traces, rich with temporal data. Especially for people living with mental health issues, like depression, the Internet provides a unique platform for health information behaviour. Mental health issues are often associated with stigma and barriers to care, why individuals suffering from mental health issues have been shown to be more likely to seek information about their problems online, making online health information behaviour in relation to mental health issues rather extensive (Ayers et al., 2013; Ayers et al., 2014a; Chan et al., 2016; Gustafson et al., 1999; Hasler, 2013, pp. 612-613 ; Hausner et al., 2008; Oliphant, 2010, p. 2; Powell and Clarke, 2006). Studying, analysing and explaining temporal variations for various health behaviours and illness responses both at an individual and population level are at the core of medical, public health, nursing and social science research (Henly et al., 2011). Yet, knowledge about many of these health related issues and threats often fall outside the scope of traditional healthcare systems, as between seventy to ninety per cent of all health related problems and issues are solved or managed without the involvement of healthcare professionals (Fox and Ward, 2006; McGowan, 2002; Oliphant, 2010, p. 2; Siepmann, 2008). This leaves a substantial amount of relevant data about different health behaviours unexplored and unknown, as these kind of preclinical health behaviours, events and aspects, including health information behaviour, are not documented within the scope of professional healthcare, in patient records or other registers. Many of these behaviours also fall outside the scope of traditional epidemiological data collection methods, such as annual surveys, polls or focus groups (Eysenbach, 2009). This leaves, because of methodological reasons, much of the temporal variations and rhythms,

apart from some annual or seasonal ones, outside the scope of data collection and therefore only provides snapshots of different health behaviours and their rhythmicity (Anker et al., 2011; Ayers et al., 2014). Yet, these health behaviours relate to all aspects of everyday life, from work and social relations to lifestyle choices and education (Svalastog et al., 2017). Moreover, health status varies and changes across an individual's lifespan and is influenced by a number of factors. These include lifestyle choices, structural factors like access to healthcare, as well as natural and contextual factors, like time, temporal structures and rhythms as well as changes that relate to these.

Today, a majority of this health information behaviour happens online, and internet use as a source for information related to health or illness in Finland is, and has, for quite some time already been ubiquitous (Ek, Eriksson-Backa and Niemelä, 2013; Tilastokeskus, 2018). Because of the digital developments during the last decades, online health information is now accessible anywhere and at any time. Laptops, tablets and smartphones allow people to act on health related information needs whenever and wherever the needs arise (Ren et al., 2019). This behaviour leaves vast amounts of digital traces behind, especially in search engines, social media and on websites (Hausner et al., 2008; Kim and Oh, 2011; Lee et al., 2016). This justifies the study of different aspects of online health information behaviour, including rhythmicity and temporal aspects.

Studying and systematically analysing this rising amount of health related digital traces, or web data, from a temporal perspective can reveal evidence of behavioural rhythmicity, and provide insights to temporal patterns, trends and variations of online health information behaviour (Ayers et al., 2014; Eysenbach 2011). This type of research or approach has been dubbed infodemiology, or information epidemiology, a novel multidisciplinary approach defined as *the science of distribution and determinants of information in an electronic medium, specifically the Internet, or in a population with the ultimate aim to inform public health and public policy* (Eysenbach 2011, p. 155). Infodemiology is rooted in the idea that there is a relationship between population health on one hand, and information and communication patterns on the internet, on the other (Eysenbach, 2009).

Within information science in general, the aspect of rhythmicity or the context of time has not received much attention, in theory building or practice, as focus has favoured and emphasised questions like *why*, *where* and *how* (Adam, 1990, pp. 10-13; McKenzie and Davies, 2002; Solomon, 1997). Overall, research on rhythmicity, or the temporal aspects of

information behaviour, are scarce, even if several information behaviour researchers have highlighted the importance of studying the aspect of time (Dervin, 1992, p. 66-71; McKenzie and Davies, 2002; Savolainen, 2006a; Savolainen, 2018; Solomon, 1997). Turning the focus towards time and the question of *when* can provide evidence for both rhythmicity of health information behaviour as well as the relationship between time and health. This kind of research can both complement epidemiological data, as well as provide new insights for preclinical health related behaviours and their rhythmicity. This in turn can help in painting a more holistic picture of health, well-being the temporal aspect of behaviours related to these. Studying these rhythms and temporal variations has previously been difficult, and knowledge about them can therefore provide new insights into health and health information behaviour.

1.1. Purpose and aim

Time and temporal aspects have previously been neglected in information science, as well as in health information behaviour research. Yet, these aspects are of great significance, as health is a dynamic process bound by rhythmicity and temporal structures. The aim of this thesis is therefore to examine the rhythmicity and temporal variations of online health information behaviour. This is done both in relation to health issues in general, as well as specifically for depression, a significant health issue in Finland and globally. The objective is to find out if rhythmicity and temporal variations can be observed on different time-scales, from daily (circadian) and weekly (circaseptan) to monthly (circalunar) and seasonal (circannual). Circadian rhythms are rhythms of observed activity with a period length close to 24 hours, associated with the rotation of the earth around its axis. Circaseptan rhythms describe the cyclic seven-day phenomena, whereas circalunar or circa-monthly rhythms are associated by the complete lunar orbit around earth every thirty days. Circannual, or seasonal rhythms, again, are entrained by seasonal and annual changes caused by the rotation of earth around its sun (Dunlap et al., 2004, p. 107; Reinberg et al., 2017). The fundamental idea of rhythmicity merely conveys regular predictable occurrence of an event, and all non-random events are, by definition, rhythmic phenomena (Dunlap et al., 2004, p. 15). In this thesis, the rhythmical nature of health information behaviour is studied by analysing web data from search engines and social media. Finland, with its northern location and intense variations in both seasonal temperature and daylight makes an interesting subject to study health related phenomena from a

temporal aspect, as variations in external conditions are particularly noticeable (Basnet et al., 2016; Dunlap et al., 2004, p. 107). The purpose is therefore to turn attention to the aspect of time, and study the temporal nature and rhythmicity, or *whens'* of health information behaviour. More specifically, the thesis aims to answer if the infodemiology approach and metrics can be utilised to study rhythmicity of health information behaviour in relation to different health issues, from more general to more specific. This overreaching aim is answered by four more specific research questions:

1. What are the strengths and weaknesses of utilising the infodemiology approach to study the context of time in online health information behaviour? (Study IV)
2. Does online health information behaviour in general follow rhythmicity? (Study IV and V)
3. Does online health information seeking for depression related health information show annual, seasonal, weekly or daily rhythmicity? (Study I and II)
4. Does depression related online health information behaviour within a large discussion forum show temporal variations on different time-scales? (Study III)

Answering these questions not only allows us to gain insights into a new contextual aspect of health information behaviour, it also allows us to understand the multidimensional relationship between time and health. Individuals and their health status are different at different times, and health related behaviours follow these rhythmical and cyclical patterns and variations. Moreover, through its rhythmicity, life becomes predictable, as events and interactions occur and reoccur on foreseeable times. Focus on time can thus allow us to see the invisible (Adam, 1990, pp. 74-75; Roy et al., 2004).

1.2. Thesis structure

The thesis is structured in the following way. Chapter 1 – Introduction introduces the reader to the thesis subject and research area. It then presents the aims and objectives as well as the research questions of the thesis. Chapter 2 and 3 provides the reader with the necessary background as well as theoretical foundations needed to understand what is studied and why. Chapter 4 introduces Infodemiology, its aims, approaches, method and metrics, as well as the limitations of the approach. Methods and data-

analysis of the different studies included in this thesis are described in chapter 5. Chapter 6 presents the results of studies I-IV in detail. The results, implications, limitations and further directions are discussed in chapter 7. A concluding chapter (8) then follows the discussion.

2. Background

The following literature review is structured to give the reader a relevant background to the thesis topic by first providing a short review of health. This is followed by a chapter on time, rhythmicity and health, information behaviour and online health information behaviour. The aim of this chapter is also to help spot the gap or neglected aspect of a topic, or identify approaches on a topic that are deficient, in this case in online health information behaviour and its rhythmicity and temporal variations. Identifying a research gap as well as relevant previous and recent literature and research on temporal aspects of online health information behaviour, as well as infodemiology, has been done thoroughly by utilising several different databases, such as EBSCO, ScienceDirect and PubMed. Keyword such as information behavio(u)r, health information behavio(u)r, infodemiology, time, temporality, temporal aspects and chronobiology have been used in different combinations. Searches have also been complemented with queries utilising the same keywords in Google Scholar. Moreover, the research gap within time and online health information behaviour has been identified with the help of previous research.

2.1. Health

Health is the fundamental and desirable relative condition from which individuals can deviate and to which they can return (Emson, 1987; Svalastog et al., 2017). Amzat and Razum (2014, pp. 21-23) state, that the concept of health has been subject to various conceptualisations and interpretations, and that disagreements about the meaning and definition are common because of the many different components, from political and medical to social and economic, that it entails. The World Health Organization (WHO) defines health as *a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity.* (Emson, 1987). As popular and as central to the debate on the meaning of health as this definition is, it has received some criticism for being too vague, inflexible, unrealistic, of limited practical use and even composed of undefinable terms (Amzat and Razum, 2014, pp. 22-23; Emson, 1987). Moreover, as Amzat and Razum (2014, p. 23) point out, the inclusion of the word "complete" in the definition makes health both unachievable and unrealistic for any individual for a reasonable period of time. The reason for this is that it is

difficult or even impossible to gain complete contentment in physical, mental, and social well-being. Also Card (2017) sees the definition as problematic, as health is interpreted as dichotomous, making individuals either healthy or unhealthy. Instead of this dichotomy, health is situated on a continuum in time between more healthy and less healthy.

However, more recent concepts and definitions of health recognise health as more than the absence of disease, with an enhanced emphasis on the capacity and adaptability of the individual (Svalastog et al., 2017). The ability and capacity of the individual to adapt and self-manage in the face of social, physical, and emotional challenges is in today's digital society an important factor (Erikainen et al., 2019; Huber et al., 2011; Petrakaki et al., 2018). As Amzat and Razum (2014, pp. 24-25) point out, the adaptive capacity that this definition introduces makes the definition not only more realistic, but also takes into account the shift in health challenges in the twenty first century, with an increasing burden of chronic diseases. From a self-care and self-management perspective, it also emphasises the enhanced empowerment, responsibility, and possibilities, of the individual in the digital era (Erikainen et al., 2019; Petrakaki et al., 2018).

There is also a distinction between perspectives on health and illness, and disease and absence of disease, as defined by the theory of salutogenesis and pathogenesis (Svalastog et al., 2017). Salutogenesis studies what creates health and what factors support health, as opposed to the conventional, or biomedical approach of pathogenesis that studies the factors that cause disease (Svalastog et al., 2017). Health and illness are dynamic processes and each person is located on a graduated scale or continuous spectrum ranging from wellness and optimal functioning in every aspect of one's life, at one end, to illness culminating in death, at the other (Antonovsky, 1996; Svalastog et al., 2017). The salutogenesis model views individuals at a given point in time, on a continuum, and is, according to Antonovsky (1996), a more powerful and more accurate conception of reality. On the other end of the health spectrum lies illness, which presence is often communicated by complaint (Jennings, 1986). The health and illness as well as the disease and absence of disease spectrum differ from each other, as the health and illness spectrum of experiences is continuous.

The previous focus on absence of disease describes the focus and emphasis on biomedicine, as most early definitions of health have tended to focus on physiological and biological dysfunction and negative aspects that need to be overcome to return to normal functioning (Johnson and Case, 2012, pp. 6-7). A part from this, investigation of health and illness lies

outside the domain of biomedicine, because its study ultimately depends directly on phenomenological analysis of experienced, or conscious, suffering through subjective, individual self-reports and behaviours (Amzat and Razum, 2014, p. 24; Jennings, 1986). As an example of this difference, an individual can have a serious disease without being ill, as in the case of silent hypertension or a brain aneurysm. Of this follows, that an individual can also be seriously ill without having a disease. Thus, health and well-being are, as Blaxter (2010, pp. 19-20) and Card (2017) state, highly relative and subjective, conceptual concepts, as individuals assess, judge and define whether they are healthy or not differently. According to Blaxter (1990, p. 40), the *most usual way of measuring self-perceived illness, as distinct from the presence or absence of disease, is by means of symptoms lists.* Therefore, as Jennings (1986) argues, pain, suffering and distress are dimensions of illness, not of disease, and the absence of symptoms means health (Amzat and Razum, 2014, p. 25). This has also traditionally led to a distinction between physicians and patients, where physicians are more interested in the underlying disease, while patients are more concerned with their illness or suffering (Jennings, 1986). However, biomedicine enables us to empirically separate illnesses into two mutually exclusive categories: those arising from disease or injury, or medical illnesses, and those arising from other personal difficulties in living, or nonmedical/existential illnesses. The separation of these is in practice executed by pathological diagnosis of any underlying disease (Jennings, 1986). Illness is in close relation to consciousness, as both kinds of illnesses, both medical and nonmedical, vanish or can be modified if consciousness vanishes or is modified, for instance with medication or therapy (Jennings, 1986).

Health and illness are aspects that relate to all other aspects of everyday life, from work and social relations to lifestyle choices and education (Svalastog et al., 2017). In a more practical way, the enactment of health becomes concrete when considering what people do to maintain their health (Amzat and Razum, 2014, p. 25). The focus should however not only lie on the individual, but also on groups, communities, and the interaction between individuals, as advocated by a public health approach (Eysenbach, 2011; Svalastog et al., 2017). The different health concepts, particularly concepts we use to explain, treat, heal or cope with illness and disease, Svalastog et al. (2017) argue, are words we also use as search tools.

The experience of health and illness is dynamic and changing, and is best described as a movement between the two ends of the spectrum (Henly et al., 2011; Jennings, 1986). From this follows that health varies across an

individual's lifespan, and the experiences of illness or disease have been described as interruptions, intrusions or gaps in the normal flow of life (Charmaz, 1991, p. 13; Rier, 2010). Aspects of health and illness are influenced by a number of factors, including lifestyle, structural factors as access to healthcare, and contextual factors, like time or the natural rhythms and temporal structures, such as the shift between night and day, or the changing of seasons (Dunlap et al., 2004, p. 325-326; Reinberg et al., 2017; Roenneberg, 2012, pp.1-3; Svalastog et al., 2017).

2.1.1. Depression

One specific health related issue, which is characterised by the above-mentioned dynamic changes, recurrent symptoms and variations in illness severity in relation to time, is depression. Depression is defined by the World Health Organisation (Marcus et al., 2012, p. 6) as follows:

Depression is a common mental disorder, characterized by sadness, loss of interest or pleasure, feelings of guilt or low self-worth, disturbed sleep or appetite, feelings of tiredness and poor concentration. Moreover, depression often comes with symptoms of anxiety. These problems can become chronic or recurrent and lead to substantial impairments in an individual's ability to take care of his or her everyday responsibilities.

Depression is the leading cause of disability worldwide, and a considerable health issue in Finland as well, where at least five per cent of the adult population is estimated to suffer from depression every year (Patana, 2014). Depression is both a symptom and an illness with a broad spectrum from chronic to episodic depression. There is an amass of research showing that rhythmicity and temporal aspects influence individuals' tendency to suffer from not only depression, but psychological health issues in general (Antypa et al., 2016). Depression is also characterised by cyclic occurrence, and if long lasting and severe enough, it can turn into a serious health condition (Maes et al., 1993; WHO, 2018). There is also a significant risk of suicide associated with depression, as more than half of all suicide victims have suffered from depression (Brådvik, 2002; Isometsä, 2014). Even if depression can be treated effectively, only less than half of those suffering from depression receive treatment. Reasons for this is both the stigma and the reluctance to seek treatment in fear of becoming labelled as a mental patient (Aromaa et al., 2011). This complexity and stigma is also suggested to be the reason why a considerable percentage of people either interested in or suffering from depression, and mental health in general, are more likely to seek and share

answers to information needs online (Ayers et al., 2014; Chan et al., 2016; Gustafson et al., 1999; Hausner et al., 2008; Powell and Clarke, 2006).

2.2. Health from the perspective of time and rhythmicity

Time and rhythmicity are important concepts in health, as there is a multidimensional relationship between time and health (Roy et al., 2004). However, shifting the focus from static to dynamic theorising about health and illness is challenging because time and temporal concepts in information science as well as nursing science are novel (Henly et al., 2011). Time itself is a fascinating concept that has been described as a paradoxical riddle, difficult to define (Dunlap et al., 2004, p. 3-14). The definitions of time in the Oxford (2008) dictionary are descriptive of the complex nature of time. Of all seven definitions with their sub senses, three fall into the category of temporal aspects and rhythmicity: (1) *the indefinite continued progress of existence and events in the past present and future, regarded as a whole*, (2) *a point in time as measured in hours and minutes past midnight or noon or a period characterized by particular events or circumstances*, and (3) *an instance of something happening or being done*. Merriam-Webster's dictionary also provides similar definitions, as it defines time partly as: (1) *the measured or measurable period during which an action, process, or condition exists or continues*; (2) *the grouping of the beats of music* and (3) *a moment, hour, day, or year as indicated by a clock or a calendar* (Merriam-Webster, 2019). Human awareness of time, or the daily, monthly, and seasonal cycles and rhythms date back to prehistoric civilisations. Rhythmic cycles of different human behaviour, as well as human awareness of these rhythms and cycles, have for long existed in human culture, as marked by astrological and calendrical tables, stone henges and pyramids, as well as documented medical texts. Early human cultures, from the Egyptians to the Greeks, were intrigued by how different symptoms and illnesses displayed daily and seasonal variations, and how health was a cyclic and rhythmic phenomenon (Dunlap et al., 2004, pp. 3-14). For instance, Hippocrates argued that seasons had a profound effect on the mind and body, resulting in different types of predominant diseases during the winter and summer (Lloyd, 1983). Ever since, the importance of biological timing, clocks and rhythms have been recognised in human health and wellbeing, and since the mid-20th century, these rhythms are key characters in the field of chronobiology (Adam, 1990, p. 72; Dunlap et al., 2004, pp. 11-14; Roenneberg, 2012, p. 5).

The biology of all life forms is organised by temporal structures, which are defined by endogenous rhythms of several period domains,

corresponding to various environmental cycles. The source for nature's rhythmicity can be found in the movement of earth and its moon in relation to the sun. The sun, with its radiant energy, thus influences the cyclical nature of life, and governs all living beings, which regulate their cycles of activity and sleep in reference to it (Adam, 1990, p. 72; Dunlap et al, 2004, p. 325-326; Roenneberg, 2012, p. 2). Therefore, rhythmicity can be seen as a universal phenomenon, and has traditionally been categorised into circadian rhythms, circa-monthly and seasonal or circannual ones. The circadian rhythms are associated with the 24h light cycle associated with the rotation of the earth around its axis. The circa-monthly rhythm is associated with the rotation of the moon around the earth every 30 (29.53) days. The seasonal rhythm again, is entrained by the seasonal and annual changes in the daily photoperiod caused by rotation of the earth around its sun (Dunlap et al., 2004, pp. 107-110; Reinberg et al., 2017). Moreover, a fourth cyclical time structure or rhythmicity, with a less clear origin and biological value, is the seven day, or circaseptan, phenomenon. Researchers have questioned if this circaseptan rhythm is representative of a genetic attribute common to all living species and persistent in constant environmental conditions or, if it instead is socially constructed (Ayers et al., 2014; Reinberg et al., 2017). However, as Reinberg et al. (2017) point out, a wide variety of these circaseptan rhythms are reported in human beings residing in the usual environment of today, and biological oscillations have been demonstrated in various different physiological functions (for a thorough review of these, see Reinberg et al., 2017). Therefore, the human 7-day rhythms seems to be, at least to some extent, endogenous in origin, and can therefore affect behaviours as well as health and wellbeing. Therefore, all these temporal structures are important to study.

These time structures are relevant for all living things on Earth, even if the natural clues play a much less important role for modern citizens than in the lifestyle of previous rural cultures or of earlier civilisations. In the past, humans must have responded to these cycles when it came to hunting, to seeking safety during hours of darkness, and to choosing times for irrigating land and planting food. Today, a majority of humans live in a highly managed, round-the-clock environment, driven by the clock more than by any other factor. This 24-hour demand affects many human behaviours, and expectations of services is present around the clock (Dunlap et al., 2004, p. 291; Roenneberg, 2012, p.). Factors like work schedules, classes, and school are also major factors in rhythmicity. Weekdays are dominated by early alarm clock jolts, a dash to school or the office, and a weary return home in

the evening for recuperation and sleep (Dunlap et al., 2004, p. 291). Although weekends are more relaxed, scheduling is still highly important. Moreover, the social environment is a major determinant of behavioural rhythmicity, affecting many behaviours and their rhythmicity and cycles, for instance the sleep-wake cycles that can be even asynchronous with the external light-dark cycles (Dunlap et al., 2004, pp. 291-292).

This rhythmical character becomes observable when focusing on human behaviours and experiences such as eating, sleeping, breathing, thinking, communicating or interacting, as all our physiological processes are temporally organised (Adam, 1990, pp. 72-76). Many behavioural, physiological and psychological rhythms in humans are in synchrony with daily light-dark cycles and seasonal changes in these (Dunlap et al., 2004, pp. 325-326; Mogilner et al., 2018; Reinberg et al., 2017; Roenneberg, 2012, p. 2). Many, if not all, bodily functions have been shown to exhibit a circadian rhythmicity with notable precision and stability in healthy individuals, and some aspects of human daily rhythms probably result from behavioural changes associated with the light-dark and sleep-activity cycles (Dunlap et al., 2004, pp. 325-328; Mogilner et al., 2018; Reinberg et al., 2017; Roenneberg, 2012). Many, if not all, bodily functions from body temperature, heart rate, and blood pressure to hormone regulation have been shown to exhibit circadian rhythmicity with notable precision and stability in healthy individuals (Dunlap et al., 2004, pp. 341-349). Some aspects of human daily rhythms probably result from behavioural changes associated with the light-dark and sleep-activity cycles that peak either during the day or during night (Dunlap et al., 2004, pp. 341-349; Mogilner et al., 2018; Reinberg et al., 2017; Roenneberg, 2012). Some illnesses and health issues also show considerable daily variations, either on onset or in symptom severity. For instance, cerebral strokes and cardiovascular diseases occur more often between early morning and noon, and the incidence of asthma attacks is greatest at night (Dunlap et al., 2004, pp. 341-342). People with rheumatoid arthritis tend to experience their worst symptoms in waking in the morning. Circadian pathology also affects circadian mood disorders, or so called affective disorders (Dunlap et al. 2004, pp. 341-342). Circadian, or diurnal variations of depression and symptoms of depression is a characteristic of depressive mood, and have been associated with early-morning worsening, afternoon slumps or evening worsening (Wirz-Justice, 2008).

On a seasonal level, previous research has found associations between seasonal variations and various illnesses and health issues. Studies have found that more than ninety per cent of the general population show

seasonal variations in relation to mood, social activity, sleep, appetite, weight gain or energy level (Basnet et al., 2016). More specifically, mood and psychiatric disorders, like depression, anxiety, bulimia nervosa, anger, hostility and irritability, panic disorder, obsessive-compulsive disorder and substance abuse have been shown to follow seasonal variations, with peaks in incidence and severity varying from autumn to mid-winter to early spring (Basnet et al., 2016; Christensen and Dowrick, 1983; Oyane et al., 2008; Stordal et al., 2008). However, both circannual, or seasonal, as well as circadian, or diurnal, variations have in previous research been shown to be rather conflicting, which makes both the presence and direction of these variations unclear (Hasler, 2013, pp. 612-613; Powell and Clarke, 2006; Wirz-Justice, 2008). One suggested reason for the inconclusive results are the methodological limitations in studying and monitoring depression and depression symptoms and their temporal variations in populations (Ayers et al., 2013; Harmatz et al., 2000).

Seasonal changes also affect common non-communicable diseases such as diabetes, cardiovascular diseases, cancer and rheumatic diseases (Basnet et al., 2016). In addition, metabolic parameters like weight, cholesterol level, blood pressure, and blood glucose have shown seasonal variation. Moreover, morbidity and mortality of non-communicable diseases have repeatedly been demonstrated to be higher during winter than during summer, with a sixteen per cent higher mortality during winter than during summer, in European countries (Basnet et al., 2016).

Not only do these rhythms create a communal present, their existence also facilitates the prediction or anticipation of the future through recurring cycles (Adam, 1990, pp. 74-75). There are however, variances in the intensity and quality of these recurring patterns that differentiate them from each other or from other times. The timing and severity of several health disorders and issues are as, can be seen, strongly modulated by circadian and seasonal rhythms. As a result of the modern 24-hour day economy and lifestyle, individuals can demonstrate a behavioural or physiological dichotomy between day and night, where deviations from this regular diurnal rhythm of day and night have been suggested to cause most circadian malfunctions (Dunlap et al., 2004, pp. 325-329). In contrast to the abundance of information on circadian regulation in humans, considerably less is known about longer rhythmicity, such as circannual or seasonal rhythms (Dunlap et al. 2004, pp. 107-112). Human chronobiology issues and behaviour relating to natural rhythms have received growing attention in the past few decades in disciplines such as psychology, physiology,

neurosciences, and medicine. As the direct applicability of these chronobiologic principles of rhythmicity to human health and wellbeing, and behaviour related to them, becomes evermore evident, research interest in related fields grow (Dunlap et al. 2004, pp.13-14).

2.3. Health information behaviour

Wilson (1999, p. 249) defines information behaviour as *activities a person may engage in when identifying his or her own needs for information, searching for such information in any way, and using or transferring that information.* The definition encompasses information seeking behaviour, which Wilson (2000, p. 49) sees as *the purposive seeking of information as a consequence of a need to satisfy some goal*. As Wilson states, this behaviour involves interaction with either manual information systems, or computer-based systems, such as the Internet (Wilson, 2000). Health information behaviour in this thesis is thus defined as all human information behaviour in relation to health. This encompasses health information needs, seeking and use, which also includes information sharing (Wilson, 2000). Health information behaviour can also be approached as how people seek, obtain, evaluate, categorise and use health-related information in relation to their health and health threats (Ek, 2013). Wilson's definition on information behaviour characterises human behaviour in relation to dealing with information. A definition of the concept of information again is somewhat more troublesome, and the concept has been defined differently and the meaning of the word is quite proliferated (Case, 2012, pp. 48-49). Finding a single, widely accepted, universal definition or attribute for the concept of information seems not only difficult, but in this case somewhat unnecessary, as it falls outside the scope of this thesis (a thorough review of the definition and concept of information can be found in Case, 2012, pp. 45-75). In this thesis, information is treated as *a difference that makes a difference to a conscious, human mind*, as defined by Bateson (in Case, 2012, p. 46). Although Bateson's definition is broad, it does, as Case (2012, p. 46-47) points out, rule out the option of information existing independently of a knowing (human) mind. The definition is also useful for health information behaviour, as it is part of human information behaviour, where information either reduces uncertainty or changes one's image of reality (Case, 2012, p. 46-71).

Information behaviour is an important part of being human, and something that everyone engages in on a regular basis, either consciously or unconsciously. Yet for some people and in some situations, the incentives are considerably higher (Case, 2012, pp. 77-91). This is especially evident in

health related information behaviour. Motives for engaging in information behaviour are many, for instance a knowledge gap or information need, signalled by anxiety and or a need to act (Dervin, 1992, pp. 70-73). Information needs are the cause, or underlying motivation for information seeking. According to Case (2012, p. 78), information needs are typically considered inner motivational conditions or states, that give birth to thought and usually lead to action with the goal to alleviate anxiety, make decisions or understand something. These inner states can vary from wanting, believing and doubting, to fearing or expecting (Case, 2012, p. 78). Dervin (1992) again, sees information needs as a compulsion to make sense of a current situation, or knowledge gap, with the use of information. In the context of health, it can be argued that information needs fall under the category of basic human needs, in contrast to many other information needs, which can be classified as secondary (Case, 2012, pp. 78-79). Health related problems and issues have for instance been shown to cause fear and anxiety from which information needs arise (Lambert and Loiselle, 2007). Wilson (1997) states that defining information needs is somewhat unnecessary, as most studies, instead of studying information needs which are difficult to observe or are even described as a black-boxed factors assumed to trigger and drive information seeking, actually study information seeking behaviours (Savolainen, 2012). Yet, one of the contexts in which information needs are conceptualised in information seeking studies is the situation of action. This encompasses the temporal and spatial constituents that can be seen as significant, as they make it understandable how information needs may change within and between situations (Savolainen, 2012). However, as Savolainen (2012) states, these temporal and spatial constituents become meaningful only if they are anchored in concrete actions, such as information behaviours.

Information seeking has, as Case (2012, pp. 89-90) notes, received far less attention definitionwise than information needs. This is probably because the meaning is seen as obvious, and should not be overanalysed. Most of the definitions of information seeking typically describe it as an intentional behavioural action or a process in reaction to an information need (Case, 2012, pp. 90-91). Therefore, the above cited definition by Wilson (1999, in Case, 2012, p. 90) of information behaviour is more suitable, as it encompasses passive behaviours as well as information use and sharing.

Information use, according to Wilson (2000) consists of the involved acts, both physical and mental, for incorporating the found information into the knowledge base and constructing knowledge as well as altering the

knowledge structure of the person involved in these acts. Information use is, according to Fidel (2012, p. 35), a highly complex process which when studied presents challenges. However, the challenges in studying information use are not that relevant, as actual information use is not in focus within this thesis.

Information sharing, exchange, receiving, or giving is fundamentally a social activity, with an underlying assumption that people come together in different, varying, ways to share information with the goal of making sense, accomplishing a task or solving a problem in relation to common interests (Pilerot, 2012). Information sharing is a collaborative and interactive process that entails the behaviour of seeking advice or expertise from others, and sharing already acquired information with others (Talja and Hansen, 2006; Veinot, 2009). Health and illness is a significant motivation for this kind of collaborative information behaviour, or information sharing, as people with issues related to health and illnesses tend first to seek help or information from people like themselves (Neal and McKenzie, 2006). This kind of information sharing is common in various online support groups and resources, such as social media and discussion forums (Almanea, Bath and Sbaffi, 2018; Neal and McKenzie, 2006).

Information sharing can take place in different settings or contexts, and sharing usually involves some common interest, like a shared political engagement or shared experiences of an illness or health threat. In the context of health, information sharing includes providing advice to individuals with for instance similar issues and experiences and sharing questions and concerns around treatment options (Veinot, 2009; De Choudhury et al., 2014).

2.3.1. Studying online health information behaviour

As the aim of this thesis is to study rhythmicity of online health information behaviour, the emphasis of this chapter lies on how health information behaviour as well as the aspect of time have been studied in previous research, especially in an online context. Describing health information behaviour from a more qualitative and offline aspect therefore falls outside the scope of this thesis. Moreover, as Case (2012, p. 341) states, the literature on health information behaviour has grown too vast to review comprehensively. Consequently, the focus in this chapter lies on online health information behaviour.

The internet is, as mentioned, important as a source for health related information in Finland. A nation-wide survey conducted in 2009 showed

that already then, nearly 70% of 18-65-year-old Finns had visited some health-related site during the past 12 months. The most used sources were health portals, visited by 45% (Ek, Eriksson-Backa and Niemelä, 2013). Recent statistics show that in 2018, 65 per cent of the Finnish population in the ages 16-89 years had sought for health information online. There are, however, variations between age groups. As many as 79 per cent of those aged 25-34 years, and 81 per cent in the age category 35-44, engaged online health information. The use decreases in the older demographic, as only 49 per cent of those aged 65-74, and 24 per cent of people in the age group 75-89 had sought health information online during the last 3 months (Ek and Niemelä, 2010; Tilastokeskus, 2018). Online health information also shows differences in other demographics, as women, students, people with higher education and people living in urban settings show higher percentages for online health information seeking in Finland (Tilastokeskus, 2018). Similar figures are found elsewhere, for example in the UK 69% searched health information online in 2013 (Dutton and Blank, 2013). In the US, one in three adults use the internet to either diagnose or learn more about a health related topic or issue (Fox and Duggan, 2013). However, the role of the internet as a source for health information shows global variation, as an older survey conducted in Hong Kong showed that only 44 per cent had looked for health information on the internet (Yan, 2010). In China, a mere 26 per cent of the entire population used the Internet to access health information in 2016 (Ren et al., 2019).

A key characteristic in studying health information behaviour of the general public is employing a cross-sectional study design. Methodologically, this means that this behaviour is usually studied and documented by either interviews or surveys (Anker et al., 2011; Ramsey et al., 2017). Online health information behaviour in different settings, such as social networking sites or social media, is no different, which highlights the strong methodological traditions of studying health information behaviour (Anker et al., 2011; Kim and Syn, 2014; Zhao and Zhang, 2017). This type of research design provides what Giddens (1979, p. 202) identifies as a "snapshot" of health information behaviour, characterising the behaviour of a dynamic issue only at a specific point in time (Anker et al., 2011). An alternative approach has also been to employ a retrospective study design, which relies on, and compares, datasets and cohort designs collected at different points of time (Anker et al., 2011). Apart from these, health information behaviour researchers have also utilised naturalistic interventions or observations that take place in a natural situation or

environment, characterising information seeking behaviours of a specific group or demographic, for instance patients with specific ailments or illnesses, or within a specific age group (Anker et al., 2011; Kim and Syn, 2014; Zhao and Zhang, 2017). Among the least applied methods for studying health information behaviour, Anker et al. (2011) identified longitudinal studies.

As with research in general, all employed research methods to study health information behaviour also have their strengths as well as suffer from limitations. From a temporal perspective, the use of either surveys or interviews proposes a significant limitation, as it only provides insights into a behaviour at a single point in time, in a non-naturalistic setting (Anker et al., 2011). In relation to this, expenses associated to time and labour limit how often data can be collected (Huang, 2017). Another limitation is that survey-based data gathering methods also suffer from recall and information bias. Notwithstanding, health information behaviour studies would greatly benefit from utilising a longitudinal study design, to examine health information behaviour for an extended period of time, which could reveal the rhythmicity and temporal nature of health information behaviour (Anker et al., 2011). Longitudinal studies of health information behaviour could potentially provide a much broader and more extensive picture of this behaviour, especially in relation to time and the temporal aspects, which have previously been neglected within research on this area, and information science in general. Some of these limitations could be dealt with, especially in light of recent developments in innovative research approaches and methods.

Time, of the natural sort, has found significant application in information retrieval, for instance in developing methods for processing explicit temporal expressions in search queries and from document content, as well as in storing documents for long-term preservation and access (Gorichanaz, 2016).

Within research on online health information behaviour, the factors affecting use of online resources have often been in focus. These factors include socio-demographic variables, such as gender, education level and age (Bidmon and Terlutter, 2015; Ek and Niemelä, 2010; Flynn et al., 2006; Hallyburton and Evarts, 2014; Renahy et al., 2010; Tilastokeskus, 2017; Torrent-Sellens et al., 2016; Yan, 2010; Ybarra and Suman, 2008). Health status as a variable affecting online health information behaviour has also been studied frequently (Flynn et al., 2006; Renahy et al., 2010; Torrent-Sellens et al., 2016). Moreover, trust as well as eHealth literacy skills as

factors that can facilitate or hinder engagement in online health information behaviour have also been studied by Sbaffi and Rowley, (2017) and Lee et al. (2014). Pendry and Salvatore (2015) again examined outcomes of, and motivations for, using online discussion forums as a means for health information behaviour, while Marton and Choo (2012) studied how attributes of websites affect use. Marton and Choo (2012) also note that most studies seem to produce descriptive statistics, and only reporting the extent and nature of online health information seeking. Time, the temporal aspects or rhythmicity of online health information behaviour, however, does not seem to occur in previous studies.

In this thesis, online health information behaviour is studied from a temporal perspective, or the temporal organisation of activities (Reddy, Dourish and Pratt, 2006) to identify and analyse rhythmicity and temporal aspects within online health information behaviour.

3. Theoretical perspectives on time in social sciences

The following chapter provides an overview of the theoretical assumptions and perspectives of this thesis. It also provides a brief overview of how time has been theoretically approached within social sciences in general, as well as in information science.

3.1. Theory and metatheory

The most common meaning of theory is that it is an explanation or generalisation of observed phenomena and regularities (Bryman, 2016, p. 18; Case, 2012, p. 168). Theories can aid in making implications and deductions and guide or influence how empirical evidence is collected to link theory to real world phenomena (Bryman, 2016, pp. 18-19; Case, 2012, pp. 168-169). Theory in this thesis is *an interrelated set of definitions, axioms, and propositions* (Case, 2012, p. 169). Case (2012) argues that this is a typical definition in social research and especially for human behaviour.

Metatheory again, can be defined as the philosophy behind theory, or the fundamental set of ideas about how a phenomenon should be thought about and examined (Bates, 2005, p. 2). Metatheory overlaps with paradigm, which again can be seen as the metatheory, theory, methodology and ethos within a field, all combined, as defined by Kuhn (Bates, 2005, p. 2). From a metatheoretical viewpoint, this thesis takes a nomothetic constructivist approach, as it is concerned with the discovery of general laws underlying the active constructing and understanding of the worlds and contexts, in this case health and illness, in which individuals operate (Bates, 2005, pp. 8-9). Constructivist theories assume that the individual mind is the most important arena of knowledge creation. Therefore, constructivist theories deal with how individuals with specific states of knowledge, or gaps there within, interact with information (Dervin, 2005, p. 27; Talja, Tuominen and Savolainen, 2005). These gaps, states of knowledge, as well as uncertainty is in this thesis studied through online health information behaviour, with the underlying assumption that rhythmicity and temporal variations govern this behaviour. Within information science, the cognitive approach has been developed and applied foremost by Dervin's, as exemplified by the Sense-making approach (Bates, 2005, p. 11; Talja, Tuominen and Savolainen, 2005). Uncertainty is a significant concept in constructivism, both in relation to the cognitive and affective states of the individual. Even this emphasises the individual understanding, current Sense-making views are extensive in

their considerations of the dynamic influences of time, space, cognition, affect, power and culture, as well as individual and collaborative Sense-making (Tidline, 2005). Therefore, applying a bird's eye view on individual behaviour and mental models can reveal collective patterns.

3.2. Time in social sciences

Time is constitutive of human life in nature and society and touches every dimension and aspect of human life. Therefore, the study of time belongs within no single field of study. Time and its theoretical perspectives confronts every field of speciality, and is therefore a genuinely transdisciplinary category, involving physics and mathematics as well as social and health sciences (Bender and Wellbery, 1991, p. 1). The notion of time as a framework for life is, according to Bender and Wellbery (1991, p. 1), the defining quality of the modern world. Time has in different disciplines been approached from different theoretical viewpoints. The physical view and physical theories see time as that which is measured by the clock and calendar (Gorichanaz, 2016). This view sees time as one-dimensional, conceptually inextricable from space and pointing like an arrow toward entropy, or uncertainty and disorder (Gorichanaz, 2016). In social sciences, time has usually been approached as a social construction constituted of social activity and events. Natural, or biological time again, in contrast to social time, emphasises that past, present and future and the qualitative experience of time are characteristics of all living things in nature, not only characteristic of human social life. This biological aspect of time is also of interest in social science, as human beings are biological clocks and organic beings that follow natural rhythms and cycles of light and dark, growth and decay (Adam, 1990, pp. 70-72). Time in social sciences is understood as a characteristic part of life in society, and is, regardless of the fact that we are surrounded by it, usually taken for granted. Time is so deeply rooted in our existence that it is almost invisible, a reason for why time is ignored or neglected and why it continues to elude us (Adam 1990, pp. 1-8: Šubrt, 2001; Roenneberg, 2012, p. 2). This is the underlying cause of the problems related to both understanding and investigating time. According to the physical view of time, the duration of human life is clearly oriented along what Šubrt (2001) calls the arrow of time, moving steadily from a beginning or birth, to an end, death. However, in contrast to this, the world we live in surrounds and influences us by sequences of recurring phenomena, characterised by rhythmicity, periodic changes or repetitions, like the daily light and dark cycles or the seasonal changes (Adam, 1990, pp.

70-76; Roenneberg, 2012, p. 2; Šubrt, 2001). For social scientists, time is one of the essential determinations of change to be measured, and is usually associated with the qualitative experience, or reckoning, of time (Adam, 1990, pp. 11-13; Castoradis, 1991, pp. 42-43).

From a natural science perspective, the qualitative experience of time, past, present and future, the temporally based uniqueness, and even the aspiration to beauty are according to Adam (1990, p. 70) characteristics of all living nature and not the sole preserve of human social life. From this theoretical perspective, living beings are practising centres of action rather than agents of static behaviours. This in turn suggest that the world is orchestrated by rhythms of varying speed and intensity, characteristic to and important for the temporal nature of our environment, and a concept that links biological or natural analyses to those of social sciences, or human organisation and culture. An understanding of this rhythmicity is highly important for social scientists on both a substantive and theoretical level (Adams, 1990, pp. 70-76). This rhythmicity, as exemplified by the work of chronobiologists, has a direct bearing on the lives and social organisation of members within modern societies, and establishes a link between time, or rhythmicity, and health and illness of individuals (Adam, 1990, pp. 72-75). Applying this view on time and rhythmicity also shifts the emphasis away from philosophical speculations of time, to an understanding of time, rhythmicity, and temporal aspects and their effect on life supported by empirical evidence (Adam, 1990, pp. 70-76). This is highly significant, as each new discovery of the rhythmic nature, temporal construction as well as variant repetitive behaviours of human life takes us closer to redefining life and ourselves in temporal terms, which can have enormous and far-reaching social implications (Rifkin, 1987, p. 30). This means that understanding the rhythmicity of nature and applying this in research can be rewarding, as well as have implications for both empirical and theoretical domains (Adam, 1990, p. 73).

In social sciences, a dominant theoretical focus has usually been placed on clock time, duration and sequence of events and their effect on social lives and institutions. This view overlooks, or even neglects, that individuals could be seen as timepieces that, as Adam (1990, p. 75) puts it, *beat to the multiple pulses of our earth and oscillate in synchrony with nature's rhythms*. While the emphasis of clock time is spatial, the rhythmicity of humans is fundamentally temporal. Adam (1990, p. 13) argues that we are able to *grasp time and the complexity of it only if we seek the relations between time, temporality, tempo and timing, between clock time, chronology, social time and time-*

consciousness, between motion, process, change, continuity and the temporal modalities of past, present and future, between time as a resources, as ordering principle and as becoming of the possible, or between any combination of these. This could, as Adam (1990, p. 13) proposes, make time less of an enigma, or mystery, but requires a wider perspective on time and the temporal aspects. What this implies, is that time needs to be temporally conceptualised when the subject matter is life. To fully appreciate the perspective of time and times of human life, human time needs to be explored in both its continuity with and distinctiveness from, natural time. However, as Adam (1990, p. 13) points out, there is a challenge in existing social science approaches for this kind of exploration, and new kinds of ground need to be broken.

3.3. Time in information science

The conceptual issues of time and temporal factors have traditionally received little attention in information science. Reasons for that have been suggested to be the ambiguity of temporal factors and complicated problems of fundamental ontology, as time and temporal factors tends to be everywhere and all human action is embedded in time (Savolainen, 2006a). Another reason is that temporal and spatial factors usually tend to be discussed together, and the temporal factors are implicitly treated and their roles usually not elaborated (Savolainen, 2006a). In general, studies also usually favour and emphasise spatial issues (McKenzie and Davies, 2002). However, time is according to Savolainen (2006a) one of the main contextual factors of information seeking, and has with the rise of the Internet, received an enhanced focus compared to that in the past. From a contextual point of view, time has usually referred to questions like duration, frequency or regularity of access and use as well as in the exploration of the temporal issues within processes (Savolainen, 2006a). From a conceptual viewpoint, Sonnenwald and Iivonen (1999, p. 436) have suggested that time is the non-spatial continuum wherein actions and events occur. Time can also be identified as an episode, interval or an eon (Sonnewald and Iivonen, 1999). These vary in length, from short to longer.

Within information science, time has, according to Savolainen (2006a; 2018), been approached as a qualifier of access to information, as an indicator of the information-seeking process and as a fundamental attribute of situation or context of information seeking, or the general time concepts surrounding life and situations therein. It touches, as Savolainen (2018) states, on Dervin's Sense-making approach.

3.4. Dervin's Sense-making approach

In general, sense making as a theoretical direction has been explored in a many different domains and disciplines, ranging from organisational behaviour to education and human–computer interaction. Within organisational behaviour, sense making is conceptualised as a process through which individuals make sense of complex social dynamic environments and phenomena, to construct their own roles and stories within their organisations (Weick, 2001, p. 11). Similarly, sense making has been employed to study how paediatric critical care nurses make sense of dramatic changes in a patient's status. In sense making research in general, a focus shift from professionals to everyday non-expert sense making is clearly visible (Mamykina et al., 2015). This focus on individuals and their goal to find structure in everyday situations is also essential for Dervin's Sense Making approach.

Dervin's Sense-making approach focuses on the making and unmaking of sense, and the practices and procedures that relate to this. The approach provides a framework for exploring what Wilson (1999, p. 257) calls *the totality of information behaviour*, ranging from the exploration of the context in which information needs arise to the means whereby that need is satisfied, either actively or passively (Wilson, 1999). The approach is implemented in terms of four different constituent elements, situation, gap, outcome and bridge (Dervin 1999; Wilson, 2000). A situation in time and space is the context in which information problems arise and consists of variables such as histories, experiences, identities, past and present horizons, as well as barriers and constraints (Dervin, 2005, p. 28). The gap identifies the difference between the contextual situation and the desired situation, and is influenced by questions and confusion. Moreover, outcome is the consequence, impact and effect of the sense-making process, while bridge is the means by which the gap has been closed (Dervin, 2005, p. 28; Wilson, 2000). Sense-making in this thesis draws mainly on situation, as this refers to the temporal aspect, or situation in time, of knowledge gaps in relation to health. Dervin's Sense-making approach has been described as a metatheoretic tool, or a theoretic net, containing a set of assumptions and propositions as well as methods to study human sense-making (Dervin, 1999; Savolainen, 2006b; Wilson, 2000). Sense-making draws on the constructivist metatheoretical approach, which stresses the importance of the individual actor, and when adopting the Sense-making approach, focus should lie on the user. (Dalrymple, 2001; Dervin, 1999). Dervin's Sense-making is rooted in the assumption that humans are potentially capable of

translating unarticulated emotional, spiritual and embodied feelings to articulated dialogue. This allows humans to communicate and verb how they make and unmake their worlds (Dervin, 1999). In Dervin's Sense-making information is viewed as a situational subjective construction that is rooted in time-space, and created by human observers (Dalrymple, 2001).

Dervin's Sense-making approach has been employed within many different disciplines, such as media studies, education and pedagogy and health communication (Pettigrew et al., 2001). Within the field of information science, the Sense-making approach is associated with the shift in research emphasis from system centred or system oriented studies to user centred studies. The user centred approach was developed by Dervin and Zweizig, and emphasises the information user, or person, as a finder, creator and interpreter of information instead of information sources and their use Dalrymple, 2001; Tidline, 2005, p. 113). As Dervin (1986) states, focus on behaviour, or what people do, has been ignored as individuals have not been perceived as sense-making beings. The emphasis on behaviour in Sense-making, both internal and external, allows movement in time-space for the individual (Savolainen, 2006b). Dervin's Sense-making approach has been used to study information behaviour in a multitude of settings and services, and more recently, the methodology has been focused toward understanding contexts and processes of information behaviour. Sense-making often emphasises individual rather than collective behaviour, which can be explained by the lack of understanding how to operationalise the methodology for studying collective activity. However, sense-making offers a structure for deciphering collective and collaborative information behaviour, and also allows accommodation of multiple variables, including the often overlooked aspect of time (Tidline, 2005, pp. 113-117). The aim of the Sense-making approach is to yield data on behaviour where individuals take steps to construct new sense in life and situations therein, that are useful for information and communication practice (Dervin, 1986).

3.5. Time in Dervin's Sense-making approach

As opposite to the view of individuals and their behaviour as static across time-space, Dervin's Sense-making approach suggests that reality, much like health or illness, is discontinuous, gap filled and changeable across time and space. It is within this time-space that individuals need to make sense of the world (Dervin, 1986; 1992). Within the Sense-making approach lies a notion that life, from time to time, is an encounter with knowledge gaps, problems and discontinuities, which affect basic issues of human welfare

(Case, 2012, 84-85; Dervin, 1986; 1992; McKenzie and Davies, 2002; Savolainen, 2006b). These gaps, or the underlying dissatisfaction with the existing situation stops the sense-maker within the situation (Dervin 1986; Ormandy, 2010). In the Sense-making approach, an individual identifies a need for information upon encountering these knowledge gaps that prevents the individual from continuing the journey through time-space (Dervin, 1992, pp. 62-69). Savolainen (2006b) argues that the journey metaphor suggests that thinking and information seeking can be approached as physical actions, but should not be interpreted as a literal description, but rather a methodological tool, a way of looking. Information seeking is only one response to these knowledge gaps, and Dervin's Sense-making also emphasises other information behaviour, like seeking reassurance or peer-support, expressing or sharing feelings and connecting with others (Case, 2012, p. 92). This broader scope could be compared to Wilson's (1999) definition of information behaviour as those activities a person may engage in when identifying his or her own needs for information, searching for such information in any way, and using or transferring that information. Dervin (2005, p. 27) sees this type of gap-bridging as responsive and impervious to changing conditions or situations. By placing the practices, instead of the persons, in focus, Dervin (1999) argues, that we allow addressing movement, gap, space and time. Dervin also advocates that the study of these gaps has to take place in context, for instance in time (Ormandy, 2010). Different kinds of time as well as the temporal aspects not only receive attention within the Sense-making approach, but are fundamental to the approach, and the conceptualisation of time is central in predicting or explaining information behaviour (Dervin, 1999). This mandate to study temporal aspects and time allows for the possibility of conceptualising information behaviour as temporally patterned. This way, the approach opens up for the possibility to observe patterns of information behaviour that change in response to changing temporal conditions or contexts, like the different hours of the day or changing seasons (Dervin 1999; Dervin 2005, p. 28). This, again, opens up for new theorising. Dervin (1999) notes that Sense-making aims to understand human sense-making so that it can be used in the design of services, systems and procedures. Or, to be more concrete, getting the right information to the right people at the right time (Dervin, 1976). Therefore, the temporal aspect of these knowledge gaps and situations is an important factor, especially on a collective level, as it can provide evidence of rhythmicity for some behavioural patterns. This turns focus on the time and temporal aspects, and

emphasises that *when* an individual seeks information is as important as what he or she seeks (McKenzie and Davies, 2002). In this thesis, Sense-making provides a metatheoretical standpoint and positions how rhythmicity and temporal aspects of online health information behaviour can be examined. It also provides a theoretical view on individuals being different at different times, similar to the discontinuous and cyclical nature of health and illness. This view provides the opportunity to study rhythmicity and temporal aspects of online health information behaviour by analysing this behaviour in relation to two fundamental sense-making constituents, *situation* and *gap*, and the temporal aspects of these situations and knowledge gaps, in which information problems arise.

Dervin (1998) urges to look to the gap and the person-in-situation, as this is where the action in sense-making, or the sense-making instance can be found. In sense-making, time is discontinuous and stops when a gap is identified. The specific time, or context of this situation or gap can be identified in today's online society, as almost all actions leave digital traces that can be aggregated and analysed, with novel approaches.

4. Infodemiology

Infodemiology, a portmanteau of information epidemiology, is defined as *the science of distribution of and determinants of information in an electronic medium, specifically the Internet, or in a population with the ultimate aim to inform public health and public policy* (Eysenbach 2011, p. 155). Infodemiology lies in the crossroads of public health informatics, consumer health informatics, infometrics and web analytics. Infodemiology is highly interdisciplinary emerging approach that involves and requires the collaboration of information scientists, computer scientists, epidemiologists, medical experts, public health informatics experts, behavioural scientists, and statisticians (Eysenbach, 2009; 2011). The user-generated data on the internet, in search engines, on social media and websites, can be seen as a new source of health data for public health surveillance (Ayers et al., 2014a; Eysenbach, 2002; Zeraatkar and Ahmadi, 2018). This has, as Eysenbach (2011, p. 154) states, also made *measurable what was previously immeasurable*, and allows to systematically aggregate and analyse data on preclinical events and behavioural patterns, including online health information behaviour and it's rhythmicity and temporality (Eysenbach, 2011). While traditional epidemiological data gathering methods remain invaluable, the infodemiology approach can provide real-time data on online health information behaviour, free from recall or social desirability biases, and thus complement traditional epidemiology data (Eysenbach, 2009). This user-generated content contains personal experiences, various health information behaviour as well as knowledge sharing (Zeraatkar and Ahmadi, 2018). Infodemiology is rooted in the idea that in the age of the Internet, there is a relationship between population health on one hand, and information and communication patterns on the internet, on the other (Eysenbach, 2009). Thus, the changes in peoples' health behaviour, public attention and attitudes, or health status are echoed in changes in information and communication patterns, and vice versa, both on a population as well as on an individual level (Eysenbach, 2011). Thus, changes in health information behaviour patterns on the Internet can be early symptoms of changes in population health (Eysenbach, 2009; 2011; Zeraatkar and Ahmadi, 2018). However, the arrow of causation can also, as Eysenbach (2011) points out, be bidirectional, as health information can also have a negative impact on population health. For example the paper by Wakefield et al (1998) and its effect on MMR vaccination and a reduction in vaccination rates. The term infodemiology is widely used, but other describing terms for this kind of

research have also been proposed, such as techno-social predictive analysis (Boulos et al., 2010), digital disease detection (Santillana et al., 2014), web data or digital disease surveillance (Althouse et al., 2014; Ayers et al., 2014a; Guy et al., 2012), digital health research (Palmarella et al., 2018) and e-epidemiology (Bexelius, 2009, p. 1). A closely related term is infoveillance, which is defined as *the longitudinal tracking of infodemiology metrics for surveillance and trend analysis* (Eysenbach, 2011, p. 155). However, the term infodemiology itself conveys a key concept as well as emphasis, which is suitable for this thesis and the user-centred paradigm within information science, as infodemiology data are user-generated and the concept itself takes a user-perspective, albeit on a population level (Eysenbach, 2011).

4.1. History of infodemiology

Infodemiology was originally defined as *the study of the determinants and distribution of health information and misinformation* (Eysenbach, 2002, p. 763). The aim of the approach was initially to measure what was being published on the Internet, as much of the discussion in the end of the millennium concerned the quality of health information on the World Wide Web, and the effects this information had on public health (Eysenbach, 2009). Infodemiology as a research discipline and methodology was suggested to be convenient in guiding consumers, patients and health professionals to evidence-based, high quality health information, but also to identify misinformation as well as knowledge gaps in different health behaviours and the evidence behind them (Eysenbach, 2002; 2011). This was also reflected in the early infodemiology studies that analysed websites for inaccuracies, flawed or fraudulent health information (Eysenbach, 2009; Zeraatkar and Ahmadi, 2018).

Infodemiology was in 2006 further defined as *an emerging set of methods which studies the determinants and distribution of health information for public health purposes*. (Eysenbach 2006, p. 247-248). Eysenbach (Eysenbach 2006, p, 248) further stated that *the development of infodemiology metrics based on automated tracking and analysis of the distribution and determinants of health information (both supply and need) in a population and/or information space is possible and can provide important clues and evidence for public health policy and practice*. Since 2002, an increasing number of people in developed countries had started to utilise the internet for health information seeking and often first consulted the internet before going to a doctor (Eysenbach, 2006). Research utilising this health information seeking data as well as more unconventional research methods had been published. The first paper to

analyse health information seeking data, in this case Google search data for syndromic surveillance, was published in 2006, describing a model for predicting an influenza outbreak during the flu season 2004-2005 in Canada, based on changes in Internet search activity for flu-related information (Eysenbach, 2006). This type of health information demand tracking on the Internet was described as a novel and promising method for public health surveillance, which could complement the traditional surveillance approaches. Eysenbach also stated that the science of infodemiology was needed for methodology development and measuring patterns and trends for general health information, knowledge gaps, as well as predictive disease and syndromic surveillance in real-time (Eysenbach, 2006).

The earlier infodemiological research showed that the Internet made health information seeking and use measurable, and that trends and knowledge gaps could with the help of infodemiological data be traced over time, something that was previously difficult to measure (Eysenbach, 2006). Since then, infodemiological research has showed a significant rise, and the amount of publications since 2002 surpasses one thousand (Zeraatkar and Ahmadi, 2018).

4.2. Aim of infodemiology

Infodemiology is still an emerging discipline that tackles a modern societal problem that it is not about the availability of information anymore, but its aggregation and analysis. The rapidly rising amount of user-generated data from social media, search engines and webpage traffic opens up for the possibility to study health behaviour in new ways. The underlying aim and idea of infodemiology is to measure public opinion, attention, behaviour, knowledge, and attitudes by analysing health information behaviour on the Internet (Eysenbach, 2011). Infodemiological studies have the potential to provide reliable and meaningful indicators, temporal trends and rhythms for health information needs, seeking, use and sharing. Analysing health information behaviour on the Internet, including information needs, seeking and use, can provide valuable insights to health-related behaviours of populations, including health status and health problems and knowledge about health-related conditions (Eysenbach, 2009; 2011). A deeper understanding of these behavioural trends and patterns and their temporality and rhythmicity can provide additional information to public health officials and foster our understanding on how to maximise the use of the Internet to improve health, public health and health related outcomes (Eysenbach 2006; 2009; 2011). This could be achieved for instance by

tailoring and targeting health communication and education strategies (Eysenbach, 2011). Tailoring the right health information at the right time, and targeting the right audience segment based on health information need or behaviour in general, can be beneficial for health behaviour, communication and promotion (Enwald, 2013, pp. 57-58). It needs to be emphasised that utilising the infodemiology approach and metrics for research in different health related areas can complement traditional health research and epidemiology as well as the methods employed within these fields, and should in no way be seen as a replacement (Eysenbach, 2006; 2009; 2011; Zeraatkar and Ahmadi, 2018).

4.3. Demand- and supply-based infodemiology applications

The infodemiology approach and its framework has been divided into supply- and demand based applications, depending on what is studied, and what data, or metrics is employed (Eysenbach, 2009). Infodemiology data, or metrics, are usually derived from unstructured, textual, openly accessible user generated content (Eysenbach, 2009; Zeraatkar and Ahmadi, 2018). Supply-based infodemiology analyses what is being published online, on webpages or in social media, whilst demand-based infodemiology analyses health information seeking behaviour, such as search engine queries or website navigation, or click behaviour (Eysenbach, 2011; Zeraatkar and Ahmadi, 2018). The first infodemiology studies, which were concerned with the quality of health information on the internet, could be categorised as supply-based approaches, as they analysed information prevalence and information occurrence ratios, the most basic infodemiologic supply indicators. Longitudinal tracking of supply-based infodemiology metrics, for instance analysing changes over time in the absolute or relative number of internet postings for a certain keyword or concept is particularly useful, as it allows seeing changes in relation to temporal variations on different time-scales (Eysenbach, 2009).

Demand-based infodemiology and analysis of health information seeking behaviour again, provides valuable insights into information needs. For instance, supply-based indicators, such as search engine query data, or the number of clicks on a website, can reveal temporal trends and patterns for a specific health related topic in a given geographical region (Eysenbach, 2009).

Both supply- and demand-based infodemiology employ similar methodological workflows and face similar problems (Eysenbach, 2009). The similarities in both methods and problems relate, as Eysenbach (2009)

states, to the selection and filtering of health information from large textual datasets, attempts to interpret the information semantically and employing descriptive and analytical statistical methods to detect trends and patterns. Regardless of the approach, both supply and demand based infodemiology analyse and describe patterns of health information behaviour in electronic media. Analysing infodemiology metrics on how and when people engage in online health information behaviour, including how they seek, communicate and share information, can provide novel and valuable insights into health-related behaviour and can inform public health officials (Eysenbach, 2009). It also needs to be noted, that while the Internet has been, and still is by far the most popular source of infodemiological data, any consumer health application including personal health records or wearable devices can be harnessed for infodemiology and infoveillance research (Eysenbach 2011). The infodemiology approach provides unmatched opportunities for the management of health related data and information generated by the users, as well as unique opportunities for insights into online health information behaviour (Zeraatkar and Ahmadi, 2018).

4.4. Previous infodemiological research

A recent review of published infodemiology research found that a majority of infodemiology studies fall in the demand-based infodemiology category, while supply-based studies as well as articles combining both demand and supply based infodemiology are fewer in number (Zeraatkar and Ahmadi, 2018). Even if the majority of previous infodemiological research, especially the research that has caught the attention of media, has focused on disease outbreaks, such as influenza or different viruses, infodemiology should not be misunderstood as having only practical applications in the context of infectious diseases (Eysenbach, 2009; 2011; Zeraatkar and Ahmadi, 2018). Monitoring and combating various behavioural risk factors as well as health behaviours and phenomena related to these are equally important application areas within infodemiology (Eysenbach, 2011). As already mentioned, more than one thousand infodemiology articles had been published by the end of 2016 (Zeraatkar and Ahmadi, 2018). However, it needs to be noted, that not all studies utilising what could be called infodemiology metrics are labelled under infodemiology, and because alternative describing terms are, as already mentioned, used the number of infodemiology-like publications could be higher. Also based on this, the results of the scoping review of Zeraatkar and Ahmadi (2018) are somewhat questionable. However, it provides a road map to subjects studied as well

as utilised data and methods. Most articles identified as infodemiology research have a descriptive aim, while the surveillance, or infoveillance, aspect is slightly less studied. The most common data source for both demand- and supply-based studies is Google Trends, a web-based database that provides information-seeking trends based on what people use the Google search engine for, followed by the microblogging social media Twitter. As far as analysis types goes, the majority of research employ time series analysis, followed by cross-sectional analysis, or both combined. From a geographical perspective, the majority of infodemiological research is related to developed countries, such as the United States, Italy and United Kingdom, and a clear majority concern developing countries (Zeraatkar and Ahmadi, 2018).

The infodemiology approach has been utilised to study many different health related phenomena. These include mining tweets or search engine query data for pandemics, different ailments and public health issues (Chew and Eysenbach, 2010; Eysenbach, 2006; Paul and Dredze, 2011; Wilson and Brownstein, 2009), internet search trends for a multitude of illnesses and health issues, as well as epidemiological research (Abedi et al., 2015; Brigo and Trinka, 2015; Carneiro and Mylonakis, 2009; Seifter et al., 2010) and accessing and sharing information related to different topics (Wong et al., 2013; Matsuda et al., 2017). The approach has also been used in identifying and monitoring what is being published on the internet, for example news articles, outbreak reports or anti-vaccination sites (Eysenbach, 2002). Health awareness campaigns and their effectiveness, as well as public interest for different health related topics have also been studied using infodemiology metrics (Vasconcellos-Silva et al., 2017). Moreover, infodemiology metrics have been used to construct patient-centred research tools by extracting user-generated health outcomes data to monitoring different health related issues, like drug side effects or treatment adherence (Wicks et al., 2009; 2010; Frost et al., 2011)

Rhythmicity, temporal variations and patterns have also been investigated using infodemiology metrics, mostly for specific health issues or ailments. These range from mental health problems (Arendt and Scherr, 2017; Ayers et al., 2013; Chen et al., 2018), and somatic diseases like diabetes (Tkachenko et al., 2017) to disease and influenza outbreaks (Bragazzi et al., 2017; Kraut et al., 2017; Ortiz-Martínez and Jiménez-Arcia, 2017; Osuka et al., 2018; Seo and Shin, 2017). Online health information behaviour on prescription drug abuse, smoking cessation and e-cigarettes have also been studied from a rhythmical viewpoint (Zeraatkar and Ahmadi, 2018).

However, research on more general health related temporal patterns utilising infodemiology metrics is scarce. Ayers et al. (2014) investigated the circaseptan rhythms of health considerations by monitoring internet search queries, while Jadhav et al. (2014) touched upon the temporal patterns, such as morning, afternoon, evening and night, while analysing how online health information seeking behaviour may differ by accessing device. Scrutton and Stevenson (2018) studied information seeking behaviour in relation to nutrition and food, and if human appetitive behaviour shows rhythmicity on daily or seasonal level. Infodemiology studies relating to Finland are also scarce. Apart from the studies in this thesis, only a few infodemiological studies have been conducted in Finland. Gencoglu and Ermes (2018) utilised social media data (Instagram) to predict influenza like illness in Finland, while Kluger and Bouchard (2019) conducted a comparative study on search engine query trends in interest for melanoma, breast cancer and prostate cancer. Pesälä et al. (2019) studied whether health care professionals' medical database queries on oseltamivir, an antiviral medicine, and influenza could complement epidemiological disease surveillance when detecting influenza epidemics, and their seasonal trends. A similar research setting was applied in Pesälä et al. (2017), with the aim to compare the general public's visits on articles about Lyme disease to healthcare professionals' website behaviour of a physician's databases articles to evaluate temporal and seasonal information-seeking trends, and the effects behind these variations. To the best of the authors' knowledge, the studies in this thesis are the first where infodemiology metrics are used to examine rhythmicity and the temporal variations and patterns of various health related phenomena, as well as depression, in Finland.

4.5. Infodemiology limitations

As with all other approaches, methods, and metrics, infodemiology also presents certain limitations. As one of the aims of this thesis is to evaluate the strengths and weaknesses of utilising the infodemiology approach to study the context of time in online health information behaviour, both the strengths and limitations are presented and discussed in chapters 6.1.1 and 7.5.

5. Methods and data-analysis

This chapter presents an overview of the research methodology and data utilised in the different studies included in the thesis. As an introduction, the general research paradigm, process and strategies of the thesis will be presented. This is then followed by a description of data, data collection and data analysis applied in each study.

5.1. Methodological considerations

Research methodology is usually guided by the ways in which scientists perceive the connection between different viewpoints about the nature of reality and how to examine or inspect it (Bryman, 2016, pp. 149-150). Research data are, as Bryman (2016, p. 10) reminds us, invariably collected in relation to something, for instance a social problem or a theory. Sometimes data are also collected for simple fact-finding, or when a specific opportunity arises (Bryman, 2016, p. 10). However, research data become significant when viewed in relation to theoretical standpoints, emphasising the relation between theory and research. This also raises the question of why data are collected or used, to test or to build theories (Bryman, 2016, p. 21). Research can therefore be done either to answer questions posed by theoretical considerations, *deductive* research, or to view theory as something that emerges from the collection and analysis of the data, or *inductive* research (Bryman, 2016, pp. 21-23).

All five studies included in this thesis fall in the category of deductive research, a strategy often associated with quantitative research (Bryman, 2016, pp. 21-24). However, as Bryman (2016, p. 23) reminds us, the process is not clear-cut and should be considered as a general orientation, and deduction entails elements of induction, just as induction entails elements of deduction. Some of these iterative strategies between data and theory have also been present in this thesis.

From a research strategy point of view, this thesis utilises a quantitative research approach, as it emphasises quantification in both collection and analysis of data (Bryman, 2016, pp. 149-151). In a sense, this thesis is methodologically positioned somewhere between a positivist and phenomenological epistemology, as it relies on objective observations to explain an aspect of behaviour of phenomenological nature, that is the temporal variations of health related sense-making on a collective level (Bryman, 2016, pp. 24-27). This is also an aspect often raised by Dervin's

Sense-making approach, which has been characterised as an approach in-between**.** The studies in this thesis are merely preliminary attempts of explaining an aspect of behaviour in relation to time that in the future needs to be studied more closely, by applying a qualitative research perspective as well.

5.2. Methodology

All rhythmic events, ranging from phenomena like ocean waves to human behaviours can be plotted from a series of regularly recorded descriptors called a time series. Rhythmic events typically occur at regular intervals with predictable waveform, amplitude, and period, without an element of causation necessarily known (Dunlap et al., 2004, p. 15). All studies in this thesis are methodologically similar, as all are based on univariate analysis with one variable, time, common for all studies. In all studies, time is categorised as an interval/ratio variable, as the distance between the categories are identical across the range of categories. The temporal variations in distribution of search engine query volumes and discussion forum messages are calculated by both arithmetic means in Study I and II as well as relative frequency distribution in Studies III, IV and V. The arithmetic mean, or simply mean, is the sum of all values in a distribution divided by the number of values (Bryman, 2016, p. 688). This is suitable, as the variable in this thesis, time, is an interval/ratio variable. Frequency distribution again is simply put the way the data are distributed along the scale of the variable of interest. Although simple, it is still a concept that provides the essentials in descriptive statistics (Riffenburgh, 2012). When it comes to the relative frequency distribution or rate of occurrence, this is more suitable than absolute frequency itself, as relative frequencies such as rates or ratios provide the proportion of the sample bin, or interval (Eysenbach, 2009; Riffenburgh, 2012). Intervals in all studies within this thesis consist of time periods, from hours and days, to weeks, months and years. A thorough description of the data sources and data-analysis methods for the different studies in this thesis are presented in chapters 5.3-5.5.

5.3. Google Trends

The data in studies I and II consist of search engine data, in this case from the search engine Google, the most popular search engine in the world, including Finland (Statista, 2019; StatCounter, 2019). Data from Google has been shown to be more unlikely to suffer from social censoring or bias, as individuals using the search engine are online and likely alone, factors that

support expressing stigmatising or taboo thoughts (Kreuter et al., 2009). Search engine data are common infodemiological metrics, utilised in many different infodemiological studies as mentioned previously (Zeraatkar and Ahmadi, 2018). Search engine data fall under the demand-based infodemiology methods and applications, useful for identifying trends, patterns and rhythmicity in information needs and information seeking from big data, or millions of searches (Eysenbach, 2009; Stephens-Davidowitz, 2013, p. 4). The data for both studies were downloaded from Google Trends (available at http://www.google.com/trends). Google Trends is a public, open web-based database that provides time series data, of information seeking trends, and how often specific keywords, subjects and phrases have been enquired in the Google search engine over a chosen time period and geographical region. This is also called measurement of demand prevalence indicators (Eysenbach, 2009). Data from Google Trends are provided as relative search volumes (RSV), which in contrast to what raw search volumes would provide, correct for variations in search volumes due to term popularity, Internet availability or disposable time (Althouse et al., 2014; Ayers et al. 2014). Google Trends works by analysing a portion of queries to calculate how many searches have been conducted for the query term in question, relative to the total number of queries in Google over the same time period across chosen regions of the world. This relative search volume is the query share of a particular term for a given location and time period, normalised by the highest query share of that term over the time-series. A relative search value of 100 is thus the period with the highest search proportion for queries, and 50 is 50% of the highest search proportion. Queries from information seeking activity on Google are disaggregated to geographical and temporal units, ranging from years to minutes, and de-identified from any identifying information to protect user privacy.

In Study I, data collection is based on five Finnish queries relating to depression: *masennus* (depression), *masennus oireet* (depression symptoms), *masennustesti* (depressiontest), *depression* and *depression test* (the Swedish as well as English term for depression and depression test). A sixth search query, *masennus testi* (depression test) was added to the second (II) study to provide the alternative spelling for depression test. Query terms in Study I and II were not used in combination and without quotes around the search terms. Supplementary terms to the root term *masennus* (depression) were identified using the top related search queries function in Google Trends. The suggestions for terms are based on identification of either content or users' search behaviour. The search terms in Study I and II were selected to

characterise a wide sample of queries related to depression, in general as well as for symptoms and treatment, in both official languages in Finland, Finnish and Swedish. The settings were also geographically specified to Finland to prevent the inclusion of search queries that originated from other regions.

Data for Study I were downloaded for three different time periods: January 2004 to December 2015, January 2010 to December 2015, and monthly for the year 2015. These different time periods were chosen to get comprehensive data on temporal variation of depression related information seeking. Rising search volumes as well as the reported improvements in the Google Trends algorithm also motivated to select three different time periods for data gathering (Phelan et al. 2016).

Data for Study II were downloaded on a weekly basis from Google Trends in March 2017. Weekly downloads allowed obtaining hourly relative search volume data for each day in March. The reason for Study I and Study II being separate is methodological, as the daily relative search volumes are only available for the past seven days. The data collection resulted in a relative search volume for all 744 hours between 1st and 31st of March for the six different search terms. In Finland, the prevalence of depression has been shown to peak in spring (Näyhä et al., 1994). Therefore, data collection was conducted in March.

5.4. Suomi24

Studies III, IV and V all utilise the same data from the Suomi24 discussion forum. Within infodemiology, the use of data from websites, discussion forums and social media fall under the category of supply-based infodemiology methods and applications (Eysenbach, 2009). Suomi24 (https://www.suomi24.fi/) is the largest, and most popular, free and anonymous discussion forum in Finland. The Suomi24 discussion forum data, the whole of which is available for research purposes, provides infodemiologic supply indicators such as information prevalence (Eysenbach, 2009), as it contains the entire discussion forum database including all posted messages in all categories. The data ranges from the 1st of January 2001 to the 31st of December 2017, which allows for longitudinal tracking of online health information behaviour. However, in the third study (III, in review) the Suomi24 data was still limited to the time period 1.1.2001 to 14.5.2015. The updated data set, as well as the older one, can be downloaded for research purposes through FinCLARIN's Kielipankki Korp-interface (http://urn.fi/urn:nbn:fi:lb-2019010801).

The Suomi24 data is an unusually comprehensive data set, with over 82 million (82,858,608) messages between a time period ranging from 1.1.2001 ot 31.12.2017 (53,366,704 messages during 1.1.2001-14.5.2015). The Suomi24 discussion forum had 832,000 unique visitors per week in 2015, which can be considered quite high in a country with a population of 5.5 million. Each of whom spent an average of 7 minutes on the forum per session. The total amount of page loads per week was over 17 million in 2015 (Aller, 2016). The discussion forum had a monthly reach of over 2 million internet users during the first half of 2018, ranking it the 8th most popular Finnish internet site (FIAM, 2018). Statistics show that in 2014, roughly a quarter (24%) of the Finnish population within the age demography 16-89 had engaged in posting something on a discussion forum during the last three months. In 2017, the percentage was clearly lower, at 8%. Participation was evidently more active for younger demographics, as over 40% of 16-34 year olds, and 35% of 35-44 year olds, had posted messages in a discussion forum in 2014. The corresponding numbers were also lower for these demographics in 2017, as 12.5% of 16-34 year olds, and 12% of 35-44 year olds posted messages in a discussion forum (Tilastokeskus, 2017). The Suomi24 discussion forum covers several different topics such as pets, fashion and beauty, tourism, cars and vehicles and health. The health (*fi. terveys*) category with 3.8 million messages (2.5 million during the time period 1.1.2001-14.5.2015) is divided into sixteen sub-topics, ranging from plastic surgery and diseases, to medication and mental health. Sub-topics are then further divided into a different number of sub-categories, relating to the sub-topic. The content of the messages is not included in any of the studies (III-V).

5.5. Data-analysis

All five studies in this thesis are observational studies with univariate analysis, with a focus on explaining human behaviour, in this case the rhythmicity and temporal variations of online health information behaviour. In this thesis data analysis mainly consists of univariate analysis, with one variable, time, as the key nominator for all studies.

5.5.1. Study I

To get an average relative search volume (RSV) for the different weekdays, data for each month of 2015 was downloaded. These monthly data contain an average RSV for each weekday. Data were analysed in MS Excel 2016. For all five queries during the different time periods an average relative search

volume per month was calculated, and an overall average RSV per month for all query terms per month was then determined. An average RSV for each weekday in 2015 was also calculated. Differences in popularity based on relative search volumes between the five different search terms are also noted. All data-analysis was conducted in MS Excel.

5.5.2. Study II

In Study II, data analysis of the hourly data was conducted by both a continuous wavelet transform and wavelet power spectra to focus on the diurnal variations. Both analyses were conducted with MathWorks Matlab R2017b. Longer term trends were removed by discarding components with a period of longer than 32 hours. This was done to extract the diurnal variation. Components with a period of less than 4 hours were considered noise, and similarly eliminated. Thus, the calculations provide a reconstruction of the variations without trends and noise. The reconstructed data are superimposed with their arithmetic mean, separately for each query. By applying the Tukey-Kramer procedure for multiple comparisons, an analysis of variance was conducted to determine the statistically significant difference between the interests of the corresponding hours.

Data in Study II were also trichotomised and analysed in three time blocks: 8.00 to 15.59, 16.00 to 23.59, and 0.00 to 7.59. This trichotomisation of the day allows distinguishing and comparing structured time (office hours), unstructured time (leisure time) as well as night-time. All hourly relative search volume values for each search term were added to the time blocks and the differences between the three time blocks were then calculated. Analysis of variance with the Bonferroni correction was conducted to calculate statistical significant differences in the distributions between the different time blocks. The trichotomisation analyses for were performed with SPSS (version 24.0; SPSS Inc., Chicago, Illinois, USA).

5.5.3. Study III

Message metadata included in the data analysis in Study III have been extracted from messages posted in the sub-category “Depression” of the Suomi24 discussion forum. The data consist of metadata, in this case time stamps, of 116 781 messages, of which 15 386 are original posts, 57 682 replies, and 43 713 comments to replies. Messages range from January 1st 2001 to May 14th 2015. The data was processed in JSON format, and analysed using Python and the Python Data Analysis Library "pandas". The Python Data Analysis Library "pandas" provides the functionality to aggregate the

raw data and calculate the statistics of interest. The timestamps are converted from UTC to local time in Finland to account for daylight savings time. To evaluate the statistical significance of observed differences in the distribution of posts in different categories, we use the statistical library in SciPy to conduct a chi-squared test of independence. As this results in a multiple hypothesis scenario, the p-values are corrected for multiple comparisons by the step down method using Sidak adjustments from the StatsModels library. The standard significance level of 0.05 is used for p-values after the multiple hypothesis correction. A relative hourly distribution of messages was also calculated, and a triangular window rolling mean of the monthly distribution of all forum messages, as well as messages in the health category and messages in the depression subcategory, was calculated over a year.

5.5.4. Study IV and V

The updated Suomi24 discussion forum data, ranging from Janyary 1st 2001 to December 31st 2017, are provided in custom verticalised text format. The data represent a parsed version of the text for each message in the forum. The univariate analysis studies IV and V examines the metadata of the posts, in both cases the time stamp of the message, as well as the topic, sub-topic and sub-category where the message has been posted. The content of messages is excluded in both studies. The data files in study IV and V are processed using Python with the Python Data Analysis Library "pandas" and plotting library "matplotlib". The analysed messages in both studies are binned and counted according to the forum category or sub-category, as well as the hour and date of the posted message. Calculating this at different aggregation levels reveals rhythmicity and temporal variations and patterns in the data. E.g., when binning the posts by hour, and aggregating overall days in the data, the resulting distribution represents the average daily variation in posting intensity. Similarly, the posts are separately binned by date, and aggregated over the 17 years. To smooth out the effect of random variations and to reveal robust patterns, a rolling triangular window of 15 days is applied. The relative shares of messages within any bin is also calculated to eliminate the general variation in posting activity during the years 2001 and 2017. This highlights the differences between categories of interest and measures by how much the categories are over-represented or under-represented, compared to the overall posting rate in the Health category.

5.6. Reliability and validity

The reliability of research is mainly concerned with the question of whether the results of a study are repeatable, or if the measures are consistent (Bryman, 2016, p. 41). Validity in turn is concerned with the integrity of the conclusions that are drawn from research. Validity is usually divided into internal and external validity, ecological validity, inferential validity as well as measurement validity. The last one relates primarily to quantitative research. Validity is related to reliability, as unreliable measures cannot provide valid measures for a concept (Bryman, 2016, p. 41).

Search engine data have repeatedly been criticised for lack of validity and reliability, because tools that provide search engine data (services like Google Trends) are not primarily created for scientific purposes (Tran et al., 2017). Lazer et al. (2014) argue that the use of search engine data might lead to problems of overfitting, in cases where a large number of possible predictors or variables (different search terms) are used to predict only a small number of cases. In the use of search engine data, the validity of reported results, often in relation to correlation of occurrences for different health related phenomena are reported cautiously, with an emphasis on the need for further research, follow up and independent verification (Tran et al., 2017). However, from a validity perspective, this issue is not as concerning for this study, as the aim is not to find correlations to occurrence, but to explore the temporal variation of online health information behaviour, or interest in a specific issue, not actual incidence.

More generally, for search engine data, it remains unclear how users actually perform their searches and what factors influence this behaviour and to what extent (Tran et al., 2017). There is therefore a challenge in understanding the semantics of Google queries, as reasons individual search queries do not necessarily reflect the actual mood state of the user (Yang et al., 2010). However, as the aim of study I and II is to explore when people engage in health information seeking in relation to depression, the underlying causes and factors fall outside the scope of this study. Moreover, in the case of search engine data, Google Trends only provides overall normalised search volumes, with limited information about the method by which search data are generated, and the specific algorithms employed to analyse it. This does not allow differentiating search volumes by user demographics. The number of language minorities in a country have also been suggested to affect the probability of bias in Google Trends data (Tran et al., 2017). This issue has been taken into account as both Finnish, Swedish and English search queries have been included into Study I and II, but

exclusion of other language minorities in Finland may have had an overall impact on the results.

From a reliability perspective, the only study that does not deploy a longitudinal perspective is Study II, where data gathering was concentrated on a one-month period. Testing the stability of the results in Study II would propose a rather complex research design, as data would be required to be downloaded every 7 days during an extended period of time. For the rest of the studies, the longitudinal perspective takes into account both time and societal changes, making the reliability solid from a stability perspective. As all studies involve only one variable, time, and the distribution of online health information behaviour over different time periods, internal reliability, or inter-rater reliability is not an issue in this thesis.

As is common in social sciences, measured concepts are often abstract, which makes establishing validity difficult (Bryman, 2016, pp. 158-162). From a measurement validity perspective, the standpoint in this thesis is that both health information seeking in search engines, and health information seeking and sharing in a discussion forum actually reflect online health information behaviour, and that the metadata, in this case timestamps, actually measure engagement in this behaviour during a specific time of day, week, month or year. The hypothesis in this study is deduced from both previous research and theory, which adds construct validity. However, to establish a more robust convergent validity, Studies III, IV and V could have analysed message content to establish if messages actually did express health information needs or sharing. This validity issue has been raised in previous research as a need to ensure the validity of web data and its reflection of real-world outcomes and how to accurately interpret what is shared online (Yeung, 2018). As Yeung (2018) states, what people share online may constitute *a form of self-presentation or performance to a specific audience, rather than convey true feelings.* Another noteworthy aspect is that this thesis only examines online health information behaviour, and the results in this study are not validated with offline data.

5.7. Ethical considerations

The data and data-analysis in all studies are based on public meta-data that do provide neither information about the race, gender, age, or any other identifying information of the person entering either a search term in a search engine or posting a message to a discussion forum. Analysis only consists of time stamps for search queries and discussion forum messages. None of the studies involve any intervention in the integrity of a person. The

data itself can be seen as public data, as it is collected from the internet, which is claimed as a freely accessible public domain intended for the public. Moreover, the Suomi24 data as well as the Google Trends data do not belong to the private sphere of communication, and may therefore be utilised for research purposes (Savolainen, 2011). Even if there are several ethical concerns in utilising web data for research, these are of less concern within these studies, as analysis only consists of time stamps. Moreover, as all research within this dissertation is based on public information, registry and documentary data or archive data, no institutional board review was required, as proposed by the Finnish National Advisory Board on Research Integrity (Kohonen, Kuula-Luumi and Spoof, 2019, p. 62). However, on a more general note, it needs to be emphasised, as also highlighted by Ahmed et al. (2017, pp. 2-3), Eysenbach (2009) and Mamiya et al. (2017), that the use of openly accessible infodemiology metrics, or web data in general, for studying health and behaviour, highlights novel issues and threats to privacy as well as informed consent. There is a need to accompany the utilisation of user-generated web data as well as the applications in relation to these with a clear ethical framework to avoid causing serious distress among Internet users (Mamiya et al., 2017).

6. Results

The purpose of this dissertation is to examine if online health information behaviour follows temporal patterns and variations on different biological time-scales. Further, the aim is to study this rhythmicity for different health related phenomena by utilising the infodemiology approach. The underlying assumption is that, as with many other behavioural traits, also health information behaviour is governed by time, rhythmicity and temporal aspects. To address this assumption, five different studies were conducted. As the different studies answer different research questions, the results are presented in relation to these.

6.1. Infodemiology as an approach

The aim of Study IV is to answer both research question 1: *What are the strengths and weaknesses of utilising the infodemiology approach to study the context of time in online health information behaviour?* as well as research question 2: *Does online health information behaviour in general follow rhythmicity?* In Study IV, the review of the strengths and limitations of the infodemiology approach to study online health information behaviour, and its temporal aspects, are not based on a systematic review, but instead identified and evaluated based on the literature cited in the study. Moreover, in Study IV the infodemiology approach is exemplified by presenting some empirical findings on rhythmicity of online health information behaviour based on longitudinal metrics from in the Suomi24 discussion forum. The most interesting findings from the larger sub-topics within the health category of the forum were chosen because of the robust and recurring circadian, circaseptan and circannual rhythmicity and variations that they displayed.

On a more general level, the number of messages in the Suomi24 forum and its heath category show a clear peak in activity during 2006. After this, the number of messages on an annual level have been decreasing steadily, as can be seen in Figure 1.

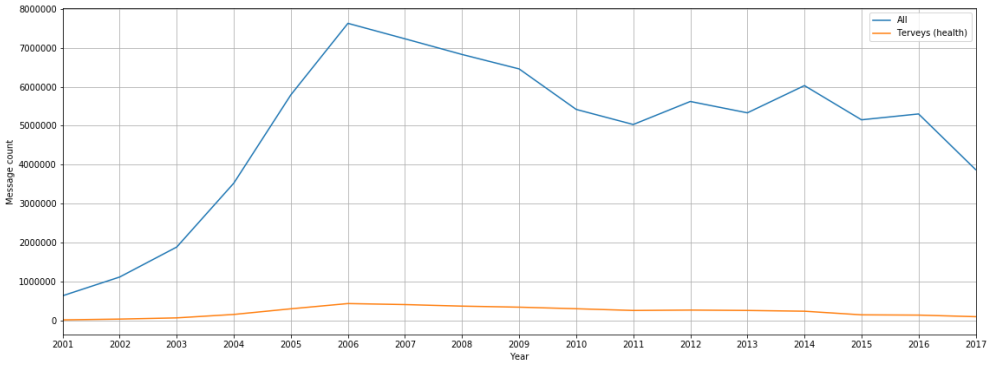

Figure 1. Distribution of the number of messages in the health (fi. Terveys) category compared to all other messages in the discussion forum between 1.1.2001 and 31.12.2017.

A closer comparison of message distribution in the Suomi24 discussion forum on a circadian level reveals a clear unimodal curve both for all the messages in the discussion forum and the health category. The relative frequency of the message distribution starts to rise in the morning, and peaks in the evening (Fig. 2). However, a closer examination of the distribution of messages in various sub-topics within the health category, reveals clear temporal variations on a circadian level, as can be seen in the results of Study V (see chapter 6.2).

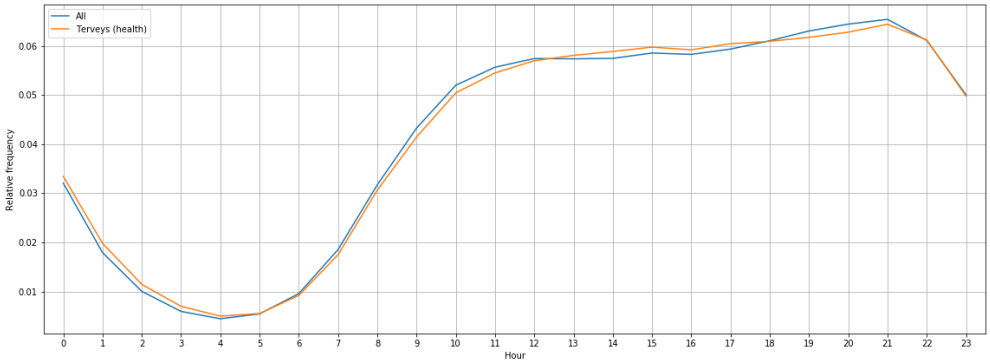

Figure 2. Comparison of hourly distribution of messages in the health (fi. Terveys) category compared to all other messages in the discussion forum.

When analysing the relative frequency in message distribution on a circaseptan level, the results also show a clear and recurring rhythmicity. As can be seen in Figure 3, both the discussion forum as a whole as well as the health category follow a rhythmical pattern with higher activity in the beginning of the week, which then decreases to a clear trough on Saturday.

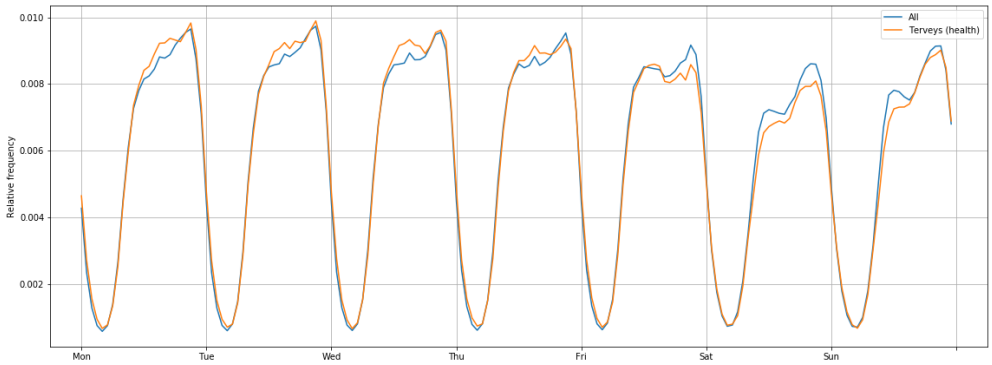

Figure 3. Message distribution in the discussion forum and the health (fi. Terveys) category on a circaseptan level.

Again, a closer examination of relative message frequency within the health category expose clear temporal variations in online health information behaviour. An example where differences become evident is shown in Figure 4. As can be seen, discussion forum activity on a circaseptan level for alcohol (*fi. alkoholi*) and birth control (*fi. ehkäisy*) are each other's opposite, with birth control presentation amplified activity during the beginning of the week, a time when alcohol related discussions are less active. This pattern alters mid-week and during the weekend, as alcohol related discussion start to increase, while activity for birth control is decreasing.

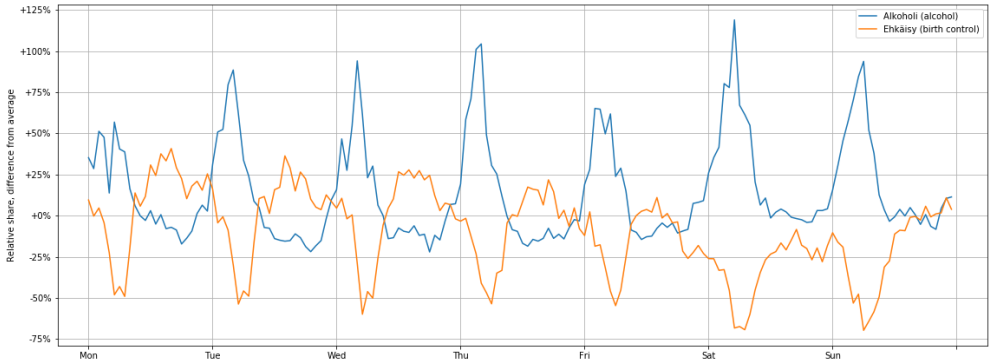

Figure 4. Comparisons between message volumes for alcohol (fi. Alkoholi) and birth control (fi. Ehkäisy) on a circaseptan time-scale.

A comparison of the relative frequency of messages between the health category and all other messages shows that rhythmicity on a seasonal level are quite similar (Fig. 5). Activity in both the health category and the forum as a whole follow a bimodal curve, with peaks in the beginning of the year,

spring, late summer and autumn. However, the peaks during the first half of the year, as well as August and September are slightly higher for health related discussions. Clear troughs in discussion forum activity are visible during summer and the end of the year, in December.

Figure 5. Circannual comparison between all messages and messages in the health (fi. Terveys) category in the discussion forum.

Similar to differences in online health information behaviour on a circadian level, seasonal variations are also visible within health related discussions in the Suomi24 discussion forum, as presented in Study V.

6.1.1. Strengths and limitations of the infodemiology approach

Despite some limitations in utilising the infodemiology approach with its methods and metrics to study online health information behaviour, the implications and strengths that they bring outweigh many of the weaknesses. Using large, or big data sets such as the Suomi24 data allows for more longitudinal studies on temporal variations of online health information behaviour, in contrast to "snapshots" provided by surveys or questionnaires. This allows identifying patterns of interaction, which, as Giddens (1979, p. 202) state, cannot be identified or revealed unless examined over time. As an increasing amount of digital data are gathered in relation to health and health information behaviour, a clearer picture of how people interact and behave in relation to their health, is emerging. Studying and analysing this data can yield in a more holistic view of health and health behaviours, and even reveal completely new trends, patterns and correlations. Identifying clear and robust patterns and rhythms also facilitates the prediction of the future, making precautionary action possible within health promotion and prevention (Adam, 1990, pp. 74-75; Ayers et al., 2014). Discovering and understanding these health related temporal

rhythms for the Finnish population therefore has strong potential for behavioural gains as well as improving public health. Health promotion campaigns could immediately be made more cost effective by targeting the population at the right time. Considering the temporal aspects and patterns of online health information behaviour on different time-structures, at times when more individuals are contemplating their health and engaging in health related behaviour could have substantial impact on interventions (Ayers et al., 2014).

Utilising infodemiology metrics also has another methodological strength, as it avoids the interviewer bias, whereby what is said during an interview is influenced by the researcher (McKee, 2013). Data from both Google Trends and the Suomi24 discussion forum involve no direct contact between the subjects and the researcher. Time stamps for both queries as well as messages are generated automatically, making data collection and analysis rapid, even real-time, with the appropriate tools.

One major limitation of the infodemiology approach, with its methods and metrics is the lack of demographic data in infodemiology metrics. This is often due to the de-identified web data that is provided for researchers (Ayers et al., 2014a; Eysenbach, 2009; 2011). This prohibits and limits researchers from drawing conclusions related to different demographics characteristics in relation to online health information behaviour. It is imperative that this limitation is kept in mind, as socio-demographic variables such as age, gender and level of education have, as previously noted, an impact on online health information behaviour. Relating to this, the infodemiology approach only provides data on, and tracks the segment of the population that engage in online health information behaviour, and disregards the user segment that does not use the internet for health related issues (Eysenbach, 2011; Zeraatkar and Ahmadi, 2018). In a Finnish context, the limitation of representativeness is of lesser concern, as a majority of the Finnish population use the internet for health related concerns (Tilastokeskus, 2018). However, internet adoption is still limited in some demographic segments, which can lead to the underrepresentation of these segments and their health information behaviour. This is especially evident for discussion forums, where statistics show that participation has been decreasing, and today, less than eight per cent of the population contribute to discussion forums in Finland, compared to twenty-four per cent in 2014 (Tilastokeskus, 2017). Thus, this methodological limitation needs to be accounted for (Mamiya et al., 2017; Guy et al., 2012).

Another key limitation arising from conducting quantitative analysis is the limitation of semantics (Eysenbach, 2011; Mavragani et al., 2018). Without conducting qualitative analysis, or analysing intent or content, there is no way of knowing the underlying reasons for seeking health related information in a search engine, or posting something on social media or in a discussion forum. Therefore, conducting only quantitative analysis makes the purpose of health information behaviour unclear, and we cannot know why an individual engages in health information seeking or sharing, at a specific time (Eysenbach, 2009; 2011). In some cases, we cannot know whether the individual engaging in health information behaviour is suffering from a health related problem or issue, or if he or she is for instance doing research on the subject, or if the topic has been featured in the news recently (Ayers et al., 2013; Eysenbach, 2011). Relating to this, it needs to be noted that the data used in the different studies in this thesis, search engine queries and discussion forum messages, do not present actual illness or symptom incidence. However, as earlier research on health information behaviour has shown, it is reasonable to assume that the reason individuals engage in online health information behaviour for a specific health related issue, is either because they themselves, or people they know, may be experiencing, or have experienced the symptoms or issues in question (Yang et al., 2010; Savolainen, 2011). Moreover, the purpose of this study is not to find out the whys, hows or whats of online health information behaviour, but rather to bring focus on the question when. Therefore, qualitative analysis of message content is left for further research.

Even if demographic factors or user intent remains unknown, the use of de-identified web data is still useful to identify trends and rhythmicity (Eysenbach, 2009). As the aim of this thesis is to study the question of when, and the temporal variations and rhythms of online health information behaviour, the whys, hows or whats and the qualitative analysis needed for those is left for further research.

Another novel issue rising from the use of web data and digital traces are the issues of anonymity, ethics and privacy (McKee, 2013). Infodemiology metrics are, as already mentioned, usually de-identified, and thus the aggregation and analysis of these published, public, open online data sources do not generally raise privacy issues or the approvals of ethics boards. However, these issues still need to be taken into account when utilising the infodemiology approach, especially if the data sources contain personally identifiable information, such as usernames, geographic origin of either posts or visits, or internet protocol address (Eysenbach, 2011). These

issues are also more evident when conducting research of qualitative nature, for instance analysing health related textual content in social media or other digital platforms, where users expect that their privacy issues and anonymity be respected (Eysenbach, 2011). Moreover, online anonymity can be seen as fundamental for the approach, as the internet provides users with protection against stigma or labelling when engaging in behaviour relating to health (Mamiya et al., 2017). This limitation, as relevant as it is, is less important for the Studies included in this thesis, as all analysed data are public, and neither search engine query volumes, or discussion forum messages, and the analysis of them, contain any identifying factors or personal data (McKee, 2013).

6.2. Rhythmicity of online health information behaviour in a discussion forum

The purpose of Study V is to answer research question 2: *Does online health information behaviour in general follow rhythmicity?* The purpose is to analyse *when* Finnish people are engaged in online health information behaviour, and to bring focus and attention to the aspect of time within online health information behaviour. The results in the study show that health information behaviour for different health related issues in the Suomi24 discussion forum follow seasonal and daily variations and patterns. The largest seasonal variations in online health information behaviour on a seasonal and daily timescale can be seen in Table 1, where the ratio is calculated based on the difference between the highest and the lowest relative share of messages.

Table 1. Largest variations in yearly and daily rhythm for the different sub-topics and sub-categories based on the ratio between maximum and minimum relative share.

Largest variation in yearly rhythm	**Ratio**		**Largest variation in daily rhythm**	**Ratio**
Loneliness (*fi. Yksinäisyys*)	2.132096		Loneliness (*fi. Yksinäisyys*)	3.343914
Narcissism (*fi. Narsismi*)	1.945516		Depression (*fi. Masennus*)	2.697666
Weight loss (*fi. Laihdutus*)	1.833515		General mental health (*fi. Yleistä mielenterveydestä*)	2.546249
Eating disorders (*fi. Syömishäiriöt*)	1.818933		Weight loss (*fi. Laihdutus*)	2.516494
Alcohol (*fi. Alkoholi*)	1.811670		Drugs (*fi. Huumeet*)	2.513675
Weight control (*fi. Painonhallinta*)	1.784124		Alternative treatments (*fi. Vaihtoehtoiset hoidot*)	2.486138
Drugs (*fi. Huumeet*)	1.771270		Birth control (*fi. Ehkäisy*)	2.477815
Alternative treatments (*fi. Vaihtoehtoiset hoidot*)	1.738822		Mental health and well-being (*fi. Henkinen hyvinvointi ja mielenterveys*)	2.443714
Womens' health (*fi. Naisten terveys*)	1.682895		Weight control (*fi. Painonhallinta*)	2.191275
Internal medicine (*fi. Sisätaudit*)	1.632709		Eating disorders (*fi. Syömishäiriöt*)	1.968000

The sub-category with the largest difference in seasonal variation is Loneliness. Online health information behaviour in relation to loneliness follow a bimodal curve, with peaks throughout Christmas and the summer. Narcissism, the third most popular sub-category within the same sub-topic as Loneliness, also shows clear variations during the year, with increased activity in early April, followed by smaller monthly recurring activity peaks until late October. A noticeable decrease in online health information behaviour activity is then observable from November until March. These results are visible in Figure 6.

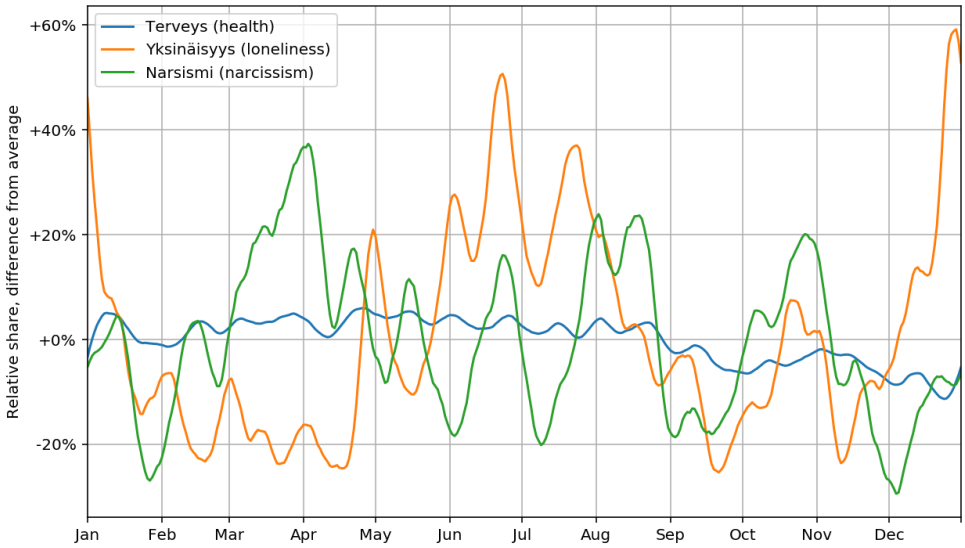

Figure 6. Yearly distribution of messages within the Loneliness (fi. Yksinäisyys) and Narcissism (fi. Narsismi) sub-categories, compared to the overall rate in the Health (fi. Terveys) category.

Online health information behaviour in relation to weight control and weight loss follow a unimodal seasonal pattern, with increased activity during the first months of the year, and a clear peak in late March (Figure 7). Activity then decreases steadily towards the end of the year, with the lowest activity around Christmas. Similar seasonal activity, with a peak in March, is found in health information behaviour related to Eating disorders, a sub-category within the sub-topic Internal Medicine. However, contrary to weight control and loss, the activity for eating disorders increases during the summer and early autumn.

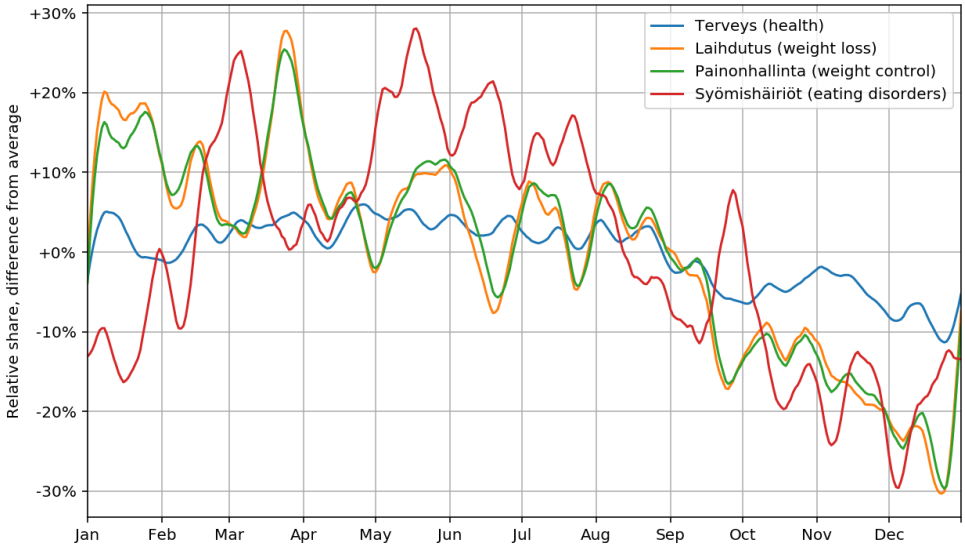

Fig 7. Yearly distribution of messages of the Weight control (fi. Painonhallinta) sub-topic and Weight loss (fi. Laihdutus) and Eating disorders (fi. Syömishäiriöt) sub-categories.

Online health information behaviour in relation to alcohol and drugs also show clear seasonal rhythmicity. As can be seen in Figure 8, discussions in the Alcohol sub-category follows a bimodal pattern, with the highest activity during Christmas, New Year and mid-July. Decreased activity is visible in February, from mid-May to mid-June and from the middle of October to the middle of November. The activity in the Drugs sub-category also follows a bimodal seasonal pattern, but with somewhat different peak times. Drug related health information behaviour activity peaks in spring, more closely mid-April to mid-May, continued by smaller peaks in activity in summer and early autumn. A decrease in activity is then observable from late autumn until early spring.

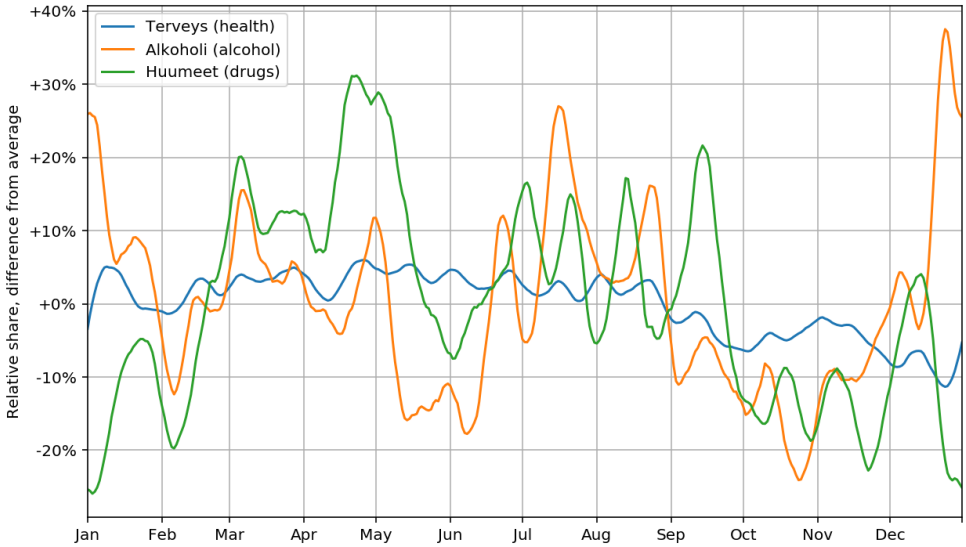

Fig 8. Yearly distribution of messages of the Alcohol (fi. Alkoholi) and Drugs (fi. Huumeet) sub-categories.

Alternative treatments, Women's health and Internal medicine likewise display clear seasonal variations (Figure 9). The yearly distribution of messages in the Alternative treatments sub-category shows the clearly highest peak after mid-April. A second, smaller peak follows in the beginning of June. After a trough in the middle of summer, two peaks are found in September and October. The yearly distribution of messages in Women's health again follows a unimodal curve with a peak season from mid-April to early August and troughs from autumn to early spring. Topics relating to internal medicine on the other hand peak in late winter and spring. The discussions follow a bimodal curve with decreases in activity during summer and winter, especially mid-July and Christmas.

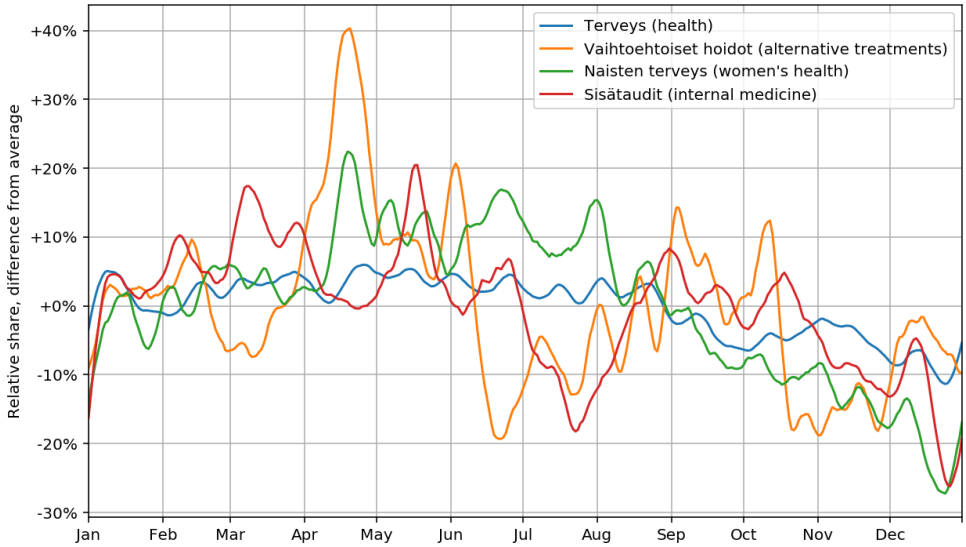

Figure 9. Yearly distribution of messages in the Alternative treatments (fi. Vaihtoehtoiset hoidot), Women's health (fi. Naisten terveys) and Internal medicine (fi. Sisätaudit) sub-categories.

On a daily time-scale, the three categories with the largest variations in online health information behaviour are Loneliness, Depression and General mental health. All three categories follow a similar 24-hour unimodal pattern, with clear night-time peaks and daytime troughs (Figure 10.) The same pattern can be found in the Mental health and wellbeing sub-topic. The differences in relative shares are considerable, even though the Health category in general follow a similar curve. Night-time peaks can also be observed within the sub-categories Drugs and Eating disorders, although the pattern for Eating disorders shows a slightly earlier peak time, between 01.00 and 02.00 at night, as illustrated in Figure 11.

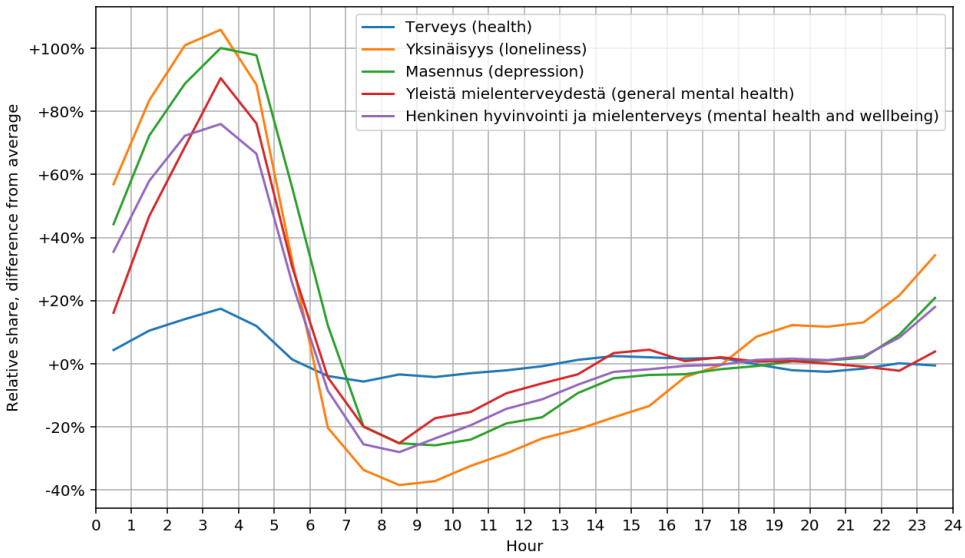

Figure 10. Daily distribution of messages in the Loneliness (fi. Yksinäisyys), Depression (fi. Masennus), General mental health (fi. Yleistä mielenterveydestä) sub-categories and the Mental health and wellbeing (fi. Henkinen hyvinvointi ja mielenterveys) sub-topic.

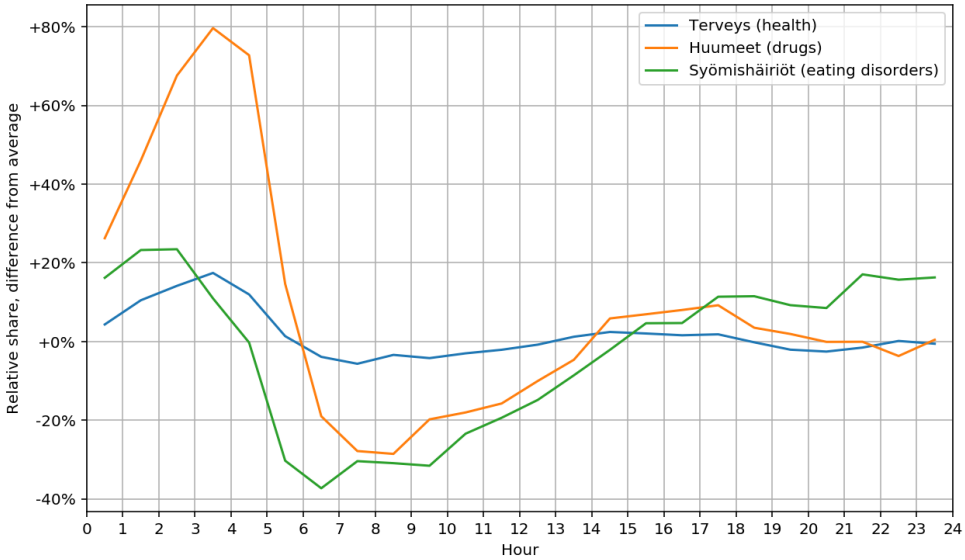

Figure 11. Daily distribution of messages in the Drugs (fi. Huumeet) and Eating disorders (fi. Syömishäiriöt) sub-categories.

The sub-topic Weight control and sub-category Weight loss show an opposite curve (Fig. 12), with daytime peaks, between 06 and 16 followed by night-time troughs, a trend that is opposite to the Health category in general.

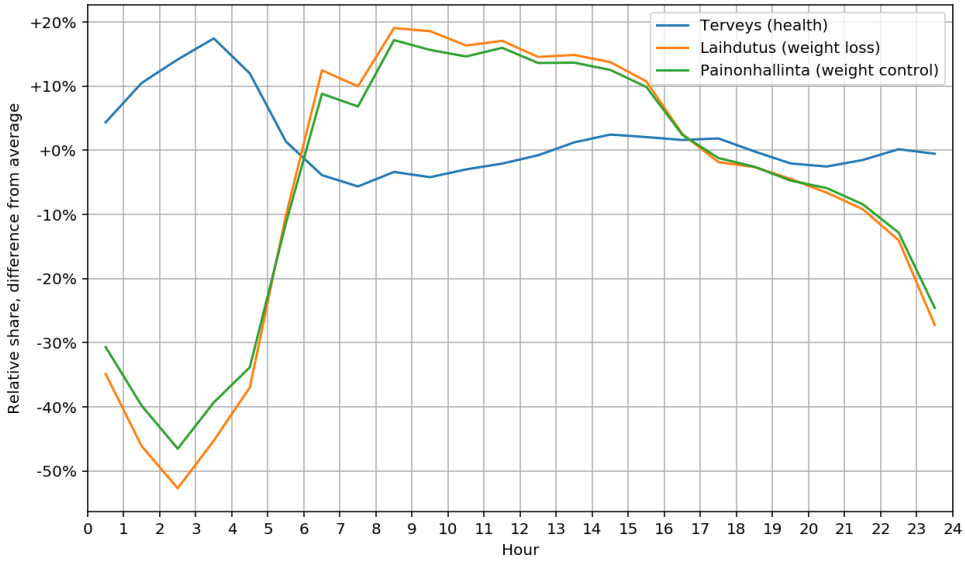

Figure 12. Daily distribution of messages in the Weight loss (fi. Laihdutus) sub-category and Weight control (fi. Painonhallinta) sub-topic.

Similar increased daytime discussion activity can be found in the Alternative treatments sub-category and the Birth control sub-topic, as health information behaviour in relation to Alternative treatments follow a similar daily distribution as for Weight loss and Weight control. However, as can be seen in Figure 13, birth control related online health information behaviour shows peaks slightly later during the day, in the afternoon.

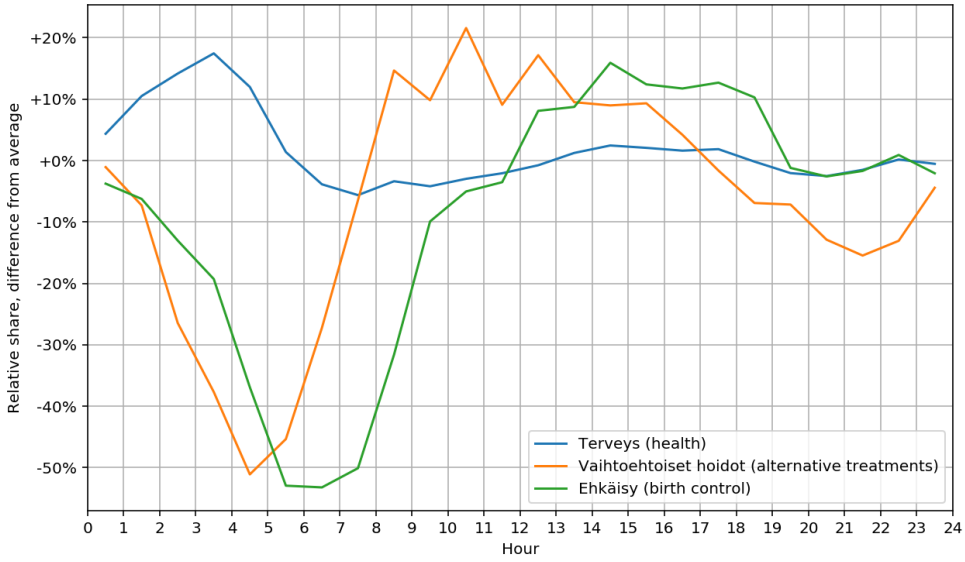

Figure 13. Daily distribution of messages in the Alternative treatments (fi. Vaihtoehtoiset hoidot) sub-category and Birth control (fi. Ehkäisy) sub-topic.

6.3. Temporal variations of depression related health information seeking

Articles I and II, both exploratory infodemiological studies, aim at answering the research question 3: *Does health information seeking for depression related health information show annual, seasonal, weekly or daily rhythmicity?* The results from these studies reveal that health information seeking for depression related information follows a clear rhythmicity on monthly or seasonal, weekly as well as daily time-scales. An annual trend is also visible.

As Study I shows, search engine query data reveals that depression related information seeking follows a bimodal pattern with seasonal variations. There are increased relative search volumes during spring and autumn, and troughs during midwinter and summer (Fig. 14)

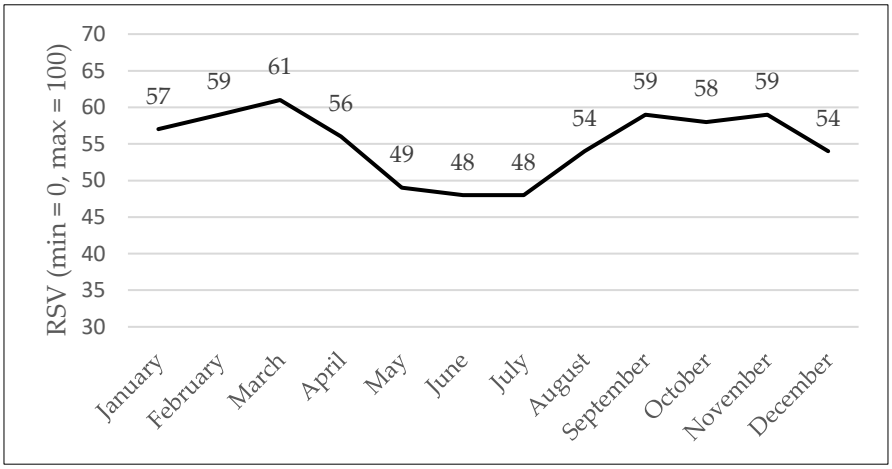

Fig. 14. Combined average monthly RSV for five depression related search queries during the time periods 2004-2015, 2010-2015, and 2015 shows bimodal peaks, in spring and in autumn.

On a weekly, or circaseptan level, depression related health information seeking, measured by relative search engine query volumes, shows highest volumes during Sunday (average RSV 61), and lowest during Friday (average RSV 52). The average relative search volumes for each weekday for the year 2015 can be seen in Table 2.

Table 2. Average relative search volume (RSV) per weekday for all query terms during the year 2015.

Sunday	*61*
Monday	*56*
Wednesday	*56*
Saturday	*56*
Tuesday	*55*
Thursday	*54*
Friday	*52*

Moreover, on an annual level, depression related health information seeking shows a slight increase in relative search volumes (Fig. 15).

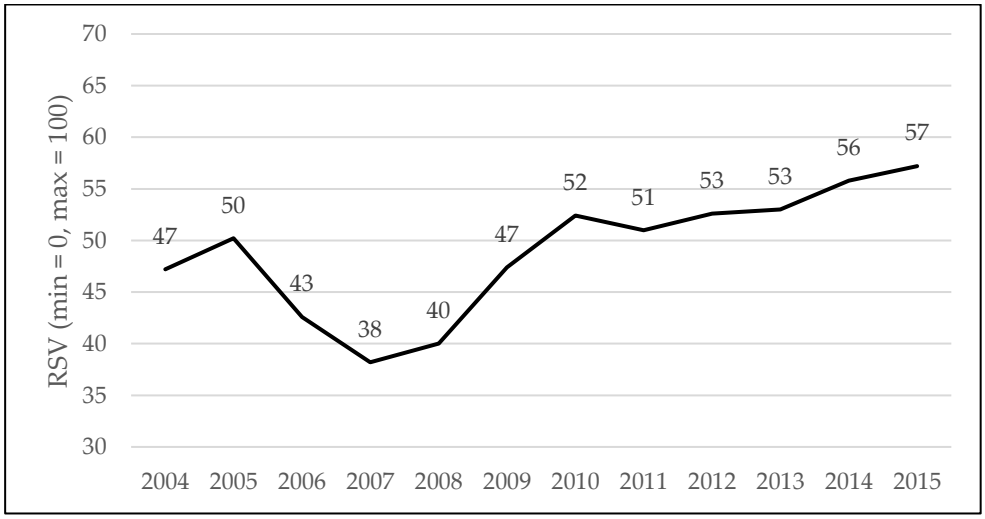

Fig. 15. Combined average RSV for five depression related search queries 2004-2015 shows a slight increase in Internet search query volume since 2007.

The main objective in Study II is to identify daily variations in depression related health information seeking in Finland by analysing search engine data on an hourly level. As in Study I, clear patterns of depression related health information seeking can be observed, in this case on a daily, or diurnal, level. For all search queries included in the study, a daily variability is identifiable, and is most clearly noticeable for the search query *masennus* (depression). On a daily or diurnal level, the relative search volumes for all depression related query terms show a unimodal regular pattern, with clear peaks during the night-time hours between 23.00 and 04.00, and troughs between 05.00 and 22.00 (Fig. 16). The analysis of variance conducted in study II reveals that search interest during night-time hours, between 23.00

and 03.00 is greater than the interest during daytime hours between 08.00 and 16.00, with some variation for different queries. The variations in night-time and daytime depression related health information seeking are also evident in the means of the continuous wavelet transform reconstructed data.

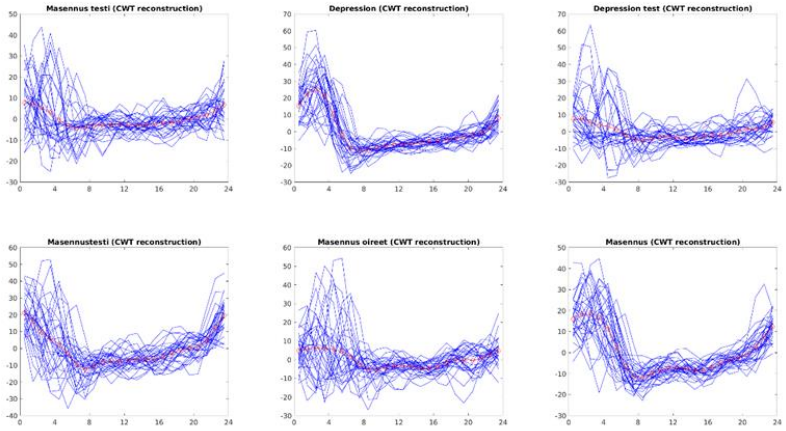

Figure 16. Each 24h period (in blue) of the reconstructed signal superimposed with their arithmetic mean (red), separately for each depression related query for the time period studied.

Variation between peaks and troughs are most evidently visible for the query term *depression*, followed by *masennustesti* and *masennus*. *Depression test*, *masennus testi* and *masennus oireet* showed less significant differences in variation during the day. The analysis of mean information seeking activity between the trichotomised time blocks showed peaks in the first time block (00.00 - 07.59) for all six terms. The mean search volumes then decreased significantly during the second time block (08.00 - 15.59) for the terms *masennus oireet* ($P < .001$), *masennus* ($P = .001$), *depression* ($P = .005$), and for *depression test* ($P = .004$). Thereafter, the mean search volume between the second and the third time block (16.00 - 23.59) rose again for the term *masennus oireet* ($P = .04$). Furthermore, there was a tendency for higher mean search volumes for the terms *masennus* ($P = .14$), *masennustesti* ($P = .07$), and *depression test* ($P = .10$) between the second and third time blocks.

6.4. Exploring temporal variations of depression-related online health information behaviour in a discussion forum.

The objective of Study III (manuscript) is to answer research question 4: *Does depression related online health information behaviour within a large discussion forum show temporal variations on different time-scales?* The study investigates the temporal variations and patterns of depression related online health information behaviour in a Finnish discussion forum, Suomi24. The results show that on a diurnal level, the messages in the depression sub-category of the discussion forum are over-represented by up to +115% during night-time hours, between 00.00 and 05.00, and somewhat under-represented by down to -30% during morning and mid-day, between 07.00 and 12.00. This is exemplified in Figure 17.

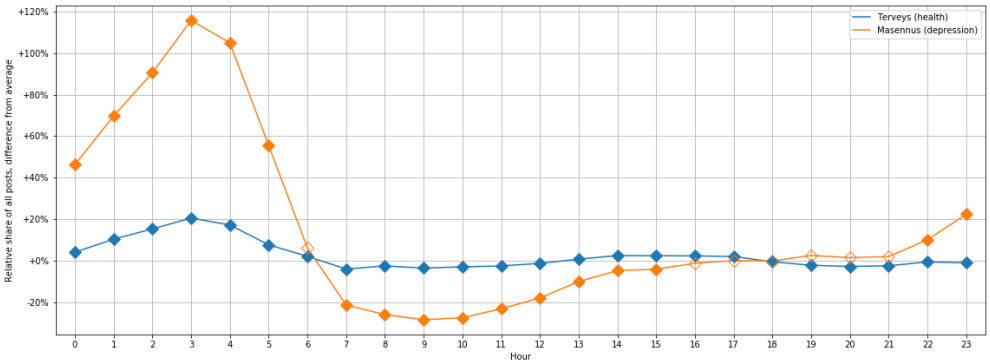

Figure 17. Relative hourly distribution of messages in the health (=terveys) category (blue) and messages in the depression (=masennus) subcategory (orange) compared to all messages. Filled diamonds denote cases where the difference of the relative share from the overall average is statistically significant, and empty diamonds where it is not significant.

Compared to messages in the health category in general the same effect is featured to a considerably weaker extent (+20% and -5%, respectively). Compared to all messages, as well as the messages in the health category, the relative amount of posts in the depression sub-category clearly shows statistically significant peaks for night-time, as indicated in figure 17. On a weekly time-scale, the hourly distribution of messages reiterates during all days of the week (Figure 18). However, clear peaks are visible for Sunday and the beginning of the week (Monday and Tuesday).

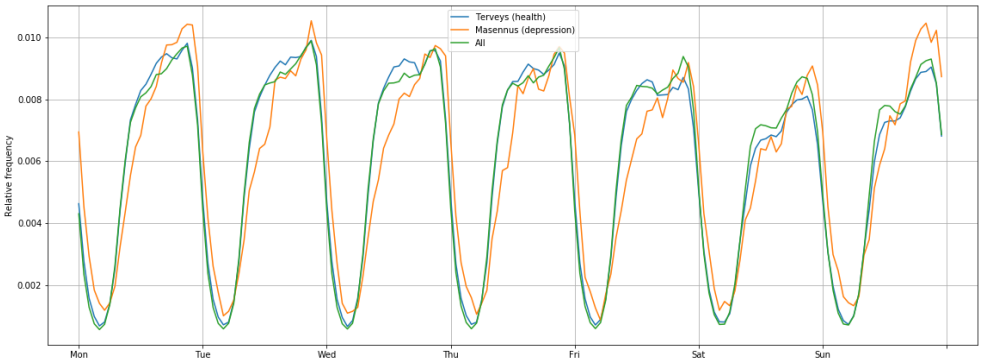

Figure 18. Hourly distribution for each day of week of all messages (green), messages in the health (fi. Terveys) category (blue) and messages in the depression (fi. Masennus) subcategory (orange).

A comparison of the relative number of messages between the health category and the depression sub-category compared to all messages of the forum clearly shows that depression-related messages follow recurring patterns of higher peaks and lower troughs on a weekly, or circaseptan, level. These peaks in the depression sub-category are relatively higher during night-time compared to the health category in general, as illustrated in Figure 19.

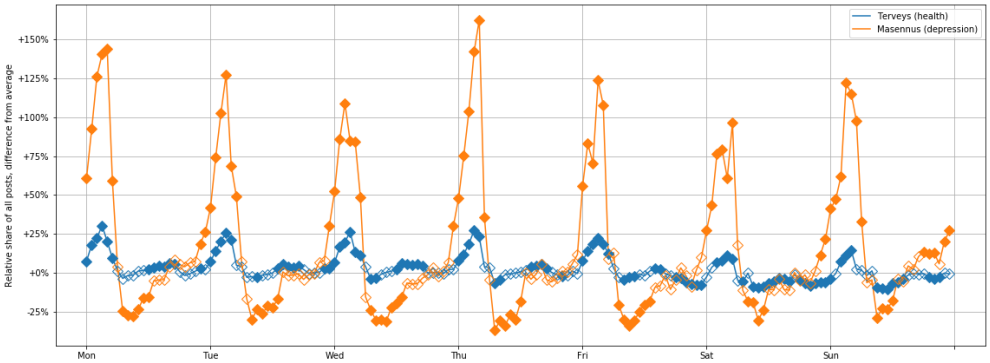

Figure 19. Relative hourly distribution for each day of week of messages in the health (=terveys) category (blue) and messages in the depression (=masennus) subcategory (orange) compared to all messages. Filled diamonds denote cases where the difference of the relative share from the overall average is statistically significant, and empty diamonds where it is not significant.

Depression related online health information behaviour also follows seasonal variations, with clear peaks in spring and autumn. The trend of messages posted in the depression sub-category reveals a bimodal pattern (Figure 20), where the peaks for messages in spring and autumn are significantly higher compared to all messages as well as the messages in the health category. The peaks are most visible in March and May, as well as in

September, November and December. Even a smaller peak is visible in the middle of summer (mid-July). The trough in messages in the depression sub-category in the beginning of the year, mainly January and February, clearly differs from the January peak in the discussion forum as a whole, but also for the messages in the health category. The daily number of messages in the depression sub-category shows a somewhat more irregular pattern for the messages on a circannual level.

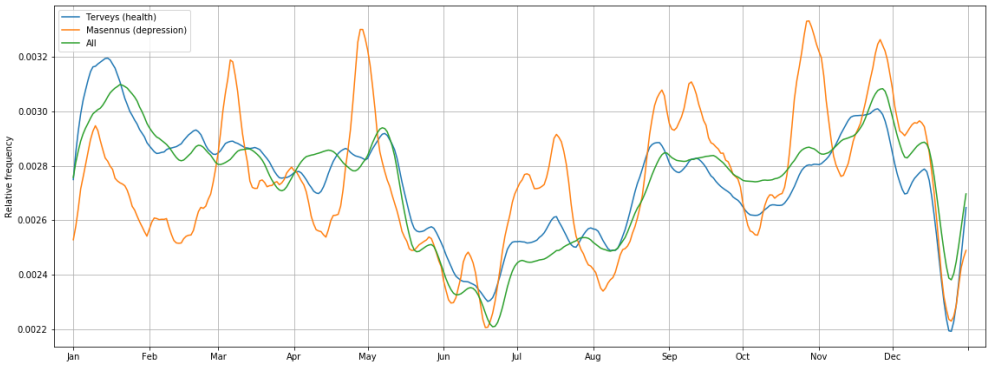

Figure 20. Triangular window rolling mean of the monthly distribution over a year of all messages (green), messages in the health (fi. Terveys) category (blue) and messages in the depression (fi. Masennus) subcategory (orange).

A closer examination and comparison of the daily amount of all messages posted in the forum, messages posted in the health category and messages posted in the depression sub-category throughout the year shows a somewhat more irregular pattern for the messages in the depression sub-category (Figure 21). It also reveals clear peak days in the depression sub-category that stand out throughout the year. The highest seasonal peaks are distributed on the following dates: the 6-7th of March, 27-29th of April, 29-30th of August, 9th of September, 22nd and 27th of October, and 2nd and 3rd of November. A more thorough examination of the depression-related messages on a thread level reveals that replies to threads are significantly over-represented by up to +70 per cent during morning hours between six and eight, compared to top-level posts.

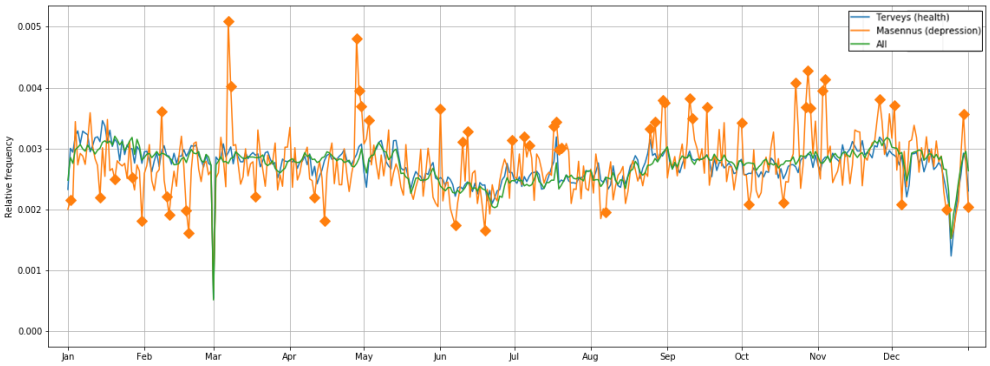

Figure 21. Daily distribution over a year of all messages (green), messages in the health (fi. Terveys) category (blue) and messages in the depression (fi. Masennus) subcategory (orange). Filled diamonds denote cases where the difference of the relative share from the overall average is statistically significant. For clarity, this is only shown for the depression category and when it is significant.

7. Discussion

The aim of this dissertation is to study online health information behaviour and its rhythmicity and temporal nature on different time-scales by utilising the infodemiology approach. The purpose is to find out if online health information behaviour in relation to different health issues is governed by time, rhythmicity and temporal aspects. This over-arching aim was divided into four more specific research questions, which are discussed below. As Study I, II and III, as well as IV and V are of similar nature, they are discussed together.

7.1. Rhythmicity of online health information behaviour in general

Research question 1 was set as follows: *What are the strengths and weaknesses of utilising the infodemiology approach to study the context of time in online health information behaviour?*. This research question was answered by Study IV. As Study IV, together with Study V, also answers research question 2: *Does online health information behaviour in general follow rhythmicity?*, the findings of these studies will be discussed together.

As shown in Study IV and V, online health information behaviour in the Suomi24 discussion forum shows clear and robust rhythmicity, providing knowledge about when people engage in health information behaviour on the internet. This is evident on all the different time-scales studied. The findings in these studies endorse the view of health as a dynamic process, as well as the multidimensional relationship between time and health (Eriksson, 1993, pp. 54-56; Henly et al., 2011; Mogilner et al., 2018; Reinberg et al., 2017; Roenneberg, 2012, p. 2; Roy et al., 2004; Svalastog et al., 2017). The findings also deliver evidence for rhythmicity and temporal variations of health behaviours that are left outside traditional epidemiological research, and can thus complement and provide a more holistic picture of health. However, this also presents challenges in relating the findings to previous research, as only a few previous studies have investigated the rhythmicity and temporal aspects of online health information behaviour in relation to some of the health issues and phenomena recognised in Studies IV and V. Therefore, the findings of Study IV and V can serve as initial steps to investigating these temporal variations and phenomena more thoroughly.

Some of the rhythmicity and temporal variations as well as patterns where online health information behaviour is more active during

beginnings, such as the beginning of the day, month or year, could be related to what has been dubbed the "fresh start effect" which indicates aspirational behaviour for a fresh, new and healthy start (Dai et al., 2014; Gabarron et al., 2015). This effect is most obvious for the circaseptan and circannual time-scales for health information behaviour, and especially in relation to weight control, and suggests that the beginning of the workweek, as well as the beginning of a new year are times when people pursue healthier behaviours (Dai et al., 2014; Gabarron et al., 2015).

The findings also show that on a circadian, or daily level, there seems to be a clear difference between the rhythmicity for mental health related issues compared to physical health issues. This is exemplified by the fact that for instance mental health related online health information behaviour activity is night-time oriented, while activity for issues such as weight control, birth control or alternative treatments is highest during daytime. This mental health related night-time peak is, especially in the case of depression related health information behaviour, in line with the findings in Study II and III of this thesis. Substance abuse issues, in this case alcohol, show clear seasonal variations which is in line with previous findings, as Mustonen et al. (2012) and Poikolainen et al. (2002) found that alcohol consumption, much like the alcohol related discussions, show peaks during Christmas, in spring and in the middle of summer. Also the findings for loneliness related health information, that presents the largest variations in rhythmicity on both a seasonal as well as daily level, are in line with previous research where loneliness was found to be the most common reason for stressor of patients evaluated in a psychiatric emergency service during Christmas season (Velamoor et al., 1992). Feelings of depression are also associated with loneliness, which is also visible in the night-time activity for both loneliness and depression. On a seasonal level, loneliness follows a similar pattern as that which has been found for suicide, for which loneliness is a primary cause (Sansone and Sansone, 2011). In Finland, peaks in suicide statistics peak in late spring and during summer (Hakko et al., 1998; Holopainen et al., 2013), which is similar to the online health information behaviour for loneliness related discussions.

7.2. Rhythmicity and temporal variations of depression related online health information behaviour

Research question 3: *Does online health information seeking for depression related health information show annual, seasonal, weekly or daily rhythmicity?* and 4: *Does depression related online health information behaviour within a large*

discussion forum show temporal variations on different time-scales? are as noted of similar nature. Studies I, II and III all aim at answering these two research questions. Study I and II by utilising health information seeking data from Google Trends, and Study III by analysing the depression category of the Suomi24 discussion forum.

The findings of Study I, II and III are of similar nature. The results in all three studies show that depression related online health information behaviour follows rhythmicity and temporal variations and patterns on daily, weekly and seasonal level.

There are clear and recurring peaks during autumn and spring, while the summer and winter months present troughs. This seasonal, bimodal pattern is visible in both search engine query volumes as well as in discussion forum activity. These findings show similarities with previous research on health information seeking, as Ayers et al. (2013), Yang et al. (2010), and Nimrod (2012) all identified similar seasonal patterns when comparing the northern and southern hemispheres with each other. However, their findings are somewhat different, as they identified the highest online health information behaviour activity during the darkest months of the year. However, Nimrods (2012) data were limited to one year only, and Ayers et al. (2013) as well as Yang et al. (2010) compared large geographical areas by a few keywords only. The findings in Study I, II and II, with peaks in spring and autumn, are also in line with the seasonality in diagnosed depression, hospital admissions and suicide rates in the Scandinavian countries and the northern hemisphere (Christensen and Dowrick, 1983; Hakko et al., 1998; Stordal et al., 2008).

The highest relative search volume as well as discussion forum activity for the different weekdays is in both Study I and III found during Sunday and Monday. These findings support previous research, stating that Sundays and Mondays are days of lower mood or when depression symptoms are more prevalent (Akay and Martinsson, 2014; Brådvik, 2002; Tumen and Zeydanli, 2014). Sundays are also, according to Brådvik (2002), the weekday when most suicides are committed. Sundays, as well as evening or night-time, are for most people lower in activity and less constrained by external demands, compared to days and hours filled with work. This, as well as loneliness, break from work, lack of energy, or other interactions between light/darkness and the nervous system, may trigger inner tension and worrying as well as negative thinking, and lead to mood disorders (Brådvik, 2002; De Choudhury, Counts and Horvitz, 2013; Rusting and Larsen, 1998).

The notion of structured versus unstructured time could also be the reason for the diurnal variations of depression related online health information behaviour, as proven by the findings in all three studies (I, II, and III). All three studies present findings with similar and significantly higher activity in behaviour during night-time. This night-time peak for depression has also been identified in Twitter posts by Chen et al. (2018) and De Choudhury, Counts and Horvitz (2013). These findings of night-time rhythmicity also supports the notion of the evening-worse diurnal mood variation for depression (Rusting and Larsen, 1998) as well as the notion of poor or disturbed sleep quality and the causal relationship between insomnia and depression, in individuals suffering from symptoms of depression (Marcus et al., 2012; Tsuno et al., 2005). The higher response rate for depression related discussions during early-morning hours in the Suomi24 discussion forum could thus be an indicator of disturbed sleep patterns.

In Study I, an interesting finding is that the relative search volume for depression related health information seeking is slightly increasing, which is contradictory with the fact that the incidence of diagnosed depression has decreased during the time-period 2004 - 2015 in Finland (Wahlbäck et al., 2015). A possible explanation for this is that people try to treat health problems themselves, especially in stigmatising health issues, by seeking information about their symptoms and possible treatments (McGowan, 2002; Oliphant, 2010, pp. 1-2; Siepmann, 2008). Another explanation that Oliphant (2010, p. 2) mentions, is that the popularity of alternative and complementary treatments are getting a stronger foothold of depression treatment.

7.3. Theoretical implications

Rhythmicity is fundamental to nature. We breathe, sleep, eat and use energy in a rhythmic way. This rhythmicity is, however, not limited to our physiology and physiological functions, but also permeates our life in general, including behavioural issues, including issues related to health and illness (Adam, 1990, pp. 70-76). This confirms what Adam (1990, p. 72), McKenzie and Davis (2002) as well as Savolainen (2006; 2018) state, that time is an important aspect in information science, and social science in general. In social sciences, time should therefore, as Adam (1990, pp. 72-73) proposes, be considered in a rhythmic fashion, which entails cycles, structures and patterns with variation and repetition. However, as previously stated, time

as an aspect for information behaviour has previously been overlooked, and this study is one of the first to focus on time and online health information behaviour.

This thesis contributes to theory on several levels. Primarily, it emphasises the temporal aspects and rhythmical nature of online health information behaviour, and that these aspects need to gain more attention within information science. A more thorough exploration of the context, in this case the temporal structures, in which information behaviour happens can have significant theoretical contributions. Given the abundance of theories and models in the field of information science, introducing new ones may, as Wilson (1999) points out, seem unhelpful. However, adding new aspects and layers to existing theories can be beneficial, and therefore the temporal aspects should be implemented in various theories of information behaviour. The findings in this thesis highlights the changing and dynamic context of health and illness, and that people are different during different times. Turning the focus to the temporal aspects of health information behaviour brings new and relevant evidence of rhythmicity for different health related phenomena, and allows us to understand the relationship between time and information behaviour. Dervin's Sense-making approach already touches upon time as a situational factor, and emphasises the individual as changeable in time-space. The results in this thesis support the notion that time is a situational factor of sense-making, and that people are different at different times. The assumption that natural time and the temporal structures that govern life on earth, affects online health information behaviour, and that this is visible through individuals sense-making, especially in the context of situation in which sense-making happens, can be confirmed based on studies in this thesis. Thus, explicitly adding the aspect of time to the various situational variables in Dervin's sense-making, can help predict, understand and explain why certain health issues rise at certain times. Breaking down the context of situation, and investigating the temporal aspects of it, allows examining information behaviour from a much more detailed perspective. This in turn can help complement and strengthen theories on health, illness and health behaviour, as well as contribute to public health related matters. Sense-making and the aspect of time can easily be studied with novel data sources, such as search engine or discussion forum data. This kind of temporal sense-making, especially from the perspective of situation, confirms the theory of collective sense-making, and provides methodological aids for conceptualising information behaviour as temporally patterned (Dervin, 2005, p. 28).

Adding the aspect of time to models of information behaviour could also be beneficial, as models are useful in pre-theory stages, and provide aids for understanding a phenomenon (Bates, 2005, p 3). Wilson's various models of information behaviour all place the user or person in context, but this context is either the user's life world, environment or role (Wilson, 2005, pp. 32-34). Adding a temporal aspect to this context could be highly beneficial, as users or persons would be seen as less static from a temporal perspective, emphasising the influence of time on individuals.

A growing number of people engage in online health information behaviour, providing vast amounts of data and metrics, which can help advance our understanding of the complex processes underlying health and illness (Bath, 2004). Analysing these behavioural data and applying the infodemiology approach not only reveals insights into the temporal nature of information and communication patterns on the internet, it also complements research and data on public health (Eysenbach, 2009). Applying not only time as an aspect to theory, but also the advantages that the infodemiology methods and metrics offer, can provide researchers in other fields and contexts with new insights into the interaction of time and different behaviours.

7.4. Practical implications

Despite some limitations in utilising the infodemiology approach to study online health information behaviour and its rhythmicity, the implications and strengths that a deeper understanding of the rhythmicity and the temporal aspects of online health information behaviour provides outweigh the limitations.

The findings in the studies of this thesis, as well as knowledge about the rhythmicity and temporal nature of online health information behaviour, can aid service design and design of novel social computing systems in creating optimal health information interventions (Huang, 2017). Knowledge of different health related rhythms could also be used to help recognise and target populations or individuals at higher risk of unhealthy behaviours to encourage maintenance of healthy behaviours at trigger times and situations (Huang, 2017). This is highly relevant and challenging in what Erikainen et al. (2019) call the *digital era*, which is characterised by unprecedented levels of interconnectivity and information access at an increasing speed, spread and dynamism in relation to knowledge and information turnover. The need for contextual understanding of health behaviours is also critical towards the development of what Huang (2017)

calls "just-in-time" interventions, especially for behaviours like smoking and drinking to support cessation or impulse control. Moreover, tools and services for enabling self-care and self-management have been shown to be most successful when their design is grounded in behavioural theories (Mamykina et al., 2015). Providing help for experienced health issues and threats at times when individuals are focusing their attention on health threats, engaging in health related behaviour and in need of, as well as receptive to, support can aid early and effective intervention, which in turn has been shown to be beneficial for positive health outcomes (Lambert and Loiselle, 2007; Wilson, 1999). The advantages of guiding individuals to optimal sources before symptoms get worse is especially evident for mental health issues, namely depression, where identifying early symptoms, such as the evening-worse patterns that is associated with milder depressive symptoms, can have an impact on treatment response (Gordijn et al., 1994; Hasler, 2013, pp. 612-613). Thus, for instance health promotion campaigns and precautionary actions could be made more cost effective by targeting the population at the right time with right and relevant information. This can also have substantial impact on health interventions (Ayers et al., 2014). Discovering and understanding rhythmicity and temporal aspects of health information behaviour therefore has strong potential for behavioural gains, as well as improving both public and individual health.

Recent developments in what has been called the post-truth era, with a growing abundance of misinformation and "fake news" (Lewandowsky et al., 2017; Veinot, 2017) also provides another implication for infodemiology research. As Eysenbach (2011) states, the arrow of causation in information and communication patterns is bidirectional, therefore, the infodemiology approach can aid in creating strategies to ensure that individuals, especially those who do not seek treatment or cannot be reached by traditional means, are provided with evidence-based health information, aid and interventions. The amount of misinformation today is significant, and this can have a negative impact on individuals, public health and society, as for example with anti-vaccination campaigns (Eysenbach, 2011; Lewandowsky et al., 2017).

As noted earlier, longitudinal data have rarely been utilised in online health information research (Anker et al., 2011). This prohibits identifying rhythms and temporal patterns or variations, as the methods usually employed only gather data that provides snapshots of situations in a specific moment in time (Giddens, 1979, p. 202). However, analysing large longitudinal data sets such as the Suomi24 discussion forum data, or search

engine data for an extended period of time, allows the examination of online health information behaviour over time. This, again, can reveal not only rhythmicity and temporal patterns of this behaviour on different time-scales, but even a more holistic picture of how individual interact and behave in relation to their health. Studying and analysing the rising amount of health related digital traces that are left online can even reveal completely new insights for different health trends, patterns and correlations. This is especially relevant for preclinical cases that fall outside the scope of traditional healthcare, or for stigmatising health issues, such as mental health problems, where individuals are less likely to seek professional help.

Another implication that studying rhythmicity and temporal aspects brings, is the possibility to predict and anticipate future change in health, as well as discontinuities and knowledge gaps in relation to health. Even if this is somewhat limited because of the lack of demographic data, identifying clear and robust patterns and rhythms of health information behaviour can facilitate precautionary efforts within health promotion and prevention (Adam, 1990, p. 75; Ayers et al., 2014).

For methodology, using infodemiology metrics provides a way of avoiding some methodological biases, as it for instance avoids the interviewer bias, whereby what is said during the interviews is influenced by the researcher (McKee, 2013). Much of infodemiology metrics are generated without the involvement or direct contact between the subjects and the researcher. Some data, like time stamps for messages are even generated automatically, making data aggregation and analysis rapid, even enabling real-time analysis using the right tools (Eysenbach, 2011).

As can be seen, the field of information science, and health information behaviour research, has a key role within infodemiology research, by bringing evidence for health related issues that more traditional research methods, such as questionnaires and surveys, have not been able to capture. Turning the focus to time and studying *when* health information behaviour happens can therefore have far-reaching consequences, also for information science.

7.5. Limitations

Limitations and critical assessment of the infodemiology approach in general, including methods and metrics, have already been presented in chapter 6.1.1. Apart from these general infodemiology limitations, there are a few limitations specific to the studies included in this thesis, mainly relating to reliability and validity as well as data sources that need to be

noted. As already mentioned, participants in internet based studies consist of the segment of the population that uses the internet, in this case for health information behaviour. In all studies included in this thesis, these are restricted to users of a specific search engine, Google, and a specific discussion forum, Suomi24.

For the studies in this thesis that utilise the Suomi24 discussion forum data, a further limitation is that the analyses are limited to one discussion forum. There are other active discussion forums, similar to Suomi24 in Finland, with various topics and objectives, even specifically related to different health and illness issues. In addition to this, some social media sites, like Facebook, Twitter and YouTube, have also become more and more essential for health information behaviour (Zhao and Zhang, 2017). However, from a data access viewpoint, Suomi24 provides all discussion forum data as open data, which allows health information behaviour to be studied longitudinally, dating back to times before social networking sites like Facebook. Furthermore, as already mentioned, the Suomi24 discussion forum is still the largest and most popular discussion forum in Finland (FIAM, 2018). Another limiting factor is that analysis is limited to posted messages only, while all online forum readers who do not engage in message posting fall outside the scope of analysis. However, as most site visitors to the Suomi24 discussion forum arrive from Google searches (Lagus et al., 2016), studies I and II could very well be seen as complementing this gap.

From a selection bias standpoint, the use of Google Trends data does not present a similar problem, as Google is clearly the most used search engine in Finland, with a market share of ninety-seven percent in 2019 (StatCounter, 2019). A key limitation with the use of Google Trends data and database, is the lack of detailed methodological information on how Google Trends generates search data and the specific algorithms utilised to analyse search queries (Nuti et al., 2014).

7.6. Further research

Research conducted for this thesis has merely answered a few limited questions on the temporal aspects of health information behaviour. As most temporal aspects of health information behaviour, or information behaviour in general have been overlooked, further research on this topic is not difficult to identify. In general, the data and methods utilised in the research for this thesis could be adopted to study many different phenomena within information science, also outside the context of health.

All studies in this thesis concentrate on analysing online health information behaviour on previous happenings. Applying and combining the metrics utilised in this study to infoveillance research, or real-time surveillance, could provide an extensive index on different health related matters in Finland to supplement public health surveillance in order to identify early warning signals. Identifying clear temporal patterns of health information behaviour also facilitates the prediction of the future, and allowing precautionary efforts (Adam, 1990, pp. 74-75). This could be particularly fruitful for stigmatising health issues, such as depression, or other matters related to mental health, where help seeking from professionals is low, and early or timely optimal intervention has a been shown to be of relevance in treatment response (Ayers et al., 2014b; Gordijn et al., 1994).

When analysing search engine data, more variables, such as more and different search queries, as well as varying geographical locations, could be added to the scope of analysis. This could not only give more comprehensive results, but would also allow comparing similarities in temporal trends and variations between different health related issues for different geographical locations.

A potential pathway for future studies utilising the Suomi24 data could also be utilising novel language processing, sentiment analysis or machine learning, to analyse the language or the emotional tone in textual messages, discover themes or classify topics in the messages, obtain indicators on changes over time, or detect emotions such as fear, or expressed peer-support (Ahmed et al., 2018; Biyani et al., 2014; Dredze and Paul, 2014; Eysenbach, 2009). This could help identify, measure and detect various health related symptoms and issues expressed in messages on discussion forums. Such research could also help construct early detection systems for various social media services and platforms on the internet.

One very specific, but particularly interesting, phenomenon that should be examined more meticulously is the anomaly of higher rate of replies for depression related messages in the Suomi24 depression sub-category, as identified in Study III. Here qualitative research methods have the potential to find an explanation and help shed some light on this interesting phenomenon. Identifying similar anomalies in the discussion forum in general would also be interesting.

The Suomi24 dataset is overall extremely extensive and longitudinal, and contains other topics besides health that could be useful and relevant to analyse from a time perspective. Broadening the scope of research to other

topics and subjects discussed in the Suomi24 forum could also provide useful insights into the temporal variations of information behaviour on topics unrelated to health. More generally, conducting content analysis on the messages posted in the Suomi24 discussion forum would be useful, in order to investigate message content and how content varies and differs during different timescales. A categorisation of messages into, for instance, information needs and information sharing could also provide insights into the rhythmicity and temporal variations of more specific online health information behaviour. Applying more variables, such as geolocation in the case of Google Trends data, thread title or view count in the case of Suomi24 data, to the scope of analysis could also yield new insights that have previously been unidentified.

Studying information prevalence, or what is published on the internet, versus disease incidence for some of the health issues could also be useful to illustrate or point out information deficits for the topics in question. Conducting qualitative research on, for instance, different social media platforms could reveal how and what users communicate about different health related issues, such as disease outbreaks (Ahmed et al., 2019). A more extensive coverage or organising of high quality information would be especially useful for emerging health related issues, which have a high disease burden, and where early intervention could have substantial impact (Ahmed et al., 2019; Eysenbach, 2009).

As more and more health related data are being produced on a myriad of digital platforms, a future opportunity could lie in utilising all digital traces to establish an effective, reliable health prediction index for different health related issues and emergencies.

8. Conclusions

Rhythms and temporal variations govern all aspects of life and behaviours within. As has been shown, this is also the case for online health information behaviour. The clear and robust rhythmical patterns and variations relating to different health related issues identifiable in both search engine as well as discussion forum data show that neither health nor the situations within sense-making are static processes. Individuals face health related knowledge gaps at different times, but when studied from a collective perspective with the infodemiology approach and metrics, patterns of rhythmicity, as well as temporal variations, emerge. As stated in the beginning, rhythmicity makes life and behaviours predictable, and provides a window to the future. These patterns and variations can have practical implications for many disciplines, and the approach itself can have far-reaching consequences, especially in an era where individuals are expected, and provided by an abundance of digital means, to take greater responsibility for their health. As the barriers to study rhythmicity information behaviour on different time-scales have been reduced, the temporal aspects of all information behaviour, and the focus on *when* should be taken into consideration. Researchers who fail to attend to the rhythmical nature, or temporal issues of health information behaviour run the risk of missing significant contextual dimensions on both individual and collective levels. The field of information science, and health information behaviour research in particular, have a significant role in infodemiology, by providing both behavioural data and evidence for health related issues that traditional research has been unable to capture, or has overlooked. Focusing on time and studying when online health information behaviour happens can thus have important value for a deeper understanding of health and its behavioural aspects.

References

Abedi, V., Mbaye, M., Tsivgoulis, G., Male, S., Goyal, N., Alexandrov, A. V. and Zand, R. (2015). Internet-based information-seeking behavior for transient ischemic attack. *International Journal of Stroke, 10*(8), 1212-1216.

Adam, B. (1990). *Time and Social Theory*. Polity Press, Cambridge.

Ahmed, W., Bath, P. A. and Demartini, G. (2017). Using Twitter as a Data Source: An Overview of Ethical, Legal, and Methodological Challenges. In *The Ethics of Online Research* (pp. 79-107). Emerald Publishing Limited.

Ahmed, W., Bath, P. A., Sbaffi, L. and Demartini, G. (2018, July). Moral panic through the lens of Twitter: An analysis of infectious disease outbreaks. In *Proceedings of the 9th International Conference on Social Media and Society* (pp. 217-221). ACM.

Ahmed, W., Bath, P. A., Sbaffi, L. and Demartini, G. (2019). Novel insights into views towards H1N1 during the 2009 Pandemic: a thematic analysis of Twitter data. *Health Information & Libraries Journal, 36*(1), 60-72.

Akay, A. and Martinsson, P. (2009). Sundays are blue: Aren't they? The day-of-the-week effect on subjective well-being and socio-economic status. *IZA Discussion Paper No. 4563*.

Aller, (2016). Suomi24 Media Card. Available at https://www.aller.fi/wp-content/uploads/2016/01/Suomi24-Tuotekortti-1_2016.pdf (accessed 4 January 2019).

Almanea, A., Bath, P. and Sbaffi, L. (2018). The Role of Online Support Groups in Empowering People with Type 2 Diabetes. *iConference 2018 Proceedings*.

Althouse, B. M., Allem, J. P., Childers, M. A., Dredze, M. and Ayers, J. W. (2014). Population health concerns during the United States' great recession. *American Journal of Preventive Medicine, 46*(2), 166-170.

Amzat, J., and Razum, O. (2014). *Medical sociology in Africa*. Springer.

Anker, A. E, Reinhart, M. R. and Feeley, T. H. (2011). Health information seeking: A review of measures and methods. *Patient Education and Counselling, 82*(3), 346-354.

Antonovsky, A. (1996). The salutogenic model as a theory to guide health promotion. *Health Promotion International, 11*(1), 11-18.

Antypa, N., Vogelzangs, N., Meesters, Y., Schoevers, R. and Penninx, B. W. (2016). Chronotype associations with depression and anxiety disorders in a large cohort study. *Depression and Anxiety, 33*(1), 75-83.

Arendt, F. and Scherr, S. (2017). Optimizing online suicide prevention: A search engine-based tailored approach. *Health Communication, 32*(11), 1403-1408.

Aromaa, E., Tolvanen, A., Tuulari, J., and Wahlbeck, K. (2011). Personal stigma and use of mental health services among people with depression in a general population in Finland. *BMC Psychiatry, 11(1)*, 52.

Ayers, J. W., Althouse, B. M., Allem, J. P., Rosenquist, J. N. and Ford, D. E. (2013). Seasonality in seeking mental health information on Google. *American Journal of Preventive Medicine, 44*(5), 520-525.

Ayers, J. W., Althouse, B. M., Johnson, M., Dredze, M. and Cohen, J. E. (2014b). What's the healthiest day?: Circaseptan (weekly) rhythms in healthy considerations. *American Journal of Preventive Medicine, 47*(1), 73-76.

Ayers, J. W., Althouse, B. M., Johnson, M. and Cohen, J. E. (2014). Circaseptan (weekly) rhythms in smoking cessation considerations. *JAMA Internal Medicine, 174*(1), 146-148.

Ayers, J. W., Althouse, B. M. and Dredze, M. (2014a). Could behavioral medicine lead the web data revolution?. *JAMA, 311*(14), 1399-1400.

Basnet, S., Merikanto, I., Lahti, T., Männistö, S., Laatikainen, T., Vartiainen, E. and Partonen, T. (2016). Seasonal variations in mood and behavior associate with common chronic diseases and symptoms in a population-based study. *Psychiatry Research, 238*, 181-188.

Bates, M.J. (2005) An introduction to metatheories, theories and models. In Fisher, K.E., Erdelez S. and McKechnie L. (eds.), *Theories of Information Behavior* (pp. 25-30). Medford, New Jersey: Information Today, Inc.

Bath, P. A. (2004). Data mining in health and medical information. *Annual Review of Information Science and Technology, 38*(1), 331-369.

Bender, J. B. and Wellbery, D. E. (1991). *Chronotypes: The Construction of Time.* Stanford, CA: Stanford University Press.

Bexelius, C. (2009). *E-epidemiology: Adapting epidemiological methods for the 21st century*. Department of Medical Epidemiology and Biostatistics. Karolinska Institutet, Stockholm, Sweden.

Bidmon, S. and Terlutter, R. (2015). Gender differences in searching for health information on the internet and the virtual patient-physician relationship in Germany: exploratory results on how men and women differ and why. *Journal of Medical Internet Research, 17*(6), e156.

Biyani, P., Caragea, C., Mitra, P. and Yen, J. (2014). Identifying emotional and informational support in online health communities. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, 827-836.

Blaxter, M. (1990). *Health and lifestyles*. NewYork: Routledge.

Blaxter, M. (2010). *Health*. Cambridge: Polity.

Brady, E., Segar, J. and Sanders, C. (2016). "You get to know the people and whether they're talking sense or not": Negotiating trust on health-related forums. *Social Science & Medicine, 162*, 151-157.

Bragazzi, N. L., Alicino, C., Trucchi, C., Paganino, C., Barberis, I., Martini, M., and Icardi, G. (2017). Global reaction to the recent outbreaks of Zika virus: Insights from a Big Data analysis. *PloS one, 12*(9), e0185263.

Brigo, F. and Trinka, E. (2015). Google search behavior for status epilepticus. *Epilepsy & Behavior, 49*, 146-149.

Boulos, K., Sanfilippo, A., Corley, C. and Wheeler, S. (2010). Social Web mining and exploitation for serious applications; Technosocial Predictive Analytics and related technologies for public health, environmental and national security surveillance. *Computer Methods Programs Biomed, 100*(1), 16-23.

Bryman, A. (2016). *Social research methods.* Oxford University Press, Oxford.

Brådvik, L. (2002). The occurrence of suicide in severe depression related to the months of the year and the days of the week. *European Archives of Psychiatry and Clinical Neuroscience, 252*(1), 28-32.

Card, A. J. (2017). Moving Beyond the WHO Definition of Health: A New Perspective for an Aging World and

the Emerging Era of Value-Based Care. *World Medical & Health Policy, 9*(1), 127-137.

Carneiro, H. A. and Mylonakis, E. (2009). Google trends: a web-based tool for real-time surveillance of disease outbreaks. *Clinical Infectious Diseases, 49*(10), 1557-1564.

Case, Donald. (2012). *Looking for information: A survey of research on information seeking, needs, and behavior*. Emerald Group Publishing.

Castoradis, C. 1991. Time and creation. In Bender and Wellbery (eds.) *Chronotypes: The Construction of Time*. Stanford, CA: Stanford University Press.

Chan, J. K., Farrer, L. M., Gulliver, A., Bennett, K. and Griffiths, K. M. (2016). University students' views on the perceived benefits and drawbacks of seeking help for mental health problems on the Internet: a qualitative study. *JMIR Human Factors, 3*(1).

Chan, J. K., Farrer, L. M., Gulliver, A., Bennett, K. and Griffiths, K. M. (2016). University students' views on the perceived benefits and drawbacks of seeking help for mental health problems on the Internet: a qualitative study. *JMIR Human Factors, 3*(1), e3.

Charmaz, K. (1991). *Good Days, Bad Days: the Self in Chronic Illness and Time*. Rutgers University Press, New Brunswick, NJ.

Chen, X., Sykora, M. D., Jackson, T. W. and Elayan, S. (2018). What about Mood Swings: Identifying Depression on Twitter with Temporal Measures of Emotions. In *Companion of the The Web Conference 2018 on The Web Conference 2018* (pp. 1653-1660). International World Wide Web Conferences Steering Committee.

Chew, C. and Eysenbach, G. (2010). Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PloS one, 5*(11), e14118.

Christensen, R. and Dowrick, P. W. (1983). Myths of mid-winter depression. *Community Mental Health Journal, 19*(3), 177-186.

Dai, H., Milkman, K. L. and Riis, J. (2014). The fresh start effect: temporal landmarks motivate aspirational behavior. *Management Science, 60*(10), 2563-2582.

De Choudhury, M., Ringel Morris, M., and White, R. (2014). Seeking and Sharing Health Information Online: Comparing Search Engines and Social Media. *CHI*, 2014.

De Choudhury, M., Counts, S. and Horvitz, E. (2013, May). Social media as a measurement tool of depression in populations. In *Proceedings of the 5th Annual ACM Web Science Conference*, 47-56.

Dervin, B. (1976). Strategies for dealing with human information needs; information or communication?. *Journal of Broadcasting & Electronic Media, 20*(3), 323-333.

Dervin, B. (1992). From the mind's eye of the user: The sense-making qualitative-quantitative methodology. In Glazier, J. D. and Powell, R. P. (eds.) *Qualitative research in information management* (pp. 61-84). Engelwood, Colorado: Libraries Unlimited.

Dervin, B. (1998) Sense-making theory and practice: an overview of user interests in knowledge seeking and use. *Journal of Knowledge Management*, 2(2), 36-46.

Dervin, B. (1999). On studying information seeking methodologically: the implications of connecting metatheory to method. *Information Processing & Management, 35*(6), 727-750.

Dervin, B. 2005. What methodology does to theory: sense-making methodology as exemplar. In Fisher, K.E., Erdelez S. and McKechnie L. (eds.), *Theories of information behavior* (pp. 25-30). Medford, New Jersey: Information Today, Inc.

Dredze, M. and Paul, M. J. (2014). Natural language processing for health and social media. *IEEE Intelligent Systems, 29*(2), 64-67.

Dunlap J.C., Loros J.J. and DeCoursey P.J. (eds.). (2004). *Chronobiology: biological timekeeping*. Sunderland MA: Sinauer Associates, 2004.

Ek, S. (2013). Gender differences in health information behaviour: a Finnish population-based survey. *Health Promotion International, 30*(3), 736-745.

Ek, S. and Niemelä, R. (2010). Onko internetistä tullut suomalaisten tärkein terveystiedonlähde? Deskriptiivistä tutkimustietoa vuosilta 2001 ja 2009. *Informaatiotutkimus, 29*(4).

Ek, S., Eriksson-Backa, K. and Niemelä, R. (2013). Use of and trust in health information on the Internet: a nationwide eight-year follow-up survey. *Informatics for Health and Social Care, 38*(3), 236-245.

Emson, H. E. (1987). Health, disease and illness: matters for definition. *Canadian Medical Association Journal, 136*(8), 811-813.

Enwald, H. (2013). Tailoring health communication: the perspective of information users health information behaviour in relation to their physical health status. *Acta Universitatis Ouluensis. Series B, Humaniora, 118*.

Erikainen, S., Pickersgill, M., Cunningham-Burley, S. and Chan, S. (2019). Patienthood and participation in the digital era. *Digital Health, 5*, 1–10.

Eriksson, K. (1993). *Hälsans idé.* Liber/Almqvist & Wiksell medicin.

Eysenbach, G., Powell, J., Kuss, O. and Sa, E. (2002). Empirical studies assessing the quality of health information for consumers on the World Wide Web: a systematic review. *JAMA 287*, 2691-2700.

Eysenbach, G. (2006). Infodemiology: tracking flu-related searches on the Web for Syndromic surveillance. *AMIA 2006 Symposium proceedings*, 244-248.

Eysenbach, G. (2009). Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *Journal of Medical Internet Research, 11*(1), e11.

Eysenbach, G. (2011). Infodemiology and infoveillance: tracking online health information and cyberbehavior for public health. *American Journal of Preventive Medicine, 40*(5), S154-S158.

FIAM - Finnish Internet Audience Measurement, (2018). Tulokset. Available at: http://fiam.fi/tulokset/ (accessed 21 December 2018)

Fidel, R. (2012). *Human information interaction: An ecological approach to information behavior*. Cambridge, MA: MIT Press.

Flynn, K. E., Smith, M. A. and Freese, J. (2006). When do older adults turn to the internet for health information? Findings from the Wisconsin longitudinal study. *Journal of General Internal Medicine, 21*(12), 1295-1301.

Fox, S. and Duggan, M. (2013). Health online 2013. *Health, 2013*, 1-55. Retrieved May 8th, 2019. Available at: http://www.pewinternet.org/2013/01/15/health-online-2013/

Fox, N. and Ward, K. (2006). Health identities: from expert patient to

resisting consumer. *Health, 10*(4), 461-479.

Frost, J., Okun, S., Vaughan, T., Heywood, J. and Wicks, P. (2011). Patient-reported outcomes as a source of evidence in off-label prescribing: analysis of data from PatientsLikeMe. *Journal of Medical Internet Research, 13*(1), e6.

Gabarron, E., Lau, A. Y. and Wynn, R. (2015). Is there a weekly pattern for health searches on Wikipedia and is the pattern unique to health topics?. *Journal of Medical Internet Research, 17*(12), e286.

Gencoglu, O. and Ermes, M. (2018). Predicting the Flu from Instagram. *arXiv preprint arXiv:1811.10949*.

Giddens, A. (1979). *Central problems in social theory – action, structure and contradiction in social analysis.* Macmillan, London.

Gordijn, M. C. M., Beersma, D. G. M., Bouhuys, A. L., Reinink, E. and Van den Hoofdakker, R. H. (1994). A longitudinal study of diurnal mood variation in depression; characteristics and significance. *Journal of Affective Disorders, 31*(4), 261-273.

Gorichanaz, T. (2016). Documents and Time. *Proceedings from the Document Academy, 3*(1), article 7.

Gustafson, D. H., Hawkins, R., Boberg, E., Pingree, S., Serlin, R. E., Graziano, F. and Chan, C. L. (1999). Impact of a patient-centered, computer-based health information/support system. *American Journal of Preventive Medicine, 16*(1), 1-9.

Guy, S., Ratzki-Leewing, A., Bahati, R. and Gwadry-Sridhar, G. (2012). Social Media: A Systematic Review to Understand the Evidence and Application in Infodemiology. In Kostkova, P, Szomszor, M. and Fowler, D (eds.) *eHealth 2011*, LNICST 91, 1-8.

Hakko H., Räsänen P. and Tiihonen J. (1998). Seasonal variation in suicide occurrence in Finland. *Acta Psychiatrica Scandinavica, 98*(2), 92-97.

Halder, K., Poddar, L. and Kan, M. Y. (2017). Modeling Temporal Progression of Emotional Status in Mental Health Forum: A Recurrent Neural Net Approach. In *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis* (pp. 127-135).

Hallyburton, A. and Evarts, L. (2014). Gender and online health information seeking: a five survey meta-analysis. *Journal of Consumer Health on the Internet, 18*(2), 128-142.

Hasler, B.P. Diurnal mood variation. In: Gellman M., Turner R.J. (eds.) *Encyclopedia of Behavioral Medicine* (pp. 612-614). New York: Springer-Verlag.

Hausner, H., Hajak, G. and Spießl, H. (2008). Gender differences in help-seeking behavior on two internet forums for individuals with self-reported depression. *Gender Medicine, 5*(2), 181-185.

Henly, S. J., Wyman, J. F. and Findorff, M. J. (2011). Health and illness over time: The trajectory perspective in nursing science. *Nursing Research, 60*(3 Suppl), S5.

Holopainen, J., Helama, S., Björkenstam, C. and Partonen, T. (2013). Variation and seasonal patterns of suicide mortality in Finland and Sweden since the 1750s. *Environmental Health and Preventive Medicine, 18*(6), 494.

Huang, T. (2017). High-resolution Temporal Representations of Alcohol and Tobacco Behaviors from Social Media Data. *Proc. ACM Hum.-Comput. Interact. 1*, CSCW, 54.

Isometsä, E. (2014). Suicidal behaviour in mood disorders—who, when, and

why?. *The Canadian Journal of Psychiatry*, *59*(3), 120-130.

Jadhav, A., Andrews, D., Fiksdal, A., Kumbamu, A., McCormick, J. B., Misitano, A., ... and Pathak, J. (2014). Comparative analysis of online health queries originating from personal computers and smart devices on a consumer health information portal. *Journal of Medical Internet Research*, *16*(7), e160.

Jennings, D. (1986). The confusion between disease and illness in clinical medicine. *Canadian Medical Association Journal*, *135*(8), 865-870.

Johnson, J. D. and Case, D. O. (2012). *Health information seeking*. Peter Lang: New York.

Kim, S. M. and Oh, J. Y. (2011). Health information acquisition online and its influence on intention to visit a medical institution offline. *Information Research*, *16*(4).

Kreuter, F., Presser, S. and Tourangeau, R. (2008). Social desirability bias in cati, ivr, and web surveysthe effects of mode and question sensitivity. *Public Opinion Quarterly*, *72*(5), 847-865.

Kluger, N. and Bouchard, L. J. (2019). A Comparative Study of Google Search Trends for Melanoma, Breast Cancer and Prostate Cancer in Finland. *Dermatology*, 1-2.

Kohonen I., Kuula-Luumi, A. and Spoof, S-K (eds.). *The ethical principles of research with human participants and ethical review in the human sciences in Finland Finnish National Board on Research Integrity TENK guidelines 2019*. Publications of the Finnish National Board on Research Integrity TENK 3/2019.

Kraut, R. Y., Snedeker, K. G., Babenko, O. and Honish, L. (2017). Influence of school year on seasonality of norovirus outbreaks in developed countries. *Canadian Journal of Infectious Diseases and Medical Microbiology*, 2017.

Lagus, K., Ruckenstein, M., Pantzar, M. and Ylisiurua, M. (2016). Suomi24: muodonantoa aineistolle. *Valtiotieteellisen tiedekunnan julkaisuja, 10*, Helsingin yliopisto, Helsinki. Retrieved from https://blogs.helsinki.fi/citizenmindscapes/files/2016/05/257383_HY_VALT_suomi24_muodonantoa_aineistolle.pdf

Lambert, S. D. and Loiselle, C. G. (2007). Health information—seeking behavior. *Qualitative Health Research*, *17*(8), 1006-1019.

Lazer, D., Kennedy, R., King, G. and Vespignani, A. (2014). The parable of Google Flu: traps in big data analysis. *Science*, *343*(6176), 1203-1205.

Lewandowsky, S., Ecker, U. K. and Cook, J. (2017). Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of Applied Research in Memory and Cognition*, *6*(4), 353-369.

Lloyd, G. E. R. (1983). *Hippocratic writings*. Penguin books, Middlesex, England.

Maes, M., Meltzer, H. Y., Suy, E. and De Meyer, F. (1993). Seasonality in severity of depression: relationships to suicide and homicide occurrence. Acta *Psychiatrica Scandinavica*, *88*(3), 156-161.

Mamiya, H., Shaban-Nejad, A. and Buckeridge, D. L. (2017). Online public health intelligence: ethical considerations at the big data era. In Shaban-Nejad, A., Brownstein, J. and Buckeridge, D. L. (eds.), *Public Health Intelligence and the Internet* (pp. 129-148), Springer, Cham.

Mamykina, L. (2015). Adopting the sensemaking perspective for chronic disease self-management. *Journal of Biomedical Informatics*, 56, 406–417.

Marcus, M., Yasamy, M. T., van Ommeren, M., Chisholm, D. and Saxena, S. (2012).

Depression: A global public health concern. *WHO Department of Mental Health and Substance Abuse, 1*, 6-8.

Marton, C. and Wei Choo, C. (2012). A review of theoretical models of health information seeking on the web. *Journal of Documentation, 68*(3), 330-352.

Matsuda, S., Aoki, K., Tomizawa, S., Sone, M., Tanaka, R., Kuriki, H. and Takahashi, Y. (2017). Analysis of patient narratives in disease blogs on the internet: an exploratory study of social pharmacovigilance. *JMIR Public Health and Surveillance, 3*(1), e10.

McGowan, P. (2002). Self-care behavior. *Encyclopedia of Public Health*. Retrieved June 1, 2019 from Encyclopedia.com: https://www.encyclopedia.com/education/encyclopedias-almanacs-transcripts-and-maps/self-care-behavior

McKee, R. (2013). Ethical issues in using social media for health and health care research. *Health Policy, 110*, 298-301.

Mogilner, C., Hershfield, H. E. and Aaker, J. (2018). Rethinking time: Implications for well-being. *Consumer Psychology Review, 1*(1), 41-53.

McKenzie, P. J. and Davies, E. (2002). Time is of the essence: Social theory of time and its implications for LIS research. In *Proceedings of the Annual Conference of CAIS/Actes du congrès annuel de l'ACSI 2002*, 1-13.

Mustonen, H., Metso, L., and Mäkelä, P. (2010). Milloin suomalaiset juovat. In Mäkelä, P., Mustonen, H. and Tigerstedt, C. (eds.). *Suomi juo - Suomalaisten alkoholinkäyttö ja sen muutokset 1968–2008* (pp. 55-69), Terveyden ja hyvinvoinnin laitos, Helsinki.

Naslund, J. A., Aschbrenner, K. A., Marsch, L. A. and Bartels, S. J. (2016). The future of mental health care: peer-to-peer support and social media. *Epidemiology and Psychiatric Sciences, 25*(2), 113-122.

Neal, D. M. and McKenzie, P. J. (2011). Putting the pieces together: endometriosis blogs, cognitive authority, and collaborative information behavior. *Journal of the Medical Library Association, 99*(2), 127-134.

Nimrod, G. (2012). From knowledge to hope: Online depression communities. *International Journal on Disability and Human Development, 11*(1), 23-30.

Nimrod, G. (2013). Online depression communities: members' interests and perceived benefits. *Health Communication 28*(5), 425-434.

Nuti, S. V., Wayda, B., Ranasinghe, I., Wang, S., Dreyer, R. P., Chen, S. I. and Murugiah, K. (2014). The use of google trends in health care research: a systematic review. *PloS one, 9*(10), e109583.

Oliphant T. (2010). *The Information Practices of People Living with Depression: Constructing Credibility and Authority*. Faculty of Information & Media Studies University of Western Ontario, London. Retrieved from http://ir.lib.uwo.ca/etd/35

Ormandy, P. (2011). Defining information need in health–assimilating complex theories derived from information science. *Health Expectations, 14*(1), 92-104.

Ortiz-Martínez, Y. and Jiménez-Arcia, L. F. (2017). Yellow fever outbreaks and Twitter: Rumors and misinformation. *American Journal of Infection Control, 45*(7), 816-817.

Osuka, H., Hall, A. J., Wikswo, M. E., Baker, J. M. and Lopman, B. A. (2018). Temporal Relationship Between Healthcare-Associated and Nonhealthcare-Associated Norovirus Outbreaks and Google Trends Data in

the United States. *Infection Control & Hospital Epidemiology, 39*(3), 355-358.

Oyane, N. M., Bjelland, I., Pallesen, S., Holsten, F. and Bjorvatn, B. (2008). Seasonality is associated with anxiety and depression: the Hordaland health study. *Journal of Affective disorders, 105*(1-3), 147-155.

Palmarella, G., McGrath, C., Solomon, S., Dupuis, R. and Cannuscio, C. (2018). Advances in Digital Health Research. *CPHI Data Briefs, 1*.

Patana, P. (2014). Mental health analysis profiles (MhAPs): Finland. *OECD Health Working Papers, 72*, OECD Publishing.

Paul, M. J. and Dredze, M. (2011). You are what you Tweet: Analyzing Twitter for public health. *Icwsm, 20*, 265-272.

Paul, M. J., Chisolm, M. S., Johnson, M. W., Vandrey, R. G. and Dredze, M. (2016). Assessing the validity of online drug forums as a source for estimating demographic and temporal trends in drug use. *Journal of Addiction Medicine, 10*(5), 324-330.

Pendry, L. F. and Salvatore, J. (2015). Individual and social benefits of online discussion forums. *Computers in Human Behavior, 50*, 211-220.

Pesälä, S., Virtanen, M. J., Sane, J., Mustonen, P., Kaila, M. and Helve, O. (2017). Health Information–Seeking Patterns of the General Public and Indications for Disease Surveillance: Register-Based Study Using Lyme Disease. *JMIR Public Health and Surveillance, 3*(4), e86.

Pesälä, S., Virtanen, M. J., Mukka, M., Ylilammi, K., Mustonen, P., Kaila, M. and Helve, O. (2019). Healthcare professionals' queries on oseltamivir and influenza in Finland 2011-2016—Can we detect influenza epidemics with specific online searches?. *Influenza and Other Respiratory Viruses, 2019*, 1–8.

Petrakaki, D., Hilberg, E. and Waring, J. (2018). Between empowerment and self-discipline: Governing patients' conduct through technological self-care. *Social Science & Medicine, 213*, 146-153.

Pettigrew, K. E., Fidel, R. and Bruce, H. (2001). Conceptual frameworks in information behavior. Annual review of information science and technology *Annual Review of Information Science and Technology, 35*, 43-78.

Phelan, N., Davy, S., O'Keeffe, G. W. and Barry, D. S. (2017). Googling in anatomy education: Can google trends inform educators of national online search patterns of anatomical syllabi?. *Anatomical Sciences Education, 10*(2), 152-159.

Pilerot, O. (2012). LIS research on information sharing activities – people, places, or information. *Journal of Documentation, 68*(4), 559-581.

Poikolainen, K., Leppänen, K. and Vuori, E. (2002). Alcohol sales and fatal alcohol poisonings: a time-series analysis. *Addiction, 97*(8), 1037-1040.

Powell, J. and Clarke, A. (2006). Internet information-seeking in mental health: Population survey. *The British Journal of Psychiatry, 189*(3), 273-277.

Reddy, M. C., Dourish, P. and Pratt, W. (2006). Temporality in medical work: Time also matters. *Computer Supported Cooperative Work, 15*(1), 29-53.

Reinberg, A. E., Dejardin, L., Smolensky, M. H. and Touitou, Y. (2017). Seven-day human biological rhythms: An expedition in search of their origin, synchronization, functional advantage, adaptive value and clinical relevance. *Chronobiology International, 34*(2), 162-191.

Ren, C., Deng, Z., Hong, Z. and Zhang, W. (2019). Health information in the digital age: an empirical study of the

perceived benefits and costs of seeking and using health information from online sources. *Health Information & Libraries Journal* (Online Version of Record before inclusion in an issue).

Renahy, E., Parizot, I. and Chauvin, P. (2010). Determinants of the frequency of online health information seeking: results of a web-based survey conducted in France in 2007. *Informatics for Health and Social Care, 35*(1), 25-39.

Rier, D. (2010). The patient's experience of illness. In Bird et al. (eds.) *Handbook of medical sociology*, p. 163-178.

Roenneberg, T. (2012*). Internal time: Chronotypes, social jet lag, and why you're so tired*. Harvard University Press.

Roy, K. M., Tubbs, C. Y. and Burton, L. M. (2004). Don't have no time: Daily rhythms and the organization of time for low-income families. *Family Relations, 53*(2), 168-178.

Rusting, C. L. and Larsen, R. J. (1998). Diurnal patterns of unpleasant mood: Associations with neuroticism, depression, and anxiety. *Journal of Personality, 66*(1), 85-103.

Sansone, R. A. and Sansone, L. A. (2011). The Christmas effect on psychopathology. *Innovations in Clinical Neuroscience, 8*(12), 10.

Santillana, M., Zhang, D. W., Althouse, B. M. and Ayers, J. W. (2014). What can digital disease detection learn from (an external revision to) Google Flu Trends?. *American Journal of Preventive Medicine, 47*(3), 341-347.

Savolainen, R. (2006a). Time as a context of information seeking. *Library & Information Science Research, 28*(1), 110-127.

Savolainen, R. (2006b). Information use as gap-bridging: The viewpoint of sense-making methodology. *Journal of the American Society for Information Science and Technology, 57*(8), 1116-1125.

Savolainen, R. (2012). Conceptualizing information need in context. *Information Research, 4*(17).

Savolainen, R. (2018). Information-Seeking Processes as Temporal Developments: Comparison of Stage-based and Cyclic Approaches. *Journal of the Association for Information Science and Technology, 69*(6), 787-797.

Scrutton, A. N. and Stevenson T.J. (2018). Appetitive information seeking behaviour reveals robust daily rhythmicity for Internet-based food-related keyword searches. *Royal Society Open Science, 5*:172080.

Seifter, A., Schwarzwalder, A., Geis, K. and Aucott, J. (2010). The utility of "Google Trends" for epidemiological research: Lyme disease as an example. *Geospatial health, 4*(2), 135-137.

Seo, D. W. and Shin, S. Y. (2017). Methods using social media and search queries to predict infectious disease outbreaks. *Healthcare Informatics Research, 23*(4), 343-348.

Siepmann, M. (2008). Health behavior. In Wilhelm Kirch (Ed.), *Encyclopedia of Public Health* (pp. 515-521). Dordrecht: Springer Science & Business Media.

Solomon, P. (1997). Discovering information behavior in sense making. I. Time and timing. *Journal of the American Society for Information Science, 48*(12), 1097-1108.

Sonnenwald, D. H. and Iivonen, M. (1999). An integrated human information behavior research framework for information studies. *Library & Information Science Research, 21*(4), 429-457.

StatCounter. (2019). Search Engine Market Share Finland. Retreived May 17th, 2019 from:

http://gs.statcounter.com/search-engine-market-share/all/finland

Statista. (2019). Worldwide desktop market share of leading search engines from January 2010 to January 2019. Retreived May 17th, 2019 from: https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/

Stephens-Davidowitz, S. I. (2013). *Essays using Google data*. Doctoral dissertation, Harvard University.

Stordal, E., Morken, G., Mykletun, A., Neckelmann, D. and Dahl, A. A. (2008). Monthly variation in prevalence rates of comorbid depression and anxiety in the general population at 63–65 North: The HUNT study. *Journal of Affective Disorders, 106*(3), 273-278.

Šubrt, J. (2001). The Problem of Time from the Perspective of the Social Sciences. *Czech Sociological Review, 9*(2), 211-224.

Svalastog, A. L., Donev, D., Kristoffersen, N. J. and Gajović, S. (2017). Concepts and definitions of health and health-related values in the knowledge landscapes of the digital society. *Croatian Medical Journal, 58*(6), 431.

Talja, S. and Hansen, P. (2006). Information sharing. In A. Spink and C. Cole (eds.) *New directions in human information behavior* (pp. 113-134). Dordrecht: Springer.

Tidline, T. (2007). Dervin's Sense-Making. In Fisher, K.E., Erdelez S. and McKechnie L. (eds.), *Theories of information behavior* (pp. 113-117). Medford, New Jersey: Information Today, Inc.

Tilastokeskus. (2017). Väestön tieto- ja viestintätekniikan käyttö 2017. Available at: https://www.stat.fi/til/sutivi/2017/13/sutivi_2017_13_2017-11-22_fi.pdf (accessed 19 November 2018)

Tilastokeskus, (2018). Väestön tieto- ja viestintätekniikan käyttö 2018. Available at: http://www.stat.fi/til/sutivi/2018/sutivi_2018_2018-12-04_tau_018_fi.html (accessed 14 May 2019).

"Time". Oxford English Dictionary. (2008). Oxford University Press, New York.

"Time". 2019. Merriam-Webster.com. Retrieved May 3, 2019 from www.merriam-webster.com/dictionary/time.

Tkachenko, N., Chotvijit, S., Gupta, N., Bradley, E., Gilks, C., Guo, W., ... and Jarvis, S. (2017). Google trends can improve surveillance of type 2 diabetes. *Scientific Reports, 7*(1), 4993.

Torrent-Sellens, J., Díaz-Chao, Á., Soler-Ramos, I. and Saigí-Rubió, F. (2016). Modelling and predicting eHealth usage in Europe: a multidimensional approach from an online survey of 13,000 European Union internet users. *Journal of Medical Internet Research, 18*(7), e188.

Tran, U. S., Andel, R., Niederkrotenthaler, T., Till, B., Ajdacic-Gross, V. and Voracek, M. (2017). Low validity of Google Trends for behavioral forecasting of national suicide rates. *PloS one, 12*(8), e0183149.

Tsuno, N., Besset, A. and Ritchie, K. (2005). Sleep and depression. *The Journal of Clinical Psychiatry, 66*(10), 1254-1269.

Tumen, S. and Zeydanli, T. (2014). Day-of-the-week effects in subjective well-being: does selectivity matter?. *Social Indicators Research, 119*(1), 139-162.

Wahlbeck, K., Nordmyr, J., Reini, K., Forsman, A.K., Aromaa, E., Nyqvist, F. and Herberts, K. Western Finland Mental Health Survey 2014: Survey methods. National Institute for Health and Welfare. Discussion paper 34/2014.

Wakefield, A.J., Murch, S.H., Anthony, A., Linnell, J., Casson, D.M., Malik, M., Berelowitz, M., Dhillon, A.P., Thomson, M.A., Harvey, P. and Valentine, A., 1998. RETRACTED: Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children.

Walther, J. B. and Boyd, S. (2002). Attraction to computer-mediated social support. In Lin CA and Atkin D (eds) *Communication Technology and Society: Audience Adoption and Uses*. (pp 153-188). Cresskill, NJ: Hampton Press.

Vasconcellos-Silva, P. R., Carvalho, D. B. F., Trajano, V., de La Rocque, L. R., Sawada, A. C. M. B. and Juvanhol, L. L. (2017). Using Google Trends data to study public interest in Breast Cancer screening in Brazil: why not a Pink February?. *JMIR Public Health and Surveillance, 3*(2), e17.

Weick, K. (2001). *Sensemaking in Organizations*. Malden, MA: Blackwell Publishing Ltd.

Veinot, T. (2009). Interactive Acquisition and Sharing: Understanding the Dynamics of HIV/AIDS Information Networks. *Journal of the American Society for Information Science and Technology, 60*(11), 2313–2332.

Vernon, J. L. (2017). Science in the post-truth era. *American Scientist, 105*(1), 2-3.

WHO. (2018). Depression fact sheet. Retrieved 22nd of May, 2019 from: https://www.who.int/news-room/fact-sheets/detail/depression

Wicks, P., Massagli, M., Kulkarni, A. and Dastani, H. (2011). Use of an online community to develop patient-reported outcome instruments: the Multiple Sclerosis Treatment Adherence Questionnaire (MS-TAQ). *Journal of Medical Internet Research, 13*(1), e12.

Wicks, P., Massagli, M., Frost, J., Brownstein, C., Okun, S., Vaughan, T., ... and Heywood, J. (2010). Sharing health data for better outcomes on PatientsLikeMe. *Journal of Medical Internet Research, 12*(2), e19.

Wilson, K. and Brownstein, J. S. (2009). Early detection of disease outbreaks using the Internet. *Cmaj, 180*(8), 829-831.

Wilson, T. D. (1999). Models in information behaviour research. *Journal of Documentation, 55*(3), 249-270.

Wilson, T. D. (2000). Human information behavior. *Informing Science, 3*(2), 49-56.

Wilson, T.D. (2005). Evolution in Information Behaviour Modeling: Wilson's Model. In Fisher, K.E., Erdelez S. and McKechnie L. (eds.), *Theories of information behavior* (pp. 31-36). Medford, New Jersey: Information Today, Inc.

Wilson, T.D. (2010). Information sharing: an exploration of the literature and some propositions. *Information Research, 15*(4), n4.

Wirz-Justice, A. (2008). Diurnal variation of depressive symptoms. *Dialogues in Clinical Neuroscience, 10*(3), 337.

Wong, P. W. C., Fu, K. W., Yau, R. S. P., Ma, H. H. M., Law, Y. W., Chang, S. S. and Yip, P. S. F. (2013). Accessing suicide-related information on the internet: a retrospective observational study of search behavior. *Journal of Medical Internet Research, 15*(1), e3.

Yang, A. C., Huang, N. E., Peng, C. K. and Tsai, S. J. (2010). Do seasons have an influence on the incidence of depression? The use of an internet search engine query data as a proxy of human affect. *PloS one, 5*(10), e13728.

Ybarra, M. and Suman, M. (2006). Reasons, assessments and actions taken: sex and age differences in uses

of Internet health information. *Health Education Research, 23*(3), 512-521.

Yeung, D. (2018). Social Media as a Catalyst for Policy Action and Social Change for Health and Well-Being: Viewpoint. *Journal of Medical Internet Research, 20*(3), e94.

Zeraatkar, K. and Ahmadi, M. (2018). Trends of infodemiology studies: a scoping review. *Health Information & Libraries Journal*, 35(2), 91-120.

Jonas Tana

Infodemiology

Studying rhythmicity in online health information behaviour

This thesis examines the rhythmicity and temporal variations of online health information behaviour. These aspects are of great significance, as health is a dynamic process bound by rhythmicity and temporal structures. Today a majority of health information behaviour happens online, leaving vast amounts of digital traces behind. Aggregating and analysing these digital traces to study online health behaviours has been dubbed infodemiology, and is rooted in the idea that there is a relationship between population health on one hand, and information and communication patterns on the internet, on the other. Utilizing the infodemiology approach not only allows us to gain insights into new contextual aspects of online health information behaviour, it also allows us to understand the multidimensional relationship between time and health.

ISBN 978-952-12-3877-2