The effects of exercise in women diagnosed with breast cancer at different stages of the cancer treatment continuum

Edward N. Stanhope

Faculty of Education, Health and Wellbeing
University of Wolverhampton

A thesis submitted for the degree of
Doctor of Philosophy

TBC 2024

Abstract

INSERT ABSTRACT HERE

The effects of exercise in women diagnosed with breast cancer at different stages of the cancer treatment continuum

Edward N. Stanhope

Faculty of Education, Health and Wellbeing

University of Wolverhampton

A thesis submitted for the degree of

Doctor of Philosophy

TBC 2024

DEDICATIONS GO HERE

Acknowledgements

INCLUDE THE ACKNOWLEDGEMENTS HERE

Edward N. Stanhope
Faculty of Education, Health and Wellbeing, University of Wolverhampton
INSERT DATE HERE

Abstract

INSERT ABSTRACT HERE

Contents

List of Figures

List of Tables

List of Abbreviations

TNF . . . . . Tumour Necrosis Factor

Introduction

0.1 Background

Cancer is the one of the world's leading health problems, with one in five men and one in six women developing the disease, and one in eight men and one in eleven women losing their life to it (GLOBOCAN). Cancer occurs due to the continual upregulated proliferation of cancer cells, which grow and divide at an uncontrolled rate, rather than responding to signals that control cellular behaviour (Feitelson et al., 2015; Hanahan & Weinberg, 2011). The abnormal proliferation of cells (tumour) can occur in a variety of cells within the body and there are therefore hundreds of distinct types of cancer, all of which behave and respond to treatment differently. A tumour may be benign, in which the cells remain confined to their original location, or malignant whereby cells invade surrounding tissues and are able to spread throughout the body (metastasis). The confinement of benign tumours makes surgical intervention possible, whereas the spread of malignant cells often makes them resistant to such localised treatment. The original hallmarks of cancer describe six processes that lead to the development and progression of cancer, self-sufficiency in growth signals, insensitivity to anti-growth signals, evading apoptosis, limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis (Hanahan & Weinberg, 2000). Since, technological advancements have led to a superior understanding of cancer hallmarks and inflammation, genomic instability and mutation, deregulation of cellular energetics and avoidance of immune destruction were recognised as 'emerging' cancer hallmarks (Figure 1) (Hanahan & Weinberg, 2011).

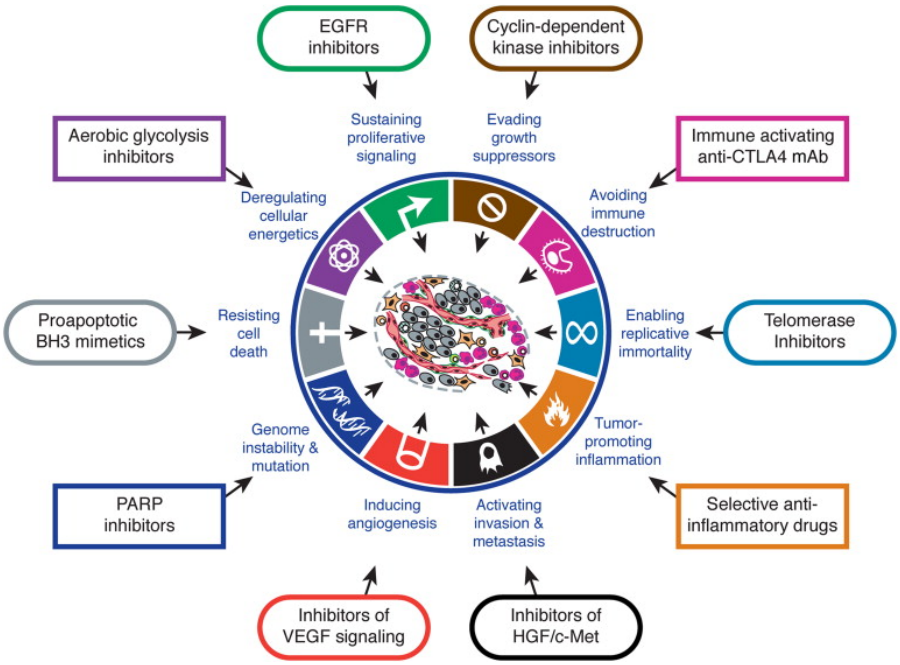

Figure 1: Ten Hallmarks of Cancer

0.2 Cancer Development and Progression

0.2.1 Hallmark One: Replicative Immortality

A 'normal' cell in the body is limited by the number of successive cell cycles it can go through, known as the Hayflick limit (Hayflick, 1965; Hayflick & Moorhead, 1961). This is due, in part, to the shortening of specialised chromatin structures at the end of chromosomes (telomeres) after each mitotic division, which protect against the end to end chromosome fusions that threaten cell viability. Each replication typically loses 50 to 100 nucleotides of DNA telomere. Once the telomere shortens beyond a specific limit, usually the 'uncapping' of telomere ends, it triggers the p53 pathway, preventing further proliferation. Though, the reasons for this progressive shortening is unclear (Reddel, 2000). In addition, oncogenic and mitogenic signals also induce senescence, establishing a barrier against tumorigenesis.

Telomere length is an important factor in the development and progression of many human diseases, specifically those associated with age, where there is increased telomere degradation (Blasco, 2005). It is generally accepted that cancer cells have unlimited replicative potential and are able to overcome senescence and

subsequent cell death (Hanahan & Weinberg, 2011; Reddel, 2000). An array of evidence suggests that cancer cells are able to elongate their chromosome telomeres through specialised DNA polymerase (telomerase) that adds repeated segments to the telomere (Blasco, 2005; Bodnar et al., 1998; Dimri, 2005; Hanahan & Weinberg, 2011). Kim and colleagues (1994) reported more than 90% of tumours contained increased telomerase activity (Kim et al., 1994). Although, it has since been discovered that not all cancers have an active telomerase, which lead researchers to explore alternative explanations for the increased lengthening of telomeres. One suggestion has been that it may be possible for tumour formation to occur, without the presence telomerase, if the telomere has a sufficient reservoir of telomere repeats before it reaches the 'limit' triggering the DNA damage response (Dagg et al., 2017; Viceconte et al., 2017). The development of cancer is a complex process and the abrogation of senescence alone does not lead to tumour development.

0.2.2 Hallmark Two: Genome Instability

Replicative immortality and many of the other cancer hallmarks depend on the successive adaptation of the tumour genome. Yet, genomes are highly refined to detect and resolve defects in the DNA to limit spontaneous mutation. Cancer cells are able to increase the rate of mutation and perform tumourigenesis by heightening mutogenic sensitivity and causing breakdown of the genomic maintenance machinery (Hanahan & Weinberg, 2011; Negrini, Gorgoulis, & Halazonetis, 2010). An array of genomic maintenance defects have been reported in the literature and can be categorised by whether they impact the detection of DNA damage, the repair of damaged DNA, or prevent DNA damage by intercepting mutogenic molecules (Ciccia & Elledge, 2010; Hanahan & Weinberg, 2011; Jackson & Bartek, 2009; Negrini et al., 2010).

The inactivation of genomic maintenance machinery leads to genomic instability, of which the main form is chromosomal instability (CIN); the increased rate of chromosome structure change compared to with normal cells (Negrini et al., 2010). Two hypotheses have been offered to explain genomic instability, the first relates

to mutation and the second to oncogene-induced DNA replication. In hereditary cancers, it appears that genomic instability is a consequence of DNA repair gene mutation, however, sequencing studies in non-hereditary cancers have failed to detect mutations and there remains a degree of uncertainty in this hypothesis (Negrini et al., 2010). Alternatively, patterns in tumour suppressor genes (TP53/p53, ATM and CDKN2A/p16INK4A and P14ARF) and oncogenes (EGFR and RAS) support the oncogene-induced DNA replication hypothesis. However, the accumulation of cells is only possible if cells can avoid being detected by the surveillance system and subsequent senescence or apoptosis, which is discussed later in this chapter (Jackson & Bartek, 2009).

0.2.3 Hallmark Three: Evasion of Growth Suppressor Signals

Tumour suppressor genes contribute to the fidelity of the cell cycle replication process by limiting cell growth, inhibiting cell transformation and preventing proliferation (Hanahan & Weinberg, 2011). The genes are responsible for regulating the cell cycle, apoptosis, differentiation, signal transduction and cellular adhesion, genomic instability surveillance and DNA repair. Tumour suppressor genes are commonly lost or deleted in cancer patients, allowing cells to replicate and grow. p53 or TP53 and retinoblastoma associated (Rb) proteins govern the decisions of a cell to proliferate, or activate apoptosis, and are the most commonly mutated genes in cancer patients. The mechanisms by which tumour suppressor genes are inactivated depends on the gene involved and therefore vary tumour to tumour, for instance the mutation BRCA1 and BRCA2 predisposes women to breast and ovarian cancers (Deng & Scott, 2000).

p53 (TP53), Rb1, BRCA1 and BRCA2 are all examples of intragenic mutations which are able to down regulate tumour suppressor function by inappropriate expression of transcriptional activators and the promotion of methylation, repressing tumour suppressors. Conceptually, tumour suppressor genes can be classified into two categories; gatekeepers and caretakers (Oliveira, Ross, & Fletcher, 2005; Zhou,

Zhang, Liu, Yang, & He, 2011). Caretaker tumour suppressors are genes whose loss of function is not directly associated with tumour development. Whereas, gatekeeper tumour suppressor genes (p53, Rb) directly inhibit tumour growth and promote cell death, therefore, inactivity of these genes facilitates cancer formation and proliferation. The retinoblastoma (Rb) protein determines whether a cell should proceed to grow and divide, or not. The gene plays a vital role in cell cycle checkpoint control, whereby active Rb1 binds to transcription factor E2F and renders it inactive. Retinoblastoma (Rb) proteins largely function extracellularly, whereas, TP53 (p53) operates within the cell (intracellularly) to detect signals of stress or abnormality and is deleted or mutated in 50% of all cancers (Oliveira et al., 2005). When the gene detects abnormal genome damage or excessive nucleotide pools, growth promoting signal, glucose or suboptimal oxygenation TP53 (p53) ceases further cell-cycle progression or initiates apoptosis (Hanahan & Weinberg, 2011). Therefore, any down regulation of these gene pathways makes cell proliferation and tumour growth possible.

TP53 mutation is more common in BRCA1 and BRCA2 breast cancer though the prevalence of TP53 mutation is lower in breast cancer generally than many other cancers and is typically associated with more aggressive conditions (Taylor & Stark, 2001). Though there appears to be some inconsistency in the literature with regard to the relationship between Rb1 and TP53 tumour suppressor expression and breast cancer prognosis (Oliveira et al., 2005).

0.2.4 Hallmark Four: Resistance to cell death

Apoptosis (programmed cell death) serves as a natural barrier to cancer development and progression and plays an essential role in development, immunity and disease where removal of superfluous cells promotes organismal health (Adams & Cory, 2007; Lowe, Cepero, & Evan, 2004). Cancer cells evade apoptosis due to signalling imbalances that result in elevated levels of oncogene. Apoptosis is governed by both upstream regulators and downstream effectors. Regulators are either extracellular death inducing signals (FAS ligand/FAS receptor) which form the extracellular

apoptotic program, or intracellular sensors which form the intracellular apoptotic program (Hanahan & Weinberg, 2011). The cells decision to undergo apoptosis is largely determined by intracellular cysteine proteases and is primarily regulated by the BCL-2 protein family that respond to DNA damage, gamma irradiation, oncogene activation and the withdrawal of growth factor (Adams & Cory, 2007). BCL-2 proteins BCL-XL, MCL-1 and BFL1 are pro-survival proteins that regulate protein turnover that respond to cell death stimuli (Campbell & Tait, 2018), it is therefore of no surprise that there is increased expression of these proteins in many cancer types. The upregulation of pro-survival BCL-2 proteins can occur through chromosomal translocation, gene amplification and increased gene expression and consequently increasing cell death resistance. Whereas, BCL-2 proteins BIM, BAD, BID, NOXA, PUMA, BMF, HRK and BIK are pro-apoptotic. Pro-apoptotic BCL-2 proteins in normal cells respond to stress to trigger cell death. Studies have revealed that effector proteins BAC, BAK and BOK conform to similar pro-survival BCL-2 proteins (Suzuki, Youle, & Tjandra, 2000). However, similarly to the effects of upregulating pro-survival proteins, the downregulation of pro-apoptotic proteins can lead to cell death resistance.

Although, the cellular conditions that trigger apoptosis are yet to be fully determined, several abnormalities can be identified in the role of tumour development. Primarily tumour cells are able to circumvent apoptosis by hindering the TP53 tumour suppressor responsible for regulating NOXA and PUMA in the presence of DNA damage. Additionally, tumours may increase the expression of pro-survival proteins or decrease the expression of pro-apoptotic proteins, as previously discussed.

0.2.5 Hallmark Five: Sustained Proliferation

Normal tissues are able to tightly control and regulate the production and release of growth promoting signals, in order to maintain tissue homeostasis. On the other hand, cancer cells possess the ability to downregulate these signals enabling the cell to proliferate due to replicative immortality, genomic instability, evasion of growth suppressor signals and their resistance to cell death previously discussed.

The mitogenic signalling in cancer cells is better understood than the proliferative signals functioning within normal tissue and can occur in a variety of ways (Lemmon & Schlessinger, 2010; Witsch, Sela, & Yarden, 2010). The cell may produce growth factor ligands and stimulate autocrine proliferation, resulting in transactivation of epidermal growth factor receptor (EGFR) and other pathways that contribute to the proliferation of tumour cells. Cancer cells also have the capacity to stimulate normal tissue cells within the stroma/microenvironment (extracellular matrix (ECM)), which in turn promotes cancer cell growth factors, giving rise to a vicious cycle. In addition receptor proteins on the cancer cell surface can be elevated, enhancing their responsiveness to ligand growth factor and downregulating receptor cell signalling (Hanahan & Weinberg, 2011).

Furthermore, cancer cell proliferation is increased in the presence of oncogene mutation. Mutations in either the upstream of downstream Ras signalling pathways are associated with cancer progression and it appears that the deregulation of Ras signalling is essential for tumourgenesis (Fernández-Medarde & Santos, 2011). Although, Ras mutations (H-Ras, N-Ras or K-Ras) are less common in high incidence cancers such as breast, prostate and liver, high levels of H-Ras in breast cancer indicate a worse prognosis (Watson et al., 1991).

0.2.6 Hallmark Six: Altered Metabolism

Sustaining uncontrolled cell proliferation requires alterations in energy metabolism in order to fuel cell growth and division. Cancer cells can alter their energy production by limiting metabolism essentially to glycolysis despite oxygen availability (Hanahan & Weinberg, 2011). It may seem counterintuitive for cancer cells to alter their metabolism in this way, since glycolysis is known to be less efficient than mitochondrial oxidative phosphorylation, however, cancer cells compensate by upregulating glucose transporters (GLUT1) (Jones & Thompson, 2009). Fuelling in such a way activates oncogenes (RAS & MYC) and mutant tumour suppressors (TP53/p53), enabling the capabilities of other cancer hallmarks. Furthermore, the hypoxic environment creates an even greater reliance on glycolysis and subsequent

upregulation of glucose transporters and glycolytic enzymes. Consequentially, lactate is produced and is a preferential source of energy in a subpopulation of oxygenated cells that utilize the citric acid cycle for energy production, to fuel growth and division. Both increased oncoprotein (RAS) and hypoxia independently increase levels of transcription factor (HIF1alpha & HIF2alpha) and activate glycolysis (Hanahan & Weinberg, 2011; Jones & Thompson, 2009).

0.2.7 Hallmark Seven: Avoiding Immune Destruction

Immune surveillance is responsible for recognising and eliminating foreign matter, therefore, for tumour cells to develop they must go undetected. Early studies in genetically engineered mice have alluded that deficiencies in CD+8 cytotoxic T lymphocytes (CTLs), CD4+Th1 helper T cells or natural killer (NK) cells increased tumour incidence, suggesting that deficiencies in innate and/or active immunity contributes to tumour development (R. Kim, Emi, & Tanabe, 2007; Teng, Swann, Koebel, Schreiber, & Smyth, 2008).

0.2.8 Hallmark Eight: Tumour Promoting Inflammation

The presence of immune cells in most cancer types was historically believed to represent the bodies anti-tumour response, and in many cases this is true. Though, evidence suggests that the tumour associated inflammatory response paradoxically accelerates tumourgenesis and assisting in the numerous hallmark capabilities discussed by supplying bioactive molecules to the tumour microenvironment (DeNardo, Andreu, & Coussens, 2010; Grivennikov, Greten, & Karin, 2010; Hanahan & Weinberg, 2011; Qian & Pollard, 2010). It is understood that inflammation acts throughout all stages of tumourgenesis and may contribute to tumour initiation through mutation, genomic instability, and epigenetic modification. Inflammation initiates the repair response, aids in the proliferation of cells including premalignant cells, and improves their likelihood of survival.

The most frequently identified immune cell in the tumour microenviornemnt is tumour associated macrophages (TAMs) and T cells. The presence of TAM

has been shown to increase tumour growth and higher concentrations of TAM is associated with poorer prognosis (Murdoch, Muthana, Coffelt, & Lewis, 2008). While T cells can either be tumour suppressive or promoting, therefore, in some cancers increased T cells is associated with improved survival, however, many of the T cell subsets (CD8+, IFNgamma producing TH-1 and Th17) are involved in tumour promotion (Grivennikov et al., 2010).

Some oncogenes, namely the RAS and MYC family, induce a transcriptional program that recruits leukocytes, increases tumour promoting chemokine and cytokine expression and triggers the angiogenic 'switch', all of which lead to the remodelling of the tumour microenvironment (Grivennikov et al., 2010). At some point during a tumours development it becomes starved of oxygen and nutrients, resulting in necrotic cell death and the release of proinflammatory mediators (IL-1 and HMGB1) (Vakkila & Lotze, 2004). The proinflammatory mediators produced by this process encourage neoangiogenesis and increase the availability of growth factor for surviving cancer cells.

In addition to the cancer cells and surrounding stroma (fibroblasts, endothelial cells, and mesenchymal cells) the tumour microenvironment, as a result of inflammation responses, contains innate (macrophages, neutrophils, mast cells, natural killer T cells (NKT) among others) and adaptive (T and B Lymphocytes) immune cells (De Visser, Eichten, & Coussens, 2006). These cells are able to communicate with one another via cytokine or chemokine production and/or direct contact and it is the balance of these cells which dictates whether pro- or anti-tumour immunity ensues.

0.2.9 Hallmark Nine: Induction of Angiogenesis

Tumour cells, like any cell, require nutrients and oxygen to survive. To meet the demand for nutrients, oxygen and immune cells, and in order to remove waste products that are associated with cancer metabolism, the tumour develops its own vascular network, this process is referred to as angiogenesis. The process of angiogenesis is triggered by the increased nutrient and oxygen requirements of tumour tissue and occurs in stages. Initially, the basement membrane becomes

injured and hypoxic. In response, angiogenic factors activate endothelial cells which assemble into tubes (vasculogenesis) and begin to sprout new vessels from the existing ones. Though, the up regulation of angiogenic factors alone is not enough for angiogenesis of the neoplasm, there must also be simultaneous down regulation of inhibitors (Nishida, Yano, Nishida, Kamura, & Kojiro, 2006). Angiogenic factors continue to influence the angiogenic process, leading to normal vasculature dormancy. In normal cells angiogenesis is required for tissue repair but is capable of "switching off", where in cancer cells it remains "on" (Hanahan & Folkman, 1996).

The up regulation of vascular endothelial growth factor-A (VEGF-A) has been well-established as an angiogenesis inducer, other proteins capable of inducing angiogenesis include basic fibroblast growth factor (bFGF), transforming growth factor (TGF) alpha and beta, tumour necrosis factor (TNF)-alpha, interleukin-8, epidermal growth factors among others. Ligands encoded by VEGF-A genes are partially responsible for the development of new blood and over expression of VEGF can be caused by both hypoxia and oncogene signalling (Ferrara, 2009; Gabhann & Popel, 2008). Hypoxia occurs due to the growing distance between tumour cells and capillaries. In response, the tumour cells secrete VEGF into the surrounding tissue activating endothelial cells to produce metalloproteinases (MMPs), which break down the extracellular matrix and fills the void between cells (Nelson, Fingleton, Rothenberg, & Matrisian, 2000; Nishida et al., 2006). After endothelial cells divide and migrate, they form tubes that, with the help of integrin alpha and beta, mature into new blood vessels (Nelson et al., 2000). A number of studies have reported a strong correlation between VEGF expression and cancer prognosis (Nishida et al., 2006), with the following studies demonstrating this in breast cancer (Kinoshita et al., 2001; Kurebayashi et al., 1999; Skobe et al., 2001). In addition to increased VEGF expression, upregulated proangiogenic signals, such as fibroblast growth factor, sustain angiogenesis (Baeriswyl & Christofori, 2009).

As previously discussed, the upregulation of VEGF and other proangiogenic signals is not enough to maintain angiogenesis, there must be simultaneous down

regulation of angiogenesis inhibitors. Thrombospondin-1 (TSP-1), an angiogenesis inhibitor, is down regulated when it binds to transmembrane receptors displayed by the endothelial cells (Hanahan & Weinberg, 2011). Downregulation of other inhibitors may include angiostatin, endostatin, interferon, platelet factor 4, prolactin 16 kd fragment, and tissue inhibitor of metalloproteinase-1, -2, and -3 (Nishida et al., 2006).

0.2.10 Hallmark Ten: Activation of Invasion and Metastasis

A series of successive cellular changes, often termed the invasion-metastasis cascade, leads to the invasion and metastasis of cancer cells. The multistep process can be broken down into the following steps: invasion, intravasation, surviving systematic circulation and extravasation. Epithelial-mesenchymal transition (EMT) has been identified as a regulatory program that possesses the ability to invade, resist apoptosis and disseminate transformed epithelial cells (Klymkowsky & Savagner, 2009; Polyak & Weinberg, 2009; Thiery, Acloque, Huang, & Nieto, 2009; Yilmaz & Christofori, 2009). Intracellular signalling pathways (TGF-beta, growth factors that bind with RTK and WNT ligands) are triggered when ligands bind to transmembrane receptors. These signals are expressed in the tumour microenvironment, where tumour cells undergo EMT (Ye & Weinberg, 2015). Genetic variation may also contribute to tumour cell EMT as evidenced by Neve and colleagues (2006) who found mesenchymal features in the absence of extrinsic signalling (Neve et al., 2006). In both instances activating EMT transcription factors that repress epithelial cell to cell adhesion whilst simultaneously promoting mesenchymal factors that alter cytoskeleton, protrusion formation and cell migration (Pearson, 2019).

The suppression of cell to cell adhesion protein E-cadherin, which has both invasive traits and metastatic properties, is the most established mechanism associated with cell migration (Pearson, 2019; Zaidel-Bar & Geiger, 2010). The reduced E-cadherin in transformed epithelial cells is enough to induce EMT, invasion and metastasis, however, the loss alone is not substantial enough to promote EMT in all circumstances (Khalil et al., 2017; Onder et al., 2008). EMT transcription

factors (EMT-TF) also induce mesenchymal genes that are capable of altering cell morphology, migration and ECM remodelling (Thiery et al., 2009). Changes in cell morphology allow EMT cells through the ECM via existing tracks created by non-tumour populations in the microenvironment (Yamaguchi, Wyckoff, & Condeelis, 2005), the EMT program itself allows tumour cells to remodel the ECM, facilitating initial invasion from ductal structures, migration and intravasation into blood vessels (Pearson, 2019).

The signalling that occurs between cancer cells and connective tissue cells of the organ (stroma) are capable of stimulating invasive behaviours. Mesenchymal stem cells (MSCs) in the tumour stroma release CCL5/RANTES in response to these signals, which has been found to encourage changes in the phenotype of cancer cells and lead to metastatic spread (Karnoub et al., 2007). Matrix degrading enzymes, found in the periphery of the tumour can further facilitate invasion (Hanahan & Weinberg, 2011). In metastatic breast cancer, epidermal growth factor (EGF) is supplied to the breast cancer cells by tumour-associated macrophages (TAMs), in response cancer cells syntheses CF-1 which stimulates macrophage production and the subsequent intravasation and dissemination of cancer cells (Qian & Pollard, 2010; Wyckoff et al., 2007). Macrophages also secrete protein, rich in cysteine (SPARC/osteonectin), that promote invasion and adhesion to other components of the ECM, and are required for spontaneous metastasis (Sangaletti et al., 2008). This colocalization of macrophages, tumour cells, and endothelial cells (tumour microenvironment of metastasis (TMEM)), has been shown to be a prognostic marker for poor survival in breast cancer (Robinson et al., 2009).

Metastasis occurs in two distinct phases; the physical dissemination of cells from a local region to distant tissue and the adaptation of cells to foreign microenvironments, that results in successful colonization. Colonization is a process of great complexity and while in some cancers the suppressor factors released by the tumour render metastases dormant until resection (Demicheli, Retsky, Hrushesky, Baum, & Gukas, 2008), in others, such as breast cancer, metastases may develop decades

after initial treatment (surgical or pharmacological treatment) (Barkan, Green, & Chambers, 2010).

0.3 Epidemiology

Population growth and age distribution are factors associated with the rapidly increasing cancer incidence and mortality worldwide, as well as changes in the prevalence of risk factors, many of which are associated with socioeconomic development. The Global Cancer Observatory estimated the worldwide population in 2018 to be more than 7.5 billion, of which over 43.5 million people were affected by cancer (5-year prevalence), 18 million new cases of cancer were reported, and more than 9.5 million deaths occurred. The most common cancers are lung, breast, colorectal and prostate, cumulatively these account for over half of new cancers (42.7%) (Bray et al., 2018). The majority of new cancer cases were in men (52%), where lung cancer was the most commonly occurring (15.5%), followed closely by prostate (14.5%) and colorectal (11.4%) cancers.

Forty-eight percent of new cancer cases in 2018 were in in women, with the most common cancer being breast cancer, which accounted for 25.4% of all cancer cases and was over twice as common as its closest rivals; colorectal (9.7%) and lung cancer (8.8%) (Bray et al., 2018).

Approximately two million cases of breast cancer are diagnosed every year, the equivalent of one new case every 16 seconds. Breast cancer is the most frequently diagnoses cancer in women across all continents, ranking highest in 154 of 185 countries, despite differences in per capita income (Bray et al., 2018). Incidence rates vary depending on regional income, with higher income countries experiencing more breast cancer cases.

Having increased year on year, the European incidence rate reached 136.0 per 100,000 in 2018, with over five hundred newly diagnosed cases, an incidence rate five times greater than that of breast cancer incidence in Africa (26.2 per 100,000), three times greater than Asia (41.1 per 100,000) and double Latin America and the

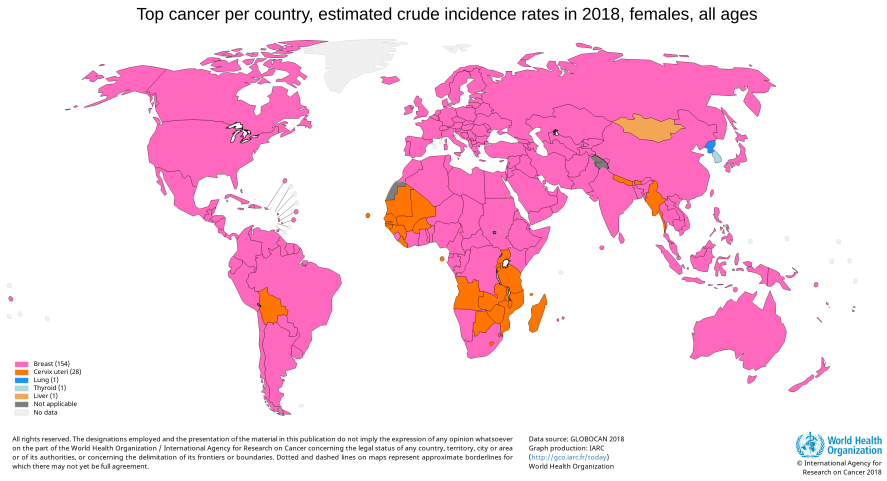

Figure 2: Top Cancer per Country, Estimated Crude Incidence Rates in 2018, Females, All Ages
(#fig:Global Cancer Incidence)

Caribbean (60.6 per 100,000). The only continent to experience greater incidence rates was North America, where the incidence rate is marginally bigger at 142.8 per 100,000. The highest incidence rates are seen in Belgium (203.7 per 100,000), The Netherlands (188.9 per 100,000), Italy (187.8 per 100,000), Luxembourg (175.4 per 100,000) and Germany (172.2 per 100,000). The United Kingdom registered 55,439 new breast cancer cases in 2018 and ranked ninth globally with an incidence rate 164.5 per 100,000. The lowest reported incidence is in Bhutan with 16 new cases and an incidence rate of 4.2 per 100,000, a 48.5-fold difference between the highest and lowest ranking countries. Breast cancer is also the leading cause of cancer mortality in over 100 countries (103 of 185) (Figure 4) (Bray et al., 2018).

Breast cancer accounts for 15% of all cancer related deaths worldwide and mortality rates are greatest in Europe (35.8 per 100,000) and North America (25.6 per 100,000). There appears to be less variability in mortality rates compared to incidence rates globally, with a 30-fold difference between the highest (Barbados; 62.9 per 100,000) and the lowest (Bhutan; 2.1 per 100,000) mortality rate. The United Kingdom reported 11,849 deaths in 2018 and ranked twenty-fourth in breast cancer mortality globally, with a mortality rate of 246.2 per 100,00.

Medical advancements in breast cancer screening and treatment have seen drastic

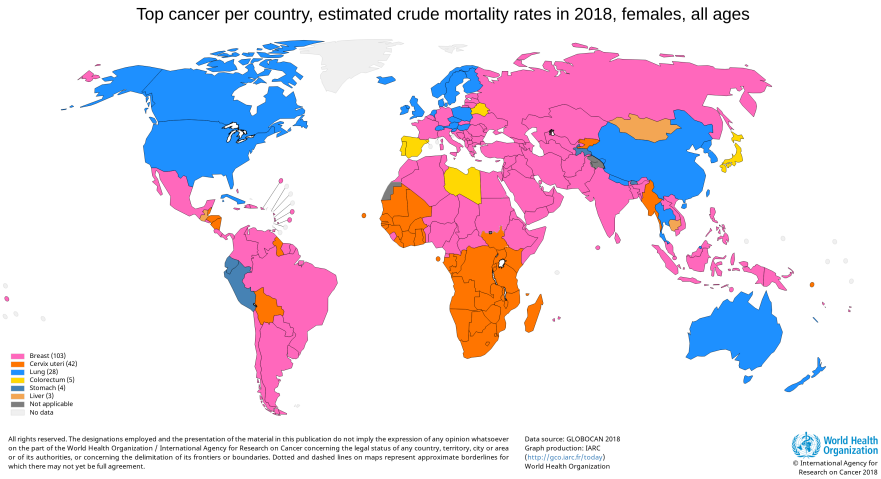

Figure 3: Top Cancer per Country, Estimated Crude Mortality Rates in 2018, Females, All Ages
(#fig:Global Cancer Mortality)

improvements in breast cancer survival rates, and as such increased breast cancer prevalence (Table 1). The breast cancer survival rate varies globally; with an estimated 5-year survival rate in developed countries of 80%, whereas in middle and low-income countries it is 60% and 40% respectively, which may be related to screening programs, early detection, and modern medical care factors (Maajani et al., 2019). It is not surprising that the stage of breast cancer diagnosis greatly influences the survival rate of patients, with one year survival decreasing by 3% (95% CI 0, 8%), 11% (95% CI 22%, 31%) and 40% (95% CI 39%, 41%) in those with stage II, stage III, and stage IV diagnosis respectively. This trend was consistent in the five year survival rates, with a rate of survival in stage I patients of 0.86 (95% CI 0.82, 0.88), stage II patients of 0.69 (95% CI 0.63, 0.74), stage III patients of 0.51 (0.45, 0.57) and stage IV patients of 0.32 (95% CI 0.23, 0.42) and ten year survival rates, with a rate of survival in stage I patients of 0.82 (95% CI 0.70, 0.92), stage II patients of 0.67 (95% CI 0.59, 0.74), stage III patients of 0.41 (95% CI 0.35, 0.47) and stage IV patients of 0.26 (95% CI 0.10, 0.45).

Hormonal, immunologic, and etiologic differences associated with aging are also prognostic factors for survival, with one, three, five and ten year survival being lower in women over the age of 50 years (1-year: 0.92 (95% CI 0.90, 0.93), 3-year: 0.80

(95% CI 0.75, 0.85), 5-year: 0.61 (95% CI 0.66, 0.73), 10-Year: 0.48 (95% CI 0.39, 0.58) compared with younger patients (1-year: 0.97 (95% CI 0.95, 0.98), 3-year: 0.87 (95% CI 0.84, 0.90), 5-year: 0.69 (95% CI 0.66, 0.73), 10-year: 0.61 (95% CI 0.55, 0.67) (Maajani et al., 2019). Survival is also greater in patients with hormone receptor positive and HER2-positive tumours, largely due to the effectiveness of aromatase inhibitors and trastuzumb treatments (Berry et al., 2005).

0.4 Breast Cancer Pathophysiology

Immunohistochemistry markers – such as estrogen receptor (ER), progesterone receptor (PR) and epidermal growth factors receptor 2 (HER-2) – mediate cell growth in breast cancers. Breast tumours can be categorised according to the arrangement of these biomarkers; [ER+|PR+|HER2- (ER or PR positive, and HER2 negative), [ER+|PR+]HER2+ (ER or PR positive, and HER2 positive), ER-PR-HER2+ (ER and PR negative, and HER2 positive), and ER-PR-HER2- (ER, PR, HER2 negative).

0.4.1 Sustained Proliferation

Estrogen receptor (ER), plays a critical role in carcinogenesis and is consider the most important biomarker in breast cancer classification. It is estimated that ER positive tumours account for 75% of all breast cancers and compared with ER negative tumours are associated with better surgical outcomes (Dunnwald, Rossing, & Li, 2007; Putti et al., 2005). ER positive and ER negative tumours differ substantially in transcriptional level, complexity of genetic aberrations, and pathways (Dai, Xiang, Li, & Bai, 2016). It remains the single most formidable predicative factor in breast cancer patients. In comparison, the predicative role of progesterone receptor (PR) positive tumours has been brought into question (Dai et al., 2016).

PR positive tumours are induced by endocrine and are therefore indicative of ER signalling; it comes as no surprise then, that breast tumours are rarely ER negative (0.2-10.0% depending on detection methods). Approximately 55% to 65% of breast

cancers have the double positive arrangement of ER and PR (ER+PR+) (Dunnwald et al., 2007; Rakha et al., 2007), 40% are ER positive PR negative (ER+PR-)(Rakha et al., 2007) and 18% to 25% are double negative (ER-PR-) (Dunnwald et al., 2007). Typically, patients with ER+PR+ tumours are older and have a smaller, lower grade tumour. In comparison, ER-PR- tumours are more aggressive in terms of size, grade, stage, outcome and hormonal therapy response (Anderson, Chu, Chatterjee, Brawley, & Brinton, 2001; Elledge et al., 2000; Kinne et al., 1987; Parl, Schmidt, Dupont, & Wagner, 1984; Ravdin et al., 1992). The arrangement of ER+/- and PR+/- may be combined with human epidermal growth factor receptor 2 (HER2). The over expression of HER2 occurs in approximately 13% to 20% of invasive ductal breast cancers; half of which are ER-PR- (Dai et al., 2016).

It is widely accepted that increased cell proliferation is a key determinant and hallmark of cancer. KI67 is the most commonly used marker of proliferation in breast cancer. It is often used to predict the response to neo- and adjuvant (chemotherapy or endocrine) treatment, in which KI67 positive tumours are associated with increased cell proliferation and poorer prognosis regardless of treatment modality (Cheang et al., 2009). TOP2A is associated with KI67 and frequently expressed in HER2 positive breast cancers and has been suggested as a potential biomarker (Dai et al., 2016). Beside KI67 and TOP2A, cell cycle genes are strongly correlated with proliferation and there over expression associated with poorer prognosis. Thus, genes involved in the cell cycle can be used to classify [ER+|PR+] tumours into distinct groups; those likely to respond to hormone therapies.

0.4.2 Activation of Invasion and Metastasis

ER-PR-HER2- can be dichotomised into two subtypes that differ in behaviour, outcome and treatment response; ER-PR-HER2- basal and non-basal phenotype (Dai et al., 2016). A number of immunohistochemistry markers can be used to detect basal phenotypes, including cytokeratins (CK 5/6, 14, 17, 8/18) and epidermal growth factor receptor (EGFR) (Carey et al., 2006; Matos, Dufloth, Alvarenga, Zeferino, & Schmitt, 2005; Neilson, Friedenreich, Brockton, & Millikan, 2009;

Rakha & Ellis, 2009; Rakha et al., 2009). While differences in the expression of these immunohistochemistry markers presents a variety of combinations and basal subtypes, the generally accepted basal subtype is 'ER-PR-HER2- tumours with positive CK5/6 and EGFR' (Carey et al., 2006). Proteins of the mesenchymal cell regulate cytokeratins and cell adhesion molecules which are associated with invasion and metastasis. Possessing one, or more, basal keratins are more likely to contain a dysfunctional breast cancer type 1 susceptibility protein (BRCA1) (Turner et al., 2007). Although, basal subtypes are typically expressed in ER-PR-HER2- tumours, basal markers are present in a small proportion (1% to 18%) of [ER+|PR+] and HER2 positive tumours and are associated with poorer prognosis (Dai et al., 2016).

Epithelial to mesenchymal transition (EMT), previously discussed, is a reversible process necessary in tumour progression and metastasis. In breast cancer, EMT markers include VIM, SNAI1, SNAI2, TWIST1, TWIST2, ZEB1, ZEB2, CDH1, and claudin (CLDN3, CLDN4, CLDN7) (West et al., 2005); the proportion of these markers determines whether a breast cancer (typically ER-PR-HER2-) is considered claudin-low or metaplastic. While the two share many similarities, metaplastic cancers are distinguished by PIK3CA, AKT or KRAS mutations and account for up to 1% of breast tumours; whereas, low-claudin cancers have, unsurprisingly, low CDH1, CLDN3, CLDN4, CLDN7 expression, and account for approximately 7% to 14% of breast tumours (Dai et al., 2016; Hennessy et al., 2009).

0.4.3 Avoiding Immune Destruction

Neque porro quisquam est qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit...

There is no one who loves pain itself, who seeks after it and wants to have it, simply because it is pain...

— Cicero's *de Finibus Bonorum et Malorum.*

1

R Markdown Basics: The Markdown syntax

Contents

Here is a brief introduction to using *R Markdown*. *Markdown* is a simple formatting syntax for authoring HTML, PDF, and MS Word documents and much, much more. *R Markdown* provides the flexibility of *Markdown* with the implementation of **R** input and output. For more details on using *R Markdown*

see `http://rmarkdown.rstudio.com`.

Be careful with your spacing in *Markdown* documents. While whitespace largely is ignored, it does at times give *Markdown* signals as to how to proceed. As a habit, try to keep everything left aligned whenever possible, especially as you type a new paragraph. In other words, there is no need to indent basic text in the Rmd document (in fact, it might cause your text to do funny things if you do).

1.1 Markdown basic syntax

1.1.1 Italics and bold

- *Italics* are done like *this* or _this_
- **Bold** is done like **this** or ___this___
- ***Bold and italics*** is done like ***this***, _____this_____, or (the most transparent solution, in my opinion) **_this_**

1.1.2 Inline code

- `Inline code` is created with backticks like `` `this` ``

1.1.3 Sub and superscript

Sub_2 and super^2 script is created like this~2~ and this^2^

1.1.4 Strikethrough

- ~~Strikethrough~~ is done ~~like this~~

1.1.5 'Escaping' (aka "What if I need an actual asterisk?")

- To include an actual *, _ or \, add another \ in front of them: *, _, \\

1.1.6 Endash (–), emdash (—)

- – and — with -- and ---

1.1.7 Blockquotes

Do like this:

> Put a > in front of the line.

1.1.8 Headings

- are done with #'s of increasing number, i.e.
  - # First-level heading
  - ## Second-level heading
  - ### Etc.

In PDF output, a level-five heading will turn into a paragraph heading, i.e. `\paragraph{My level-five heading}`, which appears as bold text on the same line as the subsequent paragraph.

1.1.9 Lists

Unordered list by starting a line with an * or a -:

- Item 1
- Item 2

Ordered lists by starting a line with a number:

1. Item 1
2. Item 2

Notice that you can mislabel the numbers and *Markdown* will still make the order right in the output.

To create a sublist, indent the values a bit (at least four spaces or a tab):

1. Item 1
2. Item 2

3. Item 3

    - Item 3a
    - Item 3b

1.1.10 Line breaks

The official *Markdown* way to create line breaks is by ending a line with more than two spaces.

Roses are red. Violets are blue.

This appears on the same line in the output, because we didn't add spaces after red.

Roses are red.
Violets are blue.

This appears with a line break because I added spaces after red.

I find this is confusing, so I recommend the alternative way: Ending a line with a backslash will also create a linebreak:

Roses are red.
Violets are blue.

To create a new paragraph, you put a blank line.

Therefore, this line starts its own paragraph.

1.1.11 Hyperlinks

- This is a hyperlink created by writing the text you want turned into a clickable link in `[square brackets followed by a](https://hyperlink-in-parentheses)`

1.1.12 Footnotes

- Are created[1] by writing either ^[my footnote text] for supplying the footnote content inline, or something like `[^a-random-footnote-label]` and

[1] my footnote text

supplying the text elsewhere in the format shown below [2]:

```
[^a-random-footnote-label]: This is a random test.
```

1.1.13 Comments

To write comments within your text that won't actually be included in the output, use the same syntax as for writing comments in HTML. That is, <!-- this will not be included in the output -->.

1.1.14 Math

The syntax for writing math is stolen from LaTeX. To write a math expression that will be shown **inline**, enclose it in dollar signs. - This: $A = \pi*r^{2}$ Becomes: $A = \pi * r^2$

To write a math expression that will be shown in a block, enclose it in two dollar signs.
This: $$A = \pi*r^{2}$$

Becomes:

$$A = \pi * r^2$$

To create numbered equations, put them in an 'equation' environment and give them a label with the syntax `(\#eq:label)`, like this:

```
\begin{equation}
  f\left(k\right) = \binom{n}{k} p^k\left(1-p\right)^{n-k}
  (\#eq:binom)
\end{equation}
```

Becomes:

$$f\left(k\right) = \binom{n}{k} p^k\left(1-p\right)^{n-k} \tag{1.1}$$

For more (e.g. how to theorems), see e.g. the documentation on bookdown.org

[2] This is a random test.

1.2 Additional resources

- *R Markdown: The Definitive Guide* - `https://bookdown.org/yihui/rmarkdown/`
- *R for Data Science* - `https://r4ds.had.co.nz`

2
Adding code

Contents

The magic of R Markdown is that we can add code within our document to make it dynamic.

We do this either as *code chunks* (generally used for loading libraries and data, performing calculations, and adding images, plots, and tables), or *inline code* (generally used for dynamically reporting results within our text).

2.1 Code chunks

The syntax of a code chunk is shown in Figure 2.1.

Common chunk options include (see e.g. bookdown.org):

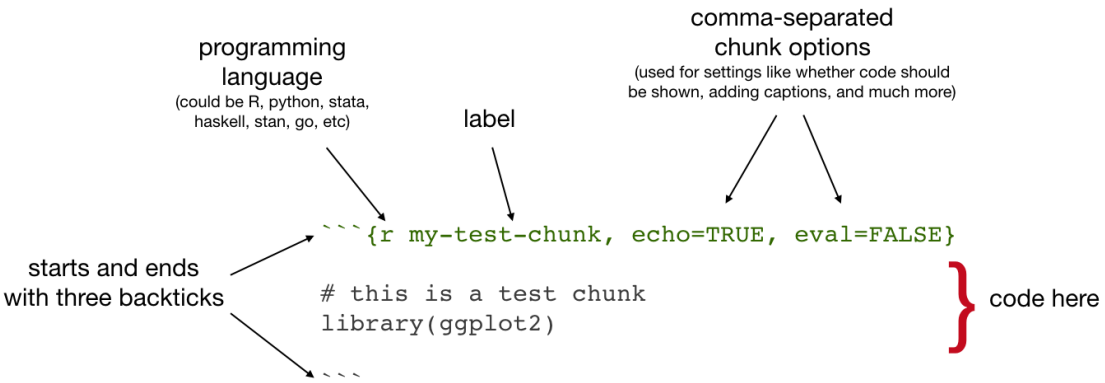

Figure 2.1: Code chunk syntax

- `echo`: whether or not to display code in knitted output
- `eval`: whether or to to run the code in the chunk when knitting
- `include`: wheter to include anything from the from a code chunk in the output document
- `fig.cap`: figure caption
- `fig.scap`: short figure caption, which will be used in the 'List of Figures' in the PDF front matter

IMPORTANT: Do *not* use underscoores in your chunk labels - if you do, you are likely to get an error in PDF output saying something like "! Package caption Error: \caption outside float".

2.1.1 Setup chunks

An R Markdown document usually begins with a chunk that is used to **load libraries**, and to **set default chunk options** with `knitr::opts_chunk$set`.

In your thesis, this will probably happen in **index.Rmd** and/or as opening chunks in each of your chapters.

````
```{r setup, include=FALSE}
# don't show code unless we explicitly set echo = TRUE
knitr::opts_chunk$set(echo = FALSE)
````
````

Figure 2.2: Oxford logo

```
library(tidyverse)
```
‘‘‘

2.1.2 Including images

Code chunks are also used for including images, with `include_graphics` from the `knitr` package, as in Figure 2.2

```
knitr::include_graphics("figures/beltcrest.png")
```

Useful chunk options for figures include:

- `out.width` (use with a percentage) for setting the image size
- if you've got an image that gets waaay to big in your output, it will be constrained to the page width by setting `out.width = "100%"`

Figure rotation

You can use the chunk option `out.extra` to rotate images.

The syntax is different for LaTeX and HTML, so for ease we might start by assigning the right string to a variable that depends on the format you're outputting to:

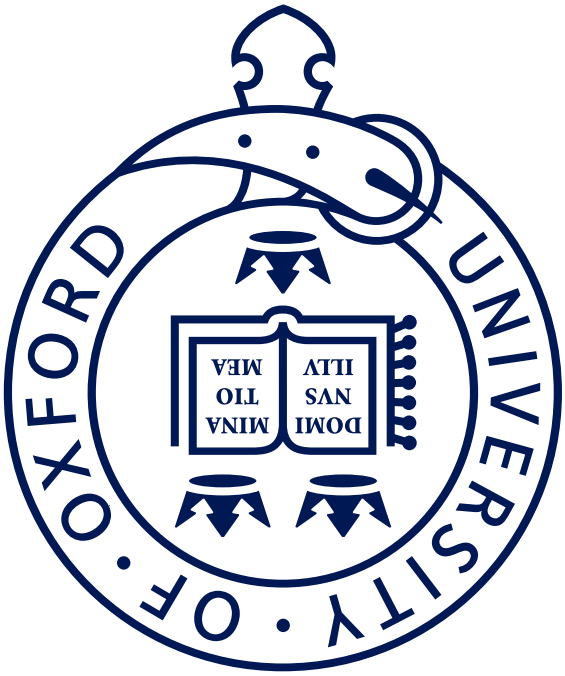

Figure 2.3: Oxford logo, rotated

```
if (knitr::opts_knit$get('rmarkdown.pandoc.to') == 'latex'){
  rotate180 <- "angle=180"
} else {
  rotate180 <- "style='transform:rotate(180deg);'"
}
```

Then you can reference that variable as the value of `out.extra` to rotate images, as in Figure 2.3.

2.1.3 Including plots

Similarly, code chunks are used for including dynamically generated plots. You use ordinary code in R or other languages - Figure 2.4 shows a plot of the `cars` dataset of stopping distances for cars at various speeds (this dataset is built in to **R**).

```
cars %>%
  ggplot() +
    aes(x = speed, y = dist) +
    geom_point()
```

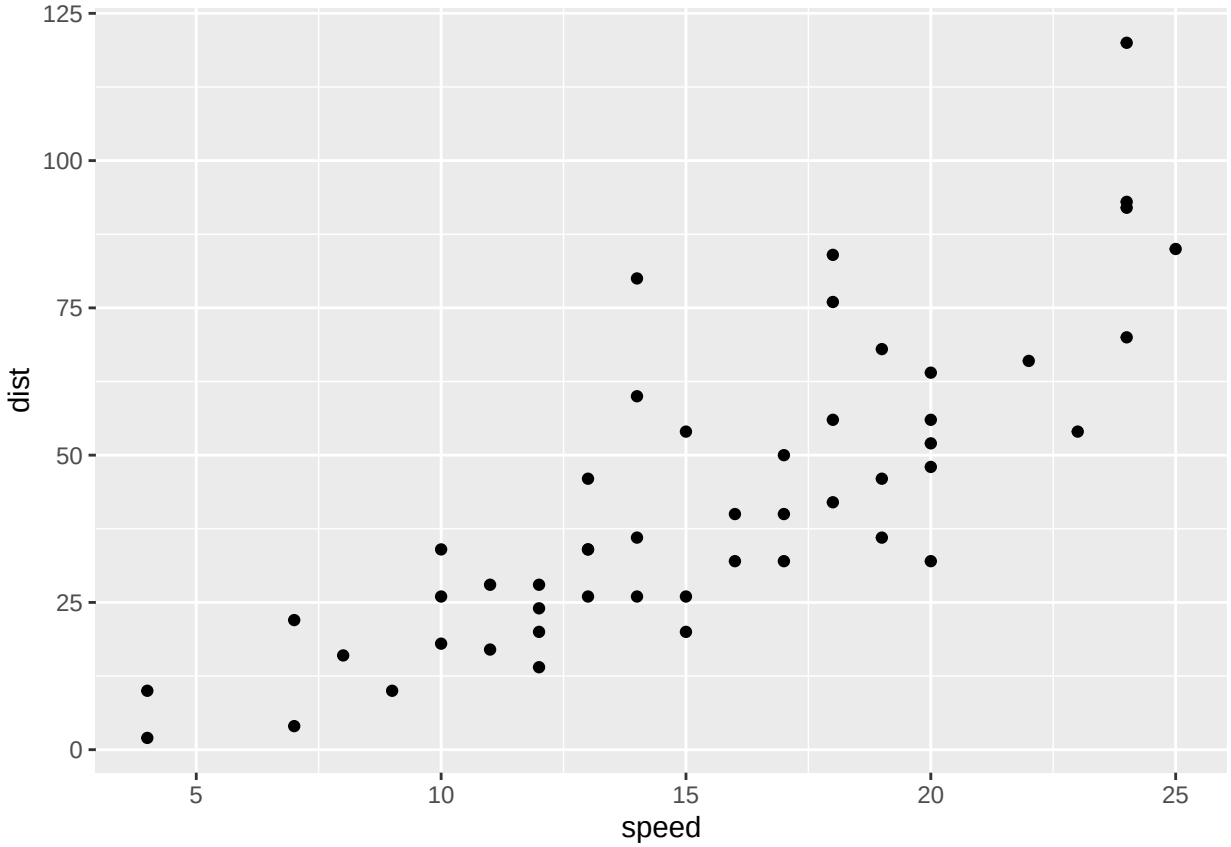

Figure 2.4: A ggplot of car stuff

Under the hood, plots are included in your document in the same way as images - when you build the book or knit a chapter, the plot is automatically generated from your code, saved as an image, then included into the output document.

2.1.4 Including tables

Tables are usually included with the `kable` function from the `knitr` package.

Table 2.1 shows the first rows of that cars data - read in your own data, then use this approach to automatically generate tables.

```
cars %>%
  head() %>%
  knitr::kable(caption = "A knitr kable table")
```

- Gotcha: when using `kable`, captions are set inside the `kable` function
- The `kable` package is often used with the `kableExtra` package

Table 2.1: A knitr kable table

speed	dist
4	2
4	10
7	4
7	22
8	16
9	10

2.1.5 A note on content positioning

One thing that may be annoying is the way *R Markdown* handles "floats" like tables and figures.

In your PDF output, LaTeX will try to find the best place to put your object based on the text around it and until you're really, truly done writing you should just leave it where it lies.

When the time comes for you to make final tweaks to content positioning, read the relevant R Markdown documentation to see if there are easy ways to do what you want.

If you have very specific needs, you might have to read up on LaTeX (`https://en.wikibooks.org/wiki/LaTeX/Floats,_Figures_and_Captions`) for your PDF output and/or on how to style HTML documents with CSS for your gitbook output.

2.2 Inline code

'Inline code' simply means inclusion of code inside text.

The syntax for doing this is `` `r R_CODE` ``

For example, `` `r 4 + 4` `` would output 8 in your text.

You will usually use this in parts of your thesis where you report results - read in data or results in a code chunk, store things you want to report in a variable, then insert the value of that variable in your text.

For example, we might assign the number of rows in the `cars` dataset to a variable:

```
num_car_observations <- nrow(cars)
```

We might then write:

"In the `cars` dataset, we have `` `r num_car_observations` `` observations."

Which would output:

"In the `cars` dataset, we have 50 observations."

2.2.1 Referring to results computed in other languages than R

I've commented the below section out, to avoid compilation errors from the `reticulate` package being unable to find a python installation (after I installed MacOS Catalina, `reticulate` was unable to select a python version on my system, and I had to set it manually with `use_python`).

If you need to use other langauges, have a look at the content I commented out by the end of the **02-rmd-basics-code.Rmd** file, which gives an example of using Python in your R Markdown file.

3

Citations and cross-references

Contents

3.1 Citations

The usual way to include citations in an *R Markdown* document is to put references in a plain text file with the extension **.bib**, in **BibTex** format.[1] Then reference the path to this file in **index.Rmd**'s YAML header with `bibliography: example.bib`.

[1]The bibliography can be in other formats as well, including EndNote (**.enl**) and RIS (**.ris**), see rmarkdown.rstudio.com/authoring_bibliographies_and_citations.

Most reference managers can create a .bib file with you references automatically. However, the **by far** best reference manager to use with *R Markdown* is Zotero with the Better BibTex plug-in, because the `citr` plugin for RStudio (see below) can read references directly from your Zotero library!

Here is an example of an entry in a **.bib** file:

```
@article{Shea2014,
  author =        {Shea, Nicholas and Boldt, Annika},
  journal =       {Trends in Cognitive Sciences},
  pages =         {186--193},
  title =         {{Supra-personal cognitive control}},
  volume =        {18},
  year =          {2014},
  doi =           {10.1016/j.tics.2014.01.006},
}
```

In this entry highlighed section, 'Shea2014' is the **citation identifier**. To default way to cite an entry in your text is with this syntax: `[@citation-identifier]`.

So I might cite some things (Shea et al. 2014; Lottridge et al. 2012).

3.1.1 PDF output

In PDF output, the bibliography is handled by the OxThesis LaTeX template. If you set `bib-humanities: true` in **index.Rmd**, then in-text references will be formatted as author-year; otherwise references will be shown as numbers.

If you choose author-year formatting, a number of variations on the citation syntax are useful to know:

- Put author names outside the parenthesis
  - This: `@Shea2014 says blah.`
  - Becomes: Shea et al. (2014) says blah.
- Include only the citation-year (in parenthesis)

  - This: `Shea et al. says blah [-@Shea2014]`
  - Becomes: Shea et al. says blah (2014)

- Add text and page or chapter references to the citation

  - This: `[see @Shea2014, pp. 33-35; also @Wu2016, ch. 1]`
  - Becomes: Blah blah (see Shea et al. 2014, pp. 33-35; also Wu 2016, ch. 1).

3.1.2 Gitbook output

In gitbook output, citations are by default inserted in the Chicago author-date format.

To change the format, add `csl: some-other-style.csl` in **index.Rmd**'s YAML header. You can browse through and download styles at zotero.org/styles.

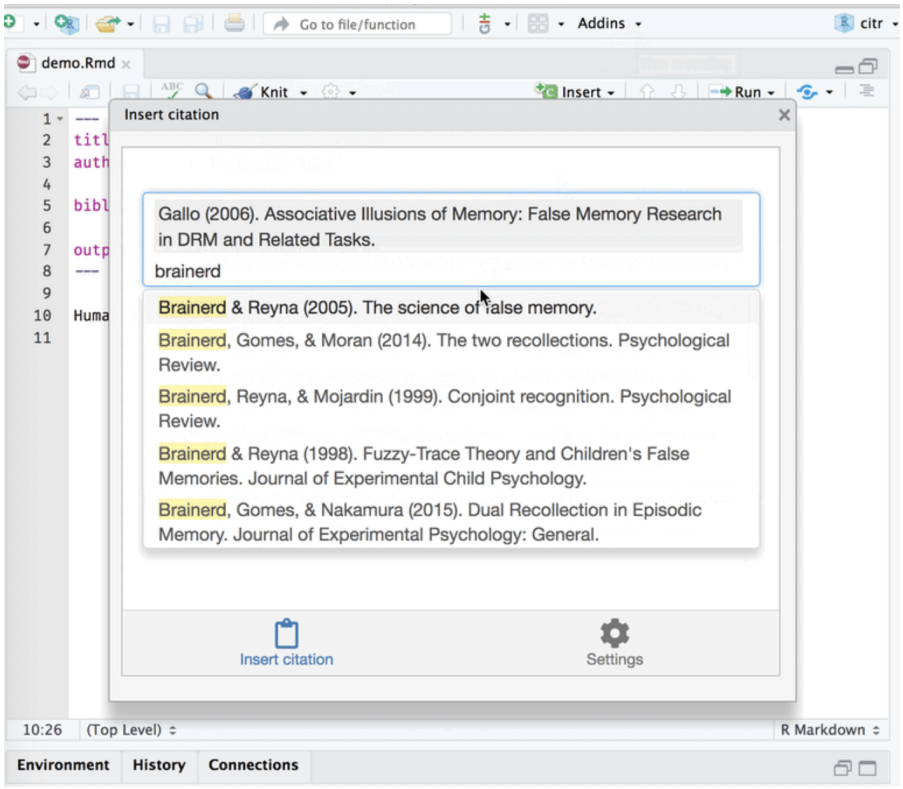

Figure 3.1: The 'citr' add-in

3.1.3 Insert references easily with the `citr` add-in

For an easy way to insert citations, try the `citr` RStudio add-in (Figure 3.1). You can install this add-in by typing `install.packages("citr")` in the R Console.

3.2 Cross-referencing

We can make cross-references to **sections** within our document, as well as to **figures** (images and plots) and **tables**.

The general cross-referencing syntax is **`\@ref(label)`**

3.2.1 Section references

Headers are automatically assigned a reference label, which is the text in lower caps separated by dashes. For example, `# My header` is automatically given the label `my-header`. So `# My header` can be referenced with `\@ref(my-section)`

Remember what we wrote in section 3.1?

We can also use **hyperlink syntax** and add # before the label, though this is only guaranteed to work properly in HTML output:

- So if we write `Remember what we wrote up in [the previous section](#citations)?`
- It becomes Remember what we wrote up in the previous section?

Creating custom labels

It is a very good idea to create **custom labels** for our sections. This is because the automatically assigned labels will change when we change the titles of the sections - to avoid this, we can create the labels ourselves and leave them untouched if we change the section titles.

We create custom labels by adding `{#label}` after a header, e.g. `# My section {#my-label}`. See our chapter title for an example. That was section 3.

3.2.2 Figure (image and plot) references

- To refer to figures (i.e. images and plots) use the syntax `\@ref(fig:label)`
- **GOTCHA**: Figures and tables must have captions if you wish to cross-reference them.

Let's add an image:

```
knitr::include_graphics("figures/captain.jpeg")
```

We refer to this image with `\@ref(fig:captain)`. So Figure 3.2 is this image.

And in Figure 2.4 we saw a cars plot.

3.2.3 Table references

- To refer to tables use the syntax `\@ref(tab:label)`

Let's include a table:

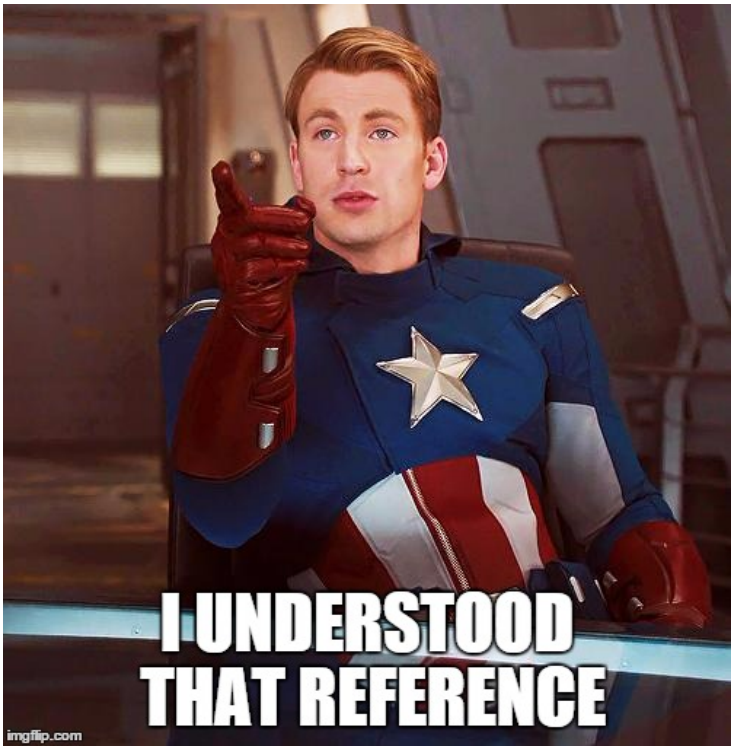

Figure 3.2: A marvel-lous meme

Table 3.1: Stopping cars

speed	dist
4	2
4	10
7	4
7	22
8	16

```
knitr::kable(cars[1:5,],
             caption="Stopping cars")
```

We refer to this table with `\@ref(tab:cars-table2)`. So Table 3.1 is this table.

And in Table 2.1 we saw more or less the same cars table.

3.2.4 Including page numbers

Finally, in the PDF output we might also want to include the page number of a reference, so that it's easy to find in physical printed output. LaTeX has a command for this, which looks like this: `\pageref{fig/tab:label}` (note: curly

braces, not parentheses)

When we output to PDF, we can use raw LaTeX directly in our .Rmd files. So if we wanted to include the page of the cars plot we could write:

- This: `Figure \@ref(fig:cars-plot) on page \pageref(fig:cars-plot)`
- Becomes: Figure 2.4 on page 29

Include page numbers only in PDF output

A problem here is that LaTeX commands don't display in HTML output, so in the gitbook output we'd see simply "Figure 2.4 on page".

One way to get around this is to use inline R code to insert the text, and use an `ifelse` statement to check the output format and then insert the appropriate text.

- So this: `` `r ifelse(knitr::is_latex_output(), "Figure \\@ref(fig:cars-plot) on page \\pageref{fig:cars-plot}", "")` ``
- Inserts this (check this on both PDF and gitbook): Figure 2.4 on page 29

Note that we need to escape the backslash with another backslash here to get the correct output.

3.3 Customising your thesis' front matter 'n stuff

3.3.1 Shorten captions shown in the list of figures (PDF)

You might want your list of figures (which follows the table of contents) to have shorter (or just different) figure descriptions than the actual figure captions.

Do this using the chunk option `fig.scap` ('short caption'), for example `{r captain-image, fig.cap="A very long and descriptive (and potentially boring) caption that doesn't fit in the list of figures, but helps the reader understand what the figure communicates.", fig.scap="A concise description for the list of figures"`

3.3.2 Shorten captions shown in the list of tables (PDF)

You might want your list of tables (which follows the list of figures in your thesis front matter) to have shorter (or just different) table descriptions than the actual table captions.

If you are using `knitr::kable` to generate a table, you can do this with the argument `caption.short`, e.g.:

```
knitr::kable(mtcars,
             caption = "A very long and descriptive (and potentially
             boring) caption that doesn't fit in the list of figures,
             but helps the reader understand what the figure
             communicates.",
             caption.short = "A concise description for the list of tables")
```

3.3.3 Shorting the running header (PDF)

You might want a chapter's running header (i.e. the header showing the title of the current chapter at the top of page) to be shorter (or just different) to the actual chapter title.

Do this by adding the latex command `\chaptermark{My shorter version}` after your chapter title.

For example, this chapter's running header is simply 'Cites and cross-refs', because it begins like this:

```
# Citations and cross-references {#cites-and-refs}
\chaptermark{Cites and cross-refs}
```

There is grandeur in this view of life, with its several powers, having been originally breathed into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved.

— Charles Darwin (Darwin 1859)

4

Final Notes on The OxThesis template and on collaboration

Contents

4.1 Beginning chapters with quotes

The OxThesis LaTeX template lets you inject some wittiness into your thesis by including a block of type `savequote` at the beginning of chapters. To do this, use the syntax ```` ```{block type='savequote'} ````.[1]

Add the reference for the quote with the chunk option `quote_author="my author name"`. You will also want to add the chunk option `include=knitr::is_latex_output()` so that quotes are only included in PDF output.

[1] For more on custom block types, see the relevant section in *Authoring Books with R Markdown*.

It's not possible to use markdown syntax inside chunk options, so if you want to e.g. italicise a book name in the reference use a 'text reference': Create a named piece of text with '(ref:label-name) My text', then point to this in the chunk option with `quote_author='(ref:label-name)'`.

4.2 Highlighting corrections

For when it comes time to do corrections, you may want to highlight changes made when you submit a post-viva, corrected copy to your examiners so they can quickly verify you've completed the task. You can do so like this:

4.2.1 Short, inline corrections

Highlight **short, inline corrections** by doing `[like this]{.correction}` — the text between the square brackets will then be highlighted in blue in the output.

4.2.2 Blocks of added or changed material

Highlight entire **blocks of added or changed material** by putting them in a block of type `correction`, using the syntax ` ```{block type='correction'} `.[2] Like so:

For larger chunks, like this paragraph or indeed entire figures, you can use the `correction` block type. This environment **highlights paragraph-sized and larger blocks** with the same blue colour.

4.2.3 Stopping corrections from being highlighted in the output

For **PDF** output, go to **index.Rmd** and (i) set `corrections: false` under `params` in the YAML header (stops block of corrections from being highlighted), (ii) comment out `pandoc_args: ["--lua-filter=scripts_and_filters/correction_filter.lua"]` (stops inline corrections from being highlighted).

[2] In the **.tex** file for PDF output, this will put the content between `\begin{correction}` and `\end{correction}`; in gitbook output it will be put between `<div class="correction">` and `</div>`.

For **gitbook** output, go to **style.css** and comment out the styling for `.correction`.

4.3 Diving in to the OxThesis LaTeX template

For LaTeX minded people, you can read through **templates/template.tex** to see which additional customisation options are available as well as **templates/ociamthesis.cls** which supplies the base class. For example, **template.tex** provides an option for master's degree submissions, which changes identifying information to candidate number and includes a word count. At the time of writing, you must set this directly in **template.tex** rather than from the YAML header in **index.Rmd**.

4.4 Collaborative writing

Best practices for collaboration and change tracking when using R Markdown are still an open question. In the blog post **One year to dissertate** by Lucy D'Agostino, which I highly recommend, the author notes that she knits .Rmd files to a `word_document`, then uses the `googledrive` R package to send this to Google Drive for comments / revisions from co-authors, then incorporates Google Drive suggestions *by hand* into the .Rmd source files. This is a bit clunky, and there are ongoing discussions among the *R Markdown* developers about what the best way is to handle collaborative writing (see issue #1463 on GitHub, where CriticMarkup is among the suggestions).

For now, this is an open question in the community of R Markdown users. I often knit to a format that can easily be imported to Google Docs for comments, then go over suggested revisions and manually incorporate them back in to the .Rmd source files. For articles, I sometimes upload a near-final draft to Overleaf, then collaboratively make final edits to the LaTeX file there. I suspect some great solution will be developed in the not-to-distant future, probably by the RStudio team.

5

Customisations and extensions

This chapter describes a number of possible customizations to the `oxforddown` thesis.

5.1 Embedding PDF documents as chapters

You may want to embed existing PDF documents into the thesis, for example if your department allows a 'portfolio' style thesis and you need to include an existing typeset publication as a chapter.

In gitbook output, you can simply use `knitr::include_graphics` and it should include a scrollable (and downloadable) PDF. You will probably want to set the chunk options `out.width='100%'` and `out.height='1000px'`:

```
knitr::include_graphics("figures/pdf_example/Lyngs2020_FB.pdf")
```

In LaTeX output, however, this approach can cause odd behaviour. Therefore, when you build your thesis to PDF, split the PDF into an alphanumerically sorted sequence of **single-page** PDF files (you can do this automatically with the package `pdftools`). You can then use the appropriate LaTeX command to insert them, as shown below (for brevity, in the `oxforddown` PDF sample content we're only including two pages) *Note that the chunk option* ***`results='asis'`*** *must be set.* You may also want to remove margins from the PDF files, which you can do with Adobe Acrobat (paid version) and likely other software.

```
# install.packages(pdftools)
# split PDF into pages stored in figures/pdf_example/split/
# pdftools::pdf_split("figures/pdf_example/Lyngs2020_FB.pdf",
# output = "figures/pdf_example/split/")

# grab the pages
pages <- list.files("figures/pdf_example/split", full.names = TRUE)

# set how wide you want the inserted PDFs to be:
# 1.0 is 100 per cent of the oxforddown PDF page width;
# you may want to make it a bit bigger
pdf_width <- 1.2

# for each PDF page, insert it nicely and
# end with a page break
cat(stringr::str_c("\\newpage \\begin{center}
    \\makebox[\\linewidth][c]{\\includegraphics[width=", pdf_width,
    "\\linewidth]{", pages, "}} \\end{center}"))
```

'I Just Want to Hack Myself to Not Get Distracted': Evaluating Design Interventions for Self-Control on Facebook

Ulrik Lyngs[1], **Kai Lukoff**[2], **Petr Slovak**[3], **William Seymour**[1], **Helena Webb**[1], **Marina Jirotka**[1], **Jun Zhao**[1], **Max Van Kleek**[1], **Nigel Shadbolt**[1]
[1]Department of Computer Science, University of Oxford, UK, {first.last}@cs.ox.ac.uk
[2]Human Centered Design & Engineering, University of Washington, Seattle, US, kai1@uw.edu
[3]Department of Informatics, King's College London, UK, petr.slovak@kcl.ac.uk

ABSTRACT
Beyond being the world's largest social network, Facebook is for many also one of its greatest sources of digital distraction. For students, problematic use has been associated with negative effects on academic achievement and general wellbeing. To understand what strategies could help users regain control, we investigated how simple interventions to the Facebook UI affect behaviour and perceived control. We assigned 58 university students to one of three interventions: goal reminders, removed newsfeed, or white background (control). We logged use for 6 weeks, applied interventions in the middle weeks, and administered fortnightly surveys. Both goal reminders and removed newsfeed helped participants stay on task and avoid distraction. However, goal reminders were often annoying, and removing the newsfeed made some fear missing out on information. Our findings point to future interventions such as controls for adjusting types and amount of available information, and flexible blocking which matches individual definitions of 'distraction'.

Author Keywords
Facebook; problematic use; self-control; distraction; ICT non-use; addiction; focus; interruptions

CCS Concepts
•**Human-centered computing** → **Empirical studies in HCI;**

INTRODUCTION

Research on 'Problematic Facebook Use' (PFU) has investigated correlations between Facebook use and negative effects on outcomes such as level of academic achievement [35] and subjective wellbeing [58, 57]. A cross-cutting finding is that negative outcomes are associated with difficulty at exerting self-control over use, as well as specific use patterns including viewing friends' wide-audience broadcasts rather than receiving targeted communication from strong ties [13, 58].

CHI '20, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6708-0/20/04.
http://dx.doi.org/10.1145/3313831.3376672

Much of this work has focused on self-control over Facebook use in student populations [2, 44, 46], with media multitasking research finding that students often give in to use which provides short-term 'guilty pleasures' over important, but aversive academic tasks [76, 88, 60]. In the present paper, we present a mixed-methods study exploring how two interventions to Facebook — goal reminders and removing the newsfeed — affect university students' patterns of use and perceived control over Facebook use. To triangulate self-report with objective measurement, our study combined usage logging with fortnightly surveys and post-study interviews.

We found that both interventions helped participants stay on task and use Facebook more in line with their intentions. In terms of use patterns, goal reminders led to less scrolling, fewer and shorter visits, and less time on site, whereas removing the newsfeed led to less scrolling, shorter visits, and less content 'liked'. However, goal reminders were often experienced as annoying, and removing the newsfeed made some participants fear missing out on information. After the study, participants suggested a range of design solutions to mitigate self-control struggles on Facebook, including controls for filtering or removing the newsfeed, reminders of time spent and of use goals, and removing features that drive engagement. As an exploratory study, this work should be followed by confirmatory studies to assess whether our findings replicate, and how they may generalise beyond a student population.

RELATED WORK

Struggles with Facebook use

Whereas many uses of Facebook offer important benefits, such as social support, rapid spread of information, or facilitation of real-world interactions [78], a substantial amount of research has focused on negative aspects [58]. For example, studies have reported correlations between patterns of Facebook use and lower academic achievement [77, 86], low self-esteem, depression and anxiety [51], feelings of isolation and loneliness [2], and general psychological distress [15]. Such 'Problematic Facebook Use' (PFU) has been studied under various names (including 'Facebook dependence' [87] and 'Facebook addiction'[5]), but a recent review summarised a common definition as 'problematic behaviour characterised by addictive-like symptoms and/or self-regulation difficulties related to Facebook use leading to negative consequences in personal and social life' [58].

REFERENCES

[1] Alexander T. Adams, Jean Costa, Malte F. Jung, and Tanzeem Choudhury. 2015. Mindless Computing: Designing Technologies to Subtly Influence Behavior. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 719–730. DOI: http://dx.doi.org/10.1145/2750858.2805843

[2] Sami Abdo Radman Al-Dubai, Kurubaran Ganasegeran, Mustafa Ahmed Mahdi Al-Shagga, Hematram Yadav, and John T. Arokiasamy. 2013. Adverse Health Effects and Unhealthy Behaviors among Medical Students Using Facebook. https://www.hindawi.com/journals/tswj/2013/465161/. (2013). DOI:http://dx.doi.org/10.1155/2013/465161

[3] All Party Parliamentary Group on Social Media and Young People's Mental Health and Wellbeing. 2019. *#NewFilters to Manage the Impact of Social Media on Young People's Mental Health and Wellbeing*. Technical Report. UK Parliament.

[4] Hunt Allcott, Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow. 2019. *The Welfare Effects of Social Media*. Working Paper 25514. National Bureau of Economic Research. DOI: http://dx.doi.org/10.3386/w25514

[5] Cecilie Schou Andreassen, Torbjørn Torsheim, Geir Scott Brunborg, and Staale Pallesen. 2012. Development of a Facebook Addiction Scale. *Psychological Reports* 110, 2 (apr 2012), 501–517. DOI: http://dx.doi.org/10.2466/02.09.18.PR0.110.2.501-517

[6] Yummy Apps. 2019. Todobook. (May 2019).

[7] Albert Bandura. 1982. Self-efficacy mechanism in human agency. *American Psychologist* 37, 2 (1982), 122–147. DOI: http://dx.doi.org/10.1037/0003-066x.37.2.122

[8] Fanni Bányai, Ágnes Zsila, Orsolya Király, Aniko Maraz, Zsuzsanna Elekes, Mark D. Griffiths, Cecilie Schou Andreassen, and Zsolt Demetrovics. 09-Jan-2017. Problematic Social Media Use: Results from a Large-Scale Nationally Representative Adolescent Sample. *PLOS ONE* 12, 1 (09-Jan-2017), e0169839. DOI: http://dx.doi.org/10.1371/journal.pone.0169839

[9] Elliot T Berkman, Cendri A Hutcherson, Jordan L Livingston, Lauren E Kahn, and Michael Inzlicht. 2017. Self-Control as Value-Based Choice. *Current Directions in Psychological Science* 26, 5 (2017), 422–428. DOI: http://dx.doi.org/10.1177/0963721417704394

[10] Walter R. Boot, Daniel J. Simons, Cary Stothart, and Cassie Stutts. 2013. The Pervasive Problem with Placebos in Psychology. *Perspectives on Psychological Science* 8, 4 (jul 2013), 445–454. DOI: http://dx.doi.org/10.1177/1745691613491271

[11] Amara Brook. 2011. Ecological Footprint Feedback: Motivating or Discouraging? *Social Influence* 6, 2 (April 2011), 113–128. DOI: http://dx.doi.org/10.1080/15534510.2011.566801

[12] Gharad Bryan, Dean Karlan, and Scott Nelson. 2010. Commitment Devices. *Annual Review of Economics* 2, 1 (Sept. 2010), 671–698. DOI:http://dx.doi.org/10.1146/annurev.economics.102308.124324

[13] Moira Burke and Robert E. Kraut. 2016. The Relationship Between Facebook Use and Well-Being Depends on Communication Type and Tie Strength. *Journal of Computer-Mediated Communication* 21, 4 (2016), 265–281. DOI: http://dx.doi.org/10.1111/jcc4.12162

[14] Moira Burke, Cameron Marlow, and Thomas Lento. 2010. Social Network Activity and Social Well-Being. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 1909–1912. DOI: http://dx.doi.org/10.1145/1753326.1753613

[15] Wenhong Chen and Kye-Hyoung Lee. 2013. Sharing, Liking, Commenting, and Distressed? The Pathway between Facebook Interaction and Psychological Distress. *Cyberpsychology, Behavior and Social Networking* 16, 10 (oct 2013), 728–734. DOI: http://dx.doi.org/10.1089/cyber.2012.0272

[16] Justin Cheng, Moira Burke, and Elena Goetz Davis. 2019. Understanding Perceptions of Problematic Facebook Use: When People Experience Negative Life Impact and a Lack of Control. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, 199:1–199:13. DOI: http://dx.doi.org/10.1145/3290605.3300429

[17] Jacob Cohen. 1992. A Power Primer. *Psychological Bulletin* 112, 1 (1992), 155–159. DOI: http://dx.doi.org/10.1037/0033-2909.112.1.155

[18] Anna L Cox, Sandy J J Gould, Marta E Cecchinato, Ioanna Iacovides, and Ian Renfree. 2016. Design Frictions for Mindful Interactions: The Case for Microboundaries. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. ACM, New York, NY, USA, 1389–1397. DOI: http://dx.doi.org/10.1145/2851581.2892410

[19] Helen Creswick, Liz Dowthwaite, Ansgar Koene, Elvira Perez Vallejos, Virginia Portillo, Monica Cano, and Christopher Woodard. 2019. "... They don't really listen to people". *Journal of Information, Communication and Ethics in Society* 17, 2 (May 2019), 167–182. DOI: http://dx.doi.org/10.1108/jices-11-2018-0090

[20] Angela L. Duckworth, Katherine L. Milkman, and David Laibson. 2018. Beyond Willpower: Strategies for Reducing Failures of Self-Control. *Psychological Science in the Public Interest* 19, 3 (Dec. 2018), 102–129. DOI: http://dx.doi.org/10.1177/1529100618821893

5.2 Customizing referencing

5.2.1 Using a .csl file with pandoc instead of biblatex

The `oxforddown` package uses biblatex in latex for referencing. It is also possible to use pandoc for referencing by providing a .csl file in the YAML header of **index.Rmd** (likely requiring commenting out the biblatex code in **templates/template.tex**). This may be helpful for those who have a .csl file describing the referencing format for a particular journal. However, note that this approach does not support chapter bibliographies (see Section 5.2.2).

```
csl: ecology.csl
```

5.2.2 Customizing biblatex and adding chapter bibliographies

This section provides one example of customizing biblatex. Much of this code was combined from searches on Stack Exchange and other sources (e.g. here).

In **templates/template.tex**, one can replace the existing biblatex calls with the following to achieve referencing that looks like this:

(Charmantier and Gienapp 2014)

Charmantier, A. and P. Gienapp (2014). Climate change and timing of avian breeding and migration: evolutionary versus plastic changes. Evolutionary Applications 7(1):15–28. doi: 10.1111/eva.12126.

```
\usepackage[backend=biber,
    bibencoding=utf8,
    refsection=chapter, % referencing by chapter
    style=authoryear,
    firstinits=true,
    isbn=false,
    doi=true,
    url=false,
    eprint=false,
```

```
    related=false,
    dashed=false,
    clearlang=true,
    maxcitenames=2,
    mincitenames=1,
    maxbibnames=10,
    abbreviate=false,
    minbibnames=3,
    uniquelist=minyear,
    sortcites=true,
    date=year
]{biblatex}
\AtEveryBibitem{%
  \clearlist{language}%
  \clearfield{note}
}

\DeclareFieldFormat{titlecase}{\MakeTitleCase{#1}}

\newrobustcmd{\MakeTitleCase}[1]{%
  \ifthenelse{\ifcurrentfield{booktitle}\OR\ifcurrentfield{booksubtitle}%
    \OR\ifcurrentfield{maintitle}\OR\ifcurrentfield{mainsubtitle}%
    \OR\ifcurrentfield{journaltitle}\OR\ifcurrentfield{journalsubtitle}%
    \OR\ifcurrentfield{issuetitle}\OR\ifcurrentfield{issuesubtitle}%
    \OR\ifentrytype{book}\OR\ifentrytype{mvbook}\OR\ifentrytype{bookinbook}%
    \OR\ifentrytype{booklet}\OR\ifentrytype{suppbook}%
    \OR\ifentrytype{collection}\OR\ifentrytype{mvcollection}%
    \OR\ifentrytype{suppcollection}\OR\ifentrytype{manual}%
    \OR\ifentrytype{periodical}\OR\ifentrytype{suppperiodical}%
    \OR\ifentrytype{proceedings}\OR\ifentrytype{mvproceedings}%
```

```
    \OR\ifentrytype{reference}\OR\ifentrytype{mvreference}%
    \OR\ifentrytype{report}\OR\ifentrytype{thesis}}
    {#1}
    {\MakeSentenceCase{#1}}}

% \renewbibmacro{in:}{}
% suppress "in" for articles
%
\renewbibmacro{in:}{%
  \ifentrytype{article}{}{\printtext{\bibstring{in}\intitlepunct}}}
%-- no "quotes" around titles of chapters/article titles
\DeclareFieldFormat[article, inbook, incollection, inproceedings, misc, thesis,
{title}{#1}
%-- no punctuation after volume
\DeclareFieldFormat[article]
{volume}{{#1}}
%-- puts number/issue between brackets
\DeclareFieldFormat[article, inbook, incollection, inproceedings, misc, thesis,
{number}{\mkbibparens{#1}}
%-- and then for articles directly the pages w/o any "pages" or "pp."
\DeclareFieldFormat[article]
{pages}{#1}
%-- for some types replace "pages" by "p."
\DeclareFieldFormat[inproceedings, incollection, inbook]
{pages}{p. #1}
%-- format 16(4):224--225 for articles
\renewbibmacro*{volume+number+eid}{
  \printfield{volume}%
  \printfield{number}%
```

```
  \printunit{\addcolon}
}
```

If you would like chapter bibliographies, in addition insert the following code at the end of each chapter, and comment out the entire REFERENCES section at the end of template.tex.

```
\printbibliography[segment=\therefsection,heading=subbibliography]
```

5.3 Customizing the page headers and footers

The following code, when it replaces the existing correpsonding code block in **ociamthesis.cls**, puts chapter number and title centered in the header and page number in the footer, centered. This may be desirable particularly when inserting PDF chapters, as the margins of the PDF may not exactly align with the left and right margins of the page, demarcated by the existing header and footer text. In the following code block, the original code is commented out where replaced.

```
\usepackage{fancyhdr}
\setlength{\headheight}{15pt}
\fancyhf{} % clear the header and footers
\pagestyle{fancy}
\renewcommand{\chaptermark}[1]{\markboth{\thechapter. #1}{\thechapter. #1}}
% \renewcommand{\chaptermark}[1]{\markboth{\thechapter. #1}{}}
\renewcommand{\sectionmark}[1]{\markright{\thesection. #1}}
\renewcommand{\headrulewidth}{0pt}
\fancyhead[CO]{\emph{\leftmark}}
\fancyhead[CE]{\emph{\rightmark}}
% \fancyhead[LO,RE]{}
% \fancyhead[LE,RO]{}
\fancyfoot[CO,CE]{\emph{\thepage}}

\fancypagestyle{plain}{\fancyhf{}\fancyfoot[C]{\emph{\thepage}}}
```

```
% JEM fix header on cleared pages for openright
\def\cleardoublepage{\clearpage\if@twoside \ifodd\c@page\else
   \hbox{}
   % \fancyhead[RE,LO]{}
   \fancyhead[CE,CO]{}
   \newpage
   \if@twocolumn\hbox{}\newpage
   \fi
   % \fancyhead[LO]{\emph{\leftmark}}
   % \fancyhead[RE]{\emph{\rightmark}}
   \fancyhead[CO]{\emph{\leftmark}}
   \fancyhead[CE]{\emph{\rightmark}}
   \fi\fi}
```

Alles Gescheite ist schon gedacht worden.
Man muss nur versuchen, es noch einmal zu denken.

All intelligent thoughts have already been thought;
what is necessary is only to try to think them again.

— Johann Wolfgang von Goethe (von Goethe 1829)

Conclusion

If we don't want Conclusion to have a chapter number next to it, we can add the `{-}` attribute.

More info

And here's some other random info: the first paragraph after a chapter title or section head *shouldn't be* indented, because indents are to tell the reader that you're starting a new paragraph. Since that's obvious after a chapter or section title, proper typesetting doesn't add an indent there.

Appendices

A The First Appendix

This first appendix includes an R chunk that was hidden in the document (using `echo = FALSE`) to help with readibility:

In 02-rmd-basics-code.Rmd

```
library(tidyverse)

knitr::include_graphics("figures/chunk-parts.png")
```

And here's another one from the same chapter, i.e. Chapter 2:

```
knitr::include_graphics("figures/beltcrest.png")
```

B

The Second Appendix, for Fun

References

Darwin, Charles (1859). *On the Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray.

Lottridge, Danielle et al. (2012). "Browser design impacts multitasking". In: *Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting*. DOI: 10.1177/1071181312561289.

Shea, Nicholas et al. (2014). "Supra-personal cognitive control and metacognition". In: *Trends in Cognitive Sciences* 18.4, pp. 186–193. DOI: 10.1016/j.tics.2014.01.006. URL: http://dx.doi.org/10.1016/j.tics.2014.01.006.

Von Goethe, Johann Wolfgang (1829). *Wilhelm Meisters Wanderjahre oder die Entsagenden*. de. Cotta.

Wu, Tim (2016). *The Attention Merchants: The Epic Scramble to Get Inside Our Heads*. Knopf Publishing Group.