

Using longitudinal linked electronic health records to investigate lifetime obesity prevalence in systemic anticancer therapy patients in England



Perletta V,¹ Kulkarni S,¹ Wang Z,² Tomlinson JW,³ Hippisley-Cox J,⁴ Collins GS,⁵ Clift AK,⁶ Dodwell D,² Lord SR¹

¹ Department of Oncology, University of Oxford; ² Nuffield Department of Population Health, University of Oxford; ³ Oxford Centre for Diabetes, Endocrinology & Metabolism and NIHR Oxford Biomedical Research Centre, University of Oxford; ⁴ Wolfson Institute of Population Health, Queen Mary University of London; ⁵ Department of Applied Health Sciences and NIHR Birmingham Biomedical Research Centre, University of Birmingham; ⁶ Department of Surgery & Cancer, Imperial College London.

BACKGROUND

- Obesity is an established risk factor for several cancer types, but its impact on cancer treatment outcomes remains uncertain.^{1,2}
- Most studies rely on body mass index (BMI) measured at treatment initiation, potentially underestimating lifetime exposure to obesity.^{2,3}
- How obesity prevalence at first treatment compares to lifetime obesity has not been well characterised.²

STUDY AIM

To use longitudinal linked electronic health records (EHR) with repeated BMI to estimate at first treatment and lifetime obesity prevalence in 58,756 individuals receiving systemic anticancer therapy in England in 2013-2023.

METHODS

DATA SOURCES

QResearch primary care records, containing a representative subset of the English population, linked with four national sources of EHR (Fig. 1).⁴⁻⁸

STUDY SAMPLE

Adults with a first diagnosis of one of thirteen cancer types recorded in NCRAS, HES or QResearch and first systemic therapy recorded in SACT during 1-Jan-2013 to 31-May-2023.

OBESITY DEFINITION

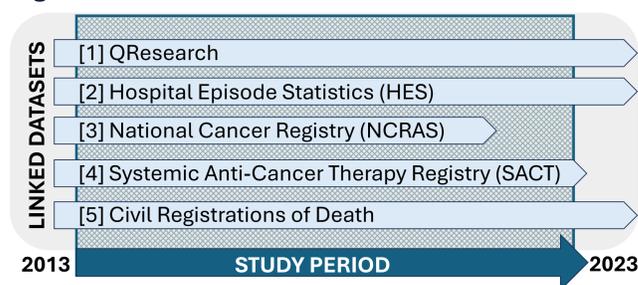
Body mass index (BMI) ≥ 30 kg/m² recorded on the date of first systemic treatment in SACT and/or from historic records in QResearch.

STATISTICAL ANALYSIS

Observed at first treatment and lifetime obesity prevalence were estimated as a percentage with 95% confidence interval (% [95%CI]).

Age-standardised prevalence was calculated by direct method (age groups 20-44, 45-54, 55-64, 65-74 and ≥ 75 years) using the mid-year 2022 population statistics for England.⁹

Fig 1. Data sources



References

1. Lauby-Secretan et al. N Engl J Med. 2016 Aug 25;375(8):794-8.
2. Lennon et al. Curr Oncol Rep. 2016 Sep;18(9):56.
3. Vaidya et al. JAMA Netw Open. 2022. 5(10): p. e2234445.
4. Hippisley-Cox et al. Inform Prim Care. 2004;12(1):49-50.
5. Hospital Episode Statistics. <https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics> [accessed 21-Jan-2026].
6. Henson et al. Int J Epidemiol. 2020 Feb 1;49(1):16-16h.
7. Bright et al. Int J Epidemiol. 2020 Feb 1;49(1):15-15L.
8. Civil Registrations of Death. <https://digital.nhs.uk/services/data-access-request-service/dars-products-and-services/data-set-catalogue/civil-registrations-of-death> [accessed 21-Jan-2026].
9. Population estimates for England and Wales: mid-2022. <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/populationestimatesforenglandandwales/mid2022> [accessed 21-Jan-2026].

Acknowledgements

We acknowledge the contribution of EMIS practices who contribute to QResearch® and NHS England, Public Health England and the Office for National Statistics for providing the HES, Cancer Registry, SACT and Death Registry datasets used in this analysis. Funding for grant IIG_FULL_2023_007 was obtained from World Cancer Research Fund (WCRF), as part of the World Cancer Research Fund International grant programme

Email: victoria.perletta@oncology.ox.ac.uk

INTERPRETATION & IMPACT

KEY FINDINGS

- Approximately 1 in 4 cancer patients were living with obesity at treatment initiation, and half had a history of obesity.
- BMI at treatment initiation vastly underestimated lifetime exposure to obesity, with implications for research into the relationship between obesity and cancer outcomes.
- Failure to account for historic obesity introduces exposure misclassification, likely limiting the accuracy of prognostic cancer models.
- Linking general practice records with national cancer registries enables finer assessment of individual risk factors, including life-course trajectories, and is key to the realisation of personalised cancer medicine.

LIMITATIONS

- BMI can be an imprecise measure of excess adiposity.
- While only a minority of patients had missing BMI, this disproportionately affected certain cancer types.
- Results may also be specific to England.

FUTURE RESEARCH

Future research will address how lifetime obesity associates with cancer survival, taking into account potential confounders in the exposure-outcome relationship, as well as the development of a prognostic model aimed at predicting and improving outcomes for patients across all levels of BMI.

RESULTS

Study sample characteristics

- We identified 67,049 eligible systemic therapy patients and 58,756 (87.6%) with complete BMI were included. BMI completeness ranged from 65.9% complete in prostate cancer to 96.6% in pancreatic cancer.
- Breast cancer was the most frequent cancer in the sample, and uterine cancer the least frequent (Table 1).
- Median age at first treatment was 66.8 years overall, and ranged from 57.5 years (breast) to 72.5 years (prostate) across cancer types.
- While 53.2% patients overall were female, the majority of patients with non-sex specific cancers were male.

Table 1. Study sample characteristics at first treatment

	SACT patients n (% col)	Age at first treatment median (IQR)	Female sex n (% row)
All patients	58,756 (100)	66.8 (17.1)	31,275 (53.2)
Breast	13,845 (23.6)	57.5 (18.5)	13,755 (99.3)
Bowel	11,009 (18.7)	66.5 (15.9)	4,613 (41.9)
Lung	9,217 (15.7)	69.6 (13.1)	4,240 (46.0)
Non-Hodgkin Lymphoma	5,469 (9.3)	69.3 (17.9)	2,351 (43.0)
Prostate	4,751 (8.1)	72.5 (11.0)	--
Gastroesophageal	4,470 (7.6)	68.5 (14.4)	1,166 (26.1)
Pancreas	2,335 (4.0)	68.7 (14.4)	1,050 (45.0)
Ovarian	2,125 (3.6)	67.1 (18.4)	2,125 (100)
Bladder	1,890 (3.2)	72.0 (12.9)	454 (24.0)
Malignant melanoma	1,581 (2.7)	68.4 (18.6)	671 (42.4)
Kidney	1,272 (2.2)	67.4 (15.6)	398 (31.3)
Hepatocellular	434 (0.7)	68.0 (13.0)	94 (21.7)
Uterine	358 (0.6)	67.3 (14.2)	358 (100)

Abbreviations: n – number of patients, SACT – systemic anticancer therapy, % col – column percentage, % row – row percentage, IQR – interquartile range

Obesity prevalence

- Age-standardised obesity prevalence at first treatment was 27.3% [26.7-27.8%] for all cancers, ranging from 14.1% [11.5-16.7%] for pancreatic cancer to 36.2% [28.4-43.9%] for uterine cancer (Table 2).
- Lifetime obesity prevalence was 55.0% [54.6-55.4%] overall and exceeded 50% for all cancer types.
- Lifetime prevalence was higher than at first treatment prevalence by 29.1 [95%CI: 28.7-29.5] percentage points (pp) overall, ranging from a 19.9 [19.2-20.5] pp increase for breast cancer to a 43.8 [41.8-45.8] pp increase for pancreatic cancer (Fig 2).

Fig 2. Observed at first treatment vs. lifetime obesity prevalence

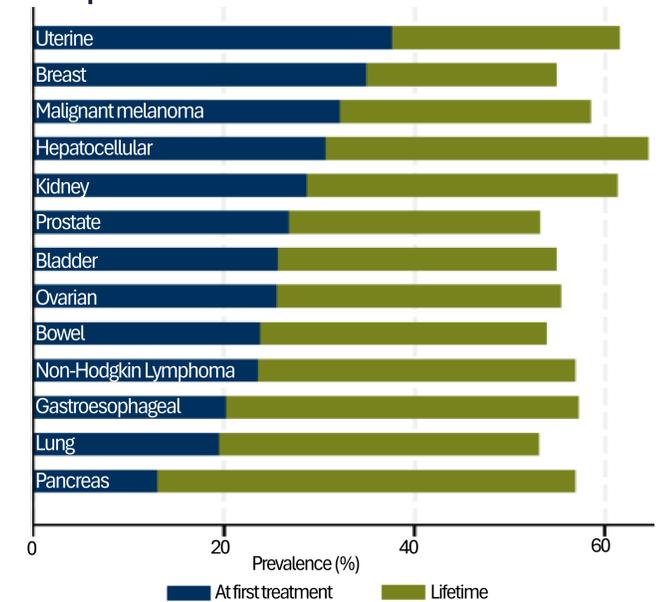


Table 2. Observed and age-standardised at first treatment vs. lifetime obesity prevalence

	At first treatment obesity prevalence			Lifetime obesity prevalence	
	n	Observed, % [95%CI]	Age-standardised, % [95%CI]	n	% [95%CI]
All patients	15,220	25.9% [25.5-26.3%]	27.3% [26.7-27.8%]	32,326	55.0% [54.6-55.4%]
Uterine	135	37.7% [32.7-42.7%]	36.2% [28.4-43.9%]	220	61.5% [56.4-66.5%]
Malignant melanoma	509	32.2% [29.9-34.5%]	34.4% [31.0-37.8%]	924	58.4% [56.0-60.9%]
Breast	4,838	34.9% [34.1-35.7%]	33.8% [33.0-34.7%]	7,591	54.8% [54.0-55.7%]
Kidney	366	28.8% [26.3-31.3%]	30.8% [26.8-34.8%]	778	61.2% [58.5-63.8%]
Prostate	1,275	26.8% [25.6-28.1%]	28.0% [22.7-33.3%]	2,524	53.1% [51.7-54.5%]
Ovarian	545	25.6% [23.8-27.5%]	27.4% [24.7-30.1%]	1,176	55.3% [53.2-57.5%]
Bladder	485	25.7% [23.7-27.6%]	26.7% [22.3-31.1%]	1,037	54.9% [52.6-57.1%]
Hepatocellular	133	30.6% [26.3-35.0%]	25.7% [18.6-32.8%]	280	64.5% [60.0-69.0%]
Non-Hodgkin Lymphoma	1,292	23.6% [22.5-24.7%]	25.3% [23.5-27.0%]	3,108	56.8% [55.5-58.1%]
Bowel	2,628	23.9% [23.1-24.7%]	24.4% [23.2-25.6%]	5,927	53.8% [52.9-54.8%]
Lung	1,804	19.6% [18.8-20.4%]	20.1% [18.4-21.7%]	4,880	52.9% [51.9-54.0%]
Gastroesophageal	906	20.3% [19.1-21.4%]	19.7% [17.7-21.7%]	2,554	57.1% [55.7-58.6%]
Pancreas	304	13.0% [11.7-14.4%]	14.1% [11.5-16.7%]	1,327	56.8% [54.8-58.8%]

Abbreviations: n – number of patients with record of obesity (body mass index ≥ 30 kg/m²), 95%CI – 95% confidence interval

Notes: Age-standardised prevalence was calculated by direct method (age groups 20-44, 45-54, 55-64, 65-74 and ≥ 75 years) using the mid-year 2022 population statistics for England.