BMJ Open

A protocol for prospective study of vitamin D obesity, and leptin in relation to bladder cancer, incidence and survival.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-019309
Article Type:	Protocol
Date Submitted by the Author:	24-Aug-2017
Complete List of Authors:	Gislefoss, Randi; Cancer Registry og Norway, Dept of Research Stenehjem, Jo; Cancer Registry of Norway, Dept. of Research Andreassen, Bettina Kulle ; Kreftregisteret Langseth, Hilde; Kreftregisteret Axcrona, Karol; Akershus Universitetssykehus HF Weiderpass, Elisabete; Kreftregisteret Mondul, Alison; University of Michigan, Scool of Public Health Robsahm, Trude; Cancer Registry of Norway, Research
Primary Subject Heading:	Oncology
Secondary Subject Heading:	Urology
Keywords:	Bladder cancer, Biobank, vitamin D, obesity

A protocol for prospective study of vitamin D, obesity, and leptin in relation to bladder cancer, incidence and survival

Authors: Randi E Gislefoss[1], Jo S Stenehjem[1], Bettina Andreassen[1], Hilde Langseth[1], Karol Axcrona[2], Elisabete Weiderpass[1], Alison Mondul[3], Trude E Robsahm[1]

[1]Department of Research, Cancer Registry of Norway, Oslo, Norway

[2]Department of Urology, Akershus University Hospital, Lørenskog, Norway

[3]University of Michigan, School of Public Health, Ann Arbor, MI, USA

Correspondence to:

Dr. Randi Elin Gislefoss: randi.gislefoss@kreftregisteret.no

Word count: 2434

Numbers of tables and figures: 2

Abstract

Introduction: Bladder cancer (BC) (including renal pelvis, ureter and urethra) is one of the most common urogenital cancers and the fourth most frequent cancer in men in the United States. In Norway, the incidence of BC has increased over the last decades. The age-standardized incidence rates per 100,000 for 2011-15 were 50.4 in men and 15.5 in women. Compared to the 5-year period 2006-2010, the percentage increase in incidence was 9.4% in men and 8.9% in women. The recurrence rate of BC is over 50%, the highest recurrence rate of any malignancy. Smoking and occupational exposure to aromatic amines are recognized as the major risk factors. Recently, low serum level of 25(OH)D and obesity have been suggested to increase the BC risk and leptin, which is important in weight regulation, may be involved in bladder carcinogenesis. More knowledge on potential risk factors for BC is necessary for planning and implementing primary prevention measures.

Methods and analyses: Cohort and nested case-control studies will be carried out using the population-based Janus Serum Bank Cohort consisting of pre-diagnostic sera, clinical measurement data (body height and weight, blood pressure, cholesterol and triglycerides) and self-reported information on lifestyle factors (smoking, physical activity). Participants were followed from cohort inclusion (1972-2003) through 2014. The cohort will be linked to the Cancer Registry of Norway (cancer data), the national Cause of Death Registry (date and cause of death), National Population Registry (vital status) and Statistic Norway (education and occupation). Serum samples will be analyzed for 25(OH)D, vitamin D binding protein, leptin, albumin, calcium and parathyroid hormone.
Cox regression and conditional logistic regression models will be used to estimate association between the exposures and BC.

Ethics and dissemination: The study has been approved by the Regional Committee for Medical Research Ethics and is funded by the Norwegian Cancer Society. Results will be published in peer-reviewed journals, at scientific conferences and through press releases.

Strengths and limitations of this study

- Use of a large sample set of more than 2300 incident BC cases
- Pre-diagnostic serum samples assure the temporality of the relationship between exposure and BC, limiting the possibility of reverse causality.
- Use of unique personal identification number for linkages between multiple data sources, to establish a virtually complete study file and complete ascertainment of follow-up.
- Reviewing and characterizing T-stage for all BC.
- Cases of the premalignant stages (papillary transitional carcinoma and carcinoma in situ) are not included in sub-study II.

Introduction

Rationale and evidence gaps

Urinary bladder cancer (BC) (including renal pelvis, ureter and urethra) is the most common urogenital cancer after prostate cancer and is the fourth most frequent cancer in men in the United States[1]. BC has over a 50 % recurrence rate, the highest of any malignancy, and is one of the most expensive cancers to treat on a per-patient basis [2]. In Norway, the incidence of cancer of the bladder has been increasing over the last decades. In 2015, the age-standardized incidence rates were 53.7 and 16.5 in Norwegian men and women, respectively. Compared to the 5-year period 2006-2010, the increase in incidence has been 6.1 % in men and 12.3 % in women. Up to 50 % of all BC cases have been ascribed to smoking [3], and 5–25 % of the cases have been attributable to occupational exposures [4]; still, the etiology of up to 45 % of BC remains unexplained. Low serum level of 25(OH)D and obesity have been suggested to increase BC risk, and the hormone leptin, which is important in weight regulation, may be involved in its carcinogenenic process [5 6].

25(OH)D is converted to its active hormonal form, 1-25-dihydroxyvitamin D ($1,25(OH)_2D$), by 1-α-hydroxylase, which is present in most tissues in the body [7]. PTH and calcium level are important factors as they affect the enzymatic conversion from 25(OH)D to active $1, 25-(OH)D_3$ in the kidney and may be involved in non-classical synthesis. Measurement of circulating 25(OH)D is considered the gold standard measurement of vitamin D status as it integrates vitamin D exposure from oral intake from diet or supplements, as well as from

exposure to ultraviolet radiation (UVR) [8 9]. Despite being the gold standard, total circulating 25(OH)D may not be the best measure of 25(OH)D exposure for all tumors. The "free hormone hypothesis" suggests that only unbound, free hormones can have biologic effects on target tissues [10]. To date, few studies have examined the role of vitamin D binding protein (DBP) also known as group-specific component or Gc-globulin, in the association between 25(OH)D and various cancer processes. DBP transports both 25(OH)D and $1,25(OH)_2D$ in circulation. This protein carries 88% of 25(OH)D and 85% of $1,25(OH)_2D$; an additional 12% of 25(OH)D and 15% of $1,25(OH)_2D$ circulate bound to albumin. Clinical laboratory assays of circulating 25(OH)D that are currently in use measure total 25(OH)D without differentiating between the bound and free forms. Thus, it remains unclear whether total or free 25(OH)D is more biologically relevant with respect to risk of BC. Two previous studies have examined free, in addition to total 25(OH)D, in relation to risk of BC. One found an inverse association between total 25(OH)D and bladder cancer that appeared to be restricted to participants with low DBP, suggesting that free 25(OH)D might be more strongly associated with risk of bladder cancer than total 25(OH)D [11]. The other study found no association between 25(OH)D overall or at any level of DBP concentration [12].

Body mass index (weight/height^2, BMI) is a reliable indicator of body fatness and can be categorized as underweight (<18.5), healthy weight (18.5-24.9), overweight (25-29.9) and obese (> 30). The prevalence of obesity in Norway has risen steeply the last decades. In the 1970s, about 15% of men and women were overweight (www.fhi.no/) while in 2013 the proportions were 58.4% and 47.3%, respectively [13]. Two meta-analyses, including 15 [14] and 11 [15] cohort studies conclude that obesity significantly increases the risk of BC. Leptin, a hormone involved in weight regulation [16], may be involved in this potential association. High leptin levels have been shown to impact development of several forms of cancer [17]. A study of Yuan et al. [6] shows that leptin receptors are aberrantly expressed in BC tissue and a high leptin level has been associated with BC carcinogenesis [5 6].

Low 25(OH)D levels are more frequent in obese persons, suggesting that 25(OH)D deficiency is associated with obesity and vice versa. 25(OH)D deficiency is suggested to be associated with obesity [18-20], and both low 25(OH)D and obesity are suggested to contribute to development of BC.

Aims and hypotheses

Better-targeted BC primary and tertiary prevention (risk and survival) is warranted. The interplay between 25(OH)D and obesity and their associations with BC risk are poorly understood. We propose a study aiming to examine BMI, serum levels of total and free 25(OH)D, and leptin levels in relation to BC risk and survival by using samples from Janus Serum Bank and associated data from population-based registries and surveys.

We hypothesize that

1 a. Obesity is associated with increased BC risk;

1b. Obesity is associated with reduced BC survival;

2a. Low free and total 25(OH)D level and high serum leptin levels (>4.1 ng/mL are associated with increased risk of BC;

2b. Low free and total 25(OH)D level and high serum leptin levels (>4.1 ng/mL are associated with reduced BC survival.

METHODS AND ANALYSIS

Study population and data sources

The Janus Serum Bank Cohort

The study will be carried out using the Janus Serum Bank Cohort, a population-based biobank for prospective cancer studies, containing serum samples and questionnaire data from 292,851 Norwegians who participated in one or more of five Norwegian Regional Health Studies in the period 1972–2003. Detailed description of the samples and data included in the Janus Serum Bank Cohort has been published elsewhere [21 22]. The quality aspects of long-term stored samples have been of high priority in the Janus Serum Bank and component stability for a large number of hormones, proteins, metabolites and electrolytes has been investigated [23-25], including for both 25(OH)D, and leptin [26-29]. A unique 11-digit personal identification number (PIN), assigned to al Norwegian residents, will be used to link the Janus Cohort with population-based registries and surveys.

Population-based registries and surveys

The Cancer Registry of Norway (CRN) has collected notifications on cancer at a national level since 1953. Cancer reporting is mandatory by law, and reports from various sources ensures high quality and completeness (98.8%) [30]. The reporting system, based on pathology and cytology reports, clinical records, and death certificates, provides information about site, histological type and stage of disease at the time of diagnosis. CRN has been involved in the Janus Serum Bank operation since establishment in the early 1970s and has been responsible for the data handling; in 2004 the serum bank was integrated into the CRN. The following information is available for cancer cases: month and year of diagnosis, tumor site (International Classification of Diseases 7th revision [ICD-7 codes] converted into ICD-10 codes), histology (codes from ICD-Oncology 2nd and 3rd revision), clinical stage (local = no metastases, regional = metastasis in regional lymph nodes or surrounding area, distant = distant metastasis). In addition, all BC diagnoses in the Janus cohort will be reviewed and assigned a pathological T-stage, according to the AJCC 8th ed.[31].

The Norwegian Institute of Public Health (NIPH) has been responsible for conducting the national health surveys, upon which the Janus Serum Bank is partly based. All participants have completed questionnaires for assessment of lifestyle factors (i.e. smoking habits, alcohol use), at the time of serum collection. A database has been established, including data from these questionnaires, as well as measured body height and weight, blood pressure, cholesterol and triglycerides [32 33]. The Janus Cohort includes participants from five of the health studies: The Oslo Study I (1972–73), The Norwegian Counties Study (1974–78, 1977–83, and 1985–88), The Age 40 Program- Oslo (1981–99), The National Age 40 Program (1985–99) and The TROFINN Health Study (2001–03). A set of about 50 variables has been harmonized and standardized due to slightly different wording in the questionnaires [22]. Available variables include: height (cm), weight (kg), BMI (kg/m^2, categorized as 12–18.49, 18.5–24.9, 25.0–29.9, ≥30), smoking status (never, former, current), cigarettes per day (1–9, 10–14, ≥15), years of smoking (1–9, 10–29, ≥30), time since smoking cessation (<3mos, 3mos–1yr, 1–5yrs, >5yrs), and total physical activity (inactive, low, medium, high), based on leisure time activity.

The Cause of death Registry has registered death certificates for all deaths in Norway since 1951. Cause of death registration is mandatory by law.

National Population Registry contains information on vital status (alive, emigrated or dead) of everyone that resides or has resided in Norway.

Statistics Norway has the responsibility of covering the needs for official statistics on the Norwegian population including individual data on settlements, migration, occupation and level of education.

Using the 11-digit PIN number we will link data from four different sources to set up the research file, illustrated in Figure 1.

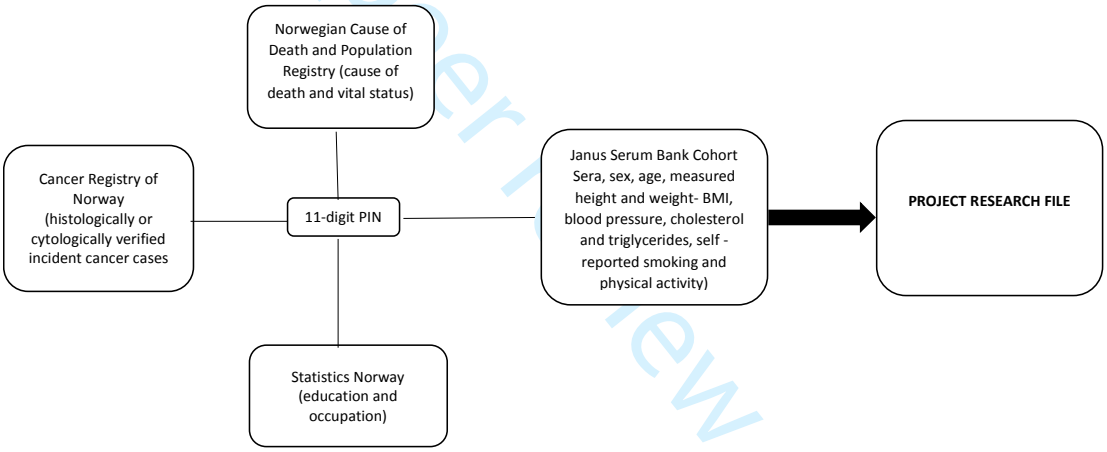

Figure 1. Data collection from different sources using PIN

Study design

Sub-study I: a prospective cohort study

In a prospective cohort study among all individuals in the Janus Serum Bank Cohort (n = 292,851) (Figure 1), we will explore baseline BMI in relation to bladder cancer risk. Among the included BC cases we will investigate baseline BMI and BC survival. By 2014, the cohort included 2 347 BC cases (ICD-10: C67). BC cases of both muscle invasive and non-invasive urothelial cell carcinoma, will be included in the study. Educational level, occupation, age, sex, physical activity, smoking habits, cholesterol, triglycerides and blood pressure will be included in the statistical analyses as confounders.

Sub-study II: a nested case-control study

The study will be nested within the prospective cohort described above (Study I), including a) 400 BC cases of muscle invasive urothelial cancer, and b) 400 controls alive and without a cancer diagnosis at the time of the BC diagnosis of the cases, matched 1:1 on sex, age (+/-1 year) and date of serum sampling (+/-1 month). The serum samples will be analysed for total 25(OH)D, vitamin D binding protein and leptin. As parathyroid hormone (PTH) and albumin-adjusted calcium level affect the enzymatic conversion from 25(OH)D to active 1,25-(OH)D_3 and might be involved in the non-classical synthesis as well, measurement of these components will be taken into account.

Statistical methods

In the cohort study, Cox regression models will be used to estimate hazard ratios (HR) with 95% confidence intervals (CI) of BC and survival after BC, taking into account stage at diagnosis, and BMI and including adjusting for season of blood collection for vitamin D. In the nested case-control study, conditional logistic regression models will be used to estimate the odds ratio (OR) with 95% CI.

As we have a number of potential confounding variables, we will use directed acyclic graphs to select variables to include in the statistical models. Confounding variables will be included in the models and tests of interaction effects will be performed when relevant.

All tests will be two-sided and $p<0.05$ will be considered statistically significant. All statistical analyses will be conducted using Stata version 14.1 (StataCorp, College Station, TX, USA).

Laboratory analyses

The serum samples (aliquots of 400 µL) will be analysed for total 25(OH) D, DBP, leptin, PTH, albumin, and calcium. The Hormone laboratory at Aker hospital, Oslo, Norway will analyze 25(OH)D, DBP, PTH and leptin. The Hormone Laboratory is accredited by the Norwegian Accreditation as a testing laboratory and complies with the requirements of the NS-EN ISO/IEC 17025 standards. Albumin and calcium will be analyzed at Department of Medical Biochemistry, Oslo University Hospital accredited by the Norwegian Accreditation reg. no TEST 103 that complies with the requirements of the NS-EN ISO 1518. The sample donors, identity and case control status will be blinded for the laboratory staff. Quality control (QC) samples from the biobank will be included in every batch to examine inter-batch and intra-batch variability.

Power and sample size

Sub-study I Within the prospective cohort (n = 292,851), there are more than 2,300 BC cases reported to the cancer registry until the end of follow up time. Assuming the risk of developing bladder cancer is 1% for the normal weight group and 5% for the obese group [14], one would need a sample size of 586 to have 80% power. Since the sample size here is significantly larger, one can safely determine that the study has adequate statistical power.

Sub-study II In the case-control study nested into the prospective cohort, the statistical power will depend on: i) proportion of exposure in the population, ii) sample size (cases and controls) and iii) the minimum difference that is possible to detect.

Table 1 shows the smallest detectable OR according to proportion of controls exposed to low vitamin D (25(OH)D and high leptin levels, for different sample sizes. The power is 0.80 and a significance level of 0.05 (www.krothman.hostbyet2. com/Episheet.xls).

Table 1. Odds ratio based on proportion of exposed controls and sample size

		Study case:control= 1:1
Proportion		Number of cases

of exposed controls		n=500	n=400	n=300
55%*	OR	1.44	1.50	1.6
45%*	OR	1.43	1.49	1.58
30%**	OR	1.45	1.51	1.62

Exposure = 25(OH)D deficiency (25(OH)D < 50 nmol /L); **Exposure = high serum leptin levels (>4.1 ng/mL)

Based on results in the table above, we consider 400 matched case-control pairs as a sufficient sample size for the case-control study.

Data analysis plan

The following analyses will be conducted to test our hypotheses:

• Hypothesis 1.a: A prospective cohort analysis of pre-diagnostic BMI and other anthropometric measures in relation to BC risk using the complete Janus Cohort (n = 292,851)

• Hypothesis 1.b: A prospective cohort analysis of pre-diagnostic BMI and other anthropometric measures in relation to survival after a BC, using all BC cases in the Janus Cohort (n ≈ 2,650)

• Hypothesis 2.a: A nested case-control analysis of BC risk according to pre-diagnostic serum levels of total and free 25(OH)D, and leptin in 400 matched case-control pairs.

• Hypothesis 2.b: A prospective analysis of survival after a BC (n=400) according to pre-diagnostic serum levels of 25(OH)D and leptin.

Study strengths and limitations

A major strength of the large sample set of more than 2,300 incident BC cases. Also use of individual PIN for linkages between multiple data sources, to establish a virtually complete study file, with exception of data on histopathology, is a strength. The data sources are high

quality population-based registries, with high degree of completeness. The bladder cancer diagnoses are coded according to ICD-0. To get information on staging, and control the data quality, all histopathological information will be reviewed and characterized by tumor-stage (T-stage). Another strength of this study is that the public health data has been quality assured, structured and harmonized [22]. The use of pre-diagnostic samples assure the proper temporality of the relationship between exposure and BC, limiting the possibility of reverse causality.

It is a limitation that carcinoma in situ (Tis) regarded as a precursor lesion and the premalignant stage of papillary translational cancer cases (Ta) and are not included in the sub-study II. A high number (40-80 %) of patients with Tis stage will develop high grade muscle invasive cancer if untreated, especially if associated with papillary tumors [34] .

ETHICS AND DISSEMINATION

The Regional Committee for Medical and Health Research Ethics has approved the study. The different data registries have approved that the use of a de-identified dataset. An ID-key, consisting of the 11-digit PIN and a study-specific ID number, will be stored and governed by a third party unavailable to the research team.

All results will be published in relevant peer- reviewed international scientific journals and presented at conferences, nationally and internationally. Results of importance will be directly communicated to health authorities and to clinicians where the annual national oncology conference “Onkologisk forum” can serve as a platform for knowledge distribution. Results of importance will also be disseminated through press releases and to user groups like the Norwegian Cancer Society. The CRN website is a potential channel to reach patients organizations and the public.

Authors' contributions

REG prepared the study. TER, JSS, HL, BA, KA, EW and AM contributed to the study design and reviewed and revised the protocol critically for important intellectual content, and approved the final versions. REG is the guarantor.

Data sharing

Requests for data sharing/case pooling may be directed to the corresponding author. This study uses third-party data derived from State government registries, which are ultimately governed by their ethics committees and data custodians. Thus, any requests to share these data will be subject to formal approval from each data source used in this study.

Funding

The research study has been reviewed and granted funding by the Norwegian Cancer Society (no. 182308-2016) and the Cancer Registry of Norway Research Fund

Conflicts of interest

None declared

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* 2016;66(1):7-30. doi: 10.3322/caac.21332 [published Online First: 2016/01/09]
2. Svatek RS, Hollenbeck BK, Holmang S, et al. The Economics of Bladder Cancer: Costs and Considerations of Caring for This Disease. *European urology* 2014 doi: 10.1016/j.eururo.2014.01.006 [published Online First: 2014/01/30]
3. Freedman ND, Silverman DT, Hollenbeck AR, et al. Association between smoking and risk of bladder cancer among men and women. *Jama* 2011;306(7):737-45. doi: 10.1001/jama.2011.1142 [published Online First: 2011/08/19]
4. Olfert SM, Felknor SA, Delclos GL. An updated review of the literature: risk factors for bladder cancer with focus on occupational exposures. *Southern medical journal* 2006;99(11):1256-63. doi: 10.1097/01.smj.0000247266.10393.72 [published Online First: 2007/01/02]
5. Ishii N, Wei M, Kakehashi A, et al. Enhanced Urinary Bladder, Liver and Colon Carcinogenesis in Zucker Diabetic Fatty Rats in a Multiorgan Carcinogenesis Bioassay: Evidence for Mechanisms Involving Activation of PI3K Signaling and Impairment of p53 on Urinary Bladder Carcinogenesis. *Journal of toxicologic pathology* 2011;24(1):25-36. doi: 10.1293/tox.24.25 [published Online First: 2012/01/25]
6. Yuan SS, Chung YF, Chen HW, et al. Aberrant expression and possible involvement of the leptin receptor in bladder cancer. *Urology* 2004;63(2):408-13. doi: 10.1016/j.urology.2003.08.038 [published Online First: 2004/02/20]
7. Bikle D. Nonclassic actions of vitamin D. *The Journal of clinical endocrinology and metabolism* 2009;94(1):26-34. doi: 10.1210/jc.2008-1454 [published Online First: 2008/10/16]
8. Giovannucci E. The epidemiology of vitamin D and cancer incidence and mortality: a review (United States). *Cancer causes & control : CCC* 2005;16(2):83-95. doi: 10.1007/s10552-004-1661-4 [published Online First: 2005/05/04]
9. Holick MF. Vitamin D status: measurement, interpretation, and clinical application. *Annals of epidemiology* 2009;19(2):73-8. doi: 10.1016/j.annepidem.2007.12.001 [published Online First: 2008/03/11]
10. Pike JW FD, Glorieux FH. Vitamin D. San Diego: Academic Press, 1997.

11. Mondul AM, Weinstein SJ, Virtamo J, et al. Influence of vitamin D binding protein on the association between circulating vitamin D and risk of bladder cancer. *British journal of cancer* 2012;107(9):1589-94. doi: 10.1038/bjc.2012.417 [published Online First: 2012/09/20]
12. Mondul AM, Weinstein SJ, Horst RL, et al. Serum vitamin D and risk of bladder cancer in the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening trial. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2012;21(7):1222-5. doi: 10.1158/1055-9965.epi-12-0439 [published Online First: 2012/05/25]
13. Ng M, Fleming T, Robinson M, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet (London, England)* 2014;384(9945):766-81. doi: 10.1016/s0140-6736(14)60460-8 [published Online First: 2014/06/02]
14. Sun JW, Zhao LG, Yang Y, et al. Obesity and risk of bladder cancer: a dose-response meta-analysis of 15 cohort studies. *PloS one* 2015;10(3):e0119313. doi: 10.1371/journal.pone.0119313 [published Online First: 2015/03/25]
15. Qin Q, Xu X, Wang X, et al. Obesity and risk of bladder cancer: a meta-analysis of cohort studies. *Asian Pacific journal of cancer prevention : APJCP* 2013;14(5):3117-21. [published Online First: 2013/06/28]
16. Klok MD, Jakobsdottir S, Drent ML. The role of leptin and ghrelin in the regulation of food intake and body weight in humans: a review. *Obesity reviews : an official journal of the International Association for the Study of Obesity* 2007;8(1):21-34. doi: 10.1111/j.1467-789X.2006.00270.x [published Online First: 2007/01/11]
17. Garofalo C, Surmacz E. Leptin and cancer. *Journal of cellular physiology* 2006;207(1):12-22. doi: 10.1002/jcp.20472 [published Online First: 2005/08/20]
18. Foss YJ. Vitamin D deficiency is the cause of common obesity. *Medical hypotheses* 2009;72(3):314-21. doi: 10.1016/j.mehy.2008.10.005 [published Online First: 2008/12/05]
19. Holick MF. Vitamin D deficiency. *The New England journal of medicine* 2007;357(3):266-81. doi: 10.1056/NEJMra070553 [published Online First: 2007/07/20]
20. Mai XM, Chen Y, Camargo CA, Jr., et al. Cross-sectional and prospective cohort study of serum 25-hydroxyvitamin D level and obesity in adults: the HUNT study. *American journal of epidemiology* 2012;175(10):1029-36. doi: 10.1093/aje/kwr456 [published Online First: 2012/02/09]
21. Langseth H, Gislefoss RE, Martinsen JI, et al. Cohort Profile: The Janus Serum Bank Cohort in Norway. *International journal of epidemiology* 2016 doi: 10.1093/ije/dyw027 [published Online First: 2016/04/12]
22. Hjerkind KV, Gislefoss RE, Tretli S, et al. Cohort Profile Update: The Janus Serum Bank Cohort in Norway. *International journal of epidemiology* 2017 doi: 10.1093/ije/dyw302 [published Online First: 2017/01/15]
23. Gislefoss RE, Grimsrud TK, Morkrid L. Long-term stability of serum components in the Janus Serum Bank. *Scandinavian journal of clinical and laboratory investigation* 2008;68(5):402-9. doi: 10.1080/00365510701809235 [published Online First: 2008/08/30]
24. Gislefoss RE, Grimsrud TK, Morkrid L. Stability of selected serum proteins after long-term storage in the Janus Serum Bank. *ClinChemLab Med* 2009;47(5):596-603.
25. Gislefoss RE, Grimsrud TK, Morkrid L. Stability of selected serum hormones and lipids after long-term storage in the Janus Serum Bank. *Clinical biochemistry* 2015;48(6):364-9. doi: 10.1016/j.clinbiochem.2014.12.006 [published Online First: 2014/12/20]
26. Stattin P, Kaaks R, Johansson R, et al. Plasma leptin is not associated with prostate cancer risk. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2003;12(5):474-5. [published Online First: 2003/05/17]

27. Stattin P, Lukanova A, Biessy C, et al. Obesity and colon cancer: does leptin provide a link? *International journal of cancer* 2004;109(1):149-52. doi: 10.1002/ijc.11668 [published Online First: 2004/01/22]
28. Tretli S, Schwartz GG, Torjesen PA, et al. Serum levels of 25-hydroxyvitamin D and survival in Norwegian patients with cancer of breast, colon, lung, and lymphoma: a population-based study. *Cancer causes & control : CCC* 2012;23(2):363-70. doi: 10.1007/s10552-011-9885-6 [published Online First: 2011/12/24]
29. Meyer HE, Robsahm TE, Bjorge T, et al. Vitamin D, season, and risk of prostate cancer: a nested case-control study within Norwegian health studies. *The American journal of clinical nutrition* 2013;97(1):147-54. doi: 10.3945/ajcn.112.039222 [published Online First: 2012/11/30]
30. Larsen IK, Smastuen M, Johannesen TB, et al. Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness. *European journal of cancer (Oxford, England : 1990)* 2009;45(7):1218-31. doi: 10.1016/j.ejca.2008.10.037 [published Online First: 2008/12/19]
31. Amin MB, Edge, S., Greene, F., Byrd, D.R., Brookland, R.K., Washington, M.K., Gershenwald, J.E., Compton, C.C., Hess, K.R., Sullivan, D.C., Jessup, J.M., Brierley, J.D., Gaspar, L.E., Schilsky, R.L., Balch, C.M., Winchester, D.P., Asare, E.A., Madera, M., Gress, D.M., Meyer, L.R. . AJCC Cancer Staging Manual 8 th edition. New York: Springer International Publishing 2017.
32. Bjartveit K, Foss OP, Gjervig T, et al. The cardiovascular disease study in Norwegian counties. Background and organization. *Acta medica Scandinavica Supplementum* 1979;634:1-70. [published Online First: 1979/01/01]
33. PG L-L. Arven fra Statens skjermbildefotografering (SSF)/Statens helseundersøkelser (SHUS) en introduksjon til databanken. *Tidsskr Norsk Forening for Epidemiol* 2003;13:14.
34. Althausen AF, Prout GR, Jr., Daly JJ. Non-invasive papillary carcinoma of the bladder associated with carcinoma in situ. *The Journal of urology* 1976;116(5):575-80. [published Online First: 1976/11/01]

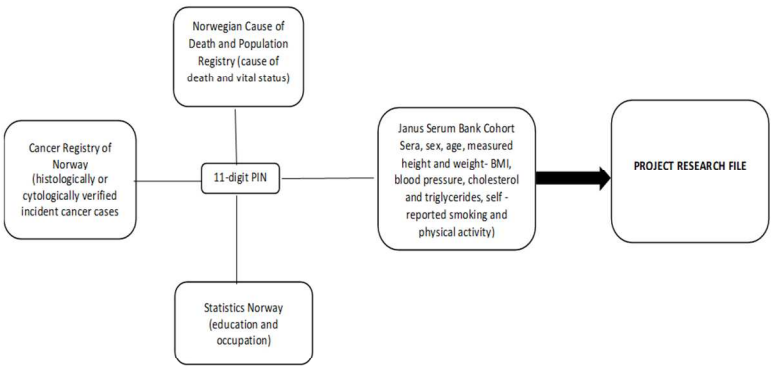

Figure 1. Data collection from different sources using PIN

338x190mm (96 x 96 DPI)

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract P 1
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found P 2
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported P 3-4
Objectives	3	State specific objectives, including any prespecified hypotheses P 5
Methods		
Study design	4	Present key elements of study design early in the paper P 7
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection P 5-7
Participants	6	(a) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up P 5-7
		(b) For matched studies, give matching criteria and number of exposed and unexposed P 7-8
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
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 P 5-7, 9
Bias	9	Describe any efforts to address potential sources of bias P 8
Study size	10	Explain how the study size was arrived at P 9-10
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding P 8
		(b) Describe any methods used to examine subgroups and interactions P 8
		(c) Explain how missing data were addressed
		(d) If applicable, explain how loss to follow-up was addressed
		(e) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed P 8
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram P 7
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders P 7-8
		(b) Indicate number of participants with missing data for each variable of interest
		(c) Summarise follow-up time (eg, average and total amount)
Outcome data	15*	Report numbers of outcome events or summary measures over time Not applicable
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

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion Not applicable		
Key results	18	Summarise key results with reference to study objectives
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
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
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 P12

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

BMJ Open

A protocol for prospective study of vitamin D, obesity, and leptin in relation to bladder cancer, incidence and survival.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-019309.R1
Article Type:	Protocol
Date Submitted by the Author:	19-Dec-2017
Complete List of Authors:	Gislefoss, Randi Elin; Cancer Registry og Norway, Dept of Research Stenehjem, Jo; Cancer Registry of Norway, Dept of Research Hektoen, Helga; Cancer Registry of Norway, Dept. of Research Andreassen, Bettina Kulle ; Cancer Registry of Norway, Dept of Research Langseth, Hilde; Cancer Registry of Norway, Dept of Research Axcrona, Karol; Akershus Universitetssykehus HF Weiderpass, Elisabete; Cancer Registry of Norway, Dept of Research Mondul, Alison; University of Michigan, Scool of Public Health Robsahm, Trude; Cancer Registry of Norway, Research
Primary Subject Heading:	Oncology
Secondary Subject Heading:	Urology
Keywords:	Bladder cancer, Biobank, vitamin D, obesity

A protocol for prospective study of vitamin D, obesity, and leptin in relation to bladder cancer, incidence and survival

Authors: Randi E Gislefoss[1], Jo S Stenehjem[1], Helga H Hektoen[1], Bettina Andreassen[1], Hilde Langseth[1], Karol Axcrona[2], Elisabete Weiderpass[1], Alison Mondul[3], Trude E Robsahm[1]

[1]Department of Research, Cancer Registry of Norway, Oslo, Norway

[2]Department of Urology, Akershus University Hospital, Lørenskog, Norway

[3]University of Michigan, School of Public Health, Ann Arbor, MI, USA

Correspondence to:

Dr. Randi Elin Gislefoss: randi.gislefoss@kreftregisteret.no

Word count: 2794

Numbers of tables and figures: 2

Abstract

Introduction: Bladder cancer (BC) (including renal pelvis, ureter and urethra) is one of the most common urogenital cancers and the fourth most frequent cancer in men in the United States. In Norway, the incidence of BC has increased over the last decades. The age-standardized incidence rates per 100,000 for 2011-15 were 50.4 in men and 15.5 in women. Compared to the 5-year period 2006-2010, the percentage increase in incidence was 9.4% in men and 8.9% in women. The recurrence rate of BC is over 50%, the highest recurrence rate of any malignancy. Smoking and occupational exposure to aromatic amines are recognized as the major risk factors. Recently, low serum level of 25 hydroxy vitamin D (25(OH)D) and obesity have been suggested to increase the BC risk, and leptin, which is important in weight regulation, may be involved in bladder carcinogenesis. More knowledge on potential risk factors for BC is necessary for planning and implementing primary prevention measures.

Methods and analyses: Cohort and nested case-control studies will be carried out using the population-based Janus Serum Bank Cohort consisting of pre-diagnostic sera, clinical measurement data (body height and weight, blood pressure, cholesterol and triglycerides) and self-reported information on lifestyle factors (smoking, physical activity). Participants were followed from cohort inclusion (1972-2003) through 2014. The cohort will be linked to the Cancer Registry of Norway (cancer data), the national Cause of Death Registry (date and cause of death), National Population Registry (vital status) and Statistic Norway (education and occupation). Serum samples will be analyzed for 25(OH)D, vitamin D binding protein, leptin, albumin, calcium and parathyroid hormone.
Cox regression and conditional logistic regression models will be used to estimate association between the exposures and BC.

Ethics and dissemination: The study has been approved by the Regional Committee for Medical Research Ethics and is funded by the Norwegian Cancer Society. Results will be published in peer-reviewed journals, at scientific conferences and through press releases.

Strengths and limitations of this study

- Use of a large sample set of more than 2300 incident BC cases
- Pre-diagnostic serum samples assure the temporality of the relationship between exposure and BC, limiting the possibility of reverse causality.
- Use of unique personal identification number for linkages between multiple data sources, to establish a virtually complete study file and complete ascertainment of follow-up.
- Reviewing and characterizing T-stage for all BC.
- Lack of treatment data

Introduction

Rationale and evidence gaps

Urinary bladder cancer (BC) (including renal pelvis, ureter and urethra) is the most common urogenital cancer after prostate cancer and is the fourth most frequent cancer in men in the United States[1]. BC has over a 50 % recurrence rate, the highest of any malignancy, and is one of the most expensive cancers to treat on a per-patient basis [2]. In Norway, the incidence of cancer of the bladder has been increasing over the last decades. In 2015, the age-standardized incidence rates were 53.7 and 16.5 in Norwegian men and women, respectively. Compared to the 5-year period 2006-2010, the increase in incidence has been 6.1 % in men and 12.3 % in women. Up to 50 % of all BC cases have been ascribed to smoking [3], and 5–25 % of the cases have been attributable to occupational exposures [4]; still, the etiology of up to 45 % of BC remains unexplained. Low serum level of 25 hydroxy vitamin D 25(OH)D and obesity have been suggested to increase BC risk, and the hormone leptin, which is important in weight regulation, may be involved in its carcinogenetic process [5 6].

25(OH)D is converted to its active hormonal form, 1-25-dihydroxyvitamin D ($1,25(OH)_2D$), by 1-α-hydroxylase, which is present in most tissues in the body [7]. Parathyroid hormone (PTH) and calcium level are important factors as they affect the enzymatic conversion from 25(OH)D to active 1, 25-$(OH)D_3$ in the kidney and may be involved in non-classical synthesis. Measurement of circulating 25(OH)D is considered the gold standard measurement of

vitamin D status as it integrates vitamin D exposure from oral intake from diet or supplements, as well as from exposure to ultraviolet radiation (UVR) [8 9]. Despite being the gold standard, total circulating 25(OH)D may not be the best measure of 25(OH)D exposure for all tumors. The "free hormone hypothesis" suggests that only unbound, free hormones can have biologic effects on target tissues [10]. To date, few studies have examined the role of vitamin D binding protein (DBP) also known as group-specific component or Gc-globulin, in the association between 25(OH)D and various cancer processes. DBP transports both 25(OH)D and $1,25(OH)_2D$ in circulation. This protein carries 88% of 25(OH)D and 85% of $1,25(OH)_2D$; an additional 12% of 25(OH)D and 15% of $1,25(OH)_2D$ circulate bound to albumin. Clinical laboratory assays of circulating 25(OH)D that are currently in use measure total 25(OH)D without differentiating between the bound and free forms. Thus, it remains unclear whether total or free 25(OH)D is more biologically relevant with respect to risk of BC. Two previous studies have examined free, in addition to total 25(OH)D, in relation to risk of BC. One found an inverse association between total 25(OH)D and bladder cancer that appeared to be restricted to participants with low DBP, suggesting that free 25(OH)D might be more strongly associated with risk of bladder cancer than total 25(OH)D [11]. The other study found no association between 25(OH)D overall or at any level of DBP concentration [12].

Body mass index (weight/height^2, BMI) is a reliable indicator of body fatness and can be categorized as underweight (<18.5), healthy weight (18.5-24.9), overweight (25-29.9) and obese (> 30). The prevalence of obesity in Norway has risen steeply the last decades. In the 1970s, about 15% of men and women were overweight (www.fhi.no/) while in 2013 the proportions were 58.4% and 47.3%, respectively [13]. Two meta-analyses, including 15 [14] and 11 [15] cohort studies conclude that obesity significantly increases the risk of BC. Leptin, a hormone involved in weight regulation [16], may be involved in this potential association. High leptin levels have been shown to impact development of several forms of cancer [17]. A study of Yuan et al. [6] shows that leptin receptors are aberrantly expressed in BC tissue and a high leptin level has been associated with BC carcinogenesis [5 6].

Low 25(OH)D levels are more frequent in obese persons, suggesting that 25(OH)D deficiency is associated with obesity and vice versa. 25(OH)D deficiency is suggested to be associated

with obesity [18-20], and both low 25(OH)D and obesity are suggested to contribute to development of BC.

Aims and hypotheses

Better-targeted BC primary and tertiary prevention (risk and survival) is warranted. The interplay between 25(OH)D and obesity and their associations with BC risk are poorly understood. We propose a study aiming to examine anthropometric data (BMI, height, weight, body surface area and weight change over time) and serum levels of leptin, total and free 25(OH)D, in relation to BC risk and survival by using samples from Janus Serum Bank and associated data from population-based registries and surveys.

We hypothesize that

1 a. Obesity is associated with increased BC risk;

1b. Obesity is associated with reduced BC survival;

2a. Low free and total 25(OH)D level and high serum leptin levels (>4.1 ng/mL are associated with increased risk of BC;

2b. Low free and total 25(OH)D level and high serum leptin levels (>4.1 ng/mL are associated with reduced BC survival.

METHODS AND ANALYSIS

Study population and data sources

The Janus Serum Bank Cohort

The study will be carried out using the Janus Serum Bank Cohort, a population-based biobank for prospective cancer studies, containing serum samples and questionnaire data from 292,851 Norwegians who participated in one or more of five Norwegian Regional Health Studies in the period 1972–2003. Detailed description of the samples and data included in the Janus Serum Bank Cohort has been published elsewhere [21 22]. The quality aspects of long-term stored samples have been of high priority in the Janus Serum Bank and

component stability for a large number of hormones, proteins, metabolites and electrolytes has been investigated [23-25], including for both 25(OH)D, and leptin [26-29]. A unique 11-digit personal identification number (PIN), assigned to al Norwegian residents, will be used to link the Janus Cohort with population-based registries and surveys.

Population-based registries and surveys

The Cancer Registry of Norway (CRN) has collected notifications on cancer at a national level since 1953. Cancer reporting is mandatory by law, and reports from various sources ensures high quality and completeness (98.8%) [30]. The reporting system, based on pathology and cytology reports, clinical records, and death certificates, provides information about site, histological type and stage of disease at the time of diagnosis. CRN has been involved in the Janus Serum Bank operation since establishment in the early 1970s and has been responsible for the data handling; in 2004 the serum bank was integrated into the CRN. The following information is available for cancer cases: month and year of diagnosis, tumor site (International Classification of Diseases 7th revision [ICD-7 codes] converted into ICD-10 codes), histology (codes from ICD-Oncology 2nd and 3rd revision), clinical stage (local = no metastases, regional = metastasis in regional lymph nodes or surrounding area, distant = distant metastasis). In addition, all BC diagnoses in the Janus cohort will be reviewed and assigned a pathological T-stage, according to the American Joint Committee on Cancer (AJCC) staging manual 8th ed.[31].

The Norwegian Institute of Public Health (NIPH) has been responsible for conducting the national health surveys, upon which the Janus Serum Bank is partly based. All participants have completed questionnaires for assessment of lifestyle factors (i.e. smoking habits, alcohol use), at the time of serum collection. A database has been established, including data from these questionnaires, as well as measured body height and weight, blood pressure, cholesterol and triglycerides [32 33]. The Janus Cohort includes participants from five of the health studies: The Oslo Study I (1972–73), The Norwegian Counties Study (1974–78, 1977–83, and 1985–88), The Age 40 Program- Oslo (1981–99), The National Age 40 Program (1985–99) and The TROFINN Health Study (2001–03). A set of about 50 variables has been harmonized and standardized due to slightly different wording in the questionnaires [22].

Available variables include: height (cm), weight (kg), BMI (kg/m^2, categorized as 12–18.49, 18.5–24.9, 25.0–29.9, ≥30), smoking status (never, former, current), cigarettes per day (1–9, 10–14, ≥15), years of smoking (1–9, 10–29, ≥30), time since smoking cessation (<3months, 3months–1yr, 1–5yrs, >5yrs), and total physical activity (inactive, low, medium, high), based on leisure time activity.

The Cause of death Registry has registered death certificates for all deaths in Norway since 1951. Cause of death registration is mandatory by law.

National Population Registry contains information on vital status (alive, emigrated or dead) of everyone that resides or has resided in Norway.

Statistics Norway has the responsibility of covering the needs for official statistics on the Norwegian population including individual data on settlements, migration, occupation and level of education.

Using the 11-digit PIN number we will link data from four different sources to set up the research file, illustrated in Figure 1.

Study design

Sub-study I: a prospective cohort study

In a prospective cohort study among all individuals in the Janus Serum Bank Cohort (n = 292,851) (Figure 1), we will explore baseline BMI in relation to bladder cancer risk. Among the included BC cases we will investigate baseline BMI and BC survival. By 2014, the cohort included 2 347 BC cases (ICD-10: C67). BC cases of both muscle invasive and non-invasive urothelial cell carcinoma, will be included in the study. Educational level, occupation, age, sex, physical activity, smoking habits, cholesterol, triglycerides and blood pressure will be included in the statistical analyses as confounders.

Sub-study II: a nested case-control study

The study will be nested within the prospective cohort described above (Study I), including a) 400 BC cases of high grade tumors, including muscle invasive (T2-T4) and non muscle invasive (Ta, T1 and carcinoma in situ (Tis) cancer cases, and b) 400 controls alive and without a cancer diagnosis at the time of the BC diagnosis of the cases, matched 1:1 on sex, age (+/-1 year) and date of serum sampling (+/-1 month). Minimum time from blood draw to diagnosis will be 5 years. The serum samples will be analysed for total 25(OH)D, vitamin D binding protein and leptin. As PTH and albumin-adjusted calcium level affect the enzymatic conversion from 25(OH)D to active 1,25-$(OH)D_3$ and might be involved in the non-classical synthesis as well, measurement of these components will be taken into account.

In sub-study I, we will investigate the association between BMI and BC, and for sub –study II we will in addition include vitamin D levels. Overall we will focus on disentangle the relationship of vitamin D, BMI and BC. This will be done in two ways:

1. By implementing regression models including and interaction effect of vitamin D and BMI on BC.
2. By testing the hypothesis whether the effect of BMI on BC is mediated by vitamin D using mediation analysis[34].

Statistical methods

In the cohort study, Cox regression models will be used to estimate hazard ratios (HR) with 95% confidence intervals (CI) of BC and survival after BC, taking into account stage at diagnosis, and BMI and including adjusting for season of blood collection for vitamin D. In the nested case-control study, conditional logistic regression models will be used to estimate the odds ratio (OR) with 95% CI.

In order to find out whether Vitamin D acts (totally or partly) as a mediator, modern causal inference theory will be used to estimate different types of effects[34]. The analysis will be done by using the mediation R package [35]

As we have a number of potential confounding variables, we will use directed acyclic graphs to select variables to include in the statistical models. Confounding variables will be included in the models and tests of interaction effects will be performed when relevant.

All tests will be two-sided and p<0.05 will be considered statistically significant. All statistical analyses will be conducted using Stata version 14.1 (StataCorp, College Station, TX, USA).

Laboratory analyses

The serum samples (aliquots of 400 µL) will be analysed for total 25(OH) D, DBP, leptin, PTH, albumin, and calcium. The Hormone laboratory at Aker hospital, Oslo, Norway will analyze 25(OH)D, DBP, PTH and leptin. The Hormone Laboratory is accredited by the Norwegian Accreditation as a testing laboratory and complies with the requirements of the NS-EN ISO/IEC 17025 standards. Albumin and calcium will be analyzed at Department of Medical Biochemistry, Oslo University Hospital accredited by the Norwegian Accreditation reg. no TEST 103 that complies with the requirements of the NS-EN ISO 1518. The sample donors, identity and case control status will be blinded for the laboratory staff. Quality control samples from the biobank will be included in every batch to examine inter-batch and intra-batch variability.

Power and sample size

Sub-study I Within the prospective cohort (n = 292,851), there are more than 2,300 BC cases reported to the cancer registry until the end of follow up time. Assuming the risk of developing bladder cancer is 1% for the normal weight group and 5% for the obese group [14], one would need a sample size of 586 to have 80% power. Since the sample size here is significantly larger, one can safely determine that the study has adequate statistical power.

Sub-study II In the case-control study nested into the prospective cohort, the statistical power will depend on: i) proportion of exposure in the population, ii) sample size (cases and controls) and iii) the minimum difference that is possible to detect.

Table 1 shows the smallest detectable OR according to proportion of controls exposed to low vitamin D (25(OH)D and high leptin levels, for different sample sizes. The power is 0.80

and a significance level of 0.05 (www.krothman.hostbyet2. com/Episheet.xls). The expected proportions of exposed controls were based on previous studies on serum samples from the Janus Cohort. For 25(OH)D, a study on prostate cancer reported that 4.4% and 30.6% of the controls had 25(OH)D levels below 30 nmol/L and 50 nmol/L, respectively. [29] For leptin, a study on colon cancer reported that 20% of the controls had a leptin level of 4.1 ng/mL or higher.[27]

Table 1. Odds ratio (OR) based on proportion of exposed controls and sample size

		Study case:control= 1:1		
Proportion of exposed controls		Number of cases		
		n=500	n=400	n=300
55%*	OR	1.44	1.50	1.6
45%*	OR	1.43	1.49	1.58
30%**	OR	1.45	1.51	1.62

*Exposure = 25(OH)D deficiency (25(OH)D < 50 nmol /L);

**Exposure = high serum leptin levels (>4.1 ng/mL)

Based on results in the table above, we consider 400 matched case-control pairs as a sufficient sample size for the case-control study.

Data analysis plan

The following analyses will be conducted to test our hypotheses:

• Hypothesis 1.a: A prospective cohort analysis of pre-diagnostic BMI and other anthropometric measures in relation to BC risk using the complete Janus Cohort (n = 292,851)

• Hypothesis 1.b: A prospective cohort analysis of pre-diagnostic BMI and other anthropometric measures in relation to survival after a BC, using all BC cases in the Janus Cohort (n ≈ 2,650)

• Hypothesis 2.a: A nested case-control analysis of BC risk according to pre-diagnostic serum levels of total and free 25(OH)D, and leptin in 400 matched case-control pairs.

• Hypothesis 2.b: A prospective analysis of survival after a BC (n=400) according to pre-diagnostic serum levels of 25(OH)D and leptin.

Study strengths and limitations

A major strength of the large sample set of more than 2,300 incident BC cases. Also use of individual PIN for linkages between multiple data sources, to establish a virtually complete study file, with exception of data on histopathology, is a strength. The data sources are high quality population-based registries, with high degree of completeness. The bladder cancer diagnoses are coded according to ICD-0. To get information on staging, and control the data quality, all histopathological information will be reviewed and characterized by tumor-stage (T-stage). Another strength of this study is that the public health data has been quality assured, structured and harmonized [22]. The use of pre-diagnostic samples assure the proper temporality of the relationship between exposure and BC, limiting the possibility of reverse causality.

Treatment data is of importance when evaluating the survival analyses. These data are missing and will be a limitation of this study

ETHICS AND DISSEMINATION

The Regional Committee for Medical and Health Research Ethics has approved the study. The different data registries have approved that the use of a de-identified dataset. An ID-key, consisting of the 11-digit PIN and a study-specific ID number, will be stored and governed by a third party unavailable to the research team.

All results will be published in relevant peer- reviewed international scientific journals and presented at conferences, nationally and internationally. Results of importance will be

directly communicated to health authorities and to clinicians where the annual national oncology conference “Onkologisk forum” can serve as a platform for knowledge distribution. Results of importance will also be disseminated through press releases and to user groups like the Norwegian Cancer Society. The CRN website is a potential channel to reach patients organizations and the public.

Authors' contributions

REG prepared the study. TER, JSS, HHH, HL, BA, KA, EW and AM contributed to the study design and reviewed and revised the protocol critically for important intellectual content, and approved the final versions. REG is the guarantor.

Data sharing

Requests for data sharing/case pooling may be directed to the corresponding author. This study uses third-party data derived from State government registries, which are ultimately governed by their ethics committees and data custodians. Thus, any requests to share these data will be subject to formal approval from each data source used in this study.

Funding

The research study has been reviewed and granted funding by the Norwegian Cancer Society (no. 182308-2016) and the Cancer Registry of Norway Research Fund

Conflicts of interest

None declared

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* 2016;66(1):7-30. doi: 10.3322/caac.21332 [published Online First: 2016/01/09]
2. Svatek RS, Hollenbeck BK, Holmang S, et al. The Economics of Bladder Cancer: Costs and Considerations of Caring for This Disease. *European urology* 2014 doi: 10.1016/j.eururo.2014.01.006 [published Online First: 2014/01/30]
3. Freedman ND, Silverman DT, Hollenbeck AR, et al. Association between smoking and risk of bladder cancer among men and women. *Jama* 2011;306(7):737-45. doi: 10.1001/jama.2011.1142 [published Online First: 2011/08/19]
4. Olfert SM, Felknor SA, Delclos GL. An updated review of the literature: risk factors for bladder cancer with focus on occupational exposures. *Southern medical journal* 2006;99(11):1256-63. doi: 10.1097/01.smj.0000247266.10393.72 [published Online First: 2007/01/02]

5. Ishii N, Wei M, Kakehashi A, et al. Enhanced Urinary Bladder, Liver and Colon Carcinogenesis in Zucker Diabetic Fatty Rats in a Multiorgan Carcinogenesis Bioassay: Evidence for Mechanisms Involving Activation of PI3K Signaling and Impairment of p53 on Urinary Bladder Carcinogenesis. *Journal of toxicologic pathology* 2011;24(1):25-36. doi: 10.1293/tox.24.25 [published Online First: 2012/01/25]
6. Yuan SS, Chung YF, Chen HW, et al. Aberrant expression and possible involvement of the leptin receptor in bladder cancer. *Urology* 2004;63(2):408-13. doi: 10.1016/j.urology.2003.08.038 [published Online First: 2004/02/20]
7. Bikle D. Nonclassic actions of vitamin D. *The Journal of clinical endocrinology and metabolism* 2009;94(1):26-34. doi: 10.1210/jc.2008-1454 [published Online First: 2008/10/16]
8. Giovannucci E. The epidemiology of vitamin D and cancer incidence and mortality: a review (United States). *Cancer causes & control : CCC* 2005;16(2):83-95. doi: 10.1007/s10552-004-1661-4 [published Online First: 2005/05/04]
9. Holick MF. Vitamin D status: measurement, interpretation, and clinical application. *Annals of epidemiology* 2009;19(2):73-8. doi: 10.1016/j.annepidem.2007.12.001 [published Online First: 2008/03/11]
10. Pike JW FD, Glorieux FH. Vitamin D. San Diego: Academic Press, 1997.
11. Mondul AM, Weinstein SJ, Virtamo J, et al. Influence of vitamin D binding protein on the association between circulating vitamin D and risk of bladder cancer. *British journal of cancer* 2012;107(9):1589-94. doi: 10.1038/bjc.2012.417 [published Online First: 2012/09/20]
12. Mondul AM, Weinstein SJ, Horst RL, et al. Serum vitamin D and risk of bladder cancer in the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening trial. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2012;21(7):1222-5. doi: 10.1158/1055-9965.epi-12-0439 [published Online First: 2012/05/25]
13. Ng M, Fleming T, Robinson M, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet (London, England)* 2014;384(9945):766-81. doi: 10.1016/s0140-6736(14)60460-8 [published Online First: 2014/06/02]
14. Sun JW, Zhao LG, Yang Y, et al. Obesity and risk of bladder cancer: a dose-response meta-analysis of 15 cohort studies. *PloS one* 2015;10(3):e0119313. doi: 10.1371/journal.pone.0119313 [published Online First: 2015/03/25]
15. Qin Q, Xu X, Wang X, et al. Obesity and risk of bladder cancer: a meta-analysis of cohort studies. *Asian Pacific journal of cancer prevention : APJCP* 2013;14(5):3117-21. [published Online First: 2013/06/28]
16. Klok MD, Jakobsdottir S, Drent ML. The role of leptin and ghrelin in the regulation of food intake and body weight in humans: a review. *Obesity reviews : an official journal of the International Association for the Study of Obesity* 2007;8(1):21-34. doi: 10.1111/j.1467-789X.2006.00270.x [published Online First: 2007/01/11]
17. Garofalo C, Surmacz E. Leptin and cancer. *Journal of cellular physiology* 2006;207(1):12-22. doi: 10.1002/jcp.20472 [published Online First: 2005/08/20]
18. Foss YJ. Vitamin D deficiency is the cause of common obesity. *Medical hypotheses* 2009;72(3):314-21. doi: 10.1016/j.mehy.2008.10.005 [published Online First: 2008/12/05]
19. Holick MF. Vitamin D deficiency. *The New England journal of medicine* 2007;357(3):266-81. doi: 10.1056/NEJMra070553 [published Online First: 2007/07/20]
20. Mai XM, Chen Y, Camargo CA, Jr., et al. Cross-sectional and prospective cohort study of serum 25-hydroxyvitamin D level and obesity in adults: the HUNT study. *American journal of epidemiology* 2012;175(10):1029-36. doi: 10.1093/aje/kwr456 [published Online First: 2012/02/09]
21. Langseth H, Gislefoss RE, Martinsen JI, et al. Cohort Profile: The Janus Serum Bank Cohort in Norway. *International journal of epidemiology* 2016 doi: 10.1093/ije/dyw027 [published Online First: 2016/04/12]

22. Hjerkind KV, Gislefoss RE, Tretli S, et al. Cohort Profile Update: The Janus Serum Bank Cohort in Norway. *International journal of epidemiology* 2017 doi: 10.1093/ije/dyw302 [published Online First: 2017/01/15]
23. Gislefoss RE, Grimsrud TK, Morkrid L. Long-term stability of serum components in the Janus Serum Bank. *Scandinavian journal of clinical and laboratory investigation* 2008;68(5):402-9. doi: 10.1080/00365510701809235 [published Online First: 2008/08/30]
24. Gislefoss RE, Grimsrud TK, Morkrid L. Stability of selected serum proteins after long-term storage in the Janus Serum Bank. *ClinChemLab Med* 2009;47(5):596-603.
25. Gislefoss RE, Grimsrud TK, Morkrid L. Stability of selected serum hormones and lipids after long-term storage in the Janus Serum Bank. *Clinical biochemistry* 2015;48(6):364-9. doi: 10.1016/j.clinbiochem.2014.12.006 [published Online First: 2014/12/20]
26. Stattin P, Kaaks R, Johansson R, et al. Plasma leptin is not associated with prostate cancer risk. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2003;12(5):474-5. [published Online First: 2003/05/17]
27. Stattin P, Lukanova A, Biessy C, et al. Obesity and colon cancer: does leptin provide a link? *International journal of cancer* 2004;109(1):149-52. doi: 10.1002/ijc.11668 [published Online First: 2004/01/22]
28. Tretli S, Schwartz GG, Torjesen PA, et al. Serum levels of 25-hydroxyvitamin D and survival in Norwegian patients with cancer of breast, colon, lung, and lymphoma: a population-based study. *Cancer causes & control : CCC* 2012;23(2):363-70. doi: 10.1007/s10552-011-9885-6 [published Online First: 2011/12/24]
29. Meyer HE, Robsahm TE, Bjorge T, et al. Vitamin D, season, and risk of prostate cancer: a nested case-control study within Norwegian health studies. *The American journal of clinical nutrition* 2013;97(1):147-54. doi: 10.3945/ajcn.112.039222 [published Online First: 2012/11/30]
30. Larsen IK, Smastuen M, Johannesen TB, et al. Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness. *European journal of cancer (Oxford, England : 1990)* 2009;45(7):1218-31. doi: 10.1016/j.ejca.2008.10.037 [published Online First: 2008/12/19]
31. Amin MB, Edge, S., Greene, F., Byrd, D.R., Brookland, R.K., Washington, M.K., Gershenwald, J.E., Compton, C.C., Hess, K.R., Sullivan, D.C., Jessup, J.M., Brierley, J.D., Gaspar, L.E., Schilsky, R.L., Balch, C.M., Winchester, D.P., Asare, E.A., Madera, M., Gress, D.M., Meyer, L.R. . AJCC Cancer Staging Manual 8 th edition. New York: Springer International Publishing 2017.
32. Bjartveit K, Foss OP, Gjervig T, et al. The cardiovascular disease study in Norwegian counties. Background and organization. *Acta medica Scandinavica Supplementum* 1979;634:1-70. [published Online First: 1979/01/01]
33. PG L-L. Arven fra Statens skjermbildefotografering (SSF)/Statens helseundersøkelser (SHUS) en introduksjon til databanken. *Tidsskr Norsk Forening for Epidemiol* 2003;13:14.
34. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychol Methods* 2010;15(4):309-34. doi: 10.1037/a0020761 [published Online First: 2010/10/20]
35. Tingley D YT, Hirsoe K, Imai K. Mediation: R package for causal mediation anakysis. *Journal of Statistical Software* 2014;59(5):1-34.

Figure legends

Figure 1. Data collection from different sources using PIN

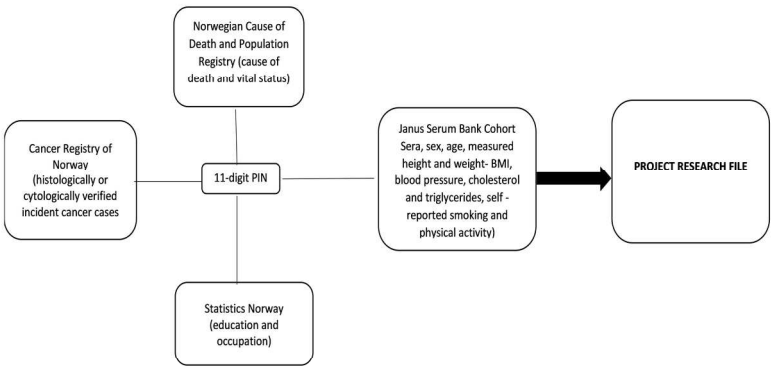

Figure 1. Data collection from different sources using PIN

190x107mm (300 x 300 DPI)

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract P 1
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found P 2
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported P 3-4
Objectives	3	State specific objectives, including any prespecified hypotheses P 5
Methods		
Study design	4	Present key elements of study design early in the paper P 7
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection P 5-7
Participants	6	(a) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up P 5-7
		(b) For matched studies, give matching criteria and number of exposed and unexposed P 7-8
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
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 P 5-7, 9
Bias	9	Describe any efforts to address potential sources of bias P 8
Study size	10	Explain how the study size was arrived at P 9-10
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding P 8
		(b) Describe any methods used to examine subgroups and interactions P 8
		(c) Explain how missing data were addressed
		(d) If applicable, explain how loss to follow-up was addressed
		(e) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed P 8
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram P 7
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders P 7-8
		(b) Indicate number of participants with missing data for each variable of interest
		(c) Summarise follow-up time (eg, average and total amount)
Outcome data	15*	Report numbers of outcome events or summary measures over time Not applicable
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

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion Not applicable		
Key results	18	Summarise key results with reference to study objectives
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
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
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 P12

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

BMJ Open

A protocol for prospective study of vitamin D, obesity, and leptin in relation to bladder cancer, incidence and survival.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-019309.R2
Article Type:	Protocol
Date Submitted by the Author:	06-Feb-2018
Complete List of Authors:	Gislefoss, Randi Elin; Cancer Registry og Norway, Dept of Research Stenehjem, Jo; Cancer Registry of Norway, Dept of Research Hektoen, Helga; Cancer Registry of Norway, Dept. of Research Andreassen, Bettina Kulle ; Cancer Registry of Norway, Dept of Research Langseth, Hilde; Cancer Registry of Norway, Dept of Research Axcrona, Karol; Akershus Universitetssykehus HF Weiderpass, Elisabete; Cancer Registry of Norway, Dept of Research Mondul, Alison; University of Michigan, Scool of Public Health Robsahm, Trude; Cancer Registry of Norway, Research
Primary Subject Heading:	Oncology
Secondary Subject Heading:	Urology
Keywords:	Bladder cancer, Biobank, vitamin D, obesity

A protocol for prospective study of vitamin D, obesity, and leptin in relation to bladder cancer, incidence and survival

Authors: Randi E Gislefoss[1], Jo S Stenehjem[1], Helga H Hektoen[1], Bettina Andreassen[1], Hilde Langseth[1], Karol Axcrona[2], Elisabete Weiderpass[1], Alison Mondul[3], Trude E Robsahm[1]

[1]Department of Research, Cancer Registry of Norway, Oslo, Norway

[2]Department of Urology, Akershus University Hospital, Lørenskog, Norway

[3]University of Michigan, School of Public Health, Ann Arbor, MI, USA

Correspondence to:

Dr. Randi Elin Gislefoss: randi.gislefoss@kreftregisteret.no

Word count: 2852

Numbers of tables and figures: 2

Abstract

Introduction: Bladder cancer (BC) (including renal pelvis, ureter and urethra) is one of the most common urogenital cancers and the fourth most frequent cancer in men in the United States. In Norway, the incidence of BC has increased over the last decades. The age-standardized incidence rates per 100,000 for 2011-15 were 50.4 in men and 15.5 in women. Compared to the 5-year period 2006-2010, the percentage increase in incidence was 9.4% in men and 8.9% in women. The recurrence rate of BC is over 50%, the highest recurrence rate of any malignancy. Smoking and occupational exposure to aromatic amines are recognized as the major risk factors. Recently, low serum level of 25 hydroxy vitamin D (25(OH)D) and obesity have been suggested to increase the BC risk, and leptin, which is important in weight regulation, may be involved in bladder carcinogenesis. More knowledge on potential risk factors for BC is necessary for planning and implementing primary prevention measures.

Methods and analyses: Cohort and nested case-control studies will be carried out using the population-based Janus Serum Bank Cohort consisting of pre-diagnostic sera, clinical measurement data (body height and weight, body surface area (BSA) and weight change over time, blood pressure, cholesterol and triglycerides) and self-reported information on lifestyle factors (smoking, physical activity). Participants were followed from cohort inclusion (1972-2003) through 2014. The cohort will be linked to the Cancer Registry of Norway (cancer data), the national Cause of Death Registry (date and cause of death), National Population Registry (vital status) and Statistic Norway (education and occupation). Serum samples will be analyzed for 25(OH)D, vitamin D binding protein, leptin, albumin, calcium and parathyroid hormone.
Cox regression and conditional logistic regression models, and mediation analysis will be used to estimate association between the exposures and BC.

Ethics and dissemination: The study has been approved by the Regional Committee for Medical Research Ethics and is funded by the Norwegian Cancer Society. Results will be published in peer-reviewed journals, at scientific conferences and through press releases.

Strengths and limitations of this study

- Use of a large sample set of more than 2300 incident BC cases
- Pre-diagnostic serum samples assure the temporality of the relationship between exposure and BC, limiting the possibility of reverse causality.
- Use of unique personal identification number for linkages between multiple data sources, to establish a virtually complete study file and complete ascertainment of follow-up.
- Reviewing and characterizing T-stage for all BC.
- Lack of treatment data

Introduction

Rationale and evidence gaps

Urinary bladder cancer (BC) (including renal pelvis, ureter and urethra) is the most common urogenital cancer after prostate cancer and is the fourth most frequent cancer in men in the United States[1]. BC has over a 50 % recurrence rate, the highest of any malignancy, and is one of the most expensive cancers to treat on a per-patient basis [2]. In Norway, the incidence of cancer of the bladder has been increasing over the last decades. In 2015, the age-standardized incidence rates were 53.7 and 16.5 in Norwegian men and women, respectively. Compared to the 5-year period 2006-2010, the increase in incidence has been 6.1 % in men and 12.3 % in women. Up to 50 % of all BC cases have been ascribed to smoking [3], and 5–25 % of the cases have been attributable to occupational exposures [4]; still, the etiology of up to 45 % of BC remains unexplained. Low serum level of 25 hydroxy vitamin D 25(OH)D and obesity have been suggested to increase BC risk, and the hormone leptin, which is important in weight regulation, may be involved in its carcinogenetic process [5 6].

25(OH)D is converted to its active hormonal form, 1-25-dihydroxyvitamin D ($1,25(OH)_2D$), by 1-α-hydroxylase, which is present in most tissues in the body [7]. Parathyroid hormone (PTH) and calcium level are important factors as they affect the enzymatic conversion from 25(OH)D to active 1, 25-$(OH)D_3$ in the kidney and may be involved in non-classical synthesis. Measurement of circulating 25(OH)D is considered the gold standard measurement of

vitamin D status as it integrates vitamin D exposure from oral intake from diet or supplements, as well as from exposure to ultraviolet radiation (UVR) [8 9]. Despite being the gold standard, total circulating 25(OH)D may not be the best measure of 25(OH)D exposure for all tumors. The "free hormone hypothesis" suggests that only unbound, free hormones can have biologic effects on target tissues [10]. To date, few studies have examined the role of vitamin D binding protein (DBP) also known as group-specific component or Gc-globulin, in the association between 25(OH)D and various cancer processes. DBP transports both 25(OH)D and $1,25(OH)_2D$ in circulation. This protein carries 88% of 25(OH)D and 85% of $1,25(OH)_2D$; an additional 12% of 25(OH)D and 15% of $1,25(OH)_2D$ circulate bound to albumin. Clinical laboratory assays of circulating 25(OH)D that are currently in use measure total 25(OH)D without differentiating between the bound and free forms. Thus, it remains unclear whether total or free 25(OH)D is more biologically relevant with respect to risk of BC. Two previous studies have examined free, in addition to total 25(OH)D, in relation to risk of BC. One found an inverse association between total 25(OH)D and bladder cancer that appeared to be restricted to participants with low DBP, suggesting that free 25(OH)D might be more strongly associated with risk of bladder cancer than total 25(OH)D [11]. The other study found no association between 25(OH)D overall or at any level of DBP concentration [12].

Body mass index (weight/height^2, BMI) is a reliable indicator of body fatness and can be categorized as underweight (<18.5), healthy weight (18.5-24.9), overweight (25-29.9) and obese (> 30). The prevalence of obesity in Norway has risen steeply the last decades. In the 1970s, about 15% of men and women were overweight (www.fhi.no/) while in 2013 the proportions were 58.4% and 47.3%, respectively [13]. Two meta-analyses, including 15 [14] and 11 [15] cohort studies conclude that obesity significantly increases the risk of BC. Leptin, a hormone involved in weight regulation [16], may be involved in this potential association. High leptin levels have been shown to impact development of several forms of cancer [17]. A study of Yuan et al. [6] shows that leptin receptors are aberrantly expressed in BC tissue and a high leptin level has been associated with BC carcinogenesis [5 6].

Low 25(OH)D levels are more frequent in obese persons, suggesting that 25(OH)D deficiency is associated with obesity and vice versa. 25(OH)D deficiency is suggested to be associated

with obesity [18-20], and both low 25(OH)D and obesity are suggested to contribute to development of BC.

Aims and hypotheses

Better-targeted BC primary and tertiary prevention (risk and survival) is warranted. The interplay between 25(OH)D and obesity and their associations with BC risk are poorly understood. We propose a study aiming to examine anthropometric data (BMI, height, weight, body surface area (BSA) and weight change over time) and serum levels of leptin, total and free 25(OH)D, in relation to BC risk and survival by using samples from Janus Serum Bank and associated data from population-based registries and surveys.

We hypothesize that

1 a. Obesity, BSA and weight change over time are associated with increased BC risk;

1b. Obesity, BSA and weight change over time are associated with reduced BC survival;

2a. Low free and total 25(OH)D level and high serum leptin levels (>4.1 ng/mL are associated with increased risk of BC;

2b. Low free and total 25(OH)D level and high serum leptin levels (>4.1 ng/mL are associated with reduced BC survival.

METHODS AND ANALYSIS

Study population and data sources

The Janus Serum Bank Cohort

The study will be carried out using the Janus Serum Bank Cohort, a population-based biobank for prospective cancer studies, containing serum samples and questionnaire data from 292,851 Norwegians who participated in one or more of five Norwegian Regional Health Studies in the period 1972–2003. Detailed description of the samples and data included in the Janus Serum Bank Cohort has been published elsewhere [21 22]. The quality aspects of long-term stored samples have been of high priority in the Janus Serum Bank and

component stability for a large number of hormones, proteins, metabolites and electrolytes has been investigated [23-25], including for both 25(OH)D, and leptin [26-29]. A unique 11-digit personal identification number (PIN), assigned to al Norwegian residents, will be used to link the Janus Cohort with population-based registries and surveys.

Population-based registries and surveys

The Cancer Registry of Norway (CRN) has collected notifications on cancer at a national level since 1953. Cancer reporting is mandatory by law, and reports from various sources ensures high quality and completeness (98.8%) [30]. The reporting system, based on pathology and cytology reports, clinical records, and death certificates, provides information about site, histological type and stage of disease at the time of diagnosis. CRN has been involved in the Janus Serum Bank operation since establishment in the early 1970s and has been responsible for the data handling; in 2004 the serum bank was integrated into the CRN. The following information is available for cancer cases: month and year of diagnosis, tumor site (International Classification of Diseases 7th revision [ICD-7 codes] converted into ICD-10 codes), histology (codes from ICD-Oncology 2nd and 3rd revision), clinical stage (local = no metastases, regional = metastasis in regional lymph nodes or surrounding area, distant = distant metastasis). In addition, all BC diagnoses in the Janus cohort will be reviewed and assigned a pathological T-stage, according to the American Joint Committee on Cancer (AJCC) staging manual 8th ed.[31].

The Norwegian Institute of Public Health (NIPH) has been responsible for conducting the national health surveys, upon which the Janus Serum Bank is partly based. All participants have completed questionnaires for assessment of lifestyle factors (i.e. smoking habits, alcohol use), at the time of serum collection. A database has been established, including data from these questionnaires, as well as measured body height and weight, blood pressure, cholesterol and triglycerides [32 33]. The Janus Cohort includes participants from five of the health studies: The Oslo Study I (1972–73), The Norwegian Counties Study (1974–78, 1977–83, and 1985–88), The Age 40 Program- Oslo (1981–99), The National Age 40 Program (1985–99) and The TROFINN Health Study (2001–03). A set of about 50 variables has been harmonized and standardized due to slightly different wording in the questionnaires [22].

Available variables include: height (cm), weight (kg), BMI (kg/m^2, categorized as 12–18.49, 18.5–24.9, 25.0–29.9, ≥30), smoking status (never, former, current), cigarettes per day (1–9, 10–14, ≥15), years of smoking (1–9, 10–29, ≥30), time since smoking cessation (<3months, 3months–1yr, 1–5yrs, >5yrs), and total physical activity (inactive, low, medium, high), based on leisure time activity. Estimated variables include BSA (m2) using the DuBois' equation (weight ^{0,4253} X height ^{07253} X 3 0.007184)[20]; and weight change calculated by subtracting the 1985–88 weight measure from the 1974– 78 measure (median time between the weight measurements of 10 years). Weight change will only be possible for a subgroup with repeated measurement of weight.

The Cause of death Registry has registered death certificates for all deaths in Norway since 1951. Cause of death registration is mandatory by law.

National Population Registry contains information on vital status (alive, emigrated or dead) of everyone that resides or has resided in Norway.

Statistics Norway has the responsibility of covering the needs for official statistics on the Norwegian population including individual data on settlements, migration, occupation and level of education.

Using the 11-digit PIN number we will link data from four different sources to set up the research file, illustrated in Figure 1.

Study design

Sub-study I: a prospective cohort study

In a prospective cohort study among all individuals in the Janus Serum Bank Cohort (n = 292,851) (Figure 1), we will explore baseline BMI in relation to bladder cancer risk. Among the included BC cases we will investigate baseline BMI and BC survival. By 2014, the cohort included 2 347 BC cases (ICD-10: C67). BC cases of both muscle invasive and non-invasive urothelial cell carcinoma, will be included in the study. Educational level, occupation, age,

sex, physical activity, smoking habits, cholesterol, triglycerides and blood pressure will be included in the statistical analyses as confounders.

Sub-study II: a nested case-control study

The study will be nested within the prospective cohort described above (Study I), including a) 400 BC cases of high grade tumors, including muscle invasive (T2-T4) and non muscle invasive (Ta, T1 and carcinoma in situ (Tis) cancer cases, and b) 400 controls alive and without a cancer diagnosis at the time of the BC diagnosis of the cases, matched 1:1 on sex, age (+/-1 year) and date of serum sampling (+/-1 month). Minimum time from blood draw to diagnosis will be 5 years. The serum samples will be analysed for total 25(OH)D, vitamin D binding protein and leptin. As PTH and albumin-adjusted calcium level affect the enzymatic conversion from 25(OH)D to active 1,25-$(OH)D_3$ and might be involved in the non-classical synthesis as well, measurement of these components will be taken into account.

In sub-study I, we will investigate the association between BMI and BC, and for sub –study II we will in addition include vitamin D levels. Overall we will focus on disentangle the relationship of vitamin D, BMI and BC. This will be done in two ways:

1. By implementing regression models including and interaction effect of vitamin D and BMI on BC.
2. By testing the hypothesis whether the effect of BMI on BC is mediated by vitamin D using mediation analysis[34].

Statistical methods

In the cohort study, Cox regression models will be used to estimate hazard ratios (HR) with 95% confidence intervals (CI) of BC and survival after BC, taking into account stage at diagnosis, and BMI and including adjusting for season of blood collection for vitamin D. In the nested case-control study, conditional logistic regression models will be used to estimate the odds ratio (OR) with 95% CI.

In order to find out whether Vitamin D acts(totally or partly) as a mediator, modern causal inference theory will be used to estimate different types of effects[34].

As we have a number of potential confounding variables, we will use directed acyclic graphs to select variables to include in the statistical models. Confounding variables will be included in the models and tests of interaction effects will be performed when relevant.

The tests will be two-sided and $p<0.05$ will be considered statistically significant. Statistical analyses will be conducted using R package and[35] Stata version 14.1 (StataCorp, College Station, TX, USA).

Laboratory analyses

The serum samples (aliquots of 400 µL) will be analysed for total 25(OH) D, DBP, leptin, PTH, albumin, and calcium. The Hormone laboratory at Aker hospital, Oslo, Norway will analyze 25(OH)D, DBP, PTH and leptin. The Hormone Laboratory is accredited by the Norwegian Accreditation as a testing laboratory and complies with the requirements of the NS-EN ISO/IEC 17025 standards. Albumin and calcium will be analyzed at Department of Medical Biochemistry, Oslo University Hospital accredited by the Norwegian Accreditation reg. no TEST 103 that complies with the requirements of the NS-EN ISO 1518. The sample donors, identity and case control status will be blinded for the laboratory staff. Quality control samples from the biobank will be included in every batch to examine inter-batch and intra-batch variability.

Power and sample size

Sub-study I Within the prospective cohort (n = 292,851), there are more than 2,300 BC cases reported to the cancer registry until the end of follow up time. Assuming the risk of developing bladder cancer is 1% for the normal weight group and 5% for the obese group [14], one would need a sample size of 586 to have 80% power. Since the sample size here is significantly larger, one can safely determine that the study has adequate statistical power.

Sub-study II In the case-control study nested into the prospective cohort, the statistical power will depend on: i) proportion of exposure in the population, ii) sample size (cases and controls) and iii) the minimum difference that is possible to detect.

Table 1 shows the smallest detectable OR according to proportion of controls exposed to low vitamin D (25(OH)D and high leptin levels, for different sample sizes. The power is 0.80 and a significance level of 0.05 (www.krothman.hostbyet2. com/Episheet.xls). The expected proportions of exposed controls were based on previous studies on serum samples from the Janus Cohort. For 25(OH)D, a study on prostate cancer reported that 4.4% and 30.6% of the controls had 25(OH)D levels below 30 nmol/L and 50 nmol/L, respectively. [29] For leptin, a study on colon cancer reported that 20% of the controls had a leptin level of 4.1 ng/mL or higher.[27]

Table 1. Odds ratio (OR) based on proportion of exposed controls and sample size

		Study case:control= 1:1		
Proportion of exposed controls		Number of cases		
		n=500	n=400	n=300
55%*	OR	1.44	1.50	1.6
45%*	OR	1.43	1.49	1.58
30%**	OR	1.45	1.51	1.62

*Exposure = 25(OH)D deficiency (25(OH)D < 50 nmol /L);

**Exposure = high serum leptin levels (>4.1 ng/mL)

Based on results in the table above, we consider 400 matched case-control pairs as a sufficient sample size for the case-control study.

Data analysis plan

The following analyses will be conducted to test our hypotheses:

• Hypothesis 1.a: A prospective cohort analysis of pre-diagnostic BMI and other anthropometric measures in relation to BC risk using the complete Janus Cohort (n = 292,851)

• Hypothesis 1.b: A prospective cohort analysis of pre-diagnostic BMI and other anthropometric measures in relation to survival after a BC, using all BC cases in the Janus Cohort (n ≈ 2,650)

• Hypothesis 2.a: A nested case-control analysis of BC risk according to pre-diagnostic serum levels of total and free 25(OH)D, and leptin in 400 matched case-control pairs.

• Hypothesis 2.b: A prospective analysis of survival after a BC (n=400) according to pre-diagnostic serum levels of 25(OH)D and leptin.

Study strengths and limitations

A major strength of the large sample set of more than 2,300 incident BC cases. Also use of individual PIN for linkages between multiple data sources, to establish a virtually complete study file, with exception of data on histopathology, is a strength. The data sources are high quality population-based registries, with high degree of completeness. The bladder cancer diagnoses are coded according to ICD-0. To get information on staging, and control the data quality, all histopathological information will be reviewed and characterized by tumor-stage (T-stage). Another strength of this study is that the public health data has been quality assured, structured and harmonized [22]. The use of pre-diagnostic samples assure the proper temporality of the relationship between exposure and BC, limiting the possibility of reverse causality.

Treatment data is of importance when evaluating the survival analyses. These data are missing and will be a limitation of this study

ETHICS AND DISSEMINATION

The Regional Committee for Medical and Health Research Ethics has approved the study. The different data registries have approved that the use of a de-identified dataset. An ID-

key, consisting of the 11-digit PIN and a study-specific ID number, will be stored and governed by a third party unavailable to the research team.

All results will be published in relevant peer- reviewed international scientific journals and presented at conferences, nationally and internationally. Results of importance will be directly communicated to health authorities and to clinicians where the annual national oncology conference “Onkologisk forum” can serve as a platform for knowledge distribution. Results of importance will also be disseminated through press releases and to user groups like the Norwegian Cancer Society. The CRN website is a potential channel to reach patients organizations and the public.

Authors' contributions

REG prepared the study. TER, JSS, HHH, HL, BA, KA, EW and AM contributed to the study design and reviewed and revised the protocol critically for important intellectual content, and approved the final versions. REG is the guarantor.

Data sharing

Requests for data sharing/case pooling may be directed to the corresponding author. This study uses third-party data derived from State government registries, which are ultimately governed by their ethics committees and data custodians. Thus, any requests to share these data will be subject to formal approval from each data source used in this study.

Funding

The research study has been reviewed and granted funding by the Norwegian Cancer Society (no. 182308-2016) and the Cancer Registry of Norway Research Fund

Conflicts of interest

None declared

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* 2016;66(1):7-30. doi: 10.3322/caac.21332 [published Online First: 2016/01/09]

2. Svatek RS, Hollenbeck BK, Holmang S, et al. The Economics of Bladder Cancer: Costs and Considerations of Caring for This Disease. *European urology* 2014 doi: 10.1016/j.eururo.2014.01.006 [published Online First: 2014/01/30]
3. Freedman ND, Silverman DT, Hollenbeck AR, et al. Association between smoking and risk of bladder cancer among men and women. *Jama* 2011;306(7):737-45. doi: 10.1001/jama.2011.1142 [published Online First: 2011/08/19]
4. Olfert SM, Felknor SA, Delclos GL. An updated review of the literature: risk factors for bladder cancer with focus on occupational exposures. *Southern medical journal* 2006;99(11):1256-63. doi: 10.1097/01.smj.0000247266.10393.72 [published Online First: 2007/01/02]
5. Ishii N, Wei M, Kakehashi A, et al. Enhanced Urinary Bladder, Liver and Colon Carcinogenesis in Zucker Diabetic Fatty Rats in a Multiorgan Carcinogenesis Bioassay: Evidence for Mechanisms Involving Activation of PI3K Signaling and Impairment of p53 on Urinary Bladder Carcinogenesis. *Journal of toxicologic pathology* 2011;24(1):25-36. doi: 10.1293/tox.24.25 [published Online First: 2012/01/25]
6. Yuan SS, Chung YF, Chen HW, et al. Aberrant expression and possible involvement of the leptin receptor in bladder cancer. *Urology* 2004;63(2):408-13. doi: 10.1016/j.urology.2003.08.038 [published Online First: 2004/02/20]
7. Bikle D. Nonclassic actions of vitamin D. *The Journal of clinical endocrinology and metabolism* 2009;94(1):26-34. doi: 10.1210/jc.2008-1454 [published Online First: 2008/10/16]
8. Giovannucci E. The epidemiology of vitamin D and cancer incidence and mortality: a review (United States). *Cancer causes & control : CCC* 2005;16(2):83-95. doi: 10.1007/s10552-004-1661-4 [published Online First: 2005/05/04]
9. Holick MF. Vitamin D status: measurement, interpretation, and clinical application. *Annals of epidemiology* 2009;19(2):73-8. doi: 10.1016/j.annepidem.2007.12.001 [published Online First: 2008/03/11]
10. Pike JW FD, Glorieux FH. Vitamin D. San Diego: Academic Press, 1997.
11. Mondul AM, Weinstein SJ, Virtamo J, et al. Influence of vitamin D binding protein on the association between circulating vitamin D and risk of bladder cancer. *British journal of cancer* 2012;107(9):1589-94. doi: 10.1038/bjc.2012.417 [published Online First: 2012/09/20]
12. Mondul AM, Weinstein SJ, Horst RL, et al. Serum vitamin D and risk of bladder cancer in the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening trial. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2012;21(7):1222-5. doi: 10.1158/1055-9965.epi-12-0439 [published Online First: 2012/05/25]
13. Ng M, Fleming T, Robinson M, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet (London, England)* 2014;384(9945):766-81. doi: 10.1016/s0140-6736(14)60460-8 [published Online First: 2014/06/02]
14. Sun JW, Zhao LG, Yang Y, et al. Obesity and risk of bladder cancer: a dose-response meta-analysis of 15 cohort studies. *PloS one* 2015;10(3):e0119313. doi: 10.1371/journal.pone.0119313 [published Online First: 2015/03/25]
15. Qin Q, Xu X, Wang X, et al. Obesity and risk of bladder cancer: a meta-analysis of cohort studies. *Asian Pacific journal of cancer prevention : APJCP* 2013;14(5):3117-21. [published Online First: 2013/06/28]
16. Klok MD, Jakobsdottir S, Drent ML. The role of leptin and ghrelin in the regulation of food intake and body weight in humans: a review. *Obesity reviews : an official journal of the International Association for the Study of Obesity* 2007;8(1):21-34. doi: 10.1111/j.1467-789X.2006.00270.x [published Online First: 2007/01/11]
17. Garofalo C, Surmacz E. Leptin and cancer. *Journal of cellular physiology* 2006;207(1):12-22. doi: 10.1002/jcp.20472 [published Online First: 2005/08/20]
18. Foss YJ. Vitamin D deficiency is the cause of common obesity. *Medical hypotheses* 2009;72(3):314-21. doi: 10.1016/j.mehy.2008.10.005 [published Online First: 2008/12/05]

19. Holick MF. Vitamin D deficiency. *The New England journal of medicine* 2007;357(3):266-81. doi: 10.1056/NEJMra070553 [published Online First: 2007/07/20]
20. Mai XM, Chen Y, Camargo CA, Jr., et al. Cross-sectional and prospective cohort study of serum 25-hydroxyvitamin D level and obesity in adults: the HUNT study. *American journal of epidemiology* 2012;175(10):1029-36. doi: 10.1093/aje/kwr456 [published Online First: 2012/02/09]
21. Langseth H, Gislefoss RE, Martinsen JI, et al. Cohort Profile: The Janus Serum Bank Cohort in Norway. *International journal of epidemiology* 2016 doi: 10.1093/ije/dyw027 [published Online First: 2016/04/12]
22. Hjerkind KV, Gislefoss RE, Tretli S, et al. Cohort Profile Update: The Janus Serum Bank Cohort in Norway. *International journal of epidemiology* 2017 doi: 10.1093/ije/dyw302 [published Online First: 2017/01/15]
23. Gislefoss RE, Grimsrud TK, Morkrid L. Long-term stability of serum components in the Janus Serum Bank. *Scandinavian journal of clinical and laboratory investigation* 2008;68(5):402-9. doi: 10.1080/00365510701809235 [published Online First: 2008/08/30]
24. Gislefoss RE, Grimsrud TK, Morkrid L. Stability of selected serum proteins after long-term storage in the Janus Serum Bank. *ClinChemLab Med* 2009;47(5):596-603.
25. Gislefoss RE, Grimsrud TK, Morkrid L. Stability of selected serum hormones and lipids after long-term storage in the Janus Serum Bank. *Clinical biochemistry* 2015;48(6):364-9. doi: 10.1016/j.clinbiochem.2014.12.006 [published Online First: 2014/12/20]
26. Stattin P, Kaaks R, Johansson R, et al. Plasma leptin is not associated with prostate cancer risk. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2003;12(5):474-5. [published Online First: 2003/05/17]
27. Stattin P, Lukanova A, Biessy C, et al. Obesity and colon cancer: does leptin provide a link? *International journal of cancer* 2004;109(1):149-52. doi: 10.1002/ijc.11668 [published Online First: 2004/01/22]
28. Tretli S, Schwartz GG, Torjesen PA, et al. Serum levels of 25-hydroxyvitamin D and survival in Norwegian patients with cancer of breast, colon, lung, and lymphoma: a population-based study. *Cancer causes & control : CCC* 2012;23(2):363-70. doi: 10.1007/s10552-011-9885-6 [published Online First: 2011/12/24]
29. Meyer HE, Robsahm TE, Bjorge T, et al. Vitamin D, season, and risk of prostate cancer: a nested case-control study within Norwegian health studies. *The American journal of clinical nutrition* 2013;97(1):147-54. doi: 10.3945/ajcn.112.039222 [published Online First: 2012/11/30]
30. Larsen IK, Smastuen M, Johannesen TB, et al. Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness. *European journal of cancer (Oxford, England : 1990)* 2009;45(7):1218-31. doi: 10.1016/j.ejca.2008.10.037 [published Online First: 2008/12/19]
31. Amin MB, Edge, S., Greene, F., Byrd, D.R., Brookland, R.K., Washington, M.K., Gershenwald, J.E., Compton, C.C., Hess, K.R., Sullivan, D.C., Jessup, J.M., Brierley, J.D., Gaspar, L.E., Schilsky, R.L., Balch, C.M., Winchester, D.P., Asare, E.A., Madera, M., Gress, D.M., Meyer, L.R. . AJCC Cancer Staging Manual 8 th edition. New York: Springer International Publishing 2017.
32. Bjartveit K, Foss OP, Gjervig T, et al. The cardiovascular disease study in Norwegian counties. Background and organization. *Acta medica Scandinavica Supplementum* 1979;634:1-70. [published Online First: 1979/01/01]
33. PG L-L. Arven fra Statens skjermbildefotografering (SSF)/Statens helseundersøkelser (SHUS) en introduksjon til databanken. *Tidsskr Norsk Forening for Epidemiol* 2003;13:14.
34. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychol Methods* 2010;15(4):309-34. doi: 10.1037/a0020761 [published Online First: 2010/10/20]
35. Tingley D YT, Hirsoe K, Imai K. Mediation: R package for causal mediation anakysis. *Journal of Statistical Software* 2014;59(5):1-34.

Figure legends

Figure 1. Data collection from different sources using PIN

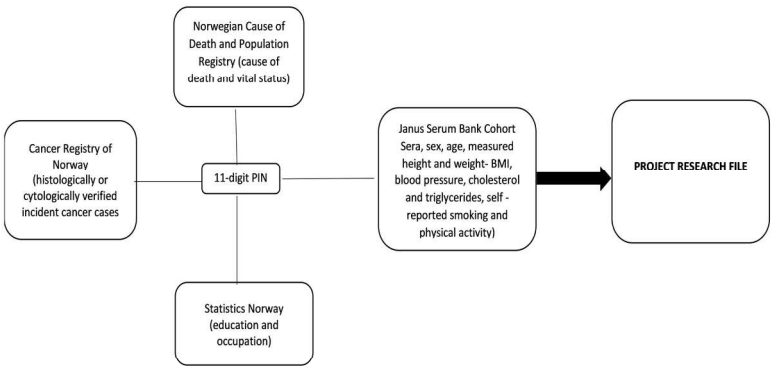

Figure 1. Data collection from different sources using PIN

190x107mm (300 x 300 DPI)

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract P 1
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found P 2
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported P 3-4
Objectives	3	State specific objectives, including any prespecified hypotheses P 5
Methods		
Study design	4	Present key elements of study design early in the paper P 7
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection P 5-7
Participants	6	(a) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up P 5-7
		(b) For matched studies, give matching criteria and number of exposed and unexposed P 7-8
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
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 P 5-7, 9
Bias	9	Describe any efforts to address potential sources of bias P 8
Study size	10	Explain how the study size was arrived at P 9-10
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding P 8
		(b) Describe any methods used to examine subgroups and interactions P 8
		(c) Explain how missing data were addressed
		(d) If applicable, explain how loss to follow-up was addressed
		(e) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed P 8
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram P 7
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders P 7-8
		(b) Indicate number of participants with missing data for each variable of interest
		(c) Summarise follow-up time (eg, average and total amount)
Outcome data	15*	Report numbers of outcome events or summary measures over time Not applicable
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

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion Not applicable		
Key results	18	Summarise key results with reference to study objectives
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
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
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 P12

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.