

HSC-PBPP End of Project Reports – January 2024

Application Reference (click on reference for EPR Summary)	Applicant	Applicant Organisation	Title and Purpose of study	Date of Approval
2122-0054	Duncan Buchanan	Research Data Scotland	Equality Protected Characteristics Dataset	10/06/2020
2122-0018	Anna Santarsieri	Cambridge University Hospitals NHS Foundation Trust	Toxicities and strategies to reduce them in blood cancer patients treated in the non-trial setting	12/08/2021

Appendix: End of Project Report Summaries

2122-0018 Anna Santarsieri

Toxicities and strategies to reduce them in blood cancer patients treated in the non-trial setting

End of Project Report

1	Aims	
	What did the study set out to achieve?	To investigate whether a modified chemotherapy protocol (escalated BEACOPDac) is as effective as standard chemotherapy (escalated BEACOPP) in Hodgkin lymphoma therapy for younger adults (16-60y)
2	Public Benefit Impact	
	How will these outcomes directly result in benefit for the public? Please give details. This should be the main section answered.	Our project has generated evidence for the clinical and genomic health benefits of escalated BEACOPDac, compared with escalated BEACOPP. As a result of presentation of our research at international conferences, many centres across the world (in the UK, France, Switzerland, the whole of Sweden, Australia) have switched from using escalated BEACOPP to escalated BEACOPDac. This has benefitted hundreds of patients with Hodgkin lymphoma across the world.
3	Data	

	<p>What data were received/processed/collected?</p> <p>Was it as expected? Please give brief details.</p>	<p>Retrospective data were collected on month/year of diagnosis, vital status at last follow-up, month/year of relapse and death (where applicable) and on specific toxicities (e.g. haematological, biochemical), we will be able to assess how efficacious and how toxic modified therapies are compared with standard treatment. Data collection was as expected.</p>
4	Methodology	
	How did you collect the data?	<p>Eligible patients were identified by Principal Investigator (Haematologist) at local NHS hospital. Patient identifiers removed. Assigned unique study ID and data collected in spreadsheet. Pseudonymised data from participating NHS Health Board were sent in password-protected file from nhs.scot to nhs.net account at Cambridge University Hospitals.</p>
	How did you process the data?	<p>Data were stored at Cambridge University Hospitals. Data were collated and analysed to compare efficacy and toxicity endpoints. Data analysis was performed using R software. Statistical tests used included Mann-Whitney U, Fisher and T-tests. A survival analysis was performed using Kaplan-Meier estimators.</p>
	How did you provision/publish the information?	<p>The data has been presented at international conferences as detailed above. We plan to write up and publish the data in 203.</p>
	Did your study scope change from its original aims? Please give brief details.	<p>The study scope has changed from its original aims in that it has focused predominantly on the treatment of Hodgkin lymphoma in younger adults.</p>
5	Outcomes:	
	The outcomes / results of your proposal. Please give brief details.	<p>With 27 months median follow-up, escalated BEACOPDac (modified treatment) appears as effective as escalated BEACOPP (standard treatment) in terms of curing younger adults (16-60y) with Hodgkin lymphoma. Patients who received escalated BEACOPDac required</p>

		<p>significantly fewer units of blood transfusion and spent significantly fewer days admitted to hospital. All women under 35y who received escalated BEACOPDac recovered their menstrual periods after treatment and had a significantly earlier return of their menstrual periods than those who had escalated BEACOPP.</p>
6	<p>Future Questions:</p>	
	<p>Have the processes / results raised further questions for future exploration? Please give brief details.</p>	<p>Reduced toxicity was observed in patients who received the modified treatment, escalated BEACOPDac. In this modified regimen dacarbazine has been substituted for procarbazine, which is known to be a relatively stem cell toxic drug. The results have raised the question whether this toxicity may reflect an excess somatic mutation burden in stem cells exposed to this drug. To explore this hypothesis further we have examined the mutational burden and mutational spectrum in haematopoietic stem and progenitor cells following treatment with either escalated BEACOPP or escalated BEACOPDac.</p>

Equality Protected Characteristics Dataset

End of Project Report

1	Aims	
	<p>What did the study set out to achieve?</p>	<p>The overall aim of the project proposal is to develop a securely held population-wide equalities dataset based on linking data on individuals across a number of administrative databases. This reference dataset would then be available to provide equalities data in future projects looking to report on the equality of outcomes and service provision for specific services where equality data is not available or adequate.</p> <p>This current application covers the first phase of the project to build and test the dataset</p> <p>The reference dataset collates data related to the protected characteristics listed under the Equality Act 2010, such as age, sex, ethnicity, religion, national identity, maternity/pregnancy, marital status and disability. The new dataset and the data are pseudonymised multiple times making it difficult to identify where data originated.</p>
2	Public Benefit Impact	

	<p>How will these outcomes directly result in benefit for the public? Please give details. This should be the main section answered.</p>	<p>During the pandemic data held on protected characteristics in administrative and clinical databases, such as ethnicity was found to be incomplete and often of low quality. This hindered attempts to produce timely evidence on how the pandemic impacted on different groups and communities within the population. This lack of complete data and evidence has the potential to discriminate groups due to lack of evidence on which to base interventions and actions. While there are attempts to address data quality in administrative and clinical databases these may take time to be realised. One approach to mitigate the risks of lack of complete data is to make maximum use of the data that is already collected by different public organisations. The current project aims to address this by linking datasets with a wide coverage across the demography and geography of Scotland and combining records on protected characteristics to aid future research projects provide timely evidence on equality of public services.</p> <p>The public benefit will be in assisting public sector organisations with evaluation of their services to ensure they are inclusive and equitable. Where inequalities in service provision or outcome are identified by research, providers can consider improvements to be made to address them.</p>
3	Data	

	<p>What data were received/processed/collected?</p> <p>Was it as expected? Please give brief details.</p>	<p>Data that were processed came from the following sources:</p> <ol style="list-style-type: none"> 1) Scottish Census 2011(from NRS) 2) School Pupil Census for 2011-19 (from SG) 3) SMR00/SMR01/ SMR02/SMR04 for 2011-22 (from PHS) <p>Only the following variables were collated from these sources and combined using on a set of business rules: Date of birth, sex, ethnicity, religion (census only), national identity (census/pupil census only), maternity episodes (SMR02 only), marital status (census/SMR only), disability (Census only) and disability-related student support needs (pupil census only).</p> <p>In addition, a pseudonymised index number was provided by NRS, who retained master index file that links this index number to the population spine.</p> <p>The data received and processed were as expected and specified in the original application.</p>
4	Methodology	
	How did you collect the data?	All data was sourced from existing national datasets collected and maintained by Scottish Government (pupil census), National Records of Scotland (Census) and Public Health Scotland (SMR records)
	How did you process the data?	The source datasets from the original data controllers were pseudonymised by National Records of Scotland which involved replacing person identifying information on the source datasets with an anonymous index number. These pseudonymised datasets were then linked separately by EPCC within the National Safe Haven and combined to create a single value for each protected characteristic using pre-agreed business rules. Analysis and testing were carried out on the final dataset within the NSH to test the application of the business rules and create summary statistics of the distributions of values.
	How did you provision/publish the information?	A Summary report describing the Dataset has been prepared (and is attached). This has been shared with researchers and others who may have an interest in using the dataset or for reviewing and providing feedback in relation to research requirements.
	Did your study scope change from its original aims? Please give brief details.	The study scope did not change from the original aims.

5	Outcomes:	
	The outcomes / results of your proposal. Please give brief details.	<p>The final dataset has captured a wide range of data on many of the protected characteristics of over 6.2 million individuals across the full distribution of age. This includes individuals who have been born or moved to Scotland since the 2011 census. It was not possible to capture information on sexual orientation or gender reassignment since these were not captured in any of the data sources. The dataset can be linked to data on individuals from public sector organisations to allow analysis of equality issues such as uptake and outcomes across a range of characteristics and is sufficiently large to allow analysis of minorities at a granular level as well as intersectional analysis.</p>
6	Future Questions:	
	Have the processes / results raised further questions for future exploration? Please give brief details.	<p>Future issues to be addressed from the first phase of the project are:</p> <ul style="list-style-type: none"> • Fitness for purpose for use as a reference dataset by public sector organisations and researchers • Retention of data on different protected characteristics sourced from varying data sources at different times, included some heavily reliant on census 2011. • Frequency and efficient methods of updating the values to reflect changing population and identities. • Methods of suitable public engagement on the use of the as a dataset as a standing resource for research. • Ensuring consistency and lack of duplication with other data improvement initiatives, especially around ethnicity data.