

## Article

# Genetically Raised Circulating Bilirubin Levels and Risk of Ten Cancers: A Mendelian Randomization Study

Nazlisadat Seyed Khoei <sup>1,†</sup>, Robert Carreras-Torres <sup>2,†</sup>, Neil Murphy <sup>3</sup>, Marc J. Gunter <sup>3</sup>, Paul Brennan <sup>4</sup>, Karl Smith-Byrne <sup>4</sup>, Daniela Mariosa <sup>4</sup>, James Mckay <sup>4</sup>, Tracy A. O'Mara <sup>5</sup> on behalf of ECAC group, Ruth Jarrett <sup>6</sup>, Henrik Hjalgrim <sup>7,8</sup>, Karin E. Smedby <sup>9,10</sup>, Wendy Cozen <sup>11</sup>, Kenan Onel <sup>12</sup>, Arjan Diepstra <sup>13</sup>, Karl-Heinz Wagner <sup>1,†</sup> and Heinz Freisling <sup>3,\*‡</sup>

- <sup>1</sup> Department of Nutritional Sciences, Faculty of Life Sciences, University of Vienna, 1090 Vienna, Austria; nazlisadat.seyedkhoei@univie.ac.at (N.S.K.); karl-heinz.wagner@univie.ac.at (K.-H.W.)
  - <sup>2</sup> Colorectal Cancer Group, ONCOBELL Program, Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, 8908 Barcelona, Spain; rcarreras@idibell.cat
  - <sup>3</sup> Nutrition and Metabolism Branch, International Agency for Research on Cancer (IARC-WHO), 69008 Lyon, France; MurphyN@iarc.fr (N.M.); gunterm@iarc.fr (M.J.G.)
  - <sup>4</sup> Genomic Epidemiology Branch, International Agency for Research on Cancer (IARC-WHO), 69008 Lyon, France; BrennanP@iarc.fr (P.B.); smith-byrne@fellows.iarc.fr (K.S.-B.); mariosad@iarc.fr (D.M.); mckay@iarc.fr (J.M.)
  - <sup>5</sup> Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane 4006, Queensland, Australia; Tracy.O'Mara@qimrberghofer.edu.au
  - <sup>6</sup> Institute of Infection, Immunity and Inflammation, MRC-University of Glasgow Centre for Virus Research, Glasgow G61 1QH, UK; Ruth.jarrett@glasgow.ac.uk
  - <sup>7</sup> Department of Epidemiology Research, Statens Serum Institut, 2300 Copenhagen, Denmark; HHJ@ssi.dk
  - <sup>8</sup> Department of Hematology, Finsen Centre, 2100 Copenhagen, Denmark
  - <sup>9</sup> Department of Medicine Solna, Division of Clinical Epidemiology, Karolinska Institutet, 171 77 Stockholm, Sweden; Karin.Ekstrom.Smedby@ki.se
  - <sup>10</sup> Department of Hematology, Karolinska University Hospital, S-141 86 Stockholm, Sweden
  - <sup>11</sup> Departments of Preventive Medicine and Pathology, Keck School of Medicine of USC, University of Southern California, Los Angeles, CA 90033, USA; wcozen@usc.edu
  - <sup>12</sup> Department of Genetics and Genomics Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 60637, USA; kenan.onel@mssm.edu
  - <sup>13</sup> Department of Pathology and Medical Biology, University of Groningen, University Medical Center Groningen, 9713 Groningen, The Netherlands; a.diepstra@umcg.nl
- † These authors contributed equally to this work, Shared first authorship.  
‡ These authors contributed equally to this work, Shared last authorship.  
\* Correspondence: Freislingh@iarc.fr; Tel.: +33-472-738-664

**Citation:** Seyed Khoei, N.; Carreras-Torres, R.; Murphy, N.; Gunter, M.J.; Brennan, P.; Smith-Byrne, K.; Mariosa, D.; Mckay, J.; O'Mara, T.A. on behalf of ECAC group; Jarrett, R.; et al. Genetically Raised Circulating Bilirubin Levels and Risk of Ten Cancers: A Mendelian Randomization Study. *Cells* **2021**, *10*, 394. <https://doi.org/10.3390/cells10020394>

Academic Editor: Frank Schnütgen

Received: 7 January 2021

Accepted: 5 February 2021

Published: 15 February 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Bilirubin, an endogenous antioxidant, may play a protective role in cancer development. We applied two-sample Mendelian randomization to investigate whether genetically raised bilirubin levels are causally associated with the risk of ten cancers (pancreas, kidney, endometrium, ovary, breast, prostate, lung, Hodgkin's lymphoma, melanoma, and neuroblastoma). The number of cases and their matched controls of European descent ranged from 122,977 and 105,974 for breast cancer to 1200 and 6417 for Hodgkin's lymphoma, respectively. A total of 115 single-nucleotide polymorphisms (SNPs) associated ( $p < 5 \times 10^{-8}$ ) with circulating total bilirubin, extracted from a genome-wide association study in the UK Biobank, were used as instrumental variables. One SNP (rs6431625) in the promoter region of the uridine-diphosphoglucuronate glucuronosyltransferase1A1 (*UGT1A1*) gene explained 16.9% and the remaining 114 SNPs (non-*UGT1A1* SNPs) explained 3.1% of phenotypic variance in circulating bilirubin levels. A one-standard-deviation increment in circulating bilirubin ( $\approx 4.4 \mu\text{mol/L}$ ), predicted by non-*UGT1A1* SNPs, was inversely associated with risk of squamous cell lung cancer and Hodgkin's lymphoma (odds ratio (OR) 0.85, 95% confidence interval (CI) 0.73–0.99,  $P$  0.04 and OR 0.64, 95% CI 0.42–0.99,  $p$  0.04, respectively), which was confirmed after removing potential pleiotropic SNPs. In contrast, a positive association was observed with the risk of breast cancer after removing potential pleiotropic SNPs (OR 1.12, 95% CI

1.04–1.20,  $p$  0.002). There was little evidence for robust associations with the other seven cancers investigated. Genetically raised bilirubin levels were inversely associated with risk of squamous cell lung cancer as well as Hodgkin’s lymphoma and positively associated with risk of breast cancer. Further studies are required to investigate the utility of bilirubin as a low-cost clinical marker to improve risk prediction for certain cancers.

**Keywords:** bilirubin; UGT1A1; Mendelian randomization; cancer risk

## 1. Introduction

Cancer is a major cause of morbidity and mortality globally, and the number of new cancer cases is expected to increase further over the next decades (CDC, 2020). In 2018, there were over 18 million new cancer cases and nine million cancer-related deaths [1].

Cancer-promoting inflammation is an enabling characteristic of cancer development, and inflammatory cells can also release reactive oxygen species [2]. A major cause of cancer is damage to DNA as a result of oxidative stress, mainly due to excess reactive oxygen species, antioxidants depletion, or both [3].

Bilirubin, a metabolic by-product of hemoglobin breakdown, is one of the most potent endogenous antioxidants of the human body and also has substantial anti-inflammatory properties [4–7]. Therefore, bilirubin may play a preventive role in cancer development. Blood levels of bilirubin are under genetic control via expression of uridine-diphosphoglucuronate glucuronosyltransferase1A1 (UGT1A1) in the liver, which converts insoluble bilirubin into a more water-soluble form for renal and biliary excretion [8]. Individuals homozygous for seven thymine–adenine (TA)-repeats (7/7) at the *UGT1A1\*28* locus have decreased enzyme activity, which leads to a less effective glucuronidation and moderately higher than normal blood levels of bilirubin (known as Gilbert’s syndrome, GS) [9,10].

The seven TA-repeats allele of *UGT1A1\*28* polymorphism underlying GS was investigated in relation to cancers of the endometrium [11], ovary [12], lung [13,14], breast [15], and prostate [16]. However, results of these studies were inconclusive, did not specifically investigate bilirubin as a putative cancer risk factor, and had limited sample size (range of number of cases 129 to 310), with the exception of Horsfall et al., where an inverse association was observed between genetically raised bilirubin levels and lung cancer risk among current smokers [14].

In this study, we investigated whether genetically raised circulating bilirubin levels are causally associated with risk of ten cancers using a Mendelian randomization (MR) approach. This technique uses genetic variation as instrumental variables [17], and in the absence of pleiotropy, an association between the genetic instruments and the outcome implies that the risk factor of interest may have a causal role in disease etiology (here: cancer risk) [18]. MR addresses unmeasured confounding (e.g., by smoking), which is a major limitation of observational studies [17]. The ten cancer types were investigated in large international consortia and selected based on previous evidence (cancers of the lung, ovary, endometrium, breast, and prostate) [11–16] or biological plausibility (pancreatic cancer, renal cell cancer, Hodgkin’s lymphoma, melanoma, and neuroblastoma).

## 2. Materials and Methods

In a two-sample MR approach, we first identified 115 single-nucleotide polymorphisms (SNPs) that were genome-wide associated ( $p < 5 \times 10^{-8}$ ) with circulating total bilirubin levels in a genome-wide association study (GWAS) that included 317,639 individuals of European ancestry (white British) from the UK Biobank (UKB) (non-British white, South Asian, African, and East Asian GWASs were excluded based on a combination of self-identification and refinement using population-specific genotype principal

components) [19]. The UKB project, a long-term prospective cohort study, recruited about 500,000 people aged between 40 and 69 years in 2006–2010 from across the UK [20]. The raw total bilirubin levels were adjusted for age, sex, the top 40 principal components for population stratification, recruitment center, indicators of socioeconomic status, and potential technical confounders (blood and urine sampling time, fasting time, and sample dilution factor) [19]. SNPs were independently associated with total bilirubin levels, which were reflected by measures of linkage disequilibrium ( $LD R^2 < 0.001$ ). The SNPs with ambiguous strand codification (T-A or guanine-cytosine, G-C) were replaced by SNPs in  $LD R^2 > 0.8$  in European populations using the *proxysnps* R package. The summary statistics for the associations of SNP allele dosage with standardized bilirubin levels are shown in Table S1.

One SNP (rs6431625) in the promoter region of the *UGT1A1* gene in chromosome 2 explained 16.9% of phenotypic variance in circulating total bilirubin levels, which was estimated as a function of the effect size for the risk factor in standard deviation units and the minor allele frequency [21]. This SNP was in strong linkage disequilibrium ( $LD R^2 = 0.74$ ) with the *UGT1A1*\*28 promoter TA-repeats polymorphism (rs3064744) [22]. The remaining 114 SNPs (non-*UGT1A1* SNPs) explained 3.1% of the total bilirubin variance [22] and provided an F-statistic for the strength of the relationship between the genetic instrument and the bilirubin levels of 89.1. The F-statistic is an estimation of the magnitude of the instrument bias (e.g., F-statistic <10 for the weak instruments) [23].

Second, ten cancer types were analyzed using genetic data that together summed up to a total of 336,110 cancer cases and 589,467 controls of European ancestry (Table 1). Cases and controls were individually matched in the original GWAS for each cancer, including pancreatic cancer (overall and sex subgroups) [24–26], renal cell cancer (overall and sex subgroups) [27], lung cancer (overall, ever and never smokers subgroups, and histological subtypes of adenocarcinoma, squamous cell, and small cell) [28], ovarian cancer (overall and serous subgroup) [29], endometrial cancer [30], breast cancer (overall and estrogen receptor (ER) subgroups) [31], prostate cancer [32], Hodgkin’s lymphoma [33], melanoma, [34] and neuroblastoma [35]. To prevent weak instrument bias, these genetic data did not include samples from the UKB. Each contributing study was approved by the appropriate institutional review board/ethics committee. All participants provided informed consent.

**Table 1.** Summary information of cancer GWAS samples and power assessment.

Cancer Type	Subtype	N Cases	N Controls	SNP Set	Minimum Detectable OR
Pancreatic cancer	overall	7110	7264	<i>UGT1A1</i> SNP	1.12/0.89
				Non- <i>UGT1A1</i> SNPs (n=113)	1.30/0.77
	men	3861	4056	<i>UGT1A1</i> SNP	1.17/0.86
				Non- <i>UGT1A1</i> SNPs (n=113)	1.43/0.70
	women	3252	3268	<i>UGT1A1</i> SNP	1.18/0.84
				Non- <i>UGT1A1</i> SNPs (n=113)	1.48/0.67
Renal cell cancer	overall	10,784	20,406	<i>UGT1A1</i> SNP	1.08/0.92
				Non- <i>UGT1A1</i> SNPs (n=111)	1.21/0.83
	men	3227	4916	<i>UGT1A1</i> SNP	1.17/0.86
				Non- <i>UGT1A1</i> SNPs (n=109)	1.43/0.70
	women	1992	3095	<i>UGT1A1</i> SNP	1.22/0.82
				Non- <i>UGT1A1</i> SNPs (n=109)	1.58/0.63
Lung cancer	overall	29,266	56,450	<i>UGT1A1</i> SNP	1.05/0.95
				Non- <i>UGT1A1</i> SNPs (n=109)	1.12/0.89
	ever smokers	23,223	16,964	<i>UGT1A1</i> SNP	1.07/0.93
				Non- <i>UGT1A1</i> SNPs (n=109)	1.17/0.85
	never smokers	2355	7504	<i>UGT1A1</i> SNP	1.17/0.85

				Non-UGT1A1 SNPs (n=109)	1.46/0.69
	adenocarcinoma	11,273	55,483	UGT1A1 SNP	1.07/0.93
				Non-UGT1A1 SNPs (n=109)	1.18/0.85
	squamous cell	7426	55,627	UGT1A1 SNP	1.09/0.92
				Non-UGT1A1 SNPs (n=109)	1.22/0.82
	small cell	2664	21,444	UGT1A1 SNP	1.15/0.87
				Non-UGT1A1 SNPs (n=109)	1.39/0.72
	overall	25,509	40,941	UGT1A1 SNP	1.06/0.95
				Non-UGT1A1 SNPs (n=111)	1.14/0.88
Ovarian cancer	serous	16,003	40,941	UGT1A1 SNP	1.07/0.94
				Non-UGT1A1 SNPs (n=111)	1.16/0.86
Breast cancer	overall	122,977	105,974	UGT1A1 SNP	1.03/0.97
				Non-UGT1A1 SNPs (n=112)	1.07/0.94
Endometrial cancer	overall	12,906	108,979	UGT1A1 SNP	1.07/0.94
				Non-UGT1A1 SNPs (n=110)	1.16/0.86
Prostate cancer	overall	79,194	61,112	UGT1A1 SNP	1.04/0.96
				Non-UGT1A1 SNPs (n=107)	1.09/0.92
Hodgkin's lymphoma	overall	1200	6417	UGT1A1 SNP	1.24/0.81
				Non-UGT1A1 SNPs (n=91)	1.65/0.61
Melanoma	overall	1804	1026	UGT1A1 SNP	1.31/0.77
				Non-UGT1A1 SNPs (n=75)	1.86/0.54
Neuroblastoma	overall	1627	3254	UGT1A1 SNP	1.23/0.81
				Non-UGT1A1 SNPs (n=57)	1.62/0.62

Abbreviations: N Number, OR odds ratio, SNP single nucleotide polymorphism.

SNPs summary estimates ( $\beta_{\text{cancer}}$ ) were retrieved from these recently published GWAS results, which were obtained from: their cancer genetic consortia [27–32,36], the Genotypes and Phenotypes database (dbGaP) [37], public web repositories, and the MR-Base platform [38].

Imputed SNPs were restricted based on imputed accuracy, and only SNPs with high imputation quality ( $R^2 > 0.8$ ) were selected for our analyses. SNPs summary statistics for genetic associations with risk of the ten cancers are shown in Table S1.

### Statistical Analyses

*A priori* power calculations were performed for MR associations of nominal statistical significance ( $\alpha < 0.05$ ) between both the *UGT1A1* SNP and non-*UGT1A1* SNPs, respectively, and cancer risk; given the explained variance and the sample sizes for the different cancers, using the method proposed by Burgess et al. [39].

Estimated risk effects were obtained for each genetic variant, named Wald estimate (genetic effect on cancer risk [ $\beta_{\text{cancer}}$ ]/ genetic effect on total bilirubin levels [ $\beta_{\text{bilirubin}}$ ]).

As the main MR approach used in the analyses of the large SNPs set, excluding the *UGT1A1* SNP, the Wald estimates were combined in a single estimate through a likelihood-based MR approach. This approach is considered to be the most robust under the general assumption for all MR methods regarding linear relationship between the exposure and the outcome [40], which we assumed in the range of bilirubin variation reflected by these SNPs.

The initial step in the sensitivity analyses was to apply the inverse-variance weighted (IVW) MR approach [41] and to assess the presence of outlier observations among the SNP Wald estimates using the MR pleiotropy residual sum and outlier (MR-PRESSO) test [42]. The MR-PRESSO approach identifies heterogeneity between SNP effects ( $p_{\text{Global}}$ ) as an evidence of directional horizontal pleiotropy, identifies outlier SNPs, and tests if the presence of outliers is biasing the estimation of risk ( $p_{\text{Distortion}}$ ).

To evaluate the extent to which directional pleiotropy may affect the risk estimates, we used the intercept test within a MR-Egger weighted linear regression approach [43]. Moreover, the weighted median [21] and the modal-based estimate MR approaches [44] were applied to estimate the weighted median and the mode of the density distribution of the SNP estimates. Both methods are less sensitive to the presence of potentially invalid SNPs. Finally, we assessed whether pleiotropic SNPs, thus potentially violating the exclusion restriction (horizontal pleiotropy) and the independence assumptions (no confounders), were driving the association estimates. We looked up the genetic association results of bilirubin SNPs with other phenotypes in the GWAS Catalog database [45] and obtained MR estimates and performed sensitivity analyses excluding the SNPs reaching a genome-wide threshold of association with other phenotypes.

Analyses were performed stratified by sex for pancreatic and renal cell cancers, also by subtypes for lung, ovarian, and breast cancer. We did not account for multiple testing, since we had a strong prior hypothesis based on biological plausibility and applying a strict multiple testing correction would likely have been overly conservative given the non-independence of risk for many of the cancers tested [46].

Scatter plots were used to depict the genetic association on total bilirubin levels and cancer risk. All statistical analyses and plots were performed using Stata SE14 (Stata Corporation, College Station, TX, USA) and R (*MR-PRESSO*, *Two-Sample MR*, *gwasrapidd*, and *ggplot2*; The R project). All statistical tests were two-tailed.

### 3. Results

The minimum odds ratios (OR) that our analyses were able to detect for the *UGT1A1* SNP and non-*UGT1A1* SNPs, respectively, and each cancer are shown in Table 1.

Each standard deviation ( $SD \approx 4.4 \mu\text{mol/L}$ ) increment in bilirubin levels predicted by rs6431625 in the *UGT1A1* gene was not associated with risk of pancreatic cancer (OR per one-standard deviation, 1-SD 1.02; 95% of confidence interval (CI) 0.95–1.11), whereas higher bilirubin levels predicted by non-*UGT1A1* SNPs showed an inverse association with pancreatic cancer overall (OR 0.74; 95% CI 0.61–0.89), with similar risk estimates among men and women (OR 0.75; 95% CI 0.58–0.96, and OR 0.73; 95% CI 0.55–0.97, per 1-SD increment, respectively) (Figure 1A). However, after removing SNPs ( $n = 22$ ) with potential pleiotropy, these associations were attenuated toward the null (Table S2). The scatter plot depicting the genetic associations of these instruments with bilirubin levels and with the risk of pancreatic cancer overall, and among men and women, including the likelihood-based MR estimate and its 95% CI, are shown in Figure S1A–C.

Neither the *UGT1A1* SNP nor the non-*UGT1A1* SNPs were associated with risk of renal cell cancer in men and women combined (Figure 1A and Figure S1D–F); however, a suggestive inverse association was observed between bilirubin levels predicted by the non-*UGT1A1* SNPs and renal cell cancer in men (OR 0.76; 95% CI 0.57–1.00). After removing potential pleiotropic SNPs ( $n = 22$ ), these associations were attenuated toward the null (Table S2).

Overall, we did not observe an association between bilirubin levels predicted by either the *UGT1A1* SNP or the non-*UGT1A1* SNPs and lung cancer risk (OR per 1-SD 0.98; 95% CI 0.94–1.02, by *UGT1A1* SNP; and OR 0.91; 95% CI 0.83–1.00, by non-*UGT1A1* SNPs) (Figure 1B). In stratified analyses, genetically raised bilirubin levels predicted by the non-*UGT1A1* SNPs, but not the *UGT1A1* SNP, were inversely associated with lung cancer risk among individuals who ever smoked, squamous cell, and small cell lung cancer subtypes with ORs equal to 0.86; 95% CI 0.76–0.96, 0.85; 95% CI 0.73–0.99, and 0.77; 95% CI 0.61–0.97, per 1-SD increment, respectively (Figure 1B). The scatter plots for bilirubin levels and lung cancer risk are shown in Figure S2. Among squamous cell carcinoma, the inverse associations were robust in terms of effect size to sensitivity analyses and after removing SNPs ( $n = 22$ ) with potential pleiotropy (Table S2).

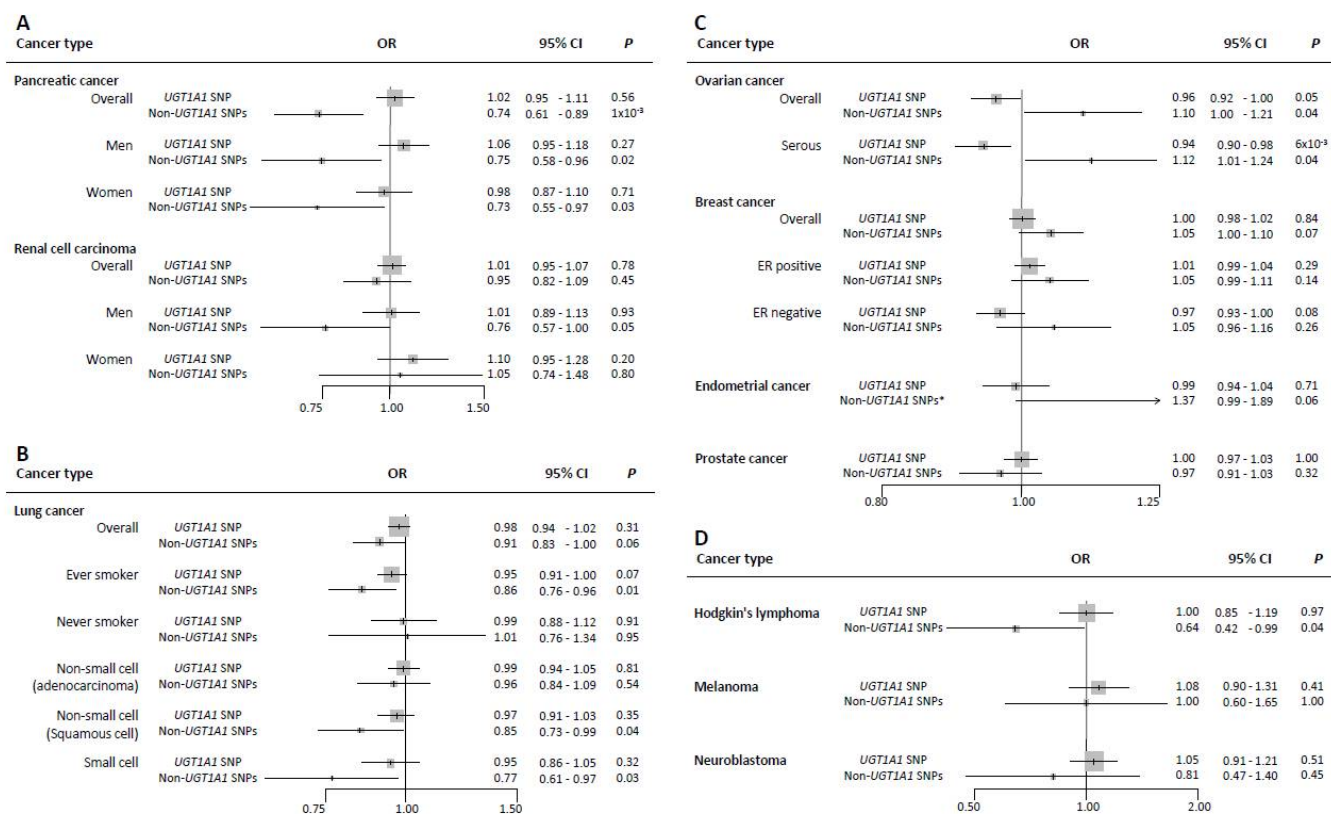
Higher bilirubin levels predicted by the *UGT1A1* SNP were weakly inversely associated with risk of ovarian cancer overall and serous ovarian cancer with ORs per 1-SD

increment equal to 0.96 (95% CI 0.92–1.00) and 0.94 (95% CI 0.90–0.98), respectively. In contrast, we observed a positive association between bilirubin levels predicted by non-*UGT1A1* SNPs and risk of ovarian cancer overall and serous ovarian cancer with ORs per 1-SD increment equal to 1.10 (95% CI 1.00–1.21) and 1.12 (95% CI 1.00–1.24), respectively (Figure 1C, Figure S3A,B). However, these associations were attenuated toward the null after removing SNPs ( $n = 22$ ) with potential pleiotropy. Furthermore, the MR-Egger Simex approach did not confirm the positive association suggested by the non-*UGT1A1* SNPs (Table S2).

Suggestive positive associations were observed between bilirubin levels, genetically predicted by non-*UGT1A1* SNPs, and risk of breast cancer (Figure 1C).

There was little evidence for associations between genetically raised bilirubin levels by either instrument and cancers of the endometrium (Figure 1C and Figure S3C–F) or prostate (Figure 1C and Figure S4).

Finally, higher bilirubin levels predicted by the non-*UGT1A1* SNPs were inversely associated with risk of Hodgkin’s lymphoma (OR 0.64, 95% CI 0.42–0.99) (Figure 1D and Figure S5A), while null results were observed for melanoma or neuroblastoma risk (Figure 1D and Figure S5B–C).



**Figure 1.** Forest plot of associations between genetically-predicted bilirubin levels and risk of ten cancers (per one-standard deviation (1-SD) increment in circulating total bilirubin levels equivalent to  $\approx 4.4 \mu\text{mol/L}$ ). (A): The association between 1-SD increment in bilirubin levels predicted by *UGT1A1* SNP and the non-*UGT1A1* SNPs with risks of pancreatic cancer and renal cell carcinoma (overall, men, and women). (B): The association between 1-SD increment in bilirubin levels predicted by *UGT1A1* SNP and the non-*UGT1A1* SNPs with risk of lung cancer (overall, ever smoker, never smoker, adenocarcinoma, squamous cell, and small cell lung cancers). (C): The association between 1-SD increment in bilirubin levels predicted by *UGT1A1* SNP and the non-*UGT1A1* SNPs with risks of ovarian cancer (overall and serous), breast cancer (overall, ER positive, and ER negative), endometrial cancer, and prostate cancer. (D): The association between 1-SD increment in bilirubin levels predicted by *UGT1A1* SNP and the non-*UGT1A1* SNPs with risks of Hodgkin’s lymphoma, melanoma, and neuroblastoma. The results are provided by a likelihood-based MR test, \*with the exception of endometrial

cancer, which results are provided by the Egger-MR test. Abbreviations: OR odds ratio, CI confidence interval,  $p$   $p$ -value, ER estrogen receptor.

#### *Sensitivity Analyses*

The IVW risk estimates performed almost identical to the main likelihood-based risk estimates, as both methods rely on the same assumptions and suffer similarly from pleiotropy. Outlier SNPs were detected by the MR-PRESSO test in some analyses; however, their presence was not biasing the initially estimated risk effects ( $p_{\text{Distortion}} > 0.25$ ) (Table S2). Additionally, the MR-Egger intercept test detected overall directional pleiotropy only in the case of endometrial cancer ( $p_{\text{intercept}} = 4 \times 10^{-4}$ ) and returned a potential positive association between bilirubin levels, predicted by non-*UGT1A1* SNPs, and endometrial cancer (OR 1.37; 95% CI 0.99–1.89) (Figure 1C and Table S2). The weighted median and modal-based approaches provided similar risk estimates as the MR-Egger test in the case of endometrial cancer and as the likelihood-based MR method for the other tested cancer types (Table S2). Finally, we identified a group of bilirubin SNPs that were genome-wide associated with other phenotypes, such as educational attainment, body mass index, and mean corpuscular volume of red blood cells. The MR analyses after removing these SNPs with potential pleiotropy ( $n = 20$  to  $22$  depending on GWAS data for specific cancer types) attenuated associations of most of the ten cancers investigated, except for squamous cell lung cancer (OR 0.80, 95% CI 0.65–0.99), breast cancer (OR 1.12, 95% CI 1.04–1.20), and Hodgkin's lymphoma (OR 0.61, 95% CI 0.32–1.14) (Table S2).

## **4. Discussion**

In this hypothesis-driven two-sample MR study, we investigated potential causal associations between genetically raised circulating bilirubin levels, a purported endogenous antioxidant, and risk of cancers of the pancreas, renal cell, endometrium, ovary, breast, prostate, lung, Hodgkin's lymphoma, melanoma, and neuroblastoma. We found that genetically raised bilirubin concentrations were inversely associated with risk of squamous cell lung cancer and Hodgkin's lymphoma, which is compatible with the antioxidant hypothesis of bilirubin, but positively associated with risk of breast cancer.

### *4.1. Lung Cancer*

The observed inverse association between genetically raised bilirubin levels and lung cancer risk among ever smokers, but not among never smokers, in our study are congruent with a prospective study in the UKB [14]. Similar inverse associations were also observed between serum bilirubin levels and the risk of lung cancer among male smoker in a Korean cohort ( $N$  cases = 240) [47] and in a prospective cohort in US ( $N$  cases = 386) [48]. Furthermore, in our study, we observed a robust inverse association with risk of squamous cell lung cancer subtype (OR 0.80, 95% CI 0.65–0.99 after excluding SNPs with potential pleiotropy), which is known to be strongly related to smoking. Taken together, genetically raised bilirubin levels may confer an advantage in terms of protecting people exposed to smoke oxidants against lung cancer [14,49].

### *4.2. Hodgkin's Lymphoma*

We observed robust inverse associations between genetically raised bilirubin levels and risk of Hodgkin's lymphoma.

Our study is the first linking bilirubin metabolism to Hodgkin's lymphoma. Given that one of the hypothesized root causes of Hodgkin's lymphoma is infection by Epstein-Barr virus [50], bilirubin might play a role in inhibiting replication of the virus as suggested for hepatitis C virus [51] and/or by balancing oxidative stress induced by Epstein-Barr virus infection [52].

### *4.3. Breast Cancer and Other Hormone-Related Cancers*

There was a suggestive positive association between bilirubin levels, genetically predicted by non-*UGT1A1* SNPs, and risk of breast cancer, in particular of the ER positive subtype (OR 1.12, 95% CI 1.03–1.22 after removing potential pleiotropic SNPs). These associations were relatively robust to a range of sensitivity analyses (Table S2). A meta-analysis of retrospective case-control studies ( $N$  cases = 5746,  $N$  controls = 8365) suggested that the *UGT1A1*\*28 allele 7/7 genotype is a potential risk factor for breast cancer in Caucasians [53]. Given that higher circulating levels of bilirubin can inhibit glucuronidation of estrogens (estradiol) into water-soluble molecules for excretion [54] suggests that higher bilirubin may affect these hormone-related cancers indirectly by interacting with the estrogen metabolism pathway.

In contrast to breast cancer, bilirubin levels predicted by the *UGT1A1* SNP were inversely associated with risk of ovarian cancer overall and serous ovarian cancer. If the inverse associations were genuine, then these findings are potentially consistent with a second-line antioxidant defense of raised bilirubin levels in the epithelial lining of the ovaries. Similar to lung cancer, oxidative stress is a critical factor in the initiation and development of ovarian cancer [55]. We are not aware of other studies investigating either circulating bilirubin levels or the *UGT1A1* polymorphism in relation to ovarian cancer risk, and further studies are warranted.

#### 4.4. Pancreatic Cancer

The findings for pancreatic cancer resemble the results of our previous two-sample MR study on the role of genetically raised bilirubin levels in colorectal cancer (CRC) using the same set of SNPs as instrumental variables [56]. We showed that bilirubin levels predicted by instrumental variables excluding the *UGT1A1* SNP were inversely associated with risk of CRC, supporting our hypothesis of anti-oxidative and anti-inflammatory properties of bilirubin. However, among men, bilirubin levels predicted by the *UGT1A1* SNP were positively associated with risk of CRC, and we argued that this could indicate either pleiotropic effects but potentially also pro-oxidative effects of an elevated bilirubin distribution among individuals with GS.

#### 4.5. Other Cancers (Renal Cell Cancer, Prostate Cancer, Melanoma, and Neuroblastoma)

We found no strong evidence for an association between genetically raised bilirubin levels and the risk of renal cell cancer, prostate cancer, melanoma, and neuroblastoma.

Our study has several strengths. First, we used thousands of cases and controls from several large genetic consortia and published GWAS, which signified the largest and most comprehensive MR study on genetically raised bilirubin levels and cancer risk. Second, our MR assumptions were met, which were supported by our sensitivity methods and pleiotropic SNPs exclusion. Finally, common sources of bias in observational studies, including residual confounding and reverse causation, were likely reduced. This study has some limitations, first, a potentially under-powered sample size for the non-*UGT1A1* instruments to detect small effects in some cancers. For these cancers, associations can occur by chance, especially when using weak instruments and small samples, which is a phenomenon known as weak instruments bias [57]. Second, we cannot completely rule out chance in explaining the weaker observed associations. However, we had a strong prior hypothesis based on biological plausibility. Third, we also stress that the genetic instruments for bilirubin do not necessarily reflect life-long exposure. However, assuming that the association between the instruments and bilirubin is constant over time, then the MR estimate could be interpreted as an estimate of the averaged cumulative effect of bilirubin on cancer within the age range at inclusion. Fourth, although we applied several strategies to account for horizontal pleiotropy, we cannot test and exclude the possibility that the main *UGT1A1* SNP affects cancer risk through pathways other than elevated bilirubin levels. There is a large region of linkage disequilibrium across the *UGT1A* locus that includes functional polymorphisms in UGT [58,59], which aside from bilirubin, also metabolize several xenobiotic and endogenous substances including (e.g., heterocyclic aromatic



amines, in well-done red meat) [56] or sex hormones [58]. Therefore, the observed differential associations of the *UGT1A1* SNP across cancer types may also reflect the modulated metabolism of such substances with carcinogenic potential. We also acknowledge that GWAS of associations of genetic variants on chronic diseases can be prone to selection bias from surviving competing risk. Methods to assess genetic effects on chronic diseases are needed to account for competing risk before recruitment [60].

Finally, we were not able to analyze non-linear associations, which would necessitate individual-level data. However, a non-linear relationship between bilirubin levels and cancer was not previously observed [56].

## 5. Conclusions

Genetically raised bilirubin levels were inversely associated with risk of squamous cell lung cancer as well as Hodgkin's lymphoma, and positively association with risk of breast cancer. These findings should help in setting priorities in future research on bilirubin levels and cancer risk. If confirmed in other studies, bilirubin could be a promising marker to risk stratify individuals for more frequent screening of selected cancers.

**Supplementary Materials:** The following are available online at [www.mdpi.com/2073-4409/10/2/394/s1](http://www.mdpi.com/2073-4409/10/2/394/s1). **Table S1** Summary statistics for genetic association of total bilirubin levels and risk of ten cancers. **Table S2** Results of sensitivity Mendelian randomization methods for total bilirubin levels and risk of ten cancers. **Figure S1** Scatter plots depicting the genetic associations between bilirubin levels and risk of pancreatic cancer and renal cell cancer, overall and sex subgroups. **Figure S2** Scatter plots depicting the genetic associations between bilirubin levels and lung cancer risk, overall, smoking subgroups, and histological subtypes. **Figure S3** Scatter plots depicting the genetic associations between bilirubin levels and risk of ovarian, endometrial, and breast cancers, overall and subtypes. **Figure S4** Scatter plots depicting the genetic association between bilirubin levels and prostate cancer risk. **Figure S5** Scatter plots depicting the genetic associations between bilirubin levels and risk of Hodgkin's lymphoma, melanoma, and neuroblastoma. Funding Sources and Acknowledgments.

**Author Contributions:** Conceptualization, N.S.K., R.C.-T., H.F., and K.-H.W.; Methodology, R.C.-T., K.S.-B., N.S.K., and H.F.; Software, R.C.-T. and K.S.-B.; Validation, R.C.-T. and K.S.-B.; Formal Analysis, R.C.-T. and K.S.-B.; Investigation, N.S.K., R.C.-T., H.F., and K.-H.W.; Resource, P.B., J.M., and T.A.O.; Data Curation, R.C.-T.; Writing—Original Draft Preparation, N.S.K. and R.C.-T.; Writing—Review and Editing, all listed authors; Supervision, H.F. and K.-H.W.; Project Administration, H.F. and K.-H.W.; Funding Acquisition, H.F. and K.H.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the French National Cancer Institute (INCa, grant nr. 2016-043), the Austrian Science Fund (FWF, grant nr. P 29608), and the European Commission, BBMRI-LPC (FP7, grant nr. 313010). Robert Carreras-Torres was funded by the European Union's Horizon 2020 research and innovation programs under the Marie Skłodowska-Curie grant agreement No 796216. We would also like to thank all other funding supports given in Supplementary Funding Sources and Acknowledgements. Study-specific funding sources are given in Supplementary Funding Sources and Acknowledgements. "Open Access Funding by the Austrian Science Fund (FWF)".

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and all studies providing data had approval from their respective institutional review boards. The UK Biobank has approval from the North West Multi-centre Research Ethics Committee, the National Information Governance Board for Health and Social Care in England and Wales, and the Community Health Index Advisory Group in Scotland. In addition, an independent Ethics and Governance Council was formed in 2004 to oversee UK Biobank's continuous adherence to the Ethics and Governance Framework that was developed for the study (<http://www.ukbiobank.ac.uk/ethics/>).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author freislingh@iarc.fr. Requests for the data require formal approval by the principal investigators of each genetic consortium.

**Acknowledgments** The authors thank all the participants who took part in this research, the numerous institutions, researchers, clinicians, and technical and administrative staff who have made possible the many studies contributing to this work.

**Conflicts of Interest** The authors declare no potential conflict of interest.

**Ethics Statement:** All participants provided written, informed consent.

**Disclaimer:** Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

**Collaborators: The authors' names and affiliations in ECAC GROUP:**

**Tracy A. O'Mara**<sup>1</sup>: Deborah J. Thompson<sup>2</sup>, Amanda B. Spurdle<sup>1</sup>, Frederic Amant<sup>3</sup>, Daniela Annibaldi<sup>3</sup>, Katie Ashton<sup>4,6</sup>, John Attia<sup>4,7</sup>, Paul L. Auer<sup>8,9</sup>, Matthias W. Beckmann<sup>10</sup>, Amanda Black<sup>11</sup>, Louise Brinton<sup>11</sup>, Daniel D. Buchanan<sup>12-15</sup>, Stephen J. Chanock<sup>16</sup>, Chu Chen<sup>17</sup>, Maxine M. Chen<sup>18</sup>, Timothy H.T. Cheng<sup>19</sup>, Linda S. Cook<sup>20,21</sup>, Marta Crous-Bous<sup>18,22</sup>, Kamila Czene<sup>23</sup>, Immaculata De Vivo<sup>18,22</sup>, Jeroen Depreuw<sup>3,24,25</sup>, Jennifer Anne Doherty<sup>26</sup>, Thilo Dörk<sup>27</sup>, Sean C. Dowdy<sup>28</sup>, Alison M. Dunning<sup>29</sup>, Matthias Dürst<sup>30</sup>, Douglas F. Easton<sup>2,29</sup>, Arif B. Ekici<sup>31</sup>, Peter A. Fasching<sup>10,32</sup>, Brooke L. Fridley<sup>33</sup>, Christine M. Friedenreich<sup>21</sup>, Montserrat García-Closas<sup>16,34</sup>, Mia M. Gaudet<sup>35</sup>, Graham G. Giles<sup>13,36</sup>, Dylan M. Glubb<sup>1</sup>, Ellen L. Goode<sup>37</sup>, Maggie Gorman<sup>19</sup>, Christopher A. Haiman<sup>38</sup>, Per Hall<sup>23,39</sup>, Susan E. Hankinson<sup>22,40</sup>, Patricia Hartge<sup>11</sup>, Catherine S. Healey<sup>29</sup>, Alexander Hein<sup>10</sup>, Peter Hillemanns<sup>27</sup>, Shirley Hodgson<sup>41</sup>, Erling Hoivik<sup>42,43</sup>, Elizabeth G. Holliday<sup>4,7</sup>, David J. Hunter<sup>18,44,45</sup>, Angela Jones<sup>19</sup>, Peter Kraft<sup>18,44</sup>, Camilla Krakstad<sup>42,43</sup>, Diether Lambrechts<sup>25,46</sup>, Loic Le Marchand<sup>47</sup>, Xiaolin Liang<sup>48</sup>, Annika Lindblom<sup>49,50</sup>, Jolanta Lissowska<sup>51</sup>, Jirong Long<sup>52</sup>, Lingeng Lu<sup>53</sup>, Anthony M. Magliocco<sup>54</sup>, Lynn Martin<sup>55</sup>, Mark McEvoy<sup>7</sup>, Roger L. Milne<sup>13,36,56</sup>, Miriam Mints<sup>57</sup>, Rami Nassir<sup>58</sup>, Irene Orlow<sup>48</sup>, Geoffrey Otton<sup>59</sup>, Claire Palles<sup>19</sup>, Paul D.P. Pharoah<sup>2,29</sup>, Loreall Pooler<sup>38</sup>, Jennifer Prescott<sup>22</sup>, Tony Proietto<sup>59</sup>, Timothy R. Rebbeck<sup>60,61</sup>, Stefan P. Renner<sup>62</sup>, Harvey A. Risch<sup>53</sup>, Matthias Rübner<sup>62</sup>, Ingo Runnebaum<sup>30</sup>, Carlotta Sacerdote<sup>63,64</sup>, Gloria E. Sarto<sup>65</sup>, Fredrick Schumacher<sup>66</sup>, Rodney J. Scott<sup>4,6,67</sup>, V. Wendy Setiawan<sup>38</sup>, Mitul Shah<sup>29</sup>, Xin Sheng<sup>38</sup>, Xiao-Ou Shu<sup>52</sup>, Melissa C. Southey<sup>12,56</sup>, Emma Tham<sup>49,68</sup>, Ian Tomlinson<sup>19,55</sup>, Jone Trovik<sup>42,43</sup>, Constance Turman<sup>18</sup>, David Van Den Berg<sup>38</sup>, Adriaan Vanderstichele<sup>69</sup>, Zhaoming Wang<sup>11</sup>, Penelope M. Webb<sup>70</sup>, Nicolas Wentzensen<sup>11</sup>, Henrica M.J. Werner<sup>42,43</sup>, Stacey J. Winham<sup>71</sup>, Lucy Xia<sup>38</sup>, Yong-Bing Xiang<sup>72</sup>, Hannah P. Yang<sup>11</sup>, Herbert Yu<sup>47</sup>, Wei Zheng<sup>52</sup>

<sup>1</sup> Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia.

<sup>2</sup> Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.

<sup>3</sup> Department of Obstetrics and Gynecology, Division of Gynecologic Oncology, University Hospitals KU Leuven, University of Leuven, Leuven, Belgium.

<sup>4</sup> Hunter Medical Research Institute, John Hunter Hospital, Newcastle, New South Wales, Australia.

<sup>5</sup> Centre for Information Based Medicine, University of Newcastle, Callaghan, New South Wales, Australia.

<sup>6</sup> Discipline of Medical Genetics, School of Biomedical Sciences and Pharmacy, Faculty of Health, University of Newcastle, Callaghan, New South Wales, Australia.

<sup>7</sup> Centre for Clinical Epidemiology and Biostatistics, School of Medicine and Public Health, University of Newcastle, Callaghan, New South Wales, Australia.

<sup>8</sup> Cancer Prevention Program, Fred Hutchinson Cancer Research Center, Seattle, WA, USA.

<sup>9</sup> Zilber School of Public Health, University of Wisconsin-Milwaukee, Milwaukee, WI, USA.

<sup>10</sup> Department of Gynecology and Obstetrics, Comprehensive Cancer Center ER-EMN, University Hospital Erlangen, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany.

- <sup>11</sup> Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA.
- <sup>12</sup> Department of Clinical Pathology, The University of Melbourne, Melbourne, Victoria, Australia.
- <sup>13</sup> Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia.
- <sup>14</sup> Genetic Medicine and Family Cancer Clinic, Royal Melbourne Hospital, Parkville, Victoria, Australia.
- <sup>15</sup> University of Melbourne Centre for Cancer Research, Victorian Comprehensive Cancer Centre, Parkville, Victoria, Australia.
- <sup>16</sup> Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD, USA.
- <sup>17</sup> Epidemiology Program: Fred Hutchinson Cancer Research Center, Seattle, WA, USA.
- <sup>18</sup> Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
- <sup>19</sup> Wellcome Trust Centre for Human Genetics and Oxford NIHR Biomedical Research Centre, University of Oxford, Oxford, UK.
- <sup>20</sup> University of New Mexico Health Sciences Center, University of New Mexico, Albuquerque, NM, USA.
- <sup>21</sup> Department of Cancer Epidemiology and Prevention Research, Alberta Health Services, Calgary, AB, Canada.
- <sup>22</sup> Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA.
- <sup>23</sup> Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.
- <sup>24</sup> Vesalius Research Center, VIB, Leuven, Belgium.
- <sup>25</sup> Laboratory for Translational Genetics, Department of Human Genetics, University of Leuven, Leuven, Belgium.
- <sup>26</sup> Huntsman Cancer Institute, Department of Population Health Sciences, University of Utah, Salt Lake City, UT, USA.
- <sup>27</sup> Gynaecology Research Unit, Hannover Medical School, Hannover, Germany.
- <sup>28</sup> Department of Obstetrics and Gynecology, Division of Gynecologic Oncology, Mayo Clinic, Rochester, MN, USA.
- <sup>29</sup> Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, UK.
- <sup>30</sup> Department of Gynaecology, Jena University Hospital-Friedrich Schiller University, Jena, Germany.
- <sup>31</sup> Institute of Human Genetics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany.
- <sup>32</sup> David Geffen School of Medicine, Department of Medicine Division of Hematology and Oncology, University of California at Los Angeles, Los Angeles, CA, USA.
- <sup>33</sup> Department of Biostatistics, Kansas University Medical Center, Kansas City, KS, USA.
- <sup>34</sup> Division of Genetics and Epidemiology, Institute of Cancer Research, London, UK.
- <sup>35</sup> Behavioral and Epidemiology Research Group, American Cancer Society, Atlanta, GA, USA.
- <sup>36</sup> Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, Victoria, Australia.
- <sup>37</sup> Department of Health Science Research, Division of Epidemiology, Mayo Clinic, Rochester, MN, USA.
- <sup>38</sup> Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.

- <sup>39</sup> Department of Oncology, Södersjukhuset, Stockholm, Sweden.
- <sup>40</sup> Department of Biostatistics & Epidemiology, University of Massachusetts, Amherst, Amherst, MA, USA.
- <sup>41</sup> Department of Clinical Genetics, St George's, University of London, London, UK.
- <sup>42</sup> Centre for Cancerbiomarkers, Department of Clinical Science, University of Bergen, Bergen, Norway.
- <sup>43</sup> Department of Gynecology and Obstetrics, Haukeland University Hospital, Bergen, Norway.
- <sup>44</sup> Program in Genetic Epidemiology and Statistical Genetics, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
- <sup>45</sup> Nuffield Department of Population Health, University of Oxford, Oxford, UK.
- <sup>46</sup> VIB Center for Cancer Biology, VIB, Leuven, Belgium.
- <sup>47</sup> Epidemiology Program: University of Hawaii Cancer Center, Honolulu, HI, USA.
- <sup>48</sup> Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY, USA.
- <sup>49</sup> Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden.
- <sup>50</sup> Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden.
- <sup>51</sup> Department of Cancer Epidemiology and Prevention, M. Sklodowska-Curie Cancer Center, Oncology Institute, Warsaw, Poland.
- <sup>52</sup> Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, TN, USA.
- <sup>53</sup> Chronic Disease Epidemiology, Yale School of Public Health, New Haven, CT, USA.
- <sup>54</sup> Department of Anatomic Pathology, Moffitt Cancer Center & Research Institute, Tampa, FL, USA.
- <sup>55</sup> Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, UK.
- <sup>56</sup> Precision Medicine: School of Clinical Sciences at Monash Health, Monash University, Clayton, Victoria, Australia.
- <sup>57</sup> Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden.
- <sup>58</sup> Department of Biochemistry and Molecular Medicine, University of California Davis, Davis, CA, USA.
- <sup>59</sup> School of Medicine and Public Health, University of Newcastle, Callaghan, New South Wales, Australia.
- <sup>60</sup> Harvard T.H. Chan School of Public Health, Boston, MA, USA.
- <sup>61</sup> Dana-Farber Cancer Institute, Boston, MA, USA.
- <sup>62</sup> Department of Gynaecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany.
- <sup>63</sup> Center for Cancer Prevention (CPO-Peimonte), Turin, Italy.
- <sup>64</sup> Human Genetics Foundation (HuGeF), Turino, Italy.
- <sup>65</sup> Department of Obstetrics and Gynecology, School of Medicine and Public Health, University of Wisconsin, Madison, WI, USA.
- <sup>66</sup> Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH, USA.
- <sup>67</sup> Division of Molecular Medicine, Pathology North, John Hunter Hospital, Newcastle, New South Wales, Australia.
- <sup>68</sup> Clinical Genetics: Karolinska Institutet, Stockholm, Sweden.

<sup>69</sup> Division of Gynecologic Oncology, Department of Obstetrics and Gynaecology and Leuven Cancer Institute, University Hospitals Leuven, Leuven, Belgium.

<sup>70</sup> Population Health Department, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia.

<sup>71</sup> Department of Health Science Research, Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN, USA.

<sup>72</sup> State Key Laboratory of Oncogene and Related Genes & Department of Epidemiology, Shanghai Cancer Institute, Renji Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China.

### Abbreviations

<b>CI</b>	confidence interval
<b>CRC</b>	colorectal cancer
<b>ER</b>	estrogen receptor
<b>GS</b>	Gilbert's syndrome
<b>GWAS</b>	genome-wide association studies
<b>IVW</b>	inverse-variance weighted
<b>MR</b>	Mendelian randomization
<b>1-SD</b>	one-standard deviation
<b>OR</b>	odds ratio
<b>SNP</b>	single-nucleotide polymorphism
<b>TA</b>	thymine–adenine
<b>UGT1A1</b>	uridine-diphosphoglucuronate glucuronosyltransferase1A1
<b>UKB</b>	UK Biobank

### References

- Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2018**, *68*, 394–424.
- Hanahan, D.; Weinberg, R.A. Hallmarks of Cancer: The Next Generation. *Cell* **2011**, *144*, 646–674, doi:10.1016/j.cell.2011.02.013.
- Waris, G.; Ahsan, H. Reactive oxygen species: Role in the development of cancer and various chronic conditions. *J. Carcinog.* **2006**, *5*, 14, doi:10.1186/1477-3163-5-14.
- Stocker, R.; Yamamoto, Y.; McDonagh, A.F.; Glazer, A.N.; Ames, B.N. Bilirubin is an antioxidant of possible physiological importance. *Sci.* **1987**, *235*, 1043–1046, doi:10.1126/science.3029864.
- Bock, K.W. Human AHR functions in vascular tissue: Pro- and anti-inflammatory responses of AHR agonists in atherosclerosis. *Biochem. Pharmacol.* **2019**, *159*, 116–120, doi:10.1016/j.bcp.2018.11.021.
- Lee, Y.; Kim, H.; Kang, S.; Lee, J.; Park, J.; Jon, S. Bilirubin Nanoparticles as a Nanomedicine for Anti-inflammation Therapy. *Angew. Chem. Int. Ed.* **2016**, *55*, 7460–7463, doi:10.1002/anie.201602525.
- Zelenka, J.; Dvořák, A.; Alán, L.; Zadinová, M.; Haluzík, M.; Vitek, L. Hyperbilirubinemia Protects against Aging-Associated Inflammation and Metabolic Deterioration. *Oxidative Med. Cell. Longev.* **2016**, *2016*, 1–10, doi:10.1155/2016/6190609.
- Deming, S.L.; Zheng, W.; Xu, W.-H.; Cai, Q.; Ruan, Z.; Xiang, Y.-B.; Shu, X.-O. UGT1A1 Genetic Polymorphisms, Endogenous Estrogen Exposure, Soy Food Intake, and Endometrial Cancer Risk. *Cancer Epidemiology Biomarkers Prev.* **2008**, *17*, 563–570, doi:10.1158/1055-9965.epi-07-0752.
- Johnson, A.D.; Kavousi, M.; Smith, A.V.; Chen, M.-H.; Dehghan, A.; Aspelund, T.; Lin, J.-P.; Van Duijn, C.M.; Harris, T.B.; Cupples, L.A.; et al. Genome-wide association meta-analysis for total serum bilirubin levels. *Hum. Mol. Genet.* **2009**, *18*, 2700–2710, doi:10.1093/hmg/ddp202.
- Stender, S.; Frikke-Schmidt, R.; Nordestgaard, B.G.; Grande, P.; Tybjaerg-Hansen, A. Genetically elevated bilirubin and risk of ischaemic heart disease: Three Mendelian randomization studies and a meta-analysis. *J. Intern. Med.* **2012**, *273*, 59–68, doi:10.1111/j.1365-2796.2012.02576.x.
- Duguay, Y.; McGrath, M.; Lépine, J.; Gagné, J.-F.; Hankinson, S.E.; Colditz, G.A.; Hunter, D.J.; Plante, M.; Têtu, B.; Bélanger, A.; et al. The Functional UGT1A1 Promoter Polymorphism Decreases Endometrial Cancer Risk. *Cancer Res.* **2004**, *64*, 1202–1207, doi:10.1158/0008-5472.can-03-3295.
- Holt, S.K.; Rossing, M.A.; Malone, K.E.; Schwartz, S.M.; Weiss, N.S.; Chen, C. Ovarian Cancer Risk and Polymorphisms Involved in Estrogen Catabolism. *Cancer Epidemiology Biomarkers Prev.* **2007**, *16*, 481–489, doi:10.1158/1055-9965.epi-06-0831.
- Nishikawa, Y.; Kanai, M.; Narahara, M.; Tamon, A.; Brown, J.B.; Taneishi, K.; Nakatsui, M.; Okamoto, K.; Uneno, Y.; Yamaguchi, D.; et al. Association between UGT1A1\*28\*28 genotype and lung cancer in the Japanese population. *Int. J. Clin. Oncol.* **2016**, *22*, 269–273, doi:10.1007/s10147-016-1061-2.
- Horsfall, L.J.; Burgess, S.; Hall, I.; Nazareth, I. Genetically raised serum bilirubin levels and lung cancer: A cohort study and Mendelian randomisation using UK Biobank. *Thorax* **2020**, *75*, 955–964, doi:10.1136/thoraxjnl-2020-214756.

15. Tsezou, A.; Tzetis, M.; Giannatou, E.; Gennatas, C.; Pampanos, A.; Kanavakis, E.; Kitsiou-Tzeli, S. Genetic Polymorphisms in the UGT1A1 Gene and Breast Cancer Risk in Greek Women. *Genet. Test.* **2007**, *11*, 303–306, doi:10.1089/gte.2007.0020.
16. Karatzas, A.; Giannatou, E.; Tzortzis, V.; Gravas, S.; Aravantinos, E.; Moutzouris, G.; Melekos, M.; Tsezou, A. Genetic polymorphisms in the UDP-glucuronosyltransferase 1A1 (UGT1A1) gene and prostate cancer risk in Caucasian men. *Cancer Epidemiology* **2010**, *34*, 345–349, doi:10.1016/j.canep.2010.02.009.
17. Davey Smith, G., & Hemani, G. Mendelian randomization: Genetic anchors for causal inference in epidemiological studies. *Human molecular genetics* **2014** 23(R1), R89–R98.
18. Burgess, S.; Butterworth, A.S.; Thompson, J.R. Beyond Mendelian randomization: How to interpret evidence of shared genetic predictors. *J. Clin. Epidemiology* **2016**, *69*, 208–216, doi:10.1016/j.jclinepi.2015.08.001.
19. Sinnott-Armstrong, N.; Gen, F.; Tanigawa, Y.; Amar, D.; Mars, N.; Benner, C.; Aguirre, M.; Venkataraman, G.R.; Wainberg, M.; Ollila, H.M.; et al. Genetics of 35 blood and urine biomarkers in the UK Biobank. *Nat. Genet.* **2021**, *53*, 185–194, doi:10.1038/s41588-020-00757-z.
20. Collins, R. What makes UK Biobank special?. *Lancet* **2012**, *379*, 1173–1174, doi:10.1016/s0140-6736(12)60404-8.
21. Bowden, J.; Smith, G.D.; Haycock, P.C.; Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet. Epidemiology* **2016**, *40*, 304–314, doi:10.1002/gepi.21965.
22. National Cancer Institute. <https://ldlinknc.nih.gov/?tab=ldpair>. (February 2020).
23. Burgess, S.; Thompson, S.G.; CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. *Int. J. Epidemiology* **2011**, *40*, 755–764, doi:10.1093/ije/dyr036.
24. Petersen, G.M.; Amundadottir, L.; Fuchs, C.S.; Kraft, P.; Stolzenberg-Solomon, R.Z.; Jacobs, K.B.; Arslan, A.A.; Bueno-De-Mesquita, H.B.; Gallinger, S.; Gross, M.D.; et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. *Nat. Genet.* **2010**, *42*, 224–228, doi:10.1038/ng.522.
25. Childs, E.J.; Mocci, E.; Campa, D.; Bracci, P.M.; Gallinger, S.; Goggins, M.; Li, D.; Neale, R.E.; Olson, S.H.; Scelo, G.; et al. Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. *Nat. Genet.* **2015**, *47*, 911–916, doi:10.1038/ng.3341.
26. Amundadottir, L.; Kraft, P.; Stolzenberg-Solomon, R.Z.; Fuchs, C.S.; Petersen, G.M.; Arslan, A.A.; Bueno-De-Mesquita, H.B.; Gross, M.D.; Helzlsouer, K.J.; Jacobs, E.J.; et al. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. *Nat. Genet.* **2009**, *41*, 986–990, doi:10.1038/ng.429.
27. Scelo, G.; Purdue, M.P.; Brown, K.M.; Johansson, M.; Wang, Z.; Eckel-Passow, J.E.; Ye, Y.; Hofmann, J.N.; Choi, J.; Foll, M.; et al. Genome-wide association study identifies multiple risk loci for renal cell carcinoma. *Nat. Commun.* **2017**, *8*, 15724, doi:10.1038/ncomms15724.
28. McKay, J.D.; SpiroMeta Consortium; Hung, R.J.; Han, Y.; Zong, X.; Carreras-Torres, R.; Christiani, D.C.; Caporaso, N.E.; Johansson, M.; Xiao, X.; et al. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nat. Genet.* **2017**, *49*, 1126–1132, doi:10.1038/ng.3892.
29. Phelan, C.M.; AOCs Study Group; Kuchenbaecker, K.B.; Tyrer, J.P.; Kar, S.P.; Lawrenson, K.; Winham, S.J.; Dennis, J.; Pirie, A.; Riggan, M.J.; et al. Identification of 12 new susceptibility loci for different histotypes of epithelial ovarian cancer. *Nat. Genet.* **2017**, *49*, 680–691, doi:10.1038/ng.3826.
30. O'Mara, T.A.; Glubb, D.M.; Amant, F.; Annibaldi, D.; Ashton, K.; Attia, J.; Auer, P.L.; Beckmann, M.W.; Black, A.; Bolla, M.K.; et al. Identification of nine new susceptibility loci for endometrial cancer. *Nat. Commun.* **2018**, *9*, 1–12, doi:10.1038/s41467-018-05427-7.
31. Michailidou, K.; Collaborators, N.; Lindström, S.; Dennis, J.; Beesley, J.; Hui, S.; Kar, S.; Lemaçon, A.; Soucy, P.; Glubb, D.; et al. Association analysis identifies 65 new breast cancer risk loci. *Nature* **2017**, *551*, 92–94, doi:10.1038/nature24284.
32. Schumacher, F.R.; Al Olama, A.A.; Berndt, S.I.; Benlloch, S.; Ahmed, M.; Saunders, E.J.; Dadaev, T.; Leongamornlert, D.; Anokian, E.; Cieza-Borrella, C.; et al. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat. Genet.* **2018**, *50*, 928–936, doi:10.1038/s41588-018-0142-8.
33. Urayama, K.Y.; Jarrett, R.F.; Hjalgrim, H.; Diepstra, A.; Kamatani, Y.; Chabrier, A.; Gaborieau, V.; Boland, A.; Nieters, A.; Becker, N.; et al. Genome-Wide Association Study of Classical Hodgkin Lymphoma and Epstein–Barr Virus Status–Defined Subgroups. *J. Natl. Cancer Inst.* **2012**, *104*, 240–253, doi:10.1093/jnci/djr516.
34. Amos, C.I.; Wang, L.-E.; Lee, J.E.; Gershenwald, J.E.; Chen, W.V.; Fang, S.; Kosoy, R.; Zhang, M.; Qureshi, A.A.; Vattathil, S.; et al. Genome-wide association study identifies novel loci predisposing to cutaneous melanoma. *Hum. Mol. Genet.* **2011**, *20*, 5012–5023, doi:10.1093/hmg/ddr415.
35. Capasso, M.; Diskin, S.J.; Totaro, F.; Longo, L.; De Mariano, M.; Russo, R.; Cimmino, F.; Hakonarson, H.; Tonini, G.P.; Devoto, M.; et al. Replication of GWAS-identified neuroblastoma risk loci strengthens the role of BARD1 and affirms the cumulative effect of genetic variations on disease susceptibility. *Carcinog.* **2013**, *34*, 605–611, doi:10.1093/carcin/bgs380.
36. Cozen, W.; Timofeeva, M.N.; Li, D.; Diepstra, A.; Hazelett, D.J.; Delahayesourdeix, M.; Edlund, C.K.; Franke, L.; Rostgaard, K.; Berg, D.J.V.D.; et al. A meta-analysis of Hodgkin lymphoma reveals 19p13.3 TCF3 as a novel susceptibility locus. *Nat. Commun.* **2014**, *5*, 1–10, doi:10.1038/ncomms4856.
37. Genotypes and Phenotypes database <https://www.ncbi.nlm.nih.gov/gap>. (February 2020).
38. Online platform for MR analyses. <http://www.mrbase.org>. (February 2020).
39. Burgess, S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. *Int. J. Epidemiology* **2014**, *43*, 922–929, doi:10.1093/ije/dyu005.

40. Burgess, S.; Dudbridge, F.; Thompson, S.G. Combining information on multiple instrumental variables in Mendelian randomization: Comparison of allele score and summarized data methods. *Stat. Med.* **2016**, *35*, 1880–1906, doi:10.1002/sim.6835.
41. Burgess, S.; Butterworth, A.; Thompson, S.G. Mendelian Randomization Analysis With Multiple Genetic Variants Using Summarized Data. *Genet. Epidemiology* **2013**, *37*, 658–665, doi:10.1002/gepi.21758.
42. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nature genetics*. 2018;50(5):693-8.
43. Bowden, J.K.; Smith, G.D.; Burgess, S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *Int. J. Epidemiology* **2015**, *44*, 512–525, doi:10.1093/ije/dyv080.
44. Hartwig, F.P.; Smith, G.D.; Bowden, J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int. J. Epidemiology* **2017**, *46*, 1985–1998, doi:10.1093/ije/dyx102.
45. Magno, R.; Maia, A.-T. gwasrapidd: An R package to query, download and wrangle GWAS Catalog data. *Bioinform.* **2019**, *36*, 649–650, doi:10.1093/bioinformatics/btz605.
46. Zheng, J.; Baird, D.; Borges, M.-C.; Bowden, J.; Hemani, G.; Haycock, P.; Evans, D.M.; Smith, G.D. Recent Developments in Mendelian Randomization Studies. *Curr. Epidemiology Rep.* **2017**, *4*, 330–345, doi:10.1007/s40471-017-0128-6.
47. Lim, J.-E.; Kimm, H.; Jee, S.H. Combined Effects of Smoking and Bilirubin Levels on the Risk of Lung Cancer in Korea: The Severance Cohort Study. *PLoS ONE* **2014**, *9*, e103972, doi:10.1371/journal.pone.0103972.
48. Wen, C.-P.; Zhang, F.; Liang, D.; Wen, C.; Gu, J.; Skinner, H.D.; Chow, W.-H.; Ye, Y.; Pu, X.; Hildebrandt, M.A.; et al. The Ability of Bilirubin in Identifying Smokers with Higher Risk of Lung Cancer: A Large Cohort Study in Conjunction with Global Metabolomic Profiling. *Clin. Cancer Res.* **2015**, *21*, 193–200, doi:10.1158/1078-0432.ccr-14-0748.
49. Freisling, H.; Khoei, N.S.; Viallon, V.; Wagner, K.-H. Gilbert’s syndrome, circulating bilirubin and lung cancer: A genetic advantage? *Thorax* **2020**, *75*, 916–917, doi:10.1136/thoraxjnl-2020-215642.
50. Carbone, A.; Gloghini, A. Epstein Barr Virus-Associated Hodgkin Lymphoma. *Cancers* **2018**, *10*, 163, doi:10.3390/cancers10060163.
51. Zhu Z, Wilson AT, Luxon BA, Brown KE, Mathahs MM, Bandyopadhyay S.; et al. Biliverdin inhibits hepatitis C virus nonstructural 3/4A protease activity: Mechanism for the antiviral effects of heme oxygenase? *Hepatology* (Baltimore, Md). 2010;52(6):1897-905.
52. Hu, J.; Li, H.; Luo, X.; Li, Y.; Bode, A.; Cao, Y. The role of oxidative stress in EBV lytic reactivation, radioresistance and the potential preventive and therapeutic implications. *Int. J. Cancer* **2017**, *141*, 1722–1729, doi:10.1002/ijc.30816.
53. Yao, L.; Qiu, L.-X.; Yu, L.; Yang, Z.; Yu, X.-J.; Zhong, Y.; Yu, L. The association between TA-repeat polymorphism in the promoter region of UGT1A1 and breast cancer risk: A meta-analysis. *Breast Cancer Res. Treat.* **2010**, *122*, 879–882, doi:10.1007/s10549-010-0742-1.
54. Zhou, J.; Tracy, T.S.; Rimmel, R.P. Correlation between Bilirubin Glucuronidation and Estradiol-3-Glucuronidation in the Presence of Model UDP-Glucuronosyltransferase 1A1 Substrates/Inhibitors. *Drug Metab. Dispos.* **2010**, *39*, 322–329, doi:10.1124/dmd.110.035030.
55. Saed, G.M.; Fletcher, N.M.; Jiang, Z.L.; Abu-Soud, H.M.; Diamond, M.P. Dichloroacetate Induces Apoptosis of Epithelial Ovarian Cancer Cells Through a Mechanism Involving Modulation of Oxidative Stress. *Reprod. Sci.* **2011**, *18*, 1253–1261, doi:10.1177/1933719111411731.
56. Seyed Khoei N, Jenab M, Murphy N, Banbury BL, Carreras-Torres R., Viallon V.; et al. Circulating bilirubin levels and risk of colorectal cancer: Serological and Mendelian randomization analyses. *BMC medicine*. 2020;18(1):229.
57. Haycock, P.C.; Burgess, S.; Wade, K.H.; Bowden, J.; Relton, C.L.; Smith, G.D. Best (but oft-forgotten) practices: The design, analysis, and interpretation of Mendelian randomization studies. *Am. J. Clin. Nutr.* **2016**, *103*, 965–978, doi:10.3945/ajcn.115.118216.
58. Wells PG, Mackenzie PI, Chowdhury JR, Guillemette C, Gregory PA, Ishii Y; et al. Glucuronidation and the UDP-glucuronosyltransferases in health and disease. *Drug metabolism and disposition: The biological fate of chemicals*. 2004;32(3):281-90.
59. Liu, W.; Innocenti, F.; Ratain, M.J. Linkage disequilibrium across the UGT1A locus should not be ignored in association studies of cancer susceptibility. *Clin. Cancer Res.* **2005**, *11*, 1348.
60. Schooling CM. Biases in GWAS—the dog that did not bark. *bioRxiv*. 2019:709063. (July 20, 2019).