Aus der Abteilung SHIP-KEF

Leiter: Prof. Dr. med. Henry Völzke

des Instituts für Community Medicine der Universitätsmedizin Greifswald

Frequency of biopsies after the disclosure of incidental findings from whole-body research MRI

Inaugural – Dissertation

zur

Erlangung des akademischen

Grades

Doktor der Medizin

(Dr. med.)

der

Universitätsmedizin

der

Universität Greifswald

2021

vorgelegt von:

Elizabeth Sierocinski

geb. am: 12.06.1991

in: Toronto, Kanada

Dekan:	Prof. Dr. med. Karlhans Endlich
1. Gutachter:	Prof. Dr. rer. med. Dr. phil. Carsten Oliver Schmidt
2. Gutachter:	Prof. Dr. med. Martin Scherer
Tag der Disputation:	15.11.2021

Table of Contents

1 Summary

Large-scale, population-based cohort studies gather a range of data from participants over extended periods of time with the goal of providing researchers with information regarding the health status, prevalence of disease, and risk factors in a regional population. Examinations conducted in the context of population-based studies include imaging and laboratory testing and may yield abnormal results, also called incidental findings. According to predetermined disclosure policies, incidental findings may be disclosed to study participants. Evidence shows that the disclosure of incidental findings results in medical follow-up as research participants and their physicians seek to clarify the significance of findings.

This work examined the effect of disclosing incidental findings from whole-body MRI (wb-MRI) on the frequency and organ system of biopsies in participants in the Study of Health in Pomerania (SHIP), a population-based cohort study in Mecklenburg-West Pomerania. As most of the incidental wb-MRI findings involved unclear masses and lesions, we hypothesized that the disclosure of wb-MRI findings would lead to an increase in diagnostic biopsies. Based on current data showing that the outcomes of incidental imaging findings are frequently clinically irrelevant, we further hypothesized that an increase in biopsies would not translate to a clinically relevant increase in diagnoses of malignancies. We also took disclosed laboratory findings into account, as they were disclosed to all SHIP participants and may play a role in the decision to pursue a biopsy.

We found that the rate of biopsies increased after participation in SHIP and disclosure of incidental MRI and laboratory findings. Overall, most biopsies showed nonmalignant findings, indicating likely overdiagnosis and overtesting resulting from the disclosure of incidental findings in our cohort. However, subgroups of participants with disclosed MRI findings had a higher proportion of biopsies revealing premalignant or malignant diagnoses after SHIP, indicating that the applied decision rules for disclosure of MRI findings led to the identification of individuals with an elevated risk for premalignant or malignant diagnoses. The clinical relevance of these diagnoses is unclear and overdiagnosis cannot be ruled out.

In summary, we recommend more restrictive disclosure policies for incidental imaging findings in research to protect research participants from overtesting and to reduce bias. Further studies regarding the long-term morbidity and mortality of participants are needed to better understand the therapeutic impact of the disclosure of incidental wb-MRI findings in the research setting.

2 Introduction

2.1 Population-based research and cohort studies

Population-based research seeks to provide information regarding health outcomes of large groups of individuals, often inhabitants of a specific region. Data on the prevalence of diseases, risk factors and other determinants of health can be generalized to guide the improvement of regional health care by influencing health policy, spending and public health initiatives (1). The improved understanding of risk factors and disease processes resulting from population-based research can also influence the development and implementation of disease treatment and prevention strategies (2).

Population-based cohort studies are longitudinal and as such gather comprehensive health data from participants over an extended period of time. By analyzing data from cohort studies, researchers hope to better understand the relationships between risk factors and disease, as well as the prevalence of both (3). Prominent examples of population-based cohort studies include the Framingham Heart Study, which through years of data collection redefined the known risk factors for cardiovascular disease, thus influencing public health guidelines for cardiovascular disease prevention (4) and the Rotterdam Study, which seeks to clarify risk factors contributing to a variety of diseases in the elderly such as cardiovascular disease, neurological diseases, diabetes and cancer (5). In Germany, the German National Cohort (GNC) studies etiologic factors of major chronic diseases on a national level in large samples of adults from the general German population (6). On a regional scale, the Study of Health in Pomerania (SHIP) investigates the connections between risk factors, diseases and subclinical conditions in the northeast German state of Mecklenburg-West Pomerania.

2.1.1 Examinations in the SHIP cohort

Data is collected in cohort studies via participant interviews, questionnaires, laboratory and physical examinations, functional testing and imaging studies via high-powered modalities such as magnetic resonance imaging (MRI). An example of the latter is SHIP, which in 2008 was the first population-based cohort study to offer participants the option to undergo whole-body MRI (wb-MRI). The goals of SHIP wb-MRI included establishing population-based reference parameters for the sizes of body structures and determining the prevalence of abnormalities such as kidney cysts and lung nodules in the adult population (7). Data from wb-MRI and other examinations in SHIP help to identify risk markers for the development of disease or prediction

of treatment outcomes as part of personalized medicine, and in doing so improve regional health care (8).

2.1.2 Importance of valid results

The observational nature of population-based cohort studies mandates investigators to gather data without intervening in the course of the outcomes being measured. The goal of population-based observational studies is to generate generalizable conclusions which hold true not only for study participants but also for the underlying populations. For example, SHIP seeks to uncover information that can be applied to northeastern Germany. The extent to which the conclusions of observational studies can be generalized to populations beyond the study participants is referred to as external validity (9).

The generalizability of observational results has several preconditions such as avoiding manipulation of the study participants to ensure that data collected represents the natural course of events. Any intervention to the participants on the part of the investigators can cause a cohort study to resemble an *interventional* setting rather than an *observational* one. Thus, interventions represent a threat to the external validity of data in cohort studies. As a result, avoiding interventions and their consequences are a priority for investigators in studies such as SHIP.

2.2 Incidental findings and study validity

A challenge to the external validity of studies like SHIP is presented by incidental findings from various study examinations, the management of which involves potential disclosure and health-related advice to study participants. In general, incidental findings are defined as unintended abnormalities yielded from diagnostic medical examinations (10). These may be generated by all types of medical diagnostics including genetic and laboratory testing or imaging.

In a clinical setting, findings are considered incidental when they fall outside of the indication or clinical question guiding the examination. For example, in a patient undergoing a diagnostic computed tomography (CT) scan as part of the work-up for abdominal pain, any findings beyond those which could be attributed to the patient's symptoms (e.g., a kidney cyst) would be considered incidental. In a population research setting such as SHIP, examinations take place outside the clinical context. Subjects are often asymptomatic and examinations take place without clinical indication; as such, any abnormalities discovered have been defined as incidental findings (10).

2.2.1 Ethical handling of incidental findings

There exists a strong ethical basis for the disclosure of incidental findings to the affected individual. Providing health information to an individual enables independent decision-making (ethical concept of autonomy); moreover, it is the duty of the examiner to prevent harm (non-maleficence) to the examinee (11). In the research setting, the aforementioned ethical principles as well as the duty of researchers to promote an exchange marked by trust between participants and investigators (reciprocity) mandate the disclosure of potentially significant findings to participants (11, 12). Furthermore, study participants expect disclosure of findings resulting from research testing; in fact, participants in SHIP reported the desire to gain information regarding their health as a chief motivator to participate in population-based research (13).

2.2.2 Balancing ethics and study validity

The management of incidental findings in population-based cohort studies places key methodological and ethical priorities at odds with one another. On the one hand, reducing the obstacles to valid findings and maximizing generalizability is one of the main goals of population-based research. As mentioned above, this is dependent on the investigators' success in minimizing manipulation of the study participants. The disclosure of incidental findings of potential or uncertain clinical relevance may trigger clinical actions such as diagnostic and therapeutic interventions which would not have taken place if these findings had not been made known to participants (14). As such, disclosures of incidental findings may be reasonably expected to affect the healthcare-seeking behavior of participants and thus represent a form of intervention on the part of researchers. By altering the natural course of healthcare-seeking behavior in participants, such interventions threaten external validity and possibly even cause the cohort study setting to resemble that of an interventional one (15).

On the other hand, investigators are ethically mandated to avoid harming participants. If left undisclosed and thus untreated, significant incidental imaging findings such as an acute fracture or lung infection could almost certainly harm the affected participant. Conversely, the disclosure of clinically irrelevant incidental findings which trigger unnecessary invasive actions such as biopsies can also cause harm to participants. Withholding non-acute but still potentially significant findings such as kidney cysts or very small lung nodules could play a role in causing future diagnostic uncertainty (e.g., the finding is described in subsequent imaging and the lack of previous mention in earlier imaging may lead a clinician to falsely conclude that the cyst or nodule is newly grown) (16).

The disclosure of incidental findings to study participants simultaneously represents a departure from the methodological research goal of avoiding intervention as well as an ethical measure to prevent harm to participants. Balancing the ethical mandate with the threat to external validity of study data represents a conundrum to investigators as they seek to generate usable data and limit harm to their participants (13).

2.2.3 Incidental findings in research imaging: a particular challenge for SHIP

Due to their unprecedented nature and the immense logistical challenges involved, the management of incidental research wb-MRI findings presented a unique challenge to investigators in SHIP.

In the *clinical* setting, recommendations exist for the classification and handling of incidental imaging findings. For example, the American College of Radiology's Incidental Findings Committee publishes regular updates on this topic (17). The disclosure of all incidental findings from clinical imaging regardless of relevance is expected; however, radiologists are encouraged to clearly state the inconsequential nature of benign findings in order to avoid prompting unnecessary diagnostic steps or distress for patients (18). Medicolegal difficulties may result if incidental findings found in a clinical setting are withheld from patients (16, 19).

In the research setting, the management of incidental research imaging findings is the topic of numerous current publications (12, 20-23). However, at the time wb-MRI was implemented in SHIP (between 2008 and 2012) and given that SHIP was a pioneer in its use of wb-MRI in a research setting, little published information was available to guide the management of incidental research wb-MRI findings. For multiple reasons, existing clinical recommendations for the management of incidental MRI findings were of limited utility in the research setting. First, research images are produced in the absence of clinical indication; as such, important contextual elements for evaluating the potential significance of findings were missing. Second, vital aspects of individual patient history such as risk factors and specific symptoms may not have been gathered to the necessary level of detail in study questionnaires. Third, the need to balance methodological and ethical priorities unique to the research setting necessitated a research-specific management protocol for incidental findings separate from the available purely clinical recommendations. As a result, SHIP investigators designed a protocol to categorize incidental wb-MRI findings and to present unclear findings for discussion in an interdisciplinary advisory board (7). According to this protocol, experienced radiologists read each research scan and sorted 13,455 incidental wb-MRI findings from approximately 2500

participants into categories of clinical relevance, ranging from findings of no consequence such as anatomical variants to those requiring urgent medical attention (24).

Following categorization, a further challenge for investigators in SHIP was determining which incidental imaging findings should be disclosed to participants. Given the nonspecific nature of many incidental findings from research imaging and the limited data regarding their clinical outcomes (8, 24), the proper definition and management of *potentially relevant findings* was unknown (25-27). Based on the predefined categories of incidental imaging findings in SHIP, only 1330 of the 13,455 incidental wb-MRI findings were deemed by the interdisciplinary advisory board to be potentially clinically relevant and disclosed to participants (7). This relatively small number of potentially significant findings is reflected in the currently available literature. Today, preliminary evidence shows that the majority of incidental findings from research imaging are of no clinical significance (22, 24, 28).

2.2.4 Consequences of disclosing incidental findings in research imaging

When deciding which findings to disclose to participants, investigators in SHIP needed to weigh the potential benefits and risks of disclosure of incidental findings to participants. To avoid worrying participants over clinically insignificant findings and to prevent unnecessary further diagnostics (29), it is now well-established that irrelevant research findings should not be disclosed. Only significant medical findings or potentially clinically relevant findings in which further diagnostic clarification is warranted should be disclosed to participants (29). Based on available data, between 50 and 80% of these potentially relevant findings are suspicious for malignancy (24, 26). Conceivably, the disclosure of such findings could be beneficial for the affected participant by enabling the detection of serious disease at a treatable stage, thus improving or preserving quality or length of life (30, 31).

However, current limited data suggests that the majority of incidental MRI findings concerning for malignancy turn out to be clinically insignificant (26). If this is the case, participants may undergo costly (32) and potentially invasive diagnostic procedures which find that the incidental finding in question is a false-positive or clinically insignificant, thus yielding no benefit or even harm for the individual (13). Serious conditions discovered too late for treatment could cause affected individuals distress without the benefit of effective treatment (13). Another serious negative consequence is any number of unnecessary therapeutic interventions resulting from overdiagnosis and overtreatment (33) from which no improvement or even a decrease in

quality of life for the affected individual results. Participants may also experience psychological distress when faced with unclear incidental findings, regardless of the final outcome (21, 34).

2.2.4.1 Medical follow-up

Participants in SHIP wb-MRI who received disclosure of potentially clinically relevant incidental findings were encouraged in writing to seek physician care to clarify the nature of the findings (14). In the absence of clear clinical guidelines for the management of incidental findings from research imaging (26), physicians and patients may embark on so-called cascades of care, in which multiple diagnostic tests are used to clarify the nature of findings (35). A national survey of physicians in the United States revealed that doctors judged 69% of the cascades to be ultimately unnecessary and that some form of harm resulted to patients in 68% of cases (14). Despite this, surveyed physicians also report feeling obligated to pursue cascades of care due to uncertainty regarding the relevance of incidental findings, patient expectations, community norms and fear of potential medicolegal consequences if findings are not worked up (35, 36).

In the context of the high percentage of tumor-related findings found in MRI, physicians may choose to refer their patients for a biopsy of the detected lesion. Histological examinations of biopsied tissue represent the most effective but also the most invasive way to exclude serious pathology following abnormal imaging. By undergoing biopsy, a participant may reap the clinical benefits of diagnosis of treatable disease. However, the subject is exposed to rare but possible complications of biopsy such as distress, infection, bleeding, pain, damage to nearby structures or tumor seeding (37-40). There is currently no available data regarding the frequency and outcomes of biopsies after the disclosure of incidental findings from research imaging.

3 Research questions and hypotheses

We targeted the effect of the disclosure of incidental wb-MRI findings on biopsies as these represent invasive diagnostic procedures which carry potential risks for the affected individuals.

In this study, we assessed how the disclosure of incidental wb-MRI findings in the population-based cohort study SHIP altered the (1) frequency and (2) organ system distribution of biopsies among participants in the two years following study participation.

We additionally assessed the (3) outcomes of biopsies, analyzing whether the disclosure of incidental findings contributed to the detection of newly diagnosed malignancies.[1]

To address potential confounding, we took laboratory findings into account as they were gathered and disclosed to all participants and may have been relevant in diagnostic actions such as biopsies.

Based on knowledge from the current literature, we hypothesized that more biopsies would be performed after SHIP participation and disclosure of incidental wb-MRI findings, but that the clinical relevance would be minimal.

[1] Research questions 1 and 3 are addressed in our published work *Richter, Sierocinski 2020* and question 2 is briefly mentioned in the Online Appendix of the published work. The detailed results and discussion of research question 2 as well as additional subgroup analyses concerning research question 3 are unique to this dissertation.

4 Methods

This prospective observational study utilized data from SHIP and histology data from Greifswald University Hospital Department of Pathology. It complies with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement (**Appendix A**, (41)).

4.1 SHIP Cohort

SHIP is a population-based study consisting of two independent cohorts, SHIP and SHIP-TREND. Participants were selected from the counties of Nord- and Ostvorpommern and the cities of Greifswald and Stralsund in northeastern Germany (3). Of 6265 eligible individuals, 4308 (2192 women, response 68.8%) participated in the baseline examination (SHIP-0) which was performed from 1997-2001 (42). 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-0) was established in 2008, for which a stratified sample of 10000 was drawn from the central population registry. After exclusion of deceased and relocated participants, a net sample of 8826 remained, of which 4420 (2275 women, response 50.1%) participated. Of the SHIP-2 and SHIP-Trend participants who were invited to participate, 3,371 individuals underwent wb-MRI and 3,382 declined participation. A detailed overview of this process is provided in the study flow-chart (**Figure 1**). All analyses in this project are based on SHIP-2 and SHIP-Trend-0, which were conducted simultaneously.

4.2 SHIP Examinations

The examinations in SHIP-2 and SHIP-Trend-0 are presented in **Table 1**. Detailed descriptions of the examinations are presented elsewhere (3).

Table 1. SHIP-2 and SHIP-Trend-0 Examinations.

Personal interviews:	Ultrasound:
-Medical history	-Thyroid gland, carotid arteries, liver, gallbladder, kidneys, pancreas
-Sociodemographic information	-Echocardiography
-Psychometric information	Bioelectrical impedance examinations
Laboratory examination:	Body plethysmography
-Blood, urine, tongue/nasal/throat swabs, saliva, stool	Cardiopulmonary exercise testing
Basic medical examinations:	Bone density
-Height, weight, blood pressure	Dental and ophthalmological examinations
-Body measurements, ECG	Whole-body MRI (optional)

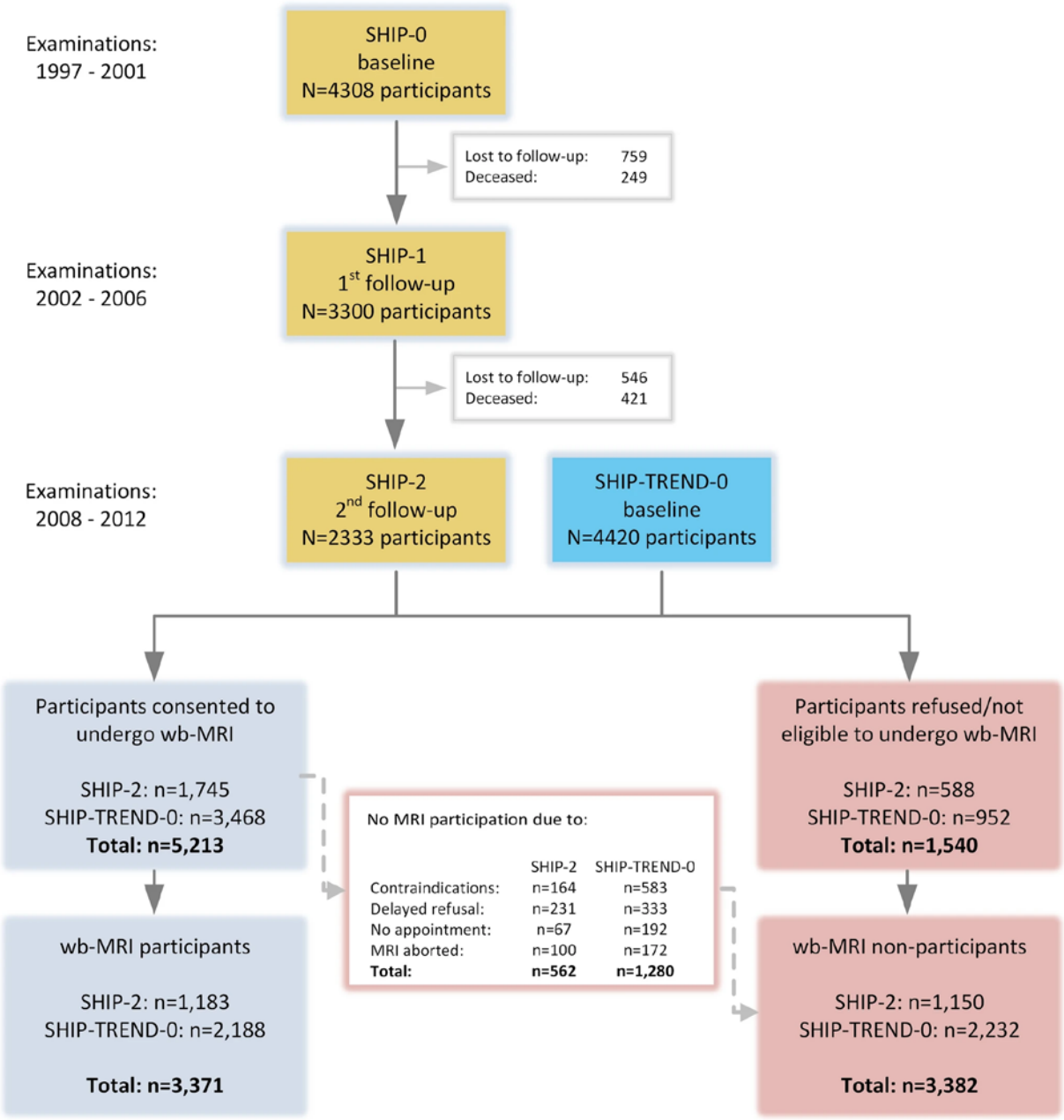


Figure 1. Study flow chart adapted from *Richter, Sierocinski et al. 2020.*

4.2.1 Whole-body MRI and reading of research scans

The methodology of wb-MRI scans and reading is adapted from our previous publication *Richter, Sierocinski et al. 2020*: "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 are described in other publications (7, 24).

Findings and anatomical variants were documented in a standardized 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 with 15 years of experience, resolved disagreements."

4.2.2 Laboratory examinations

All participants in SHIP-2 and SHIP-Trend-0 underwent comprehensive laboratory examinations. Serum and plasma values included DNA, complete blood counts, coagulation parameters, complete metabolic panels including electrolytes, creatinine, urea, and liver function tests, cholesterol panels, glycosylated hemoglobin, serum hormone and vitamin levels, thyroid antibody levels, and the inflammatory marker C-reactive protein. Urine iodine, thiocyanate, nitrate, albumin, leukocytes, erythrocytes, glucose, urobilinogen, bilirubin and creatinine were measured. Serum aliquots were stored at -80C. The laboratory in charge for SHIP blood samples participated in the official German external quality proficiency testing program. Assays were calibrated using international references (3).

4.2.2.1 Laboratory values potentially leading to biopsy

Parameters with the potential to trigger a biopsy were chosen on clinical grounds by JFC and ES (**Table 2**). When placed in the clinical context available to participants' primary care physicians, abnormalities in the selected values were judged as potentially important in the decision to pursue a biopsy.

Table 2. Laboratory parameters as potential triggers of biopsy.

Alanine transaminase (ALT)	Serum leukocytes (WBC)
Aspartate transaminase (AST)	Serum platelets (PLT)
Gamma-glutamyl transpeptidase (GGT)	Thyroid stimulating hormone (TSH)
Lipase	Urine erythrocyte count

A variable was generated for each laboratory parameter indicating whether the results crossed reference limits.

4.2.3 Disclosure of examination findings to participants

Health-related findings from laboratory testing and blood pressure measurements were disclosed to all participants. Each participant received a paper copy of laboratory results with values crossing reference limits highlighted as abnormal (3).

In contrast, not all findings from wb-MRI were disclosed to participants. The disclosure of wb-MRI findings was regulated by a standardized protocol that was approved by the institutional review board. Trained radiologists classified the findings into three categories based on clinical significance and level of urgency (**Table 3**) (3, 24).

Table 3. Categories of incidental findings from SHIP wb-MRI.

Category	Definition	Example
Category I	Normal or common findings in asymptomatic subjects.	Anatomical variants, old brain infarcts, sinusitis
Category II	Potentially clinically significant abnormalities requiring non-urgent medical evaluation	Lung nodules >4 mm, unclear liver/pancreatic/splenic lesions, chronic pancreatic, biliary or gallbladder disease, renal cysts Bosniak >2F, adrenal lesions >10 mm, prostatic hyperplasia >60 mL or lesion, complex ovarian cysts or lesions, breast lesions > BI-RADS 3
Category III	Urgent findings requiring immediate referral	Acute brain infarct, intracranial hemorrhage, cerebral edema, pneumothorax, pneumonia, ileus, acute urinary obstruction, acute bone fracture

Adapted from *Hegenscheid et al 2013* (24).

Category I abnormalities lacked well-defined diagnostic and therapeutic consequences according to existing clinical guidelines and best practice and were thus not disclosed to participants. Category II findings were disclosed to participants via post after approval by an interdisciplinary advisory board. Participants received notification of Category II findings approximately 6 weeks after wb-MRI examination (24). Category III findings were disclosed immediately to the participant after conclusion of the examination to facilitate immediate medical referral. More details regarding the nature of these findings are available in **Appendix B.** The flow of categorization and disclosure of findings is depicted in **Appendix C.**

4.3 Histological data from the University Hospital Department of Pathology

A total of 8576 histological (biopsy) reports dated from 2002 to 2019 were available from the database of the Greifswald University Medical Center Department of Pathology. Data processing of these reports involved linkage to SHIP participants followed by the systematic categorization of reports into categories based on organ system and clinical outcome.

4.3.1 Linkage of histological data to SHIP participants

After consent from participants was obtained, biopsy reports from the Department of Pathology were linked to SHIP participants based on last name, first name, date of birth, and sex of participants (43). In this process, biopsy records from 2002 to 2019 were linked to a total of 3489 SHIP-2 and SHIP-Trend-0 participants.

4.3.2 Classification of biopsy reports

All biopsy reports were available in unstructured, free-text format and contained varying levels of detail pertaining to clinical history or indication for biopsy, macro- and microscopic description of samples, differential diagnoses, excluded diagnoses, and final diagnosis. In complex cases or where immunochemical and genetic analyses were conducted, multiple follow-up reports for a single biopsy were present.

An initial categorization of biopsies into organ system categories took place using a keyword-based automatic categorization tool (**Appendix D**; German). Reports to be included in analyses underwent a systematic cross-check by independent reviewers.

4.3.2.1 Selection of reports for cross-check

An analysis period consisting of the two years before and after each participant's respective SHIP examination was established. All reports falling into this time period were selected for cross-check. Additionally, for participants with at least one report within the analysis period, all reports antedating this time interval were also cross-checked. These older reports provided relevant long-term contextual information pertaining to participant medical history such as prior malignancies diagnosed via biopsy. A total of 3011 biopsy reports were included for cross-check. Of these, 2271 were dated within the analysis period and 740 represented older biopsy reports.

4.3.2.2 Classification by organ system

Based on the tissue samples analyzed, biopsy reports were classified into organ system categories (**Table 4**). As needed, reports containing diverse tissue types were assigned to multiple organ system categories. For samples in which the biopsied tissue differed from site of disease origin (e.g. breast cancer metastases in the lung), the site of disease origin was used for organ system classification.

Table 4. Organ system categories for biopsy reports.

Organ system	**Description**
Gynecological	Female reproductive organs; not including breast
Breast	Female only
Gastrointestinal	Gastrointestinal tract and abdominal organs, including gallbladder, spleen, MALT lymphoma; not including liver or pancreas
Integumentary	Skin findings, including dermal tumors such as lipoma
Prostate	Prostate gland only; not including seminal vesicle and distal tracts
Ears-Nose-Throat	ENT findings, including tumors of the gums
Urological	Kidneys, bladder, and urinary tract; also including male reproductive organs
Neurological	Peripheral and central nervous system, including endocrine tumors in brain
Musculoskeletal	Joint, bone, ligament, muscle
Thyroid	Thyroid gland
Hematological	Blood and bone marrow
Cardiovascular	Cardiac and vascular biopsies
Pulmonary	Lungs, pleura
Liver and Pancreas	Liver, pancreas biopsies
Dental	Radicular cysts, periodontal disease; not including tumors of the gums
Ocular	Including trachoma, other causes of conjunctivitis
Miscellaneous	Nonspecific findings (e.g. foreign-body granulomas due to sutures) and those not fitting into above-established categories (e.g. adrenal adenoma, thymoma).

4.3.2.3 Classification by nature of finding

The outcome or diagnosis resulting from the histological examinations was also appraised using mutually exclusive categories: pre-cancer; 1st, 2nd, or 3rd malignancy; metastasis; benign tumor; follow-up of known malignant or suspicious process; no diagnosis (**Table 5**). As appropriate, reports containing diverse tissue types were assigned to multiple outcome categories.

Table 5. Classification of biopsy reports according to nature of findings.

Category	Description
No malignancy or tumor	Including hyperplasia (e.g. benign prostatic hyperplasia, hyperplastic polyps in gastrointestinal and gynecological organs, ductal hyperplasia in the breast), benign non-neoplastic conditions of the breast (e.g., fibrocystic changes, sclerosing adenosis), goiters, ganglion cysts, and keloid.
First malignancy	First malignancy diagnosis for the respective participant in our records.
Second malignancy	Second malignancy diagnosed for the respective participant; includes an independent growth of the first malignancy at a new location (e.g. skin cancer), but not metastases.
Metastases	Metastatic disease from a previously diagnosed malignancy.
Benign tumor	Including lipoma, uterine leiomyomas, benign breast tumors (e.g. fibroadenoma, papilloma), and hamartomas. This category was used in the case of follow-up of a known benign tumor.
Pre-cancerous lesion	Pre-cancerous or suspicious lesions requiring follow-up; including carcinoma in situ, colon polyps that were not hyperplastic or hamartomatous.
No diagnosis / Not classifiable	Pathologist unable to complete diagnostic evaluation of the sample. Reasons include insufficient tissue quantity or quality, or degeneration of tissue in the biopsy sample.
Third malignancy	Third malignancy diagnosed for the respective participant.
Fourth malignancy	Fourth malignancy diagnosed for the respective participant.
Fifth malignancy	Fifth malignancy diagnosed for the respective participant.
Consecutive report	Supplementary information provided by pathologist after submission of original report (e.g. genetic analyses, special staining, immunohistochemical analyses).

4.3.2.4 Consensus of independent readings

The cross-check of biopsy reports was completed using a data entry interface in which independent reviewers (ES, CH, JFC) classified free-text reports according to organ system and nature of finding (**Appendix E**). ES classified the entire set of N=3011 biopsies. Double-readings were conducted for 2510 (83.4%) of the reports: CH read n=1752 reports and JFC read n=758 reports. Dissent between reader classifications was resolved by consensus readings and consultation of a pathologist (StS). Overall, dissent between readers was observed in 239 of 2510 double-classified reports (9.5%). In total, 58 corrections of initial classifications done by ES (2.3% of 2510 double readings) were revoked by the consensus decision, yielding a predicted misclassification rate less than 3% in the 501 single-readings by ES.

Three (0.1%) autopsies were excluded and 56 (1.8%) reports were not classifiable due to missing or incomplete data. Twenty-five reports (0.8%) included two different tissue types and were assigned to two outcome categories.

4.4 Statistical analyses

Baseline participant characteristics were stratified by cohort and participation in wb-MRI. Crude event rates (per 100 observation years) of biopsies overall and in each organ system in the 2 years before and after SHIP were calculated. The frequency of biopsies was estimated using a generalized estimating equation (GEE) model with the variables age, sex, socio-demographic characteristics (education, relationship status), hospitalization within last 12 months, history of cancer, and the disclosure of incidental findings with tumor relevance. The coefficients of the model were exponentiated for interpretation as incident rate ratios (IRR). Cumulative biopsy rates according to the nature of findings were calculated using recurrent event analysis and represented graphically. The outcome categories 1st to 5th malignancies as well as metastases were summarized for this analysis. Participants were stratified into subgroups based on types of incidental findings to show the effects of incidental MRI findings vs laboratory abnormalities. Record linkage and the analysis of recurrent events were conducted using the statistical software R. For data pre-processing and the GEE models SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) was used. Missing values for all variables are provided in Table 1. Due to the minimal amount of missing data (maximum 1.9% for cancer history) complete case analyses were carried out without application of imputation techniques. All effects are presented with confidence intervals. Sensitivity analyses were performed to control for (1) the impact of known malignant conditions on the number of biopsies performed after SHIP participation and (2) potential limitation in coverage of participants living farther away from the Greifswald University Medical Center (i.e., effects of missing biopsy data with increasing distance of participant place of residence from examination center). Further details of all analyses are available in our published work (43). Statistical analyses were performed by Dr. Adrian Richter, co-first author of the published work *Richter, Sierocinski et. al. 2020*.

4.5 Ethics and Funding

All participants gave written informed consent. The Ethics Committee of Greifswald Medical University approved the study protocol (Nr. BB 39/08a). SHIP is funded by the Federal Ministry of Education and Research (Grant No. 03ZIK012), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania. Wb-MRI was supported by a joint grant from Siemens Healthcare, Erlangen, Germany, and the Federal State of Mecklenburg-Vorpommern. Dynamic contrast-enhanced MR mammography was supported by Bayer Healthcare. The investigation of the effects of incidental MRI findings was supported by the Deutsche Forschungsgemeinschaft [DFG, Grant Nrs. SCHM 2744/1-2:1/CH 921/1-2].

5 Results

5.1 Participant characteristics

Table 6 located on the following page provides an overview of participant characteristics (N=6753) stratified by cohort and MRI participation. Compared to MRI non-participants, MRI participants were younger, had undergone more years of education, were more likely to be in a relationship and more likely to be employed. MRI non-participants were more likely to have been hospitalized in the 12 months prior to study participation. The percentage of participants for which biopsy reports were available was similar in all strata, ranging from 17.2% to 18.3%. More than half of participants in all strata had laboratory anomalies in the selected parameters; MRI non-participants had more frequent laboratory anomalies. Of a total of 3371 MRI-participants, 1022 (30.3%; SHIP-2: n = 362, SHIP-TREND-0: n = 660) received disclosure of an MRI incidental finding. Of these, 851 MRI findings (83.3%) represented suspected tumors.

5.2 Outcome 1: Frequency of biopsies

A total of 2271 histological reports belonging to 1200 SHIP participants could be identified in our analysis period. These reports were dated from 2006 to 2012. Of these, 938 biopsy reports corresponding to 599 participants took place in the 2 years before SHIP. In the two years after SHIP, the number of biopsy reports increased to 1333 in 739 participants, corresponding to event rates of 6.95 biopsies per 100 observation years pre-SHIP and 9.87 post-SHIP (**Table 7**).

Table 7. Frequency of biopsies before and after SHIP.

Before SHIP		**After SHIP**	
No. biopsies	**Rate / 100 person-years**	**No. biopsies**	**Rate / 100 person-years**
938	**6.95 [95% CI: 6.51; 7.40]**	1333	**9.87 [9.35; 10.41]**

CI: Confidence interval.

In half (n=601) of the participants linked to biopsy reports, the first recorded biopsy was observed after participation in SHIPA total of 1041 (78.1%) of the 1333 biopsies processed by the Department of Pathology in the 2 years after SHIP belonged to these 601 participants.

Table 6. Study population characteristics.

Characteristics	SHIP-BASE (MRI)	SHIP-BASE (no MRI)	SHIP-TREND (MRI)	SHIP-TREND (no MRI)	All
N	1183	1150	2188	2232	6753
Age (years)					
Mean (SD)	55.7 (12.8)	59.0 (14.3)	51.2 (14.1)	52.7 (16.7)	53.8 (15.1)
Median [Min, Max]	56.0 [30.0, 90.0]	59.0 [31.0, 93.0]	52.0 [21.0, 82.0]	54.0 [20.0, 84.0]	54.0 [20.0, 93.0]
Sex					
Female	605 (51.1%)	630 (54.8%)	1113 (50.9%)	1162 (52.1%)	3510 (52.0%)
Male	578 (48.9%)	520 (45.2%)	1075 (49.1%)	1070 (47.9%)	3243 (48.0%)
Educational level					
Normal/high	856 (72.4%)	685 (59.6%)	1847 (84.4%)	1556 (69.7%)	4944 (73.2%)
Lower	327 (27.6%)	463 (40.3%)	337 (15.4%)	667 (29.9%)	1794 (26.6%)
Missing	0 (0%)	2 (0.2%)	4 (0.2%)	9 (0.4%)	15 (0.2%)
Years of education					
<10	253 (21.4%)	390 (33.9%)	344 (15.7%)	685 (30.7%)	1672 (24.8%)
10	652 (55.1%)	562 (48.9%)	1178 (53.8%)	1090 (48.8%)	3482 (51.6%)
>10	278 (23.5%)	193 (16.8%)	662 (30.3%)	448 (20.1%)	1581 (23.4%)
Missing	0 (0%)	5 (0.4%)	4 (0.2%)	9 (0.4%)	18 (0.3%)
Marital status					
Single	129 (10.9%)	114 (9.9%)	223 (10.2%)	262 (11.7%)	728 (10.8%)
In a relationship	965 (81.6%)	900 (78.3%)	1755 (80.2%)	1663 (74.5%)	5283 (78.2%)
Divorced	55 (4.6%)	76 (6.6%)	128 (5.9%)	155 (6.9%)	414 (6.1%)
Widowed	34 (2.9%)	57 (5.0%)	78 (3.6%)	143 (6.4%)	312 (4.6%)
Missing	0 (0%)	3 (0.3%)	4 (0.2%)	9 (0.4%)	16 (0.2%)
Employment status					
Unemployed	523 (44.2%)	649 (56.4%)	913 (41.7%)	1240 (55.6%)	3325 (49.2%)
Employed	658 (55.6%)	497 (43.2%)	1271 (58.1%)	980 (43.9%)	3406 (50.4%)
Missing	2 (0.2%)	4 (0.3%)	4 (0.2%)	12 (0.5%)	22 (0.3%)
Hospitalized in last 12 months					
No	1014 (85.7%)	919 (79.9%)	1905 (87.1%)	1864 (83.5%)	5702 (84.4%)
Yes	166 (14.0%)	227 (19.7%)	280 (12.8%)	360 (16.1%)	1033 (15.3%)
Missing	3 (0.3%)	4 (0.3%)	3 (0.1%)	8 (0.4%)	18 (0.3%)
Cancer history					
No	1092 (92.3%)	1071 (93.1%)	2047 (93.6%)	1957 (87.7%)	6167 (91.3%)
Yes	89 (7.5%)	78 (6.8%)	135 (6.2%)	153 (6.9%)	455 (6.7%)
Missing	2 (0.2%)	1 (0.1%)	6 (0.3%)	122 (5.5%)	131 (1.9%)
Histological data available					
No	979 (82.8%)	940 (81.7%)	1808 (82.6%)	1826 (81.8%)	5553 (82.2%)
Yes	204 (17.2%)	210 (18.3%)	380 (17.4%)	406 (18.2%)	1200 (17.8%)
Laboratory abnormalities					
No	492 (41.6%)	382 (33.2%)	832 (38.0%)	736 (33.0%)	2442 (36.2%)
Yes	691 (58.4%)	768 (66.8%)	1356 (62.0%)	1496 (67.0%)	4311 (63.8%)

Adapted from *Richter, Sierocinski et. al. 2020.*

5.2.1 Predictors for biopsies

Hospitalization within 12 months of SHIP, history of cancer and disclosure of incidental MRI findings were strong predictors for a higher number of biopsy reports (incidence rate ratios, IRR, 3.45, 2.89 and 2.17, respectively; **Table 8**). Tumor-related MRI findings were slightly stronger predictors for having a biopsy than MRI findings of any type (Table 8, Model 2). Female sex, higher age, disclosure of incidental laboratory findings and increasing time were moderate predictors for increased biopsy rates. Education, relationship status and employment status were found to have little association with biopsy rates.

Table 8: Predictors for the number of biopsy reports.

	Model 1		**Model 2**	
Predictors for biopsy reports	**IRR**	**95% CI**	**IRR**	**95% CI**
Age (per decade)	1.15	[1.08; 1.23]	1.16	[1.09; 1.23]
Sex (male vs. female)	0.73	[0.63; 0.84]	0.73	[0.63; 0.85]
Education (years, reference: 10y)				
<10y	0.99	[0.81; 1.20]	0.99	[0.82; 1.2]
>10y	0.99	[0.83; 1.17]	0.99	[0.84; 1.17]
Employed (yes vs. no)	0.92	[0.77; 1.11]	0.93	[0.77; 1.11]
Relationship status (reference: single)				
Married	1.04	[0.80; 1.35]	1.03	[0.79; 1.34]
Divorced	0.94	[0.65; 1.35]	0.92	[0.64; 1.33]
Widowed	1.17	[0.79; 1.75]	1.16	[0.77; 1.73]
Hospitalized in last 12 months (yes vs. no)	3.45	[3.01; 3.96]	3.45	[3.01; 3.96]
Known cancer history (yes vs. no)	2.89	[2.28; 3.67]	2.89	[2.28; 3.67]
Time-varying measures				
Disclosure of lab anomaly (yes vs. no)	1.37	[1.12; 1.67]	1.37	[1.12; 1.66]
Disclosure of MRI IF (yes vs. no)	2.17	[1.76; 2.68]	2.32	[1.85; 2.89]
Time (post-SHIP vs. pre-SHIP)	1.29	[1.07; 1.55]	1.30	[1.08; 1.56]

IF = incidental finding(s). Model 2 used only tumor-related IFs. Adapted from *Richter, Sierocinski et al. 2020.*

5.3 Outcome 2: Organ system distribution of biopsies

The proportion of biopsies in all organ systems except gynecological and skin biopsies increased in the 2 years after SHIP participation (**Figure 2**).

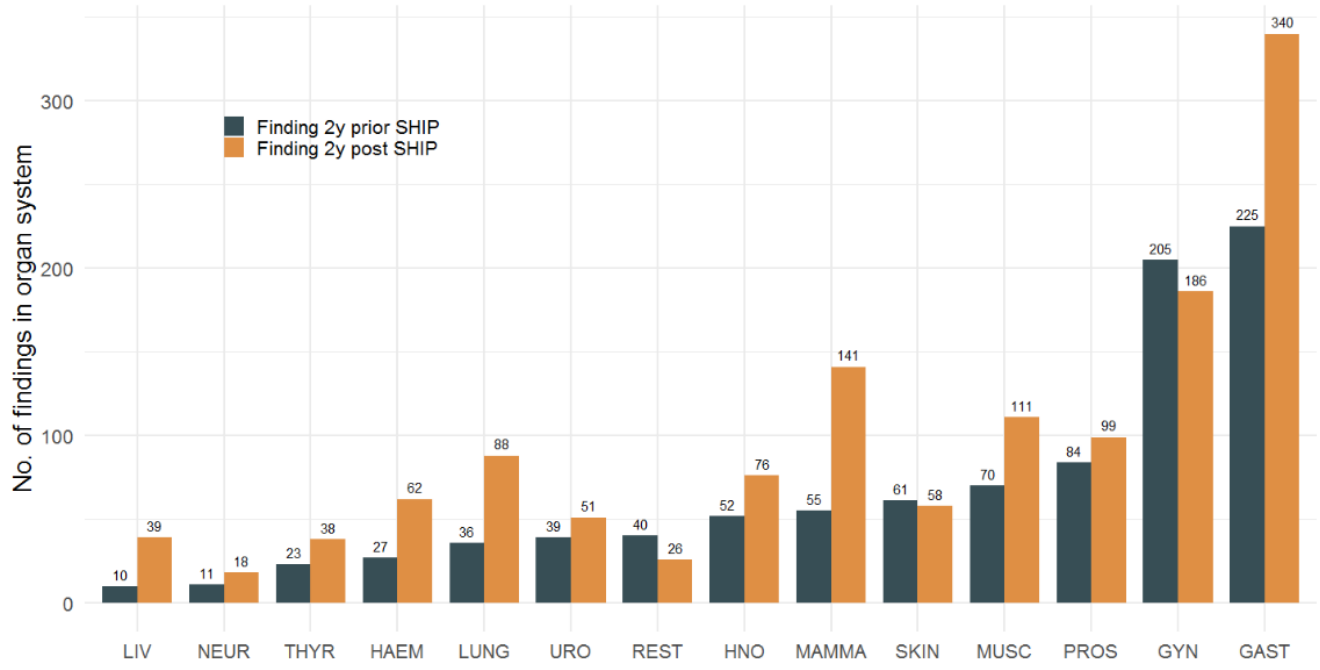


Figure 2. Frequencies of biopsies stratified for different organ systems, adapted from *Richter, Sierocinski et al 2020*, Online Supplement.
LIV = liver and pancreas, NEUR = neurological, THYR = Thyroid, HAEM = hematological, LUNG = pulmonary, URO = urological, REST = miscellaneous, HNO = ear-nose-throat, MAMMA = breast, SKIN = skin, MUSC = musculoskeletal, PROS = prostate, GYN = gynecological, GAST = gastrointestinal

The greatest increase in biopsy frequency was seen in that of biopsies of the liver and pancreas, which nearly quadrupled after SHIP (**Table 9**). A more than two-fold increase in the frequency of biopsies of the breast, lung and hematological system was observed post-SHIP. Lesser increases ranging from RR 1.46 to 1.58 were seen in biopsies of the musculoskeletal system, gastrointestinal system and ears-nose-throat. Increases between 1.18 and 1.65 which lacked statistical significance were observed in biopsies of the thyroid, neurological system, urological system and prostate. Marginal decreases were observed in skin and gynecological biopsies post-SHIP.

Table 9: Frequency of biopsies in each organ system before and after SHIP examination.

	Pre-SHIP		Post-SHIP		
Organ System	No. Biopsies	Rate / 100 person-years [95% CI]	No. Biopsies	Rate / 100 person-years	Rate Ratio (Post-/Pre-SHIP)
Liver and pancreas	10	0.07 [0.04; 0.14]	39	0.29 [0.21; 0.39]	3.85 [2.00; 8.21]
Neurological	11	0.08 [0.04; 0.15]	18	0.13 [0.08; 0.21]	1.63 [0.77; 3.58]
Thyroid	23	0.17 [0.11; 0.26]	38	0.28 [0.20; 0.39]	1.65 [0.99; 2.81]
Hematological	27	0.20 [0.13; 0.29]	62	0.46 [0.35; 0.59]	2.29 [1.47; 3.66]
Lung	36	0.27 [0.19; 0.37]	88	0.65 [0.52; 0.80]	2.44 [1.67; 3.64]
Urological	39	0.29 [0.21; 0.39]	51	0.38 [0.28; 0.50]	1.31 [0.86; 2.00]
Ear-Nose-Throat	52	0.39 [0.29; 0.50]	76	0.56 [0.44; 0.70]	1.46 [1.03; 2.09]
Breast	55	0.41 [0.31; 0.53]	141	1.04 [0.88; 1.23]	2.56 [1.89; 3.52]
Skin	61	0.45 [0.35; 0.58]	58	0.43 [0.33; 0.56]	0.95 [0.66; 1.36]
Musculoskeletal	70	0.52 [0.40; 0.65]	111	0.82 [0.68; 0.99]	1.58 [1.18; 2.15]
Prostate	84	0.62 [0.50; 0.77]	99	0.73 [0.60; 0.89]	1.18 [0.88; 1.58]
Gynecological	205	1.52 [1.32; 1.74]	186	1.38 [1.19; 1.59]	0.91 [0.74; 1.11]
Gastrointestinal	225	1.67 [1.46; 1.90]	340	2.52 [2.26; 2.80]	1.51 [1.28; 1.79]
Misc.	40	0.30 [0.21; 0.40]	26	0.19 [0.13; 0.28]	0.65 [0.39; 1.06]

5.4 Outcome 3: Biopsy outcomes by disclosure subgroups

5.4.1 Subgroup characteristics

Participants with MRI disclosures were older and more often in a relationship compared to participants without MRI disclosures (**Table 10**). Participants with MRI disclosures but without laboratory anomalies had a higher level of education than the other three subgroups. Participants with laboratory anomalies with or without MRI disclosures had higher hospitalization rates within the year prior to SHIP examination compared to participants without laboratory anomalies. A higher proportion of participants with both laboratory anomalies and MRI disclosures had a history of cancer compared to the other three subgroups. Biopsy data from the Department of Pathology were more likely to be available for participants with MRI disclosures.

Table 10. Participant characteristics in each disclosure subgroup.

Characteristics	Lab - | MRI -	Lab + | MRI -	Lab - | MRI +	Lab + | MRI +	All
N	2046	3685	396	626	6753
Age (years)					
Mean (SD)	52.6 (15.3)	53.7 (15.3)	56.6 (13.5)	57.1 (13.1)	53.8 (15.1)
Median [Min, Max]	53.0 [20.0, 88.0]	54.0 [20.0, 93.0]	58.0 [22.0, 83.0]	58.0 [23.0, 89.0]	54.0 [20.0, 93.0]
Sex					
Female	1082 (52.9%)	1889 (51.3%)	215 (54.3%)	324 (51.8%)	3510 (52.0%)
Male	964 (47.1%)	1796 (48.7%)	181 (45.7%)	302 (48.2%)	3243 (48.0%)
Educational level					
Normal/high	1543 (75.4%)	2621 (71.1%)	316 (79.8%)	464 (74.1%)	4944 (73.2%)
Lower	497 (24.3%)	1057 (28.7%)	80 (20.2%)	160 (25.6%)	1794 (26.6%)
Missing	6 (0.3%)	7 (0.2%)	0 (0%)	2 (0.3%)	15 (0.2%)
Years of education					
<10	463 (22.6%)	984 (26.7%)	71 (17.9%)	154 (24.6%)	1672 (24.8%)
10	1075 (52.5%)	1898 (51.5%)	200 (50.5%)	309 (49.4%)	3482 (51.6%)
>10	502 (24.5%)	793 (21.5%)	125 (31.6%)	161 (25.7%)	1581 (23.4%)
Missing	6 (0.3%)	10 (0.3%)	0 (0%)	2 (0.3%)	18 (0.3%)
Marital status					
Single	237 (11.6%)	419 (11.4%)	34 (8.6%)	38 (6.1%)	728 (10.8%)
In a relationship	1601 (78.3%)	2837 (77.0%)	320 (80.8%)	525 (83.9%)	5283 (78.2%)
Divorced	122 (6.0%)	234 (6.4%)	22 (5.6%)	36 (5.8%)	414 (6.1%)
Widowed	80 (3.9%)	187 (5.1%)	20 (5.1%)	25 (4.0%)	312 (4.6%)
Missing	6 (0.3%)	8 (0.2%)	0 (0%)	2 (0.3%)	16 (0.2%)
Employment status					
Unemployed	938 (45.8%)	1860 (50.5%)	201 (50.8%)	326 (52.1%)	3325 (49.2%)
Employed	1099 (53.7%)	1815 (49.3%)	194 (49.0%)	298 (47.6%)	3406 (50.4%)
Missing	9 (0.4%)	10 (0.3%)	1 (0.3%)	2 (0.3%)	22 (0.3%)
Hospitalized in last year					
No	1756 (85.8%)	3083 (83.7%)	342 (86.4%)	521 (83.2%)	5702 (84.4%)
Yes	283 (13.8%)	594 (16.1%)	53 (13.4%)	103 (16.5%)	1033 (15.3%)
Missing	7 (0.3%)	8 (0.2%)	1 (0.3%)	2 (0.3%)	18 (0.3%)
Cancer history					
No	1883 (92.0%)	3353 (91.0%)	372 (93.9%)	559 (89.3%)	6167 (91.3%)
Yes	139 (6.8%)	228 (6.2%)	24 (6.1%)	64 (10.2%)	455 (6.7%)
Missing	24 (1.2%)	104 (2.8%)	0 (0%)	3 (0.5%)	131 (1.9%)
Biopsy data available					
No	1753 (85.7%)	3051 (82.8%)	299 (75.5%)	450 (71.9%)	5553 (82.2%)
Yes	293 (14.3%)	634 (17.2%)	97 (24.5%)	176 (28.1%)	1200 (17.8%)

Lab - MRI - = no laboratory abnormalities or disclosure of MRI findings; Lab + MRI - = laboratory abnormalities, no MRI disclosures, Lab - MRI + = no laboratory abnormalities, disclosure of MRI findings, Lab + MRI + = laboratory abnormalities, disclosure of MRI findings.

5.4.2 Biopsy outcomes

Most biopsies before and after SHIP resulted in the outcome "no malignancy or tumor" (**Table 11**). The largest absolute increase in biopsies after SHIP was also found in this category (152 more biopsies post-SHIP or from 540 to 692 biopsies pre- and post-SHIP, respectively). Statistically significant increases in the rate of biopsies showing no malignancy or tumor were observed in participants with some form of incidental finding (either laboratory abnormalities or MRI disclosure) from SHIP. The greatest increases in biopsy rates post-SHIP ranging from rate ratio 1.22 to 1.87 were observed in participants with MRI disclosures, followed by participants with both laboratory abnormalities and MRI disclosures and those with laboratory anomalies only. The largest absolute number of participants entered this category after SHIP (99 additional participants post-SHIP for a total of 533 post-SHIP).

The absolute number of biopsies diagnosing benign tumors and participants receiving such biopsies increased slightly post-SHIP (4 additional biopsies and 3 additional participants for totals of 89 biopsies in 81 participants post-SHIP). The effects of SHIP participation on biopsy rate in this category did not show statistical significance. The largest of these effects was observed in the subgroup of participants with both laboratory abnormalities and MRI disclosures (RR 1.47).

Biopsies diagnosing a precancerous lesion increased from 61 pre-SHIP to 81 post-SHIP in 53 and 70 participants, respectively. No statistically significant effects on biopsy rates could be observed in any of the subgroups. The greatest potential effects were calculated for participants with MRI disclosures and without laboratory abnormalities, in which the biopsy rate tripled post-SHIP. Lesser increases of RR 1.51 and 1.30 were observed in biopsy rates for participants with laboratory anomalies and no MRI disclosures and those with both types of incidental findings, respectively.

Biopsies resulting in a newly diagnosed malignant process increased after SHIP across all participants, from 70 to 130 post-SHIP. A small number of 2nd, 3rd, 4th and 5th malignancies were diagnosed via biopsy after SHIP (**Appendix F**). Post-SHIP biopsy rates increased in all participants regardless of incidental findings. The largest increase in biopsy rate was a nearly four-fold increase in biopsies diagnosing a first malignancy in participants with incidental MRI findings but without laboratory abnormalities. Increases ranging from rate ratio 1.53 to 1.86 were observed in participants with no disclosures or laboratory abnormalities, laboratory abnormalities only, and both laboratory abnormalities and MRI disclosures post-SHIP.

The rate of biopsies containing consecutive reports increased for all participants, from 80 to 189 reports post-SHIP. A three- to six-fold increase in consecutive reports was observed in participants with MRI disclosures. The rate of follow-up biopsies of known malignant or suspicious processes increased by a rate ratio of 1.80 to 3.53 in participants with any type of incidental finding. The rate of follow-up biopsies decreased slightly in participants without MRI disclosures or laboratory anomalies. Increases in the rate of biopsies of all outcomes were observed in participants who received MRI disclosures. Overall, some type of disclosure was associated with an increase in biopsy-rate post-SHIP, with the exception of the rate of biopsies showing benign tumors in participants with laboratory abnormalities (0.82 [0.51; 1.31].

Table 11: Outcomes of biopsies stratified for disclosure of incidental findings.

Outcomes	Strata	N Participants	Before SHIP N Biopsies (Participants)	After SHIP N Biopsies (Participants)	Rate ratio of biopsies [CI]
No malignancy or tumor	Lab - | MRI -	2046	127 (109)	138 (119)	1.09 [0.85; 1.38]
	Lab + | MRI -	3685	316 (245)	385 (286)	1.22 [1.05; 1.41]
	Lab - | MRI +	396	31 (29)	58 (46)	1.87 [1.21; 2.92]
	Lab + | MRI +	626	66 (51)	111 (82)	1.68 [1.24; 2.29]
Subtotal		6753	540 (434)	692 (533)	-
Benign tumor	Lab - | MRI -	2046	18 (18)	20 (17)	1.11 [0.58; 2.13]
	Lab + | MRI -	3685	39 (36)	32 (32)	0.82 [0.51; 1.31]
	Lab - | MRI +	396	11 (10)	12 (12)	1.09 [0.47; 2.53]
	Lab + | MRI +	626	17 (14)	25 (20)	1.47 [0.79; 2.77]
Subtotal		6753	85 (78)	89 (81)	-
Pre-cancerous lesion (including carcinoma in situ)	Lab - | MRI -	2046	16 (14)	11 (9)	0.69 [0.31; 1.49]
	Lab + | MRI -	3685	33 (30)	50 (43)	1.51 [0.98; 2.37]
	Lab - | MRI +	396	2 (1)	7 (7)	3.32 [0.78; 24.59]
	Lab + | MRI +	626	10 (8)	13 (11)	1.30 [0.56; 3.06]
Subtotal		6753	61 (53)	81 (70)	-
Malignant process	Lab - | MRI -	2046	16 (16)	27 (23)	1.68 [0.91; 3.20]
	Lab + | MRI -	3685	41 (38)	63 (55)	1.53 [1.04; 2.29]
	Lab - | MRI +	396	3 (3)	12 (12)	3.85 [1.21; 17.67]
	Lab + | MRI +	626	15 (13)	28 (26)	1.86 [1.00; 3.58]
Subtotal		6753	75 (70)	130 (116)	-
Consecutive report	Lab - | MRI -	2046	24 (19)	32 (20)	1.33 [0.79; 2.29]
	Lab + | MRI -	3685	42 (31)	98 (54)	2.33 [1.63; 3.38]
	Lab - | MRI +	396	3 (3)	20 (14)	6.38 [2.17; 28.2]
	Lab + | MRI +	626	11 (7)	39 (24)	3.51 [1.86; 7.23]
Subtotal		6753	80 (60)	189 (112)	-
Follow-up of known malignant or suspicious process	Lab - | MRI -	2046	33 (17)	31 (21)	0.94 [0.57; 1.54]
	Lab + | MRI -	3685	44 (29)	80 (41)	1.82 [1.26; 2.64]
	Lab - | MRI +	396	3 (2)	11 (6)	3.53 [1.09; 16.36]
	Lab + | MRI +	626	16 (11)	29 (18)	1.80 [0.99; 3.41]
Subtotal		6753	96 (59)	151 (86)	-
Total		**6753**	**938 (599)**	**1333 (739)**	**1.42 [1.31; 1.55]**

Lab - MRI - = no laboratory abnormalities or disclosure of MRI findings; Lab + MRI - = laboratory abnormalities, no MRI disclosures, Lab - MRI + = no laboratory abnormalities, disclosure of MRI findings, Lab + MRI + = laboratory abnormalities, disclosure of MRI findings. Adapted from *Richter, Sierocinski 2020.*

5.4.3 Proportion of malignant findings by subgroup

Of the 2046 participants without laboratory abnormalities or MRI disclosures, 7.7% were biopsied before SHIP and 8.2% were biopsied after SHIP. In this group, the proportion of biopsies diagnosing malignancies before and after SHIP were 18.1% and 19.4%, respectively.

Of 3685 participants with laboratory anomalies and no MRI disclosures, 9.5% received a biopsy before SHIP and 11.3% after SHIP (**Table 12**). The proportion of biopsies diagnosing premalignant or malignant conditions increased from 17.2% before SHIP to 21.3% after SHIP.

Of the 396 participants with MRI disclosures and no laboratory abnormalities, 10.9% received a biopsy before SHIP and 19.4% were biopsied after SHIP. Prior to SHIP, 10.6% of biopsies in this subgroup diagnosed new premalignant and malignant conditions; after SHIP this proportion increased to 21.3%.

Of the subgroup of 626 participants with both laboratory abnormalities and MRI disclosures, 13.7% were biopsied prior to SHIP and 22.2% after SHIP. Prior to SHIP, 23.1% of biopsies diagnosed premalignant or malignant conditions; after SHIP, this proportion remained at 23.2%.

Table 12. Proportion of nonmalignant vs. malignant and findings in each subgroup

	Before SHIP		After SHIP	
Biopsy outcomes	**Biopsies**	**Participants***	**Biopsies**	**Participants***
Subgroup Lab- MRI- (N=2046)				
No malignancy or tumor	127	109	138	119
Benign Tumor	18	18	20	17
Pre-cancerous lesion	16	14	11	9
Malignant process	16	16	27	23
Subgroup Totals:	**177 (100%)**	**157 (7.7%)**	**196 (100%)**	**168 (8.2%)**
Nonmalignant	145 (81.9%)	127 (6.2%)	158 (80.6%)	136 (6.6%)
(Pre-)malignant	32 (18.1%)	30 (1.5%)	38 (19.4%)	32 (1.6%)
Subgroup Lab+ MRI- (N=3685)				
No malignancy or tumor	316	245	385	286
Benign Tumor	39	36	32	32
Pre-cancerous lesion	33	30	50	43
Malignant process	41	38	63	55
Subgroup Totals:	**429 (100%)**	**349 (9.5%)**	**530 (100%)**	**416 (11.3%)**
Nonmalignant	355 (82.8%)	281 (7.6%)	417 (78.7%)	318 (8.6%)
(Pre-)malignant	74 (17.2%)	68 (1.8%)	113 (21.3%)	98 (2.7%)
Subgroup Lab- MRI+ (N=396)				
No malignancy or tumor	31	29	58	46
Benign Tumor	11	10	12	12
Pre-cancerous lesion	2	1	7	7
Malignant process	3	3	12	12
Subgroup Totals:	**47 (100%)**	**43 (10.9%)**	**89 (100%)**	**77 (19.4%)**
Nonmalignant	42 (89.4%)	39 (9.8%)	70 (78.7%)	58 (14.6%)
(Pre-)malignant	5 (10.6%)	4 (1.0%)	19 (21.3%)	19 (4.8%)
Subgoup Lab+ MRI+ (N=626)				
No malignancy or tumor	66	51	111	82
Benign Tumor	17	14	25	20
Pre-cancerous lesion	10	8	13	11
Malignant process	15	13	28	26
Subgroup Totals:	**108 (100%)**	**86 (13.7%)**	**177 (100%)**	**139 (22.2%)**
Nonmalignant	83 (76.9%)	65 (10.4%)	136 (76.8%)	102 (16.3%)
(Pre-)malignant	25 (23.1%)	21 (3.4%)	41 (23.2%)	37 (5.9%)

Lab - MRI - = no laboratory abnormalities or disclosure of MRI findings; Lab + MRI - = laboratory abnormalities, no MRI disclosures, Lab - MRI + = no laboratory abnormalities, disclosure of MRI findings, Lab + MRI + = laboratory abnormalities, disclosure of MRI findings. *Percentages refer to the proportion of participants with biopsies in each subgroup.

5.4.4 Cumulative rate of selected outcomes

The rate of biopsies showing no findings (no tumor or malignancy) increased to the greatest degree in participants with MRI disclosures after SHIP (**Figure 3a**). This effect was most pronounced in participants with both MRI disclosures and laboratory anomalies (Fig. 3a, red line), followed by participants with MRI disclosures only (Fig. 4a, blue line). In contrast, no change in the rate of biopsies in participants without MRI disclosures (Fig. 3a, black and yellow lines).

The rate of biopsies showing benign tumors increased slightly in participants with MRI disclosures after SHIP (Fig. 3b, blue and red lines), but stayed unchanged in participants without MRI disclosures (Fig. 4b, black and yellow lines). Minimal to no effect could be seen in the rate of biopsies diagnosing precancerous conditions (Fig. 3c). The rate of biopsies diagnosing malignant processes increased only in the subgroup of participants with both incidental MRI findings and laboratory abnormalities (**Fig. 4d**, red line). The rate of biopsies showing malignancies did not increase for the other groups of participants (Fig. 4d, black, yellow, blue lines) after SHIP.

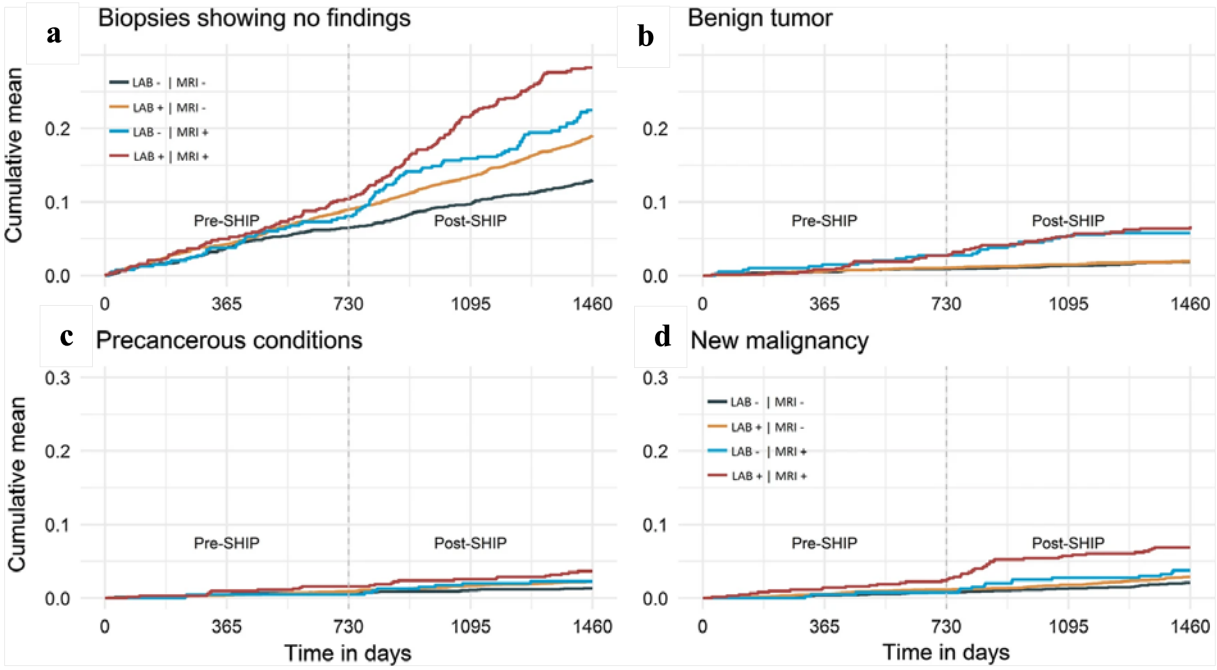


Figure 3. Cumulative rates of biopsies identifying (top left) exclusion of malignancy or benign tumor, (top right) benign tumors, (bottom left) pre-cancerous conditions (including carcinoma in situ), and (bottom right) malignancies. Adapted from *Richter, Sierocinski et al. 2020.*

5.5 Sensitivity analyses

In the first set of sensitivity analyses, our results regarding the impact of the disclosure of incidental findings on biopsy rates remained robust after excluding n = 455 participants with known malignant disease and n = 131 participants in which this information was unknown (**Appendices G-I**). Regarding coverage, MRI participation decreased slightly but inconsistently with increasing distance of participants' residence to the examination center (55–43%, **Appendix J**) and was therefore not considered in a separate analysis. Whereas biopsy data from the Department of Pathology was available for 22% of participants living near the examination site, only 7.5% of participants living 40–89 km away had available biopsy data (**Appendix K**). The coefficients for predictors of biopsy remained almost unchanged in two analyses which (1) weighted biopsies of participants living farther away and (2) excluded these participants (**Appendix L**).

6 Discussion

6.1 Summary of findings

Participation in a population-based cohort study and subsequent disclosure of incidental imaging and laboratory findings was associated with an increase in the rate of biopsies as well as a shift in the organ system distribution of biopsies performed. Incidental findings from SHIP, particularly MRI findings showing potential tumors, were contributing factors to the increase in biopsies. Most biopsies resulted in no malignancy or tumor.

6.2 Outcome 1: Frequency of biopsies

The total number of biopsies increased by 42.7% after SHIP and the rate of biopsies increased from 6.95 to 9.87 biopsies per 100 person-years. The majority of biopsies after SHIP represented first biopsies for the respective participants. Although the biopsy number can be expected to increase with the increasing age of our cohort, it is more likely that SHIP participation may have prompted first-time biopsies in many participants in the limited observation period. This is supported by the fact that each disclosure letter of incidental MRI findings included a written recommendation to consult a doctor. More than 80% of disclosed MRI findings concerned potential tumors (24), the workup of which can be expected to trigger biopsies. Indeed, the disclosure of incidental MRI findings of any type was found to be a predictor for biopsy in our study; this predictive effect was slightly stronger for MRI findings which were specifically tumor-related. Furthermore, compared to subgroups without MRI disclosures, a larger proportion of participants in subgroups with MRI disclosures was biopsied after SHIP (8-11% vs 19-22%, respectively). These results imply that MRI disclosures may have played a role in the decision to conduct a biopsy, given that this procedure is an effective way to rule out or diagnose suspected malignancies. The strongest predictor for biopsy was hospitalization within the year prior to SHIP participation with a rate ratio (RR) of 3.45, followed by known cancer history. These factors imply frequent or recent contact with the acute health care system and, particularly in the case of known cancer history, represent logical predictors for biopsy as histological examinations are regularly conducted in the context of an acute hospitalization or in cancer diagnosis and treatment.

6.3 Outcome 2: Organ system distribution of biopsies

Biopsy numbers after SHIP increased in all organ system categories except for gynecological and skin biopsies. The largest increases were observed in biopsies of the liver and pancreas, which are well-visualized on MRI and represent one of the top locations of incidentally

identified lesions in SHIP-wb-MRI (Appendix A, (24)). Marked increases were also seen in biopsies of the breast tissue, which is also well-visualized on MRI (44, 45) and represents a frequent incidental finding in SHIP. Nearly 18% of SHIP MR-mammography participants received disclosed findings related to potentially suspicious lesions of the breast rated at or above category 3 of the Breast Imaging Reporting and Data System (BI-RADS)[2] (Appendix A, (24, 45)). A possible explanation for the increase in hematological biopsies may be the disclosure of laboratory abnormalities, which may, in combination with additional clinical factors, prompt biopsies for clarification. Additionally, findings involving suspicious lymphadenopathy above normal size limits in SHIP wb-MRI were reported to participants and may also have led to biopsy (Appendix A (24)).

The decreases in the proportion of gynecological and skin biopsies after SHIP may be explained in part by their typically ambulatory setting. The tissue from e.g. ambulatory Pap-Smears as part of cervical cancer screening may have been sent to outpatient pathologists for histological examination. As a result, the data in our study, derived from the University Hospital Department of Pathology, would show fewer examinations in these organ systems. Furthermore, it is plausible that wb-MRI contributed little or no information regarding superficial findings such as moles on the skin or microscopic findings found in the cervix.

6.4 Outcome 3: Biopsy outcomes

The majority of biopsies after SHIP showed no tumor or malignancy. New malignancies were diagnosed via biopsy in a total of 116 participants (1.7%) post-SHIP. These findings correspond to results from other studies implementing wb-MRI, in which the incidence of malignancies resulting from incidental findings were 1.05% in 666 adult participants from the general population and 2.40% in 83 asymptomatic subjects (46, 47). An umbrella review of incidental findings from various imaging modalities in the clinical setting found malignancy rates ranging from less than 5% to 42%, depending on modality and organ system (25). However, the clinical studies in this review included asymptomatic oncology patients and included mostly imaging modalities other than MRI, providing potential reasons for the difference in malignancy rate when compared to our population-based research wb-MRI setting.

[2] Short-term follow-up is recommended for lesions of BI-RADS 3; biopsy is recommended for BI-RADS $\geq$ 4 (Balleyguier et al. 2007).

Biopsies of all outcomes increased most in participants with MRI disclosures, indicating that the disclosure of incidental MRI findings of unclear significance may trigger a general increase in biopsy rate. The association between incidental MRI and laboratory findings on the rates of biopsies resulting in benign and precancerous conditions was less clear. This may be a related to the low number of participants and biopsies in theses subgroups, which produced underpowered analyses. Most precancerous conditions are not expected to show visible anomalies in MRI, and neither condition is generally expected to produce laboratory abnormalities. However, some newly discovered benign and precancerous conditions may have been discovered as a result of cascades of care in which a range of diagnostics are triggered by incidental findings (14).

In the subgroup analyses, we observed the aforementioned higher proportion of participants with MRI disclosures biopsied after SHIP compared to those without MRI disclosures. Additionally, the post-SHIP proportion of malignant and premalignant biopsy results in these groups was higher than that of participants without any findings at all (21-23% vs 19%, respectively). However, although participants without MRI disclosures but with laboratory anomalies were biopsied less frequently after SHIP compared to those with MRI findings (11% vs 19-22%, respectively), this group showed the same proportion of malignant and premalignant biopsy findings after SHIP as the groups with MRI disclosures (21%). This implies firstly that the decision-making process used to determine which MRI findings to disclose to participants led to the identification of individuals who were more likely to have premalignant or malignant diseases. Indeed, after SHIP 5-6% of participants with MRI disclosures were found to have premalignant or malignant disease via biopsy, compared to 3% of those with laboratory anomalies only and 2% without either type of finding. Secondly, it appears that the disclosure of MRI findings did not necessarily increase the efficacy of biopsies in detecting malignant and premalignant conditions. The post-SHIP proportion of premalignant and malignant biopsy results was identical in the two groups with laboratory anomalies only and MRI disclosures only. This proportion remained unchanged at 23% in the subgroup with both laboratory anomalies and MRI disclosures. This illustrates that a potential ceiling effect regarding the efficacy of biopsies may have been reached. More data regarding biopsy outcomes after the disclosure of incidental findings are needed to further explore and confirm this assumption.

Only a marginal increase in the proportion of participants with malignant or premalignant diagnoses was observed in the subgroup with no MRI disclosures or laboratory anomalies. This small increase likely represents the expected small increase in malignant disease with the aging of our cohort. The proportion of participants with malignant and premalignant diagnoses in the subgroup with laboratory anomalies but without MRI disclosures rose from 2% to 3% after SHIP. This moderate increase may illustrate the increased morbidity of individuals with laboratory anomalies despite the lack of visible tumors on MRI. The subgroup with MRI disclosures but without laboratory anomalies with biopsies diagnosing malignant or premalignant disease displayed an increase from 1% to 5% of participants after SHIP. The comparatively low pre-SHIP proportion of participants with malignant and premalignant diagnoses is not explained by systematic differences between subgroups and is likely due to statistical uncertainty related to a comparatively small sample size in this subgroup. The larger post-SHIP proportion implies that having MRI findings is related to receiving malignant or premalignant biopsy results. Participants in this subgroup lacked laboratory anomalies, potentially indicating a lower level of morbidity compared to the fourth subgroup with both MRI disclosures and laboratory anomalies, which showed the highest proportion of participants with malignant or premalignant diagnoses post-SHIP (6%). This subgroup also had the highest pre-SHIP proportion of participants with malignant or premalignant biopsy findings (3%). This implies a higher level of baseline morbidity as confirmed by the comparatively higher rate of participants in this subgroup who indicated a previously known cancer history during the SHIP interview (over 10% compared to 6-7% in the other three subgroups). This constellation may illustrate the systemic effect malignancies have on an individual, causing not only imaging findings but also laboratory derangements. More research into this subgroup is required to determine its significance.

6.5 Clinical implications

6.5.1 Overtesting and overdiagnosis

The large increase in biopsies showing no tumor and no malignancy indicates possible overtesting. MRI disclosures were associated with a higher proportion of premalignant and malignant diagnoses via biopsy and most malignancies were diagnosed relatively quickly (within 6 months) after SHIP. However, it is unclear whether these diagnoses resulted in benefits for the affected participants. Based on our data, we cannot determine whether malignancies were identified at an early enough stage to implement treatment or to improve quality of life. It is well-known that the diagnosis of a malignancy does not necessarily improve

outcomes (36). Additionally, some of the malignancies diagnosed may have never become symptomatic or caused harm to the affected patients, which would be an indication of overdiagnosis (33). The large number of biopsies showing no malignancy or tumor after SHIP may represent overtesting resulting from the investigation of incidental findings (48). Additionally, the biopsies examined in our study may have been preceded or accompanied by other diagnostic steps such as additional laboratory testing or imaging studies (14). As such, we likely underestimate the extent of testing triggered by the disclosure of incidental findings from SHIP. Published data about cascades of care triggered by incidental imaging findings indicate that serious disease is only infrequently discovered (14, 22) but individuals are exposed to psychosocial and financial burdens (32, 34, 49).

6.5.2 Effects on medical practice

Imaging studies implemented in the general population such as in SHIP can be viewed as analogous to a clinical scan ordered without firm clinical indication. In the literature, incidental findings from research imaging in asymptomatic individuals have been found to be unlikely to be clinically serious (22, 28), and the clinical relevance of the premalignant and malignant conditions identified in 1.7% of our participants remains unclear. Findings from clinically non-indicated scans are also unlikely to provide relevant information. Nevertheless, physicians may order diagnostic tests which are not strictly indicated out of a desire for reassurance in situations of clinical uncertainty, a practice which has been long-criticized due to minimal or no beneficial effects for patients and increased costs to the health care system (50, 51). Educating physicians to promote a better understanding of the negative consequences of ordering unnecessary diagnostic testing has been proposed as a method of reducing the quantity of superfluous tests (52). Eye-catching acronyms such as VOMIT (Victims of Modern Imaging Technology) and SPEW (Scans Propagating Exponential Workloads) have been published to draw attention to the hardship faced by patients undergoing unnecessary and potentially harmful interventions (53) and to physicians working more than would be necessary as a result of incidental findings (54). Our study contributes to physician education by emphasizing that highly sensitive diagnostic imaging performed in asymptomatic individuals may trigger unnecessary invasive diagnostic tests. Our results serve as a warning to avoid ordering diagnostic tests without clinical indication in order to avoid causing needless harm and incurring unnecessary costs to patients and the health care system.

6.6 Research implications

6.6.1 Methodological implications

Wb-MRI and the disclosure of findings introduec bias into the longitudinal data of healthcare usage in our large cohort study. The implementation of wb-MRI and the subsequent disclosure of incidental findings represent interventions in what was intended to be an observational, non-interventional study. In fact, because SHIP implemented MRI scans in a population-based sample with a group of non-participants, this study now resembles a Phase I trial (15) with wb-MRI as a non-randomized intervention. Our findings imply that the disclosure of findings prompted a change in the healthcare-seeking behavior of MRI participants. These interventions thus limit the external validity of observations made on healthcare usage in our study; our data are thus rendered less applicable to the general population in this aspect.

6.6.2 Recommendations for disclosure policies

The methodological difficulties caused by disclosing incidental findings combined with the unclear clinical benefit (and evidence from other studies indicating psychosocial distress) for participants has led to the implementation of more restrictive disclosure policies in SHIP (55). We recommend further investigation into the effects of disclosure on participant mortality. Taking this information into account, more conservative disclosure strategies for incidental imaging findings should be implemented with the research aim to protect the integrity of study data by avoiding unnecessary interventions to participants. Such strategies also promote the wellbeing of participants by sparing them anxiety and unnecessary medical interventions (49). Furthermore, future studies implementing research imaging modalities should consider utilizing an interdisciplinary advisory board to identify potentially clinically relevant findings.

7 Strengths and Limitations

To the best of our knowledge, this was the first study to investigate the effects of the disclosure of incidental findings on the frequency and organ system distribution of routine biopsies and their diagnostic outcomes. Our cohort design, the large number of participants and presence of a control group for MRI participants in our study increases the generalizability of our evidence regarding the consequences of incidental finding disclosure in research. Our study investigated the global association between biopsies and disclosed incidental imaging findings; as such, the direct link between MRI finding location and biopsy location was not investigated in this project. Some imaging findings were already known to participants. However, a survey of SHIP

participants revealed that this was true only for a minority of participants (13.3%). Furthermore, malignant disease was already known in some participants prior to SHIP participation (586 participants total, 238 MRI participants) and some biopsies showing malignancies may have resulted from this previously known disease. However, our results showing the impact of incidental findings on biopsies remained robust in a sensitivity analysis including only SHIP participants without previously known malignancies. Participants living farther away from our examination site may have undergone biopsies at other locations. There is no indication of a systematic difference between MRI participants and nonparticipants regarding biopsy location. A sensitivity analysis including only participants living close to the University Hospital Department of Pathology did not affect our results.

8 Future areas of study

More information is needed to better understand the consequences of disclosing incidental MRI findings to research participants. Studies investigating long-term morbidity and mortality of SHIP participants will provide information on potential benefits of the disclosure of incidental findings on patient outcomes. Detailed information regarding follow-up visits, examinations and procedures resulting from disclosed incidental findings and the associated monetary costs is needed to better understand the impact of incidental findings on the health care system.

9 Conclusion

The disclosure of incidental imaging findings resulting from participation in a population-based cohort study was associated with an increase in biopsies, the majority of which showed no tumor and no malignancy. More premalignant and malignant biopsy findings were observed in subgroups of participants with incidental MRI findings; data regarding long-term mortality are needed to clarify whether a therapeutic benefit resulted from these diagnoses. There is evidence that the disclosure of incidental imaging findings leads to overtesting and possible overdiagnosis on a population level. Furthermore, the disclosure of incidental wb-MRI and laboratory findings introduced bias into our observational data regarding healthcare usage. Strict disclosure policies are recommended protect the integrity of research data and participants from unnecessary invasive, risk-associated diagnostic procedures. Furthermore, physicians considering diagnostic imaging examinations for their patients should carefully assess the indication and weigh the risks and benefits, as our study shows that high-powered scans performed without clinical indication may trigger invasive testing which does not necessarily translate to benefits for the examined individuals.

10 Appendices

Appendix A. STROBE-Checklist for Cohort studies.

	Item No	Recommendation	Page Nr
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	5
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	5
Introduction			
Background/ rationale	2	Explain the scientific background and rationale for the investigation being reported	6-11
Objectives	3	State specific objectives, including any prespecified hypotheses	12
Methods			
Study design	4	Present key elements of study design early in the paper	13
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	13-16
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	13
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	13-16
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	14-16
Bias	9	Describe any efforts to address potential sources of bias	15, 20
Study size	10	Explain how the study size was arrived at	13
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	20
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	20
		(*b*) Describe any methods used to examine subgroups and interactions	20
		(*c*) Explain how missing data were addressed	20
		(*d*) If applicable, explain how loss to follow-up was addressed	
		(*e*) Describe any sensitivity analyses	20
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	13-14
		(b) Give reasons for non-participation at each stage	Fig. 1
		(c) Consider use of a flow diagram	Fig. 1

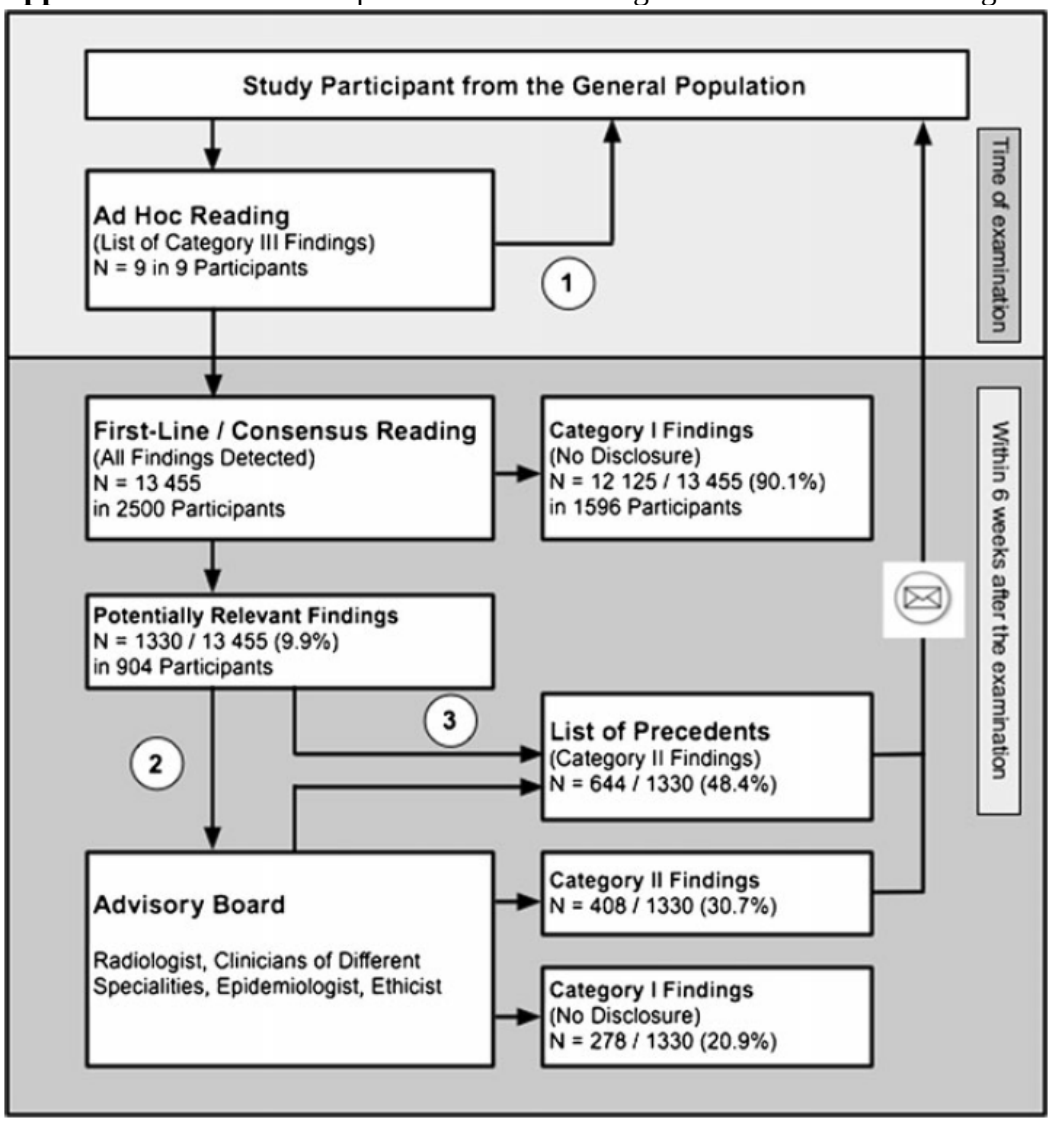
Descriptive data	14	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	22, Tab. 6
		(b) Indicate number of participants with missing data for each variable of interest	Tab. 6
		(c) Summarise follow-up time (eg, average and total amount)	22, Tab. 6
Outcome data	15	Report numbers of outcome events or summary measures over time	22, Tab. 6
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	21-32
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	28-32
Discussion			
Key results	18	Summarise key results with reference to study objectives	33
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	38-39
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	33-38
Generalisability	21	Discuss the generalisability (external validity) of the study results	34
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	20

Appendix B. Category II (n=1043) and III (n=9) findings in N=787 SHIP wb-MRI participants.

Organ	Example	No. Findings	Frequency (%)
Head	Brain tumor, subdural hematoma, normal pressure hydrocephalus, acute brain infarct	39	1.84
Neck	Pharyngeal / salivary gland / thyroid tumors, cervical lymphadenopathy >15 mm sd, goiter with tracheal compression	69	2.76
Chest	Lung nodule >4 mm, hilar / mediastinal / axillary lymphadenopathy >15 mm sd, pleural effusion, pneumonia	76	3.04
Abdomen	Chronic liver disease, liver lesions, cholestasis, chronic cholecystitis, chronic pancreatitis, pancreatic tumor, splenomegaly, splenic tumor, large herniations, abdominal lymphadenopathy >15 mm sd	170	6.80
Urinary tract	Renal cysts ≥ Bosniak 2F, renal tumors, adrenal tumors, chronic urinary obstruction, bladder tumors	170	6.80
Male reproductive	Prostatic hyperplasia or tumor >60 mL, inguinal testis, testicular / epididymal / seminal vesicle tumors	119 / 1229 male participants	9.68 of male participants
Female reproductive	Uterine / cervical tumors, complex ovarian cyst or tumor	93 / 1271 female participants	7.31 of female participants
Musculoskeletal	Spinal stenosis with myelopathy, intraspinal tumor, bone lesions, severe bone edema	150	6.00
Cardiac MRI	Heart failure, myocardial tumor, pericardial effusion	7 / 1129 participants	0.62 cardiac MRI participants
MR angiography	Intracranial aneurysm, cavernous malformation, internal carotid artery stenosis >50%, thoracic / abdominal aneurysm, thoracic / abdominal stensosis	55 /619 male participants	8.9 MR angiography participants
MR mammography	Breast lesion ≥ BI-RADS 3	97	17.83 MR mammography participants
Total		**1052**	

Findings in this table refer to all N=787 SHIP wb-MRI participants with disclosed MRI IFs.
Adapted from *Hegenscheid et al 2013*.

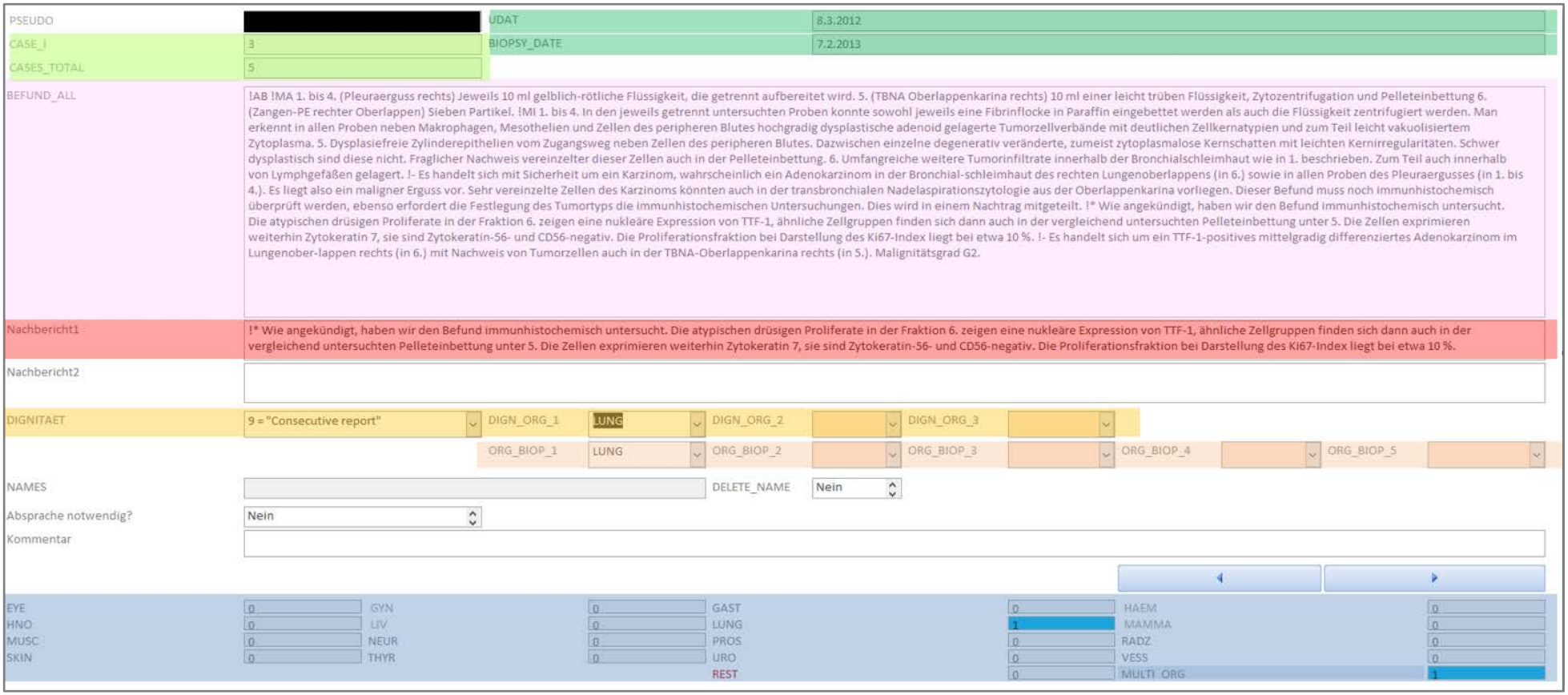
Appendix C. Standardized protocol for the management of incidental findings.



(1) Ad hoc reading: with the study subject still present in the MRI unit the radiologist checked the data set for a defined number of acute findings requiring immediate referral (category III).
(2) During firstline and a consensus reading the radiologists identified findings of potential clinical relevance that were presented monthly to an interdisciplinary advisory board. The board reached a consensus about whether to recommend further clinical work-up (category II) or not (category I). Participants were informed about the abnormality and recommendation of the advisory board by postal mail within 6 weeks of the examination.
(3) For frequent potentially relevant incidental findings the advisory board established precedents that were communicated to the participant without previous presentation to the advisory board.
Adapted from *Hegenscheid et al 2009*.

Appendix D. Keywords for automatic keyword-based categorization tool [German].

Kategorie	**Suchwort(e)**
Gynäkologie (ohne Brust)	Uterus, Zervix, Cervix, Abort, Myom, Endometri, Ovar, Abrasio, endozervikal, Leiomyom, Salpin, Adnex, Plazenta, Placenta, Blasenmole, Portio, Papanicolaou, Eileiter, Tuba uterina, fetal, Fetus, uteri, Hysterektomie, Menorrhoe, PAP-Test, PAP, Vulva, Schwangerschaft
Brust (weiblich)	Mamma, Mastopathie, Mastitis, Brustdrüse, Mastektomie, Fibroadenom, Progesteron(rezeptor), Östrogen(rezeptor), Rezeptorstatus, HER2/neu
Gastrointestinaltrakt	Ösophag, Oesophag, Dickdarm, Colon, Magen, Antrum, Kolon, Gastritis, Dünndarm, Divertikulitis, Sigma, Rektum, Korpuschleimhaut, Ileum, Appendi, Gallenblase, Anus, Hämorrhoid, Haemorrhoid, Cholecyst, Duoden(um), Jejunum, rektal, biliaris, Caecum, Zökum, Cardia, Pylorus, sigmoideum, inkarzerierter Hernie, periproktitisch, Bruchsack
Haut und Unterhaut	Melanom, Hautklinik, seborrhoisch, Naevus, Nävus, Korium, Dermatose, Basaliom, Hautexzidat, Pilonidal, Fremdkörpergranulom, Atherom, Basalzellkarzinom, Spinaliom, Lichtschaden, Narbenkeloid, Papillom, Condylomata
Prostata	Prostata, Gleason
HNO	Kontaktgranulom, Parotis, Nasennebenhöhle, Zunge, Ohrmuschel, Laryn, Mundboden, Nasenrachen, Stimmlippe, Pharyn, Rachen, Glottis, Rhinitis, Sinusitis, Mundschleimhaut. Speicheldrüse, Tonsill, Kieferhöhle, Sialolithiasis
Urologie (ohne Prostata)	Seminom , Hoden, Niere, Ureter, Hydrocele, Hydrozele, Urozystitis, Harnblase, Nebenhoden, Epididym, Penis, nephro, renal, Urethra, Urothel, Phimose, deferens
Neurologie (ZNS und PNS)	Neurinom, Glioblastom, Meningeom, Schwannom, Liquor, neuro, zerebral, Neuritis, Meningitis, Myelitis, Hypophysenadenom
Muskuloskeletal	Nucleus pulposus, Synoviali, Bandscheibe, Meniskus, (Pseud)arthrose, patella, osteo, Knochennekrose, Enchondrom, Exostose, Achillessehne, myopathisch, Bizeps, Knochenmetastase, Arthrose, Wirbelkörper, Chordom, kondylus, Osteomyelitis
Schilddrüse	Schilddrüse, Struma, thyreo, TSH, T3, T4
Blut	Non-Hodgkin-Lymphom, B-Zell-Lymphom, Ringsideroblasten, Myelodysplasie, monoklonale Gammopathie. Anämie, Anaemie, Leukozytose, Lymphom, toxische Knochenmark. Sideropenie, Erythropoese, Granulopoese, Leukämie, Plasmozytom, Knochenmark
Gefäße	Arteriitis, Angiopathie, Arteriosklerose, arteriosklerotische Plaque
Lunge	Bronchial, Pleura, Alveol, COPD, Pneumo, Bronchitis, Mesothel, Bronchus, pulmonal
Leber und Pankreas	Leberbiopsie, Leberpunktat, Pankreaskarzinom, pankrea Ascites, Leberzirrhose, Hepatitis B, periportal, Leber
Zähne	Radikuläre Zyste, dental, periapikal, Parodontitis
Auge	Trachom, Konjunktivitis
Rest	Keines der Suchworte kam vor



Appendix E. Data entry interface for categorization of findings by multiple readers.

Biopsy number, total number of biopsies for participant

Date of SHIP examination, date of biopsy

Organ system(s) assigned by automatic categorization tool

Free-text body of biopsy report

Follow-up report (e.g., immunohistochemical testing)

Nature of finding

Organ system

Appendix F: Detailed outcomes of biopsies stratified for disclosure of incidental findings.

Outcomes	Strata	N Participants	pre SHIP N Biopsies (Participants)	post SHIP N Biopsies (Participants)	Rate ratio of biopsies [CI]
No malignancy or tumor	Lab - | MRI -	2046	127 (109)	138 (119)	1.09 [0.85; 1.38]
	Lab + | MRI -	3685	316 (245)	385 (286)	1.22 [1.05; 1.41]
	Lab - | MRI +	396	31 (29)	58 (46)	1.87 [1.21; 2.92]
	Lab + | MRI +	626	66 (51)	111 (82)	1.68 [1.24; 2.29]
Subtotal		6753	540 (434)	692 (533)	-
Benign tumor	Lab - | MRI -	2046	18 (18)	20 (17)	1.11 [0.58; 2.13]
	Lab + | MRI -	3685	39 (36)	32 (32)	0.82 [0.51; 1.31]
	Lab - | MRI +	396	11 (10)	12 (12)	1.09 [0.47; 2.53]
	Lab + | MRI +	626	17 (14)	25 (20)	1.47 [0.79; 2.77]
Subtotal		6753	85 (78)	89 (81)	-
Pre-cancerous lesion (including carcinoma in situ)	Lab - | MRI -	2046	16 (14)	11 (9)	0.69 [0.31; 1.49]
	Lab + | MRI -	3685	33 (30)	50 (43)	1.51 [0.98; 2.37]
	Lab - | MRI +	396	2 (1)	7 (7)	3.32 [0.78; 24.59]
	Lab + | MRI +	626	10 (8)	13 (11)	1.3 [0.56; 3.06]
Subtotal		6753	61 (53)	81 (70)	-
1st malignancy	Lab - | MRI -	2046	15 (15)	21 (21)	1.4 [0.72; 2.77]
	Lab + | MRI -	3685	33 (33)	44 (44)	1.33 [0.85; 2.11]
	Lab - | MRI +	396	2 (2)	11 (11)	5.18 [1.37; 36.5]
	Lab + | MRI +	626	11 (11)	21 (21)	1.9 [0.93; 4.11]
Subtotal		6753	61 (61)	97 (97)	-
2nd malignancy	Lab - | MRI -	2046	1 (1)	2 (2)	1.88 [0.15; 58.99]
	Lab + | MRI -	3685	3 (3)	10 (10)	3.22 [0.97; 15.05]
	Lab - | MRI +	396	0 (0)	0 (0)	0 [0; Inf]
	Lab + | MRI +	626	2 (2)	4 (4)	1.93 [0.35; 15.61]
Subtotal		6753	6 (6)	16 (16)	-
3rd malignancy	Lab - | MRI -	2046	0 (0)	1 (1)	NA
	Lab + | MRI -	3685	1 (1)	2 (2)	1.88 [0.15; 58.99]
	Lab - | MRI +	396	1 (1)	0 (0)	0 [0; 19]
	Lab + | MRI +	626	0 (0)	1 (1)	NA
Subtotal		6753	2 (2)	4 (4)	-
4th malignancy	Lab - | MRI -	2046	0 (0)	1 (1)	NA
	Lab + | MRI -	3685	1 (1)	0 (0)	0 [0; 19]
	Lab - | MRI +	396	0 (0)	0 (0)	0 [0; Inf]
	Lab + | MRI +	626	0 (0)	0 (0)	0 [0; Inf]
Subtotal		6753	1 (1)	1 (1)	-
5th malignancy	Lab - | MRI -	2046	0 (0)	0 (0)	0 [0; Inf]
	Lab + | MRI -	3685	1 (1)	0 (0)	0 [0; 19]
	Lab - | MRI +	396	0 (0)	0 (0)	0 [0; Inf]
	Lab + | MRI +	626	0 (0)	0 (0)	0 [0; Inf]
Subtotal		6753	1 (1)	1 (1)	-
Metastasis	Lab - | MRI -	2046	0 (0)	2 (2)	NA
	Lab + | MRI -	3685	2 (1)	7 (7)	3.32 [0.78; 24.59]
	Lab - | MRI +	396	0 (0)	1 (1)	NA
	Lab + | MRI +	626	2 (1)	2 (2)	1 [0.1; 9.61]
Subtotal		6753	4 (2)	12 (12)	-
Consecutive report	Lab - | MRI -	2046	24 (19)	32 (20)	1.33 [0.79; 2.29]
	Lab + | MRI -	3685	42 (31)	98 (54)	2.33 [1.63; 3.38]
	Lab - | MRI +	396	3 (3)	20 (14)	6.38 [2.17; 28.2]
	Lab + | MRI +	626	11 (7)	39 (24)	3.51 [1.86; 7.23]
Subtotal		6753	80 (60)	189 (112)	-
Follow-up of known malignant or suspicious process	Lab - | MRI -	2046	33 (17)	31 (21)	0.94 [0.57; 1.54]
	Lab + | MRI -	3685	44 (29)	80 (41)	1.82 [1.26; 2.64]
	Lab - | MRI +	396	3 (2)	11 (6)	3.53 [1.09; 16.36]
	Lab + | MRI +	626	16 (11)	29 (18)	1.8 [0.99; 3.41]
Subtotal		6753	96 (59)	151 (86)	-
Total		**6753**	**938 (599)**	**1333 (739)**	**1.42 [1.31; 1.55]**

*IFs: Incidental findings; Lab - | MRI - = no laboratory abnormalities or disclosure of MRI IFs; Lab + | MRI - = laboratory abnormalities, no MRI IF disclosure, Lab - | MRI + = no laboratory abnormalities, disclosure of MRI IFs, Lab + | MRI + = laboratory abnormalities, disclosure of MRI IFs. Adapted from *Richter, Sierocinski 2020*.

Appendix G. Sensitivity Analysis #1: GEE results stratified for the subsample of participants without known malignancies (IRR).

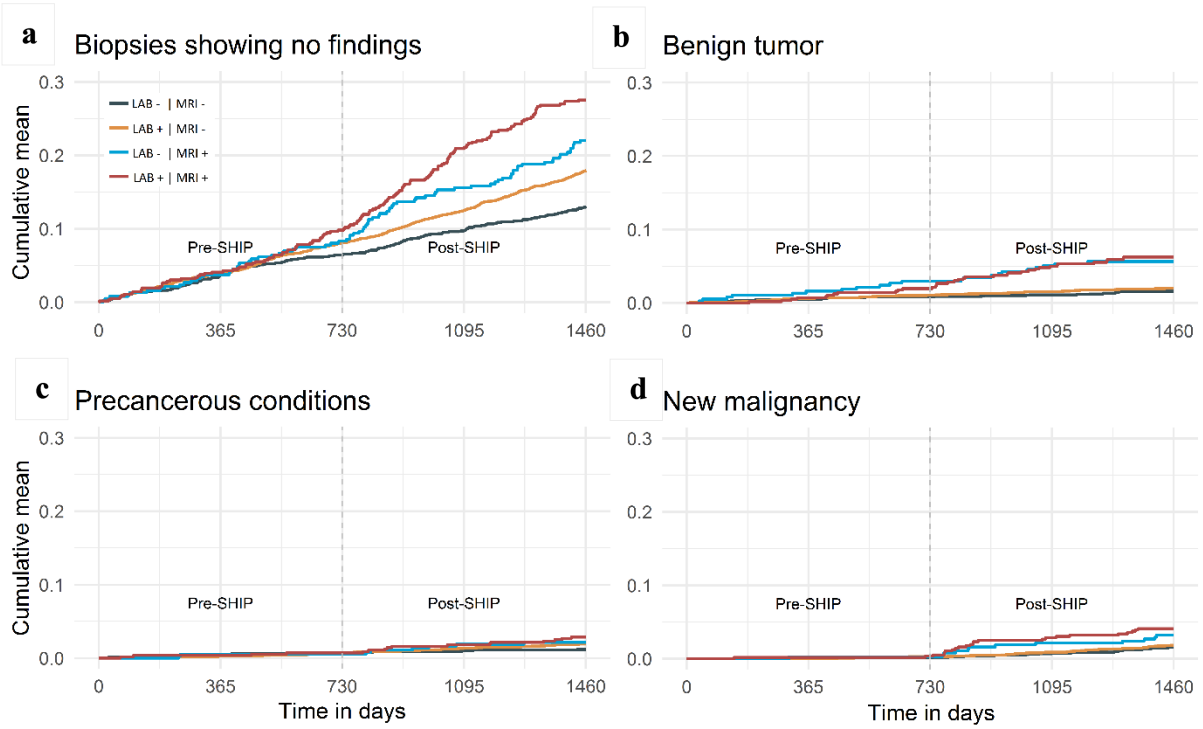
	Model 1		Model 2	
Predictors for biopsy reports	**IRR**	**95% CI**	**IRR**	**95% CI**
Age (per decade)	1.2	[1.13; 1.28]	1.21	[1.13; 1.29]
Sex (male vs. female)	0.70	[0.61; 0.82]	0.70	[0.61; 0.82]
Education (years, reference: 10y)				
<10y	1.02	[0.82; 1.25]	1.02	[0.83; 1.26]
>10y	1.03	[0.86; 1.24]	1.04	[0.87; 1.24]
Employed (yes vs. no)	0.93	[0.77; 1.12]	0.94	[0.77; 1.13]
Relationship status (reference: single)				
Married	1.01	[0.77; 1.34]	1.00	[0.76; 1.32]
Divorced	0.93	[0.63; 1.38]	0.91	[0.61; 1.35]
Widowed	1.05	[0.7; 1.59]	1.04	[0.69; 1.58]
Hospitalized in last 12 months (yes vs. no)	3.71	[3.2; 4.29]	3.70	[3.2; 4.29]
Known cancer history (yes vs. no)	NA	NA	NA	NA
Time-varying measures				
Disclosure of lab anomaly (yes vs. no)	1.38	[1.12; 1.69]	1.38	[1.12; 1.7]
Disclosure of MRI IF (yes vs. no)	2.23	[1.78; 2.79]	2.37	[1.87; 3]
Time (post-SHIP vs. pre-SHIP)	1.44	[1.18; 1.75]	1.45	[1.19; 1.76]

GEE with negative binomial distribution calculated by AR in n=6,167 participants (missing data in covariates; exclusion of those with known cancer diagnoses (n=586)).
Adapted from *Richter, Sierocinski et. al. 2020*, Online Appendix.

Appendix H. Sensitivity Analysis #1: Outcomes of biopsies stratified for disclosure of incidental findings in participants without known malignancies.

Outcome	**Combination of IFs**	**Strata size (Participants)**		**Pre SHIP N Biopsies (Participants)**	**Post SHIP N Biopsies (Participants)**	**Rate ratio biopsies [CI]**
		unselected	**selected**			
No malignancy or tumor	Lab - | MRI -	2046	**1883**	118 (100)	127 (108)	1.08 (0.84; 1.38)
	Lab + | MRI -	3685	**3353**	264 (212)	338 (248)	1.28 (1.09; 1.50)
	Lab - | MRI +	396	**372**	31 (29)	51 (41)	1.64 (1.06; 2.60)
	Lab + | MRI +	626	**559**	53 (40)	101 (74)	1.90 (1.37; 2.67)
Benign tumor	Lab - | MRI -	2046	**1883**	16 (16)	14 (11)	0.88 (0.42; 1.81)
	Lab + | MRI -	3685	**3353**	35 (33)	32 (32)	0.91 (0.56; 1.48)
	Lab - | MRI +	396	**372**	11 (10)	10 (10)	0.91 (0.38; 2.18)
	Lab + | MRI +	626	**559**	12 (10)	23 (19)	1.90 (0.96; 3.98)
Pre-cancerous lesion (including carcinoma in situ)	Lab - | MRI -	2046	**1883**	13 (11)	9 (8)	0.70 (0.28; 1.63)
	Lab + | MRI -	3685	**3353**	22 (20)	42 (36)	1.90 (1.15; 3.25)
	Lab - | MRI +	396	**372**	2 (1)	6 (6)	2.86 (0.63; 21.59)
	Lab + | MRI +	626	**559**	4 (4)	12 (10)	2.92 (1.00; 10.77)
Malignant process	Lab - | MRI -	2046	**1883**	3 (3)	26 (22)	8.28 (2.90; 36.04)
	Lab + | MRI -	3685	**3353**	9 (9)	51 (43)	5.58 (2.88; 12.21)
	Lab - | MRI +	396	**372**	1 (1)	11 (11)	9.74 (1.88; 240.04)
	Lab + | MRI +	626	**559**	2 (2)	21 (19)	9.82 (2.87; 66.48)

*IFs: Incidental findings; Lab - | MRI - = no laboratory abnormalities or disclosure of MRI IFs; Lab + | MRI - = laboratory abnormalities, no MRI IF disclosure, Lab - | MRI + = no laboratory abnormalities, disclosure of MRI IFs, Lab + | MRI + = laboratory abnormalities, disclosure of MRI IFs. Adapted from *Richter, Sierocinski 2020*



Appendix I. Sensitivity Analysis #1: Cumulative biopsy rates in participants without known malignancies (exclusion of n=586 participants).

Cumulative rates of biopsies identifying (a) exclusion of malignancy or benign tumor, (b) benign tumors, (c) pre-cancerous conditions, and (d) malignancies. Adapted from *Richter, Sierocinski et al. 2020.*

Appendix J. Coverage: MRI participation stratified by the distance of participants' residence to the SHIP examination center.

Distance to exam center	No MRI (N)	MRI (N)	Percentage of participation (%)
<10km	816	1005	55.19
10-19km	308	378	55.10
20-29km	379	282	42.66
30-39km	1289	1186	47.92
40-49km	284	216	43.20
50-89km	141	111	44.05
above 89km	165	193	53.91

Adapted from *Richter, Sierocinski et. al. 2020,* Online Appendix.

Appendix K. Coverage: Biopsy availability stratified by distance of participant residence to SHIP examination center.

Distance to SHIP	Histological data: No	Histological data: Yes	Percentage (%)	Weight	N	Prob
<10km	1421	400	21.97	1.00000000	1821	0.26965793
10-19km	549	137	19.97	1.10015023	686	0.10158448
20-29km	519	142	21.48	1.02281192	661	0.09788242
30-39km	2096	379	15.31	1.43500980	2475	0.36650378
40-49km	408	92	18.4	1.19402174	500	0.07404117
50-89km	229	23	9.13	2.40635268	252	0.03731675
above 89km	331	27	7.54	2.91379310	358	0.05301348

Adapted from *Richter, Sierocinski et. al. 2020,* Online Appendix.

Appendix L. Sensitivity Analysis #2: Results of sensitivity analyses for the coverage of biopsy reports.

	Weighted GEE			GEE restricted to participants of closest region		
	IRR	**LCL**	**UCL**	**IRR**	**LCL**	**UCL**
Intercept	0.06	0.04	0.09	0.05	0.02	0.10
Age (per decade)	1.14	1.07	1.22	1.18	1.05	1.32
Sex (males vs. females)	0.71	0.61	0.82	0.78	0.60	1.00
Education (years, reference: 10y)						
<10y	1.04	0.84	1.28	0.80	0.55	1.16
>10y	0.96	0.81	1.15	1.11	0.83	1.47
Relationship status (reference: single)						
Married	1.00	0.75	1.33	1.08	0.68	1.71
Divorced	0.86	0.59	1.25	0.96	0.52	1.74
Widowed	1.10	0.72	1.67	1.08	0.53	2.21
Work status (yes vs. no)	0.89	0.73	1.08	0.90	0.64	1.26
Hospitalization in last 12 month (yes vs. no)	3.36	2.92	3.86	3.25	2.56	4.13
Known cancer history (yes vs. no)	2.82	2.22	3.59	3.06	2.16	4.34
Disclosure of laboratory IFs (yes vs. no)	1.30	1.06	1.60	1.53	1.07	2.20
Disclosure of MRI IFs (yes vs. no)	1.94	1.57	2.41	2.32	1.62	3.31
Time (post SHIP vs. pre SHIP)	1.35	1.12	1.64	1.17	0.84	1.64

Adapted from *Richter, Sierocinski et. al. 2020*, Online Appendix.

11 Presentation of Results

1. **Richter, A***, **Sierocinski E*, Singer S, Bülow R, Hackmann C, Chenot JF, Schmidt CO** (2020). "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." European Journal of Epidemiology.
   equally contributed.
2. **Sierocinski E, Richter A, Singer S, Bülow R, Hackmann C, Chenot JF, Schmidt CO** (2020). *The potential of overtesting and overdiagnosis after disclosure of incidental findings from whole-body MRI in an observational cohort study* [Presentation]. DGEpi 15. Jahrestagung 2020; Greifswald, Germany. https://2020.dgepi.de/wp-content/uploads/2020/09/DGEpi2020_AbstractBooklet_29092020.pdf.

12 Deutsche Zusammenfassung

In bevölkerungsbezogenen Kohortenstudien werden Studienteilnehmer auf das Vorliegen von Krankheiten und Risikofaktoren untersucht. Dabei kommen zahlreiche Verfahren zur Anwendung, unter anderem bildgebende Verfahren und Laboruntersuchungen. Dies führt zur Entdeckung klinisch abnormer Befunde (Zufallsbefunde). Durch die Befundmitteilung an die Studienteilnehmer kann wiederrum eine Vielzahl weiterer, abklärender Diagnostik ausgelöst werden, was den natürlichen Verlauf von gesundheitsbezogenen Outcomes verändern kann.

Dieses Projekt untersucht die Auswirkung der Mitteilung von Zufallsbefunden aus einer Ganzkörper-MRT (gk-MRT) Untersuchung auf die Häufigkeit und Organsystemverteilung von Biopsien bei Probanden der Kohortenstudie „Study of Health in Pomerania“ (SHIP) und untersucht die Ergebnisse der Biopsien. Laboranomalitäten wurden berücksichtigt, weil sie eine mögliche Rolle in der klinischen Entscheidungsfindung zur Biopsie spielen.

Aus früheren Studien ist bekannt, dass klinisch potenziell relevante Zufallsbefunde aus dem MRT in 80% der Fälle einen möglichen Hinweis auf einen Tumor geben. Daher haben wir die Hypothese aufgestellt, dass die Befundmitteilung zu einer Zunahme in Biopsien führen würde. Nach dem aktuellen Stand der Literatur erweist sich die Mehrheit von Zufallsbefunden in Forschungs-MRT im Nachhinein allerdings als klinisch nicht bedeutsam; daher vermuteten wir, dass die Zunahme in Biopsien nicht zu einer klinisch relevanten Zunahme an diagnostizierten Malignitäten führen würde.

Insgesamt beobachteten wir eine deutliche Zunahme an Biopsien nach der SHIP Untersuchung. Die Mehrheit der Biopsien ergab keine Malignität; dies weist auf mögliche Überdiagnose und Übertestung als ungewollte Folgen der Mitteilung inzidenteller Befunde hin. Ein deutlicher Zuwachs in Biopsien, die maligne und prämaligne Erkrankungen diagnostizierten, konnte in Subgruppen von Probanden mit gk-MRT Befundmitteilungen jedoch festgestellt werden. Der therapeutische Effekt dieser Diagnosen war kein Gegenstand dieser Arbeit. Daher kann das Vorliegen von Überdiagnosen nicht ausgeschlossen werden.

Basierend auf den Ergebnissen werden strengere Rückmeldungsrichtlinien von Zufallsbefunden in der Forschung empfohlen, um Probanden vor unnötigen invasiven Eingriffen zu schützen und um die Verzerrung von Beobachtungsdaten zu reduzieren. Weitere Studien zur Beurteilung der langfristigen Morbidität und Mortalität der Probanden nach der Befundmitteilung sind notwendig, um den Nutzen für die Probanden genauer zu evaluieren.

13 Works Cited

1. Kindig D, Stoddart G. What Is Population Health? American Journal of Public Health. 2003;93(3):380-3.
2. Rose G. Sick individuals and sick populations. Int J Epidemiol. 1985;14(1):32-8.
3. Volzke H, Alte D, Schmidt CO, Radke D, Lorbeer R, Friedrich N, Aumann N, Lau K, Piontek M, Born G, Havemann C, Ittermann T, Schipf S, Haring R, Baumeister SE, Wallaschofski H, Nauck M, Frick S, Arnold A, Junger M, Mayerle J, Kraft M, Lerch MM, Dorr M, Reffelmann T, Empen K, Felix SB, Obst A, Koch B, Glaser S, Ewert R, Fietze I, Penzel T, Doren M, Rathmann W, Haerting J, Hannemann M, Ropcke J, Schminke U, Jurgens C, Tost F, Rettig R, Kors JA, Ungerer S, Hegenscheid K, Kuhn JP, Kuhn J, Hosten N, Puls R, Henke J, Gloger O, Teumer A, Homuth G, Volker U, Schwahn C, Holtfreter B, Polzer I, Kohlmann T, Grabe HJ, Rosskopf D, Kroemer HK, Kocher T, Biffar R, John U, Hoffmann W. Cohort profile: the study of health in Pomerania. Int J Epidemiol. 2011;40(2):294-307.
4. Tsao CW, Vasan RS. Cohort Profile: The Framingham Heart Study (FHS): overview of milestones in cardiovascular epidemiology. Int J Epidemiol. 2015;44(6):1800-13.
5. Ikram MA, Brusselle GGO, Murad SD, van Duijn CM, Franco OH, Goedegebure A, Klaver CCW, Nijsten TEC, Peeters RP, Stricker BH, Tiemeier H, Uitterlinden AG, Vernooij MW, Hofman A. The Rotterdam Study: 2018 update on objectives, design and main results. Eur J Epidemiol. 2017;32(9):807-50.
6. German National Cohort (GNC) Consortium. The German National Cohort: aims, study design and organization. Eur J Epidemiol. 2014;29(5):371-82.
7. Hegenscheid K, Kuhn JP, Volzke H, Biffar R, Hosten N, Puls R. Whole-body magnetic resonance imaging of healthy volunteers: pilot study results from the population-based SHIP study. Rofo. 2009;181(8):748-59.
8. Volzke H, Schmidt CO, Hegenscheid K, Kuhn JP, Bamberg F, Lieb W, Kroemer HK, Hosten N, Puls R. Population imaging as valuable tool for personalized medicine. Clin Pharmacol Ther. 2012;92(4):422-4.
9. Carlson MDA, Morrison RS. Study design, precision, and validity in observational studies. J Palliat Med. 2009;12(1):77-82.
10. Schmücker R. Incidental Findings: Definition of the Concept. In: Weckbach S, editor. Incidental Radiological Findings. Cham: Springer International Publishing; 2017. p. 3-7.
11. Knoppers BM, Deschenes M, Zawati MH, Tasse AM. Population studies: return of research results and incidental findings Policy Statement. Eur J Hum Genet. 2013;21(3):245-7.

12. Booth TC, Waldman AD, Wardlaw JM, Taylor SA, Jackson A. Management of incidental findings during imaging research in "healthy" volunteers: current UK practice. Br J Radiol. 2012;85(1009):11-21.
13. Erdmann P. Incidental Findings – Ethical Aspects. In: Weckbach S, editor. Incidental Radiological Findings. Medical Radiology. Cham: Springer International Publishing; 2016. p. 9-24.
14. Ganguli I, Simpkin AL, Lupo C, Weissman A, Mainor AJ, Orav EJ, Rosenthal MB, Colla CH, Sequist TD. Cascades of Care After Incidental Findings in a US National Survey of Physicians. JAMA Netw Open. 2019;2(10):e1913325.
15. Pinato DJ, Stavraka C, Tanner M, Esson A, Jacobson EW, Wilkins MR, Libri V. Clinical, ethical and financial implications of incidental imaging findings: experience from a phase I trial in healthy elderly volunteers. PLoS One. 2012;7(11):e49814.
16. Pandharipande PV, Herts BR, Gore RM, Mayo-Smith WW, Harvey HB, Megibow AJ, Berland LL. Rethinking Normal: Benefits and Risks of Not Reporting Harmless Incidental Findings. J Am Coll Radiol. 2016;13(7):764-7.
17. American College of Radiology Incidental Findings Committee. Publications by the ACR IF Committee. Accessed on 01.01.2021. https://publish.smartsheet.com/42d18e874a164318a0f702481f2fbb70.
18. Hoang JK. Insignificant Findings: Don’t Leave Them Questioning the Significance. Journal of the American College of Radiology. 2015;12(12, Part A):1244.
19. Berlin L. Medicolegal Aspects and Informed Consent. In: Weckbach S, editor. Incidental Radiological Findings. Cham: Springer International Publishing; 2017. p. 25-34.
20. Weckbach S. Incidental Radiological Findings: Springer International Publishing; 2017. 209 p.
21. Booth TC, Jackson A, Wardlaw JM, Taylor SA, Waldman AD. Incidental findings found in "healthy" volunteers during imaging performed for research: current legal and ethical implications. Br J Radiol. 2010;83(990):456-65.
22. Sandeman EM, Hernandez Mdel C, Morris Z, Bastin ME, Murray C, Gow AJ, Corley J, Henderson R, Deary IJ, Starr JM, Wardlaw JM. Incidental findings on brain MR imaging in older community-dwelling subjects are common but serious medical consequences are rare: a cohort study. PLoS One. 2013;8(8):e71467.
23. Wardlaw JM, Davies H, Booth TC, Laurie G, Compston A, Freeman C, Leach MO, Waldman AD, Lomas DJ, Kessler K, Crabbe F, Jackson A. Acting on incidental findings in research imaging. BMJ. 2015;351:h5190.

24. Hegenscheid K, Seipel R, Schmidt CO, Volzke H, Kuhn JP, Biffar R, Kroemer HK, Hosten N, Puls R. Potentially relevant incidental findings on research whole-body MRI in the general adult population: frequencies and management. Eur Radiol. 2013;23(3):816-26.
25. O'Sullivan JW, Muntinga T, Grigg S, Ioannidis JPA. Prevalence and outcomes of incidental imaging findings: umbrella review. BMJ. 2018;361:k2387.
26. Gibson LM, Paul L, Chappell FM, Macleod M, Whiteley WN, Salman RA, Wardlaw JM, Sudlow CLM. Potentially serious incidental findings on brain and body magnetic resonance imaging of apparently asymptomatic adults: systematic review and meta-analysis. BMJ. 2018;363:k4577.
27. Prabhu V, Chhor CM, Ego-Osuala IO, Xiao JM, Hindman NM, Rosenkrantz AB. Frequency and Outcomes of Incidental Breast Lesions Detected on Abdominal MRI Over a 7-Year Period. AJR Am J Roentgenol. 2017;208(1):107-13.
28. Gibson LM, Nolan J, Littlejohns TJ, Mathieu E, Garratt S, Doherty N, Petersen S, Harvey NCW, Sellors J, Allen NE, Wardlaw JM, Jackson CA, Sudlow CLM. Factors associated with potentially serious incidental findings and with serious final diagnoses on multi-modal imaging in the UK Biobank Imaging Study: A prospective cohort study. PLoS One. 2019;14(6):e0218267.
29. Bunnik EM, Vernooij MW. Incidental findings in population imaging revisited. Eur J Epidemiol. 2016;31(1):1-4.
30. Hiom SC. Diagnosing cancer earlier: reviewing the evidence for improving cancer survival. Br J Cancer. 2015;112 Suppl 1:S1-5.
31. Allemani C, Weir HK, Carreira H, Harewood R, Spika D, Wang X-S, Bannon F, Ahn JV, Johnson CJ, Bonaventure A, Marcos-Gragera R, Stiller C, Azevedo e Silva G, Chen W-Q, Ogunbiyi OJ, Rachet B, Soeberg MJ, You H, Matsuda T, Bielska-Lasota M, Storm H, Tucker TC, Coleman MP. Global surveillance of cancer survival 1995–2009: analysis of individual data for 25 676 887 patients from 279 population-based registries in 67 countries (CONCORD-2). The Lancet. 2015;385(9972):977-1010.
32. Orme NM, Fletcher JG, Siddiki HA, Harmsen WS, O'Byrne MM, Port JD, Tremaine WJ, Pitot HC, McFarland EG, Robinson ME, Koenig BA, King BF, Wolf SM. Incidental findings in imaging research: evaluating incidence, benefit, and burden. Arch Intern Med. 2010;170(17):1525-32.
33. Brodersen J, Kramer BS, Macdonald H, Schwartz LM, Woloshin S. Focusing on overdiagnosis as a driver of too much medicine. BMJ. 2018;362:k3494.
34. Schmidt CO, Sierocinski E, Hegenscheid K, Baumeister SE, Grabe HJ, Volzke H. Impact of whole-body MRI in a general population study. Eur J Epidemiol. 2016;31(1):31-9.

35. Ganguli I, Simpkin AL, Colla CH, Weissman A, Mainor AJ, Rosenthal MB, Sequist TD. Why Do Physicians Pursue Cascades of Care After Incidental Findings? A National Survey. J Gen Intern Med. 2019 Oct 2;2(10):e1913325.
36. Booth TC, Boyd-Ellison JM. The current impact of incidental findings found during neuroimaging on neurologists' workloads. PLoS One. 2015;10(2):e0118155.
37. Shyamala K, Girish HC, Murgod S. Risk of tumor cell seeding through biopsy and aspiration cytology. J Int Soc Prev Community Dent. 2014;4(1):5-11.
38. Leveridge MJ, Finelli A, Kachura JR, Evans A, Chung H, Shiff DA, Fernandes K, Jewett MA. Outcomes of small renal mass needle core biopsy, nondiagnostic percutaneous biopsy, and the role of repeat biopsy. Eur Urol. 2011;60(3):578-84.
39. Ghadian A. Concerns about renal mass biopsy. Can Urol Assoc J. 2014;8(7-8):230.
40. Yun S, Kang H, Park S, Kim BS, Park JG, Jung MJ. Diagnostic accuracy and complications of CT-guided core needle lung biopsy of solid and part-solid lesions. Br J Radiol. 2018;91(1088):20170946.
41. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. Int J Surg. 2014;12(12):1495-9.
42. Schmidt CO, Alte D, Volzke H, Sauer S, Friedrich N, Valliant R. Partial misspecification of survey design features sufficed to severely bias estimates of health-related outcomes. J Clin Epidemiol. 2011;64(4):416-23.
43. Richter A, Sierocinski E, Singer S, Bülow R, Hackmann C, Chenot J-F, Schmidt CO. 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. European Journal of Epidemiology. 2020;35(10):925-35.
44. Greenwood HI, Freimanis RI, Carpentier BM, Joe BN. Clinical Breast Magnetic Resonance Imaging: Technique, Indications, and Future Applications. Seminars in ultrasound, CT, and MR. 2018;39(1):45-59.
45. Balleyguier C, Ayadi S, Van Nguyen K, Vanel D, Dromain C, Sigal R. BIRADS™ classification in mammography. European Journal of Radiology. 2007;61(2):192-4.
46. Cieszanowski A, Maj E, Kulisiewicz P, Grudzinski IP, Jakoniuk-Glodala K, Chlipala-Nitek I, Kaczynski B, Rowinski O. Non-contrast-enhanced whole-body magnetic resonance imaging in the general population: the incidence of abnormal findings in patients 50 years old and younger compared to older subjects. PLoS One. 2014;9(9):e107840.

47. Ulus S, Suleyman E, Ozcan UA, Karaarslan E. Whole-Body MRI Screening in Asymptomatic Subjects; Preliminary Experience and Long-Term Follow-Up Findings. Pol J Radiol. 2016;81:407-14.
48. Brodersen J, Schwartz LM, Heneghan C, O'Sullivan JW, Aronson JK, Woloshin S. Overdiagnosis: what it is and what it isn't. BMJ Evid Based Med. 2018;23(1):1-3.
49. Schmidt CO, Hegenscheid K, Erdmann P, Kohlmann T, Langanke M, Volzke H, Puls R, Assel H, Biffar R, Grabe HJ. Psychosocial consequences and severity of disclosed incidental findings from whole-body MRI in a general population study. Eur Radiol. 2013;23(5):1343-51.
50. Vento S, Cainelli F, Vallone A. Defensive medicine: It is time to finally slow down an epidemic. World J Clin Cases. 2018;6(11):406-9.
51. Dhaliwal K, Malkhasyan V, Elhassan M. In with acute bronchitis; out with duodenal perforation: the potentially harmful cascade of over-testing. A case report. J Community Hosp Intern Med Perspect. 2018;8(1):26-8.
52. Hendee WR, Becker GJ, Borgstede JP, Bosma J, Casarella WJ, Erickson BA, Maynard CD, Thrall JH, Wallner PE. Addressing overutilization in medical imaging. Radiology. 2010;257(1):240-5.
53. Hayward R. VOMIT (victims of modern imaging technology)—an acronym for our times. BMJ. 2003;326(7401):1273.
54. Booth CM, Tannock IF. Randomised controlled trials and population-based observational research: partners in the evolution of medical evidence. Br J Cancer. 2014;110(3):551-5.
55. Richter A, Sierocinski E, Singer S, Bülow R, Hackmann C, Chenot J-F, Schmidt CO. 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. European Journal of Epidemiology. 2020;35(10):925-35.

14 Attachments

14.1 Scientific publications

European Journal of Epidemiology (2020) 35:925–935
https://doi.org/10.1007/s10654-020-00679-4

POPULATION IMAGING

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

Adrian Richter[1] · Elizabeth Sierocinski[1,2] · Stephan Singer[3,4] · Robin Bülow[5] · Carolin Hackmann[1,6] · Jean-François Chenot[2] · Carsten Oliver Schmidt[1]

Received: 3 April 2020 / Accepted: 19 August 2020 / Published online: 29 August 2020

Abstract
Magnetic resonance imaging (MRI) yields numerous tumor-related incidental findings (IFs) which may trigger diagnostics such as biopsies. To clarify these effects, we studied how whole-body MRI IF disclosure in a population-based cohort affected biopsy frequency and the detection of malignancies. Laboratory disclosures were also assessed. Data from 6753 participants in the Study of Health in Pomerania (SHIP) examined between 2008 and 2012 were utilized. All underwent laboratory examinations and 3371 (49.9%) a whole-body MRI. Electronic biopsy reports from 2002 to 2017 were linked to participants and assigned to outcome categories. Biopsy frequency 2 years pre- and post-SHIP was investigated using generalized estimating equations with a negative-binomial distribution. Overall 8208 IFs (laboratory findings outside reference limits: 6839; MRI: 1369) were disclosed to 4707 participants; 2271 biopsy reports belonged to 1200 participants (17.8%). Of these, 938 biopsies occurred pre-SHIP; 1333 post-SHIP (event rate/100 observation years = 6.9 [95% CI 6.5; 7.4]; 9.9 [9.3; 10.4]). Age, cancer history, recent hospitalization, female sex, and IF disclosure were associated with higher biopsy rates. Nonmalignant biopsy results increased more in participants with disclosures (post-/pre-SHIP rate ratio 1.39 [95% CI 1.22; 1.58]) than without (1.09 [95% CI 0.85; 1.38]). Malignant biopsy results were more frequent post-SHIP (rate ratio 1.74 [95% CI 1.27; 2.42]). Biopsies increased after participation in a population-based cohort study with MRI and laboratory IF disclosure. Most biopsies resulted in no findings and few malignancies were diagnosed, indicating potential overtesting and overdiagnosis. A more restrictive policy regarding IF disclosure from research findings is required.

Keywords Magnetic resonance imaging · Incidental findings · Biopsies · Histological examinations · Record linkage

Introduction

The challenge of managing incidental findings (IFs) in clinical practice and research is growing with the increasing accessibility of powerful imaging modalities such as MRI [1–3]. IFs frequently occur in clinical as well as in research settings [1, 2, 4]. The use of whole-body MRI (wb-MRI) in a population-based cohort of 2500 participants resulted in 13,455 IFs, of which 1330 were potentially clinically relevant and disclosed to participants [1].

Between 50 and 80% of IFs from research MRI are suspicious for malignancy [1, 3]. Such findings may enable timely treatment of a disease, offering potential improvement or preservation of quality and length of life [5, 6]. On the other hand, overtesting and overdiagnosis may result [7], incurring additional costs to the health care system as well as psychosocial costs for patients who anxiously await results or are faced with findings of unknown relevance [8, 9].

Despite the frequency of these potentially significant findings, information regarding the clinical outcomes of IFs is limited [3]. There exists uncertainty as to which IFs warrant disclosure and further investigation. Management guidelines are lacking [3, 10, 11]. Research participants with disclosed IFs may present to their physicians and invasive

Adrian Richter and Elizabeth Sierocinski have contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10654-020-00679-4) contains supplementary material, which is available to authorized users.

✉ Adrian Richter
adrian.richter@uni-greifswald.de

Extended author information available on the last page of the article

diagnostic actions such as biopsies may result to rule out serious pathology. Histological examinations of biopsied tissue are the most effective but also the most invasive way to obtain diagnostic certainty. Individuals undergoing biopsy are exposed to potential discomfort, pain, and complications such as infection, bleeding, damage to nearby structures, or tumor seeding [12–16]. Other adverse consequences include potential costs, side effects, and complications of clinically unnecessary therapeutic interventions resulting from overdiagnosis [7, 17]. Such consequences are particularly undesirable in the context of observational research designs as they may bias longitudinal study data of health related outcomes.

To the best of our knowledge, there exists no longitudinal data on the effects of IFs on the frequency and outcomes of biopsies. The aim of this study was therefore twofold. First, we assessed whether the disclosure of incidental wb-MRI findings in a population-based cohort study was associated with an increase in biopsy frequency. The number of biopsies during the 2 years prior to examination in a large population-based cohort study was compared to the number 2 years after participation, adjusted for socio-demographic and clinical characteristics. Second, we analyzed whether the disclosure of IFs contributed to the detection of new malignancies via biopsy. We accounted for the concomitant effects of disclosed laboratory results because of their potential role in the decision to biopsy.

Methods

Data sources

This study uses data from the Study of Health in Pomerania (SHIP) and histology data from the Greifswald University Medical Center Department of Pathology. SHIP design and methods are described in detail in other publications [18]; a summary is provided below.

SHIP cohorts

SHIP is a population-based project consisting of two independent cohorts, SHIP and SHIP-TREND. Participants were selected from northeastern Germany [18]. Out of 6265 eligible individuals of the first cohort, 4308 (2192 women) participated (response 68.8%) in the SHIP-0 baseline examination [19]. Baseline examinations were performed from 1997 to 2001. 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-0, was established in 2008. A stratified sample of 10,000 was drawn from the central population registry. Out of the net sample of 8826, after exclusion of deceased and relocated participants, 4420 (2275 women) participated (response 50.1%) in the baseline examination between 2008 and 2012.

All analyses are based on data from the SHIP-2 and SHIP-Trend-0 examinations, which were conducted in parallel and included among others extensive laboratory investigations, a personal interview about medical history, socio-demographics, and a whole-body MRI. The full scope of examinations is described elsewhere [18].

All SHIP-2 and SHIP Trend participants were invited to take part in a whole-body MRI examination, with the exception of 527 (SHIP-2: n = 57, SHIP-TREND-0: n = 467) who received a brief examination in remote examination centres. For the latter no appointment for MRI was arranged. A total of 3371 individuals participated in the MRI examination and 3382 did not. A detailed overview is provided in the flow-chart (Fig. 1).

Whole-body MRI

A 1·5-Tesla system (Magnetom Avanto; Siemens Medical Solutions, Erlangen, Germany) was used for wb-MRI. The wb-MRI protocol was identical for all participants and included a plain whole-body MRI with 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 [1, 20].

Abnormal findings and anatomical variants were documented in a standardized reading protocol. The radiologists reading the scans had no access to 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 with 15 years of experience, resolved disagreements [1].

Laboratory examination

Venous blood and urine samples were taken from all study participants. Serum aliquots were stored at − 80 °C. The laboratory in charge for SHIP blood samples takes part in the official German external quality proficiency testing program. All assays are calibrated against the international reference preparations. A list of all covered biomarkers is provided elsewhere [18].

Parameters with the potential to trigger a biopsy were chosen on clinical grounds by JFC and ES: alanine transaminase (ALT), aspartate transaminase (AST), gamma-glutamyl transpeptidase (GGT), lipase, serum white blood cells (WBC), serum platelets (PLT), thyroid stimulating hormone (TSH), and urine erythrocyte count. Parameters exceeding

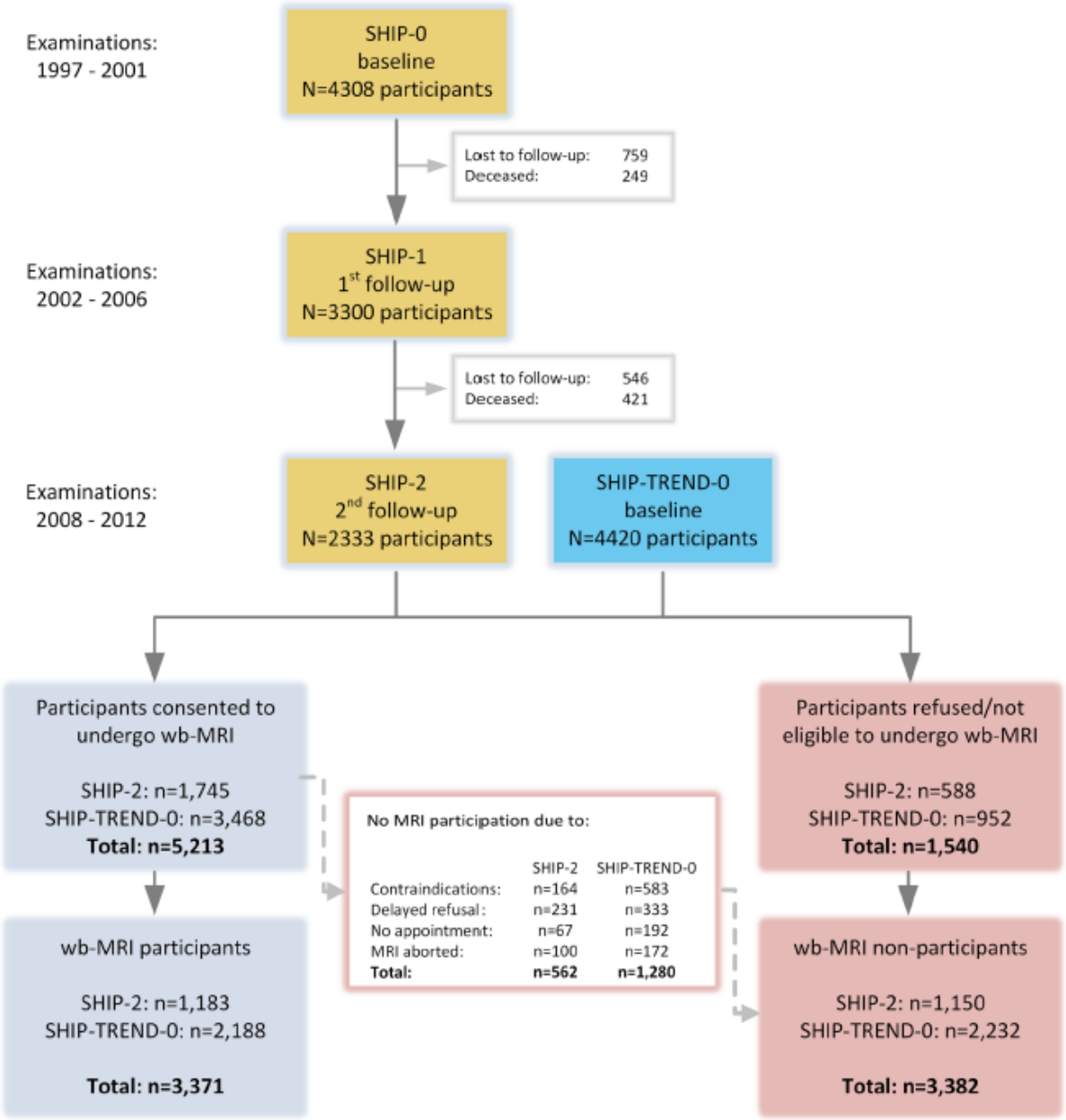


Fig. 1 Study flow chart

reference limits were summarized in a variable indicating whether one or more laboratory results exceeded the limits.

Disclosure of incidental findings from imaging and laboratory

A standardized protocol, approved by the institutional review board, regulated the handling of wb-MRI incidental findings. Findings were classified into three categories by trained radiologists: Category I comprised medically non-significant findings in asymptomatic individuals (e.g., anatomical variants, old brain infarcts). Category II findings were abnormalities needing further non-urgent medical evaluation (e.g., tumors or nodules of unclear significance). Category III included urgent findings requiring immediate referral (e.g., acute brain infarcts, fractures, lobar pneumonia). Category II

findings were disclosed to participants by means of a postal letter after approval by an interdisciplinary advisory board. This included a recommendation to seek further medical assistance. Category III findings were disclosed immediately to the participant after the examination. A more detailed description of the process and the frequencies of category II and III findings has been provided elsewhere [1, 20].

All participants received a paper copy of their laboratory results. Laboratory values crossing reference limits were highlighted [18].

Linkage of histological data with SHIP data

Data from the pathology department and the SHIP study have no common key for linkage. Therefore we applied record linkage as discussed in Vatsalan et al. [21] based on: last name, first name, date of birth, and sex given the consent of participants. The linkage process included a normalization of personal data (to upper case letters, removal of special characters), the indexing of candidate pairs based on birth date, and the comparison. Candidate pairs were compared using generalized Levenshtein distances [22]. Respective R-Code was parallelized and blocked via birth date to reduce computational costs [23, 24]. In total, 3340 SHIP participants were linked based on a Levenshtein distance of zero. In 422 participants with a Levenshtein distance > 0, a manual revision for transposition, omission, or addition of single characters in their personal data identified additional 149 matches. Out of the 3489 participants with biopsy records, 1200 had at least one biopsy report in the analysis period (±2 years from SHIP examination).

Classification of biopsy reports

From the database of the Greifswald University Medical Center Department of Pathology a total of 8576 histological reports from SHIP participants were available, dating from 2002 to 2019. The histological reports were unstructured and available in free-text format. They contained varying levels of detail regarding clinical history, macro- and microscopy, and differential/excluded/final diagnosis.

Three (0.1%) autopsies were excluded and 56 (1.8%) reports were not classifiable due to missing or incomplete data. Twenty-five reports (0.8%) included two different tissue types and were assigned to two outcome categories. In total, 3011 biopsy reports were included and classified. Of these, 2271 were dated within the analysis period, i.e., ±2 years to the SHIP examination, and 740 represented earlier biopsy reports from patients with at least one report within the analysis period. Reports from the pre-analysis period were included when a report within the analysis period occurred. This ensured the correct classification of 1st and 2nd malignancies as well as follow-up biopsies.

The outcome or diagnosis detailed in each biopsy report was classified into mutually exclusive categories: pre-cancer; 1st to 5th malignancy; metastasis; benign tumor; follow-up of known malignant or suspicious process; no tumor or malignancy; updated pathological report; no diagnosis possible.

ES classified the entire set of biopsies. Double-readings were conducted for 2510 (83.4%) biopsies (CH: n = 1752, JFC: n = 758). Dissent was resolved by consensus readings and by consulting a pathologist (StS). Overall, dissent between readers was observed in 239 of 2510 double-classified reports (9.5%). In total, 58 corrections of initial classifications done by ES (2.3% of 2510 double readings) were revoked by the consensus decision. Therefore, we assume a misclassification rate of lower than 3% in the 501 single-readings.

Statistical analyses

Baseline characteristics were stratified for the respective cohort and participation of MRI examination. Descriptive measures for location (mean, median) are shown with standard deviation (SD), minimum, and maximum. Categorical data was described using the number of events (N) and percentage. Missing values for all variables are provided.

Crude event rates (per 100 observation years) of biopsies were calculated using median unbiased estimates and exact mid-p 95% confidence intervals [25, 26]. The observation period covered 2 years before and after the SHIP examination. In addition, the frequency of biopsies (count data) was estimated using a generalized estimating equation model with a negative binomial distribution. Based on the QIC statistic we chose an independent working correlation [27]. Coefficients of the model were exponentiated for interpretation as incidence rate ratios [28]. Variable selection was based on clinical appraisal of relevance a priori. Age, sex, socio-demographic characteristics (education, relationship status), hospitalization within last 12 months, history of cancer, and the disclosure of incidental findings were included in the model. Due to the small amount of missing values (max. 1.9% for history of cancer, Table 1) no imputation techniques were applied and complete case analyses were conducted.

The cumulative rates of biopsies were investigated using recurrent event analysis. Biopsy outcomes such as "no malignant finding" or "follow up of known malignant process" could occur multiple times for each participant. The 1st to 5th malignancies as well as metastases were summarized in one category of recurrent events of the same entity. The cumulative rate of recurrent events was represented graphically using the R-package reReg [29]. The R-package reReg allows for the computation of cumulative hazard rate

Table 1 Study population characteristics

Characteristics	SHIP-BASE (MRI)	SHIP-BASE (no MRI)	SHIP-TREND (MRI)	SHIP-TREND (no MRI)	All
N	1183	1150	2188	2232	6753
Age (years)					
Mean (SD)	55.7 (12.8)	59.0 (14.3)	51.2 (14.1)	52.7 (16.7)	53.8 (15.1)
Median [Min, Max]	56.0 [30.0, 90.0]	59.0 [31.0, 93.0]	52.0 [21.0, 82.0]	54.0 [20.0, 84.0]	54.0 [20.0, 93.0]
Sex					
Female	605 (51.1%)	630 (54.8%)	1113 (50.9%)	1162 (52.1%)	3510 (52.0%)
Male	578 (48.9%)	520 (45.2%)	1075 (49.1%)	1070 (47.9%)	3243 (48.0%)
Educational level					
Normal/high	856 (72.4%)	685 (59.6%)	1847 (84.4%)	1556 (69.7%)	4944 (73.2%)
Lower	327 (27.6%)	463 (40.3%)	337 (15.4%)	667 (29.9%)	1794 (26.6%)
Missing	0 (0%)	2 (0.2%)	4 (0.2%)	9 (0.4%)	15 (0.2%)
Years of education					
< 10	253 (21.4%)	390 (33.9%)	344 (15.7%)	685 (30.7%)	1672 (24.8%)
10	652 (55.1%)	562 (48.9%)	1178 (53.8%)	1090 (48.8%)	3482 (51.6%)
> 10	278 (23.5%)	193 (16.8%)	662 (30.3%)	448 (20.1%)	1581 (23.4%)
Missing	0 (0%)	5 (0.4%)	4 (0.2%)	9 (0.4%)	18 (0.3%)
Marital status					
Single	129 (10.9%)	114 (9.9%)	223 (10.2%)	262 (11.7%)	728 (10.8%)
In a relationship	965 (81.6%)	900 (78.3%)	1755 (80.2%)	1663 (74.5%)	5283 (78.2%)
Divorced	55 (4.6%)	76 (6.6%)	128 (5.9%)	155 (6.9%)	414 (6.1%)
Widowed	34 (2.9%)	57 (5.0%)	78 (3.6%)	143 (6.4%)	312 (4.6%)
Missing	0 (0%)	3 (0.3%)	4 (0.2%)	9 (0.4%)	16 (0.2%)
Employment status					
Unemployed	523 (44.2%)	649 (56.4%)	913 (41.7%)	1240 (55.6%)	3325 (49.2%)
Employed	658 (55.6%)	497 (43.2%)	1271 (58.1%)	980 (43.9%)	3406 (50.4%)
Missing	2 (0.2%)	4 (0.3%)	4 (0.2%)	12 (0.5%)	22 (0.3%)
Hospitalized in last 12 months					
No	1014 (85.7%)	919 (79.9%)	1905 (87.1%)	1864 (83.5%)	5702 (84.4%)
Yes	166 (14.0%)	227 (19.7%)	280 (12.8%)	360 (16.1%)	1033 (15.3%)
Missing	3 (0.3%)	4 (0.3%)	3 (0.1%)	8 (0.4%)	18 (0.3%)
Cancer history					
No	1092 (92.3%)	1071 (93.1%)	2047 (93.6%)	1957 (87.7%)	6167 (91.3%)
Yes	89 (7.5%)	78 (6.8%)	135 (6.2%)	153 (6.9%)	455 (6.7%)
Missing	2 (0.2%)	1 (0.1%)	6 (0.3%)	122 (5.5%)	131 (1.9%)
Histological data available					
No	979 (82.8%)	940 (81.7%)	1808 (82.6%)	1826 (81.8%)	5553 (82.2%)
Yes	204 (17.2%)	210 (18.3%)	380 (17.4%)	406 (18.2%)	1200 (17.8%)
Laboratory abnormalities					
No	492 (41.6%)	382 (33.2%)	832 (38.0%)	736 (33.0%)	2442 (36.2%)
Yes	691 (58.4%)	768 (66.8%)	1356 (62.0%)	1496 (67.0%)	4311 (63.8%)

using the Nelson-Aalen estimator [30] without adjusting the risk set in case of an event.

Record linkage and the analysis of recurrent events were conducted using the statistical software R [31]. For data preprocessing and the GEE models SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) was used. We present all effects including confidence intervals [32, 33]. Confidence intervals and effect sizes provide orientation for the interpretation of effects in terms of clinical or epidemiological relevance.

Sensitivity analyses

In the first sensitivity analysis, we restricted the study population to participants with no known malignant diseases at the time of the SHIP examination. Participants in which this

information was unknown were also excluded. The interview in which this information was reported by SHIP participants antedates any disclosure of IFs. This analysis was conducted to control for the impact of known malignant conditions on the number of biopsies after SHIP participation.

In the second sensitivity analysis, we examined the coverage of biopsies based on participant place of residence. The University Department of Pathology is the main biopsy provider in the region, covering all inpatient and a portion of outpatient biopsies. Nevertheless, biopsies may be sent to private laboratories in the region, particularly for patients living far away from Greifswald University Medical Center. Similarly, the decision to participate in wb-MRI may have depended on the distance of participants' residence to the examination center. Both of these potentially limiting factors in our study's coverage were examined using stratification for distance to the examination center.

Results

Participant characteristics

The descriptive characteristics of participants (N = 6753) are summarized in Table 1 stratified by cohort and MRI participation. Participants who underwent MRI were younger compared to non-participants, had a higher education and employment status, and were more frequently in a relationship. The percentage of biopsy reports was similar across all four strata and ranged from 17.2 to 18.3%. Disclosure of abnormal laboratory findings was more frequent in MRI non-participants. In total, 1022 out of 3371 participants (30.3%) who underwent MRI examination (SHIP-2: n = 362, SHIP-TREND-0: n = 660) received disclosure of an incidental MRI finding. Of these 851 (83.3%) were suspected to be tumors.

Frequency of biopsies

Within ± 2 years to SHIP there were 2271 histological reports belonging to 1200 participants. The reports were dated from 2006 to 2012. Of these, 938 biopsy reports (599 participants) occurred in the 2 years prior to the SHIP examination and 1333 (739 participants) occurred during the 2 years after SHIP. A detailed list of biopsies and outcomes is shown in Table A of the Online Appendix. These numbers correspond to an event rate of biopsies of 6.9 [95% CI 6.5; 7.4] pre-SHIP and 9.9 [9.3; 10.4] post-SHIP (rate per 100 observation years).

For 601 participants (50.0%), the first recorded biopsy was observed after participation in SHIP; i.e., no biopsy had been recorded in the Department of Pathology between 2002 and the date of the SHIP examination. Of the 1333 biopsies found in the 2 years after participation in SHIP, 1041 (78.1%) belonged to these 601 participants.

Factors associated with a higher number of biopsies included higher age, female sex, hospitalization (within 12 months of SHIP examination), history of cancer, and the disclosure of IFs (laboratory and MRI), as well as time (Table 2). No or only minor effects were associated with education, relationship status and employment status. In a model stratified for sex the IRRs were comparable (Online Appendix, Table B). In model 2 only IFs from MRI with tumor relevance (Yes vs. No) were considered. This analysis shows a small increase of the effect size for MRI IFs.

Biopsy outcomes

Both before and after SHIP examination, the majority of biopsies resulted in the outcome "no malignancy or tumor" (Table 3). The largest absolute increase was found in this category. Biopsies diagnosing a malignancy increased irrespective of the disclosure of IFs in all strata, but rate ratios were highest in case of disclosed MRI findings. Participants with disclosed laboratory or MRI IFs received more biopsies in all outcome categories except benign tumors.

Cumulative rate of selected outcomes

Across all biopsy outcome categories, participants with disclosed MRI findings consistently showed the greatest increase in cumulative biopsy rates post-SHIP. Participants who received a combination of laboratory and MRI IFs were more likely to undergo a biopsy resulting in the diagnosis of a malignancy (Fig. 2).

Sensitivity analysis

In the first set of sensitivity analyses, we excluded n = 455 participants with known malignant disease and n = 131 participants in which this information was unknown. The impact of IF disclosure was found to be higher than in the results obtained in the unselected study population (Online Appendix, Table C, D and Figure A).

Regarding coverage, MRI participation decreased slightly but inconsistently with increasing distance of participants' residence to the examination center (55–43%, Online Appendix: Table E) and has therefore not been considered in separate analysis. Whereas biopsy data from the Department of Pathology was available for 22% of participants living near the examination site, only 7.5% of participants living 40–89 km away had available biopsy data (Online Appendix, Table F).

We applied two approaches to analyze the sensitivity of our results to potentially missing biopsy reports. First, we restricted the GEE to participants living close to the

Table 2 Predictors for the no. of biopsy reports

Predictors for biopsy reports	Model 1		Model 2	
	IRR	95% CI	IRR	95% CI
Age (per decade)	1.15	[1.08; 1.23]	1.16	[1.09; 1.23]
Sex (male vs. female)	0.73	[0.63; 0.84]	0.73	[0.63; 0.85]
Education (years, reference: 10y)				
< 10y	0.99	[0.81; 1.20]	0.99	[0.82; 1.2]
> 10y	0.99	[0.83; 1.17]	0.99	[0.84; 1.17]
Employed (yes vs. no)	0.92	[0.77; 1.11]	0.93	[0.77; 1.11]
Relationship status (reference: single)				
Married	1.04	[0.80; 1.35]	1.03	[0.79; 1.34]
Divorced	0.94	[0.65; 1.35]	0.92	[0.64; 1.33]
Widowed	1.17	[0.79; 1.75]	1.16	[0.77; 1.73]
Hospitalized in last 12 months (yes vs. no)	3.45	[3.01; 3.96]	3.45	[3.01; 3.96]
Known cancer history (yes vs. no)	2.89	[2.28; 3.67]	2.89	[2.28; 3.67]
Time-varying measures				
Disclosure of lab anomaly (yes vs. no)	1.37	[1.12; 1.67]	1.37	[1.12; 1.66]
Disclosure of MRI IF (yes vs. no)	2.17	[1.76; 2.68]	2.32	[1.85; 2.89]
Time (post-SHIP vs. pre-SHIP)	1.29	[1.07; 1.55]	1.30	[1.08; 1.56]

Results from GEE with a negative binomial distribution shown as incidence-rate-ratios (IRR). In model 2 only IFs from MRI with tumor relevance were used

GEE with a negative binomial distribution calculated in n = 6593 participants due to missing data in covariates

Table 3 Outcomes of biopsies stratified for disclosure of incidental findings

Outcome	Combination of IFs	Strata size (participants)	Pre SHIP N biopsies (participants)	Post SHIP N biopsies (participants)	Rate ratio biopsies [CI]	Δ participants (%)
No malignancy or tumor	Lab−IMRI−	2046	127 (109)	138 (119)	1.09 [0.85; 1.38]	10 (0.49)
	Lab+IMRI−	3685	316 (245)	385 (286)	1.22 [1.05; 1.41]	41 (1.11)
	Lab−IMRI+	396	31 (29)	58 (46)	1.87 [1.21; 2.92]	37 (9.34)
	Lab+IMRI+	626	66 (51)	111 (82)	1.68 [1.24; 2.29]	31 (4.95)
Benign tumor	Lab−IMRI−	2046	18 (18)	20 (17)	1.11 [0.58; 2.13]	− 1 (−0.05)
	Lab+IMRI−	3685	39 (36)	32 (32)	0.82 [0.51; 1.31]	− 4 (−0.11)
	Lab−IMRI+	396	11 (10)	12 (12)	1.09 [0.47; 2.53]	2 (0.51)
	Lab+IMRI+	626	17 (14)	25 (20)	1.47 [0.79; 2.77]	6 (0.96)
Pre-cancerous lesion (including carcinoma in situ)	Lab−IMRI−	2046	16 (14)	11 (9)	0.69 [0.31; 1.49]	− 5 (−0.24)
	Lab+IMRI−	3685	33 (30)	50 (43)	1.51 [0.98; 2.37]	13 (0.35)
	Lab−IMRI+	396	2 (1)	7 (7)	3.32 [0.78; 24.59]	6 (1.52)
	Lab+IMRI+	626	10 (8)	13 (11)	1.30 [0.56; 3.06]	3 (0.48)
Malignant process	Lab−IMRI−	2046	16 (16)	27 (23)	1.68 [0.91; 3.20]	7 (0.34)
	Lab+IMRI−	3685	41 (38)	63 (55)	1.53 [1.04; 2.29]	17 (0.46)
	Lab−IMRI+	396	3 (3)	12 (12)	3.85 [1.21; 17.67]	9 (2.27)
	Lab+IMRI+	626	15 (13)	28 (26)	1.86 [1.00; 3.58]	13 (2.08)

Lab−IMRI−, no disclosure of laboratory or MRI IFs; Lab+IMRI−, disclosure of laboratory IFs and no disclosure of MRI IFs; Lab−IMRI+, no disclosure of laboratory IFs and disclosure of MRI IFs; Lab+IMRI+, disclosure of laboratory; MRI IFs Δ, delta or change in the number of participants

examination center. Second, we conducted a weighted GEE, i.e., biopsies of participants living further away were given higher weights. In both approaches, changes in coefficients were marginal and relevant predictors for an increase of biopsies remained the same (Online Appendix, Table G).

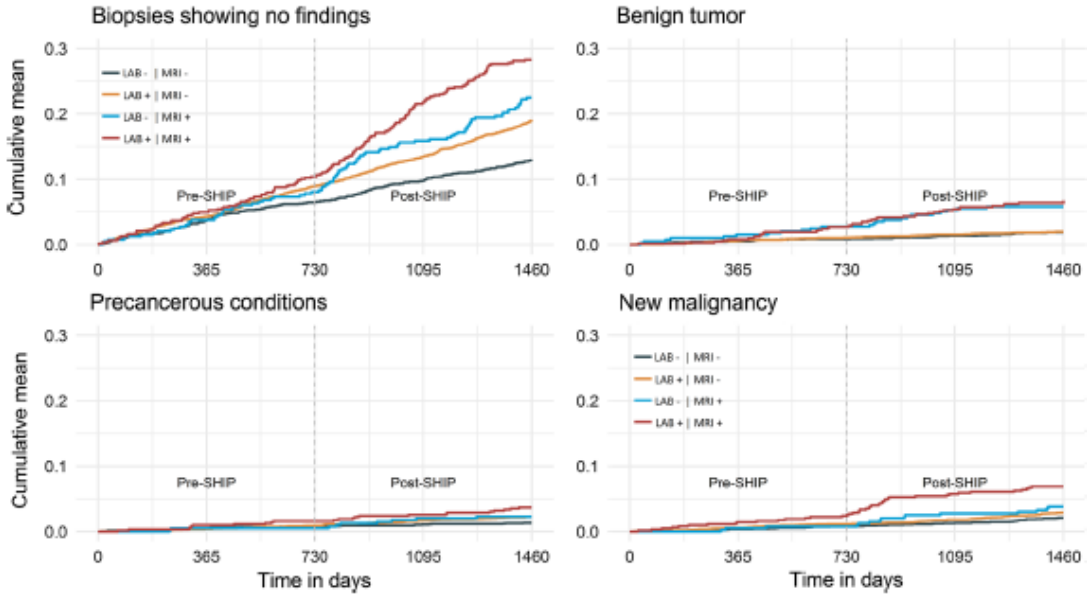


Fig. 2 Cumulative biopsy rates. Cumulative rates of biopsies identifying (top left) exclusion of malignancy or benign tumor, (top right) benign tumors, (bottom left) pre-cancerous conditions (including carcinoma in situ), and (bottom right) malignancies

Discussion

We assessed the effects of the disclosure of MRI and laboratory IFs on the frequency and outcomes of biopsies in the 2 years after examination in a large population-based cohort study. The overall number of biopsies increased by 42.7% comparing the 2 years before SHIP examination with the 2 years afterwards. The disclosure of MRI and laboratory IFs contributed strongly to this increase; in particular, MRI IFs with tumor relevance (Table 2, model 2). IF disclosures are therefore likely to reflect an undesired intervention effect in our observational study. New malignancies were most often discovered in a subgroup of participants with laboratory abnormalities and MRI disclosures. However, the absolute number of malignancies detected due to biopsies was small compared to the number of biopsies yielding no malignancy or tumor. Therefore, from a clinical perspective, overtesting and potentially overdiagnosis seem to be negative consequences associated with the disclosure of incidental findings in the setting of a general population cohort. Our findings favor a more restrictive communication policy of incidental study findings.

While the number of biopsies was expected to rise in an aging cohort, this is unlikely to account for the increase we observed. A more plausible explanation is that the disclosed IFs played a causative role in the increase in biopsies. First, the written disclosure of an MRI finding always included a recommendation to consult a doctor. Second, the vast majority of disclosed findings was related to tumors of unknown clinical significance [1] which may trigger biopsies. The overall effect of disclosed MRI IFs on biopsies was larger than that of laboratory abnormalities. This is plausible as the disclosure of laboratory findings did not entail any explicit recommendation to seek further diagnostic action.

The absolute effect of IF disclosure, both for imaging and laboratory findings, was greatest for biopsies showing nonmalignant findings. More than every fourth participant with an MRI and laboratory IF (28.1%) had at least one biopsy, and 62.1% of biopsies revealed no malignancy or tumor. In comparison, few new malignancies were detected via biopsy after SHIP. The low rate of malignancies detected in our large population-based cohort study corresponds to results from smaller studies using wb-MRI. In a study of 666 subjects the incidence of malignant lesions was 1.05%; in 83 subjects this number was 2.40% [34, 35]. An umbrella review found that malignancy rates of IF from clinical imaging range from less than 5–42% [2]. However, most of the studies in this review included asymptomatic oncology patients, explaining the higher rate of malignancies compared to our population-based study.

The effects of the disclosure of MRI and laboratory IFs on biopsies revealing benign and precancerous conditions was less consistent. This may be due to underpowered analyses given the low number of participants and events (biopsies) in the subgroups as defined by different combinations of IF types. However, increases in the detection of benign and precancerous conditions may also represent an indirect effect of IF disclosure, namely the triggering of so-called cascades of care [36]. Benign and precancerous conditions are generally not expected to cause laboratory anomalies, and small

precancerous lesions are not expected to be visible on MRI. However, the disclosure of any type of IF has been shown to initiate a range of diagnostic tests, often resulting in additional diagnoses of varying clinical significance [17, 36].

From a clinical perspective, some participants may have benefited from disclosed IFs through the early detection of new malignancies, most of which were diagnosed within 6 months after the SHIP examination. However, the clinical benefit of the earlier detection of malignancies is uncertain, as early diagnosis and intervention does not always equate to a better outcome [10]. We cannot determine from our data which of the disclosed IFs led to therapeutic benefits. Furthermore, the high rate of biopsies resulting in no malignant findings raises the risk for harm resulting from cascades of care, in which patients are exposed to the additional risks and complications of unnecessary tests and interventions [17, 36]. Overdetected conditions have the potential to increase diverse types of costs to the participant and the health care system [7]. These include, among others, psychosocial and financial burdens [8, 9, 37]. In addition, biopsies may be part of a larger cascade of care with preceding diagnostic steps such as further imaging or laboratory tests triggered by IFs [36]. As a result, we likely underestimate the burden of overtesting that resulted from IF disclosure. Clinicians and patients faced with IFs often feel compelled to pursue cascades of care to rule out serious disease out of a need for certainty [38]. Our results serve as a reminder to clinicians to critically assess the need and consequences of diagnostic tests prior to ordering them.

From a research perspective, the disclosure of IFs introduced bias into health service utilization and other outcomes [8, 37] in our population-based cohort study. This loss in validity needs to be weighed against the well-being and health of our study participants. Our results suggest that the benefit of the IF disclosure is limited at best. Moreover, many participants reported distress due to IF disclosure [9]. In the context of the limited benefits associated with IF disclosure, we recommend a restrictive disclosure policy for research studies to minimize the costs and consequences for participants and the health care system.

Strengths and limitations

To the best of our knowledge, this is the first study that has used routine biopsies as an outcome measure of IF disclosure. By combining cross-sectional research data with longitudinal routine clinical data on invasive procedures following disclosure of MRI IFs, our study contributes to a better understanding of the consequences of IF disclosure in research and clinical settings. All findings in this study are considered incidental as they were obtained outside a clinical context of routine care. Nonetheless, not all findings were new to the participants because of clinical diagnostics outside the SHIP study. Yet, only a minority of 13.3% in a surveyed SHIP-subsample reported already having had full knowledge of the communicated findings [9].

A limitation of this study is the missing direct link between an IF and the observed biopsies. Organ-specific IFs should result in specific biopsies and this association has not been examined in detail in this study (Online Appendix, Figure C). Rather, our approach examines a global association between IFs and changes in the number of biopsies. Furthermore, the presence of known malignant disease in some participants (overall: n = 586, MRI participants: n = 238) implies that some IFs cannot be considered truly incidental as the biopsies conducted after SHIP may have been a consequence of this disease. Nevertheless, the impact of IF disclosure on the number of biopsies conducted remained robust in sensitivity analyses including only participants with no previously known malignant diseases.

We are aware that certain biopsies (e.g., Pap-smears) were performed in an ambulatory setting and are missing from our data. Biopsies for patients living far away from the study site may have been more often conducted in pathology laboratories other than the Department of Pathology. However, we have no reason to assume a systematic difference between wb-MRI participants and non-participants in this regard. Our results remained robust in a sensitivity analysis including only participants living near our site.

The cumulative availability of biopsy reports from 2002 onward is consistent (Online Appendix, Figure B), but a small number of biopsy reports may have been lost in the linkage process. We have no reason to assume a systematic linkage error affecting any one subgroup. We assume that the small proportion of reports (16.6%) that did not undergo a double reading do not affect our results because misclassification in the cross-validated reports was less than 5%.

Conclusion

The disclosure of MRI and laboratory findings in a population-based cohort study is fraught with problems for both participants and the integrity of observational research. The disclosure of IFs represents an intervention that introduced bias into the natural course of health service utilization by exerting a lasting influence on biopsy rates. Researchers and clinicians should be aware that the increase in biopsy frequency after disclosure of MRI and laboratory IFs is substantial, but the rate of malignancies diagnosed is low. Our data thus supports a more restrictive disclosure policy for research MRI findings.

Author contributions AR filed the data request; programed statistical scripts for data linkage; created a data base for readings of biopsy reports, conducted the data preparation and all statistical analyses; created and revised the manuscript. ES filed the data request; designed the categorization of biopsy data; completed the reading of all biopsy data; conducted the literature research; selected laboratory markers; created and revised the manuscript. StS responsible for biopsy data acquisition; supervised the consent process of readings; revised and approved the manuscript. RB responsible for MRI data acquisition; revised and approved the manuscript. CH completed the reading of biopsy data; revised and approved the manuscript. JFC conceived the study; designed the categorization of biopsy data; completed the reading of biopsy data; selected laboratory markers; co-drafted, revised and approved the manuscript. COS conceived the study; supervised the data linkage and statistical analyses; co-drafted, revised and approved the manuscript

Funding Open Access funding provided by Projekt DEAL. SHIP is part of the Community Medicine Research network of the University of Greifswald, Germany. Examinations were funded by the Federal Ministry of Education and Research (Grant No. 03ZIK012), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania. Whole-body MR imaging was supported by a joint grant from Siemens Healthcare, Erlangen, Germany, and the Federal State of Mecklenburg-Vorpommern. Dynamic contrast-enhanced MR mammography research is part of the entire whole-body MRI study and was supported by Bayer Healthcare. The study of effects due to incidental MR findings were specifically supported by the Deutsche Forschungsgemeinschaft [DFG, Grant Nos. SCHM 2744/1-2:1/CH 921/1-2].

Availability of data and material 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 histology data utilized in this study are not publicly available due to privacy restrictions to protect personal data of research participants.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval and consent to participate All participants gave written informed consent for the examinations and 98% (n = 6626) provided consent for a linkage with the hospital records. The Ethics Committee of Greifswald University Medical Center approved the study protocol.

References

1. Hegenscheid K, et al. Potentially relevant incidental findings on research whole-body MRI in the general adult population: frequencies and management. Eur Radiol. 2013;23(3):816–26.
2. O'Sullivan JW, et al. Prevalence and outcomes of incidental imaging findings: umbrella review. BMJ. 2018;361:k2387.
3. Gibson LM, 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.
4. Prabhu V, et al. Frequency and outcomes of incidental breast lesions detected on abdominal MRI over a 7-year period. AJR Am J Roentgenol. 2017;208(1):107–13.
5. Hiom SC. Diagnosing cancer earlier: reviewing the evidence for improving cancer survival. Br J Cancer. 2015;112(Suppl 1):S1–5.
6. Allemani C, et al. Global surveillance of cancer survival 1995–2009: analysis of individual data for 25,676,887 patients from 279 population-based registries in 67 countries (CONCORD-2). The Lancet. 2015;385(9972):977–1010.
7. Brodersen J, et al. Focusing on overdiagnosis as a driver of too much medicine. BMJ. 2018;362:k3494.
8. Orme NM, et al. Incidental findings in imaging research: evaluating incidence, benefit, and burden. Arch Intern Med. 2010;170(17):1525–32.
9. Schmidt CO, et al. Psychosocial consequences and severity of disclosed incidental findings from whole-body MRI in a general population study. Eur Radiol. 2013;23(5):1343–51.
10. Booth TC, Boyd-Ellison JM. The current impact of incidental findings found during neuroimaging on neurologists' workloads. PLoS ONE. 2015;10(2):e0118155.
11. Kelly ME, et al. Incidental findings detected on emergency abdominal CT scans: a 1-year review. Abdom Imaging. 2015;40(6):1853–7.
12. Djavan B, et al. Safety and morbidity of first and repeat transrectal ultrasound guided prostate needle biopsies: results of a prospective European prostate cancer detection study. J Urol. 2001;166(3):856–60.
13. Leveridge MJ, et al. Outcomes of small renal mass needle core biopsy, nondiagnostic percutaneous biopsy, and the role of repeat biopsy. Eur Urol. 2011;60(3):578–84.
14. Ghadian A. Concerns about renal mass biopsy. Can Urol Assoc J. 2014;8(7–8):230.
15. Yun S, et al. Diagnostic accuracy and complications of CT-guided core needle lung biopsy of solid and part-solid lesions. Br J Radiol. 2018;91(1088):20170946.
16. Shyamala K, Girish HC, Murgod S. Risk of tumor cell seeding through biopsy and aspiration cytology. J Int Soc Prev Community Dent. 2014;4(1):5–11.
17. Korenstein D, et al. Development of a conceptual map of negative consequences for patients of overuse of medical tests and treatments. JAMA Intern Med. 2018;178(10):1401–7.
18. Volzke H, et al. Cohort profile: the Study of Health in Pomerania. Int J Epidemiol. 2011;40(2):294–307.
19. Schmidt CO, et al. Partial misspecification of survey design features sufficed to severely bias estimates of health-related outcomes. J Clin Epidemiol. 2011;64(4):416–23.
20. Hegenscheid K, et al. Whole-body magnetic resonance imaging of healthy volunteers: pilot study results from the population-based SHIP study. Rofo. 2009;181(8):748–59.
21. Vatsalan D, Christen P, Verykios VS. A taxonomy of privacy-preserving record linkage techniques. Inf Syst. 2013;38(6):946–69.
22. Van der Loo MP. The stringdist package for approximate string matching. R J. 2014;6(1):111–22.

23. Schmidberger M, Morgan M, Eddelbuettel D, et al. State of the Art in Parallel Computing with R. J Stat Softw. 2009;31(1):1–27.
24. Steorts R, Ventura S, Sadinle M, Fienberg S. A comparison of blocking methods for record linkage. In: Domingo-Ferrer J, editor. Privacy in statistical databases. Berlin: Springer; 2014. p. 253–68.
25. Aragon TJ. epitools: Epidemiology tools. R package version 0.5. Accessed 10 Jan 2020
26. Vollset SE. Confidence intervals for a binomial proportion. Stat Med. 1993;12(9):809–24.
27. Pan W. Akaike's information criterion in generalized estimating equations. Biometrics. 2001;57(1):120–5.
28. Long JS, Freese J. Regression models for categorical dependent variables using Stata. College Station: Stata Press; 2006.
29. Chiou SH, Huang C-Y. reReg: Recurrent event regression. R package version 1.2.1, 2019.
30. Bie O, Borgan Ø, Liestøl K. Confidence intervals and confidence bands for the cumulative hazard rate function and their small sample properties. Scand J Stat. 1987;14:221–33.
31. R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020.
32. Wasserstein RL, Schirm AL, Lazar NA. Moving to a world beyond "p < 0.05". Washington: Taylor & Francis; 2019.
33. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. Nature. 2019;567(7748):305–7.
34. Cieszanowski A, et al. Non-contrast-enhanced whole-body magnetic resonance imaging in the general population: the incidence of abnormal findings in patients 50 years old and younger compared to older subjects. PLoS ONE. 2014;9(9):e107840.
35. Ulus S, et al. Whole-body MRI screening in asymptomatic subjects; preliminary experience and long-term follow-up findings. Pol J Radiol. 2016;81:407–14.
36. Ganguli I, et al. Cascades of care after incidental findings in a US national survey of physicians. JAMA Netw Open. 2019;2(10):e1913325.
37. Schmidt CO, et al. Impact of whole-body MRI in a general population study. Eur J Epidemiol. 2016;31(1):31–9.
38. Ganguli I, et al. Why do physicians pursue cascades of care after incidental findings? A national survey. J Gen Intern Med. 2019.

Affiliations

Adrian Richter[1] · Elizabeth Sierocinski[1,2] · Stephan Singer[3,4] · Robin Bülow[5] · Carolin Hackmann[1,6] · Jean-François Chenot[2] · Carsten Oliver Schmidt[1]

Elizabeth Sierocinski
elizabeth.sierocinski1@uni-greifswald.de

Stephan Singer
stephan.singer@uni-greifswald.de;
stephan.singer@med.uni-tuebingen.de

Robin Bülow
robin.buelow@med.uni-greifswald.de

Carolin Hackmann
carolin.hackmann@charite.de

Jean-François Chenot
jchenot@uni-greifswald.de

Carsten Oliver Schmidt
carsten.schmidt@uni-greifswald.de

[1] Department SHIP-KEF, Institute for Community Medicine, Greifswald University Medical Center, Walther Rathenau Str. 48, 17475 Greifswald, Germany

[2] Department of Family Medicine, Institute for Community Medicine, Fleischmannstr. 42, 17475 Greifswald, Germany

[3] Institute for Pathology, Greifswald University Medical Center, Friedrich-Loeffler-Str. 23e, 17487 Greifswald, Germany

[4] Institute of Pathology, University Hospital Tuebingen, Liebermeisterstrasse 8, 72076 Tuebingen, Germany

[5] Department of Diagnostic Radiology and Neuroradiology, Greifswald University Medical Center Greifswald, 17475 Greifswald, Germany

[6] Institute for Hygiene and Environmental Medicine, Charité – University Medicine Berlin, Hindenburgdamm 27, 12203 Berlin, Germany

14.2 Danksagung

Meinem Betreuer Herrn Prof. Carsten Oliver Schmidt danke ich für die unermüdliche Unterstützung, die hilfreichen Hinweise, konstruktive Kritik und Diskussionen beim Anfertigen der Arbeit. Herrn Dr. Adrian Richter danke ich für die statistische Auswertung, das Feedback und den hilfreichen Austausch während des gesamten Projekts. Herrn Prof. Jean-François Chenot danke ich für die Zusammenarbeit bei der Validierung der histologischen Befunde sowie die inhaltliche Unterstützung und konstruktive Kritik. Frau Carolin Hackmann danke ich für die Teamarbeit bei der Validierung der histologischen Befunde. Allen an dem Projekt Beteiligten danke ich für die gute Zusammenarbeit.