This is a peer-reviewed, accepted author manuscript of the following paper: Laskey, J., Baillie, K., Clarke, J., Crearie, C., Dunlop, E., Kurdi, A., Mueller, T., Wu, O., & Bennie, M. (Accepted/In press). *Real World Data - Can routinely captured health care data be used to determine outcomes of SACT? Experience from the Cancer Medicines Outcomes Programme*. Abstract from British Oncology Pharmacy Association (BOPA) 23rd Annual Symposium, .

Real World Data - Can routinely captured health care data be used to determine outcomes of SACT? Experience from the Cancer Medicines Outcomes Programme

Authors: Jennifer Laskey, Kelly Baillie, Julie Clarke, Christine Crearie, Emma Dunlop, Kimberley Kavanagh, Amanj Kurdi, Tanja Mueller, Olivia Wu, Marion Bennie

Background

Real world data (RWD) is increasingly valued to bridge gaps between clinical trial findings and actual benefits of medicines in local populations. The ability to reliably and efficiently capture RWD, however, is challenging. Within NHS Scotland a Chemotherapy Electronic Prescribing and Administration System (CEPAS) is used, providing an opportunity to report outcomes from Systemic Anti-Cancer Treatments (SACT) through linkage with other routinely captured datasets. Funding was made available as part of the cancer plan 'Beating Cancer: Ambition and Action', a joint clinician / academic team was formed to deliver the Cancer Medicines Outcomes Programme (CMOP) with a vision to maximise the use of existing and evolving local and national electronic datasets to better understand SACT outcomes in the Scottish population.

Objectives:

- Explore the generic capability of existing systems to support outcome evaluations (intended and unintended consequences) of cancer medicines.
- Deploy a small number of exemplar projects to test the capability for clinical outcome evaluation using electronic record linkage (ERL).

Method

A number of exemplar studies were carried out within NHS Greater Glasgow and Clyde (NHS GGC) and the West of Scotland Cancer Network using a generic methodological approach for data collection and analysis. These retrospective cohort design studies captured information about baseline demographics, SACT administration and clinical outcome measures including overall survival (OS). Patient cohorts were identified using CEPAS. Data were collected and analysed via record linkage of routinely collected electronic data sets (ERL) using the NHS GGC Safe Haven platform and from patient case notes (individual patient level data retrieval (IPLR)) to validate ERL. Findings were compared using Chi square, Fisher exact and two-sample t-tests.

Results

In the exemplar studies carried out to date ERL has been used to reliably identify cohorts of interest, determine baseline characteristics and has proven to be a feasible approach to describe clinical outcomes such as OS (table 1). Our studies showed that baseline characteristics, available via ERL, such as poorer performance status; raised lactate dehydrogenase (melanoma); raised prostate specific antigen and low albumin (prostate) had a negative impact on OS. Other prognostic information, such as presence of brain metastases (negative impact on OS (melanoma)) was only available via IPLR and limited our ability to fully contextualise outcomes using ERL. Furthermore, data

on toxicity and reasons for discontinuation of SACT were limited as were not routinely captured via CEPAS.

Discussion / Conclusion

CMOP has demonstrated that ERL is feasible and can provide important information on the clinical effectiveness of cancer medicines in a local population. The next phase of CMOP will further explore the feasibility of this approach within a national setting and to share our data quality findings to drive improvement. Moreover, as cancer treatments become increasingly complex with an emphasis on precision medicine based on individual patient molecular/genetic characteristics there will be a need to continually re-evaluate the methodological approach and ensure that new and emerging datasets can be included.

Table 1. Summary of comparison in data availability between electronic record linkage and individual patient level data retrieval

Workstream (Patient numbers)	ERL findings comparable with IPLR?			
	Construct cohort	Baseline characteristics	Outcome measures	Comments
Prostate (abiraterone and enzalutamide for mCRPC) (271)	✓	Demographics ✓ Performance status ✓ Gleason score ✓ Metastatic disease at diagnosis X Previous therapy ✓ Charlson comorbidity score X*	OS ✓ Duration of therapy ✓ PSA response ✓ Time to next therapy ✓** Identification of trial eligible patients***	Data on metastatic disease at diagnosis unavailable for ~70% via ERL
Melanoma (locally advanced/metastatic) (362)	✓	Demographics ✓ Performance status ✓ Primary melanoma site X* Presence of metastatic sites (e.g. brain) X BRAF status X	OS ✓ Duration of therapy ✓ Adverse events ✓***	Data missing from cancer registry ~4% without entry Data on BRAF status unavailable for ~40% via ERL
Colorectal (metastatic) (295)	✓	Final results awaited		
Gynaecological (NACT for vulval, endometrial & cervical) (approx. 235)	In progress	Final results awaited		NACT episode available for cervical only; may limit cohort identification

Abbreviations: ERL – electronic record linkage; IPLR – individual patient data level retrieval; mCRPC – metastatic castration resistant prostate cancer; OS – overall survival; PSA –prostate specific antigen; NACT – neo-adjuvant chemotherapy

*Key: *Methodology for IPLR and ERL were different due to the available data sets; ** Except time to radiotherapy;*** Additional data via ERL (e.g. hospital admissions for adverse events)*