Beyond the "Angelina Effect": A Longitudinal Analysis of Celebrity Breast Cancer Disclosures' Impact on News Media and Public Online Breast Cancer Information Seeking Outcomes, 2005-2016

A Dissertation
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Susan M. LoRusso

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Rebekah H. Nagler, Marco Yzer

July 2017

Acknowledgements

My name graces the title page of this dissertation, but this work is the product of a large group of advisers, faculty, colleagues, friends, and family. I want to take the opportunity to thank all of those who have helped make this accomplishment possible.

First, I want thank my advising team, Dr. Rebekah Nagler and Dr. Marco Yzer. Dr. Nagler, as a busy assistant professor you chose to invest part of your time advising me, and for this I am grateful. Not only did you provide me with thoughtful feedback, guidance, and support, but you also helped me to see the value in my work and build my confidence as a researcher. As my first adviser, Dr. Yzer you taught me early on that integrity in research is more important than statistically significant results—one of the most important lessons I've learned in the last five years. I also want to thank my dissertation committee chair, Dr. Sarah Gollust, for your invaluable insight during the development of my dissertation project and support throughout the process. Lastly, I want to thank my dissertation committee member, Dr. Darin Erickson, who helped me to conceptualize my analytical approach for Study 2.

Other faculty and staff of the SJMC deserve thanks for their help and support over the years. These individuals influenced my research and teaching, or simply helped me navigate the challenge that is graduate school: Mary Achartz, Dr. Colin Agur, Dr. Betsy Anderson, John Blair, Dr. Valerie Belair-Gagnon, Sue Couling, Dr. Ruth DeFoster, Julie Golias, Kathleen Hanson, Dr. Jisu Huh, Dr. Sherri Katz, Dr. Seth Lewis, Dr. Meghan Manning, Christine Mollen, Dr. Amy O'Connor, Nora Paul, Rebecca Rassier, Angie Rehn, Dr. Hyejoon Rim, Dr. Amy Sanders, Dr. Dan Sullivan, Dr. Chris Terry, Dr. Albert Tims, Dr. Dan Wackman, Dr. Brendan Watson, Jennifer Welsch, and Dr. Jennifer

Williams. I would also like to acknowledge the funding support of the Joel and Laurie Kramer School of Journalism Fund, Hubbard Graduate Fellowship, Kriss Research Support Grants, and the Ralph D. Casey Dissertation Research Award. In addition, I would like to thank my coders for Study 1: Dylan Beckwith; Haley Johnson; and Sarah Smokrovich.

I have been honored to be surrounded by an amazing group of peers over the last five years. Sarah Cavanah, Jennifer Lueck, Ben Miller, Casey Carmody, Taemin Kim, Kriste Patrow, and Chelsea Reynolds—you have been among my greatest friends, sources of support, and encouragement during my graduate studies. Without you, I'm not sure how I would have gotten through this process. In addition, I would like to acknowledge all the help and support from my fellow graduate students: Cassie Batchelder Merrick, Michelle Chen, Chuqing Dong, Meghan Erkkinen, Xiaofei He, Liz Housholder, Debra Kelly, Hyejin Kim, Eunah Kim, Jisu Kim, Soojung Kim, Soyoon Kim, Brett Johnson, Jiyoon Lee, Magdalene Lee, Xinyu Lu, Anna Popkova, Jared Martin-Rogers, Michaele Myers, Chris Nettleton, Alex Pfeuffer, Missi Rossi, Weijia Shi, Wooyeol Shin, Konstantin Toropin, Sarah Wiley, Rodrigo Zamith, and Xuan Zhu. Thanks also to my friend of nearly 30 years, Meghan Snegosky. Meghan, you have been a constant source of encouragement, support, and laughter throughout my reentry into higher education.

I must also thank the University of Wisconsin-Stout community who supported my graduate school aspirations, particularly the faculty and staff of the Food and Nutrition department. In addition, I offer my thanks to Dr. Amanda Brown for taking a risk and mentoring a complete research novice. Lastly, without the mentoring and support

I received from Dr. Richard Tafalla and colleagues in the Ronald E. McNair Scholar's Program, a Ph.D. would have never been a consideration.

As described above, faculty, staff, friends, and colleagues have played a critical role in my graduate studies and this dissertation, but my family's role cannot be overstated. My entire family has been immensely supportive as I've trudged through higher ed over the last eight years, but their greatest influence was during my formative years. To my mother, Mary LoRusso, thank you for all of the love and support you have given me over a lifetime. Thank you for putting up with my endless questions about everything, and allowing me to stay up way too late reading with a flashlight under the covers. To my brothers, Matthew LoRusso and Carl LoRusso, thank you for being my first students, even though you had little choice in the matter. To my aunts, Vicki Bowler and Loretta Sondrall, thank you for unconditional kindness and for enabling me to see the world beyond Luck, WI. To the loved ones I have lost: Grandma and Grandpa—Ardyce and Reuben "Swede" Haglund—thank you for inviting us into your home, playing along with my imaginary friends, driving me to the library every week, and your love; Uncle Garry, each book, chemistry set, and fossil, that you gave me, invited me to think critically and question assumptions; Uncle Bill, thank you for always encouraging me and accepting me for the just the way that I am.

My family has expanded since childhood. I must thank my sister-in-law, Jennifer LoRusso, for your constant support over the last eight years, and my sister-in-law, Stephanie LoRusso, for helping me navigate reentry into higher ed. Additionally, I would like to thank you both (and your husbands, Matthew and Carl, respectively) for creating the five cutest nieces and nephews a person could have. Your children have provided a

much needed distraction from the stresses of graduate school, contributing to the thread of sanity that I have left. To Ben Blanchette, thank you for helping me to get through the day-to-day, see hope in the future, believing in me, and for your love. I wouldn't have finished this dissertation without you.

Finally, I would like to acknowledge and thank the 95 women and men who served as celebrities-of-interest for this dissertation. This dissertation is not a criticism of your decision to come forward. At one time, there was stigma surrounding even the discussion of breast cancer. Notable persons willing to publicly discuss their own experiences with breast cancer have helped to push that stigma aside.

Dedication

This work is dedicated to my grandmother, Ardyce Haglund. Grandma, you are responsible for my lifelong obsession for the news, and your influence is seen in these pages. Your public display of pride for my Ph.D. pursuits was often embarrassing, but appreciated nonetheless. You wanted to be here to see me earn my Ph.D.—and I wish you were—but, regardless, you played a significant role in the conclusion of this dissertation. I've often thought of throwing in the towel on this Ph.D. thing. When those thoughts crept in, I could hear your voice—full of disappointment, shaming me into completing this dissertation. I would rather make you proud, so I kept going.

Abstract

A long research tradition exists investigating the content of news coverage of celebrity breast cancer disclosures and, to a greater extent, the impact these personal health narratives have on public cancer-related outcomes. However, the bulk of this research focuses on specific, large-scale media events, such as Angelina Jolie's 2013 BRCA disclosure. The attention to individual disclosures provide insight about the specific media event, but does not further knowledge about the larger phenomenon of celebrity cancer disclosures.

To go beyond the Angelina effect, this dissertation addresses three overarching research questions: 1) What breast cancer-related messages are present in media coverage of celebrity breast cancer disclosures; 2) do these messages impact public cancer-related behavioral outcomes (i.e., online breast cancer information seeking); and 3) are there attributes of the celebrity that predict media and public outcomes?

To address these questions, first, 110 individual celebrity breast cancer disclosures between 2005 and 2016 were systematically identified. Then, two longitudinal studies were conducted. To address the first question, Study 1 used computer assisted and hand coded procedures to assess the presence of episodic frames (defined as containing information specific to the celebrity and her experience with breast cancer) and thematic frames (defined as including population and subpopulation breast cancer information [e.g., prevalance, risk, survial rates]). In addition, the presence of seven content categories classified as misinformation (defined as information which is innaccurate, misleading, or oversimplified) in news coverage was assessed. Results demonstrated that 80% of the news articles were written with an episodic frame, and 20%

were written with a thematic frame, indicating very little information beyond the celebrity's own experience with breast cancer was conveyed to the public. However, misinformation was largely absent in the news coverage—only misinformation pertaining to early breast cancer detection and mastectomy decisions was present in 10% or more of the news coverage.

Study 2 attempts to determine if news content impacts information seeking by using the framing outcomes from Study 1 to predict Google Trends search query outcomes. Due to the disparate rates in the presence of episodic and thematic frames this dissertation is unable to provide support linking content and online breast cancer information seeking outcomes. However, time series models suggest that media coverage of celebrity breast cancer disclosures in the aggregate have a distal impact on the public's breast cancer information seeking outcomes. For example, some analyses suggested effects happened as late as 17 months after news coverage of the disclosure. Yet the nature of these trends may be a function of the data.

Establishing if celebrity attributes can predict media and public outcomes was done through a moderation analysis of the results of Study 1 and Study 2. Specifically, the extent to which the presence of episodic and thematic and misinformation were present and statistically significant information seeking models were examined as a function of the celebrity's age, career type, breast cancer-event type, and level of celebrity status (defined as the degree of fame the celebrity achieved at the time of disclosure). Eighty-seven percent of thematic frames present were in news coverage of celebrities at the highest levels of fame. Specific categories in the age, career-type, and level of celebrity status variables predicted the presence of misinformation. Some

preliminary evidence suggests level of celebrity status may predict online breast cancer information seeking outcomes. The implications of the dissertation's findings for health communication research, mass media effects research, and professional health communicators are discussed.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

In September of 1974, First Lady Betty Ford publicly announced she had undergone breast cancer surgery. Nearly forty years later, actress Angelina Jolie wrote an op-ed for the *New York Times* where she disclosed her decision to have a bilateral mastectomy after learning she is BRCA1 positive. Both of these events can be defined as a celebrity breast cancer disclosure—when a well-known public figure publicly announces a breast cancer-related event (e.g., diagnosis, breast cancer treatment, preventive action) or dies due to complications from breast cancer. These two celebrity's announcements bookend the decades of research dedicated to understanding what health-related impact such disclosures may have on the public (see Noar, Willoughby, Myrick, & Brown, 2014). Given the great volume of media coverage celebrity breast cancer disclosures receive—breast cancer disclosures make up nearly 25% of annual breast cancer news coverage—and evidence that the public's decision making can be influenced by celebrities, scholarly attention is warranted (Corbett & Mori, 1999; Greenberg, Freimuth, & Bratic, 1979; Jensen, Moriarty, Hurley, & Stryker, 2010; Ohanian, 1990).

Evidence suggests that Ford's disclosure led to an increase in population-level breast cancer screenings in the months following her announcement (Fink et al., 1978). Jolie's disclosure has been credited with increases in scheduled BRCA screening appointments and online breast cancer-related information seeking, among other outcomes (e.g., Kosenko, Binder, & Hurley, 2015; Noar, Althouse, Ayers, Francis, & Ribisl, 2015). In the decades between these disclosures, similar results have been found for other public figures such as Nancy Reagan and Kylie Minogue (see Noar et al., 2014).

These findings underscore the significant impact a celebrity breast cancer disclosure can have on public health-related outcomes. However, this work is largely unsystematic. That is, most of the research in the celebrity cancer disclosure domain is concerned with a specific disclosure, the ensuing media event, and public effects directly connected to the disclosure, making it difficult to generalize outcomes beyond the particular celebrity disclosure.

In their review and research agenda on public figure cancer announcements, Noar et al. (2014) conclude that based on 19 studies reviewed, there have been "meaningful effects on a whole range of outcomes, from news volume to information seeking to choice of surgery to cancer screening behaviors" (p. 12). But, they assert that the lack of research beyond the select few celebrities or contemporary announcements limits this line of research. Thus, they call for more research in this domain. Specific areas of opportunity include examining: 1) the role the news media plays (volume of news coverage, duration for which the disclosure is covered, and actual content of the reports); 2) how that role impacts public cancer-related outcomes; and 3) what types of celebrities have the greatest impact on media and cancer-related outcomes.

It should be noted that the review and research agenda (Noar et al., 2014) was published prior to the glut of studies focused on Jolie's disclosure. At this time nearly 40 empirical studies have been published which primarily investigate effects related to her May 2013 announcement. This intense scholarly attention to this contemporary disclosure gives insight to the impact of Jolie's announcement, yet it does not truly answer Noar and colleagues' (Noar et al., 2014) call for further research.

Noar et al.'s (2014) review of public figure cancer announcement research reveals these disclosures have "effects," but the impact is considered short-term. For example, increases in screening or information seeking are observed proximal to the disclosure, but return to pre-disclosure levels within weeks or months. Noar et al. (2014) compare this outcome to those of purposive health communication campaigns—"these kinds of announcements appear to act as interventions (albeit 'naturally occurring' interventions) that achieve increased interest and engagement in the topic of cancer, with corresponding increases in actions directly related to the cancer with which the public figure was stricken," (p. 12).

These naturally occurring interventions, however, are not necessarily unique. Breast cancer is of course a top public health concern. It is second in prevalence (next to male reproductive cancers) and second in fatal outcomes for women following lung cancer as a topic receives significant media attention (National Cancer Institute, 2017). Strategic health communication interventions aimed at certain breast cancer-related behaviors (e.g., self-breast exams, and routine screening) have been a public priority for decades (Osuch et al., 2012); there has also been an abundance of news coverage related to the topic of breast cancer—breast cancer is the most reported on cancer (Corbett & Mori, 1999; Jensen et al., 2010). Breast cancer's prominence in the news media tends to be largely event driven. New scientific findings related to breast cancer, breast cancer fundraisers and events (e.g., Susan G. Komen 3-Day Walk), and the changes in breast cancer screening guidelines all have driven cycles of news coverage on the topic of breast cancer (Clarke & Everest, 2006; Corbett & Mori, 1999; Henderson & Kitzinger, 1999; Jensen et al., 2010). Unfortunately, much of the research on media coverage of breast

cancer lacks evidence as to what specific information is being communicated to the public beyond topic and source (see Jensen et al., 2010). However, what we know about strategic breast cancer communication and news coverage on breast cancer indicates that these messages often compete against each other, can be contradictory, and ultimately have the potential to confuse the public rather than promote public health (Carpenter et al., 2015; Nagler, 2014; Nagler & LoRusso, 2017; Niederdeppe & Levy, 2007; Walls, Peeters, Proietto, & McNeil, 2011). What role might celebrity breast cancer disclosures play in this greater breast cancer information environment?

The Greater Breast Cancer Information Environment

As stated above, celebrity breast cancer disclosures are one piece of the breast cancer information environment. For a study to examine media coverage and public outcomes related to contemporary celebrity disclosures, it is important to consider the *contemporary* greater information environment and the secular trends in which they exist (and potentially contribute). For example, promoting breast cancer screening for average-risk women has been a public health priority for years. Professional clinical organizations developed evidence-based screening recommendations founded on age and risk level, which have been central in strategic communication promoting screenings (American Cancer Society, 2015; U.S. Preventive Services Task Force, 2009; Osuch et al., 2012). Screening guidelines on their own have also been the topic of news coverage and included in the coverage of human interest stories or in coverage of breast cancer fundraisers and related events (see Corbett & Mori, 1999; Jensen et al., 2010).

Screening guidelines earned headlines in November of 2009 because the U.S. Preventive Services Task Force (USPSTF) issued new breast cancer screening

recommendations. Previous guidelines from the USPSTF recommended women of average risk initiate annual screening at age 40. However, new evidence indicated that not only had routine annual screenings initiated at age 40 not reduced breast cancer-related deaths but had even contributed to over-diagnosis (Harris, 2014; Welch & Frankel, 2011). Taking these results into consideration, the updated guidelines from the USPSTF increased the recommended age of routine screening for a woman at average risk to age 50 and suggested screening take place every two years instead of the previous annual recommendation (Harris & Sheridan, 2013; Wilt & Partin, 2011; U.S. Preventive Services Task Force, 2009).

The dramatic changes in recommended frequency received significant media coverage. Weeks, Friedenberg, Southwell, and Slater (2012) found that in November 2008 (the control month/year) 174 news articles covered the topic of mammograms, but in November 2009 (the month the updated USPSTF guidelines were released) a total of 670 newspaper articles covered the topic of mammograms. Weeks et al. (2012) did not investigate the actual content of coverage, but subsequent content analyses revealed that not all of the coverage was devoted to conveying the specifics of the updated guidelines or why the changes were made. Instead, much of the news coverage focused on the conflicting recommendations between USPSTF and other agencies such as the American Cancer Society (ACS) and the American College of Radiology. These organizations still recommended routine screening initiated at age 40 for women of average risk for a breast cancer diagnosis (Fowler & Gollust, 2015; Nagler, Fowler, & Gollust, 2015). In the years since the USPSTF update, additional guideline changes have expanded the volume of conflicting recommendations. As of October 2015, the ACS recommends routine

screening to begin at age 45 (American Cancer Society, 2015; National Cancer Institute, 2017; Oeffinger et al., 2015) and other organizations, such as the American Congress of Obstetricians and Gynecologists, still recommend screening initiation at age 40.

There may not be a scientific consensus, but many medical and public health professionals and researchers endorse the USPSTF guidelines (Harris & Sheridan, 2013; Hersch et al., 2014; Waller, Douglas, Whitaker, & Wardle, 2013; Welch & Frankel, 2011; Wilt & Partin, 2011). Perhaps due to the still conflicting recommendations between organizations and the long-time message of routine screening at age 40 from physicians and the media (Corbett & Mori, 1999; Hersch et al., 2011; Jensen et al., 2010; Welch & Frankel, 2011), there is evidence that women are resistant to delay routine screening until age 50 or to reduce screening frequencies to two years (Hersch et al., 2013; Hersch et al., 2011, 2015; Nagler, Fowler, & Gollust, 2017; Waller, Osborne, & Wardle, 2015; Waller et al., 2013; Yu, Nagler, Fowler, Kerlikowske, & Gollust, 2017). Such beliefs have been found to be influenced by normative beliefs about early detection (Hersch et al., 2013; Waller et al., 2015).

This brief discussion of the existing conflicting breast cancer screening recommendations, media coverage of this conflict, and reluctance by women to follow the USPSTF guidelines demonstrates the confusion surrounding just this specific breast cancer-related topic. Placing celebrity breast cancer disclosures into a media environment that at one time consistently told women that routine screening should begin at age 40, which has for the last eight years been replaced with conflicting messages about screening recommendations, may further complicate (or perhaps may offer clarity) as to what age a woman should begin screening. There is evidence that celebrity breast cancer

disclosures, such as Austrailian Pop singer Kylie Minogue and First Lady Nancy Reagan, have impacted breast cancer screening rates (Kelaher et al., 2008; Lane, Polednak, & Burg, 1989; Noar et al., 2014). However, in the case of Minogue—who was 36 at the time of diagnosis—dramatic increases in screening initiation were seen in women well under the recommended age of screening for women of average risk. Simple increases in screening rates are not the goal. If these disclosures, as Noar and colleagues (Noar et al., 2014) state, work as naturally occurring interventions, then it is important that the target population receives accurate information.

Breast cancer screenings are likely the dominant breast cancer-related topic in the news media over the last decade. But celebrities experience breast cancer-related events beyond screening and subsequent diagnoses. Personal, medical decision-making stories, from an observational standpoint, are often disclosed and discussed with the news media. Jolie's disclosure is perhaps the most well-known example, but other celebrities such as actress Christina Applegate or Florida Congresswoman Debbie Wasserman Schultz have shared similar experiences. Celebrity breast cancer disclosures exist in a cluttered media environment. Changing breast cancer screening guidelines, breast cancer fundraising events, or novel findings in medical research are often given media attention, but a celebrity breast cancer disclosure can easily outshine the other stories due to the nature of celebrity; a celebrity's testimonial or endorsement can be a powerful tool (Beck, Aubuchon, McKenna, Ruhl, & Simmons, 2014; Kelman, 1961; Ohanian, 1990; Stout & Moon, 1990; Thrall et al., 2008). A celebrity's personal experience with breast cancer could further the public's understanding of breast cancer screening guidelines, for

example, but only if media coverage makes salient how the given celebrity's experience relates to other women (those at similar risk and those of average risk).

Dissertation Objectives

To understand what role celebrity breast cancer disclosures play in the greater breast cancer media environment—and to answer Noar and colleagues' (2014) call for further research—it is necessary to go beyond studying specific disclosures and instead offer generalizable evidence. The goal of this dissertation is to provide a comprehensive analysis of media and public health-related outcomes across celebrity disclosures and to provide evidence for under what conditions specific outcomes should be expected. To do this, this dissertation answers the following questions: 1) What breast cancer-related messages are conveyed to the public in media reports of celebrity breast cancer disclosures?; 2) Do these messages impact public cancer-related behavioral outcomes?; and 3) Are there attributes of the celebrity and/or breast cancer-related event that may predict or explain media and public outcomes?

To address these questions, I conducted two longitudinal studies. Study 1 was a content analysis of the presence of breast cancer-related information in news coverage of celebrity breast cancer disclosures. Specific types of information that have the potential to reinforce misperceptions about breast cancer held by the public were identified. In addition, media frames which solely focused on the celebrity and her or his breast cancer-related event (episodic frame) or frames that included a greater public health message (thematic frame) (e.g., population and subpopulation risk information, survival rates) were also quantified. By exploring both the content and prevalence of these messages in top circulating national newspapers from 2005 through 2016, Study 1 offers evidence of

limited breast cancer-related information present in news coverage of celebrity breast cancer disclosures, as well as specific media events which propagated misleading information.

Study 1 focuses on the first overarching research question, and Study 2 was concerned with the second research question. As mentioned previously, numerous public breast cancer-related outcomes are possible in response to a celebrity breast cancer disclosure. Because this dissertation is interested in types of information present in media coverage of the disclosures, a relevant information-based outcome is breast cancer-related information seeking. A few studies have found that celebrity cancer disclosures positively correlate with cancer information seeking behaviors, but these studies focus on volume of coverage as a mechanism for search and do not consider the role content might play in information seeking behaviors (Ayers, Althouse, Noar, & Cohen, 2014; Noar et al., 2015b; Noar, Ribisl, Althouse, Willoughby, & Ayers, 2013b). Study 2 directly tests the impact news coverage plays on online information seeking results (i.e., Google Trends) and also attempts to test the impact the volume of dichotomous frames (i.e., episodic/thematic) have on specific breast cancer-related information search query domains.

Finally, a subanalysis on the results of Study 1 and Study 2 was performed to address the third overarching research question. Attribute variables which might moderate media and information seeking outcomes were constructed that categorize celebrities in groups based on age, career type, breast cancer-related event type, and level of celebrity status. These analyses provide further empirical evidence as to *who* is most

likely to prompt certain media and public breast cancer-related information seeking outcomes.

This dissertation proceeds with the following sections: First, a review of the theoretical literature and empirical findings in the natural coverage effects research tradition, sociological perspectives of fame, journalistic norms, celebrity cancer disclosure research, media effects, and concepts from the information sciences (i.e., ambiguity, misinformation, and health information seeking). Next, hypotheses and research questions, as well as the conceptual model of effects are offered. Methods and measures are a particular focus of this dissertation. A full chapter outlining the systematic approach used to determine celebrities-of-interest and the construction and measurement of celebrity attribute groups is put forward. The following chapter offers methods used to retrieve media reports (e.g., recall and precision search string testing) for Study 1, the process of codebook development, and coding procedures, and concludes with the data analysis and results. Next, the methods and results chapter on Study 2 presents steps taken to ensure content validity of search query data, the construction of breast cancer-related information seeking domains, the analytic approach for time series data analysis, and the subsequent results. The final chapter discusses the dissertation's findings, and the theoretical, methodological, and practical implications of this research.

Chapter 2

Literature Review

Beyond a Single Celebrity Disclosure Event

A celebrity cancer disclosure media event is often treated as an isolated event by researchers (e.g., public breast cancer information seeking in response to Angelina Jolie's BRCA1 disclosure). For example, Stryker's (2003) discussion on the evidence of news effects on health behaviors separates the evidence into two conceptualizations—"short-term effects of large media events" and "more gradual and cumulative effects of news coverage on long-term secular trends in health behavior," (p. 307). Stryker's "large media events" examples that fall into the "short-term" effects conceptualization include Betty Ford's breast cancer surgery disclosure and Ronald Reagan's removal of an intestinal tumor (i.e., celebrity cancer disclosures). Fishbein and Hornik (2008) also mention a celebrity health disclosure ("the announcement of HIV status by a celebrity," [p. 3]) when discussing short-term effects.

Stryker (2003) and Fishbein and Hornik (2008) were not incorrect in categorizing a celebrity health disclosure in the "short-term effects" conceptualization. Much of the research in the celebrity cancer/health domain focus on a handful of select individuals—Magic Johnson, Steve Jobs, Patrick Swayze, and Angelina Jolie (see Beck et al., 2014; Noar et al., 2014). Only in the last few years have researchers begun to study content and effects connected to celebrity cancer/health disclosures beyond the single isolated event. Of the few studies which investigate outcomes related to multiple celebrity cancer disclosures, two of these studies test the relationship between the disclosure and cancer information seeking (Niederdeppe, 2008; Noar, Althouse, Willoughby, & Ayers, 2013).

Both found proximal increases in cancer information seeking either in the aggregate (Noar et al., 2013) or through self-report (Niederdeppe, 2008) on average across the group of celebrities under study. However, increases in seeking behaviors were not necessarily found for each individual celebrity, and there were differential levels of seeking by celebrity, perhaps leaving more questions than answers.

Another study focusing on multiple celebrity cancer disclosures investigated the news media content of 17 celebrity breast cancer disclosures from 1992 to 2014 (Sabel & Dal Cin, 2016). Of the 17, Sabel and Dal Cin (2016) found that Christina Applegate's (a high profile actress) disclosure of a breast cancer diagnosis and a bilateral mastectomy received the highest rates of media coverage. This media event coincided with a significant increase in bilateral mastectomies at the University of Michigan in the same year. The spike in bilateral mastectomies does decay in the years that follow, but even through 2015 bilateral mastectomy rates remain elevated over rates prior to Applegate's disclosure. The author's coin this the "Applegate Effect." While the authors investigate press coverage of multiple celebrity breast disclosures, they took a short-term effects approach in discussing their findings. Applegate may have triggered the rise in bilateral mastectomies, but several other celebrities in the following years also disclosed bilateral mastectomy decisions (namely Guiliana Rancic, Wanda Sykes, and Kathy Bates). These additional announcements could have perpetuated sustained media coverage of bilateral mastectomies, potentially strengthening Applegate's initial effect of surgery decisions.

It should be noted that the authors do not disclose how the celebrities under study were selected, but the list reflects an unsystematic approach—as results in Chapter 4 of

this dissertation reveal, several "celebrities" who experienced a breast cancer-related event in the years under study were not included in the analysis.

The three studies (Niederdeppe, 2008; Noar et al., 2013; Sabel & Dal Cin, 2016) discussed here represent the totality of published studies investigating either content or effects of multiple celebrity cancer disclosures over time. Because numerous celebrity disclosures beyond the usual suspects are investigated, this approach takes a step forward in the celebrity cancer disclosure research. However, the unsystematic selection of celebrities, the lack of data analysis from an overall trends perspective, or only questioning *post hoc* the influence of different types of celebrities does little to provide further knowledge about the overall celebrity cancer disclosure phenomenon.

To do this, Stryker's (2003) second perspective on media effects—"more gradual and cumulative effects of news coverage on long-term secular trends in health behavior" (p. 307)—is applicable. Stryker's (2003) study analyzes long-term news coverage of marijuana effects on beliefs towards marijuana. While the topic of celebrity cancer disclosures and marijuana may seem distinct, it is likely that news coverage of marijuana is primarily event-driven as well (e.g., new laws/regulations passed; new study released on the effects of). Just as Stryker (2003) and others (e.g., Romantan, 2005; Stevens & Hornik, 2014) have approached studying event-driven news coverage with health implications (e.g., marijuana; HIV) as "longitudinal studies of the impact of news messages on health behavior" (p. 307), this dissertation treats celebrity breast cancer disclosures as a general topic—like marijuana—and tests assumptions of public behavioral outcomes (i.e., breast cancer information seeking) over time. This approach

does not have a distinct classification, but from this point forward it will be referred to as a "natural coverage effects" approach.

Natural Coverage Effects Research Tradition

Any discussion thus far on the potential effects of a celebrity cancer disclosure or cancer disclosures in the aggregate hinges on exposure to news content of celebrity cancer disclosures. Exposure could be direct—reading an article about a celebrity cancer disclosure—or indirect—talking with a friend about the disclosure. Regardless of the means by which an individual is exposed to the information, exposure must occur for any expectation of effect. Traditionally, to assess exposure, media effects scholars have directly asked people about their exposure. This is typically done through self-report surveys, where individuals may be asked to recall time they spent with a particular medium (i.e., global self-report measures) or exposure to specific content (Nagler, in press). However, asking individuals to accurately recall either type of exposure has its limitations. Nagler (in press) points out that such lines of inquiry bring with them questions of accuracy, particularly cognitive and motivational errors. Niederdeppe (2016) argues further that with greater connectivity, diffusion, and narrowcasting of content, it simply places too much of a cognitive burden on individuals to ask them to remember where they were exposed to content or how much time they spent with specific content. Nagler's (in press) comment on motivational inaccuracy can be due to particular interest in a topic, which might result in overreporting exposure, or could be due to social desirability bias and lead to underreporting (e.g., exposure to pornography).

Another issue or limitation of self-report data is the availability of data. For a study about a health topic covered in the news media for a significant period of time, it

may be impossible to have data applicable to the topic of interest. For example, there are no existing data that measures individual's exposure to celebrity breast cancer disclosures for the last decade. To circumvent the accuracy and availability issues that come with self-report data, ecological measures of exposure are used. Niederdeppe (2014) describes ecological measures as "those that characterize exposure in terms of its potential reach based on geographic or temporal variability in message availability" (p. 171). Nagler (in press) clarifies the intention of ecological measures: "[E]cological measures do not ask people questions; rather, they assess the potential or opportunity for people to be exposed" (p. 9). In campaign effects research, ecological exposure is often manipulated (messages are manipulated across time periods or regions), but in the case of routine media exposure, natural coverage is measured (Nagler, in press). Measuring natural coverage requires analyzing specific content. It is virtually impossible to content analyze all media content related to a topic, so media effects researchers rely on content which represents the media environment in the aggregate (Fishbein & Hornik, 2008; Nagler, in press; Niederdeppe, 2016). Fishbein and Hornik (2008) comment on the underlying suppositions of such an approach: "This approach assumes that media content makes its way into individual consciousness and influences behavior, either because people are directly exposed to media messages or because they are indirectly exposed when others share content that they have seen or diffuse new behaviors they have learned" (p. 2).

In the case of a longitudinal analysis, such as Styker's (2003), once the "natural coverage" is assessed—both volume and content—the coverage is typically compared with aggregate-level outcomes over the same time period. Seminal work in this line of research often used trend analyses to visually detect if there were distinct links between

media coverage and behavioral outcomes. For example, Soumerai and colleagues (Soumerai, Ross-Degnan, & Spira Kahn, 1992) investigated the potential connection between aspirin use in children and Reye's syndrome and if the lay and medical press' coverage over several years influenced professional and consumer behaviors. One of the teams' key findings was that aspirin sales declined sharply during the height of the media's reporting on the potential link with Reye's. This result was attributed to direct consumer behavior (i.e., not purchasing and administering aspirin to young children) and the actions of retailers (i.e., limiting the supply of aspirin). Other studies which explored—using visual trends analyses—behavioral outcomes correlated with extensive media coverage of consumer products include the discontinued use in the aggregate of the intra-uterine device and the contraceptive pill after adverse health effects were covered by the press (Cates, Grimes, Ory, Tyler, & Cates, 1977), and smoking cessation and initiation after intense media coverage of the health effects from smoking (Pierce & Gilpin, 2001).

In more recent research, time series analysis is used to test causal relationships (see Nagler, in press; Niederdeppe, 2016; Stevens & Hornik, 2014; Stryker, 2003). Romantan (2005) hypothesized that high volumes of news media coverage of plane crashes, particularly coverage with a "conspiracy" frame (e.g., possible terrorism, sabotage, insurance scheme) would be negatively associated with airline travel behavior. To do this, passenger boarding numbers from 1978-2001 were obtained from the Department of Transportation and media volume was estimated from *Associated Press* coverage. Through times series regression analysis, the hypothesized relationship was indeed found. In months with the greatest rates of coverage given to plane crashes,

passenger boarding numbers significantly decreased. Media reporting with a "conspiracy" frame only strengthened this relationship.

Using similar methods, Stevens and Hornik (2014) found through time series regression analysis that as newspaper coverage of HIV/AIDS increased, HIV testing behaviors for both Whites and Blacks decreased by 1.7% in the following month. This affect was even more pronounced for Blacks. In the aforementioned Stryker (2003) study, the impact of *Associated Press* coverage (i.e., volume and framing of marijuana use [PRO/CON]) was tested on adolescents' beliefs towards marijuana (i.e., personal disapproval; perceived harmfulness) and marijuana behaviors from 1977-1999. Stryker (2003) found that media coverage of marijuana use did predict abstinence behaviors. In addition, PRO coverage (stories emphasizing the negative aspects of marijuana use) were positively associated with perceived harmfulness, although CON coverage (stories emphasizing the positives of marijuana use) did not discourage adolescents' beliefs about the potential harms of marijuana.

Studies such as Romantan's (2005), Stevens and Hornik's (2014), and Stryker's (2003) use data which precludes exploration of the underlying mechanisms that produce effects, although a limitation, this approach is useful as a first step to determine if effects are detectable in the aggregate. Demonstrating real-world effects can then provide opportunities for future research to consider mechanisms with ecological validity. Furthermore, these studies which explore effects of a topic covered by the media over a prolonged period of time offer a realistic perspective to media effects research. That is, researching effects from a specific event (an isolated celebrity cancer disclosure) implicitly, as Stryker (2003) states, promotes a hypodermic needle view of effects. These

reviewed studies did not consider airline boarding numbers after one plane crash or marijuana use in adolescents after one marijuana-related news event. Examining longitudinal effects of news coverage enables identification of secular trends in health outcomes instead of immediate effects from a specific event (Hornik, 2002; Stryker, 2003; Viswanath & Finnegan, 2002). Such is the approach this dissertation takes in studying celebrity breast cancer disclosures.

Conceptualizing Celebrity

In the extant literature devoted to the content and effects of celebrity cancer disclosures, *celebrity* is often treated a theoretical primitive—little to no conceptual definition is stated by the authors, and often little to no operational criteria are offered. Even in a review and research agenda on 19 studies examining celebrity cancer announcements, Noar et al. (2014) never explicate celebrity or public figure, as they use these interchangeably. As previously discussed, the studies reviewed by the authors and other studies researching celebrity cancer/health disclosures focus on a specific celebrity (e.g., Angelina Jolie; Magic Johnson) (e.g., Basil, 1996; Borzekowski, Guan, Smith, Erby, & Roter, 2014; Kalichman & Hunter, 1992; Kosenko et al., 2015). Such exemplars likely have high name recognition, giving face validity to *celebrity*. Thus, focusing on one individual may negate the necessity to conceptualize celebrity. However, with the study of multiple celebrities, it is important to fully conceptualize *celebrity*. That is, who is considered a celebrity and why.

A conceptual definition is rarely offered, but operationalizations of celebrity can provide some insight as to how researchers regard the construct. For example, in a content analysis of newspaper coverage of cancer, Jensen et al. (2010) coded persons

with cancer as either not famous or famous. Coding for famous was operationalized as: "anyone known to a large number of people on a national or local level (e.g., athlete, politician, entertainer, etc.)" (p. 142). Niederdeppe's (2008) study investigated cancer information seeking proximate to celebrity cancer disclosures or deaths in 2005. Niederdeppe described a celebrity as someone who "was deemed sufficiently famous to need no descriptor beyond the name in the headline" (p. 428). Finally, Noar et al.'s (2013) longitudinal analysis of media and public information seeking outcomes in response to 25 public figure's pancreatic cancer announcements or deaths never conceptualize or operationalize celebrity or public figure. Instead the authors describe who the exemplars are: "Our search uncovered 25 public figures, some well-known and others lesser known, that had been diagnosed with or had died from pancreatic cancer. The list included prominent producers, singers, college presidents, chief executive officers, attorneys, singers, artists, authors, actors, Olympic medalists, and others" (p. 189). These descriptions differ to some degree. Niederdeppe implies that a celebrity is a name that is recognizable to all, while Jensen et al. (2010) and Noar et al.'s (2013) definitions hint at the possibility that celebrities may have different levels of visibility or recognition.

Research in other fields, such as advertising and public relations also fail to conceptualize celebrity (e.g., Kelman, 1961; Ohanian, 1990; Stout & Moon, 1990; Thrall et al., 2008). Sociologists, perhaps more than any other field, have spent time getting at the epistemological roots of *celebrity*. Although, as the health communication literature uses celebrity and public figure somewhat interchangeably, sociological musings on celebrity often intersect with *fame*. For example, Milner (2010) raises the question "Is

there a difference between being famous and being a celebrity?" (p. 380). He attempts to parse this out by offering the Merrian-Webster dictionary definitions of the two: "'fame' is defined as 'public estimation, reputation, popular acclaim' and 'celebrity' is defined as the 'state of being celebrated, fame, a famous or celebrated person'" (as cited in Milner, 2010). With similar inquiry, Ferri (2010) initially offers the definition of the Latin *celebritas,* which *celebrity* is rooted in: "fame, renown, or celebration" (p. 403). *Fame* is inherent in these formal definitions of celebrity, but scholars admit that the two are not one in the same. Breese (2010) delineates the two through accomplishment. That is, *fame* is earned by a particular achievement or talent and *celebrity* is cultivated through a series of newsworthy events. However we may situate the two, Milner (2010) asserts the two are intrinsically linked: "In popular contemporary usage 'celebrity' is a subcategory of famous people, referring mainly to entertainers and sports stars—but not a separate phenomenon" (pps. 380-381).

This discussion and attempt to distinguish between celebrity and fame, and the more neutral label of public figure from the health communication literature, has its merits particularly for the purposes of construct validity. While the three words or phrases might instinctually bring to mind different persons, a case can be made that any person who is considered to be on one list (e.g., famous) could just as easily be included on another list (e.g., celebrity). While these words and phrases have imperfect culturally shared meanings, there is a prerequisite of "knownness" to be thought of as a *celebrity, public figure*, or to be *famous*. That is, one must be known outside their own social circle to be labeled any of the three concepts discussed here (Milner, 2010). Several scholars include a dimension of knownness in their definitions of celebrity and fame. Ferri (2010)

states that a celebrity is "recognized by far more people than one can recognize back" (p. 363). Boorstin (1961) defines the celebrity as "someone who is well known for being well-known" (p. 28). Finally, Braudy (1986) offers this account of fame: "In its root sense, fame means to be talked about."

These numerous conceptualizations of fame and celebrity all indicate that in order to be famous or a celebrity or a public figure, others must bestow this recognition upon you. Perhaps what these different labels indicate is that both from scholars and the public, it is inherently known that fame, for example, is not a one-size-fits-all label. Therefore, from this point forward celebrity will be the construct used, as an overarching umbrella term which includes famous and public figure, but ultimately, simply means a knownness by others outside of one's social circle.

To have knownness requires others knowing who you are, which will likely occur through media, but why does media give attention to some people and not others? Most scholars agree that one's occupation plays a primary role in knownness, but other contributing factors might include relationships. Hollander (2010) puts forth this list of occupations and social positions, all fitting under the umbrella of celebrity, which tend to garner attention from the media and subsequently the public: "models and super models, fashion and interior designers, TV anchors, talk show hosts, 'TV personalities', athletes, beauty queens, famous hostesses and society ladies ('socialites'), members of the rich upper classes, a politician (or his wife), even some criminals may join their ranks" (p. 389).

Hollander's list almost mirrors Noar et al.'s (2013) list of 25 public figures who announced or died from pancreatic cancer, but Hollander's (2010) list also gives attention

to spouses, implying someone can be well-known because of their own particular social network which automatically elevates their own knownness beyond that of their social circle. Furthermore, the "TV personality" of today does not just exist on network and cable television, but instead we have to consider the rise of the internet celebrity—the "YouTuber." These are individuals who create and distribute content, via YouTube channels. As of 2016, Felix Arvid Ulf Kjelberg, "a foul-mouthed Swedish video-game commenter," (McAlone, 2016) had attracted 39.3 million subscribers. To put those numbers in perspective, the multi-platinum recording artist Adele has just under 11 million subscribers. It is likely that Adele's name is recognizable to more people than Felix the YouTuber, but we have quantitative evidence that Felix is known to at least 39.3 million people. For the purposes of conceptualizing *celebrity* Felix's fame on YouTube underscores the importance of considering the knowness of someone to niche audiences as well as widespread name recognition (e.g., Adele).

Noar et al.'s (2014) and Hollander's (2010) lists of celebrity occupations and social positions both include politicians. Perhaps decades ago a politician would not have been considered a celebrity, but as news about celebrities has moved away from the margins of journalism towards mainstream platforms and audiences, politicians are often under the glare of the entertainment media spotlight (Breese, 2010; Cohen, 2004). Cohen (2004) argues that a politician in the eyes of the public is really no different than those in the entertainment industry: "While our political leaders are very real, as are the consequences of their actions, we know them only through the media. In terms of our feelings toward them, the quality of the interaction we have with them, or their reality in our lives, is probably more similar to a television or movie star than to a family member

or co-worker" (p. 189). Of course, the recent election of real estate mogul turned reality television star Donald Trump to the US Presidency reaffirms Cohen (2004) and others' assertions about the intersection of fame and political service.

Keeping this discussion in mind, and the operational criteria used in health communication literature (Jensen et al., 2010; Niederdeppe, 2008; Noar et al., 2013), the following is used to define *celebrity* for the purposes of this dissertation:

> A *celebrity* refers to a person who is known outside of his or her own social circle to differing degrees (i.e., household name vs. known to niche audiences). The individual, likely through his or her career or a personal relationship, has gained attention from the public and the media; both play a role in creating and maintaining celebrity. The construction of celebrity is a reciprocal process between the famous individual, the media, and the public. This dynamic process includes media coverage of career and life events which garner attention from the public.

Celebrity attributes as moderators. The potential for celebrity cancer disclosures to garner attention from the media and public is well documented (see Noar et al., 2014), but no research has systematically investigated what type of celebrity or cancer disclosure is most likely to generate certain media-oriented and health-related behavioral outcomes. A handful of studies have moved this line of research forward simply by investigating outcomes of more than one disclosure (Niederdeppe, 2008; Noar et al., 2013; Sabel & Dal Cin, 2016). However, because no criteria were established *a priori* to stratify celebrities by potential commonalities, the authors were only able to theorize *post hoc* about differential outcomes. Noar et al. (2014), in their review and research agenda

on public figure cancer announcements, call for further investigation into the potential effects of celebrities on media and health-oriented outcomes. The authors state: “When looking at large-scale public health effects, it is likely that the kinds of announcements by the occasional well-known figures—perhaps only those with some level of ‘celebrity’ status—garner significant effects,” and continue with, “This remains largely an empirical question, however, and the more we know about the attributes of public figures whose announcements will result in significant effects, the better able health communicators will be to capitalize on such events for effective cancer communication and prevention” (Noar et al., 2014, p. 457).

Investigating outcomes related to a large set of celebrity cancer disclosures provides the opportunity to begin to answer Noar and colleagues’ (2014) posited question. Due to certain journalistic norms around newsworthiness, particular personal attributes, such as age and career type may garner greater attention from the media (de Leon, 2002). Pertaining to audience effects, identification—“an imaginative process through which an audience member assumes the identity, goals and perspective of a character” (Cohen, 2004, p. 261)—has been found to mediate health-related intentions and behaviors in response to celebrity health disclosures (e.g., Basil, 1996; Myrick, Noar, Willoughby, & Brown, 2014). Attributes which have been found to correlate with identification include age, ethnicity, gender, cancer history, and career type (Basil, 1996; Kosenko et al., 2015; Myrick et al., 2014; Myrick, Willoughby, Noar, & Brown, 2013). Journalistic norms or a process such as identification may help to explain and predict media and public health-related outcomes, but at the aggregate level it has been argued that effects from celebrity disclosures occur when there is “a high degree of knowledge of

the celebrity" (Brown & Basil, 1995, p. 351), or how famous or well-known the celebrity is—her or his level of celebrity status—as Noar et al. (2014) imply.

Level of celebrity status. Previously, a thorough discussion on who a celebrity might be, and how *celebrity* will be conceptualized moving forward, was offered. The final definition is rather broad and includes many individuals, but this is strategic. The previous discussion revealed that *celebrity* and its synonyms (e.g., *famous*, *public figure*) are all ambiguous terms, but instead of attempting to parse out who is a celebrity and who is just famous, for example, it is more appropriate to group many under one construct and then further refine. This refinement comes as a stratification of *celebrity*, or as I will call it moving forward, *level of celebrity status*.

On its face, stating that there are different levels of celebrity status is valid. The idea of the "A list, B list, and C list" celebrity has been pervasive in American culture for years. But, again, most of the empirical research testing effects of celebrity cancer disclosure has not spent much time conceptualizing levels of *celebrity status* nor have they considered potential differential effects from celebrities of lesser or greater status. As with the construct of *celebrity*, I believe that such lack of attention to these levels of fame is due to the strict focus, specifically in the health domain, on very specific disclosures from very famous exemplars (e.g., Angelina Jolie, Magic Johnson). However, we need to consider that lesser-known celebrities, but those who fit the conceptual definition of celebrity, generally still receive some media coverage of their disclosures (Noar et al., 2013), which has the potential to impact health outcomes for some.

Two studies lend support to this line of inquiry: First, UK reality television star Jade Goody's (starred in the UK's Big Brother Season 5) celebrity status could be

considered of equal status to many women in the United States who have also spent a season on a reality show. Goody's level of celebrity status in the UK certainly was not at the level of fame of, for example, a Kate Moss (well-known model), Victoria Beckham (former Spice Girl), or Keira Knightly (well-known actress). However, once Goody's cervical cancer diagnosis went public, the disclosure received significant attention from the British press. Metcalfe, Price, and Powell (2011) found strong positive correlations between Goody's public announcement of cervical cancer (and the announcement of her subsequent death) and aggregate level increases for cervical cancer screening for women ages 25-64 in the year following Goody's death. The study demonstrates that the British press made salient Goody's diagnosis and death, and health-related effects (i.e., screenings) were found for this mid- to lower-tiered celebrity.

Metcalfe et al.'s (2011) study does provide evidence that a lower-tiered celebrity can receive attention from the media and the public, but given the singular focus on Goody this could be an anomaly. However, Noar et al.'s (2013) study of 25 celebrity pancreatic cancer announcements and the effects on online cancer-related information seeking provides further support. The majority of the 25 celebrities did appear to stimulate pancreatic cancer and general cancer online information search queries, averaging an increase of 28% and 11%, respectively, but significant increases in search queries (more than an 100% increase) were only seen for Steve Jobs, Patrick Swayze, and Griffin Bell (former attorney general). Noar et al. (2013) acknowledge that Jobs' and Swayze's high level of fame likely influenced the rates of online information seeking. The authors expressed surprise about Bell's effect on media and information seeking outcomes, but they do not attempt to parse this out any further.

In their review and research agenda, Noar et al. (2014) revisit this study and in this context provide a bit more insight on their thoughts of *celebrity status*: "[V]irtually all of the lesser known public figures, including figures with less 'celebrity' status, such as Supreme Court Justice Ruth Bader Ginsberg, did not garner the news or search effects that were as evident with well-known figures such as Steve Jobs and Patrick Swayze" (p. 457). This excerpt highlights an implicit meaning of celebrity status—there are varying degrees or levels of fame, and those levels likely contribute to differential effects of media and health related outcomes—and underscore the practical importance of this line of inquiry.

To answer the empirical question posed by Noar et al. (2013) about which celebrities are most likely to garner effects, it is imperative to move beyond *post hoc* theorizing and instead conceptualize celebrity status *a priori*. However, celebrity status may be even more ambiguous a construct than celebrity. Celebrity is a wide umbrella term, but celebrity status gets at hierarchical levels of fame and *when* this fame occurs. van de Rijt and colleagues (van de Rijt, Shor, Ward, & Skiena, 2013) offer a succinct statement on "the most famous" phenomenon: "Fame exhibits both an extreme concentration of attention around a tiny selection of individuals and a high rate of turnover in this select group" (p. 267). Milner (2010) lends support to this statement: "Celebrity status is not stable; when a multi-million dollar athlete's performance declines significantly he is likely to lose his fans and his contract" (p. 383). Milner talks further about this decaying status of celebrity: "The erosion of status is even more likely if the celebrity's performance declines or fails to reach new heights. Decline is also made likely because many other talented or beautiful people are eager to replace established

celebrities" (p. 383). But, is there actually turnover? Does the athlete's fans truly go away or is there a shift in his level of celebrity status? I assert that *celebrity* does not simply disappear. This is evident when a celebrity—often an aging celebrity—who has not had a career or life event covered in years dies but there is an outpouring of public grief (e.g., Leonard Nimoy's, of *Star Trek* fame, death in early 2015). The celebrity had not lost their status as a celebrity, but this trigger event (e.g., death) made the celebrity salient with the media and the public once again.

If celebrity status does not completely erode how might we think of this stratification of celebrity status? Social status systems of course predate celebrity status systems (Milner, 2010). Before the silver screen starlet there were kings and queens, political and military elites, and social caste systems. Like these historical social status systems, modern day celebrities are often granted social privileges that their unknown counterparts are not subject to. However, as Milner (2010) points out, celebrity status systems may be less stable than traditional status systems: "There are many ways of expanding economic power and wealth just as there are many ways to expand political power. In contrast, celebrity status cannot be expanded in similar fashion, mainly because it is primarily a relative ranking. If one ascends in rank, then others must eventually descend" (p. 381). I believe that Ferri's (2010) use of *hierarchy* perhaps best describes the phenomenon I am interested in conceptualizing and subsequently empirically measuring:

> "There is a hierarchy of celebrity that starts with the most powerful and well-known politicians and media moguls—like Oprah Winfrey, known simply as 'Oprah'—and moves on down. Moving down the ladder are current sports and

> media stars, followed by, for example, old rock bands that road-trip across the country… The lowest rung of celebrity, especially the focus and content of media like YouTube, was first articulated by artist Andy Warhol's assertion that 'in the future everyone will be world-famous for 15 minutes'" (p. 405).

van de Rijt and collegues (2013) investigated the mobility of fame over a 20-year period by counting name mentions in 14 newspapers. They found that media coverage is allocated to a very specific group of individuals, with 60% of all newspaper coverage going to 1% of the names covered by the media. Also, significant coverage over one year was predictive of significant coverage the following year. That is, they found that once a celebrity receives high levels of media coverage he or she tends to maintain those high levels. Mobility, however, is greater at the entry levels of fame, where the first year someone enters into celebrity status they are more likely to not be covered the next year. Their data also demonstrated that "When a previously unknown individual is involved in an event that triggers a large and long enough public conversation, or reserves a place in a series of follow-up events, the name *locks in*" (p. 282). They also found that career may be predictive in duration and level of fame. Ephemeral fame (fleeting fame/high mobility) was largely present for those in the entertainment industry (e.g., authors, actors, artists) but was not present, or there was low mobility, for name in business, politics and sports, with most of these names being in the upper strata of fame. The authors explain this finding by reflecting upon the type of media coverage these particular careers seem to garner: "Authors, actors, and artists must regain popularity after each book, movie, or CD, whereas employment in professional sports or public office guarantees consistent coverage throughout a season" (p. 278). However, their data does bore out a pattern of

deterioration; lower strata names disappear quickly after the passing of the event(s) in which they emerged, and higher strata names "follow career-type patterns of growth, sustenance, and gradual decay over the course of decades" (p. 284).

Both scholars' theorizing on levels of celebrity status and empirical work substantiate the claim that there are different levels of fame. There is the *super celebrity* (e.g., Oprah), who could be a household name and their fame has persisted over a long period of time, to the lesser-known celebrity, whose time in the spotlight has been limited and fleeted quickly, and everything in between. Further, the level of one's celebrity status is not stable through the trajectory of one's status as a celebrity and can and will be mobile. Career type influences the degree of mobility (high or low). Media coverage helps to raise one to celebrity status, and in somewhat circular fashion, also indicates how important the media believes a celebrity to be. Therefore, when measuring effects from a celebrity cancer disclosure it is important to consider level of celebrity status at the particular point in time in which the announcement occurred. However, I believe that both van de Rijt and colleagues' (2013) study and some of the conceptualizations offered here place too much emphasis on the media's role in creating and maintain celebrity status for individuals. No doubt, the media plays a critical role here, but the audience's perception of how famous someone is also important. As several of the conceptual definitions offered in the section on *celebrity* point out, audience reception, knowing who someone is and arguably caring who someone is (e.g., purchasing memorabilia, telling stories) is a key determinant when conceptualizing, and operationalizing, the level of *celebrity status*. With this in mind, the following serves as the conceptual definition for level of celebrity status:

The *level of celebrity status* is conceptualized as the degree of fame a celebrity has achieved at a particular point in time. Someone with high celebrity status has been known by the media and the public for a substantial period of time and is considered to be a household name by the public. A celebrity with low celebrity status has likely been salient with the media and/or the public for a short period of time or is only known to particular niche media and public audiences. This is not a dichotomous concept— multiple levels of celebrity status are to be expected.

News Coverage of Celebrity Cancer Disclosures

Much of the natural coverage effects research considers volume and content of the studied topic. Having evidence of the general nature of news coverage of celebrity cancer disclosures is necessary in guiding future research questions and to explain media effects research findings. If coverage is simply assumed by researchers, there is great potential for misguided inquiry. For example, Borzekowski and colleagues (2014) tested public learning outcomes regarding risk and BRCA, and they found that less than 10% of the sample could accurately interpret Jolie's risk of "developing cancer relative to a woman unaffected by the BRCA gene mutation" (p. 516). A subsequent content analysis (Kamenova, Reshef, & Caulfield, 2014) found that very little news coverage of Jolie's disclosure offered risk information or any breast cancer-related information beyond Jolie's own experience. Having a more nuanced understanding of news coverage content of these disclosures may guide researchers to ask more critical questions (e.g., might there be deleterious outcomes related to the disclosures?). In Noar et al.'s (2014) review and research agenda of public figure cancer announcements, of the 19 identified studies on celebrity cancer disclosures, only five included the volume of news coverage and four

considered the content of the coverage, with only one study using the frequency/content as predictor variables (i.e., to predict increased breast cancer screening) (Chapman, McLeod, Wakefield, & Holding, 2005).

In general, the volume of coverage is high and only persists for a few days to a week after the disclosure is announced (Ayers, Althouse, Noar, & Cohen, 2014; Metcalfe et al., 2011; Noar et al., 2014). The limited evidence that is available reveals that the cancer aspect of the story and/or how the celebrity's cancer or cancer risk is relevant to the general population is not generally the focus of news coverage. Of the 1203 UK newspaper articles of Jade Goody's cervical cancer announcement and her death of cervical cancer, only 9.6% of the articles provided any information to infer methods of reducing personal cervical cancer risk (Metcalfe et al., 2011). Of the 103 newspaper stories discussing Angelina Jolie's op-ed for *The New York Times*, only 32% of the stories discuss that Jolie's gene mutation is rare (Kamenova et al., 2014).

Further substantiating this trend of little focus on the cancer-aspect of the story by the press, of the 550 UK newspaper articles reporting on Patrick Swayze's announcement of pancreatic cancer, only 180 included the words "pancreatic cancer" (Williamson & Hocken, 2010). One content analysis on the main themes in coverage of 17 Australian news broadcast segments of Kylie Minogue's breast cancer diagnosis demonstrated that media coverage may actually spread inaccurate information about breast cancer and general population risk. Chapman et al. (2005) found that news coverage of Minogue's disclosure emphasized that breast cancer can affect women of any age, early detection is important, and routine mammograms are important for women of all ages. These recommendations are at minimum an over generalization (breast cancer is more likely to

affect older women) to simply inaccurate information (routine mammograms are not recommended for women of average risk under recommended ages [depending on the recommendations, ages 40-49]). Finally, in a content analysis of 17 celebrity breast cancer disclosures from 1992–2014, Sabel and Dal Cin (2016) found that media reports of celebrities opting to undergo bilateral mastectomies dramatically increased in 2008 and 2009. They also found that media tone was more negative towards mentions of treatment decisions, such as chemotherapy and radiation, and more positive towards surgical decisions, particularly bilateral mastectomies. While not quantified, the authors do mention that most of the news articles focused on the individual celebrity's experience with cancer and not on population or subpopulation breast cancer-related information (Sabel & Dal Cin, 2016).

Origins of content in news coverage of celebrity cancer disclosures. The studies discussed above comprise the entire body of research focused on the quantity and quality of media reports of celebrity cancer and breast cancer disclosures. Due to the limited number of studies and some limitations with methodology (e.g., nonrepresentitve sampling, one specific disclosure), generalizations about media coverage cannot be made. However, these studies do begin to indicate that news coverage of celebrity cancer disclosures largely focuses on the celebrity, and there may be issues with inaccurate reporting. These reporting outcomes are likely influenced by two factors: 1) in terms of inaccuracy, overgeneralizations, or misinformation (which will all be discussed in further detail in following sections), the origins of such communication stem from the complicated nature of cancer, including breast cancer; and 2) specific journalistic norms that guide story selection and framing.

Limited scientific understanding. Changing scientific understanding of breast cancer etiology, prognosis, prevention, and screening poses challenges for breast cancer communication at both the individual and media systems levels. Awareness of breast cancer is high among the general public (Jensen et al., 2014), yet understanding of the intricacies of the disease, risk for certain subpopulations, recommendations for screening initiation, and treatment options is minimal (Borzekowski et al., 2014; see Hersch et al., 2011). While a celebrity breast cancer disclosure could provide an opportunity to better inform the public about breast cancer-related decision making, for example, the content analyses reviewed indicate this opportunity is unlikely to be realized.

Journalists have received significant critique from clinical and communication scholars about how health and science topics are covered. Some of the "issues" with reporting are a result of normative influences from the profession and practical concerns in the industry. But journalists' own scientific understanding of the disease may also be a factor in the case of overgeneralizations or inaccurate breast cancer-related information. Research is scant on the practice of health and science journalism, but a few studies provide some descriptive information about journalists' backgrounds. In a 2008 study, Viswanath and colleagues (Viswanath et al., 2008) found that nearly 70% of the respondents had a bachelor's degree, 19% had a master's degree, and 4.5% had a doctoral or M.D. degree. Over half of these held degrees in journalism and nearly one-fifth held degrees in communication. Only 8% had degrees in what was categorized as the "life sciences." Of the respondents, over 33% were working journalists for 20 years or more, but not necessarily as health or science journalists. To be clear, level of education is not

necessarily predictive one's understanding of health information, but may explain some of the inaccuracies reported in celebrity cancer disclosures.

Journalistic norms. The decision to report on a celebrity cancer disclosure may fit into three newsworthiness criteria held by health journalists: potential for public impact; ability to provide a human-interest angle; and coverage by competitors (Viswanath et al., 2008). The predominant emphasis on the celebrity's experience with cancer (as opposed to cancer-related information relevant to the general population or at-risk subpopulations) in media coverage may be influenced by structural characteristics (organizational ownership) and framing norms. Journalists from private organizations (e.g., *Associated Press*) reported that "educating people to make informed decisions is an important priority in their health reporting," (Wallington et al., 2010, p. 89), although many national news organizations are publicly held companies (e.g., *The New York* Times; *The Washington Post*) where "educating people" is reported to be far less of a priority (Wallington et al., 2010). However, journalists employed at privately owned news organizations are also "twice as likely to say that providing entertainment is important," (Wallington et al., 2010, p. 89). Regardless of institution type, journalists with less than 15 years of reporting experience also indicate that entertainment is a priority in storytelling and journalists with a bachelor's degree or less were more than two times likely to indicate that reporting with a human-interest angle is important.

The prioritization of emphasizing a human-interest component in health news or to place educating the public further down the list of priorities may seem surprising in the context of health reporting. However, most health journalists are trained as journalists before working on the health beat, and educating the public or advancing health literacy

are not values held in the journalism profession (Wallington et al., 2010). Just as in political or public affairs reporting, health journalists do not believe their job is tell people what or how to think, but to simply provide the public with information (Amend & Secko, 2012; Hallin & Briggs, 2014). Hallin and Briggs (2014) push back on the criticisms from health communication scholars and health promotion advocates who say journalists should do a better job communicating health information, including focusing less on event-driven and human interest coverage: "Reporting things that are 'novel or controversial and yet likely to be relevant to many individuals, however, is at the core what journalism is as a social practice and a form of knowledge production" (p. 93). This statement echoes what Amend and Secko (2012) found; just as political reporters do not communicate certain information simply because a particular stakeholder may want it publicized (e.g., political candidates, political institutions), health journalists resist the notion that they have a responsibility to communicate specific health information—or present the information in a particular way—that academics, health care or public health practitioners, or institutions feel they should communicate. After all, health journalists are serving the public and not these other interest groups (Amend & Secko, 2012; Halin & Briggs, 2014).

Discussion of health journalists' professional norms and perceived roles is applicable to reporting on breast cancer disclosures, however, to only focus on health journalism would be shortsighted in the context of the reporting of celebrity breast cancer disclosures. There is limited evidence, but Kamenova and colleagues' (Kamenova et al., 2014) findings indicate that health journalists may not be those primarily reporting on these disclosures. In Kamenova et al.'s (2014) content analysis of news coverage of

Jolie's disclosure, only 7.8% of the articles in the sample were in the "health" sections and only 4.9% of the stories were written by science/health journalists. Nearly 39% of the articles appeared in the news sections of newspapers and 4.9% were found on the front page. Thirty-one percent were news stories and 23.3% were editorial or opinion pieces. Over 41% of the articles were written by staff writers and 9.7% were written by the newspaper's regular columnists (Kamenova et al., 2014). If the news coverage of celebrity cancer disclosures are generally reported on by staff writers, rather than health journalists, this may shed further light on the high rates of coverage and the potential emphasis on the celebrity's disclosure and her or his own experience with cancer.

It was stated earlier that health journalists' norms are informed by professional journalistic norms, but their specialty does give them a somewhat different perspective than traditional journalists. This is pointed out by Hallin & Briggs (2014): "[R]eporters covering this 'beat' [health/medical journalism] often express more didactic and instrumental conceptions of their role than other journalists" (p. 92). General news values, as originally identified by Gans (1979), include prominence, human interest, conflict, novelty, timeliness, and proximity or local appeal. These shared news values held by editors and reporters are used when "determining priorities for reporting and newsworthiness" (Wallington et al., 2010, p. 79). The general news values cited here provide insight as to why celebrity cancer disclosures are frequently reported on and why little cancer information may be included in these reports.

A celebrity publicly announcing her or his experience with a serious illness, such as cancer, is the embodiment of several news values (e.g., prominence, human interest, timeliness). And in our current media environment the reporting of celebrity cancer

disclosures has been validated consistently—high numbers of clicks on online news websites and high online search volume for celebrities proximate to their disclosures (Dean, 2015; Noar et al., 2015, 2014). Given who is reporting on these disclosures (i.e., staff writers) and that the reporting is about a celebrity, it would be unusual to find cancer-related information beyond the celebrity's own experience contained in the story. But this is not specific to health disclosures from celebrities, and instead is how journalists have historically reported on celebrities (de Leon, 2002). Relaying the intimate details of a celebrity's life is believed to be more enticing to readers than conveying a broader contextual message (de Leon, 2002; Lerner, 2006).

A note on the public's role. It is tempting to interpret the discussion on journalists limited scientific understanding and norms as a specific critique on the practice of journalism or journalists. If only journalists had a better understanding of science…, if only journalistic norms placed an emphasis on educating the public, then the public would better understand breast cancer, for example. Unfortunately, these turn of events would not guarantee a deeper understanding of health issues by the public. The public too has a limited understanding of science. For example, data from 2007 demonstrates that U.S. adults with a "minimal level of understanding the meaning of scientific study" is at 29%, although a significant increase from only 12% in 1957 (Miller, 2010). But even "minimal levels of understanding" are likely insufficient to adequately interpret scientific (including health) information. A recent Pew Research Study found evidence of this. Their findings indicate that the public perceives little scientific consensus on topics that are largely agreed upon in the scientific community. Examples of such a disconnect include: 37% of participants believe that scientists do not agree on the presence of

climate change; 67% do not believe scientists have a strong understanding of genetically modified organisms' health effects; and 52% believe scientists are divided on "the Big Bang" theory (Pew Research Center, 2015; Nagler & LoRusso, 2017). The potential for the public's understanding of breast cancer is discussed further in the Cancer Information Seeking section of this literature review.

The framing of celebrity cancer/health news. The above discussion points to why journalists and editors select to report on celebrity cancer and health disclosures and that shared news values impact the content of the news reports. How the content is reported on or presented to the public, from an academic perspective, is how the information is *framed* (Scheufele, 2000). Scheufele (2000) states: "Mass media actively set the frames or references that readers or viewers use to interpret and discuss in public" (p. 105). Gamson and Modigliani (1989) define a media fame as a "central organizing idea or story line that provides meaning to an unfolding stream of events… The frame suggests what the controversy is about, the essence of the issues" (p. 143). Tuchman (1978) also offers: "The news frame organizes everyday reality and the news frame is part and parcel of everyday reality…[it] is an essential feature of the news" (p. 193). These definitions have subtle differences, yet they all imply the framing of a story impacts public outcomes. Essentially, the way a story is told—and the information included within it—may affect an individual's thoughts or activate certain schemas (Scheufele, 2000).

There appears to be virtual consensus that media frames are enacted in news production, but the construct of media frames in the extant literature is "fractured, fragmented, and inconsistent at best," (Entman, 1993, p. 51) primarily because many

researchers create and test their own conceptualizations of media frames and do little to add construct validity of previously tested frames (Entman, 1993; Iyengar, 1990; Matthes & Kohring, 2008; Scheufele, 2000; Tuchman, 1978). Public health researchers have likely contributed to this fractured and fragmented study of media frames, with the investigation of numerous framing dimensions that are similar conceptually but presented as distinct: Gain and loss frames; lifestyle, political, economy, and medical frames; behavioral, environmental, and systemic frames; and public health model of reporting and traditional reporting (see Hawkins & Linvill, 2010; Henry, Trickey, Huang, & Cohen, 2012; Lawrence, 2004; Sangalang, Hurley, & Tewksbury, 2015).

Coleman, Thorson, and Wilkins (2011) discuss the role of traditional reporting and the public health model of reporting: "Public health experts are not always satisfied with the way the media report health news. The focus on individuals and anecdotes at the expense of context and societal contributions to disease gives people a distorted view of the problem, they say" (p. 1). Coleman et al. (2011) contend that the public health model of reporting is how health experts would prefer journalists report on health issues. They go on to define the public health model of reporting as follows: "The *public health model* is defined as an approach that sees the causes of death and injury as preventable rather than inevitable. By studying the interaction among the victims, the agents, and the environment, this approach seeks to define risk factors, then develop and evaluate methods to prevent problems that threaten public health" (Coleman et al., 2011). They claim that while this model of reporting is ideal from a public health perspective it is rarely realized in practice—this stance is in agreement with the previous discussion on journalistic norms, values, and practices. Coleman and colleagues (2011) refer to

reporting of health-related stories that do not focus on causes of the disease, risk factors, and prevention strategies as simply "traditional reporting." To place the public health model of reporting in traditional mass communication research, Coleman et al. (2011) compare traditional reporting and the public health model of reporting to Iyengar's episodic and thematic frames (e.g., Iyengar & Kinder, 2010; Iyengar, 1987, 1990).

Iyengar's episodic and thematic frames have been used to categorize media content more than any other media frame typology (Matthes & Kohring, 2008). According to Iyengar (1987, 1990) an episodic frame is the depiction of issues in the form of concrete instances or specific events, and a thematic frame presents issues on a more abstract level which implicate general outcomes. Iyengar originally used the episodic/thematic media frame typology in the context of public affairs reporting. While research using this set of frames is still common in political communication (Gross, 2008; Jha, 2007; Matthes & Kohring, 2008; Semetko & Valkenburg, 2000; Smith, McCarthy, McPhail, & Augustyn, 2001; Zaharopoulos, 2007; Zillmann, Chen, Knobloch, & Callison, 2004), in more recent years health communication scholars have focused on these frames in the context of obesity (Gearhart, Craig, & Steed, 2012; Hatley-Major, 2009), lung cancer (Hatley-Major, 2009), fetal alcohol syndrome (Connolly-Ahern & Broadway, 2008); autism (Holton, Weberling, Clarke, & Smith, 2012; Holton, Farrell, & Fudge, 2014), and H1N1 (Lee & Basnyat, 2013). Studies which quantify the presence of episodic or thematic frames, typically find that most reporting uses episodic frames (e.g., Holton et al., 2014; Matthes & Kohring, 2008).

Coleman et al.'s (2011) discussion of traditional reporting and the public health model of reporting, and applying it to celebrity breast cancer disclosures, is in line with

how Holton, Farrel and Fudge (2014) describe episodic and thematic frames in health and science news coverage. For example, “Episodic frames… might introduce an individual to represent a particular illness or disability, helping provide insightful viewpoints that otherwise go unheard. These frames are likely to focus on personal lifestyles and to suggest individual solutions or cures” (p. 193). This conceptualization of an episodic frame in health news is an apt description of an episodic frame in news coverage of a celebrity breast cancer disclosure—it will likely convey information about the celebrity (age, career, relationships, the context of the disclosure, and the celebrities experience with breast cancer).

Also in the context of health news, Holton, Farrel and Fudge (2014) state: “Thematic frames place more emphasis on the connection between issues or events and society. In the case of autism, journalists might focus on the role of the science community or public fundraising in finding causes, treatments, or cures for autism. Thematic approaches help individuals relate to stories and understand how they connect them with other people” (p. 193). Again, this conceptualization of a thematic frame is applicable to what a thematically framed media report of a celebrity breast cancer disclosure might look like. The story will still contain much of the information found in a story with an episodic frame, but will then connect the celebrity’s breast cancer experience to a larger public health-related theme (population/subpopulation risk factors; screening guidelines; treatment options; survivorship).

The categorizing of information as either episodic or thematic in news coverage of celebrity breast cancer disclosures is admittedly simple and likely misses some of the more nuanced messages being conveyed. However, for this dissertation, while

quantifying the presence of episodic and thematic frames will in part represent the content of these disclosures, the frames will also serve as predictive evidence for breast cancer-related information seeking outcomes. To date, longitudinal studies using media content as an independent variable have only used dichotomous measures of content (Romantan, 2005; Stevens & Hornik, 2014; Stryker, 2003). It is possible that more nuanced measures of content could be used, but further stratification of content may create issues of power (Stryker, 2003). It should be noted that specific types of content will also be measured for further purposes of this study and is discussed in forthcoming sections.

Potential impact of episodic and thematic frames. Research of episodic/thematic frames often quantifies the presence of the frames in actual news content, but is not limited to descriptive findings. Iyengar (1987, 1990) hypothesized and later found that episodic frames lead the reader to attribute responsibility at the individual level and thematic frames elicit responses which attribute responsibility to societal/systemic factors (Iyengar & Simon, 1993). Iyengar's work is firmly placed in the political/public affairs news space and uses public opinion polling data to confirm hypothesized framing effects. However, as with content analysis, episodic and thematic frames linked with the concept of attribution of responsibility (or blame) has been tested numerous times in the context of health reporting and health issues (e.g., Hatley-Major, 2009; Wise & Brewer, 2010) and typically the results support Iyengar's original hypothesis. Attribution of responsibility, specifically in the context of health issues (e.g., smoking, obesity) has also been found to be predictive of policy support (thematic frames are more likely to elicit

responses in support of policy measures meant to improve conditions and/or reduce rates of the given health issue) (see Hatley-Major, 2009).

Almost no research has considered potential effects of health stories with episodic or thematic frames beyond the context of attribution of responsibility and policy support. As Coleman, Thorson, and Wilkins (2011) likened the public health model of reporting to a thematic frame, which focuses "on causes of the disease, risk factors, and prevention strategies" (p. 2), they also postulated that such information could have effects beyond the usual line of inquiry—episodic and thematic framing may "affect people's intentions to change their own health behavior," (p. 3). Across numerous public health issues (diabetes, smoking, obesity, and immigrant health), the researchers found that thematic frames were "significantly more likely to cause people to say they intended to change their own behavior" (Coleman et al., 2011, p. 8) than episodic frames.

Coleman et al. (2011) are not able to provide causal mechanism evidence as to why a thematic frame may impact behavioral intentions, but they discuss some possibilities. Thematic frames offer information about the disease, risk factors, and prevention strategies that one could easily use to ascertain specific strategies (behaviors) to lessen her or his likelihood of experiencing the illness and/or decrease morbidity. The inverse to this cognitive process is likely to occur from an episodic frame. Because the episodic frame only offers information specific to the event (i.e., celebrity breast cancer disclosure), someone exposed to such information is less likely to draw specific conclusions about how the health issue (i.e., breast cancer) might apply to herself or himself and/or what behavioral modifications one could make to reduce risk for occurrence (Coleman et al., 2011).

The general nature of a media report with an episodic frame—lacking cancer-related information (e.g., causes, risk factors, prevention strategies)—has the potential to be confusing or ambiguous to the average news consumer. Ambiguous information is defined as information that is difficult to interpret or is inherently confusing (Ellsberg, 1961; Han, Kobrin, et al., 2007; Hurley, Kosenko, & Brashers, 2011). The cognitive effect of ambiguity occurs when there is a question of "reliability, credibility, or adequacy" from the given information (Han et al., 2007). Little content analytic research has been done to quantify the presence of ambiguous health or cancer information in news content, but Hurley et al. (2011) used Babrow, Kasch, and Ford's (1998) information quality of completeness (having too little information) for one dimension in the coding of uncertainty information in health news coverage. Self-report data on the public's perception of cancer news coverage has found it to cause ambiguity due to the conflicting nature of information (e.g., varying guidelines for cancer screenings) and, in line with Hurley et al.'s research (2011), insufficient information (Clarke & Everest, 2006; Han, Moser, & Klein, 2007; Nagler, 2014; Niederdeppe, Fowler, Goldstein, & Pribble, 2010).

Ambiguity can be a driving force in deleterious health-related attitudinal, cognitive, and behavioral outcomes. For example, media exposure has been found to be positively correlated with perceived ambiguity towards cancer prevention recommendations and cancer fatalism (Han, Moser, et al., 2007; Niederdeppe et al., 2010). Nagler (2014) found that of the participants who reported greater exposure to contradictory nutrition information, they also reported higher levels of nutrition confusion. Nutrition confusion was associated with greater backlash, and the confusion

and backlash "were negatively associated with intentions to engage in healthy lifestyle behaviors" (p. 12).

No content analytic work has specifically measured episodic (or ambiguous) frames in the news coverage of celebrity breast cancer disclosures, but some findings implicate the presence of episodic frames and the potential for ambiguous information. For example, Angelina Jolie's op-ed in *The New York Times* was about her decision to have a prophylactic bilateral mastectomy after genetic testing confirmed she is BRAC1 positive, and therefore considered to be at high-risk for breast cancer. Jolie's decision to undergo genetic testing was largely influenced by her strong family history of breast and ovarian cancers. Thus, the story of Jolie's genetic testing and subsequent surgical decisions were only applicable to a very small percentage of women, and not relevant to women of average risk. However, as previously mentioned, Kamenova, Reshef & Caufield (2014) found that only 32% of the news stories covering Jolie's disclosure stated that her genetic condition is rare. Because the news coverage does not make clear what factors put a woman at greater risk for having a BRCA mutation positive result, the information, or lack thereof, can be considered ambiguous. Perhaps this ambiguous coverage was responsible for findings where 90% of participants were unable to interpret Jolie's risk for breast cancer to a woman of average risk (Borzekowski et al., 2014), or when women with no family history of breast cancer (but identified strongly with Jolie) reported intentions to have BRCA genetic testing (Kosenko et al., 2015).

It is possible that the presence of episodic frames is more common than not in media coverage of celebrity breast cancer disclosures, and that this content can be linked with data indicative of cognitive ambiguity, but thus far, no research has attempted to

make this connection. In the forthcoming cancer information seeking section, the connection between the cognitive effect of ambiguity and cancer information seeking will be explicitly discussed.

Misperceptions of breast cancer. Nyhan (2010) describes misinformation or misperceptions as "[F]alse or unsubstantiated beliefs that are confidently held by members of the public, potentially distorting their issue preferences…" (p. 2). As discussed in the introduction, breast cancer screening recommendations have historically endorsed women of average risk initiating annual screening at age 40. Considerable amounts of resources went into bringing awareness to screening initiation through strategic health communication campaigns, awareness campaigns from nonprofit organizations, and news coverage. Eight years after the change in screening recommendations from the USPSTF many women still have a firmly held belief that they should begin routine screening at age 40 (Hersch et al., 2013; Oeffinger et al., 2015; Squiers et al., 2011; Waller et al., 2013). This publicly held misperception was partially created by the medical and public health communities. Celebrity breast cancer disclosures have the potential to either reinforce or refute such misperceptions.

Misinformation or misperceptions, particularly regarding health beliefs, can be incredibly difficult to correct. Individuals often accept new information of an unknown topic quite readily, but once we have formed a belief about the topic, the introduction of disparate information often creates great skepticism, making it almost impossible to change these existing beliefs (Nyhan, Reifler, & Ubel, 2013; Nyhan, 2010; Southwell & Thorson, 2015).

In the context of celebrity breast cancer disclosures, the concepts of misinformation or misperception can be applied in two ways. First, currently held misinformed beliefs that were largely created from information that was based on scientific consensus at one time (e.g., women of average-risk for breast cancer should initiate routine annual breast cancer screenings at age 40) can then be reinforced by information provided by the media in reports of celebrity breast cancer disclosures. An anecdotal example of such reinforcement can be found in coverage of *Good Morning America's* anchor Amy Robach's experience with breast cancer. She was nearly 41 when she had a mammogram live on air. Later, she was diagnosed with breast cancer in one breast. In subsequent news coverage of her experience, Robach laments of her guilty feelings towards not getting screened as soon as she turned 40—she even says she kept putting off screening (Sulik, 2013). This coverage did not include any discussion of the varying age recommendations by multiple organizations or that the USPSTF recommendations call for informed conversations with a woman's provider. Robach is reinforcing the belief that women of average-risk for breast cancer should begin screening at age 40.

Other breast cancer-related misinformation that falls within the change of scientific discovery include statements which endorse self-breast exams—self-breast exams are no longer recommended by the USPSTF or the ACS (American Cancer Society, 2015; U.S. Preventive Services Task Force, 2009). Also, the messages that "early detection is key" to survival and that "screenings save lives" are too simplistic to be accurate (Chapman et al., 2005; Sulik, 2013). Specific breast cancer stages (0, 1, 2, 3, 4) are used as a general description of a particular cancer's size and growth. But any

emphasis on "catching" the cancer early assumes a linear progression of the cancer, which is often not the case. It also implies that if breast cancer is diagnosed in Stage 0 or 1, for example, then the breast cancer can be stopped from progressing (Welch & Frankel, 2011; Wilt & Partin, 2011). This is inaccurate—some cases of breast cancer are so aggressive that treatment will not stop the progression and other cases may be very slow growing and never progress at all. "Screening saves lives" is also a gross overstatement. Several studies have found increased rates of screening have not resulted in a reduction of population-level mortality rates (Dalton et al., 2003; Gøtzsche & Jørgensen, 2013; Kalager, Zelen, Langmark, & Adami, 2010; Welch & Frankel, 2011). Furthermore, higher rates of breast cancer screenings have led to over diagnosis and over treatment. To convey to the public that a mammogram will save a woman's life oversimplifies the complexities of the disease and treatment (Harris & Sheridan, 2013; Harris, 2014; Prasad, Lenzer, & Newman, 2016).

Other misperceptions regarding breast cancer are based on information that was never based on scientific consensus and can be attributed to the media's interpretation of statistics and focus on certain aspects of personal accounts of breast cancer. For example, the prevalence of breast cancer diagnosis is often misrepresented—the statistic "1 in 8" or a 12% probability is given (Corbett & Mori, 1999; Covello & Peters, 2002). In fact, a woman only reaches a 12% chance of being diagnosed with breast cancer in her 70s. More accurate probabilities include a .44% chance for a woman age 30, a 1.47% chance at age 40, a 2.38% at age 50, a 3.56% chance at age 60, and a 3.82% chance at age 70 (National Cancer Institute, 2017). The misrepresentation of the incidence of breast cancer

may explain why women tend to significantly overestimate the prevalence of breast cancer diagnosis (Rahib et al., 2014).

Another misrepresentation of breast cancer risk is who can be diagnosed with breast cancer. As the above probabilities indicate, women of all ages can indeed have a breast cancer occurrence; however, the younger a woman is, risk is greatly reduced. Personal narratives about experiences with breast cancer, including celebrity breast cancer disclosures, have been found to emphasize "any woman" or "even young women" can be diagnosed with breast cancer and fail to include how low that risk really is. As mentioned previously, Chapman et al. (2005) found the content of the media's coverage of Kylie Minogue's breast cancer diagnosis, who was 36 at the time, emphasized that young women do get breast cancer. In the months following Minogue's disclosure there was a 100% increase in breast cancer screenings for women aged 40-69 (in Australia) who had never been screened (Chapman et al., 2005) While this result can be seen as a positive public health behavioral response, a later study uncovered a more negative outcome. In the year following Minogue's disclosure, doctor-referred breast imaging in Australia for women aged 22-44 rose dramatically (Kelaher et al., 2008). The authors concluded that media coverage of Minogue's disclosure influenced the beliefs of women at low-risk beliefs about their own risk (Kelaher et al., 2008).

Finally, the media's focus on individual's breast cancer treatment decisions may be creating misinformed beliefs regarding surgical decisions. Evidence supporting efficacy of bilateral mastectomies when breast cancer has only been found in one breast falls under a very specific situation: A woman who has been diagnosed with breast cancer and has tested positive for a BRCA(1/2) genetic mutation. For all other breast cancer

diagnoses – breast cancer diagnosis located in one breast or BRCA(1/2)—there is no evidence that a prophylactic mastectomy will prevent mortality (Li et al., 2016). Instead, breast conserving treatment (e.g., a lumpectomy with radiation) has proven to be an effective treatment and is less invasive than a bilateral mastectomy. Sabel and Dal Cin (2016) found that media coverage of several celebrities who opted to have a prophylactic mastectomy (despite not being optimal candidates for the procedure) failed to include information about genetics, family history, risk, or efficacy of the procedure. The authors also found that an increase in media coverage of these surgical decisions coincided with dramatic increases in population level bilateral mastectomy rates. Sabel and Dal Cin (2016) conclude that this rise in bilateral mastectomies is influenced by the media coverage of celebrities' surgical decisions.

There is limited supporting evidence to conclude that celebrity breast cancer disclosures include information that may either reinforce currently held misperceptions or create new misperceptions. Most of the evidence discussed above is either anecdotal or is extracted post hoc. Empirically identifying specific passages of inaccurate breast cancer-related information would help to support these suppositions and provide a framework for future research regarding public misperceptions of breast cancer.

Cancer Information Seeking

There is strong evidence that news coverage of health issues can impact cognitive, affective, and behavioral health-related outcomes (e.g., Fishbein & Hornik, 2008; Nagler, 2014; Niederdeppe et al., 2013; Niederdeppe, Frosch, & Hornik, 2008). As has been discussed throughout this literature review, there is support linking news coverage of celebrity cancer disclosures and public cancer-related outcomes at all levels (see Noar et

al., 2014). Because this dissertation seeks to test the relationship between media coverage of celebrity breast cancer disclosures and public breast cancer-related outcomes, the decided upon outcome variable should be information-based and relevant to the numerous possible disclosure types (e.g., various breast cancer types, different courses of treatment). Breast cancer-related information seeking is an obvious and important potential outcome. Cancer information seeking is what Noar and colleagues (Noar et al., 2014) refer to as an intermediate behavioral outcome, because it is often an important mediator in the process of decision making about cancer/health-related behaviors (Lambert & Loiselle, 2007; Lee, Zhao, & Pena-y-Lillo, 2016).

The health/cancer information seeking literature has typically used self-report data to verify seeking behaviors among a variety of sources, including interpersonal and media sources (Barbour, Rintamaki, Ramsey, & Brashers, 2012; Dobransky & Hargittai, 2012; Lambert & Loiselle, 2007; Niederdeppe et al., 2007). With the changing media environment, especially with the proliferation of digital media, many researchers have turned to aggregate online search engine data to measure health information seeking. The ease and availability of online aggregate data has influenced this shift in health information seeking measurement. However, statistics on Internet use and online health information seeking behaviors make for a compelling argument of the utility of using aggregate search data. As of 2014: 1) 87% of Americans use the Internet; 2) 70% of these users have searched for health information on the Internet in the last year; and 3) 77% of those who sought health information began their search using a search engine, such as Google (Mitchell, Jurkowitz, & Olmstead, 2014).

Pathways leading to online health/cancer information seeking. Not only does much of the research on health information seeking rely on self-report data, but it also typically focuses on individuals who were diagnosed with a condition or those in certain at-risk subgroups (Gollop, 1997; Lambert & Loiselle, 2007; Myrick et al., 2014; Shim, 2008). For individuals in these groups, perceived risk and ambiguity, among others, have been found to be mechanisms of information seeking. Online aggregate search data are unable to provide information about who is seeking and why they are seeking. This opaque data does create some issues with construct validity in research—what does aggregate search query data measure? To explain aggregate information seeking outcomes, scholars either simply present descriptive findings or rely on established media effects and information seeking theories, resulting in varying conclusions on the data's meaning and significance.

A large stream of research uses aggregate online search data for epidemiological purposes. Using search data and real-world incident rates of influenza, specific cancers, and Lyme disease, several studies have evidenced Google Trends' ability to track disease prevalence in real time (Cooper, Mallon, Leadbetter, Pollack, & Peipinism, 2005; Polgreen, Chen, Pennock, Nelson, & Weinstein, 2008). Some research has used these data to create statistical models which can predict outbreaks in certain locations (Choi & Varian, 2012; Wilson & Brownstein, 2009). Again, because of the use of aggregate data, mechanisms of seeking behaviors cannot be tested, but the positive correlations found between disease incidence rates and search data indicate that the public is searching for disease symptoms slightly before diagnosis, making it possible to predict future outbreaks.

This line of research has been put into question because it does not account for other factors which may impact public online information seeking behaviors, such as exposure to news media coverage. For example, Towers and colleagues (Towers et al., 2015) investigated the potential relationship between media coverage of Ebola and U.S. Google search queries related to Ebola. While Ebola was a proximal threat in West Africa, there were only four confirmed cases of Ebola in the United States, guaranteeing that dramatic increases in search queries for Ebola could not be predictive of actual occurrence in the U.S. The researchers found that volume of media coverage was predictive of the relative search volume for general search terms related to Ebola and for search phrases that were more indicative of personal concern for the disease (e.g., "do I have Ebola"). The authors assert that fear was a motivating factor in increased search queries indicating personal concern. Towers et al. (2015) prove the point that for certain health issues, illness, and disease, aggregate online search data are not valid measures of predicting incidence rates, and research using these data must account for larger, current, social, and cultural trends.

Mass media effects concepts. The assertion of fear as a causal mechanism for the public's seeking behaviors could be accurate, but Tower's et al.'s research is not couched in media effects, communication, or the information sciences research traditions. That is, the justification for testing the relationships between media and information search query data is not based on theoretical underpinnings, and the observational conclusions do not add further insight to communication processes. Yet this approach is not isolated to Towers et al. (2015), and even communication scholars have offered similar conclusions. For example, Ayers et al. (2014) investigated online search queries related to smoking

cessation in Brazil after Brazilian President Lula da Silva's announcement of his laryngeal cancer diagnosis, which he attributed to smoking. The study found that the highest amount of relative online searches occurred eight days after cessation media coverage had spiked (Ayers et al., 2014). This same research team investigated search volume increases for queries related to breast cancer and BRCA, proximal to Angelina Jolie's breast cancer-related disclosure. Some search queries had over a 9,000% increase in the week following her disclosure (Noar et al., 2015).

These results indicate that exposure (direct or indirect) prompted online information seeking, but *why* this occurred is not explored. This line of inquiry is most similar to studies of online information seeking that have invoked the concept of agenda-setting. McCombs and Shaw (1993) described agenda-setting as the "ability [of the news media] to influence the salience of topics on the public agenda," (p. 58). Traditionally, agenda-setting research has used a combination of content analysis of news media content (dominant topics covered by the news media) (i.e., media salience) and survey data (i.e., public salience) to test for correlations between the two. The "most important problem" (MIP) question—what issues the public believes to be the greatest in the country—asked of respondents represents public salience. The economy, jobs, and immigration, are typically at the top of the list, and as the agenda-setting hypothesis posits, are also issues that frequently receive significant news coverage (McCombs & Shaw, 1993; McCombs & Shaw, 1972; Mellon, 2014).

In more recent years, online search query data has been used to represent the MIP measure in traditional agenda-setting research. For example, Lee, Kim and Scheufele (2015) found that several search queries for economic-related search terms predicted

issue salience when compared with MIP results for the same time period. Mellon (2013) found strong positive correlations for relative search volume and the percentage of Americans who indicated fuel prices ("fuel") and illegal immigration as the MIP. However, both studies had inconsistent findings. Mellon (2013), for example, found that while the economy was high on the MIP list from survey respondents, there was no correlation with the online relative search volume data.

There is debate about whether relative search volume can stand as a proxy for the MIP question in agenda-setting research. Several studies have side-stepped this issue by not following a strict interpretation of agenda-setting. Instead, researchers have tested for potential relationships between the volume of news coverage and the degree of increase in the relative search volume for specific topics. For example, over a two-year period, Ragas and Tran (2013) concluded relative search volume for "President Obama" was significantly predicted by the volume of media coverage of the President over the previous five weeks. Showing more immediate effects, Weeks and Southwell (2010) found the volume of newspaper and television news coverage on the "President Obama is secretly a Muslim" rumor was correlated with a "pulse effect" of online information search. That is, there was a strong correlation between same-day newspaper and television news coverage of the rumor and Google searches on the rumor, with a steady decline in searches each subsequent day (Weeks & Southwell, 2010). These studies invoke the concept of agenda-setting as a framework of inquiry, unlike the atheoretical Ayers et al. (2013) and Noar et al. (2015) studies, but their studies' assertions are similar—the volume of media coverage of the studied topic predicts increases in relative search volume.

Again this line of inquiry does not follow a strict interpretation of agenda-setting, and perhaps is more indicative of the concept of media priming, which is closely linked with agenda-setting. The media priming hypothesis posits that because mass media makes some issues more salient than other issues, this attention sets a standard in which the public uses for evaluation. This standard is created by the formation of schema which organizes information into categories and creates relationships among categories. Priming can also bring schema top-of-mind when activated (DiMaggio, 1997; Perse & Lambe, 2016). Because this formation of schemas and subsequent activation can be influenced by media, media priming is generally considered to be an outcome of agenda-setting (Scheufele, 2000; Scheufele & Tewksbury, 2007). Because evaluation is specific to priming—and not agenda-setting—it may be a more accurate construct to ascribe to online information seeking data. However, priming involves a process of developing or changing attitudes and beliefs, so results which are more immediate (e.g., a celebrity health disclosure media event driving a same-day dramatic increase in online search queries) could be best described as agenda-setting effects, whereas more long-term impacts on information seeking outcomes could be indicative of priming effects (e.g., relative search volume of breast cancer was significantly predicted by the volume of media coverage of a celebrity breast cancer disclosure received over the previous ten weeks). At this time there is no evidence to indicate that celebrity breast cancer disclosures have a priming effect on cancer information seeking behaviors. That is, all previous research demonstrates proximal increases in the public's online cancer-related information.

Health communication concepts. Agenda-setting and media priming as media effects concepts are most commonly applied to political and public affairs news coverage and politically oriented outcomes, as demonstrated above. Health communication scholars have historically given less consideration to the impact of news coverage on health-related public outcomes, with the natural coverage effects tradition being an exception. Specific to health information seeking, the news media has often been considered a primary source of health information for the public, but only in the last several years have the effects of news coverage of health issues on health information seeking behaviors been explored (e.g., Niederdeppe et al., 2008; Niederdeppe, 2008; Noar et al., 2014).

In the case of cancer-related information seeking in response to a celebrity breast cancer disclosure there must be exposure to the story (directly or indirectly) before information seeking can take place. I posit that the health communication concepts of information scanning and information seeking begin to help better explicate the proposed exposure → seeking relationship. Paying attention to a news story that is not purposively sought out (e.g., article posted on your Facebook newsfeed or a health-related story on the nightly news) is considered information scanning. It is defined by Niederdeppe, Frosch, and Hornik (2007) as: "Information scanning represents information acquisition that occurs within routine patterns of exposure to mediated and interpersonal sources that can be recalled with a minimal prompt" (p. 154). In contrast, information seeking "describes active efforts to obtain specific information outside of the normal patterns of exposure to mediated and interpersonal sources" (Niederdeppe et al., 2007, p. 155). These two methods of information acquisition are often studied as separate behaviors, but

the two are clearly linked. Paying attention to (or scanning) a news story can lead to information seeking in response to the original exposure (scanning). For example, Weeks et al. (2012) found strong positive correlations in volume of news coverage of the USPSTF change in mammography guidelines and relative search volume in November 2009 (the month the guidelines were released). The pathways involved in such a relationship include exposure to the change in mammography guidelines story (directly or indirectly), processing the information, and then seeking information. Similar pathways can be used to describe breast cancer-related information seeking proximal to a celebrity breast cancer disclosure.

***Ambiguity as a mechanism*.** The above discussion of possible pathways leading to online information seeking all describe the relationship but do not explain *why* this relationship occurs. The results from Mellon (2013) and Lee et al. (2015) partially support agenda-setting effects, but also indicate saliency is not the sole explanation for online information seeking, particularly in the case of complex issues like the economy or cancer. Scholars acknowledge that when a search term is queried, that action certainly means the issue is salient among the seeker, but in the case of aggregate online search data, a lack of relative search volume does not necessarily mean that the search term is not salient with the audience.

A search query is a form of information seeking, and therefore, theories of information must be considered. Indeed, information theorists posit that individuals seek information due to an information sufficiency gap, and the decision to seek more information to lessen the gap requires resources (e.g., time and energy of seeking and processing), so there must sufficient motivation to prompt the act of seeking (Atkin,

1973; Ball-Rokeach & DeFleur, 1976; Knobloch, Dillman Carpentier, & Zillmann, 2003). In the information literature, uncertainty is the common thread for information seeking behaviors. Uncertainty occurs when a situation is complex, ambiguous, or is impossible to predict because information is either unavailable, inconsistent, or uninterpretable (Atkin, 1973; Ball-Rokeach & DeFleur, 1976; Ellsberg, 1961).

The concepts of uncertainty or ambiguity get back to Towers et al.'s (2015) explanation of fear driving the increase in relative search volume for Ebola during the West African Ebola crisis. Uncertainty or ambiguity are components of fear and therefore offer a substantiated causal mechanism for the public seeking information about Ebola symptoms, for example. Retrieving information about the symptoms can reduce uncertainty and calm the purported fear. Uncertainty is in fact tied to anxiety. Seeking health information can be seen as filling an instrumental need, as Case and colleagues (Case, Andrews, Johnson, & Allard, 2005) states: "reducing uncertainty helps us not only maximize future outcomes, but also guards against emotional stress" (p. 355).

The need to reduce uncertainty or ambiguity is well documented in the information seeking literature using self-report data, but is only infrequently cited in information seeking research using online aggregate data. Using ambiguity, or in some cases the concept of "public uncertainty," limited evidence suggests that news coverage of more ambiguous topics positively correlate with significant increases in search queries for related search terms. For example, Maurer and Holbach (2015) compared the relative volume of two distinct topics: the outbreak of an epidemic of hemolytic-uremic syndrome (HUS) in Germany and unemployment. HUS is a relatively uncommon and unknown disease, and therefore characterized as ambiguous (i.e., the absence of knowledge) by

Maurer and Holbach (2015), whereas unemployment is a predictable topic not characterized by ambiguity or uncertainty. Their conceptualizations of the two topics (ambiguous vs. not ambiguous) are substantiated by the study's results—aggregate search volume for HUS was significantly higher than that of unemployment with nearly identical levels of media coverage. The authors conclude with this comment on what they describe as public uncertainty:

> "Finally, we use the concept of uncertainty to explain our findings but do not measure it. Obviously, public uncertainty cannot easily be measured in studies such as ours, which do not use survey data. However, we think that we have good reasons to assume uncertainty in the case of HUS but not in the case of unemployment. As uncertainty is defined as the absence of knowledge, it should occur after surprising events such as the HUS outbreak but not in the case of unemployment, which has been a regularly discussed issue in Germany for decades" (p. 13).

Weeks et al.'s (2012) study of online information seeking for mammography-related information in the month USPSTF changed their mammography recommendations also supports Mauer and Holbach's (2015) assertion that particular topics or stories can be considered ambiguous and therefore prompt online health-related information seeking. Because the change in mammography recommendations were quite dramatic (initiation ages for women of average risk changed from age 40 to age 50), and much of the news coverage focused on the "controversy" rather than the substance of the changes, Weeks et al. (2012) posited that the elements of this story were ambiguous.

These elements created ambiguity in the public, which led to substantial increases in online information seeking for mammography-related search terms.

The concept of public uncertainty, or ambiguity, can be applied to celebrity breast cancer disclosures. The topic of breast cancer, in general, may not automatically promote breast cancer information seeking, because most people likely know what breast cancer is. However, certain features of a celebrity breast cancer disclosure may be more analogous to ambiguity. News articles that discuss the circumstances surrounding a celebrity's diagnosis, but do not mention when an average-risk woman should begin screening, may be ambiguous. Articles that discuss the surgical procedures undergone by the celebrity, but do not discuss the generally recommended treatment options for that breast cancer diagnosis type, are likely to cause ambiguity. This type of reporting falls into the media frame category of episodic (as discussed previously). If episodic frames are the predominant frame used when reporting on celebrity breast cancer disclosures, theoretically the potential for public uncertainty in response to celebrity breast cancer disclosures is high.

Chapter 3

Hypotheses and Research Questions, Conceptual Model of Effects, and Research Overview

The previous chapter provides theoretical support for investigating news media coverage of celebrity breast cancer disclosures and the potential effects that coverage may have on public breast cancer-related information seeking outcomes. In addition, support is provided for considering what moderating influence celebrities' attributes and cancer-related events might play on the media and information seeking outcomes. In this chapter I offer the dissertation's hypotheses and research questions (brief rationales based on theoretical and empirical support are included), propose a model of effects, and briefly summarize the two studies which test specific aspects of the model.

Hypotheses and Research Questions

No research has explored journalistic decisions for reporting on celebrity cancer disclosures. Due to longstanding journalistic norms, including coverage of celebrity news (de Leon, 2002; Schudson, 2003) and from the limited evidence available from prior content analyses of media reports of celebrity cancer disclosures (Chapman et al., 2005; Sabel & Dal Cin, 2016), it is likely that journalists will not contextualize the celebrity's breast cancer disclosure into a greater discussion of breast cancer (e.g., prevalence rates, screening guidelines, risk information). Instead, it is expected that media reports will focus on the celebrity, aspects of her or his life and career, and her or his experience with breast cancer. This focus is conceptualized as an episodic frame in the literature review, whereas a thematic frame will include breast cancer-related information beyond the celebrity's experience (public levels of incidence rates, risk information, etc.).

H1: Higher proportions of episodic media frames, as opposed to thematic frames, are expected to be found in the sample of media reports on celebrity breast cancer disclosures.

Chapman and colleagues (Chapman et al., 2005) found evidence of inaccurate information—or at least information that may reinforce misperceptions held by the public about breast cancer—in the reporting of Kyle Minogue's breast cancer disclosure. Some anecdotal evidence supports this finding, but no other research has explored this possibility.

RQ1: Is misperception reinforcement information present in the media reports of celebrity breast cancer disclosures?

Noar and colleagues (2014) suggest exploration of what factors might moderate the association between celebrity cancer disclosures and media coverage. Informed by the discussion on conceptualizing celebrity and level of celebrity status (e.g., Thrall et al., 2008), the potential for audience identification to influence outcomes (e.g., Brown & Basil, 1995), and journalistic norms and news values (de Leon, 2002; Hallin & Briggs, 2014), the following research questions are designed to explore several potential moderators of media outcomes:

RQ2: Are there differences in the volume and content of news coverage of celebrity breast cancer disclosures by a) age of the celebrity; b) career type of the celebrity[1]; c) breast cancer-related event type; and d) level of celebrity status?

[1] In the literature review, several demographic attributes were mentioned as potential moderating variables which may explain or predict media and public outcomes. I chose to only focus on age and career type of the celebrity, and not ethnicity/race and sex because preliminary searches for celebrities suggested that the celebrity-of-interest sample would be predominantly and overwhelmingly Caucasian and female. The final set of celebrities-of-interest confirms these initial observations.

Previous research has found that celebrity cancer disclosures have, on average, impacted increases in online search queries for the respective cancer disclosed by the celebrity (Ayers et al., Noar et al., 2013, Noar et al., 2016). None of these studies cited or tested a particular theoretical framework, but it could be argued that in terms of health communication concepts, there is an underlying assumption of the public's exposure to news coverage of the disclosure (i.e., scanning), which then leads to online information seeking. These descriptive studies do mirror the methodology in studies such as Weeks and Southwell's (2010) where there is a strong correlation between same-day newspaper and television news coverage of the "President Obama is secretly a Muslim" rumor and Google searches for the rumor. Weeks and Southwell (2012) cite agenda-setting as the framework for such a relationship. For this dissertation, it is expected that the overall group of celebrity disclosures under study will have similar impacts on breast cancer-related information seeking in the aggregate. In addition, it is expected that the greater the media salience (i.e., volume of coverage) the greater the public salience (i.e., increased search queries).

H2: Relative search volume for breast cancer-related search queries will have significant increases proximal to media coverage of the celebrity breast cancer disclosure.

H2a: Weeks (and months) with greater volume of media reports on celebrity breast cancer disclosures will have greater increases in relative search volume for breast cancer-related search queries than weeks (and months) with lower numbers of media reports.

The sheer volume of coverage may affect breast cancer information seeking behaviors, but mere exposure may not explain all breast cancer-related information seeking behaviors. As information theorists posit, ambiguity is likely to affect information seeking behaviors (Atkin,1973; Ball-Rokeach & DeFleur, 1976; Knobloch et al., 2003). As is conceptualized in this dissertation, media reports with an episodic frame do not include breast cancer information beyond the celebrity's own experience, and therefore more likely to contain ambiguous information (Hurley et al., 2011). Thus, news coverage with episodic frames are more likely to cause public uncertainty (Maurer & Holbach, 2015) and prompt online breast cancer-related information seeking than news coverage with a thematic frame (e.g., incidence rate information, risk information).

H3: Weeks (and months) with a greater proportion of media coverage of celebrity breast cancer disclosures with episodic frames will have greater relative search volume for breast cancer-related search queries than weeks (and months) with greater proportions of coverage with thematic frames.

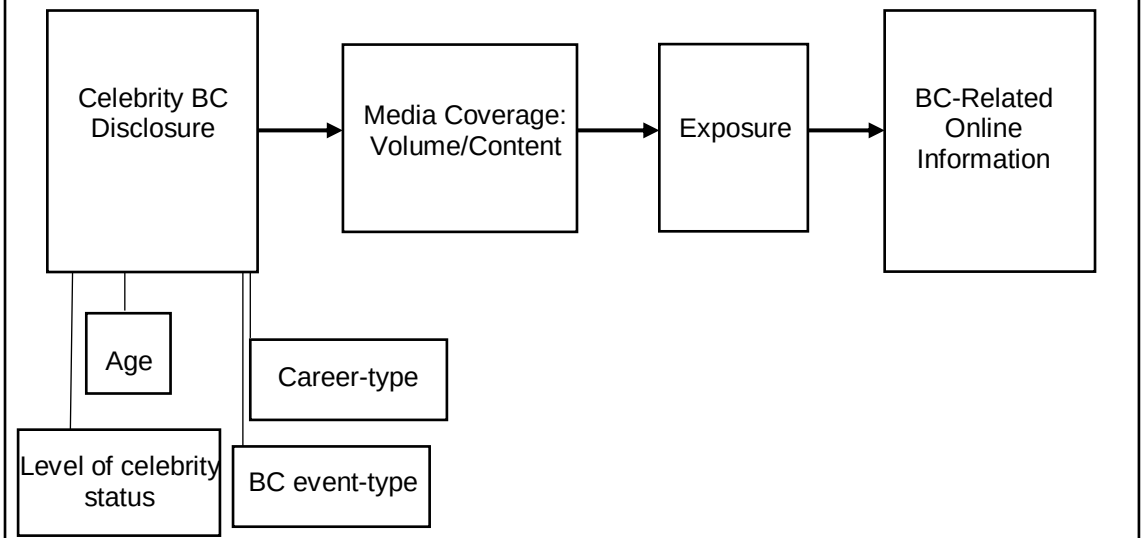
As with the media outcomes, the potential for differential breast cancer information seeking outcomes based on attributes of the celebrity and of the cancer-type will be explored:

RQ3: Are there differences in volume of breast cancer information seeking by a) age of the celebrity; b) career type of the celebrity; c) breast cancer-related event type; and d) level of celebrity status?

Conceptual Model of Effects

Figure 1 presents the proposed conceptual model of effects, giving a visual representation of the study's hypotheses and research questions. A celebrity breast cancer

Figure 1. Simplified Conceptual Model of Effects



disclosure is the starting point of the model of effects—for media coverage to occur and to observe subsequent breast cancer-related information seeking effects, a celebrity must disclose an experience with breast cancer. Once a celebrity publicly announces her or his breast cancer experience, then media coverage can/will occur. There will likely be differences in these media outcomes: 1) volume of coverage; 2) frequency of episodic and thematic frames; and 3) information that reinforces misperceptions of breast cancer held by the public. From there, this natural coverage effects project assumes exposure via the media outcomes (i.e., volume and framing). It is expected that celebrity breast cancer disclosures, on average, will increase relative search volume for breast cancer-related search queries. However, disclosures with higher rates of coverage are expected to further

increase these outcomes. Ambiguous content (i.e., episodic frames) is also expected to have a greater impact on breast cancer-related information seeking outcomes than thematic frames. In the model of effects, four potential moderators (i.e., celebrity and disease related attributes) are located under the celebrity breast cancer disclosure. These could potentially impact media and information seeking outcomes.

Overview of Research Studies

As stated in the Introduction, this dissertation asks three overarching research questions: 1) What breast cancer-related messages are conveyed to the public in media reports of celebrity breast cancer disclosures; 2) Do these messages impact public cancer-related behavioral outcomes; and 3) Are there attributes of the celebrity and/or cancer-related event that predict media and public outcomes? Two studies were conducted to answer these overarching research questions. Both studies are longitudinal analyses over a 12-year period. The rationale for the chosen time period is outlined in Chapter 4, but the decision to focus on a substantial period of time was made in order to capture a large number of celebrity breast cancer disclosures. This is a necessary condition to understand the overarching nature of news coverage and impact on subsequent public outcomes. Study 1 is a content analysis of news coverage of celebrity breast cancer disclosures and answers the first overarching research question. Study 2 focuses on answering the second question. This is done by using results (i.e., volume of coverage and volume of media frames) from Study 1 as predictor variables and aggregate breast cancer-related online search query data as outcomes variables. Finally, the third overarching research question is answered with a subanalysis in both Study 1 and Study 2. Attributes of the celebrity and her or his breast cancer-related event type are used to determine if differential media

and information seeking outcomes occur in a predictable manner based on the moderators.

There is little empirical evidence on the content of celebrity breast cancer disclosures (see Chapter 2). Study 1 seeks to begin to fill this gap in the extant literature. Study 1 has two goals: 1) to estimate the prevalence of breast cancer-related information present (i.e., thematic frames) in media reports of celebrity breast cancer disclosures; and 2) to estimate the prevalence of information that may reinforce misperceptions about breast cancer held by the public. To address the first goal, a framing analysis (i.e., episodic or thematic) was done on articles of celebrity breast cancer disclosures produced by the *Associated Press*, *The New York Times*, and *The Washington Post* over a 12-year period. The three content providers are considered to be a good proxy for the national news environment (Stryker, 2008). Using the same sample of content, the second goal was addressed by identifying specific statements which oversimplify breast cancer or are inaccurate, and have the potential to reinforce misperceptions of breast cancer held by the general public. Such statements include: "1 in 8 women will be diagnosed with breast cancer in their lifetime"; "all women should begin breast cancer screening at age 40"; and "mastectomy is the best therapy." Study 1 is detailed in Chapter 5.

The nature of news coverage on celebrity breast cancer disclosures (i.e., volume and framing) allows for estimating the potential for exposure (from a natural coverage effects perspective), and therefore, can be used to predict online breast cancer-related information seeking outcomes. Study 2 uses this framework. Study 2 has two goals: 1) to verify face and content validity of numerous breast cancer-related search term dimensions; and 2) to test the hypothesized relationships between news media predictor

variables and the aggregate online information seeking outcomes. All aggregate online information seeking/search query data was obtained through Google Trends (trends.google.com/trends/explore). Several steps were taken to confirm face and content validity and to construct theoretically informed search query dimensions. Time series analysis was used to test the hypothesized relationships. Study 2 is detailed in Chapter 6.

Finally, to answer the third overarching research question and to answer several proposed research questions, a subanalysis was performed on the results of both Study 1 and Study 2. The goal of the subanalyses is straightforward: to build a typology of celebrity/cancer attributes that can predict and explain media and public breast cancer-related outcomes. The four potential moderating variables—the celebrity's age, career-type, breast cancer event-type, and level of celebrity status—were constructed using information found in coverage of the celebrity's disclosure (age, career-type, breast cancer event-type), and media and audience salience data (i.e., media coverage counts and Google Trends search data). Chapter 4 focuses on how specific celebrity breast disclosures were selected for study and the construction of the potential moderating variables. The results from each test are in the respective study (Chapters 5 and 6).

Chapter 4

Time Period of Study, Sources of Data, and Celebrity Measures

The hypotheses and research questions outlined in Chapter 3 were tested through a longitudinal analysis of media content and online aggregate search data. Chapters 5 and 6 offer the methods and results of Studies 1 and 2, respectively. This chapter discusses elements central to both studies, including: 1) the timeline of study; 2) sources of data; 3) the operational definition and methods used to determine celebrities-of-interest; and 4) the operational definitions and measures for the four attribute variables (i.e., age, career-type, breast cancer event-type, and level of celebrity status) fundamental to the posited research questions.

Time Period of Study and Sources of Data

Google Trends (Google, n.d.) was used to retrieve all online search query data for the dissertation (i.e., audience salience, breast cancer information seeking). Google Trends was chosen as the source of search query data because: 1) Google has been the most popular online search engine for nearly a decade (Alexa, n.d.); 2) Google Trends data are publicly available; and 3) are widely used in academic research (as discussed in the literature review), indicating validity. Google Trends data are a random sample of Google search queries which Google categorizes, organizes by topic, and then removes personal information. The data provided by Google are the relative search volume (RSV) of the search term or phrase queried. According to Google (n.d.) the RSV of "each data point is divided by the total searches of the geography and time range it represents." The final RSV data point is scaled from 0 to 100. This proportion is relative to all searches for the given location and time period.

The dates (2004-current date) for which these data are available influence the timeline of the dissertation. To construct audience salience measures (see pages 90-95), some volume of data is needed prior to any celebrity breast cancer disclosures under investigation. With this in mind, one year of the available data was reserved. Therefore, the longitudinal analysis of media reports of celebrity breast cancer disclosures and breast cancer-related information seeking takes place from 2005 through 2016 (the last full calendar year available).

Sampling Frame

The media landscape underwent significant changes during the dates under study (i.e., 2005-2016). Greater amounts of news content from traditional news organizations became available online, news consumption shifted away from newspapers and television overall and at greater paces for certain demographics, media organization ownership has homogenized, and a glut of independent online news organizations have come and gone (e.g., Perse & Lambe, 2016; Shoemaker & Reese, 2013). The changes in news ownership, production, availability, and consumption poses challenges for researchers interested in capturing the national news media environment. Specific to this dissertation, consistency in coverage and national prominence and consumption during this time of flux was a priority, as well as being able to systematically retrieve the content.

With these priorities in mind, all content used for this dissertation was retrieved from two national newspapers (*The New York Times*, *The Washington Post*) and the *Associated Press* wire service. The three content providers are used as a proxy for national news coverage of celebrity breast cancer disclosures. These content providers were chosen because 1) they are among the highest circulation national newspapers

throughout the time period of study; 2) their status as agenda-setters is well documented (Golan, 2006; Kiousis & McCombs, 2004); and 3) and previous work has found that the combination of these content providers is the best indicator of the national news environment, rather than simply relying on one publication (e.g., *The Associated Press*) or including other national newspapers (e.g., *USA Today, The Los Angeles Times*) (Stryker, 2008). The LexisNexis database was used to retrieve all media content.

Celebrities-of-Interest

A celebrity-of-interest must have either taken preventive measures for breast cancer, been diagnosed with breast cancer, treated for breast cancer (e.g., surgery and/or chemotherapy/radiation), or have died from complications related to breast cancer. The celebrity either personally disclosed to the media her or his breast cancer-related event, or the media reported on the event without explicit permission or confirmation from the celebrity. Reporting of the disclosure, however, is not the sole criterion. Previous discussions of *celebrity* have included the concepts of knownness, visibility, media attention, and audience attention (See Chapter 2 for discussion). These concepts imply a degree of salience from both the media and public.

Media or public attention to an individual's breast cancer disclosure alone does not necessarily fit the parameters of the *celebrity* construct. For the purposes of this study, a celebrity should have a degree of knownness prior to her or his public announcement. In Lerner's (2006) book "When Illness Goes Public," he offers a historical analysis of 12 cases of persons, whom he calls "celebrities," who fought their given illness in the public eye. However, several of the celebrities only became well-known because of the unusual or uncommon characteristic of their illness. It is important

to recognize that such cases are not uncommon in today's media landscape and is precisely why Jensen et al. (2010) coded for famous or not famous people featured in stories about cancer. To truly expand on the celebrity cancer disclosures body of research, both in furthering theory and determining effects, it is necessary to continue to focus on celebrities who are known, to both the press and the public, prior to their illness.

With this in mind three components are necessary to identify a celebrity-of-interest for this project: 1) a breast cancer disclosure during the time under study; 2) evidence of media salience prior to the disclosure; and 3) audience salience prior to the disclosure. The following describes how *celebrities-of-interest* for the dissertation were determined.

Label of celebrity-of-interest (breast cancer disclosure). Celebrities-of-interest were first identified through an iterative process of multiple Boolean search strings in the Google search engine. Online lists of celebrities who have experienced a breast cancer-related event (e.g., preventive measures, diagnosis, death) from 2005-2016 2 (e.g., "How 8 celebrities bravely battled breast cancer") were retrieved. The Google search engine was used at this stage because of the large volume and diversity of content available. "Celebrities" and its synonyms (e.g., stars, famous people) were combined with "breast cancer" for one type of search string (e.g., "celebrities" AND "breast cancer) and some iterations of this string also included AND "list". Using the same structure (i.e., "celebrities" and its synonyms) with other breast cancer-related words ("breast cancer deaths"; "BRCA"; "breast cancer gene"; "mastectomy"; "breast cancer surgery") were also used. Again, an iteration of these search strings with AND "list" were used for retrieval. Once repetition of names was observed from several lists, I considered

saturation reached, and the initial step in celebrity-of-interest identification was concluded. This search is not necessarily exhaustive, but the goal was to identify as many celebrities as possible. The celebrity's name, date of disclosure, age at diagnosis and disclosure, breast cancer-related event (e.g., diagnosis, surgery, treatment, death), career type, and the specific media organization(s) which published the list was recorded. Using these methods, a total of 112 celebrities were identified as having disclosed a breast cancer-related event from 2005–2016. Some celebrities had multiple disclosures—127 separate disclosure events were initially identified.

The above steps established that a celebrity breast cancer disclosure occurred. In order to verify that the individual fits the operational criteria of celebrity-of-interest, an indication of media and audience salience prior to the disclosure were confirmed.

A note on celebrities, media reports, and relevancy. Before discussing the process of determining media salience for the potential celebrities-of-interest, this section addresses relevancy. Much of this dissertation involves counts of celebrities' names in media reports or analyzing media content of celebrity breast cancer disclosures. Establishing article relevancy is necessary because simply entering a given celebrity's name in a database does not guarantee the retrieved articles are relevant (i.e., about the celebrity). For example, a search in LexisNexis for "Robin Roberts," the ABC television news anchor and identified preliminary celebrity-of-interest, retrieves articles for both her and articles for *Robin Roberts*, a starting pitcher for the Philadelphia Phillies from the late 1940s through the late 1960s, and even marriage announcements for an average citizen sharing the name of the news anchor and baseball star.

To ensure reliability for the decisions of relevancy, formal definitions for each preliminary celebrity-of-interest were written to guide determining article relevancy. The relevancy definitions include the celebrity's date of birth (if the exact date of birth was not available, a general year was given or the celebrity's approximate age), her or his date of death, if applicable, why the celebrity is a considered a notable person (e.g., singer-song writer or American actress), and specific examples of her or his work. This information was obtained from each celebrity's Wikipedia page. The following is an example of a relevancy definition for a preliminary celebrity-of-interest:

> **Berzon, Betty:** (January 18, 1928-January 24, 2006). American author and psychotherapist known for her work with gay and lesbian communities. Books included: *Positively Gay*; *Permanent Partners*; *The Intimacy Dance;* and her personal memoir – *Surviving Madness: A Therapist's Own Story*.

The celebrity-of-interest definitions were tested for reliability (Stryker, Wray, Hornik, & Yanovitzky, 2006). The sample for testing the reliability of the relevancy definitions was drawn from all newspaper coverage archived in LexisNexis. For each preliminary celebrity-of-interest, three randomly selected articles were selected (N = 332).

An intercoder reliability test was conducted between the author and an independent coder. Coder training took place for approximately one hour. Following the training, the sample was double-coded. The sole variable coded for was *Relevanccy*—according to the relevancy definition of the celebrity-of-interest, is the news article relevant or not (1 = relevant; 2 = not relevant)? Using the relevancy definitions, intercoder agreement was 97.9% and a high level of reliability was reached

(Krippendorf's α = .937). The full coding protocol—including the relevancy definitions for each preliminary celebrity-of-interest—for determining relevancy of mentions of celebrity is available in Appendix D on pages 222–236.

Label of celebrity-of-interest (media salience). Having a media presence before disclosure is a necessary condition to be selected as a celebrity-of-interest for this study. To confirm this status, all persons-of-interest identified in the celebrity breast cancer search lists were entered into the LexisNexis database with the date search parameters set for any time before the disclosure. If there were any media reports about the individual before her or his disclosure, the individual was coded as a celebrity (0 = not a celebrity; 1 = celebrity). Of the 112 celebrities initially identified, there were no archived news reports for six individuals (Char Fontane; Amber Marchese; Karen Mayo-Chandler; Screechy Peach; Hollie Stevens; and Eleanor Dapkus Wolf).

Label of celebrity-of-interest (audience salience). The conceptual and operational definitions of *celebrity* recognize the audience's role in the creation and maintenance of the celebrity. To measure a degree of audience salience over the time period for this study, Google Trends search data for the United States was used. Because Google Trends data are scaled, it is unknown as to how high the volume must be for data to be retrieved, but Google does state that there must be significant search traffic for the search term to even have an RSV of 1 (Google, n. d.). For example, I entered my name ("Susan LoRusso") and then my adviser's ("Rebekah Nagler") into Google Trends. For each of us the retrieved results were: "Not enough volume to show graphs," indicating that neither one of us is salient with the public.

The RSV was retrieved for any time prior to an individual's disclosure through the week before the disclosure. An average RSV for the given time period was calculated for each potential celebrity-of-interest. Given that an RSV of 1 indicates enough search traffic to be considered "significant" by Google, any potential celebrity-of-interest with an RSV of 1 was coded as a celebrity (0 = not a celebrity; 1 = celebrity). Of the 112 celebrities initially identified, the average RSV did not reach a level of 1 for the following 12 individuals: Barbara Allen; Chris Calloway; Eleanor Dapkus Wolf; Rosalie Gaull Silberman; Angela King; Joy Langan; Yoko Sano; Soraya; Heather Stilwell; Lindalee Tracey; Karen Wynn Fonstad; Laura Ziskin.

Final measure. The final label of celebrity-of-interest was applied to all individuals whose name was included in the celebrity breast cancer media lists *and* who received codes of 1 for the two salience measures (N = 95). Of these 95 celebrities-of-interest, 110 separate disclosures were identified. A table with all celebrities-of-interest can be found in Appendix A (pps. 201-204).

Celebrity Attribute Variables

Research questions RQ2 and RQ3 seek to determine if particular attributes of the celebrity and/or her or his breast cancer-related event impact media and online breast cancer-related information seeking outcomes. Such inquiry was prompted by Noar et al.'s (2014) review and research agenda on public figure cancer announcements, where they recommend investigating the attributes of celebrities whose disclosures result in significant (or less than significant) effects. These potential moderators (age, career type, cancer-related event type, and level of celebrity status) were constructed with data obtained through media coverage of the celebrity breast cancer disclosure and Google

Trends search query data. This section offers the operational definition, describes the method used to construct the variable, and offers the final measure.

Age. Age refers to the age at which the celebrity experienced her or his breast cancer-related event. The age of each celebrity was first determined in the initial reading of celebrity breast cancer media lists. All ages were confirmed through the celebrity's Wikipedia page. Appendix A (pps. 201-204) includes the age of each celebrity-of-interest at the time of her or his breast cancer-related event.

Age was constructed as a categorical variable, as were all attribute variables. Determining how to categorize celebrities by age was given significant consideration. In respect to media-related outcomes (i.e., volume and content), I suspected that "younger" celebrities—those that experience breast cancer prior to the age range that is typically associated with breast cancer—receive greater media attention because the story aligns with the journalistic norm of novelty (Gans, 1979; Hallin & Briggs, 2014). Similarly, audience reaction (i.e., information seeking) may be more likely to occur when a celebrity's age is incongruent with schemas built around the age of people who are most likely to experience breast cancer. The median age of a woman diagnosed with breast cancer is 62 (Susan G. Komen, n.d.); however, it is not known what specific ages the public associates with breast cancer. Screening recommendations offer a baseline of age to investigate. As previously discussed, screening recommendations are provided by numerous agencies and associations. These organizations publicly promote the suggested age recommendations for breast cancer screening for women of average risk. Therefore, it is reasonable to assume that both journalists and the public would see a celebrity under the age recommended for screening as an outlier.

The timeline of study for this dissertation further complicates categorizing celebrities under or over the age recommended for breast cancer screening because of the change in guidelines from the USPSTF in November of 2009. Other agency guidelines (e.g., ACS) did not change the minimum age for screening initiation, creating conflicting guidelines[2]. Because the USPSTF change in screening recommendations received significant media attention, it is possible the newly recommended age of 50 gained some saliency with the public. Therefore, celebrities-of-interest were divided by pre- and post-recommended age for breast cancer screening initiation as recommended during the current year by the USPSTF. From January 1, 2005-December 31, 2009 (the time period prior to, or right at the time of, the release of the updated USPSTF), there were 60 celebrity breast cancer-related disclosures. Only four of those were from women under the age of 40 (Group 1) and 56 were from celebrities over the age of 40 (Group 2). From January 1, 2010 – December 31, 2016 there were 50 celebrity breast cancer disclosures. During this time period, the two age groups are almost equally distributed—26 disclosures from celebrities under 50 (Group 3) and 24 disclosures from celebrities over 50 (Group 4). Because there are only three men in the study's sample of celebrities (and no breast cancer screening guidelines for men), they were excluded from the age group analyses. See Appendix A for the specific celebrities-of-interest that comprise the four

[2] In October of 2015, the ACS began recommending initiating annual routine screening for women of average risk at age 45, instead of their previous recommendation of age 40 (American Cancer Society, 2015). Because this recommendation change comes towards the end of the time period of study it is assumed that the USPSTF recommendations would have a greater impact on secular trends in breast cancer-related information seeking then the ACS. Now that the change in ACS recommendations nearly two years old, future research should consider other ways of grouping by age than the parameters set forth here.

age group variables (Group 1, *n* = 4); (Group 2, *n* = 56); (Group 3, *n* = 26); (Group 4, *n* = 24).

Career type. Career type refers to the type of career the celebrity has which likely contributed to their knownness. While not a career type, this variable also includes a relationship with another notable person if that is how the celebrity-of-interest is primarily known. Career for the individual celebrity was first determined through initial readings of media coverage of the celebrity and her or his disclosure. For example, articles discussing Rachel Bissex's breast cancer disclosure mention her status as an award-winning American folk singer/songwriter. Bissex was originally recorded as a Singer. Once all celebrity-of-interest's careers were identified, career categories were collapsed into larger categories. Instead of having a category of Singer and a category of Band Member, these two categories were collapsed, categorizing Bissex, for example, as a Musician. This first round of collapsing consisted of 20 categories: musician; athlete; relationship with; academic/scholar; journalist; activist; actor; author; model; service person/military; filmmaker; politician; business leader; felon; artist; chef; fashion designer; comedian; TV personality; and pageant contestant. Some categories resulted in a very small number of celebrities (e.g., model, *n* = 1). The final career type categories serve as eight separate groups of the *career type* variable: actor (*n* = 25); athlete/sports-related (*n* = 6); academic/author/activist/creative (*n* = 18); journalist/news anchor (*n* = 15); musician (*n* = 15); personal affiliation (e.g., spouse of) (*n* = 7); politician/policy maker/service person (*n* = 9); and television personality (*n* = 15). See Appendix A for the specific celebrities-of-interest that comprise the eight categories of the *career type* variable.

Breast cancer-related event. The breast cancer-related event measures the type of breast cancer-related experience the celebrity disclosed. All breast cancer-related event types were categorized according to the National Cancer Institute's (NCI) cancer control continuum (National Cancer Institute, n.d.). The continuum is recognized as an organizing tool which indicates specific phases of the disease. For the purposes of this study, it also streamlines the almost innumerable possibilities for breast cancer diagnosis and breast cancer treatment, among others.

Many of the disclosures included more than one breast cancer-related event. For example, Tig Notaro disclosed that she had been diagnosed with breast cancer and subsequently underwent a mastectomy. Because most disclosures (other than preventive [e.g., BRCA testing]) include a diagnosis of breast cancer and then discuss treatment, as in the case of Notaro, selecting "diagnosis" was considered redundant. Therefore, all disclosures were treated as mutually exclusive. The continuum categories and an added category of *death* created six breast cancer-related event categories: prevention (n = 7); detection (n = 0); diagnosis (n = 4); treatment (n = 44); survivorship (n = 1); and death (n = 54*).* See Appendix A for the specific celebrities-of-interest that comprise the 6 categories of the *breast cancer-related event* variable.

Level of celebrity status. The level of celebrity status is the degree of fame the celebrity has reached at the time of her or his breast cancer disclosure. Because celebrity is a dynamic, reciprocal process which includes the media and the public, the quantified level of celebrity status includes measures of media and public salience from before the celebrity's disclosure and at the time of disclosure. Very little previous research has attempted to quantify hierarchy of fame, and those that have only used media indicators

(see Thrall et al., 2008; van de Rijt, 2013). Acknowledging and incorporating the audience into measures of a celebrity's level of fame allows this dissertation to move away from mediacentrism (Driessens, 2015). Including measures of media and audience salience over time, even imperfect measures, is a needed step forward in the empirical analysis of celebrity. As Driessens (2015) states: "[A]ccounting for both basic dimensions [media and audience] through the notion of memory might help us to gradually refine celebrity as a sensitizing concept—not by cleaning up the mess, but by dealing with it in better and more transparent ways" (p. 372).

For this dissertation the *level of celebrity status* index uses two measures for both media salience, as measured by Thrall and colleagues (2008), and audience salience, as suggested by Driessens (2015). For both media and audience salience, one measure reflects salience at the time of disclosure (i.e., current salience) and the other measure reflects the celebrity's larger social relevance (i.e., longer-term salience). The following offers the methods used to construct the four individual measures that comprise the final nine-point *level of celebrity status* index.

Media salience. Measuring media salience of a celebrity involves two time periods, *current salience* and *longer-term salience* (Thrall et al., 2008). For both measures, a media audit was done to establish the volume of media coverage relevant to each construct. To construct each measure, different approaches in data retrieval were taken. As previously discussed, entering a given celebrity-of-interest's name in a database such as *LexisNeixs* may not always retrieve articles relevant to the celebrity. Although further refinement of the name (i.e., search string) could eliminate relevant articles. Either case can affect the validity of the results (Riffe, Lacy, & Fico, 2014).

Stryker and colleagues (Stryker et al., 2006) set up calculation methods for estimating proportions of recall (ability to accurately retrieve desired items) and precision (ability to avoid extraneous ones) metrics for judging the quality of search strings (p. 413). However, the decision to prioritize recall over precision, for example, is largely dependent on the method of coding—human coding or computer-assisted.

This dissertation, including the current and longer-term media salience measures, uses both methods, although not concurrently. Therefore, separate decisions regarding the prioritization of recall and precision were made for each measure. The following sections operationalize each media salience construct, describe in further detail the search string procedure used to retrieve the necessary data, and offer the specifics of the index.

Current media salience. Current media salience (CMS) demonstrates the immediate magnitude of media attention to the celebrity-of-interest's breast cancer-related disclosure. As suggested by previous research, current media salience was measured by the number of articles published during the initial day of the disclosure through four calendar weeks following the disclosure (date parameters decision discussed in Chapter 5, p. 99) (Thrall et al., 2008).

Recall was prioritized for the current media salience measure because all articles retrieved would be human coded. The initial article relevancy test (see pps. 75-77) was not sufficient to test recall and precision, but the results clearly indicated that some celebrity-of-interest's names would be more likely to yield relevant articles than other celebrity-of-interest's names. For example, all articles coded for Christina Applegate, Kathy Bates, Judy Blume, Sheryl Crow, Shannen Doherty, Elizabeth Edwards, among others were considered relevant by both coders. Conversely, none of the articles

randomly selected for coding for Barbara Allen, Allison Chapman, Yvonne Carter, Jennifer Lyon, Heather Pick, Jean Shubert, Hollie Stevens, and Pat Stevens were considered relevant. The article relevancy rate for other celebrities fell somewhere in between.

With such results, it was clear that further steps would need to be taken in search string development. Current media salience is bound by strict date parameters (the date of the disclosure + four calendar weeks). The specific date range was believed to curtail some of the issues with data retrieval inherent to the open date range used in measuring longer-term media salience. To test this theory and to determine whether to use an "open" search string—"a search term that is developed to capture any and all relevant stories relevant to the topic of interest" (Stryker et al., 2006, p. 416)—or a "closed" search string, a relevancy test was conducted from a randomly selected subsample (16.3% [*n* = 15] of the total sample) of celebrities-of-interest.

An "open" search string is typically quite long because it includes all synonyms for the topic of interest. However, this procedure is not applicable for search strings which retrieve proper nouns, such as a celebrity's name. Names, unlike a broad topic such as cancer, have no synonyms and are inherently more precise than a topic. I am unaware of research testing the recall and precision of individuals' names, but Cavanah (2016) builds search strings based on the names of communities. To do this, the name of the community was used as the "open" search string (e.g., Alsea, OR; Chenoweth, OR). For many of the communities the "open" search string met criteria for recall and precision. However, the "open" search string for some communities retrieved high proportions of irrelevant articles. For many of the communities, more complex search

strings were built. For this preliminary test, Cavanah's (2016) procedures were followed. Using the "open" search string for 15 randomly selected celebrities-of-interest, all articles retrieved (n = 231) were relevant to the celebrity, but only 63.2% (n = 146) were about the celebrity's breast cancer-related event. Using a more "closed" search string (celebrity name AND breast cancer) all articles retrieved (n = 104) were relevant to the celebrity of interest and included information about the celebrity's breast cancer-related event. While the "closed" search string demonstrated perfect precision, 42 relevant articles were eliminated with this search string. Therefore, higher recall was prioritized. Because these units of analysis would be hand-coded, it was determined that the "open" search string would provide greater reliability for the final current media salience measure.

To retrieve media reports from the content providers used for this study, all 95 celebrity-of-interest names (i.e., open search string) were queried in LexisNexis. The date parameters were for the date of the disclosure through four weeks post-disclosure. All articles retrieved (N = 962) were coded by three independent coders (coding procedures discussed in detail in Chapter 5, pps. 103– 105). The coders were asked to determine the article's relevancy (using criteria regarding the celebrity-of-interest and of a reported on breast cancer-related event [full code/decision criteria is available in the coding protocol in Appendix E, pps. 237– 248]). The total number of relevant articles for each celebrity-of-interest (range = 0 – 117; $M = 5.42$, sd = 14.24) comprise the current media salience measure. Appendix B on pages 205-210 gives the absolute value of this measure for each celebrity-of-interest.

Longer-term media salience. Longer-term media salience (LTMS) was measured by the volume of media coverage given to the celebrity up until the date of the breast

cancer disclosure. The three content providers (i.e. *The Associated Press, The New York Times*, and *The Washington Post*) were also used as a proxy for the greater media environment.

For this measure, computer-assisted coding was used to determined volume of media coverage. Therefore, it was necessary to construct valid search strings (Riffe et al., 2014; Stryker et al., 2006). To estimate recall and precision a three-stage process is required, with the first stage focusing on: 1) defining the relevant universe; 2) defining story relevance; and 3) specifying the satisfactory levels of recall and precision (Stryker et al., 2006, p. 415). Thus far, the first two steps in this stage have been addressed. For the longer-term media salience measure, precision was prioritized because no human coding will occur with the retrieved data. As recommended by Stryker et al. (2006), the desired level of recall and precision, with a .05 confidence interval, was set at .70.

The second stage of this process requires developing and refining search strings with "a random sub-sample drawn from the universe of texts" (Stryker et al., 2006, p. 416). This was a complex process for this dissertation. While many studies are about a singular topic (e.g., mammograms or breast cancer), the stories to be retrieved for the media salience measure are about 95 separate individuals. For the preliminary recall and precision test used for the CMS measure, some celebrity names had perfect precision with the "open" search string, while other celebrity names did not draw any related stories when queried in LexisNexis.

The non-relevant texts retrieved in the preliminary tests revealed three common issues with the greater body of texts:

1. Common name/average person: Common name or average person issues occurred when the celebrity-of-interest's name was retrieved in full, but the article was about someone other than the celebrity. For example:

> The engagement of **Robin Roberts** to Andrew Mercogliano has been announced by Mr. and Mrs. George Roberts of Old Bridge, N.J., parents of the future bride.
>
> Her fiancé is the son of Mr. and Mrs. John Mercogliano of Asbury Park, N.J.

This article is clearly an engagement announcement and not about Robin Roberts, the news anchor and celebrity-of-interest for this study.

2. Places or things: Some names retrieved articles where the mentions of the names were about places or things. For example:

> Headline: Colours' creator used 'summer' palette in room. **Rose and gray** themes color room where she met clients.

In this case, the "rose" and "gray" in the headline are referring to the colors and not to the chef and celebrity-of-interest, Rose Gray.

3. Other notable persons: The final common issue when using celebrity-of-interests' names to retrieve relevant articles is when the celebrity's name is retrieved in full, but the name is referencing another notable or famous person. For example:

> At 65, looking fit enough to begin both ends of a doubleheader, righthander **Robin Roberts** has been named to the board of directors of the national baseball Hall of Fame. Already in the Hall himself as a pitcher, Robby probably also would merit Cooperstown because he was the man who picked Marvin Miller as the ballplayers' representation.

"Robin Roberts" is the MLB pitcher Robin Roberts and not the news anchor and celebrity-of-interest, Robin Roberts.

With these issues in mind, the most straightforward path was to test each celebrity's relevancy with the "open" search string (the celebrity-of-interest's name), and when relevancy was low, develop more refined search strings. While this was a somewhat informal process, 20% of articles retrieved were tested using this method. Of the 95 celebrities under study, only 11 did not meet perfect relevancy. In most cases, simply adding the celebrity's career (e.g., "Pat Steven AND actress") or her work or personal affiliation (e.g., "Rose Gray AND River Café"; "Elizabeth Edwards AND John") resulted in higher rates of relevancy.

Testing this method indicated that this approach was hitting acceptable precision levels. The third stage recommended by Stryker et al. (2006) is the formal test—the "best" search string or equation is confirmed by using a new random sample (p. 418). Because of the specificity needed for this project (names vs. topics) drawing a large random sample based on a number from all of the celebrities-of-interest would not necessarily draw a representative sample for each celebrity/search string. Following procedures suggested by Cavanah (2016), a random year was pulled for each celebrity. All but two of the search string had a score of 1 (Linda Clark and Pat Stevens exceeded .85 for both tests). The total number of articles retrieved from LexisNexis with the validated search string are the final measure of a celebrity-of-interest's long-term media salience (range = 1 – 11,744; M = 498.99, sd = 1430.11). Appendix B on pages 205-210 gives the absolute value of this measure for each celebrity-of-interest.

Audience salience. The two studies which measured stratification and hierarchy of fame (see Thrall et al., 2008; van de Rijt et al., 2013) solely relied on media indicators of fame. Given this study's conceptual definitions of *celebrity* and *level of celebrity status*, using a similar approach would be shortsighted because of the acknowledged role the public plays in creating and maintaining celebrity. This stance is substantiated in Driessen's (2015) commentary, "On the Epistemology and Operationalisation of Celebrity." Here, he criticizes the previously cited studies' singular use of media indicators, and instead recommends combining article frequency counts with indicators of public awareness and response. As already specified, this dissertation follows this recommendation in quantifying level of celebrity status. Because this study uses a retrospective longitudinal design, obtaining archival data relevant to the audiences' awareness and response to the list of celebrities-of-interest is difficult—no public opinion data exists. However, online search query data are indicators of both awareness (i.e., salience) and response (i.e., information seeking) and are available for the time period under study. Therefore, Google Trends search query data were used to determine audience salience. The celebrity's name (e.g. "Angelina Jolie") was entered into Google Trends—only search volume for the U.S. was considered. The Google RSV score proximal to the celebrity's announcement represents *current audience salience*, while the celebrity's RSV mean score prior to the disclosure represents *longer-term audience salience*.

Current audience salience. As noted above, the current audience salience (CAS) measure is the RSV score proximal to the disclosure. Depending on the duration in which data are retrieved, Google offers different time periods for the RSV data-points. That is,

if data are retrieved for a time period of 90 days or less, RSV data points are offered by calendar days, whereas data retrieved for up to a five-year period are offered in weekly RSV data points. For any time period exceeding five years, RSV data points are provided as monthly scores. As discussed previously, an RSV score is a weighted value for the location and time-period for which the data are retrieved. This weighting results in different RSV scores for the same dates when the overall time periods investigated differ.

The data-point needed for this measure is the RSV score for the week or month of the disclosure. Two time periods were chosen to explore and compare RSV scores and their potential variation—January 1, 2004–December 31, 2016 and January 1, 2004 through the month after the celebrity's disclosure[3]. For many celebrities, if not most, the date of her or his breast cancer disclosure received the highest RSV score or data point during the two time periods explored in Google Trends. However, for a few celebrities-of-interest who have had longer careers (which is typically correlated with several high RSV data points) retrieving data from 2004-2016 suppressed the RSV at the time of disclosure. For example, Angelina Jolie's RSV at the time of disclosure during the 2004-2016 time period was an 88, but for the shorter time-period (2004-one month post-disclosure) Jolie's RSV is 95. The difference in weighting is due to a higher RSV (100) in August of 2016 (divorce announcement), which changed the scale on which the search volume at the time of disclosure was measured. Other examples of subsequent RSVs

[3] Initially, the entire time period of study was used to retrieve data for the CAS measure. I assumed using the same time period for each celebrity would ensure reliability and had not considered that post-disclosure data would affect the data point at the week of the disclosure. Visual inspection of the data indicated that the weighting of these datasets may be problematic, as is discussed in the above discussion of Jolie's data. The month after the disclosure was chosen as an end-date in order to capture the disclosure and all but isolate it. The two datasets were then compared to determine which dataset would be most appropriate for the CAS measure, as is discussed in this section.

impacting the RSV score at the time of disclosure include any of the celebrities who had multiple disclosures—Elizabeth Edwards, Allyn Rose, Anastacia, Samantha Harris, Joan Lunden, Shannen Doherty, and Janice Dickensen. In these cases, the RSV at the time of the first disclosure is impacted by the RSV of the second (or third) disclosure—the second disclosure alters the weighting of the first disclosure, and so on. Given these issues in weighting, the RSVs at the time of disclosure for the 2004–month after disclosure were used for the current media salience measure.

Retrieving search query data for proper names also creates some issues. There is a general search term available and a search term that typically lists the individual's career (a personal relationship may also be listed). Determining validity for Google Trends data, specifically face and content validity, is discussed in detail in Chapter 6, but briefly, Google Trends offers a list of related search terms queried proximal to the particular search term. Mellon (2013) suggests using these lists to determine content validity—are the search query data capturing the actual search queries related to the topic under study? For many of the celebrities, both the general search term and the occupation-related search term did not differ here. However, for celebrity names that are common names or common nouns, the general search term was less reliable in the related content retrieved. For example, the related search queries to the general search term for chef Rose Gray include "rose gold," "black and gray rose," "black and gray rose tattoo," "dorian gray," among others. None of the related search queries appear to be connected to the celebrity-of-interest, Rose Gray. When "Rose Gray – Chef" is queried, related queries include "chef" and "The River Café" (Gray's restaurant), demonstrating greater content validity than the general "Rose Gray" search term. The search queries included in the data

retrieval impact the RSV scores. For example, the RSV for the week of Gray's disclosure using the general search term is 41, but for the "Rose Gray – Chef" search query the RSV rises to 97. Not all RSVs specific to names increase with the categorized data retrieval—several names resulted in a more conservative estimate. For example, the general search query "Linda Day" resulted in a 58 RSV score for the week of disclosure, while the "Linda Day – television director" resulted in an RSV of 13. The fairly low RSV at the time of her disclosure (death) is likely due to the high relative search volume in March of 2004 when she was married for the first time. The general search term is including data not relevant to the celebrity-of-interest, but when the data is isolated to the celebrity the RSV is more likely to reflect the audience salience at the time of her or his disclosure.

Therefore, for the current audience salience (CAS) measure, data was retrieved for 2004 through the week after disclosure using the occupation search term (e.g., Sheryl Crow – singer songwriter). All searches were set to "United States." The RSV score (range = 0 – 100; M = 69.34, sd = 36.10) for the week of the disclosure was recorded and is the final measure for current audience salience for each celebrity-of-interest. Appendix B on pages 205-210 gives the value of this measure for each celebrity-of-interest.

Longer-term audience salience. Longer-term audience salience (LTAS) represents the level of audience salience for the celebrity-of-interest over her or his career prior to the disclosure. As was the case for CAS, the occupation-related search terms were used to retrieve search query data. To create this measure, it was determined *a priori* to average the RSV scores for each celebrity-of-interest for this longer time period. Because LTAS is interested in the degree of audience salience prior to the disclosure, the first attempt at retrieving data focused on the time period of January 1, 2004 through the week prior to

disclosure. These weighted data, however, are not comparable from one person to the next, and the timeline of study highlighted this drawback in the data. Some celebrities-of-interest that may be considered "lesser-known," such as Wendie Jo Sperber (RSV = 32.54), Anna Moffo (RSV = 26.78), JoAnna Lund (RSV = 36.37), and Lorraine Hunt Lieberson (RSV = 34.85), had nearly as high or higher RSV scores than potentially "more known" celebrities-of-interest, such as Melissa Etheridge (RSV = 35.30), Kylie Minogue (RSV = 39.32), and Sheryl Crow (RSV = 27.21). (Media salience measures correlate with these assumptions of lesser and more known.) These abnormal findings are due to the weighted data. Those who do not receive search queries on a regular basis end up with multiple RSV data points with high scores, averaging out to the means mentioned above. Individuals who are queried on a more regular basis have a greater abundance of RSVs with low scores, which then results in similar mean scores to someone who has received less overall search traffic. Figure 2 demonstrates the visual contrast of these two types of scenarios.

The weighted nature of Google Trends data is not conducive to directly comparing RSV scores, but creating a similar scale for all or most of the celebrities-of-interest makes RSV scores more comparable. Manipulating dates was the most appropriate way to do this (i.e., using the week/month after disclosure as the end-date). The CAS measure demonstrated that for most of the celebrities-of-interest, her or his disclosure was the highest (or one of the highest) scored events. Using the CAS data and then averaging the scores prior to the week before the disclosure demonstrated more conservative mean RSV scores, and these scores across celebrities-of-interest had face validity. The LTAS mean RSV scores for Wendie Jo Sperber (RSV = 1.31), Anna

Figure 2. Visual Comparison of Search Query Data for Low and High Name Recognition Celebrities

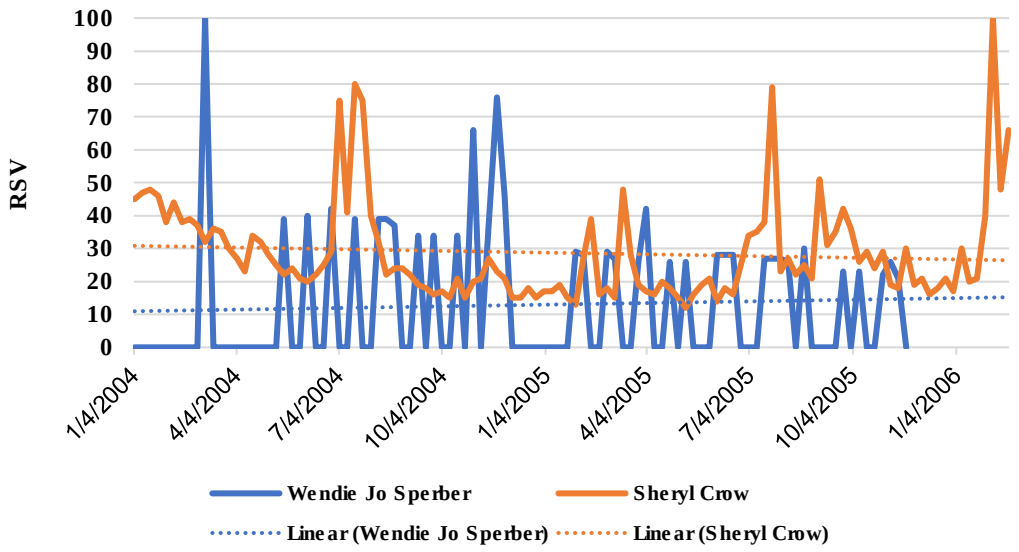


Figure 2. Google Trends search query data plotted from January 1, 2005 through the week before Sperber's and Crow's disclosures. Sperber's mean RSV score is over five points higher than Crow's, yet the visual trends demonstrate that Crow received more consistent search from the public than Sperber.

Moffo (RSV = 6.84), and Lorraine Hunt Lieberson (RSV = 3.67) now stood in stark contrast to the scores of Melissa Etheridge (RSV = 11.61), Kylie Minogue (RSV = 15.43), and Sheryl Crow (RSV = 27.91) (Crow's mean RSV score does not shift due to a high-point RSV of 100 two weeks prior to the disclosure [break-up with Lance Armstrong]). While still not completely comparable scales, the results are more proportionally in line with each other. The final mean scores representing the LTAS for each celebrity-of-interest range from 0 to 30.9 (M = 8.15, sd = 8.36). Figure 3 on page 96 demonstrates the shift in scale for the two time periods tested for this measure. Appendix B on pages 205-210 gives the value of this measure for each celebrity-of-interest.

Figure 3. Examples of Visual Shifts in Search Query Data Points between Pre- and Post-Disclosure Time- Periods

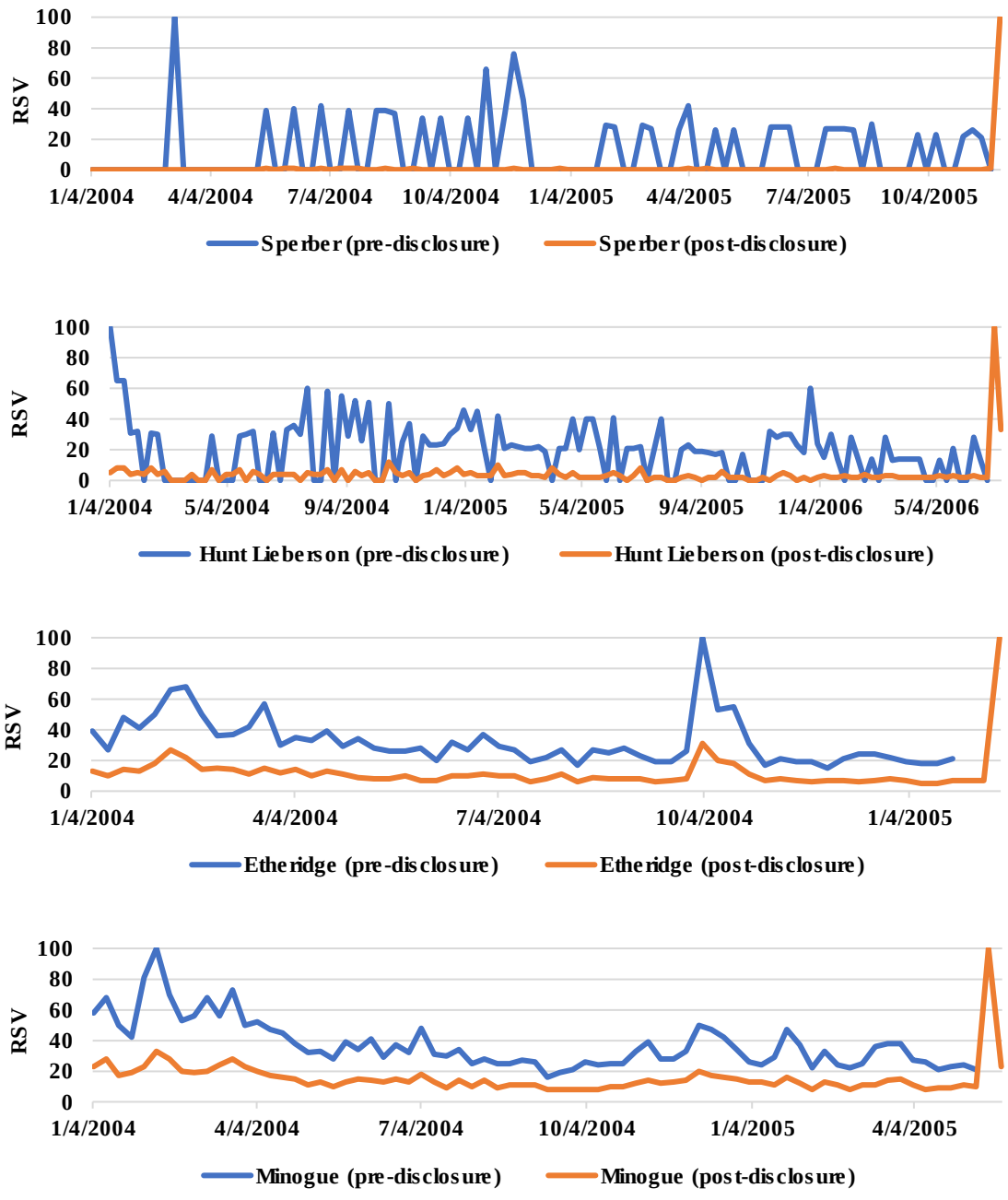


Figure 3. RSV data points plotted for four of the celebrities-of-interest. The blue lines represent the weekly RSV scores retrieved when using the time-period of January 1, 2005 through one week prior to disclosure. The orange lines represent the weekly RSV scores for the January start date through one week past disclosure. The disclosure RSV of 100 lowers the scale of the search query data prior to the disclosure. The data represented by the orange line was used for the final LTAS measure.

Final measure for level of celebrity status index. For each dimension of salience, (i.e., media and audience) the current and long-term measures were multiplied together. Previous research has added current and long-term measures (e.g., Thrall et al., 2008), but because the Google Trends data, are scaled data, magnitude was prioritized over an absolute number. The media salience measures ranged from 0 to 912,249, and the audience salience measures ranged from 0 to 2400.26. The weighted results for both measures by celebrity-of-interest are reported in Appendix B on pages 205-210. Each dimension was then split into quintiles, with celebrities-of-interest in the bottom 20% of media or audience salience scores receiving a "1" and the top 20% receiving a "5". Quintiles as opposed to quartiles, for example, were chosen because of the likelihood for smaller, more homogeneous groups. The resulting quintile numbers were added together across dimensions. The resulting integers serve as the nine points in the level of celebrity status index, with a 2 representing the lowest level of celebrity status and a 10 representing the highest level of celebrity status: 2 (n = 7); 3 (n = 9); 4 (n = 15); 5 (n = 16); 6 (n = 20); 7 (n = 17); 8 (n = 5); 9 (n = 10); 10 (n = 11) . See Appendix A and Appendix B for the specific celebrities-of-interest that comprise the nine levels in the *level of celebrity status index*[4].

[4]Some celebrities-of-interest with multiple disclosures have differing levels of celebrity status based on the specific disclosure. This is to be expected and is in line with the discussion on level of celebrity status in the literature review. How famous someone is not a stable concept, and can vary over time and at different points in time. The LTMS and LTAS measures capture salience prior to the disclosure, which will differ from Disclosure 1 to Disclosure 2, for example. The CMS and CAS measures are not dependent on time, but can differ at each disclosure.

Chapter 5

Study 1. Media Coverage of Celebrity Breast Cancer Disclosures

Study 1 investigates the media outcomes in response to celebrity breast cancer disclosures. Specifically, this study tests the hypothesis that episodic frames will be predominantly employed in media coverage of celebrity breast cancer disclosures (H1), and will also determine if misinformation or misleading information—potentially reinforcing misperceptions about breast cancer held by the public—is present in the news coverage (RQ1). Additionally, this study seeks to establish if the volume and content of the coverage of celebrity breast cancer disclosures differ according to personal attributes of the celebrity (i.e., age, career type, and level of celebrity status) or her/his breast cancer event-type (RQ2).

Sampling Frame

The sampling frame for this study is all media reports from the *The New York Times, Washington Post,* and *The Associated Press* of celebrity breast cancer disclosures from January 2005 through December 2016 (N = 962). (Methods used to select the units of analysis are outlined in Chapter 4).

Independent Variable

As per the conceptual model in Chapter 3, each celebrity breast cancer disclosure event is considered the independent variable.

Celebrity Attributes Variables

Tests were conducted on all media outcomes to determine if differences exist based on *age, career type, breast cancer-related event type,* and *level of celebrity status*

held by the celebrities-of-interest. The specific measures for each celebrity attribute variable are discussed in detail in Chapter 4.

Dependent Variables

Volume of media coverage. Volume of media coverage comprises the weekly counts of coverage for each celebrity-of-interest for four weeks following the disclosure. Previous work has found that celebrity health disclosure event-driven news coverage remains on the news media agenda for a rather short period of time, typically around one week (Kamenova et al., 2014; Noar et al., 2013; Sabel & Dal Cin, 2016). Thus, four weeks of data collection should capture the complete lifespan of the media event.

Determining relevancy of news articles. Using the procedures outlined in Chapter 4 for the *Current Media Salience* measure (pps. 84-89), the open search string (i.e., the celebrity's first and last name) was used in the *LexisNexis* database to retrieve all articles from the three content providers. The coders were asked to determine relevancy. To do this, two criteria must be met: 1) each media report analyzed must be about one of the predetermined celebrities-of-interest; and 2) the media report must contain some information about the celebrity's breast cancer related event (see the full operational criteria for relevancy in the codebook in Appendix E on p. 237).

Content categories. Content categories were developed that represent content present (or not) in the media reports of celebrity breast cancer disclosures. Some of these content categories are for descriptive purposes, and some are central to test hypotheses in Studies 1 and 2 and to answer the research questions for Study 1. All content categories were informed by previous content analytic work and published commentaries (Chapman et al., 2005; Kamenova et al., 2014; Sabel & Dal Cin, 2016) and through reading media

reports of celebrity breast cancer disclosures from media organizations not included in the sampling frame. Content categories are divided into four sections: 1) General Information (V1-V3); 2) Information about the News Article (V4-7); 3) Media Frames (V8); and 4) Statements on Breast Cancer (V9-15). Additionally, for variables V10a through V15a, coders included the article text that fit the misperception reinforcement coding criteria.

Table 1

List of Content Categories

Content Category Sections/Variable	Measure/Purpose	Hypothesis/RQ
General Information		
V1) Numerical code (unit of analysis)	case/unit of analysis	N/A
V2) Celebrity-of-Interest	relevancy	N/A
V3) Relevancy	relevancy	N/A
News Article Information		
V4) Media organization	descriptive	N/A
V5) Date of the article	descriptive/predictive (Study 2)	H2; H3
V6) Section	descriptive	N/A
V7) Word Count	descriptive	N/A
Media Frames		
V8) Media frames	episodic; thematic	H1; RQ2; H3; RQ3
Statements on Breast Cancer		
V9) Breast cancer can happen to anyone	misperception reinforcement (MR)	RQ1; RQ2
V10) Early detection	MR	RQ1; RQ2
V10a) Early detection	MR—qualitative	RQ1
V11) Screening saves lives	MR	RQ1; RQ2
V11a) Screening saves lives	MR—qualitative	RQ
V12) All women should begin screening at 40	MR	RQ1; RQ2
V12a) All women should begin screening at 40	MR—qualitative	RQ1
V13) Breast self-exam	MR	RQ1; RQ2
V13a) Breast self-exam	MR—qualitative	RQ1
V14) “1 in 8”	MR	RQ1; RQ2
V14a) “1 in 8”	MR—qualitative	RQ1
V15) Mastectomy is the best therapy	MR	RQ1; RQ2
V15a) Mastectomy is the best therapy	MR—qualitative	RQ1

Table 1 gives a brief description of each content category, what the category measures, and the hypotheses and research questions with which it corresponds. For full descriptions of each content category, see the coding protocol in Appendix E (pps. 237–248).

Media frames. For this study, the media frame present in a media report of a celebrity disclosure describes the general information present in the story. Either an episodic or a thematic frame is present—all frames were dichotomously coded.

Episodic frame. A media report can be considered written with an *episodic frame* when the primary focus of the article is on the celebrity (e.g., age, career, personal relationships) and her or his breast cancer-related event (e.g., the context of the disclosure and the celebrity's experience with breast cancer). Breast cancer information about the celebrity's own experience can and will likely be included in the article. If there is information that speaks to breast cancer in a broader public health context (e.g., who is at greater risk for breast cancer, how the celebrity's risk is comparable to other groups of women, screening guidelines, treatment efficacy), then that article does not have an *episodic frame*.

Thematic frame. A news article reporting on a celebrity breast cancer disclosure with a *thematic frame* will convey how the celebrity's breast cancer-related event is relevant in a broader public health context (e.g., risk relative to those of similar or average risk, screening guidelines, treatment efficacy). Such an article can include any and all of the information listed for the episodic frame, but will also provide breast cancer-related information beyond that of the celebrity's own experience.

Misperception reinforcement information. Misperception reinforcement information is specific information that is inaccurate, misleading, or oversimplified. This type of information has the potential to reinforce misperceptions about breast cancer held by the public. Coding categories were determined from previous content analyses and published commentaries on misinformation related to human interest stories covering breast cancer (see discussion in Chapter 2, pps. 47-51) (Chapman et al.,, 2005; Prasad, Lenzer, & Newman, 2016; Sabel & Dal Cin, 2016; Sulik, 2013; Raheb, Scitt, & Rendel, 2015). A total of seven categories were established to identify misperception reinforcement information:

- ***Breast cancer can happen to anyone.*** Specific information that either implies or explicitly states that "anyone (or any woman) can have/get breast cancer."
- ***Early detection.*** Specific information either infers or directly states that early screening/mammography/detection is important/essential for survival.
- ***Screening saves lives.*** The article directly quotes the celebrity (or makes an inference that the celebrity stated) something to the effect of "mammograms save lives" or "screenings saves lives."
- ***All women should begin breast cancer at age 40.*** The media report either explicitly states or implies that all women or most women should begin routine breast cancer screening at the age of 40.
- ***Breast self-exam.*** Specific information is contained in the media report that asserts women should conduct self-breast exams.
- ***"1 in 8."*** Content in the media report discusses a woman's probability of being diagnosed with breast cancer in her lifetime as "1 in 8," or that a woman in the

general population has about a 12.5% chance of being diagnosed with breast cancer in her lifetime.

- ***Mastectomy is the best therapy.*** When discussing a celebrity's double mastectomy/prophylactic/bilateral mastectomy, there is an assertion that the procedure greatly improves the celebrity's chance of survival.

Coder Procedures and Intercoder Reliability Testing

Coder training. Three independent coders were hired and trained to hand code the census of media reports of celebrity breast cancer disclosure. Initial training included discussing each content category, and then a practice session of coding several celebrity breast cancer disclosure media reports from outside the sampling frame (using different media organizations). The practice results were compared and discussed. Some tweaking of the original content codes took place at this time. This initial coder training session took approximately two hours.

Intercoder reliability tests. Because the census contained nearly 1000 articles, testing intercoder reliability for the entire census is impractical and inefficient. Therefore, a random sample was drawn for reliability testing. Riffe et al. (2014) recommend the following formula to determine the appropriate sample size based on population size (i.e., N), population level of agreement (i.e. P; $Q = (P - 1)$), and standard error (i.e., SE):

$$n = \frac{(N-1)(SE)^2 + PQN}{(N-1)(SE)^2 + PQ}$$

According to this formula, 97 units of analysis were randomly drawn to test for intercoder reliability—the total census consisted of 962 units, with an assumed level of 90% agreement in the population desired at a 95% confidence level.

Once the number of sampling units was determined, a random number generator was used for the random draw. After the initial coder training, the coders were given this set of media reports and the codebook to code on their own time. This intercoder reliability sample, and the complete sample, were coded using Qualtrics. Each unit of analysis (i.e., media report) was considered as "one participant" in the Qualtrics database. For both the intercoder reliability sample and the final sample, three identical surveys were created—one for each coder. At the beginning of each "survey," the coder was asked to enter the numerical code of their assigned article. Qualtrics was chosen as the platform to record coded content because: 1) I could access coded content in real time to track progress; 2) all coding was automatically saved; and 3) once all coding was done, for both the intercoder reliability test and the final sample, data was easily downloaded in a CSV or SPSS file.

The coders were given two weeks to code the initial intercoder reliability sample. All data was downloaded into a CSV file, and this file was uploaded into ReCal (Freelon, 2013), an online intercoder reliability service to calculate Krippendorf's alpha. ReCal is an efficient way of calculating reliability with varying levels of measurement and is commonly used to calculate reliability (Freelon, 2013; Riffe et al., 2014). ReCal has the capacity to provide other estimates of agreement (e.g., Scott's pi, Cohen's kappa), but Krippendorpf's alpha is recommended for accounting for chance agreement particularly for nominal variables, which many of the content categories for this study are (Riffe et al., 2014). There is no consensus on acceptable levels of agreement, but Riffe et al. (2014) suggest that an alpha of .80 indicates adequate reliability, although they also state variables with alphas as low as .667 can be considered acceptable for drawing tentative

conclusions. Ultimately, Riffe et al. (2014) recommend calculating and reporting for each variable the percentage of agreement and alpha. If both exceed .80, the variable's reliability is acceptable.

All content categories had a Krippendorpf's alpha over .80 except for media frames (V8) and mastectomy is the best therapy (V15). To reconcile these differences, a second meeting was held with the researcher and coders. The researcher worked with the coders to add clarifying information to the codebook for the content categories in question. The content in red in the coding protocol in Appendix E on pages 237-248 denotes the added language included in the codebook for the content categories with coder disagreement. These two content categories were recoded and the Krippendorpf's alpha for next round of intercoder reliability tests were over 80% coder agreement, but fell slightly short of the .80 alpha; the decision was made to proceed with coding (Table 2 breaks down the percentage of agreement and Krippendorpf's alpha for each content category). The remaining units in the sample were divided evenly between the three coders. The coders took six weeks to code the final sample.

Table 2

Study 1 Intercoder-Reliability Test

	Round 1		Round 2	
Content Category	% agreement	Krippendorf's α	% agreement	Krippendorf's α
V3. Relevancy	100%	1		
V4. Media organization	100%	.98		
V8. Media frames	71.69%	.571	82.65%	.771
V9. Breast cancer can happen to anyone	94.52%	.89		
V10. Early detection	91.35%	.929		
V11. Screening saves lives	96.35%			
V12. Screening at 40	100%	1		
V13. Breast self-exam	97.26%	.946		
V14. "1 in 8"	96.35%	.927		
V15. Mastectomy is the best therapy	73.84%	.543	85.39%	.786

Analytic Approach

To address H1—higher proportions of episodic frames will be present in the census of media reports than thematic frames—frequency analyses were performed. RQ1 (presence of misperception reinforcement information) was addressed in the same way. Frequency analyses and crosstabulations were employed to explore if differences existed among groups as outlined in RQ2. All analyses were performed using SPSS version 22.

Results

The hand-coded sample used to create the *current media salience* measure is the same sample used for this content analysis. Using the "open" search string, as described in Chapter 4, 962 media reports were originally identified. After the coders determined article relevancy, 595 (62% relevancy) were coded.

Volume of coverage. Just under 50 articles (M = 49.42; sd = 14.31) were published annually, on average. Several years fell well below this average, but 2007, 2009, and 2013 stood out as high volume years due to either a high number of disclosures (i.e., 2007 [n = 16] and 2009 [n = 16]) or individuals receiving disproportionate high volumes of coverage (i.e., Elizabeth Edwards [2009] and Angelina Jolie [2013]). Forty-four percent (n = 263) of the media reports were published by the *Associated* Press, 25.88% (n =154) by *The New York Times*, and 29.92% (n = 178) by *The Washington Post*. The majority of the media reports were news articles (94.5%), 4.2% were editorials, and 1.3% were obituaries. The media reports of celebrity breast cancer disclosures appeared in numerous sections, including: News/Main news (n = 173; 29.08%); Entertainment/entertainment news (n = 54; 9.1%); Style/fashion (n = 26; 4.4%);

Metropolitan desk (*n* = 25; 4.2%); Lifestyle (*n* = 20; 3.4%); and Arts/culture (*n* = 15; 2.5%). Only four (.7%) of the media reports were published in the health sections of the newspapers. Media reports ranged in length from 100 to 6174 words (*M = 612.83; sd = 659.22).*

Framing analysis. H1 predicts that higher proportions of episodic media frames, as opposed to thematic frames, would be present in the census of media reports of celebrity breast cancer disclosures. Of the 595 media reports analyzed, 476 (80.0%) of the media reports were written with an episodic frame and 119 (20.0%) were written with a thematic frame. Overall, the national news coverage of celebrity breast cancer disclosures focuses on the celebrity and her or his disclosure and experience with breast cancer. Media reports contextualizing the celebrity's disclosure and breast cancer-related experience with greater public health information (e.g., prevention, risk, or treatment information at a population or subpopulation level) is only present in one out of every five articles. Therefore, H1 is confirmed.

Misperception reinforcement information. RQ1 asks if information which may reinforce public misperceptions held about breast cancer is present in the media reports of celebrity breast cancer disclosures. From this point forward "misinformation" is used as shorthand to describe breast cancer-related information that might be considered misleading (e.g., 1 in 8 women will be diagnosed with breast cancer in her lifetime) or an overgeneralization (e.g., breast cancer can happen to any woman). Seven types of misinformation were identified from previous research and published commentaries on breast cancer misinformation/misperceptions and coded for within this census of media

Table 3

Misperception Reinforcement Information Present in the Census of Media Reports

Content Category	# of media reports	% of media reports
V9. Breast cancer can happen to anyone	6	1%
V10. Early detection is key	66	11.1%
V11. Screening saves lives	39	6.6%
V12. Begin screening at 40	15	2.5%
V13. Conduct self-breast exams	12	2%
V14. "1 in 8" women	13	2.2%
V15. Mastectomy is the best therapy	105	17.6%

reports. Table 3 describes the frequency and percentage of media reports with each misinformation variable present.

As Table 3 demonstrates, misinformation is not predominant in the news coverage of celebrity breast cancer disclosures. Given the lack of population and subpopulation breast cancer-related information, this initial finding is not surprising. However, for the three content categories where at least five percent of the census has this information present (V10; V11; V15), cross-tabulations reveal that misinformation is more prevalent in episodically framed media reports rather than thematically framed. Although not more likely to be present: a greater proportion of thematically framed media reports contain misinformation than episodically framed. For example, 75 (15.8%) of the 476 episodically framed media reports contain information relevant to *mastectomy is the best therapy*. But, of the 119 thematically frames media reports, 30 or 25.2% of the articles contain information that is considered relevant to *mastectomy is the best therapy* content category.

The presence of misinformation in a media report with an episodic frame largely occurs because there is no breast cancer-related information beyond that of the celebrity's own experience. In these cases, much of the misinformation is a direct quote from the

Table 4

Qualitative Evidence of "Misinformation" in Media Reports with Episodic Frames

Content Category	Example Quote of "Misinformation"
V10) Early Detection	"The 36-year-old actress is benefiting from early detection, however, and the cancer is not life-threatening, according to a statement issues by Applegate's publicist." "And the singer Sheryl Crow, interviewed on CNN Thursday night, said that like many others she couldn't help but feel a tinge of fear when she heard Edwards' story even though her own cancer was caught early and she believes she has been cured." "Elizabeth Edwards said Wednesday she feels she let down her family and the country by neglecting to get mammograms that could have caught her cancer earlier." "Without early detection on my side, I could be telling a very different story," she said. "Or not be here to tell it at all."
V11) Screening Saves Lives	"She had always been diligent about her exams, and thank God she had been diligent about her exams," the governor said, adding, "If she waited until she was 50, this would be a very different situation." "'It had the chance to migrate because I sat at home doing whatever I thought was important and didn't get mammograms,' Mrs. Edwards said."
V15) Mastectomy is the Best Therapy	"Rather than risk it, Sykes, whose mother's side of the family has a history of breast cancer, opted to have them removed." "Kathy Bates says she is recovering from a double mastectomy. She decided to have the operation 'after much consideration' and won't have to undergo radiation or chemotherapy. Her doctors have assured her she'll be around 'for a long time.'" "Lee, 48, who is Gov. Cuomo's live-in girlfriend, announced Tuesday that she had been diagnosed with breast cancer, and she would have the double mastectomy instead of a lumpectomy and radiation treatment."

celebrity speaking about her own experience. Table 4 on page 109 offers several statements from episodically framed media reports with content relevant to the *early detection is key, screening saves lives,* and *mastectomy is the best therapy* content categories. This table demonstrates what misinformation with an episodic frame looks like.

Table 5 on page 111 offers qualitative evidence of these same misinformation content categories from thematically framed media reports. Much of this misinformation arises for the same reasons that it does in episodically framed media reports. That is, while the media report was coded as having a thematic frame because it did offer some population or sub-population breast cancer-related information, the particular misinformation variable present may not have been contextualized in such a way. For example, 41.02% (n = 48) of the 117 media reports on Angelina Jolie's disclosure have a thematic frame. Of those, 37.50 % (n = 18) had the *mastectomy is the best therapy* variable present. Many of these media reports may have contained information about how many women have the BRCA genetic mutation, for example, but, the article may not have offered any information about diagnosis or survival rates between women who opt for a bilateral mastectomy and for those who do not. Therefore, that lack of information is considered misinformation.

Overall, there was very little misinformation present in the census of media reports. The two most common categories of misinformation present were *early detection is key* (11.1%) and *mastectomy is the best therapy* (17.6%). On the one hand, the lack of misinformation can be explained by the lack of the breast cancer-related information

Table 5

Qualitative Evidence of "Misinformation" in Media Reports with Thematic Frames

Content Category	Example Quote of Misinformation
V10) Early Detection	"Other reasons for the decline in the cancer mortality rate in men and women, Dr. Edwards says, include better treatment, better care, and earlier detection." "MELISSA ETHERIDGE'S story of suddenly discovering a large tumor in her breast illustrates the importance of early and regular screening for cancer for all women." "'I consider all of those circumstances unbelievably fortuitous,' she told reporters, because it allowed doctors to spot the cancer at an early stage and begin the treatments that would allow her to live 3 1/2 more years."
V11) Screening Saves Lives	"...Women need to be vigilant, continue getting mammograms and seeing their physician yearly." "I can only hope my story will do the same and inspire every woman who hears it to get a mammogram, to take a self- exam," she said. "No excuses. It is the difference between life and death." "Without screening, patients show up later, with larger tumors and potentially metastatic disease. Five-year survival for Stage 3 breast cancer is 36 percent. For Stage 4, it's 7 percent. Screening isn't about profits. It's about saving lives."
V15) Mastectomy is the best therapy	"Doctors told her she had an 87 percent chance of getting breast cancer and a 50 percent risk of ovarian cancer. She said the surgery reduced her risk of breast cancer to below 5 percent." "'My doctors estimated that I had an 87 percent risk of breast cancer and a 50 percent risk of ovarian cancer.'"After two very intense months of recalling my mom's death from ovarian cancer and her father's death from pancreatic cancer, I made the very difficult decision to have a prophylactic mastectomy to avoid breast cancer."

present in the reports, but on the other hand misinformation is more likely to occur in thematically framed articles than those with an episodic frame.

Differences by celebrity and breast cancer-related attributes. The above results describe the overall volume and content when analyzing media coverage in the aggregate. RQ2 explores if these media outcomes differ based on specific celebrity attributes (i.e., age, career type, and level of celebrity status) or by the type of breast cancer-related event the celebrity discloses.

Volume of coverage. Volume of coverage indicates the degree of media salience celebrity breast cancer disclosures receive. The following subsections offer tables and interpretation that describe the volume and percentage of disclosures, volume and percentage of media reports, and the average number of media reports by the four attribute groups.

Age. As described in Chapter 4, the celebrities-of-interest were divided into four age groups, based on the age in which the celebrity experienced her breast cancer-related event and the year of the disclosure (pre- or post-2010). Because the sample is a census, descriptive statistics are appropriate, but the disparate frequency of disclosures/media reports between groups makes it difficult to determine what results may be considered "significant."

Comparing means does reveal some potentially meaningful differences between groups. Table 6 shows that the two age groups under the recommended screening ages (pre- and post-2010) are reported on, on average, nearly two times as often as the age groups at or over the recommended screening ages. However, it should be noted that two of the celebrities-of-interest (i.e., Edwards and Jolie) make up a large proportion of the

Table 6

Volume of Media Coverage by Age Group

Age Group	Disclosures		Media Reports		Media Reports
	#	%	#	%	*M*
Before 2010					
Under 40	4	3.60%	35	5.90%	8.75 (*sd =12.55)*
Over 40	56	50.90%	267	44.90%	5.04 (*sd =11.64*)
After 2010					
Under 50	26	23.60%	221	37.10%	8.50 (*sd* = 23.34)
Over 50	24	21.90%	72	12.10%	3.00 (*sd* = 4.46)
Total	107	100%	595	100%	5.56 (*sd* = 14.42)

Note. The total number of disclosures for the age group attribute group is 107 instead of 110. As described in Chapter 4, the three men celebrities-of-interest were removed from this analysis.

age group (and of the remaining three attribute groups). Seventy-four media reports of Elizabeth Edwards' first disclosure (D1) were retrieved. Removing her from the Over 40 group lowers the group's mean to 3.69 (*sd = 6.51)* from 5.04 (*sd* = 18.64). This outlier does not change the result that "younger" women's disclosures (i.e., Under 40 and Under 50) receive more media attention, but the 117 media reports on Jolie's disclosures does shift the mean score for the Under 50 group in a meaningful way. Removing the Jolie disclosure puts the average number of reports for the celebrities-of-interest Under 50 at 4.20 (*sd* = 7.64) from 8.50 (*sd* = 23.34). Therefore, while the "younger" groups still received more media attention than the "older" groups (i.e., Over 40 and Over 50), differences in mean scores are not meaningful. (The Under 40 age group is likely too small to interpret the differences in mean scores as anything other than chance.)

Career type. Table 7 reveals that the Personal Affiliation group has the greatest mean score for volume of media reports. Yet, as in the Age groups, the high reporting rate for Edwards (D1) inflates the mean score for the group she is in (i.e., Personal Affiliation)—Edwards accounts for 92 out of the 95 media reports. In addition, Jolie's

Table 7

Volume of Media Coverage by Career Type Group

Career-type	Disclosures		Media Reports		Media Reports
	#	%	#	%	*M*
Academic/author/activist	18	16.40%	24	4.00%	1.33 (*sd* = 1.50)
Actor	25	22.70%	192	32.30%	7.68 (*sd* = 23.02)
Athlete/sports related	6	5.50%	58	9.70%	9.97 (*sd* = 14.07)
Journalist/News anchor	15	13.60%	75	12.60%	5.00 (*sd* = 6.85)
Musician	15	13.60%	67	11.30%	4.47 (*sd* = 7.43)
Personal affiliation	7	6.40%	97	16.30%	13.90 (*sd* = 27.27)
Politician/public servant	9	8.20%	22	3.70%	2.44 (*sd* = 4.16)
Television personality	15	13.60%	60	10.10%	4.00 (*sd* = 9.20)
Total	110	100%	595	100%	5.41 (*sd* = 14.25)

disclosure accounts for 117 of the 192 media reports for the Actor category, putting the groups' mean at 3.13 (*sd* = 3.44), instead of 7.68 (*sd* = 23.02). The outliers are not isolated to Jolie and Edwards in the career type category. For example, in the Athlete/Sports-related category, the number of media reports specific to Kay Yow's disclosure account for 65.50% of the group. Removing Yow from the group lowers the mean to 4 (*sd* = 2.55) from 9.97 (*sd* = 14.07) and places that category's mean in line with the other groups. Overall, there is no meaningful difference in the number of media reports, on average, written about a celebrity breast cancer disclosure based on the celebrity-of-interest's career type.

Breast cancer-related event type. Determining differences in volume of media coverage based on breast cancer-related event type suffers from the same issues as the previous categories. Jolie and Edwards make up such significant portions of their corresponding groups, 96% and 92%, respectively, that no meaningful conclusions can be drawn about impacts from the groups to which they have been assigned to. However, the similarities between the Treatment and Death groups (i.e., number of disclosures, similar

Table 8

Volume of Media Coverage by Breast Cancer-Related Event Type Group

Cancer Event-type	Disclosures		Media Reports		Media Reports
	#	%	#	%	*M*
Prevention	7	6.40%	122	20.50%	8.00 (*sd* = 43.92)
Detection	0	0.00%	0	0%	-
Diagnosis	5	4.60%	81	13.60%	13.50 (*sd* = 29.66)
Treatment	43	39.100%	216	36.30%	5.14 (*sd* = 7.76)
Survivorship	1	.90%	9	1.50%	-
Death	54	49.10%	167	28.10%	3.09 (*sd* = 6.104)
total	110	100%	595	100%	5.41 (*sd* = 14.25)

standard deviations) allow for some comparison between the two groups. Treatment disclosures did receive higher rates of media coverage, on average, than Death disclosures. (See Table 8 for results.)

Level of celebrity status. The results of stratifying the volume of coverage by level of celebrity status are displayed in Table 9. Although the disclosure group sizes are somewhat disparate (7 – 20), in general, the number of media reports increase with the level of celebrity status. Celebrity status Level 8 is an exception to this overall trend—there are only five celebrities and only 11 media reports for this group. Just as the other attribute categories had extreme outliers in specific groups, so does level of celebrity status. Media reports about Jolie's disclosure make up 47% of the total media reports for the Level 10 group, and reports on Edwards make up 61.20% of the Level 9 group. Removing these cases results in a mean of 13.20 (*sd* = 10.48) for Level 10 instead of 22.64 (*sd* = 38.84), and a mean of 5.22 (*sd* = 4.12) from 12.10 (*sd* = 22.09) for Level 9. Even with the exclusion of Jolie's disclosures, the group of celebrities-of-interest which comprise Level 10 have the greatest number of media reports than any other group. However, the removal of Edwards' cases lowers the Level 9 mean score below Levels 7 and 8.

Table 9

Volume of Media Coverage by Level of Celebrity Status Group

Level of Celebrity Status	Disclosures		Media Reports		Media Reports
	#	%	#	#	*M*
2 – lowest level	7	6.36%	0	0%	-
3	10	9.10%	1	.20%	-
4	15	13.63%	18	3.00%	1.20 (*sd* = 1.47)
5	16	14.55%	21	3.50%	1.31 (*sd* = 1.35)
6	20	18.18%	65	10.90%	6.19 (*sd* = 5.15)
7	17	15.45%	109	18.30%	6.41 (*sd* = 9.79)
8	5	4.55%	11	1.80%	2.20 (*sd* = 1.10)
9	10	9.10%	121	20.30%	12.10 (*sd* = 22.09)
10 – highest level	11	10.00%	249	41.80%	22.64 (*sd* = 32.84)
total	110	100%	595	100%	5.41 (*sd* = 14.25)

I argue that the removal of Jolie and Edwards as outliers for the level of celebrity status group is less justified than their removal from the other categories. Volume of media coverage at the time of disclosure and in the long term are two of the measures which contributed to the level of celebrity status index. Therefore, celebrities-of-interest in the highest level of celebrity status groups should theoretically have the greatest number of media reports written about their breast cancer disclosures. The results here offer preliminary evidence that a celebrity's level of celebrity status is associated with the volume of media coverage she or he receives about her or his breast cancer disclosure.

Media frames. As previously reported, there was a stark contrast between the presence of episodic and thematic frames in the media reports analyzed. Dividing the media reports based on age, career type, breast cancer-related event type, and level of celebrity status does provide further insight as to when we might expect particular frames to be present in media coverage of celebrity breast cancer disclosures.

Age. Table 10 provides the number and percentage of media reports with both the episodic and thematic frames by the four age groups. What appears to be a meaningful difference is the framing outcomes for the Under 50 group. Again, Jolie's outlying number of 117 media reports, and 48 of those with thematic frames, inflates the descriptive findings between groups. In addition, Edwards' 26 thematically framed media reports on her first disclosure (D1) conflate the findings here as well. Removing Edwards (D1) from the Over 40 group puts the percentage of thematic frames at 7.25% (*n* = 7.25) rather than 15% (*n* = 40) and episodic frames at 92.75 (*n* = 179) rather than 85% (*n* = 227). Removing Jolie's cases also decreases the percentage of thematic frames present in the Under 50 group (15.38%, *n* = 16). Given these adjustments, it is clear that no age group has a strong presence of thematic frames in the news coverage of the celebrities' breast cancer disclosures.

Table 10

Frequency of Media Frames by Age Group

Age Group	Episodic Frames		Thematic Frames		
	#	%	#	%	Total
Before 2010					
Under 40	29	82.9%	6	17.1%	35
Over 40	227	85.0%	40	15.0%	267
After 2010					
Under 50	157	71.0%	64	29.0%	221
Over 50	63	87.5%	9	12.5%	72
Total	476		119		595

Note. Percentage across rows represents the percentage of frames present in that group (e.g., *Under 40*). Because the number of disclosures (cases) are disparate between groups, comparing the proportion of frames present in each group is more meaningful than representing the percentage between groups.

Table 11

Frequency of Media Frames by Career Type

Career-type	Episodic Frames		Thematic Frames		
	#	%	#	%	total
Academic/author/activist/creative	23	95.8%	1	4.2%	24
Actor	136	70.8%	56	29.2%	192
Athlete/sports related	51	87.9%	7	12.1%	58
Journalist/News anchor	69	92.0%	6	8.0%	75
Musician	62	92.5%	5	7.5%	67
Personal affiliation	66	68.0%	31	32.0%	97
Politician/public servant	20	90.9%	2	9.1%	22
Television personality	49	81.70%	11	18.3%	60
total	476		119		595

Note. Percentage across rows represents the percentage of frames present in that group (e.g., *Actor*). Because the number of disclosures (cases) are disparate between groups, comparing the proportion of frames present in each group is more meaningful than representing the percentage between groups.

Career type. As shown in Table 11, these groups in the aggregate appear to have meaningful differences based upon the career type. However, in the three groups which have the highest rates of thematic frames, all of the thematic frames come from one celebrity (Jolie [actor, n = 48]; Edwards (D1) [Personal Affiliation, n = 26]; Sandra Lee (D1, D2) [Television Personality, n = 11]). Otherwise, a few groups have slight differences in the volume of frames, but the differences are not enough to conclude the media frames differ by career type.

Breast cancer-related event type. The number and percentage of media reports with both the episodic and thematic frames by the six dimensions of the cancer event-type are displayed in Table 12. As discussed in the *Volume of Media Coverage by Breast Cancer-Related Event Type* section, reports on Jolie and Edwards account for almost all of the reports in the Preventive and Diagnosis categories—making these two groups ineffectual for analysis. In addition, the low volume of thematic frames in the Treatment

Table 12

Frequency of Media Frames by Breast Cancer-Related Event Type

Cancer Event-Type	Episodic		Thematic		
	#	%	#	%	total
Prevention	73	59.8%	49	40.2%	122
Detection	0	0%	0	0%	0
Diagnosis	54	66.7%	27	33.3%	81
Treatment	185	85.6%	31	14.4%	216
Survivorship	8	88.9%	1	11.1%	9
Death	156	93.4%	11	6.6%	167
total	476		119		595

Note. Percentage across rows represents the percentage of frames present in that group (e.g., *Prevention*) Because the number of disclosures (cases) are disparate between groups, comparing the proportion of frames present in each group is more meaningful than representing the percentage between groups.

and Death categories makes the percentage difference between the two groups not particularly meaningful.

Level of celebrity status. Table 13 provides the number and percentage of media reports with both the episodic and thematic frames by the nine levels of celebrity status groups. In this analysis, cross-tabulations reveal that thematic frames are most prevalent in media reports of celebrity breast cancer disclosures for celebrities which occupy the three highest levels of celebrity status (8-10). These results indicate that the proportion of episodic and thematic frames present in media reports of celebrity breast cancer disclosures differ based on level of celebrity status.

Misinformation. As discussed in the general *Misinformation* section in this chapter, overall, misinformation was largely absent from this census of media reports. However, three of the seven content categories (i.e., V10 early detection; V11 screening saves lives; V15 mastectomy is the best therapy) have a large enough presence (6.6% - 17.6%) to investigate possible differences between the attribute groups. The following

Table 13

Frequency of Media Frames by Level of Celebrity Status

Level of Celebrity Status	Episodic Frames		Thematic Frames		
	#	%	#	%	total
2 – lowest level	0	0%	0	0%	0
3	1	100.00%	0	0%	1
4	17	94.44%	1	5.56%	18
5	20	95.00%	1	4.80%	21
6	59	90.80%	6	9.20%	65
7	101	92.70%	8	7.30%	109
8	8	72.70%	3	27.30%	11
9	89	73.55%	32	26.45%	121
10 – highest level	181	72.70%	68	27.30%	249
total	476		119		595

Note. Percentage across rows represents the percentage of frames present in that group (e.g., *2*). Because the number of disclosures (cases) are disparate between groups, comparing the proportion of frames present in each group is more meaningful than representing the percentage between groups.

sections discuss differences in results for these three misinformation content categories by attribute groups.

Age. Table 14 on page 121 presents the presence (quantity and percentage) of the misinformation categories *early detection*, *screening saves lives*, and *mastectomy is the best therapy* by age groups. On their face, the four age groups have similar proportions of *early detection* and *screening saves lives* misinformation present in the media reports. The Under 50 group stands out in the *mastectomy is the best therapy* category, with 39.40% (n = 87) of the media reports containing this type of misinformation. The volume of media reports about Jolie and Edwards's (D1) breast cancer disclosures remain as extreme outliers within their respective groups, although, in this analysis, the removal of Jolie and Edwards' cases is more likely to increase the proportion of misinformation

Table 14

Presence of Misinformation by Age Group

	Not Present		Present		
Content Category	#	%	#	%	Total
V10. Early Detection					
Before 2010					
Under 40	29	82.90%	6	17.10%	35
Over 40	242	90.60%	25	9.40%	267
After 2010					
Under 50	198	89.60%	23	10.40%	221
Over 50	60	83.30%	12	16.70%	72
Total	529		66		595
V11. Screening Saves Lives					
Before 2010					
Under 40	32	91.40%	3	8.60%	35
Over 40	256	95.90%	11	4.10%	267
After 2010					
Under 50	203	91.90%	18	8.10%	221
Over 50	65	90.30%	7	9.70%	72
Total	556		39		595
V15. Mastectomy is the Best Therapy					
Before 2010					
Under 40	34	97.10%	1	2.90%	35
Over 40	263	98.50%	4	1.50%	267
After 2010					
Under 50	134	60.60%	87	39.40%	221
Over 50	59	81.90%	13	18.10%	72
Total	490		105		595

present in the group than diminish it as was seen in previous analyses. This outcome is due to greater distribution of misinformation across celebrities-of-interest, whereas thematic frames, for example, were largely present in news coverage related to Jolie's and Edwards' (D1) breast cancer disclosures.

The results indicate that the presence of *early detection* misinformation is less prevalent in the Over 40 group than the other three age groups. The *screening saves lives* misinformation variable follows a similar pattern. There are no differences in prevalence

rates for the Under 40 and Over 50 groups, this information is largely absent from the Over 40 group, and is most common in the Under 50 group. Finally, for the *mastectomy is the best therapy* misinformation variable,

misinformation was all but absent in the Under 40 and Over 40 age groups. Any misinformation here is found in the Under 50 and Over 50 age groups. Even when Jolie's 117 cases are removed this misinformation is still nearly two times (35.58%) that of the Over 50 group. Overall, between the three categories of misinformation presented here, the Under 50 age group contained higher levels of misperception reinforcement information than the other three groups. Although, while at lower rates than the Under 50 group, 10 to 15 % of the media reports of breast cancer disclosures from the Over 50 group consistently contained the three types of misinformation.

Career type. Table 15 presents the results of the cross tabulations exploring potential differences between misinformation outcomes (i.e., *early detection*, *screening saves lives*, and *mastectomy is the best therapy*) and career type groups. The way in which celebrities are divided allocates the cases of misinformation in a way that does not require removing Jolie's (with the exception of *mastectomy is the best therapy)* and Edwards' cases. This is because the percent present in the groups (i.e., Actor and Personal Affiliation) is already quite low. For both the *early detection* and *screening saves lives* content categories, misinformation is present in nearly 15% to 21% of media reports for the Journalist/News Anchor and Television Personality groups. For almost all other groups, the presence of such misinformation is under 5%. Once Jolie's cases are removed from the *mastectomy is the best therapy* results, misinformation is still present in 22.67% (17 of 75) of the media reports, making that the second highest frequency of

Table 15

Presence of Misinformation by Career Type Group

	Not Present		Present		
Content Category	#	%	#	%	Total
V10. Early Detection					
Academic/author/activist/creative	24	100.00%	0	0%	24
Actor	177	92.20%	15	7.80%	192
Athlete/sports related	58	100.00%	0	0%	58
Journalist/news anchor	59	78.70%	16	21.30%	75
Musician	61	91.00%	6	9.00%	67
Personal affiliation	81	83.50%	16	16.50%	97
Politician/public servant	21	95.50%	1	4.50%	22
Television personality	48	80.00%	12	20.00%	60
Total	529		66		595
V11. Screening Saves Lives					
Academic/author/activist/creative	23	95.80%	1	4.20%	24
Actor	188	97.90%	4	2.10%	192
Athlete/sports related	56	96.60%	2	3.40%	58
Journalist/news anchor	64	85.30%	11	14.70%	75
Musician	65	97.00%	2	3.00%	67
Personal affiliation	88	90.70%	9	9.30%	97
Politician/public servant	21	95.50%	1	4.50%	22
Television personality	50	83.33%	10	16.66%	60
Total	555		40		595
V15. Mastectomy is the Best Therapy					
Academic/author/activist/creative	24	100.00%	0	0%	24
Actor	125	65.10%	67	34.90%	192
Athlete/sports related	57	98.30%	1	1.70%	58
Journalist/news anchor	71	94.70%	4	5.30%	75
Musician	66	98.50%	1	1.50%	67
Personal affiliation	95	97.90%	2	2.10%	97
Politician/public servant	22	100.00%	1	0%	22
Television personality	30	50.00%	30	50.00%	60
Total	490		105		595

the eight groups. The Television Personality group contains the highest proportion of *mastectomy is the best therapy* misinformation—50% of the 60 media reports. Overall, misinformation was most prevalent in the Television Personality and Journalist/News Anchor groups, with the exception of the Actor group for *mastectomy is the best therapy*.

Table 16

Presence of Misinformation by Breast Cancer-Related Event Type Group

	Not Present		Present		
Content Category	#	%	#	%	Total
V10. Early Detection					
Prevention	118	96.70%	4	3.30%	122
Detection	0	0%	0	0%	0
Diagnosis	66	81.50%	15	18.50%	81
Treatment	172	79.60%	44	20.40%	216
Survivorship	8	88.90%	1	11.10%	9
Death	165	98.80%	2	1.20%	167
Total	529		66		595
V11. Screening Saves Lives					
Prevention	120	98.40%	2	1.60%	122
Detection	0	0%	0	0%	0
Diagnosis	73	90.10%	8	9.90%	81
Treatment	191	88.40%	25	11.60%	216
Survivorship	7	77.80%	2	22.20%	9
Death	165	98.80%	2	1.20%	167
Total	556		39		595
V15. Mastectomy is the Best Therapy					
Prevention	63	51.60%	59	48.40%	122
Detection	0	0%	0	0%	0
Diagnosis	78	96.30%	3	3.70%	81
Treatment	179	82.90%	37	17.20%	216
Survivorship	8	88.90%	1	11.10%	9
Death	162	97.00%	5	3.00%	167
Total	490		105		595

Breast cancer-related event type. Table 16 presents the presence (quantity and percentage) of the misinformation categories *early detection*, *screening saves lives*, and *mastectomy is the best therapy* by breast cancer-related event type groups. The overall composition of the breast cancer-related event types are not conducive for cross-tabulation analysis. Media reports of Jolie and Edwards' (D1) disclosures make up over 91% of their groups (Prevention and Diagnosis, respectively). Also, the Survivorship group consists of only one disclosure. With that, percentage of misinformation present (or

Table 17

Presence of Misinformation by Level of Celebrity Status Group

Content Category	Not Present		Present		Total
	#	%	#	%	
V10. Early Detection					
2 – lowest level	0	0%	0	0%	0
3	1	100.00%	0	0%	1
4	16	88.89%	2	11.11%	18
5	21	100.00%	0	0%	21
6	55	84.62%	10	15.38%	65
7	101	92.70%	8	7.30%	109
8	9	81.82%	2	18.18%	11
9	101	83.47%	20	16.53%	121
10 – highest level	225	90.36%	24	15.15%	249
Total	529		66		595
V11. Screening Saves Lives					
2 – lowest level	0	0%	0	0%	0
3	1	100.00%	0	0%	1
4	17	94.44%	1	5.56%	18
5	21	100.00%	0	0%	21
6	59	90.77%	6	9.23%	65
7	106	97.25%	3	2.75%	109
8	11	100.00%	0	0%	11
9	109	90.08%	12	9.92%	121
10 – highest level	232	93.17%	17	6.83%	249
Total	556		39		595
V15. Mastectomy is the Best Therapy					
2 – lowest level	0	0%	0	0%	0
3	1	100.00%	0	0%	1
4	17	94.44%	1	5.56%	18
5	20	95.24%	1	4.76%	21
6	56	85.15%	9	13.85%	65
7	106	97.25%	3	2.75%	109
8	11	100.00%	0	0%	11
9	115	95.04%	6	4.96%	121
10 – highest level	164	65.86%	85	34.14%	249
Total	490		105		595

not) can be compared between the treatment and death groups. For all three variables, under 3% of the media reports contain misinformation for the Death group, whereas between 11% and 20% of the Treatment group's media reports contain some misinformation.

Level of celebrity status. Table 17 presents the results of the cross-tabulations exploring potential differences between the misinformation outcomes (i.e., *early detection*, *screening saves lives*, and *mastectomy is the best therapy*) and level of celebrity status groups. Unlike volume of media coverage and media frames, the differences in misinformation outcomes by level of celebrity of status are less systematic. The only group that consistently had misinformation present in 10% or more of the media reports was Level 6. Between the three misinformation variables, Levels 9 and 10 intermittently had misinformation present in at least 10% of the media reports. The greatest proportion of misinformation present in media reports connected to one level of celebrity status group is the 34.14% of media reports with *mastectomy is the best therapy* information in the Level 10 celebrity status group.

Summary of Study 1 Findings

Study 1 analyzes news content present in the reporting on celebrity breast cancer disclosures that are an indicator of the greater media environment from 2005 through 2016. The central finding of this study is that little population and subpopulation breast cancer-related information (i.e., thematic frame) are present in media reports of celebrity breast cancer disclosures. Of the 595 media reports analyzed, only 20% contained this type of information, while 80% focused solely on the celebrity and her own experience with breast cancer (i.e., episodic frame).

Misperception reinforcement information was largely absent in the news coverage. Only three of the seven content categories were present in 5% or more of the media reports—*early detection is key* (11.1%), *screening saves lives* (6.6%), and *mastectomy is the best therapy* (17.6%). A greater proportion of this misperception reinforcement information was present in thematically framed media reports, but existed at higher rates in episodically framed media reports.

Exploring differences in these media outcomes, including volume of coverage by age, career type, breast cancer-related event type, and level of celebrity status, provides some further insight as to which attributes may influence certain media outcomes. However, many of these analyses are affected by the disproportionate coverage of Angelina Jolie's and Elizabeth Edwards' (D1) breast cancer disclosures—media reports on Jolie's disclosure make up nearly 20% of overall coverage, and Edwards' disclosure comprises over 12% of the coverage.

After removing Jolie and Edwards' cases, there were no differences in rates of coverage by age groups. The same can be said about differences between career type groups. For breast cancer-related event type, the only comparable groups were the Treatment and Death groups. Between these two, the Treatment group did receive higher rates of coverage than the death group. Rates of media coverage of celebrity breast cancer disclosures did increase as the level of celebrity status increased.

In regards to the presence of episodic and thematic frames, no meaningful differences existed between age, career type groups, or breast cancer-related event type. However, 87% of the thematic frames present in the study's census were in media reports written about celebrities in the highest levels of celebrity status.

Misinformation was only found sporadically throughout the census, but the highest rates of it were found in media reports of celebrity breast cancer disclosures from celebrities in the Under 50 and Over 50 categories. Analyses exploring differences in the proportions of misinformation present by career type found that misinformation was most prevalent in media reports on breast cancer disclosures from Television Personalities and Journalists/News Anchors. And 50% of the media reports on those in the Actor group contained information related to *mastectomy is the best therapy*. No meaningful differences were found between breast cancer-related event type groups, but misinformation was most common in the Level 6 and Level 10 groups in the level of celebrity status category. Discussion of the results and potential implications are included in Chapter 7.

Chapter 6

Study 2. Breast Cancer-Related Information Seeking

Study 2 tests the relationships between the celebrity breast cancer disclosure, media coverage of the disclosure, and online breast cancer-related information seeking outcomes. Media coverage includes frequency and framing of news coverage. In line with the proposed hypotheses and the conceptual model of effects, the objectives for Study 2 were to test the main effects from the entire celebrity sample (H2) and the frequency (H2a) and framing (H3) of news coverage on breast cancer-related information seeking outcomes. Moderation effects on the hypothesized relationships by age, career type, breast cancer-related event type, and level of celebrity status were explored (RQ3). Due to the low volume of certain types of coverage not all hypotheses could be tested. These issues are presented in the results section of this chapter.

Independent Variables

Volume of media coverage. The final weekly and monthly volume of media coverage during the study's timeframe (January 1, 2005–December 31, 2016). The weekly-level variable was constructed from the results of Study 1. To accommodate for the monthly Google Trends scaled data (see Dependent Variable on p. 130), the weekly totals were combined into calendar month totals as well.

Media frames. The total weekly volume for each of the media frames (episodic and thematic) analyzed in Study 1. To accommodate for the monthly Google Trends scaled data, the weekly totals were combined into calendar month totals as well.

Celebrity Attribute Variables—Potential Moderator Variables

Moderation potential between media reports and breast cancer-related information seeking outcomes by *age*, *career type*, *breast cancer-related event*, and *level of celebrity status* was tested. The specific measure for each variable is discussed in detail in Chapter 4 (p. 78-97).

Dependent Variable—Online breast cancer-related information seeking

Google Trends data were used as a proxy for online breast cancer-related information seeking. All Google Trends data for this study were accessed and downloaded from https://www.google.com/trends/. The location for RSV retrieval was set to "United States." Weekly Google Trends data are only available through five-year calendar periods; beyond that the data is scaled by calendar month. However, up until August of 2016, weekly-level data was available for any length of time. Therefore, the data in almost all of the existing literature using Google Trends, is weekly-level. The formulated hypotheses are based on this prior evidence, so it is appropriate to use the same level of data that was used in the extent research. Although to capture the entire time period of study, the monthly-level data must be used. Using the two levels of data creates an opportunity to compare results across the two levels, which no prior research has done. With this, all RSV data for each search term were retrieved for the entire 12-year time period under study (2005-2016) (monthly-level data) and in four-year increments (2005-2008; 2009-2012; 2013-2016) (weekly level-data). Hypotheses and research questions were tested using both the monthly- and weekly-level RSV datasets.

Selecting valid search query terms to retrieve breast cancer-related RSV data was an iterative process. Research using Google Trends data as a proxy for online information seeking uses search terms that have face validity, but generally do not offer further

evidence of validity. One method suggested by Mellon (2013) is to use the "related searches" feature in Google Trends to select search terms based on their content validity as well as face validity. The "related searches" are search terms that were most frequently searched with the specified term in the same search session within the same country or region (Google, n.d.).

For example, two treatment-related search terms are "mastectomy" and "breast surgery" (admittedly, face validity may not be equal for both). Related searches to "mastectomy" include "breast mastectomy," "double mastectomy," and "lumpectomy." Related searches to "breast surgery" include "plastic surgery," "breast reduction," "breast augmentation," and "breast implants." The related searches demonstrate that "mastectomy" has good content validity for a study investigating search query volume related to breast cancer information seeking, and "breast surgery" has poor content validity. Content validity was assessed for each search term selected. When a few of the "related searches" put into question the content validity of a search term/phrase, the invalid term or terms were removed from the original search term. For example, for "breast cancer symptoms," "lung" was one of the related searches. To remove "lung" from the relative search volume data, the following search string was entered into Google Trends: "breast cancer symptoms -lung." The final search terms (with the validated search string and related search terms) selected for study are included in Appendix C on pages 211-224.

Just as the majority of research using Google Trends as a proxy for health information seeking has not tested search term validity beyond face validity, most research has only used one or two search terms to represent all search terms related to the

topic under study. For example, Noar et al. (2013) used the search terms "pancreatic cancer(s)" and "cancer(s)" (removing search queries for pancreatic cancers) to retrieve Google RSV data related to celebrity pancreatic cancer announcements and deaths. For the given subject, this wide net approach has face validity and likely content validity, but such general search terms do not afford a more nuanced understanding about specific cancer information seeking behavior.

Cancer, be it pancreatic cancer or breast cancer, and the possible information one may be seeking about it, is complicated. One could be seeking general information about what breast cancer is, but it is probable many are aware of what breast cancer is (at least in a general sense), and instead may be seeking information related to how to prevent or detect breast cancer, for example. Noar and colleagues (2015) hypothesized that Angelina Jolie's BRCA1 genetic mutation confirmation and her subsequent prophylactic double mastectomy disclosure would stimulate significant online breast cancer-related information seeking. Noar et al. (2015) did use the search phrase "breast cancer" to retrieve RSV data, but also used multiple search terms to construct risk assessment, genetics, and treatment dimensions.

This study used a framework similar to Noar et al.'s (2015) to quantify online breast cancer information seeking outcomes. Some of the specific search terms used by Noar et al. (2015) were used for Study 2 (e.g., breast cancer risk, breast cancer causes), but not all are applicable (e.g., nipple delay). The general cancer search phrase (i.e., "breast cancer") was used as a proxy for general breast cancer-related information seeking, and then more complex search dimensions were built based on the NCI's *Cancer Control Continuum* (CCC) (National Cancer Institute, n.d.). Using the CCC creates

conceptual continuity with the breast cancer related-event categories—celebrity-of-interests were grouped according to the dimensions of the CCC (see page 82). Search terms that conceptually mirrored concepts in each dimension of the CCC, and considered to have good face validity, were selected to retrieve Google Trends RSV data. These terms were then tested for content validity. Once content validity was established, the RSV data for each search term were downloaded.

The search terms making up a search dimension were visually inspected for potential weighting issues (as discussed in Study 1, see pages 90-96)—inflation or deflation of relative search query values due to low absolute search or a dramatic elevated outlier of search. Trend lines were added to the raw data. Data sets with trend lines contrasting with the rest of the datasets in the same dimension were treated as outliers and removed from analysis. The following description uses the initial retrieval of data for the *Treatment* dimension to illustrate this process.

Potential search topics/terms identified which might be considered a "good fit" for the *Treatment* dimension included, mastectomy, lumpectomy, chemotherapy, radiation, breast cancer surgery, and tamoxifen (a pharmaceutical hormonal therapy often used to treat some breast cancers). Initial search data retrieval for chemotherapy and radiation demonstrated poor content validity. That is, the related search queries demonstrated good content validity for "chemotherapy" but not necessarily chemotherapy related to breast cancer. This distinction is important. The top 10 related search queries were cancer, cancer chemotherapy, chemotherapy effects, chemo, side effects chemotherapy, chemotherapy treatment, after chemotherapy, what is chemotherapy, chemotherapy drugs, and radiation. In order to capture search data

relevant to chemotherapy and breast cancer, "breast cancer chemotherapy" was entered. However, that particular search string resulted in "your search doesn't have enough data to show here." Another attempt was made using "breast cancer AND chemotherapy." A similar process was followed for radiation—"breast cancer AND radiation." For both Boolean search strings, the data had high RSVs and the "related search terms" had very low percent similarity scores. These results indicate that both search phrases were rarely queried and brought into question the reliability of the data. These search terms were excluded from analysis.

Tamoxifen was also removed from analysis because the trend pattern was distinct from the remaining datasets (a striking positive trend). Mastectomy demonstrated good content validity, but due to the dramatic increase in search at the time of Jolie's disclosure in May 2013, the remaining data is scaled quite low. The inclusion of this dataset would lower a composite average score, so it too was excluded from analysis. The final dimension included the search phrases "breast cancer treatment," "lumpectomy," and "breast cancer surgery." Data retrieved from all three search terms shared visually similar trends. It should be noted that the final dimension is not void of chemotherapy, radiation, and mastectomy search query information. "Chemotherapy" and "radiation" had high percentage of related searches in the "breast cancer treatment" search query data, and "mastectomy" had a high percentage of related searches with all three search strings (see Appendix C to see the related search queries for each search term by search dimension). Similar data inspection procedures were followed for each search dimension. Figures 4 and 5 demonstrate the nature of the original sets of data and the final *Treatment* dimension used for analysis.

Figure 4. Original and Final Search Term RSV Data Included in the Treatment Search Dimension, 2005-2016

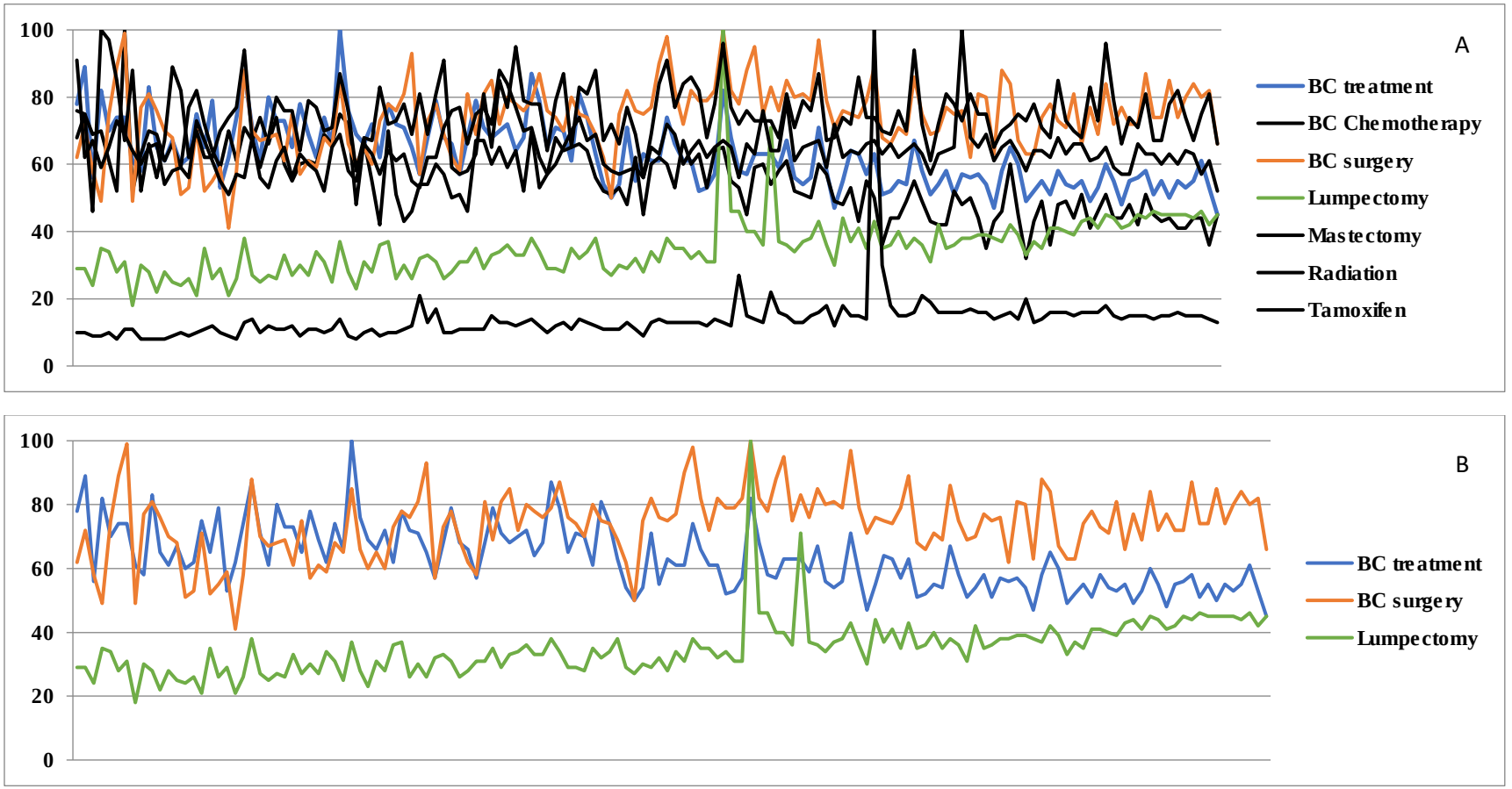


Figure 4. Figure A plots the RSV data (2005-2016) obtained for the original seven search phrases/terms for the Treatment search dimension. The black lines represent the four search phrases/terms which were removed from the final dimension. The colored lines and Figure B represent the three search phrases/terms which comprise the final Treatment search dimension.

Figure 5. Original and Final Composite Average RSV for the Treatment Search Dimension, 2005-2016

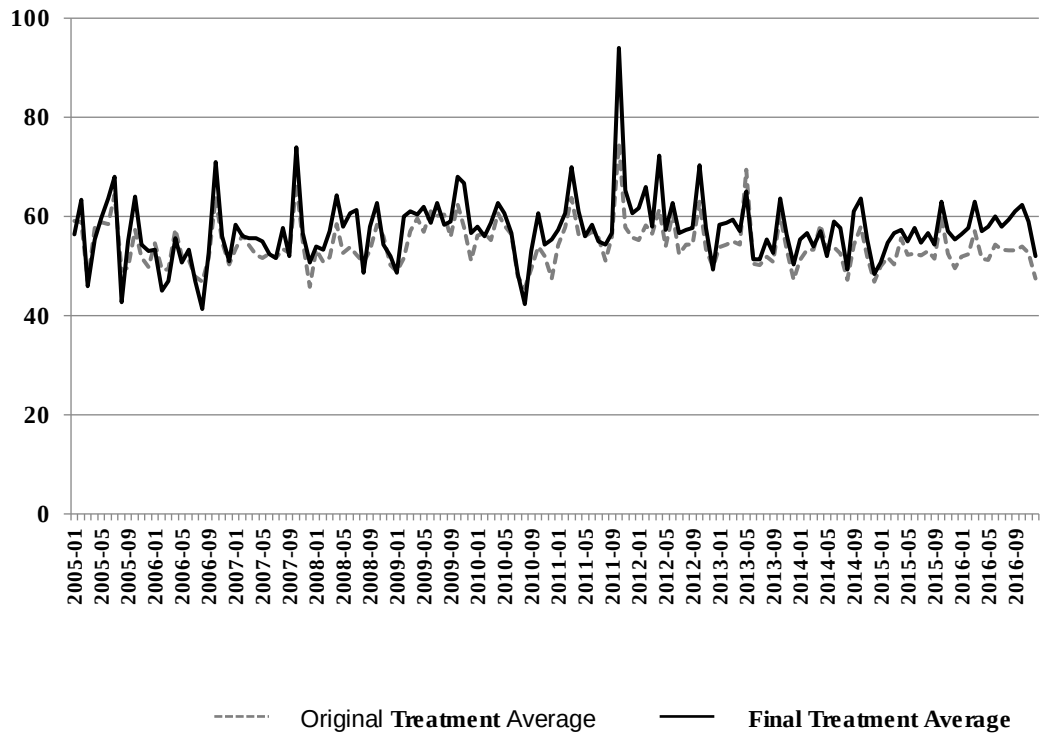


Figure 5. The plotted data demonstrates the RSV average score for the original and final *Treatment* Search Dimensions. Removing the datasets which fit the exclusion criteria elevates the final average score; the overall search pattern between the two averages remain similar.

Through this iterative process, five exclusion criteria were established:

- Outlying trend line: The trend line of the specific search phrase/term is incongruent with the other datasets for the particular search dimension.
- Poor content validity: "Related search terms" are not related to breast cancer or the specific search dimension.

- No search volume: Google Trends did not retrieve RSV scores for the search phrase/term.
- Inflated RSV: High RSV scores with no or few "related search terms."
- Deflated RSV: Due to one significant search event (e.g., Angelina Jolie), the remaining scores are significantly reduced.

The weekly and monthly RSV scores for each search phrase/term within a specific domain were averaged together. The weekly and monthly composite RSV scores were used as the final outcome variables. The following lists each domain in the *Cancer Control Continuum*, the final datasets included in each dimension for analysis, the search terms/phrases initially included, and the exclusion criteria applied. The final search dimensions, search terms, search equations, and related search queries (i.e., content validity) are included in Appendix C on pages 211-224.

Prevention. The final dimension is comprised of RSV data from the search terms/phrases: Breast cancer risk(s); Breast cancer cause(s); Breast cancer prevention; and BRCA. Data for BRCA(1)(2) and Breast cancer causes were downloaded but not included because of outlying trend lines or inflated RSV scores. No search volume was available for "breast cancer AND lifestyle" or related iterations.

Detection. The final dimension is comprised of RSV data from the search terms/phrases: Breast cancer detection; Mammography; Breast cancer screening; and Breast cancer symptoms. The following search terms/phrases were considered for inclusion in the dimension, but either no search volume or deflated RSVs excluded these terms from analysis: Breast cancer imaging; Breast MRI; Clinical breast exam; Breast cancer symptoms; Breast self exam; Breast lump; and Dense breasts.

Diagnosis. The final dimension is comprised of RSV data from the search terms/phrases: Breast biopsy; Breast cancer stages; and Breast cancer type. Outlying trend lines or no search volume excluded the following from analysis: DCIS; Triple negative breast cancer; Metastatic breast cancer; and Breast cancer spread.

Treatment. The final dimension is comprised of RSV data for the search terms/phrases: Breast cancer treatment(s); Lumpectomy; and Breast cancer surgery. Chemotherapy and radiation were excluded from analysis due to poor content validity. Mastectomy and Tamoxifen were excluded due to inflated RSV scores and outlying trend lines, respectively.

Survivorship. The final dimension is comprised of RSV data from the search terms/phrases: Surviving breast cancer; Breast cancer recurrence; and Breast cancer reconstruction. Breast cancer treatment side effects had poor content validity. No search volume was available for breast cancer AND palliative treatment or breast cancer AND pain management, and breast cancer treatment side effect had inflated RSV scores, excluding these from further analysis.

Analytic Approach

The hypotheses for Study 2 (H2; H2a; and H3) test the effects of news coverage of celebrity breast cancer disclosures on the public's level of online breast cancer-related information seeking. The hypothesized causal relationships can only be supported if: 1) there is co-variation between media coverage of celebrity breast cancer disclosures and breast cancer-related information seeking outcomes; 2) the media coverage precedes changes in the information seeking outcomes; and 3) the relationship between media coverage and the information seeking outcomes cannot be accounted for by a third

variable. Cross correlation tests and regression analysis can address points 1 and 2 from above. Accounting for other possible explanatory variables is a more difficult task, particularly given the aggregate level variables used in this study. However, because the independent and dependent variables are measured across time, time is a variable that must be controlled. Once time is controlled, statistical evidence indicates that media coverage is associated with changes in breast cancer-related online information seeking during the span of this study—not just during one or two disclosure events—then the likelihood of spurious causality is diminished.

To address the hypotheses and the issues of causality, and accounting for the impact of time, time series analysis methods were employed. Time series analysis is a broad term to describe varying methods and models where much of the methodology is aimed at using and correcting autocorrelated data (i.e., past values influence future values) and explaining this correlation, unlike typical cross correlations or regression analysis tests. Lagged endogenous variables (i.e., variables derived by the model and are explained by relationships among the model's functions) are added in the model to remove the correlation between the dataset (i.e., independent and dependent variables) at time (t) t -1, t -2 and so forth, controlling for time (Cowpertwait & Metcalfe, 2009; Shumway & Stoffer, 2010). In addition, methods in time series analysis help to remove seasonal trends, if any exist. Initial visual inspection of the general breast cancer information seeking domain clearly shows a peak in October each year (see Figure 6). This peak can likely be attributed to October's status as Breast Cancer Awareness month. Methods such as differencing in time series analysis can remove this seasonal trend from analysis. For example, if there is a seasonal trend at a weekly level, the trend is removed

by subtracting the value from the previous week. See page 146 for a discussion on the procedures taken to difference the monthly-level Breast Cancer (2005-2016) data series.

Time series analysis consists of a broad set of methods and models, but can roughly be grouped into three broad classes: autoregressive (AR), moving average (MA), and integrated models (I). An AR model regresses values on previous values from the same time series. MA models use previous forecast errors in a predictive model (regression-like). And I models apply finite differences of data points to non-stationary data (Cowpertwaid & Metcalfe, 2009; Shumway & Stoffer, 2010). Exploratory analysis is necessary to determine the appropriate methods used for the final analyses. Exploratory analysis and subsequent modeling was conducted for each search domain data set. All analyses were conducted in RStudio 3.1.1.

Visual inspection. The raw composite average of each search domain was initially plotted with the celebrity breast cancer disclosure overlaid at the respective time point. This was done to visually explore the possibility that the celebrity disclosures impact search volume. Exemplars (i.e., Breast Cancer, 2005-2016 [monthly data]; Breast Cancer, 2005-2008 [weekly data]; Diagnosis dimension, 2005-2016 [monthly data]; and Diagnosis dimension, 2005-2008 [weekly data]) of this approach are located in Figures 6 and 7, respectively. Visual inspection indicates that there are clear instances where celebrity disclosures occur concurrently with increased volumes of search. The inverse is just as clearly illustrated, making the case for further and more sophisticated analyses.

The first step to explore time series data in RStudio is to convert each data series into a time series object. For univariate and multivariate series with regularly spaced calendar time series data, the time series object allows for further analysis of the dataset

Figure 6. Celebrity Breast Cancer Disclosures and the Breast Cancer Search Dimension, 2005-2016; 2005-2008

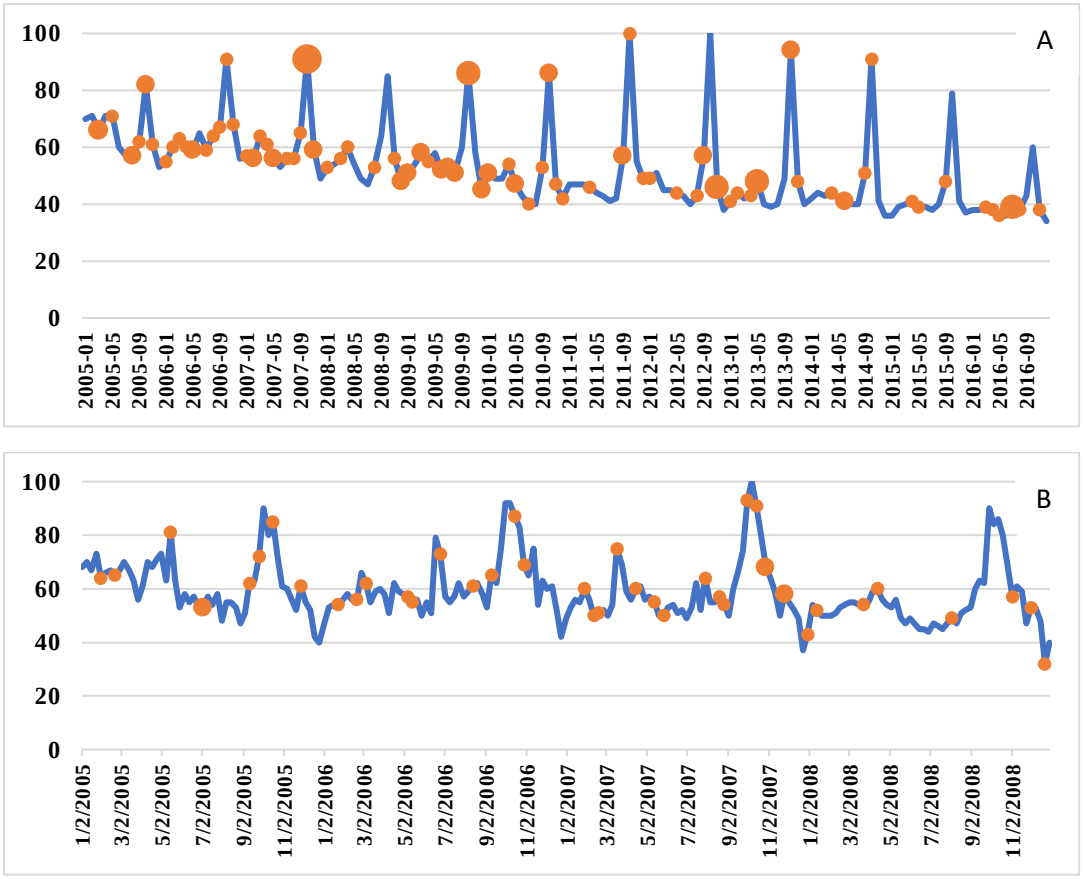


Figure 6. Figure A plots the Breast Cancer Dimension composite scores from 2005-2016 (monthly-level data). The orange dots represent celebrity breast cancer disclosures at the time of disclosure. The greater the size of the orange dot, the greater the number of disclosures. The number of disclosures ranged from 0 to 4 per month. Figure B plots the same type of information, but is specific to the Breast Cancer Dimension, 2005-2008 (weekly-level data). The number of disclosures for the weekly-level data range from 0 to 3 per week.

as a whole, rather than treating time points as individual cases. After data are converted into a time series object, the next step in time series analysis is to create a time series visual, or figure (Cowpertwait & Metcalfe, 2009). With the visualization, determining any kind of trend, seasonality, or random behavior is of concern. Three issues to look for

Figure 7. Celebrity Breast Cancer Disclosures and the Diagnosis Search Dimension, 2005-2016; 2005-2008

Figure 7. Figure A plots the Diagnosis Dimension composite scores from 2005-2016 (monthly-level data). The orange dots represent celebrity breast cancer disclosures at the time of disclosure. The greater the size of the orange dot, the greater the number of disclosures. The number of disclosures ranged from 0 to 4 per month. Figure B plots the same type of information, but is specific to the Diagnosis Dimension, 2005-2008 (weekly-level data). The number of disclosures for the weekly-level data range from 0 to 3 to week.

in the initial visualization are: 1) a constant mean, rather than a function of time; 2) homoscedasticity (variance should not be a function of time); and 3) spread should not be a function of time (covariance) (Cowpertwait & Metcalfe, 2009; Shumway & Stoffer, 2010).

These initial visualizations were conducted for the 24 sets of time series data (four data series [2005-2016; 2005-2008; 2009-2012; 2013-2016] * 6 search domains). The

results will be reported for each multivariate time series, but for illustrative purposes this section will provide the specific analytical steps taken for the general Breast Cancer Dimension, 2005-2016. First, the monthly total volume of media reports of celebrity breast cancer disclosures and the monthly RSVs for Breast Cancer were converted into time series objects. RStudio has a decompose function which estimates and plots the series' trend, seasonal trends, and the error terms of the series using a moving average method (see Figure 8 on p. 144) (Cowpertwait & Metcalfe, 2009). This process begins to address the three issues raised on pages 141-142 (i.e., constant mean, homoscedasticity, covariance).

Seasonal trends, stationarity, and normal distribution. Figure 8 visually indicates seasonal trends for both sets of data, a stochastic trend (i.e., stationary) for the volume of media reports (IV), and a downward trend for the Breast Cancer RSV (DV). Seasonal trends will be removed in the model fitting process, but the overall trends must first be statistically verified to determine if the series are stationary (i.e., mean, variance, and autocorrelation are constant over time). To do this, parametric or nonparametric tests can be done. In order to know which of these tests to run, each series needs to be assessed for normal distribution (Cowpertwait & Metcalfe, 2009; Shumway & Stoffer, 2010). To visually discern distribution, each series was fitted to a Q-Q plot. A Q-Q plot demonstrates what proportion of the series compares with the theoretical proportion of the sample's mean and standard deviation of the normal distribution model (Shumway & Stoffer, 2010). The Q-Q plots for the IV and DV are offered in Figure 9. The Q-Q plots indicate that both series are not normally distributed. To offer rigorous statistical

Figure 8. Decomposition of the Monthly Volume of Media Reports and the Breast Cancer Search Dimension 2005-2016 Time Series

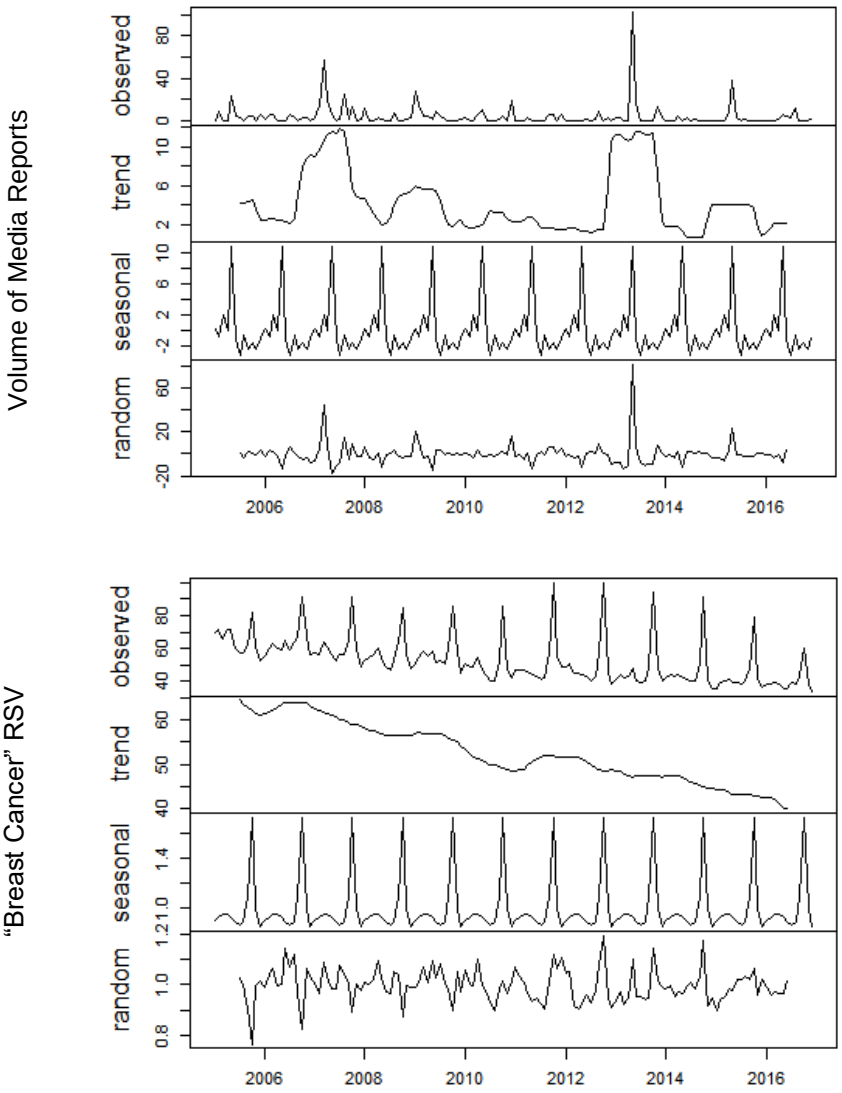


Figure 8. Both figures plot the observed time series objects for the monthly volume of media reports (IV) and Google Trends RSV scores for the Breast Cancer Dimension (DV). The remaining plots decompose the yearly trends (moving average), seasonal effects, and error terms.

Figure 9. Q-Q Plots for the Volume of Media Reports and "Breast Cancer" RSV 2005-2016 Time Series

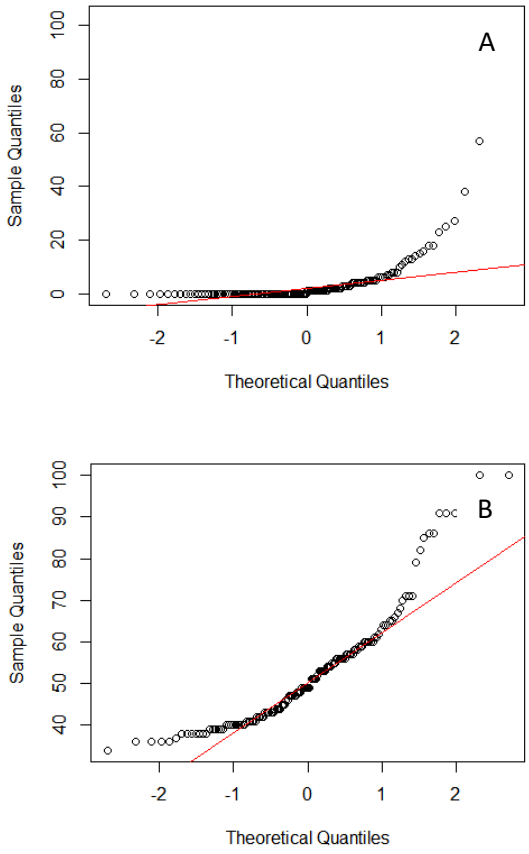


Figure 9. The original raw data (N = 144) for the volume of media reports (IV) (A) and Breast Cancer RSV (DV) (B) as a Q-Q plot. The IV follows a normal distribution except in the extreme tails. The DV series is non-normal on visual inspection.

evidence, Anderson Darling normality tests were conducted to verify the Q-Q plots (Anderson & Darling, 1954; Thode, 2002). The results of the normality test reject the hypothesis that the series are normally distributed (IV, A = 26.990, p < .001; DV, A = 4.698, p < .001).

Because neither series is normally distributed, the non-parametric test seasonal Mann-Kendall was conducted to offer evidence for trends of the 12-year period (Nghiem,

Papworth, Lim, & Carrasco, 2016; Yue, Pilon, & Cavadias, 2002). A Kendall's tau statistic τ provides evidence for the null hypothesis of no trend. The Kendall's tau is a measure of the association between two samples, which are based on ranks with the samples. As Nghiem, Papworth, Lim, and Carrasco (2016) describe, "τ is calculated as:

$$\tau = \frac{S}{1/2\, n(n-1)}$$

S is the subtraction of the discordant pairs ($x_j < x_k$ for $j>k$; x denotes variable of study and *j* and *k* denote current and future points in time) from the number of concordant ($x_j > x_k$ for $j>k$) pairs across all possible pairs in the *n* observations in the time series" (p. 3). Both the volume of media reports (IV) and "Breast Cancer" RSV (DV) series had negative tau's (tau = -.0162, $p = .009$; tau = -.591, $p < .001$, respectively). A negative τ indicates a downward trend (whereas a positive τ indicates an upward trend). The Kendall's tau statistic is reported for each data series in the results section (see Table 18, p. 152).

Cross correlations. To determine if the series are correlated in the hypothesized direction—volume of media reports as the X series and RSV as the Y series—a sample cross correlation function (CCF) was performed. However, prior to the CCF the time series must be stationary. A stationary series means that each series' mean, variance, and autocorrelations are unconnected to time (Cowpertwait & Metcalfe, 2009; Nghiem et al., 2016; Shumway & Stoffer, 2010). Both the original exploratory plots (see Figure 8) and the Mann-Kendall test demonstrated time trends in the data, making the series non-stationary. Differencing the data removes these time dependent trends, making the data stationary. First ordered differences (i.e., removing the time trends) were conducted for both series (see Figure 10). However, the transformation process introduces more noise to

Figure 10. Plotted Original Time Series and Differenced Data: Volume of Media Reports and Breast Cancer RSV, 2005-2016

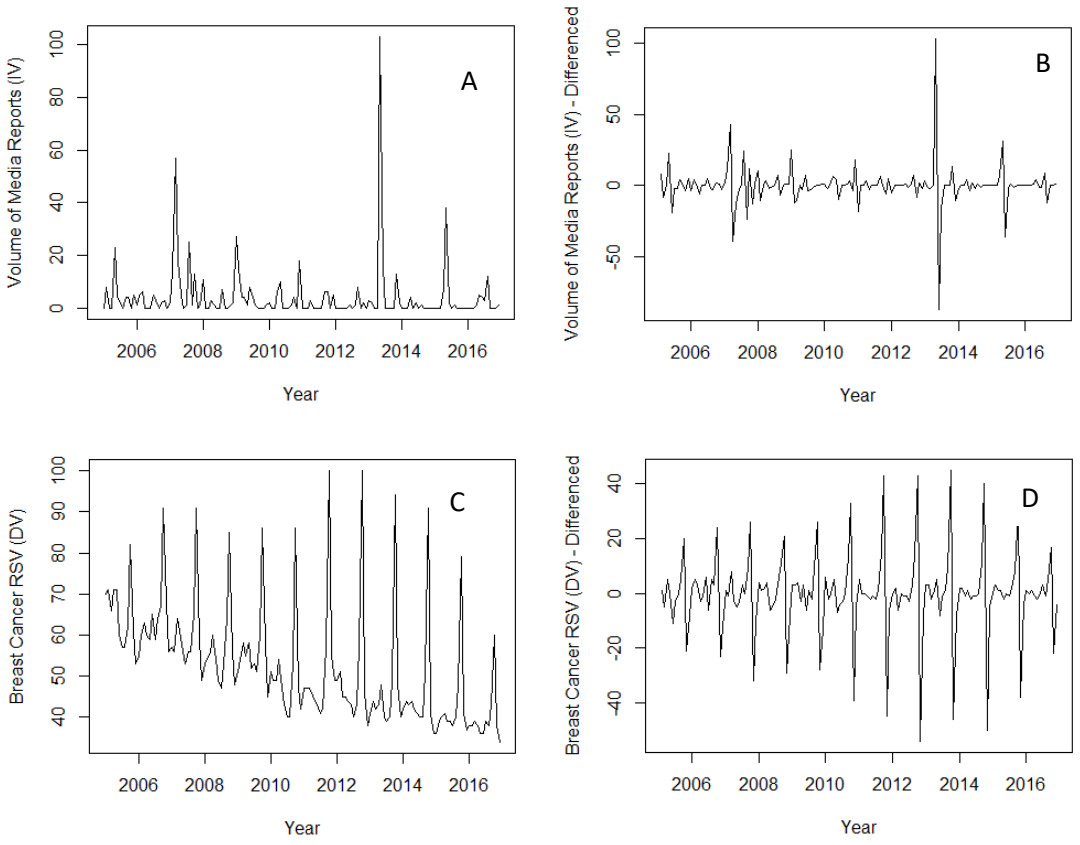


Figure 10. Plots A and C represent the original time series data. Plots B and D represent the differenced series. Time trends are removed from both series, making the data stationary.

the series. Partial least squares or prewhitening (inverse linear regression) reduces the presence of information not relevant to prediction (separates the time series from its own autocorrelation). The prewhitenting process yields CCF values for different lags. (A lag is a later point in the time series data. Time is measured at point T than T+1 is one lag later.) The lag k (k is the order of the test) value yielded by the cross correlation function

Figure 11. Cross Correlation Function Results for Volume of Media Reports and Breast Cancer RSV, 2005-2016

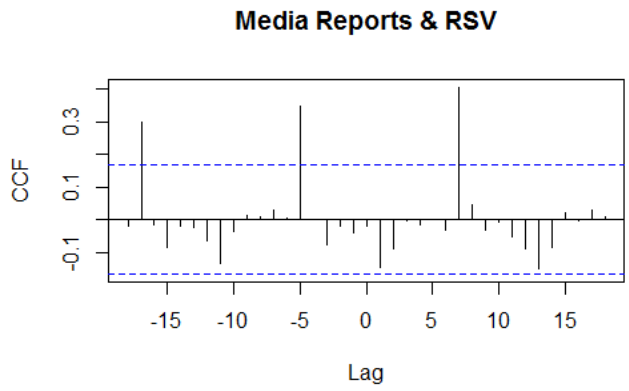


Figure 11. The cross correlation function offers evidence of potential time periods where *X* (media reports) will predict *Y* (RSV). Significant lags are found at -17, -5, and 7. The first significant lag, -17, indicates that x leads y at 17 lags. However, the small CCF value indicates the effect may be small, whereas a value closer to 0 would indicate a larger effect.

(ccf[x, y]) approximates the correlation between x[t+k] and y[t] (Anderegg & Goldsmith, 2014; Nghiem et al., 2016; Probst, Stelzenmüller, & Fock, 2012; Shumway & Stoffer, 2010). The resulting CCF (see Figure 11) shows positive CCF values at negative lags, indicating that *x* leads *y*. That is, *x* (IV) predicts *y* (DV). These procedures were followed for each time series.

Time series regression. As the previous discussion indicated, the data series are autocorrelated. Standard regression of these data would inflate statistical significance. Time series regression models correct for this potential (Cowpertwait & Metcalfe, 2009; Shumway & Stoffer, 2010; Stryker, 2003). For the main effects models (RSV will significantly increase proximal to media-related IVs), first, each series was fitted to a vector autoregression model (VAR). A VAR model (an autoregressive [AR] model) is commonly used for multivariate time series and confines the series' linear

interdependencies—each variable is explained by its past values and other variables' past values in the series system (Cowpertwait & Metcalfe, 2009). To account for the seasonal component for specific search dimension RSV data series, the data were seasonally decomposed prior to model fitting (see Figure 8 on pg. 144).

To determine if the IV (i.e., volume of media reports) causally predicts the DV (i.e., RSV), Granger causality tests were performed on all VAR models. Granger causality is not considered to be a test which determines true causality, but instead provides evidence of one variable preceding the other (Barnett & Seth, 2014; Shumway & Stoffer, 2010; Stryker, 2003). That is, time series X Granger-causes Y if a series of F-tests on lagged values of X and Y demonstrate that the X values provide statistically significant information about the future values of Y. The Granger causality test tests two competing hypotheses. For example: 1) Media reports do not Granger-cause RSV; and 2) RSV do not Granger-cause media reports. A statistically significant result ($p<.05$) rejects the null hypothesis (Barnett & Seth, 2014; Shumway & Stoffer, 2010). Significant lags identified in the CCF are used in the VAR model to determine time order of effects. The positive CCF values at negative lags already indicate that *x* leads *y*, but testing the competing hypotheses is required for Granger causality and serves to verity the CCF results. All 24 (four time periods * 6 search dimensions) bivariate models tested these competing hypotheses.

Moderation. To explore potential moderating effects based on celebrity attributes (i.e., career type; age; and level of celebrity status) and breast cancer-related event type, multiple linear regression analysis was conducted. The regression models were first fit with the data series fitted for the VAR models (detrended and seasonally adjusted). Then,

a model for each time period and search dimension was tested using a categorical variable of one of the four celebrity attribute variables (i.e., age, career type, breast cancer-related event type, and level of celebrity status). More than one celebrity breast cancer disclosure occur at certain time points (weeks or months), which means that the levels of the given variable were inconsistent at that point in time. For this analysis, the categorical levels pertaining to the celebrity disclosure with the largest volume of media reports for that time point were chosen for each moderator. For example, Angelina Jolie and Zorida Sambolin both announced a breast cancer-related event on May 13, 2013. One-hundred-and-seventeen media reports regarding Jolie's disclosure were retrieved, whereas three were for Sambolin. Both women are in the same age category, but are in different breast cancer-related event type, career type, and level of celebrity status categories. Because Jolie's disclosure dominated this date (and week and month), the categories for each moderator which corresponded with Jolie were used, rather than Sambolin.

This analytical approach is somewhat liberal—linear modeling may overestimate the effects of the model, including the moderators (Cowperwait & Metcalfe, 2009; Stryker, 2003). However, because such a large number of models are included for analysis, the somewhat limited findings from the Granger causality tests (see the Results section), and the constraints in time series analysis with categorical variables (particularly non-dichotomous variables—the four categorical variables range in four to nine levels), linear modeling was chosen in order to begin to understand any potential influence from these four celebrity attribute categories. Statistically significant results from these

analyses cannot be considered strong evidence of moderation, but instead signal which variables should be considered for future research.

Results

Trends analysis. Downward trends for the volume of media reports were found across all four time periods (2005-2016; 2005-2008; 2009-2012; 2013-2016). However, only the monthly-level datasets and the 2009-2012 weekly-level data had negative trends that were statistically significant. These results indicate that overall media reports of celebrity breast cancer disclosures have decreased over time, and the particular chunks of time studied vary in volume. These findings are consistent with the descriptive results in Study 1.

Of the 24 RSV series, 16 had downward trends and eight had upward trends. Seventeen of the series had statistically significant trends. All six of the 2005-2016 series had statistically significant trends (four downwards and two upwards). Statistically significant weekly-level datasets were more sporadic. These results may be attributed to the arbitrary four-year time periods selected—affecting the weighting of the data. Using the monthly-level data series, it can be concluded that the public's information seeking related to general breast cancer, prevention, detection, and survivorship topics had decreased in volume over the 2005 to 2016 time period, whereas information related to breast cancer diagnosis and treatment had increased during the same time period. The results of the seasonal Mann-Kendall trend tests for both the media and breast cancer-related information seeking outcomes are available in Table 18.

Table 18

Results of Seasonal Mann-Kendall Trend Test on Media Reports and RSV of Search Dimensions

	Time-Period			
Variables	2005-2016	2005-2008	2009-2012	2013-2016
IV				
Media Reports	-.16 (.01)*	-.01 (.86)	-.15 (.01)*	-.02 (.78)
DV				
Breast Cancer	-.59 (<.001)**	-.24 (<.001)**	-.22 (<.001)**	-.30 (<.001)**
Prevention	-.13 (.02)*	-.01 (.91)	-.10 (0.03)*	-.15 (.002)*
Detection	-.18 (<.001)**	-.20 (<.001)**	-.26 (<.001)**	.08 (.08)
Diagnosis	.24 (<.001)**	.20 (<.001)**	-.06 (.20)	.18 (<.001)**
Treatment	2.04 (<.001)**	.07 (<.14)	.01 (.82)	.08 (.08)
Survivorship	-.21 (<.001)**	-.21 (<.001)**	-.03 (.47)	-.20 (<.001)**

Note. The Kendall's tau statistic and 2-sided p-value are reported. Positive values indicate an upward trend and negative values indicate a downward trend.
* indicates statistically significant trends <.05; ** indicates statistically significant trends <.001.

Hypothesis 2. H2 predicts that relative search volume for breast cancer–related search queries will have significant increases proximal to media coverage of the celebrity breast cancer disclosure. To test this hypothesis, the CCF test determined lags of explanatory variables that have potential to predict search volume. All significant lags for each series' pairs are reported in Table 19. Four models had no significant lags; no further tests were conducted. For the remaining 20 models, the significant lags were entered into the VAR model. Lags that would indicate volume of media coverage have the potential to influence information seeking behaviors are positive CCF values with negative lags (*x* leads *y*). Half of the models had such CCF values and lags.

Table 19

Results of CCF Tests

Search Domain	2005-2016	2005-2008	2009-2012	2013-2016
Breast Cancer	lag = -17 (0.30) lag = -5 (0.348)	lag = 0 (1.055)	lag = -4 (0.190)	lag = -6 (0.106)
Prevention	lag = -6 (0.335)	lag = -8(0.206)	lag = 7 (0.166)	lag = -17(-0.233)
Detection	lag = -17 (-.219) lag = -5 (.206)	--	lag = 17(-0.181)	lag = -20 (0.200)
Diagnosis	lag = -17(0.263) lag = -5 (0.307)	--	lag = -2(0.252)	lag = -20 (0.216) lag = -5 (-0.143)
Treatment	--	lag = 0 (0.166)	lag = -8 (-0.145) lag = -5 (0.123)	lag = -20 (0.154) lag = 0 (0.209)
Survivorship	lag = -17 (0.194)	--	lag = 19(0.143)	lag = -5 (0.230)

Note. Reported are the lags that entered each model. Coefficients estimates are located in the brackets. All lags reported are statistically significant (<.05).

Negative lags that are large in value (e.g., 0, 1, or 2)—or at least close to 0—would indicate that the potential causal relationship is proximal. Most of the monthly-level data models have significant lags with relatively small values (-17 to -6), indicating that the media reports on the breast cancer-related information seeking outcomes was not immediate or proximal—each lag represents approximately one month—and instead offers evidence of more cumulative effects. The Granger causality tests demonstrated statistically significant results which reject the first competing Granger causality hypothesis (i.e., Media reports do not Granger-cause RSV) and confirm the second competing hypothesis (i.e., RSV do not Granger-cause media reports) for the Breast Cancer, Prevention, and Diagnosis (2005-2016) search domains (see Table 20). These results indicate that media coverage of celebrity breast cancer disclosures do impact breast cancer-related information seeking for these domains, but again, the effects are not proximal. The weekly RSV series that had statistically significant Granger causality

Table 20

Results of Granger Causality Tests for Bivariate VAR Models

Search Domain	Media Reports			RSV		
(Time-period)	*df*	*F*	*P*	*df*	*F*	*p*
2005-2016						
Breast Cancer	5, 254	3.57	<.05*	5, 254	.25	.94
Prevention	6, 248	*3.21*	<.05*	5, 254	.66	.66
Detection	6, 248	.96	.45	6, 248	.71	.64
Diagnosis	7, 242	2.21	.04*	7, 242	1.75	.10
Treatment	-	-	-	-	-	-
Survivorship	6, 248	.30	.94	6, 248	1.20	.31
2005-2008						
Breast Cancer	1, 408	.453	.50	1, 408	2.17	.14
Prevention	2, 404	3.30	.04*	2, 404	2.34	.10
Detection	-	-	-	-	-	-
Diagnosis	-	-	-	-	-	-
Treatment	1, 410	2.17	.14	1, 410	.12	.73
Survivorship	-	-	-	-	-	-
2009-2012						
Breast Cancer	4, 390	3.01	.02*	4, 390	.84	.50
Prevention	7, 372	1.10	.36	7, 372	2.78	<.05*
Detection	3, 396	.38	.77	3, 396	1.57	.20
Diagnosis	8, 366	3.05	<.05*	8, 366	.41	.91
Treatment	3, 396	.69	.56	3, 396	1.22	.30
Survivorship	3, 396	1.39	.25	3, 396	.67	.57
2013-2016						
Breast Cancer	4, 388	3.54	<.05*	4, 388	.17	.95
Prevention	2, 400	.20	.82	2, 400	.21	.81
Detection	4, 390	.73	.57	4, 390	.08	.99
Diagnosis	3, 394	1.61	.19	3, 394	.58	.63
Treatment	4, 390	1.88	.11	4, 390	.45	.78
Survivorship	1, 408	.01	.92	1, 408	0	.10

Note. The Media Reports results are from the first competing hypothesis (i.e., media reports do not Granger-cause RSV) tested by the Granger causality test. The RSV results represent the results from the second competing hypothesis (i.e., RSV do not Granger-cause media reports).

* indicates statistically significant results

results (Prevention [2005-2008]; Breast Cancer [2009-2012]; Diagnosis [2009-2012]; Breast Cancer [2013-2016]) for the first competing hypotheses (i.e., Media reports do not Granger-cause RSV)—thus rejecting it—and confirmed the second competing hypothesis (i.e., RSV do not Granger-cause media reports) had considerably shorter lags/time periods between the media reports leading the breast cancer-related information seeking outcomes. The Diagnosis (2009-2012) RSV scores had only a lag of -2 (i.e., approximately two weeks), Breast Cancer (2009-2012) a lag of -4, Breast Cancer (2013-2016) a lag of -6, and Prevention (2005-2009) a lag of -8. Granger causality was not found for the remaining 14 weekly-level VAR models. Some of the VAR models have negative lags at negative CCF values; however, only the Prevention (2009-2012) model has a statistically significant result confirming the first competing hypothesis (i.e., Media reports do not Granger-cause RSV) and rejecting the second competing hypothesis (i.e., RSV do not Granger-cause media reports). This Granger causality result and the negative lag with a negative CCF value indicates that Prevention-related search volume effects the number of media reports (a reverse causal relationship).

The results in total indicate partial support for H2—the search domains of general breast cancer, prevention, and diagnosis (generally across time periods) were positively impacted by media coverage of celebrity breast cancer disclosures (with the exception of Prevention [2009-2012] which an inverse causal relationship), but in most cases the impact was small (see discussion on CCF tests and lags) and distal rather than proximal.

Hypothesis 2a. H2a hypothesized that weeks (or months) with higher volumes of media reports on celebrity breast cancer disclosures will have greater increases in relative search volume for breast cancer-related search queries than weeks with lower numbers of

media reports. For the models where Granger causality is supported, the significant lags are generally months prior to the increase in search volume (with weekly-level data, several weeks and up to two months). Because the impact of media reports is not proximal to search volume increase, there is no evidence to support H2a.

Hypothesis 3. H3 hypothesized that weeks (or months) with higher proportions of media reports with episodic frames would see greater volumes of search than weeks (or months) with higher proportions of media reports with thematic frames. The same issues with H2a—no evidence to support proximal information seeking behaviors—are at play for H3 as well. In addition, the volume of the two types of frames were highly disproportionate (episodic frames [n = 476]; thematic frames [n = 119]). No months had higher rates of thematic frames, and only one week of the 624 weeks in the weekly-level data had a higher rate of thematic frames than episodic frames (see Figure 12). There is not enough power to test the hypothesis.

Research Question 3. RQ3 explores potential moderation by the celebrity's a) age; b) career type; c) breast cancer-related event type; and d) level of celebrity status in the hypothesized causal relationship between media reports and search query volume. Even with the tendency for liberal results with linear-regression, very few significant results were found and the results are highly inconsistent. That is, the impact of the celebrity attribute or breast cancer-related event type is dependent on the search dimension and the time period. The following offers a summary of the results based on attributes. (All statistically significant results are offered in-text. In total, 96 models were tested [24 bivariate models * 4 categorical variables], but largely with insignificant

Figure 12. Frequency of Episodic and Thematic Frames: Monthly- and Weekly-Level Measures

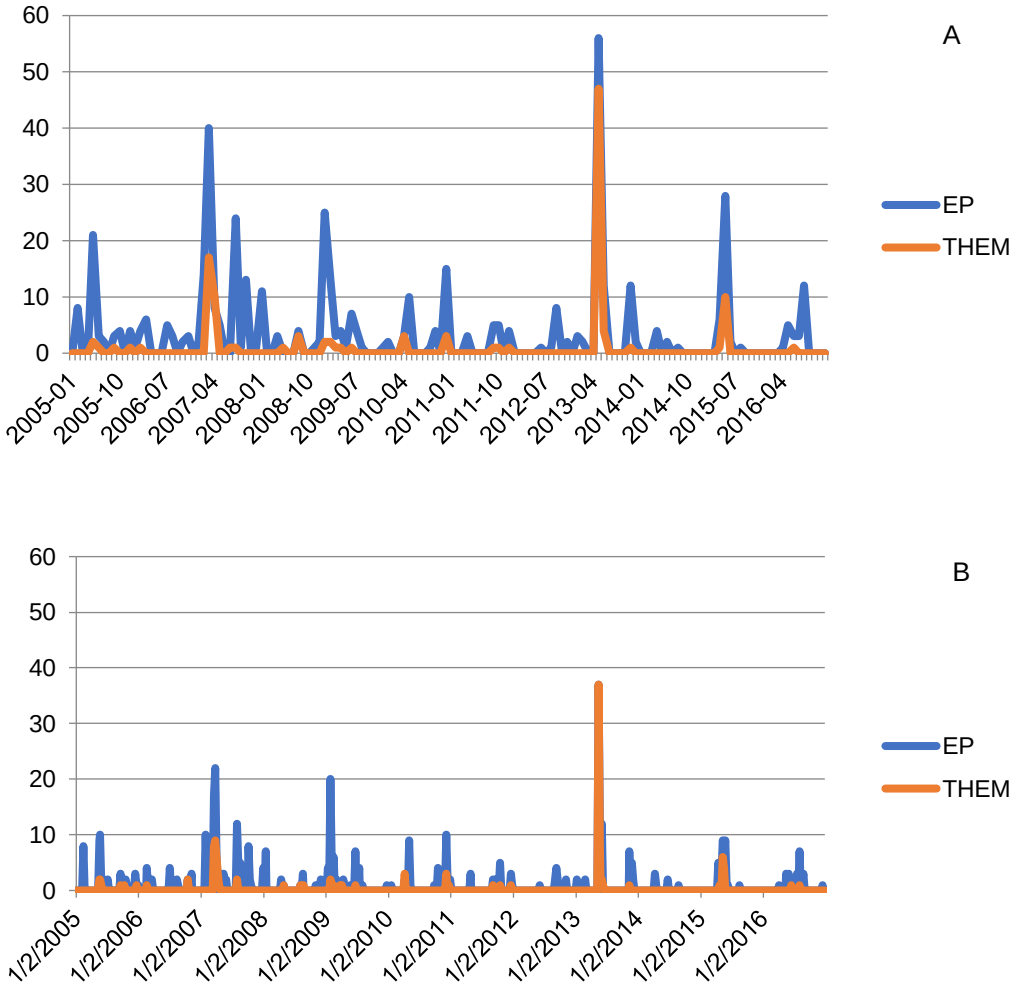


Figure 12. Figure A plots the volume of media reports with episodic and thematic for the monthly-level data from 2005-2016. Figure B also plots the volume of media reports with episodic and thematic frames, but with the weekly-level data from 2005-2016.

results. Providing a table of all of the results is a voluminous task. Results of specific models are available on request.)

Age. Only five of the models demonstrated moderation. The third level (i.e., Under 50 after 2010) of the four age group levels positively impacted RSV scores for the

Breast Cancer (2009-2012) dimension (β = 4.32, p = .023) and negatively impacted scores for the Detection (2013-2016) dimension (β = -3.95, p = .03). In regards to the impact of "older" celebrities, the Over 40 (before 2010) age group had a negative moderation effect for the Detection (2005-2008) dimension (β = -2.93, p = 0.01), whereas the Over 50 group positively moderated the effect between media reports and RSV scores for the Breast Cancer (2009-2012) (β = 4.32, p = .026) and Survivorship (2009-2012) (β = 9.74, p = .01) dimensions. The 2005-2016 and 2013-2016 (other than the Detection dimension) series were not impacted by the age variables. Overall, age was largely not a significant component of the regression models, but when it was, particular age categories moderated an increase in search volume for Breast Cancer, or a decrease in search volume for Detection.

Career type. Of the 24 models tested, nine demonstrated moderation by at least one level of the career type variable. Actors moderated the relationship between media reports and RSV scores for the Breast Cancer (2009-2012) (β = 5.17, p = 2.03) and for the Survivorship (2009-2012) (β = 10.55, p = .02) domains. The Activist/Author/Academic category moderated the RSV scores for the Breast Cancer (2005-2008) information seeking domain (β = 7.71, p = .02). The Journalist/News Anchor category moderated the RSV scores for the Diagnosis (2013-2016) dimension. The musician category moderated the results in the Survivorship (2009-2012) domain (β = 23.39, p = .02). Personal Affiliation was associated with positive moderation effects in the Prevention (2005-2008) domain (β = 10.87, p = .03) and Breast Cancer (2009-2012) domain (β =8.13, p = .02), but had a negative effect on the RSV scores for the Detection (2009-2012) (β = -6.37, p = .05) and Diagnosis (2009-2012) (β = -9.54, p = .004).

Television Personalities also had disparate directional moderating effects based on search domain. It was associated with a positive effect for the Treatment (2009-2012) dimension (β = 13.44, p = .01) and a negative effect for the Breast Cancer (2009-2012) domain (β = -9.96, p = .02). No moderation effect was found for celebrities in the Athlete/Sports and Politician/Policy Maker/Service Person categories. In addition, the monthly-level data series were not moderated by any level of career type, and of the 2013-2016 data series, only the Diagnosis dimension saw any moderation by career type. Overall, Personal Affiliation was the most common career type to moderate the relationship between media reports and RSV scores, although directional impact was split.

Cancer-related event type. Of the 24 models which tested moderation effects of the celebrity's cancer-related event, only six models had significant results. Potential moderating effects from Treatment disclosures were found in four models: Prevention (2005-2016) (β = 5.96, p = .04); Detection (2005-2008) (β = -3.50, p = .01); Treatment (2009-2012) (β = 4.94, p = .02); and Treatment (2013-2016) (β = -2.74, p = .01). Other significant results included moderation from a Diagnosis disclosure on the Diagnosis (2013-2016) dimension (β = -11.47, p = .02) and Prevention on Detection (2013-2016) (β = -9.89, p = .03) and Treatment (2013-2016) (β = -8.69, p = .03). In general, where cancer-related events moderate the relationship between media reports and RSV scores, the impact is negative. That is, as media reports increase in number for a particular breast cancer-related event type (i.e., Treatment, Diagnosis, and Prevention), relative search volume decreases.

Level of celebrity status. Unlike the other attribute moderators, level of celebrity status primarily had moderating effects on the 2005-2016 data series. The Level 6

celebrity disclosures negatively moderated the Diagnosis (β = -4.39, p = .05) and Treatment (β = -7.95, p = .001) dimensions. Media reports of the Level 9 celebrity disclosures also negatively moderated the Breast Cancer (β = -9.490, p = .01) and Diagnosis (β = -5.362, p = .03) dimensions. Level 10 celebrity disclosures positively moderated the relationship between media reports and RSV scores for the Breast Cancer (β = 7.05, p = .01) and Survivorship dimensions (β = 4.95, p = 0.01). These results indicate that only the most famous (Level 10) prompted an increase in online breast cancer-related information seeking, and for Levels 6 and 9, their disclosures coincided with a decrease in online search queries.

Summary of Study 2 Findings

Study 2 tests: 1) the relationship between media coverage of celebrity breast cancer disclosures and the public's online breast cancer-related information seeking; and 2) the moderating potential of celebrity attributes (i.e., age, career type, breast cancer-related event type, and level of celebrity status).

Table 21 (see p. 161) presents a summary of significant models and the correspondent significant lags. The results of the main effects models were inconsistent—seven of the 24 VAR models had statistically significant findings. Of the monthly-level data series (2005-2016), three of the six models were statistically significant, indicating that the media reports do affect the level of breast cancer-related information seeking. However, the influence of news coverage on information seeking is not proximal, and instead the impact is seen months to over a year later. Only four (i.e., Prevention [2005-2008]; Breast Cancer [2009-2012]; Diagnosis [2009-2012]; Breast Cancer [2013-2016]) of the 18 weekly-level data series models (2005-2008; 2009-2012; 2013-2016) had

Table 21

Summary of Statistically Significant VAR Models and Corresponding Lags

Significant Search Dimension	Significant Lag	Interpretation of Effect
Monthly-level		
Breast Cancer, 2005-2016	-17	Distal
Prevention, 2005-2016	-6	Distal
Diagnosis, 2005-2016	-17	Distal
Weekly-level		
Prevention, 2005-2088	-8	Medial
Breast Cancer, 2009-2012	-4	Proximal
Diagnosis, 2009-2012	-2	Proximal
Prevention*, 2009-2012	-6	Medial
Breast Cancer, 2013-2016	-6	Medial

* The results for this model indicate a reverse causal relationship (*y* leads *x*).

statistically significant results in the hypothesized direction. Yet, the impact of media coverage on breast cancer-related information seeking at this level is more direct than the monthly-level data series—the impact of media reports is seen two to eight weeks prior to increases in relative search volume. Possible explanations for the disparate results between the monthly- and weekly-level data series are discussed in Chapter 7 Discussion and Conclusions.

The result of the moderation analyses of the four attribute variables were highly inconsistent, based on search dimension, monthly- or weekly-level data series, and time period. The low volume of significant results limit any type of sweeping conclusions, but the categories that had the most prominent moderating effects are: 1) “older” age groups (40+ before 2010; 50+ after 2010) (Age); 2) Personal Affiliation (Career type); 3) Treatment (cancer-related event type); and 4) Level 10 (the most famous) (Level of celebrity status). Other than the level of celebrity status, the other significant moderators had mixed directional causality. Discussion of these conflicting results and the implications are included in Chapter 7 Discussion and Conclusions.

Chapter 7

Discussion and Conclusion

This project examined news coverage of celebrity breast cancer disclosures and its relationship with the public's online breast cancer-related information seeking. Additionally, the potential for the celebrities' personal attributes (i.e., age, career type, level of celebrity status) and breast cancer-related event type to explain or moderate the media and information seeking outcomes was explored. Two studies were conducted, one a content analysis of media reports of the celebrities-of-interests' breast cancer disclosures, and a study using the media outcomes as predictor variables for numerous online breast cancer-related information seeking outcomes. In the end, this dissertation offers empirical support that news coverage of celebrity breast cancer disclosures largely focus on the celebrity and her breast cancer-related disclosure (episodic frame), and breast cancer-related information beyond the celebrity's experience (thematic frame) is infrequently present in news coverage. Information which might reinforce misperceptions about breast cancer is all but absent in the coverage of these disclosures. Finally, this dissertation also provides preliminary evidence that across celebrity breast cancer disclosures, media coverage impacts some dimensions of breast cancer-related information seeking, but these effects tend to be more long-term or cumulative, rather than immediate. Although, context matters—of the six search dimensions tested, search queries were only stimulated by media coverage for the general breast cancer, prevention, and diagnosis domains.

Analyses of *who* is most likely to garner particular media and public online breast cancer-related information seeking outcomes did not yield straightforward results,

particularly for Study 2. The content analysis (Study 1), did demonstrate consistent results for age and level of celebrity status. Celebrities Under 50 did receive greater media attention than any other age group, and misperception reinforcement information was more common for this group as well. Level of celebrity status explained variation in media outcomes more than any other celebrity attribute variable. The greater the level of celebrity status, the greater the number of media reports on the celebrity's breast cancer disclosure, and a greater presence of thematic frames. Finally, while a dearth of misperception reinforcement information was present in the study's census, a third of media reports for celebrities with the highest level of celebrity status did contain misinformation.

The results for Study 2 varied greatly, making it difficult to draw overall conclusions—significant results were dependent on the search term domain and level of data (i.e., monthly- or weekly-level data). The results for the weekly-level data are highly variable, which impede meaningful conclusions. But results indicate that for the monthly-level data, celebrity status did moderate the relationship between media reports and the public's volume of breast cancer-related searches. The results indicate that celebrities who received the most media attention were also the most likely to positively impact search volume. This final chapter discusses the results of Study 1 and Study 2 in detail, and concludes with a discussion of the study's theoretical and methodological contributions and its real-world implications. Limitations and calls for future research are discussed throughout the chapter.

Study 1. Media Coverage of Celebrity Breast Cancer Disclosures

The impetus of this dissertation rested upon the assumption that celebrity breast cancer disclosures are a catalyst for news media coverage, and have been so over time. Some celebrity breast cancer disclosures result in large media events (e.g., Angelina Jolie), while others produce modest to minimal levels of reporting. Academic inquiry has all but ignored these lesser-reported disclosures. I argued that if there is consistent reporting over time, in the aggregate, the reporting of celebrity breast cancer disclosures begin to create an overarching media narrative, potentially creating or influencing, as Stryker (2003) states, "more gradual and cumulative effects of news coverage on long-term secular trends in health behavior" (p. 307). The results of Study 1 support these arguments. Media reports on two of the celebrities-of-interest (i.e., Angelina Jolie and Elizabeth Edwards) made up nearly a third of the census, and the remaining articles are the result of dozens of other breast cancer disclosures, with varying degrees of coverage. Age of the celebrity is likely a contributing factor, but level of celebrity status offers the clearest explanatory evidence for differential levels in reporting—the higher the status, the greater number of media reports of an individual's breast cancer disclosure.

Some of the disclosures under study received no media attention from the three journalistic organizations used for analysis. These omissions were found at the lowest levels of celebrity status (i.e., levels 2, 3, and 4), adding construct validity to the celebrity status index. The finding of no coverage is not reflected by the news media as a whole—the initial selection criteria of celebrities-of-interest does offer evidence that other media organizations reported on these lower-level celebrity status disclosures. However, the media salience threshold was quite low; only one media report was needed to confirm

media salience. Because all news sources on *LexisNexis* were used for this measure, the news coverage may have been from local, niche, or industry-specific publications. If hundreds of publications available through the database only result in one or two media reports of a celebrity breast cancer disclosure, then a result of zero from the combined use of the *Associated Press*, *The New York Times*, and *The Washington Post* as a proxy measure for the national news environment is certainly not invalidated.

Yet the use of these three publications can be considered a limitation. Given the study's timeframe and Stryker's (2008) validation of this approach, this national news environment proxy measure was considered the most reliable method to retrieve data. Stryker's (2008) validation study is nearly a decade old and the current media environment is arguably more fragmented. It is possible these publications are a less reliable measure of the national news environment today than they were in 2008. In addition, future research should explore the news coverage (i.e., volume and content) of entertainment and tabloid news sources. The traditional news media is not the only content provider of celebrity breast cancer disclosures. Using a more varied content sample could offer different results, but the traditional norms embedded in the journalism profession likely limit the impact of specific media organizations used for analysis (Gans, 1979; Hallin & Briggs, 2014). Research has found similarities in content across news organizations (Boczkowski & de Santos, 2007; Maiier, 2010; McCombs & Shaw, 1972), and my own data from a study in progress shows no differences in the content of reports of a celebrity of health disclosure between traditional news and entertainment organizations. However, this evidence is limited and should be further explored in the context of celebrity cancer disclosures.

Study 1 offers further evidence to support the argument that celebrity breast cancer disclosures create a persistent media narrative. The results of the framing analysis are congruent with previous research, finding that episodic frames tend to be more prevalent in general, including in stories about health-related topics (Gross, 2008; Holton et al., 2012; Holton et al., 2014; Iyengar & Simon, 1993; Matthes & Kohring, 2008). This study hypothesized that episodic frames would be found in higher proportion than thematic frames, and the results here are overwhelming. The general tendency towards reporting with an episodic frame is in line with historical precedence in celebrity news—relaying the intimate details of a celebrity's life is believed to be more enticing to readers than conveying a broader contextual message (de Leon, 2002; Lerner, 2006). Selecting to report on celebrity breast cancer disclosures reflects some core journalistic norms, particularly prominence, human interest, and timeliness (Gans, 1979; Hallin & Briggs, 2014). The sections of the newspapers in which the media reports appeared in most frequently mirror these news values: news, entertainment, style, lifestyle, and culture, and only a handful were published in health sections. From the study's descriptive results, it can be deduced that journalists and news organizations consider celebrity breast cancer disclosures to be celebrity news, and not health-related news. Future research may want to explore journalists' and news organizations' intent and approach when covering celebrity cancer and health disclosures to verify this conclusion.

Episodic frames do dominate news media coverage of celebrity breast cancer disclosures. However, by delineating media frames by level of celebrity status, the overall prevalence is put into perspective. Levels 8 through 10 see a higher rate of thematic frames than the rest of the sampled celebrities-of-interest—a two to fivefold

increase. In this upper-echelon of fame, nearly a quarter of the media reports are written with a thematic frame. Parsing this out even further, the media reports on Jolie's and Edwards's disclosures—the two standout media events—have even higher proportions of thematic frames than their status level sharing counterparts (i.e., 41% and 34%, respectively).

In general, media reports of celebrity breast cancer-related disclosures infrequently contain breast cancer information beyond that of the celebrity's own experience. Prior research has evidenced null findings for learning and individual risk assessment outcomes (e.g., Borzekowski et al., 2014). The results of the framing analysis provide some context for these findings. Given the lack of contextual breast cancer-related information, it seems unlikely that most news coverage of celebrity breast cancer disclosures provide any information that would enhance learning or personal risk assessment outcomes. However, for those celebrities in the higher tiers of celebrity status, the presence of population and subpopulation-level breast cancer-related information is more prevalent. Given that these disclosures receive elevated levels of media attention, and subsequently a greater chance of audience exposure, these findings could have important implications. For example, with the presence of breast cancer information about population risk levels or survivorship rates—among those exposed to high status disclosures—there may be justification to test for learning or individual risk assessment outcomes.

This study goes beyond prior research, and beyond the "Angelina Effect," by quantifying the presence or absence of population and sub-population breast cancer information in media reports on a large and diverse group of celebrities. Using

dichotomous frames was conducive for precision in coding and for predictive purposes for Study 2 (had there been a less disparate distribution of frames), but lack specific detail. The framing analysis gives a general sense of media coverage on celebrity breast cancer disclosures, but specific information cited such as the celebrity's age at the time of the cancer-related event, diagnosis, treatment, population risk information, screening recommendations, treatment recommendations, survivorship rates, among others, would not only provide stronger evidence on what exactly is being communicated to the public, but also offer further direction to researchers on possible effects.

In content analyses of health news coverage, there is a tendency for health communication researchers to conclude by urging journalists to use a public health model of reporting, thematic frames, or a close iteration (see Hallin & Briggs, 2014). Some experimental research demonstrates increases in knowledge or change in behavioral intentions, for example, from exposure to news with a public health angle (Coleman & Thorson, 2002; Coleman et al., 2011). However, to the best of my knowledge, experimental research testing similar effects using celebrity health disclosures has not been done. This is an important step before urging journalists to change their methods of reporting.

Limited empirical evidence from Kylie Minogue's disclosure (Chapman et al., 2005), and anecdotal evidence from Amy Robach's disclosure (Sulik, 2013) indicated the presence of misinformation about breast cancer. Findings from this study reveal that overall, misinformation is negligible in news coverage of celebrity breast cancer disclosures; all seven categories were present in less than 10% of the news coverage. Although for celebrities in the Under 50 category and the highest level of celebrity status,

the presence jumps to nearly one-third of all media reports. In part, this likely occurs due to more detailed reporting (i.e., thematic frames) for these two groups—more potential for error. The qualitative passages recorded from the media reports indicate that much of the misinformation comes from the celebrity herself. Even people who have experienced breast cancer are not necessarily more knowledgeable about the etiology of breast cancer or the science informing screening recommendations (Bickell, Weidmann, Kezhen, Lin, & Leventhal, 2009; Fagerlin et al., 2006).

Misinformation conveyed about bilateral mastectomies—the most frequently coded for category of misinformation—was common simply because the celebrity's course of treatment is not contextualized in any way. This quote in particular stands out: "Rather than risk it, Sykes, whose mother's side of the family has a history of breast cancer, opted to have them [breasts] removed." Syke's double mastectomy is presented as the best option for the celebrity, and therefore could be deduced as the right course of treatment for anyone with a similar diagnosis, or for someone with a family history. These findings are consistent with Sabel and Dal Cin's (2016). They conclude that media reporting of celebrity breast cancer disclosures demonstrate a bias that "tends to overemphasize the use of bilateral mastectomies" (p. 2800). They also found that with the rise in this bias in reporting, prophylactic bilateral mastectomies have risen at the population level as well. They reference Christina Applegate's disclosure as an exemplar of this bias. Applegate was diagnosed with DCIS, but also had a BRCA mutation and a strong family history, making her a good candidate for a bilateral mastectomy. However, most of the media reports of her disclosure did not include her genetic and family history information. Sabel and Dal Cin (2016) conclude: "By failing to highlight her elevated

risk of a second breast cancer and the importance of her genetic profile in the decision-making process, this produces another bias [by the reader]" (p. 2800). Future research should explore this potential. When a woman reads a statement such as the one about Syke's decision, how does she apply that to her own treatment decision making process, for example?

Study 2. Breast Cancer-Related Information Seeking

One of the goals of this study was to determine whether news content of celebrity breast cancer disclosures impacts public cancer-related behavioral outcomes (i.e., breast cancer-related information seeking). The hypotheses for Study 2 are related to this goal and also reflect the different theoretical frameworks or underlying assumptions used in research testing aggregate level online search query data—agenda-setting and ambiguity. Agenda-setting is most relevant when testing for proximal effects from news coverage regardless of content. Depending on the level of data (monthly or weekly), there is limited support for such effects (e.g., approximately a two-week lag for diagnosis related information seeking after a disclosure for [weekly level data, 2009-2012]) or no support (e.g., approximately a 17-month lag for breast cancer information seeking [monthly level data, 2005-2016]).

The potential for ambiguity to be a causal mechanism in information seeking outcomes could be tested by using the volume of episodic and thematic frames as predictor variables; however, the presence of thematic frames in the content analysis sample was too low to test. Given the lack of support for proximal agenda-setting effects, ambiguity or public uncertainty could be a causal mechanisms at play (e.g., Ball-Rokeach & DeFleur, 1976; Maurer & Holbach, 2015). The culmination of consistent episodic

frames has the potential to influence these cumulative breast cancer-related information seeking effects, just as Stryker (2003) found that news coverage of marijuana gradually impacted beliefs towards marijuana use in teenagers. Future research should explore the potential for episodically framed news coverage of cancer and health topics to cause ambiguity, both in relation to celebrity health disclosures and for coverage related to other media events (e.g., medical study, screening recommendations).

Because the presence of media frames was so disparate, the potential for frames to impact seeking behaviors could not be directly tested. However, the search query domains for which there were statistically significant results help to elucidate a possible content and information seeking relationship. For example, the public's increase in search for prevention-related information may be linked to news coverage of the largest media event, Angelina Jolie's preventive disclosure. In addition, six other celebrities-of-interest had similar disclosures. While Treatment and Death disclosures were the most common, prevention information for breast cancer could be perceived by the public as more ambiguous than the other types of disclosures. As the information science literature suggests (e.g., Atkin, 1973; Knobloch et al., 2003), these more novel disclosures could induce ambiguity and thus promote information seeking for terms such as BRCA, breast cancer risk, breast cancer causes, and breast cancer prevention, and may help the public to learn more about prevention strategies, thus reducing ambiguity.

Similar pathways may be at play for diagnosis and general breast cancer domain search queries—the other two significant search query domains—or alternatively the consistent mention of those words in the news coverage may activate these schemas when seeking online information, in line with longer-term priming effects. Entering

"breast cancer" into the Google search engine is likely an almost instinctual response, requiring the least cognitive effort when in search of breast cancer-related information. A diagnosis-related search query may also be top-of-mind after exposure to coverage of a celebrity breast cancer disclosure. The Diagnosis breast cancer-related event type was an infrequent designation for celebrities-of-interest; creating mutually exclusive disclosure categories introduces a limitation when contextualizing the information seeking results. Quantitatively the Diagnosis disclosures are low in number, but from countless readings of many of the media reports, news coverage of celebrities classified as having a Treatment disclosure also contained information regarding the celebrity's diagnosis. (It logically follows that for one to be treated for breast cancer, one must be diagnosed).

Media priming was discussed briefly in the literature review as a potential media effects concept that could be applied to some studies using Google Trends data as a proxy for information seeking outcomes. However, this theoretical explanation was quickly dismissed because previous research on celebrity cancer disclosures and cancer-related information seeking had found proximal effects. In retrospect, media priming may conceptually align with the results of Study 2 and is more in line with Stryker's (2003) second perspective on media effects—gradual and cumulative effects of news coverage on long-term secular trends (p. 307). Formulating a hypothesis with priming in mind would have resulted in a confirmed hypothesis.

The third question this dissertation addressed was if there are predictive celebrity attributes or breast cancer-related event types that can predict media and public behavioral outcomes. The influence of celebrity age and level of celebrity status are fairly clear with the media outcomes; this cannot be said of the information seeking outcomes.

This is due, in part, to the inconsistent results in the main effects models and then in the moderation analysis. The only attribute to have a moderating effect at the monthly-level was level of celebrity status. Celebrities with the highest level of fame (Level 10) positively moderated the relationship between coverage and the public's online information seeking in the breast cancer and survivorship domains. Interestingly, levels 6 and 9 have a negative moderating effect in the models—these levels influenced a decrease in relative search volume. Negative effects from the attribute categories were not expected prior to the analysis, but in fact, depending on the search dimension, several of the attributes have negative moderating effects. Identification informed the selection of attributes; positing that when individuals identify with the celebrity, they may be more likely to seek further information, as has been found in previous research (Basil, 1996; Brown & Basil, 2010; Myrick et al., 2014, 2013). Perhaps this bore out with the moderators with a positive moderating effect, but the inverse relationship may also support the same supposition. Concepts explicating why people search for health and cancer information is discussed at length in this dissertation, yet why people do not search is almost as robust a field. Information avoidance can occur for myriad of reasons, but fear is known to cause information avoidance, particularly in the case of a serious disease (Bawden & Robinson, 2009; Case et al., 2005; Howell & Shepperd, 2013; Lambert & Loiselle, 2007). Because it is common for people to identify with a celebrity—to consider yourself similar to the celebrity, even if demographically, for example, you are not—news of a celebrity being diagnosed with breast cancer could elicit fear and subsequent information avoidance.

The conflicting nature of the overall results for Study 2 are due, in part, to the numerous models tested. If I only tested "breast cancer" for the complete duration under study, as much of the previous research has (e.g., Ayers et al., 2014; Noar et al., 2013; Weeks et al., 2012), the results would have been straightforward: media coverage of celebrity breast cancer disclosures have gradual effects on breast cancer-related information seeking, and the highest level of celebrity status moderates these effects. (Some of the more nuanced issues with the measures and methods with Study 2 are discussed in the Conclusion section below.) To combat these issues future research could collect self-report cancer information seeking data proximal to the disclosure, as did Myrick and colleagues (2013, 2014) after the death of Steve Jobs. Additionally, using the identified celebrities-of-interest, future research could replicate this study using other aggregate breast cancer-related behavioral data (e.g., screenings, surgical procedures)—similar behavioral effects have been found related to a single disclosure (e.g., Jolie, Minogue) (Kelaher et al., 2008). While online information seeking is a discrete action, it is almost effortless. Medical appointments and surgical decisions, for example, require more effort, cognitive and behavioral. If these behaviors are impacted by media coverage of celebrity breast cancer disclosures, gradual effects over time are expected.

Conclusions

This dissertation answers Noar and colleagues' (Noar et al., 2014) call for further research on celebrity cancer disclosures—going beyond the "Angelina Effect"—particularly by investigating the attributes of public figures potential impact on corresponding media and public-related health behaviors. This study appears to be the first to provide theoretically informed conceptual and operational definitions of *celebrity*

which are used to systematically identify over 100 celebrity breast cancer disclosures across a significant period of time. It is also the first health communication study to categorize the census of celebrities by personal attributes and breast cancer-related event types. The categorizations may be imperfect measures, but the results here indicate that age and level of celebrity status may be the strongest predictors of media and breast cancer information seeking outcomes—providing empirical support for Noar et al.'s (2014) supposition that "perhaps only those with some level of 'celebrity' status – garner significant effects" (p. 457).

A core argument of this dissertation was because celebrity breast cancer disclosures are so common, researching isolated disclosures only provides a deeper understanding of that specific media event and impedes building knowledge about the greater phenomena. This dissertation offers some evidence to support this argument; 80% of the news coverage was written with an episodic frame, misinformation was all but absent from the media report census, and for certain search query domains media coverage of celebrity health disclosures have long-term cumulative effects. However, the distinct media (i.e., elevated presence of media attention, thematic frames, and misinformation) and breast cancer-related information seeking outcomes for those with the highest levels of celebrity status lends empirical support to justify scholarly attention to specific individuals. Because media and information seeking outcomes are rather consistent in response to the majority of celebrity disclosures, it is possible that outlying cases are the most likely to have proximal large-scale public health effects.

Age and level of celebrity status often intersect—many higher status celebrities are Under age 50—but this relationship is not a rule nor should it be a primary motivator

for selecting celebrities-of-interest for research. While I noted in the introduction and literature review that most of the research on celebrity breast cancer disclosures focus on the super-celebrity, I had not noticed another commonality shared by celebrity subjects: age. The seminal breast cancer disclosures, and subsequently seminal research, came from women who were in their late 50s (i.e., Betty Ford) and mid-60s (i.e., Nancy Reagan). Had a level of celebrity status index existed at this time, these two first ladies would likely have been in the upper tiers. Their disclosures were notable because it was unusual for women of such status to discuss breast cancer publicly, but being diagnosed with breast cancer at their respective ages was not. However, studies on content and effects from a single disclosure that have been released in the timeframe of this study have only focused on Jolie and Minogue, both in their 30s at time of disclosure. When including Sabel and Dal Cin's (2016) work, Etheridge, Minogue, Crow, Nixon, Roberts, Applegate, Mitchell, Rancic, Sykes, Bates, and Lunden are included. Other than Mitchell, Bates, and Lunden, all of the disclosures studied in the last decade were from women in their 30s and 40s.

Perhaps the relative youth of these celebrities makes their disclosures stick out to researchers, but I suggest that a more systematic approach be used when selecting celebrities-of-interest. Studying the effects of celebrities under 50 can be important, particularly for deleterious outcomes such as over screening for women at average risk, for example. However, at a population level, breast cancer risk increases with age—the median age of a woman diagnosed with breast cancer is 62 (National Cancer Institute, 2017; Susan G. Komen, n.d.). Disclosures from women at an age of 50 or older could potentially impact breast cancer-related outcomes from women of the ages who are most

at risk. This dissertation identifies several celebrities-of-interest at age of 50 and older who also occupy the highest levels of celebrity status (Levels 7 -10). Opportunities for studying women in these groups include Shirley Horn, Molly Ivins, Elizabeth Edwards, Dorothy Hammill, Maggie Smith, Martina Navratilova, Kim Novak, Judy Blume, and Rita Wilson. In the future, a few multiple Boolean Google searches for celebrity breast cancer disclosures, in the last year (for example), would provide a more systematic approach for choosing celebrities-of-interest, resulting in studies with reduced bias in age.

The use of systematic methods to determine celebrities-of-interest is representative of a broader approach for this dissertation: novel methodological approaches. To the best of my knowledge this is the first study to test recall and precision of article retrieval in an online news database using individuals' names. Future research can use the steps provided here when retrieving content for names. Perhaps the most novel approach in measurement was the level of celebrity status index. This is the first study to quantify level of celebrity status in the health communication literature. This study also answers Driessen's call for empirically measuring celebrity using both media and audience salience measures.

The weighting issues with the Google Trends data made the audience salience measurement process difficult. This dissertation does not answer the question of construct validity with Google Trends—salience or information seeking? I believe the behavioral action of entering a phrase into a search engine generally measures information seeking and not simply salience—although to seek, the search phrase must be salient (i.e., the celebrity). These issues likely result in an imperfect index. But, the index is a good

example of how researchers can incorporate audience measures into indices delineating hierarchy of fame. Hopefully the level of celebrity status index is a starting point for other academics to either build upon or think of new ways to capture audience salience. As Driessen (2015) states:

> "What is open for discussion and contingent upon the nature of particular research projects are the questions (a) how broadly we demarcate celebrity's representations in media and in other cultural artifacts and (b) how we measure their consumption and/or retention in certain celebrity or media cultures in space and time. Yet accounting for both basic dimensions through the notion of memory might help us to gradually refine celebrity as sensitizing concept—not by cleaning up the mess, but dealing with it in better and more transparent ways" (p. 372).

Finally, the use of several multifaceted search dimensions at different levels (i.e., monthly and weekly) to measure breast cancer-related information seeking broadens the scope in which information seeking is defined and measured. This dissertation made extensive use of Google Trends search data; including data downloaded for audience salience measures, over 1,200 data sets were retrieved. Spending significant time with these less than transparent data allowed me to garner an appreciation as to how ambiguous these data are. The discussion and figures I provide in Chapter 4 about how the relative value of the data and the trend can change by shifting the start and end dates of the time frame for data retrieval brings these issues to the forefront. The relative value of the data also change dramatically depending on the level of data (i.e., monthly or weekly). This can be seen in Chapter 6 visual trends analyses. The four-year increments of data are not just stretched out versions of the same time period in the full 2005-2016

data set. This likely influences the differences in significant lag periods found between the 2005-2016 datasets and the weekly-level datasets.

Active debate about Google Trends' construct validity has taken place over the last several years (see Maurer & Holbach, 2015; Mellon, 2014), however, the actual validity of the data has almost been ignored. Lazer, Kennedy, King, and Vespignani (2014) do raise questions about Google Flu Trends validity. Google Flu Trends was a web service by Google which used RSV scores of search terms related to influenza. Several early models were quite accurate in predicting flu outbreaks in regions of several countries based on these data, but later models proved unable to replicate such results. The service was subsequently discontinued. Perhaps the core issue with Google Trends, and other "big data," is summarized with this statement from Lazer et al. (2014): "The core challenge is that most big data that have received popular attention are not the output of instruments designed to produce valid and reliable data amenable for scientific analysis" (p. 2). Lazer et al. (2014) offer many critiques to support this statement, but the two most relevant here are, as I pointed out above, the lack of transparency in measurement, and that Google is constantly updating their algorithm. Data retrieval from one day to the next could result in different RSV scores.

While the process I used in retrieving RSV data and comparing data over different time periods was very time consuming, and ultimately concluded with inconsistent results, I would recommend such an approach for future research. Google Trends is not transparent data, but I believe researchers should be transparent in their data collection process and about inconsistencies in the data. Google Trends can still provide valuable insight to general information seeking behaviors of the public, but perhaps the

conclusions that we make should be more conservative. In addition, when possible, using both Google Trends data and self-report public opinion or information seeking data would triangulate findings and strengthen the nomological validity of studies.

This systematic multi-method longitudinal analysis supports Noar and colleagues' (2014) assertion that some level of celebrity status is necessary to see big public health effects, at least in the short-term. However, building upon the natural coverage effects research tradition, the evidence of general news media and breast cancer-related information seeking outcomes indicate that sustained news coverage of multiple celebrity breast cancer disclosures over time may also have pervasive, gradual effects. Future work can and should replicate and extend the novel methods and measures provided.

Because this dissertation exclusively uses "real world" data, contextualizing the practical importance of this dissertation is straightforward. Contrary to many public health and health communication scholars assertions, media coverage of celebrity breast cancer disclosures is typically not a "good" public health communication intervention and as scholars we should not consider these announcements to be "teachable moments." While the long-term effects on breast cancer-related information seeking could be a pro-public health good—individuals seeking out additional information—the aggregate level evidence offered here cannot tell us if those who are most in need of accurate breast cancer-related information are the ones seeking it, or if the information obtained through the Google search is reliable, accurate information. Therefore, there is potential for deleterious outcomes related to celebrity breast cancer disclosures.

To combat this potential, I offer two suggestions. First is outreach to journalists. Throughout this dissertation, I have cautioned researchers about the tendency to suggest

to journalists that they should change their practices. I will not contradict myself now. But outreach does not have to include criticizing reporting practices. As suggested earlier, there are research opportunities involving the reporting practices of celebrity cancer and health disclosures. This dissertation demonstrates that reporting practices are fairly uniform and predictable across celebrities and by level of celebrity status. Inquiry into journalists' decision making processes would provide a better understanding of why these outcomes occur, and also may provide an opportunity for journalists to think critically about their approach.

Beyond research, simple discussions with journalists and news organizations about the evidence regarding public health outcomes related to news coverage of celebrity health disclosures could begin a dialogue between the public health and journalism communities. I do believe more evidence is needed that directly connects volume and content of news coverage to negative public health outcomes before clear recommendations could be made about reporting practices. While unusual, some outside lobbying from those in the public health sector have impacted guidelines in the Associated Press Stylebook—the only concrete way to ensure journalists adopt reporting practices. After dozens of studies over several decades demonstrated copycat effects from news media coverage of suicide (see Stack, 2000), in 2015 the AP Stylebook included clear guidelines for reporting on suicide. Other than such drastic changes, for those of us who teach journalism students, discussing the news coverage of celebrity cancer and health disclosures and the potential for deleterious outcomes in class provides an opportunity for budding journalists to think critically about these issues early in their career.

The second suggestion to combat potential negative public health consequences from celebrity cancer disclosures is for strategic cancer communication. Evidence from this dissertation and other studies indicate that for certain large media events (e.g., Jolie's disclosure) effects can be immediate (see Noar et al., 2014). In cases like these, close monitoring of the information environment is necessary. If misinformation or ambiguous information is apparent, strategic messages providing clear and accurate information should be disseminated through multiple channels immediately. Outreach to the celebrity and or her or his publicist is also recommended. When Charlie Sheen announced his HIV status, his doctor appeared with him during interviews to make sure the health information conveyed was accurate. Perhaps more celebrities would be open to such a strategy when publicly discussing their health issues.

Finally, effects can be immediate in some cases, but this dissertation along with studies such as Sabel and Dal Cin (2016) offer preliminary evidence that effects may be more gradual and even impact secular trends over time (e.g., increased bilateral mastectomies). These results coupled with the conflicting and contradictory breast cancer information in the greater information environment, indicate that more sustained strategic breast cancer communication efforts are warranted. Several celebrities who are breast cancer survivors lend their voices to public relations and advertising campaigns for organizations such as the American Cancer Society, and Sheryl Crow is even a paid spokesperson for the Genius 3D Mammography Exam. Leveraging celebrity status to provide clear and accurate information about screening guidelines, BRCA testing, and treatment decision making processes could have a large impact on breast cancer knowledge and subsequently informed decision making at large.

References

Alexa. (n.d.). The top 500 sites on the web. Retrieved from http://www.alexa.com/topsites

Amend, E., & Secko, D. M. (2012). In the face of critique: A metasynthesis of the experiences of journalists covering health and science. *Science Communication*, *34*(2), 241-282. https://doi.org/10.1177/1075547011409952

American Cancer Society. (2015). American cancer society releases new breast cancer guidelines. Retrieved from https://www.cancer.org/latest-news/american-cancer-society-releases-new-breast-cancer-guidelines.html

Anderegg, W. R., & Goldsmith, G. R. (2014). Public interest in climate change over the past decade and the effects of the "climategate"media event. *Environmental Research Letters*, *9*(5), 54. Retrieved from http://iopscience.iop.org/1748-9326/9/5/054005)

Anderson, T. W., & Darling, D. A. (1954). A test of goodness of fit. *Journal of the American Statistical Association*, *49*(28), 765–769. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/01621459.1954.10501232

Atkin, C. (1973). Information utilities and information seeking. In P. Clarke (Ed.), *New models in mass communication research* (pp. 205–242). Beverly Hills, CA: SAGE.

Ayers, J. W., Althouse, B. M., Noar, S. M., & Cohen, J. E. (2014). Do celebrity cancer diagnoses promote primary cancer prevention? *Preventive Medicine*, *58*, 81–4. https://doi.org/10.1016/j.ypmed.2013.11.007

Babrow, A. S., Kasch, C. R., & Ford, L. A. (1998). The many meanings of uncertainty in illness: Toward a systematic accounting. *Health Communication*, *10*(1), 1-23. http://dx.doi.org/10.1207/s15327027hc1001_1

Ball-Rokeach, S. J., & DeFleur, M. L. (1976). A dependency model of mass-media effects. *Communication Research*, *3*(1), 3–21. https://doi.org/10.1177/009365027600300101

Barbour, J. B., Rintamaki, L. S., Ramsey, J. A., & Brashers, D. E. (2012). Avoiding health information. *Journal of Health Communication*, *17*(2), 212–29. https://doi.org/10.1080/10810730.2011.585691

Barnett, L., & Seth, A. K. (2014). The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. *Journal of Neuroscience Methods*, *223*, 50–68. https://doi.org/10.1016/j.jneumeth.2013.10.018

Basil, M. D. (1996). Identification as a mediator of celebrity effects. *Journal of Broadcasting & Electronic Media*, *40*(4), 478. http://dx.doi.org/10.1080/08838159609364370

Bawden, D., & Robinson, L. (2009). The dark side of information: Overload, anxiety and other paradoxes and pathologies. *Journal of Information Science*, *35*(2), 180–191. https://doi.org/10.1177/0165551508095781

Beck, C. S., Aubuchon, S. M., McKenna, T. P., Ruhl, S., & Simmons, N. (2014). Blurring personal health and public priorities: An analysis of celebrity health narratives in the public sphere. *Health Communication*, *29*(3), 244–56. https://doi.org/10.1080/10410236.2012.741668

Bickell, N. A., Weidmann, J., Kezhen, F., Lin, J. J., & Leventhal, H. (2009). Underuse of breast cancer adjuvant treatment: Patient knowledge, beliefs, and medical mistrust. *Journal of Clinical Oncology*, *27*(31), 5160–5167. https://doi.org/10.1200/JCO.2009.22.9773

Boczkowski, P. J., & de Santos, M. (2007). When more media equals less news: Patterns of content homogenization in Argentina's leading print and online newspapers. *Political Communication*, *24*(2), 167–180. https://doi.org/10.1080/10584600701313025

Boorstin, D. (1961). *The image: Or what happened to the American Dream*. London: Weidenfeld & Nicholson.

Borzekowski, D. L. G., Guan, Y., Smith, K. C., Erby, L. H., & Roter, D. L. (2014). The Angelina effect: Immediate reach, grasp, and impact of going public. *Genetics in Medicine*, *16*(7), 516–21. https://doi.org/10.1038/gim.2013.181

Braudy, L. (1986). *The frenzy of renown:Fame and its history*. Oxford, UK: Oxford University Press.

Breese, E. B. (2010). Reports from "backstage" in entertainment news. *Society*, *47*(5), 396–402. https://doi.org/10.1007/s12115-010-9350-2

Brown, W. J., & Basil, M. D. (1995). Media celebrities and public health: Responses to "Magic" Johnson's HIV disclosure and its impact on AIDS risk and high-risk behaviors. *Health Communication*, *7*(4), 345–370. Retrieved from 10.1207/s15327027hc0704_4

Brown, W. J., & Basil, M. D. (2010). Parasocial interaction and identification: Social change processes for effective health interventions. *Health Communication*, *25*(6-7), 601–2. https://doi.org/10.1080/10410236.2010.496830

Carpenter, D. M., Geryk, L. L., Chen, A. T., Nagler, R. H., Dieckmann, N. F., & Han, P. K. J. (2015). Conflicting health information: A critical research need. *Health Expectations*, *19*(6), 1173-1182. https://doi.org/10.1111/hex.12438

Case, D. O., Andrews, J. E., Johnson, J. D., & Allard, S. L. (2005). Avoiding versus the relationship of information seeking to avoidance, blunting, coping, dissonance, and related concepts. *Journal of the Medical Library Association*, *93*(3), 353. Retrieved from http://scholarcommons.usf.edu/cgi/viewcontent.cgi?article=1252&context=si_facpub

Cates, W., Grimes, D. A., Ory, H. W., Tyler, C. W., & Cates, B. W. (1977). Publicity and the public health: The elimination of IUD-related abortion deaths. *Family Planning Perspectives*, *9*(3), 138–140. http://www.jstor.org/stable/2134528

Cavanah, S. B. (2016). *Measuring metropolitan newspaper pullback and its effects on political participation*. https://doi.org/https://conservancy.umn.edu/handle/11299/182213

Chapman, S., McLeod, K., Wakefield, M., & Holding, S. (2005). Impact of news of celebrity illness on breast cancer screening. *Medical Journal of Australia*, *183*(5), 247–250. Retrived from https://www.mja.com.au/journal/2005/183/5/impact-news-celebrity-illness-breast-cancer-screening-kylie-minogues-breast?inline=true

Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. *Economic Record*, *88*(1), 2–9. https://doi.org/10.1111/j.1475-4932.2012.00809.x

Clarke, J. N., & Everest, M. M. (2006). Cancer in the mass print media: Fear, uncertainty and the medical model. *Social Science & Medicine (1982)*, *62*(10), 2591–600. https://doi.org/10.1016/j.socscimed.2005.11.021

Cohen, J. (2004). Parasocial break-up from favorite television characters: The role of attachment styles and relationship intensity. *Journal of Social and Personal Relationships*, *21*, 187–202. https://doi.org/10.1177/0265407504041374

Coleman, R., & Thorson, E. (2002). The effects of news stories that put crime and violence into context: Testing the public health model of reporting. *Journal of Health Communication*, *7*(5), 401–425. https://doi.org/10.1080/10810730290001783

Coleman, R., Thorson, E., & Wilkins, L. (2011). Testing the effect of framing and sourcing in health news stories. *Journal of Health Communication*, *16*(9), 941-954. https://doi.org/10.1080/10810730.2011.561918

Connolly-Ahern, C., & Broadway, S. C. (2008). “To booze or not to booze?” Newspaper coverage of fetal alcohol spectrum disorders. *Science Communication*, *29*(3), 362–385. https://doi.org/10.1177/1075547007313031

Cooper, C. P., Mallon, K. P., Leadbetter, S., Pollack, L. A., & Peipinism, L. A. (2005). Cancer Internet search activity on a major search engine, United States 2001-2003. *Journal of Medical Internet Research*, *7*(3), 36. https://doi.org/10.1111/j.1475-4932.2012.00809.x

Corbett, B. J. B., & Mori, M. (1999). Medicine, media, and celebrities: News coverage of breast cancer, 1960-1995. *Journalism & Mass Communication Quarterly*, *76*(2), 229–249. https://doi.org/10.1177/107769909907600204

Covello, V.T., & Peters, R. G. (2002). Women's perceptions of the risks of age-related diseases, including breast cancer: Reports from a 3-year research study. *Health Communication*, *14*(3), 377-395. http://dx.doi.org/10.1207/S15327027HC1403_5

Cowpertwait, P. S. P., & Metcalfe, A. V. (2009). *Introductory time series with R*. London: Springer.

Dalton, M. A, Sargent, J. D., Beach, M. L., Titus-Ernstoff, L., Gibson, J. J., Ahrens, M. B., …
Heatherton, T. F. (2003). Effect of viewing smoking in movies on adolescent smoking initiation: a cohort study. *Lancet*, *362*, 281–5. https://doi.org/10.1016/S0140-6736(03)13970-0

de Leon, C. L. P. (2002). *Self-exposure: Human-interest journalism and emergence of celebrity in America, 1890-1940*. University of North Carolina Press.

Dean, M. (2015). Celebrity health announcements and online health information seeking: An analysis of Angelina Jolie’s preventative health decision. *Health Communication*, *236*, 1–10. https://doi.org/10.1080/10410236.2014.995866

DiMaggio, P (1997). "Culture and cognition". *Annual Review of Sociology*. **23**: 263–287. https://doi.org./10.1146/annurev.soc.23.1.263

Dobransky, K., & Hargittai, E. (2012). Inquiring minds acquiring wellness: Uses of online and offline sources for health information. *Health Communication*, *27*(4), 331–43. https://doi.org/10.1080/10410236.2011.585451

Driessens, O. (2015). On the epistemology and operationalisation of celebrity. *Celebrity Studies*, *6*(3), 370–373. https://doi.org/10.1080/19392397.2015.1062651

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics*, *75*(4), 643–669. http://www.jstor.org/stable/1884324

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, *43*(4), 51–58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x

Fagerlin, A., Lakhani, I., Lantz, P. M., Janz, N. K., Morrow, M., Schwartz, K., … Katz, S. J. (2006). An informed decision?: Breast cancer patients and their knowledge about treatment. *Patient Education & Counseling*, *64*(1), 303–312. https://doi.org/10.1016/j.pec.2006.03.010

Ferri, A. J. (2010). Emergence of the entertainment age? *Society*, *47*(5), 403–409. https://doi.org/10.1007/s12115-010-9351-1

Fink, R., Roeser, R., Venet, W., Strax, P., Venet, L., & Lacher, M. (1978). Effects of news events on response to a breast cancer screening program. *Public Health Reports*, *93*(4), 318. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1431913/

Fishbein, M., & Hornik, R. (2008). Measuring media exposure: An introduction to the special issue. *Communication Methods and Measures*, *2*, 1–2. http://dx.doi.org/10.1080/19312450802095943

Fowler, E. F., & Gollust, S. E. (2015). The content and effects of politicized health controversies. *The ANNALS of the American Academy of Political and Social Science*, *658*(1), 155–171. https://doi.org/10.1177/0002716214555505

Freelon, D. (2013). ReCal OIR: Ordinal, interval, and ratio intercoder reliability as a web service. *Journal of Internet Science*, *8*(1), 10–16. Retrieved from http://dfreelon.org/utils/recalfront/

Gamson, W. A., & Modigliani, A. (1989). Media discourse and public opinion on nuclear power: A constructionist approach. *American Journal of Sociology*, *95*(1), 1–37. Retrieved from http://www.jstor.org/stable/2780405

Gans, H. J. (1979). The messages behind the news. *Columbia Journalism Review*, *17*(5), 40. Retrieved from http://search.proquest.com/openview/52ed70aac5ec0494526e31df1dc04fd2/1?pq-origsite=gscholar&cbl=1817229

Gearhart, S., Craig, C., & Steed, C. (2012). Network news coverage of obesity in two time periods: An analysis of issues, sources, and frames. *Health Communication*, *27*(7), 653–662. https://doi.org/10.1080/10410236.2011.629406

Golan, G. (2006). Inter-media agenda setting and global news coverage: Assessing the influence of the New York Times on three network television evening news programs. *Journalism Studies*, *7*(2), 323–333. http://dx.doi.org/10.1080/14616700500533643

Gollop, B. C. J. (1997). Health information-seeking behavior and older African American women. *Bulletin of the Medical Library Association*, *85*(2), 141–146. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC226241/

Google. (n.d.). Retrieved from https://trends.google.com/trends/explore

Gøtzsche, P. C., & Jørgensen, K. J. (2013). Screening for breast cancer with mammography. *The Cochrane Library*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD001877.pub5/pdf

Greenberg, R., Freimuth, V. S., & Bratic, E. (1979). A content analytic study of daily newspaper coverage of cancer. *Communication Yearbook*, *3*, 645–654.

Gross, K. (2008). Framing persuasive appeals: Episodic and thematic framing, emotional response, and policy opinion. *Political Psychology*, *29*(2), 169–192. Retrieved from http://www.jstor.org/stable/20447111

Hallin, D. C., & Briggs, C. L. (2014). Transcending the medical/media opposition in research on news coverage of health and medicine. *Media, Culture & Society*, *37*(1), 85–100. https://doi.org/10.1177/0163443714549090

Han, P. K. J., Kobrin, S. C., Klein, W. M. P., Davis, W. W., Stefanek, M., & Taplin, S. H. (2007). Perceived ambiguity about screening mammography recommendations: Association with future mammography uptake and perceptions. *Cancer Epidemiology and Prevention Biomarkers*, *16*(3), 458-466. https://doi.org/10.1158/1055-9965.EPI-06-0533

Han, P. K. J., Moser, R. P., & Klein, W. M. P. (2007). Perceived ambiguity about cancer prevention recommendations: Associations with cancer-related perceptions and behaviours in a US population survey. *Health Expectations*, *10*(4), 321–36. https://doi.org/10.1111/j.1369-7625.2007.00456.x

Harris, R. (2014). Screening is only part of the answer to breast cancer. *Annals of Internal Medicine*, *160*(12), 2014–2017. https://doi.org/10.7326/M14-0616

Harris, R., & Sheridan, S. (2013). The times they (may) be a-changin': Too much screening is a health problem. *American Journal of Preventive Medicine*, *45*(2), 248–9. https://doi.org/10.1016/j.amepre.2013.05.002

Hatley-Major, L. (2009). Break it to me harshly: The effects of intersecting news frames in lung cancer and obesity coverage. *Journal of Health Communication*, *14*, 174–188. https://doi.org/10.1080/10810730802659939

Hawkins, K. W., & Linvill, D. L. (2010). Public health framing of news regarding childhood obesity in the United States. *Health Communication*, *25*(8), 709–17. https://doi.org/10.1080/10410236.2010.521913

Henderson, L., & Kitzinger, J. (1999). The human drama of genetics: "Hard" and "soft" media representations of inherited breast cancer. *Sociology of Health & Illness*, *21*(5), 560–578. https://doi.org/10.1111/1467-9566.00173

Henry, M., Trickey, B., Huang, L. N., & Cohen, S. R. (2012). How is cancer recently portrayed in Canadian newspapers compared to 20 years ago? *Supportive Care in Cancer*, *20*(1), 49–55. https://doi.org/10.1007/s00520-010-1049-9

Hersch, J., Barratt, a., Jansen, J., Houssami, N., Irwig, L., Jacklyn, G., … McCaffery, K. (2014). The effect of information about overdetection of breast cancer on women's decision-making about mammography screening: study protocol for a randomised controlled trial. *BMJ Open*, *4*(5), e004990–e004990. https://doi.org/10.1136/bmjopen-2014-004990

Hersch, J., Barratt, A., Jansen, J., Irwig, L., McGeechan, K., Jacklyn, G., … McCaffery, K. (2015). Use of a decision aid including information on overdetection to support informed choice about breast cancer screening: a randomised controlled trial. *Lancet*, *385*, 1642–52. https://doi.org/10.1016/S0140-6736(15)60123-4

Hersch, J., Jansen, J., Barratt, a., Irwig, L., Houssami, N., Howard, K., … McCaffery, K. (2013). Women's views on overdiagnosis in breast cancer screening: A qualitative study. *BMJ*, *346*, f158–f158. https://doi.org/10.1136/bmj.f158

Hersch, J., Jansen, J., Irwig, L., Barratt, A., Thornton, H., Howard, K., & McCaffery, K. (2011). How do we achieve informed choice for women considering breast screening? *Preventive Medicine*, *53*(3), 144–6. https://doi.org/10.1016/j.ypmed.2011.06.013

Hollander, P. (2010). Why the celebrity cult? *Society*, *47*(5), 388–391. https://doi.org/10.1007/s12115-010-9348-9

Holton, A., Farrell, L. C., & Fudge, J. L. (2014). A Threatening Space ?: Stigmatization and the Framing of Autism in the News. *Communication Studies*, *65*(2), 189–207. https://doi.org/10.1080/10510974.2013.855642

Holton, A., Weberling, B., Clarke, C. E., & Smith, M. J. (2012). The blame frame: Media attribution of culpability about the MMR-autism vaccination scare. *Health Communication*, *27*(7), 690–701. https://doi.org/10.1080/10410236.2011.633158

Hornik, R. C. (2002). Introduction: Making sense of contradictory evidence. In R. C. Hornik (Ed.), *Public health communication: Evidence for behavior change* (pp. 1–22). Mahwah, NJ: Lawrence Erlbaum Associates.

Howell, J. L., & Shepperd, J. A. (2013). Reducing health-information avoidance through contemplation. *Psychological Science*, *24*(9), 16. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/0956797613478616

Hurley, R. J., Kosenko, K. A., & Brashers, D. (2011). Uncertain terms: Message features of online cancer news. *Communication Monographs*, *78*(3), 370–390. https://doi.org/10.1080/03637751.2011.565061

Iyengar, S. (1987). Television news and citizens' explanations of national affairs. *American Political Science Review*, *81*(3), 815–831. https://doi.org/10.2307/1962678

Iyengar, S. (1990). Framing responsibility for political issues: The case of poverty. *Political Behavior*, *12*(1), 19–40. https://doi.org/10.1007/BF00992330

Iyengar, S., & Kinder, D. R. (2010). *News that matters: Television and American opinion*. University of Chicago Press.

Iyengar, S., & Simon, A. (1993). News coverage of the Gulf Crisis and public opinion: A study of agenda-setting, priming, and framing. *Communication Research*, *20*, 365. https://doi.org/10.1177/009365093020003002

Jensen, J. D., Moriarty, C. M., Hurley, R. J., & Stryker, J. E. (2010). Making sense of cancer news coverage trends: A comparison of three comprehensive content analyses. *Journal of Health Communication*, *15*(2), 136–51. https://doi.org/10.1080/10810730903528025

Jensen, J. D., Scherr, C. L., Brown, N., Jones, C., Christy, K., & Hurley, R. J. (2014). Public estimates of cancer frequency: Cancer incidence perceptions mirror distorted media depictions. *Journal of Health Communication*, *19*(5), 609–24. https://doi.org/10.1080/10810730.2013.837551

Jha, S. (2007). Exploring Internet influence on the coverage of social protest. Content Analysis comparing protest coverage in 1967 and 1999. *Journalism & Mass Communication Quarterly*, *84*(1), 40–57. https://doi.org/10.1177/107769900708400104

Kalager, M., Zelen, M., Langmark, F., & Adami, H. O. (2010). Effect of screening mammography on breast-cancer mortality in Norway. *New England Journal of Medicine*, *363*(13), 1203–1210. Retrieved from http://www.nejm.org/doi/full/10.1056/Nejmoa1000727#t=article

Kalichman, S. C., & Hunter, T. L. (1992). The disclosure of celebrity HIV infection: Its effects on public attitudes. *American Journal of Public Health*, *82*(10), 1374–1376. https://doi.org/10.2105/AJPH.82.10.1374

Kamenova, K., Reshef, A., & Caulfield, T. (2014). Angelina Jolie's faulty gene: Newspaper coverage of a celebrity's preventive bilateral mastectomy in Canada, the United States, and the United Kingdom. *Genetics in Medicine*, *16*(7), 522–8. https://doi.org/10.1038/gim.2013.199

Kelaher, M., Cawson, J., Miller, J., Kavanagh, A., Dunt, D., & Studdert, D. M. (2008). Use of breast cancer screening and treatment services by Australian women aged 25-44 years following Kylie Minogue's breast cancer diagnosis. *International Journal of Epidemiology*, *37*(6), 1326–32. https://doi.org/10.1093/ije/dyn090

Kelman, H. C. (1961). Processes of opinion change. *Public Opinion Quarterly*, *25*(1), 57-78. https://doi.org/10.1086/266996

Kiousis, S., & McCombs, M. (2004). Agenda-setting effects and attitude strength: Political figures during the 1996 presidential election. *Communication Research*, *31*, 36–57. https://doi.org/10.1177/0093650203260205

Knobloch, S., Dillman Carpentier, F., & Zillmann, D. (2003). Effects of salience dimensions of informational utility on selective exposure to online news. *Journalism & Mass Communication Quarterly*, *80*(1), 91–108. https://doi.org/10.1177/107769900308000107

Kosenko, K. A., Binder, A. R., & Hurley, R. (2015). Celebrity influence and identification: A Test of the Angelina effect. *Journal of Health Communication*, *73*, 1–9. https://doi.org/10.1080/10810730.2015.1064498

Lambert, S. D., & Loiselle, C. G. (2007). Health information seeking behavior. *Qualitative Health Research*, *17*(8), 1006–19. https://doi.org/10.1177/1049732307305199

Lane, D. S., Polednak, A. P., & Burg, M. A. (1989). The impact of media coverage of Nancy Reagan's experience on breast cancer screening. *American Journal of Public Health*, *79*, 1551–1552. https://doi.org/10.2105/AJPH.79.11.1551

Lawrence, R. G. (2004). Framing obesity: The evolution of news discourse on a public health issue. *The Harvard International Journal of Press/Politics*, *9*(3), 56–75. https://doi.org/10.1177/1081180X04266581

Lazer, D. R., Kennedy, G., King, G., & Vespignani. (2014). The parable of Google Flu Trends: Traps in big data analysis. *Science*, *343*, 1203–1205. https://doi.org/10.1126/science.1248506

Lee, B., Kim, J., & Scheufele, D. A. (2015). Agenda setting in the Internet Age: The recirocity between online searches and issue salience. *International Journal of Public Opinion Research*. https://doi.org/10.1093/ijpor/edv026

Lee, C.-J., Zhao, X., & Pena-y-Lillo, M. (2016). Theorizing the pathways from seeking and scanning to mammography screening. *Health Communication*, *31*(1), 117–28. https://doi.org/10.1080/10410236.2014.942769

Lee, S. T., & Basnyat, I. (2013). From press release to news: Mapping the framing of the 2009 H1N1 A influenza pandemic. *Health Communication*, *28*(2), 119–32. https://doi.org/10.1080/10410236.2012.658550

Lerner, B. H. (2006). *When illness goes public: Celebrity patients and how we look at medicine*. Baltimore: The John Hopkins University Press.

Li, X., You, R., Wang, X., Liu, C., Xu, Z., Zhou, J., … Zou, Q. (2016). Effectiveness of prophylactic surgeries in BRCA1 or BRCA2 mutation carriers: A meta-analysis and systematic review. *Clinical Cancer Research*, *22*(15), 3971–3981. https://doi.org/10.1158/1078-0432.CCR-15-1465

Maiier, S. (2010). All the news fit to post? Comparing news content on the web to newspapers, television, and radion. *Journalism & Mass Communication Quarterly*, *87*(3), 548–562. https://doi.org/10.1177/107769901008700307

Matthes, J., & Kohring, M. (2008). The content analysis of media frames: Toward improving reliability and validity. *Journal of Communication*, *58*(2), 258–279. https://doi.org/10.1111/j.1460-2466.2008.00384.x

Maurer, M., & Holbach, T. (2015). Taking online search queries as an indicator of the public agenda: The role of public uncertainty. *Journalism & Mass Communication Quarterly*, 1–15. https://doi.org/10.1177/1077699015610072

McAlone, N. (2016). Meet the YouTube millionares: These are the highest-paid YouTube stars of 2016. Retrieved from http://www.businessinsider.com/youtube-stars-who-make-the-most-money-in-2016-2016-12

McCombs, M. E., & Shaw, D. L. (1993). The evolution of agenda-setting research: Twenty-five years in the marketplace of ideas. *Journal of Communication*, *43*(2), 58–67. https://doi.org/ 10.1111/j.1460-2466.1993.tb01262.x

McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. *Public Opinion Quarterly*, *32*(2), 176–187. https://doi.org/10.1086/267990

Mellon, J. (2013). Where and when we can use Google Trends to measure issue salience. *Political Science & Politics*, *46*, 280–290. https://doi.org/http://dx.doi.org/10.1017/S1049096513000279

Mellon, J. (2014). Internet search data and issue salience: The properties of Google Trends as a measure of issue salience. *Journal of Elections, Public Opinion, & Parties*, *24*, 45–72. https://doi.org/10.1080/17457289.2013.846346

Metcalfe, D., Price, C., & Powell, J. (2011). Media coverage and public reaction to a celebrity cancer diagnosis. *Journal of Public Health (Oxford, England)*, *33*(1), 80–5. https://doi.org/10.1093/pubmed/fdq052

Miller, J. D. (2010). Civic scientific literacy: The role of the media in the electronic era. *Science and the Media*, 44-63. Retrieved from http://www.jstor.org/stable/3088437

Milner, M. (2010). Is celebrity a new kind of status system? *Society*, *47*(5), 379–387. https://doi.org/10.1007/s12115-010-9347-x

Mitchell, A., Jurkowitz, M., & Olmstead, K. (2014). Social, search and direct:Pathways to digital news. Pew Research Center. Retrieved from http://www.journalism.org/2014/03/13/social-search-direct/

Myrick, J. G., Noar, S. M., Willoughby, J. F., & Brown, J. (2014). Public reaction to the death of Steve Jobs: Implications for cancer communication. *Journal of Health Communication*, *19*(11), 1278–95. https://doi.org/10.1080/10810730.2013.872729

Myrick, J. G., Willoughby, J. F., Noar, S. M., & Brown, J. (2013). Reactions of young adults to the death of Apple CEO Steve Jobs: Implications for cancer communication. *Communication Research Reports*, *30*(2), 115–126. https://doi.org/10.1080/08824096.2012.762906

Nagler, R. H. (2014). Adverse outcomes associated with media exposure to contradictory nutrition messages. *Journal of Health Communication*, *19*(1), 24–40. https://doi.org/10.1080/10810730.2013.798384

Nagler, R.H. (In press). Measurement of media exposure. In J. Matthes (Ed.), *International Encyclopedia of Communication Research Methods.* Hoboken, NJ: Wiley-Blackwell and the International Communication Association.

Nagler, R. H., Fowler, E. F., & Gollust, S. E. (2015). Covering controversy: What are the implications for women's health? *Women's Health Issues*, *25*(4), 318–21. https://doi.org/10.1016/j.whi.2015.04.011

Nagler, R. H., Fowler, E. F., & Gollust, S. E. (2017). Communicating about breast cancer overdiagnosis and overtreatment: Data from a population-based survey of U.S. women. In *International Communication Association Annual Conference*. San Diego, CA.

Nagler, R. H., & LoRusso, S. M. (2017). Conflicting Information and Message Competition In Health And Risk Messaging. In *Encyclopedia of Health and Risk Message Design and Processing*. Oxford University Press.

National Cancer Institute. (n.d.). Cancer control continuum. Retrieved from https://cancercontrol.cancer.gov/od/continuum.html

National Cancer Institute. (2017). Breast cancer risk in American women. Retrieved from https://www.cancer.gov/types/breast/risk-fact-sheet

Nghiem, L. T. P., Papworth, S. K., Lim, F. K. S., & Carrasco, L. R. (2016). Analysis of the capacity of Google Trends to measure Interest in conservation topics and the role of online news, 1–12. https://doi.org/10.1371/journal.pone.0152802

Niederdeppe, J. (2008). Beyond knowledge gaps: Examining socioeconomic differences in response to cancer news. *Human Communication Research*, *34*(3), 423–447. https://doi.org/10.1111/j.1468-2958.2008.00327.x

Niederdeppe, J. (2014). Conceptual, empirical, and practical issues in developing valid measures of public communication campaign exposure. *Communication Methods and Measures*, *8*(2), 138–161. https://doi.org/10.1080/19312458.2014.903391

Niederdeppe, J. (2016). Meeting the challenge of measuring communication exposure in the digital age. *Communication Methods and Measures*, *10*(2-3), 170-172. https://doi.org/10.1080/19312458.2016.1150970

Niederdeppe, J., Fowler, E. F., Goldstein, K., & Pribble, J. (2010). Does local television news coverage cultivate fatalistic beliefs about cancer prevention. *Journal of Communication*, *60*, 230–253. https://doi.org/10.1111/j.1460-2466.2009.01474.x

Niederdeppe, J., Frosch, D. L., & Hornik, R. C. (2008). Cancer news coverage and information seeking. *Journal of Health Communication*, *13*(2), 181–99. https://doi.org/10.1080/10810730701854110

Niederdeppe, J., Hornik, R. C., Kelly, B. J., Frosch, D. L., Romantan, A., Stevens, R. S., … Stevens, R. S. (2007). Examining the dimensions of cancer-related information seeking and scanning behavior. *Health Communication*, *22*(2), 153-167. https://doi.org/10.1080/10410230701454189

Niederdeppe, J., Lee, T., Robbins, R., Kim, H. K., Kresovich, A., Kirshenblat, D., … Fowler, E. F. (2013). Content and effects of news stories about uncertain cancer causes and preventive behaviors. *Health Communication*, *29*(4), 332–346. https://doi.org/10.1080/10410236.2012.755603

Niederdeppe, J., & Levy, A. G. (2007). Fatalistic beliefs about cancer prevention and three prevention behaviors. *Cancer Epidemiology Biomarkers & Prevention*, *16*(5), 998–1003. https://doi.org/10.1158/1055-9965.EPI-06-0608

Noar, S. M., Althouse, B. M., Ayers, J. W., Francis, D. B., & Ribisl, K. M. (2015). Cancer information seeking in the digital age: Effects of Angelina Jolie's prophylactic mastectomy announcement. *Medical Decision Making*, *35*(1), 16–21. https://doi.org/10.1177/0272989X14556130

Noar, S. M., Ribisl, K. M., Althouse, B. M., Willoughby, J. F., & Ayers, J. W. (2013). Using digital surveillance to examine the impact of public figure pancreatic cancer announcements on media and search query outcomes. *Journal of the National Cancer Institute. Monographs*, *2013*(47), 188–94. https://doi.org/10.1093/jncimonographs/lgt017

Noar, S. M., Willoughby, J. F., Myrick, J. G., & Brown, J. (2014). Public figure announcements about cancer and opportunities for cancer communication: A review and research agenda. *Health Communication*, *29*(5), 445–61. https://doi.org/10.1080/10410236.2013.764781

Nyhan, B. (2010). Why the "death panel" myth wouldn't die: Misinformation in the health care reform debate. *The Forum*, *8*(1). https://doi.org/10.2202/1540-8884.1354

Nyhan, B., Reifler, J., & Ubel, P. A. (2013). The hazards of correcting myths about health care reform. *Medical Care*, *51*(2), 127–32. https://doi.org/10.1097/MLR.0b013e318279486b

Oeffinger, K. C., Fontham, E. T. H., Etzioni, R., Herzig, A., Michaelson, J. S., Shih, Y.-C. T., … Wender, R. (2015). Breast cancer screening for women at average risk. *JAMA*, *314*(15), 1599. https://doi.org/10.1001/jama.2015.12783

Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractivess. *Journal of Advertising*, *19*(3), 39–52. Retrieved from http://www.jstor.org/stable/4188769

Osuch, J. R., Silk, K., Price, C., Barlow, J., Miller, K., Hernick, A., & Fonfa, A. (2012). A historical perspective on breast cancer activism in the United States: From education and support to partnership in scientific research. *Journal of Women's Health*, *21*(3), 355–362. https://doi.org/10.1089/jwh.2011.2862

Perse, E. M., & Lambe, J. (2016). *Media effects and society*. Routledge.

Pew Research Center. (2015). Public and scientists' views on science and society. Retrieved from http://pewresearch.org

Pierce, J. P., & Gilpin, E. A. (2001). News media coverage of smoking and health is associated with changes in population rates of smoking cessation but not initiation. *Tobacco Control, 102),* 145-153. https://doi.org/10.1136/tc.2005.013177

Polgreen, P. M., Chen, Y., Pennock, D. M., Nelson, F. D., & Weinstein, R. A. (2008). Using Internet searches for influenza surveillance. *Clinical Infectious Diseases*, *47*(11), 1443–1448. https://doi.org/10.1086/593098

Prasad, V., Lenzer, J., & Newman, D. H. (2016). Why cancer screening has never been shown to "save lives" - and what we can do about it. *BMJ (Online)*, 352. https://doi.org/10.1136/bmj.h6080

Probst, W. N., Stelzenmüller, V., & Fock, H. O. (2012). Using cross-correlations to assess the relationship between time-lagged pressure and state indicators: An exemplary analysis of North Sea fish population indicators. *Journal of Marine Science*, *69*(4), 670–681. https://doi.org/10.1093/icesjms/fss015

Ragas, M. W., & Tran, H. (2013). Beyond Cognitions: A longitudinal study of online search salience and media coverage of the president. *Journalism & Mass Communication Quarterly, 90*(3), 478-499. https://doi.org/10.1177/1077699013493792

Rahib, L., Smith, B. D., Aizenberg, R., Rosenzweig, A. B., Fleshman, J. M., & Matrisian, L. M. (2014). Projecting cancer incidence and deaths to 2030: The unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Research*, *74*(11), 2913–2921. https://doi.org/10.1158/0008-5472.CAN-14-0155

Riffe, D., Lacy, S., & Fico, F. (2014). *Analyzing media messages: Using quantitative content analysis in research* (3rd ed.). New York: Routledge.

Romantan, A. (2005). The social amplification of commercial aviation risks: A longitudinal analysis of media effects on air travel behavior, 1978-2001. In *International Communication Association* (pp. 1978–2001).

Sabel, M. S., & Dal Cin, S. (2016). Trends in media reports of celebrities' breast cancer treatment decisions. *Annals of Surgical Oncology, 23*(9), 2795-2801. https://doi.org/10.1245/s10434-016-5202-7

Sangalang, A., Hurley, R. J., & Tewksbury, D. (2015). Framing cancer for online news: Implications for popular perceptions of cancer. *Journal of Communication, 65*(6), 1018–1040. https://doi.org/10.1111/jcom.12183

Scheufele, D. A. (2000). Agenda-setting, priming, and framing revisited: Another look at cognitive effects of political communication. *Mass Communication & Society*, *3*(2-3), 297–316. https://doi.org/10.1207/S15327825MCS0323_07

Scheufele, D., & Tewksbury, D. (2007). Framing, agenda setting, and priming: The evolution of three media effects models. *Journal of Communication*, *57*(1), 9-20. https://doi.org/ 10.1111/j.0021-9916.2007.00326.x

Schudson, M. (2003). *The sociology of news*. New York: W.W. Norton & Company.

Semetko, H. A., & Valkenburg, P. M. (2000). Framing European politics: A content analysis of press and television news. *Journal of Communication, 50*(2), 93–109. https://doi.org/10.1111/j.1460-2466.2000.tb02843.x

Shim, M. (2008). Connecting Internet use with gaps in cancer knowledge. *Health Communication, 23*(5), 448–61. https://doi.org/10.1080/10410230802342143

Shoemaker, P. J., & Reese, S. D. (2013). *Mediating the message in the 21st century: A media sociologyperspective*. Routledge.

Shumway, R. H., & Stoffer, D. S. (2010). *Time series analysis and its applications: With R examples*. Springer Science & Business Media.

Smith, J., McCarthy, J. D., McPhail, C., & Augustyn, B. (2001). From protest to agenda building: Description bias in media coverage of protest events in Washington, D. C. *Social Forces, 79*(4). https://doi.org/10.1353/sof.2001.0053

Soumerai, S. B., Ross-Degnan, D., & Spira Kahn, J. (1992). Effects of professional and media warnings about the association between Aspirin use in children and Reye's syndrome. *The Milbank Quarterly, 70*(1), 155–182. http://www.jstor.org/stable/3350088

Southwell, B. G., & Thorson, E. A. (2015). The prevalence, consequence, and remedy of misinformation in mass media systems. *Journal of Communication*, *65*(4), 589–595. https://doi.org/10.1111/jcom.12168

Squiers, L. B., Holden, D. J., Dolina, S. E., Kim, A. E., Bann, C. M., & Renaud, J. M. (2011). The public's response to the U.S. Preventive Services Task Force's 2009 Recommendations on mammography screening. *American Journal of Preventive Medicine, 40*(5), 497–504. https://doi.org/10.1016/j.amepre.2010.12.027

Stack, S. (2000). Media impacts on suicide: A quantitative review of 293 findings. *Social Science Quarterly*, 957–971. http://www.jstor.org/stable/42864031

Stevens, R., & Hornik, R. C. (2014). AIDS in black and white: The influence of newspaper coverage of HIV/AIDS on HIV/AIDS testing among African Americans and White Americans, 1993-2007. *Journal of Health Communication*, *19*(8), 893–906. https://doi.org/10.1080/10810730.2013.864730

Stout, P. A., & Moon, Y. S. (1990). Use of endorsers in magazine advertisements. *Journalism Quarterly*, *67*(3), 536-546. https://doi.org/10.1177/107769909006700309

Stryker, J. E. (2003). Media and marijuana: A longitudinal analysis of news media effects on adolescents' marijuana use and related outcomes, 1977-1999. *Journal of Health Communication*, *8*(4), 305–28. https://doi.org/10.1080/10810730305724

Stryker, J. E. (2008). Measuring aggregate media exposure: A construct validity test of indicators of the national news environment. *Communication Methods and Measures*, *2*(1-2), 115–133. https://doi.org/10.1080/19312450802062620

Stryker, J. E., Wray, R. J., Hornik, R. C., & Yanovitzky, I. (2006). Validation of database search terms for content analysis: The case of cancer news coverage. *Journalism & Mass Communication Quarterly*, (2), 413–430. https://doi.org/10.1177/107769900608300212

Sulik, G. A. (2013). Amy Robach story spreads heartfelt misinformation. Retrieved from https://www.psychologytoday.com/blog/pink-ribbon-blues/201312/amy-robach-story-spreads-heartfelt-misinformation

Susan G. Komen. (n.d.). Age. Retrieved from http://ww5.komen.org/BreastCancer/GettingOlder.html

Thode, H. C. (2002). *Testing for normality*. CRC Press.

Thrall, a. T., Lollio-Fakhreddine, J., Berent, J., Donnelly, L., Herrin, W., Paquette, Z., … Wyatt, A. (2008). Star power: Celebrity advocacy and the evolution of the public sphere. *The International Journal of Press/Politics*, *13*(4), 362–385. https://doi.org/10.1177/1940161208319098

Towers, S., Afzal, S., Bernal, G., Bliss, N., Brown, S., Espinoza, B., … Castillo-Chavez, C. (2015). Mass media and the contagion of fear: The case of Ebola in America. *PloS One*, *10*(6), e0129179. https://doi.org/10.1371/journal.pone.0129179

Tuchman, G. (1978). *Making news: A study in the construction of reality*. London, New York: Free Press.

US Services Preventive Task Force (2009). Breast Cancer Screening. Retrieved from https://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/breast-cancer-screening

van de Rijt, a., Shor, E., Ward, C., & Skiena, S. (2013). Only 15 minutes? The social stratification of fame in printed media. *American Sociological Review, 78*(2), 266–289. https://doi.org/10.1177/0003122413480362

Viswanath, K., Blake, K. D., Meissner, H. I., Saiontz, N. G., Mull, C., Freeman, C. S., … Croyle, R. T. (2008). Occupational practices and the making of health news: A national survey of US health and medical science journalists. *Journal of Health Communication, 13*(8), 759-777. https://doi.org/10.1080/10810730802487430

Viswanath, K., & Finnegan, J. R. (2002). Reflections on community health campaigns: Secular trends and the capacity to effect change. In R. C. Hornik (Ed.), *Public health communication: Evidence for behavior change* (pp. 289–313). Mahwah, NJ: Lawrence Erlbaum Associates.

Waller, J., Douglas, E., Whitaker, K. L., & Wardle, J. (2013). Women's responses to information about overdiagnosis in the UK breast cancer screening programme: A qualitative study. *BMJ Open, 3*(4), 1–9. https://doi.org/10.1136/bmjopen-2013-002703

Waller, J., Osborne, K., & Wardle, J. (2015). Enthusiasm for cancer screening in Great Britain: a general population survey. *British Journal of Cancer, 112*(3), 562–6. https://doi.org/10.1038/bjc.2014.643

Wallington, S. F., Blake, K., Taylor-clark, K., Viswanath, K., Flynt, S., Blake, K., … Viswanath, K. (2010). Antecedents to agenda setting and framing in health news: An examination of priority, angle, source, and resource usage from a national survey of US health reporters and editors. *Journal of Health Communication, 15*(1), 76-94. https://doi.org/10.1080/10810730903460559

Walls, H. L., Peeters, A., Proietto, J., & McNeil, J. J. (2011). Public health campaigns and obesity - a critique. *BMC Public Health, 11*, 136. https://doi.org/10.1186/1471-2458-11-136

Weeks, B., Friedenberg, L. M., Southwell, B. G., & Slater, J. S. (2012). Behavioral consequences of conflict-oriented health news coverage: The 2009 mammography guideline controversy and online information seeking. *Health Communication, 27*(2), 158–66. https://doi.org/10.1080/10410236.2011.571757

Weeks, B., & Southwell, B. (2010). The symbiosis of news coverage and aggregate online search behavior: Obama, rumor, and presidential politics. *Mass Communication and Society*, *13*(4), 341–360. https://doi.org/http://doi.org/10.1080/15205430903470532

Welch, H. G., & Frankel, B. A. (2011). Likelihood that a woman with screen-detected breast cancer has had her "life saved" by that screening. *Archives of Internal Medicine*, *171*(22), 2043–6. https://doi.org/10.1001/archinternmed.2011.476

Williamson, J. M. L., & Hocken, David, B. (2010). Pancreatic cancer in the media: The Swayze shift. *Annals of the Royal College of Surgeons of England*, *92*(3), 243–5. https://doi.org/10.1308/003588410X1251883644060

Wilson, K., & Brownstein, J. S. (2009). Early detection of disease outbreaks using the Internet. *Canadian Medical Association Journal*, *180*(8), 829–831. https://doi.org/10.1503/cmaj.1090215

Wilt, T. J., & Partin, M. R. (2011). Screening: Simple messages...sometimes. *JAMA: The Journal of the American Medical Association*, *171*(22), 2046–2048. https://doi.org/10.1001/archinternmed.2011.509

Wise, D., & Brewer, P. R. (2010). Competing frames for a public health issue and their effects on public opinion. *Mass Communication and Society*, *13*(4), 435–457. https://doi.org/10.1080/15205430903296077

Yu, J., Nagler, R. H., Fowler, E. F., Kerlikowske, K., & Gollust, S. E. (2017). Women's awareness and perceived importance of the harms and benefits of mammography screening: Results from a 2016 national survey. In *Academy Health Annual Research Meeting*. New Orleans, LA.

Yue, S., Pilon, P., & Cavadias, G. (2002). Power of the Mann–Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. *Journal of Hydrology*, *259*(1), 254–271. https://doi.org/10.1016/S0022-1694(01)00594-7

Zaharopoulos, T. (2007). The news framing of the 2004 Olympic Games. *Mass Communication & Society*, *10*(2), 235–249. https://doi.org/10.1080/15205430701265752

Zillmann, D., Chen, L., Knobloch, S., & Callison, C. (2004). Effects of lead framing on selective exposure to Internet news reports. *Communication Research*, *31*(1), 58–81. https://doi.org/10.1177/0093650203260201

Appendix A

List of Celebrities-of-Interest by Date of Disclosure

	Celebrity	Date of disclosure	Age at BC event	Career type	Breast Cancer-Related Event	Level of Celebrity Status
1	Etheridge, Melissa	2/05/2005	43	6	4	10
2	Bissex, Rachel	2/20/2005	49	6	6	5
3	Minogue, Kylie	5/17/2005	37	6	4	10
4	Green, Ernie	7/2005	66	3	4	4
5	Mann, Judy	7/8/2005	61	5	6	6
6	Feldman, Sandra	9/18/2005	65	4	6	8
7	Kennedy, Joan	10/1/2005	68	7	4	4
8	Horn, Shirley	10/20/2005	71	6	6	8
9	Sperber, Wendie Jo	11/29/2005	47	2	6	3
10	Berzon, Betty	1/24/2006	78	4	6	4
11	Crow, Sheryl	2/25/2006	44	6	4	10
12	Moffo, Anna	3/9/2006	73	6	6	7
13	Nathan, Melissa	4/7/2006	37	4	6	2
14	Barnett, Lisa A.	5/2/2006	48	4	6	4
15	Lund, JoAnna	5/20/2006	61	4	6	6
16	Hunt Lieberson, Lorraine	7/3/2006	52	6	6	7
17	Frost, Kathryn	8/18/2006	57	8	6	7
18	Faithful, Marianne	9/14/2006	59	6	4	6
19	Engelberg, Miriam	10/17/2006	48	4	6	6
20	Dewar, Helen	11/4/2006	70	5	6	4
21	Ivins, Molly	1/31/2007	62	5	6	10
22	McGhee, William	2/17/2007	76	2	6	3
23	Gittings, Barbara	2/18/2007	75	4	6	7
24	Edwards, Elizabeth (D1)*	3/21/2007	58	7	3	9
25	Syler, Renee	4/16/2007	55	5	1	5
26	Saubert, Jean	5/14/2007	65	3	6	5
27	Wyler, Gretchen	5/27/2007	75	2	6	5
28	Roberts, Robin	7/31/2007	47	5	4	7
29	Paley, Grace	8/22/2007	84	4	6	9
30	Zuk, Judith D.	9/1/2007	55	4	6	5
31	Davis, JoAnn	10/6/2007	57	8	6	9
32	Kotb, Hoda	10/17/2007	43	5	4	6

	Celebrity	Date of disclosure	Age at BC event	Career type	Breast Cancer-Related Event	Level of Celebrity Status
33	Fraction, Karen	10/30/2007	49	2	6	3
34	Rabinovitch, Dina	10/30/2007	45	5	6	4
35	Baker Tharp, Carol	11/25/2007	55	8	6	2
36	Bates, Jeanne	11/28/2007	89	2	6	4
37	Hamill, Dorothy	1/4/2008	51	3	4	10
38	Frontiere, Georgia	1/18/2008	80	3	6	7
39	Smith, Maggie	3/18/2008	73	2	4	9
40	Nixon, Cynthia	4/15/2008	40	2	4	6
41	Applegate, Christina	8/4/2008	36	2	4	10
42	Pick, Heather	11/7/2008	38	5	6	5
43	Romney, Ann	12/5/2008	59	7	4	4
44	Nesler, Ellie	12/26/2008	56	9	6	3
45	van Bruggen, Coosje	1/10/2009	66	4	6	5
46	Yow, Kay	1/24/2009	66	3	6	7
47	Fiorina, Carly	3/3/2009	54	8	4	5
48	Wasserman Schultz, Debbie	3/22/2009	41	8	4	8
49	Kosofsky Sedgwick, Eve	4/12/2009	58	4	6	7
50	Nielsen, Jerri	6/23/2009	57	5	6	7
51	Forbes, Mary Lou	6/27/2009	83	8	6	3
52	Tierney, Maura	7/14/2009	44	2	4	9
53	Mendez, Olga	7/29/2009	84	8	6	4
54	Sims, Naomi	8/4/2009	61	4	6	4
55	Wexler, Anne	8/7/2009	79	8	6	5
56	Criss, Peter	10/7/2009	62	6	4	7
57	Carter, Yvonne	10/20/2009	50	8	6	3
58	Day, Linda	10/23/2009	72	4	6	5
59	Semple, Goldie	12/9/2009	56	2	6	3
60	Reed Hall, Alaina	12/17/2009	63	2	6	4
61	de Sela, Lhasa	1/1/2010	37	6	6	5
62	Lyon, Jennifer	1/19/2010	37	2	6	4
63	Gray, Rose	2/28/2010	71	4	6	6
64	Navratilova, Martina	4/7/2010	53	3	4	9
65	Redgrave, Lynn	5/2/2010	67	3	6	6
66	Stevens, Pat	5/26/2010	64	2	6	6
67	Chapman Booker, Alison	7/1/2010	47	5	6	2
68	Walker, Catherine	9/26/2010	65	7	6	7
69	Dolgin, Gail	10/7/2010	65	4	6	2

	Celebrity	Date of disclosure	Age at BC event	Career type	Breast Cancer-Related Event	Level of Celebrity Status
70	Novak, Kim	10/19/2010	77	2	4	9
71	Fonda, Jane	11/11/2010	72	2	4	6
72	Edwards, Elizabeth, (D2)	12/7/2010	61	7	6	6
73	Styrene, Poly	4/25/2011	53	6	6	5
74	Mitchell, Andrea	9/7/2011	64	5	4	10
75	Sykes, Wanda	9/23/2011	47	2	4	5
76	Rancic, Giuliana	10/17/2011	37	9	4	6
77	Ekvall, Eva	12/17/2011	28	2	6	4
78	Pence, Ellen	1/6/2012	63	4	6	5
79	Kamen Goldmark, Kathi	5/24/2012	63	4	6	6
80	Notaro, Tig	8/3/2012	41	9	4	7
81	Blume, Judy	9/5/2012	74	4	4	9
82	Bates, Kathy	9/12/2012	64	2	4	10
83	Heaton, Michelle	9/13/2012	33	6	1	2
84	Osbourne, Sharon	11/4/2012	60	9	1	6
85	Rose, Allyn (D1)	11/9/2012	27	9	1	4
86	Rose, Allyn (D2)	1/11/2013	24	9	1	4
87	Anastacia (D1)	2/28/2013	45	6	3	7
88	Gordon, Kim	4/22/2013	57	6	4	6
89	Jolie, Angelina	5/14/2013	35	2	1	10
90	Sambolin, Zorida	5/14/2013	47	5	4	6
91	Luft, Lorna	5/25/2013	60	7	4	**7**
92	Anastacia (D2)	10/1/2013	45	6	4	6
93	Brodnick, Caitlin	10/2/2013	28	9	1	3
94	Robach, Amy	11/11/2013	40	5	4	6
95	Harris, Samantha (D1)	4/9/2014	40	9	3	**6**
96	Harris, Samantha (D2)	10/24/2014	40	9	4	5
97	Lunden, Joan (D1)	6/24/2014	63	5	3	7
98	Lunden, Joan (D2)	9/24/2014	63	5	4	7
99	Harris, Samantha (D3)	10/24/2014	40	9	4	3
100	Wilson, Rita	4/14/2015	58	2	4	9
101	Lee, Sandra (D1)	5/12/2015	48	9	4	10
102	Doherty, Shannen (D1)	9/19/2015	44	2	3	8
103	Dickenson, Janice (D1)	3/28/2016	61	9	3	2
104	Lee, Sandra (D2)	4/5/2016	48	9	4	6
105	Lee, Sandra (D3)	5/18/2016	48	9	5	7
106	Rowe, Debbie	7/5/2016	57	7	4	5

	Celebrity	Date of disclosure	Age at BC event	Career type	Breast Cancer-Related Event	Level of Celebrity Status
107	Doherty, Shannen (D2)	7/20/2016	44	2	4	9
108	Dickenson, Janice (D2)	7/28/2016	61	9	4	2
109	Doherty, Shannen (D3)	8/1/2016	44	2	4	10
110	Doherty, Shannen (D4)	11/28/2016	44	2	4	8

Note. Career types are numerically coded in no particular order: 2 = actor; 3 = athlete/sports-related; 4 = activist/author/academic/creative; 5 = journalist/news anchor; 6 = musician; 7 = personal affiliation; 8 = politician/policy maker/service person; 9 = television personality.
Breast cancer-related events are numerically coded in no particular order: 1= prevention; 2 = detection; 3=diagnosis; 4= treatment; 5= survivorship; 6=death.
Level of celebrity status are numerically coded: 2 = lowest level of celebrity status; 10 = highest level of celebrity status
*Denotes the number of disclosure per celebrity. For example, D1 indicates that this disclosure is the first disclosure a celebrity made of multiple disclosures.

Appendix B

Level of Celebrity Status: Media and Audience Salience Measures

Celebrity-of-interest	Date of Disclosure	CMS	LTMS	Media Salience Total	Media Salience Rank	CAS	LTAS	Audience Salience Total	Audience Salience Rank	Level of Celebrity Status
Nathan, Melissa	4/7/2006	0	3	0	1	0	5.94	0	1	2
Baker Tharp, Carol	11/25/2007	0	1	0	1	0	0	0	1	2
Booker, Alison	7/1/2010	0	1	0	1	0	0	0	1	2
Dolgin, Gail	10/7/2010	0	20	0	1	17	2.33	39.61	1	2
Heaton, Michelle	11/2012	0	2	0	1	9	9.08	81.72	1	2
Dickenson, Janice (D1)	3/28/2016	0	12	0	1	9	2.5	22.5	1	2
Dickenson, Janice (D2)	7/28/2016	0	12	0	1	4	2.43	9.72	1	2
Sperber, Wendie Jo	11/29/2005	0	5	0	1	100	1.31	131	2	3
McGhee, William	2/17/2007	0	1	0	1	31	5.08	157.48	2	3
Fraction, Karen	10/30/2007	0	1	0	1	100	1.57	157	2	3
Nesler, Ellie	12/26/2008	0	72	0	2	0	0	0	1	3
Forbes, Mary Lou	6/27/2009	1	2	2	2	0	0	0	1	3
Carter, Yvonne	10/20/2009	0	1	0	1	32	7.31	233.92	2	3
Semple, Goldie	12/9/2009	0	7	0	1	69	1.85	127.65	2	3
Brodnick, Caitlin	10/2/2013	0	6	0	2	0	0	0	1	3
Harris, Samanath (D3)	10/24/2014	0	93	0	2	9	6.91	62.19	1	3

Celebrity-of-interest	Date of Disclosure	CMS	LTMS	Media Salience Total	Media Salience Rank	CAS	LTAS	Audience Salience Total	Audience Salience Rank	Level of Celebrity Status
Green, Ernie	7/2005	0	9	0	1	33	9.05	298.65	3	4
Kennedy, Joan	10/1/2005	1	121	121	3	1	2	2	1	4
Berzon, Betty	1/24/2006	0	1	0	1	71	4.49	318.79	3	4
Barnett, Lisa A.	5/2/2006	0	1	0	1	75	3.63	272.25	3	4
Dewar, Helen	11/4/2006	3	3	9	2	39	5.95	232.05	2	4
Rabinovitch, Dina	10/30/2007	0	1	0	1	69	4.71	324.99	3	4
Bates, Jeanne	11/28/2007	1	2	2	2	18	8.56	154.08	2	4
Romney, Ann	12/5/2008	2	76	152	3	4	6.31	25.24	1	4
Mendez, Olga	7/29/2009	0	35	0	2	26	5.16	134.16	2	4
Sims, Naomi	8/4/2009	1	19	19	3	100	0.72	72	1	4
Reed Hall, Alaina	12/17/2009	1	16	16	3	100	0.52	52	1	4
Lyon, Jennifer	1/19/2010	1	8	8	2	100	1.09	109	2	4
Ekvall, Eva	12/17/2011	5	6	30	3	100	0.44	44	1	4
Allyn Rose (D1)	11/9/2012	0	1	0	1	100	2.97	297	3	4
Allyn Rose (D2)	1/11/2013	3	2	6	2	100	1.92	192	2	4
Bissex, Rachel	2/20/2005	0	24	0	1	100	7.72	772	4	5
Syler, Rene	4/16/2007	0	25	0	2	42	7.14	299.88	3	5
Saubert, Jean	5/14/2007	3	9	27	3	39	5.94	231.66	2	5
Wyler, Gretchen	5/27/2007	2	18	36	3	100	2.15	215	2	5
Zuk, Judith D.	9/1/2007	1	22	22	3	44	4.89	215.16	2	5
Pick, Heather	11/7/2008	1	1	1	2	100	2.84	284	3	5

Celebrity-of-interest	Date of Disclosure	CMS	LTMS	Media Salience Total	Media Salience Rank	CAS	LTAS	Audience Salience Total	Audience Salience Rank	Level of Celebrity Status
van Bruggen, Coosje	1/10/2009	4	118	472	3	16	9.25	148	2	5
Fiorina, Carly	3/3/2009	2	1005	2010	4	7	10.56	73.92	1	5
Wexler, Anne	8/7/2009	0	142	0	2	100	3.58	358	3	5
Day, Linda	10/23/2009	0	28	0	2	13	22.26	289.38	3	5
de Sela, Lhasa	1/1/2010	1	6	6	2	100	3.32	332	3	5
Styrene, Poly	4/25/2011	3	9	27	3	100	1.12	112	2	5
Sykes, Wanda	9/23/2011	3	608	1824	4	19	3.63	68.97	1	5
Pence, Ellen	1/6/2012	0	5	0	1	100	6.65	665	4	5
Harris, Samanath (D2)	6/20/2014	0	93	0	2	50	7.6	380	3	5
Rowe, Debbie	7/5/2016	1	484	484	4	2	3.73	7.46	1	5
Mann, Judy	7/8/2005	2	9	18	3	60	4.61	276.6	3	6
Lund, JoAnna	5/20/2006	0	4	0	1	55	19.48	1071.4	5	6
Faithfull, Marianne	9/14/2006	0	292	0	2	39	15.38	599.82	4	6
Engelberg, Miriam	10/17/2006	2	2	4	2	100	6.29	629	4	6
Kotb, Hoda	10/17/2007	0	12	0	1	100	19.72	1972	5	6
Nixon, Cynthia	4/15/2008	3	508	1524	4	29	5.45	158.05	2	6
Gray, Rose	2/28/2010	0	25	0	2	97	7.58	735.26	4	6
Redgrave, Lynn	5/2/2010	10	1311	13110	5	100	0.59	59	1	6
Stevens, Pat	5/26/2010	0	1	0	1	43	21.09	906.87	5	6
Fonda, Jane	11/11/2010	0	4996	4996	5	17	5.9	100.3	1	6
Edwards, Elizabeth, (D2)	12/7/2010	18	927	16686	5	100	0.954	95.4	1	6

Celebrity-of-interest	Date of Disclosure	CMS	LTMS	Media Salience Total	Media Salience Rank	CAS	LTAS	Audience Salience Total	Audience Salience Rank	Level of Celebrity Status
Rancic, Giuliana	10/17/2011	5	73	365	3	100	3.33	333	3	6
Kamen Goldmark, Kathi	5/24/2012	1	13	13	3	100	4.12	412	3	6
Osbourne, Sharon	11/4/2012	2	930	1860	4	32	7.22	231.04	2	6
Gordon, Kim	4/22/2013	0	334	0	2	58	10.74	622.92	4	6
Sambolin, Zoraida	5/14/2013	3	16	48	3	100	2.99	299	3	6
Anastacia (D2)	10/1/2013	0	152	0	3	12	22.89	274.68	3	6
Robach, Amy	11/11/2013	15	65	975	4	100	2.14	214	2	6
Harris, Samanath (D1)	4/9/2014	3	92	276	3	59	7.03	414.77	3	6
Lee, Sandra (D2)	4/5/2016	1	588	588	4	14	14.06	196.84	2	6
Moffo, Anna	3/9/2006	3	107	321	3	100	6.84	684	4	7
Hunt Lieberson, Lorraine	7/3/2006	5	147	735	4	100	3.67	367	3	7
Frost, Kathryn	8/18/2006	3	6	18	3	100	7.25	725	4	7
Gittings, Barbara	2/18/2007	4	5	20	3	100	4.45	445	4	7
Roberts, Robin	7/31/2007	23	288	6624	5	100	2.32	232	2	7
Frontiere, Georgia	1/18/2008	6	455	2730	4	100	3.51	351	3	7
Yow, Kay	1/24/2009	38	636	24168	5	100	1.52	152	2	7
Kosofsky Sedgwick, Eve	4/12/2009	3	18	54	3	100	4.22	422	4	7
Nielsen, Jerri	6/23/2009	8	123	984	4	100	3.03	303	3	7
Criss, Peter	10/7/2009	0	82	0	2	100	21.09	2109	5	7
Walker, Catherine	9/26/2010	1	67	67	3	36	13.46	484.56	4	7
Notaro, Tig	8/3/2012	1	11	11	2	100	9.81	981	5	7

Celebrity-of-interest	Date of Disclosure	CMS	LTMS	Media Salience Total	Media Salience Rank	CAS	LTAS	Audience Salience Total	Audience Salience Rank	Level of Celebrity Status
Anastacia (D1)	2/28/2013	2	147	294	3	20	24.68	493.6	4	7
Luft, Lorna	5/25/2013	0	281	0	2	49	29.28	1434.72	5	7
Lunden, Joan (D1)	6/24/2014	2	864	1728	4	100	2.94	294	3	7
Lunden, Joan (D2)	9/24/2014	1	868	868	4	100	3.34	334	3	7
Lee, Sandra (D3)	5/18/2016	9	590	5310	5	16	14.57	233.12	2	7
Feldman, Sandra	9/18/2005	3	776	2328	4	100	5.11	511	4	8
Horn, Shirley	10/20/2005	3	490	1470	4	100	7.24	724	4	8
Wasserman Schultz, Debbie	3/22/2009	3	172	516	4	89	9.07	807.23	4	8
Doherty, Shannen (D1)	9/19/2015	1	971	971	4	45	16.31	733.95	4	8
Doherty, Shannen (D4)	11/28/2016	1	996	996	4	56	8.62	482.72	4	8
Edwards, Elizabeth (D1)	3/21/2007	74	270	19980	5	100	4.7	470	4	9
Paley, Grace	8/22/2007	3	264	792	4	100	10.4	1040	5	9
Davis, Jo Ann	10/6/2007	13	211	2743	5	100	5.32	532	4	9
Smith, Maggie	3/18/2008	1	1391	1391	4	79	30.16	2382.64	5	9
Tierney, Maura	7/14/2009	4	199	796	4	100	8.92	892	5	9
Navratilova, Martina	4/7/2010	6	11774	70644	5	85	7.56	642.6	4	9
Novak, Kim	10/19/2010	4	553	2212	4	100	12.65	1265	5	9
Blume, Judy	9/5/2012	2	630	1260	4	52	30.97	1610.44	5	9
Wilson, Rita	4/14/2015	11	937	10307	5	100	6.68	668	4	9
Doherty, Shannen (D2)	7/20/2016	3	980	2940	5	70	8.86	620.2	4	9
Etheridge, Melissa	2/05/2005	15	665	9975	5	100	11.61	1161	5	10

Celebrity-of-interest	Date of Disclosure	CMS	LTMS	Media Salience Total	Media Salience Rank	CAS	LTAS	Audience Salience Total	Audience Salience Rank	Level of Celebrity Status
Minogue, Kylie	5/17/2005	27	200	5400	5	100	15.43	1543	5	10
Crow, Sheryl	2/25/2006	8	1557	12456	5	86	27.91	2400.26	5	10
Ivins, Molly	1/31/2007	13	225	2925	5	100	9.58	958	5	10
Hammill, Dorothy	1/4/2008	5	654	3270	5	100	11.33	1133	5	10
Applegate, Christina	8/4/2008	7	494	3458	5	100	10.09	1009	5	10
Mitchell, Andrea	9/7/2011	4	1117	4468	5	92	9.78	899.76	5	10
Bates, Kathy	9/12/2012	6	1913	11478	5	71	17.52	1243.92	5	10
Jolie, Angelina	5/14/2013	117	7797	912249	5	75	14.7	1102.5	5	10
Lee, Sandra	5/12/2015	36	488	17568	5	89	17.27	1537.03	5	10
Doherty, Shannen (D3)	8/1/2016	11	983	10813	5	100	8.82	882	5	10

Note. The above table gives the current media salience (CAS), long-term media salience (LMTS), current audience salience (CAS), and long-term audience salience (LTAS) values. The media salience total and audience salience total columns provide the weighted results of the respective current and long-term measures. The media salience rank and audience salience rank columns provide the quartile ranking results. The level of celebrity status column results from adding the two columns together.

Appendix C

Breast Cancer-Related Information Seeking Dimensions: Content Validity

Breast Cancer-Related Information Seeking Dimensions, Search Term, Search Equation, and Related Search Queries, 2005-2016

Dimension/Search Term	Search Equation	Related Search Queries
General Breast Cancer		
Breast cancer	Breast cancer -awareness -month -walk -pink -ribbon -society	Breast cancer symptoms (100%); cancer symptoms (100%); what is breast cancer (50%); breast cancer treatment (50%); breast cancer signs (45%); symptoms of breast cancer (40%); signs of breast cancer (35%) breast pain (30%); breast cancer pain (30%); inflammatory breast cancer (30%); breast cancer lump (25%); breast cancer radiation (25%); triple negative breast cancer (25%); men breast cancer (20%); breast cancer causes (20%); breast cancer research (20%); breast cancer picture (20%); metastatic breast cancer (20%); breast cancer surgery (20%); breast cancer surgery (20%); treatment for breast cancer (20%); chemotherapy (15%); breast cancer statistics (15%); stage 4 breast cancer (15%)
Prevention		
Breast cancer risk	Breast cancer risk -lung	risk of breast cancer (100%); breast cancer risk factors (40%); risk factors for breast cancer (20%); what is breast cancer (20%); risk factors of breast cancer (15%); risk assessment (15%); breast cancer risk assessment (10%); breast cancer risk calculator (10%); breast cancer symptoms (10%); breast cancer screening (10%); symptoms of breast cancer (5%); gail model breast cancer risk (5%); gail model (5%); mammogram (5%); gail risk model (5%); breast cancer risk assessment tool (5%); causes of breast cancer (5%); breast cancer prevention (5%); mastectomy (5%); breast cancer statistics (5%); tamoxifen (5%); 22) what causes breast cancer (5%); alcohol and breast cancer risk (5%); lifetime risk of breast cancer (5%); types of breast cancer (5%)
Breast cancer causes	Breast cancer causes -awareness	causes of cancer (100%); causes of breast cancer (100%); what causes cancer (90%); what causes breast cancer (90%); what is cancer (25%); what is breast cancer (25%); what are the cause of breast cancer (5%); abortion causes breast cancer (5%); breast cancer statistics (5%)

Dimension/Search Term	Search Equation	Related Search Queries
Breast cancer prevention	Breast cancer prevention -treatment -awareness -fund -symptoms -society -angelina -lung	prevention of breast cancer (100%); tamoxifen (15%); breast cancer prevention institute (10%); breast cancer statistics (5%); causes of breast cancer (5%)
BRCA	brca	brca gene (100%); brca testing (85%); brca breast cancer (45%); breast cancer (45%); brca test (40%); brca mutation (30%); brca positive (20%); brca2 (20%); what is brca (15%); brca genetic testing (15%); breast cancer gene (15%); genetic testing (15%); ovarian cancer (10%); myriad (10%); myriad brca (10%); brca2 (10%); brca mutations (10%); brca 1 and 2 (10%); brca gene mutation (5%); brca screening (5%); brca genes (5%); brca testing cost (5%); brac (5%)
Detection		
Breast cancer detection	Breast cancer detection -society -pubmed	breast cancer early detection (100%); detection of breast cancer (90%); early detection of breast cancer (45%); breast cancer awareness (10%); breast cancer prevention (10%); breast cancer symptoms (10%); breast cancer statistics (10%); breast cancer facts (5%)
Mammography	Mammography -jobs -technologist -salary	breast mammography (100%); mammogram (80%); digital mammography (65%); mammography screening (50%); breast cancer (40%); breast cancer (40%); 3d mammography (40%); mobile mammography (30%); what is mammography (25%); mammograms (20%); mammography guidelines (20%); acr mammography (15%); acr (15%); solis mammography (15%); solis (15%); diagnostic mammography (15%); mri (10%); tomosynthesis mammography (10%); tomosynthesis (10%); fda (10%); screening mammogram (10%); what is a mammography (10%); hologic (10%); mammography recommendations (10%)
Breast cancer screening	Breast cancer screening -society -symptoms	screening for breast cancer (100%); breast cancer screening guidelines (50%); free breast cancer screening (30%); mammogram (30%); mammogram screening (30%); uspstf (20%); uspstf breast cancer screening (20%); uspstf breast cnacer (20%); breast cancer screening recommendations (15%); new breast cancer screening (10%); acog breast cancer screening (10%); breast cancer awareness (10%); breast cancer prevention (10%); acog (5%); acs breast cancer screening (5%); breast cancer statistics (5%); breast cancer screening icd (5%); planned parenthood breast cancer screening ; 19) mammogram guidelines; 20) acs breast cancer screening guidelines ; breast cancer awareness month (5%); new breast cancer screening guidelines (5%)

Dimension/Search Term	Search Equation	Related Search Queries
Breast cancer symptoms	Breast cancer symptoms	symptoms of cancer (100%); symptoms of breast cancer (100%); breast pain (20%); symptoms for breast cancer (20%); breast cancer signs symptoms (15%); breast cancer signs (15%); cancer symptoms in women (15%); breast cancer symptoms in women (15%); signs of breast cancer (10%); signs of cancer (10%); breast cancer lump symptoms (10%); what is cancer (10%); symptoms of breast cancer in women (10%); symptoms of cancer in women (10%); breast lump (10%); breast cancer signs
Diagnosis		
Breast biopsy	Breast biopsy	Biopsy of breast (100%); breast cancer biopsy (95%); breast cancer (95%); biopsy for breast (70%); needle biopsy breast (60%); needle biopsy (60%); stereotactic breast biopsy (40%); stereotactic biopsy (40%); stereotactic (40%); core biopsy (40%); breast ultrasound (30%); breast biopsy results (30%); ultrasound biopsy (30%); biopsy of the breast (30%); breast lump biopsy (20%); biopsy for breast cancer (20%); breast biopsy procedure (20%); breast biopsy cpt code (20%); what is a biopsy (20%); what is a breast biopsy (20%); breast mri (20%); breast biopsy pain (15%); needle biopsy of breast (15%)
Breast cancer stages	Breast cancer stages	stages of breast cancer (100%); stages of cancer (100%); breast cancer symptoms (15%); breast cancer treatment (15%); what is breast cancer (10%); symptoms of breast cancer (10%); what are the stages of breast cancer (10%); what are the stages of cancer; breast cancer survival rate (10%); stage 4 breast cancer (5%); early stages of breast cancer (5%); stage 2 breast cancer (5%); stage 3 breast cancer (5%); breast cancer survival rates (5%); different stages of breast cancer (5%); types of breast cancer (5%); inflammatory breast cancer (5%); how many stages of cancer are there (5%); breast cancer statistics (5%); how many stages of breast cancer are there (5%); metastatic breast cancer (5%); signs of breast cancer (5%); chemotherapy (5%)
Breast cancer type	Breast cancer type	type of breast cancer (100%); breast cancer types (15%); most common type of breast cancer (10%); types of breast cancer (10%); breast cancer symptoms (10%); symptoms of breast cancer (10%); triple negative breast cancer (5%); worst type of bc (5%)
Treatment		
Breast cancer treatment	Breast cancer treatment - symptoms	treatment for breast cancer (100%); treatment of breast cancer (75%); breast cancer and treatment (50%); radiation treatment (25%); breast cancer radiation treatment (25%); breast cancer radiation (25%); radiation (25%); what is breast cancer (15%); radiation for breast cancer (15%); chemotherapy (15%); breast cancer treatment options (10%); metastatic breast cancer treatment (10%); triple negative breast cancer treatment (10%); metastatic breast cancer (10%); triple negative breast cancer treatment (10%)

Dimension/Search Term	Search Equation	Related Search Queries
Lumpectomy	Lumpectomy -radiation	Lumpectomy (100%); mastectomy (70%); breast cancer lumpectomy (65%); breast cancer (65%); lumpectomy surgery (55%); what is lumpectomy (35%); lumpectomy recovery (35%); breast surgery (25%); mastectomy vs lumpectomy (25%); lumpectomy procedure (20%); lump (15%); lumpectomy reconstruction (10%); lumpectomy definition (10%); lumpectomy scar (10%); breast cancer surgery (10%); double lumpectomy (10%); dcis (10%); lumpectomy pictures (10%); breast reconstruction (10%); lumpectomy bras (5%); lymph nodes (5%); seroma (5%); partial mastectomy (5%); lumpectomy vs. mastectomy (5%)
Breast cancer surgery	Breast cancer surgery –jolie –games –hysterectomy –rancic –awareness –vasectomy –society	Surgery for breast cancer (100%); mastectomy (30%); mastectomy surgery (30%); brast cancer treatment (30%); plastic surgery (25%); breast reconstruction surgery (25%); breast cancer reconstruction (25%); reconstructive breast surgery (15%); chemotherapy (15%); breast cancer surgery recovery (15%); breast cancer symptoms (15%); breast implants (10%); breast reduction surgery (10%); lymph nodes (10%); breast augmentation (5%); breast cancer survival rate (5%); male breast cancer (5%); lymphedema (5%); inflammatory breast cancer (5%); statistics (5%)
Survivorship		
Surviving breast cancer	Surviving breast cancer	No related queries
Breast cancer recurrence	Breast cancer recurrence	Side effects of breast cancer (100%); breast cancer radiation (80%); radiation side effects (80%); radiation (75%) breast radiation side effects (75%); radiation side effects for breast cancer (40%); radiation for breast cancer (40%); breast cancer treatment (35%); side effects of radiation (35%); chemotherapy side effects (30%); breast cancer chemotherapy (30%); chemotherapy (30%); side effects of radiation for breast cancer (25%); chemo side effects (25%); tamoxifen side effects (25%); radiation therapy side effects (20%); radiation therapy (20%); tamoxifen (20%); breast cancer symptoms (20%); radiation treatment side effects (15%); chemotherapy for breast cancer (15%); side effects of chemotherapy (15%); radiation therapy for breast cancer (15%); side effects of radiation therapy (10%); breast cancer drugs (10%)
Breast cancer reconstruction	Breast cancer reconstruction	Mastectomy (100%); breast cancer surgery (85%); breast reconstruction surgery (85%); breast reconstruction after mastectomy (60%); breast implants (30%); breast cancer reconstruction photos (30%); breast reconstruction photos (25%); breast cancer symptoms (10%)

Breast Cancer-Related Information Seeking Dimensions, Search Term, Search Equation, and Related Search Queries, 2005-2008

Dimension/Search Term	Search Equation	Related Search Queries
General Breast Cancer		
Breast Cancer	Breast cancer -walk – foundation -susan -american -society -month -ribbon -day	symptoms breast cancer (100%); cancer symptoms (100%); breast cancer treatment (70%); inflammatory breast cancer (70%); breast cancer stage (65%); breast cancer risk (40%); symptoms of breast cancer (40%); breast cancer signs (35%); breast cancer survival (35%); cancer research (30%); breast cancer research (30%); what is breast cancer (30%); breast cancer pain (30%); breast pain (30%); breast cancer statistics (30%); breast cancer radiation (25%); chemotherapy (25%); breast cancer chemotherapy (25%); breast cancer site (25%); metastatic breast cancer (25%); signs of breast cancer (25%); breast cancer causes (20%); breast cancer pictures (20%); men breast cancer (20%)
Prevention		
Breast cancer risk	breast cancer risk	breast cancer risk factors (100%); risk factors for breast cancer (35%); risk factors of breast cancer (25%); breast cancer risk assessment (25%); fellatio may significantly decrease the risk of breast cancer in women (25%); breast cancer risk calculator (20%); breast cancer risk assessment tool (5%)
Breast cancer causes	Breast cancer causes -symptoms -inflammatory	causes of breast cancer (100%); what causes breast cancer (55%)
Breast cancer prevention	Breast cancer prevention -fund	prevention of breast cancer (100%); breast cancer prevention diet (15%)
BRCA	brca	brca gene (100%); brca testing (80%); brca test (40%); brca mutation (40%); brca 1 gene (20%); brca genetic testing (20%); brca gene testing (20%); brca genes (15%); braca (10%)
Detection		
Breast cancer detection	Breast cancer detection	Breast cancer early detection (100%); early detection of breast cancer (35%)
Mammography	Mammography -jobs	digital mammography (100%); mammogram (50%); breast cancer (40%); mammography screening (30%); mobile mammography (15%); acr mammography (15%); acr (15%); mammography guidelines (10%); national mammography day (5%); digital mammography (5%)
Breast cancer screening	Breast cancer screening	Breast cancer screening guidelines (100%)
Breast cancer symptoms	Breast cancer symptoms -ovarian -lung	symptoms of cancer (100%); symptoms of breast cancer (100%); breast pain (20%); breast cancer signs (15%); symptoms for breast cancer (15%); signs of breast cancer (10%); breast cancer signs and symptoms (10%); inflammatory breast cancer symptoms (10%);signs and symptoms of breast cancer (5%);male breast cancer (5%)

Dimension/Search Term	Search Equation	Related Search Queries
Diagnosis		
Breast biopsy	Breast biopsy	breast cancer biopsy (100%); breast cancer (95%); breast needle biopsy (70%); needle biopsy (65%); stereotactic biopsy (55%); stereotactic breast biopsy (55%); breast core biopsy (45%); breast lump biopsy (25%); breast biopsy results (20%); mri breast biopsy (15%); surgical breast biopsy (15%); breast biopsy procedure (15%); excisional breast biopsy (15%); breast calcification (10%); core biopsy of breast (10%); ultrasound guided breast biopsy (5%); mri guided breast biopsy (5%); breast biopsy procedures (5%)
Breast cancer stages	Breast cancer stages	stages of cancer (100%); stages of breast cancer (100%)
Breast cancer type	Breast cancer type	No related queries
Treatment		
Breast cancer treatment	Breast cancer treatment -society	treatment for breast cancer (100%); treatment of breast cancer (90%); breast cancer radiation treatment (30%); breast cancer symptoms (15%); radiation treatment for breast cancer (15%); breast cancer research and treatment (10%); breast cancer treatment options (10%); inflammatory breast cancer (5%); breast cancer treatment guidelines; stages of breast cancer (5%)
Lumpectomy	lumpectomy	cancer lumpectomy (50%); lumpectomy surgery (25%); lumpectomy recovery (15%); lumpectomy procedure (10%); what is a lumpectomy (5%)
Breast cancer surgery	Breast cancer surgery	No related queries
Survivorship		
Surviving breast cancer	Surviving breast cancer	No related queries
Breast cancer recurrence	Breast cancer recurrence	No related queries
Breast cancer reconstruction	Breast cancer reconstruction	No related queries

Note. The search dimensions (in bold) are the NCI Cancer Control Continuum search categories, which are the final outcome variable for Study 2. The search terms are the data retrieved to construct the search dimensions. The search equations were the word strings entered into Google Trends to retrieve data for the given search term. A negative sign (-) preceding a word indicated that that word was removed from data retrieval. The percentage (%) score following each related search term, indicates the percentage of search queries for this term in the same search session as the final search term.

Breast Cancer-Related Information Seeking Dimensions, Search Term, Search Equation, and Related Search Queries, 2009-2012

Dimension/Search Term	Search Equation	Related Search Queries
General Breast Cancer		
Breast cancer	Breast cancer -awareness -walk -month -pink -ribbon -susan -society -foundation	breast cancer symptoms (100%); cancer symptoms (95%); breast cancer treatment (40%); what is cancer (40%); what is breast cancer (40%); symptoms of breast cancer (40%); breast cancer signs (30%); inflammatory breast cancer (25%); breast pain (25%); signs of breast cancer (25%); men breast cancer (25%); radiation (20%); breast cancer statistics (20%); breast cancer radiation (20%); breast lump (20%); breast cancer pictures (20%); triple negative breast cancer (20%); chemotherapy (15%); metastatic breast cancer (15%); breast cancer surgery (15%); breast cancer rates (15%); breast cancer stages (15%); strides against breast cancer (15%); breast cancer research (15%); making strides against breast cancer (15%)
Prevention		
Breast cancer risk	Breast cancer risk -awareness	risk of breast cancer (100%); breast cancer risk factors (45%); risk factors for breast cancer (25%); risk factors of breast cancer (20%); breast cancer risk assessment (20%); breast cancer risk calculator (10%); fellatio may significantly decrease the risk of breast cancer in women (5%); breast cancer risk assessment tool (5%); gail model (5%); what are the risk factors for breast cancer (5%); alcohol and breast cancer risk (5%)
Breast cancer causes	Breast cancer causes -symptoms -pain -awareness -inflammatory -hiccups -signs	causes of breast cancer (100%); what causes breast cancer (85%); what causes cancer (80%); what is breast cancer (25%); what are the causes of breast cancer (10%); abortion causes breast cancer (5%); breast cancer facts (5%)
Breast cancer prevention	Breast cancer prevention -fund -awareness -symptoms -month	prevention of breast cancer (100%); breast cancer prevention diet (25%); breast cancer prevention institute (10%)
BRCA	brca	brca gene (100%); brca testing (75%); breast cancer (55%); brca breast cancer (55%); brca test (45%); brca mutation (35%); brca1 (20%); brca positive (20%) brca genetic testing (15%); myriad (15%); brca2 (15%); brca gene testing (15%); brca 1 and 2 (10%); brac (10%); brca gene mutation (10%); brca genes (10%); brca screening (10%); braca (5%); brca testing cost (5%); 20) brca 1 gene (5%); myriad genetics (5%); brca mutation testing (5%)

Dimension/Search Term	Search Equation	Related Search Queries
Detection		
Breast cancer detection	Breast cancer detection	breast cancer early detection (100%); early detection of breast cancer (50%); breast cancer awareness (10%)
Mammography	Mammography -jobs -training	digital mammography (100%); mammogram (90%); screening mammography (60%); breast cancer (60%); mobile mammography (35%); mammography guidelines (30%); acr mammography (25%); what is mammography (20%); acr (20%); diagnostic mammography (20%); 3d mammography (15%); mammography salary (15%); mqsa (10%); mammography recommendations (10%); mammography machine (10%); tomosynthesis mammography (10%); mammography screening guidelines (10%); holgic (5%); digital mammography vs film mammography (5%); art (5%); mobile mammography van; mammography (5%); new mammography guidelines (5%)
Breast cancer screening	Breast cancer screening -society	screening for breast cancer (100%); breast cancer screening guidelines (60%); free breast cancer screening (45%); uspstf (25%); uspstf breast cancer screening (25%); breast cancer screening recommendations (20%); new breast cancer screening (15%); planned parenthood breast cancer screening (10%)
Breast cancer symptoms	Breast cancer symptoms -lung	symptoms of breast cancer (100%); symptoms of cancer (100%); breast cancer symptoms pain (25%); breast pain (25%); symptoms for breast cancer (25%); breast cancer symptoms in women (20%); breast cancer signs (15%); symptoms of breast cancer in women (15%); breast cancer symptoms and signs (15%); signs of breast cancer (10%); what is breast cancer (10%); inflammatory breast cancer symptoms (10%); inflammatory breast cancer (10%); pain in breast (10%); what are symptoms of breast cancers (10%); breast cancer symptoms pictures (5%); signs and symptoms of breast cancer (5%); breast cancer pictures (5%); what are the symptoms of breast cancer (5%); male breast cancer symptoms (5%); male breast cancer (5%); causes of breast cancer; lump in breast (5%); breast cancer lumps (5%)
Diagnosis		
Breast biopsy	Breast biopsy	Biopsy of breast (100%); breast cancer (95%); breast cancer biopsy (90%); breast biopsy needle (65%); stereotactic breast biopsy (45%); stereotactic (45%); stereotactic biopsy (45%); core biopsy breast (40%); core biopsy (40%); breast biopsy results (30%); ultrasound breast biopsy (30%); ultrasound biopsy (30%); breast biopsy procedure (25%); biopsy on breast (25%); breast biopsy procedure (25%); biopsy on breast (25%); breast lump biopsy (20%); breast mri (20%); biopsy for breast cancer (20%); core needle biopsy breast (15%); needle biopsy of breast (15%)

Dimension/Search Term	Search Equation	Related Search Queries
Breast cancer stages	Breast cancer stages	stages of cancer (100%); stages of breast cancer (100%); breast cancer symptoms (15%); what are the stages of breast cancer (10%); breast cancer survival rates (5%); early stages of breast cancer (5%); breast cancer statistics (5%); inflammatory breast cancer (5%); types of breast cancer (5%); stage 3 breast cancer (5%); different stages of breast cancer (5%); how many stages of breast cancer (5%); stage 1 breast cancer (5%)
Breast cancer type	Breast cancer type	No related queries
Treatment		
Breast cancer treatment	Breast cancer treatment -society	treatment for breast cancer (100%); treatment of breast cancer (80%); radiation treatment (30%); breast cancer radiation treatment; what is breast cancer (15%); chemotherapy (15%); breast cancer treatment options (10%); breast cancer treatments (10%); metastatic breast cancer (10%); breast cancer research and treatment (10%); triple negative breast cancer treatment (10%); breast cancer stages (5%); triple negative breast cancer (5%); tamoxifen (5%); breast cancer treatment guidelines (5%); stage 1 breast cancer treatment (5%); stage 4 breast cancer; breast cancer treatment drugs (5%)
Lumpectomy	Lumpectomy -Rancic	Breast lumpectomy (100%); breast cancer lumpectomy (50%); breast cancer (50%); mastectomy (45%); lumpectomy surgery (35%); lumpectomy recovery (25%); what is lumpectomy (20%); radiation after lumpectomy (15%); lumpectomy procedure (15%); lumpectomy vs mastectomy (15%); what is a lumpectomy (15%); lump (10%); lumpectomy recovery time (10%); lumpectomy and radiation (10%); double lumpectomy (10%); lumpectomy picture (5%)
Treatment		
Breast cancer treatment	Breast cancer treatment -society	treatment for breast cancer (100%); treatment of breast cancer (80%); radiation treatment (30%); breast cancer radiation treatment; what is breast cancer (15%); chemotherapy (15%); breast cancer treatment options (10%); breast cancer treatments (10%); metastatic breast cancer (10%); breast cancer research and treatment (10%); triple negative breast cancer treatment (10%); breast cancer stages (5%); triple negative breast cancer (5%); tamoxifen (5%); breast cancer treatment guidelines (5%); stage 1 breast cancer treatment (5%); stage 4 breast cancer; breast cancer treatment drugs (5%)

Dimension/Search Term	Search Equation	Related Search Queries
Lumpectomy	Lumpectomy -Rancic	Breast lumpectomy (100%); breast cancer lumpectomy (50%); breast cancer (50%); mastectomy (45%); lumpectomy surgery (35%); lumpectomy recovery (25%); what is lumpectomy (20%); radiation after lumpectomy (15%); lumpectomy procedure (15%); lumpectomy vs mastectomy (15%); what is a lumpectomy (15%); lump (10%); lumpectomy recovery time (10%); lumpectomy and radiation (10%); double lumpectomy (10%); lumpectomy picture (5%); breast cancer treatment (5%); recovery from lumpectomy (5%); lumpectomy without radiation (5%); breast reconstruction after lumpectomy (5%); lumpectomy bras (5%); lumpectomy vs. mastectomy (5%)
Breast cancer surgery	Breast cancer surgery	Surgery for breast cancer (100%); breast cancer treatment (35%); breast cancer symptoms (20%); chemotherapy (15%); breast cancer survival rates (5%); male breast cancer (5%)
Survivorship		
Surviving breast cancer	Surviving breast cancer	No related queries
Breast cancer recurrence	Breast cancer recurrence	No related queries
Breast cancer reconstruction	Breast cancer reconstruction	No related queries

Note. The search dimensions (in bold) are the NCI Cancer Control Continuum search categories, which are the final outcome variable for Study 2. The search terms are the data retrieved to construct the search dimensions. The search equations were the word strings entered into Google Trends to retrieve data for the given search term. A negative sign (-) preceding a word indicated that that word was removed from data retrieval. The percentage (%) score following each related search term, indicates the percentage of search queries for this term in the same search session as the final search term.

Breast Cancer-Related Information Seeking Dimensions, Search Term, Search Equation, and Related Search Queries, 2013-2016

Dimension/Search Term	Search Equation	Related Search Queries
General Breast Cancer		
Breast cancer	Breast cancer -walk -month -ribbon -pink -shirts -foundation -awareness	breast cancer symptoms (100%); cancer symptoms (100%); what is breast cancer (60%); what is cancer (60%); breast cancer signs (50%); symptoms of breast cancer (45%); symptoms of cancer (45%); signs of breast cancer (45%); breast pain (35%); breast cancer lump (30%); breast cancer triple negative (30%); inflammatory breast cancer (25%); breast cancer radiation (25%); men breast cancer (20%); metastatic breast cancer (20%); breast cancer surgery (20%); breast cancer pictures (20%); stage 4 breast cancer (20%); treatment for breast cancer (20%); stage 4 cancer (20%); breast cancer screening (20%); breast cancer survival rate (20%); breast cancer test (15%); chemotherapy (15%)
Prevention		
Breast cancer risk	Breast cancer risk	risk of breast cancer (100%); cancer risk factors (40%); breast cancer risk factors (40%); what is breast cancer (20%); risk factors for breast cancer (20%); risk factors of breast cancer (15%); breast cancer risk assessment (15%); breast cancer risk calculator (10%); breast cancer risk assessment tool (5%); gail model (5%); lifetime risk of breast cancer (5%); gail risk model (5%); dense breast tissue cancer risk (5%); tamoxifen (5%); what are the risk factors for breast cancer (5%)
Breast cancer causes	Breast cancer causes -symptoms -pain -signs -types -inflammatory -awareness	causes of breast cancer (100%); what causes cancer (100%); what causes breast cancer (100%); what is breast cancer (30%); what are the causes of breast cancer (10%)
Breast cancer prevention	Breast cancer prevention -anagelina -symptoms -awareness -fund	prevention of breast cancer (100%); breast cancer prevention diet (15%); tamoxifen (10%); breast cancer prevention institute (5%)
BRCA	brca	brca gene (100%); brca testing (95%); brca test (45%); brca breast cancer (40%); breast cancer (40%); brca mutation ((30%); brca positive (20%); what is brca (15%); brca1 (15%); brca gene testing (15%); brca genetic testing (15%); genetic testing (15%); breast cancer gene (15%); ovarian cancer (10%); brca gene test (10%); myriad brca (10%); brca2 (10%); brca gene mutation (10%); brca testing cost (10%); brca mutations (5%); brca test cost

Dimension/Search Term	Search Equation	Related Search Queries
Breast cancer detection	Breast cancer detection	breast cancer early detection (100%); early detection of breast cancer (55%); breast cancer awareness (10%)
Detection		
Mammography	Mammography -jobs -technologist -training	breast mammography (100%); mammogram (95%); 3d mammography (70%); screening mammography (50%); digital mammography (45%); breast cancer (40%); solis (30%); mobile mammography (30%); what is mammography (30%); solis mammography (30%); mammography guidelines (20%); mammograms (20%); tomosynthesis (20%); diagnostic mammography (20%); acr (20%); acr mammography (15%); acr mammography (15%); 3d mammogram (10%); mammography recommendations (10%); mqsa (10%); hologic
Breast cancer screening	Breast cancer screening -society	screening for breast cancer (100%); breast cancer screening guidelines (55%); mammogram (30%); free breast cancer screening (25%); uspstf breast cancer screening (20%); uspstf (20%); uspstf breast cancer (20%); breast cancer screening recommendations (25%); acog breast cancer screening (10%); acs breast cancer screening (10%); breast cancer symptoms (10%); breast cancer awareness (10%); breast cancer awareness (10%); icd 10 code for breast cancer screening (5%); breast cancer awareness month (5%); new breast cancer screening guidelines (5%); nccn guidelines (5%); planned parenthood breast cancer screening (5%)
Breast cancer symptoms	Breast cancer symptoms -lung	symptoms of cancer (100%); symptoms of breast cancer (100%); breast pain (20%); symptoms for breast cancer (20%); breast cancer symptoms pain (20%); breast cancer signs (15%); breast cancer symptoms in women (15%); signs of breast cancer (10%); what is cancer (10%); what is breast cancer (10%); what are symptoms of breast cancer (10%); breast cancer symptoms and signs (10%); symptoms of breast cancer in women (10%); symptoms of cancer in women (10%); pain in breast (10%); inflammatory breast cancer symptoms (10%); inflammatory breast cancer (10%); breast cancer causes (5%); what are the symptoms of breast cancer (5%); signs and symptoms of breast cancer (5%); breast cancer treatment (5%); lump in breast (5%); men breast cancer (5%); causes of breast cancer (5%); men breast cancer symptoms (5%)

Dimension/Search Term	Search Equation	Related Search Queries
Diagnosis		
Breast biopsy	Breast biopsy	biopsy of breast (100%); breast cancer (90%); breast cancer biopsy (85%); biopsy for breast (80%); needle biopsy (55%); needle breast biopsy (55%); core biopsy breast (35%); stereotactic (35%); core biopsy (35%); stereotactic breast biopsy (35%); stereotactic biopsy (35%); biopsy on breast (35%); breast ultrasound (30%); breast biopsy results (30%); breast biopsy cpt code (25%); biopsy for breast cancer (25%); breast lump biopsy (20%); what is a biopsy (20%); what is a breast biopsy (20%); mri breast biopsy (20%); breast biopsy pain (20%); biopsy procedure (20%); breast mri (20%); needle biopsy of breast (15%); breast biopsy procedure (15%)
Breast cancer stages	Breast cancer stages	No related queries
Breast cancer type	Breast cancer type	No related queries
Treatment		
Breast cancer treatment	Breast cancer treatment -society	treatment for breast cancer (100%); treatment of breast cancer (70%); breast cancer radiation treatment (25%); radiation treatment (25%); what is breast cancer (20%); radiation treatment for breast cancer (15%); chemotherapy (10%); breast cancer treatment options (10%); triple negative breast cancer treatment (10%); triple negative breast cancer (10%); metastatic breast cancer (10%); metastatic breast cancer treatment (10%); new breast cancer treatment (10%); stage 1 breast cancer treatment (5%); stage 1 breast cancer (5%); stage 2 breast cancer treatment (5%); stage 2 breast cancer (5%); stage 4 breast cancer treatment (5%); stage 4 breast cancer (5%); breast cancer treatment guidelines (5%); mastectomy (5%); radiation therapy (5%); types of breast cancer (5%); breast cancer stages (5%); early breast cancer treatment (5%)
Lumpectomy	lumpectomy	Breast lumpectomy (100%); after lumpectomy (75%); breast cancer (45%); breast cancer lumpectomy (45%); mastectomy (40%); lumpectomy surgery (30%); what is lumpectomy (20%); lumpectomy recovery (20%); radiation after lumpectomy (20%); breast surgery (15%); what is a lumpectomy (15%); mastectomy vs lumpectomy (15%); lumpectomy recovery time (10%); lumpectomy and radiation (10%); breast biopsy (10%); lumpectomy

Dimension/Search Term	Search Equation	Related Search Queries
Breast cancer surgery	Breast cancer surgery –jolie –society –vasectomy -adams	Surgery for breast cancer (100%); breast cancer treatment (25%); mastectomy (25%); breast cancer surgery recovery (10%); breast cancer symptoms (10%); breast augmentation (5%)
Survivorship		
Surviving breast cancer	Surviving breast cancer	Chances of surviving breast cancer (100%)
Breast cancer recurrence	Breast cancer recurrence	Recurrence of breast cancer (100%); breast cancer recurrence rate (30%); triple negative breast cancer (25%); triple negative breast cancer recurrence (25%); breast cancer recurrence after mastectomy (20%)
Breast cancer reconstruction	Breast cancer reconstruction	Mastectomy (100%); breast cancer reconstruction photos (20%)

Note. The search dimensions (in bold) are the NCI Cancer Control Continuum search categories, which are the final outcome variable for Study 2. The search terms are the data retrieved to construct the search dimensions. The search equations were the word strings entered into Google Trends to retrieve data for the given search term. A negative sign (-) preceding a word indicated that that word was removed from data retrieval. The percentage (%) score following each related search term, indicates the percentage of search queries for this term in the same search session as the final search term.

Appendix D

Coding Procedure for Determining Relevancy of Mentions of Celebrity

Introduction

This coding protocol addresses news coverage of specific celebrities-of-interest over time. Much of the analysis for this dissertation will be done through human coding, however, some portions of content will be coded using computer-assisted content analysis. Therefore, this specific sub-protocol focuses on establishing reliability for determining if a sample article is about a specific celebrity-of-interest or about another topic or person that shares the same name. For example, whether an article includes the mention of Barbara Allen the Kansas State House representative (a celebrity-of-interest for this dissertation) or Barbara Allen the first female naval aviator (not a celebrity-of-interest for this dissertation).

This protocol addresses coding for relevancy for sampled returns for an open search string. This means that any story that includes a term that matches the name of a celebrity-of-interest will be included. Establishing reliability when articles are or are not about celebrities-of-interest is part of the process of creating a closed search string that accurately captures only the content needed. At this stage, coders are asked to judge only if the article includes any mention of the celebrity-of-interest.

Procedure for Article Eligibility for Study

All articles retrieved are from an open search string (the celebrity-of-interest's name), so any article with a mention of the celebrity-of-interest's name – whether or not it is actually a story about the celebrity-of-interest – is eligible. To aid in this, every time a search term appears in the article it will be **bolded in red and underlined** in the text version provided for coding. If you cannot find a **red bolded underlined term** in the article, please contact Susan LoRusso (lorus004@umn.edu), primary researcher.

Coding Procedure

All coding will be done in the Google Spreadsheet shared with you. All articles have been provided an identification number – a large bolded letter and number combination (e.g., **A2**) located in the upper-right-hand corner of each article. Coding for each article should be done in the line of the Google Spreadsheet with the corresponding letter/number combination to the article. For each celebrity-of-interest, three articles have been randomly selected. The celebrity-of-interest can be identified from the **red bolded underlined name/phrase** in the article. All articles will be dichotomously coded – (0) = not relevant and (1) = relevant.

Read headlines, decks (sub-headline), and copy for mentions of celebrities' names. Within these features of the article, look for clues to the type of person (or place or thing) the article is discussing. Any mention of the celebrity-of-interest, even if it is the third or eighth mention, that fits the relevancy rules, and means that the story should be considered relevant. There may be articles that include elements of a celebrity's name, but it is clear that these mentions are not specific to the celebrity-of-interest (or a person at all). Some examples are provided in the next section.

The reference of the celebrity-of-interest does not have to be about an ongoing or current event to be considered relevant. Also, the reference does have to be specific to something the celebrity-of-interest has done. For example, an article could be discussing a celebrity impersonator. This is a pop culture reference of the celebrity-of-interest, so therefore should be coded as relevant.

Variable Operational Definitions

V1. Relevant: An article is considered relevant if it contains at least one mention that indicated the article is talking about one of the selected celebrities-of-interest. Code:

(0) Not relevant
(1) Relevant

The attached definitions of each celebrity-of-interest (see pages 4 – 12) include information which should help you determine if the person in the article is indeed a celebrity-of-interest (and therefore *relevant*). Generally, the information provided in these "definitions" includes the celebrity's birthdate (and day of death if applicable), their career, and/or significant relationships. Use this information to judge if the person being discussed is the celebrity-of-interest (Does the celebrity-of-interest's age correspond with the age reported [in context of the year the report was made]? No mention of career or relationship – could be a report of an "average person" with the same name as a celebrity-of-interest).

Please note: articles should be coded as relevant if *any* mention meets the relevancy rules, even if other mentions do not.

When determining relevancy, consider the following three common issues in articles retrieved based on the celebrity-of-interests of names.

Common Name/Average Person: *Common name/Average person* issues are when the celebrity-of-interest's name is retrieved in full, but the name is referring to an average citizen (and not the celebrity of interest). For example:

> The engagement of **Robin Roberts** to Andrew Mercogliano has been announced by Mr. and Mrs. George Roberts of Old Bridge, N.J., parents of the future bride. Her fiancé is the son of Mr. and Mrs. John Mercogliano of Asbury Park, N.J.

This article clearly is an engagement announcement. Given the nature of the story (it doesn't mention her journalism career) you will conclude that this story is *(0) Not relevant*.

Places or Things: Some names double as words of places or things in the English language. For example:

For the search string "Peter Criss" the following was retrieved:

> But I foresee shouts of "philistine!" and "purist!" once more **criss**-crossing the restaurants and theatre foyers.

This mention of "criss" is obviously not in reference to Peter Criss, celebrity-of-interest. This article will be coded as *(0) Not relevant*.

Other examples:

> Headline: Colours' creator used 'summer' palette in room **Rose and gray** themes color room where she met clients.

In this case, "rose and gray" are referring to the colors rose and gray and not to *Rose Gray*, celebrity-of-interest. This article will be coded as *(0) Not Relevant*.

> Market Research Report on Global and Chinese **Poly**-a-Methyl **Styrene** Industry, 2009-2019 is a professional and in-depth market survey on Global and Chinese **Poly**-a-Methyl **Styrene** industry.

This article is referencing polystyrene the synthetic polymer product and not *Poly Styrene*, musician and celebrity-of-interest. This article will be coded as *(0) Not relevant.*

Other Notable Persons

The third issue likely to be included in articles selected for this analysis are articles that include the celebrity-of-interest's full name but the reference is of another notable person with the same name. For example:

> At 65, looking fit enough to begin both ends of a doubleheader, righthander **Robin Roberts** has been named to the board of directors of the national baseball Hall of Fame. Already in the Hall himself as a pitcher, Robby probably also would merit Cooperstown because he was the man who picked Marvin Miller as the ballplayers' representation.

This article is discussing the hall of fame baseball player, *Robin Roberts*, and not the journalist. An obvious clue that this article is not relevant is the use of the pronoun "he." The celebrity-of-interest identifies as a woman. Therefore, any stories about her would include the pronoun "she."

Another example:

> **Yvonne Carter** was drugged and raped by a tour guide while she was travelling with friends in the Bolivian jungle. Three years after her ordeal the 28-year -old physiotherapist has learnt that her attacker is still free and is suspected of having carried out at least three other sexual assaults on tourists.

This article is about a sexual assault survivor who received significant media coverage. But, this is not *Yvonne Carter* the British medical doctor who is a celebrity-of-interest. Another clue, the article was published in 2002 and states that Yvonne Carter is 28 years old. The celebrity-of-interest Yvonne Carter was 43 in 2002. This article will be coded as *(0) Not Relevant*.

***All celebrities-of-interest are listed on the next several pages. Some celebrities may be instantly recognizable to you, but some names are more obscure. Therefore, please double-check the names and corresponding definitions before making a relevancy decision.*

***All of the celebrities-of-interest experienced a breast cancer-related event at some point in their lives. Many articles may mention this connection. This is another "clue" that the article is about a celebrity-of-interest.*

Celebrities-of-Interest

Celebrity Name	Definition of Celebrity
Allen, Barbara	(d. 3/18/2005) Assistant Attorney General, 1985-1987; Representative, Kansas State House, 1989-2001; Kansas Senator (R)
Anastacia	(b. September 17, 1968) American singer-song writer, pop star. albums: Not that Kind; Freak of Nature
Applegate, Christina	(b. November 25, 1971) American actress, began as child actress (Kelly Bundy/Married W/Children), Notable films: Don't Tell Mom the Babysitter is Dead; The Sweetest Thing; Anchorman: The Legend of Ron Burgundy; Anchorman 2: The Legend Continues; Notable Sitcoms: Up all Night and Samantha Who?
Baker Tharp, Carol	(May 13, 1952 - November 25, 2007) American general manager and former executive director; University of Southern California School of Politics, Planning and Development as the deputy director of the Civic Engagement Initiative. Executive director of Coro in Southern California and was Community Relations director for Eugene, Oregon.
Barnett, Lisa A.	(1958 - May 2, 2006) American Lambda Literary Award winning science fiction writer/author. Novels: The Armor of Light; Point of Hopes; Point of dreams
Bates, Jeanne	(May 21, 1918 - November 28, 2007) American radio, film and television actress. Notable films: The Phantom, The Chance of a Lifetime; Death of a Salesman.
Bates, Kathy	(b. June 28, 1948) American actress. Notable films: Misery (Oscar), Fried Green Tomatoes, Delores Claiborne, and Titanic. Notable television shows: Harry's Law; American Horror Story (Coven); Two and a Half Men (Emmy winner).
Berzon, Betty	(January 18, 1928 - January 24, 2006) American author and psychotherapist known for her work with gay and lesbian communities. Books included: Positively Gay; Permanent Partners; The Intimacy Dance; and personal memoir - Surviving Madness a Therapist's Own Story.
Bissex, Rachel	(December 27, 1956 - February 20, 2005) American folk singer/songwriter. Works included "Dancing With My Mother" and "Drive All Night"
Blume, Judy	(b. February 12, 1938) American writer known for children's and young adult fiction. Notable works include: Are You There God? It's Me, Margaret. Tales of a Fourth Grade Nothing, and Blubber.
Chapman, Alison	(June 23, 1963 - July 1, 2010) British presenter and newsreader at 106 Jack FM and BBC Oxford.
Brodnick, Caitlin	Comedian, currently 30ish, from Kensington, Maryland.
Calloway, Chris	(1945-2008) Jazz singer, daughter of Cab Calloway, Santa Fe resident.
Carter, Yvonne	(April 16, 1959 - October 20, 2009) British General practitioner and Dean of the Warwick Medical School.

Celebrity Name	Definition of Celebrity
Criss, Peter	(b. December 20, 1945) American musician (KISS) and actor. Inducted into the Rock and Roll Hall of Fame
Crow, Sheryl	(b. February 11, 1932) American singer-songwriter. Garnered nine Grammy Awards. Notable albums: Tuesday Night Music Club, The Globe Sessions, Feels like Home. Dated Owen Wilson and Lance Armstrong.
Dapkus Wolf, Eleanor	(December 5, 1923 – June 6, 2011). A center fielder and pitcher who played from 1943-1950 in the All-American Girls Professional Baseball League.
Davis, Jo Ann	(June 29, 1950 - October 6, 2007) Representative in the U. S. Congress from Virginia (R).
Day, Linda	(August 12, 1938 - October 23, 2009) American television director (sitcoms).
de Sela, Lhasa	(September 27, 1972 - January 1, 2010) American-born singer-songwriter, raised in Mexico and the United States. Major success in Canada.
Dewar, Helen	(August 7, 1936 - November 4, 200) Reporter for The Washington Post for 25 years (United States Senate).
Doherty, Shannen	(b. April 12, 1971) American actress. Best known for her roles as Brenda Walsh in Beverly Hills, 90210 and her time on Charmed.
Dolgin, Gail	(April 4, 1945 - October 7, 2010) American documentary filmmaker.
Edwards, Elizabeth	(July 3, 1949 - December 7, 2010). American attorney, best selling author, health care activist, and was married to Former U.S. Senator John Edwards.
Ekvall, Eva	(March 15, 1983 - December 17, 2011). Former Miss Venezuela, Venezuelan television news anchor, author, and model.
Engelberg, Miriam	(January 7, 1958 - October 17, 2006) Graphic novelist and illustrator. Her cartoon Planet 501c3 was the first cartoon series depicting life in the nonprofit sector.
Etheridge, Melissa	(b. Mary 29, 1961) American rock singer-writer and gay rights activist. Grammy award winner.
Faithful, Marianne	(b. December 29, 1946) English singer, songwriter and actress. Lead female artists during the "British Invasion" in the United States. Had a relationship with Mick Jagger.
Feldman, Sandra	(October 13, 1939 - September 18, 2005) American civil rights activist, educator and labor leader who served as president of the American Federation of Teachers from 1997 to 2007. Before serving as president of the AFT served as a field representative (Ocean Hill-Brownsville strike) and vice president.
Fiorina, Carly	(b. September 6, 1954) American Businesswoman. Primarily known for her tenure as the Chief Executive Officer (CEO) of Hewlett Packard.

Celebrity Name	Definition of Celebrity
Fonda, Jane	(b. December 21, 1937) American actress, writer, political activist, and fashion guru. Two-time Academy Award winner. Notable films: Julia, The China Syndrome, On Golden Pond, Monster in Law, The Butler. Notable television shows include The Newsroom and Grace and Frankie.
Fontane, Char (Kaci)	(January 12, 1952 - April 1, 2007) American actress and singer. Memorable performances include ABC miniseries Pearl, The Love Boat, Love American Style, and Broadway's Grease.
Forbes, Mary Lou	(June 21, 1926 - June 27, 2009) American journalist and commentator. Won a Pulitzer Prize for her coverage of the 1958 school integration crisis in Virginia.
Fraction, Karen	(February 15, 1958 - October 30, 2007) American actress, dancer and model. Most notable roles were as Dr. Perry, the Chief Medical Officer in seaQuest 2032 and Jennifer Parker in My Brother and Me.
Frontiere, Georgia	(November 21, 1927 - January 18, 2008) American businesswoman and entertainer. She was the majority owner and chairperson of the LA/St. Louis Rams football team and the most prominent female owner in the league. She also sat on the board of the United Way, Saint Louis Symphony Orchestra, Herbert Hoover Boys and Girls Club, and the American Foundation for AIDS Research.
Frost, Kathryn	(November 7, 1948 - August 18, 2006) Commander of the United States Army and Air Force Exchange Service from August 2002 - April 2005. At the time of her retirement, she was the highest-ranking woman in the United States Army. She was also the wife of former United States Representative Martin Frost of Texas. She held several other high ranking and notable positions in the U.S. military throughout her career.
Gaull Silberman, Rosalie (Ricky)	(March 31, 1937 - February 18, 2007) American conservative activist co-founded the Independent Women's Forum. Worked for the Equal Employment Opportunity Commission from 1984 until 1995, rising to the positions of vice-chair and commissioner.
Gittings, Barbara	(July 31, 1932 - February 18, 2007) A prominent American activist for gay equality. She was part of the movement to get the American Psychiatric Association to drop homosexuality as a mental illness. She was also very involved with the American Library Association. They named an annual award for the best gay or lesbian novel.
Gordon, Kim	(b. April 28, 1953) American musician, songwriter, and visual artist. She is most known for being in the band Sonic Youth.
Gray, Rose	(January 28, 1939 - February 28, 2010) British chef and cookery writer. As Chef of the The River Cafe she won a Michelin star in 1998. She wrote a series of cookbooks and starred in The Italian Kitchen.
Green, Ernie	(b. October 15, 1938) A former American football fullback for the Cleveland Browns in the NFL. Currently he runs Ernie Green Industries, which manufactures components for the automotive industry in Dayton, OH.

Celebrity Name	Definition of Celebrity
Hamill, Dorothy	(b. July 2, 1956) Retired American figure skater. She was the 1976 Olympic and World champion. She continued skating professionally and was on season 16 of Dancing with the Stars.
Harris, Samantha	(b. November 27, 1973) American television hostess. She co-hosted seasons 2 - 9 of Dancing with the Stars and was a correspondent at Entertainment Tonight from 2010 - 2012.
Heaton, Michelle	(b. July 19, 1979) An English pop singer, actress, television personality, personal trainer and model. She was a member of the pop group Liberty X from 2001 through 2007.
Horn, Shirley	(May 1, 1934 - October 20, 2005) American jazz singer and pianist. Grammy award winner.
Hunt Lieberson, Lorraine	(March 1, 1954 - July 3, 2006) American mezzo-soprano. Performed at the Metropolitan Opera.
Ivins, Molly	(August 30, 1944 - Jan 31, 2007) American newspaper columnist, author, political commentator, and humorist. Most known for her time at The New York Times.
Jolie, Angelina	(b. June 4, 1975) American actress and humanitarian. Received an Academy Award and 3 Golden Globes. Best known for Girl Interrupted, Tomb Raider, and Mr. and Mrs. Smith. She has also had notable relationships with actors Billy Bob Thornton and Brad Pitt. Her advocacy for refugees led her to be a Special Envoy for the United Nations High Commissioner for Refugees.
Kamen Goldmark, Kathi	(August 18, 1948 - May 24, 2012) American author, columnist, publishing consultant, radio and music producer, songwriter, and musician. Best known for her novel "And My Shoes Keep Walking Back to You" and producing the radio show "West Coast Live." She was also the president of "Don't Quit Your Day Job" Productions Inc.
Kennedy, Joan	(b. September 2, 1936) American socialite, musician, author, and former model. She is best known because she was the first wife of U.S. Senator Ted Kennedy.
King, Angela	(August 28, 1938 - Feb. 5, 2007) Jamaican diplomat. Worked for the United Nations for 38 years, from 1966 to 2004, working mainly equal rights for women.
Kosofsky Sedgwick, Eve	(May 2, 1950 - April 12, 2009) American academic scholar in the fields of gender studies, queer theory, and critical theory. She was best known for groundbreaking books in the field of queer theory - Between Men: English Literature and Male Homosocial desire, Epistemology of the Closet, and Tendencies.
Kotb, Hoda	(b. August 9, 1964) An American television news anchor and TV host known as the co-host on the Today Show and as a correspondent for Dateline NBC since 1998.
Langan, Joy	(January 23, 1943 - July 30, 2009) Member of the Canadian House of Commons from 1988 to 1993.

Celebrity Name	Definition of Celebrity
Lee, Sandra	(b. July 3, 1966) American television chef and author. She has had two shows on the Food Network - "Semi-Homemade Cooking with Sandra Lee" and "Sandra's Money Saving Meals." She has released 25 books including Sand Lee Semi-Homemade: Cool Kids Cooking and a memoir, Made From Scratch. She has also has two lifestyle magazines. She is also known for her long-term relationship with New York Governor Andrew Cuomo.
Luft, Lorna	(b. November 21, 1952) American television stage and film actress and singer. She is the daughter of singer and actress Judy Garland and Sid Luft and half-sister of singer and actress Liza Minnelli.
Lund, JoAnna	(September 4, 1944 - May 20, 2006) Author of many books, including Healthy Exchanges Cookbook, HELP: Healthy Exchanges Lifetime Plan, and Make a Joyful Table. Also known to have appearances on CNN, Home Shopping Club, and QVC.
Lunden, Joan	(b. September 19, 1950) An American journalist, an author and a television host. Best known for her 17 years co-hosting ABC's Good Morning America. Currently, she is a special correspondent for NBC's Today.
Lyon, Jennifer	(February 27, 1972 - January 19, 2010) An American actress and one of the competitors in Survivor: Palau (10th Season)
Mann, Judy	(December 24, 1943 - July 8, 2005) Correspondent for The Washington Post where she wrote about women, children, and the politics of the women's movement.
Marchese, Amber	(b. 1977) American television personality. Best known for starring in the reality television series The Real Housewives of New Jersey.
Mayo-Chandler, Karen	(April 18, 1958 - July 11, 2006) British model and actress. She appeared in issues of Vogue, Harper's Bazaar and Playboy. Also known for her relationship with actor Jack Nicholson.
McGhee, William	(July 24, 1930 - February 17, 2007) A film and television actor. Best known for his roles in the films High Yellow, Curse of the Swamp Creature, Don't Look in the Basement and Drive-In. He was one of the first unionized African-American actors in Dallas with SAG.
Mendez, Olga	(February 5, 1925 - July 29, 2009) Was the first Puerto Rican woman elected to a state legislature (New York) - D in the United States mainland.
Minogue, Kylie	(b. May 28, 1968) Australian singer, songwriter, dancer and actress. She is the highest selling Australian artist of all time. She is best known in the U.S. for her singles "The Loco-Motion" and "Can't Get You Out of My Head".

Celebrity Name	Definition of Celebrity
Mitchell, Andrea	(b. October 30, 1946) American television journalist, anchor, reporter and commentator for NBC News. She has appeared on NBC Nightly News with Lest Holt, Today, and several MSNBC shows.
Moffo, Anna	(June 27, 1932 - March 9, 2006) American opera singer, television personality, and award-winning dramatic actress. She had considerable fame in Italy. But, did perform often at the Metropolitan Opera in her lifetime.
Nathan, Melissa	(June 13, 1968 - April 7, 2006) Journalist and UK author of popular "chick lit" novels in the early 200s.
Navratilova, Martina	(b. October 18, 1956) Czech and American tennis player and coach. She won 18 Grand Slam singles, 31 major women's doubles titles, and 10 mixed doubles titles. She reached the Wimbledon singles final 12 times, including nine consecutive years from 1982 through 1990, and won the women's singles title at Wimbledon a record nine times. She has been an "out" lesbian since 1981 and has been an activist for LGTB rights.
Nesler, Ellie	(1952 – 2008) Mother who shot and killed child's accused molester in a California courtroom.
Nielsen, Jerri	(March 1, 1952 - June 23, 2009) American physician who famously self-treated her breast cancer while stationed at Amundsen-Scott South Pole Station in Antarctica (1998) until she could be evacuated safely.
Nixon, Cynthia	(b. April 9, 1966) American actress. She is best known for her portrayal of Miranda Hobbes in the HBO series Sex and the City.
Notaro, Tig	(b. March 24, 1971) American stand-up comic, writer, radio contributor, and actress.
Novak, Kim	(b. February 13, 1933) American film and television actress. She is best known for her role in Vertigo. She has continued to act in small roles.
Osbourne, Sharon	(b. October 9, 1952) Media personality. Stars in The Talk and America's Got Talent. Became well-known for her family's MTV reality show "The Osbourne's". She is married to heavy metal singer Ozzy Osbourne.
Paley, Grace	(December 11, 1922 - August 22, 2007) American short story writer, poet, teacher, and political activist. She published several short story collections.
Pence, Ellen	(1948 – January 6, 2012) Scholar and social activist. She co-founded the Duluth Domestic Abuse Intervention Project (an inter-agency collaboration model used to reduce domestic violence against women in all 50 states and over 17 countries).

Celebrity Name	Definition of Celebrity
Pick, Heather	(1970 - November 7, 2008) American television news anchor and activist. Worked at television stations in Chicago and Ohio. At Ohio, she worked closely with Jack Hanna and assisted him on national television program segments. Worked with Mary Tyler Moore in support of the Juvenile Diabetes Research Foundation.
Rabinovitch, Dina	(June 9, 1962 - October 30, 2007) British journalist and writer who wrote a column for The Guardian.
Rancic, Giuliana	(b. August 17, 1974) American television personality. Best known for anchor role on E! News, E!'s Fashion Police, and her reality-television show "Giuliani and Bill". She also often co-hosts red carpet events for the Golden Globes and the Academy Awards.
Redgrave, Lynn	(March 8, 1943 - Mary 2, 2010) English actress for both stage and film. More recent appearances include Gods and Monsters, Ugly Betty, Desperate Housewives, and Law & Order: Criminal Intent.
Reed Hall, Alaina	(November 10, 1946 - December 17, 2009) American singer and actress best known for her roles as Olivia Robinson on Sesame Street and as Rose on the NBC sitcom 227.
Robach, Amy	(b. February 6, 1973) American television journalist. She has been a national correspondent for NBC News, co-host of NBC's Today and anchor on MSNBC. She currently is an anchor on Good Morning American and rotates as a 20/20 anchor.
Roberts, Robin	(b. November 23, 1960) American television broadcaster. She first became known as a sportscaster for 15 years on ESPN. She became a co-anchor on Good Morning America in 2005.
Romney, Ann	(b. April 16, 1949) Wife of American businessman and politician, Mitt Romney. She was the First Lady of Massachusetts from 203 to 2007. She was diagnosed with multiple sclerosis in 1998. To cope with this, she has taken up equestrianism and has competed professionally in dressage.
Rose, Allyn	(b. 1988) American beauty pageant titleholder and model. She won the title of Miss District of Columbia in 2012 and competed in the Miss America 2013 pageant.
Sambolin, Zoraida	(b. July 10, 1965) American television journalist. Anchored news broadcasts for local affiliates in Chicago for nine years. She is currently the host of Early Start on CNN.
Sano, Yoko	(1938 – November 5, 2010) Japanese author and illustrator of children's books.
Shubert, Jean	(May 1, 1942 - May 14, 2007) Alpine ski racer from the U.S. She won two medals in the 1964 Olympics in Austria. She was inducted into the National Ski Hall of Fame in 1976.
Semple, Goldie	(December 11, 1952 - December 9, 2009) A Canadian actress. She primarily did stage performances but appeared on television for series such as Queer as Folk and Street Legal.

Celebrity Name	Definition of Celebrity
Sims, Naomi	(March 30, 1948 - August 1, 2009) American model and author. She was the first African-American model to appear on the cover of Ladies' Home Journal and is widely considered the first African-American supermodel.
Smith, Maggie	(b. December 28, 1934) Dame Margaret Natalie Smith. An English actress. She has an extensive career in stage, film, and television spanning over 60 years. She is best known for her roles in Othello, A Room with a View, Gosford Park, Hook, Sister Act, Tea with Mussolini, and most recently Dowton Abbey.
Soraya	(March 11, 1969 - May 10, 2006) A Colombian-American singer/songwriter, guitarist, arranger and record producer. She won Latin Grammy Awards in 2004 and 2005.
Sperber, Wendie Jo	(September 15, 1958 - November 29, 2005) An American actress best known for her roles in Bachelor Party, Back to the Future, Bosom Buddies and Private Benjamin.
Stevens, Hollie	(January 4, 1982 - July 3, 2012) Stage name of American pornographic actress. She was considered a pioneer of the porn genre known as clown porn. She entered into the adult film industry in 2003 and appeared in over 180 films.
Stevens, Pat	(September 16, 1945 - May 26, 2010) An American actress and voice actress. She is best known for her role as Nurse Bake on MASH and her voice work as Velma in the cartoon series Scooby-Doo.
Stilwell, Heather	(January 26, 1944 – December 3, 2010). A Canadian political activist who was well known for her opinions opposing homosexuality, abortion, and sex education.
Styrene, Poly	(July 3, 1957 - April 25, 2011) A British musician, singer-songwriter, and front woman for the punk band X-Ray Spex.
Sykes, Wanda	(b. March 7, 1964) American comedian, writer, actress and voice artist. Best known for her roles on the New Adventures of Old Christine, HBO's Curb Your Enthusiasm, The Wanda Sykes Show, Monster-in-Law, Evan Almighty, Ice Age, and her comedy stand-up comedy specials.
Syler, Rene	(b. February 17, 1963) Television journalist. Began her career as a reporter in Reno, Nevada. She continued on to anchor local affiliate news broadcasts in Alabama and Dallas. She was an anchor on CBS News' The Early Show. She has continued to do guest appearances and host a show on the Live Well Network.
Tierney, Maura	(b. February 3, 1965) American film and television actress who is best known for her roles on NewsRadio, Liar Liar, ER, and The Affair.
Tracey, Lindalee	(May 14, 1957 - October 19, 2006) A Canadian broadcast journalist, documentary filmmaker, writer, and exotic dancer. Best known for her work in the film Not a Love Story.

Celebrity Name	Definition of Celebrity
Walker, Catherine	(June 27, 1945 - September 23, 2010) A French fashion designer based in London. She is best known for designing over 1,000 outfits for Diana Princess of Wales, who was buried in a black dress designed by Walker. Kate Middleton also wore several of her designs.
Wasserman Schultz, Debbie	(b. September 27, 1966) An American politician. She is the U.S. Representative for Florida's 23rd congressional district and a member of the Democratic Party who served as chairperson of the Democratic National Committee from 2011 to 2016. She has also served in the Florida House of Representatives and the Florida Senate.
Wexler, Anne	(February 10, 1930 - August 7, 2009) An American influential Democratic political consultant, public policy advisor and became the first woman to head a leading lobbying firm in Washington.
Wilson, Rita	(b. October 26, 1956) American actress and producer. Best known for her roles in Sleepless in Seattle, Now and Then, Jingle All the Way, and Runaway Bride and producing My Big Fat Greek Wedding. She has been married to actor Tom Hanks since 1988.
Wyler, Gretchen	(February 16, 1932 - May 27, 2007) An American actress and founder of the Genesis Awards for Animal Protection. Had roles in several major Broadway musicals from the 1950s - 1970s.
Wynn Fonstad, Karen	(April 18, 1945 - March 11, 2005) American cartographer and academic. She designed several atlases of fictional worlds.
Yow, Kay	(March 14, 1942 - January 24, 2009) An American basketball coach. She was the head coach of the NC State Wolfpack women's basketball team from 1975 to 2009. She also coached the U.S. women's basketball team to an Olympic gold medal in 1988.
Ziskin, Laura	(March 3, 1950 - June 12, 2011) An American film producer. Best known for her executive producer role of the 1990 film Pretty Woman. She was also the first woman to produce the Academy Awards.
Zuk, Judith D. (Judy)	(September 11, 1951 - September 1, 2007) An American horticulturist, author and conservationist. She served as president of the Brooklyn Botanic Garden from 1990 t0 2005.

Appendix E

Coding Protocol for Study 1

V1. News article code: Each article is labeled with an alphabetical/numerical combination code (ex. **ANWA - 35**). The code is located in the upper-right hand corner of the news article. Please record the code in the space provided in Qualtrics under **V1**.

V2. Celebrity-of-Interest: Who is the article about? It is possible that other celebrities will be mentioned in the news article (perhaps other celebrities who are breast cancer survivors), but for this variable you need to determine the celebrity-of-interest the news article focuses on - who is central to the story. If an article focuses or gives equal space to more than one celebrity, check (in Qualtrics) all that apply. Check the headline, deck, and copy for at least one mention of one of the 110 celebrities-of-interest listed on page 2. While steps have been taken to ensure that all articles will be about one of the celebrities-of-interest, it is possible than a non-relevant article was retrieved. If this is the case, select "*none*" at the bottom of the list.

******If you select "none", no further coding will be done for that news article. Qualtrics will automatically end the survey.***

V3. Relevancy: This study is concerned with news coverage of celebrity breast cancer disclosures. Therefore, each article analyzed must be about one of the celebrities-of-interest selected for this study (as determined for **V2**) AND must contain some information about the celebrity's breast cancer-related event, to be considered relevant.

- Not relevant
- Relevant

Using the same procedure as mentioned above (checking the headline, deck [sub-headline], and copy) determine if the article is discussing or mentions the celebrity's experience with breast cancer. In most cases, the article will make a direct link between the celebrity and breast cancer.

Example: HEADLINE: EXCLUSIVE: Shannen Doherty Reveals Her Breast Cancer Has Spread: 'The Unknown Is the Scariest Part'

DECK: Shannen Doherty has some devastating news about her battle with breast cancer.

Explanation: Both the headline and the deck of this story explicitly state that Shannen Doherty (a celebrity-of-interest for the current study) has experienced a breast cancer-related event.

Several celebrities-of-interest died due to complications with breast cancer. As long as the article mentions a celebrity-of-interest and that she or he died of breast cancer, or at least experienced breast cancer at some point in her or his life, this information meets the criteria for a mention of breast cancer.

Example: Hollie Stevens, best known as "The Queen of Clown Porn" died on Tuesday, July 3, 2012 in San Francisco. She fell asleep peacefully while holding the hand of her husband, comedian and artist, Eric Cash.

Hollie was diagnosed in March 2011 with Stage 3, Metastatic Breast Cancer. Within a year, it had spread to her bone, rib, liver, and brain. After an outpouring of support from a fundraising effort facilitated by her family and friends which raised $16,000 she said, "I cannot believe how many people care, how good my fans are, and how much complete strangers have helped me. This is unreal. Not everyone hates clowns after all!"

Explanation: While it takes some space to get to Steven's diagnosis of breast cancer, and does not explicitly state that she died from breast cancer, this article meets the criteria for a mention of breast cancer related to a celebrity-of-interest.

- Final note: The news article does not have to state that the celebrity was diagnosed with breast cancer. Some celebrities were never diagnosed with breast cancer, but instead, took preventive measures such as genetic/BRCA (1 or 2) testing, and/or had a preventive prophylactic mastectomy.
- Key words to look for:
  - Breast cancer, chemotherapy, radiation, BRCA (may include 1 or 2), genetic testing, mastectomy, lumpectomy, prophylactic mastectomy, preventive, breast surgery, breast conserving surgery

Final determination:

- ✓ The celebrity, or one of the persons, mentioned in the news article is listed in the celebrities-of-interest table on page 2 of this document.
- ✓ The article discusses or mentions breast cancer or a breast cancer-related event in connection with the celebrity-of-interest.
  - If you can answer "yes" to both of the above, then the news article is considered "relevant". In Qualtrics, select *Relevant* under **V3**.
  - If you answer "no" to one or both of the above criteria, than the news article is
    considered "not relevant". In Qualtircs, select *Not relevant* under **V3**.

******If you code the article as Not relevant, no further coding will be done for that news article. Qualtrics will automatically end the survey.***

<u>Section 1- Information about News Article</u>

The following 4 variables, or categories, include descriptive information about the actual news article.

V4. Media organization: The media organization is the owner/producer/distributor of the content of the news article. See the **Figure 1** for the **media organization** location.

- The three media organizations under study are the *Associated Press, The New York Times,* and *The Washington Post*.

The media organization listed in your coding sample may be an iteration of one of these organizations. These might include:

- Associated Press: AP; Associated Press Online
- The New York Times: The New York Times Blogs
- The Washington Post: Washington Post Blogs; Washington Post Magazine; Washingtonpost.com

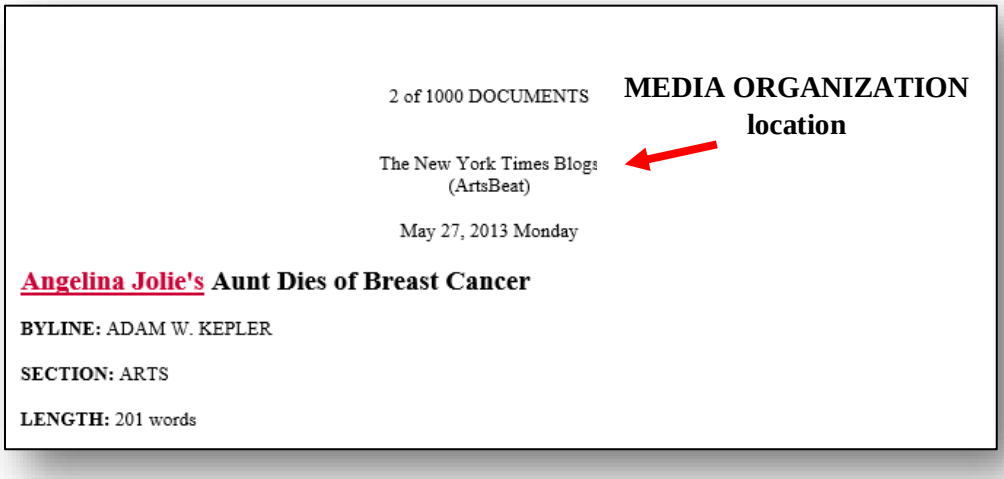


Figure 1. Media Organization Location

In Qualtrics for **V4**, select the media organization responsible for the news article. For the above figure, you would select *The New York Times*.

- *Associated Press*
- *The New York Times*
- *The Washington Post*

V5. Date of the article: The **date of the article** is the date the article was published. See **Figure 2** on the next page for the **date of the article** location. Manually enter the date in the space provided in Qualtrics under **V5**. Please use the following formatting: Month (XX)/Day (XX)/Year(XXXX). For the Figure 2, you would enter *05/27/2013*.

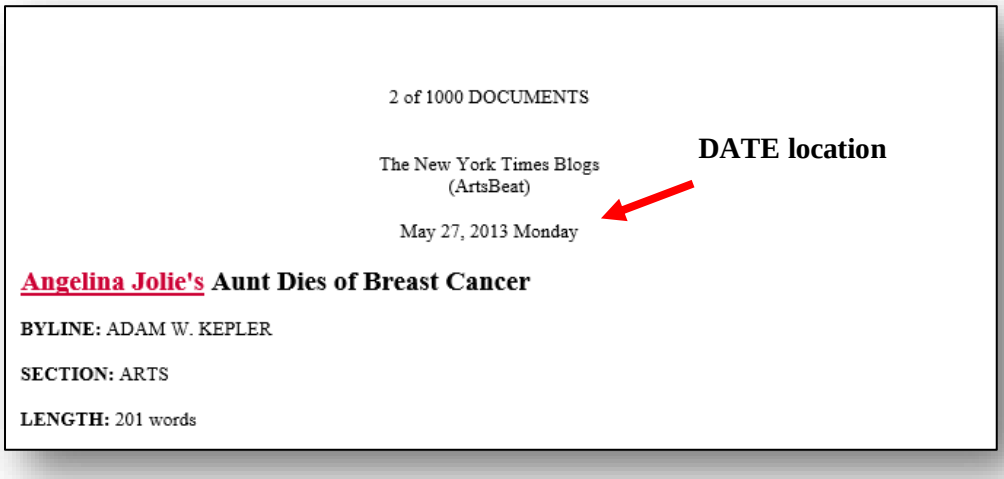


Figure 2. Date Location

V6. Section: This variable is concerned with what "section" the article appears in the newspaper or wire service. See **Figure 3** for the **section** location. If no section is listed, please select *None*. The list below and in Qualtrics provides what are likely to be the dominant sections the news articles appear in; however, it may not be an exhaustive list. If a section is listed in the news article that is not included here or in Qualtrics under **V6**, then please select *Other*. For the figure below, you would enter *Arts*.

- Arts/Arts Beat/Culture/Cultural Desk/Performing Arts
- Business News/Financial News
- Celebrities
- Classified
- Domestic News
- Entertainment/Entertainment News
- Health
- International News
- Lifestyle/Life
- Metropolitan Desk
- Music
- News/Main news
- Obituary
- Opinion/Editorial
- Political News
- Sports
- Style/Fashion
- Other
- None

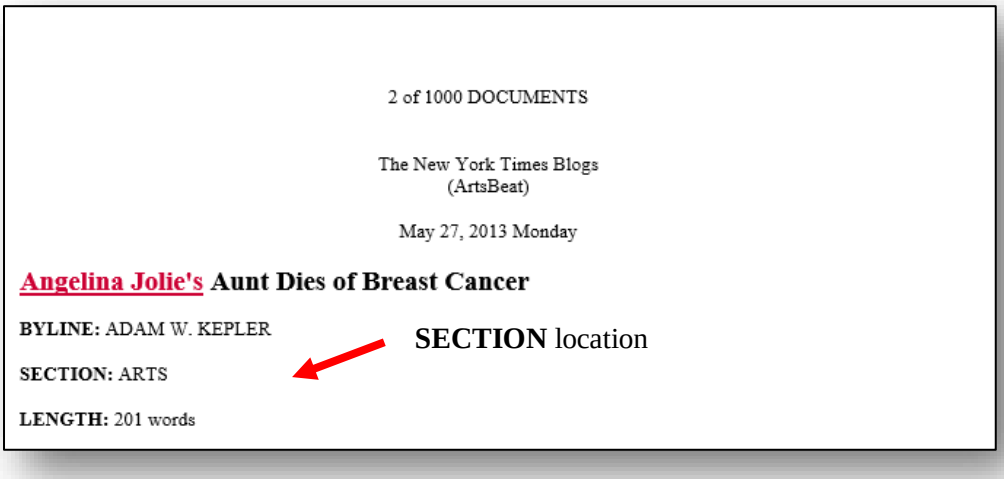


Figure 3. Section

V6a. "Other" section in newspaper: If you selected *Other* for **V6**, please manually type in the section name in Qualtrics under **V6a.**

V7. Word count: Word count refers to the number of words present in the article. See Figure 4 below for the **word count** location for the news articles. In Qualtrics for **V7**, manually type in the **word count**. For the figure below, you would enter "201".

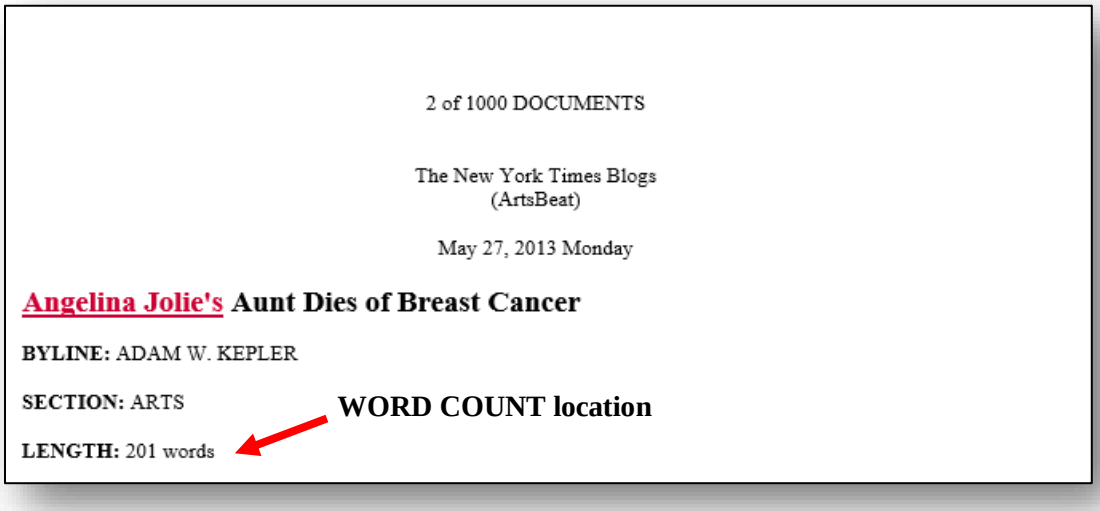


Figure 4. Word Count

Section 2 – Content

Section 2 requires you to assess how the information in the article is presented or framed.

V8. Media frames: This variable is concerned with the framing of the news article – episodic or thematic. For this variable, all news articles will be dichotomously coded (1 = Episodic frame and 2 = Thematic frame). This means that each article can only contain one frame or the other.

- Episodic Frame
- Thematic Frame

Episodic frame definition:

A news article can be considered written with an *episodic frame* when the primary focus of the article is on the celebrity and her or his breast cancer-related event. Health information about the celebrity's own experience with cancer can and will likely be included in the article. If there is information that speaks to breast cancer in a broader public health context (e.g., who is at greatest risk for breast cancer, how the celebrity's risk is comparable to other groups of women, etc.), then that article does not have an *episodic frame*.

Features that may be included in an article with an Episodic frame:

- Celebrity's current age/age at diagnosis
- Celebrity's career/what the celebrity is known for
- Personal relationships the celebrity has had
- Celebrity's family
- A statement that the celebrity announced a breast cancer-related event
- The specific breast cancer-related event the celebrity experienced (breast cancer diagnosis; genetic testing or BRCA [1 or 2]; lumpectomy; mastectomy; single or lateral mastectomy, double or bilateral mastectomy; prophylactic mastectomy; chemotherapy; radiation; nonspecific breast cancer treatment; nonspecific breast cancer surgery; nonspecific breast cancer-related drug treatment; death from complications of breast cancer)
- A statement on why the celebrity decided to go public about her or his breast cancer experience
- May list a specific breast cancer the celebrity was diagnosed with (ductal carcinoma in situ [DCIS]; invasive ductal carcinoma [IDC]; inflammatory breast cancer [IBC]; male breast cancer; Paget's disease of the nipple; breast tumor; localized breast cancer; recurrent and/or metastatic breast cancer)
- Specific stage of breast cancer the celebrity was diagnosed with (stage 0, 1, 2, 3, 4)

- Circumstances surrounding the celebrity's diagnosis, genetic testing, surgical decision, etc. (felt lump in breast; clinical exam; routine screening/mammogram; cancer found during other procedure; family history)
- Celebrity's prognosis (expected to make a full recovery; cancer free; further treatments needed; death)

This is not an exhaustive list of the information included in a news article with an episodic frame, nor do all of these points or information need to be included to have an episodic frame. What you should take from this list is that all of the information, whether it is specific to the celebrity's career, personal life, or her or his breast cancer experience, is about the celebrity and not connected to breast cancer beyond the celebrity's experience.

If a news article fits the definition of an *episodic frame*, select *Episodic frame* in Qualtrics for **V8**.

Thematic frame definition:

A news article reporting on a celebrity breast cancer disclosure with a *thematic frame* will convey how the celebrity's breast cancer-related event may be relevant to others – either those of similar or average risk. Such an article can include any and all of the information listed for the episodic frame, but will also provide breast cancer information beyond that of the celebrity's own experience.

Features that may be included in a news article with a Thematic frame:

- Will likely include statistics
- Includes information about breast cancer risk factors (e.g., women with a family history of breast cancer, women over the age 50, etc.)
- Ethnic/racial breast cancer survival disparities (e.g., Blacks and Native Americans have the highest cancer mortality and shortest survival time among all other racial/ethnic groups)
- Procedure recommendations by a medical professional or official guidelines from a medical/cancer source (mastectomy, drug treatment, chemotherapy, radiation)
- Screening/mammography guidelines/recommendations are discussed (several organization could be listed here: U.S. Preventive Service Task Force; National Comprehensive Cancer Network; American Cancer Society; American College of Obstetricians and Gynecologists; American College of Radiology; Society of Breast Imaging)
- Survival rates (from specific treatments [chemotherapy/radiation; specific breast cancer surgery]; for the celebrity's specific breast cancer event)
- Contact information for further information about breast cancer is provided (a url link/web address or phone number)

This is not an exhaustive list of the information included in a news article with a thematic frame, nor do all of these points or information need to be included to conclude the article has a thematic frame. What you need to take from this list is that a news article with a thematic frame will discuss the celebrity and her or his experience with breast cancer, and then add health/breast cancer information to the story that is not about the celebrity. Instead, this information gives population and subpopulation breast cancer information.

A final note: Once you are done reading the article, ask yourself: Could someone read this article and learn something about breast cancer (prevention, treatment, risk factors, population rates of diagnoses, population rates of mortality, etc.)? If yes, then you should select *Thematic frame* in Qualtrics.

If after reading the article, you conclude that someone reading the article would only learn about the celebrity's/individual's experience with breast cancer, then you should select *Episodic frame* in Qualtrics.

> If a news article fits the definition of a *thematic frame*, select *Thematic frame* in Qualtrics for **V8**.

Section 3 – Statements About Breast Cancer

Section 3 asks you to identify specific types of information/statements made about the celebrity's experience with breast cancer and advice she or he gives to others.

V9. Breast cancer can happen to anyone: Does the article directly quote the celebrity stating (or make an inference that the celebrity stated) or make an assertion something to the effect of "anyone (or any woman) can have/get breast cancer"?

- No
- Yes

Example: No woman is immune to breast cancer. It can happen to anyone.

Explanation: Because "it can happen to anyone" is explicitly stated, you would select *Yes* in Qualtrics.

Example: "The news will come as a shock to Fonda's fans and raise awareness that the disease can strike even superhealthy gym users."

Explanation: This is an example where "it can happen to anyone" is implicit rather than explicit. But, the general sentiment of the statement is that breast cancer can happen to anyone.

V10. Early detection: Does this article infer that early screening/mammography/detection is important/essential for survival.

- No
- Yes

Example: I got lucky by catching it early.

Explanation: This is an implicit statement. One can infer that the celebrity feels if she was diagnosed at a later point in time, she may not have survived or at least would have had a less positive outcome. You would select *Yes* for this statement.

Example: Roberts, 46, said she found the lump early and that the prognosis is good: 'My doctor expects me to be flying planes and hanging on to submarines in the middle of the Atlantic and scaling the Mayan pyramids in no time.'

Explanation: Same rationale as above. A statement like this would require selecting *Yes*.

Example: In a statement released Friday, the wife of the former Republican presidential contender said she and Mitt feel fortunate to have caught this early, before it became invasive.

Explanation: Same rational as above. A statement like this would require selecting *Yes*.

V10a. If you answered *Yes* to question V10, please copy and paste the identified statement in the space provided in Qualtrics under **V10a**.

V11. Screening saves lives: Does the article directly quote the celebrity stating (or make an inference that the celebrity stated) something to the effect of "mammograms save lives" or "screenings save lives"?

- No
- Yes

Example: Every producer, every person who urged me to do this, changed my trajectory. The doctors told me bluntly, 'That mammogram just saved your life.'

Explanation: A statement like this, which directly states that the mammogram saved her life, would require selecting *Yes*.

Example: The 40-year-old correspondent admitted she had been reluctant to have the public mammogram but went ahead after 'GMA' anchor Robin Roberts told her that if the story saved one life, it would be worth it.

'It never occurred to me that life would be mine," she [Amy Robach] said.

Explanation: Same rationale as above. A statement like this would require selecting *Yes*.

V11a. If you answered *Yes* to question V11, please copy and paste the identified statement in the space provided in Qualtrics under **V11a**.

V12. All women should begin breast cancer screening at 40: Does the article explicitly state or imply that all women or most women should begin routine breast cancer screening at the age of 40?

- No
- Yes

Examples: I'm 40 years old. I've never had a mammogram. I've avoided it. And I started thinking, 'Wow, if I've put it off, how many other people have put it off as well?'

At 40, she's at the age when it's recommended that women regularly check for breast cancer.

Explanation: This statement indicates that the celebrity should have started screening at age 40. A statement like this would require selecting *Yes*.

V12a. If you answered *Yes* to question V12, please copy and paste the identified statement in the space provided in Qualtrics under **V12a**.

V13. Breast self exam: Does the news article quote someone (the celebrity or expert, for example) or simply assert that women should conduct self breast exams?

- No
- Yes

Example: I can only hope my story will do the same and inspire every woman who hears it to get a mammogram, to take a self-exam.

Explanation: A statement like this would require selecting *Yes*.

Example: "I know it's going to save lives, because there's going to be someone who hadn't thought about doing a self-exam who, hearing that story, is going to say 'Wow! I better pay more attention,'" Bean said.

Explanation: This statement promotes self-exams. Therefore, a statement like this would require selecting *Yes*.

V13a. If you answered *Yes* to question V13, please copy and paste the identified statement in the space provided in Qualtrics under **V13a**.

V14. "1 in 8": Does the news article discuss a woman's probability of developing breast cancer in her lifetime as about "1 in 8," or that a woman in the general population has about a 12.5% chance of being diagnosed with breast cancer in her lifetime?

- Yes
- No

Example: Women born with the BRCA1 mutation have a 65 percent lifetime risk of developing cancer, compare with a 10 to 12 percent lifetime risk for women without the mutation.

- This statement is not exactly what I have written about (about 12.5% chance), but well within the "ballpark". This example or more explicit or exact phrasing (1 in 8 or 12.5% chance) should be coded as *Yes*.

V15. Mastectomy is the best therapy: If the article discusses the celebrity's double mastectomy/prophylactic/bilateral mastectomy decision, is there an assertion (may be implicit) that the mastectomy greatly improves the celebrity's chance of survival?

- No
- Yes

This might include discussion of how she/he doesn't have to worry about breast cancer anymore or she/he won't die of breast cancer now. Other possibilities include that there is an underlying assertion that a prophylactic/double/bilateral mastectomy improves the chances of survival over a lumpectomy/lumpectomy and chemotherapy.

Example: In an email to her "GMA" colleagues Friday, Robach wrote that doctors found a second tumor while performing life-saving surgery.

"I got very lucky finding the cancer through our ABC-sponsored mammogram and I got lucky choosing an aggressive approach, bilateral mastectomy, because while in surgery last week my surgeon found a second, undetected malignant tumor," said Robach, a mother of five.

Explanation: This statement indicates that the second tumor would have reduced Robach's chances of survival and therefore the bilateral mastectomy was Robach's best option. A statement like this would require selecting *Yes*.

There is one exception to this. If a celebrity was diagnosed with breast cancer AND tested positive for a BRCA (1 or 2) gene, then the prophylactic mastectomy does improve the celebrity's chance of survival. If this is the case, then select *No*.

Example: In March, Rep. Debbie Wasserman Schultz (D-Fla.), 42, revealed she had had a double mastectomy last year at the National Naval Medical Center in Bethesda. She had received a diagnosis of early-stage cancer in her right breast in December 2007 and had a lumpectomy. Then, she tested positive for the BRCA2 mutation and, after consulting with doctors and her husband, decided to have both breasts removed. She has had seven surgeries in all, including the insertion of silicone implants and having her ovaries taken out.

"The doctors said I had a 65 percent chance of a recurrence of cancer in the other breast," Wasserman Schultz said in a telephone interview. "Those odds were too high for me."

Explanation: Wasserman Schultz was diagnosed with breast cancer and tested positive for the BRCA 2 mutation. In this case, Wasserman Schultz's chances for survival were increased from the mastectomy. A statement like this would require selecting *No.*

Final Determination:

- If the article talks about the celebrity having 1) a double or prophylactic/double/bilateral mastectomy; 2) was diagnosed with breast cancer at some point; and 3) tested positive for the BRCA (1 or 2) gene (or breast cancer gene or the genetic mutation which causes breast cancer) then you should select *No.*
- If the article talks about the celebrity having a double/prophylactic/bilateral mastectomy and only one of the following, then you should select *Yes*:
    - diagnosed with breast cancer
    - tested positive for the BRCA (1 or 2) gene (or breast cancer gene or the genetic mutation which causes breast cancer)
- If the article discusses the celebrity having a double/prophylactic/bilateral mastectomy and does not discuss a breast cancer diagnosis or a BRCA confirmation, then you should select *Yes.*
- If the article simply discusses the decision and does not give any information about survival/mortality rates as compared to women who do not have a double/prophylactic/bilateral/ mastectomy (with the same diagnosis or BRCA confirmation) then you would select *Yes* in Qualtrics.
- If the article discusses the decision and does give information about survival/mortality rates as compared to women who do not have a double/prophylactic/bilateral mastectomy (with the same diagnosis or BRCA confirmation) OR gives risk rate reductions for all women (with the same diagnosis or BRCA confirmation) (not just focusing on how Angelina Jolie's risk was reduced, for example) who have the procedure, then you would select *No* in Qualtrics.
- If the article does not discuss or mention a celebrity having a double/prophylactic/bilateral mastectomy, then you should select *No.*

Final note: Other than the first or last options above (under *Final Determination*), after you read the article, ask yourself if someone who read this article would learn about under what conditions a woman would see the most benefit from a double/prophylactic/bilateral mastectomy. If your answer to that question is "no", then you would select *Yes* in Qualtrics.

V15a. If you answered *Yes* to question V15, please copy and paste the identified statement in the space provided in Qualtrics under **V15a**.