Supplementary Appendices

These appendices have been provided by the authors to give readers additional information about their work.

Supplement to: Different independent associations of depression and anxiety with survival in patients with cancer

Appendix A: Social deprivation scores

Social deprivation was calculated using the Scottish Index of Multiple Deprivation (SIMD) 2009.

This provides a relative measure of deprivation based on indicators from 7 domains – income, employment, health, education, access, housing and crime by dividing Scotland into 6,505 small geographical areas or divisions (datazones) and ranking these from the most deprived (ranked 1) to the least deprived (ranked 6,505).

Reference

Office of the Chief Statistician. Scottish Index of Multiple Deprivation 2009 Technical Report: Scottish Government; 2011.

Appendix B: Cancer groupings

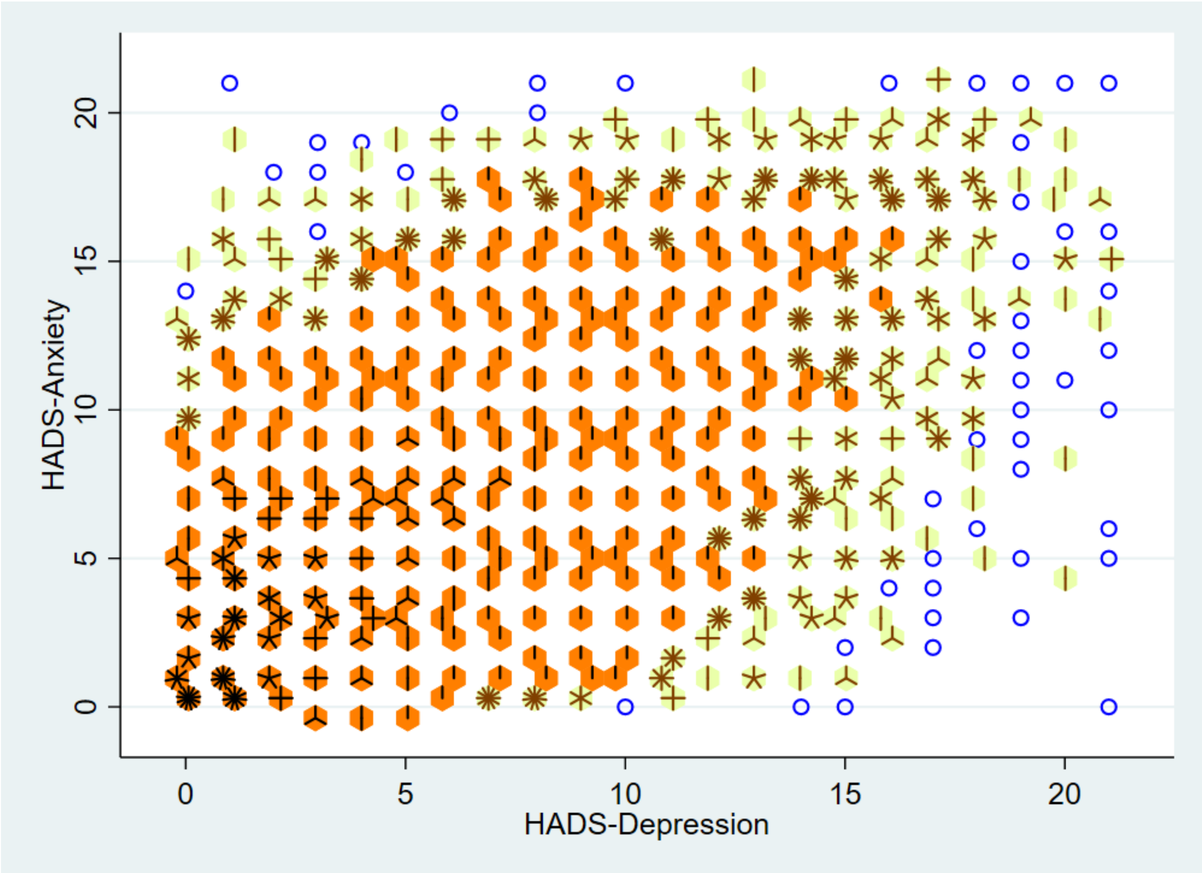
Grouping	ICD-10 codes*	Diagnoses
Breast	C500	Malignant neoplasm of nipple and areola
	C501	Malignant neoplasm of central portion of breast
	C502	Malignant neoplasm of upper-inner quadrant of breast
	C503	Malignant neoplasm of lower-inner quadrant of breast
	C504	Malignant neoplasm of upper-outer quadrant of breast
	C505	Malignant neoplasm of lower-outer quadrant of breast
	C506	Malignant neoplasm of axillary tail of breast
	C508	Malignant neoplasm, overlapping lesion of breast
	C509	Malignant neoplasm of breast, unspecified
Lung	C340	Malignant neoplasm of main bronchus
	C341	Malignant neoplasm of upper lobe, bronchus or lung
	C342	Malignant neoplasm of middle lobe, bronchus or lung
	C343	Malignant neoplasm of lower lobe, bronchus or lung
	C348	Malignant neoplasm of overlap les of bronchus & lung
	C349	Malignant neoplasm of bronchus or lung, unspecified
	C450	Mesothelioma of pleura
	C451	Mesothelioma of peritoneum
	C452	Mesothelioma of pericardium
	C457	Mesothelioma of other sites
	C459	Mesothelioma, unspecified

Colorectal	C182	Malignant neoplasm of ascending colon
	C183	Malignant neoplasm of hepatic flexure
	C184	Malignant neoplasm of transverse colon
	C185	Malignant neoplasm of splenic flexure
	C186	Malignant neoplasm of descending colon
	C187	Malignant neoplasm of sigmoid colon
	C188	Malignant neoplasm overlapping lesion of colon
	C189	Malignant neoplasm of colon, unspecified
	C19X	Malignant neoplasm of rectosigmoid junction
	C20X	Malignant neoplasm of rectum
Gynaecological	C481	Malignant neoplasm of specified parts of peritoneum
	C482	Malignant neoplasm of peritoneum, unspecified
	C510	Malignant neoplasm of labium majus
	C511	Malignant neoplasm of labium minus
	C512	Malignant neoplasm of clitoris
	C518	Malignant neoplasm of overlapping lesion of vulva
	C519	Malignant neoplasm of vulva, unspecified
	C52X	Malignant neoplasm of vagina
	C530	Malignant neoplasm of endocervix
	C531	Malignant neoplasm of exocervix
	C538	Malignant neoplasm, overlapping lesion of cervix uteri
	C539	Malignant neoplasm of cervix uteri, unspecified

	C540	Malignant neoplasm of isthmus uteri
	C541	Malignant neoplasm of endometrium
	C542	Malignant neoplasm of myometrium
	C543	Malignant neoplasm of fundus uteri
	C548	Malignant neoplasm overlapping lesion of corpus uteri
	C549	Malignant neoplasm of corpus uteri, unspecified
	C55X	Malignant neoplasm of uterus, part unspecified
	C56X	Malignant neoplasm of ovary
	C570	Malignant neoplasm of fallopian tube
	C571	Malignant neoplasm of broad ligament
	C572	Malignant neoplasm of round ligament
	C573	Malignant neoplasm of parametrium
	C574	Malignant neoplasm of uterine adnexa, unspecified
	C577	Malignant neoplasm of other specified female genital organs
	C578	Malignant neoplasm, overlapping lesion female genital organs
	C579	Malignant neoplasm of female genital organ, unspecified
	C763	Malignant neoplasm of pelvis
Prostate	C61X	Malignant neoplasm of prostate

*International Classification of Diseases 10th edition

Appendix C: Sunflower plot of HADS-Anxiety versus HADS-Depression scores



A blue circle represents one patient.

A light hexagon with: one vertical line (two radii) represents two patients, three radii represents three patients...12 radii represents 12 patients.

A dark hexagon with one radius represents 13 to 50 patients, two radii represents 51 to 100 patients etc.

Appendix D: The handling of missing data

We imputed missing data using the substantive model compatible fully conditional specification (SMCFCS) method, an extension of the more common fully conditional specification (FCS). This method imputes missing data across multiple covariates using an imputation model that is fully compatible with our substantive (intended) analysis model. For our study this may be more appropriate than FCS because we have specified non-linear associations and interactions in our regression model, which cannot be completely specified in FCS imputation. The imputation models were specified with the substantive model variables plus extra variables that, over and above those in the substantive model both (1) predict the values of the missing data and (2) predict the probability of these data being missing. We determined this using logistic regression on the complete data where the outcome is the variable in question (1) or a 0/1 indicator of its missingness (2). We add to the imputation model those variables that were statistically significant at the 5% level for both. We included tumour grade and/or clinical stage marker (as available) for each cancer type where there was evidence that these were associated with both survival and missingness. We also did not include any of these extra variables in the substantive models for survival since we wanted to use a common set of covariates throughout in order to make the cancer-specific results comparable.

<u>References</u>

Bartlett J. SMCFCS: Multiple Imputation of Covariates by Substantive Model Compatible Fully Conditional Specification. https://CRAN.R-project.org/package=smcfcs2016.

Appendix E: Primary causes of death for patients included in the analysis

Total number of deaths	**5884**
Cancer	**5386 (91.5%)**
Lung	2708
Breast	816
Gynaecological	753
Colorectal	665
Prostate	214
Other cancer	230
Circulatory	**218 (3.7%)**
Ischaemic heart disease (including acute myocardial infarction)	111
Aortic aneurysm	12
Cardiac arrhythmia	7
Cerebrovascular disease	59
Heart failure	6
Other	23
Respiratory	**93 (1.6%)**
Chronic obstructive airways disease	56
Respiratory infection	25
Interstitial pulmonary disease	6
Other	6
Gastro-intestinal	**41 (0.7%)**
Infection (non-respiratory)	**23 (0.4%)**
Injury, poisoning and external causes	**21 (0.4%)**
Fall	10
Fracture	3

Poisoning (accidental)	4
Road traffic accident	1
Drowning (undetermined intent)	1
Shooting (intentional self-harm)	1
Exposure to excessive cold	1
Neurological	**15 (0.3%)**
Renal	**12 (0.2%)**
Haematological	**12 (0.2%)**
Hepatic, pancreatic or biliary	**12 (0.2%)**
Endocrine, nutritional or metabolic	**8 (0.1%)**
Mental and behavioural	**7 (0.1%)**
Dementia	6
Alcohol dependence	1
Other	**4 (0.1%)**
Unknown	**32 (0.5%)**

Appendix F: Parameter estimates for the models relating HADS-D and HADS-A to mortality hazard

For HADS (either HADS-D or HADS-A), with knots at $k_i, i = 1,..,4$ define the restricted cubic spline variables to be created as $HADS_1$, $HADS_2$ and $HADS_3$ as follows.

$$\mathrm{HADS}_1 = \mathrm{HADS}$$

$$\mathrm{HADS}_{i+1} = \frac{(\mathrm{HADS} - k_i)_+^3 - (k_4 - k_3)^{-1}\{(\mathrm{HADS} - k_3)_+^3(k_4 - k_i) - (\mathrm{HADS} - k_4)_+^3(k_3 - k_i)\}}{(k_4 - k_1)^2}$$

for $i = 1, 2$ where $(u)_+ = u, \text{if } u > 0 \text{ or } 0, \text{if } u \le 0.$

The estimated log hazard ratios, p-values and 95% confidence intervals (CI) for each of the cancer groupings are as follows:

Prostate cancer

Predictor variable	**Log hazard ratio**	**p-value**	**95% CI**
Unadjusted (n=1531 for HADS-D and HADS-A models)			
HADS-D_1	0.46	p=0.008	0.12, 0.80
HADS-D_2	-1.82	p=0.122	-4.14, 0.49
HADS-D_3	3.52	p=0.167	-1.48, 8.51
HADS-A_1	0.07	p=0.532	-0.16, 0.30
HADS-A_2	-0.04	p=0.966	-1.76, 1.68
HADS-A_3	0.05	p=0.976	-3.12, 3.21
Adjusted (n=1531)			
HADS-D_1	0.49	p=0.006	0.14, 0.84
HADS-D_2	-1.92	p=0.108	-4.25, 0.42
HADS-D_3	3.66	p=0.154	-1.38, 8.71
HADS-A_1	-0.08	p=0.494	-0.32, 0.16
HADS-A_2	0.43	p=0.633	-1.33, 2.18
HADS-A_3	-0.70	p=0.670	-3.93, 2.52

Colorectal cancer – males

Predictor variable	Log hazard ratio	p-value	95% CI
Unadjusted (n=1573 for HADS-D and HADS-A models)			
HADS-D_1	0.09	p=0.348	-0.10, 0.28
HADS-D_2	0.15	p=0.826	-1.21, 1.51
HADS-D_3	-0.40	p=0.752	-2.87, 2.07
HADS-A_1	0.03	p=0.765	-0.15, 0.20
HADS-A_2	0.50	p=0.449	-0.79, 1.78
HADS-A_3	-1.03	p=0.392	-3.38, 1.32
Adjusted (n=1573)			
HADS-D_1	0.08	p=0.408	-0.11, 0.28
HADS-D_2	0.10	p=0.890	-1.28, 1.48
HADS-D_3	-0.26	p=0.838	-2.77, 2.25
HADS-A_1	-0.02	p=0.821	-0.20, 0.16
HADS-A_2	0.47	p=0.484	-0.84, 1.78
HADS-A_3	-0.96	p=0.431	-3.36, 1.43

Lung cancer – males

Predictor variable	Log hazard ratio	p-value	95% CI
Unadjusted (n=2299 for HADS-D and HADS-A models)			
HADS-D_1	0.10	p=0.002	0.04, 0.17
HADS-D_2	-0.11	p=0.498	-0.41, 0.20
HADS-D_3	0.17	p=0.601	-0.47, 0.81
HADS-A_1	0.05	p=0.175	-0.02, 0.11
HADS-A_2	-0.03	p=0.858	-0.32, 0.27
HADS-A_3	0.06	p=0.895	-0.79, 0.90
Adjusted (n=2299)			
HADS-D_1	0.11	p=0.001	0.04, 0.18
HADS-D_2	-0.12	p=0.433	-0.44, 0.19
HADS-D_3	0.20	p=0.540	-0.45, 0.85
HADS-A_1	-0.03	p=0.458	-0.10, 0.04
HADS-A_2	0.09	p=0.552	-0.21, 0.39
HADS-A_3	-0.24	p=0.584	-1.10, 0.62

Breast cancer

Predictor variable	Log hazard ratio	p-value	95% CI
Unadjusted (n=8467 for HADS-D and HADS-A models)			
HADS-D_1	1.36	p<0.001	1.15, 1.60
HADS-D_2	0.22	p=0.011	0.07, 0.71
HADS-D_3	24.41	p=0.013	1.95, 306.04
HADS-A_1	1.08	p=0.077	0.99, 1.18
HADS-A_2	0.74	p=0.084	0.53, 1.04
HADS-A_3	2.49	p=0.068	0.94, 6.63
Adjusted (n=8467)			
HADS-D_1	1.41	p<.001	1.19, 1.67
HADS-D_2	0.24	p=0.018	0.07, 0.78
HADS-D_3	19.18	p=0.024	1.49, 247.41
HADS-A_1	1.00	p=0.935	0.92, 1.10
HADS-A_2	0.78	p=0.156	0.56, 1.10
HADS-A_3	2.10	p=0.144	0.78, 5.65

Gynaecological cancer

Predictor variable	Log hazard ratio	p-value	95% CI
Unadjusted (n=2910 for HADS-D and HADS-A models)			
HADS-D_1	0.35	p<.001	0.18, 0.51
HADS-D_2	-1.65	p=0.004	-2.79, -0.51
HADS-D_3	2.85	p=0.007	0.79, 4.92
HADS-A_1	0.06	p=0.189	-0.03, 0.15
HADS-A_2	-0.07	p=0.718	-0.44, 0.30
HADS-A_3	0.04	p=0.946	-1.05, 1.12
Adjusted (n=2910)			
HADS-D_1	0.38	p<.001	0.21, 0.55
HADS-D_2	-1.69	p=0.004	-2.84, -0.54
HADS-D_3	2.94	p=0.006	0.86, 5.02
HADS-A_1	-0.02	p=0.666	-0.11, 0.07
HADS-A_2	0.01	p=0.973	-0.37, 0.38
HADS-A_3	-0.15	p=0.788	-1.24, 0.94

Colorectal cancer – females

Predictor variable	Log hazard ratio	p-value	95% CI
Unadjusted (n=1154 for HADS-D and HADS-A models)			
HADS-D_1	0.21	p=0.131	-0.06, 0.48
HADS-D_2	-0.60	p=0.530	-2.48, 1.27
HADS-D_3	0.90	p=0.605	-2.50, 4.29
HADS-A_1	-0.16	p=0.073	-0.33, 0.01
HADS-A_2	0.78	p=0.055	-0.02, 1.58
HADS-A_3	-1.66	p=0.062	-3.40, 0.08
Adjusted (n=1154)			
HADS-D_1	0.30	p=0.038	0.02, 0.57
HADS-D_2	-1.02	p=0.290	-2.91, 0.87
HADS-D_3	1.63	p=0.351	-1.79, 5.05
HADS-A_1	-0.23	p=0.010	-0.41, -0.06
HADS-A_2	0.89	p=0.031	0.08, 1.70
HADS-A_3	-1.85	p=0.039	-3.61, -0.09

Lung cancer – females

Predictor variable	Log hazard ratio	p-value	95% CI
Unadjusted (n=2041 for HADS-D and HADS-A models)			
HADS-D_1	0.11	p=0.003	0.04, 0.18
HADS-D_2	-0.25	p=0.142	-0.59, 0.08
HADS-D_3	0.49	p=0.170	-0.21, 1.20
HADS-A_1	0.07	p=0.041	0.00, 0.13
HADS-A_2	-0.17	p=0.114	-0.39, 0.04
HADS-A_3	0.47	p=0.143	-0.16, 1.10
Adjusted (n=2041)			
HADS-D_1	0.12	p=0.002	0.04, 0.19
HADS-D_2	-0.25	p=0.154	-0.59, 0.09
HADS-D_3	0.48	p=0.184	-0.23, 1.19
HADS-A_1	0.02	p=0.529	-0.05, 0.09
HADS-A_2	-0.14	p=0.209	-0.36, 0.08
HADS-A_3	0.39	p=0.231	-0.25, 1.02

Appendix G: Plots of the estimated hazard ratios for depression at three levels of anxiety (left panel) and estimated hazard ratios for anxiety at three levels of depression (right panel) relative to a participant with breast cancer with HADS-D/HADS-A equal to zero.

Predicted hazard ratios for depression severity (at HADS-A=0, 5 and 10) and anxiety severity (at HADS-D=0, 5 and 10) for those with breast cancer are shown below. There was no evidence that a model with all the products of the linear term for depression and cubic spline terms for anxiety and vice-versa fitted better than a simpler model with just the product of the linear terms. Therefore, we present the results from this simpler model. The figure shows a difference in the shapes of the relationship between anxiety and hazard of mortality (with a change in the association with increasing anxiety, from close to zero to protective, as depression increases). Differences in the shape of the relationship between depression and hazard of mortality across levels of anxiety were less stark, with a slightly stronger estimated relationship at lower levels of anxiety.