BMJ Open

Effects of whole-body magnetic resonance imaging on health service utilization in a general-population cohort

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-056572
Article Type:	Original research
Date Submitted by the Author:	23-Aug-2021
Complete List of Authors:	Schmidt, Carsten; University Medicine Greifswald, ICM - SHIP/KEF Sierocinski, Elizabeth; University Medicine Greifswald Baumeister, Sebastian; LMU Hegenscheid, Katrin Völzke, Henry; German Center for Diabetes Research, Greifswald; University Medicine Greifswald Chenot, J. -F.; University Medicine Greifswald, Institute for Community Medicine
Keywords:	Magnetic resonance imaging < RADIOLOGY & IMAGING, EPIDEMIOLOGY, PUBLIC HEALTH

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Effects of whole-body magnetic resonance imaging on health service utilization in a general-population cohort

Schmidt CO (PhD) [1], Sierocinski E (MD) [1,2], Baumeister SE (PhD) [3,4], Hegenscheid K (MD) [5,6], Völzke H (MD) [1,7], Chenot JF (MD, MPH) [2]

[1] Institute for Community Medicine Department SHIP-KEF, University Medicine of Greifswald; Walther Rathenau Str. 48, 17475 Greifswald, Germany (Carsten.schmidt@uni-greifswald.de)

[2] Institute for Community Medicine, Department of Family Medicine, Fleischmannstr. 42, 17475 Greifswald, Germany (jchenot@uni-greifswald.de, elizabeth.sierocinski1@uni-greifswald.de)

[3] Chair of Epidemiology, LMU München, UNIKA-T Augsburg, Augsburg, Germany

[4] Independent Research Group Clinical Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich, Germany (s.baumeister@unika-t.de)

[5] Unfallkrankenhaus Berlin, Warener Str. 7, 12683 Berlin, Germany

[6] Department of Diagnostic Radiology and Neuroradiology, University Medicine of Greifswald, 17475 Greifswald, Germany (katrin.hegenscheid@uni-greifswald.de)

[7] DZHK - German Centre for Cardiovascular Research, Partner Site Greifswald, Germany; DZD - German Centre for Diabetes Research, Site Greifswald, Germany (voelzke@uni-greifswald.de)

Running title: MRI findings and health service utilization in a general-population setting

Correspondence:

Carsten Oliver Schmidt

Study of Health in Pomerania/KEF

University Medicine of Greifswald, Institute for Community Medicine

Walther Rathenau Str. 48, D-17475 Greifswald

carsten.schmidt@uni-greifswald.de

Tel +49[0]3834/86-7713, Fax +49[0]3834/86-6684

Abstract

Objective: Whole-body magnetic resonance imaging (wb-MRI) is increasingly used in research and screening but little is known about the effects of incidental findings (IFs) on health service utilization and costs. Such effects are particularly critical in an observational study. We studied the effects of wb-MRI on ambulatory health care in a general population sample.

Design: Prospective cohort study.

Setting: General population Mecklenburg-Vorpommern, Germany.

Participants. Analyses included 5019 participants of the Study of Health in Pomerania (SHIP) with statutory health insurance data. 2969 took part in a wb-MRI examination in addition to a clinical examination program that was administered to all participants. MRI non-participants served as a quasi-experimental control group with propensity score weighting to account for baseline differences.

Primary and secondary outcome measures: Outpatient costs (total health care usage, primary care, specialist care, laboratory tests, imaging) during 24 months after the examination were retrieved from claims data. Two-part models were used to compute treatment effects.

Results: In total 1366 potentially relevant IFs were disclosed to 948 MRI participants; most concerned masses and lesions (769 participants, 81%). Costs for outpatient care during the two-year observation period amounted to an average of €2547 (95%CI: €2424-€2671) for MRI non-participants and to €2839 (95%CI: €2741-€2936) for MRI participants, yielding an

average treatment effect of €295 (95%CI: €134-€456) per participant. Imaging and specialist care related costs were the main contributors to the increase in costs.

Conclusions: Communicated findings from population-based wb-MRI substantially impacted health service utilization and costs. This introduced bias into the natural course of health care utilization and should be taken care for in any longitudinal analyses.

Strengths and limitations

1. This is the first study to report disclosed longitudinal effects of disclosed incidental whole body MRI findings on costs for outpatient services in a general population setting.
2. The longitudinal cohort design, the large sample size and the availability of a group with and without MRI participation are considerable strengths.
3. The potential impact of selection bias is reduced by the wide coverage of our participants by claims data. Claims data provide an appropriate approximation of outpatient costs as they underlie reimbursement purposes.
4. Limitations concern the scope of claims data as they do not cover medication or inpatient care, leading to an underestimation of the total healthcare-related expenditures. Furthermore, the small subgroup of privately insured participants is not represented.
5. Non randomized allocation to MRI participation in the observational cohort design reduces the comparability of both groups and statistical measures have been taken to balance potential confounders.

1 Introduction

Screening with whole-body magnetic resonance imaging (wb-MRI) may detect asymptomatic disease at an early stage thus improve treatment outcome [1], but might also cause unnecessary psychosocial distress, medical interventions, and costs due to irrelevant findings [2, 3]. Wb-MRI has evolved into a key examination tool in state-of-the-art population research [4-7]. It produces a large number of incidental findings (IFs), a large proportion of which represent potentially significant masses and lesions [2, 8-11]. Although a minority may benefit from IFs in a general population approach [1], there is a high risk of false-positives, overtesting and overdiagnosis [12, 13]. Only few malignancies were newly detected despite a large number of biopsies conducted after participation in SHIP [14] and no positive effects on quality of life were found [8].

To the best of our knowledge, there is no prospective evidence on the effects of wb-MRI IFs in a general population setting on ambulatory health service utilization. Such data is needed to guide the appropriate handling of wb-MRI screening results in research and clinical practice.

Our principal research question was therefore how participation in wb-MRI is associated with outpatient service utilization and costs. MRI non-participants served as a quasi-experimental control group with propensity score weighting to account for baseline differences. We hypothesized that participation in the MRI examination would lead to increased costs in ambulatory healthcare. Furthermore we assume that there are differential effects on different types of ambulatory services.

2 Methods

2.1 Study design and sample

SHIP is a population-based project consisting of two independent cohort studies, SHIP and SHIP-Trend. Participants were sampled from the counties of North and East Western Pomerania and the cities of Greifswald and Stralsund in Germany [7]. Participants were between 20 and 79 years of age at the date of sampling and had the target region listed as their primary place of residence. The aim in both cohorts was to recruit a representative general population sample. Clinical status was of no relevance for inclusion. Participants received no payment beyond reimbursement for travel costs.

A two-stage sampling scheme was adopted from the German World Health Organization's (WHO) MONICA Project for the first cohort [15]. Out of 6265 eligible individuals of the first cohort, 4308 (2192 women) participated at baseline (response 68.8%) [16]. Baseline examinations were performed from 1997-2001 (SHIP-0). Follow-up examinations took place between 2002 and 2006 (SHIP-1, N=3300) and between 2008 and 2012 (SHIP-2, N=2333).

A second cohort (SHIP-Trend) was established in 2008 from a stratified sample of 10000, drawn from the central population registry. Examinations took place until 2012. Of the net sample of 8826, after exclusion of deceased and relocated participants, 4420 (2275 women) participated (response 50.1%).

Of the 6753 participants in SHIP-2 and SHIP-Trend, 6312 had statutory health insurance (93%). Subjects ineligible for MRI participation for reasons such as claustrophobia, metal implants, or pregnancy were excluded from analyses [17] because treatment effects could

not be meaningfully computed for this group (n=1028). Of the resulting 5284 participants, 265 (5%) refused linkage with statutory health insurance. Our analyses included the remaining 5019 participants (Figure 1). The follow-up period for claims data was two years after SHIP participation.

The Ethics Committee of the University Medicine of Greifswald approved the study protocol (BB 39/08, BB 106/10).

[Figure 1 here]

2.2 MRI Examination

All SHIP participants were invited to take part in the MRI examination and received relevant written educational material. During a medical interview before the examination, a radiologist described the handling of IFs and conditions for disclosure [8], methods descriptions in 2.2 and 2.3 have been taken from previous SHIP publications [14, 17].

All wb-MRI were acquired on a 1.5-Tesla system (Magnetom Avanto; Siemens Medical Solutions, Erlangen, Germany). The wb-MRI protocol was identical for all participants and included a plain whole-body MRI and detailed imaging of the head, neck, chest, abdomen, pelvis, and spine. Men had the option of contrast-enhanced cardiac MRI and MR angiography, and women had the option of cardiac MRI and contrast-enhanced MR mammography. The complete imaging protocols have been described previously [8, 18].

Findings and anatomical variants were documented in a standardised reading protocol. The radiologists reading the scans had no access to the participants' clinical information. Scan

reading was performed using a digital picture archiving and communication system (IMPACS ES 5.2, AGFA Healthcare, Mortsel, Belgium). First-line reading was performed by two independent radiology residents. A third reader, a senior radiologist, resolved disagreements.

The MRI examination was entirely financed by SHIP study funding and did not contribute to billing costs represented in the claims data.

2.3 Disclosure of incidental findings

A standardised protocol regulated the handling of wb-MRI IFs. Findings were classified into three categories: Category I comprised normal or common findings in asymptomatic individuals (e.g., anatomical variants, old brain infarcts). Category II findings were abnormalities of potential clinical relevance. Category III findings required immediate referral. Category II findings were disclosed in writing via post after approval by an interdisciplinary advisory board. Category III findings were disclosed immediately to the participant. A detailed description of this protocol has been provided elsewhere [8]. Our analyses are based on Category II and III findings.

Health-related findings from other examinations such as blood testing, blood pressure measurements, somatometry, ultrasound, and cardiovascular examinations were also disclosed to SHIP participants [7].

2.4 Claims data

Claims data from the regional Association of Statutory Health Insurance Physicians included billing codes for medical and technical services (e.g. imaging) and costs for outpatient care

(excluding medication costs). Germany has a mixed billing system with capitation and fee-for-service models for specific services. We computed costs per quarter (billing period) for the two quarters prior to the examination, the quarter during which the examination took place, and the eight quarters following the examination. We distinguished between (1) total outpatient costs, (2) primary care and prevention, (3) specialist care, (4) laboratory work, and (5) imaging. All costs related to pregnancy and births were excluded, as pregnancy was an exclusion criterion for MRI participation.

2.5 Statistical analyses

We used two-part models to analyze cost variables with their zero inflated distribution [19]. Each two-part model comprised a multivariable logistic regression for any healthcare service provision and a generalized linear model with a log-link and a gamma distribution for the height of costs. Models used the predictors time, MRI participation (yes vs. no), and the interaction term to model average treatment effects. Propensity score reweighting was applied to balance the distribution of relevant covariates between MRI participants and non-participants [20, 21].

Logistic regression models were used to estimate propensity scores with the following predictors: costs for total healthcare, primary care and prevention, specialist care, laboratory work and imaging in the two quarters prior to the SHIP-2/SHIP-Trend examination, as well as age, sex, level of education (<10 years vs. ≥10 years), marital status, current employment, smoking status, and quality of life using the SF-12 mental health component summary score and physical health component summary score [22]. Resulting weights ranged from 1.0 to 13.5 (mean: 2.5; standard deviation 0.8). Standardized mean

differences (SMD) of all baseline variables after weighting were close to zero (min: -.007, 10%pct: -.004, median: 0.90%pct: .003, max: .005). Balance was also checked based on distributional properties of all baseline variables and all statistical interaction effects (SMD distribution: min: -.045, 10%pct: -.019, median: -.002, 90%pct: .009, max: .045).

Item missingness is reported in Table 1. Cost data was unavailable for 7.4% of all quarters. For 4.7% (N=238) no doctoral visit was coded at all during the observation period. Given the plausibility of a subgroup not presenting to ambulatory care regularly, we set all types of costs in these quarters to 0€ as a coded visit is a prerequisite for costs. Nevertheless, we cannot be certain that data for some participants who did present to the doctor was missing. We therefore used multiple imputation by chained equations (MICE) as a sensitivity analysis to impute missing information on costs and coded neoplasms. These results did not lead to different conclusions and are thus not presented.

Two-sided tests were applied throughout. Analyses were conducted in Stata 14 (Stata Corp., College Station, TX). Figures were generated using Microsoft PowerPoint.

2.6 Patient and public involvement

Patients, study participants and the public were not directly involved in the design, conduct and reporting of this study.

3 Results

3.1 Sample characteristics and incidental findings

MRI non-participants were on average three years older, had a lower level of education, were more often unemployed, smokers, and less often married compared to MRI participants (Table 1). Total ambulatory costs were similar in both groups and quality of life scores were slightly lower among non-MRI participants.

In total, 1366 IFs of potential clinical relevance were disclosed to 948 participants (32%). Of these, 769 participants (81%) received a finding related to masses and lesions, corresponding to 26% of all MRI participants.

[Table 1 here]

3.2 Descriptive course of outpatient costs

The course of outpatient costs is displayed stratified for MRI non-participants, MRI participants with disclosed findings, and MRI-participants who did not receive findings (Figure 2). The latter two groups were distinguished because a marked increase of costs after the SHIP examination was only expected among those with a disclosed finding. Increases occurred in all studied groups but peaks were highest in the quarter after the SHIP examination among MRI participants with disclosed findings (Figure 2 and 3). The relative increase was largest for imaging-related costs, which more than doubled (Figure 3). While decreasing after an initial peak, total costs among MRI participants with disclosed findings did not return to the initial level but remained higher (Figure 2 and 3).

[Figure 2 here]

[Figure 3 here]

3.3 MRI participation and outpatient costs

Propensity score reweighting revealed increased excess costs among MRI participants compared to non-participants that persisted over time (Figure 2 and 3).

Total weighted costs for outpatient care during the two-year observation period amounted to an average of €2547 (95%CI: €2424-€2671) for MRI non-participants and to €2839 (95%CI: €2741-€2936) for MRI participants, yielding an average treatment effect of €295 (95%CI: €134-€456) per participant (Table 2). Additional costs were higher in the second post-examination year compared to the first year. The largest contribution to excess costs resulted from specialist care, followed by imaging.The smallest effects were related to laboratory costs.

[Table 2 here]

4 Discussion

4.1 Main findings

A single wb-MRI examination sufficed to increase long-term overall outpatient costs over two years in a general population sample. Effects were larger for certain services such as specialist care and clinical imaging. From a research perspective, our results illustrate how disclosed IFs turned an observational study into an intervention. This limits the generalizability of research findings on outpatient costs and health service utilization to the underlying general population. From a clinical perspective, overtesting and overdiagnosis are likely [3, 13]. These results underscore in line with previous findings [14, 17], that restrictive communication policies seem recommendable to protect research participants and the public from questionable clinical actions and costs while safeguarding observational research aims.

4.2 Relevance from a research perspective

We conducted a cohort study without any intention to intervene. However, it was also imperative to respect our participants' health and autonomy in making health-related decisions [23, 24]. This was deemed of particular importance for wb-MRI findings because of their potential to detect asymptomatic disease at an early, potentially treatable stage [1]. The disclosure policy of wb-MRI findings in SHIP was carefully designed, but no population-based points of reference were available at that point in time [4, 8]. As a result, about 10% of all MRI findings of perceived clinical relevance in SHIP were disclosed with a recommendation for further clinical work-up to almost one third of all participants [8]. As

illustrated by our results, this effectively turned our observational cohort study into a large-scale, non-randomized intervention, converting many study participants into patients with altered outpatient care over a prolonged period. In this way, our results correspond to those of a Phase I Trial [25]. This reduces the validity of inferences about the natural course of health services utilization in the general population. Rather, we observed the course of outpatient care under the precondition of a population-based health screening.

An increase in outpatient costs occurred not only in participants with disclosed MRI findings, but also in MRI non-participants and MRI participants without disclosed findings. This increase was comparably smaller in size and likely reflects effects of disclosures from study findings other than MRI such as laboratory results, the potential impact of which has been documented [14]. Thus, carefully weighted disclosure policies for research findings are needed for all study examinations, not just MRI. However, the authors do not recommend completely withholding all research findings. No disclosure at all may prove unethical in the rare cases that research examinations uncover severe and actionable clinical conditions.

4.3 Relevance from a clinical perspective

Participants and patients may conceivably be interested in obtaining personal health information out of a desire for reassurance about health concerns or out of simple curiosity [26, 27]. However, participants and patients tend to overestimate the clinical relevance of findings, which is critical in the context of IFs. On the one hand, similarly to other studies, we observed a large number of abnormalities on wb-MRI, the majority of them being related to tumours of an unknown nature [2]. On the other hand, another study analysing biopsy results within the SHIP cohort found that despite the large numbers of tumor-related

findings, few additional malignancies were detected [14]. This is likely related to the low pre-test probability of finding severe, previously undiagnosed clinical conditions in a general population sample.

Other analyses of SHIP data found that participants experienced increased psychosocial distress after the disclosure of IFs [27], yet no effects on quality of life were found 2-3 years after the wb-MRI examination [17]. This does not rule out the possibility of benefits in individual cases. Participants may also have profited from detected category III findings requiring immediate referral such as acute brain infarction or bone fracture [8]. However, less than 1% of all findings of potential clinical relevance belonged to this category.

Other studies support our critical view on the potential benefits of disclosed MRI findings. Evidence from randomized controlled trials support screening for only a few conditions, and none involve MRI as a screening tool [28]. Moreover, the detection of a malignancy does not guarantee any clinical benefit, and may even lead to harm by overtreatment [3, 13]. This is relevant for health policy makers estimating the costs and potential benefits of wb-MRI screening.

4.4 Strengths and limitations

The longitudinal cohort design, the large sample size and the availability of a control group for MRI participants are considerable strengths. The potential impact of selection bias is reduced by the low proportion of missing claims data, and the wide coverage of our participants by claims data. Yet, the small subgroup of privately insured participants is not represented. Lack of data due to participants not having visited a doctor cannot be distinguished from missing data from corrupt linkage. Due to the low percentage of such

cases, the expected impact on our results is low. The cost of performing the wb-MRI itself is not included in our analyses.

Participants in general cohort studies have been found to show fewer unhealthy behaviors, enhancing the generalizability of our findings, and screening initiatives which cater to persons seeking health screening [16] but also to other general populations health studies.

The observational nature of SHIP limits causal inferences. Yet, the markedly different course of outpatient costs among MRI participants with disclosed findings compared to those without leaves, given the observed temporal patterns, little plausible options for alternative explanations underlying the observed increases in costs in the light of only a minority having full knowledge about disclosed findings [27].

Because available claims data do not cover medication or inpatient care [29] we likely underestimate the increase in total healthcare-related expenditures. However, it is unlikely that the inclusion of inpatient costs would have substantially altered our conclusions because diagnostics for IFs rarely justify hospital admission. Given our research setting, we likely underestimate health service utilization and costs resulting from clinically indicated wb-MRI, given the patient's right to disclosure of all IFs in a clinical setting [4].

4.5 Conclusion

Whole-body MRI examination in a general population sample has sustained effects on health service utilization and produces elevated costs. The disclosure of incidental findings in this setting may bias observational data and likely induces overdiagnosis and overtesting.

Contributors: COS conceived the study, supervised the data linkage and conducted statistical analyses. HV is PI of the SHIP study, KH was in charge of the SHIP-MRI. JFC, ES provided expertise for claims data selection and analysis. SB was involved in the study analysis and provided statistical and epidemiological expertise, and approved the manuscript. COS, JFC, and ES wrote the first draft which was revised and approved by all authors.

Funding: This work was specifically funded by the German Research Foundation [SCHM 2744/1-1/1-2 to COS and CH 921/1-2 to JFC]; the SHIP study was funded by the Federal Ministry of Education and Research [03ZIK012], the Ministry of Cultural Affairs; the Social Ministry of the Federal State of Mecklenburg-Vorpommern; Siemens Healthcare, Erlangen, Germany and Bayer Healthcare.

Disclaimer: The funders did not play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors declare that they have no conflicts of interest.

Patient consent for publication: Not required.

Ethics approval: The Ethics Committee of the University Medicine of Greifswald approved the study protocol (BB 39/08, BB 106/10).

Data availability statement: Data of the SHIP studies are available and can be applied for under https://www.fvcm.med.uni-greifswald.de/dd_service/data_use_intro.php. The claims data are not publicly available due to privacy restrictions and legal reasons.

5 References

1 Orme NM, Fletcher JG, Siddiki HA, *et al.* Incidental findings in imaging research: evaluating incidence, benefit, and burden. *Arch Intern Med* 2010;**170**:1525-32.
2 Kwee RM, Kwee TC. Whole-body MRI for preventive health screening: A systematic review of the literature. *J Magn Reson Imaging* 2019;**50**:1489-503.
3 Moynihan R, Doust J, Henry D. Preventing overdiagnosis: how to stop harming the healthy. *BMJ* 2012;**344**:e3502.
4 Booth TC, Waldman AD, Wardlaw JM, *et al.* Management of incidental findings during imaging research in "healthy" volunteers: current UK practice. *Br J Radiol* 2012;**85**:11-21.
5 Berlin L. How Do You Solve a Problem Like Incidentalomas? *Appl Radiol* 2013;**42**:10-2.
6 Bamberg F, Kauczor HU, Weckbach S, *et al.* Whole-Body MR Imaging in the German National Cohort: Rationale, Design, and Technical Background. *Radiology* 2015;**277**:206-20.
7 Volzke H, Alte D, Schmidt CO, *et al.* Cohort profile: the Study of Health in Pomerania. *Int J Epidemiol* 2011;**40**:294-307.
8 Hegenscheid K, Seipel R, Schmidt CO, *et al.* Potentially relevant incidental findings on research whole-body MRI in the general adult population: frequencies and management. *Eur Radiol* 2013;**23**:816-26.
9 Tarnoki DL, Tarnoki AD, Richter A, *et al.* Clinical value of whole-body magnetic resonance imaging in health screening of general adult population. *Radiology and oncology* 2015;**49**:10-6.
10 O'Sullivan JW, Muntinga T, Grigg S, *et al.* Prevalence and outcomes of incidental imaging findings: umbrella review. *BMJ* 2018;**361**:k2387.
11 Gibson LM, Paul L, Chappell FM, *et al.* Potentially serious incidental findings on brain and body magnetic resonance imaging of apparently asymptomatic adults: systematic review and meta-analysis. *BMJ* 2018;**363**:k4577.
12 Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst* 2010;**102**:605-13.
13 Brodersen J, Kramer BS, Macdonald H, *et al.* Focusing on overdiagnosis as a driver of too much medicine. *BMJ* 2018;**362**:k3494.
14 Richter A, Sierocinski E, Singer S, *et al.* The effects of incidental findings from whole-body MRI on the frequency of biopsies and detected malignancies or benign conditions in a general population cohort study. *Eur J Epidemiol* 2020.
15 Evans A, Tolonen H, Hense HW, *et al.* Trends in coronary risk factors in the WHO MONICA project. *IntJEpidemiol* 2001;**30 Suppl 1**:S35-S40.
16 Schmidt CO, Alte D, Volzke H, *et al.* Partial misspecification of survey design features sufficed to severely bias estimates of health-related outcomes. *J Clin Epidemiol* 2010;**64**:416-23.
17 Schmidt CO, Sierocinski E, Hegenscheid K, *et al.* Impact of whole-body MRI in a general population study. *Eur J Epidemiol* 2016;**31**:31-9.
18 Hegenscheid K, Kuhn JP, Volzke H, *et al.* Whole-body magnetic resonance imaging of healthy volunteers: pilot study results from the population-based SHIP study. *Rofo* 2009;**181**:748-59.
19 Belotti F, Deb F, Manning WG, *et al.* twopm: Two-part models. *The Stata Journal* 2015;**15**:3-20.

20 Guo S, W. FM. *Propensity Score Analysis: Statistical Methods and Applications*. Los Angeles: SAGE 2009.
21 Nichols A. Erratum and discussion of propensity score reweighting. *STATA Journal* 2008;**8**:532-9.
22 Busija L, Pausenberger E, Haines TP*, et al.* Adult measures of general health and health-related quality of life: Medical Outcomes Study Short Form 36-Item (SF-36) and Short Form 12-Item (SF-12) Health Surveys, Nottingham Health Profile (NHP), Sickness Impact Profile (SIP), Medical Outcomes Study Short Form 6D (SF-6D), Health Utilities Index Mark 3 (HUI3), Quality of Well-Being Scale (QWB), and Assessment of Quality of Life (AQoL). *Arthritis Care Res (Hoboken) 2011 Nov;63 Suppl 11:S383-412 doi: 101002/acr20541* 2011.
23 DeFrank JT, Barclay C, Sheridan S*, et al.* The psychological harms of screening: the evidence we have versus the evidence we need. *J Gen Intern Med* 2015;**30**:242-8.
24 Viberg J, Hansson MG, Langenskiold S*, et al.* Incidental findings: the time is not yet ripe for a policy for biobanks. *Eur J Hum Genet* 2014;**22**:437-41.
25 Pinato DJ, Stavraka C, Tanner M*, et al.* Clinical, ethical and financial implications of incidental imaging findings: experience from a phase I trial in healthy elderly volunteers. *PLoS ONE* 2012;**7**:e49814.
26 Townsend A, Cox SM. Accessing health services through the back door: a qualitative interview study investigating reasons why people participate in health research in Canada. *BMC medical ethics* 2013;**14**.
27 Schmidt CO, Hegenscheid K, Erdmann P*, et al.* Psychosocial consequences and severity of disclosed incidental findings from whole-body MRI in a general population study. *Eur Radiol* 2013;**23**:1343-51.
28 Saquib N, Saquib J, Ioannidis JP. Does screening for disease save lives in asymptomatic adults? Systematic review of meta-analyses and randomized trials. *Int J Epidemiol* 2015;**44**:264-77.
29 Swart E, Ihle P, Gothe H. *Routinedaten im Gesundheitswesen: Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven*. Bern: Hogrefe 2014.

Figure 1

Study flow chart

Figure 2

Course of total costs

0 is the reference quarter in which the MRI examination took place. The left column displays the descriptive course for the studied outcomes comparing MRI participants with disclosed findings, MRI participants without disclosed findings, and MRI non-participants. In the right column, average treatment effects are displayed. Propensity score reweighting was applied to balance the distribution of relevant covariates between MRI participants and non-participants. Estimates were derived from two-part models. N=5019.

Left column:

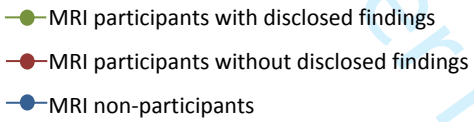


Figure 3

Course of costs by service type

0 is the reference quarter in which the MRI examination took place. The left column displays the descriptive course for the studied outcomes comparing MRI participants with disclosed findings, MRI participants without disclosed findings, and MRI non-participants. In the right column, average treatment effects are displayed. Propensity score reweighting was applied to balance the distribution of relevant covariates between MRI participants and non-participants. Estimates were derived from two-part models. N=5019.

Left column:

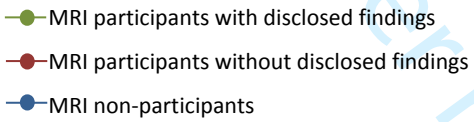


Table 1. Sample characteristics of MRI participants and MRI non-participants

	MRI participants N=2969	MRI non-participants N=2050
	N (%)	N (%)
Female	1554 (52.3)	1113 (54.3)
Low educational attainment (<10 yrs.)	634 (21.4)	743 (36.2)
Married	2004 (67.5)	1248 (60.9)
Employed	1615 (54.4)	858 (41.9)
Smoker	662 (22.3)	553 (27.0)
Coded malignant neoplasms *	108 (3.6)	95 (4.6)
	Mean (SD)	Mean (SD)
Age	53.4 (13.9)	56.3 (15.9)
SF-12 physical health component summary score	47.9 (8.2)	46.5 (9.4)
SF-12 mental health component summary score	53.0 (8.3)	51.9 (8.9)
	Mean (SD) Median (IQR)	Mean (SD) Median (IQR)
Total ambulatory care costs[a] (€)	325 (481) 190 (416)	310 (440) 191 (407)
Costs primary care, prevention[a] (€)	125 (118) 130 (188)	129 (124) 141 (190)
Costs specialized care[a] (€)	150 (373) 44 (156)	135 (324) 21 (149)
Costs imaging[a] (€)	42 (142) 0 (0)	38 (143) 0 (0)
Costs laboratory[a] (€)	18 (34) 5 (16)	17 (30) 6 (16)

SD: standard deviation. SMD: Standardized mean difference. IQR: inter-quartile range

Subjects without an exclusion criterion for MRI participation were considered eligible.

[a] Costs refer to the quarter prior to the whole body examination.

Item missingness: age, sex: 0%; educational status: 0·2%; marital status: 0·2%; employment status: 0·3%; smoking: 0·4%; SF-12: 0·3%; cost data: 7·4%· Presented are unimputed variables with the exception of cost data with a 0 imputation, as described in methods.

Table 2

Estimated total additional costs of MRI participants compared to MRI Non-participants after examination per participant in €

	Year 1 after SHIP MRI examination		Year 2 after SHIP MRI examination		Total Year 1+2 after SHIP MRI examination	
	Mean (€)	95% CI (€)	Mean (€)	95% CI (€)	Mean (€)	95% CI (€)
Total ambulatory costs	130	37; 223	164	78; 251	295	134; 456
Costs primary care, prevention	8	-16; 33	39	14; 64	48	3; 93
Costs specialized care	65	-2; 132	86	26; 146	151	40; 262
Costs Imaging	53	30; 76	46	23; 70	100	61; 139
Costs Laboratory	6	0; 13	6	-1; 12	12	1; 24

Estimates were derived from two-part models with weights to compute average treatment effects during the years 1,2, and 1+2 respectively.

Year 1: Comprises quarters 1-4 after the SHIP examination. Year 2 comprises quarters 5-8 after the SHIP examinations. Year 1+2 quarters 1-8

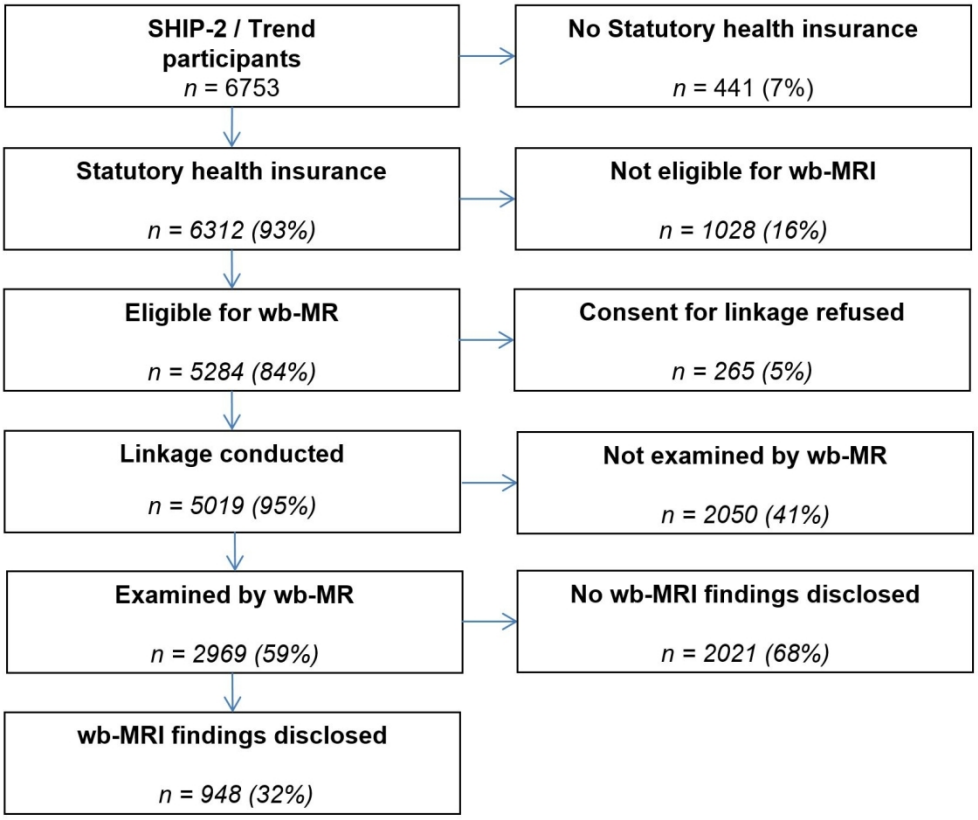
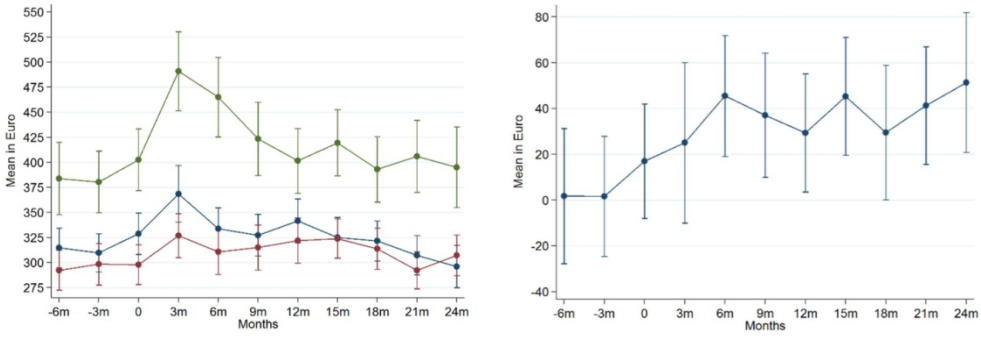


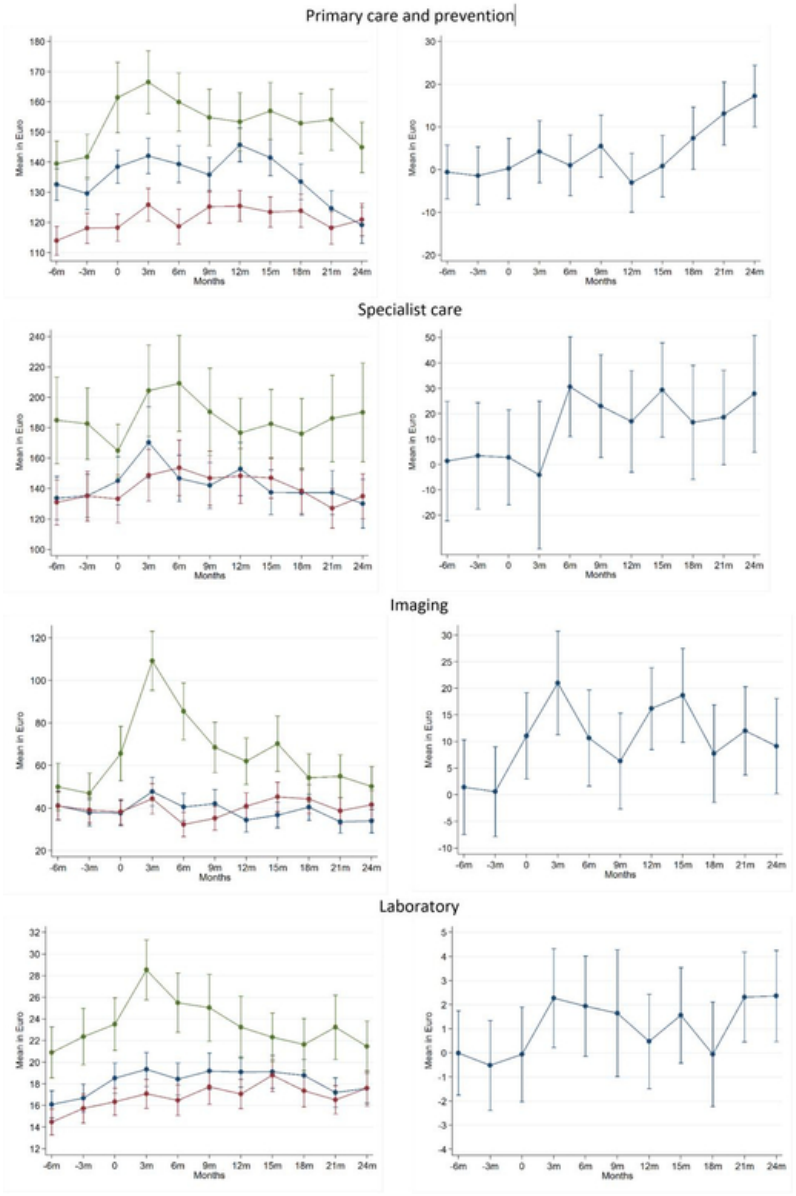
Figure 1
Study flow chart

196x165mm (216 x 216 DPI)



Course of total costs. 0 is the reference quarter in which the MRI examination took place. The left column displays the descriptive course for the studied outcomes comparing MRI participants with disclosed findings, MRI participants without disclosed findings, and MRI non-participants. In the right column, average treatment effects are displayed. Propensity score reweighting was applied to balance the distribution of relevant covariates between MRI participants and non-participants. Estimates were derived from two-part models. N=5019. Left column: green: MRI participants with disclosed findings; red: MRI participants without disclosed findings; blue: MRI non-participants.
Left column: green: MRI participants with disclosed findings; red: MRI participants without disclosed findings; blue: MRI non-participants

122x40mm (300 x 300 DPI)



Course of costs by service type. 0 is the reference quarter in which the MRI examination took place. The left column displays the descriptive course for the studied outcomes comparing MRI participants with disclosed findings, MRI participants without disclosed findings, and MRI non-participants. In the right column, average treatment effects are displayed. Propensity score reweighting was applied to balance the distribution of relevant covariates between MRI participants and non-participants. Estimates were derived from two-part models. N=5019. Left column: green: MRI participants with disclosed findings; red: MRI participants without disclosed findings; blue: MRI non-participants.

44x66mm (300 x 300 DPI)

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***
Applied to
A population-based cohort study investigating the association of incidental whole-body magnetic resonance imaging findings with outpatient costs, tumours, and mortality

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	3-4
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	6
Objectives	3	State specific objectives, including any prespecified hypotheses	6
Methods			
Study design	4	Present key elements of study design early in the paper	7
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	7-10
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	7-8
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	n. a.
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7-10
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7-10
Bias	9	Describe any efforts to address potential sources of bias	10-11
Study size	10	Explain how the study size was arrived at	7 Fig. 1
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	10-11
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	10-11
		(*b*) Describe any methods used to examine subgroups and interactions	
		(*c*) Explain how missing data were addressed	
		(*d*) If applicable, explain how loss to follow-up was addressed	
		(*e*) Describe any sensitivity analyses	
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	7 Fig. 1
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	12 Tab. 1

			(b) Indicate number of participants with missing data for each variable of interest (c) Summarise follow-up time (eg, average and total amount)	
Outcome data		15*	Report numbers of outcome events or summary measures over time	13, Fig. 2
Main results	16		(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included (*b*) Report category boundaries when continuous variables were categorized (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Fig. 2 Tab. 2,3
Other analyses	17		Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11
Discussion				
Key results	18		Summarise key results with reference to study objectives	14
Limitations	19		Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	16-17
Interpretation	20		Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	16-17
Generalisability	21		Discuss the generalisability (external validity) of the study results	14,15
Other information				
Funding	22		Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	18

BMJ Open

Effects of whole-body magnetic resonance imaging on outpatient health service costs in a general-population cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-056572.R1
Article Type:	Original research
Date Submitted by the Author:	21-Nov-2021
Complete List of Authors:	Schmidt, Carsten; University Medicine Greifswald, ICM - SHIP/KEF Sierocinski, Elizabeth; University Medicine Greifswald Baumeister, Sebastian; LMU Hegenscheid, Katrin; University Medicine Greifswald Völzke, Henry; German Center for Diabetes Research, Greifswald; University Medicine Greifswald Chenot, J. -F.; University Medicine Greifswald, Institute for Community Medicine
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Health services research
Keywords:	Magnetic resonance imaging < RADIOLOGY & IMAGING, EPIDEMIOLOGY, PUBLIC HEALTH

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Effects of whole-body magnetic resonance imaging on outpatient health service costs in a general-population cohort study

Schmidt CO (PhD) [1], Sierocinski E (MD) [1,2], Baumeister SE (PhD) [3], Hegenscheid K (MD) [4,5], Völzke H (MD) [1,6], Chenot JF (MD, MPH) [2]

[1] Institute for Community Medicine Department SHIP-KEF, University Medicine of Greifswald; Walther Rathenau Str. 48, 17475 Greifswald, Germany (carsten.schmidt@uni-greifswald.de)

[2] Institute for Community Medicine, Department of Family Medicine, Fleischmannstr. 42, 17475 Greifswald, Germany (jchenot@uni-greifswald.de, elizabeth.sierocinski1@uni-greifswald.de)

[3] Institute of Health Services Research in Dentistry, University of Muenster, 48149 Muenster, Germany (sebastian.baumeister@uni-muenster.de)

[4] Unfallkrankenhaus Berlin, Warener Str. 7, 12683 Berlin, Germany

[5] Department of Diagnostic Radiology and Neuroradiology, University Medicine of Greifswald, 17475 Greifswald, Germany (katrin.hegenscheid@uni-greifswald.de)

[6] DZHK - German Centre for Cardiovascular Research, Partner Site Greifswald, Germany; DZD - German Centre for Diabetes Research, Site Greifswald, Germany (voelzke@uni-greifswald.de)

Running title: MRI findings and health service utilization in a general-population setting

Correspondence:

Carsten Oliver Schmidt

Study of Health in Pomerania/KEF

University Medicine of Greifswald, Institute for Community Medicine

Walther Rathenau Str. 48, D-17475 Greifswald

carsten.schmidt@uni-greifswald.de

Tel +49[0]3834/86-7713, Fax +49[0]3834/86-6684

Abstract

Objective: Whole-body magnetic resonance imaging (wb-MRI) is increasingly used in research and screening but little is known about the effects of incidental findings (IFs) on health service utilization and costs. Such effects are particularly critical in an observational study. Our principal research question was therefore how participation in a wb-MRI examination with its resemblance to a population-based health screening is associated with outpatient service costs.

Design: Prospective cohort study.

Setting: General population Mecklenburg-Vorpommern, Germany.

Participants. Analyses included 5019 participants of the Study of Health in Pomerania (SHIP) with statutory health insurance data. 2969 took part in a wb-MRI examination in addition to a clinical examination program that was administered to all participants. MRI non-participants served as a quasi-experimental control group with propensity score weighting to account for baseline differences.

Primary and secondary outcome measures: Outpatient costs (total health care usage, primary care, specialist care, laboratory tests, imaging) during 24 months after the examination were retrieved from claims data. Two-part models were used to compute treatment effects.

Results: In total, 1366 potentially relevant IFs were disclosed to 948 MRI participants (32% of all participants); most concerned masses and lesions (769 participants, 81%). Costs for outpatient care during the two-year observation period amounted to an average of €2547

(95%CI: €2424-€2671) for MRI non-participants and to €2839 (95%CI: €2741-€2936) for MRI participants, indicating an increase of €295 (95%CI: €134-€456) per participant which corresponds to 11.6% (95%CI: 5.2%; 17.9%). The cost increase was sustained rather than being a short-term spike. Imaging and specialist care related costs were the main contributors to the increase in costs.

Conclusions: Communicated findings from population-based wb-MRI substantially impacted health service utilization and costs. This introduced bias into the natural course of health care utilization and should be taken care for in any longitudinal analyses.

Strengths and limitations

1. This is the first study to report disclosed longitudinal effects of disclosed incidental whole body MRI findings on costs for outpatient services in a general population setting.
2. The longitudinal cohort design, the large sample size and the availability of a group with and without MRI participation are considerable strengths.
3. The potential impact of selection bias is reduced by the wide coverage of our participants by claims data. Claims data provide an appropriate approximation of outpatient costs as they are collected for reimbursement purposes.
4. Limitations concern the scope of claims data as they do not cover medication or inpatient care. Furthermore, compared to a clinical screening scenario only selected findings were communicated. We therefore likely underestimate the total healthcare-related expenditures in a research and even more so in a clinical scenario. Furthermore, the small subgroup of privately insured participants is not represented.
5. Non randomized allocation to MRI participation in the observational cohort design reduces the comparability of both groups and statistical measures have been taken to balance potential confounders.

1 Introduction

Screening with whole-body magnetic resonance imaging (wb-MRI) may detect asymptomatic disease at an early stage thus improve treatment outcome [1], but might also cause unnecessary psychosocial distress, medical interventions, and costs due to irrelevant findings [2, 3]. Wb-MRI has evolved into a key examination tool in state-of-the-art population research [4-7]. It produces a large number of incidental findings (IFs), a large proportion of which represent masses and lesions [2, 8-11]. Although a minority may benefit from IFs in a general population approach [1], there is uncertainty around the clinical significance of a large proportion of these IFs. There is a high risk of false-positives, overtesting and overdiagnosis [12, 13]. Only few malignancies were newly detected despite a 42% increase in biopsies after participation in SHIP [14] and no detected positive effects on quality of life [15]. In contrast, there is clear evidence of short term adverse consequences such as psychological distress due to disclosed findings [16].

To the best of our knowledge, there is no prospective evidence on the effects of wb-MRI IFs in a general population setting on ambulatory health service utilization and associated costs. Such data is needed to guide the appropriate handling of wb-MRI screening results in research and clinical practice.

Our principal research question was therefore how participation in a wb-MRI with resemblance to a population-based health screening is associated with outpatient service utilization and costs. MRI non-participants served as a quasi-experimental control group with propensity score weighting to account for baseline differences. We hypothesized that participation in the MRI examination would lead to increased costs in ambulatory

healthcare. Furthermore we assume that there are differential effects on different types of ambulatory services.

2 Methods

2.1 Study design and sample

SHIP is a population-based project consisting of two independent cohort studies, SHIP and SHIP-Trend. Participants were sampled from the counties of North and East Western Pomerania and the cities of Greifswald and Stralsund in Germany [7]. Participants were between 20 and 79 years of age at the date of sampling and had the target region listed as their primary place of residence. The aim in both cohorts was to recruit a representative general population sample. Clinical status was of no relevance for inclusion. Participants received no payment beyond reimbursement for travel costs.

A two-stage sampling scheme was adopted from the German World Health Organization's (WHO) MONICA Project for the first cohort [17]. Out of 6265 eligible individuals of the first cohort, 4308 (2192 women) participated at baseline (response 68.8%) [18]. Baseline examinations were performed from 1997-2001 (SHIP-0). Follow-up examinations took place between 2002 and 2006 (SHIP-1, N=3300) and between 2008 and 2012 (SHIP-2, N=2333).

A second cohort (SHIP-Trend) was established in 2008 from a stratified sample of 10000, drawn from the central population registry. Examinations took place until 2012. Of the net sample of 8826, after exclusion of deceased and relocated participants, 4420 (2275 women) participated (response 50.1%).

Of the 6753 participants in SHIP-2 and SHIP-Trend, 6312 had statutory health insurance (93%). Subjects ineligible for MRI participation for reasons such as claustrophobia, metal implants, or pregnancy were excluded from analyses [15] because treatment effects could

not be meaningfully computed for this group (n=1028). Of the resulting 5284 participants, 265 (5%) refused linkage with statutory health insurance. Our analyses included the remaining 5019 participants (Figure 1). The follow-up period for claims data was two years after SHIP participation.

The Ethics Committee of the University Medicine of Greifswald approved the study protocol (BB 39/08, BB 106/10).

[Figure 1 here]

2.2 MRI Examination

All SHIP participants were invited to take part in the MRI examination and received relevant written educational material. During a medical interview before the examination, a radiologist described the handling of IFs and conditions for disclosure [8], methods descriptions in 2.2 and 2.3 have been taken from previous SHIP publications [14, 15].

All wb-MRI were acquired on a 1.5-Tesla system (Magnetom Avanto; Siemens Medical Solutions, Erlangen, Germany). The wb-MRI protocol was identical for all participants and included a plain whole-body MRI and detailed imaging of the head, neck, chest, abdomen, pelvis, and spine. Men had the option of contrast-enhanced cardiac MRI and MR angiography, and women had the option of cardiac MRI and contrast-enhanced MR mammography. The complete imaging protocols have been described previously [8, 19].

Findings and anatomical variants were documented in a standardised reading protocol. The radiologists reading the scans had no access to the participants' clinical information. Scan

reading was performed using a digital picture archiving and communication system (IMPACS ES 5.2, AGFA Healthcare, Mortsel, Belgium). First-line reading was performed by two independent radiology residents. A third reader, a senior radiologist, resolved disagreements.

The MRI examination was entirely financed by SHIP study funding and did not contribute to billing costs represented in the claims data.

2.3 Disclosure of incidental findings

A standardised protocol regulated the handling of wb-MRI IFs. Findings were classified into three categories: Category I comprised normal or common findings in asymptomatic individuals (e.g., anatomical variants, old brain infarcts, disc herniation, sinusitits). Category II findings were abnormalities of potential clinical relevance. Category III findings required immediate referral. Category II findings were disclosed in writing via post after approval by an interdisciplinary advisory board (e.g. breast lesion ≥ BI-RADS 3, adrenal tumour > 10 mm, lung nodule > 4 mm, chronic pancreatitis, internal carotid artery stenosis). Category III findings were disclosed immediately to the participant (e.g. acute brain infarction, intracranial haemorrhage, lobar pneumonia, bone fracture). A detailed description of this protocol has been provided elsewhere [8]. Our analyses are based on Category II and III findings.

Health-related findings from other examinations such as blood testing, blood pressure measurements, somatometry, ultrasound, and cardiovascular examinations were also disclosed to SHIP participants [7].

2.4 Claims data

Claims data from the regional Association of Statutory Health Insurance Physicians included billing codes for medical and technical services (e.g. imaging) and costs for outpatient care. Claims data do not include medication costs or services like emergency department attendances. Germany has a mixed billing system with capitation and fee-for-service models for specific services. We computed costs per quarter (billing period) for the two quarters prior to the examination, the quarter during which the examination took place, and the eight quarters following the examination. We distinguished between (1) total outpatient costs, (2) primary care and prevention, (3) specialist care, (4) laboratory work, and (5) imaging. All costs related to pregnancy and births were excluded, as pregnancy was an exclusion criterion for MRI participation.

2.5 Statistical analyses

We used two-part models to analyze cost variables with their zero inflated distribution [20]. Each two-part model comprised a multivariable logistic regression for any healthcare service provision and a generalized linear model with a log-link and a gamma distribution for the height of costs. Models used the predictors time, MRI participation (yes vs. no), and the interaction term to model average treatment effects. Propensity score reweighting was applied to balance the distribution of relevant covariates between MRI participants and non-participants [21, 22].

Logistic regression models were used to estimate propensity scores with the following predictors: costs for total healthcare, primary care and prevention, specialist care, laboratory work and imaging in the two quarters prior to the SHIP-2/SHIP-Trend

examination, as well as age, sex, level of education (<10 years vs. ≥10 years), marital status, current employment, smoking status, and quality of life using the SF-12 mental health component summary score and physical health component summary score [23]. Resulting weights ranged from 1.0 to 13.5 (mean: 2.5; standard deviation 0.8). Standardized mean differences (SMD) of all baseline variables after weighting were close to zero (min: -.007, 10%pct: -.004, median: 0.90%pct: .003, max: .005). Balance was also checked based on distributional properties of all baseline variables and all statistical interaction effects (SMD distribution: min: -.045, 10%pct: -.019, median: -.002, 90%pct: .009, max: .045).

Item missingness is reported in Table 1. Cost data was unavailable for 7.4% of all quarters. For 4.7% (N=238) no doctoral visit was coded at all during the observation period. Given the plausibility of a subgroup not presenting to ambulatory care regularly, we set all types of costs in these quarters to 0€ as a coded visit is a prerequisite for costs. Nevertheless, we cannot be certain that data for some participants who did present to the doctor was missing. We therefore used multiple imputation by chained equations (MICE) as a sensitivity analysis to impute missing information on costs and coded neoplasms. These results did not lead to different conclusions and are thus not presented.

Two-sided tests were applied throughout. Analyses were conducted in Stata 14 (Stata Corp., College Station, TX). Figures were generated using Microsoft PowerPoint.

2.6 Patient and public involvement

Patients, study participants and the public were not directly involved in the design, conduct and reporting of this study.

3 Results

3.1 Sample characteristics and incidental findings

MRI non-participants were on average three years older, had a lower level of education, were more often unemployed, smokers, and less often married compared to MRI participants (Table 1). Total ambulatory costs were similar in both groups and quality of life scores were slightly lower among non-MRI participants.

In total, 1366 IFs of potential clinical relevance were disclosed to 948 participants (32%). Of these, 769 participants (81%) received a finding related to masses and lesions, corresponding to 26% of all MRI participants. A more detailed overview of the structure and type of wb-MRI findings and the affected organs has been provided previously [8].

[Table 1 here]

3.2 Descriptive course of outpatient costs

The course of outpatient costs is displayed stratified for MRI non-participants, MRI participants with disclosed findings, and MRI-participants who did not receive findings (Figure 2). The latter two groups were distinguished because a marked increase in costs after the SHIP examination was only expected among those with a disclosed finding. Increases occurred in all studied groups but peaks were highest in the quarter after the SHIP examination among MRI participants with disclosed findings (Figure 2 and 3). The relative increase was largest for imaging-related costs, which more than doubled (Figure 3). While decreasing after an initial peak, total costs among MRI participants with disclosed findings did not return to the initial level but remained higher (Figure 2 and 3).

[Figure 2 here]

[Figure 3 here]

3.3 MRI participation and outpatient costs

Propensity score reweighting revealed increased excess costs among MRI participants compared to non-participants that persisted over time (Figure 2 and 3).

Total weighted costs for outpatient care during the two-year observation period amounted to an average of €2547 (95%CI: €2424-€2671) for MRI non-participants and to €2839 (95%CI: €2741-€2936) for MRI participants, yielding an average treatment effect of €295 (95%CI: €134-€456), which corresponds to an increase of +11.6% (5.2%; 17.9%) per participant (Table 2). Additional costs were higher in the second post-examination year compared to the first year. The largest contribution to excess costs resulted from specialist care, followed by imaging.The smallest effects were related to laboratory costs.

[Table 2 here]

4 Discussion

4.1 Main findings

A single wb-MRI examination sufficed to increase long-term overall outpatient costs over two years in a general population sample. Effects were larger for certain services such as specialist care and clinical imaging. From a research perspective, our results illustrate how disclosed IFs turned an observational study into an intervention. This limits the generalizability of research findings on outpatient costs and health service utilization to the underlying general population. From a clinical perspective, overtesting and overdiagnosis are likely [3, 13]. These results underscore in line with previous findings [14, 15], that restrictive communication policies seem recommendable to protect research participants and the public from questionable clinical actions and costs while safeguarding observational research aims.

4.2 Relevance from a research perspective

We conducted a cohort study without any intention to intervene. However, it was also imperative to respect our participants' health and autonomy in making health-related decisions [24, 25]. This was deemed of particular importance for wb-MRI findings because of their potential to detect asymptomatic disease at an early, potentially treatable stage [1]. The disclosure policy of wb-MRI findings in SHIP was carefully designed, but no population-based points of reference were available at that point in time [4, 8]. As a result, about 10% of all MRI findings of perceived clinical relevance in SHIP were disclosed with a recommendation for further clinical work-up to almost one third of all participants [8]. As

illustrated by our results, this effectively turned our observational cohort study into a large-scale, non-randomized intervention, converting many study participants into patients with altered outpatient care over a prolonged period. In this way, our results correspond to those of a Phase I Trial [26]. This reduces the validity of inferences about the natural course of health services utilization in the general population. Rather, we observed the course of outpatient care under the precondition of a population-based health screening.

An increase in outpatient costs occurred not only in participants with disclosed MRI findings, but also in MRI non-participants and MRI participants without disclosed findings. This increase was comparably smaller in size and likely reflects effects of disclosures from study findings other than MRI such as laboratory results, the potential impact of which has been documented [14]. Thus, carefully weighted disclosure policies for research findings are needed for all study examinations, not just MRI. However, the authors do not recommend completely withholding all research findings. No disclosure at all may prove unethical in the rare cases that research examinations uncover severe and actionable clinical conditions.

4.3 Relevance from a clinical perspective

Participants and patients may conceivably be interested in obtaining personal health information out of a desire for reassurance about health concerns or out of simple curiosity [16, 27]. However, practitioners [28-30], participants and patients alike tend to overestimate the clinical relevance of findings, which is critical in the context of IFs. On the one hand, similarly to other studies, we observed a large number of abnormalities on wb-MRI, the majority of them being related to tumours of an unknown nature [2]. On the other hand, another study analysing biopsy results within the SHIP cohort found that despite the large

numbers of tumor-related findings, few additional malignancies were detected [14]. This is likely related to the low pre-test probability of finding severe, previously undiagnosed clinical conditions in a general population sample.

Other analyses of SHIP data found that participants experienced increased psychosocial distress after the disclosure of IFs [16], yet no effects on quality of life were found 2-3 years after the wb-MRI examination [15]. This does not rule out the possibility of benefits in individual cases. Participants may also have profited from detected category III findings requiring immediate referral such as acute brain infarction or bone fracture [8]. However, less than 1% of all findings of potential clinical relevance belonged to this category but it is not known if improved outcomes resulted from their communication. Furthermore, Category II findings amounted to approximately only 10% of all findings [8]. In SHIP most Category I findings resulted from a highly detailed structured reading protocol that also included anatomical variants. While it is unlikely that these would have been of much interest in a screening, Category I findings also comprised clinical findings without any best practice recommendation to communicate them such as disc herniation. Nevertheless, in a health screening setting it seems likely that such findings would have been communicated to patients, thus leading to higher subsequent consultations and costs with an even elevated risk of overtesting, overdiagnosis, and overtreatment.

Other studies support our critical view on the potential benefits of disclosed MRI findings. Evidence from randomized controlled trials support screening for only a few conditions, and none involve MRI as a screening tool [31]. Moreover, the detection of a malignancy does not guarantee any clinical benefit, and may even lead to harm by overtreatment [3, 13]. This is

relevant for health policy makers estimating the costs and potential benefits of wb-MRI screening. The costs for the work up of IFs generated by our cohort study were covered by statutory health insurance. It is an issue of debate whether or not such findings should be subject to financing by a public health system. For example, the Royal College for General Practitioners argues that work up of IFs from private screening should not be the responsibility of primary care physicians [32].

4.4 Strengths and limitations

The longitudinal cohort design, the large sample size and the availability of a control group for MRI participants are considerable strengths. The potential impact of selection bias is reduced by the low proportion of missing claims data, and the wide coverage of our participants by claims data. Yet, the small subgroup of privately insured participants (7% in this sample) is not represented. This subgroup may be even more prone to cascades of subsequent health care and elevated costs due to the more favorable reimbursement of diagnostic and therapeutic measures.

Lack of data due to participants not having visited a doctor cannot be distinguished from missing data from corrupt linkage. Due to the low percentage of such cases, the expected impact on our results is low. The cost of performing the wb-MRI itself is not included in our analyses.

Participants in general cohort studies have been found to show fewer unhealthy behaviors, enhancing the generalizability of our findings, and screening initiatives which cater to persons seeking health screening [18] but also to other general populations health studies.

The observational nature of SHIP limits causal inferences. Yet, the markedly different course of outpatient costs among MRI participants with disclosed findings compared to those without leaves, given the observed temporal patterns, little plausible options for alternative explanations underlying the observed increases in costs in the light of only a minority having full knowledge about disclosed findings [16]. While being closely related to health service costs, the frequency of consultations cannot be validly inferred from German claims data. Therefore, despite being of interest, this aspect was not addressed.

We targeted the impact of communicated MRI findings and computed related treatment effects. However, other clinical findings were disclosed as well such as those from laboratory examinations. These were not examined independently in our study. Increases in health care costs in participants without disclosed MRI IFs (Figures 2,3) indicate that elevated costs were potentially related to the disclosure of findings from other SHIP examinations. However, effect sizes are much smaller compared to those in participants with disclosed MRI IFs.

Because available claims data do not cover medication costs or inpatient care as well as emergency department attendances [33] we likely underestimate the increase in total healthcare-related expenditures. However, it is unlikely that the inclusion of inpatient costs would have substantially altered our conclusions because diagnostics for IFs rarely justify hospital admission. Given our research setting, we likely underestimate health service

utilization and costs resulting from clinically indicated wb-MRI, given the patient's right to disclosure of all IFs in a clinical setting [4].

4.5 Conclusion

Whole-body MRI examination in a general population sample has sustained effects on health service utilization, leading to elevated costs that may well persist beyond the duration of the two year observation period after the wb-MRI examination. The disclosure of incidental findings in this cohort study may bias the longitudinal study of health related outcomes and likely induced overdiagnosis and overtesting.

Contributors: COS conceived the study, supervised the data linkage and conducted statistical analyses. HV is PI of the SHIP study, KH was in charge of the SHIP-MRI. JFC, ES provided expertise for claims data selection and analysis. SB was involved in the study analysis and provided statistical and epidemiological expertise, and approved the manuscript. COS, JFC, and ES wrote the first draft which was revised and approved by all authors.

Funding: This work was specifically funded by the German Research Foundation [SCHM 2744/1-1/1-2 to COS and CH 921/1-2 to JFC]; the SHIP study was funded by the Federal Ministry of Education and Research [03ZIK012], the Ministry of Cultural Affairs; the Social Ministry of the Federal State of Mecklenburg-Vorpommern; Siemens Healthcare, Erlangen, Germany and Bayer Healthcare.

Disclaimer: The funders did not play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors declare that they have no conflicts of interest.

Patient consent for publication: Not required.

Ethics approval: The Ethics Committee of the University Medicine of Greifswald approved the study protocol (BB 39/08, BB 106/10).

Data availability statement: Data of the SHIP studies are available and can be applied for under https://www.fvcm.med.uni-greifswald.de/dd_service/data_use_intro.php. The claims data are not publicly available due to privacy restrictions and legal reasons.

5 References

1 Orme NM, Fletcher JG, Siddiki HA, *et al.* Incidental findings in imaging research: evaluating incidence, benefit, and burden. *Arch Intern Med* 2010;**170**:1525-32.
2 Kwee RM, Kwee TC. Whole-body MRI for preventive health screening: A systematic review of the literature. *J Magn Reson Imaging* 2019;**50**:1489-503.
3 Moynihan R, Doust J, Henry D. Preventing overdiagnosis: how to stop harming the healthy. *BMJ* 2012;**344**:e3502.
4 Booth TC, Waldman AD, Wardlaw JM, *et al.* Management of incidental findings during imaging research in "healthy" volunteers: current UK practice. *Br J Radiol* 2012;**85**:11-21.
5 Berlin L. How Do You Solve a Problem Like Incidentalomas? *Appl Radiol* 2013;**42**:10-2.
6 Bamberg F, Kauczor HU, Weckbach S, *et al.* Whole-Body MR Imaging in the German National Cohort: Rationale, Design, and Technical Background. *Radiology* 2015;**277**:206-20.
7 Volzke H, Alte D, Schmidt CO, *et al.* Cohort profile: the Study of Health in Pomerania. *Int J Epidemiol* 2011;**40**:294-307.
8 Hegenscheid K, Seipel R, Schmidt CO, *et al.* Potentially relevant incidental findings on research whole-body MRI in the general adult population: frequencies and management. *Eur Radiol* 2013;**23**:816-26.
9 Tarnoki DL, Tarnoki AD, Richter A, *et al.* Clinical value of whole-body magnetic resonance imaging in health screening of general adult population. *Radiology and oncology* 2015;**49**:10-6.
10 O'Sullivan JW, Muntinga T, Grigg S, *et al.* Prevalence and outcomes of incidental imaging findings: umbrella review. *BMJ* 2018;**361**:k2387.
11 Gibson LM, Paul L, Chappell FM, *et al.* Potentially serious incidental findings on brain and body magnetic resonance imaging of apparently asymptomatic adults: systematic review and meta-analysis. *BMJ* 2018;**363**:k4577.
12 Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst* 2010;**102**:605-13.
13 Brodersen J, Kramer BS, Macdonald H, *et al.* Focusing on overdiagnosis as a driver of too much medicine. *BMJ* 2018;**362**:k3494.
14 Richter A, Sierocinski E, Singer S, *et al.* The effects of incidental findings from whole-body MRI on the frequency of biopsies and detected malignancies or benign conditions in a general population cohort study. *Eur J Epidemiol* 2020.
15 Schmidt CO, Sierocinski E, Hegenscheid K, *et al.* Impact of whole-body MRI in a general population study. *Eur J Epidemiol* 2016;**31**:31-9.
16 Schmidt CO, Hegenscheid K, Erdmann P, *et al.* Psychosocial consequences and severity of disclosed incidental findings from whole-body MRI in a general population study. *Eur Radiol* 2013;**23**:1343-51.
17 Evans A, Tolonen H, Hense HW, *et al.* Trends in coronary risk factors in the WHO MONICA project. *IntJEpidemiol* 2001;**30 Suppl 1**:S35-S40.
18 Schmidt CO, Alte D, Volzke H, *et al.* Partial misspecification of survey design features sufficed to severely bias estimates of health-related outcomes. *J Clin Epidemiol* 2010;**64**:416-23.

19 Hegenscheid K, Kuhn JP, Volzke H, *et al.* Whole-body magnetic resonance imaging of healthy volunteers: pilot study results from the population-based SHIP study. *Rofo* 2009;**181**:748-59.
20 Belotti F, Deb F, Manning WG, *et al.* twopm: Two-part models. *The Stata Journal* 2015;**15**:3-20.
21 Guo S, W. FM. *Propensity Score Analysis: Statistical Methods and Applications*. Los Angeles: SAGE 2009.
22 Nichols A. Erratum and discussion of propensity score reweighting. *STATA Journal* 2008;**8**:532-9.
23 Busija L, Pausenberger E, Haines TP, *et al.* Adult measures of general health and health-related quality of life: Medical Outcomes Study Short Form 36-Item (SF-36) and Short Form 12-Item (SF-12) Health Surveys, Nottingham Health Profile (NHP), Sickness Impact Profile (SIP), Medical Outcomes Study Short Form 6D (SF-6D), Health Utilities Index Mark 3 (HUI3), Quality of Well-Being Scale (QWB), and Assessment of Quality of Life (AQoL). *Arthritis Care Res (Hoboken) 2011 Nov;63 Suppl 11:S383-412 doi: 101002/acr20541* 2011.
24 DeFrank JT, Barclay C, Sheridan S, *et al.* The psychological harms of screening: the evidence we have versus the evidence we need. *J Gen Intern Med* 2015;**30**:242-8.
25 Viberg J, Hansson MG, Langenskiold S, *et al.* Incidental findings: the time is not yet ripe for a policy for biobanks. *Eur J Hum Genet* 2014;**22**:437-41.
26 Pinato DJ, Stavraka C, Tanner M, *et al.* Clinical, ethical and financial implications of incidental imaging findings: experience from a phase I trial in healthy elderly volunteers. *PLoS ONE* 2012;**7**:e49814.
27 Townsend A, Cox SM. Accessing health services through the back door: a qualitative interview study investigating reasons why people participate in health research in Canada. *BMC medical ethics* 2013;**14**.
28 Sajid IM, Parkunan A, Frost K. Unintended consequences: quantifying the benefits, iatrogenic harms and downstream cascade costs of musculoskeletal MRI in UK primary care. *BMJ Open Qual* 2021;**10**.
29 Morgan DJ, Pineles L, Owczarzak J, *et al.* Accuracy of Practitioner Estimates of Probability of Diagnosis Before and After Testing. *JAMA Intern Med* 2021;**181**:747-55.
30 Whiting PF, Davenport C, Jameson C, *et al.* How well do health professionals interpret diagnostic information? A systematic review. *BMJ Open* 2015;**5**:e008155.
31 Saquib N, Saquib J, Ioannidis JP. Does screening for disease save lives in asymptomatic adults? Systematic review of meta-analyses and randomized trials. *Int J Epidemiol* 2015;**44**:264-77.
32 Mahase E. Private screening: GPs shouldn't have to deal with results, says RCGP. *BMJ* 2019;**366**:l5707.
33 Swart E, Ihle P, Gothe H. *Routinedaten im Gesundheitswesen: Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven*. Bern: Hogrefe 2014.

Figure 1

Study flow chart

Figure 2

Course of total costs

0 is the reference quarter in which the MRI examination took place. The left column displays the descriptive course for the outcomes comparing MRI participants with disclosed findings, MRI participants without disclosed findings, and MRI non-participants. Average treatment effects are displayed in the right column. Propensity score reweighting was applied to balance the distribution of relevant covariates between MRI participants and non-participants. Estimates were derived from two-part models. N=5019.

Left column:

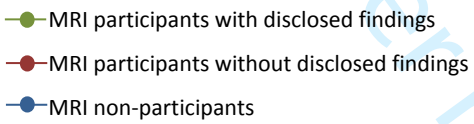


Figure 3

Course of costs by service type

0 is the reference quarter in which the MRI examination took place. The left column displays the descriptive course for the studied outcomes comparing MRI participants with disclosed findings, MRI participants without disclosed findings, and MRI non-participants. Average treatment effects are displayed in the right column. Propensity score reweighting was applied to balance the distribution of relevant covariates between MRI participants and non-participants. Estimates were derived from two-part models. N=5019.

Left column:

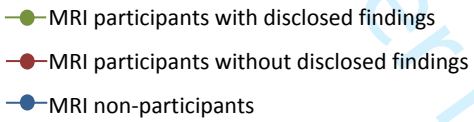


Table 1. Sample characteristics of MRI participants and MRI non-participants

	MRI participants N=2969	MRI non-participants N=2050
	N (%)	N (%)
Female	1554 (52.3)	1113 (54.3)
Low educational attainment (<10 yrs.)	634 (21.4)	743 (36.2)
Married	2004 (67.5)	1248 (60.9)
Employed	1615 (54.4)	858 (41.9)
Smoker	662 (22.3)	553 (27.0)
Coded malignant neoplasms *	108 (3.6)	95 (4.6)
	Mean (SD)	Mean (SD)
Age	53.4 (13.9)	56.3 (15.9)
SF-12 physical health component summary score	47.9 (8.2)	46.5 (9.4)
SF-12 mental health component summary score	53.0 (8.3)	51.9 (8.9)
	Mean (SD) Median (IQR)	Mean (SD) Median (IQR)
Total ambulatory care costs[a] (€)	325 (481) 190 (416)	310 (440) 191 (407)
Costs primary care, prevention[a] (€)	125 (118) 130 (188)	129 (124) 141 (190)
Costs specialized care[a] (€)	150 (373) 44 (156)	135 (324) 21 (149)
Costs imaging[a] (€)	42 (142) 0 (0)	38 (143) 0 (0)
Costs laboratory[a] (€)	18 (34) 5 (16)	17 (30) 6 (16)

SD: standard deviation. SMD: Standardized mean difference. IQR: inter-quartile range

Subjects without an exclusion criterion for MRI participation were considered eligible.

[a] Costs refer to the quarter prior to the whole body examination.

Item missingness: age, sex: 0%; educational status: 0·2%; marital status: 0·2%; employment status: 0·3%; smoking: 0·4%; SF-12: 0·3%; cost data: 7·4%· Presented are unimputed variables with the exception of cost data with a 0 imputation, as described in methods.

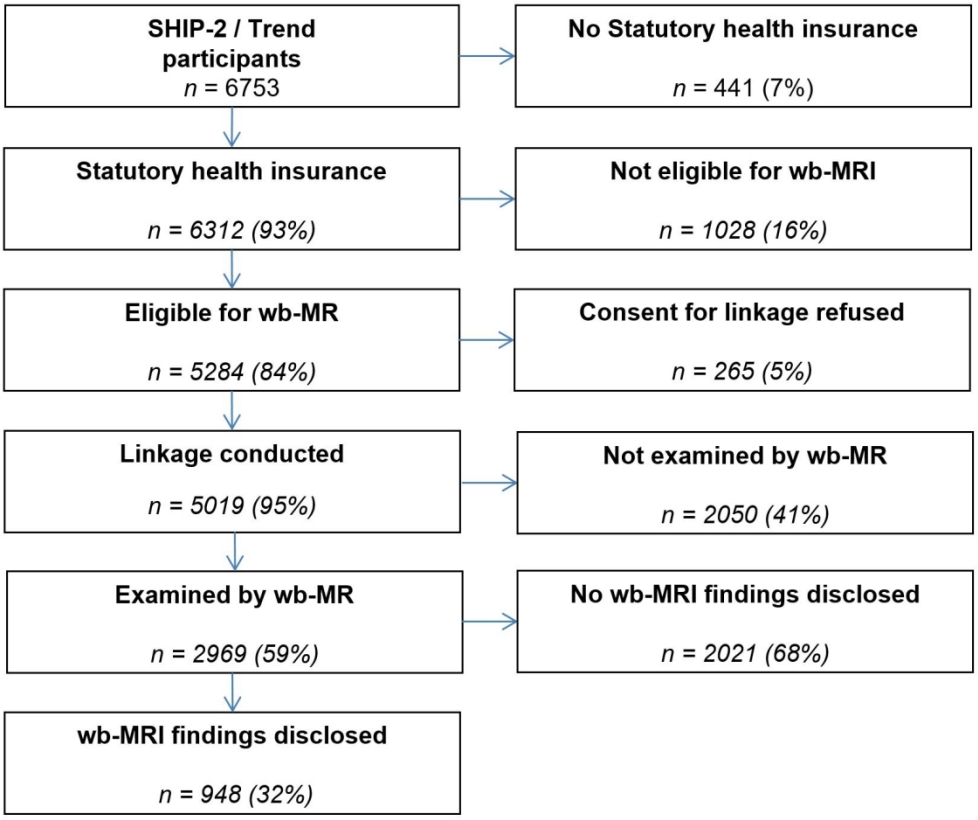
Table 2

Estimated total additional costs of MRI participants compared to MRI Non-participants after examination per participant in €

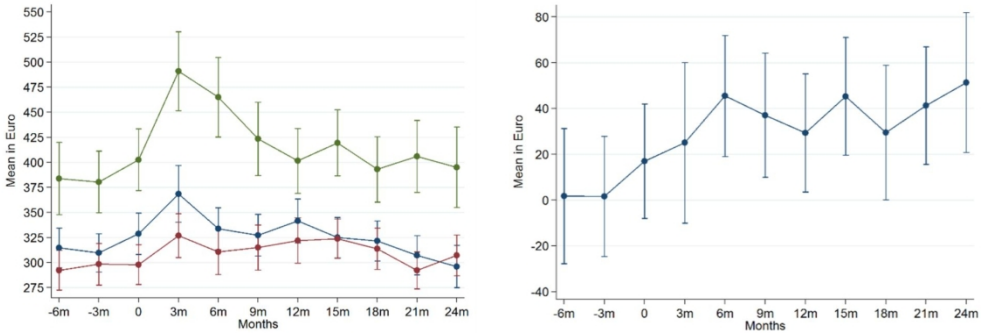
	Year 1 after SHIP MRI examination		Year 2 after SHIP MRI examination		Total Year 1+2 after SHIP MRI examination	
	Mean (€)	95% CI (€)	Mean (€)	95% CI (€)	Mean (€)	95% CI (€)
Total ambulatory costs	130	37; 223	164	78; 251	295	134; 456
Costs primary care, prevention	8	-16; 33	39	14; 64	48	3; 93
Costs specialized care	65	-2; 132	86	26; 146	151	40; 262
Costs Imaging	53	30; 76	46	23; 70	100	61; 139
Costs Laboratory	6	0; 13	6	-1; 12	12	1; 24

Estimates were derived from two-part models with weights to compute average treatment effects during the years 1,2, and 1+2 respectively.

Year 1: Comprises quarters 1-4 after the SHIP examination. Year 2 comprises quarters 5-8 after the SHIP examinations. Year 1+2 quarters 1-8

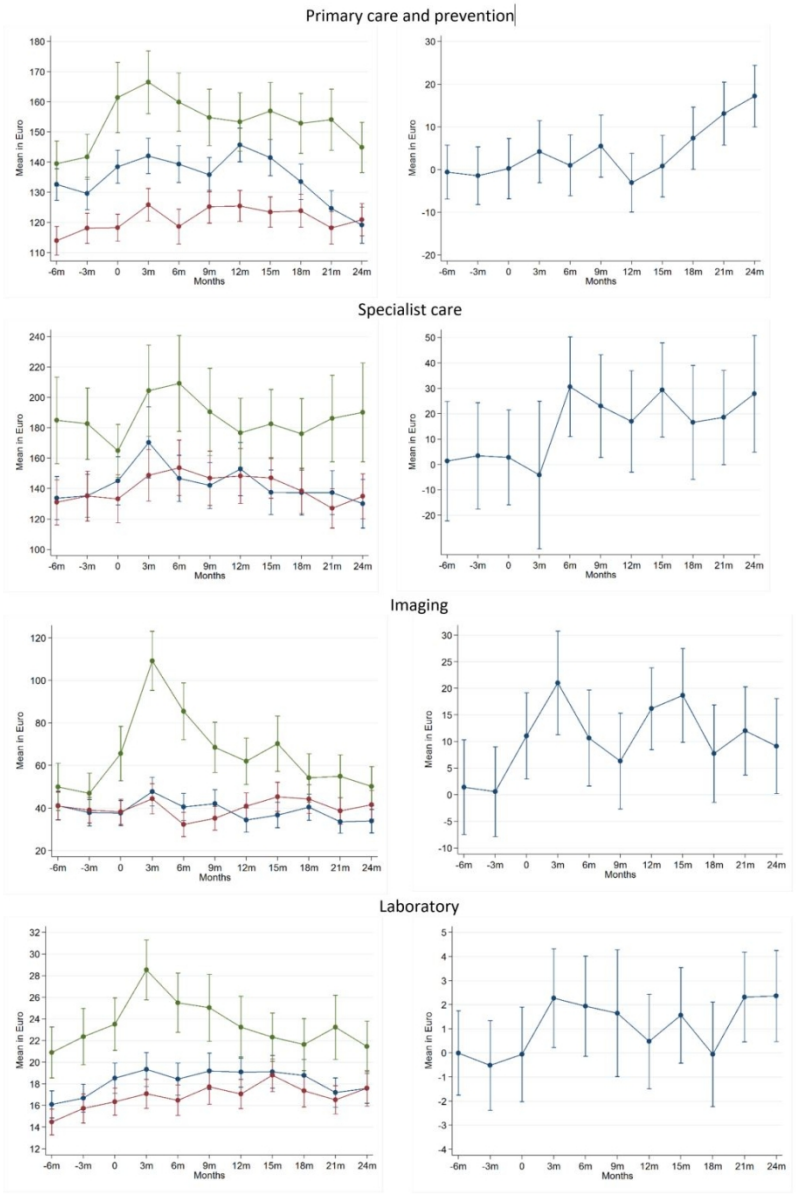


196x165mm (216 x 216 DPI)



Course of total costs. 0 is the reference quarter in which the MRI examination took place. The left column displays the descriptive course for the studied outcomes comparing MRI participants with disclosed findings, MRI participants without disclosed findings, and MRI non-participants. In the right column, average treatment effects are displayed. Propensity score reweighting was applied to balance the distribution of relevant covariates between MRI participants and non-participants. Estimates were derived from two-part models. N=5019. Left column: green: MRI participants with disclosed findings; red: MRI participants without disclosed findings; blue: MRI non-participants.
Left column: green: MRI participants with disclosed findings; red: MRI participants without disclosed findings; blue: MRI non-participants

122x40mm (400 x 400 DPI)



Course of costs by service type. 0 is the reference quarter in which the MRI examination took place. The left column displays the descriptive course for the studied outcomes comparing MRI participants with disclosed findings, MRI participants without disclosed findings, and MRI non-participants. In the right column, average treatment effects are displayed. Propensity score reweighting was applied to balance the distribution of relevant covariates between MRI participants and non-participants. Estimates were derived from two-part models. N=5019. Left column: green: MRI participants with disclosed findings; red: MRI participants without disclosed findings; blue: MRI non-participants.

44x66mm (600 x 600 DPI)

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***
Applied to
A population-based cohort study investigating the association of incidental whole-body magnetic resonance imaging findings with outpatient costs, tumours, and mortality

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	3-4
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	6
Objectives	3	State specific objectives, including any prespecified hypotheses	6
Methods			
Study design	4	Present key elements of study design early in the paper	7
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	7-10
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	7-8
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	n. a.
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7-10
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7-10
Bias	9	Describe any efforts to address potential sources of bias	10-11
Study size	10	Explain how the study size was arrived at	7 Fig. 1
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	10-11
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	10-11
		(*b*) Describe any methods used to examine subgroups and interactions	
		(*c*) Explain how missing data were addressed	
		(*d*) If applicable, explain how loss to follow-up was addressed	
		(*e*) Describe any sensitivity analyses	
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	7 Fig. 1
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	12 Tab. 1

			(b) Indicate number of participants with missing data for each variable of interest (c) Summarise follow-up time (eg, average and total amount)	
Outcome data		15*	Report numbers of outcome events or summary measures over time	13, Fig. 2
Main results	16		(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included (*b*) Report category boundaries when continuous variables were categorized (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Fig. 2 Tab. 2,3
Other analyses	17		Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11
Discussion				
Key results	18		Summarise key results with reference to study objectives	14
Limitations	19		Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	16-17
Interpretation	20		Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	16-17
Generalisability	21		Discuss the generalisability (external validity) of the study results	14,15
Other information				
Funding	22		Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	18

BMJ Open

Effects of whole-body magnetic resonance imaging on outpatient health service costs: a general-population prospective cohort study in Mecklenburg-Vorpommern, Germany

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-056572.R2
Article Type:	Original research
Date Submitted by the Author:	09-Dec-2021
Complete List of Authors:	Schmidt, Carsten; University Medicine Greifswald, ICM - SHIP/KEF Sierocinski, Elizabeth; University Medicine Greifswald Baumeister, Sebastian; LMU Hegenscheid, Katrin; University Medicine Greifswald Völzke, Henry; German Center for Diabetes Research, Greifswald; University Medicine Greifswald Chenot, J. -F.; University Medicine Greifswald, Institute for Community Medicine
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Health services research
Keywords:	Magnetic resonance imaging < RADIOLOGY & IMAGING, EPIDEMIOLOGY, PUBLIC HEALTH

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Effects of whole-body magnetic resonance imaging on outpatient health service costs: a general-population prospective cohort study in Mecklenburg-Vorpommern, Germany

Schmidt CO (PhD) [1], Sierocinski E (MD) [1,2], Baumeister SE (PhD) [3], Hegenscheid K (MD) [4,5], Völzke H (MD) [1,6], Chenot JF (MD, MPH) [2]

[1] Institute for Community Medicine Department SHIP-KEF, University Medicine of Greifswald; Walther Rathenau Str. 48, 17475 Greifswald, Germany (carsten.schmidt@uni-greifswald.de)

[2] Institute for Community Medicine, Department of Family Medicine, Fleischmannstr. 42, 17475 Greifswald, Germany (jchenot@uni-greifswald.de, elizabeth.sierocinski1@uni-greifswald.de)

[3] Institute of Health Services Research in Dentistry, University of Muenster, 48149 Muenster, Germany (sebastian.baumeister@uni-muenster.de)

[4] Unfallkrankenhaus Berlin, Warener Str. 7, 12683 Berlin, Germany

[5] Department of Diagnostic Radiology and Neuroradiology, University Medicine of Greifswald, 17475 Greifswald, Germany (katrin.hegenscheid@uni-greifswald.de)

[6] DZHK - German Centre for Cardiovascular Research, Partner Site Greifswald, Germany; DZD - German Centre for Diabetes Research, Site Greifswald, Germany (voelzke@uni-greifswald.de)

Running title: MRI findings and health service utilization in a general-population setting

Correspondence:

Carsten Oliver Schmidt

Study of Health in Pomerania/KEF

University Medicine of Greifswald, Institute for Community Medicine

Walther Rathenau Str. 48, D-17475 Greifswald

carsten.schmidt@uni-greifswald.de

Tel +49[0]3834/86-7713, Fax +49[0]3834/86-6684

Abstract

Objective: Whole-body magnetic resonance imaging (wb-MRI) is increasingly used in research and screening but little is known about the effects of incidental findings (IFs) on health service utilization and costs. Such effects are particularly critical in an observational study. Our principal research question was therefore how participation in a wb-MRI examination with its resemblance to a population-based health screening is associated with outpatient service costs.

Design: Prospective cohort study.

Setting: General population Mecklenburg-Vorpommern, Germany.

Participants. Analyses included 5019 participants of the Study of Health in Pomerania (SHIP) with statutory health insurance data. 2969 took part in a wb-MRI examination in addition to a clinical examination program that was administered to all participants. MRI non-participants served as a quasi-experimental control group with propensity score weighting to account for baseline differences.

Primary and secondary outcome measures: Outpatient costs (total health care usage, primary care, specialist care, laboratory tests, imaging) during 24 months after the examination were retrieved from claims data. Two-part models were used to compute treatment effects.

Results: In total, 1366 potentially relevant IFs were disclosed to 948 MRI participants (32% of all participants); most concerned masses and lesions (769 participants, 81%). Costs for outpatient care during the two-year observation period amounted to an average of €2547

(95%CI: €2424-€2671) for MRI non-participants and to €2839 (95%CI: €2741-€2936) for MRI participants, indicating an increase of €295 (95%CI: €134-€456) per participant which corresponds to 11.6% (95%CI: 5.2%; 17.9%). The cost increase was sustained rather than being a short-term spike. Imaging and specialist care related costs were the main contributors to the increase in costs.

Conclusions: Communicated findings from population-based wb-MRI substantially impacted health service utilization and costs. This introduced bias into the natural course of health care utilization and should be taken care for in any longitudinal analyses.

Strengths and limitations of this study

- This is the first study to report disclosed longitudinal effects of disclosed incidental whole body MRI findings on costs for outpatient services in a general population setting; the longitudinal cohort design, large sample size and the availability of groups with and without MRI participation are considerable strengths.
- Claims data provide an appropriate approximation of outpatient costs as they are collected for reimbursement purposes, and selection bias is reduced by the wide coverage of participant claims data.
- Limitations concern the scope of claims data as they do not cover medication or inpatient care, and we are likely to have underestimated the total healthcare-related expenditures compared with a clinical scenario because only selected findings were disclosed to participants.
- The small subgroup of privately insured participants is not represented.
- Non-randomized allocation to MRI participation reduces the comparability of participants and non-participants.

1 Introduction

Screening with whole-body magnetic resonance imaging (wb-MRI) may detect asymptomatic disease at an early stage thus improve treatment outcome [1], but might also cause unnecessary psychosocial distress, medical interventions, and costs due to irrelevant findings [2, 3]. Wb-MRI has evolved into a key examination tool in state-of-the-art population research [4-7]. It produces a large number of incidental findings (IFs), a large proportion of which represent masses and lesions [2, 8-11]. Although a minority may benefit from IFs in a general population approach [1], there is uncertainty around the clinical significance of a large proportion of these IFs. There is a high risk of false-positives, overtesting and overdiagnosis [12, 13]. Only few malignancies were newly detected despite a 42% increase in biopsies after participation in SHIP [14] and no detected positive effects on quality of life [15]. In contrast, there is clear evidence of short term adverse consequences such as psychological distress due to disclosed findings [16].

To the best of our knowledge, there is no prospective evidence on the effects of wb-MRI IFs in a general population setting on ambulatory health service utilization and associated costs. Such data is needed to guide the appropriate handling of wb-MRI screening results in research and clinical practice.

Our principal research question was therefore how participation in a wb-MRI with resemblance to a population-based health screening is associated with outpatient service utilization and costs. MRI non-participants served as a quasi-experimental control group with propensity score weighting to account for baseline differences. We hypothesized that participation in the MRI examination would lead to increased costs in ambulatory

healthcare. Furthermore, we assume that there are differential effects on different types of ambulatory services.

2 Methods

2.1 Study design and sample

SHIP is a population-based project consisting of two independent prospective cohort studies, SHIP and SHIP-Trend. Participants were sampled from the counties of North and East Western Pomerania and the cities of Greifswald and Stralsund in Germany [7]. Participants were between 20 and 79 years of age at the date of sampling and had the target region listed as their primary place of residence. The aim in both cohorts was to recruit a representative general population sample. Clinical status was of no relevance for inclusion. Participants received no payment beyond reimbursement for travel costs.

A two-stage sampling scheme was adopted from the German World Health Organization's (WHO) MONICA Project for the first cohort [17]. Out of 6265 eligible individuals of the first cohort, 4308 (2192 women) participated at baseline (response 68.8%) [18]. Baseline examinations were performed from 1997-2001 (SHIP-0). Follow-up examinations took place between 2002 and 2006 (SHIP-1, N=3300) and between 2008 and 2012 (SHIP-2, N=2333).

A second cohort (SHIP-Trend) was established in 2008 from a stratified sample of 10000, drawn from the central population registry. Examinations took place until 2012. Of the net sample of 8826, after exclusion of deceased and relocated participants, 4420 (2275 women) participated (response 50.1%).

Of the 6753 participants in SHIP-2 and SHIP-Trend, 6312 had statutory health insurance (93%). Subjects ineligible for MRI participation for reasons such as claustrophobia, metal implants, or pregnancy were excluded from analyses [15] because treatment effects could

not be meaningfully computed for this group (n=1028). Of the resulting 5284 participants, 265 (5%) refused linkage with statutory health insurance. Our analyses included the remaining 5019 participants (Figure 1). The follow-up period for claims data was two years after SHIP participation.

The Ethics Committee of the University Medicine of Greifswald approved the study protocol (BB 39/08, BB 106/10). All participants provided written informed consent before being recruited into SHIP.

[Figure 1 here]

2.2 MRI Examination

All SHIP participants were invited to take part in the MRI examination and received relevant written educational material. During a medical interview before the examination, a radiologist described the handling of IFs and conditions for disclosure [8], methods descriptions in 2.2 and 2.3 have been taken from previous SHIP publications [14, 15].

All wb-MRI were acquired on a 1.5-Tesla system (Magnetom Avanto; Siemens Medical Solutions, Erlangen, Germany). The wb-MRI protocol was identical for all participants and included a plain whole-body MRI and detailed imaging of the head, neck, chest, abdomen, pelvis, and spine. Men had the option of contrast-enhanced cardiac MRI and MR angiography, and women had the option of cardiac MRI and contrast-enhanced MR mammography. The complete imaging protocols have been described previously [8, 19].

Findings and anatomical variants were documented in a standardised reading protocol. The radiologists reading the scans had no access to the participants' clinical information. Scan reading was performed using a digital picture archiving and communication system (IMPACS ES 5.2, AGFA Healthcare, Mortsel, Belgium). First-line reading was performed by two independent radiology residents. A third reader, a senior radiologist, resolved disagreements.

The MRI examination was entirely financed by SHIP study funding and did not contribute to billing costs represented in the claims data.

2.3 Disclosure of incidental findings

A standardised protocol regulated the handling of wb-MRI IFs. Findings were classified into three categories: Category I comprised normal or common findings in asymptomatic individuals (e.g., anatomical variants, old brain infarcts, disc herniation, sinusitits). Category II findings were abnormalities of potential clinical relevance. Category III findings required immediate referral. Category II findings were disclosed in writing via post after approval by an interdisciplinary advisory board (e.g. breast lesion ≥ BI-RADS 3, adrenal tumour > 10 mm, lung nodule > 4 mm, chronic pancreatitis, internal carotid artery stenosis). Category III findings were disclosed immediately to the participant (e.g. acute brain infarction, intracranial haemorrhage, lobar pneumonia, bone fracture). A detailed description of this protocol has been provided elsewhere [8]. Our analyses are based on Category II and III findings.

Health-related findings from other examinations such as blood testing, blood pressure measurements, somatometry, ultrasound, and cardiovascular examinations were also disclosed to SHIP participants [7].

2.4 Claims data

Claims data from the regional Association of Statutory Health Insurance Physicians included billing codes for medical and technical services (e.g. imaging) and costs for outpatient care. Claims data do not include medication costs or services like emergency department attendances. Germany has a mixed billing system with capitation and fee-for-service models for specific services. We computed costs per quarter (billing period) for the two quarters prior to the examination, the quarter during which the examination took place, and the eight quarters following the examination. We distinguished between (1) total outpatient costs, (2) primary care and prevention, (3) specialist care, (4) laboratory work, and (5) imaging. All costs related to pregnancy and births were excluded, as pregnancy was an exclusion criterion for MRI participation.

2.5 Statistical analyses

We used two-part models to analyze cost variables with their zero inflated distribution [20]. Each two-part model comprised a multivariable logistic regression for any healthcare service provision and a generalized linear model with a log-link and a gamma distribution for the height of costs. Models used the predictors time, MRI participation (yes vs. no), and the interaction term to model average treatment effects. Propensity score reweighting was

applied to balance the distribution of relevant covariates between MRI participants and non-participants [21, 22].

Logistic regression models were used to estimate propensity scores with the following predictors: costs for total healthcare, primary care and prevention, specialist care, laboratory work and imaging in the two quarters prior to the SHIP-2/SHIP-Trend examination, as well as age, sex, level of education (<10 years vs. ≥10 years), marital status, current employment, smoking status, and quality of life using the SF-12 mental health component summary score and physical health component summary score [23]. Resulting weights ranged from 1.0 to 13.5 (mean: 2.5; standard deviation 0.8). Standardized mean differences (SMD) of all baseline variables after weighting were close to zero (min: -.007, 10%pct: -.004, median: 0.90%pct: .003, max: .005). Balance was also checked based on distributional properties of all baseline variables and all statistical interaction effects (SMD distribution: min: -.045, 10%pct: -.019, median: -.002, 90%pct: .009, max: .045).

Item missingness is reported in Table 1. Cost data was unavailable for 7.4% of all quarters. For 4.7% (N=238) no doctoral visit was coded at all during the observation period. Given the plausibility of a subgroup not presenting to ambulatory care regularly, we set all types of costs in these quarters to €0 as a coded visit is a prerequisite for costs. Nevertheless, we cannot be certain that data for some participants who did present to the doctor was missing. We therefore used multiple imputation by chained equations (MICE) as a sensitivity analysis to impute missing information on costs and coded neoplasms. These results did not lead to different conclusions and are thus not presented.

Two-sided tests were applied throughout. Analyses were conducted in Stata 14 (Stata Corp., College Station, TX). Figures were generated using Microsoft PowerPoint.

2.6 Patient and public involvement

Patients, study participants and the public were not directly involved in the design, conduct and reporting of this study.

3 Results

3.1 Sample characteristics and incidental findings

MRI non-participants were on average three years older, had a lower level of education, were more often unemployed, smokers, and less often married compared to MRI participants (Table 1). Total ambulatory costs were similar in both groups and quality of life scores were slightly lower among non-MRI participants.

In total, 1366 IFs of potential clinical relevance were disclosed to 948 participants (32%). Of these, 769 participants (81%) received a finding related to masses and lesions, corresponding to 26% of all MRI participants. A more detailed overview of the structure and type of wb-MRI findings and the affected organs has been provided previously [8].

[Table 1 here]

3.2 Descriptive course of outpatient costs

The course of outpatient costs is displayed stratified for MRI non-participants, MRI participants with disclosed findings, and MRI-participants who did not receive findings (Figure 2). The latter two groups were distinguished because a marked increase in costs after the SHIP examination was only expected among those with a disclosed finding. Increases occurred in all studied groups but peaks were highest in the quarter after the SHIP examination among MRI participants with disclosed findings (Figure 2 and 3). The relative increase was largest for imaging-related costs, which more than doubled (Figure 3). While decreasing after an initial peak, total costs among MRI participants with disclosed findings did not return to the initial level but remained higher (Figure 2 and 3).

[Figure 2 here]

[Figure 3 here]

3.3 MRI participation and outpatient costs

Propensity score reweighting revealed increased excess costs among MRI participants compared to non-participants that persisted over time (Figure 2 and 3).

Total weighted costs for outpatient care during the two-year observation period amounted to an average of €2547 (95%CI: €2424-€2671) for MRI non-participants and to €2839 (95%CI: €2741-€2936) for MRI participants, yielding an average treatment effect of €295 (95%CI: €134-€456), which corresponds to an increase of +11.6% (5.2%; 17.9%) per participant (Table 2). Additional costs were higher in the second post-examination year compared to the first year. The largest contribution to excess costs resulted from specialist care, followed by imaging.The smallest effects were related to laboratory costs.

[Table 2 here]

4 Discussion

4.1 Main findings

A single wb-MRI examination sufficed to increase long-term overall outpatient costs over two years in a general population sample. Effects were larger for certain services such as specialist care and clinical imaging. From a research perspective, our results illustrate how disclosed IFs turned an observational study into an intervention. This limits the generalizability of research findings on outpatient costs and health service utilization to the underlying general population. From a clinical perspective, overtesting and overdiagnosis are likely [3, 13]. These results underscore in line with previous findings [14, 15], that restrictive communication policies seem recommendable to protect research participants and the public from questionable clinical actions and costs while safeguarding observational research aims.

4.2 Relevance from a research perspective

We conducted a prospective cohort study without any intention to intervene. However, it was also imperative to respect our participants' health and autonomy in making health-related decisions [24, 25]. This was deemed of particular importance for wb-MRI findings because of their potential to detect asymptomatic disease at an early, potentially treatable stage [1]. The disclosure policy of wb-MRI findings in SHIP was carefully designed, but no population-based points of reference were available at that point in time [4, 8]. As a result, about 10% of all MRI findings of perceived clinical relevance in SHIP were disclosed with a recommendation for further clinical work-up to almost one third of all participants [8]. As

illustrated by our results, this effectively turned our cohort study into a large-scale, non-randomized intervention, converting many study participants into patients with altered outpatient care over a prolonged period. In this way, our results correspond to those of a Phase I Trial [26]. This reduces the validity of inferences about the natural course of health services utilization in the general population. Rather, we observed the course of outpatient care under the precondition of a population-based health screening.

An increase in outpatient costs occurred not only in participants with disclosed MRI findings, but also in MRI non-participants and MRI participants without disclosed findings. This increase was comparably smaller in size and likely reflects effects of disclosures from study findings other than MRI such as laboratory results, the potential impact of which has been documented [14]. Thus, carefully weighted disclosure policies for research findings are needed for all study examinations, not just MRI. However, the authors do not recommend completely withholding all research findings. No disclosure at all may prove unethical in the rare cases that research examinations uncover severe and actionable clinical conditions.

4.3 Relevance from a clinical perspective

Participants and patients may conceivably be interested in obtaining personal health information out of a desire for reassurance about health concerns or out of simple curiosity [16, 27]. However, practitioners [28-30], participants and patients alike tend to overestimate the clinical relevance of findings, which is critical in the context of IFs. On the one hand, similarly to other studies, we observed a large number of abnormalities on wb-MRI, the majority of them being related to tumours of an unknown nature [2]. On the other hand, another study analysing biopsy results within the SHIP cohort found that despite the large

numbers of tumor-related findings, few additional malignancies were detected [14]. This is likely related to the low pre-test probability of finding severe, previously undiagnosed clinical conditions in a general population sample.

Other analyses of SHIP data found that participants experienced increased psychosocial distress after the disclosure of IFs [16], yet no effects on quality of life were found 2-3 years after the wb-MRI examination [15]. This does not rule out the possibility of benefits in individual cases. Participants may also have profited from detected category III findings requiring immediate referral such as acute brain infarction or bone fracture [8]. However, less than 1% of all findings of potential clinical relevance belonged to this category but it is not known if improved outcomes resulted from their communication. Furthermore, Category II findings amounted to approximately only 10% of all findings [8]. In SHIP most Category I findings resulted from a highly detailed structured reading protocol that also included anatomical variants. While it is unlikely that these would have been of much interest in a screening, Category I findings also comprised clinical findings without any best practice recommendation to communicate them such as disc herniation. Nevertheless, in a health screening setting it seems likely that such findings would have been communicated to patients, thus leading to higher subsequent consultations and costs with an even elevated risk of overtesting, overdiagnosis, and overtreatment.

Other studies support our critical view on the potential benefits of disclosed MRI findings. Evidence from randomized controlled trials support screening for only a few conditions, and none involve MRI as a screening tool [31]. Moreover, the detection of a malignancy does not guarantee any clinical benefit, and may even lead to harm by overtreatment [3, 13]. This is

relevant for health policy makers estimating the costs and potential benefits of wb-MRI screening. The costs for the work up of IFs generated by our cohort study were covered by statutory health insurance. It is an issue of debate whether or not such findings should be subject to financing by a public health system. For example, the UK Royal College for General Practitioners argues that work up of IFs from private screening should not be the responsibility of primary care physicians [32].

4.4 Strengths and limitations

The longitudinal design, the large sample size and the availability of a control group for MRI participants are considerable strengths. The potential impact of selection bias is reduced by the low proportion of missing claims data, and the wide coverage of our participants by claims data. Yet, the small subgroup of privately insured participants (7% in this sample) is not represented. This subgroup may be even more prone to cascades of subsequent health care and elevated costs due to the more favorable reimbursement of diagnostic and therapeutic measures.

Lack of data due to participants not having visited a doctor cannot be distinguished from missing data from corrupt linkage. Due to the low percentage of such cases, the expected impact on our results is low. The cost of performing the wb-MRI itself is not included in our analyses.

Participants in general-population cohort studies have been found to show fewer unhealthy behaviors, enhancing the generalizability of our findings, and screening initiatives which

cater to persons seeking health screening [18] but also to other general populations health studies.

The observational nature of SHIP limits causal inferences. Yet, the markedly different course of outpatient costs among MRI participants with disclosed findings compared to those without leaves, given the observed temporal patterns, little plausible options for alternative explanations underlying the observed increases in costs in the light of only a minority having full knowledge about disclosed findings [16]. While being closely related to health service costs, the frequency of consultations cannot be validly inferred from German claims data. Therefore, despite being of interest, this aspect was not addressed.

We targeted the impact of communicated MRI findings and computed related treatment effects. However, other clinical findings were disclosed as well such as those from laboratory examinations. These were not examined independently in our study. Increases in health care costs in participants without disclosed MRI IFs (Figures 2,3) indicate that elevated costs were potentially related to the disclosure of findings from other SHIP examinations. However, effect sizes are much smaller compared to those in participants with disclosed MRI IFs.

Because available claims data do not cover medication costs or inpatient care as well as emergency department attendances [33] we likely underestimate the increase in total healthcare-related expenditures. However, it is unlikely that the inclusion of inpatient costs would have substantially altered our conclusions because diagnostics for IFs rarely justify hospital admission. Given our research setting, we likely underestimate health service utilization and costs resulting from clinically indicated wb-MRI, given the patient's right to disclosure of all IFs in a clinical setting [4].

4.5 Conclusion

Whole-body MRI examination in a general population sample has sustained effects on health service utilization, leading to elevated costs that may well persist beyond the duration of the two year observation period after the wb-MRI examination. The disclosure of incidental findings in this prospective cohort study may bias the longitudinal study of health related outcomes and likely induced overdiagnosis and overtesting.

Contributors: COS conceived the study, supervised the data linkage and conducted statistical analyses. HV is PI of the SHIP study, KH was in charge of the SHIP-MRI. JFC, ES provided expertise for claims data selection and analysis. SB was involved in the study analysis and provided statistical and epidemiological expertise, and approved the manuscript. COS, JFC, and ES wrote the first draft which was revised and approved by all authors.

Funding: This work was specifically funded by the German Research Foundation [SCHM 2744/1-1/1-2 to COS and CH 921/1-2 to JFC]; the SHIP study was funded by the Federal Ministry of Education and Research [03ZIK012], the Ministry of Cultural Affairs; the Social Ministry of the Federal State of Mecklenburg-Vorpommern; Siemens Healthcare, Erlangen, Germany and Bayer Healthcare.

Disclaimer: The funders did not play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors declare that they have no conflicts of interest.

Patient consent for publication: Not required.

Ethics approval: The Ethics Committee of the University Medicine of Greifswald approved the study protocol (BB 39/08, BB 106/10).

Data availability statement: Data of the SHIP studies are available and can be applied for under https://www.fvcm.med.uni-greifswald.de/dd_service/data_use_intro.php. The claims data are not publicly available due to privacy restrictions and legal reasons.

5 References

1 Orme NM, Fletcher JG, Siddiki HA, *et al.* Incidental findings in imaging research: evaluating incidence, benefit, and burden. *Arch Intern Med* 2010;**170**:1525-32.
2 Kwee RM, Kwee TC. Whole-body MRI for preventive health screening: A systematic review of the literature. *J Magn Reson Imaging* 2019;**50**:1489-503.
3 Moynihan R, Doust J, Henry D. Preventing overdiagnosis: how to stop harming the healthy. *BMJ* 2012;**344**:e3502.
4 Booth TC, Waldman AD, Wardlaw JM, *et al.* Management of incidental findings during imaging research in "healthy" volunteers: current UK practice. *Br J Radiol* 2012;**85**:11-21.
5 Berlin L. How Do You Solve a Problem Like Incidentalomas? *Appl Radiol* 2013;**42**:10-2.
6 Bamberg F, Kauczor HU, Weckbach S, *et al.* Whole-Body MR Imaging in the German National Cohort: Rationale, Design, and Technical Background. *Radiology* 2015;**277**:206-20.
7 Volzke H, Alte D, Schmidt CO, *et al.* Cohort profile: the Study of Health in Pomerania. *Int J Epidemiol* 2011;**40**:294-307.
8 Hegenscheid K, Seipel R, Schmidt CO, *et al.* Potentially relevant incidental findings on research whole-body MRI in the general adult population: frequencies and management. *Eur Radiol* 2013;**23**:816-26.
9 Tarnoki DL, Tarnoki AD, Richter A, *et al.* Clinical value of whole-body magnetic resonance imaging in health screening of general adult population. *Radiology and oncology* 2015;**49**:10-6.
10 O'Sullivan JW, Muntinga T, Grigg S, *et al.* Prevalence and outcomes of incidental imaging findings: umbrella review. *BMJ* 2018;**361**:k2387.
11 Gibson LM, Paul L, Chappell FM, *et al.* Potentially serious incidental findings on brain and body magnetic resonance imaging of apparently asymptomatic adults: systematic review and meta-analysis. *BMJ* 2018;**363**:k4577.
12 Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst* 2010;**102**:605-13.
13 Brodersen J, Kramer BS, Macdonald H, *et al.* Focusing on overdiagnosis as a driver of too much medicine. *BMJ* 2018;**362**:k3494.
14 Richter A, Sierocinski E, Singer S, *et al.* The effects of incidental findings from whole-body MRI on the frequency of biopsies and detected malignancies or benign conditions in a general population cohort study. *Eur J Epidemiol* 2020.
15 Schmidt CO, Sierocinski E, Hegenscheid K, *et al.* Impact of whole-body MRI in a general population study. *Eur J Epidemiol* 2016;**31**:31-9.
16 Schmidt CO, Hegenscheid K, Erdmann P, *et al.* Psychosocial consequences and severity of disclosed incidental findings from whole-body MRI in a general population study. *Eur Radiol* 2013;**23**:1343-51.
17 Evans A, Tolonen H, Hense HW, *et al.* Trends in coronary risk factors in the WHO MONICA project. *IntJEpidemiol* 2001;**30 Suppl 1**:S35-S40.
18 Schmidt CO, Alte D, Volzke H, *et al.* Partial misspecification of survey design features sufficed to severely bias estimates of health-related outcomes. *J Clin Epidemiol* 2010;**64**:416-23.

19 Hegenscheid K, Kuhn JP, Volzke H, *et al.* Whole-body magnetic resonance imaging of healthy volunteers: pilot study results from the population-based SHIP study. *Rofo* 2009;**181**:748-59.
20 Belotti F, Deb F, Manning WG, *et al.* twopm: Two-part models. *The Stata Journal* 2015;**15**:3-20.
21 Guo S, W. FM. *Propensity Score Analysis: Statistical Methods and Applications*. Los Angeles: SAGE 2009.
22 Nichols A. Erratum and discussion of propensity score reweighting. *STATA Journal* 2008;**8**:532-9.
23 Busija L, Pausenberger E, Haines TP, *et al.* Adult measures of general health and health-related quality of life: Medical Outcomes Study Short Form 36-Item (SF-36) and Short Form 12-Item (SF-12) Health Surveys, Nottingham Health Profile (NHP), Sickness Impact Profile (SIP), Medical Outcomes Study Short Form 6D (SF-6D), Health Utilities Index Mark 3 (HUI3), Quality of Well-Being Scale (QWB), and Assessment of Quality of Life (AQoL). *Arthritis Care Res (Hoboken) 2011 Nov;63 Suppl 11:S383-412 doi: 101002/acr20541* 2011.
24 DeFrank JT, Barclay C, Sheridan S, *et al.* The psychological harms of screening: the evidence we have versus the evidence we need. *J Gen Intern Med* 2015;**30**:242-8.
25 Viberg J, Hansson MG, Langenskiold S, *et al.* Incidental findings: the time is not yet ripe for a policy for biobanks. *Eur J Hum Genet* 2014;**22**:437-41.
26 Pinato DJ, Stavraka C, Tanner M, *et al.* Clinical, ethical and financial implications of incidental imaging findings: experience from a phase I trial in healthy elderly volunteers. *PLoS ONE* 2012;**7**:e49814.
27 Townsend A, Cox SM. Accessing health services through the back door: a qualitative interview study investigating reasons why people participate in health research in Canada. *BMC medical ethics* 2013;**14**.
28 Sajid IM, Parkunan A, Frost K. Unintended consequences: quantifying the benefits, iatrogenic harms and downstream cascade costs of musculoskeletal MRI in UK primary care. *BMJ Open Qual* 2021;**10**.
29 Morgan DJ, Pineles L, Owczarzak J, *et al.* Accuracy of Practitioner Estimates of Probability of Diagnosis Before and After Testing. *JAMA Intern Med* 2021;**181**:747-55.
30 Whiting PF, Davenport C, Jameson C, *et al.* How well do health professionals interpret diagnostic information? A systematic review. *BMJ Open* 2015;**5**:e008155.
31 Saquib N, Saquib J, Ioannidis JP. Does screening for disease save lives in asymptomatic adults? Systematic review of meta-analyses and randomized trials. *Int J Epidemiol* 2015;**44**:264-77.
32 Mahase E. Private screening: GPs shouldn't have to deal with results, says RCGP. *BMJ* 2019;**366**:l5707.
33 Swart E, Ihle P, Gothe H. *Routinedaten im Gesundheitswesen: Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven*. Bern: Hogrefe 2014.

Figure 1

Study flow chart

Figure 2

Course of total costs

0 is the reference quarter in which the MRI examination took place. The left column displays the descriptive course for the outcomes comparing MRI participants with disclosed findings, MRI participants without disclosed findings, and MRI non-participants. Average treatment effects are displayed in the right column. Propensity score reweighting was applied to balance the distribution of relevant covariates between MRI participants and non-participants. Estimates were derived from two-part models. N=5019.

Left column:

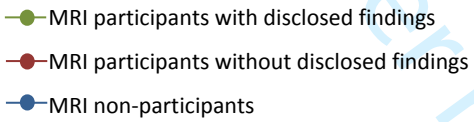


Figure 3

Course of costs by service type

0 is the reference quarter in which the MRI examination took place. The left column displays the descriptive course for the studied outcomes comparing MRI participants with disclosed findings, MRI participants without disclosed findings, and MRI non-participants. Average treatment effects are displayed in the right column. Propensity score reweighting was applied to balance the distribution of relevant covariates between MRI participants and non-participants. Estimates were derived from two-part models. N=5019.

Left column:

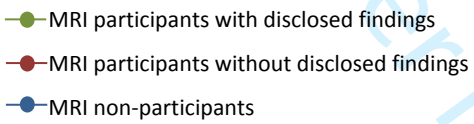


Table 1. Sample characteristics of MRI participants and MRI non-participants

	MRI participants N=2969	MRI non-participants N=2050
	N (%)	N (%)
Female	1554 (52.3)	1113 (54.3)
Low educational attainment (<10 yrs.)	634 (21.4)	743 (36.2)
Married	2004 (67.5)	1248 (60.9)
Employed	1615 (54.4)	858 (41.9)
Smoker	662 (22.3)	553 (27.0)
Coded malignant neoplasms *	108 (3.6)	95 (4.6)
	Mean (SD)	Mean (SD)
Age	53.4 (13.9)	56.3 (15.9)
SF-12 physical health component summary score	47.9 (8.2)	46.5 (9.4)
SF-12 mental health component summary score	53.0 (8.3)	51.9 (8.9)
	Mean (SD) Median (IQR)	Mean (SD) Median (IQR)
Total ambulatory care costs[a] (€)	325 (481) 190 (416)	310 (440) 191 (407)
Costs primary care, prevention[a] (€)	125 (118) 130 (188)	129 (124) 141 (190)
Costs specialized care[a] (€)	150 (373) 44 (156)	135 (324) 21 (149)
Costs imaging[a] (€)	42 (142) 0 (0)	38 (143) 0 (0)
Costs laboratory[a] (€)	18 (34) 5 (16)	17 (30) 6 (16)

SD: standard deviation. SMD: Standardized mean difference. IQR: inter-quartile range

Subjects without an exclusion criterion for MRI participation were considered eligible.

[a] Costs refer to the quarter prior to the whole body examination.

Item missingness: age, sex: 0%; educational status: 0·2%; marital status: 0·2%; employment status: 0·3%; smoking: 0·4%; SF-12: 0·3%; cost data: 7·4%· Presented are unimputed variables with the exception of cost data with a 0 imputation, as described in methods.

Table 2

Estimated total additional costs of MRI participants compared to MRI Non-participants after examination per participant in €

	Year 1 after SHIP MRI examination		Year 2 after SHIP MRI examination		Total Year 1+2 after SHIP MRI examination	
	Mean (€)	95% CI (€)	Mean (€)	95% CI (€)	Mean (€)	95% CI (€)
Total ambulatory costs	130	37; 223	164	78; 251	295	134; 456
Costs primary care, prevention	8	-16; 33	39	14; 64	48	3; 93
Costs specialized care	65	-2; 132	86	26; 146	151	40; 262
Costs Imaging	53	30; 76	46	23; 70	100	61; 139
Costs Laboratory	6	0; 13	6	-1; 12	12	1; 24

Estimates were derived from two-part models with weights to compute average treatment effects during the years 1,2, and 1+2 respectively.

Year 1: Comprises quarters 1-4 after the SHIP examination. Year 2 comprises quarters 5-8 after the SHIP examinations. Year 1+2 quarters 1-8

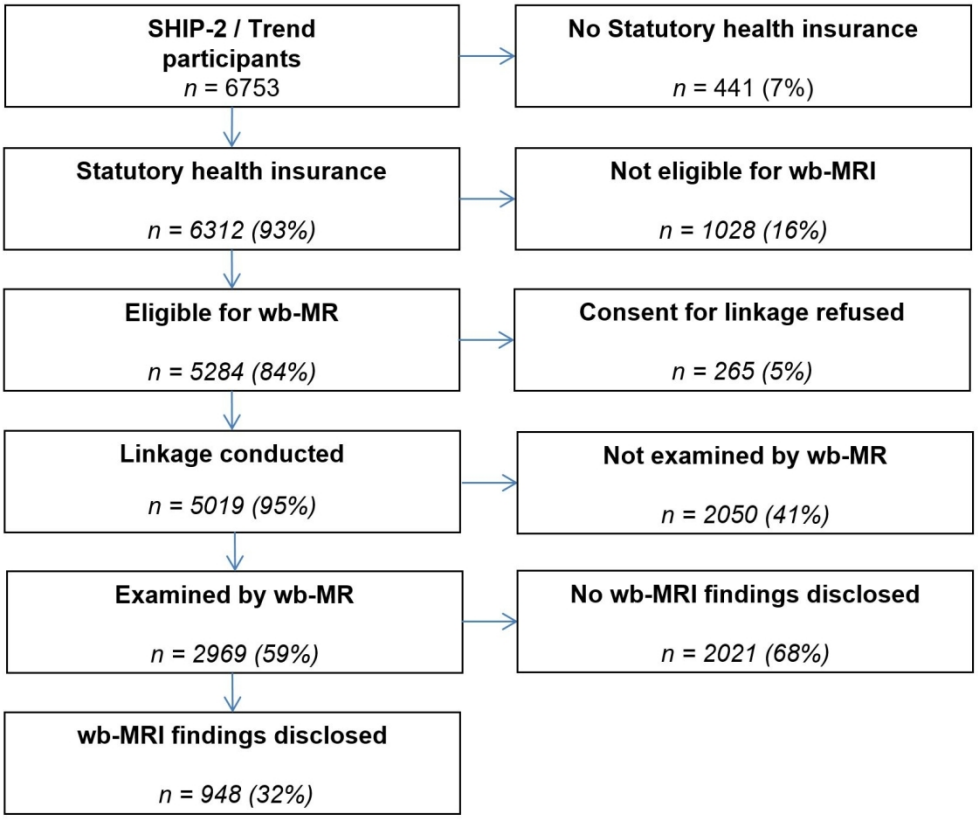


Figure 1
Study flow chart

196x165mm (216 x 216 DPI)

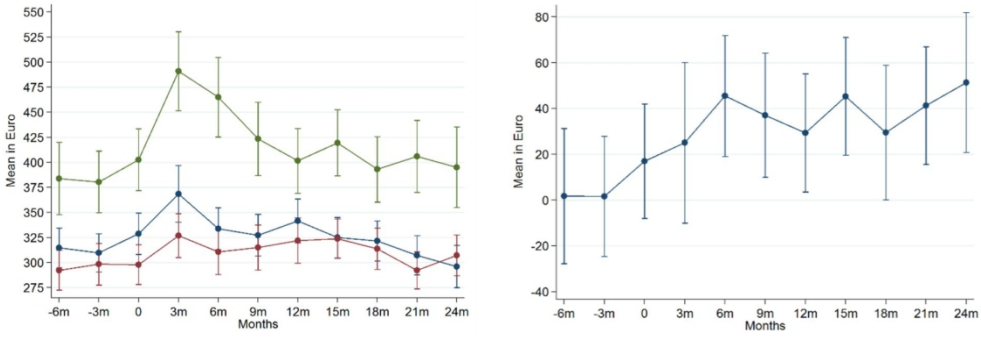


Figure 2
Course of total costs
0 is the reference quarter in which the MRI examination took place. The left column displays the descriptive course for the outcomes comparing MRI participants with disclosed findings, MRI participants without disclosed findings, and MRI non-participants. Average treatment effects are displayed in the right column. Propensity score reweighting was applied to balance the distribution of relevant covariates between MRI participants and non-participants. Estimates were derived from two-part models. N=5019.

122x40mm (400 x 400 DPI)

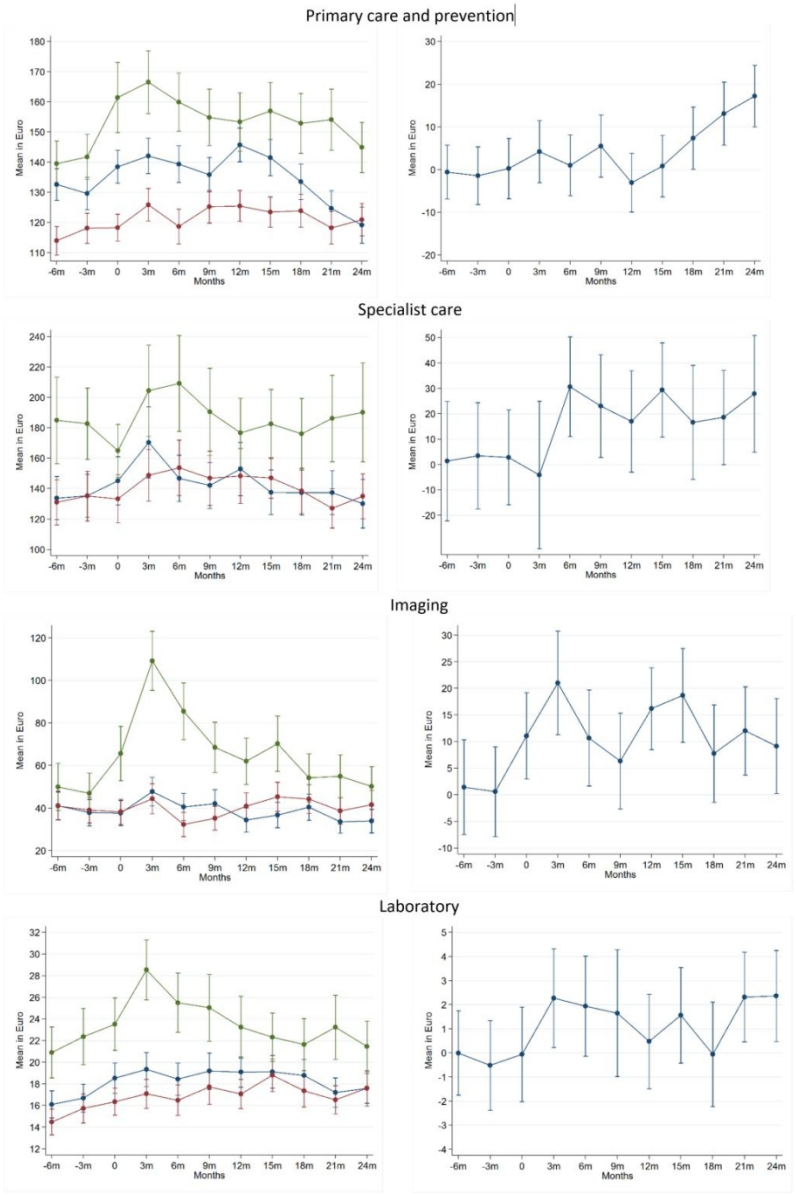


Figure 3
Course of costs by service type
0 is the reference quarter in which the MRI examination took place. The left column displays the descriptive course for the studied outcomes comparing MRI participants with disclosed findings, MRI participants without disclosed findings, and MRI non-participants. Average treatment effects are displayed in the right column. Propensity score reweighting was applied to balance the distribution of relevant covariates between MRI participants and non-participants. Estimates were derived from two-part models. N=5019.

44x66mm (600 x 600 DPI)

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***
Applied to
A population-based cohort study investigating the association of incidental whole-body magnetic resonance imaging findings with outpatient costs, tumours, and mortality

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	3-4
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	6
Objectives	3	State specific objectives, including any prespecified hypotheses	6
Methods			
Study design	4	Present key elements of study design early in the paper	7
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	7-10
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	7-8
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	n. a.
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7-10
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7-10
Bias	9	Describe any efforts to address potential sources of bias	10-11
Study size	10	Explain how the study size was arrived at	7 Fig. 1
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	10-11
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	10-11
		(*b*) Describe any methods used to examine subgroups and interactions	
		(*c*) Explain how missing data were addressed	
		(*d*) If applicable, explain how loss to follow-up was addressed	
		(*e*) Describe any sensitivity analyses	
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	7 Fig. 1
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	12 Tab. 1

			(b) Indicate number of participants with missing data for each variable of interest (c) Summarise follow-up time (eg, average and total amount)	
Outcome data		15*	Report numbers of outcome events or summary measures over time	13, Fig. 2
Main results	16		(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included (*b*) Report category boundaries when continuous variables were categorized (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Fig. 2 Tab. 2,3
Other analyses	17		Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11
Discussion				
Key results	18		Summarise key results with reference to study objectives	14
Limitations	19		Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	16-17
Interpretation	20		Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	16-17
Generalisability	21		Discuss the generalisability (external validity) of the study results	14,15
Other information				
Funding	22		Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	18