BMJ Open

Media Coverage of the Benefits and Harms of Testing the Healthy: a protocol for a descriptive study

Mary O’Keeffe,1 Alexandra Barratt,2 Christopher Maher,1 Joshua Zadro,*1 Alice Fabbri,3 Mark Jones,4 Ray Moynihan2,4

ABSTRACT

Introduction Much testing in medicine is aimed at healthy people to facilitate the early detection of health conditions. However, there is growing evidence that early detection is a double-edged sword that may cause harm in the form of overdiagnosis. The media can be seen as a major generator of consumer demand for health services. Previous research shows that media coverage tends to overstate the benefits and downplay the harms of medical interventions for the sick, and often fails to cover relevant conflicts of interest of those promoting those interventions. However, little is known about how the benefits and harms of testing the healthy are covered by media. This study will examine the media coverage of the benefits and harms of testing the healthy, and coverage of potential conflicts of interest of those promoting the testing.

Methods and analysis We will examine five tests: 3D mammography for the early detection of breast cancer; blood liquid biopsy for the early detection of cancer; blood biomarker tests for the early detection of dementia; artificial intelligence technology for the early detection of dementia; and the Apple Watch Series 4 electrocardiogram sensor for the early detection of atrial fibrillation. We will identify media coverage using Google News and the LexisNexis and ProQuest electronic databases. Sets of two independent reviewers will conduct story screening and coding. We will include English language media stories referring to any of the five tests from January 2016 to May 2019. We will include media stories if they refer to any benefits or harms of the test for our conditions of interest. Data will be analysed using categorical data analysis and multinomial logistic regression.

Ethics and dissemination No ethical approval is required for this study. Results will be presented at relevant scientific conferences and in peer-reviewed literature.

INTRODUCTION

Much testing in medicine is aimed at apparently healthy people to identify those at an increased risk of a disease or disorder.1 These ‘healthy’ people can subsequently be offered more tests, treatment(s) or preventive strategies (eg, a preventive medicine).1 The increasing popularity of testing is indicative of recent enthusiasm for early detection,2 which is part of the promise of ‘precision medicine’. That is, early detection is always better, and treatment is more effective when it is tailored to the individual.3 Apparently healthy or well people are increasingly encouraged to proactively monitor, and be vigilant about understanding their health, with testing seen as a positive step in consumer health empowerment. However, there is mounting evidence that testing can harm healthy people, and the quest for ever-earlier detection of disease can lead to unnecessary classification of the healthy as sick: overdiagnosis.3–8

Although an exact definition of overdiagnosis remains the subject of debate, particularly in the context of non-cancer conditions, overdiagnosis can be considered to occur when persons are labelled with a technically correct diagnosis that does not improve health outcomes.9 10 Key drivers of overdiagnosis have been identified.11 One is the use of more sensitive tests which can detect smaller abnormalities, many of which are benign.11 There is growing evidence demonstrating the presence of overdiagnosis, often arising through testing healthy people, across different areas of medicine. Examples include screening for cancer (eg, breast, prostate, thyroid), cardiovascular disease and dementia.7 8 12–16 Inappropriate screening in this context is likely to lead to higher healthcare spending and worse outcomes (eg, psychological effects and unnecessary and harmful treatments).17–21

Sustained promotion to the public, patients and clinicians of the importance

Strengths and limitations of this study

► This will be the first study to analyse the media coverage of the benefits and harms of tests that have the potential for overdiagnosis in healthy people.
► Media stories will not be restricted by country. The results could inform interventions to improve the quality of medical reporting in the media.
► The study will only consider media coverage of five tests.


Protocol
of early detection and testing, including via the media, is considered another driver of overdiagnosis. Uncritical coverage of new tests, without consideration of their potential downsides, contributes to the general lack of knowledge about the potential harms of getting tested when healthy. In fact, research has shown that only a small proportion of people are knowledgeable about overdiagnosis. This includes individuals offered tests where the potential for overdiagnosis is high. As such, patients (and clinicians) overestimate the benefits of testing, while underestimating the harms.

There is also concern about how changing media environments, such as the rising influence of social media, can lead to ‘junk-food news’. Indeed, previous studies on the media have identified evidence of exaggeration, inaccurate media coverage of published scientific papers, overstating of benefits of treatments, downplaying of harms and failure to report important conflicts of interest of the experts cited in the story. Concern has also been expressed about the financial closeness between journalists and industry. For example, pharmaceutical industry funding of journalism practice, awards and education, has been documented. Further, there appears to be a lack of independent medical research commentators in the media. Only one in six media reports of research published in high-impact medical journals included comments from people who were independent of the study investigators. Moreover, one in three of the independent commenters had financial conflicts of interest, most of which were not disclosed in the media stories.

Poor media coverage of medicine is a significant issue; it can influence how the public perceives the risk of health services and how patients make treatment decisions. For example, media coverage about the celebrity Kylie Minogue’s self-referral mammogram bookings led to a 20-fold increase in media coverage about breast cancer and a 40% increase in mammogram bookings during the 2-week peak after the interview. Six weeks later, media coverage was still up by 30%.

Media coverage of overtreatment has been examined in one study which examined the framing of medical overtreatment in United States (US) newspapers from January 2007 to December 2010. The study found that the media focused on the harms of overtreatment relating to cancer, but the overall media coverage may have implied that overtreatment was not seen as an issue across other health conditions.

To date, however, no studies have examined media coverage of new tests with significant potential for overdiagnosis. Furthermore, there has been little formal, rigorous evaluation of the media’s coverage of testing healthy people. An evaluation may give us an insight into how the media can contribute to overdiagnosis, and inform potential strategies to enhance media coverage of medical testing. In this study, we aim to examine the media coverage of the benefits and harms of testing the healthy, and the potential conflicts of interest of those promoting the testing, by examining the coverage of five tests.

**METHODS AND ANALYSIS**

**Overview**

We will conduct a large descriptive cross-sectional study of global English-language media coverage of five tests from January 2016 to May 2019.

**Tests and conditions of interest**

This study will focus on five tests:

1. 3D mammography for the early detection of breast cancer
2. Blood liquid biopsy tests for the early detection of cancer(s)
3. Blood biomarker tests for the early detection of dementia
4. Artificial intelligence (AI) technology for the early detection of dementia
5. Apple Watch Series 4 ECG sensor for the early detection of atrial fibrillation (AF)

We identified these tests based on the following criteria:

1. Evidence-based concerns about overdiagnosis
2. Evidence of Food and Drug Administration (FDA) approval for certain classes of each test
3. Concern that the results of these tests may not lead to improved health outcomes for some individuals; either due to the unavailability of effective treatment options (eg, dementia) or treatments that may cause more harm than benefit (eg, early mammography)
4. Identifiable groups or companies with a financial interest in promoting these tests, or maximising the markets for downstream treatments
5. Notable media coverage

One reviewer (MO) used Google News to track media coverage of tests for the healthy between April 2018 and October 2018. Results were recorded in a Microsoft Excel file. Based on these results, the same reviewer designed a series of Google Alerts (running between April 2018 and December 2018) with specific keywords related to testing. The Google Alert results were screened to identify media coverage on tests for healthy people which met the first four criteria mentioned above. Consideration was given to conducting a descriptive study of a random sample of recent media coverage for all tests. However, we felt that tracking a specific number of tests over time would give us a more comprehensive picture of how the benefits and harms of testing were reported in the media over time. Five tests that were receiving notable media coverage were included. Below we provide details of the five identified tests.

**3D mammography for the early detection of breast cancer**

3D mammography or digital breast tomosynthesis is an advancement on traditional mammography. Many companies have received FDA approval for 3D mammography, partially based on an emerging body of evidence...
on its detection capability relative to 2D mammography screening. In March 2019, the FDA announced new policies to change current mammography standards in the US. The proposed changes aim to increase the use of 3D mammography screening.

There is a high level of uncertainty surrounding the benefits of 3D mammography. For example, a systematic review and meta-analysis found that 3D mammography improves cancer detection rate and reduces recall for assessment compared with traditional mammography. However, these improvements varied by setting; compared with retrospective studies in the US where annual screening was encouraged, there were greater improvements observed in prospective studies embedded in the European biennial screening programme.

There is a large amount of research on the benefits and harms of mammography, particularly compared with the other four tests we selected. Overdiagnosis is a significant harm of screening mammography, and notwithstanding uncertainty in the data, there are estimates of its frequency ranging between 10% and 50% of screen-detected cancers. Overdiagnosis of breast cancer can lead to unnecessary surgery, radiotherapy and endocrine therapies, as these are standard treatments for cancer can lead to unnecessary surgery, radiotherapy and endocrine therapies, as these are standard treatments for cancer can lead to unnecessary surgery, radiotherapy and endocrine therapies. These treatments cause harm through physical and psychological effects that can impact on quality of life. Overall, little is known about whether the improved cancer detection rate estimated for 3D mammography screening—compared with 2D mammography screening—will have additional benefit (ie, whether it will further reduce breast cancer mortality and/or increase quality of life) or whether it will lead to harms (eg, overdiagnosis), or a combination of both.

**Liquid biopsy for the early detection of cancer**

A liquid biopsy is a blood test using genomic profiling to detect mutations or cancer cells (circulating tumour DNA cells). The FDA approved the first liquid biopsy in 2016. This class of liquid biopsy was designed for clinicians to monitor cancer status and patient response to treatment. The FDA has since (April 2018) approved another liquid biopsy test by a molecular information company called Foundation Medicine.

While liquid biopsy was initially designed for monitoring patients with cancer, there seems to be increasing interest in its use for the early detection of cancer, and that the test may eventually be used for routinely screening people and detecting cancers before they cause symptoms. In fact, there are ongoing studies assessing whether the test can detect tumours in seemingly cancer-free individuals. There is a lot of uncertainty surrounding the effectiveness of liquid biopsy for both early detection, and improvement of cancer treatment.

There are also concerns that the detection of circulating tumour DNA cells in asymptomatic populations could lead to overdiagnosis. The concerns are linked to findings that circulating tumour DNA cells and cancer-related mutations have been detected in healthy individuals who never go on to develop a cancer. It has also been mentioned that the cancer-related proteins used by liquid biopsy can reflect tissue damage common in inflammatory conditions like arthritis, in the absence of cancer.

**Blood biomarker tests for the early detection of dementia**

There is enormous interest in identifying a cheap and simple test to detect dementia in the early stages, with the hope this will improve the treatment of dementia. In recent years, there has been particular interest in blood tests to detect abnormal levels of two proteins: amyloid beta and tau. Both are considered biomarkers of dementia. The FDA granted approval to a genetic company called ‘23andMe’ in 2017 to offer direct to consumer tests for the early detection of dementia. The decision has generated controversy with commentators concerned about what they see as a lack of robust evidence to support the testing in the early detection of dementia, concerns about overdiagnosis and the implications of testing a condition that has not yet been shown to be amenable to intervention.

In January 2019, the Cochrane Dementia and Cognitive Improvement Group published a commentary expressing their concerns about the increasing use of biomarker tests in dementia. In their commentary, they referred to research demonstrating that up to 60% of healthy people over 80 years could be labelled as having dementia under new disease definitions, even though these people may never develop clinical symptoms. In the same commentary, the authors stated that reducing dementia to positive amyloid biomarkers is ‘an open invitation to overdiagnosis’. Further to this, authors refer to the data documenting the psychological, social and legal harms of overdiagnosing or overpredicting dementia. Finally, they expressed concerns about the lack of data validating the proposed biomarkers for dementia.

**Artificial Intelligence (AI) technology for the early detection of dementia**

Similarly to blood tests, there is a keen interest in using artificial intelligence (AI) technology to improve the early detection of dementia. AI refers to the use of machine algorithms (eg, computer programmes) to model intelligent behaviour with minimal human intervention. The FDA has now approved one type of AI technology to monitor brain structures in different neurological conditions, including dementia. Since February 2018, the FDA have also relaxed regulatory policies on drugs for dementia so they could approve the delivery of disease treatments to people displaying certain biological signals years before the disease shows outward signs. AI technology has been proposed as a key step in detecting these subtle biological signals of dementia (eg, small changes in brain size, metabolic changes, memory recognition, voice recognition, etc.). Researchers have expressed concerns about the lack of robust clinical research in this.
area, and the potential of AI to lead to false positives, and overdiagnosis.\textsuperscript{60}

Apple Watch Series 4 electrocardiogram (ECG) sensor for the early detection of atrial fibrillation (AF)

The Apple Watch Series 4 was released in 2018. It features an electrocardiogram (ECG) sensor that has received FDA medical device approval.\textsuperscript{61} The primary rationale for this new sensor is to facilitate the early detection of atrial fibrillation (AF). AF is the most common heart arrhythmia (irregular heartbeat), and can be associated with an increased risk of early mortality, heart failure and stroke.\textsuperscript{62–64} Because of this, a diagnosis of AF can often lead to drug therapy and in some cases, surgery.\textsuperscript{63}

The Apple Watch ECG sensor records electrical impulse patterns of a person’s heart to predict AF. The sensor will inform the user of the presence of AF and advise on the need for medical consultation. Many concerns have been expressed about testing the healthy for AF, and it has been suggested that overdiagnosis is ‘only a matter of time’ with the Apple Watch.\textsuperscript{65–67} There is a concern regarding the poor specificity of testing methods for AF.\textsuperscript{66} Furthermore, AF has a low prevalence\textsuperscript{68} and screening the healthy could potentially lead to harms in the form of false positives, overdiagnosis and overtreatment.\textsuperscript{69} In fact, some researchers state that the Apple Watch specifically could lead to a misdiagnosis of AF in ~1 million people for every 10 million screened.\textsuperscript{67} This may lead to harms from overtesting, bleeding from unnecessary anticoagulation and anxiety due to having a cardiac diagnosis.\textsuperscript{67}

There also seems to be a lack of knowledge around the natural history of AF. For example, there is uncertainty surrounding the outcome of untreated stroke risks, so the net benefit of treating AF with anticoagulants is unclear. Finally, while screening for AF leads to increased detection, office visits and prescriptions for anticoagulants,\textsuperscript{69} there is still uncertainty around its effects on health outcomes.

Inclusion and exclusion criteria

We will include media stories referring to any of our five target tests for the corresponding conditions of interest. Media stories will be included if they refer to any benefits (eg, early detection of the condition, early treatment of the condition, prevention of the condition, saves lives) or harms of the test (eg, overdiagnosis, inappropriate diagnostic testing, misdiagnosis, false alarms, false positives, false negatives, unnecessary and/or harmful treatment, psychological distress, health anxiety, costs). We will exclude media stories that only focus on tests for symptomatic people or people who already have the condition of interest (eg, mammography for monitoring the progression of breast cancer), media stories about patent approval or business issues only, press releases, conference proceedings, trade journal reports and scholarly journal articles. We will first pilot our screening process. Depending on the results of the pilot, we may add additional criteria, or provide more detail on the current inclusion and exclusion criteria.

Search strategy

We will identify all relevant English-language media stories by searching the LexisNexis and ProQuest electronic databases, using explicit keywords, from January 2016 to May 2019. In line with a previous study on media coverage of medicine,\textsuperscript{70} we will supplement this database search with a Google News search, reviewing the first 20 pages of each Google search result. Different keywords will be required for each of the five tests; therefore, five searches (one specific to each test) will be performed. A librarian/information specialist with expertise in systematic review search design will assist with the search strategy. We will not restrict articles by country. Our searches will cover all of the following media coverage: newspapers, major world publications, blogs, magazines, broadcast and podcast transcripts, wire feeds/services and webnews. These are named categories within the LexisNexis and Proquest databases.

Screening process

Sets of two independent reviewers will be involved in performing the screening of media stories for each test. We will exclude exact duplicates (same title, same outlet and same date) before starting the screening and will keep track of the number of duplicates. Reviewers will independently assess the eligibility of media reports for potential inclusion according to the predefined selection criteria. Any disagreements in judgement will be resolved by discussion to reach consensus or by consultation with a third reviewer. Syndicated studies will be included but will only be coded once. For example, if there are 10 media stories about the Apple Watch where the same or extremely similar story has been run across multiple media outlets, we will code this story once, but include the number 10 as the number of media reports about the Apple Watch.

Data extraction and coding

We will use a structured template (see table 1) to extract and code the relevant data in Research Electronic Data Capture hosted at The University of Sydney.\textsuperscript{71} The data extraction tool will be adapted from tools used in similar studies\textsuperscript{27 70 72} and an iterative design process will be used to refine the tool for the purpose of our study. Sets of two independent reviewers will extract data and code the media stories; two independent reviewers for each test. Any disagreements in extraction or coding will be resolved by discussion to reach consensus or by consultation with a third reviewer. The percentage of disagreements on each coding variable requiring resolution through use of a third reviewer will be recorded. Before formal data extraction and coding, the sets of independent reviewers will apply the data extraction tool to code 20 media stories; four for each test. Disagreements in data extraction and coding will be resolved by discussion.
Table 1  Draft coding tool

<table>
<thead>
<tr>
<th>Media story description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name of media source</td>
</tr>
<tr>
<td>Country of media source</td>
</tr>
<tr>
<td>Word count</td>
</tr>
<tr>
<td>Release date</td>
</tr>
<tr>
<td>Author name</td>
</tr>
<tr>
<td>Type of media</td>
</tr>
<tr>
<td>Test mentioned</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Mention of the health condition that the test is used for?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Context about the screening test (benefits and harms)</td>
</tr>
<tr>
<td>How benefit was described</td>
</tr>
<tr>
<td>Any benefit mentioned or implied?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Any benefit quantified?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Was anecdote or other real-life example of benefit given?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Anecdote</td>
</tr>
<tr>
<td>Celebrity</td>
</tr>
<tr>
<td>Benefits referred to as revolutionary, life-saving, breakthrough, leading to improved treatment</td>
</tr>
<tr>
<td>How harm was described</td>
</tr>
<tr>
<td>Any harm mentioned or implied?</td>
</tr>
<tr>
<td>Any harm quantified?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Was anecdote or other real-life example of harm given?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Anecdote</td>
</tr>
<tr>
<td>Celebrity</td>
</tr>
<tr>
<td>Any specific harms of screening mentioned? (eg, overdiagnosis, potential overtreatment)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Evidence of conflicts of interest</td>
</tr>
<tr>
<td>Any specific scientific study quoted or mentioned about the screening test?</td>
</tr>
<tr>
<td>Does the scientific study disclose any financial ties of the authors to the manufacturers of the screening test discussed in story?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>If yes, state verbatim</td>
</tr>
<tr>
<td>Did the media story include information about financial ties of the authors to the manufacturer of the screening test? (if relevant)</td>
</tr>
<tr>
<td>Does the scientific study mention receipt of study funding from the manufacturers of the screening test discussed in story?</td>
</tr>
<tr>
<td>Did the media story include information about receipt of study funding from the manufacturers of the relevant screening test? (if relevant)</td>
</tr>
</tbody>
</table>

Continued
Table 1  Continued

<table>
<thead>
<tr>
<th>Other sources quoted or mentioned about the screening test</th>
<th>Yes/No</th>
<th>1=Yes, 0=No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physician/provider?</td>
<td>Yes/No</td>
<td>1=Yes, 0=No</td>
</tr>
<tr>
<td>Patient quoted?</td>
<td>Yes/No</td>
<td>1=Yes, 0=No</td>
</tr>
<tr>
<td>Other source?</td>
<td>Yes/No</td>
<td>1=Yes, 0=No</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overall impressions/tone</th>
<th>Media story leaves you with sense that the screening test is</th>
<th>1=Beneficial overall, 2=Harmful overall, 3=Neutral (balanced information given about benefits and harms), 4=Unclear</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Overall tone about going for the screening test</th>
<th>Leave a comment or paste here anything interesting (even quotes)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Leave a comment or paste here anything interesting (even quotes)</td>
</tr>
</tbody>
</table>

This is a draft coding tool. The tool may be modified once the reviewers pilot test it with a sample of included media stories.

and subsequent revisions to the data extraction tool. We plan to extract information about the media story (eg, type of media, test mentioned, country of origin), benefits mentioned, harms mentioned, disclosure of conflicts of interest and the overall tone of the story (positive or negative).

**Data analysis**

Descriptive statistics (means, SD, counts and percentages) will be used to summarise the extracted data (eg, number of stories, number of countries, number reporting benefits and harms, etc.). Analysis will be performed separately for each test. Categorical data analysis will be used to investigate potential associations between overall impression of the media reports and explanatory variables including conflicts of interest, time and type of media. We plan to use multinominal logistic regression where the dependent variable is overall impression and a neutral impression is the reference category. We will report odds ratios and 95% CIs for negative impressions and positive impressions associated with the independent explanatory variables (as referred to above). Analysis will be conducted separately for each test. We will use $\chi^2$ tests to compare the distribution of categories of overall impressions across the five tests. We will also outline at least one example of a media story for each test in the Results section.

**Ethical considerations**

No ethical approval is required for this study.

**Patient or public involvement**

Patients and members of the public were not involved in the design of this study.

**Dissemination**

The results of this study will be published and presented at relevant medical conferences. We anticipate the results of this study will inform the development of an intervention to improve the quality of medical reporting in the media.

**DISCUSSION**

This study will examine media coverage reporting the benefits and harms of five medical tests that are controversially being promoted to the healthy. While other drivers, including research and professional prominence given to early detection,2 are important, sustained media coverage is likely a powerful source of influence of public attitudes towards new tests. The results will provide valuable information about the quality of media reporting on tests targeting the healthy and may help inform interventions to improve the quality of medical reporting.

**Acknowledgements**

We would like to thank Professor Andrew D. Oxman (Centre for Informed Health Choices, Norwegian Institute of Public Health, Oslo, Norway and University of Oslo, Oslo, Norway) for providing feedback on an earlier draft of this protocol.

**Contributors** MO, RM, AB and CM have been primarily responsible for study conception, design and designing the data coding approach. MJ advised on statistical analysis. JZ and AF helped pilot the search strategy and coding tool. MO drafted the first version of this manuscript. All authors provided critical evaluation and revision of the manuscript and had given final approval of the manuscript accepting responsibility for all aspects.

**Funding** This study will not receive any specific funding. MO is supported by a European Union Marie Skłodowska-Curie postdoctoral fellowship. AB and RM are investigators on a National Health and Medical Research Council of Australia (NHMRC) funded CRE grant no 1104136. RM is supported by an NHMRC grant, #1124207. AF is a postdoctoral fellow on an Australian NHMRC project grant no1122332. CM is supported by an Australian NHMRC Research Fellowship.
REFERENCES

54. Brunet M. Targets for dementia diagnoses will lead to overdiagnosis. BMJ 2014;348:g2224.