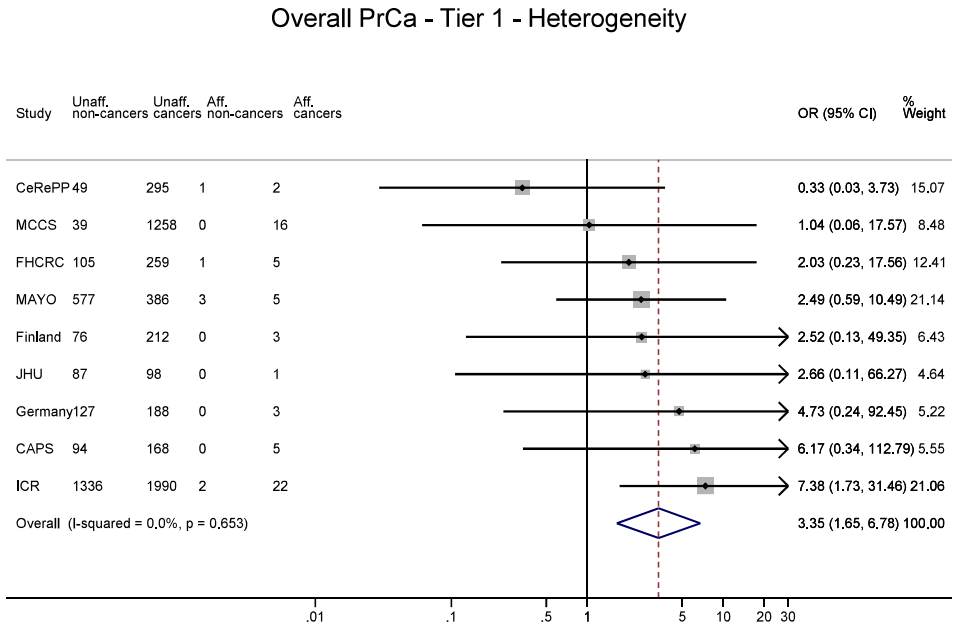
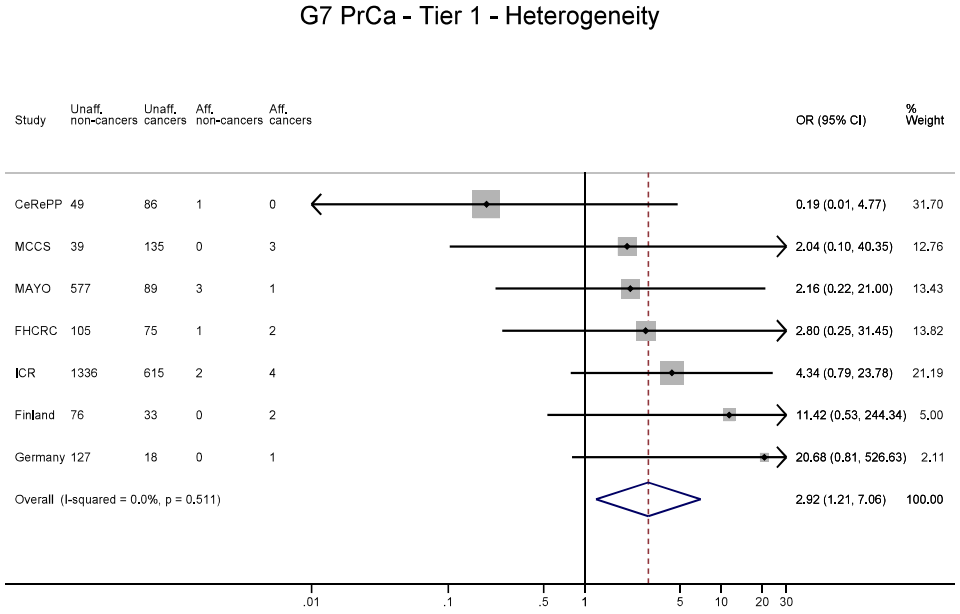
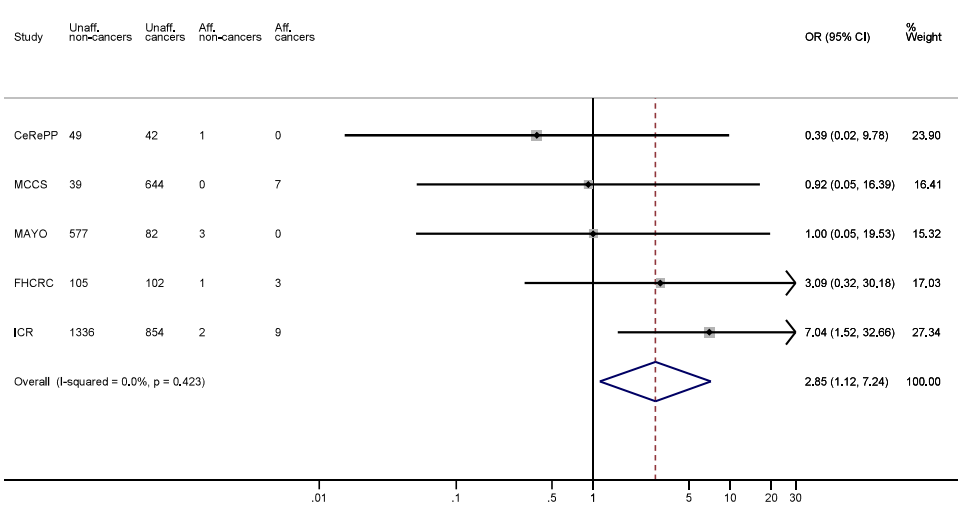
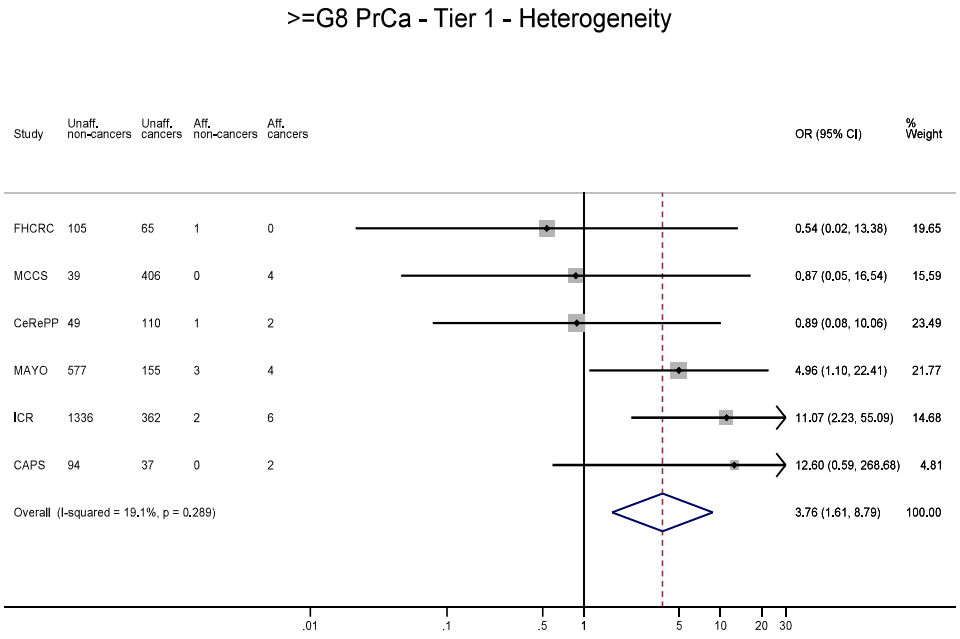
Supplemental Material

Supp. Fig 1a. Tier 1. Heterogeneity.

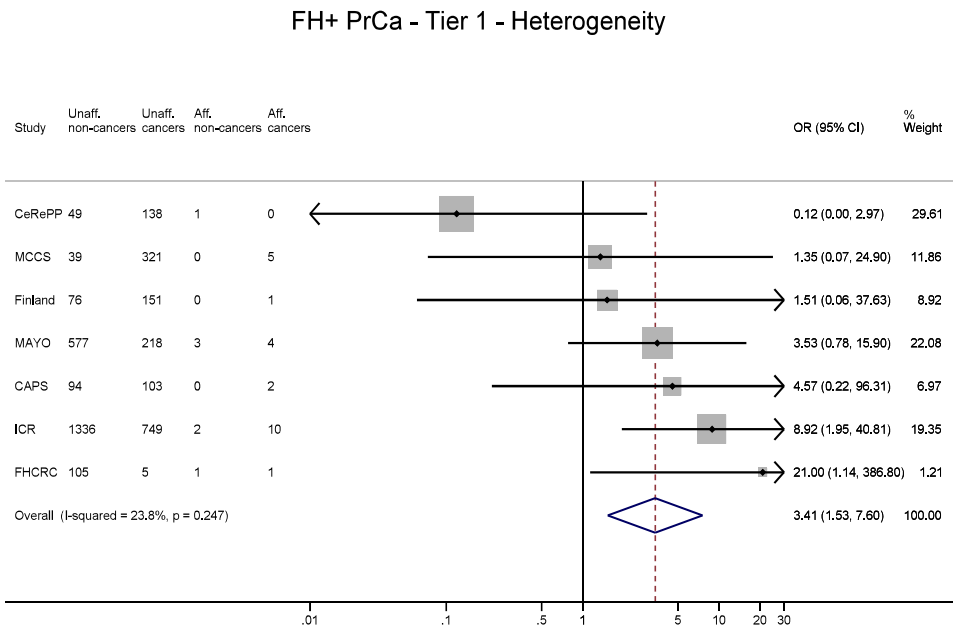


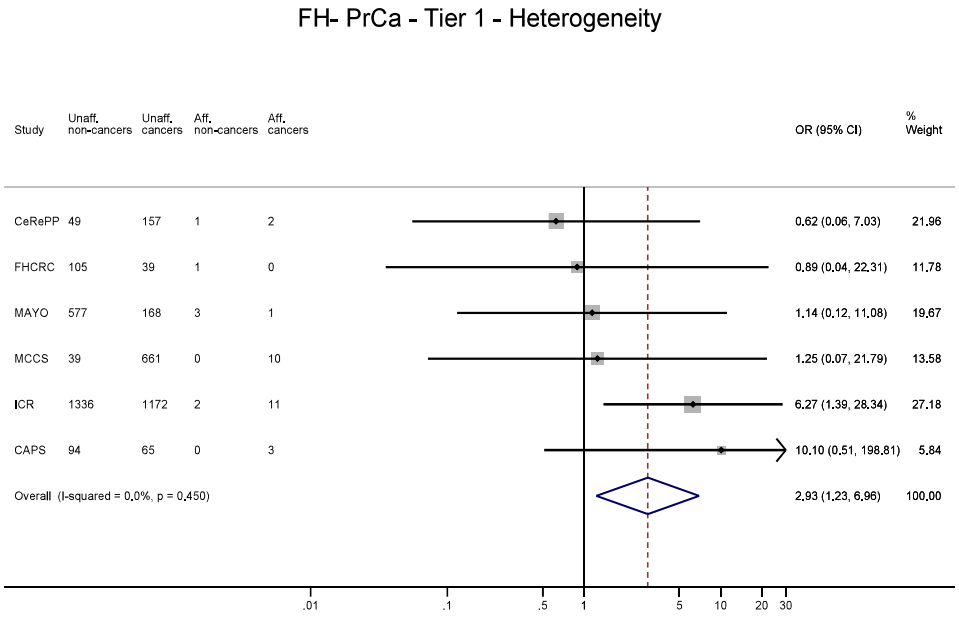


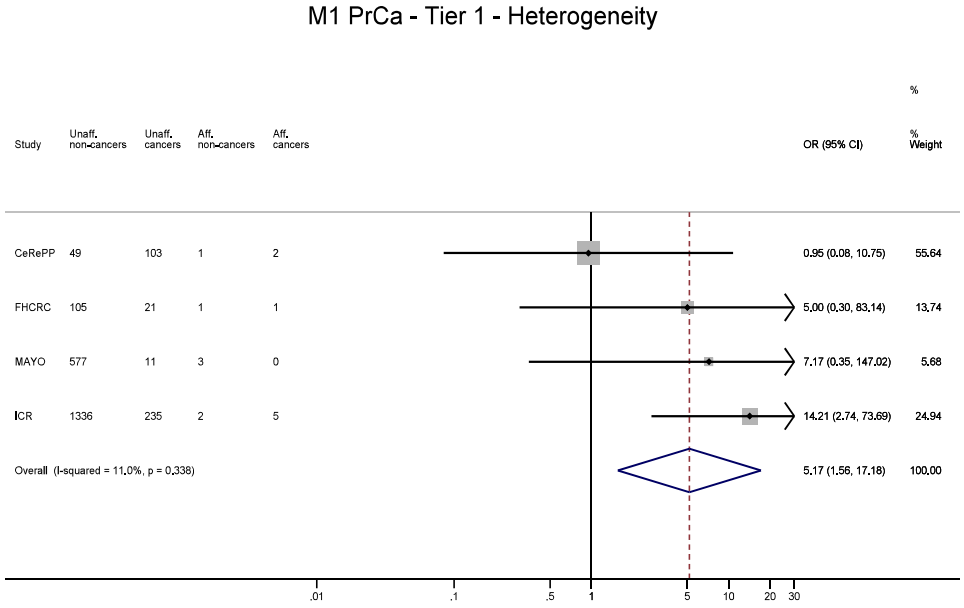


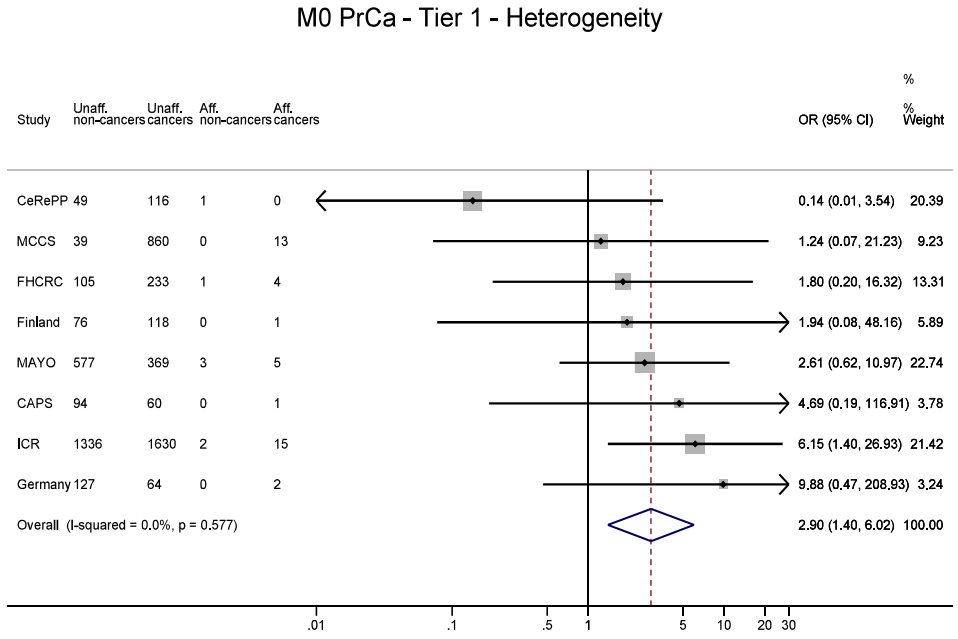


Supp. Fig 1b. Tier 1. Heterogeneity.

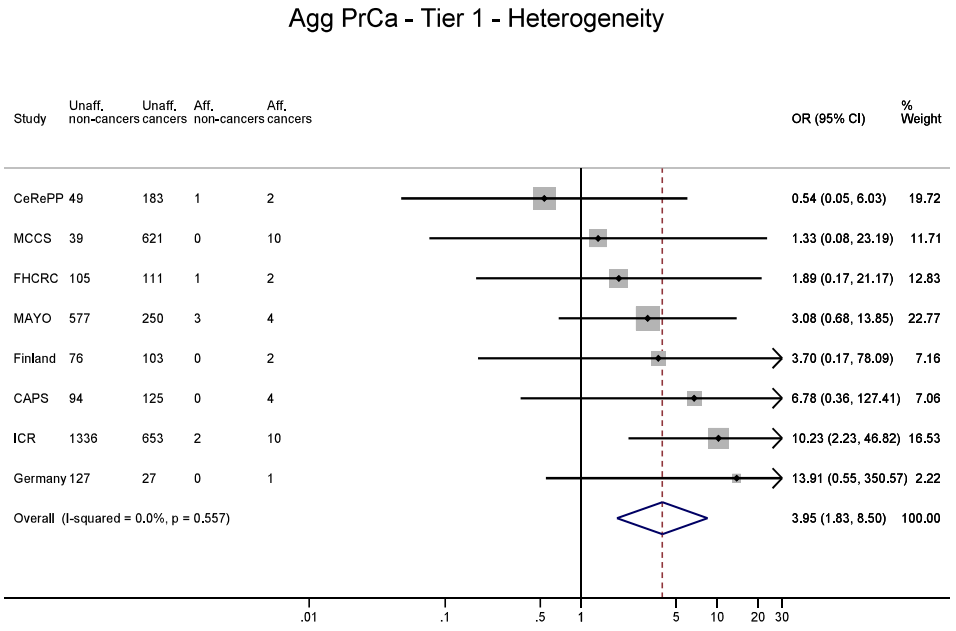


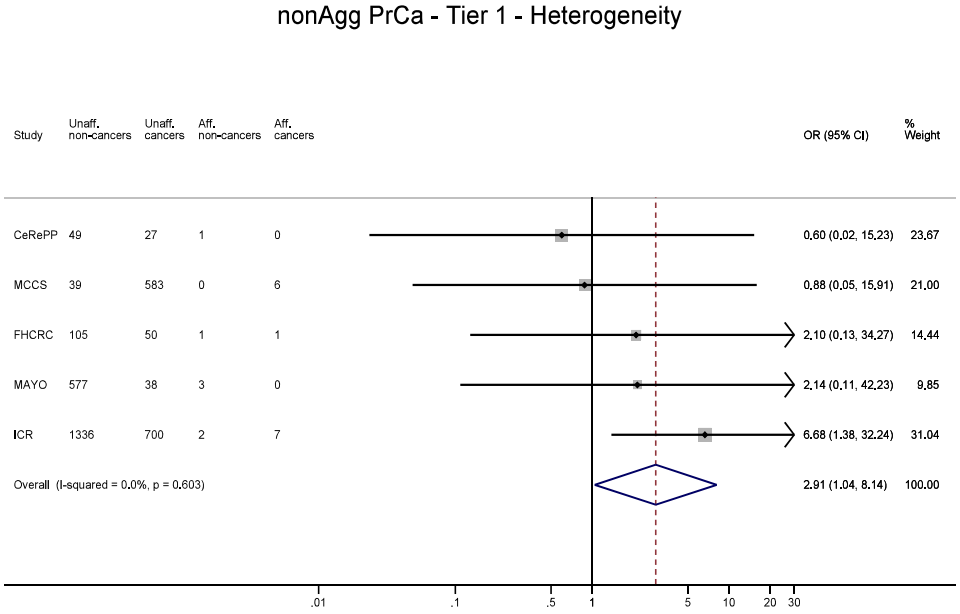


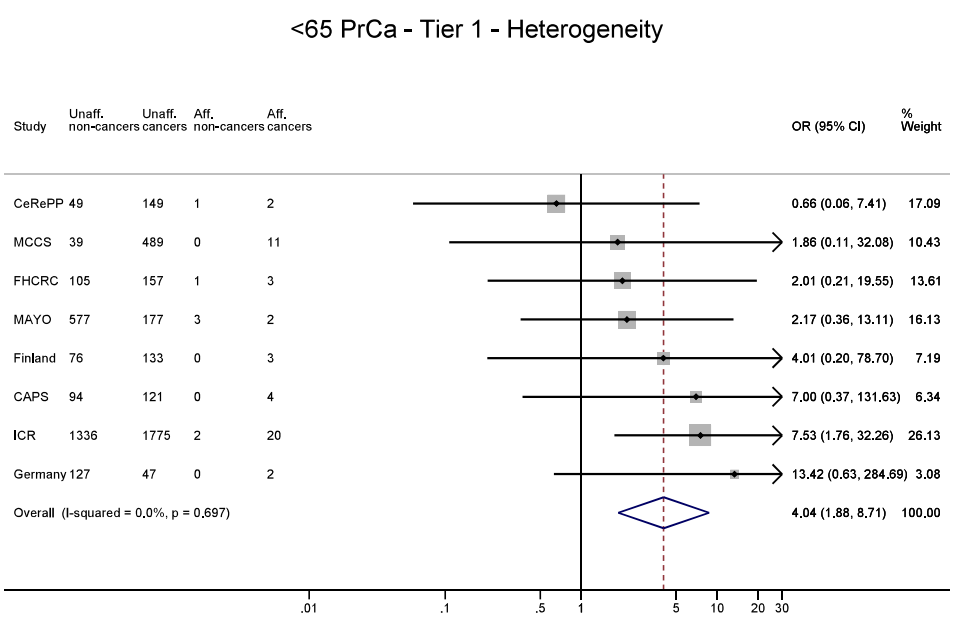


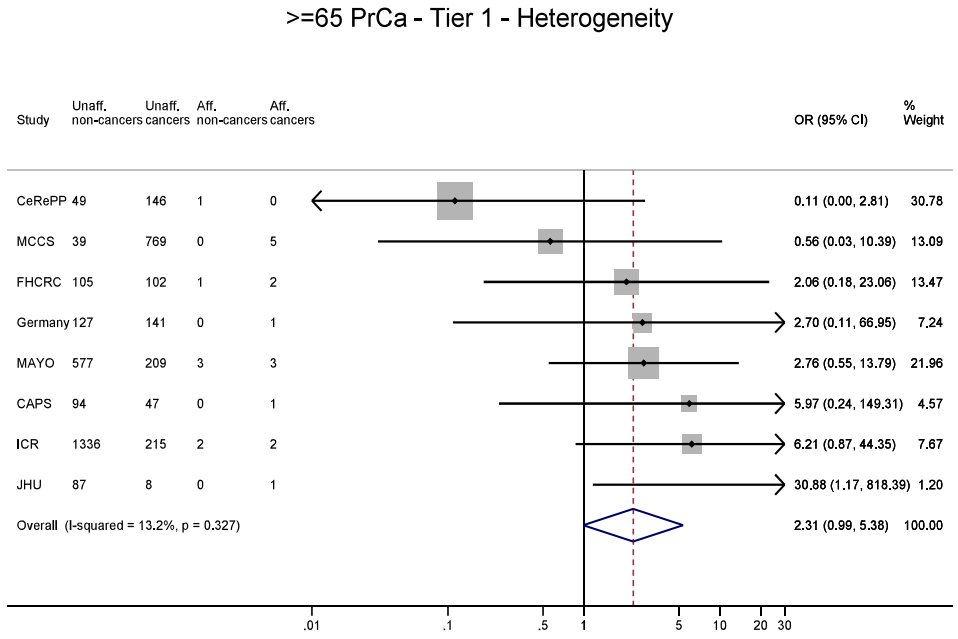


Supp. Fig 1c. Tier 1. Heterogeneity.

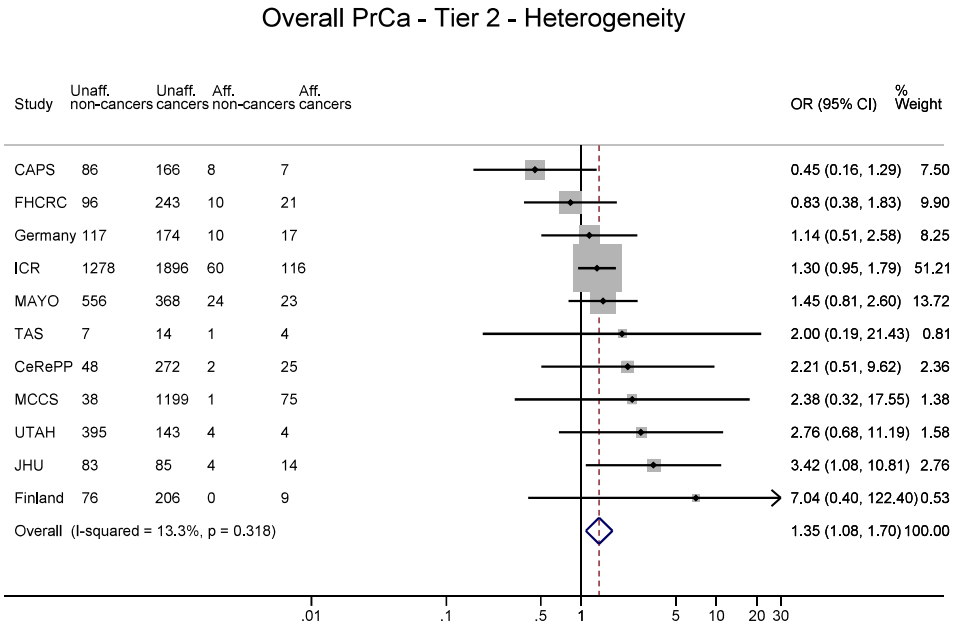


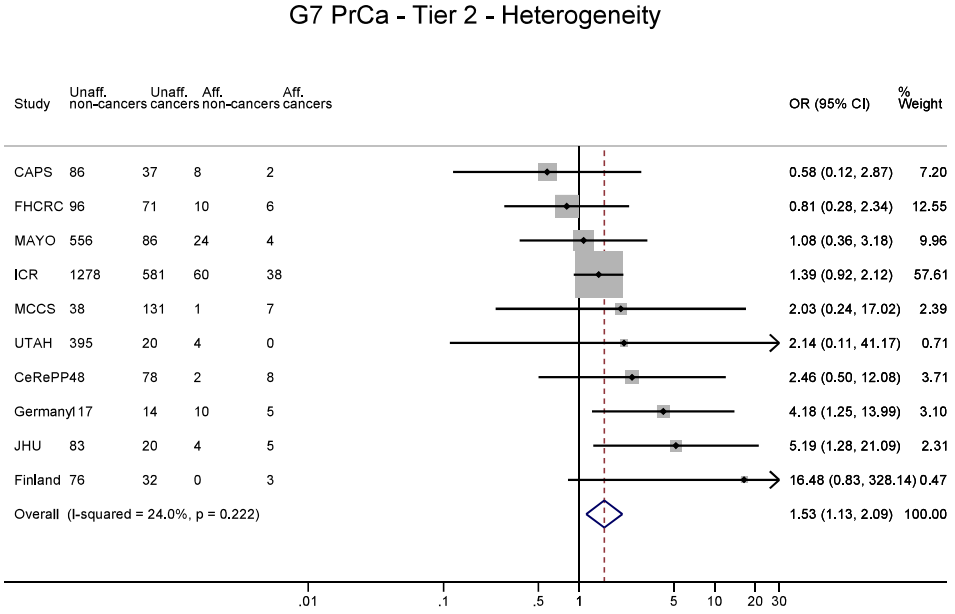


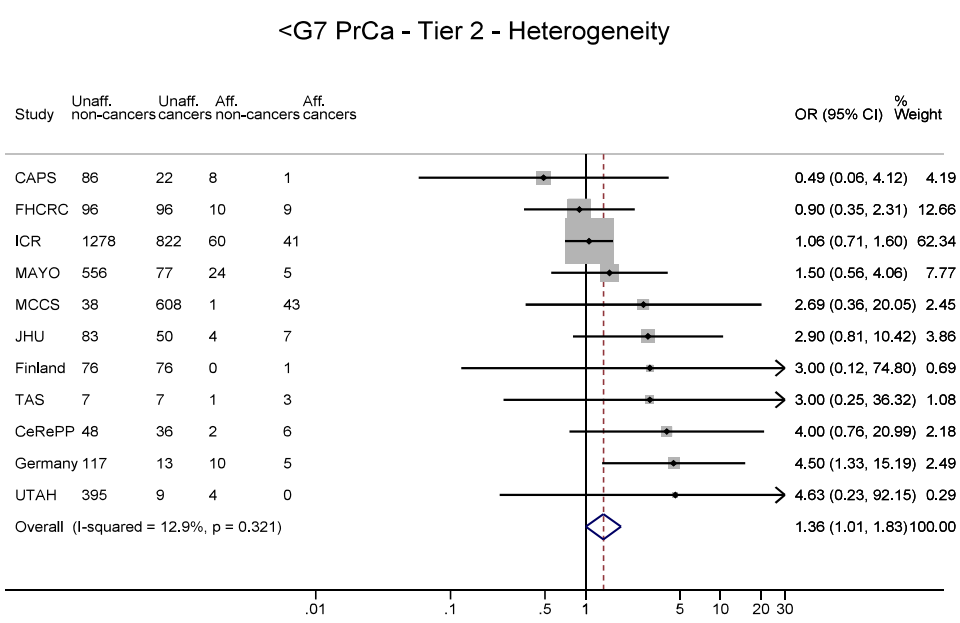


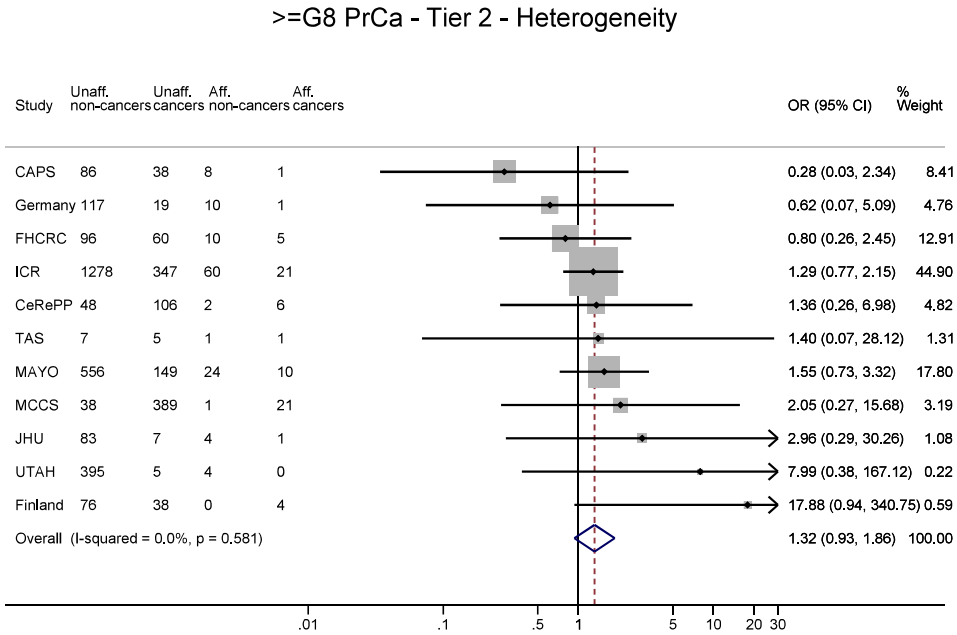


Supp. Fig 2a. Tier 2. Heterogeneity.

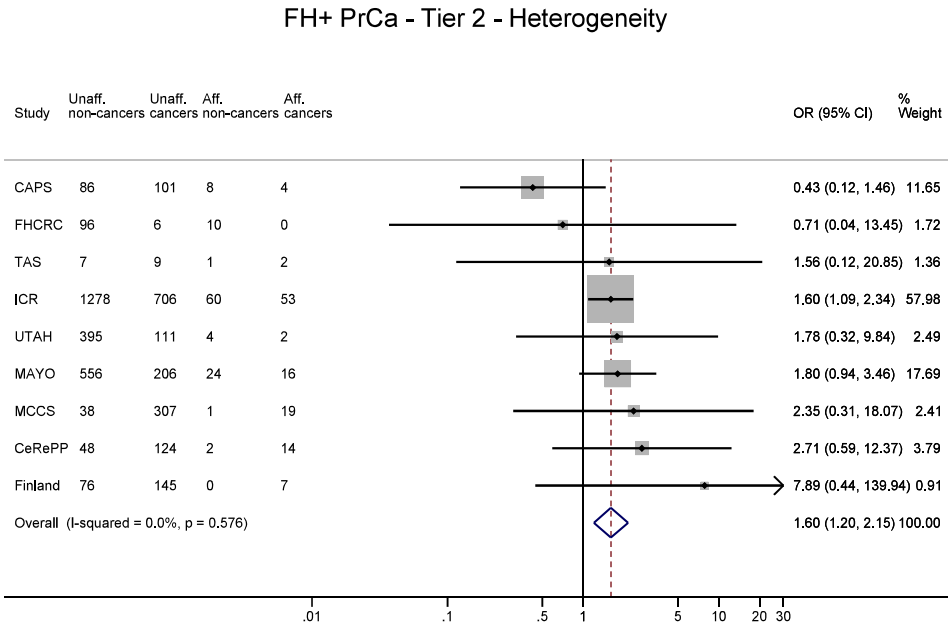


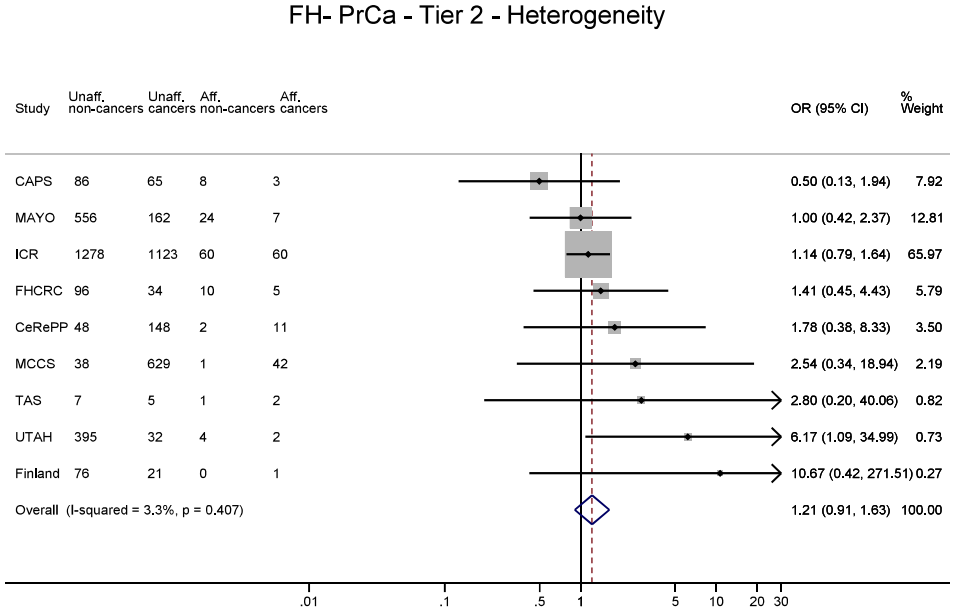


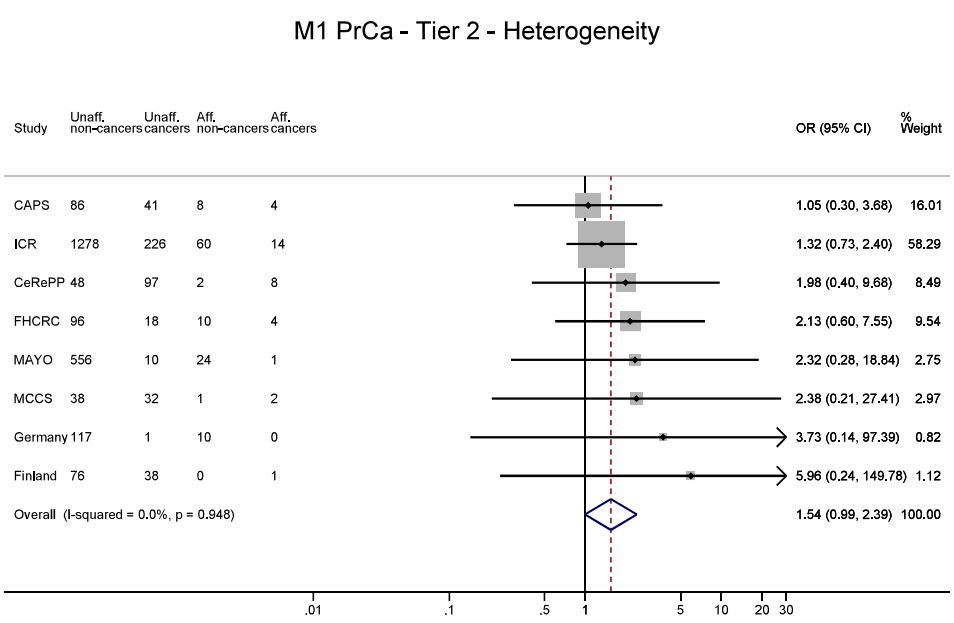


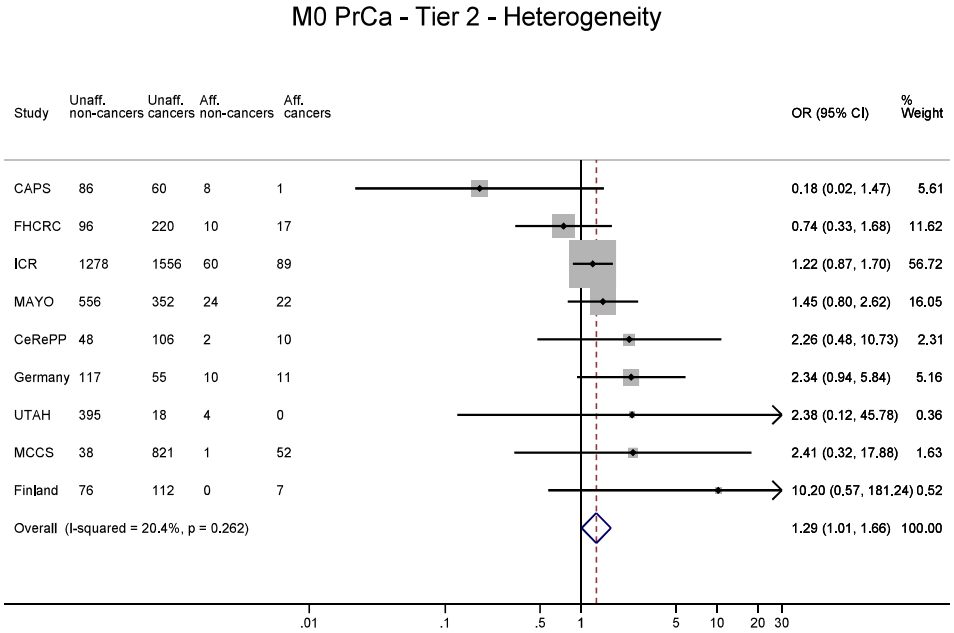


Supp. Fig 2b. Tier 2. Heterogeneity.

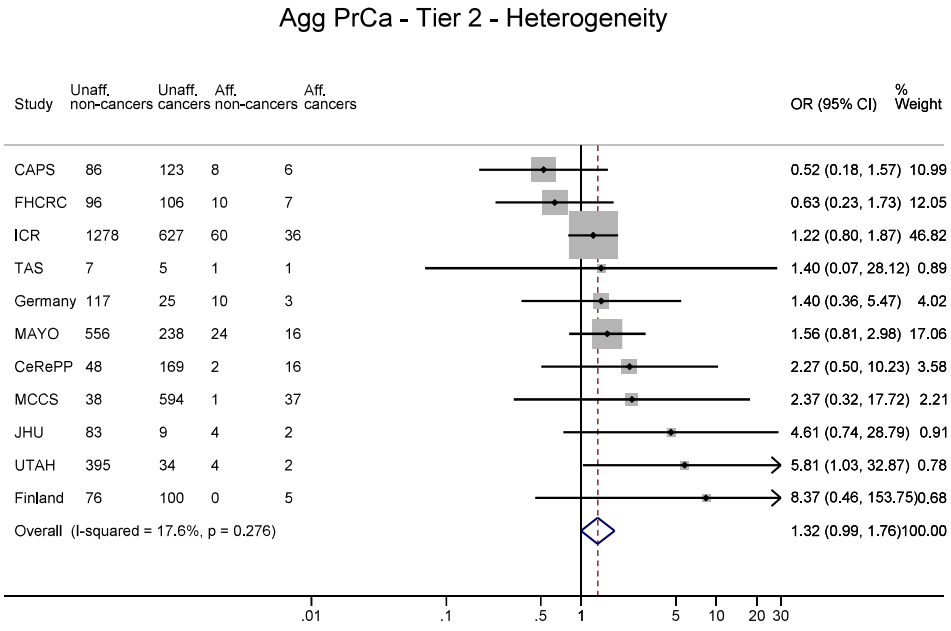


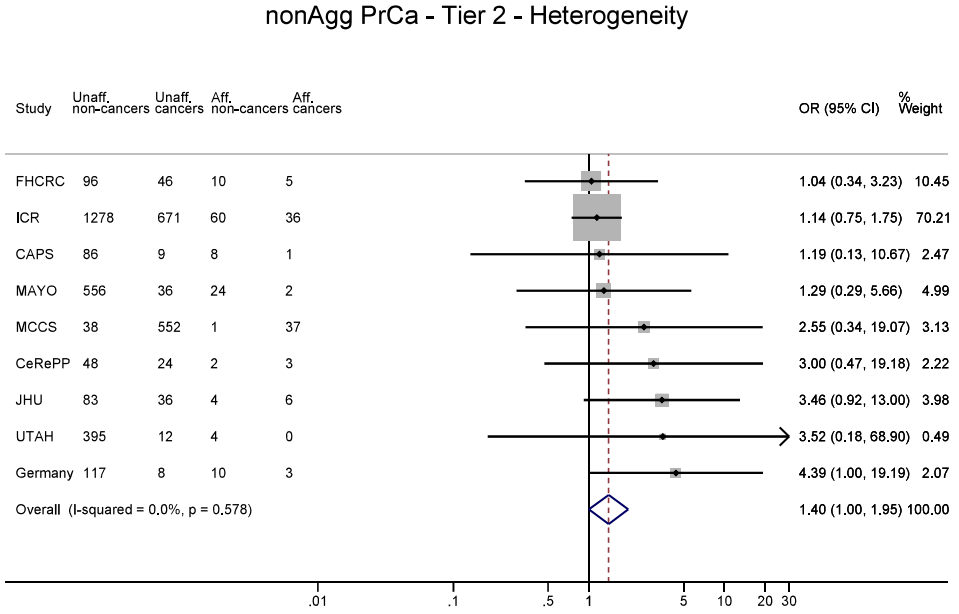


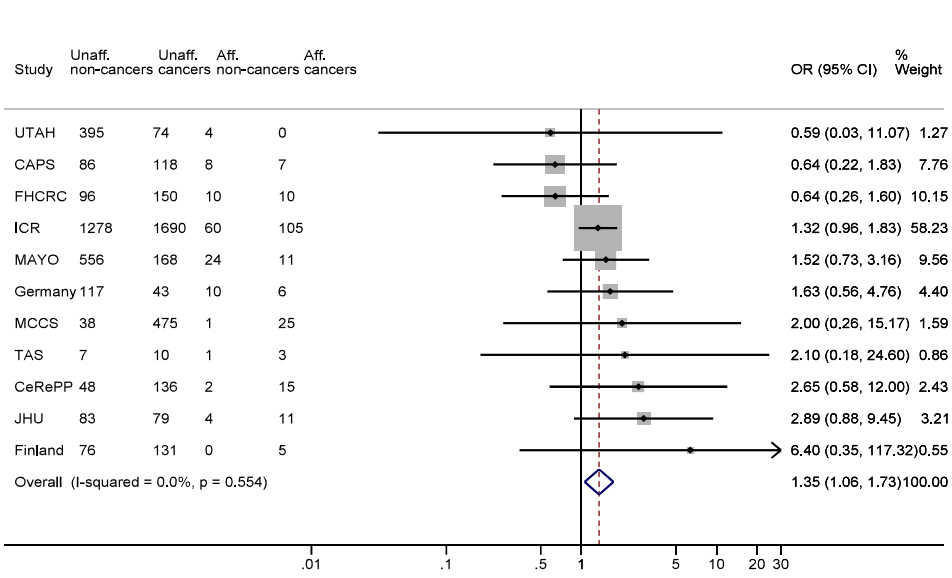


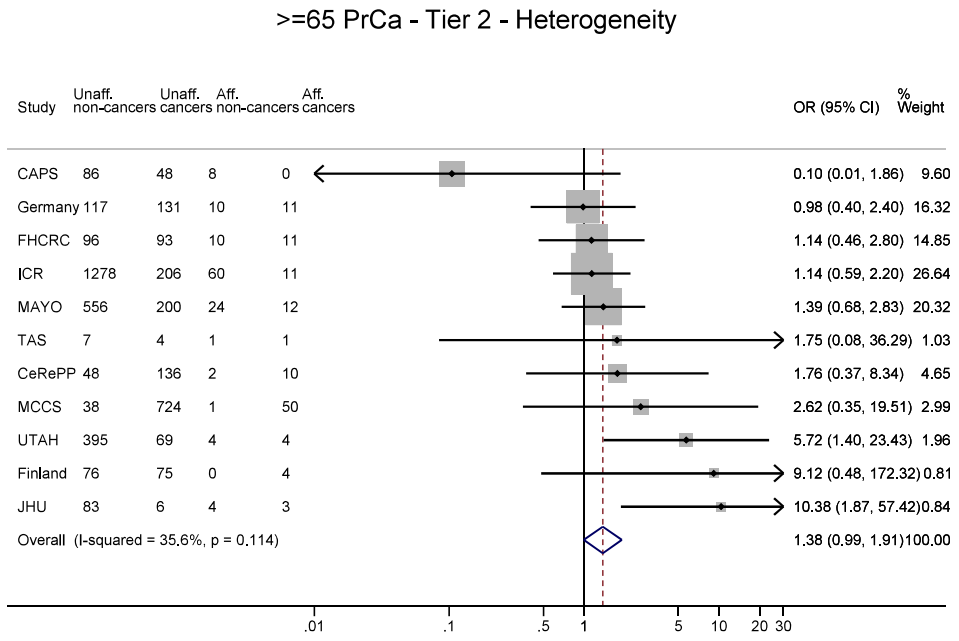


Supp. Fig 2c. Tier 2. Heterogeneity.

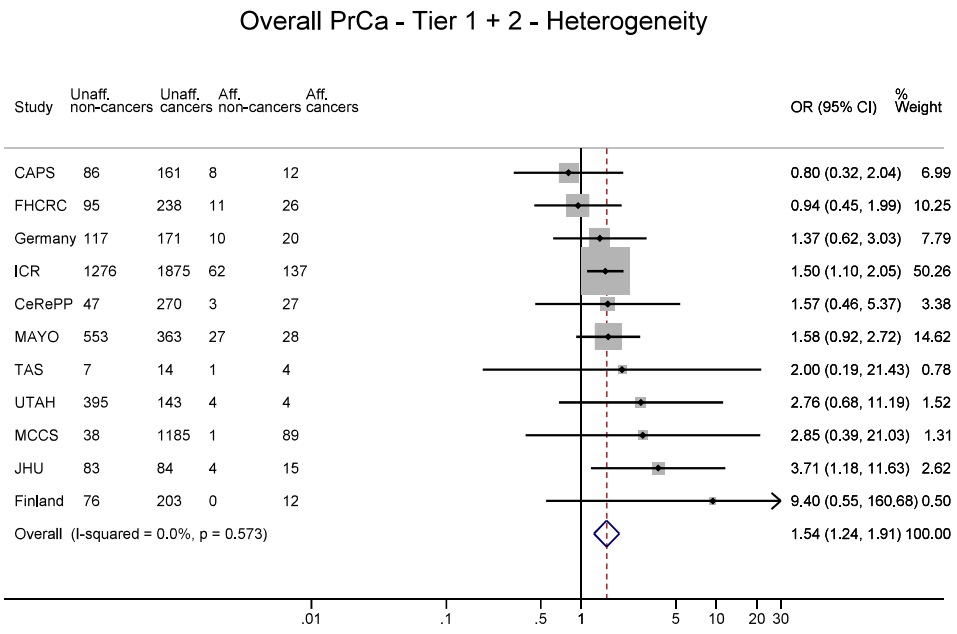


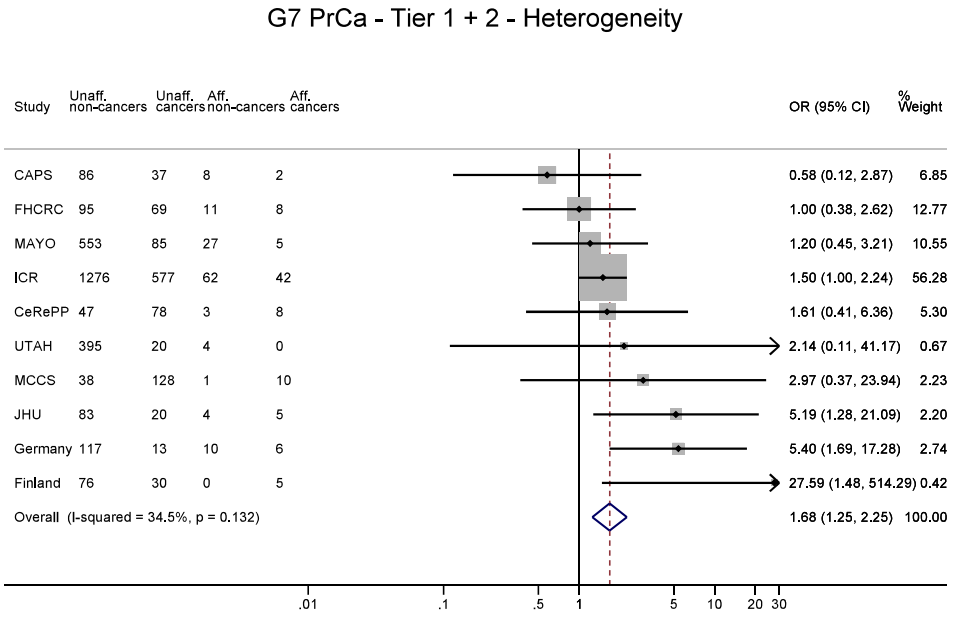


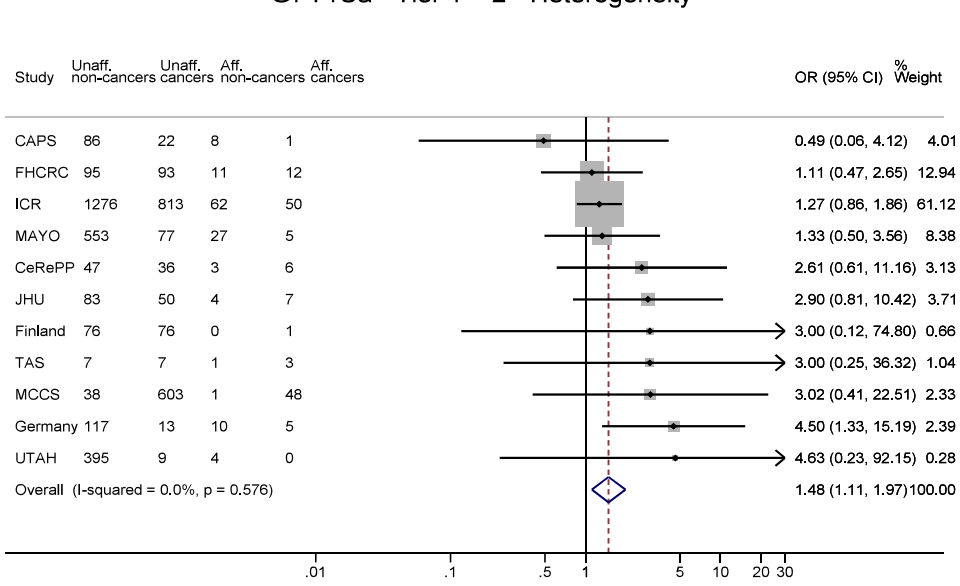


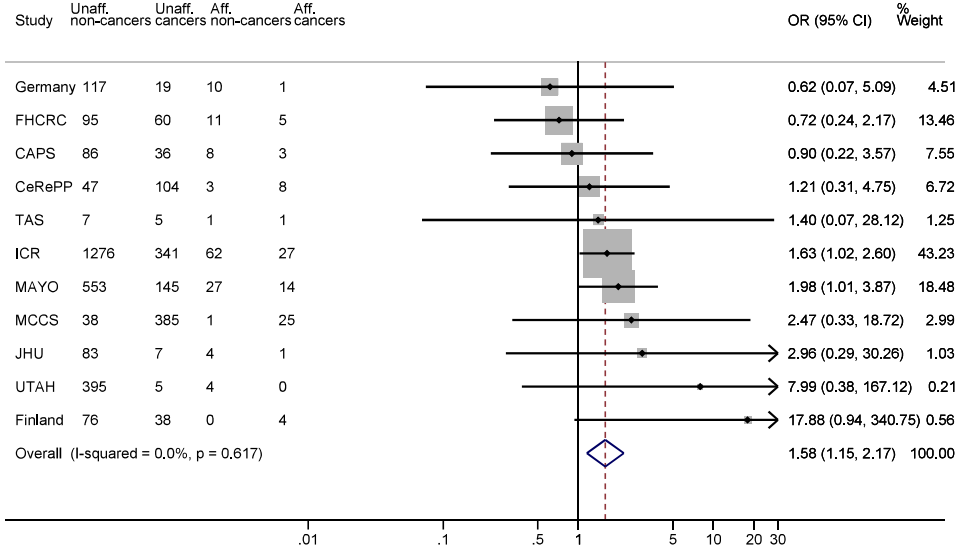


Supp. Fig 3a. Tier 1+2. Heterogeneity.

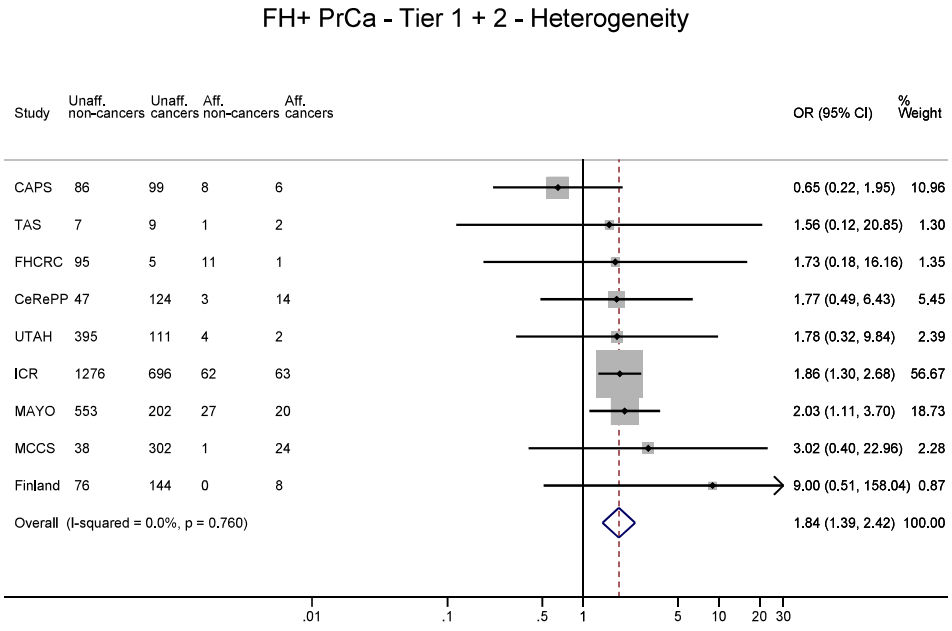


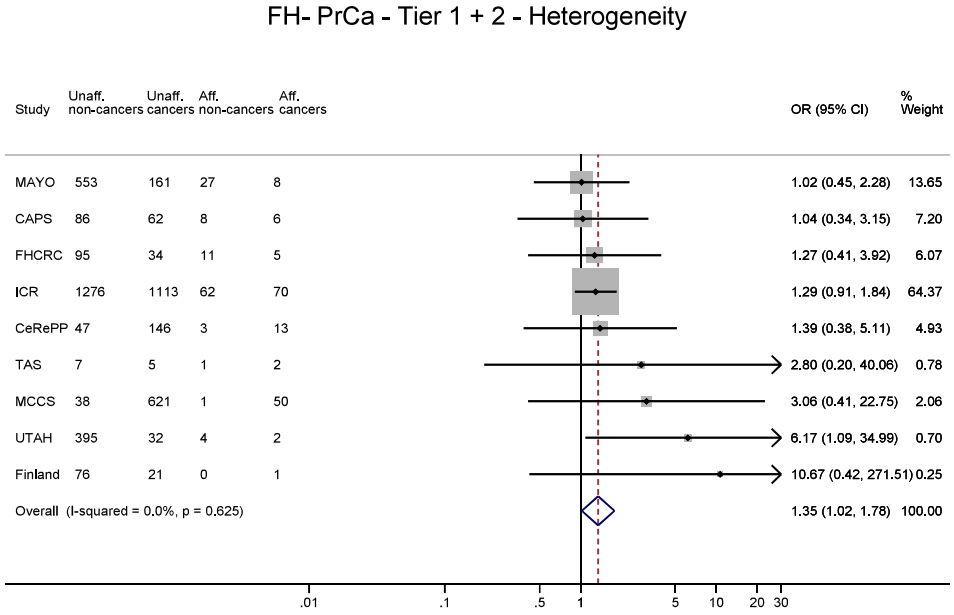


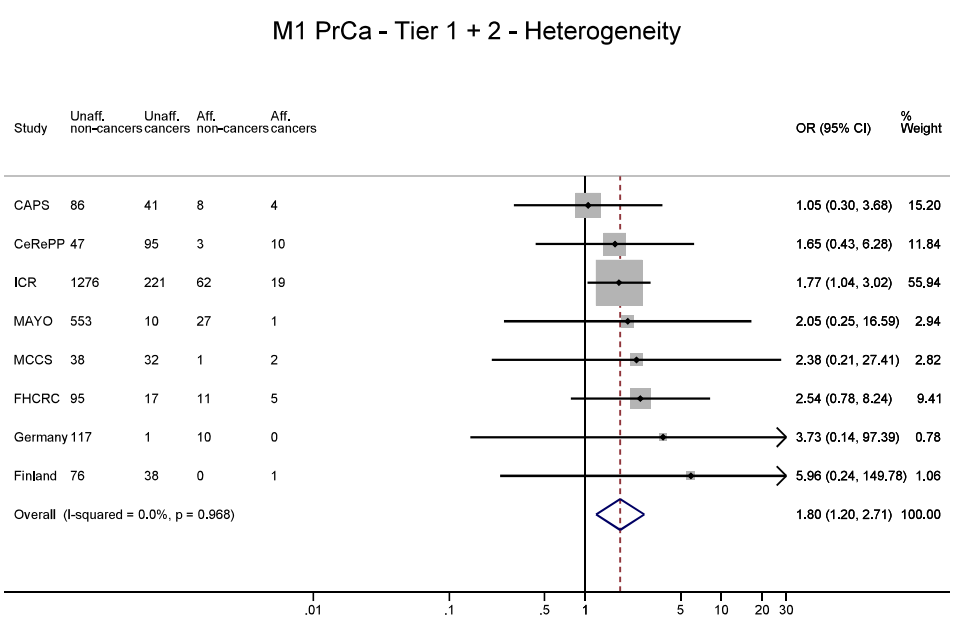


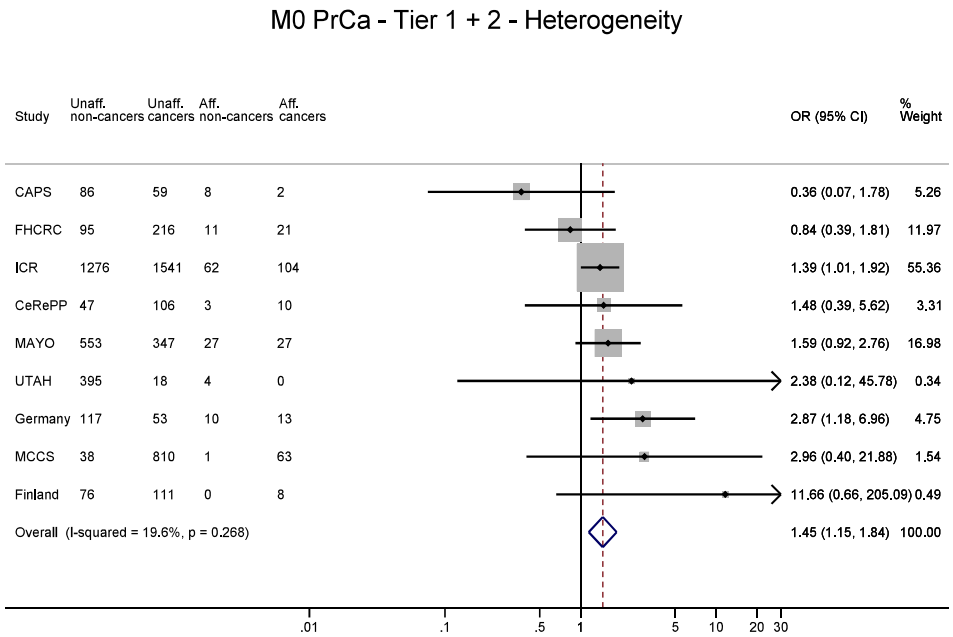


Supp. Fig 3b. Tier 1+2. Heterogeneity.

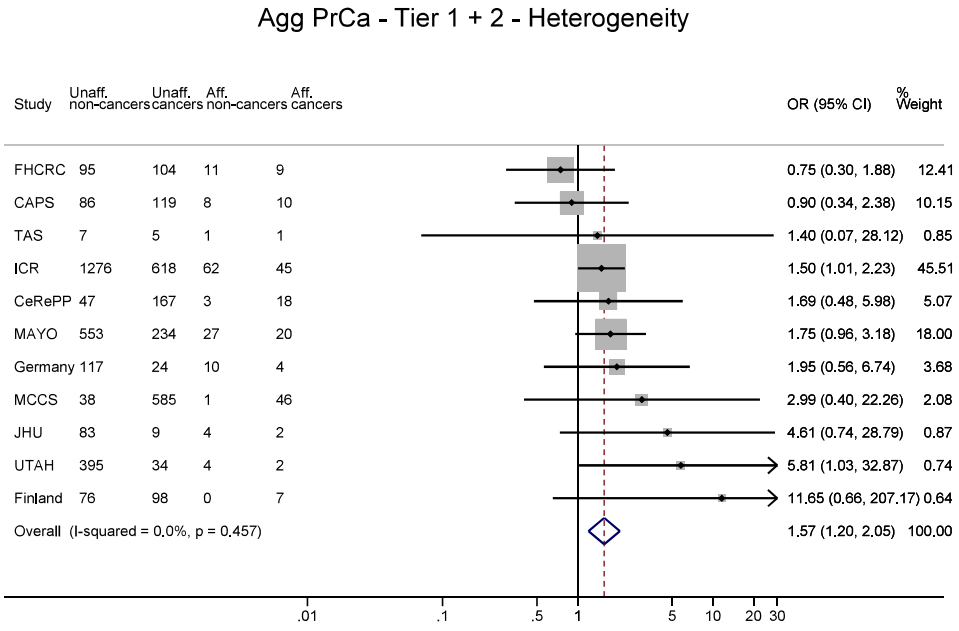


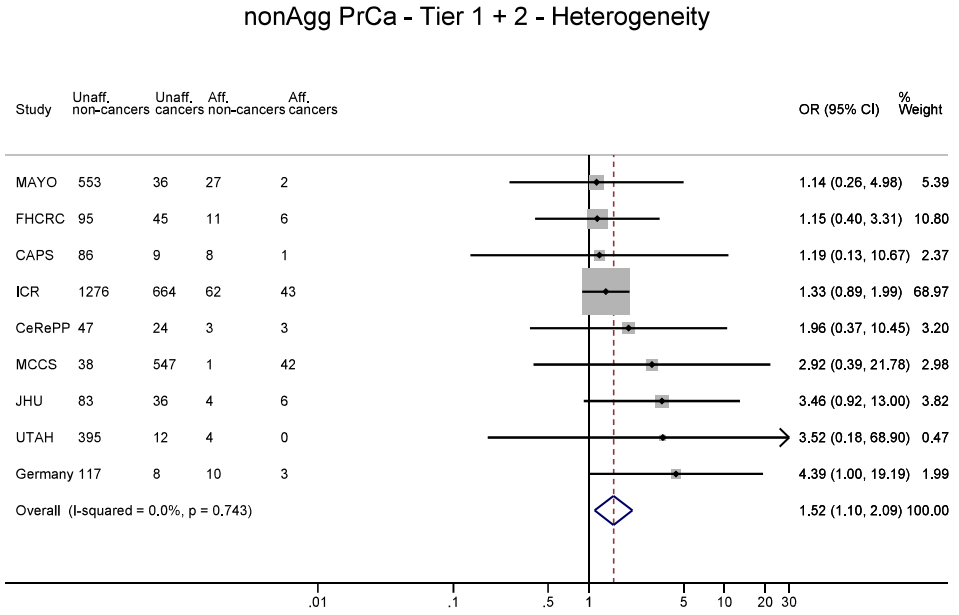


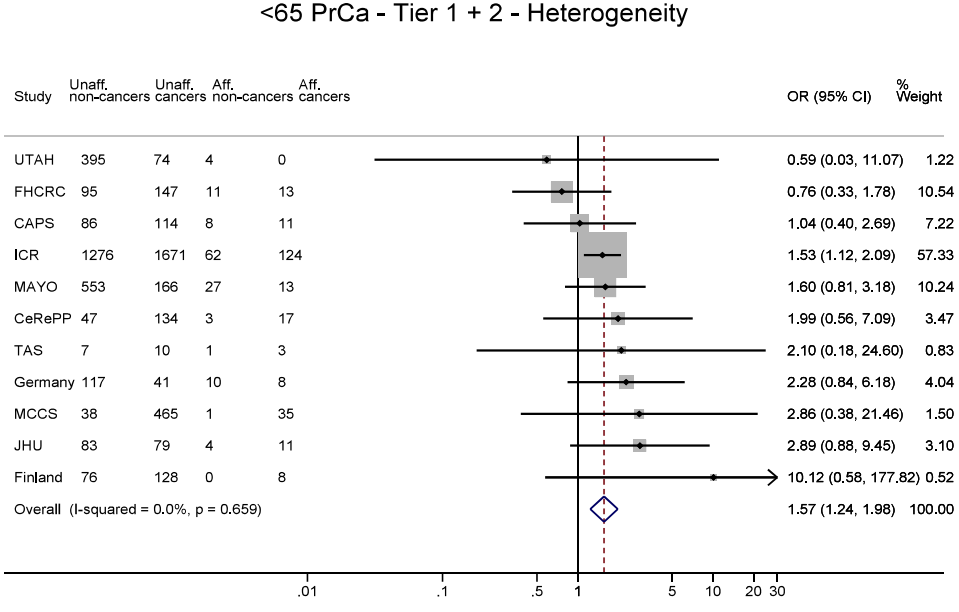


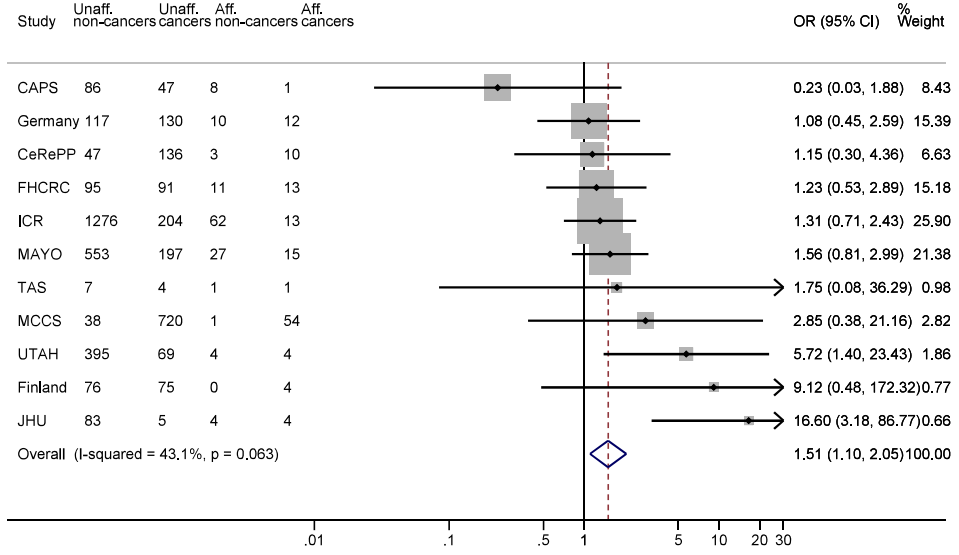


Supp. Fig 3c. Tier 1+2. Heterogeneity.









Funding and Acknowledgements

The PRACTICAL Consortium

We acknowledge support from Cancer Research UK C5047/A17528, the Bob Champion Trust and the Global Challenges Research Fund (GCRF).

CAPS

The Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden was supported by the Cancer Risk Prediction Center (CRisP; www.crispcenter.org), a Linneus Centre (Contract ID 70867902) financed by the Swedish Research Council, Swedish Research Council (grant no K2010-70X-20430-04-3), the Swedish Cancer Foundation (grant no 09-0677), the Hedlund Foundation, the Soederberg Foundation, the Enqvist Foundation, ALF funds from the Stockholm County Council. Stiftelsen Johanna Hagstrand och Sigfrid Linner's Minne, Karlsson's Fund for urological and surgical research. We thank and acknowledge all of the participants in the Stockholm-1 study. We thank Carin Cavalli-Bjoerkman and Ami Roennberg Karlsson for their dedicated work in the collection of data. Michael Broms is acknowledged for his skilful work with the databases. KI Biobank is acknowledged for handling the samples and for DNA extraction. Hans Wallinder at Aleris Medilab and Sven Gustafsson at Karolinska University Laboratory are thanked for their good cooperation in providing historical laboratory results.

CeRePP

The CeRePP study was supported by the U.S. Public Health Service, National Institutes of Health, contract grant U01CA08960 (ICPCG: International Consortium for Prostate Cancer Genetics), and the French Prostate ICGC project was founded by Institut National de la Santé et de la Recherche Médicale (INSERM) and Institut National du Cancer (INCa) through the grant INSERM CV_2011/023 (C18). We thank and acknowledge all the participants in the ProGene study.

DISCLOSE

The DISCLOSE study is supported by the National Medical Research Council Singapore (grant no. CSA/0027/2018), the Singhealth Duke-NUS Oncology Academic Clinical Programme - Proton Therapy Research Philanthropic fund (grant no. 08/FY2018/EX/(SL)/44-A85), and a research grant by Ferring Singapore.

FHCRC (Fred Hutchinson Cancer Research Center)

The FHCRC studies were supported by grants R01-CA080122, R01-CA056678, R01-CA082664, R01-CA092579, and K05-CA175147 from the US National Cancer Institute, National Institutes of Health, with additional support from the Fred Hutchinson Cancer Research Center (P30-CA015704). We thank all the individuals who participated in these studies.

Finland

The Tampere (Finland) study was supported by the Academy of Finland (310105), Cancer Foundation Finland sr, Sigrid Juselius Foundation, and the Jane and Aatos Erkko Foundation. The PSA screening samples were collected by the Finnish part of ERSPC (European Study of Screening for Prostate Cancer). TAMPERE would like to thank Riina Liikanen, Liisa Määttänen and Kirsi Talala for their work on samples and databases.

Germany / ULM

The Ulm group received funds from the German Cancer Aid (Deutsche Krebshilfe).

ICR / UKGPCS

UKGPCS would also like to thank the following for funding support: The Institute of Cancer Research and The Everyman Campaign, The Prostate Cancer Research Foundation, Prostate Research Campaign UK (now Prostate Action), The Orchid Cancer Appeal, The National Cancer Research Network UK, The National Cancer Research Institute (NCRI) UK. We are grateful for support of NIHR funding to the NIHR Biomedical Research Centre at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust. UKGPCS should also like to acknowledge the NCRN nurses, data managers and Consultants for their work in the UKGPCS. UKGPCS would like to thank all urologists and other persons involved in the planning, coordination, and data collection of the study. KM and AL were in part supported from the NIHR Manchester Biomedical Research Centre. TD was supported by the NIH (R01/81570343) grant.

JHU

MAYO

The Mayo group was supported by the US National Cancer Institute (R01CA72818)

MCCS

Melbourne Collaborative Cohort Study (MCCS) cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further augmented by Australian National Health and Medical Research Council grants 209057, 396414 and 1074383 and by infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry and the Australian Institute of Health and Welfare, including the National Death Index and the Australian Cancer Database.

MD Anderson

Porto

The IPO-Porto study was funded by Fundação para a Ciência e a Tecnologia (FCT; UID/DTP/00776/2013 and PTDC/DTP-PIC/1308/2014) and FEDER (POCI-01-0145-FEDER-028245). MC (SFRH/BD/116557/2016) and MPS (SFRH/BD/132441/2017) are research fellows from FCT. We would like to express our gratitude to all patients and families who have participated in this study.

TASPRAC

The Tasmanian Familial Prostate Cancer Study and the Tasmanian Prostate Cancer Case Control study were funded by Cancer Australia, Cancer Council Tasmania, the Royal Hobart Hospital Research Foundation, Perpetual Trustees, the Mazda Foundation and the Max Bruce Trust. An Australian Research Council Future Fellowship supported JLD in addition to supporting the whole genome sequencing. LMF is supported by a Cancer Council Tasmania/College of Health and Medicine Fellowship. Prostate cancer cases and their clinical information were ascertained through the Tasmanian Cancer Registry. We are greatly indebted to Tasmanian Cancer Registry staff, urologists and collaborating pathologists for their valuable contribution. We would particularly like to thank the participants who volunteered to take part in these studies and the Tasmanian community for their ongoing support of our research.

UTAH

UTAH: The Keith and Susan Warshaw Fund, C. S. Watkins Urologic Cancer Fund and The Tennity Family Fund supported the Utah study. The project was supported by Award Number P30CA042014 from the National Cancer Institute and DOD Proposal Number PC170413.

ATM Groups, Recruitment, Sample Selection, and Sequencing

1. CAPS

Recruitment

CAPS is a hospital-based case-control study. Identified and recruited biopsy confirmed prostate cancer cases from four out of six regional cancer registries in Sweden, diagnosed between July 2001 and October 2003. Clinical data including TNM stage, Gleason grade and PSA levels at time for diagnosis were retrieved through record linkage to the National Prostate Cancer Registry. Control subjects, who were recruited concurrently with case subjects, were randomly selected from the Swedish Population Registry and matched according to the expected age distribution of cases (groups of 5-year intervals) and geographic region. Whole blood was collected from all individuals for extraction of genomic DNA.
[CAPS et al. 2008]

Sequencing
19 cases were sequenced at the study site.
154 cases and 94 controls were sequenced at Mayo as part of ICPCG.

2. CeRePP

Recruitment
All prostate cancer cases and controls were selected among the patients included in the French Prostate Cancer Case Control Study (CeRePP), also known as ProGene study which began in July 1994 [Thomas et al 2008]. Cases were recruited from patients treated in French departments of Urology, who had histologically confirmed prostate cancer. Controls were recruited as participating in a systematic health screening program and found unaffected (normal digital rectal examination and total PSA < 4 ng/ml, or negative biopsy if PSA > 4 ng/ml). All patients provided written informed consent, consistent with local Research Ethics Board (IRB 00003835). DNA was extracted from either blood or saliva.
Among them, 15 samples from men with aggressive prostate cancer (defined by International Society of Urological Pathology ≥3) who had undergone radical prostatectomy at CHU Poitiers were analysed through the French International Cancer Genome Consortium (ICGC) program on prostate cancer [Tonon et al. 2019].

Sequencing
The 15 cases included in the French ICGC program on prostate cancer were sequenced at the study site.
Libraries were prepared using TruSeqDNA Sample Preparation Kit v2 (Illumina Inc, San Diego, CA, USA) and the KAPA Library Preparation kit (Kapa Biosystems, Roche, Bale, Switzerland). Whole-genome sequencing was performed using TruSeq SBS Kit v3-HS (Illumina Inc, San Diego, CA, USA), on a HiSeq2000 (Illumina Inc, San Diego, CA, USA) according to standard Illumina operation procedures. Primary data analysis, the image analysis, base calling, and quality scoring of the run were processed using the manufacturer's software Real Time Analysis (RTA 1.17.21.3; Illumina Inc, San Diego, CA, USA), followed by generation of FASTQ sequence files by CASAVA. Sequences were aligned to the human reference genome (GRCh37) using BWA (AUS Ref. 5). Alignments were refined using GATK v3.5.0 (McKenna et al., 2010; DePristo et al., 2011, Van der Auwera GA et al., 2013) and Picard tools (http://broadinstitute.github.io/picard/) to realign locally near indels, recalibrate bases quality scores, and remove duplicated reads.

The other 282 cases and 50 controls were sequenced at Mayo as part of ICPCG.

3. FHCRC (Fred Hutchinson Cancer Research Center)

Recruitment

Germline DNA purified from blood samples was available for sequencing from two sources: 1) The Prostate Cancer Genetic Research Study (PROGRESS), a family-based study of hereditary prostate cancer (HPC) [Stanford JL, et al., 2009]; and 2) Population-based case-control studies in Caucasian and African American residents of King County, WA [Stanford JL, et al., 1999][Agalliu I, et al., 2008]. Pedigree-based PROGRESS cases were enrolled from across North America, with families having three or more first-degree relatives diagnosed with prostate cancer, three or more generations of men diagnosed with prostate cancer, or two first-degree relatives diagnosed with prostate cancer before age 65 years. One affected individual with early onset prostate cancer or an aggressive phenotype was selected from each of 219 HPC families for sequencing. The population-based cases were ascertained through the Seattle-Puget Sound Surveillance, Epidemiology and End Results cancer registry, were 35-74 years of age at diagnosis in either 1993-1996 or in 2002-2005, and all had histologically confirmed adenocarcinoma of the prostate. Controls without a history of prostate cancer were identified through random digit telephone dialling and were frequency matched to cases on age. A total of 45 incident cases and 106 population controls was sequenced. All individuals included in these studies provided informed consent, and study protocols and procedures were approved by the FHCRC Institutional Review Board.

References:

Stanford JL, Wicklund K, McKnight B, Daling JR, Brawer MK. Vasectomy and risk of prostate cancer. *Cancer Epidemiol Bio Prev* 1999;8:881-886.

Agalliu I, Salinas CA, Hansten P, Ostrander EA, Stanford JL. Statin use and prostate cancer risk: a population-based case-control study. *Am J Epidemiol* 2008;168:250-260. PMCID:2585510

Stanford JL, FitzGerald LM, McDonnell SK, Carlson EE, McIntosh L, Deutsch K, Hood L, Ostrander EA, Schaid DJ. Dense genome-wide SNP linkage scan in 301 hereditary prostate cancer families identifies multiple regions with suggestive evidence for linkage. *Hum Mol Genet* 2009;18(10):1839-1848. PMCID:2671990

Sequencing
264 cases and 106 controls were sequenced at Mayo as part of ICPCG.

4. Finland / TAMPERE / TURKU

Recruitment
DNA samples were collected from two different sources. The Turku Prostate Cancer Consortium (Turku, Finland) provided both prostate cancer patient samples (n=29) and normal prostate tissue from bladder cancer patients (n=3). Samples were recruited at the Turku University Hospital. From the Tampere University Hospital (Tampere, Finland) n=78 prostate cancer patient samples were obtained (Schleutker et al., 2000). All subject samples were collected with written and signed informed consent.

Sequencing

107 cases and 3 controls sequenced at study site.
Both sample sets were sequenced on Illumina HiSeq (Illumina, Inc, San Diego, CA) platform. Alignment was performed based on the most recent Genome Analysis Toolkit (GATK) best practices protocol (McKenna et al., 2010; DePristo et al., 2011, Van der Auwera GA et al., 2013) in both sample sets.
115 cases and 67 controls were sequenced at Mayo as part of ICPCG.

5. Germany / ULM

Recruitment
The Familial Prostate Cancer Study is a case-control study. Cases were recruited in two different ways. Familial PrCa probands (index cases) were ascertained from all over Germany. They were advised by their attending physicians to contact the Clinic of Urology of Ulm. The positive family history was then verified by reviewing medical records or death certificates of family members. In each case, only one member of each family (e.g. the proband) was enrolled in the present study. Sporadic cases, who reported no relatives affected with PrCa, were almost exclusively collected at Ulm during their course of treatment (e.g. radical prostatectomy) in our Urology Clinic. The control group consists of 213 age-matched healthy men and 295 population controls of unknown disease status. [Luedeke et al. 2009]

Sequencing
191 cases and 127 controls were sequenced at Mayo as part of ICPCG.

6. ICR / UKGPCS

Recruitment
Cases identified through clinics at the Royal Marsden NHS Foundation Trust and nationwide NCRN hospitals
Controls from the ProtecT study were identified through invitation of subjects in the community.

Sequencing
1,990 cases and 1,336 controls were sequenced at this study site. Cases and controls selected from the UKGPCS and controls from the ProtecT study were sequenced as part of three distinct projects. 1,281 young-onset cases diagnosed ≤60 years old from the UKGPCS and 1,160 controls with no family history of PrCa or PSA <0.5ng/ml from the UKGPCS and ProtecT studies were sequenced using a custom 175 gene Agilent SureSelect XT bait library on an Illumina HiSeq 2000 instrument using v4 chemistry, aligned to the GRCh37 reference assembly using BWA 0.5.8 and genotyped using GATK v2.8-1, as described previously [Leongamornlert et al. Eur Urol, 2019]. 280 cases (139 aggressive: metastatic, diagnosed age<60; 141 non-aggressive: Gleason score <7, tumour stage T1-2b, no nodal spread or metastases, diagnosed age ≥60) from the UKGPCS were sequenced using Agilent SureSelectXT2 Human All Exon V5 baits on an Illumina HiSeq 2500 instrument using v4 chemistry, aligned to the GRCh37 reference assembly using BWA-MEM 0.7.10 and genotyped using GATK 3.5, as described previously [Mijuskovic et al. BJC, 2018]. 191 cases with family history of PrCa (≥2 relatives also diagnosed with PrCa) were sequenced using a 22 gene custom Agilent SureSelect bait library on an Illumina HiSeq 2000 instrument using v2 chemistry, aligned to the GRCh37 reference assembly using BWA 0.5.8 and genotyped using GATK v3, as described previously [Leongamornlert et al. BJC, 2014].

7. JHU

Sequencing

213 cases and 91 controls were sequenced at Mayo as part of ICPCG.

8. MAYO / ICPCG

[Schaid et al. 2020 in Press]

Recruitment

From reference provided [Schaid et al. 2020, in press]:

ICPCG is an international collaboration that has conducted numerous family-based studies.
A two-stage design was used […for the study described in the reference provided…]. In the first stage, men with PrCa were sampled from families that had at least three closely related male relatives with PrCa, and at least one affected member with DNA available.
All participants gave informed consent to utilize samples for research purposes, and the study was approved by the different Institutional Review Boards.

Sequencing

For all cases and controls sequenced at the Mayo Clinic, exome capture was performed using Agilent SureSelect Human All Exon 50Mb or V4+UTR capture kits. Samples were pooled post-capture and sequenced three to a lane using the Illumina HiSeq. All cases sequenced at Mayo Clinic were also genotyped on the Illumina Infinium OmniExpress-12 array. Externally sequenced samples were sequenced using a variety of capture kits including Illumina TruSeq and Nimblegen SeqCap. Quality control and bioinformatics processing are described in the supplemental material of [Schaid et al. 2020].

580 controls from the Mayo Biobank were included (Olson et al). These were at least 70 years old and free of PrCa.

391 cases were sequenced at Mayo as part of ICPCG.

9. MCCS

Recruitment

Participants in this study were identified from i) the Melbourne Collaborative Cohort Study (MCCS), ii) the Aggressive Prostate Cancer (APC) study, iii) the Risk Factors for Prostate Cancer Study and iv) the Early-Onset Prostate Cancer Family Study (AUS Ref. 1- 3).
Aggressive cases (n=787) were selected using the following criteria: PrCa as a cause of death (regardless of stage or Gleason score at diagnosis), or stage 4 (regardless of Gleason score) or stage 3 and Gleason score >= 8. Non-aggressive cases (n=770) were selected using the following criteria: stage 1 (T1/T2a) and Gleason score <=6 and age at diagnosis >=65 years; or stage 1 (T1/T2a) and Gleason score <=6 and age at diagnosis 55-64 years and >=10 years of follow-up; or stage 2 and Gleason score <=6, age at diagnosis >=65 years and >=10 years of follow-up. Germline DNA from these 1,557 participants was obtained from blood samples. Informed consent was obtained from all individual participants included in the study that was approved by the Human Research Ethics Committee of the Cancer Council Victoria.

Sequencing

1272 cases and 7 controls were sequenced at the study site.
Amplicon-based sequencing of the coding regions and proximal intron-exon junctions of 26 genes (including ATM (NM_000051.3), was performed using the Hi-Plex2 protocol (AUS Ref. 4Massively parallel sequencing (150 bp paired-end) was performed on the NextSeq550 platform (v2 chemistry, 2x150 bp) (Illumina, San Diego, CA, USA).

Paired-end reads were aligned to the reference genome (GRCh37) using bwa-mem 0.7.17 (AUS Ref. 5). Target coverage was then calculated using bedtools (AUS Ref. 6). Samples with ≥80% target bases covered at ≥ 50X sequencing depth were considered successfully sequenced.

274 cases and 53 controls were sequenced at Mayo as part of ICPCG.

10. MD Anderson

Recruitment

The ATM controls from MD Anderson were recruited in UT MD Anderson Cancer Center.

Sequencing

449 controls were sequenced at the study site.
The whole exome sequencing capture kit is Agilent SureSelect Clinical Research Exome v1. We sequenced each sample using Illumina HiSeq 4000 at an average raw depth of ~150X. The genome alignment, bam file clean, and genotypes calling was based on GATK's pipeline (including HaplotypeCaller v3.4 for joint genotype calling). Sample and variant level quality control was conducted using our QC tool, called XPAT, with default parameters. [Yu et al. 2018]

11. IPO Porto

Recruitment

Details regarding sample recruitment can be found in the following paper [Maia et al. 2015].
'This study comprised a total of 462 index cases (HPC samples) from families with early-onset and/or familial/hereditary PrCa, which were selected based on one of the following criteria: 1) PrCa diagnosis before the age of 56 or 2) PrCa diagnosis at any age with family history of the disease (up to fourth degree relatives) and at least one family member (the proband or a relative) with PrCa before the age of 66. Most patients were invited to participate in a study with the main purpose of identifying germline mutations associated with inherited PrCa predisposition, having as starting point all living patients registered at the North Region Cancer Registry (RORENO) with a PrCa diagnosis before the age of 66, whereas a minority of the families had been referred for genetic counselling due to early-onset or family history. All but two patients (one from the United Kingdom and another from Angola) had at least one Portuguese ancestor. No systematic PrCa screening program exists in the population under study, only opportunistic screening is offered to men over 50 years of age.'

DNA was extracted from peripheral blood leukocytes using the MagNA Pure LC DNA Isolation Kit—Large Volume (Roche Diagnostics GmbH, Penzberg, Germany) and whenever DNA was available from more than one affected relative per family, the youngest at the time of diagnosis was considered as the index case.

Sequencing

479 cases were sequenced at the study site.
IPO-Porto samples were enriched for a Custom panel using the SureSelect chemistry and run in a HiSeq platform. Submitted vcf files for the ATM project were obtained with the NextGENe software (v2.4.1; Softgenetics). Some ATM data was already reported using a different panel and pipeline [Paulo et al 2018]

12. TASPRAC

Recruitment

FitzGerald, L.M., Raspin, K., Marthick, J.R., Field, M.A., Malley, R.C., Thomson, R.J., Blackburn, N.B., Banks, A., Charlesworth, J.C., Donovan, S., et al. (2017). Impact of the G84E variant on HOXB13 gene and protein expression in formalin-fixed, paraffin-embedded prostate tumours. Sci Rep 7, 17778.

Whole genome sequencing data was available for 50 individuals, and eighteen cases and eight controls were selected from two studies; The Tasmanian Prostate Cancer Familial Study and The Tasmanian Case-Control Study. The Tasmanian Prostate Cancer Familial Study is comprised of a rare collection of 73 PCa families from the founder population of Tasmania. The number of affected men in these families range from five to over 140, and include up to five affected brothers and multiple father/son and uncle/nephew pairs. DNA samples from blood or saliva have been collected for 326 affected men and 474 male and female relatives.

The Tasmanian Case-Control Prostate Cancer Study is a population-based study, which includes blood or saliva samples from 504 cases and 332 controls. Cases were identified from the Tasmanian Cancer Registry (TCR) and considered eligible for this study if they were diagnosed under the age of 75 between the years 1996 and 2005. Controls were selected at random from the Tasmanian electoral roll and frequency matched by five-year age groups to the cases. Controls are periodically checked against the TCR for PCa diagnosis, most recently 2019.

Ethics approval for both cohort studies was obtained from the Human Research Ethics Committee Tasmania, Australia and is ongoing (ethics approval number H0017040). Written informed consent was obtained from all participating individuals.

Sequencing

Whole-genome and whole-exome sequencing data was generated for several individuals at the Illumina Genome Network, USA, on the HiSeq 2500 s or at the Kinghorn Centre for Clinical Genomics, Australia, on the Illumina HiSeq XTM Ten platform using the TruSeq Nano library preparation. Additional samples were sequenced at the Australian Genome Research Facility (AGRF) on the Illumina NovaSeq 6000.

13. UTAH

Recruitment

All prostate cancer cases were drawn from the set of sampled prostate cancer cases identified in the Utah Cancer Registry and belonging to extended Utah high-risk pedigrees who were sequenced as part of prostate cancer predisposition gene projects. The controls were selected as controls for a study of high-risk melanoma pedigrees requiring: no diagnosis of cancer in self or first degree relatives and no melanoma cases within relatives out to second cousins.

Sequencing

147 cases had sequencing performed at various sites including the Huntsman Cancer Institute Sequencing Core, Mayo Sequencing Core, and Nanthealth as part of the Heritage 1K project; 399 controls were sequenced at MD Anderson.

Sample Recruitment and Sequencing References

Schleutker J, Matikainen M, Smith J, Koivisto P, Baffoe-Bonnie A, Kainu T et al. A 325 genetic epidemiological study of hereditary prostate cancer (HPC) in Finland: 326 frequent HPCX linkage in families with late-onset disease. Clin Cancer Res 2000; 327 6(12): 4810-4815.

McKenna, A. et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing nextgeneration DNA sequencing data. Genome research 20(9), pp. 1297–1303.
DePristo, M. et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature Genetics 43, pp. 491–498.

Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine 350 A et al. From FastQ data to high confidence variant calls: the Genome Analysis 351 Toolkit best practices pipeline. Curr Protoc Bioinformatics 2013; 43: 11 10 11-33.

Schaid, McDonnell, Fitzgerald et al. Two-Stage Study of Familial Prostate Cancer by Whole-Exome Sequencing and Custom Capture Identifies 10 Novel Genes Associated with Risk of Prostate Cancer (2020 in Press?)

Tonon L et al. Mutational Profile of Aggressive, Localised Prostate Cancer from African Caribbean Men Versus European Ancestry Men. Eur Urol. 2019 Jan;75(1):11-15. doi: 10.1016/j.eururo.2018.08.026."

Thomas et al. 2008. Nat Genet. "Multiple loci identified in a genome-wide association study of prostate cancer." 2008 Mar;40(3):310-5. doi: 10.1038/ng.91. Epub 2008 Feb 10.

CAPS et al 2008 "Cumulative association of five genetic variants with prostate cancer." N Engl J Med. 2008 Feb 28;358(9):910-9

Agalliu, Stanford et al. "Statin use and risk of prostate cancer: results from a population-based epidemiologic study." Am J Epidemiol. 2008 Aug 1;168(3):250-60. doi: 10.1093/aje/kwn141. Epub 2008 Jun 12.

Stanford et al. "Vasectomy and risk of prostate cancer." Cancer Epidemiol Biomarkers Prev. 1999 Oct;8(10):881-6.

Luedeke M, et al. Predisposition for TMPRSS2-ERG fusion in prostate cancer by variants in DNA repair genes. Cancer Epidemiol Biomarkers Prev. 2009 Nov;18(11):3030-5.

AUS Ref. 1. Milne RL, Fletcher AS, MacInnis RJ, Hodge AM, Hopkins AH, Bassett JK, Bruinsma FJ, Lynch BM, Dugue PA, Jayasekara H, Brinkman MT, Popowski LV, et al. Cohort Profile: The Melbourne Collaborative Cohort Study (Health 2020). Int J Epidemiol 2017;46:1757-57i.

AUS Ref. 2. Papa NP, MacInnis RJ, English DR, Bolton D, Davis ID, Lawrentschuk N, Millar JL, Pedersen J, Severi G, Southey MC, Hopper JL, Giles GG. Ejaculatory frequency and the risk of aggressive prostate cancer: Findings from a case-control study. Urol Oncol 2017;35:530 e7-30 e13.

AUS Ref. 3. Giles GG, Severi G, Sinclair R, English DR, McCredie MR, Johnson W, Boyle P, Hopper JL. Androgenetic alopecia and prostate cancer: findings from an Australian case-control study. Cancer Epidemiol Biomarkers Prev 2002;11:549-53.

AUS Ref. 4. Hammet F, Mahmood K, Green TR, Nguyen-Dumont T, Southey MC, Buchanan DD, Lonie A, Nathanson KL, Couch FJ, Pope BJ, Park DJ. Hi-Plex2: a simple and robust approach to targeted sequencing-based genetic screening. Biotechniques 2019.

AUS Ref. 5. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754-60.

AUS Ref. 6. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010;26:841-2.

AUS Ref. 7. Lai Z, Markovets A, Ahdesmaki M, Chapman B, Hofmann O, McEwen R, Johnson J, Dougherty B, Barrett JC, Dry JR. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. Nucleic Acids Res 2016;44:e108.

AUS Ref. 8. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, Flicek P, Cunningham F. The Ensembl Variant Effect Predictor. Genome Biol 2016;17:122.

AUS Ref. 9. Paila U, Chapman BA, Kirchner R, Quinlan AR. GEMINI: integrative exploration of genetic variation and genome annotations. PLoS Comput Biol 2013;9:e1003153.

AUS Ref. 10. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285-91.

AUS Ref. 11. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405-24.

AUS Ref. 12. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, Maglott DR. ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Research 2014;42:D980-5

Yu, Yao, et al. "XPAT: a toolkit to conduct cross-platform association studies with heterogeneous sequencing datasets." Nucleic acids research 46.6 (2018): e32-e32.

Maia et al. "Identification of Two Novel HOXB13 Germline Mutations in Portuguese Prostate Cancer Patients." PLoS One. 2015 Jul 15;10(7):e0132728. doi: 10.1371/journal.pone.0132728. eCollection 2015.

Paulo et al. "Targeted next generation sequencing identifies functionally deleterious germline mutations in novel genes in early-onset/familial prostate cancer." PLoS Genet. 2018 Apr 16;14(4):e1007355. doi: 10.1371/journal.pgen.1007355. eCollection 2018 Apr.

FitzGerald et al. "Impact of the G84E variant on HOXB13 gene and protein expression in formalin-fixed, paraffin-embedded prostate tumours." Sci Rep. 2017 Dec 19;7(1):17778. doi: 10.1038/s41598-017-18217-w.

Olson et al. “The Mayo Clinic Biobank: A Building Block for Individualized Medicine.” Mayo Clinic Proceedings 88, 952-962 (2013)

Field, M. A. et al. “Reducing the search space for causal genetic variants with VASP.” Bioinformatics. 31, 2377–2379 (2015).