THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

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

OPEN ACCESS

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

Lorna M. Gibson[ID][1], John Nolan[1], Thomas J. Littlejohns[ID][2], Edouard Mathieu[3], Steve Garratt[4], Nicola Doherty[4], Steffen Petersen[ID][5,6], Nicholas C. W. Harvey[ID][7], Jonathan Sellors[2,4], Naomi E. Allen[2], Joanna M. Wardlaw[ID][8,9], Caroline A. Jackson[1], Cathie L. M. Sudlow[ID][1] *

1 Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, United Kingdom, 2 Clinical Trial Service Unit and Epidemiological Studies Unit, University of Oxford, Oxford, United Kingdom, 3 Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom, 4 UK Biobank Co-ordinating Centre, UK Biobank, Stockport, United Kingdom, 5 William Harvey Research Institute, Queen Mary University of London, London, United Kingdom, 6 Barts Heart Centre, Barts Health NHS Trust, London, United Kingdom, 7 MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton, United Kingdom, 8 UK Dementia Research Institute, University of Edinburgh, Edinburgh, United Kingdom, 9 Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, United Kingdom

* Cathie.Sudlow@ed.ac.uk

Abstract

Background

Feedback of potentially serious incidental findings (PSIFs) to imaging research participants generates clinical assessment in most cases. Understanding the factors associated with increased risks of PSIFs and of serious final diagnoses may influence individuals' decisions to participate in imaging research and will inform the design of PSIFs protocols for future research studies. We aimed to determine whether, and to what extent, socio-demographic, lifestyle, other health-related factors and PSIFs protocol are associated with detection of both a PSIF and a final diagnosis of serious disease.

Methods and findings

Our cohort consisted of all UK Biobank participants who underwent imaging up to December 2015 (n = 7334, median age 63, 51.9% women). Brain, cardiac and body magnetic resonance, and dual-energy x-ray absorptiometry images from the first 1000 participants were reviewed systematically by radiologists for PSIFs. Thereafter, radiographers flagged concerning images for radiologists' review. We classified final diagnoses as serious or not using data from participant surveys and clinical correspondence from GPs up to six months following imaging (either participant or GP correspondence, or both, were available for 93% of participants with PSIFs). We used binomial logistic regression models to investigate

associations between age, sex, ethnicity, socio-economic deprivation, private healthcare use, alcohol intake, diet, physical activity, smoking, body mass index and morbidity, with both PSIFs and serious final diagnoses. Systematic radiologist review generated 13 times more PSIFs than radiographer flagging (179/1000 [17.9%] versus 104/6334 [1.6%]; age- and sex-adjusted OR 13.3 [95% confidence interval (CI) 10.3–17.1] p<0.001) and proportionally fewer serious final diagnoses (21/179 [11.7%]; 33/104 [31.7%]). Risks of both PSIFs and of serious final diagnoses increased with age (sex-adjusted ORs [95% CI] for oldest [67–79 years] versus youngest [44–58 years] participants for PSIFs and serious final diagnoses respectively: 1.59 [1.07–2.38] and 2.79 [0.86 to 9.0] for systematic radiologist review; 1.88 [1.14–3.09] and 2.99 [1.09–8.19] for radiographer flagging). No other factor was significantly associated with either PSIFs or serious final diagnoses. Our study is the largest so far to investigate the factors associated with PSIFs and serious final diagnoses, but despite this, we still may have missed some associations due to sparsity of these outcomes within our cohort and small numbers within some exposure categories.

Conclusion

Risks of PSIFs and serious final diagnosis are substantially influenced by PSIFs protocol and to a lesser extent by age. As only 1/5 PSIFs represent serious disease, evidence-based PSIFs protocols are paramount to minimise over-investigation of healthy research participants and diversion of limited health services away from patients in need.

Introduction

Brain and body imaging is increasingly used for research, diagnostic and screening purposes and is accompanied by the risk of identifying abnormalities which are unrelated to the purposes of the imaging, so-called incidental findings (IFs) [1]. Since very few IFs turn out to represent serious disease [2], it is of limited value to feedback clearly non-serious IFs. Therefore, we focus on potentially serious IFs (PSIFs), defined as those which indicate the possibility of a condition which, if confirmed, would carry a real prospect of seriously threatening life span, or of having a substantial impact on major body functions or quality of life [2]. Feedback of PSIFs detected during research imaging generates some form of clinical assessment (e.g. general practitioner appointments and specialist referrals, or further investigations including imaging and invasive procedures) in almost all cases [2]. Information on the factors associated with increased risk of detection and feedback of a PSIF (and therefore of subsequent clinical assessment), and with increased risk of eventually receiving a serious final diagnosis may influence individuals' decisions to consent to participate in imaging research [3–5] and inform researchers' designs of appropriate PSIFs policies, which are required by major research funders [6, 7].

A small number of studies (N = 151 to 5800) which followed-up unselected participants with PSIFs suggest that PSIFs are associated with age, but not with sex. However none of these studies investigated the associations of PSIFs with PSIFs protocols, or any factors associated with serious final diagnoses [8–17].

The UK Biobank Imaging Study provides an opportunity to investigate potential risk factors for PSIFs and serious final diagnoses. In the UK Biobank Imaging Study, 100,000 of the original 500,000 participants are undergoing brain, cardiac and body magnetic resonance imaging (MRI), dual-energy X-ray absorptiometry (DXA) and carotid Doppler ultrasound;

does not alter our adherence to PLOS ONE policies on sharing data and materials.

Abbreviations: BMI, Body mass index; CI, Confidence interval; DXA, Dual-energy X-ray absorptiometry; GP, General practitioner; IFs, Incidental findings; MRI, magnetic resonance imaging; OR, Odds ratio; PSIFs, Potentially serious incidental findings.

over 32,000 participants have been imaged as of December 2018] [18]. These imaging data are linked to detailed sociodemographic, lifestyle, physical measurement, genetic and routine healthcare data generating an extensive research resource [19].

The UK Biobank Imaging Study will inevitably generate PSIFs. To inform the development of a pragmatic PSIFs protocol that aims to minimise harm to (the largely asymptomatic) 100,000 imaged participants, UK Biobank reviewed current practice, published literature and guidance, and sought advice from professional bodies, and from ethical and legal experts [2]. The protocol is based on radiographers flagging images of potential concern to a radiologist for their review [2, 20]. This approach was evaluated against a protocol involving systematic radiologist review of all images (which is more commonly used in research studies), and found to generate less harm (i.e., less unnecessary anxiety to participants and their families) and a lower burden on the publicly-funded UK National Health Service [2]. UK Biobank is continuing to evaluate this PSIFs protocol through systematic follow-up of all participants identified with a PSIF.

We aimed to determine whether, and to what extent, socio-demographic, lifestyle, other health-related factors and PSIFs protocol are associated with detection of both a PSIF and a final diagnosis of serious disease. We achieved this using data from the first 7,334 participants imaged during the first 20 months of the UK Biobank Imaging Study (including systematic follow-up of 283 participants with PSIFs).

Methods

We prepared this manuscript according to STROBE guidelines (S1 File) [21]. The statistical analysis code is available online [22]. UK Biobank obtained ethics approval for the imaging study, and evaluation of the PSIFs protocol (North West Research Ethics Committee reference numbers: 11/NW/0382; 16/NW/0274). We provided all participants with written information about the imaging study and the UK Biobank imaging IFs protocol [23]. All participants provided written consent to take part in the imaging study, and for UK Biobank to feed back any identified potentially serious IFs to them and their general practitioner (GP).

UK Biobank Imaging Study

Of 9.2 million adults aged 40–69 invited to participate in UK Biobank, 0.5 million (5.5%) participated, providing initial baseline data between 2006 and 2010 [24]. From April 2014 to December 2015, participants living within approximately 120 km of the imaging centre in Stockport were further invited to take part in the UK Biobank Imaging Study [25]. Participants were excluded if they had metal implants, penetrating metal injury, non-removable metallic items, or if they would find it difficult to complete the imaging, e.g. due to claustrophobia (Fig 1) [25].

At the imaging visit, participants underwent brain, heart and body MRI, whole-body, spine and hip DXA and carotid Doppler ultrasound [26–30]. Participants also repeated the UK Biobank baseline assessment, which involved: a touchscreen questionnaire to collect data on potentially relevant risk factors for diseases, including sociodemographic, lifestyle and medical history; an interview; and physical measurements [31].

UK Biobank PSIFs protocol

During imaging, UK Biobank radiographers may notice PSIFs and 'flag' concerning images for radiologist review; radiologist-confirmed PSIFs are then fed back to participants and their GP [2]. To evaluate this PSIFs protocol, all images from the first 1000 participants were also systematically reviewed by radiologists for PSIFs [2]. Radiographers did not flag any PSIFs in
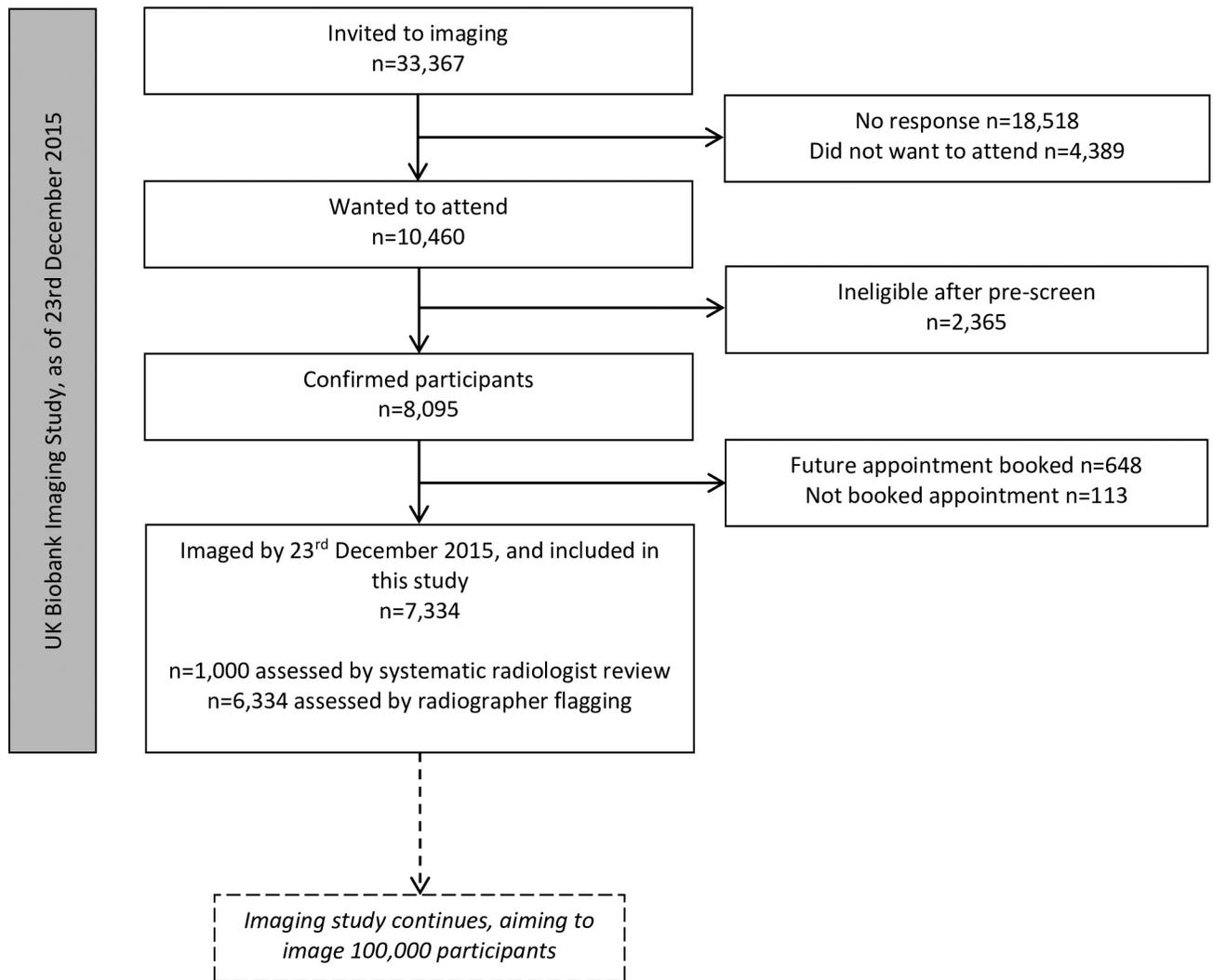
Fig 1. Participant flowchart.

addition to those detected by the radiologists within the first 1000 imaged participants [2]. Therefore, for the purposes of this present study, we classified the first 1000 imaged participants as undergoing the 'systematic radiologist review' PSIFs protocol, and subsequently imaged participants as undergoing the 'radiographer flagging' PSIFs protocol. For both protocols, to aid interpretation of images, radiologists received information on participants' age, sex, ethnicity, alcohol intake, smoking status, blood pressure, body mass index (BMI), employment status, and self-reported medical history. The legal and ethical background to the UK Biobank 'radiographer flagging' protocol has been previously described [2].

Participants with PSIFs are surveyed at six weeks and six months after receiving feedback, while their GPs are surveyed six months after feedback and asked for copies of relevant clinical correspondence; these responses include data on final diagnoses [2].

Carotid Doppler ultrasound was deemed extremely unlikely to generate PSIFs under UK Biobank's protocol [25], as asymptomatic carotid stenosis is rare and the clinical relevance is not well established, and extra-carotid abnormalities were not likely to be relevant as UK

Biobank sonographers are trained in vascular Doppler US only. As such, carotid Doppler US data are not included in this study.

Data sources and variables

PSIFs and serious final diagnoses. We extracted data on the number, types and body region of each participant's PSIF(s) from radiologists' reports. A consultant physician and an experienced clinical radiology specialty trainee independently classified final diagnoses using all available survey data and clinical correspondence; we contacted participants and GPs by telephone where these data were insufficient to classify final diagnoses [2]. We classified final diagnoses as either: serious (if they were likely to threaten life span, or have a substantial impact on quality of life or major body function); not serious (if this was not the case); or indeterminate (if there remained insufficient data to classify a final diagnosis as serious or not) [2]. A PSIF which turned out to be a known diagnosis was deemed to be very unlikely have any additional impact on the participant, and therefore we classified such findings as 'non-serious.' classified participants with more than one PSIF according to their most serious final diagnosis [2].

Participant factors. We selected variables available from UK Biobank [32] which might be associated with PSIFs or would be possible confounders. These were age, sex, ethnicity, Townsend socio-economic deprivation score (which may reduce access to healthcare, increase the risk of disease and reduce opportunities for disease detection prior to research imaging), use of private healthcare (which may be associated with reduced risk of PSIFs if it increases prior knowledge of disease), alcohol intake [33], fruit and vegetable intake [34], physical activity [35], smoking status, body mass index (BMI) [36] and morbidity. We measured the latter using the Elixhauser Index calculated using Hospital Episode statistics data from two years before the date of imaging, and defined morbidity as ≥ 1 Elixhauser Index health conditions [37–39].

Statistical analyses

Since our previous study showed that the 'systematic radiologist review' protocol produced approximately ten times more PSIFs compared with the 'radiographer flagging' protocol [2], all analyses were stratified by PSIFs protocol to control for potential confounding. We compared characteristics between participants with and without PSIFs, and with and without serious final diagnoses, and calculated age- and sex-adjusted odds ratios (ORs) with 95% confidence intervals (CIs) using binomial logistic regression models.

We tested for normal distributions of continuous variables by visual inspection of graphed data and Kolmogorov-Smirnov goodness-of-fit tests. We attempted to normalise non-normally distributed data using log transformations, and if this failed, recoded variables into categories, aiming for similar numbers of participants in each category to optimise statistical efficiency. We used non-parametric tests to compare distributions of non-normally distributed variables between two groups. We considered data to be missing if participants did not respond, or if they responded 'do not know' or 'prefer not to answer'; such participants were excluded only from the relevant analyses. We present summary statistics of the characteristics of the whole UK Biobank cohort only to inform on the likely influence of factors associated with PSIFs or serious final diagnoses which may occur as the UK Biobank imaging study continues; these cannot be compared directly to the imaged sub-cohort included in this study due to lack of independence of these two samples. The majority of variables had no, or only small proportions ($< 3\%$) of missing data. In total, 460/7334 (6.3%) participants had missing data for at least one variable. We performed all analyses using SPSS version 22.

Results

Participants

By 23rd December 2015, 7,334 of 33,367 invited participants (22.0%) had been imaged and were included in this study (Fig 1). Median age of the imaged participants was 63 (interquartile range 56–68) years and 3,804 (51.9%) were women (Table 1).

Compared to the entire UK Biobank cohort, this imaged sub-cohort included lower proportions of women, people of minority ethnicity groups, and people with less healthy lifestyles, including those with harmful alcohol intake, current smokers, low physical activity levels, or those who were overweight or obese. Conversely, a higher proportion of the imaged sub-cohort had one or more health conditions as measured using the Elixhauser Index compared to the whole cohort (Table 1).

PSIFs and final diagnoses

PSIFs were detected in 283/7,334 (3.9%) people: 179 of the first 1000 (17.9%) by systematic radiologist review; 104 of the subsequent 6,334 (1.6%) by radiographer flagging (OR for systematic radiologist review versus radiographer flagging: 13.3, 95% CI 10.3–17.1, p<0.001, Table 2). The majority of PSIFs were finally diagnosed as clinically non-serious (229/283, 80.9%). Serious final diagnoses occurred in 54/7,334 (0.7%) participants: 21 of the first 1000 (2.1%) undergoing the systematic radiologist review protocol and 33 of the 6,334 (0.5%) undergoing the radiographer flagging protocol (OR 4.2, 95% CI 2.4–7.4, p<0.001, Table 2). Radiographer flagging thus resulted in a higher proportion of PSIFs with serious final diagnoses than radiologist review (33/104 [31.7%] versus 21/179 [11.7%] respectively). The most common serious final diagnoses were tumours and vascular diseases (Tables 3, 4 and 5). The two doctors agreed on the initial classification of final diagnoses in 270/283 (95.4%) of cases, and readily resolved the 13 cases of disagreement through discussion.

Systematic radiologist review generated higher proportions of PSIFs on all imaged body regions (OR range 4.3–21.3, all p<0.001, Table 2) compared to radiographer flagging. Radiologists more commonly detected PSIFs on cardiac (8.1%) and body MRI (8.3%) compared to brain MRI (2.3%) or DXA (1.4%). In contrast, radiographer flagging generated similar proportions of PSIFs across body regions (range 0.3–0.6%, Table 2). Serious final diagnoses occurred most commonly on cardiac MRI assessed by systematic radiologist review (13/1000, 1.3%, Table 2).

Factors associated with PSIFs and serious final diagnoses

Across the relatively narrow age range of the included participants, older age was associated with an increased odds of PSIFs and of serious final diagnoses under both protocols, albeit not statistically significant for serious final diagnoses under systematic radiologist review (sex-adjusted ORs [95% CI] for oldest [67–79 years] versus youngest [44–58 years] participants for PSIFs and serious final diagnoses respectively: 1.59 [1.07–2.38] and 2.79 [0.86 to 9.0] for systematic radiologist review; 1.88 [1.14–3.09] and 2.99, 95% CI [1.09–8.19] for radiographer flagging) (Figs 2 and 3). Of the participants with PSIFs, those with serious final diagnoses were older than those with non-serious final diagnoses (median ages [range minimum-maximum] in years: 66 [50–76] versus 64 [44–76] respectively, p = 0.021).

Of participants assessed by radiographer flagging, overweight participants had reduced odds of serious final diagnoses compared to those of normal or underweight BMI (age- and sex-adjusted OR 0.21, 95% CI 0.08–0.58, p = 0.003, Fig 3), but the number of overweight participants was very small (n = 5).

Table 1. Characteristics of the UK Biobank cohort and the imaged sub-cohort included in this study.

	Entire UK Biobank cohort (of whom 100,000 will be imaged) (N = 502,205)[1] n (%)	Imaged UK Biobank sub-cohort included in this study (N = 7,334)[2] n (%)
Sociodemographics		
Age[3]		
Median (IQR)	63 (55–68)	63 (56–68)
Sex[4]		
Female	273,224 (54.4)	3804 (51.9)
Male	228,981 (45.6)	3530 (48.1)
Ethnicity[5]		
White	472,493 (94.1)	7023 (95.8)
Minority ethnicity groups	27,012 (5.4)	225 (3.1)
TDI[5]		
Median (IQR)	-2.1 (-3.6–0.6)	-2.5 (-3.9 –-0.5)
Private healthcare[5]		
Never used	120,934 (70.1)	5377 (73.3)
Ever used	49,980 (29.0)	1850 (25.2)
Lifestyle		
Alcohol[5,6]		
None	19,942 (14.1)	848 (11.6)
Moderate	73,886 (52.3)	4124 (56.2)
Hazardous	34,980 (24.8)	1854 (25.3)
Harmful	9,084 (6.4)	376 (5.1)
Smoking[5]		
Never	273,400 (54.4)	4350 (59.3)
Previous	172,980 (34.4)	2575 (35.1)
Current	52,947 (10.5)	319 (4.3)
Fruit and vegetable portions/day[5,7]		
< 5	342,833 (68.3)	5028 (68.6)
≥ 5	144,064 (28.7)	2141 (29.2)
Days/week of moderate physical activity[5,8]		
0–2	169,162 (33.7)	2149 (29.3)
3–4	118,615 (23.6)	1967 (26.8)
5–7	187,251 (37.3)	2983 (40.7)
Other factors		
Morbidity[9]		
None	457,301 (91.1)	6422 (87.6)
≥1 condition	44,904 (8.9)	912 (12.4)
BMI[5,10]		
Underweight	2625 (0.5)	47 (0.6)
Normal	162,348 (32.3)	2733 (37.3)
Overweight	212,064 (42.2)	3061 (41.7)

(Continued)

Table 1. (Continued)

	Entire UK Biobank cohort (of whom 100,000 will be imaged) (N = 502,205)[1] n (%)	Imaged UK Biobank sub-cohort included in this study (N = 7,334)[2] n (%)
Obese	122,228 (24.3)	1454 (19.8)

IQR = interquartile range, TDI = Townsend Deprivation Index (higher score indicates greater deprivation), BMI = body mass index

[1]. Data collected at recruitment visit, unless otherwise indicated.

[2]. Data collected at the imaging visit, unless otherwise indicated.

[3]. Age on 30th April 2014, i.e. the start of the imaging study, for the entire cohort, and the imaged cohort.

[4]. Sex data were only available from the recruitment visit.

[5]. Data were missing for ethnicity (2,700/502,205 [0.5%], 86/7,334 [1.2%]), TDI (627/502,205 [0.1%], 0/7,334 [0.0%]), private healthcare use (1,694/172,608 [1.0%, questions on private healthcare were introduced partway through the recruitment period on 29th April 2009, thus giving a smaller denominator], 107/7,334 [1.5%]), alcohol (3,357/141,149 [2.3%, questions on subtypes of alcoholic drinks were introduced partway through the recruitment period on 29th August 2009, thus giving a smaller denominator], 132/7,334 [1.8%]), smoking (2,878/502,205 [0.6%], 90/7,334 [1.2%]), fruit and vegetable intake (15,308/502,205 [3.0%], 165/7,334 [2.2%]), physical activity (27,177/502,205 [5.4%], 235/7,334 [3.2%]), BMI (2,940/502,205 [0.6%], 39/7,334 [0.5%]), from the whole UK Biobank cohort versus the imaged sub-cohort respectively.

[6]. We calculated alcohol intake in units per week and categorised these using British Medical Association guidelines (women: moderate $> 0 < 14$, hazardous 14–35, harmful > 35; men: moderate $>0 < 21$, hazardous 21–50, harmful > 50) [33].

[7]. We calculated portions of fruit and vegetable intake per day, and categorised these into five or more portions per day, or not. [34]

[8]. Participants were asked 'in a typical week, on how many days did you do 10 minutes or more of moderate physical activities like carrying light loads, cycling at normal pace (do not include walking)?' [35].

[9]. We calculate morbidity using an Elixhauser Index score [37, 38] based on two-years of routinely collected Hospital Episode Statistics data, looking back from date of recruitment for the entire UK Biobank cohort, and the date of imaging for the imaged sub-cohort. Routinely collected health data are used to calculate payments for providers for services delivered for different conditions. The system for applying prices to healthcare services changed in 2012 [39], therefore the numbers of conditions coded in health records may not be directly comparable between the entire cohort, and the imaged cohort.

[10]. We defined BMI categories as underweight, normal, overweight and obese as BMIs of <18.5, $\geq 18.5 < 25.0$, $\geq 25.0 < 30.0$, ≥ 30.0 respectively [36].

https://doi.org/10.1371/journal.pone.0218267.t001

Table 2. Odds ratios for potentially serious incidental findings (PSIFs) and serious final diagnoses comparing two protocols.

	Systematic radiologist review (N = 1000) n (%)[1]	Radiographer flagging (N = 6334) n (%)[1]	OR (95% CI)[2] systematic radiologist review versus radiographer flagging	p-value[3]
PSIFs	179 (17.9)	104 (1.6)	13.3 (10.3–17.1)	<0.001
Brain MRI	23 (2.3)	35 (0.6)	4.3 (2.5–7.3)	<0.001
Cardiac MRI	81 (8.1)	29 (0.5)	19.7 (12.8–30.2)	<0.001
Body MRI	83 (8.3)	27 (0.4)	21.3 (13.7–33.0)	<0.001
DXA	14 (1.4)	16 (0.3)	5.8 (2.8–11.9)	<0.001
Serious final diagnoses	21 (2.1)	33 (0.5)	4.2 (2.4–7.4)	<0.001
Brain MRI	4 (0.4)	13 (0.2)	2.0 (0.7–6.2)	0.221
Cardiac MRI	13 (1.3)	10 (0.2)	8.5 (3.7–19.5)	<0.001
Body MRI	3 (0.3)	5 (0.1)	4.1 (1.0–17.1)	0.056
DXA	1 (0.1)	5 (0.1)	1.3 (0.2–11.0)	0.818

OR = odds ratio, CI = confidence interval, PSIFs = potentially serious incidental findings, MRI = magnetic resonance imaging, DXA = dual energy X-ray absorptiometry

[1]. Numerators are the number of participants with at least one PSIF per region. Multiple PSIFs occurred in four participants (who had two PSIFs each) under radiographer flagging, and in 33 (28 had two and five participants had three PSIFs each) under systematic radiologist review, giving a total of 325 PSIFs; therefore the sums of the body region PSIFs are greater than the 104 and 179 participants with at least one PSIF respectively. No participant had more than one serious final diagnosis.

[2]. Age- and sex-adjusted ORs for PSIFs and serious final diagnoses.

[3]. p-value from Wald test.

https://doi.org/10.1371/journal.pone.0218267.t002

Table 3. Serious final diagnoses.

Image modality	Serious final diagnoses	Systematic radiologist review (N = 1000) n participants	Radiographer flagging (N = 6334) n participants
Brain MRI	Arachnoid cyst with hydrocephalus	1	-
	Arteriovenous malformation	-	1
	Cavernoma	-	1
	Meningioma requiring surgery	1	3
	Normal pressure hydrocephalus	-	1
	Pituitary tumour	2	4
	Pleomorphic adenoma requiring surgery	-	1
	Vestibular schwannoma	-	2
Cardiac MRI	Atrial fibrillation	1	1
	Cardiomyopathy	2	3
	Coronary heart disease	1	-
	Heart block and LV impairment	1	-
	Lung tumour	3	-
	Mesothelioma	-	1
	Myxoma	-	1
	Severe valve disease	-	2
	Thoracic aortic aneurysm	5	2
Body MRI: Abdomen	Abdominal aortic aneurysm > 5 cm	1	1
	Colonic tumour	-	1
	Gastrointestinal stromal tumour	1	-
	Pancreatic tumour	1	1
	Renal tumour	-	2
DXA	Osteoporotic crush fracture	1	5
All modalities: serious final diagnoses		**21**	**33**

MRI = magnetic resonance imaging, LV = left ventricular, DXA = dual-energy X-ray absorptiometry,— = zero

No significant associations were found between PSIFs or serious final diagnoses and any other investigated factor for participants assessed by either PSIFs protocol (Figs 2 and 3).

Discussion

Systematic radiologist review of images resulted in approximately 13 times more PSIFs, and four times more serious final diagnoses than the radiographer flagging protocol; these effect sizes are larger than those of any other risk factor assessed for association with either PSIFs or serious final diagnoses. Most (80%) PSIFs did not turn out to represent serious disease. The odds of PSIFs and of serious final diagnoses increased with age, regardless of PSIFs protocol. There were no clear associations between either PSIFs or serious final diagnoses and sex, ethnicity, socio-economic deprivation, use of private healthcare, alcohol intake, diet, physical activity, smoking status, BMI or morbidity among participants assessed using either PSIFs protocol.

Our study confirms and updates our previous findings from the first 1000 imaged UK Biobank participants [2]: compared to systematic radiologist review, radiographer flagging resulted in substantially fewer participants with potentially serious IFs and a higher proportion of these had serious final diagnoses. We also confirm the findings of the above-mentioned smaller cohort [2], that around 80% of PSIFs do not turn out to represent serious disease.

Table 4. Non-serious final diagnoses.

Image modality	Non-serious final diagnoses	Systematic radiologist review (N = 1000) n participants	Radiographer flagging (N = 6334) n participants
Brain MRI	Already known diagnosis	1	3
	Benign cyst/lesion	15	10
	Hydrocephalus (not serious)	-	2
	Suspected lesion not confirmed	3	3
Cardiac MRI	Already known cardiac diagnosis	7	5
	Already known lung diagnosis	2	1
	Already under investigation	-	1
	Cardiac diagnosis—not serious	8	8
	Lung diagnosis—not serious	28	2
	Other non-serious diagnosis	10	1
	Suspected lesion not confirmed	18	1
Body MRI: Abdomen	Abdominal aortic aneurysm < 5cm	2	1
	Already known diagnosis	4	3
	Benign lesion (e.g. cyst)	57	14
	Other non-serious diagnosis	4	-
	Suspected lesion not confirmed	13	2
Body MRI: Leg	Already known diagnosis	1	-
	Bone/soft tissue diagnosis—not serious	5	-
	Suspected lesion not confirmed	2	-
DXA	Already known diagnosis	5	5
	Non-serious diagnosis	5	3
	Suspected lesion not confirmed	2	2
All modalities: non-serious final diagnoses		**192**	**67**

MRI = magnetic resonance imaging, DXA = dual-energy X-ray absorptiometry,— = zero

Previous studies, mostly of brain MRI, found that PSIFs were associated with increased age [10, 13, 15, 17], but not clearly associated with sex [8–16]. We have further confirmed these findings in participants undergoing multimodal imaging of multiple body regions, and shown this to be independent of the IFs protocol. Previous studies did not demonstrate any

Table 5. Uncertain final diagnoses.

Image modality	Uncertain final diagnoses	Systematic radiologist review (N = 1000) n participants	Radiographer flagging (N = 6334) n participants
Brain MRI	Lesion, unclear nature	-	4
Cardiac MRI	Lung consolidation, unclear nature	1	1
	Lung nodule, unclear nature	2	-
Body MRI: Abdomen	Cysts, unclear nature	-	2
DXA	Crush fracture T11, unclear relevance	1	-
	Fractures, unclear cause	-	1
All modalities: uncertain final diagnoses		**4**	**8**

MRI = magnetic resonance imaging, DXA = dual-energy X-ray absorptiometry,— = zero

	Systematic radiologist review (N=1000)					Radiographer flagging (N=6334)				
	PSIF (N=179) n (%)	No PSIF (N=821) n (%)	OR (95% CI)¹ for PSIF	p-value²		PSIF (N=104) n (%)	No PSIF (N=6230) n (%)	OR (95% CI)¹ for PSIF	p-value²	
Sociodemographics										
Age tertile										
44-58 years	53 (29.6)	300 (36.5)	1.00			24 (23.1)	2020 (32.4)	1.00		
59-66 years	59 (33.0)	279 (34.0)	1.21 (0.80-1.81)	0.365		34 (32.7)	2216 (35.6)	1.28 (0.76-2.17)	0.359	
67-79 years	67 (37.4)	242 (29.5)	1.59 (1.07-2.38)	0.022		46 (44.2)	1994 (32.0)	1.88 (1.14-3.09)	0.013	
Sex										
Female	99 (55.3)	425 (51.8)	1.00			45 (43.3)	3235 (51.9)	1.00		
Male	80 (44.7)	396 (48.2)	0.84 (0.61-1.17)	0.304		59 (56.7)	2995 (48.1)	1.35 (0.91-2.00)	0.136	
Ethnicity³										
White	170 (97.1)	777 (95.7)	1.00			100 (97.1)	5876 (97.0)	1.00		
Minority ethnicity	5 (2.9)	35 (4.3)	0.74 (0.28-1.95)	0.544		3 (2.9)	182 (3.0)	1.11 (0.36-3.55)	0.862	
TDI tertile³										
Least deprived	50 (27.9)	247 (30.1)	1.00			38 (36.5)	2106 (33.8)	1.00		
Middle tertile	46 (25.7)	239 (29.1)	0.97 (0.63-1.51)	0.906		32 (30.8)	2127 (34.2)	0.86 (0.53-1.38)	0.526	
Most deprived	83 (46.4)	334 (40.7)	1.30 (0.88-1.93)	0.187		34 (32.7)	1994 (32.0)	1.02 (0.64-1.63)	0.949	
Private healthcare³										
Never used	126 (71.6)	605 (74.8)	1.00			81 (78.6)	4565 (74.4)	1.00		
Ever used	50 (28.4)	204 (25.2)	1.20 (0.83-1.73)	0.325		22 (21.4)	1574 (25.6)	0.79 (0.49-1.27)	0.333	
Lifestyle										
Alcohol³,⁴										
None	24 (14.0)	90 (11.1)	1.00			11 (10.7)	723 (11.8)	1.00		
Moderate	98 (57.0)	464 (57.4)	0.81 (0.49-1.34)	0.416		64 (62.1)	3498 (57.2)	1.15 (0.60-2.21)	0.666	
Hazardous	40 (23.3)	204 (25.2)	0.76 (0.43-1.33)	0.335		23 (22.3)	1587 (25.9)	0.90 (0.44-1.87)	0.783	
Harmful	10 (5.8)	50 (6.2)	0.80 (0.35-1.81)	0.587		5 (4.9)	311 (5.1)	1.05 (0.36-3.06)	0.929	
Smoking³										
Never	106 (60.6)	471 (57.9)	1.00			62 (61.4)	3711 (60.3)	1.00		
Previous	61 (34.9)	303 (37.3)	0.88 (0.62-1.25)	0.465		35 (34.7)	2176 (35.4)	0.87 (0.57-1.33)	0.528	
Current	8 (4.6)	39 (4.8)	1.00 (0.45-2.21)	0.998		4 (4.0)	268 (4.4)	0.93 (0.33-2.58)	0.882	
Fruit and vegetable portions/day³,⁴										
< 5	120 (69.0)	539 (67.5)	1.00			68 (67.3)	4301 (70.6)	1.00		
≥ 5	54 (31.0)	260 (32.5)	0.93 (0.65-1.33)	0.681		33 (32.7)	1794 (29.4)	1.16 (0.76-1.76)	0.503	
Days/week of moderate physical activity³,⁴										
0-2	54 (32.0)	227 (28.6)	1.00			31 (30.4)	1837 (30.5)	1.00		
3-4	37 (21.9)	202 (25.4)	0.75 (0.47-1.19)	0.221		33 (32.4)	1695 (28.1)	1.10 (0.67-1.81)	0.702	
5-7	78 (46.2)	367 (46.1)	0.84 (0.57-1.24)	0.375		38 (37.3)	2500 (41.4)	0.86 (0.53-1.39)	0.533	
Other factors										
Morbidity³,⁴										
None	147 (82.1)	710 (86.5)	1.00			87 (83.7)	5478 (87.9)	1.00		
≥1 condition	32 (17.9)	111 (13.5)	1.30 (0.84-2.02)	0.239		17 (16.3)	752 (12.1)	1.33 (0.78-2.25)	0.295	
BMI³,⁴										
Normal or underweight	82 (46.3)	351 (42.9)	1.00			42 (40.4)	2305 (37.2)	1.00		
Overweight	68 (38.4)	312 (38.1)	0.95 (0.66-1.37)	0.782		36 (34.6)	2645 (42.7)	0.68 (0.43-1.08)	0.100	
Obese	27 (15.3)	155 (18.9)	0.75 (0.46-1.20)	0.226		26 (25.0)	1246 (20.1)	1.11 (0.68-1.83)	0.675	

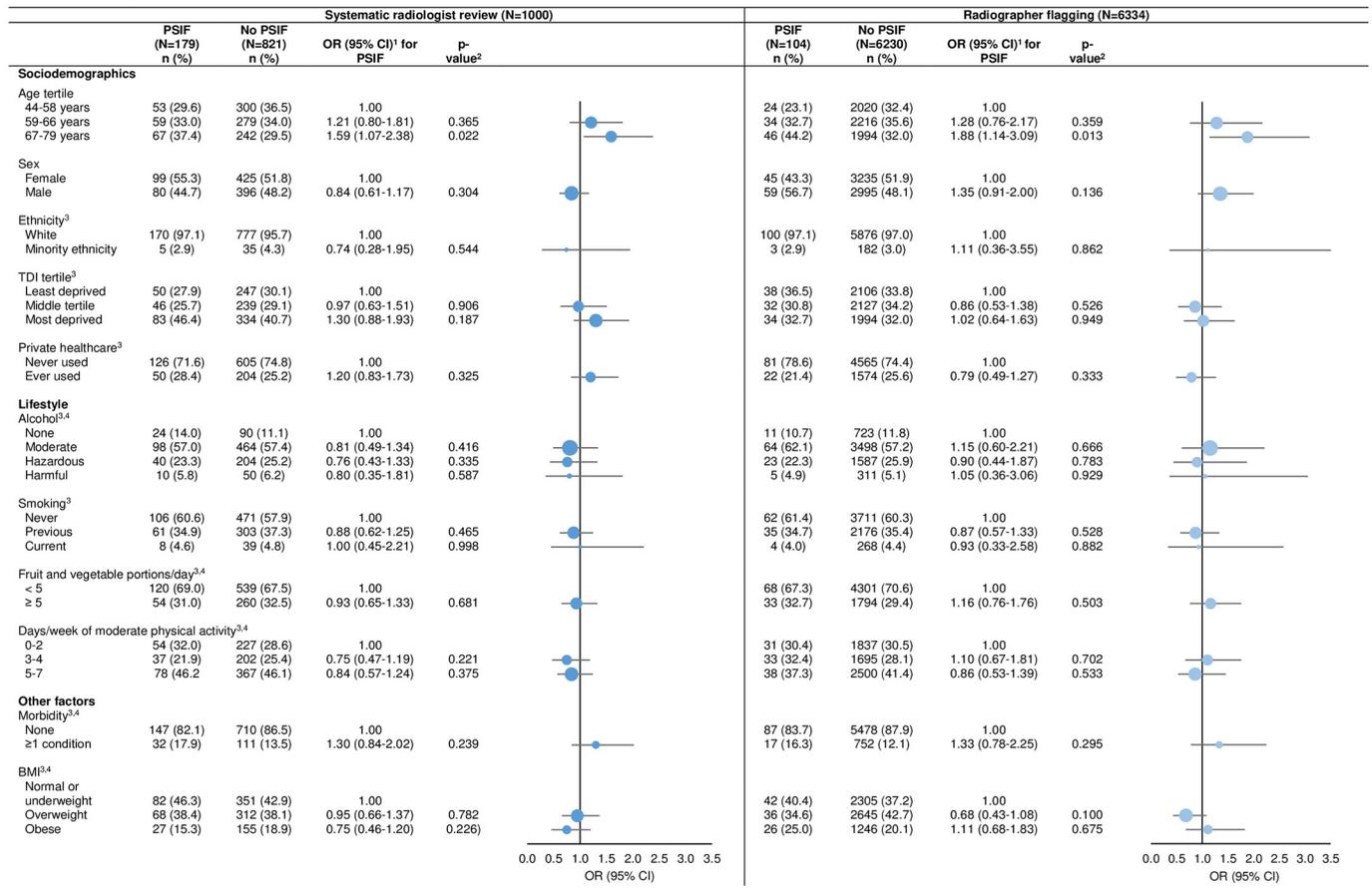OR (95% CI) axis: 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5

Fig 2. Age- and sex-adjusted odds ratios for potentially serious incidental findings (PSIFs) stratified by PSIFs protocol. PSIFs = potentially serious incidental findings, OR = odds ratio, CI = confidence interval, TDI = Townsend Deprivation Index, BMI = body mass index. Circles are weighted by the proportion of participants within a category. 1. Age- and sex-adjusted ORs, except age tertiles which are adjusted for sex only, and sex which is adjusted for age only. 2. p-value from Wald test. 3. Data were missing for ethnicity (13/1000 [1.3%], 73/6334 [1.2%]), TDI (1/1000 [0.1%], 3/6334 [<0.0%]), private healthcare use (15/1000 [1.5%], 92/6334 [1.5%]), alcohol (20/1000 [2.0%], 112/6334 [1.8%]), smoking (12/1000 [1.2%], 78/6334 [1.2%]), fruit and vegetable intake (27/1000 [2.7%], 138/6334 [2.2%]), physical activity (35/1000 [3.5%], 200/6334 [3.2%]) and BMI (5/1000 [0.5%], 34/6334 [0.5%]), for participants assessed by systematic radiologist review and by radiographer flagging respectively. 4. We calculated alcohol intake, fruit and vegetable intake, physical activity, morbidity and BMI as described in the footnotes to Table 1.

associations with PSIFs and medical history of cardiac disease [10], psychotic episodes [40] or human-immunodeficiency virus [41]. Given the varying nature of PSIFs (tumours, aneurysms etc.), a common biological risk factor seems unlikely. Instead, we captured morbidity using the Elixhauser Index, which comprises 30 conditions [37, 38]. There was no convincing association between morbidity and either PSIFs or serious final diagnoses, but sparse data on both of these outcomes and exposure data on morbidity (which may be secondary to healthy volunteer bias and a relatively short period of retrospective capture within linked hospital admissions data, chosen to limit any bias that may arise from changes in healthcare record coding practices in 2012 [39]) may have attenuated any true association. Furthermore, different definitions of morbidity may well produce different results.

Large studies are needed to investigate the factors associated with PSIFs and with serious final diagnoses, as these outcomes are relatively rare, particularly under a protocol of radiographer flagging. Our study is the largest so far to investigate the factors associated with PSIFs, and the first to investigate factors associated with serious final diagnoses, in unselected, healthy
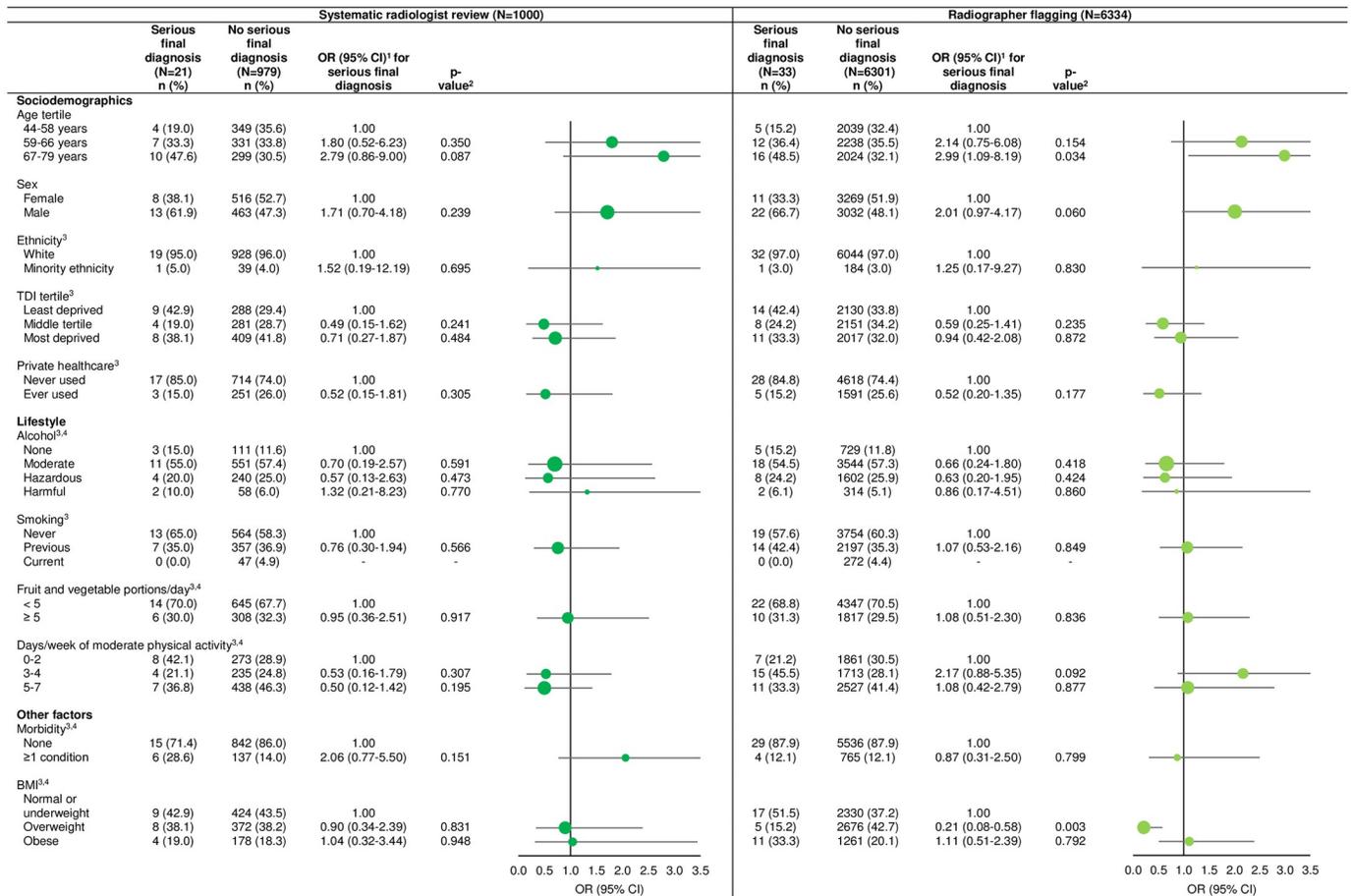
	Systematic radiologist review (N=1000)				Radiographer flagging (N=6334)			
	Serious final diagnosis (N=21) n (%)	No serious final diagnosis (N=979) n (%)	OR (95% CI)¹ for serious final diagnosis	p-value²	Serious final diagnosis (N=33) n (%)	No serious final diagnosis (N=6301) n (%)	OR (95% CI)¹ for serious final diagnosis	p-value²
Sociodemographics								
Age tertile								
44-58 years	4 (19.0)	349 (35.6)	1.00		5 (15.2)	2039 (32.4)	1.00	
59-66 years	7 (33.3)	331 (33.8)	1.80 (0.52-6.23)	0.350	12 (36.4)	2238 (35.5)	2.14 (0.75-6.08)	0.154
67-79 years	10 (47.6)	299 (30.5)	2.79 (0.86-9.00)	0.087	16 (48.5)	2024 (32.1)	2.99 (1.09-8.19)	0.034
Sex								
Female	8 (38.1)	516 (52.7)	1.00		11 (33.3)	3269 (51.9)	1.00	
Male	13 (61.9)	463 (47.3)	1.71 (0.70-4.18)	0.239	22 (66.7)	3032 (48.1)	2.01 (0.97-4.17)	0.060
Ethnicity³								
White	19 (95.0)	928 (96.0)	1.00		32 (97.0)	6044 (97.0)	1.00	
Minority ethnicity	1 (5.0)	39 (4.0)	1.52 (0.19-12.19)	0.695	1 (3.0)	184 (3.0)	1.25 (0.17-9.27)	0.830
TDI tertile³								
Least deprived	9 (42.9)	288 (29.4)	1.00		14 (42.4)	2130 (33.8)	1.00	
Middle tertile	4 (19.0)	281 (28.7)	0.49 (0.15-1.62)	0.241	8 (24.2)	2151 (34.2)	0.59 (0.25-1.41)	0.235
Most deprived	8 (38.1)	409 (41.8)	0.71 (0.27-1.87)	0.484	11 (33.3)	2017 (32.0)	0.94 (0.42-2.08)	0.872
Private healthcare³								
Never used	17 (85.0)	714 (74.0)	1.00		28 (84.8)	4618 (74.4)	1.00	
Ever used	3 (15.0)	251 (26.0)	0.52 (0.15-1.81)	0.305	5 (15.2)	1591 (25.6)	0.52 (0.20-1.35)	0.177
Lifestyle								
Alcohol³,⁴								
None	3 (15.0)	111 (11.6)	1.00		5 (15.2)	729 (11.8)	1.00	
Moderate	11 (55.0)	551 (57.4)	0.70 (0.19-2.57)	0.591	18 (54.5)	3544 (57.3)	0.66 (0.24-1.80)	0.418
Hazardous	4 (20.0)	240 (25.0)	0.57 (0.13-2.63)	0.473	8 (24.2)	1602 (25.9)	0.63 (0.20-1.95)	0.424
Harmful	2 (10.0)	58 (6.0)	1.32 (0.21-8.23)	0.770	2 (6.1)	314 (5.1)	0.86 (0.17-4.51)	0.860
Smoking³								
Never	13 (65.0)	564 (58.3)	1.00		19 (57.6)	3754 (60.3)	1.00	
Previous	7 (35.0)	357 (36.9)	0.76 (0.30-1.94)	0.566	14 (42.4)	2197 (35.3)	1.07 (0.53-2.16)	0.849
Current	0 (0.0)	47 (4.9)	-	-	0 (0.0)	272 (4.4)	-	-
Fruit and vegetable portions/day³,⁴								
< 5	14 (70.0)	645 (67.7)	1.00		22 (68.8)	4347 (70.5)	1.00	
≥ 5	6 (30.0)	308 (32.3)	0.95 (0.36-2.51)	0.917	10 (31.3)	1817 (29.5)	1.08 (0.51-2.30)	0.836
Days/week of moderate physical activity³,⁴								
0-2	8 (42.1)	273 (28.9)	1.00		7 (21.2)	1861 (30.5)	1.00	
3-4	4 (21.1)	235 (24.8)	0.53 (0.16-1.79)	0.307	15 (45.5)	1713 (28.1)	2.17 (0.88-5.35)	0.092
5-7	7 (36.8)	438 (46.3)	0.50 (0.12-1.42)	0.195	11 (33.3)	2527 (41.4)	1.08 (0.42-2.79)	0.877
Other factors								
Morbidity³,⁴								
None	15 (71.4)	842 (86.0)	1.00		29 (87.9)	5536 (87.9)	1.00	
≥1 condition	6 (28.6)	137 (14.0)	2.06 (0.77-5.50)	0.151	4 (12.1)	765 (12.1)	0.87 (0.31-2.50)	0.799
BMI³,⁴								
Normal or underweight	9 (42.9)	424 (43.5)	1.00		17 (51.5)	2330 (37.2)	1.00	
Overweight	8 (38.1)	372 (38.2)	0.90 (0.34-2.39)	0.831	5 (15.2)	2676 (42.7)	0.21 (0.08-0.58)	0.003
Obese	4 (19.0)	178 (18.3)	1.04 (0.32-3.44)	0.948	11 (33.3)	1261 (20.1)	1.11 (0.51-2.39)	0.792

(Forest plot x-axes: OR (95% CI), scale 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 for both panels)

Fig 3. Age- and sex-adjusted odds ratios for serious final diagnoses stratified by PSIFs protocol. OR = odds ratio, CI = confidence interval, TDI = Townsend Deprivation Index, BMI = body mass index. Circles are weighted by the proportion of participants within a category. 1. Age- and sex-adjusted ORs, except age tertiles which are adjusted for sex only, and sex which is adjusted for age only. 2. p-value from Wald test. 3. Data were missing as described in Fig 2, footnote 3. 4. We calculated alcohol intake, fruit and vegetable intake, physical activity, morbidity and BMI as described in the footnotes to Table 1.

https://doi.org/10.1371/journal.pone.0218267.g003

participants undergoing MRI of any body region. Our sample is approximately 25% larger than the largest previous study of factors associated with PSIFs detected on brain MRI (N = 5,800) [8] and 50 times larger than the largest previous such study of multi-region MRI (N = 148) [16]. We systematically followed-up 50% more participants for data on final diagnoses compared to the largest previous study (N = 188) [8]. Despite the size of our study, we still may have missed associations with PSIFs or final diagnoses due to sparsity of these outcomes within our cohort and small numbers within some exposure categories (e.g. minority ethnicity groups). Healthy volunteer selection bias likely affects the UK Biobank cohort, as participants are less deprived than non-participants and less likely to be obese, smoke, drink alcohol daily or have self-reported medical conditions compared to the general population [24]. The imaged cohort are then further selected, with lower proportions of people having more 'unhealthy' lifestyles; imaged participants have survived and also remain healthy enough to travel to the imaging centres and undergo the imaging assessment. As with all epidemiological studies which use self-reported data, our data on exposures may be further limited by reporting bias; participants may have inaccurately reported alcohol intake, smoking habits, physical activity and diet. The apparently reduced odds of serious final diagnoses in overweight participants may be

spurious, secondary to data sparsity of both the outcome and the exposure. The direction of an association (if any) between increased BMI and PSIFs is unclear. The associations between increased BMI and certain cancers [42] may lead to increased risk of PSIFs and serious final diagnoses; alternatively, risks may be reduced if people with increased BMI tend not to complete all MRI sequences, or imaging of all body regions.

Our classifications of 'serious' final diagnoses are based on clinical judgement using data collected up to six months after feedback of a PSIF. Reaching final diagnoses of some PSIFs may take longer [2]. Feedback of PSIFs may impact on non-medical domains such as emotional wellbeing, insurance and finances and work and activities, regardless of the health-related severity of the final diagnosis [2]. 'Severity' of a final diagnosis is therefore inherently difficult to judge, though we did show good agreement between two independent physicians' classifications using a medical-based definition.

By deliberately focusing our study on participants with PSIFs and serious final diagnoses our results inform on factors associated with findings which are likely to generate clinical assessment, and those with serious health consequences, respectively. While our cohort is not representative of the general population, exposure-outcome associations can be generalised to other populations [24, 43, 44], to inform the design of appropriate IFs handling policies, which are required by major funders,[7] and of materials to facilitate the informed consent of potential research participants.

Compared to sociodemographic, lifestyle and health-related factors, the protocol for identifying PSIFs protocol has by far the largest influence on the generation of PSIFs and serious final diagnoses. As the majority of PSIFs do not turn out to be serious, but feedback generates clinical assessments and negative impacts on emotional wellbeing, insurance and finances and work and activities [2], our study suggests that researchers have the opportunity to greatly influence (for better or worse) the potential harms done to participants and the burden on publicly-funded health services. There remain many unanswered questions on the impacts of different methodologies to feedback research results to participants [45]; to inform future policy design, evaluations of the impacts of different protocols are paramount.

PSIFs are rare, and few are finally diagnosed as serious disease; hence large studies are needed to investigate the associated factors. This study represents the largest such cohort so far. Furthermore, since 100,000 participants will complete the UK Biobank imaging assessment over the next few years, it will in due course be possible to update these analyses with a substantially larger sample size, providing more comprehensive and statistically better powered estimates of the factors associated with PSIFs and with serious final diagnoses.

Supporting information

S1 File. STROBE checklist.
(DOCX)

Acknowledgments

Imperial College London

Prof Paul Matthews (Edmond and Lily Safra Chair and Head of Brain Sciences)
   Dr Tony Goldstone (Senior Clinical Research Fellow)

Newcastle University

Prof Andrew Blamire (Professor of Magnetic Resonance Physics)

Queen Mary University of London

Prof Steffen Petersen (Professor of Cardiovascular Medicine)

University of Edinburgh

Prof Cathie Sudlow (Chair of Neurology and Clinical Epidemiology)
   Dr Lorna Gibson (Clinical Research Fellow and Specialty Trainee in Clinical Radiology)

University of Manchester

Prof Alan Jackson (Professor of Radiology)

University of Oxford

Prof Naomi Allen (Associate Professor in Epidemiology)
   Prof Sir Rory Collins (BHF Professor of Medicine and Epidemiology)
   Prof Paul Leeson (Professor of Cardiovascular Medicine)
   Prof Karla Miller (Professor of Biomedical Engineering)
   Prof Stefan Neubauer (Professor of Cardiovascular Medicine)
   Prof Stephen Smith (Professor of Biomedical Engineering)

University of Southampton

Prof Nicholas Harvey (Professor of Rheumatology and Clinical Epidemiology)

University of Westminster

Prof Jimmy Bell (Professor)

Author Contributions

Conceptualization: Lorna M. Gibson, Naomi E. Allen, Joanna M. Wardlaw, Cathie L. M. Sudlow.

Data curation: Lorna M. Gibson, John Nolan, Thomas J. Littlejohns, Edouard Mathieu, Steve Garratt, Cathie L. M. Sudlow.

Formal analysis: Lorna M. Gibson, John Nolan.

Funding acquisition: Lorna M. Gibson, Naomi E. Allen, Joanna M. Wardlaw, Cathie L. M. Sudlow.

Investigation: Lorna M. Gibson, John Nolan.

Methodology: Lorna M. Gibson, Caroline A. Jackson, Cathie L. M. Sudlow.

Project administration: Nicola Doherty.

Software: Lorna M. Gibson, John Nolan.

Supervision: Joanna M. Wardlaw, Caroline A. Jackson, Cathie L. M. Sudlow.

Visualization: Lorna M. Gibson.

Writing – original draft: Lorna M. Gibson.

Writing – review & editing: John Nolan, Thomas J. Littlejohns, Edouard Mathieu, Steve Garratt, Steffen Petersen, Nicholas C. W. Harvey, Jonathan Sellors, Naomi E. Allen, Joanna M. Wardlaw, Caroline A. Jackson, Cathie L. M. Sudlow.

References

1. Wolf SM, Lawrenz FP, Nelson CA, Kahn JP, Cho MK, Clayton EW, et al. Managing incidental findings in human subjects research: analysis and recommendations. The Journal of law, medicine & ethics: a journal of the American Society of Law, Medicine & Ethics. 2008; 36(2):219–48.

2. Gibson LM, Littlejohns TJ, Adamska L, Garratt S, Doherty N, Group. UBIW, et al. Impact of detecting potentially serious incidental findings during multi-modal imaging [version 3; referees: 2 approved, 1 approved with reservations]. Wellcome Open Res. 2018; 2:114.

3. Hegedus P, von Stackelberg O, Neumann C, Selder S, Werner N, Erdmann P, et al. How to report incidental findings from population whole-body MRI: view of participants of the German National Cohort. European radiology. 2019. Epub 2019/03/28. https://doi.org/10.1007/s00330-019-06077-z PMID: 30915558.

4. Vernooij MW, Ikram MA, Tanghe HL, Vincent AJ, Hofman A, Krestin GP, et al. Incidental findings on brain MRI in the general population. The New England journal of medicine. 2007; 357(18):1821–8. Epub 2007/11/06. 17978290. https://doi.org/10.1056/NEJMoa070972 PMID: 17978290

5. Kirschen MP, Jaworska A, Illes J. Subjects' expectations in neuroimaging research. Journal of magnetic resonance imaging: JMRI. 2006; 23(2):205–9. Epub 2006/01/18. https://doi.org/10.1002/jmri.20499 PMID: 16416438.

6. Medical Research Council, Wellcome Trust. Framework on the feedback of health-related findings in research. 2014 [1st November 2017]. www.mrc.ac.uk/documents/pdf/mrc-wellcome-trust-framework-on-the-feedback-of-health-related-findings-in-researchpdf/.

7. Farrar J, Savill J. UK funders' framework for health-related findings in research. The Lancet. 2014; 383 (9933):1954–5. https://doi.org/10.1016/S0140-6736(14)60545-6

8. Bos D, Poels MMF, Adams HHH, Akoudad S, Cremers LGM, Zonneveld HI, et al. Prevalence, clinical management, and natural course of incidental findings on brain MR images: the population-based Rotterdam Scan Study. Radiology. 2016; 281(2):507–15. https://doi.org/10.1148/radiol.2016160218 PMID: 27337027.

9. Yue NC, Longstreth WT Jr., Elster AD, Jungreis CA, O'Leary DH, Poirier VC. Clinically serious abnormalities found incidentally at MR imaging of the brain: data from the Cardiovascular Health Study. Radiology. 1997; 202(1):41–6. Epub 1997/01/01. https://doi.org/10.1148/radiology.202.1.8988190 PMID: 8988190.

10. Tsushima Y, Taketomi-Takahashi A, Endo K. Prevalence of abnormal findings on brain magnetic resonance (MR) examinations in adult participants of brain docking. BMC neurology. 2005; 5:18. Epub 2005/10/07. https://doi.org/10.1186/1471-2377-5-18 PMID: 16207376.

11. Håberg AK, Hammer TA, Kvistad KA, Rydland J, Müller TB, Eikenes L, et al. Incidental intracranial findings and their clinical impact: the HUNT MRI study in a general population of 1006 participants between 50–66 years. PLoS ONE. 2016; 11(3):e0151080. https://doi.org/10.1371/journal.pone.0151080

12. Sandeman EM, Hernandez MDCV, Morris Z, Bastin ME, Murray C, Gow AJ, et al. 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. https://doi.org/10.1371/journal.pone.0071467 PMID: 23967214

13. Brugulat-Serrat A, Rojas S, Bargallo N, Conesa G, Minguillon C, Fauria K, et al. Incidental findings on brain MRI of cognitively normal first-degree descendants of patients with Alzheimer's disease: a cross-sectional analysis from the ALFA (Alzheimer and Families) project. BMJ open. 2017; 7(3):e013215. Epub 2017/03/28. https://doi.org/10.1136/bmjopen-2016-013215 PMID: 28341686.

14. Kumar R, Sachdev PS, Price JL, Rosenman S, Christensen H. Incidental brain MRI abnormalities in 60- to 64-year-old community-dwelling individuals: data from the Personality and Total Health Through Life study. Acta neuropsychiatrica. 2008; 20(2):87–90. Epub 2008/04/01. https://doi.org/10.1111/j.1601-5215.2008.00273.x PMID: 25385468.

15. Illes J, Rosen AC, Huang L, Goldstein RA, Raffin TA, Swan G, et al. Ethical consideration of incidental findings on adult brain MRI in research. Neurology. 2004; 62(6):888–90. Epub 2004/03/24. https://doi.org/10.1212/01.wnl.0000118531.90418.89 PMID: 15037687.

16. Morin SHX, Cobbold JFL, Lim AKP, Eliahoo J, Thomas EL, Mehta SR, et al. Incidental findings in healthy control research subjects using whole-body MRI. Eur J Radiol. 2009; 72(3):529–33. https://doi.org/10.1016/j.ejrad.2008.08.006 PMID: 18818038

17. Hartwigsen G, Siebner HR, Deuschl G, Jansen O, Ulmer S. Incidental findings are frequent in young healthy individuals undergoing magnetic resonance imaging in brain research imaging studies: a prospective single-center study. Journal of computer assisted tomography. 2010; 34(4):596–600. Epub 2010/07/27. https://doi.org/10.1097/RCT.0b013e3181d9c2bb PMID: 20657230.

18. UK Biobank. UK Biobank Imaging Study 2018 [26/09/18]. https://imaging.ukbiobank.ac.uk/.

19. Matthews PM, Sudlow C. The UK Biobank. Brain. 2015; 138(12):3463–5. https://doi.org/10.1093/brain/awv335 PMID: 26598488

20. Gibson LM, Sellors J, Sudlow CLM. Management of incidental findings on multimodal imaging in UK Biobank. In: Weckbach S, editor. Incidental Radiological Findings. Switzerland: Springer; 2016. p. 71–8.

21. 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. Journal of Clinical Epidemiology. 2008; 61(4):344–9. https://doi.org/10.1016/j.jclinepi.2007.11.008 PMID: 18313558

22. Gibson LM, Nolan J. Factors associated with potentially serious incidental findings and with serious final diagnoses on multimodal imaging in the UK Biobank Imaging Study: a prospective cohort study, [software]: University of Edinburgh, Usher Institute of Population Health Sciences and Informatics,; 2018. https://datashare.is.ed.ac.uk/handle/10283/3112.

23. UK Biobank. UK Biobank imaging assessment visit: incidental findings 2016 [26/09/18]. http://www.ukbiobank.ac.uk/wp-content/uploads/2017/11/incidental-findings-2.pdf.

24. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with the general population. Am J Epidemiol. 2017; 186:1026–34. Epub 2017/06/24. https://doi.org/10.1093/aje/kwx246

25. UK Biobank. UK Biobank: report on feasibility of imaging assessments in 100,000 participants. 2015.

26. UK Biobank. UK Biobank brain imaging documentation 2016 [14th March 2018]. http://biobank.ctsu.ox.ac.uk/crystal/docs/brain_mri.pdf.

27. UK Biobank. UK Biobank imaging modality cardiovascular magnetic resonance (CMR) 2015 [14th March 2018]. http://biobank.ctsu.ox.ac.uk/crystal/docs/cardiac_mri_explan.pdf.

28. UK Biobank. UK Biobank body (abdominal) MRI scan 2015 [14th March 2018]. http://biobank.ctsu.ox.ac.uk/crystal/docs/body_mri_explan.pdf.

29. UK Biobank. UK Biobank imaging modality DXA 2015 [14th March 2018]. http://biobank.ctsu.ox.ac.uk/crystal/docs/DXA_explan_doc.pdf.

30. UK Biobank. UK Biobank imaging modality: carotid ultrasound 2015 [14th March 2018]. http://biobank.ctsu.ox.ac.uk/crystal/docs/carult_explan_doc.pdf.

31. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015; 12(3):e1001779. https://doi.org/10.1371/journal.pmed.1001779 PMID: 25826379.

32. UK Biobank. List of available fields 2017 [14th March 2018]. http://biobank.ctsu.ox.ac.uk/crystal/util/ukb_field.txt.

33. British Medical Association. Alcohol: guidelines on sensible drinking. London, UK: 1995.

34. Cassidy S, Chau JY, Catt M, Bauman A, Trenell MI. Cross-sectional study of diet, physical activity, television viewing and sleep duration in 233,110 adults from the UK Biobank; the behavioural phenotype of cardiovascular disease and type 2 diabetes. BMJ open. 2016; 6(3):e010038. Epub 2016/03/24. https://doi.org/10.1136/bmjopen-2015-010038 PMID: 27008686.

35. UK Biobank. UK Biobank touch-screen questionnaire: final version 2017 [18th March 2018]. http://biobank.ctsu.ox.ac.uk/crystal/docs/TouchscreenQuestionsMainFinal.pdf.

36. World Health Organisation. Obesity: preventing and managing the global epidemic 2000 [14th March 2018]. http://apps.who.int/iris/bitstream/10665/42330/1/WHO_TRS_894.pdf.

37. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Medical care. 1998; 36(1):8–27. Epub 1998/02/07. PMID: 9431328.

38. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Medical care. 2005; 43(11):1130–9. Epub 2005/10/15. PMID: 16224307.

39. Health and Social Care Information Centre. HRG4+ companion. 2013.

40. Sommer IE, de Kort GA, Meijering AL, Dazzan P, Hulshoff Pol HE, Kahn RS, et al. How frequent are radiological abnormalities in patients with psychosis? A review of 1379 MRI scans. Schizophrenia bulletin. 2013; 39(4):815–9. Epub 2012/03/15. https://doi.org/10.1093/schbul/sbs037 PMID: 22416264.

41. Loy A, Morgan R, O'Dea S, Daly C, Mulcahy F. Clinically significant extra-cardiac findings in asymptomatic HIV-positive men undergoing cardiac magnetic resonance imaging. International journal of STD & AIDS. 2015; 26(5):346–51. Epub 2014/05/30. https://doi.org/10.1177/0956462414538005 PMID: 24872375.

42. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UK adults. Lancet. 384 (9945):755–65. https://doi.org/10.1016/S0140-6736(14)60892-8

43. Manolio TA, Collins R. Enhancing the feasibility of large cohort studies. JAMA: the journal of the American Medical Association. 2010; 304(20):2290–1. https://doi.org/10.1001/jama.2010.1686 PMID: 21098774

44. Collins R. What makes UK Biobank special? The Lancet. 2012; 379(9822):1173–4. http://dx.doi.org/10.1016/S0140-6736(12)60404-8.

45. Wong CA, Hernandez AF, Califf RM. Return of research results to study participants: uncharted and untested. Jama. 2018; 320(5):435–6. Epub 2018/06/23. https://doi.org/10.1001/jama.2018.7898 PMID: 29931289.