Identifying progressive cancers in head and neck cancer patients using routine health datasets in England

James Charnock[1], Joanna Pethick[1], Rebecca Girdler[2], Vivian Mak[2], Hannah McConnell[3], Rachel White[3], Zi Wei Liu[4], Matthew Fittall[5], Matt Williams[6]

[1]National Cancer Registration and Analysis Service (NCRAS)/ Macmillan Cancer Support, [2]NCRAS, [3]Macmillan Cancer Support, [4]Great Ormond Street Hospital, [5]Guy's and St Thomas' Hospital, [6]Imperial College Hospital

In partnership with WE ARE MACMILLAN. CANCER SUPPORT

INTRODUCTION

Historically there has been limited national information on progressive cancer (recurrence, subsequent primary cancers or metastatic disease). Macmillan Cancer Support and Public Health England's National Cancer Registration and Analysis Service (NCRAS) are collaborating with academics and clinicians by using patient-level national cancer datasets to build our understanding of progressive cancer at a population level. We have developed an algorithm to identify patients in England with primary head and neck cancers who have subsequently developed progressive cancer, using routinely collected electronic healthcare data.

METHODS

Records of patients diagnosed with head and neck cancer (ICD10 codes: C00-C14, C30-C32) in 2013 in England, were extracted from Public Health England's cancer registration data (Cancer Analysis System – CAS) using the criteria detailed in **Figure 1**. These records were linked to the Hospital Episode Statistics (HES) data, Systemic Anti-Cancer Therapy (SACT) data and the Radiotherapy Dataset (RTDS).

Patients who received a radical cancer treatment after 90 days from a previous radical treatment were identified as having a recurrence as per *Ricketts et al* 2014[1]. Treatments within the 90 day window were classed as part of a treatment plan and not considered a recurrence (**Figure 2 – Patient A**).

Alongside treatment information being used to potentially identify a recurrence there are also a group of patients who could suffer a recurrence but never come into the health system again for treatment. For this reason, a patient who died over 90 days after the final radical treatment was considered as a patient of interest, as they potentially fall into this group (**Figure 2 – Patient B**).

The underlying cause of death of deceased head and neck patients as depicted in **Figure 2 – Patient B** was investigated, shown in **Figure 4**.

The results of the algorithm will be validated in two phases. We first validate this against the data in both the head and neck audit (DAHNO) and the cancer waiting times (CWT) datasets. The second round of validation uses patient records data held by selected hospital trusts, namely Imperial College and North Cumbria University Hospitals, to identify patients with a recurrence.
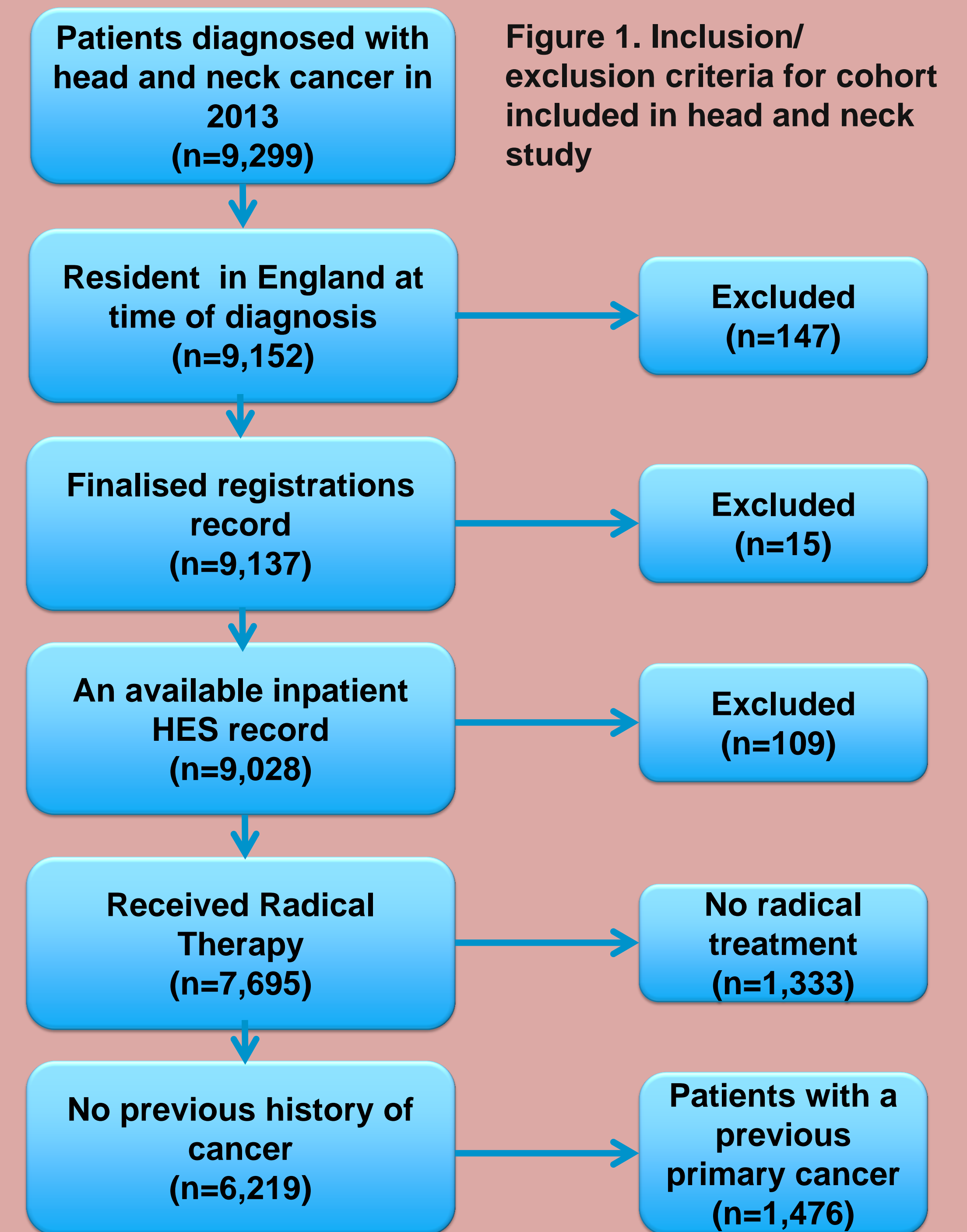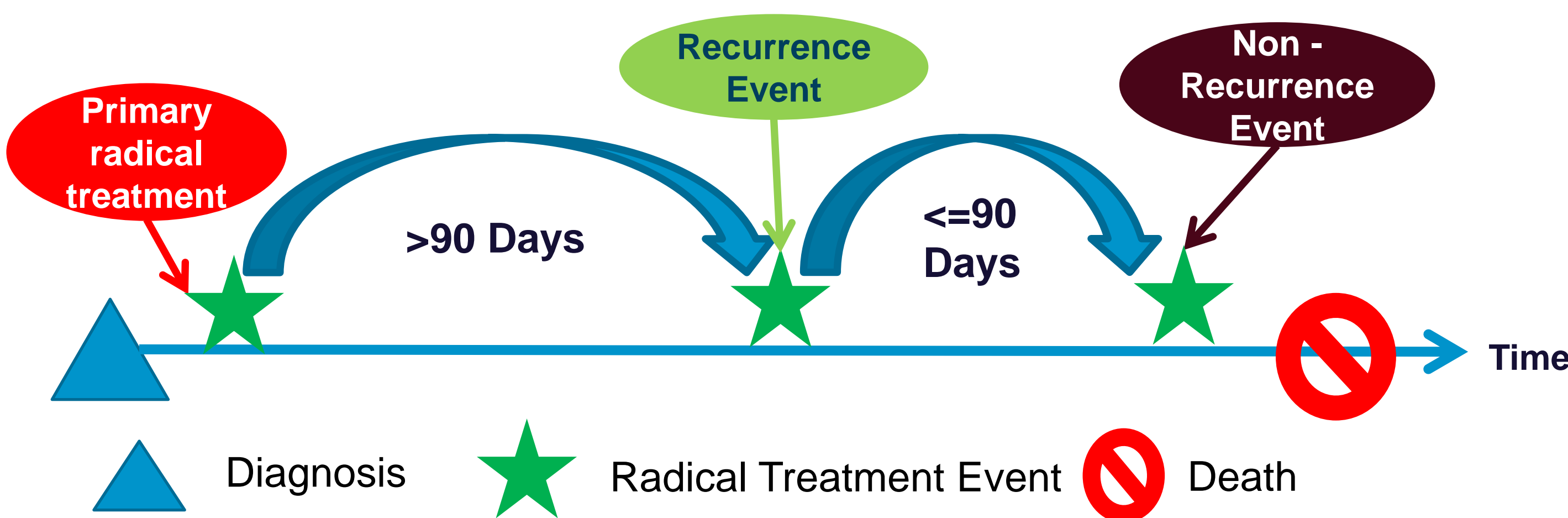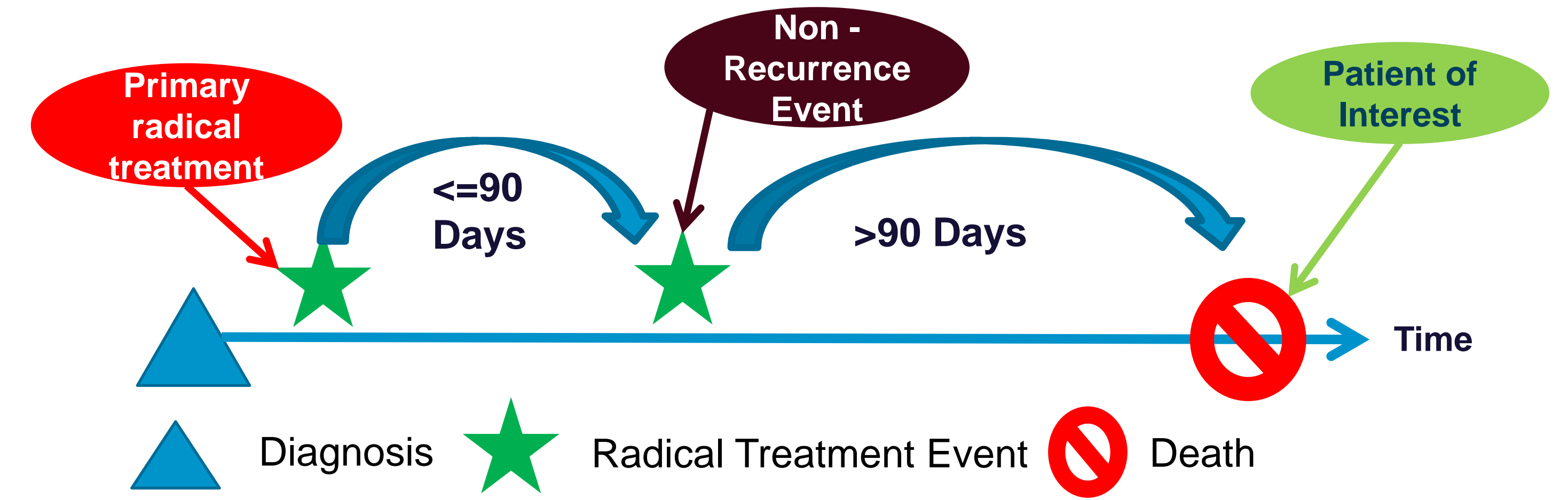
Figure 1. Inclusion/ exclusion criteria for cohort included in head and neck study

- Patients diagnosed with head and neck cancer in 2013 (n=9,299)
- Resident in England at time of diagnosis (n=9,152) → Excluded (n=147)
- Finalised registrations record (n=9,137) → Excluded (n=15)
- An available inpatient HES record (n=9,028) → Excluded (n=109)
- Received Radical Therapy (n=7,695) → No radical treatment (n=1,333)
- No previous history of cancer (n=6,219) → Patients with a previous primary cancer (n=1,476)

Figure 2 – Patient A
Primary radical treatment → >90 Days → Recurrence Event → <=90 Days → Non - Recurrence Event → Death
Diagnosis | Radical Treatment Event | Death

Figure 2 – Patient B
Primary radical treatment → <=90 Days → Non - Recurrence Event → >90 Days → Patient of Interest → Death
Diagnosis | Radical Treatment Event | Death

PRELIMINARY RESULTS

Figure 3 shows:
- The majority of patients (66%) with no previous history of cancer are classified as "early stage" at diagnosis - this is based on patients presenting with stage 1-4a disease at diagnosis
- Only a small number of patients present with "Advanced" disease (5%) – stage 4b or higher
- 25% of patients are recorded with unknown stage at diagnosis
- Over 70% of patients in the cohort were still alive as of January 2016
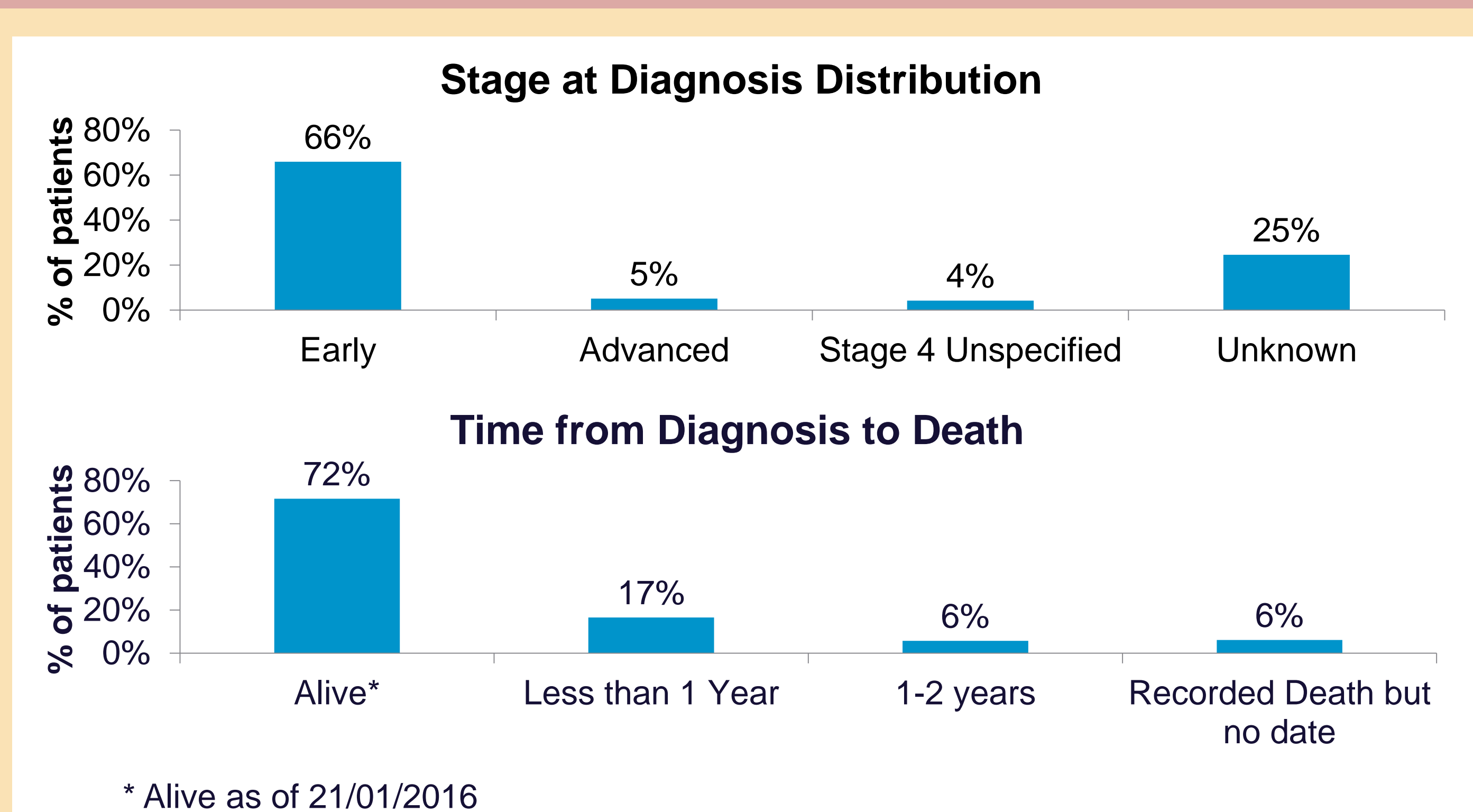

Stage at Diagnosis Distribution
- Early 66%
- Advanced 5%
- Stage 4 Unspecified 4%
- Unknown 25%

Time from Diagnosis to Death
- Alive* 72%
- Less than 1 Year 17%
- 1-2 years 6%
- Recorded Death but no date 6%

* Alive as of 21/01/2016

Figure 3. Distribution of cohort by stage at diagnosis and time from diagnosis

Total Cohort n=6,219 → Potential recurrence based on treatment n=448 (7.2%)

Patient of interest - based on death n=692 (11.1%)

Cause of Death
- Head and Neck Cancer 74%
- Other Cancer (C00-C97) 16%
- Non-Cancer 9%

Figure 4. Initial algorithm results for recurrence and 'patients of interest'
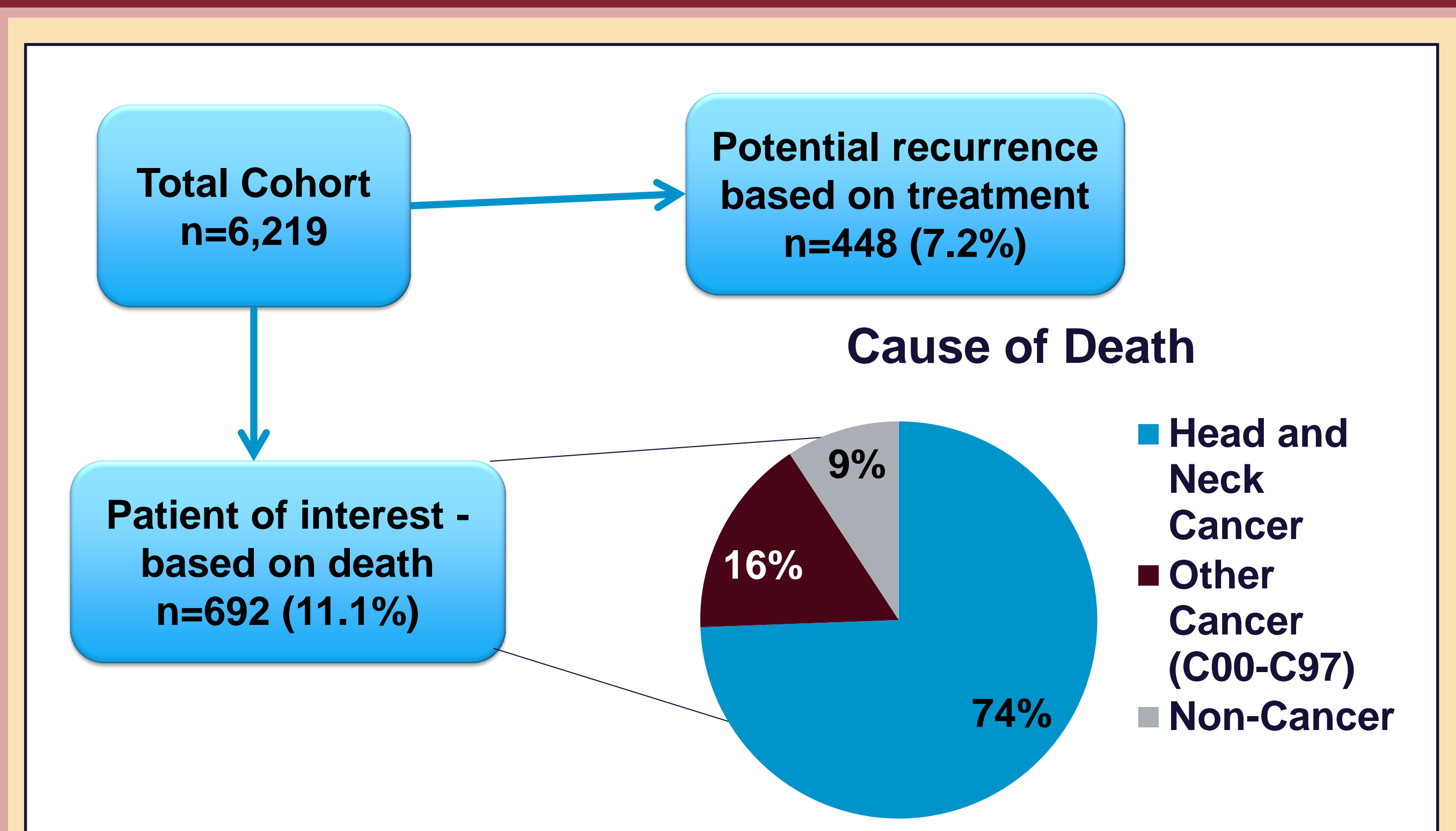
Figure 4 shows:
- 7.2% of the 6,219 patients were identified by the algorithm as having a recurrence based from their treatment data as described in **Figure 2 – Patient A**
- 11.1% of the patients in the cohort were identified as 'patients of interest' as they died over 90 days from their final radical treatment. The majority (74%) of these deaths were due to head & neck cancers (C00-C14, C30-C32)

DISCUSSION

Due to a lack of reported information on recurrence, this algorithm attempts to identify possible markers for recurrence, initially utilising directly coded treatment information from HES, SACT and RTDS. This will then form part of a larger project aimed at identifying a population of cancer patients with "progressive" disease.

The algorithm is currently showing that only 7% of the cohort have suffered a recurrence, this is most likely an underestimate as there will be a group of patients that never receive treatment for their recurrence. These patients could potentially be identified in the 'patients of interest group', however validation of the data will help to understand this.

Data on date and underlying cause of death can also be used to calculate progression free survival within the cohort and will help inform the picture of progressive cancer.

An attempt to use DAHNO data to validate this result was made, however due to the poor quality of recurrence data available in DAHNO this approach was deemed an insufficient means of validation. Further validation work utilising CWT and interrogation of patient records will help in refining the algorithm and identifying patients suffering a cancer recurrence.

ACKNOWLEDGEMENTS

REFERENCES

1. K. Ricketts, M. Williams, Z.-W. Liu, A. Gibson, Automated estimation of disease recurrence in head and neck cancer using routine healthcare data, *Computer Methods and Programs in Biomedicine* 117 (2014) 412-424.