

Missing Environmental Data

Number of Pairs Odds Ratio Confidence Interval

XPD Lys751Gln

original data set	202	1.90	(1.20 – 3.00)
-------------------	-----	------	-----------------

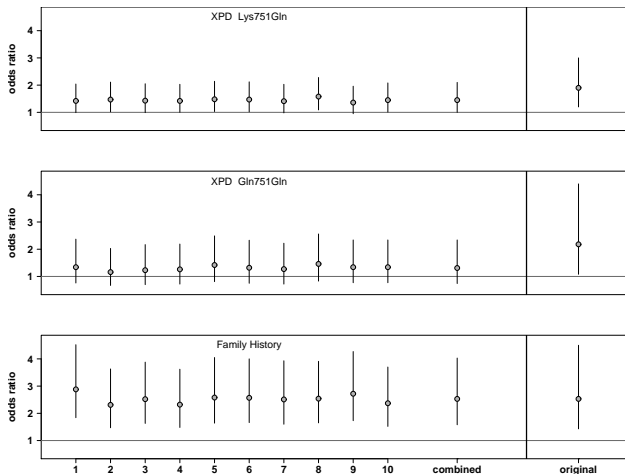
XPD Gln751Gln

original data set	202	2.18	(1.08 – 4.40)
-------------------	-----	------	-----------------

Positive Family History

original data set	202	2.53	(1.43 – 4.50)
-------------------	-----	------	-----------------

Missing Environmental Data



The missing data were imputed using decision trees.

Dai et al (2006) *Imputation methods to improve inference...* Genetic Epidemiology, 30(8): 690-702.

Missing Environmental Data

Number of Pairs Odds Ratio Confidence Interval

XPD Lys751Gln

original data set	202	1.90	(1.20 – 3.00)
-------------------	-----	------	-----------------

XPD Gln751Gln

original data set	202	2.18	(1.08 – 4.40)
-------------------	-----	------	-----------------

Positive Family History

original data set	202	2.53	(1.43 – 4.50)
-------------------	-----	------	-----------------

Missing Environmental Data

Number of Pairs Odds Ratio Confidence Interval

XPD Lys751Gln

original data set	202	1.90	(1.20 – 3.00)
multiple imputations	321	1.45	(1.00 – 2.10)

XPD Gln751Gln

original data set	202	2.18	(1.08 – 4.40)
multiple imputations	321	1.31	(0.74 – 2.34)

Positive Family History

original data set	202	2.53	(1.43 – 4.50)
multiple imputations	321	2.53	(1.58 – 4.03)

Missing Environmental Data

	Family History not complete				Family History complete			
	AA	AC	CC	na	AA	AC	CC	na
	raw numbers							
case	43	54	5	5	61	121	25	7
control	35	57	12	3	90	102	22	0
	percentages							
case	40.2	50.5	4.7	4.7	28.5	56.5	11.7	3.3
control	32.7	53.3	11.2	2.8	42.1	47.7	10.3	0.0

Brewster et al (2006) *Polymorphisms of the DNA repair genes XPD and...* Breast Cancer Res Treat, 95(1): 73-80.