UNIVERSITY OF
LINCOLN

EAST MIDLANDS INTEGRATED LIFESTYLE (ILS)

DATABASE - FEASIBILITY STUDY

Final Report – May 2021

THE RESEARCH TEAM

Dr Joseph Akanuwe, Research Assistant, School of Health and Social Care, University of Lincoln; https://staff.lincoln.ac.uk/jakanuwe

Dr Ros Kane, Associate Professor and Director of Research, School of Health and Social Care, University of Lincoln; https://staff.lincoln.ac.uk/rkane

Dr Hannah Henderson, Senior Lecturer, School of Sport and Exercise Science, University of Lincoln; https://staff.lincoln.ac.uk/hhenderson

THE STEERING GROUP

The following steering group members met fortnightly to discuss and monitor progress of work:

- Dr Ros Kane, Associate Professor and Director of Research, School of Health and Social Care, University of Lincoln
- Sally Bassett, Public Health Speciality Manager, National Institute of Health Research East Midlands Clinical Research Network (NIHR EM CRN)
- Mr David Clark, Public Health Programme Manager, Lincolnshire County Council
- Dr Joseph Akanuwe, Research Assistant, School of Health and Social Care, University of Lincoln

For further details please contact:

Dr Ros Kane
School of Health and Social Care
University of Lincoln
Brayford Pool
Lincoln, LN6 7TS
Tel: 01522 83 7326
Email: rkane@lincoln.ac.uk

Dr Joseph Akanuwe
School of Health and Social Care
University of Lincoln
Brayford Pool
Lincoln, LN6 7TS
Tel: 01522 88 6293
Email: jakanuwe@lincoln.ac.uk

ACKNOWLEDGEMENTS

Special thanks to the Directors of Public Health in the East Midlands and particularly Professor Derek Ward, Director of Public Health in Lincolnshire and Chair of the group of East Midlands Directors of Public Health, for his support with the project. We are also grateful to all of the key stakeholders who participated in the consultation exercise and willingly shared their views to inform this report. Thanks also to Dr. Sam Cooke for his comments on a draft of the report and final proof-reading.

FUNDING

This project was funded by the National Institute for Health Research East Midlands Clinical Research Network and the College of Social Science, University of Lincoln.

CONTENT

FIGURES AND TABLES

EXECUTIVE SUMMARY

Background

A regional integrated database could serve as a rich data source for in-depth analysis in research studies across key Public Health lifestyle areas in the East Midlands. This could inform Public Health policy, service delivery and commissioning decisions. Unfortunately, existing datasets are poorly aligned across the four key Public Health lifestyle areas examined in this study: physical activity, smoking cessation, reduction in alcohol consumption, and diet and weight management. This feasibility study was therefore commissioned by the East Midlands Directors of Public Health Group chaired by Professor Derek Ward, Director of Public Health in Lincolnshire, with funding from the NIHR East Midlands Clinical Research Network and the College of Social Science, University of Lincoln. Public Health researchers in the Mental Health, Health and Social Care Research Group (MH2aSC) at the University of Lincoln were invited to carry out the study to explore the feasibility of developing and implementing an integrated lifestyle database across the East Midlands Region.

Methods

A scoping review for available evidence was conducted to inform decisions about feasibility of the proposed integrated lifestyle database. This was followed by a consultation exercise with 18 stakeholders, predominantly in the East Midlands, from September 2020 to February 2021. The consultation exercise sought to gather the views of stakeholders, purposively invited to take part due to their role in public health, about the potential feasibility of an integrated database. Stakeholders were identified and invited by email to participate in the consultation meetings which took place via Microsoft Teams. A topic guide, designed specifically for this study, was used to guide the discussion. The meetings were recorded, transcribed, and analysed thematically.

Results

The scoping literature review revealed potential benefits but also barriers to the development of an integrated lifestyle dataset, and highlighted the need to consider local factors which need to be better understood prior to implementation. These findings from the literature were supported by rustults from the subsequent consultation exercise.

Stakeholders for the most part, welcomed the idea of an integrated East Midlands lifestyle database because of its potential benefits for research and to produce evidence to inform service development and commissioning decisions.

However, concerns were expressed by some providers including anxieties around revealing their business strategies to rival organisations also involved in the provision of lifestyles services, the cost

of setting up and running the proposed integrated database, and the complexities involved in information sharing and governance arrangements which would need to be established.

Conclusion

In view of the findings the following options should be explored while taking into consideration the barriers and facilitators expressed by stakeholders:

1. A fully integrated **individual level** lifestyle dataset across the whole East Midlands covering all four lifestyle areas, with governance and access controlled by one institution (possibly a Local Authority or a university) that will house and maintain the database.

2. A fully integrated **individual level** dataset for all four lifestyle areas, within just one geographical area to start with, which is owned by the service provider. There is a need to consider how to make this available more widely, as the providers only report collated data back to the commissioners.

3. A fully integrated **individual level** dataset initially starting with one health area (possibly smoking which already has a standardised Key Performance Indicators (KPI) across the whole region, (to be rolled out later subject to success), with governance and access controlled by the institution (either a Local Authority or a local university) that will house the database.

4. An integrated **aggregated level** dataset covering all four lifestyle areas (reporting similar KPIs as is done currently by service providers who report back to their commissioners), across the whole East Midlands, with governance and access controlled by one institution (possibly a Local Authority or a university) that will house and maintain the database.

5. A fully integrated **aggregated level** dataset for all four lifestyle areas, within just one geographical area to start with, as we have in Lincolnshire, which is owned by the service provider. There is a need to consider how to make this more widely available, as the providers only report collated data back to the commissioners. This is the model already used in Lincolnshire.

6. An integrated **aggregated level** dataset initially starting with one health area (possibly smoking which already has a standardised KPI) across the whole region, (to be rolled out later subject to success), with governance and access controlled by the institution (either a Local Authority or a local university) that will house the database.

1.0 CONTEXT

We have been funded by the NIHR East Midlands Clinical Research Network, supported by additional funding from the College of Social Science, University of Lincoln, to explore the feasiblility of developing and implementing an integrated database, pertaining to lifestyle services commissioned and/or delivered by Local Authorities . We are working with local policy makers and commissioners in Public Health to explore the opportunities and barriers to developing and sharing an integrated database capturing public health lifestyle data focusing on four key areas: physical activity, smoking cessation, reduction in alcohol consumption, and diet and weight management.

A regional integrated lifestyle dataset has the potential to be of great importance for data sharing and secondary analysis of existing data to maximise research potential. This could allow for geographical comparisons, increased sample size and power within datasets, development of existing resources and could provide greater returns on research investments (Doiron et al., 2013; Pisani & AbouZahr, 2010; Piwowar et al., 2008). An integrated dataset could also allow access to a large sample of individual level data on the behaviours relating to smoking, alcohol consumption, diet and exercise, by Directors of Public health and other public health and research personel, which could be analysed to inform commissioning and decisions around service provision. This could also inform research into the risk factors for and determinants of key lifestyle bahaviours and outcomes.

Under the steer of the group of regional Directors of Public Health, chaired by Professor Derek Ward, the following key aims were agreed and form the basis of this report (see Gantt chart in Appendix 1):

- To conduct a scoping literature review
- To conduct a consultation exercise with key stakeholders to assess the feasibility of developing and implementing a regional integrated lifestyle dataset
- To devise a communication plan for dissemination of findings
- To produce a position statement on delivery options
- To develop a subsequent NIHR funding bid to investigate the key research questions that will emerge from this feasibility study.

2.0 BACKGROUND

Lifestyle behaviours such as, smoking, poor diet, lack of exercise and excessive alcohol consumption contribute to poor health. Evidence shows that low levels of physical activity and high levels of sedentary behaviour are important risk factors for non-communicable diseases (Biswas et al., 2015; WHO, 2016; Wilmot et al., 2012). Smoking is responsible for about a fifth of all cancer cases in the UK, and more than a quarter of all cancer deaths (Parkin et al., 2011). Fortunately, smoking reduction services are effective in tackling smoking related poor health (Bauld et al., 2010). Obesity resulting from lack of exercise and an unhealthy diet is also an important public health problem. Being overweight or obese increases the risk of developing chronic diseases such as type 2 diabetes, heart disease, stroke and some cancers which in turn creates an increased demand on health and care services (WHO, 2009). Similarly, excessive alcohol consumption is linked with chronic ill-health including heart disease, cancer and digestive disorders (WHO, 2014; Rehm et al., 2010).

There is a national drive towards preventing illness by tackling unhealthy behaviours and supporting people to remain in good health. The Health and Social Care Act (2012) places specific duties on county councils to protect and promote health and reduce health inequalities. Local Authorities provide interventions which reduce risks to health and the impact of disease across primary, secondary and tertiary levels, and are exploring innovative ways of evaluating and improving their services; for example, through sharing data and secondary analysis of existing data to identify risk factors and trends in patterns of behaviour.

A regional integrated database could potentially be a powerful tool to inform public health policy, service delivery and commissioning decisions and could provide a rich data source for more in-depth analysis in future research studies (with geographical comparisons of determinants and prevalence of lifestyle behaviours and comparisons between and across the four lifestyle areas). The number of interventions and individual records that can be harnessed in a regional database could be significant for driving research and helping to improve health outcomes. Indeed, the Lincolnshire lifestlye service (OneYou Lincolnshire) alone received around 1000 referrals per month and has around 3000 service users registered at any one time. Accumulating data over time and across the whole East Midlands region therefore has the potential to create a resource of significant scale to allow in-depth statistical analysis and stratification by multiple variables, not currently possible (due to unlinked data and small sample sizes) with the existing separate datasets.

Building an integrated lifestyle database could also enhance East Midlands regional and national efforts to promote health and support post COVID recovery and return to business as usual, which aligns with the government's new obesity strategy that seeks to health And wellbeing, and to protect against COVID-19, reducing the social and economic pressure placed on the NHS (Department of

Health and Social Care, 2020). It also aligns with the new government white paper (Working together to improve health and social care for all), seeking to put in place targeted improvements for the delivery of public health and social care interventions to support local systems to deliver higher-quality care to their communities (Department of Health and Social Care, 2021).

3.0 SCOPING LITERATURE REVIEW

We developed a protocol (see Appendix 2) to guide the scoping literature review for evidence to inform the potential development and implementation of an integrated public health lifestyle database. We employed the framework recommended by Arksey and O'Malley, 2005, which allows for consultation with stakeholders and comprises the following steps: *identifying the research question; identifying relevant studies; study selection; charting the data; collating, summarizing and reporting the results; and consultation*.

3.1 Search strategy

Using the search terms (public health AND (lifestyle OR "life style" OR life-style) AND (database* OR dataset* OR "data set*") AND (feasib* OR develop* OR implement* OR use) AND (stop smoking* OR smoking cessation*) AND (weight management* OR weight control*) AND (healthy diet* OR diet control*) AND (alcohol abuse* OR increased alcohol consumption*) AND (health outcome* OR health impact* OR health input*). We searched five electronic databases: Medline, CINAHL, Cochrane, Scopus, Psych-INFO. The reference lists of articles found through the searches were also checked for relevant studies.

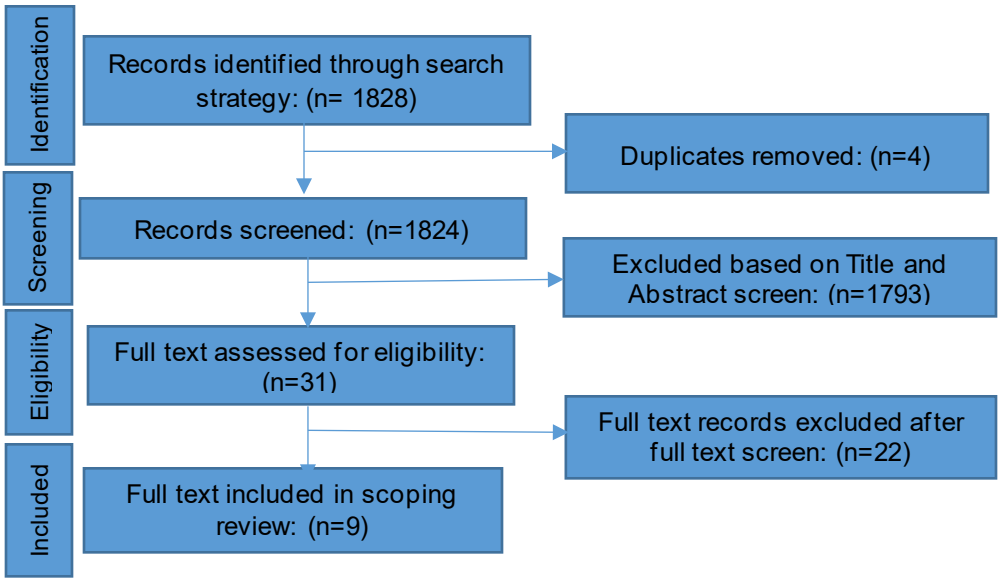
3.2 Study selection

To select the relevant studies, we followed the PRISMA Extension for Scoping Reviews (PRISMA-ScR). Two reviewers independently screened the studies by title and abstract, guided by the inclusion criteria. Any disagreement between the two reviewers over the eligibility of particular studies was resolved through discussion with a third reviewer. After the title and abstract screening, full texts of the eligible articles were retrieved and screened against the inclusion criteria:

- Any year of publication. No limits to the publication dates were enforced since we were unsure of the extent of the evidence available
- English language only
- Peer-reviewed academic literature
- Study design: quantitative, qualitative and mixed methods
- Interventions pertaining to any age group (adults and children)
- Focus on any type of public health lifestyle data set
- Any geographical location

The study selection process is presented in the PRISMA flowchart (Fig 1) below.

Figure 1: PRISMA flow chart representing the study screening process



Screening and selection of studies and extraction of data from selected studies were guided by the protocol which was designed by the research team and agreed with the funders. Studies that did not meet the inclusion criteria were excluded (see Appendix 3 for a list of excluded studies). Pending a full literature review, an annotated bibliography is provided below, covering the 9 studies that resulted from the study selection process .

Included studies after screening 31 full text studies for eligibility

1.Saunders P, Mathers J, Parry J et al. "Identifying 'non-medical' datasets to monitor community health and well-being." ***Journal of Public Health*****. 2001; 23;2: 103-108.**

In this study, a stakeholder discussion was conducted involving public and environmental health academics and professionals in the West Midlands region of England, inaddition to examining the Office for National Statistics databases. The aim was to identify routinely collected 'non-medical' datasets containing information on the physical environment, crime, housing and homelessness, social services, the socio-economic environment including employment, lifestyles, education, leisure and culture, transport and accidents.

Although the authors collected information on a variety of datasets, only the lifestyle data (particularly smoking and alcohol consumption data from the Health Survey for England dataset) are relevant to our study. Saunders et al. (2001) suggested that intersectoral working and multi-agency involvement at the local level are central to improving the quality of many datasets and can promote their use in the measurement and monitoring of community health.

Lesson learnt: Encouraging intersectoral working such as Local Authority collaboration in data collection and sharing can increase the feasibility of developing and implementing an integrated lifestyle database for the East Midlands region.

2. Zwisler AD, Rossau HK, Nakano A, et al. The Danish Cardiac Rehabilitation Database. *Clinical Epidemiology*. 2016; 8:451-456.

This study reports on the development of the Danish Cardiac Rehabilitation Database (DHRD), an online, clinical quality database that aims to provide higher quality cardiac rehabilitation for patients with coronary heart disease (CHD) in Denmark. The DHRD systematically monitors the quality of Cardiac Rehabilitation provision across programmes over time and data can be accessed for research related to both the outcome and organisation of cardiac rehabilitation. In the process of data collection, patient-level data are registered by clinicians at the time of entry to Cardiac Rehabilitation, directly into an online system and simultaneously linked to other central patient registers. Follow-up data are entered after 6 months.The variables recorded in the DHRD include smoking status, exercise capacity, height, weight and blood pressure; and data on performed diagnostic tests (eg, diabetes and depression), along with individual plans for rehabilitation (eg, training sessions, dietary treatment, and/or smoking cessation).

Lessons learnt: The process of developing the DHRD can potentially inform the development and implementation of an East Midlands integrated lifestyle database. Specifically, we can learn from how patient-level or individual service user data are registered by clinicians at the time of entry to Cardiac Rehabilitation, and simultaneously linked online to other central patient registers. The follow-up data that are entered after 6 months can also be a useful lesson.

3. Lakervield J, Loyen A, Ling FCM*, et al.* Identifying and sharing data for secondary data analysis of physical activity, sedentary behaviour and their determinants across the life course in Europe: general principles and an example from DEDIPAC. *British Medical Journal Open.* 2017; 7: e017489.

Lakervield et al. (2017) described the development of a comprehensive European dataset and the process towards cross-European secondary analyses of pooled data on physical activity and sedentary behaviour. The authors applied the Findable; Accessible; Interoperable; Reusable (FAIR) framework to provide guidance in the discovery and reuse of data for further investigation, and followed a five-step methodology: (1) identification of relevant datasets across Europe, (2) development of a compendium including details on the design, study population, measures and level of accessibility of data from each study, (3) definition of key topics and approaches for secondary analyses, (4) process of gaining access to datasets and (5) pooling and harmonisation of the data

and the development of a data harmonisation platform. However, barriers such as limited potential for reuse, variation in assessment methods and operationalisation of outcome variables across current European studies hampered data harmonisation.

Lessons learnt: Improving data collection and management through consistent data collection methodologies and application of an appropriate model of implementation such as the FAIR principles, could help address potential barriers to the developemt and implementation of an integrated lifestyle database in the East Midlands.

4. Clarke A & Steele R. "Summarized data to achieve population-wide anonymized wellness measures," 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA. 2012; 2158-2161.

Clarke and Steele (2012) discussed some of the technologies that increase the ease and capability of gathering quantitative wellness data via smartphones, how specific and detailed the data should be for public health use, and the challenges associated with such anonymised data collection. They then proposed a framework to facilitate the collection of non-identifying data; this is based on increased local processing so that only the required information is submitted to avoid the risk of re-identification.

Lessons learnt: To inform the collection and integration of anonymised lifestyle datasets in the East Midlands, lessons relating to service user protection can be drawn from the use of technology based conceptual framework which comprises of four layers: sensors, mobile application, communications and analysis server. In particular, we can learn from the communications and analysis component of the framework, whereby data are shared and analysed without identification of the data source.

5. Sarkar C, Webster C, Gallacher J. UK biobank urban morphometric platform (UKBUMP) - a nationwide resource for evidence-based healthy city planning and public health interventions. Annals of GIS. 2015; 21:135-48.

Sarkar et al. (2015) report the development of a seminal UK-wide baseline spatial database that will function as a platform for evidence-based healthy city planning that will facilitate the construction of models to explicitly decipher health impacts from the genetic to micro built environment scales for half-a-million Britons. Together, the conceptual and empirical data models provide a basis for multilevel urban planning, health policies and intervention strategies at both individual and population levels, allowing for a much greater accuracy of evidence-based policy-making.

Lessons learnt: While this study focuses on built environment, it reinforces the idea that the proposed integrated lifestyle database for East Midlands can be used for a multilevel lifestyle data analysis to inform public health policy and practice.

6. Li S, Zhang L, Liu S et al. "Surveillance of Noncommunicable Disease Epidemic Through the Integrated Noncommunicable Disease Collaborative Management System: Feasibility Pilot Study Conducted in the City of Ningbo, China." ***Journal of Medical Internet Research. 2020;*** **22 (7) e17340.**

This was a pilot study conducted in Ningbo city by the Chinese Center for Disease Control and Prevention (CDC) with the aim of developing an innovative model for non-communicabe disease (NCD) surveillance and management: the integrated noncommunicable disease collaborative management system (NCDCMS). The Ningbo model was designed and developed through a 3-level (county/district, municipal, and provincial levels) direct reporting system based on a regional health information platform. The uniform data standards and interface specifications were established to connect different platforms and conduct data exchanges, allowing for automatic NCDs data exchanging and sharing. According to the authors, the NCDCMS completely reshaped the process of NCD surveillance reporting and had unique advantages, which include reducing the work burden of different stakeholders by data sharing and exchange, reducing the amount of underreporting, and structuring population-based cohorts.

Lessons learnt: The Ningbo model is a milestone in NCDs surveillance, control, and prevention in China, but applicable in other countries. Taking local factors into consideration, the Ningbo model which allows for automatic NCDs data exchanging and sharing could inform the development and implementation of an integrated lifestyle database for East Midlands.

7. Lewer D, Bourne T, George A et al. Data Resource: the Kent Integrated Dataset (KID). International Journal of Population Data Science. 2018; 3:6.

The Kent Integrated Dataset (KID) uses pseudonymisation to link patient-level records from services including general practices, hospitals, community health services and social care. Data are refreshed monthly with processes in place to monitor data quality. For each episode of care, the KID includes date of the episode, the type of service accessed, the cost of the episode and clinical information such as the health condition being treated and results of diagnostic tests. The KID contains three types of data:

- Demographics - a dynamic list of Kent and Medway's registered patient population from NHS Digital, comprising of NHS numbers, age, gender, Lower Super Output Area (LSOA) and

Unique Property Reference Number (UPRN) starting from April 2014. This represents the 'hub' of the KID in which all other local datasets are linked to each other.

- Service activity and costs based on data flows from approximately 220 local organisations across primary care, acute, community and mental health, social care, fire and rescue, etc.
- At least five population segmentation tools, e.g. MOSAIC, ACORN and Electronic Frailty Index.

The KID is a unique and rich dataset available to researchers who are investigating a broad range of public health questions. It provides system-level insight into patient journeys and care utilisation and supports commissioning based on patient needs.

Numerous analytical projects have been carried out supporting public health commissioning, Kent and Medway STP and Joint Strategic Needs Assessment. Examples include designing capitated budgets for multimorbid patients, matched control impact evaluation of Kent Fire and Rescue Safe and Well Visits for the frail elderly, health profiling of population at risk of fuel poverty, levels of social isolation across different risk groups, and equity audit of NHS health checks programme.

Lessons learnt: The KID is a unique and rich dataset for public health research. The KID development process uses pseudonymisation to link patient-level records and data refreshed monthly with processes in place to monitor data quality. This KID development process could be used to inform the proposed integrated lifestyle database for East Midlands Public Health Lifestyle Services**.**

8. Bottle, A, Cohen, C, Lucas, A, et al. How an electronic health record became a real-world research resource: comparison between London's Whole Systems Integrated Care database and the Clinical Practice Research Datalink. *BMC Medical Informatics and Decision Making*. 2020; 20:71.

In this paper key features of the 'Discover' database, including scope, architecture and governance; and descriptive analyses are used to compare the population demographics and chronic disease prevalences with those in the Clinical Practice Research Datalink (CPRD), a popular UK research database, also based on linked primary care records. Like the CPRD, the Discover database population matches the age, sex and ethnic distribution of the UK. But unlike CPRD, Discover has identifiable care organisations and postcodes, allowing mapping and linkage to healthcare provider variables such as staffing, and includes contacts with social, community and mental health care.

Lessons learnt: Both the CPRD and Discover models are valuable research tools and the usefulness of these reinforce the idea of an integrated lifestyle database for research and public health services in the East Midlands. But the Discover model may need to be modified as it has identifiable care

organisations and postcodes – service users and providers of lifestyle datasets may not want their personal data to be identifiable.

9. Trefan L, Akbari A, Paranjothy S et al. Electronic Longitudinal Alcohol Study in Communities (ELAStiC) Wales – protocol for platform development. ***International Journal of Population Data Science.*** **2019; 4(1).**

This Electronic Longitudinal Alcohol Study in Communities (ELAStiC) provides a description of existing key datasets integrated with existing, routinely collected electronic health data on a secure platform, with relevant derived variables linked to enable population-based research on alcohol-related harm in Wales. The routinely-collected available data included hospital admissions, general practice, socioeconomic descriptors and mortality data that can be used in other epidemiological studies.

Lessons learnt: The ELAStiC is a population-based secure platform of routinely collected longitudinal data that can support research and further understanding of alcohol-related disease and health trajectories across the life course. It can inform the facilitation ofsimilar initiatives like the integrated lifestyle database for East Midlands.

3.3 Key findings

The following key findings emerged from the scoping literature review:

- Several databases have been developed or implemented in and outside the UK.
- Using an appropriate model such as the Findable, Accessible, Interoperable and Reusable (FAIR) principles and taking into consideration local factors, could facilitate the development and implementation of a database like the proposed Integrated Lifestyle Database for the East Midlands.
- The KID presents a useful model to inform the construction of an Integrated Lifestyle Database.
- The KID has been used in health profiling of populations at risk of fuel poverty, levels of social isolation across different risk groups, and equity audit of NHS health checks programme. This supports the idea that the proposed East Midlands Integrated Lifestyle Database could be used to assess the impact of lifestyle service on health outcomes.
- One study (Bottle et al., 2020) used descriptive analyses to compare the population demographics and chronic disease prevalences in two databases, the Discover and CPRD, which are both valuable research tools that can serve as models for other databases.
- The ELAStiC study (Trefan et al., 2019) is a population-based secure analytic platform with longitudinal data that can support research to further our understanding of alcohol-related

disease and health trajectories across the life course; and this supports the idea that the proposed Lifestyle Integrated Database may be used to assess the impact of Public Health lifestyle interventions on health outcomes in East Midlands.

- Barriers such as limited potential for reuse of data, variation in assessment methods and operationalisation of outcome variables could hamper data harmonisation
- There is a need to consider local factors in the development and implementation of a database.

In view of the findings from the scoping review, particularly the need to consider local factors, we proceeded to consult and gather the views of stakeholders associated with East Midlands Public Health Services about the feasibility of developing and implementing an Integrated Lifestyle Database.

A range of existing databases, uncovered via both the scoping review and subsequent consultation exercise (reported on below), which have the potential to inform the construction of the proposed Integrated Lifestyle Database for East Midlands are presented in Appendix 4.

4.0 CONSULTATION EXERCISE WITH KEY STAKEHOLDERS TO ASSESS THE FEASIBILITY AND IMPLEMENTATION OF AN INTEGRATED LIFESTYLE DATABASE

Working in collaboration with the East Midlands Clinical Research Network, we conducted a consultation exercise with key stakeholders to gather their views on the feasibility of developing and implementing an East Midlands integrated lifestyle service database. This included examining any potential issues with data access/collation, data sharing and the potential usability of an integrated lifestyle dataset. Prelimiary work involved identifying the key personnel and organisations across the five Local Authorities in the East Midlands, and any other appropriate colleagues, that were to be invited to take part in the consultation.

4.1 Ethical approval

The project was formally logged with the Universtiy of Lincoln Ethics system on 16th September 2020 (Ethics review REF number: 2020_3761).

4.2 Topic guide for the consultation exercise

A topic guide, informed by a framework for feasibility studies (Bowden et al., 2009) to guide the content of the consultations was developed and sent to a representative of the Intergrated Lifestyle Service in Lincolnshire for comment. This informed slight modification of the tool which was then agreed with the funders before being utlised for the consultation meetings which were recorded and later transcribed. A copy is included in this report as Appendix 5.

4.3 Details of stakeholders who participated in the consultation exercise

An invitation was sent out via email, and although there were many competing priorities due to the Covid-19 pandemic, 18 stakeholders (see Table 1 below) responded and participated in the consultation exercise. This included in some cases the provision of additional information and response to questions on the interview guide and followed up by a virtual meeting via MS TEAMS.

Table 1: Stakeholders who attended the consultation exercise

Local Authority/other organisation	Stakeholder	Date of consultation
Lincolnshire County Council	Head of service, Thrive Tribe/OYL	22/10/2020
Nottinghamshire County Council	3 participants: A Consultant in Public Health, a Senior Public Health and Commissioning Manager and a health improvement principal	02/11/2020
Nottingham City Council	A Consultant in Public Heath	10/12/20
Leicestershire county Council	Integrated Lifestyle Manager - Public Health	02/11/2020
Derby City Council	Completed questions on interview guide	11/11/2020
Health Improvement Directorate, Public Health England, East Midlands	Associate Director, Local Knowledge & Intelligence Service	14/12/2020
Northamptonshire County Council	2 stakeholders: A Public Health Intelligence and Insight Manager & Performance Analyst; and a Data health checks and smoking cessation officer	15/12/2020
Lincolnshire's Population Health Management Implementation	2 stakeholders: County Manager - Performance & Intelligence LCC Lead and Lincolnshire's ILS service Lead	15/12/2020
Population Health Improvement Team, Public Health England, East Midlands	Health and Wellbeing Programme Lead	15/01/21
Health Improvement Directorate, Public Health England	Head of Population Health Analysis and Head of Nutrition, Diet, Obesity, Physical Activity and MSK	28/01/2021
Thrive Tribe/OYL, Lincolnshire	Service User	05/02/2021
NHS Lincolnshire CCG (Health Inequalities/Population Health	Health Inequalities/Population Health work Lead	26/02/2021

Note: 1. Derbyshire County Council did not respond
2. Leicester City Council did not respond

The stakeholders above included lifestyle service providers and the details of services, which are presented in Appendix 6.

4.4 Analysis of the data and themes arising from the consultations

Guided by concepts from the selected framework for feasibility studies (Bowden et al., 2009) we conducted a framework analysis (Richie & Spencer, 1994) with the qualitative data from the consultation exercise with stakeholders in East Midlands. The key points that emerged are demonstrated by the direct quotations in Table 2:

Table 2: Themes from consultations

Framework codes (Bowen et al. 2009)	Themes	Descriptive codes	Quotes
Acceptability /Demand *(based on potential benefits/enablers /usefulness/ desirability)*	Opportunities – as it is a potentially useful resource for research, service commissioning and ease of information access	•A potentially rich and well powered resource particularly for research.	*"Yeah, I think it is more of a potential research resource"* (Associate Director, Local Knowledge & Intelligence). *"Actually, if you're relating all the provider datasets, that would be a bigger job and it would be potentially more useful for research purposes because it's got those individual level variables"* (Consultant in Public Health).
		•Potential for geographical comparisons in both KPIs (from a commissioner-level dataset) and individual level lifestyle/socio-demographic data (provider-level dataset).	*"Just in terms of sharing best practice and looking at makeup of the community or the geographical sort of, you know….and comparing similar areas to ourselves. And looking at performance in terms of what are they doing, what are they doing differently that we could perhaps adapt and use locally?"* (Lifestyle service manager). *"It would give you the opportunity to compare how well things are doing in Lincolnshire compared to say Nottinghamshire because we would have both sets of information, yeah"* (Service User). *Potentially help understand more about clients and good practice whilst also creating a space for greater dialogue between services across the region* (Public Health Manager).
		•Potential for comparison across and between the 4 key public health lifestyle areas (would allow for comparison between individually commissioned vs integrated lifestyle services.	*"Good to compare the performance and effectiveness of our service against others as this is a new model based on emerging evidence. I know this could be tricky as each area will have a slightly different model but all we can do at the moment is compare against our separate behaviour change services"* (Health Improvement Principal). *"….better understanding of performance in locality versus nationally, an opportunity to share best practice"* (Head of lifestyle service).
		•Potential to ease or speed lifestyle service information access or searching.	*"Where to go to for your particular needs - you've got five minutes and you want to know about smoking cessation services in Derbyshire. Where would you go to? You've got half an hour to do a bit deeper dive. Where would you go to?"* (Health & Wellbeing Programme Lead).
		• Some stakeholders would fully embrace the concept of having	*"Having a national collection and reporting process like that of the DOH smoking reports would be useful" (Lifestyle service manager).*

		an integrated and shared dataset.	*"As a service provider we would embrace a shared dataset across the East Midlands – there are several considerations; however, I believe it would add value to service delivery and best practice" (Head of Lifestyle service).* *"That could help towards the delivery of a better course. You know, you could have courses for specific people. People with heart attacks, stroke, muscular skeletal injuries. Mine was a very extreme goal. I wanted to lose 5 stone. If someone only want to lose, say half stone or a stone or whatever, and the courses can be tailored more towards them" (Service User).*
Adaptation/ Expansion (based on *existing datasets/other resources etc that can be adapted or expanded into an integrated database)*	Strengths - as there are *existing datasets/other resources etc that can be adapted or expanded into an integrated database*	•Datasets already exist in each of the Local Authority areas across the East Midlands and within each provider organisation.	*"For the commissioned lifestyle service two data sets exist, one which is used for smoking cessation and another which is used for lifestyle"* (Public Health Manager). *"We have a number of integrated patient databases on the health side of things, and one, the care portal which is starting to link to adult social care records. None currently includes primary care data or data from providers of services such as lifestyle services – they will eventually though, primary care first. However, we have a huge programme of work in train to have an entirely linked identifiable, and pseudonymised, data base in Lincolnshire, as most other ICSs will have, as part of our move to a Population Health Management approach. The first part of this is currently underway"* (Performance & Intelligence Lead).
		•Data are already routinely collected.	*"The data is part of the service reviews which are undertaken annually to improve service delivery"* (Public Health Manager).
		•Some alignment with Information Technology (IT) systems already in place.	*"Our external partners use a variety of software packages that are imported into our master system through CSV files"* (Head of lifestyle service). *"This would need work with our provider as they are the data controller but as mentioned above they have found a way to share with the GPRCC system"* (Health Improvement Principal).
		•Some standardisation in reporting mechanisms and variables collected already in place (particularly in relation to smoking cessation).	*"But you know it's quite sort of standardized that people actually think this has got potential"* (Lifestyle service manager).

Practicality	Weaknesses - due to wide variety of commissioned services, unstandardised collection of variables, unaligned IT systems and related cost	•Collection of variables not completely standardised across the region.	*"Yeah, so I suppose it's about standardising of what people would commission, and another service which may be a private provider. Or it might be an NHS provider. So, provided you've got those two levels of understanding from a Commissioner point of view. And then separately understanding from the provider point of view"* (Health Improvement Principal). "*As we have a personalised data set for our area and there isn't a consistent data collection or reporting nationally, there isn't currently an easy fit to standardise collection from other healthy lifestyle services nationally or locally"* (Head of Lifestyle Service).
		•IT systems and processes not fully aligned across the region.	*"Even just trying to change the system onto a version that everyone across the region is quite a nightmare with all the current protection controls. I don't think it could set with everything from the same region in the same database"* (Public Health Intelligence and Insight Manager).
		•Wide variety of currently commissioned providers of public health services currently in place.	*"What I would add is that for smoking cessation in particular, our service is not the only one in the city, so we have stop smoking service, that's run by our GP Alliance and those commissioned by the local authority. The GP Alliance have decided that they would use some services. We have a CCG commissioned a smoking cessation service that sits within"* (Consultant in Public Health). *"NHS England have just announced that they are going to provide weight management services for those with diabetes or with high blood pressure. And that's going to be a digital weight management offer. And there was some consultation, but there's not too many details yet as to how they'll be rolled out, but that will take some of the people who would have been attending our weight management because"* (Consultant in Public Health).
		•Areas vary in the profile of provision (some offer fully integrated services encompassing all 4 key lifestyle areas) others have separately delivered services.	*"We don't have a central database, it's so individual because we commission separately for each one. So if you went, for example, somewhere like Derbyshire or Leicester City who have integrated lifestyle services, they probably hold an individual level data on most of those people, and therefore probably can quite easily collate it"* (Consultant in Public Health).
		•Some stakeholders are sceptical about the practicalities and utility	*"I think you need to talk to providers directly about some of this and Commissioners directly about it; I think there are insurmountable barriers that cost money" (*Associate Director, Local Knowledge & Intelligence).

		of having an integrated and shared dataset.	*"Time and capacity to support any implementation as well as any budgetary or procurement processes or other local arrangements"* (Public Health Manager). *"…to be honest ….I have concerns about both the feasibility of a shared regional dataset and whether it would present any tangible benefits"* (Public Health Manager). *"And the requirements in the format. So we get different things for each service and the performance indicators. And unless it's all commissioned together, I don't see how we can really integrate it fully"* (Public Health Intelligence and Insight Manager). *"But I can imagine a lot of people who've not had the exposure I've had would be worried by it. I can remember a number of ladies that were involved were worried about the data sharing, and where the information would go and who had access to it"* (Service User).
		•Expensive to standardise software.	*"It is expensive to standardise software and approaches"; "I wouldn't know the full cost here but to standardise and implement this in the first instance I would imagine the costs for this would be substantial and front loaded":* (Lifestyle service manager)
		•Reluctance to share commercial products.	**"***reluctance of companies to share their private commercial products"* (Lifestyle Service Manager) *"Yes, and I thought you'd be stepping into an interesting territory in terms of data protection. Yeah, figure out how you get, even if it's pseudonymized data sets from all of them and whether they were willing to share that"* (Consultant in Public Health).
		•Commissioner expectations and demands vary.	*"commissioner expectations and demands vary wildly from boroughs"* (Lifestyle Service Manager).
	Threats – relating to sensitivity of data, governance and data access issues and reluctance of providers to share business strategies with competitors	•Data are sensitive and there is variation in the extent to which organisations are prepared to fully share their data.	"*A concern here when considering partners would be the reluctance in sharing commercially sensitive information – the data we collect and the way that we designed the processes are key in securing future business and are therefore not something that we want readily available to competitors in the public domain*" (Head of Lifestyle Service). *"Our alcohol treatment data in particular is highly sensitive and must be compliant with the core National Drug Treatment Monitoring System (NDTMS) dataset and confidentiality guidelines – it would not be appropriate to include in this sort of exercise"* (Public Health Intelligence and Insight Manager).

			"Well, I would be a bit concerned because I know that all the names and addresses are attached to that information" (Service User).
		•Difficulties in fully establishing data sharing agreements across all parties. •Unanswered questions about Information Governance and location/security of dataset. •Mechanisms for data access and permissions might be difficult to establish.	*"If you're going to make it useful, then you have to have a way for us to access it whilst maintaining data protection and ensuring that we're using it in a way that you agreed with providers that it will be used. Anyway, a whole lot of issues about holding it in data protection"* (Consultant in Public Health). *"…and especially with data sharing agreements that would need to be put in place between all the different providers"* (Consultant in Public Health). *"Datasets and data collection methods vary greatly between services and there would be significant information governance issues with trying to combine them at a local level, let alone regional"* (Public Health Intelligence and Insight Manager). *"I think when you engage providers, I think you are going to have quite a clear view on what you're trying to achieve with the data set"* (Consultant in Public Health).
		•Need to fully understand user involvement and consent to use personal (albeit anonymised) data for research/commissionin g purposes.	*"…and you have to have, I guess consent from the service users as well, it's going to be shared in a different way. You know, it depends on how the providers want it shared, because many of them will likely have some agreement already"* (Consultant in Public Health). *"Data sharing would have to be done with client consent and how reasonable is it to ask clients to agree to their data being shared with numerous other stakeholders with no direct involvement in whatever intervention they're receiving?"* (Public Health Intelligence and Insight Manager).
		•May be costly initially (*to migrate from current dataset to a new integrated dataset)* with cost benefits not emerging until dataset is fully established (as high risk difficult to accurately project/model).	*"This would require buy in from several disparate contractors or national guidance in line with something similar to the DOH smoking data. Implementation in our service would require expensive alterations to collection and reporting software and this would need to be considered when making alterations to fixed term contracts"* (Head of Lifestyle Service).

		•Lifestyle service providers seem to be in competition for contracts and may not want to data share because of fear of revealing their service provision strategy.	*"The data we collect and the way that we designed the processes are key in securing future business and are therefore not something that we want readily available to competitors in the public domain"* (Head of Lifestyle Service). *"That makes sense because yeah, I imagine some of them may have objections potentially, if they think they're going to be compared with other areas of the effectiveness of other services"* (Consultant in Public Health).

5.0 DECISION ABOUT FEASIBILITY

Alongside the consultation exercise, we applied a Traffic Light system (Avery et al., 2017) to support our decision about feasibility, and to determine whether the project should move into a future dataset development phase.

5.1 The Trafic light system

The Traffic Light system (Avery et al., 2017) suggests that instead of employing simply stop/go criteria, it is more useful to employ red/amber/green Traffic Light progression criteria, outlined as follows:

- stop/red (when there are intractable issues that cannot be resolved);
- amend/amber (where there are remediable issues in which modifications may be needed before progressing); and
- continue/green (where no concerns that threaten the success of the proposed intervention - in this case the development of the dataset have arisen).

Table 3: Summary of feasibility findings showing progression decision and proposed modifications for each outcome

Outcome	Decision	Summary of opinions from the consultation exercise
Acceptability/demand	Amend/amber 	Stakeholders would embrace the concept of having an integrated database because there are potential benefits e.g: • A potentially rich and well powered resource for commissioning but particularly for research. • Potential for geographical comparisons in KPIs from a commissioner-level dataset) and individual level lifestyle/socio-demographic data (individual-level dataset). • Potential to ease access to lifestyle service information . Some providers questioned the acceptability of the concept especially where they are in competition for contracts which causes a reluctiance to data share because of fear of revealing their service provision strategy.
Practicality	Amend/amber 	A number of concerns and practical issues (barriers) were highlighted by stakeholders: • Collection of variables not completely standardised across the region. • IT systems and processes not fully aligned across the region. • Areas vary in the profile of provision: some offer fully integrated services encompassing all four key lifestyle areas) others have separately delivered services. • There is an initial need to map out provision across the region. • Some stakeholders are sceptical about the practicalities and utility of having an integrated and shared dataset. • Expensive to standardise software. • Reluctance to share commercial products. • Commissioner expectations and demands vary. • Data are sensitive and there is variation in the extent to which organisations are prepared to fully share their data. • Difficulties in fully establishing data sharing agreements across all parties. • Unanswered questions about Information Governance and location/security/ownership of dataset. • Mechanisms for data access and permissions might be difficult to establish. • Need to fully understand user involvement and consent to use personal (albeit anonymised) data for research/commissioning purposes. • May be costly initially (to migrate from current dataset to a new integrated dataset) with cost benefits not emerging until dataset is fully established (as high risk difficult to accurately project/model).
Adaptation/expansion	Amend/amber 	Existing datsets can be integrated/expanded when barriers are addressed.

Note: The outcomes here (acceptability/demand, practicality and adaptation/expansion) are concepts from the feasibility framework (Bowen et al., 2009) that were relevant and therefore informed the data analysis in Table 2 above. Other concepts of the framework (implementation, integration and limited efficacy) did not directly match our data and were excluded from the analysis.

The findings from the consultation exercise with stakeholders (see Table 2 above) included potential benefits and concerns to the development and implementation of the proposed integrated database as highlighted in the Strengths, Weakness, Opportunities and Threats (SWOT) analysis in Table 4 below.

Table 4: SWOT analysis of data from scoping review and consultation exercise with public health professionals

Strengths	**Weaknesses**
• Datasets already exist in each of the Local Authority areas across the East Midlands and within each provider organisation. • Data are already routinely collected. • Some alignment with Information Technology (IT) systems already in place. • Some standardisation in reporting mechanisms and variables collected already in place (particularly in relation to smoking cessation). • Evidence that some stakeholders would fully embrace the concept of having an integrated and shared dataset.	• Collection of variables not completely standardised across the region. • IT systems and processes not fully aligned across the region. • Wide variety of currently commissioned providers of public health services currently in place. • Areas vary in the profile of provision (some offer fully integrated services encompassing all 4 key lifestyle areas) others have separately delivered services. • There is an initial need to map out provision across the region. • Evidence to suggest that some stakeholders are sceptical about the practicalities and utility of having an integrated and shared dataset. • Expensive to standardise software and approaches. • Reluctance of companies to share their private commercial products. • Commissioner expectations and demands vary wildly across localities.
Opportunities	**Threats**
• Great potential for a rich and well powered resource for research and commissioning decisions. • Potential for geographical comparisons in both KPIs (from commissioner-level dataset) and individual level lifestyle/socio-demographic data (provider-level dataset). • Potential for comparison across and between the 4 key public health lifestyle areas (would allow for comparison between individually commissions vs integrated lifestyle services.	• Data are sensitive and there is variation in the extent to which organisations are prepared to fully share their data. • Difficulties in fully establishing data sharing agreements across all parties. • Unanswered questions about Information Governance and location/security of dataset. • Mechanisms for data access and permissions might be difficult to establish. • Need to fully understand user involvement and consent to use personal (albeit anonymised) data for research/commissioning purposes. • May be costly initially (*to migrate from current dataset to a new integrated dataset)* with cost benefits not emerging until dataset is fully established high risk as difficult to accurately project/model). • Lifestyle service providers seem to be in competition for contracts and may not want to data share because of fear of revealing their service provision strategy.

Guided by the specific findings from the consultation exercise stated in Table 2 and the SWOT Analysis Table 4 above, and following the Traffic Light System, a decision of 'Amend/amber was made. There is a need to conduct further and indepth consultations with more stakeholders, particularly service users and providers and of particular importance, IT Governance and Data Protection experts, to explore how best barriers to the database construction and implementation can be overcome.

6.0 POSITION STATEMENT AND OPTIONS TO APPRAISE FOR FUTURE DELIVERY MODEL

The evidence from the literature review suggests that, guided by an appropriate model and taking into consideration local factors, an Integrated East Midlands Lifestyle Database could be developed and implemented. Findings from the consultation exercise suggest that stakeholders would also welcome it because of the potential benefits including provision of a rich and well powered resource for commissioning and research, and facilitating access to lifestyle data.

However, as presented in Table 3, key concerns were raised and would need to be addressed prior to the successful development and implementation of an integrated dataset across the East Midlands.

Given the findings from the stakeholder consultation, we are proposing the following six options (summarised in Table 5 below) which may be examined as a first stage to establishing the dataset:

Option 1: A fully integrated lifestyle **individual level** dataset across the whole East Midlands covering all four lifestyle areas, with governance and access controlled by one institution (possibly a Local Authority or a university) that would house and maintain the database.

In terms of strengths, this will be a potentially useful resource for research, service commissioning and ease of information access.

> *"Just in terms of sharing best practice and looking at makeup of the community or the geographical sort of, you know….and comparing similar areas to ourselves. And looking at performance in terms of what are they doing, what are they doing differently that we could perhaps adapt and use locally?"* (Lifestyle service manager).
>
> *"It would give you the opportunity to compare how well things are doing in Lincolnshire compared to say Nottinghamshire because we would have both sets of information, yeah"* (Service User).

> "*Where to go to for your particular needs - you've got five minutes and you want to know about smoking cessation services in Derbyshire. Where would you go to? You've got half an hour to do a bit deeper dive. Where would you go to?"* (Health & Wellbeing Programme Lead).

However, there may be barriers to this option as there are a wide variety of commissioned services, unstandardised collection of variables, unaligned IT systems and related cost.

> *"Yeah, …. it's about the standardising of what people would commission, and another service which may be a private provider. Or it might be an NHS provider. So, provided you've got those two levels of understanding ……"* (Health Improvement Principal).

> *"I think you need to talk to providers…… and Commissioners directly about it; I think there are insurmountable barriers that cost money"* (Associate Director, Local Knowledge & Intelligence).

> *"Time and capacity to support any implementation as well as any budgetary or procurement processes or other local arrangements"* (Public Health Manager).

Option 2: A fully integrated **individual level** dataset for all four lifestyle areas, within just one geographical area to start with, as is currently in place in Lincolnshire, which is owned by the service provider.

The strengths in option 1 also apply to this option in terms of supporting research, service commissioning and ease of information access within one geographical area or local authority. However, there may be barriers relating to how to make this available more widely, as the providers only report collated data back to the commissioners. There is also a need to fully understand user involvement and consent to use personal (albeit anonymised) data for research/commissioning purposes. For example, a person accessing just one element of a lifestyle service would need to consent to their data being used in a merged dataset and this would need to be secured at point of service entry.

> *"…and you have to have, I guess consent from the service users as well, it's going to be shared in a different way. You know, it depends on how the providers want it shared, because many of them will likely have some agreement already*" (Consultant in Public Health).

> *"Data sharing would have to be done with client consent and how reasonable is it to ask clients to agree to their data being shared with numerous other stakeholders with no direct involvement in whatever intervention they're receiving?"* (Public Health Intelligence and Insight Manager).

Option 3: A fully integrated **individual level** dataset initially starting with one lifestyle area (possibly smoking which already has a standardised KPI) across the whole region, (to be rolled out later subject to success), with governance and access controlled by the institution (either a Local Authority or a local university) that will house the database. As in option 2, this option also has the strength of supporting research, service commissioning and ease of information access. The barriers relating to service user consent before sharing their data also apply to this option.

Option 4: An integrated **aggregate level** dataset covering all four lifestyle areas (reporting similar KPIs as is done currently by service providers who report back to their commissioners), across the whole East Midlands, with governance and access controlled by one institution (possibly a Local Authority or a university) that will house and maintain the database. As in options 1, 2 and 3, the option has the benefit of supporting research, service commissioning and ease of information access. The aggregate data in this option is also unlikely to have easily identifiable service user personal information. However, this option may not fully support a more detailed research involving stratified individual service user data.

Option 5: An integrated **aggregate level** dataset for all four lifestyle areas, within just one geographical area to start with, as we have in Lincolnshire, which is owned by the service provider. There is a need to consider how to make this available more widely, as the providers only report collated data back to the commissioners. This is the model already used in Lincolnshire.
This option also shares with the other options, the strength of supporting research, service commissioning and ease of data access. Since the data are aggregated, there is also the barrier of not fully supporting research work involving a detailed and stratified individual service user data analysis. There is also the barrier of first exploring the willingness of service users for their data to be shared with other users, apart from their service provider and commissioner, even within the same Local Authority.

Option 6: An integrated **aggregate level** dataset initially starting with one lifestyle area (possibly smoking which already has a standardised KPI) across the whole region, (to be rolled out later subject to success), with governance and access controlled by the institution (either a Local Authority or a local university) that will house the database.

As with the other five options, option 6 has the benefit of supporting research, service commissioning and ease of information access. There is also the barrier of not supporting a more detailed and stratified individualised research data analysis since the data are aggregated. In addition, service user consent needs to be explored before data sharing as already highlighted in the previous five options.

Table 5: Summary of options for consideration

	Individual level	Aggregate level
All four lifestyle areas across all geographical regions	Option 1	Option 4
All four lifestyle areas across one geographical region	Option 2	Option 5 (OYL model)
One lifestyle area across all geographical regions	Option 3	Option 6

The relative merits of each option are furtuer compared below in Table 6.

Table 6: A comparison between all options

	Considerations and comparisons						
Option for delivery	**Initial Cost**	**Complexity of IT Governance**	**Potential benefit as research & commissioning resource**	**Maintenance costs**	**Timescales**	**Complexity in Database housing & controller**	**Complexity of Ethics & accessing requirements**
Option 1 • A fully integrated lifestyle **individual level** dataset across the East Midlands • Covering all 4 lifestyle areas • Governance and access controlled by an institution (Local Authority or a university) that will house and maintain the database	**++++ +**	**+++++**	**+++++**	**+++++**	**+++++**	**+++++**	**+++++**
Option 2 • A fully integrated **individual level** dataset for all 4 lifestyle areas, within just one geographical area to start with • As in Lincolnshire - owned by the service provider. • Need to consider how to make this available more widely, as providers only report collated data back to commissioners	**++++**	**++++**	**++++**	**++++**	++++	**++++**	**++++**
Option 3 • A fully integrated **individual level** dataset starting with one health area (possibly smoking - already has a standardised KPI) across the whole region • To be rolled out later subject to success • Governance and access controlled by an institution (Local Authority or a university) that will house the database	**++++**	**++++**	**++++**	**++++**	++++	**++++**	**++++**
Option 4 • An integrated **aggregate level** dataset covering all 4 lifestyle area (reporting similar KPIs as is done currently by service providers who report back to their commissioners), across the whole East Midlands • Governance and access controlled by one institution (Local Authority or a university) that will house and maintain the database	**+++**	**+++**	**+++**	**+++**	**+++**	**+++**	**+++**

Option 5 • An integrated **aggregate level** dataset for all four lifestyle areas, within just one geographical area to start with • As we have in Lincolnshire, which is owned by the service provider. • Need to consider how to make this available more widely, as the providers only report collated data back to the commissioners	**+++**	**+++**	**+++**	**+++**	**+++**	**+++**	**+++**
Option 6 • An integrated **aggregate level** dataset starting with one lifestyle area (possibly smoking which already has a standardised KPI) across the whole region • To be rolled out later subject to success • Governance and access controlled by the institution (Local Authority or a university) that will house the database	**++**	**++**	**++**	**++**	**++**	**++**	**++**

7.0 Next steps

The project to date has started to reveal the complexity of the ambition to develop and implement a fully integrated lifestyle data set pertaining to the four key areas of public health (smoking, physical activity, alcohol consumtion, diet/weight management) examined in this report. As such, a number of further research questions, worthy of in-depth examination are emerging.

In response to this, efforts have already been made to secure funding to develop this work further:

A Bid was developed and submitted to the NIHR by Dr Ros Kane, Professor Derek Ward, Professor Graham Law (a Medical Statistician at the University of Lincoln), Dr Joseph Aaknuwe and a service user representative who was a co-applicant. This was unsuccessful but helpful feeback was received and this is informing the resubmission of the bid to alternative funding sources, which are currently being explored.

Dr Ros Kane and Dr Joseph Aaknuwe also secured £10,066 funding from the university to support this project which was utlised to extend this current study to allow a slightly wider consultation exercise.

For dissemination of the findings, the communication plan below will be followed.

8.0 Recommendations

The proposed options need to be considered for the development and implementation of an integrated lifestyle database for the East Midlands, and a much greater understanding of the costs and governance issues are needed. There is also a need to consider the issues uncovered about who would own and maintain the data and how data will be accessed. Further and indepth consultations with more stakeholders, particularly service users and providers, and of particular importance, IT Governance and Data Protection experts, to explore how best barriers to the database construction and implementation can be overcome. Further work should seek to:

- Examine, under what conditions people would be willing to work collaboratively with a shared dataset across the East Midlands region
- Examine any potential facilitators (benefits) and barriers to data access/collation and sharing
- Assess the potential usability of such a dataset to Local Authorities and academics
- Produce a toolkit, outlining the key things to be considered (how these could be addressed) in the construction of such a regional dataset.

REFERENCES

Arksey, H. and O'Malley, L. (2005) Scoping studies: towards a methodological framework. Int J Soc ResMethodol 8(1):19–32.

Avery, K.N.L., Williamson, P.R., Gamble, C. et al. (2017) Informing efficient randomised controlled trials: exploration of challenges in developing progression criteria for internal pilot studies. BMJ Open;7: e013537.

Bauld, L., Bell, K., McCullough, L. et al. (2010) The effectiveness of NHS smoking cessation services: a systematic review. J Public Health (Oxf); 32(1).

Biswas, A., Oh, P.I., Faulkner, G.E. et al. (2015) Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults. Ann Intern Med. 162:123.

Bottle, A., Cohen, C., Lucas, A. et al. (2020) How an electronic health record became a real-world research resource: Comparison between London's Whole Systems Integrated Care database and the Clinical Practice Research Datalink. BMC Med Inform Decis Mak; 20: 71–71.

Clarke, A. and Steele, R. (2012) "Summarized data to achieve population-wide anonymized wellness measures," 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, pp. 2158-2161.

Department of Health and Social Care (2021) Integration and innovation: working together to improve health and social care for all. Policy paper. Available at https://www.gov.uk/government/publications/working-together-to-improve-health-and-social-care-for-all/integration-and-innovation-working-together-to-improve-health-and-social-care-for-all-html-version. Accessed: 03/03/2021.

Department of Health and Social Care (2020) New obesity strategy unveiled as country urged to lose weight to beat coronavirus (COVID-19) and protect the NHS. ttps://www.gov.uk/government/news/new-obesity-strategy. Accessed: 12/01/2020.

Doiron, D., Burton, P., Marcon, Y. et al. (2013) Data harmonization and federated analysis of population-based studies: the BioSHaRE project. Emerg Themes Epidemiol;10:12

Health and Social Care Act (2012) c.7. Accessed: 20/09/2020

Lakervield, J., Loyen, A., Ling, F.C.M. et al. (2017) Identifying and sharing data for secondary data analysis of physical activity, sedentary behaviour and their determinants across the life course in Europe: general principles and an example from DEDIPAC. *BMJ Open*;7: e017489.

Lewer, D., Bourne, T., George, A. et al. (2018) Data Resource: the Kent Integrated Dataset (KID). Internat J Population Data Sci.; 3:6.

Li, S., Zhang, L., Liu, S., Hubbard, R., and Li, H. (2020). Surveillance of Noncommunicable Disease Epidemic Through the Integrated Noncommunicable Disease Collaborative Management System: Feasibility Pilot Study Conducted in the City of Ningbo, China. Journal of medical Internet research, 22(7), e17340. https://doi.org/10.2196/17340.

Parkin, D.M., Boyd, L., Walker, L.C. (2011) The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. British Journal of Cancer; 105:S77-S81.

Pisani, E., AbouZahr, C. (2010) Sharing health data: good intentions are not enough. Bull World Health Organ; 88:462-6.

Piwowar, H.A., Becich, M.J., Bilofsky, H. et al. (2008) Towards a data sharing culture: recommendations for leadership from academic health centers. PLoS Med; 5:e183-9.

Rehm, J., Baliunas, D., Borges, G.L. et al. (2010) The relation between different dimensions of alcohol consumption & burden of disease: an overview external icon. Addiction; 105(5):817-43.

Ritchie, J., Spencer, E. (1994) Qualitative data analysis for applied policy research In Bryman A, Burgess RG, Eds. Analyzing Qualitative Data. London: Routledge.

Saunders, P., Mathers, J., Parry, J., Stevens, A. (2001). "Identifying 'non-medical' datasets to monitor community health and well-being." Journal of Public Health 23(2): 103-108.

Sarkar, C., Webster, C., Gallacher, J. (2015) UK biobank urban morphometric platform (UKBUMP) - a nationwide resource for evidence-based healthy city planning and public health interventions. Ann GIS. 21:135–48.

Trefan, L., Akbari, A., Paranjothy, S. et al. (2019) Electronic Longitudinal Alcohol Study in Communities (ELAStiC) Wales – protocol for platform development. Int J Popul Data Sci.;4(1).

WHO (2016) Physical activity factsheet N°385. Accessed: 24/11/20.

Wilmot, E.G., Edwardson, C.L., Achana, F.A. et al. (2012) Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis. Diabetologia; 55:2895-905.

WHO (2009) Global health risks: mortality and burden of disease attributable to selected major risks.

WHO (2014) Global status report on alcohol and health-2014 external icon. Geneva.

Zwisler, A.D., Rossau, H.K., Nakano, A. et al. (2016) The Danish Cardiac Rehabilitation Database. *Clin Epidemiol*. 2016; 8:451-456.

APPENDICES

Appendix 1: Initial Gantt Chart to steer project

Task No	Activity	Month/Week : 01/09/20 - 18/03/2021																												
		Sept Week: 1-4				Oct Week: 5-8				Nov Week: 9-12				Dec Week: 13-16				Jan Week: 17-20				Feb Week: 21-24				Mar Week 25-26				
1	Project initial meeting University of Lincoln Research office will scope potential funding opportunities from the beginning of the project (September 2020).																													
2	• Developing plan of work • Project registration application																													
Meeting with Prof Derek Ward/steering group on 02st October																														
3	Scoping literature search of existing lifestyle research nationally (for draft of main key points of the review). If sufficient evidence is located this will be worked up into a publishable article by the end of the full study period – 18/03/2020.																													
	Write interim report: indications of feasibility and emerging research questions – 16/11/2020 -15/12/2020																													
4	Consultation with key stakeholders to assess the feasibility of developing an East Midlands Integrated Lifestyle service Database including any potential issues with data access/collation and sharing and the potential usability of such a dataset to Local Authorities. • Initial activity will be to identify and agree the individual personnel and organisations to be invited to participate in the consultations and the number of consultations required/feasible. • A template to guide the content of the interviews will be agreed with the funders. • Discussions will take place via telephone conversations, emails and/or meetings on MS Teams – End date: 30/11/2020																													
5	Develop a research bid to NIHR Local Authority or Rapid Funding Programme to support the implementation of the data-set or address issues arising from the consultation exercise– End Date: 31/12/2020																													
Progress report/meeting with steering group/DsPH – 18th Dec 2020																														
6	Develop (and have approved) a communication plan to disseminate key findings form the scoping literature search and consultation exercise – End Date: 31/01/2021																													
7	Position statement, delivery model and implementation plan – End Date: 31/01/2021																													
8	Disseminate to key stakeholders regionally (as will be agreed in the communication plan above): February- April 2021.																													

Appendix 2: Protocol for screening, selection of studies and extraction of data from selected studies

Introduction

The social care legislation (Health and Social Care Act, 2012) places specific duties on county councils to protect and promote health, and reduce health inequalities. Local authorities across the East Midlands deliver lifestyle services to communities, which include smoking cessation, improvement in diet, reduction in physical inactivity and reduction in alcohol consumption. While some authorities offer stand-alone services, others have implemented innovative integrated service models.

In the pursuit of improved health outcomes authorities are evaluating their services through a range of approaches. The national evidence base needs enhancing to drive improvement and efficiency. This could be done by exploring the potential to have one integrated dataset bringing together information on public health lifestyle interventions across the East Midlands. This review will explore the literature for evidence of whether it is feasible to develop and implement an integrated public health lifestyle data set in the East Midlands. The objectives of the review will include:

- To map the existing peer-reviewed literature on feasibility of implementing a public health lifestyle data set
- To determine the type and extent of the evidence available
- To identify any gaps for further research

METHODS

We will follow Arksey and O'Malley's framework for this scoping review: identifying the research question; identifying relevant studies; selecting the studies; charting the data; collecting, summarising and reporting results (Arksey & O'Malley, 2005).

Identifying the research question

The review will address the question: What is known about the development and implementation of a public health lifestyle data set?

Identifying relevant studies (Searches)

The following inclusion and exclusion criteria will guide the search strategy.

Inclusion criteria:

- There will be no limits to the publication dates since we are not sure of the extent of the evidence available.

- Only studies in the English language will be included due to lack of funds for translating studies in other languages.
- Peer-reviewed academic literature
- Study design will be quantitative, qualitative and mixed methods
- Any age group (adults and children)
- The focus will be on any public health lifestyle data set
- Any geographical location

Exclusion criteria:

- Studies not in the English language
- Non-peer reviewed literature
- Non-public health lifestyle data set

Using the search terms (public health AND (lifestyle OR "life style" OR life-style)) AND (database* OR dataset* OR "data set*") AND (feasib* OR develop* OR implement* OR use), we will search the following electronic databases (Medline, CINAHL, Cochrane, Scopus, Psych-INFO).The reference lists of studies found through the electronic database searches will be checked for relevant studies using Google Scholar search.

Selecting the studies

We will follow the PRISMA Extension for Scoping Reviews (PRISMA-ScR). Two reviewers will independently screen all titles and abstracts in line with the inclusion and exclusion criteria. Any disagreement between the two reviewers over the eligibility of particular studies will be resolved through discussion with a third reviewer. After the title and abstract screening, full texts of the eligible articles will be retrieved and screened. The study selection process will be presented in a PRISMA flowchart.

Charting the data

The data will be charted according to an analytical framework that will facilitate sorting the material into a data extraction table. The table will be created by the research team to meet the study objectives. Data will be charted by one researcher (JA) and will be checked by two members of the research team (RK and HH). Standard information such as authors, year of publication, study setting, aim, methods, study population, findings and country/location will be extracted from all included articles and charted.

Collating, summarising and reporting results

The findings from the included studies will be collated, summarised into themes using a thematic analysis approach. This will initially be done by one reviewer (JA) and then checked by two members of the research team (RK and HH). We will not conduct a quality appraisal of the included studies since scoping reviews usually aim to only provide a descriptive account of the evidence (Coughlan & Cronin, 2017).

References

Arksey, H, O'Malley, L. (2005) Scoping studies: towards a methodological framework. Int J Soc ResMethodol 8(1):19–32.

Coughlan, M., Cronin, P. (2017) Doing a literature review in nursing, health and social care, 2nd edn. SAGE, London.

Health and Social Care Act 2012, c.7. Available at: https://www.legislation.gov.uk/ukpga/2012/7/contents/enacted (Accessed: 20/09/2020).

Appendix 3: Details of 22 studies excluded from the review

Excluded studies after full text screening of 31 studies for eligibility

The 22 studies listed below were generated from the systematic search strategy, but these did not relate to datasets involving any of the lifestyle of interest: smoking, alcohol consumption, physical activity and diet and exercise.

1. Kinsner-Ovaskainen, A.,, Lanzoni, M., Garne, E., Loane, M., Morris, J., Neville, A., Nicholl, C., Rankin, J., Rissmann, A., Tucker, D., Martin, S. (2018) A sustainable solution for the activities of the European network for surveillance of congenital anomalies: EUROCAT as part of the EU Platform on Rare Diseases Registration. European journal of medical genetics.

2. Ordoñana, J.R., Carrillo, E., Colodro-Conde, L., García-Palomo, F.J., González-Javier, F., Madrid-Valero, J.J., Martínez Selva, JM., Monteagudo, O., Morosoli, J.J., Pérez-Riquelme, F., Sánchez-Romera, J.F. (2019) An Update of Twin Research in Spain: The Murcia Twin Registry. Twin research and human genetics: the official journal of the International Society for Twin Studies.

3. Thompson, M.L., Miller, R.S., Williams, M.A. (2007) Construction and characterisation of a longitudinal clinical blood pressure database for epidemiological studies of hypertension in pregnancy. Paediatric and perinatal epidemiology.

4. Godderis, L., Mylle, G., Coene, M., Verbeek, C., Viaene, B., Bulterys, S., Schouteden, M. (2015) Data warehouse for detection of occupational diseases in OHS data. Occupational medicine (Oxford, England).

5. Sugiyama, T., Miyo, K,. Tsujimoto, T., Kominami, R., Ohtsu, H., Ohsugi, M., Waki, K., Noguchi, T., Ohe, K., Kadowaki, T., Kasuga, M., Ueki, K., Kajio, H. (2017) Design of and rationale for the Japan Diabetes compREhensive database project based on an Advanced electronic Medical record System (J-DREAMS). Diabetology international.

6. Ogushi, Y., Haruki, Y., Okada, Y., Takahashi, M., Shimizu, M., Izumi, Y., Watabe, T., Kobayashi, S., Okuyama, J., Kurita, Y. (1998) Development and evaluation of regional health database systems. Studies in health technology and informatics,book chapter

7. Austin, M.A., Harding, S., McElroy, C (2003). Genebanks: a comparison of eight proposed international genetic databases. Community genetics.

8. Baus, A., Wood, G., Pollard, C., Summerfield, B., White, E. (2013) Registry-based diabetes risk detection schema for the systematic identification of patients at risk for diabetes in West Virginia primary care centers. Perspectives in health information management.

9. Howell, N.A., Tu, J.V., Moineddin, R., Chen, H., Chu, A., Hystad, P., Booth, GL. (2019) The probability of diabetes and hypertension by levels of neighbourhood walkability and traffic-related air pollution across 15 municipalities in Southern Ontario, Canada: A dataset derived from 2,496,458 community dwelling-adults. Data in brief.

10. Hwee, L.J.,Witarsyah, D., Kasim, S., Fudzee, M.F.M. (2020) Healthy food intake advisor using decision support system.

11. Armstrong, J., Rudkin, J.K., Allen, N., Crook, D.W., Wilson, D.J., Wyllie, D.H., O'connell, A.M. (2020) Dynamic linkage of covid-19 test results between public health England's second generation surveillance system and uk biobank.

12. Prosperi, M., Min, J.S., Bian, J., Modave, F. (2018) Big data hurdles in precision medicine and precision public health. BMC Medical Informatics and Decision Making.

13. Yoo, H., Chung, K. (2018) Mining-based lifecare recommendation using peer-to-peer dataset and adaptive decision feedback. Peer-to-Peer Networking and Applications.

14. Doiron, D., Burton, P., Marcon, Y., Gaye, A., Wolffenbuttel, B.H.R., Perola, M., Stolk, R.P., Foco L., Minelli, C., Waldenberger, M., Holle, R., Kvaløy, K., Hillege, H.L., Tassé, A.-M., Ferretti, V., Fortier I. (2013) Data harmonization and federated analysis of population-based studies: The BioSHaRE project 2013. Emerging Themes in Epidemiology.

15. Sak, J., Pawlikowski, J., Goniewicz, M., Witt, M. (2012) Population biobanking in selected European countries and proposed model for a Polish national DNA bank.Journal of Applied Genetics.

16. Malmberg, G., Nilsson, L.G., Weinehall, L. (2020) Longitudinal data for interdisciplinary ageing research. Design of the Linnaeus Database. Scandinavian Journal of Public Health.

17. Hansen, W., Kalapasev, N., Gillespie, A., Singler, M., Ball, M. (2009) Development of a pedestrian walkability database of northern Kentucky using Geographic Information Systems (GIS).. Journal of Physical Activity and Health.

18. Mauri, D., Pazarlis, P., Mauri, J., Altinoz, H., Flores, F.J., Karentzou, I., Proiskos, A., Lakiotis, V., Maragkaki, A., Terzoudi, E., Dambrosio, M., Spiliopoulou, A., Varsami, A., Alexandropoulou, P., Tolis, C., Pavlidis, N., Vittoraki, A. (2004) SESy-Europe: A multi-language database dedicated to cancer screening monitoring.Journal of Experimental and Clinical Cancer Research.

19. Godard, B., Marshall, J., Laberge, C., Knoppers, B.M.)2004) Strategies for Consulting with the Community: The cases of four large-scale genetic databases.Science and Engineering Ethics.

20. Austin, M.A., Harding, S., McElroy, C. (2003) Genebanks: A comparison of eight proposed international genetic databases. Community Genetics.

21. Stern, R.M., Tarkowski, S. (1990) The need for a unified European environmental health database. Information Services and Use.

22. Ng, N; Van Minh, H; Tesfaye, F; Bonita, R; Byass, P; Stenlund, H; Weinehall, L; Wall, S. (2006) Combining risk factors and demographic surveillance: potentials of WHO STEPS and INDEPTH methodologies for assessing epidemiological transition. Scandinavian Journal of Public Health..

Appendix 4: Existing and relevant databases (in primary care, secondary care and public health)

No	Database	Type of data	Data controller	Governance	IT system	Data access	Strengths	Limitations
1	Kent integrated dataset (KID)	Data from council adult social care, GP practices, mental health services, public health & community health at an individual, but de-personalised level	Each organisation remains as DC for its own data e.g GP surgeries, Acute Trusts, Adult Social Care etc	General Data Protection Regulation 2016	Built from existing systems, using a SQL-server data warehouse (The Kent and Medway Health Informatics Services) and a purchased business intelligence tool	Licensed access available for research that benefits the Kent and Medway rgion	1. Person level data linking routinely collected administrative activity and cost data from almost all NHS providers across Kent and many non NHS organisations 2. Same NHS number throughout the dataset so each contact with a service is traceable 3. Personal data is anonymised	1. Completeness of data 2. Sensitivity of partnership arrangements 3. Access control policy 4. Resource constraints
2	QResearch	A large consolidated database derived anonymised health records of over 45 million patients from over 1800 general practices	Individual General practices BUT University of Oxford is the sole data controller for individual and the resulting database including the linked datasets	University of Oxford is sole data controller for the individual and linked datasets Has continued ethical approval from East Midlands - Derby Research Ethics Committee	EMIS clinical computer system	At a cost to keep database running	Following review by the QResearch Advisory Board, East Midlands - Derby Research Ethics Committee and the Ethics and Confidentiality Committee of the National Information Governance Board, the entire database has been linked to cause of death data, cancer and hospital data at individual patient level with linkages extending back as far as 1993	All projects need to be funded to cover the costs of accessing the QResearch database

3	Clinical Practice Research Datalink (CPRD)	High quality anonymous primary and secondary health care data for Public Health & Medical Research	UK GP Practices	The Medicines and Healthcare products Regulatory Agency Researchers seeking access to CPRD data must undergo a rigorous approvals process including ethics	Patient electronic health record system	Subject to protocol approval by an advisory body	A real-world research service supporting retrospective and prospective public health and clinical studies	
4	The Health Improvement Network (THIN)	UK Primary Care database containing anonymised longitudinal patient records for approximately 6% of the UK Population (nationally representative)	UK Primary care	General Data Protection Regulation (GDPR) Regulated by an Advisory Committee Has ethics approval from Department of Health's South East Multicentre Research Ethical Committee	Patient electronic record system	Contact THIN team	Anonymised data which goes back to 1994 and is nationally representative of the UK population	
5	Fingertips - Local Authority Health Profiles	Population level local authority data	Local authorities			Freely available at: https://fingertips.phe.org.uk/profile/health-profiles	Freely available online	Data not at individual person level

6	NHS Digital		-Sourced from surveys -The Strategic Data Collection Service (SDCS) is a secure data collection system used by health and social care organisations to submit data to NHS Digital			Freely available at: https://digital.nhs.uk/data-and-information/areas-of-interest/public-health/lifestyles	Freely available online	Data not available at individual person level
7	Oxford RCGP Research and Surveillance Centre					https://www.phc.ox.ac.uk/research/covid-19/projects/oxford-rcgp-research-and-surveillance-centre	Freely available	
8	UK Data Archive					https://www.data-archive.ac.uk/resources/	Certified to ISO 27001, the highest international security standard	
9	BioBank					https://www.ukbiobank.ac.uk/		
10	University of Bath Research Data			https://researchdata.bath.ac.uk/information.html		https://researchdata.bath.ac.uk/cgi/search/simple?q=demographics&_action_search=Search&_action_search=Search&_order=bytitle&basic_srcht	Openly discoverable/downloadable	

						ype=ALL&_satisfyall=ALL		
11	ONS					https://www.ons.gov.uk/	Available online	
12	UCL longitudinal data					https://cls.ucl.ac.uk/data-access-training/		
13	Oxford Research Archive					https://ora.ox.ac.uk/		
14	Experian dataset					https://www.experian.co.uk/business/data-management/data-purchasing/data-sets/	Broad range of data available	Data access is at a cost
15	The Secure Anonymised Information Linkage (SAIL) Databank at Swansea University							
16	General Practice Repository for ClinicAL Care (GPRCC)				eHealthScope		GPRCC gives eHealthScope the data to be able to inform GPs of specific care needs in their area, so they can create solutions such as drop in services to address healthcare needs before they become a problem.	

Appendix 5 Data collection tool utlised during the consultation exercise

UNIVERSITY OF
LINCOLN

East Midlands Integrated Lifestyle (ILS) service Database- feasibility study

Topic guide

September 2020

Dr Joseph Akanuwe, Dr Ros Kane and Dr Hannah Henderson
School of Social Sciences
University of Lincoln
Brayford Pool
Lincoln
LN6 7TS
jakanuwe@lincoln.ac.uk/rkane@lincoln.ac.uk/hhenderson@lincoln.ac.uk

This project is a feasibility study to explore the development of an overarching Public Health Lifestyle dataset to support the services delivered by Local Authorities across the East Midlands, with the aim of improving lifestyle and health outcomes. The project will consider the options for constructing such a dataset and will develop a delivery model for putting it in place. Focusing on four key lifestyle areas (smoking cessation, reduction in alcohol consumption, reduction in physical inactivity and diet/weight management, the following questions will guide a consultation exercise with stakeholder.

Please write your responses in the space provided if you wish to

Background questions

1. Your local authority ……
2. Your current role .....
3. We have identified the following key public health lifestyle interventions:

- Smoking cessation
- Reduction in physical inactivity
- Reduction in alcohol consumption
- Diet/weight management

Are there any other key lifestyle areas that may be added?

Questions on datasets

4. How do you currently use lifestyle intervention datasets? (*Implementation*)

5. Are you willing to share details of the template (a blank lists of variables) used to collect data across the lifestyle areas, so we can determine whether the variables can be linked to form one database? (*Expansion*)

6. Do you have any examples of datasets that can be shared - to show how the variables are mapped out/collected? (Expansion, adaptation)

7. What datasets exist in your area/that you are aware of? (*Practicality*)

8. Are the datasets in your area separate (for different lifestyles: smoking cessation, alcohol consumption, reduction in physical inactivity and diet/weight management) or integrated? (*Integration*)

9. How are the datasets used to inform service development? (*Demand, Implementation*)

10. How can a shared dataset fit into your lifestyle service? (*Integration*)

11. How can an existing dataset be modified for shared use across East Midlands? (*Adaptation*)

Questions about ownership/storage/sharing agreements with datasets

12. How useful would a shared data set across the whole East Midlands be? (*Acceptability, demand, limited efficacy*)

13. What are the advantages and disadvantages of a shared dataset? (*Practicality*)

14. Are there any barriers to creating an East Midlands wide dataset? If yes, what are they? (*Practicality*)

15. What factors will facilitate the implementation of a shared dataset across East Midlands? (*Practicality*)

16. Who are the datasets used by? (*Acceptability, Demand*)

17. How cost effective will a shared dataset across East Midlands be? (*Practicality*)

18. How is evidence base used (how can it be used) to steer decision making in your organisation (*practicality)*?

19. Who else should we be speaking to in your organisation?

20. Where are data shared already and with whom?

21. Any further comments on developing and implementing a shared dataset in East Midlands?

THANK YOU FOR YOUR TIME AND SUPPORT

Appendix 6 Details of lifestyle services and providers in the East Midlands

County Council	District/Borough/City Councils	Health area	Lifestyle service delivered (Y/N)	Provider
Lincolnshire County Council www.lincolnshire.gov.uk	East Lindsey District Council West Lindsey District Council Boston Borough Council South Holland District Council South Kesteven District Council North Kesteven District Council Lincoln City Council	Physical activity Alcohol Smoking cessation Weight management	Y Y Y Y	OneYou Lincolnshire
Nottinghamshire County Council www.nottscc.gov.uk	Ashfield District Council Bassetlaw District Council Broxtowe Borough Council Gedling Borough Council Mansfield District Council Newark and Sherwood District Council Rushcliffe Borough Council Nottingham City Council	Physical activity	Y	1. Your Health, Your Way - Physical activity 2. Base 51 – Gym 3. Chilwell Olympia Karate Club 4. Mansfield InStep Health Walks 5. Running Club - Holme Pierrepont 6. Southwell Running Club 7. West Bridgford Tai Chi - Chen style 8. Wheatley Tennis Club
		Alcohol	Y	1.Your Health, Your Way - Alcohol reduction 2. Addiction Therapy Care
		Smoking cessation	Y	1.Your Health, Your Way - Smoking cessation 2. Giving Up Smoking
		Weight management	Y	1.Your Health, Your Way - Weight Management 2. Academy of Life - Personal Health Management 3. Change4Life
Leicestershire County Council www.leics.gov.uk	Harborough District Council Charnwood Borough Council Blaby District Council North West Leicestershire District Council Oadby and Wigston Borough Council Melton Borough Council Hinckley and Bosworth Borough Council Leicester City Council	Physical activity Alcohol Smoking cessation Weight management	Y Y Y Y	Live Well Leicester (Integrated Lifestyle Serevice)

Northamptonshire County Council www.northamptonshire.gov.uk	Corby Borough Council Daventry District Council Kettering Borough Council East Northamptonshire District Council South Northamptonshire Council Wellingborough Borough Council Northampton Borough Council	Physical activity	Y	1.Northamptonshire Sport Get Active 2. Activity on Referral (Referral by GP) 3. Run In England (self referral) 4. Walking for health (Self referral) 5. First for Wellbeing
		Alcohol	Y	Drug & Alcohol Advice Service – Substance to Solutions
		Smoking cessation	Y	Northamptonshire Stop Smoking Service (First for Wellbeing) www.smokefree.nhs.uk
		Weight management	Y	1.Healthy Weight Change4life 2. Weightwatchers (self referral or by GP)
Derbyshire County Council www.derbyshire.gov.uk	Amber Valley Borough Council Chesterfield Borough Council Derbyshire Dales District Council High Peak Borough Council North East Derbyshire District Council Peak District National Park Authority Bolsover District Council South Derbyshire District Council Erewash Borough Council Derby City Council	Physical activity	Y	1.Live Life Better Derbyshire 2. Exercise by Referral Derbyshire Scheme 3. Active Derbyshire -Walking for heath scheme -Jog Derbyshire
		Alcohol	Y	Live Life Better Derbyshire
		Smoking cessation	Y	Live Life Better Derbyshire
		Weight management	Y	1.Live Life Better Derbyshire 2. Heart of Derbyshire
Rutland County Council www.rutland.gov.uk		Physical activity	Y	1. Active Rutland 2. Rutland Healthy Walks
		Alcohol Smoking cessation Weight management		Unable to ascertain

Appendix 7 Routinely collected lifestyles variables

Table 3: Variables routinely collected at individual level

Lifestyle variables routinely collected	Lincolnshire County Council	Nottingham-shire County Council	Leicester-shire County Council	Rutland County Council	Northamp-tonshire County Council	Derby-shire County Council
Smoking cessation variables collected (please add more rows as required):	Please tick (√)	Please tick (√)	Please tick (√)	Please tick (√)	Please tick (√)	Please tick (√)
Do you smoke in the home/car?	√	√				
WEMWBS score	√	√				
How ready are you to quit?	√	√			√	
Have you got friends and family to help you in your quit attempt?	√	√			√	
How soon after you wake up do you smoke your first cigarette?	√	√			√	
Do you find it difficult to stop smoking in no-smoking areas?	√	√			√	
Which cigarette would you hate most to give up?	√	√			√	
How many cigarettes per day do you usually smoke?	√	√			√	
Do you smoke more frequently in the first hours after waking than during the rest of the day?	√	√			√	
Do you smoke if you are so ill that you are in bed most of the day?	√	√			√	
Audit C questionnaire	√	√			√	
Do any members of your household smoke?	√	√			√	
Ask if the smoker would like an internal referral to following service - Weight management, Physical activity, Falls, Alcohol	√	√			√	
Ask about past quit attempts	√	√			√	
Ask the client to Set the Quit Date	√	√			√	
Have you any potential high-risk situations in the coming week?	√	√			√	
How committed are you to follow the programme?	√	√			√	
Do you smoke an e-cigarette?	√	√			√	
Diet and Weight management variables collected:						
How many days a week do you eat breakfast?	√	√			√	
How many days a week do you eat lunch?	√	√			√	
How many days a week do you eat dinner?	√	√			√	
Do you plan your meals?	√	√			√	
How many snacks a day on average?	√	√			√	
How many portions of fruit do you eat a day?	√	√			√	

How many portions of vegetables do you eat a day?	√	√			√	
How many sugary drinks do you have a day?	√	√			√	
How many glasses of water or non sugary drinks do you drink a day?	√	√			√	
What do you think your portion sizes are?	√	√			√	
Have you tried to reduce your portion sizes?	√	√			√	
How many days a week do you eat out?	√	√			√	
Do you read food labels?	√	√			√	
Do you follow a special diet or follow any specific foods?	√	√			√	
Regarding your current diet and eating pattern, can you run me through a typical day?	√	√			√	
Physical activity:						
During the last 7 days, on how many days did you do vigorous physical activities like heavy lifting, digging, aerobics, or fast bicycling?	√	√				
How much time in total did you usually spend on one of those days doing vigorous physical activities?	√	√				
During the last 7 days, on how many days did you do moderate physical activities like carrying light loads, bicycling at a regular pace, or doubles tennis? Do not include walking	√	√				
How much time in total did you usually spend on those days doing moderate physical activities?	√	√				
During the last 7 days, on how many days did you walk for at least 10 minutes at a time? This includes walking at work and at home, walking to travel from place to place, and any other walking that you did solely for recreation, sport, exercise or leisure?	√	√				
How much time in total did you usually spend walking on one of those day?	√	√				
During the last 7 days, how much time in total did you usually spend sitting on a week day?	√	√				
BORG Scale 0-10 Level of exertion	√	√				
Alcohol reduction:						
How often do you have a drink containing alcohol? (ALL CLIENTS)	√	√				
How many units of alcohol do you drink on a typical day when you are drinking? (ALL CLIENTS)	√	√				
How often have you had 6 or more units if female, or 8 or more if male, on a single occasion in the last year? (ALL CLIENTS)	√	√				
How often during the last year have you found that you were not able to stop drinking once you had started?	√	√				
How often during the last year have you failed to do what was normally expected from you because of your drinking?	√	√				

How often during the last year have you needed an alcoholic drink in the morning to get yourself going after a heavy drinking session?	√	√				
How often during the last year have you had a feeling of guilt or remorse after drinking?	√	√				
How often during the last year have you been unable to remember what happened the night before because you had been drinking?	√	√				
Have you or somebody else been injured as a result of your drinking?	√	√				
. Has a relative or friend, doctor or other health worker been concerned about your drinking or suggested that you cut down?	√	√				

Note: Further in-depth investigation is needed to fully confirm the accuracy of the above table as the process of identifying this information was very complex.
Unable to collect variables from Derbyshire county and Leicester City Council due to non-response to invitations to participate in stakeholder consultation.
Derby City council not represented at stakeholder consultation, but responses received by email.