

Appendix S1

SUPPLEMENTARY MATERIALS

Immunosuppressive drugs

Immunosuppressive drugs included the Danish health care classification system codes (Sundhedsvæsenets Klassifikationssystem, SKS) ML04* and BOHJ* (excluding antibody treatments) and Anatomical Therapeutic Chemical Classification System (ATC) codes L04*. Prednisolone included the SKS codes (MA07EA01, MH02AB06, MC05AA04, MD07AA03, MD07XA02, MR01AD02, MS01CB02, MS01BA04, MS02BA03, MS01CB02, MS03BA02, MR01AD52, MA01AC54, MD07BA01, MD07CA03, MS01CA02, MV03AB05, BWAHA159, BWAHA171, BWAHA134, BWAHA137, BWAHA168) and ATC codes (H02AB06 and A07AE01). Methotrexate included the SKS codes (BWAHA115, BWAHA122, BWAHA130, BWAHA117, BWAHA161, BWAHA185, BWAHA219, BKHE8A) and ATC codes (L01BA01 and L04AX03). Azathioprine included SKS codes (BWHB83) and ATC codes (L04AX01). Ciclosporin included SKS codes (BOHJ20) and ATC codes (L04AD01 and S01XA18). Imiquimod included SKS (MD06BB10 and BNHG01A) and ATC (D06BB10). Photodynamic therapy (PDT) included SKS (BWHD and BNGC1) and ATC (L01XD03 and L01XD04). Fluorouracil cream included the ATC (L01BC02).

Determining the number of clusters.

To determine the number of clusters, three different evaluation metrics were used: the elbow method, the silhouette score and cluster stability. The elbow method was estimated on the whole dataset while the silhouette score was estimated from a 50,000 random sample of the population due to high computation time of calculating the distance matrix (Figure S1 and S2). The cluster stability was assessed via bootstrapping (20 iterations), with the mean Adjusted Rand Index (ARI) calculated for each number of clusters (Figure S3). Based on these three-evaluation metrics, excluding choices with 2 and 3 clusters, the optimal number of clusters was determined to be 6 or 7. Seven clusters were selected over six clusters due to the finer distinction between those with high comorbidity score and moderate comorbidity score. The 6-cluster choice and corresponding figures (S4 and S5) and tables (Table S1–S3) are presented below. This solution reproduces similar cluster to the 7-cluster choice, especially cluster 3–5, which are almost identical between both cluster choices. Furthermore, cluster 1 in the 6-cluster choice was young, high-income individuals with high proportion of BCCs, similarly to cluster 1 in the 7-cluster solution, however, did differentiate by having slightly more males than females in contrast to cluster 1 in the 7-cluster choice. The 6-cluster choice only has one highly comorbid cluster and a much larger cluster 2.

The 2- and 3-cluster choice

The 2 and 3 cluster choices were not chosen as the broad split did not represent the heterogeneous keratinocyte carcinoma patient group well and the clinical interpretation between the few clusters were less refined and clinically meaningful than the 6 or 7 cluster choice. Briefly, the 2-cluster choice mainly differentiated on younger and high-income patients versus older and lower-income individuals. The 3-cluster choice mainly differentiated again on a younger, high-income patient cluster versus two older patient clusters which differentiated from each other on number of prior skin cancers before index.

Figure S1: Showing the loss with respect to the number of clusters.

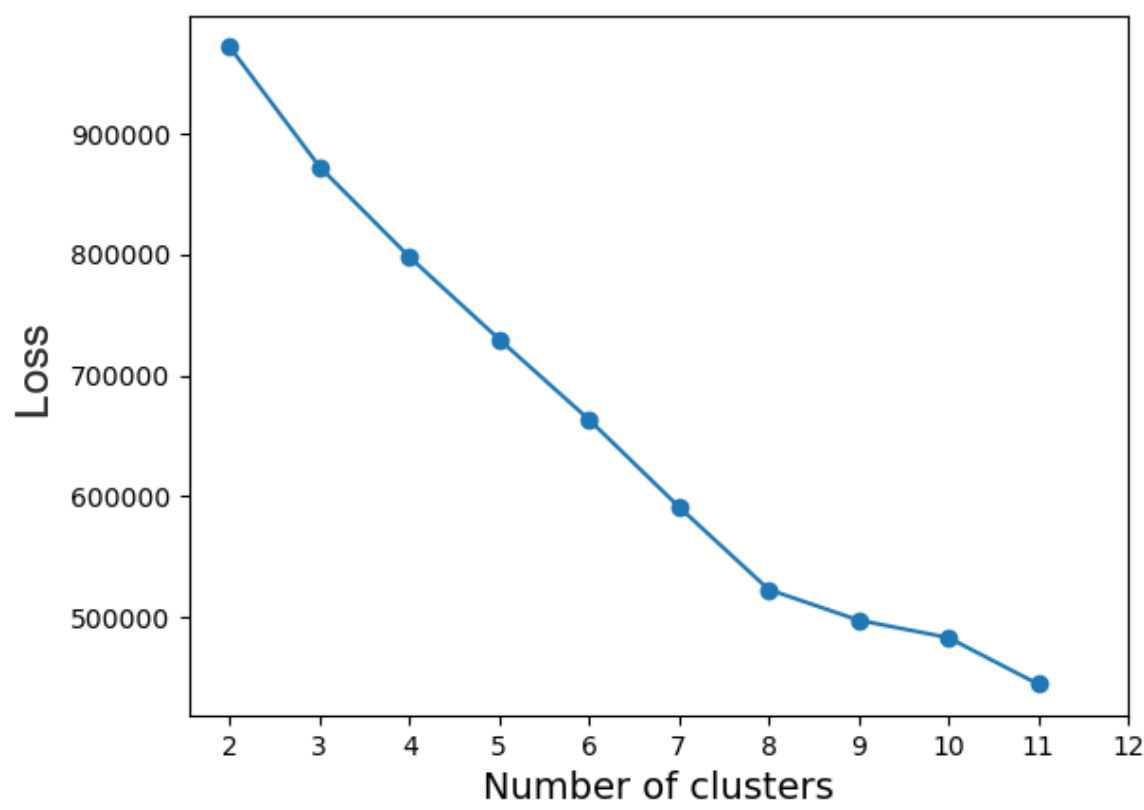


Figure S2: Showing the silhouette score with respect to the number of clusters estimated from a 50,000 random sample of the population.

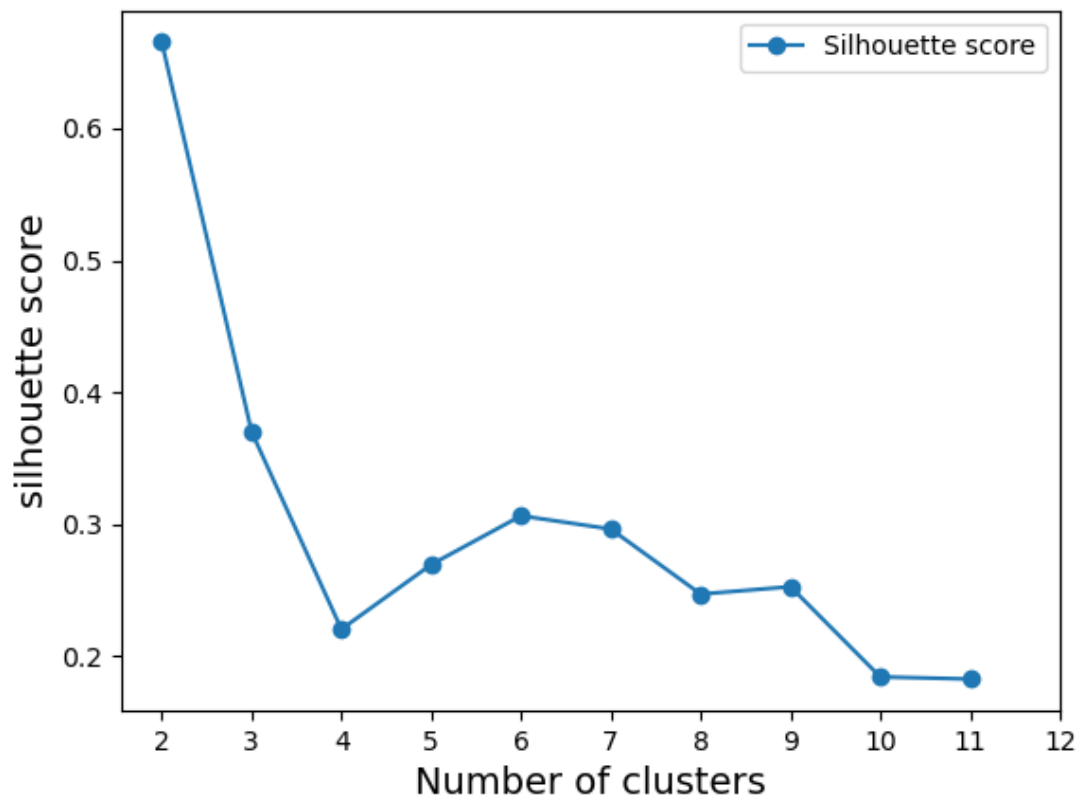


Figure S3: Showing the cluster stability with respect to the number of clusters using adjusted rand index as cluster stability score.

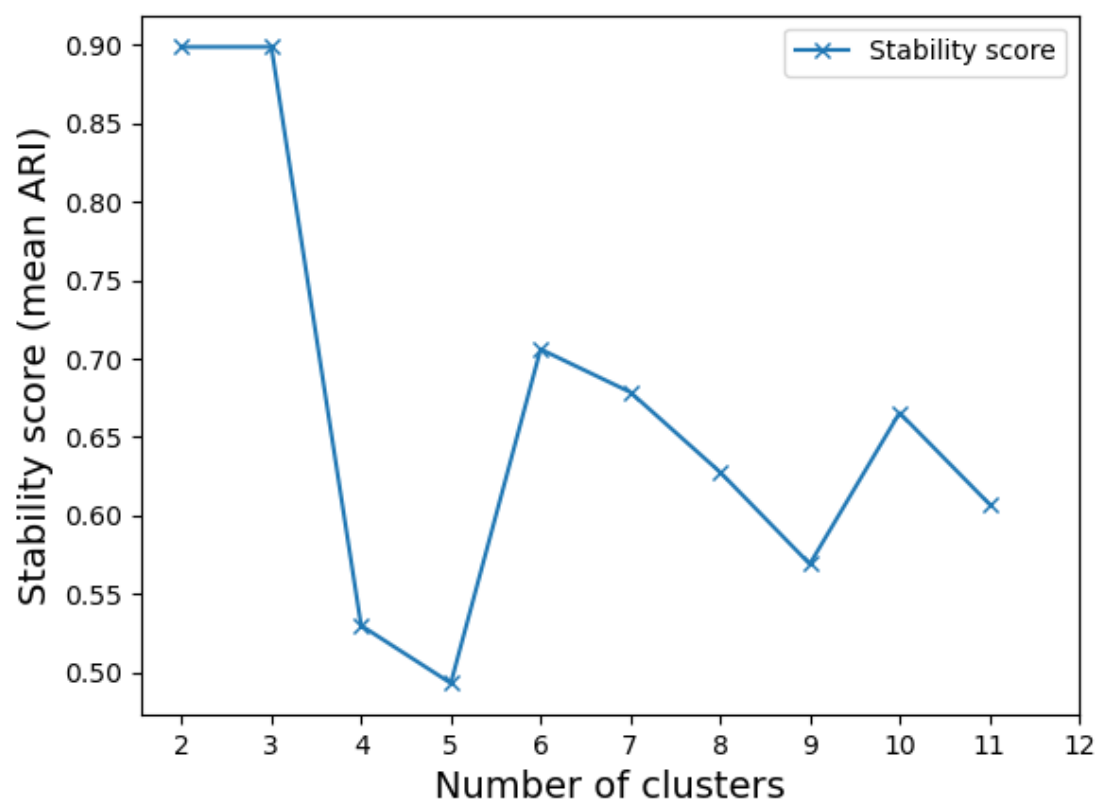


Figure S4: Histogram over percentage of different comorbidities in each cluster for the 6-cluster solution.

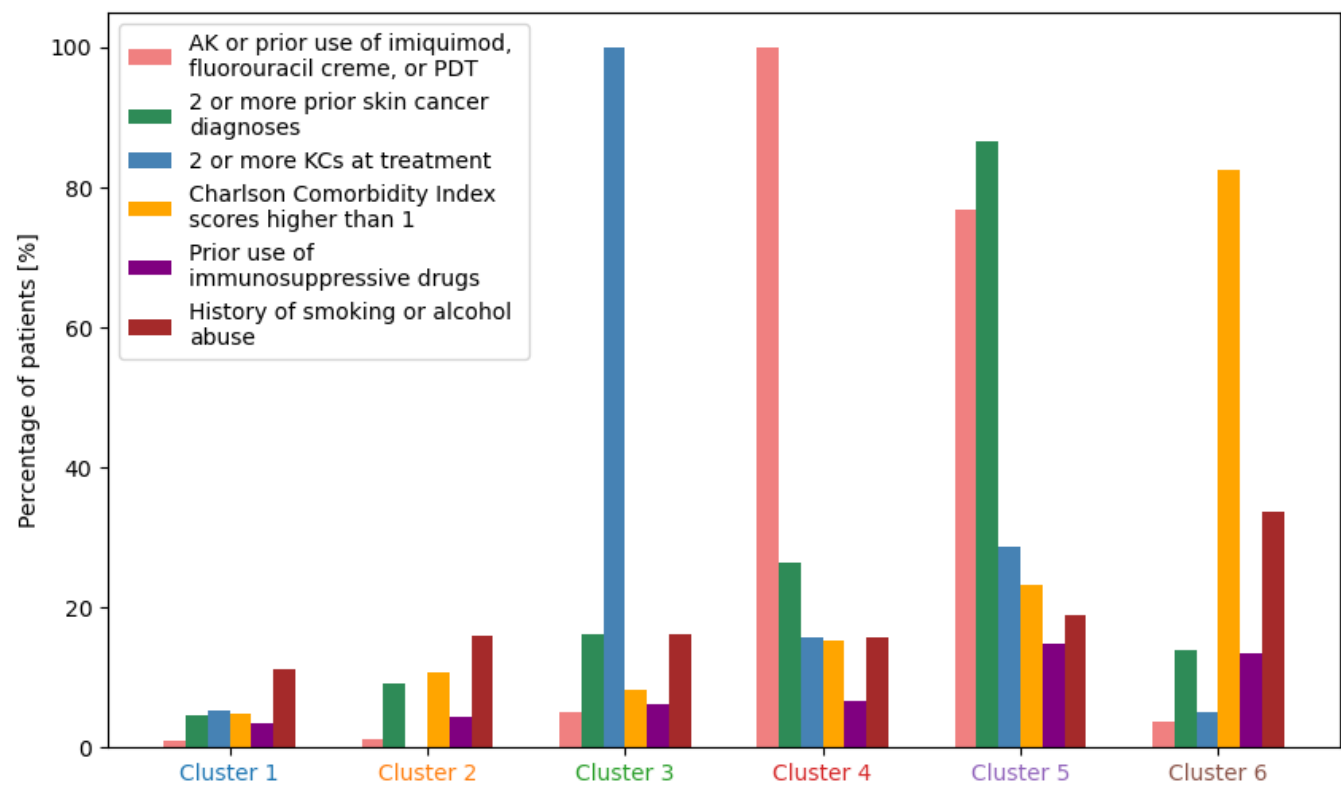


Figure S5: Heatmap over percentage point difference between each cluster compared with the rest of the KC population for the 6-cluster solution.

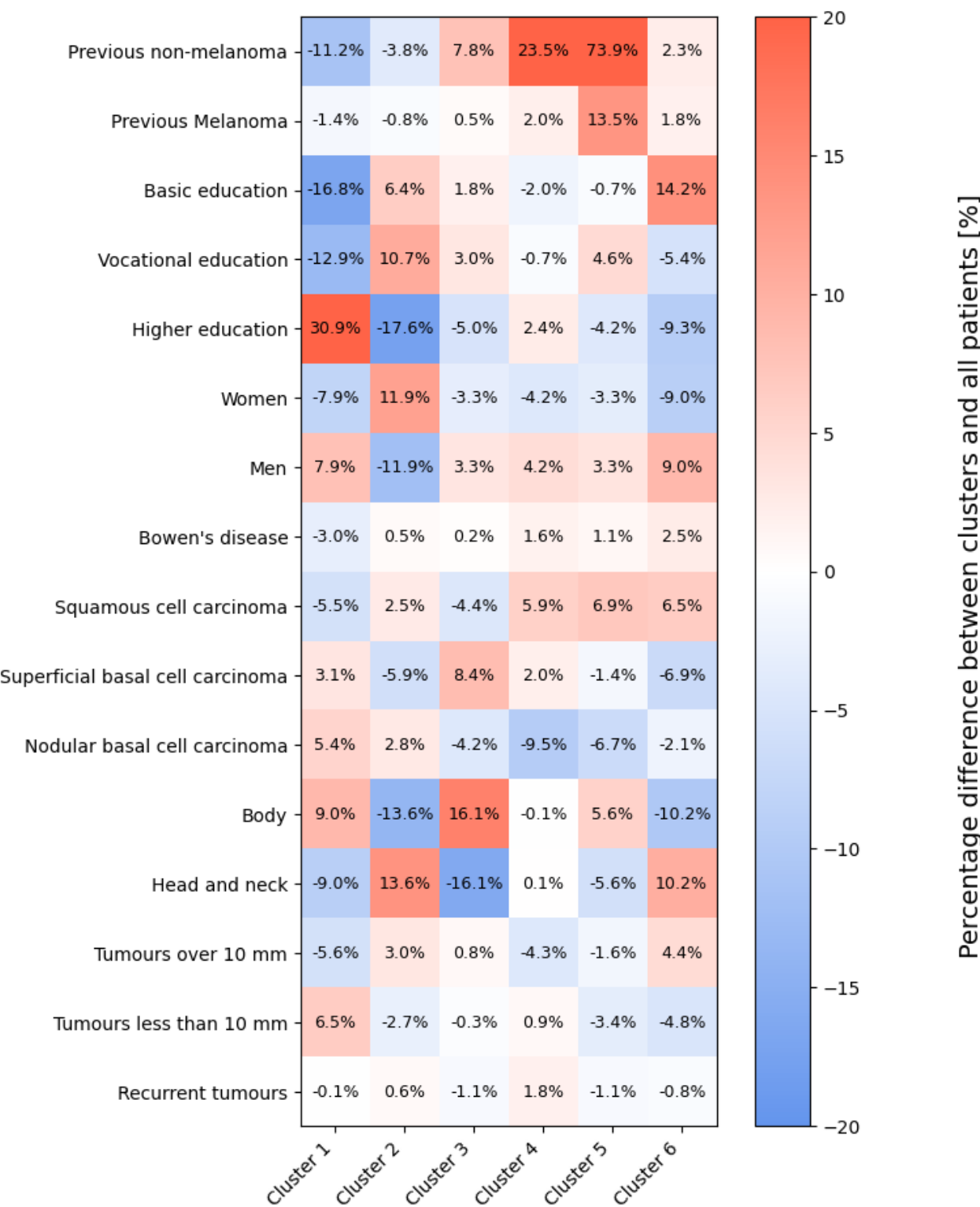


Table S1: Patient characteristics for the 6-cluster choice

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Overall
Total number of patients	25590 (24.0%)	48994 (46.0%)	11082 (10.4%)	7993 (7.5%)	567 (0.5%)	12264 (11.5%)	106490 (100%)
Sex, n(%)							
Male	13779 (53.8)	20279 (41.4)	5627 (50.8)	4138 (51.8)	290 (51.1)	6842 (55.8)	50955 (47.8)
Female	11811 (46.2)	28715 (58.6)	5455 (49.2)	3855 (48.2)	277 (48.9)	5422 (44.2)	55535 (52.2)
Age at treatment, median	56.4 (49.1–63.6)	75.1 (69.9–80.9)	74.5 (68.0–80.8)	75.3 (68.4–81.3)	76.6 (70.8–83.0)	77.9 (72.1–83.7)	72.5 (63.5–79.4)
Disposable Income [DKK]	788632	319826	348485	387331	338749	324127	387567
at index date, median	1574183–	1237336–	1252306–	1274282–	1254429–	1238801–	1266729–
Number of hospital visits	11 (6–20)	17 (10–27)	18 (9–30)	21 (12–34)	39 (23–58.5)	48 (34–65)	18 (9–31)
Highest achieved education, n(%)							
Higher education	14019 (54.8)	10679 (21.8)	2968 (26.8)	2682 (33.6)	154 (27.2)	2831 (23.1)	33333 (31.3)
Basic education	3916 (15.3)	15449 (31.5)	3288 (29.7)	2096 (26.2)	155 (27.3)	4982 (40.6)	29886 (28.1)
Vocational education	7469 (29.2)	21941 (44.8)	4618 (41.7)	3064 (38.3)	247 (43.6)	4200 (34.2)	41539 (39.0)
Unknown education	186 (0.7)	925 (1.9)	208 (1.9)	151 (1.9)	11 (1.9)	251 (2.0)	1732 (1.6)

Table S2: Tumour characteristics for the 6-cluster choice

Table S3: Comorbidities and medication prior inclusion date for 6-cluster choice

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Overall
Medication, n(%)							
Immunosuppressive drugs	880 (3.4)	2108 (4.3)	673 (6.1)	530 (6.6)	84 (14.8)	1638 (13.4)	5913 (5.6)
Prednisolone	3278	9379 (19.1)	2312 (20.9)	1748 (21.9)	168 (29.6)	4696 (38.3)	21581
Superficial basal cell carcinoma	5518	1104 (14.7)	6409 (64.7)	1847 (19.3)	133 (19.6)	1339 (11.3)	22741 (18.1)
Ciclosporin	47 (0.2)	92 (0.2)	39 (0.4)	25 (0.3)	16 (2.8)	77 (0.6)	296 (0.3)
Azathioprine	310 (1.2)	420 (0.9)	144 (1.3)	148 (1.9)	31 (5.5)	293 (2.4)	1346 (1.3)
Squamous cell carcinoma	811 (3.0)	4349 (8.9)	998 (13.9)	1187 (12.8)	113 (14.2)	1097 (9.5)	9125 (8.3)
Methotrexate	456 (1.8)	1444 (2.9)	407 (3.7)	311 (3.9)	54 (9.5)	914 (7.5)	3586 (3.4)
Medication related to actinic keratosis or skin cancer, n(%)							
Number of tumours at treatment, Imiquimod	1 (1-1)	1 (0-1)	2 (3-3)	6 (18-28)	25 (1-24.3)	185 (1-1.5)	747 (1-20)
Fluorouracil cream	5 (0.0)	0 (0.0)	8 (0.1)	523 (6.5)	16 (2.8)	8 (0.1)	560 (0.5)
Tumour size, n(%)							
PDT	5 (0.0)	0 (0.0)	27 (0.2)	1433 (17.9)	265 (46.7)	30 (0.2)	1760 (1.7)
Summation of Imiquimod, fluorouracil	0 (0-0)	0 (0-0)	0 (0-0)	1 (1-1)	1 (0-1)	0 (0-0)	0 (0-0)
10-15 mm	3021	6734 (13.7)	3343 (12.9)	992 (10.7)	92 (11.6)	1935 (15.0)	16117 (12.9)
16-20 mm	1460 (5.5)	4502 (9.2)	2205 (8.0)	600 (6.5)	62 (7.0)	1202 (10.0)	10222 (8.2)
Total number of skin cancer prior,	0 (0-0)	0 (0-0)	0 (0-0)	0 (0-2)	8 (4-18)	0 (0-0)	0 (0-0)
Anatomical locations, n(%)							
Melanoma	587 (2.3)	1435 (2.9)	423 (3.8)	414 (5.2)	95 (16.8)	600 (4.9)	3554 (3.3)
Head and neck (Face, ear, scalp or neck)	1345	3230 (6.6)	1051 (3.8)	1031 (13.2)	335 (25.8)	600 (4.9)	6600 (5.1)
Non-melanoma skin cancer	1637 (6.4)	6290 (12.8)	2424 (21.9)	2926 (36.6)	501 (88.4)	2072 (16.9)	15850
Actinic keratosis	9554	197 (0.8)	544 (1.1)	768 (1.7)	1077 (13.5)	222 (1.5)	5225 (2.2)
Comorbidities							
Modified Charlson Comorbidity Index, scale of neck	1839 (6.8)	4219 (8.6)	2030 (7.9)	969 (10.4)	94 (12.8)	1237 (9.9)	10721 (8.4)
Diabetes, n(%)	167 (0.7)	701 (1.4)	86 (0.8)	153 (1.9)	16 (2.8)	1653 (13.5)	2776 (2.6)
Myocardial infarction, n(%)	307 (1.2)	629 (2.5)	187 (1.2)	248 (3.1)	14 (2.5)	1056 (10.4)	3624 (3.1)
Congestive heart failure, n(%)	121 (0.5)	813 (1.7)	125 (1.1)	184 (2.3)	23 (4.1)	1794 (14.6)	3060 (2.9)
Peripheral vascular disease, n(%)	190 (0.7)	1206 (2.5)	128 (1.2)	430 (5.4)	19 (3.4)	488 (4.1)	3675 (3.5)
Cancer, n(%)	839 (3.3)	3401 (6.9)	629 (5.7)	720 (9.0)	59 (10.4)	5364 (43.7)	11012
Metastases, n(%)	1836 (6.8)	3913 (8.0)	2350 (21.3)	851 (10.7)	83 (15.6)	687 (5.5)	8517 (8.1)
Kidney disease, n(%)	5 (0.0)	0 (0.0)	<3	>33	16 (2.8)	1472 (12.0)	1531 (1.4)
Leukaemia, n(%)	69 (0.3)	150 (0.3)	45 (0.4)	75 (0.9)	12 (2.1)	488 (4.0)	839 (0.8)
Lymphoma, n(%)	86 (0.3)	156 (0.3)	44 (0.4)	50 (0.6)	6 (1.1)	553 (4.5)	895 (0.8)
Smoking ever, n(%)	2224 (8.7)	6607 (13.5)	1491 (13.5)	1056 (13.2)	94 (16.6)	3677 (30.0)	15149
Alcohol abuse, n(%)	831 (3.2)	1793 (3.7)	439 (4.0)	271 (3.4)	22 (3.9)	903 (7.4)	4259 (4.0)