MANCHESTER 1824
The University of Manchester

Evaluating the NHS Diabetes Prevention Programme (NHS DPP): the DIPLOMA research programme (Diabetes Prevention – Long term Multimethod Assessment)

Protocol

Evaluating the NHS Diabetes Prevention Programme (NHS DPP): the DIPLOMA research programme (**Di**abetes **P**revention – **Lo**ng term **M**ultimethod **A**ssessment)

Summary of Research

The NHS Diabetes Prevention Programme (NHS DPP) is an ambitious programme to deliver an evidence-based behavioural change intervention ('the NHS DPP intervention') to patients at risk of developing diabetes, to encourage behaviour change and reduce risk of diabetes. The size of the programme and the complexity of its implementation present significant challenges for any evaluation.

DIPLOMA research programme is a mixed methods evaluation delivered by an experienced multidisciplinary team, which is designed to provide:

(a) feedback regularly to NHS DPP stakeholders on the delivery and outcomes of the programme to support ongoing development and quality improvement

(b) a rigorous longer-term assessment of the success of the NHS DPP in meeting the aim of reducing diabetes prevalence in a way that is cost-effective and sustainable for the NHS.

We have the necessary skills and experience to deliver this evaluation. Our expertise includes quasi-experimental evaluations of population health initiatives,[1-4] use of routine datasets to support evaluation of policy,[5-12] mixed methods evaluation of regional and national policy,[13-27] design and evaluation of behaviour change interventions,[28-36] implementation science [37 38] and economic evaluation[39 40] as well as experience of the evaluation of local diabetes prevention schemes.[41 42]

The DIPLOMA research programme will have 8 work packages:

Work package 1 Access and equity will use quantitative analysis of national survey and administrative data to explore the characteristics of patients who enter the NHS DPP, and those who are eligible and do not enter. This analysis will be complemented by qualitative research on the process of accessing the NHS DPP, and the experience of patients and professionals.

Work package 2 Implementation will explore the process of implementation of the NHS DPP in selected sites, exploring the local organisation of the programme, workforce, funding, and pathways.

Work package 3 Service Delivery and Fidelity will analyse data on uptake and adherence to the NHS DPP intervention, combined with a detailed analysis of the degree to which providers adhere to the specification for the intervention.

Work package 4 Outcomes will use data collected by the NHS DPP to assess effects on patient outcomes within the programme. Data will also be linked to data from 'Implementation' and 'Service Delivery and Fidelity' work packages to explore factors influencing outcomes and 'active ingredients.'

Work package 5 Comparative Effectiveness will use quantitative analysis of administrative data to explore whether the NHS DPP leads to a reduction in the prevalence of diabetes and other outcomes, using quasi-experimental methods.

Work package 6 Validation sample will use a patient survey to evaluate the risk of confounding on participation in the NHS DPP, allowing adjustment for the effects of confounding.

Work package 7 Comparative Long Term Cost Effectiveness will use data from other work packages, the international literature and administrative data to create an economic model. This will compare NHS DPP costs with the benefits (reductions in diabetes prevalence, improvements in quality of life and mortality and reductions in health care utilisation) over the long-term

Work package 8 Programme management will include management of the programme, co-ordination of the research work packages, liaison with stakeholders and dissemination of findings to internal (NHS DPP) and external academic, service and patient audiences.

Background and Rationale

Non-diabetic hyperglycaemia is a term which is used to describe the decreased ability of the body to regulate glucose effectively, such as impaired glucose regulation (IGR), impaired glucose tolerance (IGT) and impaired fasting glucose (IFG). It accounts for conditions where blood glucose levels are above the normal range but are not high enough for a diagnosis of type 2 diabetes mellitus. People with non-diabetic hyperglycaemia often have no symptoms, but 5-10% of those with non-diabetic hyperglycaemia will go on to develop type 2 diabetes every year if left untreated.

The health implications of type 2 diabetes are serious, with poor control (i.e. high blood pressure / low medication adherence) resulting in loss of vision, nerve pain, and in severe cases, limb amputation. Further, type 2 diabetes carries with it a high-risk of developing other cardiovascular health complications.[43] Type 2 diabetes is thought to cost the NHS £10 billion per year, around nine per cent of the total NHS budget.

These figures highlight the importance of diabetes prevention as a national public health concern. The main causes of non-diabetic hyperglycaemia are behavioural in nature, including poor diet and limited physical activity.[44] Indeed, evidence suggests that making changes to lifestyle behaviours which reduce weight, such as increasing physical activity, can decrease the risk of non-diabetic hyperglycaemia developing into type 2 diabetes by 50%.[45] However, the asymptomatic nature of non-diabetic hyperglycaemia means that people often go undiagnosed and untreated, therefore remaining at a higher risk of developing type 2 diabetes.[46]

To tackle this problem, Diabetes Prevention Programmes have been developed and implemented worldwide, including the USA[47] and Finland.[48] Such programmes aim to reduce the incidence of diabetes by targeting dietary and physical activity behaviours of those considered at risk of developing type 2 diabetes. Large randomised controlled trials of these programmes have demonstrated that lifestyle interventions can reduce the risk of developing diabetes by up to 58%, through a relatively modest weight loss of 5-7%.[49] This illustrates the importance of weight loss, as the risk of diabetes was found to reduce by 16% for each kilogram of weight lost.[48] It also highlights the role of obesity in the rise of diabetes and supports the targeting of weight reduction for the prevention of diabetes.[44]

The NHS DPP intervention will be delivered by 4 providers procured by NHS England based on a published specification. Although some variability is expected, the core intervention will involve a predominantly group-based model delivered in person across a minimum of 9 months, including at 16 hours contact time, with a focus on diet, physical activity and weight loss. People with non-diabetic hyperglycaemia will be identified through NHS Health Checks and registers in primary care. There will be a focus on systematic data collection to support the programme.

We have previous experience with the delivery and evaluation of services for patients with non-diabetic hyperglycaemia, including the Salford Care Call (telephone based) approach for people with non-diabetic hyperglycaemia.[42] The results suggest that this approach may be effective at sustaining long-term behaviour change and improved outcomes, thus reducing risk of type 2 diabetes. However, these were small observational studies. Members of our team are also currently involved in the independent evaluation of the Salford Care Call as part of the NHS DPP demonstrator site work and we are keen to utilise the skills, tacit knowledge and understanding from the previous research in Salford to carry out an in depth, robust and valued evaluation of the national programme.

Evidence explaining why this research is needed now

NHS England, Public Health England and Diabetes UK have recently initiated a national diabetes prevention programme in England. This is a significant investment, based on the wider international evidence on the effectiveness of diabetes prevention programmes and the results of early demonstrator projects.

However, maintaining the effectiveness of interventions from the particular environment of clinical trials during roll-out to the context of routine NHS delivery remains a significant challenge. High quality evaluations are required, both to provide ongoing, independent feedback to the programme on the success of the roll-out,[50] to explore the impact of context and variability in delivery on effectiveness,[51] and to provide a longer term, rigorous assessment of the degree to which the new programme is meeting its aims compared to usual care.

Aims and objectives

The overall aim of the DIPLOMA research programme is to provide a comprehensive assessment of the implementation, delivery and outcomes of the NHS DPP to inform commissioning.

The individual work package aims are as follows:

Work package 1 Access and Inequality

To assess whether sociodemographic factors influence access to the NHS DPP, and to explore the experience of patients and professionals in accessing the NHS DPP

Work package 2 Implementation

To assess the process of implementation of the NHS DPP, and explore the barriers and facilitators that affect the implementation of the NHS DPP

Work package 3 Service Delivery and Fidelity

Aims: To assess the theory, techniques and content of the NHS DPP, examine variation in delivery, and report the extent to which the NHS DPP is delivered with fidelity

Work package 4 Outcomes

To assess what outcomes participants achieve in the NHS DPP, and whether outcomes vary by services delivered and patient characteristics

Work package 5 Comparative Effectiveness

To assess whether the NHS DPP is more effective than usual care in reducing conversion of non-diabetic hyperglycaemia to diabetes, eventually reducing diabetes prevalence in England

Work package 6 External validation

To assess the risk of confounding in participation in NHS DPP and allow adjustment in other work packages

Work package 7 Comparative Cost effectiveness

To assess whether the NHS DPP is cost-effective compared to usual care in terms of long-term costs and benefits

Work package 8 Programme Management

To deliver the research to time, target and budget, and to ensure that the results are disseminated to key stakeholders and inform NHS decision-making

We have structured our comprehensive research programme into separate work packages, although our management of the programme will take advantage of the synergies between them. We are aware that the resources required for this programme are significant, and in line with advice from HS&DR, we have costed each work package as accurately as possible to allow judgement of their individual contribution and value. To help meet our aims, we will draw on a range of theoretical frameworks. RE-AIM was developed to guide the development and determine the potential impact of public health interventions,[52 53] and has been widely to evaluate intervention implementation and impact in a variety of settings including other diabetes prevention programmes.[54 55] We will also draw on relevant social science frameworks relating to access and candidacy,[56] behaviour change theories from health psychology,[57] and normalisation process theory for implementation.[58]

Research Plan / Methods

Work package 1 Access and inequality (Whittaker, Chandola and Sanders)

Aims: To assess whether sociodemographic factors influence access to the NHS DPP, and to explore the experience of patients and professionals in accessing the NHS DPP

Rationale: Inequalities in health have been documented in England,[59] and are of global concern.[60] Inequalities may exist between several population groups, and providers of health care are subject to the Equality Act (2010) concerning equality of opportunity across protected characteristics: age, disability, gender reassignment, pregnancy and maternity, race, religion or belief, sex, and sexual orientation. This is in line with the NHS Constitution which seeks to deliver a service based on need.

Inequalities in health may manifest where inequalities in access exist.[61] Understanding the effect that the NHS DPP may have on reducing inequalities in diabetes requires a thorough assessment of inequalities in access. We will draw on the 'reach' aspect of the Re-AIM theoretical framework, defined as 'the absolute number, proportion, and representativeness of individuals who are willing to participate in a given initiative.[62] Reach depends on effective access among those at risk. Access to the NHS DPP intervention involves a process:

(a) identification of patients 'at risk' in contexts such as NHS Health Checks
(b) a blood test which show they have non-diabetic hyperglycaemia ('NHS DPP eligible' patients)
(c) decision-making by professionals and patients about referral to the NHS DPP
(d) invitation to the NHS DPP
(e) initial attendance at the NHS DPP intervention ('NHS DPP attenders' and 'non-attenders')
(f) completion of the NHS DPP intervention ('NHS DPP completers')

An overarching definition of access concerns empowerment of an individual to use services when needed.[63] Access may be related to multiple factors: availability of the NHS DPP in the area; awareness of NHS DPP among patients and professionals, acceptability of NHS DPP to professionals and patients; and the perceived 'costs' of NHS DPP (such as taking time off work). Each domain of access may be correlated with protected characteristics and social class.

To understand any issues in access to the NHS DPP and the impacts of the programme on diabetes incidence, we will assess whether inequalities in protected characteristics exist for:

1. the identification of patients eligible for the programme ('NHS DPP eligible' patients)
2. the referral of patients to the programme ('NHS DPP attenders' and 'non-attenders')
3. programme delivery and completion (working with work package 3)
4. the effectiveness of the programme (working with work package 4)

The experiences of those targeted to participate in the programme, as well as those delivering it, will also be crucial to its success. Previous research has demonstrated that factors such as age, gender, ethnicity, socio-economic status, employment status and relationship status, have an impact upon experience of living with or being at risk of type 2 diabetes, attendance at NHS Health Checks, as well as other aspects of help-seeking, health behaviour, and self-management.[56 64 65]

Overview of methods: Work package 1 will use a combination of quantitative analysis of existing data sources, and qualitative research. The first research question will use data from the NHS DPP and existing administrative data sources to explore access.

Research questions 2 and 3 will link with work packages 3 and 4 to apply considerations of inequality to uptake, delivery and outcomes of the NHS DPP intervention.

Research question 4 will explore access is influenced by the understandings and experiences of participants

Research question WP 1.1: Are there inequalities in the identification of patients eligible for the NHS DPP? Is there an under-representation of participants with protected characteristics in the NHS DPP compared to the general population of adults at risk of diabetes?

Methods: We will compare the characteristics of patients identified with non-diabetic hyperglycaemia in GP patient records (using the Clinical Practice Research Datalink (CPRD) - see work package 5 for a detailed description of CPRD) to the characteristics of patients identified with non-diabetic hyperglycaemia in representative survey data for England.

The English Longitudinal Study of Ageing; Understanding Society (ELSA), and the Health Survey for England (HSE) record HbA1c test results (enabling an assessment of non-diabetic hyperglycaemia) and contain comparative measures of social class (sourced via Lower Super Output Area (LSOA) IMD score with LSOA and/or IMD score obtained via special access license), ethnicity, sex, and disability status. Multivariable probability analysis will be used to test if the probability of having non-diabetic hyperglycaemia differs across surveys for particular patient characteristics.[66 67] This analysis will identify whether 'NHS DPP eligible' patients in primary care records differ from those in nationally representative surveys. We will use:

a) Understanding Society waves 2 and 3 (2010-2012). All adult respondents are surveyed but blood samples were taken of 13,107/35,937 eligible (12,162 instances of HbA1c (mmol/mol) being recorded), all ages 16 and over, 964 participants with HbA1c between 42-47 mmol/mol

b) ELSA - adults aged 50+ with 457 wave 2 participants, 1,297 wave 4 participants and 1,315 wave 6 participants with HbA1c between 42-47 mmol/mol

c) CPRD (comparative observable factors e.g. age, gender, IMD, comorbidity; Read code for non-diabetic hyperglycaemia)

Research question WP 1.2: Are there inequalities in the referral of patients to the NHS DPP? Once a diagnosis of non-diabetic hyperglycaemia has been made, the GP may or may not refer the patient to NHS DPP. This decision to refer may be influenced by both the GP and patient, and may be related to patient or professional perceptions of the value of the service.

Method: We will compare the observable characteristics of referred and non-referred patients with non-diabetic hyperglycaemia. This analysis will be conducted in CPRD, using Read codes (hypoglycaemic - non-diabetic; referral to NHS DPP; and referral declined) and data on patient characteristics.

Research question WP 1.3: Are there inequalities in attendance and completion of the NHS DPP?

Following a positive identification of non-diabetic hyperglycaemia, patient uptake and adherence to the NHS DPP intervention may vary. We can examine this in CPRD, using relevant Read codes for 'referral to the NHS DPP intervention', 'intervention started', 'declined' and 'completed'.

Method: Using CPRD data (observable patient characteristics e.g. age, gender, IMD, co-morbidities; and Read codes). for those referred and on the programme we will compare the distribution of observable protected characteristics e.g. age, gender, ethnicity, disability; and IMD, of those completing and not completing. Linking with work package 3, multivariable probability analysis will be performed to obtain differences conditional on observable differences that may predict participation and be correlated with protected characteristics and IMD.

Research question WP 1.4 Are there inequalities in outcomes of the NHS DPP intervention?

Linking with work package 4, we will explore whether effects of the NHS DPP intervention on outcomes vary by protected characteristics observed in CPRD (e.g. age, gender, ethnicity, and IMD) and in the minimum data set.

Research question WP1.5 What is the experience of patients and professionals in accessing the NHS DPP?

Access to care is a complex concept requiring a range of methods to understand the mechanisms that might underlie inequities in access.[68 69] Focused qualitative work will be used to explore the experience of participants (and potential participants) and providers within the NHS DPP, to build a more comprehensive model of the 'reach' of the NHS DPP and to provide feedback to NHS DPP stakeholders on effective practice to support the ongoing programme.[70 71]

Methods: We will sample a 6-10 practices to reflect diverse populations at local level, including practices in areas with high and low deprivation, and high levels of BME patients.

We will conduct observations of relevant consultations between practice staff and patients with non-diabetic hyperglycaemia ('NHS DPP eligible' patients) to investigate how risk factors are discussed with attendees. Although opportunistic identification of such consultations in routine practice is likely to be logistically demanding, our experience during our pilot work in Salford is that many practices will set up specific clinics for such patients, enhancing the feasibility of such observational work.

Consultations will be recorded for detailed analysis (using audio- or video-tape according to feasibility and participant preference). We will aim to record 50 consultations with 10 professionals. We aim to achieve a maximum variation sample of patients according to protected characteristics such as age, gender, disability, and ethnicity and other factors like socio-economic status.

We will also recruit GPs/nurses in practices to participate in a one-to-one interview or focus group interview about their experiences of DPP and their views surrounding diabetes prevention

Some GP practices notify patients of their pre-diabetes diagnosis via letter only. If this is the case, patients are directed to phone the NHS DPP call centre if they are interested in going on the course. We therefore also intend to interview and observe the calls made to the DPP call centre. The patient side of the conversation will not be listened to as obtaining written consent will be difficult. The research team will only observe the call centre worker's responses and ask them questions about the sort of questions they are asked by patients.

Our recruitment strategy is based on two previous studies from the University of Manchester: WISE (REC 07/H1011/96) and Diabetes prevention in primary care (14/SC/0097)

To recruit patients, we will use two methods: (a) Researchers will attend the practices on agreed days for study recruitment (b) Practice staff recruit and consent patients

We know that some practices may want to recruit and consent patients into the study without the presence of a researcher. This method of recruitment was suggested by GPs involved in another diabetes prevention study (14/SC/0097), as diabetes prevention advice is sometimes provided ad hoc

In this method, primary care staff will answer any questions about the study, go through the Patient Information Sheet, and consent patients once they have understood the study.

Using the above strategies, patients will be provided with all necessary information about the study so they can make sure participation is right for them. Also, patients will be made aware of their right to withdraw their data and participation from the study at any time.

We are aware that some patients may enter the NHS DPP through routes other than primary care, such as response to adverts. If many sites are using referral options other than primary care, we will work with NHS DPP providers to recruit these patients to understand their experience of accessing the course. In these cases, we would ask the NHS DPP providers to send patients an invitation letter and a Patient Information Sheet, and ask those patients to return a 'consent to contact' form if they are willing to be interviewed

We will interview all professionals (n=10) after consultations and a proportion of patients (approximately 20) to gauge understanding of levels of risk and the meaning of relevant diagnostic information, as well as both perspectives on any decisions about referral and need for the NHS DPP intervention. We will follow the approach of our previous research[72 73] to analyse observations alongside interviews to explore influences associated with access and use of the NHS DPP.

Additionally, we will interview people with a Read code for non-diabetic hyperglycaemia ('NHS DPP eligible' patients), but who have not been referred to the DPP. We will also include people for whom referral has been offered but declined ('NHS DPP attenders'). We will work with practices to identify such patients from their primary care records. Practices will send patients an invitation letter and patients who agree to participate will complete and return 'consent to contact form'.

Final sample sizes will be contingent on iterative analysis to inform sampling until saturation of data is achieved.

For observational data, we will ensure detailed transcription to allow a focus on language and interaction between patients and professionals within the consultation drawing on a conversation analysis approach. This will enable us to focus on how risk and test results are discussed with consequences for subsequent decisions and action. This is an approach we have previously adopted to investigate how practice within consultations supports or undermines self-management and diagnosis in primary care.[72 73]

Interviews will be audio-recorded with consent, transcribed and thematically analysed, using a modified framework approach[74], and drawing on relevant theoretical frameworks such as candidacy,[56] cultural competence,[75] and the impact of practical and structural barriers.[65 76 77]

We will ensure we include some patients who are non-English speakers, and translators will be used. We will also address variations in health literacy via this maximum variation sample.

Work package 2 - Implementation (Wilson)

Aims: To assess the process of implementation of the NHS DPP, and explore the barriers and facilitators that affect the implementation of the NHS DPP

Rationale: The NHS DPP has an ambitious plan to roll out nationally, using a staged approach across different 'areas' and building on work within initial pilots and demonstrators. Work in implementation science has highlighted the importance of context in the success or failure of health care innovations.[51 78] This work package will explore the local context surrounding wider implementation of the NHS DPP.

Overview of methods: To meet our aims we propose a descriptive survey to develop a sampling frame, followed by qualitative case study research, to answer 3 questions.

Research question 2.1 What is the local context for implementation of the NHS DPP?

For the roll out of areas commencing April 2017, we will undertake an initial scoping phase to understand the organisational context for implementation.

Methods: we will contact the designated local lead and NHS DPP provider for each area. We will seek to obtain all available pilot contracts, and then undertake a rapid synthesis of these contracts to gather data on:

(a) area specific aims and objectives
(b) key performance indicators to measure 'success'
(c) details of any incentive structures
(d) targeted patient groups
(e) workload and anticipated outcomes.

Where provider contracts are not accessible, we will supplement this data with short semi-structured telephone interviews with relevant commissioning leads. Data generated in this scoping phase will be used to develop a typology of areas to act as a sampling frame for detailed case study exploration.

It is anticipated that we will generate a purposive sample of up to 4 case sites for each of the 4 main providers of the NHS DPP intervention (n=16). We will recruit a mix of areas, varying on characteristics such as rural and urban locations, populations with different socio -economic characteristics, areas with particular GP recruitment and retention challenges, and those utilising a range of approaches to patient identification and referral.

Research question 2.2 What are the barriers and facilitators to the implementation of the NHS DPP within areas?

We will analyse the process of implementation in our selected sites, and explore barriers and facilitators to implementation.

Methods: In this second phase, we will conduct longitudinal interviews with the designated leads for each local area. Semi structured interviews will be conducted twice (at 3-6 months and 9-12) and will utilise a topic guide informed by data from the scoping phase.

Initial interviews will explore the process of local implementation of the NHS DPP, including the local organisation of the programme, expectations of and attitudes to the NHS DPP, funding, target populations and referral and clinical pathways.

Later interviews will explore reflections on implementation and sustainability as well any recruitment challenges and unintended consequences (such as who is actually referred, or whether pre-existing services for lifestyle change or diabetes prevention are displaced or foregone).

Research question 2.3 What are the barriers and facilitators to the implementation of the NHS DPP within practices?

To complement the area-level analysis, we plan to explore in-depth the development and implementation the NHS DPP at the level of the individual general practice. We will examine the organisational processes implemented to identify and refer patients to the NHS DPP intervention (such as case finding, NHS Health Checks, or opportunistic screening). This work will complement qualitative work within Work Package 1 looking at the experience of patients and professionals in accessing the NHS DPP.

Methods: Data generated from the area level exploration will be used to identify specific practice level examples of approaches to patient identification and referral. In up to 8 general practice sites, we will explore in-depth, the development and implementation of practice-level processes to identify and refer patients to the NHS DPP intervention. In each practice, we will conduct longitudinal telephone interviews with the practice manager and (where appropriate) the lead clinician.

Early interviews (around 3 months) will focus on the development and implementation of local practice-level strategies and processes to identify and refer patients to the DPP intervention. Later work (9-12 months) will focus on the extent to which 'case-finding' processes have become embedded into routine practice systems. In doing so, we will draw upon a theoretical approach known as Normalisation Process Theory[58 79] which facilitates understanding of the extent to which new processes become part of normal practice. NPT proposes that for a complex intervention (in this instance case finding) to become part of routine care (where there will be competing priorities), we need to consider the following mechanisms: coherence ('what is the work'), cognitive participation ('who does the work'), collective action ('how does the work get done') and reflexive monitoring ('how is the work understood'). By focusing early and late practice interviews around these four constructs we believe we will be better able to understand any recruitment challenges and/or unintended consequences that arise from the practice-level processes as they develop over time.

At both time points, we will explore perceptions and attitudes towards the NHS DPP scheme, identify any recruitment challenges and any unintended consequences from practice-level processes or the NHS DPP intervention itself. Analysis of data from this work package will be undertaken in close collaboration with qualitative work within Work Package 1 to take advantage of synergies between the organisational perspectives adopted here, and those of patients and professionals in work package 1. Interviews will be audio-recorded with consent, transcribed and thematically analysed, using a modified framework approach.[74]

Work package 3 – Service Delivery and Fidelity (French)

Aims: To assess the theory, techniques and content of the NHS DPP, examine variation in delivery, and report the extent to which the NHS DPP is delivered with fidelity

Rationale: The NHS DPP intervention is based on a systematic review and meta-regression identifying effective components of diabetes prevention programmes, which included the length of the intervention period, engagement of social support, targeting diet and physical activity, and use of well-established behaviour change techniques.[49] To achieve the potential benefits outlined in the review, the NHS DPP intervention will have to be delivered with fidelity to those components, across multiple sites and practitioners - a major challenge for any complex intervention.[53 80 81]

Intervention fidelity as defined by the NIH Behaviour Change Consortium (NIH-BCC)[57] includes a number of elements:

(a) Study design – is the intervention congruent with relevant theory and best practice?
(b) Training - have practitioners been properly trained to deliver the intervention?
(c) Delivery – has the intervention been delivered as designed?
(d) Receipt – do patients understand the intervention and perform key skills during delivery?
(e) Enactment – do patients perform relevant skills in real life setting?

Overview of methods: This work package use a combination of document review, observation and interviews to assessing whether the NHS DPP intervention is delivered with fidelity to the content and design principles specified. Work package 3 has 4 research questions.

Research question WP 3.1 Study design - what are the explicit theoretical principles, behaviour change technique (BCT) content and mode of delivery of the NHS DPP intervention as exemplified in (a) intervention protocol/manual, and (b) training materials? What is the observed variation in these across providers, sites and settings?

Methods: We will collect intervention protocols, manuals and training materials from at least the 4 main NHS DPP providers, and probably from 4 sites per provider, that are purposively selected to cover variation in socio-demographic status and ethnicity. We will code these materials in terms of BCT content, using the CALORE BCT taxonomy or BCTv1[82] and the Theory Coding Scheme.[83] The CALO-RE BCT taxonomy provides a list of BCTs (e.g. setting behavioural goals, prompting self-monitoring) with definitions and guidance on how to assess whether a BCT is present. We will also use the TiDieR framework to describe the broader nature of the NHS DPP interventions.[84] The TiDieR framework provides a checklist of key intervention features, including mode of delivery, but also who delivered, where, and in what dose.

The primary analyses will report the specific BCTs designed into the intervention, key features of intervention delivery, and the extent to which the justification is explicitly based on theory. We will report these (broken down by provider and site) as well as for overall programme, where such documents exist.

> *Research question WP 3.2.* Training - To what extent does the training of NHS DPP staff address elements of theory and BCT content? What is the variation across providers, sites and settings?

Methods: We will observe training sessions from at least the 4 main NHS DPP providers, and probably from 4 sites per provider, purposively selected to cover variation in socio-demographic status and ethnicity. These sessions will be coded using schemes described in WP 3.1, but will involve observation of training sessions instead of documents.

> *Research question WP 3.3.* Delivery - To what extent is the NHS DPP intervention delivered with fidelity to intervention protocols and manuals?

Methods: We will digitally record or otherwise capture a selection of NHS DPP intervention sessions, again from at least the 4 main NHS DPP providers, and probably from 4 sites per provider. We will aim to capture 5 sessions per site, to provide a total of 80 sessions for coding. We will not code the

use of theory, as it is not appropriate to provide detailed descriptions of theory when providing patients with intervention instructions.

For observation of group sessions, we will start with an overall sampling frame of DPP providers and sites that are in place during the evaluation period, and purposively sample on the basis of these, to get maximum variation in patient SES and ethnicity, as well as geographical location. Within sites, we will purposively sample sessions to provide to provide variation in terms of patient SES and ethnicity, as well as geographical location, and variation in duration that the site has been running and times of year.

Where possible, we will work with NHS DPP providers to send patients a Patient Information Sheet and consent form prior to their course starting.Our researchers will attend the course on the first date, and will be present to go through any outstanding consent forms, and answer any questions about the research prior to the course starting. We accept that failure to gain consent from any patients on the course means that we will be unable to observe or record any of the course. Eligible patients may attend a session with a partner or a carer who has not been invited to take part in the DPP programme. If this is the case, they will also be consented to participate in the study. If they refuse consent, this would also mean we would be unable to observe or record the course.
Together, these analyses (3.1, 3.2 and 3.3) will provide information on the extent to which the interventions as specified in intervention design contain specific BCTs, are theoretically based, as well as details of their mode of delivery. We can then examine variation across providers, sites and people delivering the intervention. Also, we can examine the extent to which there is a loss of fidelity to the key principles underlying the intervention in delivery, and where such loss occurs.

> *Research question WP 3.4.* Receipt - To what extent is the content of the NHS DPP intervention understood by recipients as intended by providers?

Methods: We will undertake qualitative interviews with people receiving the NHS DPP intervention, to assess what they think the NHS DPP intervention is trying to do ('intervention receipt'). These interviews will also explore issues beyond the specific scope of the NIH-BCC, such as participant experience of the NHS DPP intervention, their experience of the nature of interaction,[72] the experience of different delivery modes, issues of cultural acceptability,[75 77] and the impacts of the wider social network on access to and adherence to the intervention.[85 86]

Again, we will sample from at least the 4 main NHS DPP providers, and from 4 sites per provider, interviewing 2 people per site, giving 32 interviews.

We will aim to sample patients in WP3.4 from those sessions that we are recording to assess intervention delivery in WP3.3. Given this, we will aim to recruit all patients to the interview study at the beginning of the group intervention, and then be able to sample from a larger group of patients than the two patients per site that we are aiming to interview. We will explicitly ask for permission to contact patients if they drop out of the intervention, and therefore be able to contact those patients who subsequently drop out as well as those who complete. Although people who drop out of interventions are usually less willing to take part in interview studies, we have previously recruited people who withdrew from a weight management service using similar procedures.

Interviews will be audio-recorded with consent, transcribed and thematically analysed, using a modified framework approach for assessment of 'intervention receipt',[74] and grounded theory for the exploration of wider issues around participant experience.

Work package 4 –Outcomes (Cotterill)

Aims: To assess what outcomes participants achieve in the NHS DPP, and whether outcomes vary by services delivered and patient characteristics

Rationale: High quality data collection and measurement of outcomes is increasingly seen as a critical aspect of effective health care.[87-89] The NHS DPP will put in place a comprehensive protocol for data collection. Detailed analysis of this data will be critical to assessing the effectiveness of the programme in achieving improvements in well-being, behaviour change and lifestyle among those attending. Such data may provide useful feedback for providers on performance, and variation between providers, sites and professionals.

Overview of methods: This work package seeks to assess what outcomes participants achieve in the NHS DPP, and to identify which components of the NHS DPP intervention are effective, and for whom. It will identify the sources of any variation in outcomes among those who attend the NHS DPP intervention, and explore drivers of that variation.

Work package 4 will focus primarily on comparisons within the programme, and will link closely with analyses from work package 3 on Service Delivery and Fidelity. WP4 will analyse the individual level data set collected by NDPP providers to look at variation in participation, service delivery and outcomes. There will not be a comparator group, as the comparative analysis will be done more robustly in WP5. Analyses against external comparators will be considered by Work package 5.

Research question WP 4.1 What services are delivered by NHS DPP and what is the extent of participation in the NHS DPP intervention?

> *Research question WP 4.2* How does service delivery and participation in the programme vary by (a) the 4 NHS DPP providers (and any variation within provider by area) and the associated variation in content and delivery (b) between patient subgroups?
>
> *Research question WP 4.3* – What are the outcomes of patients in the NHS DPP intervention, including wellbeing score, weight change, HbA1c, and mortality?
>
> *Research question WP 4.4* – How do outcomes vary by (a) the 4 NHS DPP providers (and any variation within provider by area) and the associated variation in content and delivery (using data from work package 3) (b) between patient subgroups?

Methods: All four research questions will be addressed using a minimum individual patient data set. NHS DPP have already stipulated that NHS DPP providers will be required to collect and supply this data, which will then be passed in anonymised form to the research team. The individual dataset will include 48 items, including referrer, delivery organisation, personal/demographic details, process data (including records of each session) and outcome data (Warwick-Edinburgh Mental Wellbeing Scale, weight change, and HbA1c).

Based on the NHS DPP announcement that it expects to provide 100,000 places per year by 2020 (https://www.england.nhs.uk/ourwork/qual-clin-lead/diabetes-prevention), we have made the assumption that there will be 50,000 individuals in 2017/18; 80,000 in 2018/19 and 100,000 in 2019/20, a total of 230,000 individuals.

We will match this individual patient data to service-level data on what is being provided by the providers in different areas, collected by the 'Implementation' and 'Service and Fidelity' work packages. This will be a rich individual level longitudinal dataset, that will allow patients be followed over time, through the steps of invitation, baseline assessment, enrolment, service delivery, completion, and end-of-service outcomes.

We anticipate that, like any administrative dataset, there will be error and missing data. As with other aspects of the evaluation, the DIPLOMA team will work closely with the NHS DPP to ensure that suitable processes are in place to ensure data collection is as accurate and complete as possible. We will provide feedback on missing data through our initial analyses, and follow up with providers if problems continue.

We will use descriptive statistics to assess overall outcomes, and multivariable regression methods to estimate the relationship between patient, provider and service characteristics and participation (question 4.2) or outcomes (question 4.4). For analysis, we will use multilevel modelling of individual characteristics (focussing on health inequalities – WP1) and service delivery (number of sessions attended; area/provider; and important elements of service delivery identified in WP3 (e.g. behaviour change techniques) on the outcomes collected by NDPP providers (weight, HbA1c, WEMWBs etc). We will work closely with providers to minimise missing data, and we will report the extent of missing data. We will consider multiple imputation methods, if we assess that the data meets the assumption of NMAR.

WP4 will use the insights/models developed in the WP1-3 to influence the choice of variables in our statistical models: on health inequalities (WP1) and service delivery/fidelity (WP3). We will incorporate these to estimate their effect on participation, service delivery and outcomes for NDPP attenders.

We will produce regular reports to inform ongoing NHS DPP service delivery, drawing on exemplars such as the IAPT dataset.[87] To ensure the relevance of these reports, a detailed analysis plan will be written in advance, in negotiation with the NHS DPP, and this can be re-negotiated over the course of the project.

Timing of analysis: Analysis will be undertaken at three time points: Summer 2018 (using year 1 data, around 50,000 patients); summer 2019 (using Year 1 and 2 data, around 130,000 patients), summer 2020 (using complete dataset, around 230,000 patients). The earlier reports will focus on reporting who is invited to attend, baseline measures and enrolment, but as time progresses, we will be able to say more about the services delivered, completion rates and end-of-service outcomes.

Work package 5 Comparative Effectiveness (Kontopantelis and Reeves)

Aims: To assess whether the NHS DPP is more effective than usual care in reducing conversion of non-diabetic hyperglycaemia to diabetes, eventually reducing diabetes prevalence in England

Rationale: Although the NHS DPP is based on a strong international evidence base,[49] justifying the commissioning of such a large and complex programme requires rigorous scientific evidence that the programme is achieving benefits beyond those delivered by current prevention services. The roll-out of the programme makes formal randomised evaluation problematic. This work package will utilise administrative data and a range of complex statistical techniques to provide a rigorous estimate of the success of the programme in reducing conversion of non-diabetic hyperglycaemia to diabetes (incidence), and reducing the overall numbers of cases of diabetes (prevalence). Outputs from work packages 4 and 5 will complement each other and provide a complete evaluation of the outcomes of the scheme.
Overview of methods: For these analyses, we will use numerous sources of administrative data. First, we will use the Clinical Practice Research Datalink (CPRD), a large database of administrative primary care data. The database has been active since the 1980s with high quality data becoming available after 2000 with the introduction of the Quality and Outcomes Framework (QOF).[90] Complete data on all aspects of care (diagnoses, referrals, treatments, tests) have been collected from over 500 practices each financial year, covering approximately 7% of the UK population.[91] The CPRD population is generally representative of the UK population, especially in terms of practice and patient deprivation, although it is largely tied to a single clinical computer system (Vision),[92] and as a consequence the North-East of England is somewhat under-represented.[93] The CPRD data can be linked to Hospital Episode Statistics (HES) and Office for National Statistics (ONS) data, allowing the construction of a more complete patient journey through primary and secondary care and the mapping of causes of mortality (e.g. diabetes, cardiovascular).

The second major source of routine data will be the general practice diabetes registers as collected for the QoF,[94] and related information about the general practices themselves. We will make use of free public health datasets from the ONS and the Health & Social Care Information Centre (HSCIC), including geographical data, 2011 census based population estimates, deprivation and rurality information,[95] to map diabetes at a low population level and also to scale up our findings to a national level. We will also obtain point of interest data from the Ordnance Survey (OS).[96] All these datasets will be linked at a low population level and will be used to map diabetes and non-diabetic hyperglycaemia prevalence rates and their associations with area and population characteristics nationally. Non-diabetic hyperglycaemia data in practice registers should become available after implementation of the NHS DPP nationally, and will be linked with existing datasets.

We have negotiated with another provider of GP data (ResearchOne) who have offered access to their data at no cost. This will provide far greater coverage of English general practice than CPRD alone (essentially doubling our sample size) and improves coverage in areas where ResearchOne is dominant (such as the North East, East Midlands and Yorkshire).

We are aware that data from the National Diabetes Audit may be of relevance to work package 5, but at present the utility of that source cannot be confirmed, as it does not yet collect data on non-diabetic hyperglycaemia. We will explore the possibilities for use of this data as the research progresses.

> *Research question WP 5.1* – What is the current epidemiology of non-diabetic hyperglycaemia and diabetes?

This work package will use the CPRD to investigate rates and patterns of non-diabetic hyperglycaemia and its conversion into a diabetes diagnosis, to provide critical background information for the interpretation of findings related to the effects of the NHS DPP.

Methods: We will investigate the prevalence and incidence of non-diabetic hyperglycaemia using a specific list of Read codes which are routinely used in UK primary care.[97] A preliminary analysis using CPRD shows that the use of the codes has increased considerably over time. For the latest period (April 2015 to March 2016), at least one relevant code was found in the records of 71,521 patients (1.57%), implying around 900,000 people nationally with non-diabetic hyperglycaemia.

We will also estimate the association between diabetes and non-diabetic hyperglycaemia prevalence at the practice level. We expect a strong correlation which would indicate that diabetes data can be used as a reasonable proxy in the absence of data on non-diabetic hyperglycaemia data.

From 2016-17 onwards, we will use these historical codes and the new codes that will be introduced with the NHS DPP programme to calculate and report detailed information on non-diabetic hyperglycaemia. Excluding cases with an existing diabetes diagnoses, we will calculate both the prevalence and incidence of non-diabetic hyperglycaemia over time, nationally and regionally (such as CCGs or NHS England Local Area Teams). To better understand regional variation, we will calculate age-sex standardised prevalence and incidence rates. We will use longitudinal multivariable mixed-effects regression models to identify predictors of non-diabetic hyperglycaemia from a list of relevant covariates: age, sex, deprivation and QOF comorbidities (excluding diabetes mellitus).[98 99] We will investigate changes in the characteristics of the non-diabetic hyperglycaemia populations over time, by including time interaction terms in the regression models.

We will use an interrupted time-series design[2] to quantify the overall effect of the introduction of the NHS DPP on the prevalence and incidence recording of non-diabetic hyperglycaemia (overall and regionally). This is a quasi-experimental design which we have applied in a number of previous studies, most notably to evaluate the impact of the QOF.[93 100 101] Using a mixed-effects regression model we will control for any changes in the population demographics over time, to obtain a more reliable effect estimate.[102] This design assumes a linear pre-intervention trend, which seems plausible according to our initial modelling. Nevertheless, we will statistically assess linearity and if rejected we will use alternative models with non-linear terms for prevalence of non-diabetic hyperglycaemia.

More inclusive criteria in the Health Survey for England have returned much higher prevalence estimates for non-diabetic hyperglycaemia, and we will explore implementing them in the CPRD to obtain more speculative estimates through which to conduct sensitivity analyses (i.e. based on HbA1c values and relevant Read codes).[97] However, for the purposes of all planned analyses in this and other work-packages, this is a relatively minor issue (except regarding the true non-diabetic hyperglycaemia prevalence rate). For example, the predictors of non-diabetic hyperglycaemia should be consistent across the two analyses (of conservative or more speculative cases)) since we would not expect systematic bias in relation to the research questions (i.e. the predictors of non-diabetic hyperglycaemia should be the same whether a large percentage of patients is missed, or not). Even in there is some form of systematic bias (for example if Read code cases are more 'severe'), the findings from the conservative analysis would still be relevant and important. Finally, it might be the case that a conservative approach where we focus on definite non-diabetic hyperglycaemia cases which have been coded as such may be more desirable.

> *Research question WP 5.2* - What is the effectiveness of the NHS DPP at reducing the conversion of non-diabetic hyperglycaemia to diabetes?

The primary objective of the NHS DPP is to reduce, or at least slow, the rate at which patients with non-diabetic hyperglycaemia go on to develop a full diabetes diagnosis. In line with this, the main outcome in these analyses will be the conversion of non-diabetic hyperglycaemia to diabetes. The analyses assume that the intervention will be widely implemented across the UK in order to ensure that a substantial number of the practices contributing data to the CPRD and ResearchOne will be within the scheme.[91]

Methods: The primary analysis for assessing the effectiveness of the NHS DPP will use the CPRD. For robustness, we will address the question using two different research designs.

WP 5.2 design 1: We will use all practices participating in the NHS DPP, and compare the conversion rate from non-diabetic hyperglycaemia to diabetes in patients prior to the start of the NHS DPP to the rate afterwards in equivalent patients. To accomplish this, we will match pre-intervention cases of non-diabetic hyperglycaemia[97] to post-intervention cases at the same practice and with the same practitioner if possible (excluding cases with a previous diabetes diagnoses). We will then compare the rates of conversion to diabetes within 1 year and 2 years between these groups.

Earlier, we confirmed that risk-of-diabetes Read codes are already routinely being used. As the NHS DPP-specific Read codes are not available for patients prior to the scheme, we will conduct an initial exercise to identify the group of pre-existing codes that best identify patients classed as eligible using

the NHS DPP-specific codes. We will then use the identified codes to classify patients as scheme-eligible both pre- and post-NHS DPP, to ensure comparability. We will use propensity score methods for the matching. In the logistic regression model to calculate the score we will include: age, sex, region, deprivation, QOF comorbidities and all available biological parameters (e.g. body mass index and HbA1c levels at baseline). Data will be complete for all covariates except the biological parameters, for which we expect a very high level of completeness for this group of patients. Nevertheless, we will use appropriate multiple imputation methods for longitudinal data if needed.[103]

We will compare the conversion rates to diabetes between the two matched groups. Since practices are unlikely to refer all their cases of non-diabetic hyperglycaemia to the scheme (because of capacity limits), we will also compare those patients who were referred to their matched pre-intervention controls, while recognising that this comparison will be subject to confounding with any selection bias.

For the analyses we will use both logistic regression models to compare 1-year and 2-year conversion rates (and over longer time-periods if the data are available) and more appropriate Cox proportional hazards and competing-risks survival regression models to account for censoring and competing risks (e.g. deaths). We will include a region covariate in the models to assess the heterogeneity of the effect, i.e. to investigate if the observed effect varies greatly across regions

Sample size: We estimate 26,581 invites will be needed nationally to achieve 90% power to detect an intention to treat risk reduction of 25%. The NHS DPP aims to invite 100,000 patients each year by 2020. We assume a balanced design, alpha level of 5%, a conservative baseline 1-year conversion rate to diabetes of 7.5%,[104] and an intention to treat risk reduction of 25% (i.e. 1-year conversion rate of 5.625% or OR=0.735).[49] On this basis, we would need a total of 7331 patients to achieve 90% power to detect that level of risk reduction. As demonstrated earlier, we will have many times that number for the pre-intervention group (71,521 cases associated with non-diabetic hyperglycaemia in financial year 2015-16). For the post intervention group we would need 3,666 patients, a figure we would expect to be available in the CPRD if 52,371 invitations to the scheme are administered nationally (since the CPRD covers approximately 7% of the UK population). Assuming a 5 (pre-intervention) to 1 (NHS DPP) design and with all other assumptions unchanged, we would need a total of 11,164 to achieve 90% power, or 1,861 intervention patients. Only 26,581 invites nationally would be required to identify this number of invites in the CPRD.

We will explore a number of secondary outcomes. For hospitalisation, we will use all cases of hospitalisation and cases where the main reason for hospitalisation was diabetes, within 1 year and 2 years of the index date (referral to the scheme). For primary care visits, we will use all visits to primary care within 1 year and 2 years of the index date. For biological parameters (HbA1c and BMI) we will use the last available measurement within 1 year and 2 years of the index date (we have also developed relevant methodologies for BMI prediction and have used other methodological tools for multiple imputation of missing data, which we will consider using in this context).

All analyses will be repeated for the secondary outcomes with small changes in the analytic models: Poisson regressions for hospitalisations and primary care visits; linear regressions for HbA1c levels and weight/BMI; Cox proportional hazards regression for deaths. Sensitivity analyses will be used to assess the robustness of the results to different assumptions about patient eligibility, choice of co-variates, and modelling options (e.g. multivariable regression instead of matching).

WP 5.2 design 2: For this analysis, we will only use data from the post-intervention period and compare non-diabetic hyperglycaemia-to-diabetes conversion rates between: (i) patients referred to the scheme versus matched patients not referred, within the same practice and practitioner if possible; and (ii) practices and practitioners referring to the scheme versus those that are not referring. As before we will use propensity score matching to select control (non-referred) patients, 5 to 1 if possible to increase power, using the same covariates as for research design 1. For the comparison of referring versus non-referring practices and practitioners we will again propensity-score match, using practice- and practitioner-level variables, including list size, region, patient population characteristics and practitioner age and sex. The propensity score matching approaches will also inform us whether practitioners and practices with certain characteristics are more likely to refer patients to the NHS DPP.

Analyses will closely resemble those in the first design for both the primary (conversion to diabetes) and secondary outcomes (hospitalisation, primary care visits, HbA1c and BMI/weight levels, and death), and we will also investigate effect heterogeneity across regions in this design as well. The

power considerations for an intention to treat analysis (comparing not invited and invited) are the same as in the first design. The analyses are potentially subject to confounding due to selection bias (at the patient and practice level, respectively), but no more than other types of comparisons based on CPRD or routinely collected health records. There are a variety of methods available for assessing/adjusting for unmeasured confounding,[105-108] and we will apply selected methods to determine the likelihood that any results could be accounted for by this.

Research question WP 5.3 - What is the long-term impact of the NHS DPP on diabetes prevalence?

If successful, the scheme is designed to lead to a long-term reduction in the population-level prevalence of type 2 diabetes. Therefore we will also undertake a population-level investigation using national datasets of aggregate data. We calculate the power to detect a population-level effect for the NHS DPP. However, these analyses are secondary, since they rest on numerous assumptions. The main analyses will be at the patient-level, using CPRD data.

As noted in WP 5.1, we will describe the national prevalence of diabetes over time, with routinely collected data from general practices used for calculating QOF payments. We have summarised currently available information, for both type 1 and type 2 diabetes. Although the data are aggregated across both types, 90% of diabetes cases are type 2[109] (around 95%).[110-112]

The investigation of non-diabetic hyperglycaemia at the population level will require that non-diabetic hyperglycaemia practice registers will become available after the full implementation of the NHS DPP (e.g. like QOF condition registers). Assuming this is the case, we will combine this dataset with ONS data on rurality, deprivation and other covariates we will identify as relevant and measured in the 2011 census. Using methodology which we have developed to link practice level data to low-level population statistics,[113] we will explore the associations between the population characteristics at the Lower Super Output Area level (LSOA: low-level geographies containing an average of 1,500 people). We will also estimate the age-sex adjusted prevalence rates in each LSOA and using both adjusted and unadjusted rates, we will investigate for regional clusters of high prevalence of non-diabetic hyperglycaemia and rapid prevalence increase over time, compared to the national average. This analysis will identify hotspots of non-diabetic hyperglycaemia and will further inform on where best to target capacity for maximum impact. We will use the same approach to map diabetes prevalence, for which data are readily available. In the absence of non-diabetic hyperglycaemia register data we will use the strength of the association between non-diabetic hyperglycaemia and diabetes at the practice level (see WP2) to decide whether the diabetes register can act as a reasonable proxy in identifying hotspots of non-diabetic hyperglycaemia with this methodology.

To statistically assess the short- and long-term effects of the intervention on diabetes prevalence we will use an interrupted time series analysis.[114] The method compares the level and trend in an outcome (e.g. diabetes prevalence) post-intervention, with a prediction made from the pre-intervention level and trend, to look for a significant change. Typically, the pre-intervention trend is assumed to be linear, but this can be tested and curvilinear pre-intervention trends also modelled. For each of the approximately 8,000 GP practices contributing QOF data, we will use interrupted time series to predict diabetes prevalence (i.e. the practice's QOF diabetes register) in each year subsequent to the start of the NHS DPP and compare these predictions to the actual prevalence rates in those years.

Numbers of patients receiving the intervention will be relatively small initially, but will increase over time and after a few years a cumulative population-level effect should begin to show, provided the scheme is successful. To assess the power that this design would have to detect a change in the national diabetes prevalence rate, we set up a simulation making the following assumptions:

A large pool of patients with hotspots of non-diabetic hyperglycaemia to receive the intervention.[97]

30,000 NHS DPP interventions in 2016/17, 50,000 in 2017/18, 80,000 in 2018/19 and 100,000 in 2019/20, from a random sample of practices and 1% of patients in the NHS DPP intervention. Risk for developing diabetes in a year in those with hotspots of non-diabetic hyperglycaemia, varying 5-10% at the practice level.

A 25% risk reduction for those receiving the intervention (intention to treat).[104 115]

A cumulative effect over time, i.e. those who do not develop diabetes because of the intervention in a specific year never do, and their numbers accumulate over time.[46]

That the change in prevalence of diabetes over time (from 2007/08 to 2019/20) at the practice level would be linear, if the NHS DPP was not implemented.

Random noise at the practice level post-intervention in the form of variability in the diabetes register (between -5% to +5% of the register in the previous year).

After 1000 simulations (the code is available from the applicants), we estimate the power (at alpha=5%) to detect the cumulative effect on national diabetes prevalence in 2017/18 at 0.0%, in 2018/19 at 39.9%, in 2019/20 at 66.6% and in 2019/20 at 92.4%. The total number of patients at risk of diabetes who received the intervention and did not ever progress to diabetes was 4,853 under these simulation settings (averaged across the 1000 iterations).

A key assumption in the simulation is that the underlying trend in diabetes prevalence is linear and that any deviation from this would be due to the NHS DPP and not any other factors. However, even small changes in the prevalence due to external parameters (e.g. less severe cases being diagnosed, driven by the NHS DPP itself) could compromise the population-level investigation. In the actual analysis, we will test and adjust for any non-linearity in the pre-NHS DPP period.

It is also the case that not all practices will be referring to the NHS DPP in the first few years, and the introduction of the scheme will be gradual. Provided that information on participating practices becomes available, we will use it to compare trends between practices that are, and are not, accessing the NHS DPP, under the interrupted time series analysis. This will allow us to use the non-participating practices as a comparator group to better adjust the analyses for potential post-NHS DPP trend changes, not attributed to the NHS DPP, and conduct additional tests on our assumptions.

Post-NHS DPP, once we have completed our analyses, we will be able to revisit this section and scrutinize the assumptions made. We will combine our estimates of the incidence rate of non-diabetic hyperglycaemia and rate of conversion to diabetes, and the impact of the NHS DPP with population-level data, to make forward projections of:

(i) future rates of diabetes incidence and prevalence, under differing assumptions about the numbers of NHS DPP places on offer

(ii) what these trends would have been without the NHS DPP, and the reduction attributable to the NHS DPP.

This analysis will help to determine the intervention capacity required to produce a sustainable reversal in the current trend of increasing rates of both non-diabetic hyperglycaemia and diabetes.

The risks of not obtaining sufficient data for these analyses are closely linked to practice participation to the scheme, a parameter that is out of our control. With the CPRD and ResearchOne databases we have a combined resource that covers over 10% of UK primary care and is representative in terms of deprivation and geography. If the uptake of the scheme is very poor , we may not be able to identify enough cases in the databases. Another potential risk would be if the practices contributing data to the databases were less likely to participate in the scheme. This is highly unlikely, however, since these are generally large and well organised practices and we would expect them to be keen to increase their income through participation to the scheme.

Data linkage
Within the dataset we will have access to information on personal characteristics, diagnoses of non-diabetic hyperglycaemia (NDH - and its predecessors impaired glucose regulation – IGR - and impaired glucose tolerance - IGT), exposure to the NHS-DPP and the primary outcome (progression to diabetes).

There will be no need for us to link patient-level variables from different data sources. CRPD data has already been successfully linked to Hospital Episode Statistics (HES) and Office for National Statistics (ONS) data, allowing the construction of a more complete patient journey through primary and secondary care and the mapping of causes of mortality (e.g. diabetes, cardiovascular).

To complement the patient level analyses and deliver a more complete evaluation of the scheme, we will also use various national databases of aggregate data, at the practice or low geographical level. The aggregate datasets will be linked at the practice level using the NHS practice ID, and then all relevant data will be linked using the NHS attribution dataset (which links primary care patients to ONS lower super output areas using their residence postcode) to low-level geographical areas where they will be combined with deprivation and relevant census covariates. The patient-level data and the aggregate data databases will not be linked.

Work package 6 Validation study (Reeves and Cotterill)

Aims: To assess the risk of confounding in participation in NHS DPP and allow adjustment in other work packages

Rationale: The comparison in the effectiveness (WP 5) and cost effectiveness (WP 7) work packages are between those who do and do not participate in NHS DPP. However, in this observational comparison the influence of unmeasured confounders (that is, factors that have an effect on both DPP participation and progression to diabetes), cannot be excluded. To better evaluate the risk of confounding and adjust for it, we propose a validation study to collect data on the key potential barriers to access and confounders unavailable from the routine records, to inform the analysis and increase the robustness of our findings.

Methods:

Step 1: Initial scoping study

The purpose of the scoping study is to identify factors that may influence access to NHS DPP, to help us design a questionnaire for the validation study. Method: we will use data from the qualitative interviews with patients and health professionals in the Access work package to identify important potential influences on access. We will use this to draw up suitable questions, in discussion with patients and clinicians on the Research Advisory Group. We anticipate that this is likely to include factors such as health status, comorbidities, motivation to improve health, health literacy and competing demands (e.g. employment, caring responsibilities).

Step 2: Questionnaire survey

We will undertake a short questionnaire survey of 400 patients (using post, telephone and online methods), to collect data on key patient factors (identified from the scoping study) that are likely to influence the offer and take-up of referral to the NHS DPP (i.e. whether 'NHS DPP referred' patients are referred to and attend the programme), but that are unavailable in routine datasets.

The purpose of the survey is:

(a) To delineate the role psychological and social factors play in access inequalities, and identify factors potentially addressable by targeted interventions or programme modifications;

(b) To determine the relative influence of and overall influence of these factors on participation;

(c) To combine the results with analyses of CPRD, the minimum dataset, and other data sources to assess the extent to which associations between referral, programme attendance, and development of diabetes are confounded with patient characteristics and to adjust the associations for those factors.

The sample will consist of up to 30 patients from each of 20 practices. A suitable approach to sample size estimation for building regression models is to include 20 cases per parameter to be estimated. Our questionnaire includes 20 predictors (some of the questions will be combined into composites), giving us a recruitment target sample size of 400. To achieve this we are approaching a total of 600 patients.

Practices will produce a list of people at risk of diabetes who have been referred to DPP at least 3 months ago. After checking the list, practices will use a randomisation list provided by researchers to randomly select patients in a 1:1 ratio according to whether they are:
(a) 'NHS DPP non-attenders', referred to NHS DPP but who did not attend any sessions
(b) NHS DPP attenders.
In practices where DPP attendance is not recorded we will send questionnaires to referred patients and ask about NDPP attendance as part of the questionnaire, to place each individual responder in the relevant attenders or non-attenders group. On average, 50% of people referred to DPP attend the programme, so this approach should result in roughly equal groups. Additionally, in practices which hold clinics to review patients at risk of diabetes, study information can be provided to eligible patients.
A high recruitment rate is essential to minimise bias in the validation sample. The questionnaire will be as short as practicable. We will use intensive recruitment and collection methods, working closely with

professionals and the NIHR CRN locally, piloting the questionnaire prior to use with patients, and using financial incentives to patients and practitioners,[116 117] aiming for a 70%-80% recruitment rate.

The high response rate required means that this work-package is 'high-risk'. If we manage to reach this target then our findings in this section will augment the quality of the analyses in work-package 5 and 7. However, the analyses in those work packages are not dependent on the success or otherwise of this survey. If a lower response rate is evident, the data will still be of considerable value, as it will allow us to examine the role of psychological and social factors in access inequalities, and study the relative influence of GPs and patient self-selection on NDPP participation.

Analysis: Descriptive statistics will be used to characterise the groups of people who attend, and do not attend, the DPP across a range of demographic, attitudinal and psychological factors; and comparisons will be made to available national data (eg from the DPP "minimum dataset"). We will use univariate and multivariate logistic regression methods to estimate and compare the role of the various factors in influencing attendance at the NHS DPP intervention. We will also compute a propensity score for each individual representing their probability of attending given their particular combination of factors, to provide a picture of the extent to which these factors as a whole are associated with attendance. The individual factor and propensity scores will further be used in combination with analyses of CPRD, the minimum dataset, and other data sources to assess the extent to which any observed associations between DPP referral and attendance and subsequent development of diabetes are confounded with demographic differences and to adjust the associations for those differences. These results will further feed into the health economics analysis of the cost-effectiveness of the programme, particularly in relationship to estimating effectiveness in patient sub-populations.

Work package 7 Comparative Long-term Cost effectiveness (Sutton and Meacock)

Aims: To assess whether the NHS DPP is cost-effective compared to usual care in terms of long-term costs and benefits

Rationale: As with work package 5, commissioning of the NHS DPP will also be dependent on a rigorous demonstration of cost-effectiveness. Diabetes is associated with compromised quality and length of life and significant long-term health care utilisation. Effective prevention has the potential to generate health benefits and significant cost-savings over the longer term. This work package will utilise economic modelling to explore the longer term costs and benefits of the NHS DPP.

Overview of methods: This work package will estimate the overall cost effectiveness of the NHS DPP taking account of the costs incurred by commissioners to deliver the programme and the long-term health benefits for participants and non-participants. It will draw together the findings of the other work packages and use an economic model to summarise the overall cost-effectiveness of the programme as implemented and identify major opportunities for enhancing its cost effectiveness.

> *Research questions WP 6.1* What are the additional costs of implementing and providing NHS DPP to the range of commissioning agencies involved?

Methods: We will estimate the additional costs of implementing and providing NHS DPP using information from the national commissioning agencies and the Implementation work package. These costs will include the contracted amounts paid to the NHS DPP providers and the costs to general practices of identifying, referring and following-up on 'NHS DPP eligible' patients. We will need to identify the additional costs incurred by general practices due to the introduction of the NHS DPP, excluding the costs that they would incur anyway in identifying and managing this patient group. We will also examine whether general practices transfer the costs of providing services to the target group as a result of the introduction of a targeted scheme for these patients paid for by other means.

> *Research questions WP 6.2* What are the short-term health benefits of NHS DPP to participants in the scheme, and what are the cost consequences of the short-term changes in health service utilisation for participants in the NHS DPP?

Methods: We will use data from the Outcomes work package to estimate the average increases in health-related quality of life reported by NHS DPP participants. We will model the expected effects on quality of life using mapping to the EQ5D 5L. We will use data from participants on health service utilisation and figures from NHS Reference Costs and the Unit Costs of Health and Social Care to estimate the changes in costs. We will focus initially on the short-term effects reported by participants.

> *Research questions WP 6.3* What are the expected long-term health benefit consequences of the introduction of NHS DPP? What are the expected long-term cost consequences of the NHS DPP?

Methods: NHS DPP is designed to increase individual awareness and help them to reduce their risks through lifestyle changes. We will use the ELSA (2004-2014) data to: a) estimate the benefits of lifestyle changes on the diagnosis of diabetes; and b) estimate the effect of diagnosis itself on lifestyle changes. We will use multivariable non-linear models to predict the benefit of lifestyle changes (measured by physical activity, BMI, consumption of fruit and vegetables) on the likelihood of developing diabetes for those who are at risk of diabetes in the first observation in ELSA (using glucose level measurements in ELSA). We will estimate these benefits by age and gender.

We will then determine the wider behavioural effects of the NHS DPP that arise from a non-diabetic hyperglycaemia diagnosis. International evidence suggests that individuals improve their lifestyles in response to diagnosis.[119 120] However, patients with non-diabetic hyperglycaemia could also decide to indulge in less healthy lifestyles anticipating increased and effective treatment in the future. We will estimate this behavioural response using dynamic models of lifestyle choices and comparing diabetics to a statistically defined group of people at 'risk' of developing diabetes using the propensity score model[121] using a similar US survey, the Health and Retirement Survey.

> *Research questions WP 6.4* Is the overall NHS DPP cost-effective compared to usual care? How does incorporation of equity consequences affect the overall cost effectiveness of NHS DPP? What changes to the NHS DPP would improve its short and long-term cost effectiveness?

Methods: It is expected that the NHS DPP will increase costs in the short term as participants make greater use of health services to better identify and manage their newly-identified needs. Health-related quality of life may also deteriorate in the short term as participants adjust to the discovery that they are at higher risk. The health benefits and cost reductions that are expected from the NHS DPP are likely to accrue in the longer-term. There are established simulation models for future consequences of diabetes (http://www.mthooddiabeteschallenge.com/). We will feed the measured effects of NHS DPP into these models to simulate the expected long-term costs and benefits.

A model to support the DPP has been developed by a University of Sheffield (ScHARR) team. We have already started discussions and we will continue to engage with them to maximise consistency. In the analysis, we will separately identify the effects of changes to the model structure from changes to the inputs into the model.

Reflecting the importance of the distributional consequences of NHS DPP, we will use appropriate methods to consider how taking equity into account affects the overall judgement on the cost effectiveness of NHS DPP through the simulation model, linking with work package 1.[122] This analysis will also identify where priority should be focused to improve contribution to social value.

We will use the simulation model to identify the main drivers of cost-effectiveness. We will use scenario analysis to generate proposals for ways in which the cost effectiveness of the programme can be increased. We will consider, for example, which groups generate the highest costs and are at most risk of harmful effects of diabetes.

National coverage within DIPLOMA

In terms of primary data collection (WP 1, 2, 3 and 6), we will seek to ensure that our data collection captures variation which will enable us to understand the national picture, in terms of key geographical factors such as deprivation and rurality. We expect that at least one site for these work packages will include London to ensure that we understand any specific issues.

CPRD covers 6.9% of the UK population (http://ije.oxfordjournals.org/content/44/3/827.full.pdf+html). Since submitting the original bid we have had discussion with another provider of GP electronic data, Research One, which will increase our coverage further (we estimate to over 10% of the English population).

As NDPP wave 1 covers some 26 million patients, we are confident there will be sufficient coverage across databases to enable our planned analyses to progress. All NHS-DPP participants will be coded as such in their GP electronic record, and we will be able to pick these up retrospectively in the CPRD (and identify participating and non-participating practices).

CPRD is a nationally representative group of practices in terms of deprivation, geography and ethnicity (although ethnicity is not coded reliably in the database prior to 2006). UK primary care has been computerised for some time now and we are confident on the completeness of electronic health records. In particular, completeness and reliability increased post 2004 when the QOF was introduced as this incentivised recording of processes and treatments for all major chronic conditions, which enables a comprehensive assessment of co-morbidity. Disability status will be inferred from relevant Read codes concerning BMI. Although the CPRD is adequate, access to the ResearchOne data will improve the representativeness of our data even further and add approximately 5% of the UK population to our sample. Learning disabilities will be measured using recommended Read codes as specified in the 2015/16 GMS contract. Comorbidity will be measured using Read codes concerning major chronic conditions incentivized by the QOF (and hence well recorded)

Dissemination and projected outputs

We will make regular reports to NHS DPP stakeholders, especially NHS England and Public Health England, to ensure that our emerging findings can influence the delivery of the NHS DPP services. We will disseminate the projects outputs through a variety of media, including conference presentations and conventional academic publications, seminars and short accessible reports for stakeholders, and plain English summaries and podcasts for patients and the public. We will work closely with our Stakeholder Advisory Group and the Research Advisory Group to maximise the utility of our dissemination. We have costed in time for patient and public contribution to dissemination.

Plan of investigation and timetable

Pre-funding, we will prepare ethics and other approvals; initiate staff recruitment; negotiate external data access; and strengthen relationships with NHS DPP. **Year 1** we will conduct induction and training (WP 1,2,3,5&7), finalise access to external and NHS DPP data; select sites for case studies (WP2) and for qualitative research (WP1-3). We will set up Stakeholder Advisory, Research Advisory and Study Steering Groups and hold first meetings. We will begin work on WP1, 2 and 5, and discuss options for additional nested studies. **Year 2** we will begin work on all remaining WPs, with a focus on generating and disseminating findings from WP1, 2, 3 and 4. We will begin the WP6 validation study. **Year 3** we will complete the main analyses for WP1, 2, 3, 4 and 6, and start the main analysis for WP5 and 7. **Year 4** we will complete WP5, 6 and 7, complete final reports and publications.

We plan to start the programme on 1st April 2017. We will be in regular contact with key DPP stakeholders to ensure that we communicate early findings in summer 2017 during planned meetings.

Project management

DIPLOMA is a large and complex undertaking. We propose the following project management structure to meet our aims of delivering the research to time, target and budget, and to ensure that the results are disseminated to key stakeholders and inform NHS decision-making.

Given the size and complexity of the programme, we have identified programme management as a specific work package that will be jointly led by Sutton and Bower). This work package will have 5 functions:

Ensuring progress on the research work packages, and effective co-ordination between work packages on common data and issues.

Ensuring suitable patient and public involvement and engagement in the research project, via relevant training and support, ongoing dialogue, and use of opportunities for engagement of contributors at critical points (such as development of patient facing materials, and dissemination).

Reporting to the NIHR HS&DR Study Steering Committee, who will provide supervision on behalf of the Sponsor and Funder and ensure work is conducted to the standards in the Research Governance Framework for Health and Social Care and the Guidelines for Good Clinical Practice. The NIHR HS&DR Study Steering Committee will meet every six months.

Engagement with external stakeholders (Department of Health, NHS England, Public Health England, Diabetes UK, local NHS providers and commissioners and NHS DPP providers), feeding back on progress with the project, sharing of interim findings with stakeholders, and as a forum for sharing other information.

Engagement with NHS DPP. Certain work packages are dependent on data being collected effectively. For example, it is imperative that the referral to the NHS DPP code is used consistently and an in unbiased way by GPs. This work package will ensure that every effort is made to ensure accurate recording through regular audit and feedback.

Each of the work packages has a lead academic. Each work package will set up monthly meetings and will meet the Programme Management leads (Sutton and Bower) quarterly.

An advisory group will be created to provide important clinical and 'expert by experience' input in relation to key areas of the NHS DPP. The advisory group will meet every six months, and will be available for ad hoc advice and support throughout.

Patient and Public Involvement

We recruited 6 people (who were either diagnosed with diabetes, with family history, or 'at risk'), varying in age, gender and ethnicity. We sent our research plans to them and also asked what they would like to know about NHS DPP. Their responses confirmed our view of the core questions: the effect of NHS DPP on diabetes prevention and other outcomes; cost effectiveness; equal access; implementation and service delivery.

The PPI contributors also raised many other interesting issues. This has led to some changes in the content of the research. The contributors raised the issue of what patients might think of the use of commercial providers, and we will explore this issue in the WP1 qualitative interviews. Other areas they felt that it was important to explore included checking on the level of patient and public involvement in NHS DPP itself (WP2 interviews); issues such as confidentiality and links to GP (WP1 interviews); patient choice in NHS DPP, and how patients cope with identification of diabetes and associated issues of risk to health (WP1 and 3 interviews).

The group also felt that it was important to explore the impact of wider social networks on uptake. We have experience of work in social networks,[127 128] and we will explore all these issues in the WP1 qualitative interviews and observations. Finally, our contributors highlighted the importance of patient experience of the consent process and accuracy of information provided during the referral process , and availability of services outside NHS DPP. Again, we will explore these issues through WP1 and 2.

Sanders will act as lead for PPI within the DIPLOMA programme. Co-ordination of PPI will be a key role of the programme manager funded through DIPLOMA. We have budgeted £10,000 for fees to contributors (66 days at INVOLVE rate of £150 per day) and £2,000 for public contributors to the programme to attend relevant conferences alongside the team.

An additional focus on public engagement would be highly relevant for this programme. Engagement is defined as 'the myriad of ways in which the activity and benefits of higher education and research can be shared with the public. Engagement is by definition a two-way process, involving interaction and listening, with the goal of generating mutual benefit'.[1] We will work alongside our PPI partners to create accessible web-based materials explaining the NDPP programme and the DIPLOMA evaluation, to ensure that patients and the public are informed about ongoing developments through the life of the programme.

The Patient and Public Involvement team at Salford Royal Foundation Trust will facilitate two-way public engagement through their existing links with the Research for the Future project and in particular their Help BEAT Diabetes programme. Research for the Future is an NIHR CRN Greater Manchester campaign which aims to encourage people to get more involved with NHS research. It consists of a series of 'Help BEAT' campaigns. There are currently 4200+ people on the Help BEAT Diabetes database, and they are expanding their remit to people at risk of diabetes. We will complement this resource by engagement work in other sites involved in DIPLOMA to ensure national coverage.

Expertise and justification of support required

The project has two co-principal investigsators (Bower and Sutton). Bower will be the lead on the management of the research study (including ethics and governance).

The academic team is listed below. A team of research staff and project managers will be recruited once contracts are secured.

Matthew Sutton will be the Principal Investigator jointly with **Bower**, co-ordinating work packages (with a focus on the quasi-experimental WP5 and 7) and liaising with the NIHR Study Steering Committee, the Research Advisory Group and with stakeholders. He will also lead work package 7. Sutton is a Professor of Health Economics and NIHR Senior Investigator with a worldwide reputation in non-randomised studies of policy and health interventions. He co-authored MRC guidance on natural experiments.[129] His work includes high impact papers on financial incentives and pay for performance,[9] resource allocation) and 7 day services, for which he was an expert witness to the House of Commons Health Select Committee. **Peter Bower** will act as co-PI, assisting with co-ordination (with a focus on the patient and organisation level health services research in WP1-4 and 6) and liaison with stakeholders. Bower is a Professor of Health Services Research with experience of design and delivery of policy-relevant evaluations, including the national evaluation of the Expert Patients Programme, the Whole Systems Demonstrators and complex interventions related to multimorbidity and integrated care. **William Whittaker** will co-lead the Access and Inequality work package with Chandola. Whittaker is a Research Fellow in Health Economics with experience with survey and administrative data. His work includes assessments of access to care, inequalities in health, resource allocation formulae,[130-132] evaluating dentistry services,[133] 7-day primary care,[134] and care inequalities by sexual orientation.[135] **Tarani Chandola** is a Professor of Medical Sociology. His research is primarily on the social determinants of health, focusing on health inequalities and psychosocial factors, and the analysis of longitudinal cohort studies. Much of his research is on stress at work and its effects on health. His current research projects include the MRC funded FRAILL study, the ESRC funded International Centre for Life course Studies in Society and Health and a work stress intervention study funded by the NIHR. **Paul Wilson** will lead the Implementation work package. He has considerable experience in evidence synthesis, knowledge transfer and implementation. He is part of the NIHR CLAHRC Greater Manchester and Deputy Editor in Chief of the journal Implementation Science. **David French** will lead the Fidelity work package. French is a Professor of Health Psychology and member of the Manchester Centre for Health Psychology. He has extensive experience of the design and evaluation of complex interventions to target health behaviours such as exercise, as well as experience of utilising theories of health behaviour change. **Sarah Cotterill** will lead the Effectiveness work package. Cotterill is a Research Fellow with research interests in public health, behaviour change and methodology. She has undertaken evaluation of diabetes prevention programmes, leading an evaluation of an NHS DPP pathfinder site in Salford. She is an NIHR Research Design Service adviser and part of NIHR CLAHRC Greater Manchester. **Caroline Sanders** will lead patient experience work in the Access and Inequality, Fidelity and Effectiveness work packages. Sanders is Senior Lecturer in Medical Sociology and highly experienced in conducting qualitative work using a variety of methods on clinical trials[136 137] and large scale national evaluations, including the Whole Systems Demonstrators.[27] **Evan Kontopantelis** will lead the Effectiveness work package. Kontopantelis is a Reader in Statistics with world-leading expertise in the use of large scale databases to develop and evaluate health policy. His work includes analyses of the impact of the Quality and Outcomes scheme on mortality and inequality, as well as methodological work on optimal functioning of the scheme. **David Reeves** will work within the Effectiveness work package and lead the External Validation work package. Reeves is a Reader in Statistics and a health services researcher with extensive experience in the design of evaluations of relevance to health policy, as well as analyses of clinical databases to support health policy evaluation. **Rachel Meacock** will work with Sutton on the Comparative Long-Term Cost Effectiveness work package with. Meacock is a Research Fellow in Health Economics. **Simon Heller** is Professor of Clinical Diabetes and National Diabetes Speciality Lead for the NIHR CRN and will provide clinical expertise