

# Building our understanding of progressive cancer using routine national datasets

Authors  
L Irvine[1]; H McConnell[2]  
[1] Macmillan-NCIN Senior Data Fellow, NCIN, [2] Macmillan Cancer Support

## Background

Historically there has been limited national information on progressive cancers (recurrence, second cancers and metastatic disease), as data on these have not been routinely collected. As part of a joint programme, Macmillan and NCIN are using patient-level national cancer datasets to build our understanding of patterns of progressive cancer at a population level. In order to meet the needs of people living with cancer beyond their initial treatment and to facilitate timely re-introduction to the healthcare system we need to understand how many people have progressive cancers and their touch points on the health system. This will inform the services we develop to meet their needs and allow health professionals to spot triggers which could indicate progressive cancer.

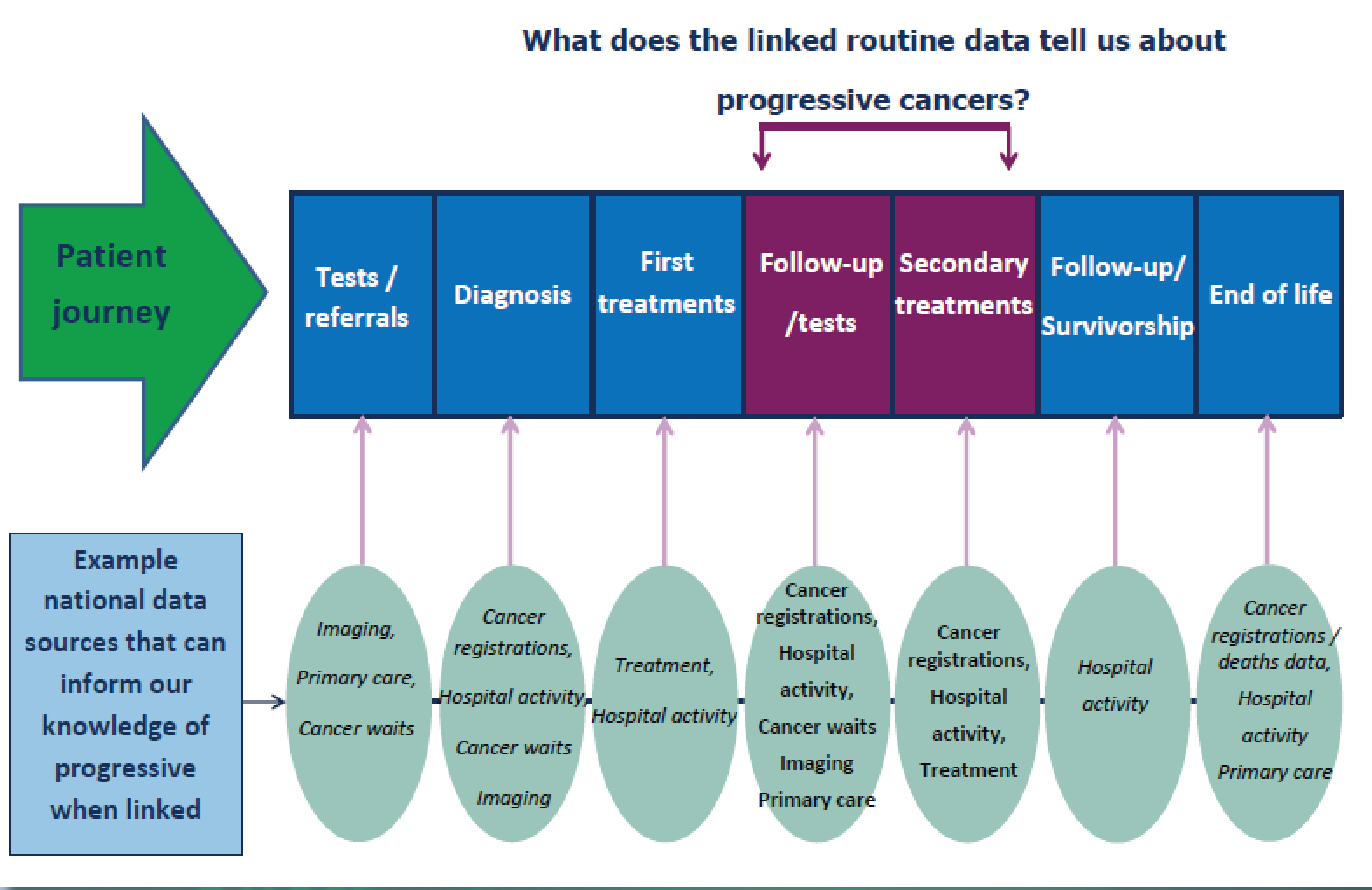
## Method

The study assesses the quality of routine national datasets for analysing progressive cancer, and determines how these data can be used to develop a methodology to identify cancer progression. The project is in the early stages, but initial analysis will be complete later this year.

There are a number of national datasets which, when linked, can inform our understanding of how cancers progress after the first diagnosis. These datasets will be used to identify activity patterns in the patient pathways over time. These patterns along with clinical input will be used to identify trigger points which we will use to develop proxy algorithms to identify recurrence, second cancers and metastatic cancers.

The work will use patient level national datasets in the National Cancer Data Repository (NCDR). Initial analysis will be for England using data that are currently available, including cancer registrations data, hospital activity data (HES), cancer waiting times and radiotherapy (Radiotherapy Dataset – RTDS). We then plan to incorporate other datasets into the analysis, when they become available, including Diagnostics Imaging Datasets, Chemotherapy and Primary care.

We will focus our initial analysis on four cancer types. Throughout we will seek clinical advice on the approach for analysis and interpretation of findings. We will discuss and compare our work with other progressive cancers related analysis, which use local, trust level or specific data sources.



## Conclusions

Used in isolation current national data sources do not provide us with enough evidence to understand progressive cancers. We aim to identify proxy triggers to classify progressive cancers using linked national routine datasets. We will validate these proxies against more detailed local data sources where available and seek to replicate analysis across the UK adjusting for known variations between available sources where possible. Linking national data sources will provide a fuller picture of care pathways and help us identify patterns of cancer progression.