Original research

Protein-truncating and rare missense variants in *ATM* and *CHEK2* and associations with cancer in UK Biobank whole-exome sequence data

Toqir K Mukhtar [1,2] Naomi Wilcox,[1] Joe Dennis,[1] Xin Yang [1] Marc Naven,[1] Nasim Mavaddat,[1] John R B Perry,[3] Eugene Gardner,[4] Douglas F Easton[1,5]

ABSTRACT

Background Deleterious germline variants in *ATM* and *CHEK2* have been associated with a moderately increased risk of breast cancer. Risks for other cancers remain unclear.

Methods Cancer associations for coding variants in *ATM* and *CHEK2* were evaluated using whole-exome sequence data from UK Biobank linked to cancer registration data (348 488 participants), and analysed both as a retrospective case-control and a prospective cohort study. Odds ratios, hazard ratios, and combined relative risks (RRs) were estimated by cancer type and gene. Separate analyses were performed for protein-truncating variants (PTVs) and rare missense variants (rMSVs; allele frequency <0.1%).

Results PTVs in *ATM* were associated with increased risks of nine cancers at p<0.001 (pancreas, oesophagus, lung, melanoma, breast, ovary, prostate, bladder, lymphoid leukaemia (LL)), and three at p<0.05 (colon, diffuse non-Hodgkin's lymphoma (DNHL), rectosigmoid junction). Carriers of rMSVs had increased risks of four cancers (p<0.05: stomach, pancreas, prostate, Hodgkin's disease (HD)). RRs were highest for breast, prostate, and any cancer where rMSVs lay in the FAT or PIK domains, and had a Combined Annotation Dependent Depletion score in the highest quintile.

PTVs in *CHEK2* were associated with three cancers at p<0.001 (breast, prostate, HD) and six at p<0.05 (oesophagus, melanoma, ovary, kidney, DNHL, myeloid leukaemia). Carriers of rMSVs had increased risks of five cancers (p<0.001: breast, prostate, LL; p<0.05: melanoma, multiple myeloma).

Conclusion PTVs in *ATM* and *CHEK2* are associated with a wide range of cancers, with the highest RR for pancreatic cancer in *ATM* PTV carriers. These findings can inform genetic counselling of carriers.

INTRODUCTION

Gene panel testing for cancer susceptibility is now an important part of clinical practice. To provide reliable genetic counselling, it is important to be able to provide risks for all cancer types. Many susceptibility genes, notably *BRCA1*, *BRCA2*, and the mismatch repair genes, are associated with multiple types of cancer,[1 2] but in general, cancer pleiotropy is poorly understood. Here we consider rare variants in two genes, Checkpoint Kinase 2 (*CHEK2*) and Ataxia-Telangiectasia Mutated (*ATM*), that are included on many cancer gene panels. Variants in these genes are known to be associated with moderately increased risks of breast cancer, but the risks of many other cancer types are unclear.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ While previous research shows there is evidence for association between variants in *ATM* or *CHEK2* and multiple cancer types in individual smaller studies, the associations have not been consistently evaluated across all cancer types and, with the exception of breast cancer, the strengths of association are unclear.

WHAT THIS STUDY ADDS

⇒ We examined data from a large cohort study to derive relative and absolute risks for all cancer types for carriers of protein-truncating variants (PTVs) and rare missense variants (rMSVs) in *ATM* and *CHEK2*.
⇒ *ATM* PTVs were associated with significantly increased risk for 12 of 24 sites examined (nine at p<0.001), with the relative risk (RR) being highest for pancreatic cancer (approximately sevenfold). Carriers of rMSVs had increased risks of four cancers, with a RR of approximately 1.5.
⇒ For *CHEK2* PTVs, statistically significant risks were observed for nine of the 24 sites examined (three at p<0.001). Carriers of rMSVs had increased risks of five cancers with the risk being highest for lymphoid leukaemia (approximately twofold).

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ *ATM* and *CHEK2* are included on many cancer gene panels used in family cancer clinics, and the risk estimates from these analyses can inform genetic counselling for carriers.
⇒ The estimated absolute risks for pancreatic cancer in *ATM* PTV carriers (11% in males and 8% in females by age 85) are notably higher than for other major pancreatic susceptibility genes including *BRCA2*, *CDK2NA,* and *PALB2*. Our findings can also inform National Institute for Health and Care Excellence guidelines for pancreatic cancer, which do not currently include *ATM*.

CHEK2 encodes a protein kinase, activated in response to DNA damage,[3] and with a role in cell-cycle arrest, DNA repair, and cell death.[4 5] The most common protein-truncating variant (PTV) in Western European populations is the *CHEK2*1100delC*. Previous research has shown an association between the 1100delC deletion and risk of breast cancer in women, with the estimated risk being approximately twofold.[6 7] A similar association has been seen for other PTVs in aggregate.[6 8] Rare *CHEK2* missense variants (MSVs), in aggregate, have also been associated with breast cancer risk, with a relative risk (RR) of ~1.4.[6] There is evidence that the breast cancer risk differs by variant, using in silico and functional classifications of deleteriousness.[9]

Previous studies have found some evidence of an association between PTVs and cancer other than breast cancer, with most evidence being based on the 1100delC variant: these include increased risks of prostate cancer,[10] colorectal cancer,[11 12] malignant melanoma, gastric cancer, and bladder cancer.[13] Previous studies have reported evidence that carriers of the I157T variant have increased risk of prostate,[10] colorectal,[14] ovarian,[15] pancreatic,[16] gastric,[17] thyroid,[18] kidney,[19] and colon cancer.[19] There is some evidence of a decreased risk of lung cancer.[20] *CHEK2* PTVs have also been shown to influence ovarian ageing in women through inhibited DNA damage sensing.[21]

ATM[22] encodes a protein kinase involved in DNA repair, apoptosis, and regulation of cell-cycle checkpoints.[22 23] It is recruited and activated by double-strand DNA breaks. *ATM* activates *CHEK2* through phosphorylation. Deleterious mutations in *ATM* cause the rare autosomal recessive disorder Ataxia-Telangiectasia (A-T). Patients with A-T have a large excess risk of cancers in childhood, in particular, lymphomas and leukaemias, but are also at increased risk of adult cancers, including breast cancer. There is evidence both from studies of families of A-T patients,[24] and from large case-control studies, of an association between pathogenic variants in *ATM* and increased risk of breast cancer,[6] with an estimated RR of approximately twofold. Previous studies indicate subsets of rare missense variants (rMSVs) from the PIK, FAT, and FATC domains are also associated with an excess risk of breast cancer.[25]

For cancers other than breast, there is evidence for an association between mutations in *ATM* and cancer of the prostate[26] and pancreas.[27] Some studies have found evidence for a higher risk of colorectal cancer,[24] melanoma,[28] gastric,[29] ovarian, and lung cancer.[30]

While there is evidence for association between variants in *ATM* or *CHEK2* and multiple cancer types, the associations have not been consistently evaluated across all cancer types and, with the exception of breast cancer, the strengths of association are unclear. In this paper, we examine data from a large cohort study to derive relative and absolute risks for all cancer types, for carriers of PTVs and rMSVs in *CHEK2* and *ATM*. We also examine the evidence for variation in risk for rMSVs according to previously defined in silico classifications.

MATERIALS AND METHODS
Dataset
UK Biobank is a population-based prospective study of approximately 500 000 participants in the UK.[31] Participants were enrolled between 2006 and 2010 and were aged between 40 and 69 on recruitment. Whole-exome sequencing (WES) data were available at the time of analysis for 454 756 UK Biobank participants.[32] Details of the WES variant calling, filtering, and classification are given in the online supplemental appendix. In *ATM*, rMSVs were categorised by functional protein domain[33]
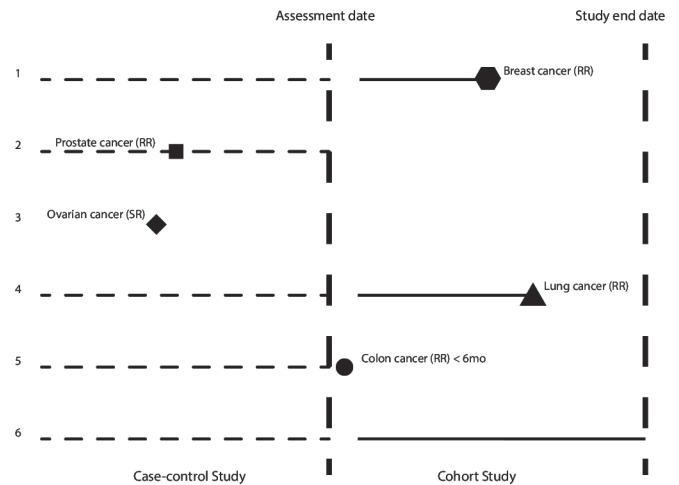

Figure 1 Case-control and prospective cohort studies generated from UK Biobank data. Start of observation for the prospective study 6 months after the assessment date. Person 1: registry-reported breast cancer 2 years after the day of first assessment, and no prior cancers; contributes to both studies. Person 2: registry-reported prostate cancer prior to the first day of assessment; contributes to case-control study only. Person 3: self-reported ovarian cancer prior to the first day of assessment, and no registry-reported cancers; excluded from both studies. Person 4: registry-reported lung cancer 3 years after the day of first assessment; contributes to both studies. Person 5: registry-reported colon cancer 4 months after the start of observation; contributes to case-control study only. Person 6: no self-reported nor registry-reported cancers; contributes to both studies. RR: registry-reported cancer. SR: self-reported cancer.

and pathogenicity score predicted by the Combined Annotation Dependent Depletion (CADD; version 1.6) algorithm.[34] In *CHEK2*, rMSVs were classified as high or low risk using the deleteriousness score given by the Helix algorithm (version 4.4.1).[35] These classifications were used as they provided the best discrimination for breast cancer risk in the analysis of the large BRIDGES dataset.

Cancers occurring in study participants were identified through linkage to the national cancer registry data for England, Wales, and Scotland.[36] Invasive cancers (codes C00-C97) and breast carcinoma in situ (CIS; code D05), defined using the 10th Revision of the International Classification of Diseases,[37] were included as endpoints. Self-reported cancers were not included.

Study design
UK Biobank includes both a retrospective (based on data at study entry) and a prospective component (figure 1). The retrospective analyses were based on cases diagnosed with cancer before entry to the study. Controls were individuals who had neither a self-reported nor registry-reported cancer before entry, other than non-melanoma skin cancer. A total of 348 488 males and females were included, of whom 18 838 were diagnosed with at least one cancer before the day of first assessment.

The prospective analysis included individuals without a self-reported cancer or registry-reported cancer before the start of follow-up, or a registry-reported cancer within 6 months of follow-up (except for non-melanoma skin cancer). Thus, follow-up started 6 months from the day of first assessment for the study, to avoid inadvertent inclusion of cases where onset was prior to entry. The event indicator was first diagnosis of any cancer by type, and follow-up was censored at the earliest of first diagnosed cancer of any type, date of death, or a last follow-up

date dependent on the location of the assessment centre attended by participants (for England, this was 31 December 2020; Scotland, 30 November 2021; and Wales, 31 December 2016). A total of 328 919 individuals were included, with 37 802 cancers diagnosed over 3 484 613 person-years of follow-up.

Statistical analyses

Since most variants were too rare to estimate the associated risk, we conducted burden analyses which evaluated the risks associated with carrying any one of a set of variants: PTVs or rMSVs. rMSVs were further subdivided by functional protein domain and prediction scores. We present results for each variant type and gene for 23 invasive cancer types with at least 400 occurrences in the dataset, plus Hodgkin's disease (HD) and CIS. We also conducted exploratory analyses for 65 rarer cancer types.

In the retrospective dataset, logistic regression was used to estimate the OR for each cancer type associated with carrying either a PTV or rMSV. For the prospective analyses, Cox proportional hazards regression models were used to estimate the HR for any cancer, and for each cancer type, associated with carrying either a PTV or rMSV. All models included sex and 10 ancestry informative principal components[35] as covariates. Since the retrospective and prospective components are essentially independent, we also computed overall RR estimates combining the retrospective and prospective estimates (see online supplemental appendix). There were too few cases of male breast cancer to analyse separately; the estimates based on female alone and males and females combined were identical. We investigated whether the HRs varied by age by fitting HRs by age group (<50, 50–59, 60–69, 70–79, and ≥80 years) and by fitting models in which the log(HR) varied linearly with age, using the tt() function in R.

Cumulative risks for cancer were estimated by combining HR estimates from the prospective analyses with population cancer incidences in the UK, 2016–2018,[38] using the same approach as that used by Schmidt et al.[39] Where there was evidence of a trend in the HR by age, time-dependent HR risk estimates were used, otherwise a constant HR was assumed.

Binary logistic regression and Cox proportional hazards regression analyses were conducted using R version 4.1.0. Meta-analyses were also conducted in R using the meta package version 6.5-0, and the metagen function. All p values were two-sided, with p<0.05 considered significant.

RESULTS

PTVs in ATM

In the pooled analysis (online supplemental table 1), PTVs in *ATM* were associated with increased risks of 12 of 24 cancers: oesophagus (RR: 3.90 (95% Confidence Interval (CI) 1.85 to 8.20, p=3.43×10^{-4})); colon (2.18 (1.37 to 3.48, p=1.04×10^{-3})); rectosigmoid junction (3.43 (1.10 to 10.66, p=0.033)); pancreas (7.35 (4.27 to 12.66, p=6.53×10^{-13})); lung (2.64 (1.61 to 4.32, p=1.18×10^{-4})); melanoma (2.28 (1.44 to 3.59, p=4.18×10^{-4})); breast (2.27 (1.77 to 2.91, p=1.18×10^{-10})); ovary (3.20 (1.66 to 6.16, p=5.12×10^{-4})); prostate (2.35 (1.78 to 3.11, p=2.41×10^{-9})); bladder (3.74 (1.94 to 7.21, p=8.30×10^{-5})); diffuse non-Hodgkin's lymphoma (DNHL) (3.18 (1.51 to 6.69, p=2.28×10^{-3})); and lymphoid leukaemia (LL) (4.42 (2.11 to 9.29, p=8.67×10^{-5})). PTVs were also associated with CIS (3.22 (1.96 to 5.28, p=4.05×10^{-6})). Among rarer cancer types, we observed some evidence for increased risk for cancers of the small intestine (3 cases, RR: 6.12 (1.97 to 18.98, p=0.0017)), gallbladder (2 cases, RR: 10.00 (2.50 to 39.99,

p=8.84×10^{-5})), and other and unspecified parts of the biliary tract (6 cases, RR: 20.77 (9.11 to 47.34, p=5.31×10^{-13})).

Pooled analyses also showed increased risk for any cancer in females (2.23 (1.88 to 2.64, p=1.86×10^{-20})) and males (2.33 (1.96 to 2.77, p=3.81×10^{-22})). These RRs remained elevated (though slightly reduced in females), once breast and prostate cancer were excluded (2.06 (1.63 to 2.59, p=1.03×10^{-9}) and 2.30 (1.83 to 2.88, p=9.90×10^{-13})), respectively).

The OR and HR estimates from the retrospective and prospective analyses were similar for each of the associated cancers, and the differences were not statistically significant for any cancer type. The estimates for pancreatic cancer were higher in the retrospective analysis (20.64 (4.97 to 85.76, p=3.10×10^{-5})) than in the prospective analysis (6.15 (3.39 to 11.14, p=2.18×10^{-9})); however, carrier counts were low, and CIs correspondingly wide.

There was some evidence of a decline in the HR with age for breast cancer (RR per year 0.95 (0.92 to 0.99, p=0.012)) and any cancer in females (0.97 (0.95 to 0.99, p=0.013)), with the highest HR for breast cancer under age 50 years (3.92 (2.16 to 7.10, p=6.64×10^{-6})). There was no evidence for a trend in the HR by age for prostate nor pancreatic cancer (online supplemental tables 5 and 9a).

rMSVs in ATM

Pooled analyses (online supplemental table 2) showed that rMSV carriers had higher risks of cancer of the stomach (1.54 (95% CI 1.07 to 2.23, p=0.022)), pancreas (1.60 (1.20 to 2.13, p=1.27×10^{-3})), prostate (1.15 (1.04 to 1.27, p=7.21×10^{-3})), and HD (1.72 (1.02 to 2.90, p=0.043)). There was some increased risk for any cancer in males (1.13 (1.06 to 1.21, p=2.60×10^{-4})); removing prostate cancer made little difference to the estimate.

There was no evidence for a trend in the HR by age for breast cancer, prostate cancer, pancreatic cancer, or any cancer in females or males (online supplemental table 6).

rMSV subset by functional protein domain and CADD score

Pooled analyses (online supplemental table 11a–f) showed that RRs were highest for breast cancer (1.52 (95% CI 1.14 to 2.04, p=4.85×10^{-3})) and prostate cancer (1.67 (1.23 to 2.29, p=1.20×10^{-3})) where variants lay inside the FAT or PIK domains, and had a CADD score within quintile 5 (>3.74) (compared with RRs for other variants: p-diff <0.0001 for both breast and prostate cancer). Similarly, the RRs for all cancers combined for variants in this category were higher than for all other variants combined, both in females (1.40 (1.15 to 1.71, p=8.74×10^{-4}); p-diff <0.0001) and males (1.34 (1.09 to 1.65, p=6.57×10^{-3}); p-diff <0.0001).

For pancreatic cancer, the RR was highest where variants lay outside the FAT and PIK domains (1.56 (1.11 to 2.20, p=0.011)), but not significantly different from variants inside these domains (p-diff=0.65).

PTVs in CHEK2

In the pooled analyses (online supplemental table 3), PTVs were associated with an increased risk of nine cancer types: oesophagus (RR: 2.13 (95% CI 1.10 to 4.10, p=0.025)); melanoma (1.45 (1.00 to 2.09, p=0.049)); breast (2.44 (2.08 to 2.86, p=3.61×10^{-28})); ovary (2.33 (1.37 to 3.97, p=1.73×10^{-3})); prostate (1.92 (1.59 to 2.32, p=1.18×10^{-11})); kidney (1.83 (1.06 to 3.18, p=0.032)); HD (4.26 (1.89 to 9.64, p=4.93×10^{-4})); DNHL (1.97 (1.09 to 3.57, p=0.025)); and myeloid leukaemia (ML) (2.81 (1.33 to 5.94, p=6.75×10^{-3})). PTVs were also

associated with CIS (2.05 (1.35 to 3.13, p=8.48×10^{-4})). The OR and HR estimates from the retrospective and prospective studies were broadly similar for all the associated cancers and did not differ significantly for any cancer type.

The risk of all cancers combined was elevated in both females (1.81 (1.58 to 2.06, p=1.07×10^{-18})) and males (1.57 (1.37 to 1.79, p=4.21×10^{-11})). After excluding prostate cancer, the risk for males was reduced but still elevated (1.35 (1.12 to 1.62, p=1.99×10^{-3})), while the excess in females was no longer significant after excluding breast cancer (1.18 (0.96 to 1.46, p=0.124) and 1.28 (1.11 to 1.46, p=6.40×10^{-4}) for both sexes combined).

We examined the variation in the HR by age for breast, prostate, pancreatic, and any cancer. There was evidence of association in all age groups for breast and prostate cancer, and no evidence of differences in the HR by age for any cancer (online supplemental tables 7 and 10a).

rMSVs in *CHEK2*

In the pooled analyses (online supplemental table 4), rMSVs were associated with a higher risk of five cancer types: melanoma (1.32 (95% CI 1.01 to 1.71, p=0.041)); breast cancer (1.50 (1.31 to 1.72, p=1.02×10^{-8})); prostate cancer (1.41 (1.21 to 1.64, p=1.43×10^{-5})); multiple myeloma (1.86 (1.13 to 3.05, p=0.014)); and LL (2.23 (1.42 to 3.52, p=5.22×10^{-4})). rMSVs also were associated with CIS (1.55 (1.13 to 2.14, p=6.82×10^{-3})).

Combined estimates also showed increased risk of any cancer in females (1.26 (1.15 to 1.39, p=1.09×10^{-6})), but the risk was no longer significant after excluding breast cancer (1.05 (0.91 to 1.21, p=0.501)). Risk of any cancer was also elevated for males (1.27 (1.14 to 1.40, p=4.80×10^{-6})) and after exclusion of prostate cancer (1.17 (1.02 to 1.34, p=0.022)).

Statistically significant risk estimates from the retrospective and prospective analyses were broadly similar, and the differences were not statistically significant for any cancer type. There was no evidence of a trend in the HR by age for breast cancer, prostate cancer, pancreatic cancer, or any cancer in females or males (online supplemental table 8).

rMSV Helix score

rMSVs were categorised into high (>0.5) or low (≤0.5) Helix scores (online supplemental table 12a–f). In the pooled analyses, the RR for any cancer in females was higher in the Helix-high group (1.43 (95% CI 1.23 to 1.65, p=1.60×10^{-6})) than in the Helix-low group (1.20 (1.06 to 1.36, p=3.72×10^{-3}); p-diff <0.0001). A higher RR for breast cancer was also found in the Helix-high group (1.62 (1.31 to 2.01, p=9.53×10^{-6})) versus 1.42 (1.20 to 1.69, p=6.37×10^{-5}); p-diff=0.031) than in the Helix-low group. A similar pattern was observed for any cancer in males (Helix-high: 1.35 (1.15 to 1.60, p=3.61×10^{-4}) and Helix-low: 1.22 (1.06 to 1.39, p=4.31×10^{-3}); p-diff=4.34×10^{-4}), but not for prostate cancer.

Absolute risks

Online supplemental figure 1 shows the cumulative risks of female breast cancer by gene and variant type. By age 80, the cumulative risk of breast cancer was 30% (95% CI 16% to 41%) for *ATM* PTV carriers and 25% (19% to 30%) for *CHEK2* PTV carriers, compared with 12% in the general population. The corresponding average cumulative risks for rMSV carriers were
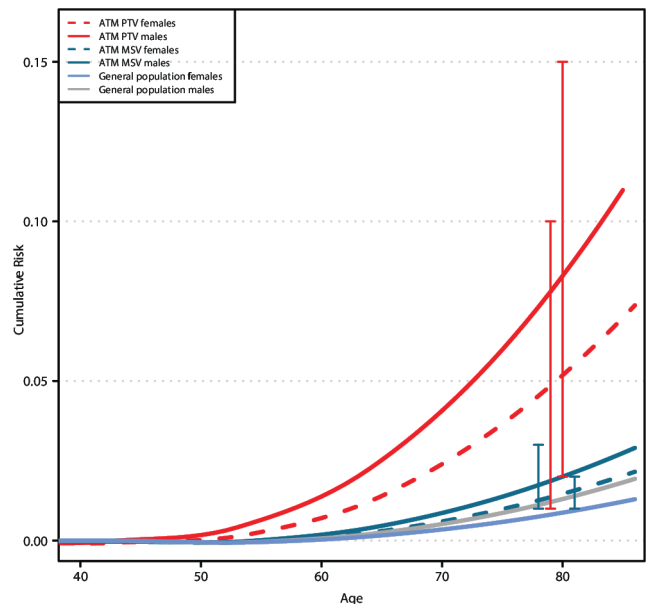

Figure 2 Cumulative absolute pancreatic cancer risks for *ATM* protein-truncating and missense variant carriers, and in the general population, by age and sex. ATM, Ataxia-Telangiectasia Mutated; MSV, missense variant; PTV, protein-truncating variant.

18% (15% to 21%) for *CHEK2* and 13% (12% to 15%) for *ATM*.

For prostate cancer (online supplemental figure 2), the cumulative risks by age 80 were 31% (23% to 38%) for *ATM* PTV carriers and 25% (20% to 29%) for *CHEK2* PTV carriers. The corresponding average cumulative risk for rMSV carriers was 19% (16% to 21%) for *CHEK2* and 15% (13% to 16%) for *ATM*.

The risks of pancreatic cancer for *ATM* mutation carriers by sex and variant type are shown in figure 2. By age 80, the estimated cumulative risk of pancreatic cancer in *ATM* PTV carriers was 8% (2% to 15%) in males and 6% (1% to 10%) in females. The cumulative risks for rMSV carriers were 2% (1% to 3%) in males and 2% (1% to 2%) in females.

The cumulative risk of developing any cancer by the age of 80 for *ATM* PTV carriers was 64% (95% CI 46% to 76%) for females and 69% (61% to 76%) for males, compared with 33% for females and 39% for males in the general population. The corresponding cumulative risks for *CHEK2* PTV carriers were 49% (43% to 54%) for females and 52% (47% to 58%) for males (online supplemental figures 3 and 4, respectively).

DISCUSSION

The pooled analyses show clear evidence of association between carrying a PTV in either gene and a higher risk of multiple cancers. For *ATM*, the RR for breast cancer (2.27 in the combined analysis) was consistent with previous case-control and family-based analyses.[6] The RR for CIS (3.22 (95% CI 1.96 to 5.28)) was consistent with that for invasive breast cancer, and a similar RR was found for prostate cancer (2.35). The excess risk of other (non-breast and prostate) cancer was more marked than for *CHEK2* PTVs, with a pooled RR of approximately twofold in both males and females. *ATM* PTVs were associated with significantly increased risk for 12 of 24 sites examined (9 at p<0.001), and (as for *CHEK2*) no sites had a significantly reduced risk. A previously reported association for pancreatic cancer was confirmed,[30] with the RR for pancreatic cancer being the highest

for any cancer type. There was some similarity with the spectrum of cancers associated with *CHEK2* PTVs (eg, oesophageal cancer, DNHL, and melanoma being common), presumably reflecting their related roles in DNA damage response; notable differences are of the association with colon and lung cancer in *ATM* PTV carriers not seen in *CHEK2*.

For *ATM* rMSVs, there was only a small excess cancer risk overall (RR 1.06 in females and 1.13 in males). Previous research has shown that breast cancer risk appears to be largely restricted to a subset of evolutionary conserved MSVs in the FAT and PIK domains.[9] When analyses were restricted to these variants, risks for breast cancer, prostate cancer, any cancer in females, and any cancer in males were higher for variants which lay within the FAT or PIK domains, with a high CADD score, and largely restricted to these variants. For pancreatic cancer, there was no evidence of a difference in RR by in silico risk category, but the confidence limits were wide.

For *CHEK2* PTVs, the estimated RR for breast cancer (approximately twofold) is similar to that previously reported for recent large case-control studies.[6 8] The RR for prostate cancer (1.92) is somewhat lower than some previous estimates,[19 40] but here we provide much stronger evidence for the association with prostate cancer risk, and more precise estimates. The association with CIS, with a RR similar to that for invasive disease, is also consistent with that reported previously for c.1100delC.[39]

There was also evidence for an excess risk of other cancers in aggregate in *CHEK2* PTV carriers, although breast and prostate cancer explained more than 50% of the excess cancer risk, and the remaining excess was only statistically significant in males. Statistically significant increased risks were observed for seven of the other 22 sites considered: oesophagus, ovary, kidney, HD, DNHL, ML, and melanoma. It is notable that no common sites were associated with a reduced risk—this and the clear association with overall cancer risk after excluding breast and prostate cancer suggests that many of these associations are genuine. While some other sites (eg, kidney) have been previously suggested to be associated,[19] none have been firmly established. We did not, however, find evidence to support the previous observation of an increased risk of colorectal cancer.

Rare MSVs in *CHEK2* were also associated with increased risk of both breast and prostate cancer, with RRs of approximately 1.4, comparable to previous estimates for breast cancer,[6 25] and to estimates for the association of the I157T variant and prostate cancer.[10 40] The RR for breast cancer was higher for variants classified as likely deleterious by Helix score. These results confirm, with similar RR estimates, previous analyses showing that Helix scores are predictive of risk, and that variants with low scores are also associated with risk. A similar pattern was seen for all cancers in both males and females, but not for prostate cancer. The latter result might, however, be due to chance, given the results for all cancers combined. A significant excess risk in rMSV carriers was observed for several other cancer types (melanoma, multiple myeloma, and LL), which were also associated with PTVs. Taken together, this is broadly consistent with the hypothesis that *CHEK2* MSVs confer increased, but on average lower, risks of a similar spectrum of cancers to PTVs, with risks being mutation dependent. High-throughput functional assays to classify the degree of *CHEK2* abrogation are becoming available and may be able to define variant-specific risks more precisely.[41–43] Previous studies have shown that the *CHEK2**I157T variant is associated with a reduced risk for lung cancer.[44] There were too few carriers of this variant to investigate individually, but there was no evidence of a reduced risk associated with rMSVs in aggregate for this site.

Strengths and limitations

A major strength of this study is that, as a very large cohort study, it was possible to assess the risks for many cancer types simultaneously. Moreover, as individuals were genotyped independently of phenotype, the analyses should free from biases due to differential genotyping of cases and controls. In contrast to family-based studies, the risk estimates should be broadly applicable to the population. A potential limitation is that UK Biobank, in common with similar cohorts, is biased towards healthy volunteers: total cancer incidence for the age-group 70–74 is 12–18% lower than in the general population.[45] However, while this affects the power to detect associations, it is unlikely that the RRs will be materially biased. Further, we generated absolute risk estimates by applying these RRs to national incidence rates. A significant limitation is that more than 90% of the cohort is of European ancestry, and we therefore restricted the analyses to this population. Other analyses indicate that the RRs conferred by cancer susceptibility genes tend to be similar across populations of different ancestries[6 46]; nevertheless, extrapolation to other ancestries may be less reliable (particularly if there are ancestry-specific variants associated with different risks).

Implications

ATM and *CHEK2* are included on many cancer gene panels used in family cancer clinics, and the risk estimates from these analyses can inform genetic counselling for carriers. The estimated absolute risks for pancreatic cancer in *ATM* PTV carriers (11% in males and 8% in females by age 85) are notably higher than for other major pancreatic susceptibility genes including *BRCA2*, *CDK2NA,* and *PALB2,* although the confidence limits are wide. Current National Institute for Health and Care Excellence guidelines for pancreatic cancer management do not include *ATM*, though it is included in other guidelines. Given the frequency of *ATM* variants, *ATM* PTVs would explain ~2% of pancreatic cancer cases (with potentially a higher proportion due to MSVs). Prognosis for pancreatic cancer is poor, with only 7% of patients diagnosed with pancreatic cancer between 2013 and 2017 in England surviving 5 years.[47] *ATM* variant carriers may provide a suitable population in which to evaluate new methods of early detection of this disease. Moreover, given the high excess lifetime risk for all cancers, carriers may be a good target population for methods, such as those based on circulating tumour DNA, than can be used for multiple cancer types.[48]

It is important to note that cancer risks are determined by a combination of multiple factors, including rare gene variants such as those in *ATM* and *CHEK2* studied here, but also lifestyle factors, and commoner genetic variants. Ideally, counselling should be based on synthesising all available data, using risk models such as BOADICEA/CanRisk that have been developed for breast, ovarian, and prostate cancer.[49 50] It would be relatively straightforward to extend BOADICEA/CanRisk to incorporate the ovarian and prostate cancer risks associated with *ATM* and *CHEK2* PTVs. An ongoing challenge is to extend such models to consider a wider range of cancers.

Our analyses demonstrate that *ATM* and *CHEK2* PTVs are associated with increased risks of a wide range of cancers, with the overall cancer risk being higher for *ATM*. *CHEK2* rMSVs are also associated with smaller risks of a similar range of cancers; for *ATM* (but not *CHEK2*), the risks appear to be restricted to a small subset of rMSVs. The absolute magnitudes of these risks (with more than half of PTV carriers being affected by age 80) are significant and may influence guidelines for counselling and management.

ORCID iDs
Toqir K Mukhtar http://orcid.org/0009-0001-4983-495X
Xin Yang http://orcid.org/0000-0003-0037-3790

REFERENCES

1. Kuchenbaecker KB, Hopper JL, Barnes DR, *et al*. Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. *JAMA* 2017;317:2402–16.
2. Niskakoski A, Pasanen A, Lassus H, *et al*. Molecular changes preceding endometrial and ovarian cancer: a study of consecutive endometrial specimens from Lynch syndrome surveillance. *Mod Pathol* 2018;31:1291–301.
3. Matsuoka S, Huang M, Elledge SJ. Linkage of ATM to cell cycle regulation by the Chk2 protein kinase. *Science* 1998;282:1893–7.
4. Chehab NH, Malikzay A, Appel M, *et al*. Chk2/hCds1 functions as a DNA damage checkpoint in G(1) by stabilizing p53. *Genes Dev* 2000;14:278–88.
5. Bartek J, Falck J, Lukas J. CHK2 kinase--a busy messenger. *Nat Rev Mol Cell Biol* 2001;2:877–86.
6. Breast Cancer Association Consortium, Dorling L, Carvalho S, *et al*. Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women. *N Engl J Med* 2021;384:428–39.
7. Consortium C-C. CHEK2*1100delC and Susceptibility to Breast Cancer: A Collaborative Analysis Involving 10,860 Breast Cancer Cases and 9,065 Controls from 10 Studies. *Am J Hum Genet* 2004;74:1175–82.
8. Hu C, Hart SN, Gnanaolivu R, *et al*. A Population-Based Study of Genes Previously Implicated in Breast Cancer. *N Engl J Med* 2021;384:440–51.
9. Dorling L, Carvalho S, Allen J, *et al*. Breast cancer risks associated with missense variants in breast cancer susceptibility genes. *Genome Med* 2022;14:51.
10. Cybulski C, Wokołorczyk D, Huzarski T, *et al*. A large germline deletion in the Chek2 kinase gene is associated with an increased risk of prostate cancer. *J Med Genet* 2006;43:863–6.
11. de Jong MM, Nolte IM, Te Meerman GJ, *et al*. Colorectal cancer and the CHEK2 1100delC mutation. *Genes Chromosomes Cancer* 2005;43:377–82.
12. Xiang H, Geng X, Ge W, *et al*. Meta-analysis of CHEK2 1100delC variant and colorectal cancer susceptibility. *Eur J Cancer* 2011;47:2546–51.
13. Weischer M, Heerfordt IM, Bojesen SE, *et al*. CHEK2*1100delC and risk of malignant melanoma: Danish and German studies and meta-analysis. *J Invest Dermatol* 2012;132:299–303.
14. Kilpivaara O, Alhopuro P, Vahteristo P, *et al*. CHEK2 I157T associates with familial and sporadic colorectal cancer. *J Med Genet* 2006;43:e34.
15. Szymanska-Pasternak J, Szymanska A, Medrek K, *et al*. CHEK2 variants predispose to benign, borderline and low-grade invasive ovarian tumors. *Gynecol Oncol* 2006;102:429–31.
16. Obazee O, Archibugi L, Andriulli A, *et al*. Germline BRCA2 K3326X and CHEK2 I157T mutations increase risk for sporadic pancreatic ductal adenocarcinoma. *Int J Cancer* 2019;145:686–93.
17. Teodorczyk U, Cybulski C, Wokołorczyk D, *et al*. The risk of gastric cancer in carriers of CHEK2 mutations. *Fam Cancer* 2013;12:473–8.
18. Siołek M, Cybulski C, Gąsior-Perczak D, *et al*. CHEK2 mutations and the risk of papillary thyroid cancer. *Int J Cancer* 2015;137:548–52.
19. Cybulski C, Górski B, Huzarski T, *et al*. CHEK2 is a multiorgan cancer susceptibility gene. *Am J Hum Genet* 2004;75:1131–5.
20. Wang Y, McKay JD, Rafnar T, *et al*. Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. *Nat Genet* 2014;46:736–41.
21. Ruth KS, Day FR, Hussain J, *et al*. Genetic insights into biological mechanisms governing human ovarian ageing. *Nature New Biol* 2021;596:393–7.
22. Shiloh Y, Ziv Y. The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nat Rev Mol Cell Biol* 2013;14:197–210.
23. Savitsky K, Bar-Shira A, Gilad S, *et al*. A single ataxia telangiectasia gene with a product similar to PI-3 kinase. *Science* 1995;268:1749–53.
24. Thompson D, Duedal S, Kirner J, *et al*. Cancer risks and mortality in heterozygous ATM mutation carriers. *J Natl Cancer Inst* 2005;97:813–22.
25. Decker B, Allen J, Luccarini C, *et al*. Rare, protein-truncating variants in ATM, CHEK2 and PALB2, but not XRCC2, are associated with increased breast cancer risks. *J Med Genet* 2017;54:732–41.
26. Karlsson Q, Brook MN, Dadaev T, *et al*. Rare Germline Variants in ATM Predispose to Prostate Cancer: A PRACTICAL Consortium Study. *Eur Urol Oncol* 2021;4:570–9.
27. Hsu F-C, Roberts NJ, Childs E, *et al*. Risk of Pancreatic Cancer Among Individuals With Pathogenic Variants in the ATM Gene. *JAMA Oncol* 2021;7:1664–8.
28. Dombernowsky SL, Weischer M, Allin KH, *et al*. Risk of cancer by ATM missense mutations in the general population. *J Clin Oncol* 2008;26:3057–62.
29. Helgason H, Rafnar T, Olafsdottir HS, *et al*. Loss-of-function variants in ATM confer risk of gastric cancer. *Nat Genet* 2015;47:906–10.
30. Laitman Y, Nielsen SM, Bernstein-Molho R, *et al*. Cancer risks associated with heterozygous ATM loss of function and missense pathogenic variants based on multigene panel analysis. *Breast Cancer Res Treat* 2022;196:355–61.
31. UK biobank. UK biobank: protocol for a large-scale prospective epidemiological resource. 2007.
32. Backman JD, Li AH, Marcketta A, *et al*. Exome sequencing and analysis of 454,787 UK Biobank participants. *Nature New Biol* 2021;599:628–34.
33. Lau WCY, Li Y, Liu Z, *et al*. Structure of the human dimeric ATM kinase. *Cell Cycle* 2016;15:1117–24.
34. Kircher M, Witten DM, Jain P, *et al*. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014;46:310–5.
35. Vroling B, Heijl S. White paper: the helix pathogenicity prediction platform. *arXiv preprint* [Preprint] 2021.
36. UKBiobank. Cancer data: linkage from national cancerregistries2023 11.09.23; (Version 2.0). Available: https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/CancerLinkage.pdf
37. World Health Organisation. The International Classification of Diseases and Related Health Problems, Tenth Revision. Geneva, Switzerland, 1992.
38. Cancer Research UK. Cancer incidence statistics. n.d. Available: https://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence
39. Schmidt MK, Hogervorst F, van Hien R, *et al*. Age- and Tumor Subtype-Specific Breast Cancer Risk Estimates for CHEK2*1100delC Carriers. *J Clin Oncol* 2016;34:2750–60.
40. Cybulski C, Huzarski T, Górski B, *et al*. A novel founder CHEK2 mutation is associated with increased prostate cancer risk. *Cancer Res* 2004;64:2677–9.
41. Kleiblova P, Stolarova L, Krizova K, *et al*. Identification of deleterious germline CHEK2 mutations and their association with breast and ovarian cancer. *Int J Cancer* 2019;145:1782–97.
42. Boonen RACM, Wiegant WW, Celosse N, *et al*. Functional Analysis Identifies Damaging CHEK2 Missense Variants Associated with Increased Cancer Risk. *Cancer Res* 2022;82:615–31.

43 Stolarova L, Kleiblova P, Zemankova P, *et al*. ENIGMA CHEK2gether Project: A Comprehensive Study Identifies Functionally Impaired CHEK2 Germline Missense Variants Associated with Increased Breast Cancer Risk. *Clin Cancer Res* 2023;29:3037–50.

44 Cybulski C, Masojc B, Oszutowska D, *et al*. Constitutional CHEK2 mutations are associated with a decreased risk of lung and laryngeal cancers. *Carcinogenesis* 2008;29:762–5.

45 Fry A, Littlejohns TJ, Sudlow C, *et al*. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am J Epidemiol* 2017;186:1026–34.

46 Ahearn TU, Choudhury PP, Derkach A, *et al*. Breast Cancer Risk in Women from Ghana Carrying Rare Germline Pathogenic Mutations. *Cancer Epidemiol Biomarkers Prev* 2022;31:1593–601.

47 Office of National Statistics. Cancer Survival in England: adults diagnosed between 2013 and 2017 and followed up to 2018, 2023. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancersurvivalratescancersurvivalinenglandadultsdiagnosed

48 Aravanis AM, Lee M, Klausner RD. Next-Generation Sequencing of Circulating Tumor DNA for Early Cancer Detection. *Cell* 2017;168:571–4.

49 Lee A, Mavaddat N, Wilcox AN, *et al*. BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genet Med* 2019;21:1708–18.

50 Nyberg T, Brook MN, Ficorella L, *et al*. CanRisk-Prostate: A Comprehensive, Externally Validated Risk Model for the Prediction of Future Prostate Cancer. *J Clin Oncol* 2023;41:1092–104.